Spectral Efficiency Analysis of Cell-Free Distributed Massive MIMO Systems With Imperfect Covariance Matrix

Feng Ye, Student Member, IEEE, Jiamin Li, Member, IEEE, Pengcheng Zhu, Member, IEEE, Dongming Wang, Member, IEEE, Hao Wu, and Xiaohu You, Fellow, IEEE

Abstract—In this article, the impacts of imperfect channel covariance matrix on the spectral efficiency (SE) of cell-free distributed massive multiple-input–multiple-output (MIMO) systems are analyzed. We propose to estimate the channel covariance matrix by alternately using the assigned pilots and their phase-shifted pilots in different coherent blocks, which improves the accuracy of channel estimation with imperfect covariance matrix and reduces pilot overhead. Under this scheme, the closed-form expressions of SE with maximum ratio combination (MRC) and zero-forcing (ZF) receivers are derived, which enable us to select key parameters for better system performance. Simulation results verify the effectiveness of the proposed channel estimation method and the accuracy of the derived closed-form expressions. When more coherent blocks are used to estimate the covariance matrix, we can get better system performance. Moreover, some insightful conclusions are arrived from the SE comparisons between different receiving schemes (ZF and MRC) and different pilot allocation schemes (orthogonal pilot and pilot reuse).

Index Terms—Cell-free distributed massive multiple-input–multiple-output (MIMO), channel estimation, covariance matrix estimation, pilot contamination, spectral efficiency.

I. INTRODUCTION

Due to the increasing requirements for high system throughput, low latency, ultrareliability, and near-instant connection, more innovative technologies are needed to pave the way for the development of B5G and 6G technologies [1]. A large number of access points (APs) distributed in an area serve all users in the same time-frequency resource [2], [3].

Manuscript received 2 May 2021; revised 12 October 2021; accepted 23 January 2022. Date of publication 17 February 2022; date of current version 9 December 2022. This work was supported in part by the National Key Research and Development Program of China under Grant 2020YFB1806600, and by the National Natural Science Foundation of China (NSFC) under Grant 61971127, Grant 61871465, and Grant 61871122. (Corresponding author: Jiamin Li.)

Feng Ye and Pengcheng Zhu are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: yefeng@seu.edu.cn; p.zhu@seu.edu.cn).

Jiamin Li, Dongming Wang, and Xiaohu You are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China, and also with the Purple Mountain Laboratories, Nanjing 211111, China (e-mail: jiaminli@seu.edu.cn; wangdm@seu.edu.cn; xhyu@seu.edu.cn).

Hao Wu is with the State Key Laboratory of Mobile Network and Mobile Multimedia Technology, ZTE Corporation, Shenzhen 518055, China (e-mail: wu.hao89@zte.com.cn).

Digital Object Identifier 10.1109/JSYST.2022.3147516

Cell-free distributed massive multiple-input–multiple-output (MIMO) can get all the benefits from massive MIMO (which facilitates propagation and channel strengthening when multiple antennas are used at the APs [4]) and network MIMO (which increases macrodiversity gain). Therefore, it has a very high spectral efficiency (SE), energy efficiency, and coverage. In addition, the cell-free structure can solve the problem that user performance and user location are related in the cellular network. Compared with the centralized system, cell-free distributed massive MIMO system has advantages such as channel diversity, no switching, higher coverage, and no need deploying cells in specific areas [5]. Therefore, in recent years, cell-free distributed massive MIMO has aroused widespread research interest [6]–[10].

Obtaining channel state information (CSI) is essential for taking advantage of the cell-free distributed massive MIMO system. There have been some studies on channel estimation methods. In most of these studies, the minimum mean square error (mmse) channel estimation method is extended to the cell-free distributed MIMO system, or make certain improvements on this basis [6], [11]–[14]. There may be other channel estimation methods, such as least-squares (LS) channel estimation, which result in very poor system performance [15]. Based on the estimated CSI, the performance analysis of cell-free massive MIMO system has also been well investigated in literature [16]. However, these studies are based on the large-scale fading information that is perfectly available at the base station (BS), and the estimation is based on the small-scale fading information. They ignore that the covariance matrix is imperfect and the mobility of the user will change the covariance matrix in practice. Therefore, the estimation of the covariance matrix is a supplement to the existing channel estimation techniques in the cell-free massive MIMO architecture, and the nonideal factors of imperfect covariance matrix are considered in the actual situation. In addition, pilot contamination is more serious in massive MIMO systems than in traditional MIMO systems. Because the covariance matrix contains the CSI of the uplink channel, estimation of the channel covariance matrix can effectively reduce the impact of pilot contamination and improve the accuracy of channel estimation [17]. Therefore, the estimation of covariance matrix is not only a requirement of the actual system but also an effective means to improve system performance.
Generally speaking, there are six typical methods to obtain the covariance matrix as follows.

1) Use a specific phase for learning the covariance matrix, where every user in the network uses a unique orthogonal pilot, from which a sample correlation matrix can be formed, possibly by using regularization to get robustness [18].

2) Eliminate the interference by sending pilots from other interfering users to obtain the covariance matrix of the target user [17].

3) Add additional pilot sequences for covariance estimation in the coherent blocks to obtain accurate channel estimation [19].

4) Instead of reserving a specific phase for learning the covariance matrix, the pilot assignment is changed between different coherence blocks [20].

5) Use the mmse estimator of the structure model as the blueprint of the neural network structure, and a similarly effective (but suboptimal) estimator is obtained [21].

6) Phase-rotate the pilot sequence in each coherence block to approximately decorrelate the channel observations at the APs [22].

However, these methods are faced with the problems of low estimation accuracy and high computation complexity. Additionally, these methods are all covariance matrix acquisition methods in a centralized cellular scenario, and there is no relevant research in a cell-free distributed massive MIMO scenario.

To meet the needs of super-dense user scenarios, the total number of antennas in cell-free massive MIMO systems needs to be increased greatly. With the increase of the total number of antennas in the system, the spatial multiplexing gain of the system will be greatly improved, and the number of users who can serve in the same frequency will also be greatly increased.

However, the increase of system capacity is established when the CSI is known. If orthogonal pilots are used to estimate the CSI, the pilot overhead will increase linearly with the increasing of the number of antennas and will ultimately restrict the improvement of the overall transmission efficiency. For uplink wireless transmission, although the pilot overhead is only proportional to the number of users and independent of the number of antennas, it is still too overhead if the completely orthogonal pilot is adopted considering the large number of users the cell-free distributed massive MIMO needs to support. Therefore, it is necessary to adopt a pilot multiplexing scheme to improve the utilization of pilot resources, which will result in pilot contamination [23]–[26].

Pilot contamination makes the estimation of the covariance matrix more complicated, because the users using the same pilot sequences will cause interference, so the channel estimation derived from covariance matrix estimation is itself contaminated. Naively utilizing the contaminated channel estimations in a sample covariance estimator will result in the target user covariance matrix estimation containing the covariance matrices of the interference users, which in turn leads to inaccurate estimation of the covariance matrix and ultimately affects the overall transmission efficiency.

In this article, we propose a new covariance matrix estimation method and derive the closed-form expressions for the uplink achievable rates in a cell-free distributed massive MIMO system with both maximum ratio combination (MRC) and ZF receivers under pilot contamination. The main contributions of this article are summarized as follows.

