

Uplink Precoding Design for Cell-Free Massive MIMO with Iteratively Weighted MMSE

Zhe Wang, Jiayi Zhang, *Senior Member, IEEE*, Hien Quoc Ngo, *Senior Member, IEEE*,
Bo Ai, *Fellow, IEEE* and Mérouane Debbah, *Fellow, IEEE*

Abstract

In this paper, we investigate a cell-free massive multiple-input multiple-output system with both access points and user equipments equipped with multiple antennas over the Weichselberger Rayleigh fading channel. We study the uplink spectral efficiency (SE) for the fully centralized processing scheme and large-scale fading decoding (LSFD) scheme. To further improve the SE performance, we design the uplink precoding schemes based on the weighted sum SE maximization. Since the weighted sum SE maximization problem is not jointly over all optimization variables, two efficient uplink precoding schemes based on Iteratively Weighted sum-Minimum Mean Square Error (I-WMMSE) algorithms, which rely on the iterative minimization of weighted MSE, are proposed for two processing schemes investigated. Furthermore, with maximum ratio combining applied in the LSFD scheme, we derive novel closed-form achievable SE expressions and optimal precoding schemes. Numerical results validate the proposed results and show that the I-WMMSE precoding schemes can achieve excellent sum SE performance with a large number of UE antennas.

Index Terms

Cell-free massive MIMO, uplink precoding, weighted sum-rate maximization, spectral efficiency.

I. INTRODUCTION

Cell-free massive multiple-input multiple-output (CF mMIMO) has attracted a lot of research interest and is regarded as a promising technology for future wireless communications, for its ability to achieve

This article was presented in part at IEEE International Conference on Communications 2022 [1].

Z. Wang and J. Zhang are with the School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China, and also with the Frontiers Science Center for Smart High-speed Railway System, Beijing Jiaotong University, Beijing 100044, China (e-mail: {zhewang_77, jiayizhang}@bjtu.edu.cn).

H. Q. Ngo is with the Institute of Electronics, Communications, and Information Technology, Queen's University Belfast, BT3 9DT Belfast, U.K. (email: hien.ngo@qub.ac.uk).

B. Ai is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China, also with the Frontiers Science Center for Smart High-Speed Railway System and the Henan Joint International Research Laboratory of Intelligent Networking and Data Analysis, Zhengzhou University, Zhengzhou 450001, China, and also with the Research Center of Networks and Communications, Peng Cheng Laboratory, Shenzhen 518066, China (e-mail: boai@bjtu.edu.cn).

M. Debbah is with the Technology Innovation Institute, Abu Dhabi, United Arab Emirates, and also with CentraleSupélec, University Paris-Saclay, 91192 Gif-sur-Yvette, France (e-mail: merouane.debbah@tii.ae).

uniformly high spectral efficiency (SE) [2]–[7]. Basically, a large number of access points (APs), arbitrarily distributed in a wide coverage area and connected to one or several central processing units (CPUs), jointly serve all user equipments (UEs) on the same time-frequency resource. Compared with the traditional cellular mMIMO system, the CF mMIMO system operates with no cell boundaries and many more APs than UEs [8]–[10]. Relying upon the prominent network topology of CF mMIMO, four uplink (UL) signal processing schemes, distinguished from levels of the mutual cooperation between all APs and the assistance from the CPU, can be implemented as [5]. Among these schemes, the “*Level 4*” and “*Level 3*” are viewed as efficient processing techniques. The so-called *Level 4* is a fully-centralized processing scheme where all the pilot and data signals received at APs are transmitted to the CPU via the fronthaul links and the CPU performs channel estimation and data detection. The similar scheme was also investigated in [11]–[13]. The so-called *Level 3* stands for a two layer decoding scheme: in the first layer, each AP estimates channels and decodes the UE data locally by applying an arbitrary combining scheme based on the local channel state information (CSI); in the second layer, all the local estimates of the UE data are gathered at the CPU in which they are linearly weighted by the optimal large-scale fading decoding (LSFD) coefficient to obtain the final decoding data. The LSFD scheme has been widely investigated in [14]–[17] since it can make full use of the prominent network topology for CF mMIMO and achieve excellent performance.

To promote the practical implementation of the CF mMIMO network, a new framework of scalable CF mMIMO system and its respective processing algorithms were proposed in [9] by exploiting the dynamic cooperation cluster (DCC) concept. Besides, the scalability aspects in a realistic scenario with multiple CPUs were considered in [18], where the data processing, network topology and power control strategies with multiple CPUs were discussed. Moreover, the authors of [19] considered the uplink of a radio-strip-based CF mMIMO network architecture with sequential fronthaul links between APs and proposed MMSE-based sequential processing schemes, which significantly reduced the fronthaul requirement. However, when the CF mMIMO network is operated in practice, a more practical capacity-constrained fronthaul network would have a great effect on the system performance. The authors of [20] and [21] discussed the uplink performance of a CF mMIMO system with limited capacity fronthaul links. Furthermore, it is worth noting that the CF mMIMO architecture has been co-designed with another promising future wireless technology: Reconfigurable Intelligent Surface (RIS) [22], [23], which would undoubtedly provide vital tutorials for the future wireless network design.

The vast majority of scientific papers on CF mMIMO focus on the scenario with single-antenna UEs. However, in practice, contemporary UEs with moderate physical sizes have already been equipped with multiple antennas to achieve higher multiplexing gain and boost the system reliability. The authors of [24] investigated the UL performance of a CF mMIMO system with multi-antenna UEs over maximum ratio (MR) combining and zero-forcing (ZF) combining. The authors of [25] considered a user-centric (UC) approach for CF mMIMO with multi-antenna UEs and proposed power allocation strategies for either sum-rate maximization or minimum-rate maximization. Besides, the authors of [26] analyzed the downlink SE performance for a CF mMIMO system with multi-antenna UEs and computed SE expressions in closed-form. Then, the SE performance for a CF mMIMO system with multi-antenna UEs and low-resolution DACs was investigated in [27]. Nevertheless, these works only investigated a simple distributed processing scheme and are based on the overly idealistic assumption of independent and identically distributed (i.i.d.) Rayleigh fading channels, neglecting the spatial correlation that has a significant impact on practical CF mMIMO systems [15], [16]. The authors of [28] considered a CF mMIMO system with multi-antenna UEs over the jointly-correlated Weichselberger model [29] and analyzed four UL processing schemes.

As observed in [26], [28], increasing the number of antennas per UE may not always benefit the SE performance. The SE would reach the maximum value with particular number of antennas per UE, then decrease with the increase of number of antennas per UE. One main reason for this phenomenon is that the UEs cannot make full use of the benefit of equipping with multiple antennas to achieve higher SE performance without UL precoding schemes. So it is undoubtedly vital to design the UL precoding scheme to further improve the performance of systems. However, it is worth noting that the design of UL precoding for CF mMIMO has not been investigated. For the traditional mMIMO or MIMO systems, one popular optimization objective for the uplink/downlink precoding design is to maximize the weighted sum rate (WSR) [30]–[33]. The authors of [30] and [32] discussed the equivalence between the WSR maximization problem and the Weighted sum-Minimum Mean Square Error (WMMSE) problem in MIMO systems and proposed an iteratively downlink transceiver design algorithm for the WSR maximization. Note that the algorithm relies on the iterative minimization of weighted MSE since the WMMSE problem are not jointly convex over all optimization variables. Moreover, the authors of [31] investigated the UL precoding scheme optimization based on [30] under sum-power-constraint or individual-power-constraint.

Motivated by the above observations, we investigate a CF mMIMO system with both multi-antenna APs and UEs over the Weichselberger Rayleigh fading channel. Two pragmatic processing schemes: 1)

the fully centralized processing scheme; 2) the large-scale fading decoding scheme are implemented. The main contributions are given as follows.

- We design an efficient UL precoding scheme to maximize the WSR for the fully centralized processing scheme based on an iteratively WMMSE (I-WMMSE) algorithm. Note that the design of I-WMMSE precoding scheme for the fully centralized processing scheme is implemented at the CPU and based on the instantaneous CSI.
- For the LSFD processing scheme, we derive a UL precoding scheme for the WSR maximization based on an iteratively WMMSE algorithm. The design of I-WMMSE precoding scheme for the LSFD scheme is implemented at the CPU but based only on channel statistics. More importantly, we compute achievable SE expressions and optimal precoding schemes in novel closed-form for the LSFD scheme with MR combining.
- We analyze the practical implementation and computation complexity for the proposed I-WMMSE precoding schemes. It is found that the proposed I-WMMSE precoding schemes can be guaranteed to converge. More importantly, the proposed UL precoding schemes are efficient to achieve excellent sum SE/rate performance and the average rate benefits from the multiple antennas at the UE-side, which undoubtedly provides vital insights for the practical implementation of multi-antenna UEs.

Note that this paper differs from the conference version [1] in the following aspects: *i)* we investigate the fully centralized processing/LSFD schemes and design their respective UL precoding schemes, while only the LSFD scheme was considered in [1]; *ii)* we provide details for the derivation of the I-WMMSE precoding schemes, which are omitted in [1] due to the lack of space; *iii)* we analyze the practical implementation and convergence behavior of the proposed precoding schemes. More importantly, numerical results show vital insights for the CF mMIMO system with the proposed UL precoding schemes.

The rest of this paper is organized as follows. In Section II, we consider a CF mMIMO system with the Weichselberger Rayleigh fading channel, and describe the channel estimation and data detection. Then, Section III introduces the fully centralized processing and LSFD processing schemes, and provides their respective achievable SE expressions. Novel closed-form SE expressions for the LSFD scheme with MR combining are derived. More importantly, based on the achievable SE expressions, we propose UL I-WMMSE precoding schemes for two processing schemes. Then, Section IV provides some insights for the practical implementation and computation complexity of proposed I-WMMSE precoding schemes. In Section V, numerical results and performance analysis for the I-WMMSE precoding schemes are provided.

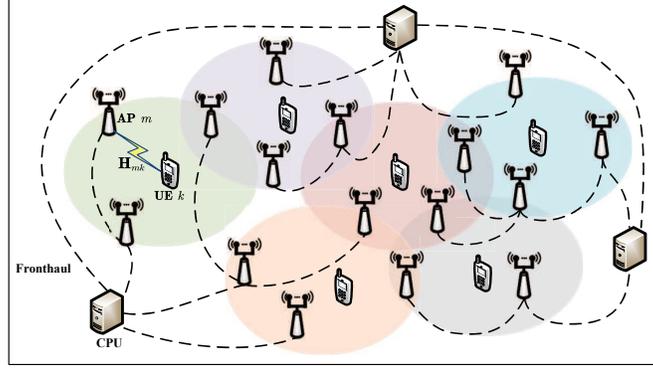


Fig. 1. A cell-free massive MIMO system.

Finally, the major conclusions and future directions are drawn in Section VI.

Notation: Lowercase letters \mathbf{x} and boldface uppercase letters \mathbf{X} denote the column vectors and matrices, respectively. $\mathbb{E}\{\cdot\}$, $\text{tr}\{\cdot\}$ and \triangleq are the expectation operator, the trace operator, and the definitions, respectively. $|\cdot|$, $\|\cdot\|$ and $\|\cdot\|_F$ are the determinant of a matrix or the absolute value of a number, the Euclidean norm and the Frobenious norm, respectively. $\text{vec}(\mathbf{A})$ denotes a column vector formed by the stack of the columns of \mathbf{A} . The $n \times n$ identity matrix is represented by $\mathbf{I}_{n \times n}$. The Kronecker products and the element-wise products are denoted by \otimes and \odot , respectively. Finally, $\mathbf{x} \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{R})$ is a circularly symmetric complex Gaussian distribution vector with correlation matrix \mathbf{R} .

II. SYSTEM MODEL

In this paper, we investigate a CF mMIMO system consisting of M APs and K UEs, where all APs are connected to one or several CPUs via fronthaul links as shown in Fig. 1. For simplicity, there is only one CPU and all APs serve all UEs¹. The numbers of antennas per AP and UE are L and N , respectively. A standard block fading model is investigated, in which the channel response is constant and frequency flat in a coherence block of τ_c -length (channel uses). Let τ_p and $\tau_c - \tau_p$ denote channel uses dedicated for the channel estimation and data transmission, respectively. We denote by $\mathbf{H}_{mk} \in \mathbb{C}^{L \times N}$ the channel response between AP m and UE k . We assume that \mathbf{H}_{mk} for different AP-UE pairs are independent.

A. Channel Model

Based on the jointly-correlated (also known as the Weichselberger model [29]) Rayleigh fading channel², \mathbf{H}_{mk} is modeled as

$$\mathbf{H}_{mk} = \mathbf{U}_{mk,r} \left(\tilde{\boldsymbol{\Omega}}_{mk} \odot \mathbf{H}_{mk,\text{iid}} \right) \mathbf{U}_{mk,t}^H \quad (1)$$

¹As shown in Fig. 1, a more practical network topology is with multiple CPUs and dynamic cooperation clusters, where each UE is only served by a cluster of APs and the APs are grouped into cell-centric clusters. Each cell-centric cluster is connected to a particular CPU.

²Note that the Rayleigh fading channel is a special case of the Rician fading channel. And the performance gap between the Rician channel and the Rayleigh channel is small [34]. However, the focus of this paper is not on the channel model but on the UL precoding scheme design. So for the simplicity of analysis, we investigate an essential Rayleigh fading channel by assuming there is no line-of-sight (LoS) link between each UE and AP.

where $\mathbf{U}_{mk,r} = [\mathbf{u}_{mk,r,1}, \dots, \mathbf{u}_{mk,r,L}] \in \mathbb{C}^{L \times L}$ and $\mathbf{U}_{mk,t} = [\mathbf{u}_{mk,t,1}, \dots, \mathbf{u}_{mk,t,N}] \in \mathbb{C}^{N \times N}$ are the eigenvector matrices of the one-sided correlation matrices $\mathbf{R}_{mk,r} \triangleq \mathbb{E}[\mathbf{H}_{mk} \mathbf{H}_{mk}^H]$ and $\mathbf{R}_{mk,t} \triangleq \mathbb{E}[\mathbf{H}_{mk}^T \mathbf{H}_{mk}^*]$, and $\mathbf{H}_{mk,\text{iid}} \in \mathbb{C}^{L \times N}$ is composed of i.i.d. $\mathcal{N}_{\mathbb{C}}(0, 1)$ random entries, respectively. Besides, we denote by $\boldsymbol{\Omega}_{mk} \triangleq \tilde{\boldsymbol{\Omega}}_{mk} \odot \tilde{\boldsymbol{\Omega}}_{mk} \in \mathbb{R}^{L \times N}$ the ‘‘eigenmode coupling matrix’’ with the (l, n) -th element $[\boldsymbol{\Omega}_{mk}]_{ln}$ specifying the average amount of power coupling from $\mathbf{u}_{mk,r,l}$ to $\mathbf{u}_{mk,t,n}$. \mathbf{H}_{mk} can also be formed as $\mathbf{H}_{mk} = [\mathbf{h}_{mk,1}, \dots, \mathbf{h}_{mk,N}]$ with $\mathbf{h}_{mk,n} \in \mathbb{C}^L$ being the channel between AP m and n -th antenna of UE k . By stacking the columns of \mathbf{H}_{mk} on each other, we define $\mathbf{h}_{mk} \triangleq \text{vec}(\mathbf{H}_{mk}) = [\mathbf{h}_{mk,1}^T, \dots, \mathbf{h}_{mk,N}^T]^T \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{R}_{mk})$, where $\mathbf{R}_{mk} \triangleq \mathbb{E}\{\mathbf{h}_{mk} \mathbf{h}_{mk}^H\}$ is the full correlation matrix

$$\mathbf{R}_{mk} = (\mathbf{U}_{mk,t}^* \otimes \mathbf{U}_{mk,r}) \text{diag}(\text{vec}(\boldsymbol{\Omega}_{mk})) (\mathbf{U}_{mk,t}^* \otimes \mathbf{U}_{mk,r})^H. \quad (2)$$

Moreover, note that \mathbf{R}_{mk} can be structured into the block form as [28] with the (n, i) -th submatrix being $\mathbf{R}_{mk}^{ni} = \mathbb{E}\{\mathbf{h}_{mk,n} \mathbf{h}_{mk,i}^H\}$. Besides, the large-scale fading coefficient β_{mk} can be extracted from \mathbf{R}_{mk} as $\beta_{mk} = \frac{1}{LN} \text{tr}(\mathbf{R}_{mk}) = \frac{1}{LN} \|\boldsymbol{\Omega}_{mk}\|_1$. It is worth mentioning that the motivations for adopting the Weichselberger channel model are: 1) The Weichselberger model investigated in (1) not only captures the correlation features at both the AP-side and UE-side but models the joint correlation dependence between each AP-UE pair through the coupling matrix; 2) The coupling matrix $\boldsymbol{\Omega}_{mk}$ reflects the practical spatial arrangement of scattering objects between AP m and UE k . More significantly, the Weichselberger model can reduce to most channel models of great interest by adjusting the coupling $\boldsymbol{\Omega}_{mk}$ to particular formulation, such as the Kronecker model and i.i.d. Rayleigh fading model [28], [29]; 3) Compared with other stochastic channel models, the Weichselberger model displays significantly less modeling error, which is validated based on the practical measurement in [29].

