

Research Article

Downlink Linear Precoders Based on Statistical CSI for Multicell MIMO-OFDM

Ebrahim Baktash,¹ Chi-Heng Lin,² Xiaodong Wang,³ and Mahmood Karimi¹

¹School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran

²Department of Statistics, Columbia University, New York, NY 10027, USA

³Department of Electrical Engineering, Columbia University, New York, NY 10027, USA

Correspondence should be addressed to Xiaodong Wang; wangx@ee.columbia.edu

Received 21 September 2017; Accepted 6 December 2017; Published 31 December 2017

Academic Editor: Mostafa Zaman Chowdhury

Copyright © 2017 Ebrahim Baktash et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With 5G communication systems on the horizon, efficient interference management in heterogeneous multicell networks is more vital than ever. This paper investigates the linear precoder design for downlink multicell multiple-input multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) systems, where base stations (BSs) coordinate to reduce the interference across space and frequency. In order to minimize the overall feedback overhead in next-generation systems, we consider precoding schemes that require statistical channel state information (CSI) only. We apply the random matrix theory to approximate the ergodic weighted sum rate of the system with a closed form expression. After formulating the approximation for general channels, we reduce the results to a more compact form using the Kronecker channel model for which several multicarrier concepts such as frequency selectivity, channel tap correlations, and intercarrier interference (ICI) are rigorously represented. We find the local optimal solution for the maximization of the approximate rate using a gradient method that requires only the covariance structure of the MIMO-OFDM channels. Within this covariance structure are the channel tap correlations and ICI information, both of which are taken into consideration in the precoder design. Simulation results show that the rate approximation is very accurate even for very small MIMO-OFDM systems and the proposed method converges rapidly to a near-optimal solution that competes with networked MIMO and precoders based on instantaneous full CSI.

1. Introduction

Multicell multiple-input multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) is a promising technology for next-generation telecommunication networks. Both MIMO and OFDM are known to facilitate the wireless networks with excellent capabilities; when MIMO and OFDM are combined together, MIMO-OFDM achieves tremendous capacities that will be enjoyed by 5G systems in the near future. While MIMO-OFDM systems have superb performance, several challenges invoked by MIMO and OFDM characteristics are involved in the system design [1–6]. Other competing technologies in 5G networks are filtered-OFDM (F-OFDM) and universal filtered multicarrier (UFMC) that are receiving attention in the literature [7, 8].

MIMO wireless cellular systems can achieve excellent data-rates if proper coordination among base stations (BSs) is

employed to suppress the intercell and intracell interference. The highest system capacity is achieved by networked MIMO where both the channel state information (CSI) and the data streams of all users in different cells are made available at all BSs [9, 10]. A more practical and less complex form of coordination is through joint linear precoding with only the CSI of the users shared among BSs [11–13]. In this case, each BS obtains the CSI of its intracell users and shares it with its adjacent BSs. Such precoding approach is very efficient for small number of BS antennas and it can alleviate the interference to a significant degree. As the number of BS antennas increases, the amount of signaling overhead for passing the CSI among BSs becomes taxing [14]. This problem is aggravated in frequency-selective channels employing OFDM where the CSI is different on each subcarrier. To overcome this difficulty, the precoders can be designed based on statistical CSI at the transmitter rather than instantaneous CSI [3, 15–20]. Since

the statistical CSI changes much slower than instantaneous CSI, the BSs need to be updated less frequently, hence, much less signaling overhead. While there is a performance loss due to the lack of exact CSI, precoder design based on statistical CSI for large-scale MIMO systems is an efficient approach because the performance loss is negligible [18, 20].