1) We propose to estimate the channel covariance matrix by using alternate assigned pilots and their phase-shifted pilots in different coherent blocks, which obtains accurate channel estimation with low complexity and without extra pilots.

2) It is essential to study the performance of a massive MIMO system for a more realistic case, where the covariance matrices are unknown and need to be estimated. Given the estimated channel covariance matrix, we derive accurate closed-form expressions of the uplink achievable rates with both MRC and ZF receivers which enables the SE analysis of cell-free distributed massive MIMO systems with imperfect covariance matrix.

3) Numerical simulations are performed to corroborate our theoretical analysis, and insightful conclusions are drawn from the comparison between system performance with perfect and imperfect covariance matrix using different receivers with different pilot allocation schemes and the analysis of the relationship between the number of coherent blocks used to calculate the covariance estimation and the SE, which can be used to choose parameters based on SE requirements.

The remainder of this article is organized as follows. In Section II, we describe the system model including the system configuration, channel model, and channel estimation with pilot contamination in the case of perfect or imperfect covariance matrix. Section III contains pilot structure design and covariance estimation. Section IV provides the SE analysis of cell-free distributed massive MIMO systems with both MRC and ZF receivers which enables the SE analysis of cell-free distributed massive MIMO systems with imperfect covariance matrix. Section V introduces numerical results and drawn conclusions. Section VI concludes this article.

**Notation:** Boldface letters stand for matrices (upper case) or vectors (lower case). The transpose and conjugate transpose are denoted by \((\cdot)^T\) and \((\cdot)^H\), respectively. \(I_{M \times N}\) stands for the \(M \times N\) identity matrix, and \(CN(\mu, \sigma^2)\) denotes the circularly symmetric complex Gaussian distribution with mean \(\mu\) and variance \(\sigma^2\), while \(\mathcal{W}(N, R)\) denotes Wishart matrix with \(N\) degrees of freedom and \(R\) is the covariance matrix that corresponds to underlying Gaussian random vectors.

## II. System Model

This article considers a cell-free distributed massive MIMO system. There are \(M\) APs equipped with \(N\) antennas and \(K\) single-antenna users. Assuming that the system adopts a time division duplex (TDD) mode, the BS uses the reciprocity of the uplink and downlink channels to perform downlink multiuser precoding based on the uplink channel estimation results.

### A. Channel Model

Considering a block fading channel, and the channel vector from user \(k\) to all APs can be modeled as [27]

\[
g_k = \Lambda_k^{1/2} h_k \in \mathbb{C}^{M \times 1}
\]  

(1)
with

\[
\Lambda_k = \text{diag}(\lambda_{k,1}, \cdots, \lambda_{k,M})^T \otimes I_N \in \mathbb{C}^{MN \times MN}
\]

(2)

where \(\lambda_{km} \triangleq d_{k,m}^{-\zeta} s_{k,m}\) represents the large-scale fading between user \(k\) and AP \(m\); \(d_{k,m}\) is the distance from user \(k\) to AP \(m\); \(\zeta\) is the path loss exponent; and \(s_{k,m}\) is a log-normal shadow fading variable. In addition, \(h_k \in \mathbb{C}^{MN \times 1}\) models small-scale fast fading, and each element follows \(CN(0, 1)\). So the channel \(g_k\) satisfies \(\mathcal{N}_c(0, \Lambda_k)\), that is \(E[g_k] = 0, \text{cov}(g_k, g_k) = \Lambda_k\).

The uplink received signal model is

\[
y = \sqrt{\rho} \sum_{k=1}^{K} g_k x_k + n
\]

(3)

where \(y\) represents information received from all users; \(\rho\) is the sending power; \(x_k\) is the sending information; \(n \sim \mathcal{N}_c(0, \sigma^2)\) is the noise; and \(\sigma^2\) is the noise power. We assume that \(g_k\) and \(n\) are independent of each other.

**B. Uplink Channel Estimation**

As widely recognized in most of articles, we first assume that the large-scale fading \(\Lambda_k\) is known to the BS, and the small-scale fading \(h_k\) needs to be estimated at the BS. Then, based on the actual situation that the covariance matrix is imperfect, so the large-scale fading \(\Lambda_k\) is also unknown to the BS, and both the large-scale fading \(\Lambda_k\) and the small-scale fading \(h_k\) need to be estimated at the BS.

1) Case I: Perfect Covariance Matrix: It is assumed that there are \(P\) orthogonal pilot sequences of length \(\tau\) to be used for uplink channel estimation, which is a constant independent of \(K\), and the pilots are allocated to \(K\) users when accessing the network. When \(P < K\), there will be multiple users using the same pilot, resulting in pilot contamination, which leads to the deterioration of the accuracy of channel estimation. Let \(\mathcal{U}_p \subset \{1, \ldots, K\}\) denote the subset of users using the pilot \(p\), all users using the \(p\)th pilot transmit the pilot signal \(X_p\), on the same time-frequency resource, and satisfy \(X_p X_p^H = I_P\). After correlating the received training signal with the conjugate of the pilot sequence, the BS estimates the channel based on the observations

\[
y_p = \sqrt{\rho} \sum_{i \in \mathcal{U}_p} g_i + n_p.
\]

Since \(g_k\) satisfies \(\mathcal{N}_c(0, \Lambda_k)\), for \(k \in \mathcal{U}_p\), the mmse estimate of channel \(g_k\) should be

\[
\hat{g}_k = \mathbb{E}[g_k] + \text{cov}(g_k, y_p) \text{cov}(y_p, y_p)^{-1}(y_p - \mathbb{E}[y_p])
\]

\[
= \frac{\Lambda_k}{\sqrt{\rho}} \left( \sum_{i \in \mathcal{U}_p} \Lambda_i + \frac{\sigma^2}{\rho} I_{MN} \right)^{-1} y_p
\]

(5)

with

\[
\Sigma_p = \text{cov}(y_p, y_p) = \rho \left( \sum_{i \in \mathcal{U}_p} \Lambda_i + \frac{\sigma^2}{\rho} I_{MN} \right)
\]

(6)

where \(\Sigma_p\) is the covariance matrix of the received signal.

Therefore, for \(k \in \mathcal{U}_p\), the channel estimation expression with perfect covariance matrix is

\[
\hat{g}_k = \sqrt{\rho} \Lambda_k \Sigma_p^{-1} y_p.
\]

(7)

2) Case II: Imperfect Covariance Matrix: In practice, the covariance matrix is usually imperfect, and the aforementioned channel estimation under the perfect covariance matrix does not meet actual requirements. At this point, both \(\Lambda_k\) and \(\Sigma_p\) are unknown parameters, so the channel estimation expression is transformed into

\[
\hat{g}_k = \hat{\Lambda}_k \Sigma_p^{-1} y_p.
\]

(8)

Hence, if we want to obtain the accurate channel estimation with imperfect covariance matrix, we must accurately estimate the covariance matrix. At the same time, it can also be seen in both (6) and (8) that the covariance matrix cannot be calculated directly by simply using the contaminated channel, and the covariance matrix needs to be estimated separately.