B. Channel Estimation

For the channel estimation, mutually orthogonal pilot matrices are constructed and each pilot matrix is composed of N mutually orthogonal pilot sequences. We denote by Φ_k the pilot matrix assigned to UE k with $\Phi_k^H \Phi_l = \tau_p \mathbf{I}_N$, if $l = k$ and $\mathbf{0}$ otherwise. And \mathcal{P}_k is the index subset of UEs using the same pilot matrix as UE k including itself. When all UEs transmit their pilot matrices, the received signal at AP m $\mathbf{Y}_{mk}^p \in \mathbb{C}^{L \times \tau_p}$ is $\mathbf{Y}_m^p = \sum_{k=1}^K \mathbf{H}_{mk} \mathbf{F}_{k,p} \Phi_k^T + \mathbf{N}_m^p$, where $\mathbf{F}_{k,p} \in \mathbb{C}^{N \times N}$ is the precoding matrix for UE k under the phase of pilot transmission, $\mathbf{N}_m^p \in \mathbb{C}^{L \times \tau_p}$ is the additive noise at AP m with independent $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ entries and σ^2 being the noise power, respectively. The pilot transmission should be implemented under the power constraint as $\text{tr}(\mathbf{F}_{k,p} \mathbf{F}_{k,p}^H) \leq p_k$, where p_k is the maximum transmit power for UE k . To derive sufficient statistics for \mathbf{h}_{mk} , AP m projects \mathbf{Y}_{mk}^p onto Φ_k^* as $\mathbf{Y}_{mk}^p = \mathbf{Y}_m^p \Phi_k^* =$

$\sum_{l=1}^K \mathbf{H}_{ml} \mathbf{F}_{l,p} (\Phi_l^T \Phi_k^*) + \mathbf{N}_m^p \Phi_k^* = \sum_{l \in \mathcal{P}_k} \tau_p \mathbf{H}_{ml} \mathbf{F}_{l,p} + \mathbf{Q}_{mk}^p$, where $\mathbf{Q}_{mk}^p \triangleq \mathbf{N}_m^p \Phi_k^*$. Then, following the standard MMSE estimation steps in [35] and [36], AP m can compute the MMSE estimation of \mathbf{h}_{mk} as

$$\hat{\mathbf{h}}_{mk} = \text{vec}(\hat{\mathbf{H}}_{mk}) = \mathbf{R}_{mk} \tilde{\mathbf{F}}_{k,p}^H \Psi_{mk}^{-1} \mathbf{y}_{mk}^p, \quad (3)$$

where $\hat{\mathbf{H}}_{mk}$ is the MMSE estimation of \mathbf{H}_{mk} , $\tilde{\mathbf{F}}_{k,p} = \mathbf{F}_{k,p}^T \otimes \mathbf{I}_L$, $\mathbf{y}_{mk}^p \triangleq \text{vec}(\mathbf{Y}_{mk}^p) = \sum_{l \in \mathcal{P}_k} \tau_p \tilde{\mathbf{F}}_{l,p} \mathbf{h}_{ml} + \mathbf{q}_m^p$, $\mathbf{q}_m^p = \text{vec}(\mathbf{Q}_{mk}^p)$ and $\Psi_{mk} = \sum_{l \in \mathcal{P}_k} \tau_p \tilde{\mathbf{F}}_{l,p} \mathbf{R}_{ml} \tilde{\mathbf{F}}_{l,p}^H + \sigma^2 \mathbf{I}_{LN}$, respectively. Note that the estimate $\hat{\mathbf{h}}_{mk}$ and estimation error $\tilde{\mathbf{h}}_{mk} = \mathbf{h}_{mk} - \hat{\mathbf{h}}_{mk}$ are independent random vectors distributed as $\hat{\mathbf{h}}_{mk} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \hat{\mathbf{R}}_{mk})$ and $\tilde{\mathbf{h}}_{mk} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{C}_{mk})$, where $\hat{\mathbf{R}}_{mk} \triangleq \tau_p \mathbf{R}_{mk} \tilde{\mathbf{F}}_{k,p}^H \Psi_{mk}^{-1} \tilde{\mathbf{F}}_{k,p} \mathbf{R}_{mk}$ and $\mathbf{C}_{mk} \triangleq \mathbf{R}_{mk} - \hat{\mathbf{R}}_{mk}$. We can also form $\hat{\mathbf{R}}_{mk}$ and \mathbf{C}_{mk} in the block structure with the (n, i) -th submatrix being $\hat{\mathbf{R}}_{mk}^{ni} = \mathbb{E}\{\hat{\mathbf{h}}_{mk,n} \hat{\mathbf{h}}_{mk,i}^H\}$ and $\mathbf{C}_{mk}^{ni} = \mathbb{E}\{\tilde{\mathbf{h}}_{mk,n} \tilde{\mathbf{h}}_{mk,i}^H\}$, respectively.

C. Data Transmission

For the data transmission, all antennas of all UEs simultaneously transmit their data symbols to all APs. The received signal $\mathbf{y}_m \in \mathbb{C}^L$ at AP m is

$$\mathbf{y}_m = \sum_{k=1}^K \mathbf{H}_{mk} \mathbf{s}_k + \mathbf{n}_m, \quad (4)$$

where $\mathbf{n}_m \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2 \mathbf{I}_L)$ is the independent receiver noise. The transmitted signal from UE k $\mathbf{s}_k \in \mathbb{C}^N$ can be constructed as $\mathbf{s}_k = \mathbf{F}_{k,u} \mathbf{x}_k$, where $\mathbf{x}_k \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_N)$ is the data symbol for UE k and $\mathbf{F}_{k,u} \in \mathbb{C}^{N \times N}$ is the precoding matrix for the data transmission which should satisfy the power constraint of UE k as $\text{tr}(\mathbf{F}_{k,u} \mathbf{F}_{k,u}^H) \leq p_k$.

III. SPECTRAL EFFICIENCY ANALYSIS AND I-WMMSE PRECODING DESIGN

In this section, we investigate two promising signal processing schemes, called ‘‘fully centralized processing’’ and ‘‘LSFD processing’’, and analyze their corresponding SE performance and design respective iteratively WMMSE precoding schemes³.

A. Fully Centralized Processing

1) *Spectral Efficiency Analysis*: For the fully centralized processing scheme, all M APs send all the received pilot signals and data signals to the CPU. Indeed, both the channel estimation and data detection are implemented at the CPU. The collective channel $\mathbf{h}_k \in \mathbb{C}^{MLN}$ for UE k can be constructed as $\mathbf{h}_k = [\text{vec}(\mathbf{H}_{1k})^T, \dots, \text{vec}(\mathbf{H}_{Mk})^T]^T \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{R}_k)$ with $\mathbf{R}_k = \text{diag}(\mathbf{R}_{1k}, \dots, \mathbf{R}_{Mk}) \in \mathbb{C}^{MLN \times MLN}$ being the

³We only optimize the precoding matrices for the phase of data transmission $\mathbf{F}_{k,u}$. The optimization of $\mathbf{F}_{k,p}$ is left for future research. Although we do not design $\mathbf{F}_{k,p}$ in this paper, we try to keep the derived equations more generalized. So a scenario with arbitrary $\mathbf{F}_{k,p}$ instead of limiting $\mathbf{F}_{k,p}$ to a particular form is investigated. It is worth noting that all equations in this paper hold for any $\mathbf{F}_{k,p}$, so undoubtedly provide some important guidelines for the investigation of optimization design for $\mathbf{F}_{k,p}$ in the future work.

whole block-diagonal correlation matrix for UE k . Similar to (3), the CPU can derive the channel estimate for UE k as⁴ $\hat{\mathbf{h}}_k \triangleq [\hat{\mathbf{h}}_{1k}^T, \dots, \hat{\mathbf{h}}_{Mk}^T]^T \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \tau_p \mathbf{R}_k \bar{\mathbf{F}}_{k,p}^H \Psi_k^{-1} \bar{\mathbf{F}}_{k,p} \mathbf{R}_k)$ where $\bar{\mathbf{F}}_{k,p} = \text{diag}(\underbrace{\tilde{\mathbf{F}}_{k,p}, \dots, \tilde{\mathbf{F}}_{k,p}}_M)$ and $\Psi_k^{-1} = \text{diag}(\Psi_{1k}^{-1}, \dots, \Psi_{Mk}^{-1})$. The channel estimation error is $\tilde{\mathbf{h}}_k \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{C}_k)$ where $\mathbf{C}_k \triangleq \mathbf{R}_k - \tau_p \mathbf{R}_k \bar{\mathbf{F}}_{k,p}^H \Psi_k^{-1} \bar{\mathbf{F}}_{k,p} \mathbf{R}_k$. Moreover, the received data signal at the CPU can be denoted as

$$\underbrace{[\mathbf{y}_1^T, \dots, \mathbf{y}_M^T]^T}_{=\mathbf{y}} = \sum_{k=1}^K \underbrace{[\mathbf{H}_{1k}^T, \dots, \mathbf{H}_{Mk}^T]^T}_{=\mathbf{H}_k} \mathbf{F}_{k,u} \mathbf{x}_k + \underbrace{[\mathbf{n}_1^T, \dots, \mathbf{n}_M^T]^T}_{=\mathbf{n}}, \quad (5)$$

or a compact form as $\mathbf{y} = \sum_{k=1}^K \mathbf{H}_k \mathbf{F}_{k,u} \mathbf{x}_k + \mathbf{n}$.

Under the setting of “fully centralized processing”, we assume that UL precoding matrices ($\mathbf{F}_{k,u}$ and $\mathbf{F}_{k,p}$) are available at the CPU. Based on the collective channel estimates, the CPU designs an arbitrary receive combining matrix $\mathbf{V}_k \in \mathbb{C}^{LM \times N}$ for UE k to detect \mathbf{x}_k as

$$\check{\mathbf{x}}_k = \mathbf{V}_k^H \mathbf{y} = \mathbf{V}_k^H \hat{\mathbf{H}}_k \mathbf{F}_{k,u} \mathbf{x}_k + \mathbf{V}_k^H \tilde{\mathbf{H}}_k \mathbf{F}_{k,u} \mathbf{x}_k + \sum_{l \neq k}^K \mathbf{V}_k^H \mathbf{H}_l \mathbf{F}_{l,u} \mathbf{x}_l + \mathbf{V}_k^H \mathbf{n}, \quad (6)$$

and the conditional MSE matrix for UE k is

$$\begin{aligned} \mathbf{E}_{k,(1)} &= \mathbb{E}\{(\mathbf{x}_k - \check{\mathbf{x}}_k)(\mathbf{x}_k - \check{\mathbf{x}}_k)^H | \{\hat{\mathbf{H}}_k\}, \{\mathbf{F}_{k,u}\}\} \\ &= \mathbf{I}_N - \mathbf{V}_k^H \hat{\mathbf{H}}_k \mathbf{F}_{k,u} - \mathbf{F}_{k,u}^H \hat{\mathbf{H}}_k^H \mathbf{V}_k + \mathbf{V}_k^H \left(\sum_{l=1}^K \left(\hat{\mathbf{H}}_l \bar{\mathbf{F}}_{l,u} \hat{\mathbf{H}}_l^H + \mathbf{C}'_l \right) + \sigma^2 \mathbf{I}_{ML} \right) \mathbf{V}_k \end{aligned} \quad (7)$$

where $\bar{\mathbf{F}}_{l,u} \triangleq \mathbf{F}_{l,u} \mathbf{F}_{l,u}^H$, $\mathbf{C}'_l \triangleq \text{diag}(\mathbf{C}'_{1l}, \dots, \mathbf{C}'_{Ml}) \in \mathbb{C}^{ML \times ML}$ and $\mathbf{C}'_{ml} = \mathbb{E}\{\tilde{\mathbf{H}}_{ml} \bar{\mathbf{F}}_{l,u} \tilde{\mathbf{H}}_{ml}^H\} \in \mathbb{C}^{L \times L}$ with the (j, q) -th element of \mathbf{C}'_{ml} being $[\mathbf{C}'_{ml}]_{jq} = \sum_{p_1=1}^N \sum_{p_2=1}^N [\bar{\mathbf{F}}_l]_{p_2 p_1} [\mathbf{C}'_{ml}]_{j p_1 q}$.

By implementing the per-user-basis minimum mean-squared error-based successive interference cancellation (MMSE-SIC) detector while treating co-user interference as uncorrelated Gaussian noise, we derive the achievable SE for UE k as follows.

Corollary 1. *An achievable for UE k under the setting of “fully centralized processing” with the MMSE estimator is*

$$\text{SE}_{k,(1)} = \left(1 - \frac{\tau_p}{\tau_c} \right) \mathbb{E} \left\{ \log_2 \left| \mathbf{I}_N + \mathbf{D}_{k,(1)}^H \Sigma_{k,(1)}^{-1} \mathbf{D}_{k,(1)} \right| \right\}, \quad (8)$$

where $\mathbf{D}_{k,(1)} \triangleq \mathbf{V}_k^H \hat{\mathbf{H}}_k \mathbf{F}_{k,u}$ and $\Sigma_{k,(1)} \triangleq \mathbf{V}_k^H \left(\sum_{l=1}^K \hat{\mathbf{H}}_l \bar{\mathbf{F}}_{l,u} \hat{\mathbf{H}}_l^H - \hat{\mathbf{H}}_k \bar{\mathbf{F}}_{k,u} \hat{\mathbf{H}}_k^H + \sum_{l=1}^K \mathbf{C}'_l + \sigma^2 \mathbf{I}_{ML} \right) \mathbf{V}_k$.

⁴Note that the pilot signals received at the APs are first transmitted to the CPU and then the CPU estimates the channels, where $\tau_p ML$ complex scalars are sent from the APs to the CPU at each coherence block. Alternatively, all APs can first estimate the channels as (3), and then send their channel estimates to the CPU, where MKN complex scalars are sent from the APs to the CPU at each coherence block. Since the pilot contamination is investigated ($\tau_p < KN$) in this paper, we consider the first transmission protocol due to its lower fronthaul overhead.

The expectations are with respect to all sources of randomness.

Proof. The proof follows from the similar approach as [28, Corollary 1] and is therefore omitted. \square

We notice that Corollary 1 holds for any combining schemes. One promising combining scheme is the MMSE combining as

$$\mathbf{V}_k^{\text{MMSE}} = \left(\sum_{l=1}^K \left(\hat{\mathbf{H}}_l \bar{\mathbf{F}}_{l,u} \hat{\mathbf{H}}_l^H + \mathbf{C}'_l \right) + \sigma^2 \mathbf{I}_{ML} \right)^{-1} \hat{\mathbf{H}}_k \mathbf{F}_{k,u}, \quad (9)$$

which can minimize the mean-squared error $\text{MSE}_{k,(1)} = \text{tr}(\mathbf{E}_{k,(1)})$. With the MMSE combining scheme, the conditional MSE matrix in (7) is

$$\mathbf{E}_{k,(1)}^{\text{opt}} = \mathbf{I}_N - \mathbf{F}_{k,u}^H \hat{\mathbf{H}}_k^H \left(\sum_{l=1}^K \left(\hat{\mathbf{H}}_l \bar{\mathbf{F}}_{l,u} \hat{\mathbf{H}}_l^H + \mathbf{C}'_l \right) + \sigma^2 \mathbf{I}_{ML} \right)^{-1} \hat{\mathbf{H}}_k \mathbf{F}_{k,u} \quad (10)$$

More importantly, the MMSE combining in (9) can also maximize the achievable SE in (8) as follows.

Corollary 2. *The achievable SE for UE k in (8) can be maximized by the MMSE combining scheme in (9) with the maximum value*

$$\text{SE}_{k,(1)}^{\text{opt}} = \left(1 - \frac{\tau_p}{\tau_c} \right) \mathbb{E} \left\{ \log_2 \left| \mathbf{I}_N + \mathbf{F}_{k,u}^H \hat{\mathbf{H}}_k^H \left(\sum_{l=1}^K \left(\hat{\mathbf{H}}_l \bar{\mathbf{F}}_{l,u} \hat{\mathbf{H}}_l^H + \mathbf{C}'_l \right) - \hat{\mathbf{H}}_k \bar{\mathbf{F}}_{k,u} \hat{\mathbf{H}}_k^H + \sigma^2 \mathbf{I}_{ML} \right)^{-1} \hat{\mathbf{H}}_k \mathbf{F}_{k,u} \right| \right\}. \quad (11)$$

Proof. The proof can be found in [28, Appendix B] and is therefore omitted. \square

2) *Iteratively WMMSE Precoding Design:* In this part, we design the uplink precoding scheme for the “fully centralized processing”. One popular weighted sum-rate maximization problem is investigated as⁵

$$\begin{aligned} & \max_{\{\mathbf{F}\}} \sum_{k=1}^K \mu_{k,(1)} \text{SE}_{k,(1)} \\ & \text{s.t. } \|\mathbf{F}_{k,u,(1)}\|^2 \leq p_k \quad \forall k = 1, \dots, K \end{aligned} \quad (12)$$

where $\mu_{k,(1)}$ represents the priority weight of UE k and $\text{SE}_{k,(1)}$ is given by (8).

⁵The notation \mathbf{F} is short for $\{\mathbf{F}_{k,u}\}_{k=1,\dots,K}$, denoting all variables $\mathbf{F}_{k,u}$ with $k = 1, \dots, K$. Similar definitions are applied for \mathbf{V} , \mathbf{A} , \mathbf{W} , \mathbf{S} in the following. In this section, we denote by $\mathbf{F}_{k,u,(1)}$ and $\mathbf{F}_{k,u,(2)}$ the UL precoding matrix of UE k for the fully centralized processing and LSF scheme, respectively.

As [30] and [32], the matrix-weighted sum-MSE minimization problem as

$$\begin{aligned} \min_{\{\mathbf{F}, \mathbf{V}, \mathbf{W}\}} \sum_{k=1}^K \mu_{k,(1)} \left[\text{tr}(\mathbf{W}_{k,(1)} \mathbf{E}_{k,(1)}) - \log_2 |\mathbf{W}_{k,(1)}| \right] \\ \text{s.t. } \|\mathbf{F}_{k,u,(1)}\|^2 \leq p_k \quad \forall k = 1, \dots, K \end{aligned} \quad (13)$$

is equivalent to the weighted sum-rate maximization problem (12), where $\mathbf{W}_{k,(1)} \in \mathbb{C}^{N \times N}$ is the weight matrix for UE k . We notice that (13) is convex over each optimization variable \mathbf{F} , \mathbf{V} , \mathbf{W} but is not jointly convex over all optimization variables. Following the method in [30], we can solve (13) by sequentially fixing two of the three optimization variables \mathbf{F} , \mathbf{V} , \mathbf{W} and updating the third.

Fixing the other variables, the update of \mathbf{V}_k is given by the MMSE solution as (9). Under the MMSE combining, the MSE matrix is given by (10). Then, note that optimal $\mathbf{W}_{k,(1)}$ for (13) is

$$\mathbf{W}_{k,(1)}^{\text{opt}} = \mathbf{E}_{k,(1)}^{-1}, \quad (14)$$

which can be easily derived through the first order optimality condition for $\mathbf{W}_{k,(1)}$ by fixing \mathbf{F} and \mathbf{V} .

Remark 1. When the MMSE combining $\mathbf{V}_k^{\text{MMSE}}$ and $\mathbf{W}_{k,(1)}^{\text{opt}}$ for all UEs are implemented in (13), we have $\text{tr}(\mathbf{W}_{k,(1)} \mathbf{E}_{k,(1)}) - \log_2 |\mathbf{W}_{k,(1)}| = \text{tr}(\mathbf{I}_N) - \log_2 |(\mathbf{E}_{k,(1)}^{\text{opt}})^{-1}|$. So the matrix-weighted sum-MSE minimization problem in (13) would reduce to the equivalent optimization problem of (12) as⁶:

$$\begin{aligned} \max_{\{\mathbf{F}\}} \sum_{k=1}^K \mu_{k,(1)} \log_2 \left| \left(\mathbf{E}_{k,(1)}^{\text{opt}} \right)^{-1} \right| \\ \text{s.t. } \|\mathbf{F}_{k,u,(1)}\|^2 \leq p_k \quad \forall k = 1, \dots, K \end{aligned} \quad (15)$$

which is a well-known relationship between $\mathbf{E}_{k,(1)}^{\text{opt}}$ and $\text{SE}_{k,(1)}^{\text{opt}}$.

Finally, fixing \mathbf{V} and \mathbf{W} , the update of $\mathbf{F}_{k,u,(1)}$ for (13) results in the optimization problem as⁷

⁶Note that ‘‘SE’’ is equivalent to ‘‘rate’’ except from having one scaling factor $(\tau_c - \tau_p)/\tau_c$. Since τ_c and τ_p are constants in this paper, so we ignore the difference between SE and rate in the optimization problem.