In order to design MIMO precoders with statistical CSI, the ergodic rate, which is a function of the covariance matrices of the MIMO channels, must be maximized over the precoders [16]. To facilitate the maximization, the ergodic rate can be approximated by a compact deterministic expression when the MIMO system is large-scale. The optimization can then be carried out on this deterministic approximation which is a function of the statistics of the channels [20–23]; thus, the solution is a set of precoders that are functions of the statistical CSI. Deterministic approximation to the Shannon rate can be achieved based on techniques such as Bai and Silverstein [22, 24, 25], Gaussian method [26, 27], Lindeberg principle [20, 26, 28, 29], or a combination of them [22, Theorem 6.9], [26]. All of these methods approximate the Shannon transform of large dimensional random matrices, but they are based on slightly different assumptions.

The deterministic equivalents of the rates for large-scale MIMO systems that are derived in recent works [20–23, 25, 26, 29] are functions of the channel statistics, so they can be used as objective functions for precoder design with statistical CSI. This has been done in [21–23] where closed form optimal precoders are derived for simple single-user scenarios. In [20], an uplink multiuser multicell MIMO system is considered and suboptimal linear precoders are found based on the Lindeberg approximation to the ergodic rate. We shall extend this methodology to the problem of downlink multiuser multicell MIMO-OFDM linear precoder design with statistical CSI. Considering the downlink leads to a somewhat different power constraint compared to the uplink, and the OFDM assumption gives rise to concepts such as frequency selectivity [17, 30, 31], channel tap correlations [17, 31–34], and intercarrier interference (ICI) [1, 35–38], all of which must be taken into account in the precoder design. Our derivation of statistical precoding for the downlink multicell MIMO-OFDM system is a unification of previous works [20, 29], and while several details are different from these two works, this derivation is not the main contribution. The main point of this work is applying the already established results of random matrix theory to the MIMO-OFDM scenario in order to study the impact of frequency selectivity, tap correlations, and ICI on the statistical precoder design and system performance.

We shall study the downlink multicell MIMO-OFDM linear precoder design with statistical CSI using the deterministic approximation of the Shannon rate, a problem studied before only under instantaneous CSI [12]. In particular, with instantaneous CSI, the channel statistics are not used in precoder design and one may try to optimize the instantaneous rate, as is done in [12] where it has been shown that such an optimization problem is nonconvex and a suboptimal solution has been proposed using the Karush-Kuhn-Tucker (KKT) conditions. In this work, however, we assume that there is only statistical CSI available at the BSs and we design

linear precoders that maximize the ergodic weighted sum rate. As opposed to the instantaneous case, for analysis of the ergodic rate, we need a statistical model for the MIMO channel which is usually determined by a covariance structure [23, 39]. We first consider the general correlated channel model with arbitrary probability distribution described in [23] that includes several statistical models such as the independent and identically distributed (i.i.d.) Rayleigh fading channel and the Kronecker channel [39]. Then we reduce the results to a simpler form for the Kronecker channel model.

To maximize the ergodic rate, inspired by [20, 23, 29], we find a deterministic approximation to the Shannon rate of the MIMO-OFDM system using the random matrix theory [21, 40]. The methodology will be applied to the MIMO-OFDM system with tap correlations and intercarrier interference. Then, assuming that the BSs have the statistical CSI of all users in the form of covariance matrices, we form a weighted sum rate maximization problem and propose a suboptimal solution using the KKT conditions along with the gradient descent method. Our simulations show that the approximation is quite accurate even for small size MIMO-OFDM systems and the proposed algorithm converges rapidly to a maximum which has a substantial improvement over isotropic precoding. The results only slightly underperform those obtained with perfect instantaneous CSI [12]. We then extend to the case where the frequency-selective channel suffers from correlation among channel taps and we show that the precoders become frequency dependent under tap correlations. We study the effects of tap correlation on the precoder design and system performance. Next, we allow ICI among OFDM subcarriers which is caused by carrier frequency offset due to synchronization errors and Doppler shifts [1]. The ICI introduces a new source of interference in addition to the intercell and intracell interference. We find the deterministic approximation to the rate under ICI and then study its impact on the precoder design and system performance. Our simulations show that while spatial correlations, tap correlations, and ICI decrease the system sum rate, our method alleviates this performance loss by incorporating the correlation information and ICI intensity information into the precoder design. It must be noted that our statistical CSI based method is applicable to networked MIMO with full cooperation where the BSs share the channel statistics and the transmit data.