Assuming that the channel is a block fading model, it remains unchanged within a certain coherence bandwidth \(B_c\) and a certain coherence time \(T_c\), that is, the channel remains unchanged within \(\tau_s = B_c T_c\). The covariance matrix is constant in the transmission bandwidth. Compared with the rapid change of the small-scale fading, it changes slowly in time, so it can be reasonably assumed that it remains unchanged within \(\tau_s = B_c T_c\). The coherence blocks, where \(B_s\) is the system bandwidth and \(T_s\) is the time interval. Therefore, channel estimation is required for each coherent block, but the covariance matrix only needs to be estimated once in \(\tau_s\) coherent block. According to this property, we estimate the covariance matrix and channel by using a special pilot structure.

In the following section, we propose a method of using different pilots in adjacent coherent blocks. Compared with the common method of adding extra pilots for covariance matrix estimation, this scheme can greatly reduce the pilot overhead without losing estimation accuracy.

**III. PROPOSED PILOT STRUCTURE AND METHOD FOR ESTIMATING COVARIANCE MATRIX**

In this section, we first describe the pilot structure design. Then based on the structure, we estimate the sample covariance matrix and individual covariance matrix, respectively.

**A. Pilot Structure Design**

Considering the block fading model, multiple coherent blocks can be used to estimate the covariance matrix. Since each coherent block sends pilots for channel estimation, in order not to add additional pilots, we consider using the channel estimation pilots for simultaneous covariance matrix estimation. However, due to pilot contamination, in order to obtain interference-free observations for each user, we should change the pilot structure. This pilot structure needs to be able to eliminate interference for accurate covariance matrix estimation without affecting channel estimation, so we can change the pilot by introducing random phase shift.
In addition, in order to have a low complexity system, we should use a relatively simple pilot structure. So we consider a use of alternate pilots to estimate the covariance matrix, that is, each user alternately sends allocated pilot $X_p$ and phase-shifted pilot $\Phi_k$ in adjacent coherent blocks. $\Phi_k$ is the pilot after the pilot $X_p$ undergoes a random phase shift

$$
\Phi_k[n] = e^{j\theta_k,n}X_p[2n - 1], n = 1, 2, \cdots,
$$

(9)

where $n$ is the index of coherent blocks, $\theta_k,n$ is a random phase shift, which is irrelevant to the channel vector and noise and satisfies $\mathbb{E}[e^{j\theta_k,n}] = 0$. So this random phase-shifted pilot $\Phi_k$ is irrelevant to the channel and noise. The detailed structure of the pilot is depicted in Fig. 1.

It is assumed that different pilots using the same pilot has different random phase shift $\theta_k,n$. By observing the received signals of the alternate pilot and using the irrelevance of phase shift term to remove the interference and noise of other users, an accurate covariance matrix estimation can be obtained. In practical applications, $\theta_k,n$ can be obtained by a pseudorandom sequence generator, and it is assumed that these phase shifts are known at the central processing unit (CPU).

### B. Sample Covariance Matrix Estimation

Based on the pilot structure, the sample correlation matrix is used to approximate $\Sigma_p$. We select the coherent blocks that send pilot $X_p$, that is, the interval coherent blocks are used to estimate $\Sigma_p$. Assuming that the pilots of $N_{\Sigma} \leq \tau_s/2$ coherent blocks are received, the sample covariance matrix can be obtained as

$$
\hat{\Sigma}_p = \frac{1}{N_{\Sigma}} \sum_{n=1}^{N_{\Sigma}} y_p[2n - 1](y_p[2n - 1])^H
$$

(10)

where $y_p[2n - 1]$ represents the received signal of the $(2n - 1)$th coherent block. When $N_{\Sigma}$ is larger, the estimated $\hat{\Sigma}_p$ will be more accurate.

In fact, it is not necessary to select coherent blocks sending pilot $X_p$. As long as the interval coherence blocks are selected to estimate the sample correlation matrix, the same result will be obtained.

### C. Individual Covariance Matrix Estimation

In order to estimate individual covariance matrix $\hat{\Lambda}_k$, adjacent coherent blocks are needed to remove the correlation. Based on the pilot sequence $\Phi_k$, the pilot signal received by the APs from the user is

$$
y_p^{(\Phi)} = \sqrt{\rho} \sum_{i \in \mathcal{U}_p} g_i e^{j\theta_i,2n} + n_p.
$$

(11)

According to the received signals of the alternate pilots, for $k \in \mathcal{U}_p$, the observed values of the received pilot signals of the two adjacent coherent blocks (i.e., coherent block pairs) are, respectively,

$$
\hat{h}_k^{(1)}[2n - 1] = y_p^{(\Phi)} = \sqrt{\rho} \sum_{i \in \mathcal{U}_p} g_i e^{j\theta_i,2n}
$$

$$
\hat{h}_k^{(2)}[2n] = y_p^{(\Phi)} = \rho \sum_{i \neq k, i \in \mathcal{U}_p} g_i e^{j(\theta_i,2n - \theta_k,2n)} + n_p e^{-j\theta_k,2n}
$$

(13)

where $\hat{h}_k^{(1)}[2n - 1]$ is obtained by (4).

For $k \in \mathcal{U}_p$, the channel covariance matrix estimated from the observations of the adjacent alternate pilot signals is

$$
cov(\hat{h}_k^{(1)}[2n - 1], \hat{h}_k^{(2)}[n]) = \mathbb{E} \left[ \left( \hat{h}_k^{(1)}[2n - 1] \right) \left( \hat{h}_k^{(2)}[n] \right)^H \right]
$$

$$
= \mathbb{E} \left[ \rho g_k g_k^H \right]
$$

$$
= \rho \Lambda_k.
$$

(14)

The equation is established because it is assumed that $g_k$ and $n_p$ are independent of each other, and because the random phase $\theta_k,2n$ is introduced, the two noises are not correlated and interference terms from other users using the same pilot can also be eliminated. Therefore, the estimate of $\Lambda_k$ can be obtained as

$$
\hat{\Lambda}_k = \frac{1}{N_{\Lambda}N_{\Sigma}} \sum_{n=1}^{N_{\Lambda}} \hat{h}_k^{(1)}[2n - 1] \left( \hat{h}_k^{(2)}[2n] \right)^H.
$$

(15)

However, for a finite $N_{\Lambda}$, the estimate $\hat{\Lambda}_k$ in (15) is not necessarily Hermitian symmetric. Therefore, this matrix can be regularized by approximating it with a positive semidefinite matrix. We approximate $\Lambda_k$ with the positive semidefinite matrix closest to the Frobenius norm, i.e., $\Lambda_k^{PSD} = \mathbf{U} \Lambda_k^H \mathbf{U}^H$, where $\mathbf{D}_+$ is a diagonal matrix that contains only the positive eigenvalues of the symmetric part of $\Lambda_k$, i.e., $\mathbf{D}_+ \triangleq (\Lambda_k + \Lambda_k^H)/2$, and $\mathbf{U}$ contains the corresponding eigenvectors. Therefore, $\Lambda_k$ is written in the form of $\Lambda_k^{PSD}$, which is written as

$$
\hat{\Lambda}_k = \frac{1}{2N_{\Lambda}N_{\Sigma}} \sum_{n=1}^{N_{\Lambda}} \left( \hat{h}_k^{(1)}[2n - 1] \left( \hat{h}_k^{(2)}[2n] \right)^H + \left( \hat{h}_k^{(1)}[2n - 1] \right)^H \hat{h}_k^{(2)}[2n] \right)
$$

(16)
where \( N_A \) represents the number of coherent block pairs required to estimate \( \Lambda_k \). When \( N_A \to \infty \), the estimated covariance matrix converges to the real covariance matrix, i.e., \( \hat{\Lambda}_k \to \Lambda_k \).