⁷It is worth mentioning that the updates of optimization variables are based on the preliminary of fixing the other optimization variables. For instance, when updating $\mathbf{F}_{k,u,(1)}$, we should fix the other optimization variables (\mathbf{V}_k and $\mathbf{W}_{k,(1)}$) but not only limited to their respective optimal solutions $\mathbf{V}_k^{\text{MMSE}}$ and $\mathbf{W}_{k,(1)}^{\text{opt}}$. So we update $\mathbf{F}_{k,u,(1)}$ based on (16) with generalized \mathbf{V}_k and $\mathbf{W}_{k,(1)}$ instead of (15) with optimal $\mathbf{V}_k^{\text{MMSE}}$ and $\mathbf{W}_{k,(1)}^{\text{opt}}$.

$$\begin{aligned}
\min_{\{\mathbf{F}\}} & \sum_{k=1}^K \mu_{k,(1)} \text{tr} \left[\mathbf{W}_{k,(1)} \left(\mathbf{I}_N - \mathbf{V}_k^H \hat{\mathbf{H}}_k \mathbf{F}_{k,u,(1)} \right) \left(\mathbf{I}_N - \mathbf{V}_k^H \hat{\mathbf{H}}_k \mathbf{F}_{k,u,(1)} \right)^H \right] \\
& + \sum_{k=1}^K \mu_{k,(1)} \text{tr} \left[\mathbf{W}_{k,(1)} \mathbf{V}_k^H \left(\sum_{l \neq k}^K \hat{\mathbf{H}}_l \mathbf{F}_{l,u,(1)} \mathbf{F}_{l,u,(1)}^H \hat{\mathbf{H}}_l^H \right) \mathbf{V}_k \right] - \sum_{k=1}^K \mu_{k,(1)} \log_2 |\mathbf{W}_{k,(1)}| \\
& + \sum_{k=1}^K \mu_{k,(1)} \text{tr} \left[\mathbf{W}_{k,(1)} \mathbf{V}_k^H \left(\sum_{l=1}^K \mathbb{E} \left\{ \tilde{\mathbf{H}}_l \mathbf{F}_{l,u,(1)} \mathbf{F}_{l,u,(1)}^H \tilde{\mathbf{H}}_l^H \mid \mathbf{F} \right\} + \sigma^2 \mathbf{I}_{ML} \right) \mathbf{V}_k \right] \\
\text{s.t.} & \left\| \mathbf{F}_{k,u,(1)} \right\|^2 \leq p_k \quad \forall k = 1, \dots, K
\end{aligned} \tag{16}$$

which is a convex quadratic optimization problem. So the classic Lagrange multipliers methods and Karush-Kuhn-Tucker (KKT) conditions can be applied to derive an optimal solution. The Lagrange function of (16) is

$$\begin{aligned}
f(\mathbf{F}_{1,u,(1)}, \dots, \mathbf{F}_{K,u,(1)}) & = \sum_{k=1}^K \mu_{k,(1)} \text{tr} \left[\mathbf{W}_{k,(1)} \left(\mathbf{I}_N - \mathbf{V}_k^H \hat{\mathbf{H}}_k \mathbf{F}_{k,u,(1)} \right) \left(\mathbf{I}_N - \mathbf{V}_k^H \hat{\mathbf{H}}_k \mathbf{F}_{k,u,(1)} \right)^H \right] \\
& + \sum_{k=1}^K \mu_{k,(1)} \text{tr} \left[\mathbf{W}_{k,(1)} \mathbf{V}_k^H \left(\sum_{l \neq k}^K \hat{\mathbf{H}}_l \mathbf{F}_{l,u,(1)} \mathbf{F}_{l,u,(1)}^H \hat{\mathbf{H}}_l^H \right) \mathbf{V}_k \right] \\
& + \sum_{k=1}^K \mu_{k,(1)} \text{tr} \left[\mathbf{W}_{k,(1)} \mathbf{V}_k^H \left(\sum_{l=1}^K \mathbb{E} \left\{ \tilde{\mathbf{H}}_l \mathbf{F}_{l,u,(1)} \mathbf{F}_{l,u,(1)}^H \tilde{\mathbf{H}}_l^H \mid \mathbf{F} \right\} + \sigma^2 \mathbf{I}_{ML} \right) \mathbf{V}_k \right] \\
& + \sum_{k=1}^K \lambda_{k,(1)} \left[\text{tr}(\mathbf{F}_{k,u,(1)} \mathbf{F}_{k,u,(1)}^H) - p_k \right]
\end{aligned} \tag{17}$$

Finally, we derive the optimal precoding scheme as the following theorem.

Theorem 1. *By fixing other optimization variables and applying the first-order optimality condition of (17) with respect to each $\mathbf{F}_{k,u,(1)}$, the optimal precoding scheme is given by*

$$\begin{aligned}
\mathbf{F}_{k,u,(1)}^{\text{opt}} & = \mu_{k,(1)} \left[\sum_{l=1}^K \mu_{l,(1)} \left(\hat{\mathbf{H}}_k^H \mathbf{V}_l \mathbf{W}_{l,(1)} \mathbf{V}_l^H \hat{\mathbf{H}}_k + \mathbb{E} \left\{ \tilde{\mathbf{H}}_k^H \mathbf{V}_l \mathbf{W}_{l,(1)} \mathbf{V}_l^H \tilde{\mathbf{H}}_k \mid \mathbf{V}, \mathbf{W} \right\} \right) + \lambda_{k,(1)} \mathbf{I}_N \right]^{-1} \hat{\mathbf{H}}_k^H \mathbf{V}_k \mathbf{W}_{k,(1)} \\
& = \mu_{k,(1)} \left[\sum_{l=1}^K \mu_{l,(1)} \left(\hat{\mathbf{H}}_k^H \mathbf{V}_l \mathbf{W}_{l,(1)} \mathbf{V}_l^H \hat{\mathbf{H}}_k + \bar{\mathbf{C}}_{kl} \right) + \lambda_{k,(1)} \mathbf{I}_N \right]^{-1} \hat{\mathbf{H}}_k^H \mathbf{V}_k \mathbf{W}_{k,(1)},
\end{aligned} \tag{18}$$

where $\lambda_{k,(1)} \geq 0$ is the Lagrangian multiplier and the (i, n) -th element of $\bar{\mathbf{C}}_{kl} \triangleq \mathbb{E} \left\{ \tilde{\mathbf{H}}_k^H \mathbf{V}_l \mathbf{W}_{l,(1)} \mathbf{V}_l^H \tilde{\mathbf{H}}_k \mid \mathbf{V}, \mathbf{W} \right\} \in \mathbb{C}^{N \times N}$ is $[\bar{\mathbf{C}}_{kl}]_{in} = \text{tr}(\bar{\mathbf{V}}_l \mathbb{E} \{ \tilde{\mathbf{h}}_{k,n} \tilde{\mathbf{h}}_{k,i}^H \}) = \text{tr}(\bar{\mathbf{V}}_l \mathbf{C}_{k,ni})$ with $\bar{\mathbf{V}}_l \triangleq \mathbf{V}_l \mathbf{W}_{l,(1)} \mathbf{V}_l^H$ and $\mathbf{C}_{k,ni} \triangleq \mathbb{E} \{ \tilde{\mathbf{h}}_{k,n} \tilde{\mathbf{h}}_{k,i}^H \} =$

$\text{diag}(\mathbf{C}_{1k}^{ni}, \dots, \mathbf{C}_{Mk}^{ni}) \in \mathbb{C}^{ML \times ML}$. According to the KKT condition, $\lambda_{k,(1)}$ and $\mathbf{F}_{k,u,(1)}$ should also satisfy

$$\|\mathbf{F}_{k,u,(1)}\|^2 \leq p_k, \quad \lambda_{k,(1)} \left(\|\mathbf{F}_{k,u,(1)}\|^2 - p_k \right) = 0, \quad \lambda_{k,(1)} \geq 0. \quad (19)$$

Proof: The proof is given in Appendix D. ■

We denote by $\mathbf{F}_{k,u,(1)}(\lambda_{k,(1)})$ the right-hand side of (18), when $\sum_{l=1}^K \mu_{l,(1)} (\hat{\mathbf{H}}_k^H \mathbf{V}_l \mathbf{W}_{l,(1)} \mathbf{V}_l^H \hat{\mathbf{H}}_k + \bar{\mathbf{C}}_{kl})$ is invertible and $\text{tr}[\mathbf{F}_{k,u,(1)}(0) \mathbf{F}_{k,u,(1)}(0)^H] \leq p_k$, then $\mathbf{F}_{k,u,(1)}^{\text{opt}} = \mathbf{F}_{k,u,(1)}(0)$, otherwise we have

$$\text{tr}[\mathbf{F}_{k,u,(1)}(\lambda_{k,(1)}) \mathbf{F}_{k,u,(1)}(\lambda_{k,(1)})^H] = p_k$$

to satisfy (19).

Corollary 3. $\text{tr}[\mathbf{F}_{k,u,(1)}(\lambda_{k,(1)}) \mathbf{F}_{k,u,(1)}(\lambda_{k,(1)})^H]$ is a monotonically decreasing function of $\lambda_{k,(1)}$.

Proof: Let $\mathbf{D}\mathbf{\Lambda}\mathbf{D}^H$ denote the eigendecomposition of $\sum_{l=1}^K \mu_{l,(1)} (\hat{\mathbf{H}}_k^H \mathbf{V}_l \mathbf{W}_{l,(1)} \mathbf{V}_l^H \hat{\mathbf{H}}_k + \bar{\mathbf{C}}_{kl})$. Following the method in [30], we define $\mathbf{\Phi} = \mu_{k,(1)}^2 \mathbf{D}^H \hat{\mathbf{H}}_k^H \mathbf{V}_k \mathbf{W}_{k,(1)}^2 \hat{\mathbf{H}}_k \mathbf{V}_k^H \mathbf{D}$ and we have

$$\begin{aligned} \text{tr}[\mathbf{F}_{k,u,(1)}(\lambda_{k,(1)}) \mathbf{F}_{k,u,(1)}(\lambda_{k,(1)})^H] &= \text{tr} \left\{ (\mathbf{D}\mathbf{\Lambda}\mathbf{D}^H + \lambda_{k,(1)} \mathbf{I}_N)^{-1} \mathbf{D}\mathbf{\Phi}\mathbf{D}^H (\mathbf{D}\mathbf{\Lambda}\mathbf{D}^H + \lambda_{k,(1)} \mathbf{I}_N)^{-1} \right\} \\ &= \text{tr} \left\{ (\mathbf{D}\mathbf{\Lambda}\mathbf{D}^H + \lambda_{k,(1)} \mathbf{I}_N)^{-2} \mathbf{D}\mathbf{\Phi}\mathbf{D}^H \right\} = \text{tr} \left\{ (\mathbf{\Lambda} + \lambda_{k,(1)} \mathbf{I}_N)^{-2} \right\} = \sum_{n=1}^N \frac{[\mathbf{\Phi}]_{nn}}{([\mathbf{\Lambda}]_{nn} + \lambda_{k,(1)})^2}, \end{aligned} \quad (20)$$

so $\text{tr}[\mathbf{F}_{k,u,(1)}(\lambda_{k,(1)}) \mathbf{F}_{k,u,(1)}(\lambda_{k,(1)})^H]$ is a monotonically decreasing function of $\lambda_{k,(1)}$. ■

Based on Corollary 3, optimum $\lambda_{k,(1)}$ (denoted by $\lambda_{k,(1)}^{\text{opt}}$) can be easily obtained by a one-dimensional (1-D) bisection algorithm so we derive the solution for $\mathbf{F}_{k,u,(1)}(\lambda_{k,(1)}^{\text{opt}})$. Furthermore, an iterative optimization algorithm for $\mathbf{F}_{k,u,(1)}$, called ‘‘iteratively WMMSE (I-WMMSE) algorithm’’, is summarized in Algorithm 1⁸. The convergence of Algorithm 1 is proven in [30, Theorem 3].

Remark 2. Note that the design of $\mathbf{F}_{k,p}$ is a valuable future direction to further improve the system performance. One valuable optimization problem is to minimize the total MSE of the channel estimators of all UEs as

$$\begin{aligned} \min_{\{\mathbf{F}_{k,p}\}} & \sum_{k=1}^K \text{tr}(\mathbf{C}_k) \\ \text{s.t.} & \|\mathbf{F}_{k,p}\|^2 \leq p_k \quad \forall k = 1, \dots, K \end{aligned} \quad (21)$$

where the optimization goal is only based on the statistical knowledge so $\mathbf{F}_{k,p}$ is also based on the statistical knowledge.

⁸To balance the efficiency and the computational complexity of the proposed algorithm, we also include the stopping criterion ‘‘ $R_{(1)}^{(i)} < R_{(1)}^{(i-1)}$ ’’. Moreover, the I-WMMSE precoding scheme is derived at iteration $(i-1)$, which may achieve higher sum SE than the one at iteration i .

Algorithm 1: I-WMMSE Algorithm for the Design of $\mathbf{F}_{k,u,(1)}$

- Input:** Collective channel estimates $\hat{\mathbf{H}}_k$ for all UEs; Estimation error covariance matrices \mathbf{C}_{ml} for all possible pairs; UE weights $\mu_{k,(1)}$ for all UEs;
- Output:** Optimal precoding matrices $\mathbf{F}_{k,u,(1)}$ for all UEs ($\mathbf{F}_{k,u,(1)}^{(i)}$ for the first or third stopping criterion and $\mathbf{F}_{k,u,(1)}^{(i-1)}$ for the second stopping criterion);
- 1 **Initiation:** $i = 0$, $\mathbf{F}_{k,u,(1)}^{(0)}$ and $R_{(1)}^{(0)} = \sum_{k=1}^K \mu_{k,(1)} \text{SE}_{k,(1)}^{(0)}$ for all UEs; maximum iteration number $I_{(1),\max}$ and threshold $\varepsilon_{(1)}$;
 - 2 **repeat**
 - 3 $i = i + 1$
 - 4 Update the MMSE combining scheme $\mathbf{V}_k^{(i)}$ with $\mathbf{F}_{k,u,(1)}^{(i-1)}$ based on (9);
 - 5 Update optimal MSE matrix $\mathbf{E}_{k,(1)}^{(i)}$ with $\mathbf{F}_{l,u,(1)}^{(i-1)}$ based on (10), and update $\mathbf{W}_{k,(1)}^{(i)}$ based on (14);
 - 6 Update optimal precoding matrix $\mathbf{F}_{l,u,(1)}^{(i)}$ with $\mathbf{V}_k^{(i)}$ and $\mathbf{W}_{k,(1)}^{(i)}$ based on (18), where $\lambda_{k,(1)}^{(i)}$ is found by a bisection algorithm;
 - 7 Update sum weighted rate $R_{(1)}^{(i)} = \sum_{k=1}^K \mu_{k,(1)} \text{SE}_{k,(1)}^{(i)}$;
 - 8 **until** $\left| R_{(1)}^{(i)} - R_{(1)}^{(i-1)} \right| / R_{(1)}^{(i-1)} \leq \varepsilon_{(1)}$ or $R_{(1)}^{(i)} < R_{(1)}^{(i-1)}$ or $i \geq I_{(1),\max}$;
-

B. Large-Scale Fading Decoding

1) *Spectral Efficiency Analysis:* Another promising processing scheme is “large-scale fading decoding”, which is a two-layer decoding scheme to decode the data symbol. Note that UL precoding matrices ($\mathbf{F}_{k,u}$ and $\mathbf{F}_{k,p}$) are assumed to be available at all APs and the CPU. In the first layer, AP m applies an arbitrary combining matrix $\mathbf{V}_{mk} \in \mathbb{C}^{L \times N}$ to derive local detection of \mathbf{x}_k as

$$\tilde{\mathbf{x}}_{mk} = \mathbf{V}_{mk}^H \mathbf{y}_m = \mathbf{V}_{mk}^H \mathbf{H}_{mk} \mathbf{F}_{k,u} \mathbf{x}_k + \sum_{l=1, l \neq k}^K \mathbf{V}_{mk}^H \mathbf{H}_{ml} \mathbf{F}_{l,u} \mathbf{x}_l + \mathbf{V}_{mk}^H \mathbf{n}_m. \quad (22)$$

We notice that \mathbf{V}_{mk} is designed based on local channel estimates at AP m and one handy choice is MR combining $\mathbf{V}_{mk} = \hat{\mathbf{H}}_{mk}$. Moreover, local MMSE (L-MMSE) combining

$$\mathbf{V}_{mk} = \left(\sum_{l=1}^K \left(\hat{\mathbf{H}}_{ml} \bar{\mathbf{F}}_{l,u} \hat{\mathbf{H}}_{ml}^H + \mathbf{C}'_{ml} \right) + \sigma^2 \mathbf{I}_L \right)^{-1} \hat{\mathbf{H}}_{mk} \mathbf{F}_{k,u}, \quad (23)$$

is also regarded as a promising scheme, since (23) can minimize $\mathbb{E}\{\|\mathbf{x}_k - \mathbf{V}_{mk}^H \mathbf{y}_m\|^2 \mid \{\hat{\mathbf{H}}_{mk}\}, \{\mathbf{F}_{k,u}\}\}$.

In the second layer, the “LSFD” method is implemented at the CPU [5]. The CPU weights all the local estimates $\tilde{\mathbf{x}}_{mk}$ from all APs by the LSFD coefficient matrix as

$$\hat{\mathbf{x}}_k = \sum_{m=1}^M \mathbf{A}_{mk}^H \tilde{\mathbf{x}}_{mk} = \sum_{m=1}^M \mathbf{A}_{mk}^H \mathbf{V}_{mk}^H \mathbf{H}_{mk} \mathbf{F}_{k,u} \mathbf{x}_k + \sum_{m=1}^M \sum_{l=1, l \neq k}^K \mathbf{A}_{mk}^H \mathbf{V}_{mk}^H \mathbf{H}_{ml} \mathbf{F}_{l,u} \mathbf{x}_l + \mathbf{n}'_k, \quad (24)$$

where $\mathbf{A}_{mk} \in \mathbb{C}^{N \times N}$ is the complex LSFD coefficient matrix for AP m -UE k and $\mathbf{n}'_k = \sum_{m=1}^M \mathbf{A}_{mk}^H \mathbf{V}_{mk}^H \mathbf{n}_m$.

Moreover, we can rewrite $\hat{\mathbf{x}}_k$ in a more compact form as

$$\hat{\mathbf{x}}_k = \mathbf{A}_k^H \mathbf{G}_{kk} \mathbf{F}_{k,u} \mathbf{x}_k + \underbrace{\sum_{l=1, l \neq k}^K \mathbf{A}_k^H \mathbf{G}_{kl} \mathbf{F}_{l,u} \mathbf{x}_l + \mathbf{n}'_k}_{\tilde{\mathbf{x}}_k} = \mathbf{A}_k^H \left(\mathbf{G}_{kk} \mathbf{F}_{k,u} \mathbf{x}_k + \sum_{l=1, l \neq k}^K \mathbf{G}_{kl} \mathbf{F}_{l,u} \mathbf{x}_l + \tilde{\mathbf{n}}'_k \right) \quad (25)$$

where $\mathbf{A}_k \triangleq [\mathbf{A}_{1k}^T, \dots, \mathbf{A}_{Mk}^T]^T \in \mathbb{C}^{MN \times N}$, $\mathbf{G}_{kl} \triangleq [\mathbf{V}_{1k}^H \mathbf{H}_{1l}; \dots; \mathbf{V}_{Mk}^H \mathbf{H}_{Ml}] \in \mathbb{C}^{MN \times N}$ and

$$\tilde{\mathbf{n}}'_k = [\mathbf{V}_{1k}^H \mathbf{n}_1; \dots; \mathbf{V}_{Mk}^H \mathbf{n}_M] \in \mathbb{C}^{MN \times N}.$$

Note that the CPU does not have the knowledge of channel estimates and is only aware of channel statistics [5]. The conditional MSE matrix for UE k $\mathbf{E}_{k,(2)} \triangleq \mathbb{E} \{ (\mathbf{x}_k - \hat{\mathbf{x}}_k) (\mathbf{x}_k - \hat{\mathbf{x}}_k)^H | \Theta \}$ is

$$\mathbf{E}_{k,(2)} = \mathbf{I}_N - \mathbf{F}_{k,u}^H \mathbb{E} \{ \mathbf{G}_{kk}^H \} \mathbf{A}_k - \mathbf{A}_k^H \mathbb{E} \{ \mathbf{G}_{kk} \} \mathbf{F}_{k,u} + \mathbf{A}_k^H \left(\sum_{l=1}^K \mathbb{E} \{ \mathbf{G}_{kl} \bar{\mathbf{F}}_{l,u} \mathbf{G}_{kl}^H \} + \sigma^2 \mathbf{S}_k \right) \mathbf{A}_k, \quad (26)$$

where Θ denotes all the channel statistics and $\mathbf{S}_k = \text{diag}(\mathbb{E} \{ \mathbf{V}_{1k}^H \mathbf{V}_{1k} \}, \dots, \mathbb{E} \{ \mathbf{V}_{Mk}^H \mathbf{V}_{Mk} \}) \in \mathbb{C}^{MN \times MN}$.