The remainder of the paper is organized as follows. In Section 5, we provide the system model and formulate the optimization problem. In Section 3, we obtain the deterministic approximation to the ergodic rate function. In Section 4, we give a gradient-descent-based algorithm to obtain the suboptimal precoders. In Section 5, we discuss the channel model and how the formulas simplify for the separable channel models. In Section 6, we extend to the case where there is ICI. In Section 7, we present simulation results, and finally, conclusions are given in Section 8.

We denote matrices, vectors, and scalars by upper-case bold letters as in \mathbf{X} , lower-case bold letters as in \mathbf{x} , and non-bold letters as in X and x , respectively. Moreover, \mathbf{X}^* , \mathbf{X}^T , and \mathbf{X}^H denote conjugate, transpose, and conjugate-transpose, respectively. The element on the i th row and j th column of

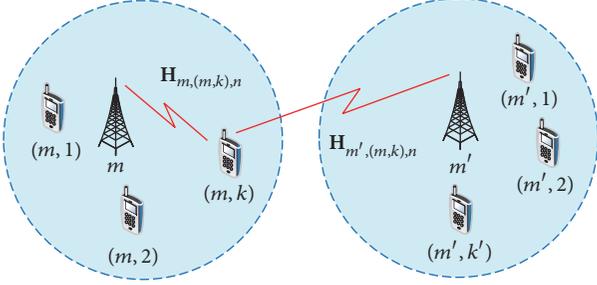


FIGURE 1: Multicell MIMO-OFDM system model.

a matrix is denoted by $[\mathbf{X}]_{i,j}$, and the i th element of a vector is denoted by $[\mathbf{x}]_i$. Vectorization, trace, and expected value operators are denoted by $\text{vec}\{\cdot\}$, $\text{tr}\{\cdot\}$, and $\mathbb{E}\{\cdot\}$, respectively. The all-one and all-zero vectors of size N and identity matrix of size $N \times N$ are denoted by $\mathbf{1}_N$, $\mathbf{0}_N$, and \mathbf{I}_N , respectively. The operators $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ represent real part and imaginary part, respectively.

2. System Model and Problem Formulation

2.1. System Model. We consider a downlink multicell MIMO-OFDM wireless network shown in Figure 1, with M BSs each serving K users over N subcarriers. All BSs have N_B transmitting antennas while each user has N_U receiving antennas with $N_B \geq N_U$. The k th user under BS m is denoted by (m, k) . The data vector to be precoded and transmitted by BS m to its k th user, that is, the data vector for user (m, k) , over subcarrier n is denoted by $\mathbf{x}_{(m,k),n} \in \mathbb{C}^{N_U}$, and the corresponding linear precoding matrix is denoted by $\mathbf{F}_{(m,k),n} \in \mathbb{C}^{N_B \times N_U}$. The downlink channel matrix from BS m' to user (m, k) over subcarrier n is denoted by $\mathbf{H}_{m',(m,k),n} \in \mathbb{C}^{N_U \times N_B}$. The thermal noise at user (m, k) on subcarrier n is denoted by $\mathbf{z}_{(m,k),n} \in \mathbb{C}^{N_U}$. Based on the above notations, the received signal at user (m, k) on subcarrier n is given by

$$\begin{aligned} \mathbf{y}_{(m,k),n} &= \mathbf{H}_{m,(m,k),n} \mathbf{F}_{(m,k),n} \mathbf{x}_{(m,k),n} \\ &+ \sum_{k' \neq k}^K \mathbf{H}_{m,(m,k),n} \mathbf{F}_{(m,k'),n} \mathbf{x}_{(m,k'),n} \\ &+ \sum_{m' \neq m}^M \sum_{k'=1}^K \mathbf{H}_{m',(m,k),n} \mathbf{F}_{(m',k'),n} \mathbf{x}_{(m',k'),n} \\ &+ \mathbf{z}_{(m,k),n}, \end{aligned} \quad (1)$$