In summary, we substitute the estimated \( \hat{\Lambda}_k \) and \( \hat{\Sigma}_p \) into (8), and we can get the channel estimation when the covariance matrix is imperfect

\[
\hat{g}_k = \hat{\Lambda}_k \hat{\Sigma}_p^{-1} y_p = \hat{W}_k y_p
\]

where

\[
\hat{W}_k = \hat{\Lambda}_k \hat{\Sigma}_p^{-1}
\]

is the deterministic matrix for the estimators.

**Remark 1:** We select alternate pilots with different phase shifts for individual covariance matrix estimation, and then use interval phase-shift-free pilots for sample covariance estimation. In this scheme, we can obtain accurate covariance matrix estimation and accurate channel estimation without adding additional pilots in the case of low system complexity.

**Remark 2:** In practical applications, the BS transmits phase sequence information to the pseudorandom sequence generator in the initial stage. Each user then sends the assigned pilot multiplied by the phase sequence in each coherent block. After the signal is received, the covariance matrix was estimated once for the time-frequency interval of \( \tau_c \) coherent blocks, and \( N_S \) coherent blocks and \( N_A \) coherent block pairs were selected to track the channel changes in real time. The proposed method also have some software and hardware implementation challenges in practical applications, such as the compromise between accuracy of channel estimation and operating speed of system, the memory capacity and computational resources of CPU and the synchronization between the transmitting pilot and the pseudorandom sequence generator.

### IV. UPLINK PERFORMANCE WITH IMPERFECT CHANNEL COVARIANCE INFORMATION

In this section, we analyze the SE of the system based on the above channel estimation method. To derive the closed-form expressions of the achievable rates, we first give the ideal lower boundary of the user achievable rates and then use random matrix theory to perform expectation transformation.

The signal model (3) received by the APs can be rewritten as [28]

\[
w_k^H y_p = \sqrt{p} E\left[ w_k^H g_k \right] + \sqrt{p} \left( w_k^H g_k - E\left[ w_k^H g_k \right] \right) + \sum_{i \neq k} \sqrt{p} w_i^H g_i + w_i^H n_p
\]

where \( w_k \) represents the receiving vector, which can be defined as

\[
w_k = \left\{ \begin{array}{ll} \hat{g}_k, & \text{for MRC} \\ \hat{z}_k, & \text{for ZF} \end{array} \right.
\]

where \( \hat{z}_k \) is the \( k \)-th column of \( \hat{G}(\hat{G}^H \hat{G})^{-1} \), \( \hat{G} = [\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K] \), and \( \hat{g}_k \) is the estimation of \( g_k \).

The lower boundary of the user achievable rate can be expressed as [19]

\[
R_k = \left( 1 - \frac{P}{\tau_c} \right) \log_2 \left( 1 + \gamma_k \right)
\]

where \( P/\tau_c \) represents the symbols occupied by the pilot. According to (19), the signal-to-noise ratio (SNR) \( \gamma_k \) is defined as

\[
\gamma_k = \frac{\left| E\left[ w_k^H g_k \right]^2 \right|}{\sum_{i=1}^{K} E\left[ |w_i^H g_i|^2 \right] - |E\left[ w_k^H g_k \right]|^2 + \sigma^2 E\left[ w_k^H w_k \right]}
\]

### A. MRC

When MRC receiver is used, the receiver vector is equal to the channel estimate. Substituting (8) into (22) and based on Lemma 1 shown below, we can get the closed-form expression of uplink achievable rates.

**Lemma 1:** For a complex Wishart matrix, \( W \) satisfies \( \forall \left( n, I \right) \) and \( n > m \), it has the property [29]

\[
E\left[ \text{tr} \left\{ W^{-1} \right\} \right] = \frac{m}{n - m}.
\]

For \( n > m + 1 \)

\[
E\left[ \text{tr} \left\{ W^{-2} \right\} \right] = \frac{mn}{(n-m)^2 - (n-m)}
\]

\[
E\left[ \left| \text{tr} \left( W^{-1} A \right) \right|^2 \right] = \frac{\left| \text{tr} (A)^2 \right| + \frac{1}{n-m} \text{tr} (AA^H)}{(n-m)^2 - 1}
\]

where \( A \in C^{m \times m} \) is a random matrix.

**Theorem 1:** When MRC receiver is used, the closed-form approximation of SNR is given by

\[
\gamma_k^{MRC} = \sum_{i=1}^{K} T_i^{EX} + \sum_{i \neq i_p} T_i^{IN} - \frac{N_k}{N_{\Sigma-M \times N}} \text{tr}(\hat{W}_k^H \Lambda_k) + \sigma^2 N_k
\]

where

\[
T_i^{EX} = \frac{\mu_1 M N}{2N_A} \text{tr}(\Lambda_i \Sigma_i^p) + \frac{\mu_2}{N_A} \text{tr}(\Sigma_i^p \Lambda_i A_i \Lambda_k) + \mu_1 \text{tr}(\hat{W}_i^H \hat{A}_i A_i \hat{A}_k) + \frac{\mu_2}{N_A} \text{tr}(\hat{W}_k^H \hat{A}_i^2 W_i) + \frac{\mu_2}{N_A} \text{tr}(\hat{W}_i^H \hat{A}_i \Sigma_i^p W_i^H) + \frac{\mu_2}{N_A} \text{tr}(\hat{W}_i^H \hat{A}_i \Sigma_i^p W_i^H)
\]

\[
T_i^{IN} = \frac{\mu_1 M N}{2N_A} \text{tr}(\Sigma_i^p \Lambda_k) + \frac{\mu_2}{N_A} \text{tr}(\Sigma_i^p \Lambda_i A_i \Lambda_k) + \mu_1 \text{tr}(\hat{W}_i^H \hat{A}_i A_i \hat{A}_k) + \frac{\mu_2}{N_A} \text{tr}(\hat{W}_k^H \hat{A}_i^2 W_i) + \frac{\mu_2}{N_A} \text{tr}(\hat{W}_i^H \hat{A}_i \Sigma_i^p W_i^H) + \frac{\mu_2}{N_A} \text{tr}(\hat{W}_i^H \hat{A}_i \Sigma_i^p W_i^H)
\]

\[
N_k = \mu_1 \text{tr}(\hat{W}_i^H \hat{A}_i) + \frac{\mu_1 M N}{2N_A} \text{tr}(\Sigma_i^p) + \frac{\mu_2}{N_A} \text{tr}(\hat{W}_i^H \hat{A}_i W_i^H) + \text{tr}(\hat{W}_i^H \hat{A}_i W_i^H) + \frac{\mu_2}{N_A} \text{tr}(\hat{W}_i^H \hat{A}_i W_i^H)
\]

and \( \mu_2 = \frac{N_k^2}{(N_{\Sigma-M \times N})^2 - 1} \).

**Proof:** Please refer to Appendix A.

In addition, when the total number of antennas \( MN \to \infty \), the limit rate of each user is given in (27), shown at the bottom of the next page.