Then, we apply classical use-and-then-forget bound to obtain the following ergodic achievable SE.

Corollary 4. *For the “LSFD” scheme, an achievable SE for UE k can be written as*

$$\text{SE}_{k,(2)} = \left(1 - \frac{\tau_p}{\tau_c} \right) \log_2 \left| \mathbf{I}_N + \mathbf{D}_{k,(2)}^H \Sigma_{k,(2)}^{-1} \mathbf{D}_{k,(2)} \right|, \quad (27)$$

where $\Sigma_{k,(2)} = \sum_{l=1}^K \mathbf{A}_k^H \mathbb{E} \{ \mathbf{G}_{kl} \bar{\mathbf{F}}_{l,u} \mathbf{G}_{kl}^H \} \mathbf{A}_k - \mathbf{D}_{k,(2)} \mathbf{D}_{k,(2)}^H + \sigma^2 \mathbf{A}_k^H \mathbf{S}_k \mathbf{A}_k$ and $\mathbf{D}_{k,(2)} = \mathbf{A}_k^H \mathbb{E} \{ \mathbf{G}_{kk} \} \mathbf{F}_{k,u}$.

Proof. The proof follows similar steps as the proof of [28, Corollary 2] and is therefore omitted. \square

Note that \mathbf{A}_k can be optimized by the CPU based on channel statistics to maximize the achievable SE in (27). Based on the theory of optimal receivers as in [37], we derive the optimal LSFD coefficient matrix, which not only maximizes the achievable SE but minimizes the conditional MSE, as follows.

Corollary 5. *The achievable SE in (27) is maximized by*

$$\mathbf{A}_k^{\text{opt}} = \left(\sum_{l=1}^K \mathbb{E} \{ \mathbf{G}_{kl} \bar{\mathbf{F}}_{l,u} \mathbf{G}_{kl}^H \} + \sigma^2 \mathbf{S}_k \right)^{-1} \mathbb{E} \{ \mathbf{G}_{kk} \} \mathbf{F}_{k,u}, \quad (28)$$

leading to the maximum value as

$$\begin{aligned} & \text{SE}_{k,(2)}^{\text{opt}} \\ &= \left(1 - \frac{\tau_p}{\tau_c} \right) \log_2 \left| \mathbf{I}_N + \mathbf{F}_{k,u}^H \mathbb{E} \{ \mathbf{G}_{kk} \} \left(\sum_{l=1}^K \mathbb{E} \{ \mathbf{G}_{kl} \bar{\mathbf{F}}_{l,u} \mathbf{G}_{kl}^H \} - \mathbb{E} \{ \mathbf{G}_{kk} \} \bar{\mathbf{F}}_{k,u} \mathbb{E} \{ \mathbf{G}_{kk}^H \} + \sigma^2 \mathbf{S}_k \right)^{-1} \mathbb{E} \{ \mathbf{G}_{kk} \} \mathbf{F}_{k,u} \right|. \end{aligned} \quad (29)$$

Note that the optimal LSF D coefficient matrix in (28) can also minimize the conditional MSE for UE k $\text{MSE}_{k,(2)} = \text{tr}(\mathbf{E}_{k,(2)})$.

Proof. The proof is given in Appendix B. \square

If the optimal LSF D coefficient matrix is applied, the MSE matrix for UE k can be written as

$$\mathbf{E}_{k,(2)}^{\text{opt}} = \mathbf{I}_N - \mathbf{F}_{k,u}^H \mathbb{E} \{ \mathbf{G}_{kk}^H \} \left(\sum_{l=1}^K \mathbb{E} \{ \mathbf{G}_{kl} \bar{\mathbf{F}}_{l,u} \mathbf{G}_{kl}^H \} + \sigma^2 \mathbf{S}_k \right)^{-1} \mathbb{E} \{ \mathbf{G}_{kk} \} \mathbf{F}_{k,u}. \quad (31)$$

Furthermore, if MR combining $\mathbf{V}_{mk} = \hat{\mathbf{H}}_{mk}$ is applied, we derive closed-form SE expressions as follows.

Theorem 2. For MR combining $\mathbf{V}_{mk} = \hat{\mathbf{H}}_{mk}$, (27) can be computed in closed-form as

$$\text{SE}_{k,(2),c} = \left(1 - \frac{\tau_p}{\tau_c} \right) \log_2 \left| \mathbf{I}_N + \mathbf{D}_{k,(2),c}^H \boldsymbol{\Sigma}_{k,(2),c}^{-1} \mathbf{D}_{k,(2),c} \right|, \quad (32)$$

where $\boldsymbol{\Sigma}_{k,(2),c} = \mathbf{A}_k^H (\sum_{l=1}^K \mathbf{T}_{kl,(1)} + \sum_{l \in \mathcal{P}_k} \mathbf{T}_{kl,(2)}) \mathbf{A}_k - \mathbf{D}_{k,(2),c} \mathbf{D}_{k,(2),c}^H + \sigma^2 \mathbf{A}_k^H \mathbf{S}_{k,c} \mathbf{A}_k$ and $\mathbf{D}_{k,(2),c} = \mathbf{A}_k^H \mathbf{Z}_k \mathbf{F}_{k,u}$, with $\mathbb{E} \{ \mathbf{G}_{kk} \} = \mathbf{Z}_k = [\mathbf{Z}_{1k}^T, \dots, \mathbf{Z}_{Mk}^T]^T$ and $\mathbf{S}_{k,c} = \text{diag}(\mathbf{Z}_{1k}, \dots, \mathbf{Z}_{Mk})$ with the (n, n') -th element of $\mathbf{Z}_{mk} \in \mathbb{C}^{N \times N}$ being $[\mathbf{Z}_{mk}]_{nn'} = \text{tr}(\hat{\mathbf{R}}_{mk}^{n'n})$. Moreover, $\mathbf{T}_{kl,(1)} \triangleq \text{diag}(\boldsymbol{\Gamma}_{kl,1}^{(1)}, \dots, \boldsymbol{\Gamma}_{kl,M}^{(1)}) \in \mathbb{C}^{MN \times MN}$ and $\mathbf{T}_{kl,(2)}^{mm'} = \boldsymbol{\Gamma}_{kl,m}^{(2)} - \boldsymbol{\Gamma}_{kl,m}^{(1)}$ if $m = m'$ and $\boldsymbol{\Lambda}_{mkl} \bar{\mathbf{F}}_{l,u} \boldsymbol{\Lambda}_{m'lk}$ otherwise, where $\mathbf{T}_{kl,(2)}^{mm'}$ denotes (m, m') -submatrix of $\mathbf{T}_{kl,(2)} \in \mathbb{C}^{MN \times MN}$, the (n, n') -th element of $N \times N$ -dimension complex matrices $\boldsymbol{\Lambda}_{mkl}$, $\boldsymbol{\Lambda}_{m'lk}$, $\boldsymbol{\Gamma}_{kl,m}^{(1)}$ and $\boldsymbol{\Gamma}_{kl,m}^{(2)}$ are $[\boldsymbol{\Lambda}_{mkl}]_{nn'} = \text{tr}(\boldsymbol{\Xi}_{mkl}^{n'n})$, $[\boldsymbol{\Lambda}_{m'lk}]_{nn'} = \text{tr}(\boldsymbol{\Xi}_{m'lk}^{n'n})$, $[\boldsymbol{\Gamma}_{mkl}^{(1)}]_{nn'} = \sum_{i=1}^N \sum_{i'=1}^N [\bar{\mathbf{F}}_{l,u}]_{i'i} \text{tr}(\mathbf{R}_{ml}^{i'i} \hat{\mathbf{R}}_{mk}^{n'n})$ and $[\boldsymbol{\Gamma}_{kl,m}^{(2)}]_{nn'}$ given by

$$\begin{aligned} [\boldsymbol{\Gamma}_{kl,m}^{(2)}]_{nn'} &= \sum_{i=1}^N \sum_{i'=1}^N [\bar{\mathbf{F}}_l]_{i'i} \left\{ \text{tr} \left(\mathbf{R}_{ml}^{i'i} \mathbf{P}_{mkl,(1)}^{n'n} \right) \right. \\ &\quad \left. + \tau_p^2 \sum_{q_1=1}^N \sum_{q_2=1}^N \left[\text{tr} \left(\tilde{\mathbf{P}}_{mkl,(2)}^{q_1 n} \tilde{\mathbf{R}}_{ml}^{i' q_2} \tilde{\mathbf{R}}_{ml}^{q_2 i} \tilde{\mathbf{P}}_{mkl,(2)}^{n' q_1} \right) + \text{tr} \left(\tilde{\mathbf{P}}_{mkl,(2)}^{q_1 n} \tilde{\mathbf{R}}_{ml}^{i' q_2} \right) \text{tr} \left(\tilde{\mathbf{P}}_{mkl,(2)}^{n' q_2} \tilde{\mathbf{R}}_{ml}^{q_2 i} \right) \right] \right\} \end{aligned} \quad (33)$$

with $\boldsymbol{\Xi}_{mkl} = \tau_p \mathbf{R}_{ml} \tilde{\mathbf{F}}_{l,p}^H \boldsymbol{\Psi}_{mk}^{-1} \tilde{\mathbf{F}}_{k,p} \mathbf{R}_{mk}$, $\boldsymbol{\Xi}_{m'lk} = \tau_p \mathbf{R}_{m'k} \tilde{\mathbf{F}}_{k,p}^H \boldsymbol{\Psi}_{m'k}^{-1} \tilde{\mathbf{F}}_{l,p} \mathbf{R}_{m'l}$, $\mathbf{P}_{mkl,(1)} = \tau_p \mathbf{S}_{mk} (\boldsymbol{\Psi}_{mk} - \tau_p \tilde{\mathbf{F}}_{l,p} \mathbf{R}_{ml} \tilde{\mathbf{F}}_{l,p}^H) \mathbf{S}_{mk}^H$, $\mathbf{S}_{mk} = \mathbf{R}_{mk} \tilde{\mathbf{F}}_{k,p}^H \boldsymbol{\Psi}_{mk}^{-1}$, $\mathbf{P}_{mkl,(2)} = \mathbf{S}_{mk} \tilde{\mathbf{F}}_{l,p} \mathbf{R}_{ml} \tilde{\mathbf{F}}_{l,p}^H \mathbf{S}_{mk}^H$, $\tilde{\mathbf{R}}_{ml}^{ni}$ and $\tilde{\mathbf{P}}_{mkl,(2)}^{ni}$ being (n, i) -submatrix of $\mathbf{R}_{ml}^{\frac{1}{2}}$ and $\mathbf{P}_{mkl,(2)}^{\frac{1}{2}}$, respectively. Furthermore, the optimal LSF D coefficient matrix in (28) and MSE matrix in (31) can also be computed in closed-form as

$$\begin{cases} \mathbf{A}_{k,c}^{\text{opt}} = \left(\sum_{l=1}^K \mathbf{T}_{kl,(1)} + \sum_{l \in \mathcal{P}_k} \mathbf{T}_{kl,(2)} + \sigma^2 \mathbf{S}_{k,c} \right)^{-1} \mathbf{Z}_k \mathbf{F}_{k,u}, \\ \mathbf{E}_{k,(2),c}^{\text{opt}} = \mathbf{I}_N - \mathbf{F}_{k,u}^H \mathbf{Z}_k^H \left(\sum_{l=1}^K \mathbf{T}_{kl,(1)} + \sum_{l \in \mathcal{P}_k} \mathbf{T}_{kl,(2)} + \sigma^2 \mathbf{S}_{k,c} \right)^{-1} \mathbf{Z}_k \mathbf{F}_{k,u}. \end{cases} \quad (34)$$

Proof: The proof is given in Appendix C. \blacksquare

2) *Iteratively WMMSE Precoding Design*: For the LSFD scheme, we also investigate a weighted sum-rate maximization problem as

$$\begin{aligned} & \max_{\{\mathbf{F}\}} \sum_{k=1}^K \mu_{k,(2)} \text{SE}_{k,(2)} \\ & \text{s.t. } \|\mathbf{F}_{k,u,(2)}\|^2 \leq p_k \quad \forall k = 1, \dots, K \end{aligned} \quad (35)$$

where $\mu_{k,(2)}$ represents the priority weight of UE k for the ‘‘LSFD’’ scheme and $\text{SE}_{k,(2)}$ is given in (27) with arbitrary combining structure in the first decoding layer.

Similarly, the matrix-weighted sum-MSE minimization problem as⁹

$$\begin{aligned} & \min_{\{\mathbf{F}, \mathbf{A}, \mathbf{W}, \mathbf{G}, \mathbf{S}\}} \sum_{k=1}^K \mu_{k,(2)} \left[\text{tr}(\mathbf{W}_{k,(2)} \mathbf{E}_{k,(2)}) - \log_2 |\mathbf{W}_{k,(2)}| \right] \\ & \text{s.t. } \|\mathbf{F}_{k,u,(2)}\|^2 \leq p_k \quad \forall k = 1, \dots, K \end{aligned} \quad (36)$$

is equivalent to the weighted sum-rate maximization problem (35), where $\mathbf{W}_{k,(2)}$ is the weight matrix for UE k . Note that (36) is convex over each optimization variable \mathbf{F} , \mathbf{A} , \mathbf{W} , \mathbf{G} , \mathbf{S} but is not jointly convex over all optimization variables. So we can solve (36) by sequentially fixing four of the five optimization variables \mathbf{F} , \mathbf{A} , \mathbf{W} , \mathbf{G} , \mathbf{S} and updating the fifth.¹⁰

The update of \mathbf{A}_k and $\mathbf{E}_{k,(2)}$ are given by the optimal LSFD scheme (28) and MSE matrix with optimal LSFD scheme (31). Note that optimal $\mathbf{W}_{k,(2)}$ for (36) is $\mathbf{W}_{k,(2)}^{\text{opt}} = \mathbf{E}_{k,(2)}^{-1}$.

Remark 3. When $\mathbf{A}_k^{\text{opt}}$ and $\mathbf{W}_{k,(2)}^{\text{opt}}$ for all UEs are applied in (36), we notice that (36) becomes to the equivalent optimization problem of (35) as

$$\begin{aligned} & \max_{\{\mathbf{F}, \mathbf{G}, \mathbf{S}\}} \sum_{k=1}^K \mu_{k,(2)} \log_2 \left| \left(\mathbf{E}_{k,(2)}^{\text{opt}} \right)^{-1} \right| \\ & \text{s.t. } \|\mathbf{F}_{k,u,(2)}\|^2 \leq p_k \quad \forall k = 1, \dots, K \end{aligned} \quad (37)$$

which is a well-known relationship between $\mathbf{E}_{k,(2)}^{\text{opt}}$ and $\text{SE}_{k,(2)}^{\text{opt}}$ and proven in Appendix E.

Last but not least, fixing other variables, the update of $\mathbf{F}_{k,u,(2)}$ for (36) results in the optimization

⁹The notation \mathbf{G} denotes all \mathbf{G} -relevant variables, like $\mathbb{E}\{\mathbf{G}_{kl} \bar{\mathbf{F}}_{l,u,(2)} \mathbf{G}_{kl}^H\}$ and $\mathbb{E}\{\mathbf{G}_{kk}\}$, etc.