It can be seen that when the MRC receiver is used, even if the AP is equipped with a large number of antennas, the SNR of the receiver still tends to be a constant related to the covariance matrix, and its performance still has a bottleneck. It can also be seen that the SNR of the receiver can be improved by reasonably allocating users using the same pilot.
B. ZF

After ZF detection, the received signal can be written as

$$\hat{G}^H \hat{G}^{-1} \hat{G}^H y = x + \left( \hat{G}^H \hat{G}^{-1} \hat{G}^H \hat{G} x + n \right)$$  \hspace{1cm} (28)

where $\hat{G} = [\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K]$, $\hat{g}_k$ represents the error of channel estimation

$$\hat{g}_k = g_k - \hat{g}_k.$$  \hspace{1cm} (29)

Therefore, the signal to interference plus noise ratio (SINR) $\gamma_k$ can be expressed as

$$\gamma_k^{ZF} = \frac{1}{e_k^H \left( \hat{G}^H \hat{G}^{-1} \hat{G}^H \hat{G} - 1 \right) e_k}$$  \hspace{1cm} (30)

where $\tilde{\Gamma}$ is defined as the interference item

$$\tilde{\Gamma} = \text{COV} (\hat{G} x + n)$$

$$= \sum_{k=1}^{K} \left( A_k - A_k \Sigma_p^{-1} A_k \Sigma_p \right) + \sigma^2 I_{MN}.$$  \hspace{1cm} (31)

**Lemma 2:** For $k \in \mathcal{U}_p$, the channels between the users using the same pilot and the APs have the following relationship:

$$\frac{1}{MN} \hat{G}^H \hat{G} \xrightarrow{a.s.} 0, p \neq i$$  \hspace{1cm} \hspace{1cm} (32)

$$\frac{1}{MN} \hat{G}^H \hat{G} \xrightarrow{a.s.} 0, p \neq i$$  \hspace{1cm} \hspace{1cm} (33)

$$\frac{1}{MN} \hat{G}^H \hat{G} - \frac{1}{MN} \Sigma_p \xrightarrow{a.s.} 0$$  \hspace{1cm} \hspace{1cm} (34)

$$\frac{1}{MN} \hat{G}^H \hat{G} - \frac{1}{MN} \Sigma_p \xrightarrow{a.s.} 0$$  \hspace{1cm} \hspace{1cm} (35)

where, for $i \in \mathcal{U}_p$ has

$$[\Sigma_p]_{q,j} = \text{tr} \left( A_k \Sigma_p^{-1} \Sigma_p^{-1} A_k \Sigma_p \right)$$  \hspace{1cm} \hspace{1cm} (36)

$$[\tilde{\Sigma}_p]_{q,j} = \text{tr} \left( A_k \Sigma_p^{-1} \Sigma_p^{-1} A_k \Sigma_p \tilde{\Gamma} \right).$$  \hspace{1cm} \hspace{1cm} (37)

Based on the above expressions, when $MN \to \infty$, (30) can be transformed into

$$\gamma_k^{ZF} = \frac{N}{e_k^H \tilde{\Sigma}_p^{-1} e_k e_k}.$$  \hspace{1cm} (38)

Next, by approximating the estimates, the theoretical expression of SINR of the ZF receiver can be obtained.

**Theorem 2:** When ZF receiver is used, the closed-form approximation of SINR is given by

$$\gamma_k^{ZF} = \frac{N}{e_k^H \tilde{\Sigma}_p^{-1} e_k e_k}$$  \hspace{1cm} (39)

where

$$[\tilde{\Sigma}_p]_{q,j} = \mu_1 \text{tr} \left( A_k \Sigma_p^{-1} \Lambda_k \right) + \mu_1 \frac{M N}{2N_A} \text{tr} \left( \Sigma_p \right)$$

$$+ \frac{\mu_1}{2N_A} \text{tr} \left( A_k \Sigma_p^{-1} \Lambda_k \right).$$  \hspace{1cm} (40)

$$[\tilde{\Sigma}_p]_{q,j} = \mu_1 \text{tr} \left( A_k \Sigma_p^{-1} \Lambda_k \tilde{\Gamma} \right) + \mu_1 \frac{M N}{2N_A} \text{tr} \left( \Sigma_p \tilde{\Gamma} \right)$$

$$+ \frac{\mu_1}{2N_A} \text{tr} \left( A_k \tilde{\Gamma} \right)$$  \hspace{1cm} (41)

and (42), shown at the bottom of this page, where $\tilde{\Sigma}_p$ is a Wishart matrix, and it satisfies $\mathcal{W}(N_\Sigma, I_{MN})$.

**Proof:** Please refer to Appendix B.

It can be seen that unlike MRC, using ZF receiver, when the number of antennas of each APs $N \to \infty$, the SINR increases linearly as (39). At the same time, it can be known that the characteristic of matrix $\Sigma_p$ is an important factor that restricts system performance, and it is closely related to pilot contamination. When $\Sigma_p$ is rank-deficient, the performance of the ZF receiver will also be seriously degraded.

In summary, the closed-form expressions of SINR in both MRC and ZF receivers are related to the number of coherent blocks $N_\Sigma$ and $N_A$ to calculate the covariance matrix. Better estimation of the covariance matrix can lead to higher SNR, and, thus, improve the system performance.

V. SIMULATION RESULTS

In this section, we fit the closed-form expressions with the simulation values, and compare them with the case, where the covariance matrix is perfect. The accuracy of channel estimations with perfect and imperfect covariance matrices are also discussed. With the numerical results, we analyze the accuracy of channel estimation with imperfect covariance matrix.

The simulation system model is shown in Fig. 2; we consider a cell-free distributed massive MIMO system in which blue triangles represent APs and red dots represent users. To meet
the situation of massive MIMO $MN \gg K[19]$, we set $M = 5$ APs with $N = 50$ antennas uniformly distributed in a circle with a diameter of 0.5 km and $K = 10$ users randomly distributed in the area. We consider a uniform linear array, a model used in a large number of MIMO literature to show the effects of strong spatial correlations and there may be several decibels channel gain variations between antennas in practical applications [30].

The channel model in (1) is used, where the path loss exponent $\zeta$ is 3.7 in an outdoor scene of massive MIMO [31], the reference distance is 1 and the variance of shadow fading satisfies exponential normal distribution. We consider a high mobility scenario with the system bandwidth $B_s = 40$ MHz. According to [17], the coherence bandwidth is $B_c = 200$ kHz and the coherence time is $T_c = 1$ ms, which means the number of symbols in a coherent block is $\tau_c = 200$ and allows for 270 km/h mobility (meet the speed requirement of high-speed rail) at a 2 GHz carrier frequency. Measurements in [32] show that large-scale fading is constant over time intervals about 100 times longer than the coherence time, hence we need a time interval $T_s = 0.2$ s to meet the requirement of $\tau_s = 20000$. For the number of pilot symbols, a rule of thumb is that $\tau$ should always be smaller than $\tau_c/2[33]$, so we choose $\tau = 10$. For comparison, we consider the following three aspects.

1) Channel Estimation With Perfect Covariance Matrix (Cov Known): The covariance matrix satisfies (2), and we perform the mmse estimation of the channel to get the sum of user rates.

2) Simulated Value of Channel Estimation With Imperfect Covariance Matrix (Simulated): We directly use the covariance matrix estimation results of (10) and (16) for channel estimation, and get the sum of user rates to judge system performance.

3) Theoretical Value of Channel Estimation with Imperfect Covariance Matrix (Theoretical): We derive the Theorem 1 and the Theorem 2, and obtain the closed expression of the SINR to get the sum of user rates.

Based on the above three aspects, we simulate MRC and ZF, respectively.

Fig. 3 shows the sum of user rates when $N = 1000$ with completely orthogonal pilots. It can be seen from Fig. 3 that the simulated value fits well with the theoretical value. Therefore, the correctness of our derived closed-form expressions is proved.

When using MRC and ZF, the sum of user rates with perfect covariance are close to the situation when the covariance is imperfect, and MRC is closer than ZF because ZF is more sensitive to noise. We can even see that when using MRC receivers, our covariance estimation scheme can achieve 95% or higher SE with a perfect covariance matrix. This proves that the covariance matrix estimation method we used can bring good system performance. But when using ZF, the sum of user rates are higher than using MRC. This is because the ZF receiver can eliminate inter-user interference. In addition, the theoretical value and the simulation value of ZF is not fitting well than MRC. This is because, when deriving theoretical expressions, in the case of ZF, conversion is required first and then derivation. The errors brought by two approximations are larger, and the errors in derivation may become more obvious in the form of product.

From the curve of the sum of user rates with perfect covariance matrix, it can be seen that the sum of user rates increase as the number of coherent blocks $N_A$ used to calculate the individual covariance estimate increases. This result is due to the fact that the covariance estimation is not accurate when $N_A$ is small, and the inaccurate channel estimation results in a lower sum of user rates. As $N_A$ continues to increase, the accuracy of the channel estimation improves, and the sum of user rates increase accordingly.

Fig. 4 shows the sum of user rates when $N = 1000$ with $P = 8$ pilots. Because of the large number of users in a dense distributed MIMO system, the situation of each user using completely orthogonal pilots cannot be satisfied. So it is necessary to consider the situation of pilot contamination, that is, multiple users use the same pilot. It can be seen that pilot contamination has a certain impact on the performance of MRC and ZF. Pilot multiplexing can improve system resource utilization, and a good pilot allocation method can achieve a tradeoff between resource utilization and system performance. Fig. 5 shows the normalized mean square error (NMSE) of channel estimation with completely orthogonal pilots. The NMSE of channel estimation is defined as

$$\text{NMSE} = \frac{\sum_{k=1}^{K} E\left[|\hat{g}_k - g_k|^2\right]}{\sum_{k=1}^{K} E\left[|g_k|^2\right]}.$$ (43)
It is proved that when the covariance matrix is imperfect, the channel estimation obtained by using our covariance matrix estimation method is close to the channel estimation with perfect covariance matrix. In addition, it can be seen that the number of coherent blocks $N_{\Sigma}$ used to estimate the covariance matrix $\Sigma_p$ of the received signal will actually affect the accuracy of channel estimation, and a more accurate estimation of $\Sigma_p$ will have better system performance. The increase of $N_{\Sigma}$ is only a better estimate of $\Sigma_p$ under the existing structure. At this time, the channel estimation with imperfect covariance matrix is closer to the perfect situation. This is at the cost of calculation. In actual situations, a good CPU is needed to increase $N_{\Sigma}$ for better system performance.

VI. CONCLUSION

This article proposes a method to estimate the covariance matrix by alternating pilots in adjacent coherent blocks. This method improves the channel estimation method without extra pilot overhead to achieve better system performance when the covariance matrix is imperfect. We derived the closed-form expression of uplink SE with MRC and ZF receivers and verified the closed-form expressions. We also analyze the influence of the covariance matrix calculation with different number of coherent blocks on the system performance. The simulation results show that the channel estimation algorithm with imperfect channel covariance information can achieve high estimation accuracy. At the same time, a better estimation of the covariance matrix of the received signal leads to better system performance. In addition, for different receivers, pilot contamination can decrease system performance. Our work will be particularly important and the application will be extremely versatile especially in the future of a larger number of users and higher mobility scenarios.

APPENDIX A

PROOF OF THEOREM 1

Based on the random matrix theory (see [29]), we provide the following proof which is divided into two steps. First, if the combining vector $w_k = \hat{W}_k g_k$ is used, we need to convert the expression to $\hat{W}_k$. For $E[w_k^H g_k]$

$$E[w_k^H g_k] = E[y_p^H \hat{W}_k^H g_k] \overset{(a)}{=} E_W \left[ \text{tr} \left( \hat{W}_k^H \Lambda_k \right) \right]$$

where (a) is obtained because the noise $n_p$ and the channel $g_k$ are independent of each other, and $x^H y = \text{tr}(yx^H)$ for any vectors $x$ and $y$. By using the same trace rule as above and identifying $\Sigma_p$, for $E[w_k^H w_k]$, we have

$$E[w_k^H w_k] = E[y_p^H \hat{W}_k^H \hat{W}_k y_p] = E_W \left[ \text{tr} \left( \hat{W}_k y_p \hat{W}_k^H y_p \right) \right] = E_W \left[ \text{tr} \left( \hat{W}_k \Sigma_p \hat{W}_k^H \right) \right].$$

Finally, for $E[w_k^H g_i]^2$

$$E[w_k^H g_i]^2 = E[w_k^H g_i \Sigma_p^{-1} w_k].$$

If $i \notin U_p$, then

$$E[w_k^H g_i^H w_k] = E[w_k^H \Lambda_i w_k] = E_W \left[ \text{tr}(\hat{W}_k E[y_p y_p^H] \hat{W}_k^H \Lambda_i) \right] = E_W \left[ \text{tr} \left( \hat{W}_k \Sigma_p \hat{W}_k^H \Lambda_i \right) \right].$$

If $i \in U_p$, $w_k$ and $g_i$ are not independent of each other, they will be coupled together, and, thus, cannot be solved according to the above method. So, we have

$$E[w_k^H g_i^H w_k] \overset{(a)}{=} E \left[ (w_k - \hat{W}_k g_i)^H g_i \right]^2 + E \left[ g_i^H \hat{W}_k g_i \right]^2$$

$$= E_W \left[ \text{tr}(\hat{W}_k (\Sigma_p - \Lambda_i) \hat{W}_k^H \Lambda_i) + \text{tr} \left( \hat{W}_k^H \Lambda_i \right)^2 \right] + E_W \left[ \text{tr} \left( \hat{W}_k \Lambda_i \hat{W}_k^H \Lambda_i \right) \right]$$

$$= E_W \left[ \text{tr} \left( \hat{W}_k \Sigma_p \hat{W}_k^H \Lambda_i \right) \right] + E_W \left[ \text{tr} \left( \hat{W}_k^H \Lambda_i \right)^2 \right].$$ (48)
where (a) is because \( w_k - \hat{W}_k g \) and \( g \) are independent.