¹⁰As for \mathbf{G} and \mathbf{S} , if L-MMSE combining scheme applied, $\mathbb{E}\{\mathbf{G}_{kk}\}$ and \mathbf{S}_k are relevant to $\mathbf{F}_{k,u,(2)}$ so we should also update them. On the contrary, $\mathbb{E}\{\mathbf{G}_{kk}\}$ and \mathbf{S}_k with MR combining structure are irrelevant to \mathbf{F} so we only need to update $\mathbb{E}\{\mathbf{G}_{kl} \bar{\mathbf{F}}_{l,u,(2)} \mathbf{G}_{kl}^H\}$.

problem as

$$\begin{aligned}
\min_{\{\mathbf{F}\}} & \sum_{k=1}^K \mu_{k,(2)} \left[\text{tr} \left(\mathbf{W}_k (\mathbf{I}_N - \mathbf{F}_{k,u,(2)}^H \mathbb{E} \{ \mathbf{G}_{kk}^H \} \mathbf{A}_k) (\mathbf{I}_N - \mathbf{F}_{k,u,(2)}^H \mathbb{E} \{ \mathbf{G}_{kk}^H \} \mathbf{A}_k)^H \right) \right] \\
& + \sum_{k=1}^K \mu_{k,(2)} \left[\text{tr} \left(\mathbf{W}_{k,(2)} \mathbf{A}_k^H \left(\sum_{l \neq k}^K \mathbb{E} \{ \mathbf{G}_{kl} \bar{\mathbf{F}}_{l,u,(2)} \mathbf{G}_{kl}^H \} + \sigma^2 \mathbf{S}_k \right) \mathbf{A}_k \right) \right] \\
\text{s.t.} & \quad \|\mathbf{F}_{k,u,(2)}\|^2 \leq p_k \quad \forall k = 1, \dots, K
\end{aligned} \tag{38}$$

which is a convex quadratic optimization problem. Thus, we can also derive the optimal precoding scheme by applying classic Lagrange multipliers methods and KKT conditions. The Lagrange function of (38) is

$$\begin{aligned}
f(\mathbf{F}_{1,u,(2)}, \dots, \mathbf{F}_{K,u,(2)}) &= \sum_{k=1}^K \mu_{k,(2)} \left[\text{tr} \left(\mathbf{W}_{k,(2)} (\mathbf{I}_N - \mathbf{F}_{k,u,(2)}^H \mathbb{E} \{ \mathbf{G}_{kk}^H \} \mathbf{A}_k) (\mathbf{I}_N - \mathbf{F}_{k,u,(2)}^H \mathbb{E} \{ \mathbf{G}_{kk}^H \} \mathbf{A}_k)^H \right) \right] \\
& + \sum_{k=1}^K \mu_{k,(2)} \left[\text{tr} \left(\mathbf{W}_{k,(2)} \mathbf{A}_k^H \left(\sum_{l \neq k}^K \mathbb{E} \{ \mathbf{G}_{kl} \bar{\mathbf{F}}_{l,u,(2)} \mathbf{G}_{kl}^H \} + \sigma^2 \mathbf{S}_k \right) \mathbf{A}_k \right) \right] \\
& + \sum_{k=1}^K \lambda_{k,(2)} (\text{tr}(\mathbf{F}_{k,u,(2)} \mathbf{F}_{k,u,(2)}^H) - p_k).
\end{aligned} \tag{39}$$

Theorem 3. *By applying the first-order optimality condition of (39) with respect to each $\mathbf{F}_{k,u,(2)}$ and fixing other optimization variables, we obtain the optimal precoding scheme as*

$$\mathbf{F}_{k,u,(2)}^{\text{opt}} = \mu_{k,(2)} \left(\sum_{l=1}^K \mu_{l,(2)} \mathbb{E} \{ \mathbf{G}_{lk}^H \mathbf{A}_l \mathbf{E}_{l,(2)}^{-1} \mathbf{A}_l^H \mathbf{G}_{lk} \} + \lambda_{k,(2)} \mathbf{I}_N \right)^{-1} \mathbb{E} \{ \mathbf{G}_{kk}^H \} \mathbf{A}_k \mathbf{E}_{k,(2)}^{-1}, \tag{40}$$

where $\lambda_{k,(2)} \geq 0$ is the Lagrangian multiplier during the phase of ‘‘LSFD’’ scheme. According to the KKT condition, $\lambda_{k,(2)}$ and $\mathbf{F}_{k,u,(2)}$ should also satisfy

$$\|\mathbf{F}_{k,u,(2)}\|^2 \leq p_k, \quad \lambda_{k,(2)} \left(\|\mathbf{F}_{k,u,(2)}\|^2 - p_k \right) = 0, \quad \lambda_{k,(2)} \geq 0. \tag{41}$$

Note that when $\sum_{l=1}^K \mu_{l,(2)} \mathbb{E} \{ \mathbf{G}_{lk}^H \mathbf{A}_l \mathbf{E}_{l,(2)}^{-1} \mathbf{A}_l^H \mathbf{G}_{lk} \}$ is invertible and $\text{tr}[\mathbf{F}_{k,u,(2)}(0) \mathbf{F}_{k,u,(2)}(0)^H] \leq p_k$, then $\mathbf{F}_{k,u,(2)}^{\text{opt}} = \mathbf{F}_{k,u,(2)}(0)$, otherwise we must have $\text{tr}[\mathbf{F}_{k,u,(2)}(\lambda_{k,(2)}) \mathbf{F}_{k,u,(2)}(\lambda_{k,(2)})^H] = p_k$. Following the similar method in Corollary 3, we notice that $\lambda_{k,(2)}$ can be easily found by a 1-D bisection algorithm since $\text{tr}[\mathbf{F}_{k,u,(2)}(\lambda_{k,(2)}) \mathbf{F}_{k,u,(2)}(\lambda_{k,(2)})^H]$ is a monotonically decreasing function of $\lambda_{k,(2)}$.

Moreover, if MR combining $\mathbf{V}_{mk} = \hat{\mathbf{H}}_{mk}$ is applied in the first layer, we can compute expectations in (40) in closed-form as following theorem.

Theorem 4. *With MR combining $\mathbf{V}_{mk} = \hat{\mathbf{H}}_{mk}$ and the optimal LSFD scheme applied, we can compute*

$\mathbb{E}\{\mathbf{G}_{kk}^H\}$, $\mathbf{A}_k^{\text{opt}}$, and $\mathbf{E}_{k,(2)}^{\text{opt}}$ in closed-form as Theorem 2. Moreover, we have $\bar{\mathbf{T}}_{lk} = \mathbb{E}\{\mathbf{G}_{lk}^H \mathbf{A}_l \mathbf{E}_{l,(2)}^{-1} \mathbf{A}_l^H \mathbf{G}_{lk}\} \in \mathbb{C}^{N \times N}$ where the (i, n) -th element of $\bar{\mathbf{T}}_{lk}$ is $\text{tr}(\bar{\mathbf{A}}_l \bar{\mathbf{G}}_{lk,ni})$ with $\bar{\mathbf{A}}_l \triangleq \mathbf{A}_l \mathbf{E}_{l,(2)}^{-1} \mathbf{A}_l^H$ and the $[(m-1)N + p, (m'-1)N + p']$ -th (or $[o, j]$ -th briefly) entry of $\bar{\mathbf{G}}_{lk,ni} \triangleq \mathbb{E}\{\mathbf{g}_{lk,n} \mathbf{g}_{lk,i}^H\} \in \mathbb{C}^{MN \times MN}$ being

$$\mathbb{E}\{\mathbf{g}_{lk,n} \mathbf{g}_{lk,i}^H\}_{oj} = \begin{cases} 0, & l \notin \mathcal{P}_k, m \neq m' \\ \text{tr}(\mathbf{R}_{mk}^{ni} \hat{\mathbf{R}}_{ml}^{p'p}), & l \notin \mathcal{P}_k, m = m' \\ \text{tr}(\mathbf{\Xi}_{mlk}^{np}) \text{tr}(\mathbf{\Xi}_{m'kl}^{p'i}), & l \in \mathcal{P}_k, m \neq m' \\ \text{tr}\left(\mathbf{R}_{mk}^{ni} \mathbf{P}_{mlk,(1)}^{p'p}\right) + \tau_p^2 \sum_{q_1=1}^N \sum_{q_2=1}^N \text{tr}\left(\tilde{\mathbf{P}}_{mlk,(2)}^{q_1p} \tilde{\mathbf{R}}_{mk}^{nq_2} \tilde{\mathbf{R}}_{mk}^{q_2i} \tilde{\mathbf{P}}_{mlk,(2)}^{p'q_1}\right) \\ \quad + \tau_p^2 \sum_{q_1=1}^N \sum_{q_2=1}^N \text{tr}\left(\tilde{\mathbf{P}}_{mlk,(2)}^{q_1n} \tilde{\mathbf{R}}_{mk}^{nq_1}\right) \text{tr}\left(\tilde{\mathbf{P}}_{mlk,(2)}^{p'q_2} \tilde{\mathbf{R}}_{mk}^{q_2i}\right), & l \in \mathcal{P}_k, m = m' \end{cases} \quad (42)$$

where $\mathbf{\Xi}_{mlk} = \tau_p \mathbf{R}_{mk} \tilde{\mathbf{F}}_{k,p}^H \mathbf{\Psi}_{mk}^{-1} \tilde{\mathbf{F}}_{l,p} \mathbf{R}_{ml}$, $\mathbf{\Xi}_{m'kl} = \tau_p \mathbf{R}_{m'l} \tilde{\mathbf{F}}_{l,p}^H \mathbf{\Psi}_{m'l}^{-1} \tilde{\mathbf{F}}_{k,p} \mathbf{R}_{m'k}$, $\mathbf{S}_{ml} = \mathbf{R}_{ml} \tilde{\mathbf{F}}_{l,p}^H \mathbf{\Psi}_{ml}^{-1}$, $\mathbf{P}_{mlk,(1)} = \tau_p \mathbf{S}_{ml} (\mathbf{\Psi}_{ml} - \tau_p \tilde{\mathbf{F}}_{k,p} \mathbf{R}_{mk} \tilde{\mathbf{F}}_{k,p}^H) \mathbf{S}_{ml}^H$ and $\mathbf{P}_{mlk,(2)} = \mathbf{S}_{ml} \tilde{\mathbf{F}}_{k,p} \mathbf{R}_{mk} \tilde{\mathbf{F}}_{k,p}^H \mathbf{S}_{ml}^H$. Plugging the derived results into (40), we can compute $\mathbf{F}_{k,u,(2)}^{\text{opt}}$ in closed-form as

$$\mathbf{F}_{k,u,(2),c}^{\text{opt}} = \mu_{k,(2)} \left(\sum_{l=1}^K \mu_{l,(2)} \bar{\mathbf{T}}_{lk} + \lambda_{k,(2)} \mathbf{I}_N \right)^{-1} \mathbf{Z}_k^H \mathbf{A}_{k,c}^{\text{opt}} \mathbf{E}_{k,(2),c}^{\text{opt},-1}. \quad (43)$$

Proof: The proof is given in Appendix F. ■

Furthermore, an iterative optimization algorithm for $\mathbf{F}_{k,u,(2)}$ is summarized in Algorithm 2. The convergence of Algorithm 2 is proven in [30, Theorem 3].

Remark 4. Relying on the iterative minimization of weighted MSE, two efficient uplink I-WMMSE precoding schemes to maximize the weighted sum SE are proposed. The I-WMMSE precoding schemes for the “FCP” and “LSFD” schemes are investigated in Algorithm 1 and Algorithm 2, respectively. Note that the design of I-WMMSE precoding scheme for the FCP/LSFD is based on instantaneous CSI/channel statistics, respectively. More importantly, we can compute I-WMMSE precoding schemes in novel closed-form only for the LSFD scheme with MR combining based on Theorem 4.

IV. PRECODING IMPLEMENTATION AND COMPLEXITY ANALYSIS

In this section, we discuss the practical implementation and analyze computational complexity for the UL precoding schemes investigated in Section III.

A. Precoding Implementation

1) *Precoding Characteristics:* As described above, we investigate a standard block fading model, where the channel response is constant and frequency flat in a coherence block, which contains τ_c channel

Algorithm 2: I-WMMSE Algorithm for the Design of $\mathbf{F}_{k,u,(2)}$

Input: Channel statistics Θ for all possible pairs; UE weights $\mu_{k,(2)}$ for all UEs;

Output: Optimal precoding matrices $\mathbf{F}_{k,u,(2)}$ for all UEs ($\mathbf{F}_{k,u,(2)}^{(i)}$ for the first or third stopping criterion and $\mathbf{F}_{k,u,(2)}^{(i-1)}$ for the second stopping criterion);

- 1 **Initiation:** $i = 0$, $\mathbf{F}_{k,u,(2)}^{(0)}$ and $R_{(2)}^{(0)} = \sum_{k=1}^K \mu_{k,(2)} \text{SE}_{k,(2)}^{(0)}$ for all UEs; maximum iteration number $I_{(2),\max}$ and threshold $\varepsilon_{(2)}$;
 - 2 **repeat**
 - 3 $i = i + 1$
 - 4 Update channel statistics $\Theta^{(i)}$, such as $\mathbb{E}\{\mathbf{G}_{kk}^{(i)}\}$, $\mathbb{E}\{\mathbf{G}_{kl}^{(i)} \bar{\mathbf{F}}_{l,u}^{(i-1)} (\mathbf{G}_{kl}^{(i)})^H\}$ and $\mathbf{S}_k^{(i)}$;
 - 5 Update optimal LSFD matrix $\mathbf{A}_k^{(i)}$ with $\mathbf{F}_{l,u,(2)}^{(i-1)}$ and $\Theta^{(i)}$ based on (28);
 - 6 Update optimal MSE matrix $\mathbf{E}_{k,(2)}^{(i)}$ with $\mathbf{F}_{l,u,(2)}^{(i-1)}$, $\mathbf{A}_k^{(i)}$ and $\mathbb{E}\{\mathbf{G}_{kk}^{(i)}\}$ based on (31) and update $\mathbf{W}_{k,(2)}^{(i)}$;
 - 7 Update optimal precoding matrix $\mathbf{F}_{k,u,(2)}^{(i)}$ with $\mathbf{A}_k^{(i)}$, $\mathbf{W}_{k,(2)}^{(i)}$ and $\Theta^{(i)}$ based on (40), where $\lambda_{k,(2)}^{(i)}$ is found by a bisection algorithm;
 - 8 Update sum weighted rate $R_{(2)}^{(i)} = \sum_{k=1}^K \mu_{k,(2)} \text{SE}_{k,(2)}^{(i)}$;
 - 9 **until** $|R_{(2)}^{(i)} - R_{(2)}^{(i-1)}|/R_{(2)}^{(i-1)} \leq \varepsilon_{(2)}$ or $R_{(2)}^{(i)} < R_{(2)}^{(i-1)}$ or $i \geq I_{(2),\max}$;
-

uses. For the “fully centralized processing” scheme, we notice that the I-WMMSE precoding design is implemented at the CPU based on the instantaneous CSI as (18). Moreover, to guarantee the convergence of Algorithm 1, only MMSE combining as (9) is advocated to detect the UL data since the equivalent relationship between $\mathbf{E}_{k,(1)}^{\text{opt}}$ and $\text{SE}_{k,(1)}^{\text{opt}}$, which only satisfies with MMSE combining, should be guaranteed. As for the LSFD scheme, the optimal design of $\mathbf{F}_{k,u,(2)}^{\text{opt}}$ as (40) can only be implemented at the CPU, but relies only on channel statistics. Besides, L-MMSE or MR combining can be applied at each AP. When MR combining is applied, all terms in Algorithm 2 can be computed in closed-form as Theorem 4.

2) *Fronthaul Requirements:* For the FCP scheme with the I-WMMSE precoding, in each coherence block, all APs should relay their received signals to the CPU and the CPU requires precoding matrices $\mathbf{F}_{k,u,(1)}$ feedback to all UEs. All APs need to send $\tau_c ML$ complex scalars ($\tau_p ML$ complex scalars for the pilot signals and $(\tau_c - \tau_p) ML$ complex scalars for the received data signals). Besides, the full correlation matrices $\{\mathbf{R}_{mk}\}$ are available at the CPU, which contains $MKL^2N^2/2$ complex scalars for each realization of the AP/UE locations/statistics¹¹. Moreover, the CPU transmits optimal precoding matrices to all UEs, which are described by KN^2 complex scalars per coherence block. In summary, for the FCP scheme with the I-WMMSE precoding implemented, total $\tau_c MLN_r + MKL^2N^2/2 + KN^2N_r$ complex scalars are transmitted via fronthaul links for each realization of the AP/UE locations. For comparison, when

¹¹Note that the channel statistics remain constant for each realization of the AP/UE locations and each realization of the AP/UE locations contains N_r channel realizations (coherence blocks).

the FCP scheme without the I-WMMSE precoding is implemented, all APs should also transmit $\tau_c ML$ complex scalars for the received signals to the CPU in each coherence block and $MKL^2N^2/2$ complex scalars for $\{\mathbf{R}_{mk}\}$ to the CPU for each realization of the AP/UE locations. So for the CFP scheme without the I-WMMSE precoding, total $\tau_c MLN_r + MKL^2N^2/2$ complex scalars are transmitted via fronthaul links for each realization of the AP/UE locations.

As for the LSFD scheme with the I-WMMSE precoding, all APs transmit their local data estimates $\tilde{\mathbf{x}}_{mk}$, described by $(\tau_c - \tau_p)MKN$ complex scalars, to the CPU per coherence block. Besides, $\mathbb{E}\{\mathbf{G}_{kk}\} \in \mathbb{C}^{MN \times N}$, described by MKN^2 complex scalars for each realization of the AP/UE locations, are also required at the CPU. As for $\mathbb{E}\{\mathbf{G}_{kl}\bar{\mathbf{F}}_{l,u,(2)}\mathbf{G}_{kl}^H\} \in \mathbb{C}^{MN \times MN}$, following the formulation method investigated in Appendix F, the optimization of $\bar{\mathbf{F}}_{l,u,(2)}$ requires the knowledge of $\{\mathbb{E}[\mathbf{v}_{ml,p}^H \mathbf{h}_{mk,n} \mathbf{h}_{m'k,i}^H \mathbf{v}_{m'l,p'}]\}$, described by $M^2K^2N^4/2$ complex scalars for each realization of the AP/UE locations, where $\mathbf{v}_{ml,p}$ denotes the p -th column of \mathbf{V}_{ml} . Moreover, the CPU requires optimal precoding matrices $\mathbf{F}_{k,u,(2)}$ feedback to all APs and UEs only for each realization of the AP/UE locations, which are KN^2 complex scalars. As for the LSFD scheme without the I-WMMSE precoding, local data estimates $\tilde{\mathbf{x}}_{mk}$, described by $(\tau_c - \tau_p)MKN$ complex scalars per coherence block, $\mathbb{E}\{\mathbf{G}_{kk}\}$, described by MKN^2 complex scalars for each realization of the AP/UE locations, and $\mathbb{E}\{\mathbf{G}_{kl}\mathbf{G}_{kl}^H\} \in \mathbb{C}^{MN \times MN}$, described by $M^2K^2N^2/2$ complex scalars for each realization of the AP/UE locations, are required. That is total $(\tau_c - \tau_p)MKNN_r + MKN^2 + M^2K^2N^2/2$ complex scalars transmitted via fronthaul links for each realization of the AP/UE locations.

3) *Practical Implementation:* Note that the basic motivation of the investigated I-WMMSE precoding schemes is to achieve as good the sum uplink SE performance as possible so we ignore some practical issues, which are vital for the realistic implementation of the investigated precoding schemes. When the precoding schemes are implemented in practice, these realistic issues should be considered.

- Capacity-constrained fronthaul network

As discussed above, the I-WMMSE precoding require more fronthaul requirements than the case without the I-WMMSE precoding. It is quite vital to consider a more practical capacity-constrained fronthaul network [38]. Moreover, the wireless fronthaul [39], which is more flexible than the conventional wire fronthaul, would also be regarded as a promising solution to boost the practical implementation of the I-WMMSE precoding.

- Scalability aspects with dynamic cooperation clusters

When the precoding schemes are implemented in practice, a more realistic network architecture with

multiple CPUs and dynamic cooperation clusters should be advocated, where each UE is only served by a cluster of APs (that is a user-centric cluster) and the APs are grouped into cell-centric clusters as shown in Fig. 1. Note a user-centric cluster might consist of APs connecting with different CPUs. Based on the signal processing schemes in [9], [18], the analytical framework in this paper can be implemented in a scalable paradigm where the fronthaul requirements and computational complexity can be relieved with an anticipated modest performance loss compared with canonical architecture. The I-WMMSE precoding design with these two practical aspects is left in future work. To bring valuable technical insights for the study of I-WMMSE precoding schemes with the DCC strategy and the capacity-constrained fronthaul link, we provide two tutorials for the FCP and LSFD in Fig. 2 based on [9], [10], [38].