Next, continue to solve the expectation of \( \hat{W}_k \) and convert the above expressions into closed-form expressions. For any random vector \( g \) that is independent of each other and each elements obeying the \( \mathcal{N}(0, 1) \) distribution

\[
E[g g^H A g g^H] = A + I tr (A)
\]

(49)
define

\[
\Sigma_p = N_\Sigma \left( \Sigma_p^{-1/2} \Sigma_p \Sigma_p^{-1/2} \right).
\]

(50)

Then \( \Sigma_p \) is a Wishart matrix, and it satisfies \( W(N_\Sigma, 1) \). Hence, for \( E_W[tr(\hat{W}_k^H A_k)] \), we have that

\[
E_W \left[ tr \left( \hat{W}_k^H A_k \right) \right] = E_W \left[ tr \left( \hat{\Sigma}_p^{-1} \hat{A}_k A_k \right) \right]
\]

(51)

where (a) and (b) are obtained by substituting (50) and (23), respectively. In addition, for \( E_W[tr(\hat{W}_k \Sigma_p \hat{W}_k^H)] \)

\[
E_W \left[ tr \left( \hat{W}_k \Sigma_p \hat{W}_k^H \right) \right] = E_W \left[ tr \left( \hat{A}_k \Sigma_p^{-1} \hat{\Sigma}_p \Sigma_p^{-1} \hat{A}_k \right) \right]
\]

(52)

where (a), (b), and (c) result from (50), (24), and (49), respectively. And for \( E_W[tr(\hat{W}_k \Sigma_p \hat{W}_k^H A_k)] \)

\[
E_W \left[ tr \left( \hat{W}_k \Sigma_p \hat{W}_k^H A_k \right) \right] = E_W \left[ tr \left( \hat{A}_k \Sigma_p^{-1} \hat{\Sigma}_p \Sigma_p^{-1} \hat{A}_k A_k \right) \right]
\]

(53)

where (a), (b), and (c) are obtained by substituting (50), (24), and (49), respectively. Finally, for \( E_W[tr(\hat{W}_k^H A_k)]^2 \)

\[
E_W \left[ tr \left( \hat{W}_k^H A_k \right) \right]^2 = E_W \left[ tr \left( \hat{\Sigma}_p^{-1} \hat{A}_k A_k \right) \right]^2
\]

(54)

where (a), (b), and (c) are obtained by substituting (50), (25), and (49), respectively. This completes the proof.

**APPENDIX B PROOF OF THEOREM 2**

Similarly, using random matrix theory to solve (39) parameters. For \( \Xi_{p}^{th} \), we have

\[
\Xi_{p}^{th} = \left[ \begin{array}{c}
\left( \hat{A}_k \Sigma_p^{-1} \Sigma_p^{-1} \hat{A}_k \Sigma_p \right) \\
\left( \hat{A}_k \Sigma_p^{-1/2} N_\Sigma \Sigma_p^{-1/2} \hat{A}_k \Sigma_p \right) \\
N_\Sigma^2 \left( \hat{A}_k \Sigma_p^{-1/2} N_\Sigma \Sigma_p^{-1/2} \hat{A}_k \Sigma_p \right)
\end{array} \right]
\]

(55)

where (a), (b), and (c) result from (50), (24), and (49), respectively. Follow the same steps as above, we have

\[
\Xi_{p}^{th} = \left[ \begin{array}{c}
\left( \hat{A}_k \Sigma_p^{-1} \Sigma_p^{-1} \hat{A}_k \Sigma_p \hat{\Gamma} \right) \\
\left( \hat{A}_k \Sigma_p^{-1/2} N_\Sigma \Sigma_p^{-1/2} \hat{A}_k \Sigma_p \hat{\Gamma} \right) \\
N_\Sigma^2 \left( \hat{A}_k \Sigma_p^{-1/2} N_\Sigma \Sigma_p^{-1/2} \hat{A}_k \Sigma_p \hat{\Gamma} \right)
\end{array} \right]
\]

(56)
\[ \begin{align*}
&= \mu_1 \text{tr} \left( A_k \Sigma_p^{-1} \hat{A} \hat{\Gamma} \right) + \mu_1 \frac{MN}{2A} \text{tr} \left( \Sigma_p \hat{\Gamma} \right) \\
&\quad + \frac{\mu_1}{2A} \left( A_k \hat{\Gamma} \right) \text{tr} \left( \Sigma_p^{-1} A_k \Sigma_p \right) \\
&= \sum_{k=1}^{K} \left[ A_k - \mathcal{E}^{\text{cov}} \right] + \sigma^2 \mathbf{I}_{MN} 
\end{align*} \]

where (a) is obtained by substituting (50), (b) results from (24), and (c) results from (49). Here, we regard the interference \( \hat{\Gamma} \) as an independent variable, and we also can get approximately accurate results. So the variable \( \hat{\Gamma} \) is

\[ \hat{\Gamma} = \sum_{k=1}^{K} [A_k - \mathcal{E}^{\text{cov}}] + \sigma^2 \mathbf{I}_{MN} \]

(a) results from (49), and (b) from (50).

This completes the proof.

**REFERENCES**

[1] M. Shafi et al., “5G: A tutorial overview of standards, trials, challenges, deployment, and practice,” IEEE J. Sel. Areas Commun., vol. 35, no. 6, pp. 1201–1221, Jun. 2017.

[2] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, “Cell-free massive MIMO versus small cells,” IEEE Trans. Wireless Commun., vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[3] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, “Ubiquitous cell-free massive MIMO communications,” EURASIP J. Signal Process. Image Process., vol. 2019, no. 1, 2019, Art. no. 197.

[4] Z. Chen and E. Björnson, “Can we rely on channel hardening in cell-free massive MIMO?”, in Proc. IEEE Globecom Workshops, 2017, pp. 1–6.

[5] S. Buzzi and C. D’Andrea, “Cell-free massive MIMO: User-centric approach,” IEEE Wireless Commun. Lett., vol. 6, no. 6, pp. 706–709, Dec. 2017.

[6] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, “On the total energy efficiency of cell-free massive MIMO,” IEEE Trans. Green Commun. Netw., vol. 2, no. 1, pp. 25–39, Mar. 2017.

[7] J. Zhang, Y. Wei, E. Björnson, Y. Han, and L. Li, “Spectral and energy efficiency of cell-free massive MIMO systems with hardware impairments,” in Proc. IEEE 9th Int. Conf. Wireless Commun. Signal Process., 2017, pp. 1–6.

[8] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, and M. Debbah, “Cell-free massive MIMO with limited backhaul,” in Proc. 2018 IEEE Int. Conf. Commun., 2018, pp. 1–7.

[9] L. D. Nguyen, T. Q. Duong, H. Q. Ngo, and K. Tourki, “Energy efficiency in cell-free massive MIMO with zero-forcing precoding design,” IEEE Commun. Lett., vol. 21, no. 8, pp. 1871–1874, Aug. 2017.

[10] T. M. Hoang, H. Q. Ngo, T. Q. Duong, H. D. Tuan, and A. Marshall, “Cell-free massive MIMO networks: Optimal power control against active eavesdropping,” IEEE Trans. Commun., vol. 66, no. 10, pp. 4724–4737, Oct. 2018.

[11] T. C. Mai, H. Q. Ngo, and T. Q. Duong, “Cell-free massive MIMO systems with multi-antenna users,” in Proc. IEEE Glob. Conf. Signal Inf. Process., 2018, pp. 828–832.

[12] H. T. Dao and S. Kim, “Effective channel gain-based access point selection in cell-free massive MIMO systems,” IEEE Access, vol. 8, pp. 108127–108132, 2020.