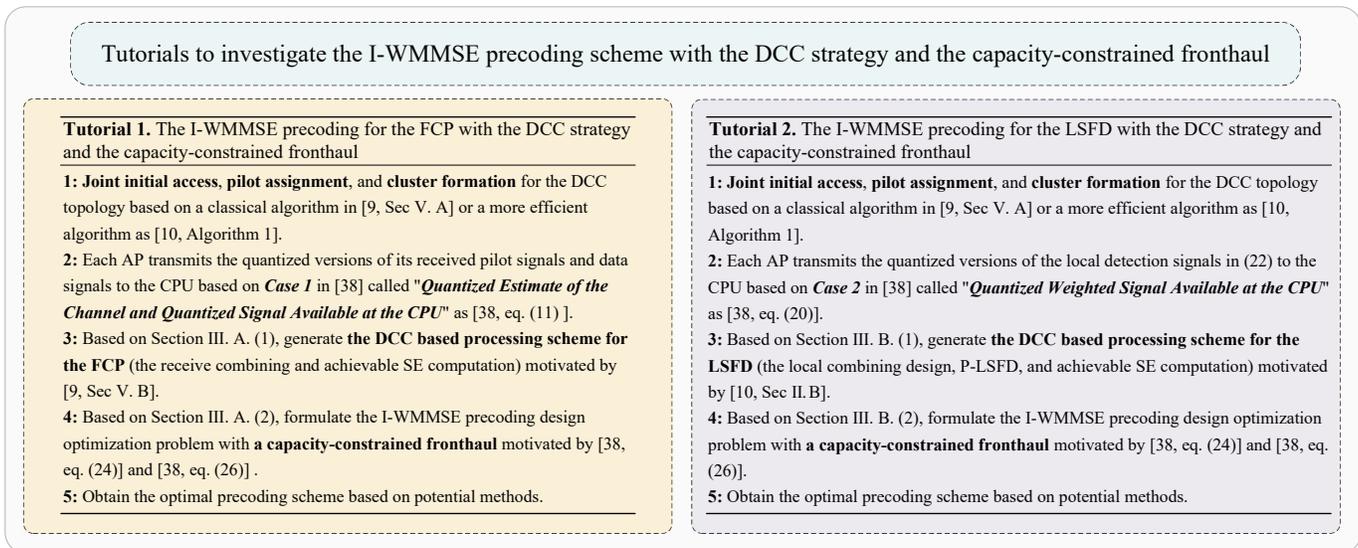


Fig. 2. Two tutorials to investigate the I-WMMSE precoding schemes with the DCC strategy and the capacity-constrained fronthaul.

B. Complexity Analysis

In this subsection, we analyze the computational complexity of two precoding schemes investigated. Since the bisection step for $\lambda_{k,\{(1),(2)\}}$ generally takes few iterations compared with other steps, we ignore bisection steps for $\lambda_{k,\{(1),(2)\}}$ in the complexity analysis. For the fully centralized processing scheme and each realization of the AP/UE locations, the per-iteration complexity of iterative optimization is $\mathcal{O}(M^3 K^2 N^5 N_r)$. For the LSFD scheme and each realization of the AP/UE locations, the per-iteration complexity of iterative optimization based on L-MMSE combining with the Monte-Carlo method, MR combining with the Monte-Carlo method and MR combining with the closed-form expressions are $\mathcal{O}(M^2 K^2 N^3 N_r)$, $\mathcal{O}(M^2 K^2 N^3 N_r + M^3 K N^3)$ and $\mathcal{O}(M^3 K^2 N^5)$, respectively. To further reduce the computation complexity, it's quite necessary to apply the asymptotic analysis method [40], [41] to compute the terms, which cannot be computed in closed-form, in approximation results.

TABLE I

COMPARISON OF TWO PRECODING SCHEMES IN THIS PAPER. THE NUMBER OF COMPLEX SCALARS IS COMPUTED FOR EACH REALIZATION OF THE AP/UE LOCATIONS. THE SUM SE IMPROVEMENT IS COMPUTED WITH $M = 20$, $K = 10$, $L = 1$ AND $N = 4$.

	FCP	LSFD
CSI	Instantaneous CSI	Statistical CSI
Detection scheme	MMSE combining	L-MMSE/MR combining + Optimal LSFD scheme
Number of complex scalars sent from APs to the CPU with I-WMMSE precoding	$\tau_c MLN_r + MKL^2 N^2 / 2$	$(\tau_c - \tau_p)MKNN_r + MKN^2 + M^2 K^2 N^4 / 2$
Number of complex scalars sent from APs to the CPU without I-WMMSE precoding	$\tau_c MLN_r + MKL^2 N^2 / 2$	$(\tau_c - \tau_p)MKNN_r + MKN^2 + M^2 K^2 N^2 / 2$
Number of complex scalars feedback sent from the CPU	$KN^2 N_r$	KN^2
Per-iteration computational complexity	$\mathcal{O}(M^3 K^2 N^5 N_r)$	L-MMSE: $\mathcal{O}(M^2 K^2 N^3 N_r)$ MR (Monte-Carlo): $\mathcal{O}(M^2 K^2 N^3 N_r + M^3 KN^3)$ MR (Analytical): $\mathcal{O}(M^3 K^2 N^5)$
Sum SE improvement	28.93%	L-MMSE: 46.74% MR: 15.13%

V. NUMERICAL RESULTS

In this paper, a CF mMIMO system is investigated, where all APs and UEs are uniformly distributed in a 1×1 km² area with a wrap-around scheme [42]. The pathloss and shadow fading are modeled similarly as [28]. In practice, $\mathbf{U}_{mk,r}$, $\mathbf{U}_{mk,t}$ and $\mathbf{\Omega}_{mk}$ are estimated through measurements [29]. However, we generate them randomly in this paper, where the coupling matrix $\mathbf{\Omega}_{mk}$ consists of one strong transmit eigendirection capturing dominant power [43]¹². Besides, we have $\mathbf{F}_{k,p} = \mathbf{F}_{k,u,\{(1),(2)\}}^{(0)} = \sqrt{\frac{p_k}{N}} \mathbf{I}_N$. As for Algorithm 1 and Algorithm 2, balancing the convergence and accuracy, we assume that $I_{(1),\max} = I_{(2),\max} = 20$, $\varepsilon_{(1)} = \varepsilon_{(2)} = 5 \times 10^{-4}$, and weights for all UEs are equal ($\mu_{k,(1)} = \mu_{k,(2)} = 1$) without losing generality, respectively. Moreover, we consider communication with 20 MHz bandwidth and $\sigma^2 = -94$ dBm noise power. All UEs transmit with 200 mW power constraint. Each coherence block contains $\tau_c = 200$ channel uses and $\tau_p = KN/2$. Besides, a pilot assignment approach similar as that in [28] is investigated.

Figure 3 shows the cumulative distribution function (CDF) of the achievable sum SE over different realizations of the AP/UE locations for two processing schemes investigated (we shortly call ‘‘fully centralized processing’’ as ‘‘FCP’’ in the following) over ‘‘I-WMMSE precoding’’ or ‘‘w/o precoding’’¹³. We notice that the FCP scheme undoubtedly achieves higher SE than that of the LSFD scheme since the

¹²In this paper, we choose one eigendirection capturing dominant channel power (randomly accounting for 80% ~ 95% of the total channel power) and other eigendirections contain the remaining power.

¹³The ‘‘w/o precoding’’ scenario denotes that identity precoding matrices $\mathbf{F}_{k,u,\{(1),(2)\}} = \sqrt{\frac{p_k}{N}} \mathbf{I}_N$ are implemented without optimization.

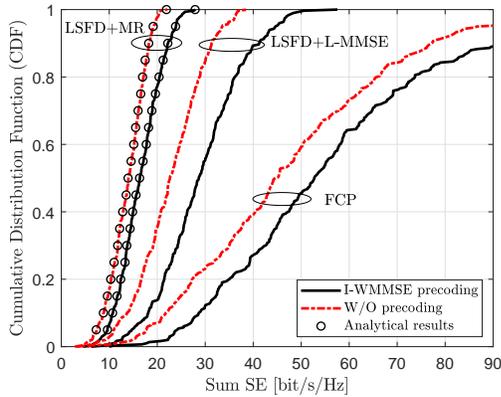


Fig. 3. CDF of the sum SE over different processing schemes and precoding schemes with $M = 20$, $K = 10$, $L = 2$, and $N = 4$.

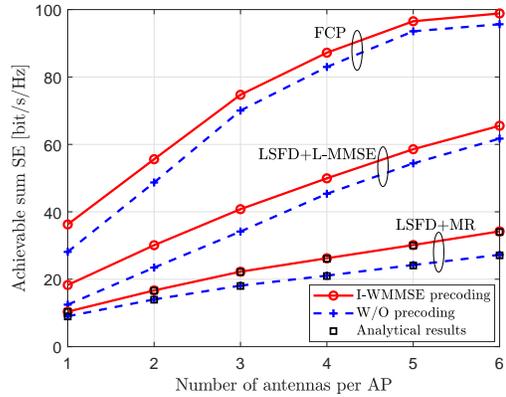


Fig. 4. Sum SE against the number antennas per AP L over different processing schemes and precoding schemes with $M = 20$, $K = 10$, and $N = 4$.

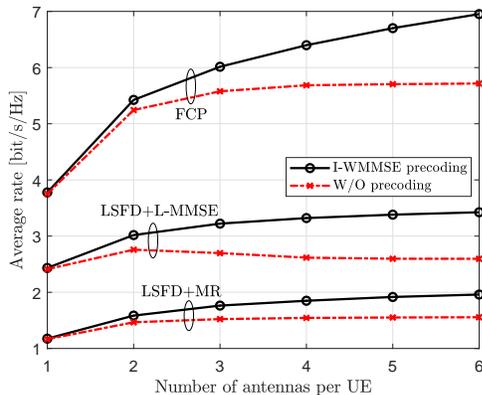


Fig. 5. Average rate against the number of antennas per UE N over different processing schemes and precoding schemes with $M = 20$, $K = 10$, and $L = 2$.

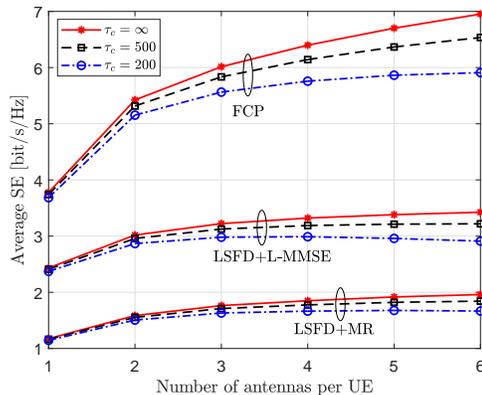


Fig. 6. Average SE with I-WMMSE precoding schemes against the number of antennas per UE N over different τ_c with $M = 20$, $K = 10$, and $L = 2$.

FCP with MMSE combining is a competitive scheme in CF mMIMO [5]. More importantly, the proposed I-WMMSE schemes are efficient to improve the respective achievable sum SE performance, e.g., 12.78%, 19.54% and 28.13% sum SE improvement for the FCP, the LSFDF with MR combining and the LSFDF with L-MMSE combining, respectively. Besides, for the LSFDF scheme with MR combining, markers “o” generated by analytical results overlap with the curves generated by simulations, respectively, validating our derived closed-form expressions.

Figure 4 shows the achievable sum SE as a function of the number of antennas per AP with two processing schemes investigated and different precoding schemes¹⁴. We notice that, for the FCP or LSFDF with (L-)MMSE combining, the performance gap between the “I-WMMSE” and “w/o precoding” becomes smaller with the increase of L , which implies that (L-)MMSE combining can use all antennas on each

¹⁴Note that the achievable sum SE investigated is the average sum SE value taken over many AP/UE locations.

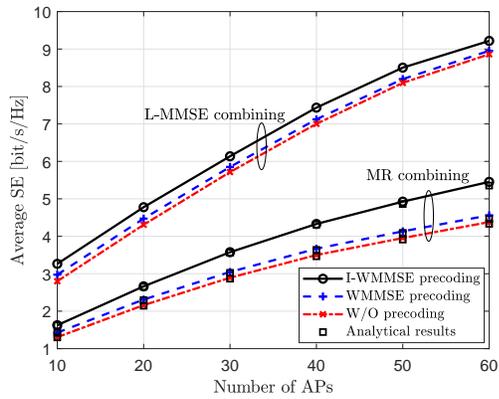


Fig. 7. Average SE against the number of APs M for the LSFD scheme with $K = 10$, $L = 4$, and $N = 4$.

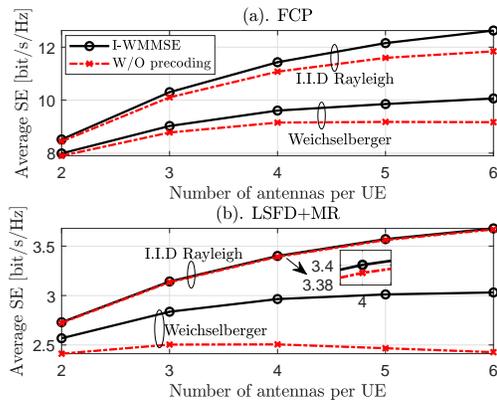


Fig. 8. Average SE against the number of antennas per UE N for different channel models with $M = 40$, $K = 8$, and $L = 2$.

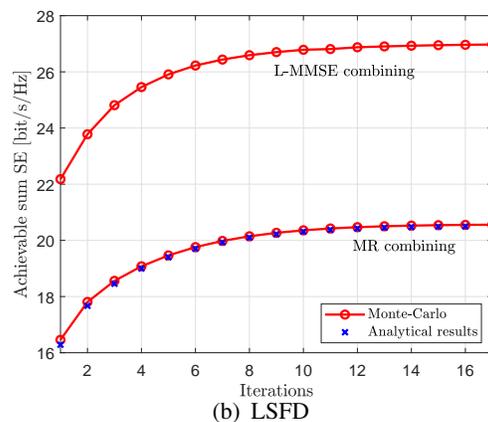
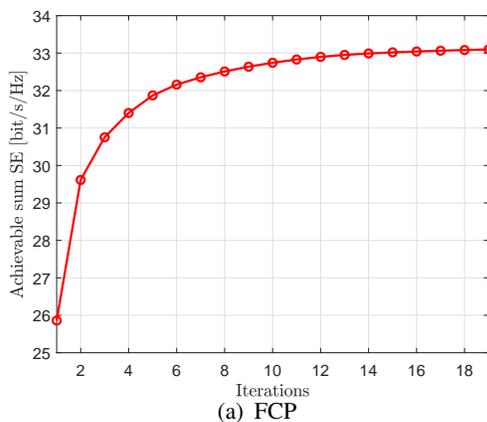


Fig. 9. Convergence examples of the I-WMMSE algorithm for the FCP and LSFD with $M = 20$, $K = 10$, $L = 2$, and $N = 4$.

AP to suppress interference and achieve excellent SE performance even without any precoding scheme. For instance, the performance gap between the “I-WMMSE” and “w/o precoding” for the LSFD with L-MMSE combining is 46.74% and 6.17% over $L = 1$ and $L = 6$, respectively. Meanwhile, for the LSFD with MR combining, the performance gap between the “I-WMMSE” and “w/o precoding” becomes large with the increase of L , e.g. 15.13% and 25.48% for $L = 1$ and $L = 6$, respectively. Besides, for the LSFD scheme with MR combining, markers “ \square ” generated by analytical results overlap with the curves generated by simulations, respectively, validating our derived closed-form expressions.

To further show the advantage of the proposed I-WMMSE precoding schemes, Fig. 5 shows the average rate¹⁵ as a function of the number of antennas per UE. We find that the average rates for all schemes with I-WMMSE precoding schemes grow with N and the average rates for the case without UL precoding may also suffer the degradation with the increase of N . The implementation of the I-WMMSE precoding

¹⁵Note that one main reason for the phenomenon that additional UE antennas may give rise to the SE degradation is that increasing N will increase the channel estimation overhead and reduce the pre-log factor “ $(\tau_c - \tau_p)/\tau_c$ ” in all SE expressions [26], [28]. So we investigate “the average rate” in Fig. 5, ignoring the effect of “ $(\tau_c - \tau_p)/\tau_c$ ”.

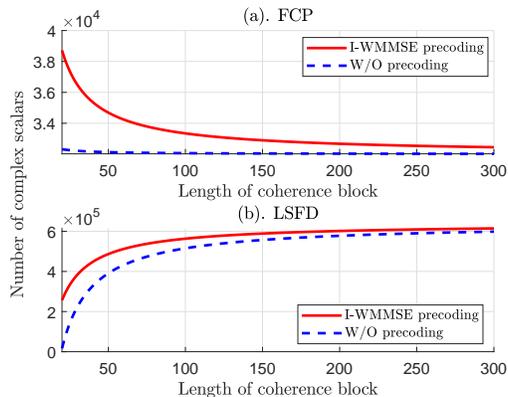


Fig. 10. Total number of complex scalars sent via the fronthaul per channel use for each realization of the AP/UE locations with $M = 20$, $K = 10$, $L = 2$, and $N = 4$.

schemes undoubtedly makes UEs benefit from multiple antennas and achieve excellent rate performance. Moreover, we observe that the I-WMMSE precoding schemes perform more efficiently with a larger number of UE antennas. For instance, the average rate improvements achieved by the I-WMMSE precoding for the LSF with L-MMSE combining are 31.91% and 9.43% for $N = 6$ and $N = 2$, respectively. However, the average SE (with scaling factor $(\tau_c - \tau_p)/\tau_c$) with I-WMMSE precoding implemented may also degrade with the increase of N as the Fig. 2 in [1] since, with the increase of N , the prerequisite of “mutually orthogonal pilot matrices” still requires huge channel uses for the pilot transmission and the inter-user interference also increases. So the design of non-orthogonal pilot matrices and per-antenna power control scheme are quite necessary, which are regarded as promising ways to reduce the cost of pilot transmission and further improve the SE performance [44].

Figure 6 discusses the average SE with I-WMMSE precoding schemes against N over different τ_c . Note that Fig. 5 can be viewed as a special case in Fig. 6 with the coherence block with infinite length $\tau_c = \infty$. We observe that the average SE with I-WMMSE precoding schemes increases with N over $\tau_c = 500$ or ∞ , which means the SE performance can benefit from having additional UE antennas when the coherence block resource is abundant.

Figure 7 investigates the average SE as a function of M for the LSF scheme over different precoding schemes¹⁶. For MR combining, markers “□” generated by analytical results overlap with the curves generated by simulations, respectively, validating our derived closed-form expressions again. Besides, the I-WMMSE algorithm is more efficient to improve the SE performance for MR combining than that of L-MMSE combining for the scenario over large L and M , e.g., 4.03% and 24.21% SE improvement for L-MMSE combining and MR combining with $M = 60$, respectively, implying that the L-MMSE combining

¹⁶The “WMMSE precoding” denotes the precoding schemes generated by the I-WMMSE algorithm with only single iteration.

based on large L and M can achieve excellent SE performance even without any precoding scheme and the proposed I-WMMSE precoding scheme is handy to mitigate the weakness of MR combining¹⁷.

Figure 8 considers the average SE as a function of N over the i.i.d. and the Weichselberger Rayleigh fading channel. As observed, the proposed I-WMMSE precoding schemes are more efficient over the Weichselberger Rayleigh fading channel. For instance, 24.89% and 9.77% average SE improvement can be achieved when $N = 6$ over the “Weichselberger” scenario for the LSFD scheme with MR combining and the FCP scheme, respectively, but only 0.29% and 6.63% average SE improvement can be achieved for “I.I.D. Rayleigh channel”. Moreover, compared with Fig. 5, we notice that the I-WMMSE precoding scheme for the FCP scheme is more efficient in the highly loaded system (the scenario in Fig. 5) where the number of total AP-antennas is comparable with the number of total UE-antennas.

Figure 9 illustrates the convergence behavior of the I-WMMSE algorithms for the FCP scheme and the LSFD scheme with L-MMSE/MR combining. Note the convergence example in Fig. 9 (a) for the FCP is given by a particular channel realization and the convergence example for the LSFD in Fig. 9 (b) is given by a particular realization of the AP/UE locations. Note that the algorithms investigated can be guaranteed to converge and are efficient to achieve excellent sum SE performance. Besides, Fig. 9 (b) for the LSFD scheme with MR combining validates our derived closed-form expressions in Algorithm 2.

Figure 10 investigates the total number of complex scalars sent via the fronthaul per channel use against τ_c for each realization of the AP/UE locations. As observed, total number of complex scalars per channel use for the FCP/LSFD scheme becomes smaller/larger, which can also be easily found from Table I. Besides, the LSFD scheme requires more fronthaul signaling than the FCP scheme since APs under the LSFD scheme need to transmit all received data signals to the CPU, which requires a huge fronthaul load. More importantly, with the increase of τ_c , the gap between “I-WMMSE precoding” and “W/O precoding” becomes smaller for either the FCP scheme or the LSFD scheme. Considering the SE performance improvement of the I-WMMSE precoding, additional fronthaul loads can be acceptable, especially when the coherence resource is abundant. Although the computational complexity of Algorithm 1 for the FCP scheme is much higher than that of Algorithm 2 for the LSFD scheme, the FCP scheme needs much less fronthaul signaling than that of the LSFD scheme and can achieve better SE performance. So two processing schemes and their respective precoding schemes can be chosen based on different requirements.

¹⁷MR combining is a simple combining scheme but cannot efficiently suppress the interference.

VI. CONCLUSION

We consider a CF mMIMO system with both APs and UEs equipped with multiple antennas over the Weichselberger Rayleigh fading channel. The FCP scheme and LSFD scheme are implemented. To further improve the sum SE performance, efficient UL precoding schemes based on iteratively WMMSE algorithms are investigated to maximize weighted sum SE for the two processing schemes. Note that we compute achievable SE expressions and optimal precoding schemes in novel closed-form for the LSFD scheme with MR combining. Numerical results show that the investigated I-WMMSE precoding schemes are efficient to achieve excellent sum SE performance. More importantly, it can be seen that the proposed I-WMMSE precoding schemes are more efficient with a larger number of UE antennas, which means the I-WMMSE precoding schemes can achieve excellent performance even with a large number of UE antennas. The derived results undoubtedly provides vital insights for the practical implementation of multi-antenna UEs in CF mMIMO systems. In future work, we will investigate the design of UL precoding scheme for the phase of pilot transmission and consider the practical implementation of the investigated I-WMMSE precoding schemes with capacity-constrained fronthaul network and dynamic cooperation clusters. Moreover, the non-orthogonal pilot matrix design will also be considered to further improve the performance for the CF mMIMO system with multi-antenna UEs. Last but not least, the UL precoding performance over a more practical Rician fading channel with phase-shifts will also be analyzed.

APPENDIX A

SOME USEFUL LEMMAS

Lemma 1. Let $\mathbf{X} \in \mathbb{C}^{M \times N}$ be a random matrix and \mathbf{Y} is a deterministic $M \times M$ matrix. The (n, i) -th element of $\mathbb{E}\{\mathbf{X}^H \mathbf{Y} \mathbf{X}\}$ is $\text{tr}(\mathbf{Y} \cdot \mathbb{E}\{\mathbf{x}_i \mathbf{x}_n^H\})$ where \mathbf{x}_i and \mathbf{x}_n are the i -th and n -th column of \mathbf{X} .

Lemma 2. For matrices $\mathbf{A} \in \mathbb{C}^{N_1 \times N_1}$, $\mathbf{B} \in \mathbb{C}^{N_1 \times N_2}$, $\mathbf{C} \in \mathbb{C}^{N_2 \times N_2}$, and $\mathbf{D} \in \mathbb{C}^{N_2 \times N_1}$, we have $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{D} \mathbf{A}^{-1} \mathbf{B} + \mathbf{C}^{-1})^{-1} \mathbf{D} \mathbf{A}^{-1}$, which is a well-known matrix inversion lemma [36, Lemma B.3].

APPENDIX B

PROOF COROLLARY 5

Since the CPU is only aware of channel statistics, we need to treat $\mathbb{E}\{\mathbf{G}_{kk}\} \mathbf{F}_{k,u}$ as the true deterministic channel and rewrite $\tilde{\mathbf{x}}_k$ in (25) as $\tilde{\mathbf{x}}_k = \mathbb{E}\{\mathbf{G}_{kk}\} \mathbf{F}_{k,u} \mathbf{x}_k + \underbrace{(\mathbf{G}_{kk} \mathbf{F}_{k,u} - \mathbb{E}\{\mathbf{G}_{kk}\} \mathbf{F}_{k,u}) \mathbf{x}_k + \sum_{l=1, l \neq k}^K \mathbf{G}_{kl} \mathbf{F}_{l,u} \mathbf{x}_l + \mathbf{n}'_k}_{\mathbf{v}}$ where \mathbf{v} is a complex circular symmetric noise with an invertible covariance matrix $\mathbf{\Xi}_k = \mathbb{E}\{\mathbf{v} \mathbf{v}^H | \Theta\} =$

$\sum_{l=1}^K \mathbb{E}\{\mathbf{G}_{kl}\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\mathbf{G}_{kl}^H\} - \mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\mathbb{E}\{\mathbf{G}_{kk}^H\} + \sigma^2\mathbf{S}_k$. Firstly, we whiten the noise as $\Xi_k^{-\frac{1}{2}}\hat{\mathbf{x}}_k = \Xi_k^{-\frac{1}{2}}\mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}\mathbf{x}_k + \tilde{\mathbf{v}}$, where $\tilde{\mathbf{v}} \triangleq \Xi_k^{-\frac{1}{2}}\mathbf{v}$ becomes white. Next, we project $\Xi_k^{-\frac{1}{2}}\hat{\mathbf{x}}_k$ in the direction of $\Xi_k^{-\frac{1}{2}}\mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}$ to obtain an effective scalar channel as

$$\left(\Xi_k^{-\frac{1}{2}}\mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}\right)^H \Xi_k^{-\frac{1}{2}}\hat{\mathbf{x}}_k = \left(\mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}\right)^H \Xi_k^{-1}\mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}\mathbf{x}_k + \left(\mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}\right)^H \Xi_k^{-1}\mathbf{v}. \quad (44)$$

Based on theories of optimal receivers [37], we derive optimal LSFD matrix $\mathbf{A}_k = \Xi_k^{-1}\mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}$ as

$$\mathbf{A}_k = \left(\sum_{l=1}^K \mathbb{E}\{\mathbf{G}_{kl}\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\mathbf{G}_{kl}^H\} - \mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\mathbb{E}\{\mathbf{G}_{kk}^H\} + \sigma^2\mathbf{S}_k\right)^{-1} \mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}. \quad (45)$$

Moreover, based on the the standard results of matrix derivation in [45], we can easily obtain the LSFD matrix minimizing the conditional MSE for UE k $\text{MSE}_k^{(2)} = \text{tr}(\mathbf{E}_k^{(2)})$ as

$$\mathbf{A}_k = \left(\sum_{l=1}^K \mathbb{E}\{\mathbf{G}_{kl}\bar{\mathbf{F}}_{l,u}\mathbf{G}_{kl}^H\} + \sigma^2\mathbf{S}_k\right)^{-1} \mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}. \quad (46)$$

We notice that the LSFD matrix in (45) is equivalent to the LSFD matrix in (46), except from having another scaling matrix $\mathbf{I}_N - (\mathbf{C}^H\mathbf{B}^{-1}\mathbf{C} + \mathbf{I}_N)^{-1}\mathbf{C}^H\mathbf{B}^{-1}\mathbf{C}$ on the right side, which would not affect the value of (27), where $\mathbf{B} = \sum_{l=1}^K \mathbb{E}\{\mathbf{G}_{kl}\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\mathbf{G}_{kl}^H\} + \sigma^2\mathbf{S}_k$ and $\mathbf{C} = \mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}$. So the LSFD matrix in (46) cannot maximize the achievable SE but minimize the MSE for UE k .

APPENDIX C

PROOF OF THEOREM 2

In this part, we compute terms of (27) in closed-form for the LSFD scheme with MR combining $\mathbf{V}_{mk} = \hat{\mathbf{H}}_{mk}$. For the first term $\mathbf{D}_{k,(2)} = \mathbf{A}_k^H \mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u}$, we have $\mathbb{E}\{\mathbf{G}_{kk}\} = [\mathbb{E}\{\mathbf{V}_{1k}^H\mathbf{H}_{1k}\}; \dots; \mathbb{E}\{\mathbf{V}_{Mk}^H\mathbf{H}_{Mk}\}] = [\mathbf{Z}_{1k}^T, \dots, \mathbf{Z}_{Mk}^T]^T \triangleq \mathbf{Z}_k$, where $\mathbf{Z}_{mk} = \mathbb{E}\{\mathbf{V}_{mk}^H\mathbf{H}_{mk}\} = \mathbb{E}\{\hat{\mathbf{H}}_{mk}^H\hat{\mathbf{H}}_{mk}\} \in \mathbb{C}^{N \times N}$ and the (n, n') -th element of \mathbf{Z}_{mk} can be denoted as $[\mathbf{Z}_{mk}]_{nn'} = \mathbb{E}\{\hat{\mathbf{h}}_{mk,n}^H\hat{\mathbf{h}}_{mk,n'}\} = \text{tr}(\hat{\mathbf{R}}_{mk}^{n'n})$. So we derive the closed-form for $\mathbf{D}_{k,(2)}$ as $\mathbf{D}_{k,(2),c} = \mathbf{A}_k^H \mathbf{Z}_k \mathbf{F}_{k,u}$. As for the second term $\mathbf{S}_k \in \mathbb{C}^{MN \times MN}$, we have $\mathbf{S}_k = \text{diag}(\mathbb{E}\{\mathbf{V}_{1k}^H\mathbf{V}_{1k}\}, \dots, \mathbb{E}\{\mathbf{V}_{Mk}^H\mathbf{V}_{Mk}\}) = \text{diag}(\mathbf{Z}_{1k}, \dots, \mathbf{Z}_{Mk})$. For $\mathbb{E}\{\mathbf{G}_{kl}\bar{\mathbf{F}}_{l,u}\mathbf{G}_{kl}^H\}$, we notice that the (m, m') -submatrix of $\mathbb{E}\{\mathbf{G}_{kl}\bar{\mathbf{F}}_{l,u}\mathbf{G}_{kl}^H\}$ is $\mathbb{E}\{\mathbf{V}_{mk}^H\mathbf{H}_{ml}\bar{\mathbf{F}}_{l,u}\mathbf{H}_{m'l}^H\mathbf{V}_{m'k}\}$.

Based on [28], we compute $\mathbb{E}\{\mathbf{V}_{mk}^H\mathbf{H}_{ml}\bar{\mathbf{F}}_{l,u}\mathbf{H}_{m'l}^H\mathbf{V}_{m'k}\}$ for four possible AP-UE combinations. For “ $m \neq m', l \notin \mathcal{P}_k$ ”, we have $\mathbb{E}\{\mathbf{V}_{mk}^H\mathbf{H}_{ml}\bar{\mathbf{F}}_{l,u}\mathbf{H}_{m'l}^H\mathbf{V}_{m'k}\} = 0$ for the independence between \mathbf{V}_{mk} and \mathbf{H}_{ml} . For “ $m \neq m', l \in \mathcal{P}_k$ ”, we have $\mathbb{E}\{\mathbf{V}_{mk}^H\mathbf{H}_{ml}\bar{\mathbf{F}}_{l,u}\mathbf{H}_{m'l}^H\mathbf{V}_{m'k}\} = \mathbb{E}\{\mathbf{V}_{mk}^H\mathbf{H}_{ml}\}\bar{\mathbf{F}}_{l,u}\mathbb{E}\{\mathbf{H}_{m'l}^H\mathbf{V}_{m'k}\} = \mathbf{\Lambda}_{mkl}\bar{\mathbf{F}}_{l,u}\mathbf{\Lambda}_{m'l'k}$, where the (n, n') -th element of $N \times N$ -dimension complex matrices $\mathbf{\Lambda}_{mkl} \triangleq \mathbb{E}\{\mathbf{V}_{mk}^H\mathbf{H}_{ml}\}$,

$\Lambda_{m'l} \triangleq \mathbb{E}\{\mathbf{H}_{m'l}^H \mathbf{V}_{m'l}\}$ are $[\Lambda_{mkl}]_{nn'} = \mathbb{E}\{\hat{\mathbf{h}}_{mk,n}^H \hat{\mathbf{h}}_{ml,n'}\} = \text{tr}(\Xi_{mkl}^{n'n})$ and $[\Lambda_{m'lk}]_{nn'} = \mathbb{E}\{\hat{\mathbf{h}}_{m'l,n}^H \hat{\mathbf{h}}_{mk,n'}\} = \text{tr}(\Xi_{m'lk}^{n'n})$ with $\Xi_{mkl} \triangleq \mathbb{E}\{\hat{\mathbf{h}}_{ml} \hat{\mathbf{h}}_{mk}^H\} = \tau_p \mathbf{R}_{ml} \tilde{\mathbf{F}}_{l,p}^H \Psi_{mk}^{-1} \tilde{\mathbf{F}}_{k,p} \mathbf{R}_{mk}$, $\Xi_{m'lk} \triangleq \mathbb{E}\{\hat{\mathbf{h}}_{m'l} \hat{\mathbf{h}}_{m'l}^H\} = \tau_p \mathbf{R}_{m'l} \tilde{\mathbf{F}}_{k,p}^H \Psi_{m'l}^{-1} \tilde{\mathbf{F}}_{l,p} \mathbf{R}_{m'l}$. For “ $m = m', l \notin \mathcal{P}_k$ ”, we define $\Gamma_{mkl}^{(1)} \triangleq \mathbb{E}\{\mathbf{V}_{mk}^H \mathbf{H}_{ml} \bar{\mathbf{F}}_{l,u} \mathbf{H}_{ml}^H \mathbf{V}_{mk}\} \in \mathbb{C}^{N \times N}$ with the (n, n') -th element $[\Gamma_{mkl}^{(1)}]_{nn'} = \sum_{i=1}^N \sum_{i'=1}^N [\bar{\mathbf{F}}_{l,u}]_{i'i} \mathbb{E}\{\hat{\mathbf{h}}_{mk,n}^H \mathbf{h}_{ml,i'} \mathbf{h}_{ml,i}^H \hat{\mathbf{h}}_{mk,n'}\}$ being

$$[\Gamma_{mkl}^{(1)}]_{nn'} = \sum_{i=1}^N \sum_{i'=1}^N [\bar{\mathbf{F}}_{l,u}]_{i'i} \text{tr}(\mathbb{E}\{\mathbf{h}_{ml,i'} \mathbf{h}_{ml,i}^H\} \mathbb{E}\{\hat{\mathbf{h}}_{mk,n'} \hat{\mathbf{h}}_{mk,n}^H\}) = \sum_{i=1}^N \sum_{i'=1}^N [\bar{\mathbf{F}}_{l,u}]_{i'i} \text{tr}(\mathbf{R}_{ml}^{i'i} \hat{\mathbf{R}}_{mk}^{n'n}) \quad (47)$$

since $\hat{\mathbf{H}}_{mk}$ and \mathbf{H}_{ml} are independent. Finally, for “ $m = m', l \in \mathcal{P}_k$ ”, $\hat{\mathbf{H}}_{mk}$ and \mathbf{H}_{ml} are no longer independent. We define $\Gamma_{mkl}^{(2)} \triangleq \mathbb{E}\{\mathbf{V}_{mk}^H \mathbf{H}_{ml} \bar{\mathbf{F}}_{l,u} \mathbf{H}_{ml}^H \mathbf{V}_{mk}\} \in \mathbb{C}^{N \times N}$ whose (n, n') -th element is $[\Gamma_{mkl}^{(2)}]_{nn'} = \sum_{i=1}^N \sum_{i'=1}^N [\bar{\mathbf{F}}_{l,u}]_{i'i} \mathbb{E}\{\hat{\mathbf{h}}_{mk,n}^H \mathbf{h}_{ml,i'} \mathbf{h}_{ml,i}^H \hat{\mathbf{h}}_{mk,n'}\}$. We follow the similar method in [28] and derive $[\Gamma_{kl,m}^{(2)}]_{nn'} = \sum_{i=1}^N \sum_{i'=1}^N [\bar{\mathbf{F}}_{l,u}]_{i'i} \text{tr}(\mathbf{R}_{ml}^{i'i} \mathbf{P}_{mkl,(1)}^{n'n}) + \tau_p^2 \sum_{q_1=1}^N \sum_{q_2=1}^N [\bar{\mathbf{F}}_{l,u}]_{i'i} [\text{tr}(\tilde{\mathbf{P}}_{mkl,(2)}^{q_1 n} \tilde{\mathbf{R}}_{ml}^{i' q_2} \tilde{\mathbf{R}}_{ml}^{q_2 i} \tilde{\mathbf{P}}_{mkl,(2)}^{n' q_1})] + \tau_p^2 \sum_{q_1=1}^N \sum_{q_2=1}^N [\bar{\mathbf{F}}_{l,u}]_{i'i} \text{tr}(\tilde{\mathbf{P}}_{mkl,(2)}^{q_1 n} \tilde{\mathbf{R}}_{ml}^{i' q_2}) \text{tr}(\tilde{\mathbf{P}}_{mkl,(2)}^{n' q_2} \tilde{\mathbf{R}}_{ml}^{q_2 i})$, where $\mathbf{P}_{mkl,(1)} \triangleq \tau_p \mathbf{S}_{mk} (\Psi_{mk} - \tau_p \tilde{\mathbf{F}}_{l,p} \mathbf{R}_{ml} \tilde{\mathbf{F}}_{l,p}^H) \mathbf{S}_{mk}^H$, $\mathbf{S}_{mk} \triangleq \mathbf{R}_{mk} \tilde{\mathbf{F}}_{k,p}^H \Psi_{mk}^{-1}$ and $\mathbf{P}_{mkl,(2)} \triangleq \mathbf{S}_{mk} \tilde{\mathbf{F}}_{l,p} \mathbf{R}_{ml} \tilde{\mathbf{F}}_{l,p}^H \mathbf{S}_{mk}^H$, respectively. Besides, $\tilde{\mathbf{R}}_{ml}^{ni}$ and $\tilde{\mathbf{P}}_{mkl,(2)}^{ni}$ denote (n, i) -submatrix of $\mathbf{R}_{ml}^{\frac{1}{2}}$ and $\mathbf{P}_{mkl,(2)}^{\frac{1}{2}}$, respectively.

In summary, combining all the cases, we have $\mathbb{E}\{\mathbf{G}_{kl} \bar{\mathbf{F}}_{l,u} \mathbf{G}_{kl}^H\} = \mathbf{T}_{kl,(1)} + \mathbf{T}_{kl,(2)}$ if $l \in \mathcal{P}_k$ and $\mathbf{T}_{kl,(1)}$ otherwise, where $\mathbf{T}_{kl,(1)} \triangleq \text{diag}(\Gamma_{kl,1}^{(1)}, \dots, \Gamma_{kl,M}^{(1)}) \in \mathbb{C}^{MN \times MN}$ and $\mathbf{T}_{kl,(2)} = \Gamma_{kl,m}^{(2)} - \Gamma_{kl,m}^{(1)}$ if $m = m'$ and $\Lambda_{mkl} \bar{\mathbf{F}}_{l,u} \Lambda_{m'lk}$ otherwise. Plugging the derived results into (28) and (31), we can easily compute the optimal LSFDF coefficient matrix and MSE matrix in closed-form as (34). So we have finished the proof of Theorem 2. For more details on the derived expression, please refer to [28, Appendix D].