[13] Y. Jin, J. Zhang, S. Jin, and B. Ai, “Channel estimation for cell-free mmWave massive MIMO through deep learning,” IEEE Trans. Veh. Technol., vol. 68, no. 10, pp. 10325–10329, 2019.

[14] Z. Wang, J. Zhang, E. Björnson, and B. Ai, “Uplink performance of cell-free massive MIMO over spatially correlated Ricean fading channels,” IEEE Commun. Lett., vol. 25, no. 4, pp. 1348–1352, 2020.

[15] S. Mishra, R. Prusty, and P. K. Hota, “Analysis of Levenberg-Marquardt and scaled conjugate gradient training algorithms for artificial neural network based LS and MMSE estimated channel equalizers,” in Proc. IEEE Int. Conf. Man Mach. Interfacing, 2015, pp. 1–7.

[16] P. Liu, K. Luo, D. Chen, and T. Jiang, “Spectral efficiency analysis of cell-free massive MIMO systems with zero-forcing detector,” IEEE Trans. Wireless Commun., vol. 19, no. 2, pp. 795–807, Feb. 2019.

[17] E. Björnson, L. Sanguinetti, and M. Debbah, “Massive MIMO with imperfect channel covariance information,” in Proc. IEEE 50th Asilomar Conf. Signals Syst. Comput., 2016, pp. 974–978.

[18] 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.

[19] K. Upadhyya and S. A. Vorobyov, “Covariance matrix estimation for massive MIMO,” IEEE Signal Process. Lett., vol. 25, no. 4, pp. 546–550, Apr. 2018.

[20] 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.

[21] D. Neumann, T. Wiese, and W. Utschick, “Learning the MMSE channel estimator,” IEEE Trans. Signal Process., vol. 66, no. 11, pp. 2905–2917, Jun. 2018.

[22] G. Interdonato, P. Frenger, and E. G. Larsson, “Self-learning detector for the cell-free massive MIMO uplink: The line-of-sight case,” in Proc. IEEE 21st Int. Workshop Signal Process. Adv. Wireless Commun., 2020, pp. 1–5.

[23] S. Haghhatshoar and G. Caire, “Massive MIMO pilot decontamination and channel interpolation via wideband sparse channel estimation,” IEEE Trans. Wireless Commun., vol. 16, no. 12, pp. 8316–8332, Dec. 2017.

[24] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” IEEE Trans. Wireless Commun., vol. 9, no. 11, pp. 3950–3960, Nov. 2010.

[25] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, “Pilot contamination and precoding in multi-cell TDD systems,” IEEE Trans. Wireless Commun., vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[26] A. Pitarokoilis, E. Björnson, and E. G. Larsson, “On the effect of imperfect timing synchronization on pilot contamination,” in Proc. IEEE Int. Conf. Commun., 2017, pp. 1–6.

[27] J. Cao, D. Wang, J. Li, Q. Sun, and Y. Hu, “Uplink spectral efficiency analysis of multi-cell multi-user massive MIMO over correlated Ricean channel,” Science China Inf. Sci., vol. 61, no. 8, 2018, Art. no. 082305.

[28] T. L. Marzetta and H. Q. Ngo, Fundamentals of Massive MIMO. New York, NY, USA: Cambridge Univ. Press, 2016.

[29] A. M. Tulino, S. Verdú, and S. Verdu, Random Matrix Theory and Wireless Communications. New York, NY, USA: Now Publishers Inc., 2004.

[30] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, “Massive MIMO in real propagation environments: Do all antennas contribute equally?,” IEEE Trans. Commun., vol. 63, no. 11, pp. 3917–3928, Nov. 2015.

[31] E. Björnson, J. Hoydis, and L. Sanguinetti, “Massive MIMO networks: Spectral, energy, and hardware efficiency,” Foundations Trends Signal Process., vol. 11, no. 3/4, pp. 154–655, 2017.

[32] I. Viering, H. Hofstetter, and W. Utschick, “Spatial long-term variations in urban, rural and indoor environments,” in Proc. 5th Meeting COST273, Lisbon, Portugal, Sep. 2002.

[33] E. Björnson, E. G. Larsson, and M. Debbah, “Massive mimo for maximal spectral efficiency: How many users and pilots should be allocated?,” IEEE Trans. Wireless Commun., vol. 15, no. 2, pp. 1293–1308, Feb. 2015.
Jiamin Li (Member, IEEE) received the B.S. and M.S. degrees in communication and information systems from Hohai University, Nanjing, China, in 2006 and 2009, respectively, and the Ph.D. degree in information and communication engineering from Southeast University, Nanjing, China, in 2014.

He joined the National Mobile Communications Research Laboratory, Southeast University, in 2014, where he has been an Associate Professor, since 2019.

His research interests include cell-free distributed massive MIMO, massive ultrareliable low-latency communications, artificial intelligence and its applications in future mobile communications.

Pengcheng Zhu (Member, IEEE) received the B.S and M.S. degrees in electrical engineering from Shandong University, Jinan, China, in 2001 and 2004, respectively, and the Ph.D. degree in communication and information science from Southeast University, Nanjing, China, in 2009.

He has been a Lecturer with the National Mobile Communications Research Laboratory, Southeast University, since 2009. His research interests include communication and signal processing, limited feedback techniques, and distributed antenna systems.

Dongming Wang (Member, IEEE) received the B.S. degree from Chongqing University of Posts and Telecommunications, Chongqing, China, the M.S. degree from Nanjing University of Posts and Telecommunications, Nanjing, China, and the Ph.D. degree from the Southeast University, Nanjing, China, in 1999, 2002, and 2006, respectively, all in electrical engineering.

He joined the National Mobile Communications Research Laboratory, Southeast University, in 2006, where he has been an Associate Professor, since 2010.

His research interests include turbo detection, channel estimation, distributed antenna systems, and large-scale MIMO systems.

Hao Wu received the B.S. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2010, and the Ph.D. degree from Southeast University, Nanjing, China, in 2015, both in electrical engineering.

He is currently with the State Key Laboratory of Mobile Network and Mobile Multimedia Technology, ZTE Corporation, Shenzhen, China, where he is a Senior Expert on wireless communication research and standardization. From 2011 to 2012, he was a Visiting Student with Columbia University, New York, NY, USA. Since 2016, he has been a delegate representing ZTE Corporation in 3GPP RAN and RAN1, to which he has submitted numerous contributions on 4G and 5G technologies, including MIMO, UE power saving, and positioning. His research interests include MIMO wireless communications, antenna array systems, and signal processing.

Xiaohu You (Fellow, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Nanjing Institute of Technology, Nanjing, China, in 1982, 1985, and 1989, respectively.

From 1987 to 1989, he was with Nanjing Institute of Technology, as a Lecturer. Since 1990, he has been with Southeast University, Nanjing, China, first as an Associate Professor and later as a Professor. He is also the Chief of the Technical Group of China 3G/B3G Mobile Communication R&D Project. His research interests include mobile communications, adaptive signal processing, and artificial neural networks with applications to communications and biomedical engineering.

Dr. You was the recipient of the Excellent Paper Prize from the China Institute of Communications, in 1987, the Elite Outstanding Young Teacher awards from Southeast University, in 1990, 1991, and 1993, and the 1989 Young Teacher Award of Fok Ying Tung Education Foundation, State Education Commission of China.