APPENDIX D

PROOF OF THEOREM 1

When other optimization variables are fixed, we derive the partial derivative of (17) w.r.t $\mathbf{F}_{k,u}^{(1)}$ as

$$\begin{aligned} \frac{\partial f(\mathbf{F}_{1,u,(1)}, \dots, \mathbf{F}_{K,u,(1)})}{\partial \mathbf{F}_{k,u,(1)}} &= \sum_{l=1}^K \mu_{l,(1)} \left(\hat{\mathbf{H}}_k^H \mathbf{V}_l \mathbf{W}_{l,(1)} \mathbf{V}_l^H \hat{\mathbf{H}}_k + \mathbb{E}\left\{ \tilde{\mathbf{H}}_k^H \mathbf{V}_l \mathbf{W}_{l,(1)} \mathbf{V}_l^H \tilde{\mathbf{H}}_k \middle| \mathbf{V}, \mathbf{W} \right\} \right) + \lambda_{k,(1)} \mathbf{I}_N \\ &\quad - \mu_{k,(1)} \hat{\mathbf{H}}_k^H \mathbf{V}_k^H \mathbf{W}_{k,(1)}. \end{aligned} \quad (48)$$

By applying the first-order optimality condition and setting $\frac{\partial f(\mathbf{F}_{1,u,(1)}, \dots, \mathbf{F}_{K,u,(1)})}{\partial \mathbf{F}_{k,u,(1)}} = 0$, we can easily obtain the optimal precoding scheme. Besides, $\lambda_{k,(1)}$ and $\mathbf{F}_{k,u,(1)}$ should also satisfy KKT condition as (19).

As for $\bar{\mathbf{C}}_{kl} \triangleq \mathbb{E}\{\tilde{\mathbf{H}}_k^H \mathbf{V}_l \mathbf{W}_{l,(1)} \mathbf{V}_l^H \tilde{\mathbf{H}}_k | \mathbf{V}, \mathbf{W}\} \in \mathbb{C}^{N \times N}$, by applying Lemma 1, the (i, n) -th element of $\bar{\mathbf{C}}_{kl}$ is $\text{tr}(\bar{\mathbf{V}}_l \mathbb{E}\{\tilde{\mathbf{h}}_{k,n} \tilde{\mathbf{h}}_{k,i}^H\})$ where $\bar{\mathbf{V}}_l \triangleq \mathbf{V}_l \mathbf{W}_{l,(1)} \mathbf{V}_l^H$ and $\tilde{\mathbf{h}}_{k,n} = [\tilde{\mathbf{h}}_{1k,n}^T, \dots, \tilde{\mathbf{h}}_{Mk,n}^T]^T \in \mathbb{C}^{ML}$ is the n -th column of $\tilde{\mathbf{H}}_k$. Finally, we derive $\mathbf{C}_{k,ni} \triangleq \mathbb{E}\{\tilde{\mathbf{h}}_{k,n} \tilde{\mathbf{h}}_{k,i}^H\} = \text{diag}(\mathbf{C}_{1k}^{ni}, \dots, \mathbf{C}_{Mk}^{ni}) \in \mathbb{C}^{ML \times ML}$ since $\tilde{\mathbf{h}}_{mk,n}$

and $\tilde{\mathbf{h}}_{m'k,n}$ for $m \neq m'$ are independent and both have zero mean. So $\mathbf{C}_{k,ni}$ is a block-diagonal matrix with the square matrices $\mathbf{C}_{1k}^{ni} = \mathbb{E}\{\tilde{\mathbf{h}}_{1k,n}\tilde{\mathbf{h}}_{1k,i}^H\}, \dots, \mathbf{C}_{Mk}^{ni} = \mathbb{E}\{\tilde{\mathbf{h}}_{Mk,n}\tilde{\mathbf{h}}_{Mk,i}^H\}$ on the diagonal.

APPENDIX E

PROOF OF (15)

For the LSFD scheme, the conditional MSE matrix for UE k can be written as (26). Based on [28, Appendix C], we prove that (28) can also minimize $\text{MSE}_{k,(2)} = \text{tr}(\mathbf{E}_{k,(2)})$. With (28) implemented, $\mathbf{E}_{k,(2)}$ is given by (31). Then, by applying Lemma 2, we have

$$\begin{aligned} \left(\mathbf{E}_{k,(2)}^{\text{opt}}\right)^{-1} &= \mathbf{I}_N + \mathbf{F}_{k,u,(2)}^H \mathbb{E}\{\mathbf{G}_{kk}^H\} \left(\sum_{l=1}^K \mathbb{E}\{\mathbf{G}_{kl}\bar{\mathbf{F}}_{l,u,(2)}\mathbf{G}_{kl}^H\} - \mathbb{E}\{\mathbf{G}_{kk}\}\bar{\mathbf{F}}_{k,u,(2)} \mathbb{E}\{\mathbf{G}_{kk}^H\} + \sigma^2\mathbf{S}_k \right)^{-1} \\ &\quad \times \mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u,(2)}, \end{aligned}$$

where $\mathbf{A} \triangleq \mathbf{I}_N$, $\mathbf{B} \triangleq -\mathbf{F}_{k,u,(2)}^H \mathbb{E}\{\mathbf{G}_{kk}^H\}$, $\mathbf{C} \triangleq (\sum_{l=1}^K \mathbb{E}\{\mathbf{G}_{kl}\bar{\mathbf{F}}_{l,u,(2)}\mathbf{G}_{kl}^H\} + \sigma^2\mathbf{S}_k)^{-1}$ and $\mathbf{D} \triangleq \mathbb{E}\{\mathbf{G}_{kk}\}\mathbf{F}_{k,u,(2)}$, respectively. We show the equivalence between $\text{SE}_{k,(2)}^{\text{opt}}$ and $\log_2 |(\mathbf{E}_{k,(2)}^{\text{opt}})^{-1}|$ without a factor $(1 - \tau_p/\tau_c)$.

APPENDIX F

PROOF OF THEOREM 4

When MR combining $\mathbf{V}_{mk} = \hat{\mathbf{H}}_{mk}$ and the optimal LSFD scheme applied, we can easily compute $\mathbb{E}\{\mathbf{G}_{kk}^H\}$, $\mathbf{A}_k^{\text{opt}}$, and $\mathbf{E}_{k,(2)}^{\text{opt}}$ in closed-form as Theorem 2. Furthermore, by applying Lemma 1, the (i, n) -th entry of $\bar{\mathbf{T}}_{lk} = \mathbb{E}\{\mathbf{G}_{lk}^H \mathbf{A}_l \mathbf{E}_{l,(2)}^{-1} \mathbf{A}_l^H \mathbf{G}_{lk}\} \in \mathbb{C}^{N \times N}$ can be denoted as $\text{tr}(\bar{\mathbf{A}}_l \mathbb{E}\{\mathbf{g}_{lk,n} \mathbf{g}_{lk,i}^H\})$, where $\bar{\mathbf{A}}_l \triangleq \mathbf{A}_l \mathbf{E}_{l,(2)}^{-1} \mathbf{A}_l^H$ and $\mathbf{g}_{lk,n} \in \mathbb{C}^{MN}$ is the n -th column of \mathbf{G}_{lk} . Note that the $(m-1)N + p$ -th element of $\mathbf{g}_{lk,n}$ is $\hat{\mathbf{h}}_{ml,p}^H \mathbf{h}_{mk,n}$ so the $[(m-1)N + p, (m'-1)N + p']$ -th (or $[o, j]$ -th briefly) entry of $\bar{\mathbf{G}}_{lk,ni} \triangleq \mathbb{E}\{\mathbf{g}_{lk,n} \mathbf{g}_{lk,i}^H\} \in \mathbb{C}^{MN \times MN}$ can be denoted as $\mathbb{E}\{\hat{\mathbf{h}}_{ml,p}^H \mathbf{h}_{mk,n} \mathbf{h}_{m'k,i}^H \hat{\mathbf{h}}_{m'l,p'}\}$, which can be computed for four AP-UE combinations as Theorem 2.

For “ $l \notin \mathcal{P}_k, m \neq m'$ ”, we have $\mathbb{E}\{\hat{\mathbf{h}}_{ml,p}^H \mathbf{h}_{mk,n} \mathbf{h}_{m'k,i}^H \hat{\mathbf{h}}_{m'l,p'}\} = 0$. For “ $l \in \mathcal{P}_k, m \neq m'$ ”, we have $\mathbb{E}\{\hat{\mathbf{h}}_{ml,p}^H \mathbf{h}_{mk,n} \mathbf{h}_{m'k,i}^H \hat{\mathbf{h}}_{m'l,p'}\} = \text{tr}(\mathbf{R}_{mk}^{ni} \hat{\mathbf{R}}_{ml}^{p'p})$. For “ $l \notin \mathcal{P}_k, m = m'$ ”, we have $\mathbb{E}\{\hat{\mathbf{h}}_{ml,p}^H \mathbf{h}_{mk,n} \mathbf{h}_{m'k,i}^H \hat{\mathbf{h}}_{m'l,p'}\} = \mathbb{E}\{\hat{\mathbf{h}}_{ml,p}^H \mathbf{h}_{mk,n}\} \mathbb{E}\{\mathbf{h}_{m'k,i}^H \hat{\mathbf{h}}_{m'l,p'}\} = \text{tr}(\Xi_{mlk}^{np}) \text{tr}(\Xi_{m'kl}^{p'i})$, where $\Xi_{mlk} = \tau_p \mathbf{R}_{mk} \tilde{\mathbf{F}}_{k,p}^H \Psi_{mk}^{-1} \tilde{\mathbf{F}}_{l,p} \mathbf{R}_{ml}$ and $\Xi_{m'kl} = \tau_p \mathbf{R}_{m'l} \tilde{\mathbf{F}}_{l,p}^H \Psi_{m'l}^{-1} \tilde{\mathbf{F}}_{k,p} \mathbf{R}_{m'k}$. For “ $l \in \mathcal{P}_k, m = m'$ ”, we obtain $\mathbb{E}\{\hat{\mathbf{h}}_{ml,p}^H \mathbf{h}_{mk,n} \mathbf{h}_{m'k,i}^H \hat{\mathbf{h}}_{m'l,p'}\} = \text{tr}(\mathbf{R}_{mk}^{ni} \mathbf{P}_{mkl,(1)}^{p'p}) + \tau_p^2 \sum_{q_1=1}^N \sum_{q_2=1}^N \text{tr}(\tilde{\mathbf{P}}_{mlk,(2)}^{q_1p} \tilde{\mathbf{R}}_{mk}^{nq_2} \tilde{\mathbf{R}}_{mk}^{q_2i} \tilde{\mathbf{P}}_{mlk,(2)}^{p'q_1}) + \tau_p^2 \sum_{q_1=1}^N \sum_{q_2=1}^N \text{tr}(\tilde{\mathbf{P}}_{mlk,(2)}^{q_1n} \tilde{\mathbf{R}}_{mk}^{nq_1}) \text{tr}(\tilde{\mathbf{P}}_{mlk,(2)}^{p'q_2} \tilde{\mathbf{R}}_{mk}^{q_2i})$, where $\mathbf{S}_{ml} = \mathbf{R}_{ml} \tilde{\mathbf{F}}_{l,p}^H \Psi_{ml}^{-1}$, $\mathbf{P}_{mkl,(1)} = \tau_p \mathbf{S}_{ml} (\Psi_{ml} - \tau_p \tilde{\mathbf{F}}_{k,p} \mathbf{R}_{mk} \tilde{\mathbf{F}}_{k,p}^H) \mathbf{S}_{ml}^H$ and $\mathbf{P}_{mkl,(2)} = \mathbf{S}_{ml} \tilde{\mathbf{F}}_{k,p} \mathbf{R}_{mk} \tilde{\mathbf{F}}_{k,p}^H \mathbf{S}_{ml}^H$ with $\tilde{\mathbf{R}}_{mk}^{ni}$ and $\tilde{\mathbf{P}}_{mkl,(2)}^{ni}$ being (n, i) -submatrix of $\mathbf{R}_{mk}^{\frac{1}{2}}$ and $\mathbf{P}_{mkl,(2)}^{\frac{1}{2}}$, respectively. We can compute $\mathbb{E}\{\mathbf{g}_{lk,n} \mathbf{g}_{lk,i}^H\}$ o_j in closed-form as (42) and $\mathbf{F}_{k,u,(2)}^{\text{opt}}$ in closed-form as (43).

REFERENCES

- [1] Z. Wang, J. Zhang, H. Q. Ngo, B. Ai, and M. Debbah, "Iteratively weighted MMSE uplink precoding for cell-free massive MIMO," in *Proc. IEEE ICC*, May 2022, pp. 1–6.
- [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] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5G," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1637–1660, Jun. 2020.
- [4] S. Chen, J. Zhang, J. Zhang, E. Björnson, and B. Ai, "A survey on user-centric cell-free massive MIMO systems," *Digit. Commun. Netw.*, 2021.
- [5] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 77–90, Jan. 2019.
- [6] S. Elhoushy, M. Ibrahim, and W. Hamouda, "Cell-free massive MIMO: A survey," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 492–523, 1st Quart. 2022.
- [7] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-free massive MIMO: A new next-generation paradigm," *IEEE Access*, vol. 7, pp. 99 878–99 888, Jul. 2019.
- [8] H. Q. Ngo, L. 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. 2018.
- [9] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4247–4261, Jul. 2020.
- [10] S. Chen, J. Zhang, E. Björnson, J. Zhang, and B. Ai, "Structured massive access for scalable cell-free massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1086–1100, Aug. 2021.
- [11] A. Papazafeiropoulos, E. Björnson, P. Kourtessis, S. Chatzinotas, and J. M. Senior, "Scalable cell-free massive MIMO systems: Impact of hardware impairments," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 9701–9715, Oct. 2021.
- [12] J. Zheng, J. Zhang, E. Björnson, Z. Li, and B. Ai, "Cell-free massive MIMO-OFDM for high-speed train communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 2823–2839, Oct. 2022.
- [13] M. Attarifar, A. Abbasfar, and A. Lozano, "Subset MMSE receivers for cell-free networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 4183–4194, Jun. 2020.
- [14] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and B. D. Rao, "Performance of cell-free massive MIMO systems with MMSE and LSFD receivers," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 203–207.
- [15] Ö. Özdogan, E. Björnson, and J. Zhang, "Performance of cell-free massive MIMO with Rician fading and phase shifts," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5299–5315, Nov. 2019.
- [16] Z. Wang, J. Zhang, E. Björnson, and B. Ai, "Uplink performance of cell-free massive MIMO over spatially correlated Rician fading channels," *IEEE Commun. Lett.*, vol. 25, no. 4, pp. 1348–1352, Apr. 2021.
- [17] J. Zhang, J. Zhang, E. Björnson, and B. Ai, "Local partial zero-forcing combining for cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8459–8473, Dec. 2021.
- [18] G. Interdonato, P. Frenger, and E. G. Larsson, "Scalability aspects of cell-free massive MIMO," in *Proc. IEEE ICC*, 2019, pp. 1–6.
- [19] Z. H. Shaik, E. Björnson, and E. G. Larsson, "MMSE-optimal sequential processing for cell-free massive MIMO with radio stripes," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7775–7789, Nov. 2021.

- [20] H. Masoumi and M. J. Emadi, "Performance analysis of cell-free massive MIMO system with limited fronthaul capacity and hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1038–1053, Feb. 2020.
- [21] M. Bashar, P. Xiao, R. Tafazolli, K. Cumanan, A. G. Burr, and E. Björnson, "Limited-fronthaul cell-free massive MIMO with local MMSE receiver under Rician fading and phase shifts," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1934–1938, Sep. 2021.
- [22] T. Van Chien, H. Q. Ngo, S. Chatzinotas, M. Di Renzo, and B. Ottersten, "Reconfigurable intelligent surface-assisted cell-free massive MIMO systems over spatially-correlated channels," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5106–5128, Jul. 2022.
- [23] E. Shi, J. Zhang, S. Chen, J. Zheng, Y. Zhang, D. W. Kwan Ng, and B. Ai, "Wireless energy transfer in RIS-aided cell-free massive MIMO systems: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 60, no. 3, pp. 26–32, Mar. 2022.
- [24] T. C. Mai, H. Q. Ngo, and T. Q. Duong, "Cell-free massive MIMO systems with multi-antenna users," in *IEEE GlobSIP*, Nov. 2018, pp. 828–832.
- [25] S. Buzzi, C. D'Andrea, A. Zappone, and C. D'Elia, "User-centric 5G cellular networks: Resource allocation and comparison with the cell-free massive MIMO approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1250–1264, Feb. 2020.
- [26] T. C. Mai, H. Q. Ngo, and T. Q. Duong, "Downlink spectral efficiency of cell-free massive MIMO systems with multi-antenna users," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4803–4815, Apr. 2020.
- [27] M. Zhou, L. Yang, and H. Zhu, "Sum-SE for multigroup multicast cell-free massive MIMO with multi-antenna users and low-resolution DACs," *IEEE Wireless Commun. Lett.*, vol. 10, no. 8, pp. 1702–1706, May 2021.
- [28] Z. Wang, J. Zhang, B. Ai, C. Yuen, and M. Debbah, "Uplink performance of cell-free massive MIMO with multi-antenna users over jointly-correlated Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7391–7406, Sep. 2022.
- [29] W. Weichselberger, M. Herdin, H. Ozelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.
- [30] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Apr. 2011.
- [31] J. Shin and J. Moon, "Weighted-sum-rate-maximizing linear transceiver filters for the K-user MIMO interference channel," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 2776–2783, Sep. 2012.
- [32] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, 2008.
- [33] X. Li, X. Zhang, Y. Zhou, and L. Hanzo, "Optimal massive-MIMO-Aided clustered base-station coordination," *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2699–2712, Feb. 2021.
- [34] H. Q. Ngo, H. Tataria, M. Matthaiou, S. Jin, and E. G. Larsson, "On the performance of cell-free massive MIMO in Rician fading," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Oct. 2018, pp. 980–984.
- [35] E. Björnson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbance," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1807–1820, Nov. 2010.
- [36] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.
- [37] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.
- [38] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, M. Debbah, and P. Xiao, "Max-min rate of cell-free massive MIMO uplink with optimal uniform quantization," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6796–6815, Oct. 2019.
- [39] B. Hu, C. Hua, C. Chen, and X. Guan, "Joint beamformer design for wireless fronthaul and access links in C-RANs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2869–2881, May 2018.

- [40] J. Hoydis, S. ten Brink, and M. Debbah, “Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [41] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, “Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback,” *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, Jul. 2012.
- [42] J. Zhang, J. Zhang, D. W. K. Ng, S. Jin, and B. Ai, “Improving sum-rate of cell-free massive MIMO with expanded compute-and-forward,” *IEEE Trans. Signal Process.*, vol. 70, pp. 202–215, 2021.
- [43] A. Tulino, A. Lozano, and S. Verdu, “Impact of antenna correlation on the capacity of multiantenna channels,” *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jun. 2005.
- [44] Y. Wu, S. Ma, and Y. Gu, “A unified framework of non-orthogonal pilot design for multi-cell massive MIMO systems,” *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7623–7633, 2020.
- [45] A. Hjørungnes, *Complex-valued matrix derivatives: with applications in signal processing and communications*. Cambridge University Press, 2011.