where the second and third terms represent the intracell and intercell interference, respectively. We assume that $\mathbf{x}_{(m,k),n}$ and $\mathbf{z}_{(m,k),n}$ are i.i.d. Gaussian with $\mathbb{E}[\mathbf{x}_{(m,k),n} \mathbf{x}_{(m,k),n}^H] = \mathbf{I}$ and $\mathbb{E}[\mathbf{z}_{(m,k),n} \mathbf{z}_{(m,k),n}^H] = \mathbf{I}$. We also assume that MIMO channels are independent across users and BSs.

Each user is assumed to know its own instantaneous CSI for detecting the data [16, 20, 22, 23, 29], which leads to the

instantaneous data-rate for user (m, k) on subcarrier n given by

$$\begin{aligned} r_{(m,k),n} &= \log \left| \mathbf{I} + \mathbf{H}_{m,(m,k),n} \mathbf{F}_{(m,k),n} \mathbf{F}_{(m,k),n}^H \mathbf{H}_{m,(m,k),n}^H \mathbf{\Theta}_{(m,k),n}^{-1} \right|, \end{aligned} \quad (2)$$

where $\mathbf{\Theta}_{(m,k),n}$ denotes the total interference plus noise on subcarrier n given by

$$\begin{aligned} \mathbf{\Theta}_{(m,k),n} &= \mathbf{I} + \sum_{k' \neq k}^K \mathbf{H}_{m,(m,k),n} \mathbf{F}_{(m,k'),n} \mathbf{F}_{(m,k'),n}^H \mathbf{H}_{m,(m,k),n}^H \\ &+ \sum_{m' \neq m}^M \sum_{k'=1}^K \mathbf{H}_{m',(m,k),n} \mathbf{F}_{(m',k'),n} \mathbf{F}_{(m',k'),n}^H \mathbf{H}_{m',(m,k),n}^H, \end{aligned} \quad (3)$$

where the first, second, and third terms represent the noise power, intracell interference power, and intercell interference power, respectively.

2.2. Problem Formulation. Our goal is to design precoder matrices based only on the second-order statistics of the MIMO channels. While the second-order statistics fully describe the statistical CSI for zero mean Gaussian channels, it is only partial statistical CSI for non-Gaussian channels. Defining $\mathcal{F} = \{\mathbf{F}_{(m,k),n}\}$ as the set of precoder matrices, we formulate the weighted sum rate optimization problem as

$$\begin{aligned} \max_{\mathcal{F}} \quad & \mathbb{E}_{\mathbf{H}} \left[\sum_{m=1}^M \sum_{k=1}^K w_{(m,k)} \sum_{n=1}^N r_{(m,k),n} \right] \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{n=1}^N \text{tr}(\mathbf{F}_{(m,k),n} \mathbf{F}_{(m,k),n}^H) = P_m, \quad \forall m, \end{aligned} \quad (\text{P1})$$

where $\mathbb{E}_{\mathbf{H}}[\cdot]$ denotes expectation over the channel matrices $\{\mathbf{H}_{m',(m,k),n}\}$ and $w_{(m,k)}$ is the maximization weight associated with user (m, k) . The constraint is due to the fact that BS m is subject to the transmit power limit of P_m . Note that the power constraint introduces coupling across subcarriers, while the rate function given by (2) is independent for each subcarrier and coupled across users and BSs only, a feature of OFDM systems.

Define the covariance matrix of $\mathbf{H}_{m',(m,k),n}$ as

$$\mathbf{\Omega}_{m',(m,k),n} = \mathbb{E}_{\mathbf{H}} \left[\text{vec} \{ \mathbf{H}_{m',(m,k),n} \} \text{vec}^H \{ \mathbf{H}_{m',(m,k),n} \} \right]. \quad (4)$$

Unlike designing precoders based instantaneous CSI as in [12], the precoders in \mathcal{F} are restricted to be functions of $\{\mathbf{\Omega}_{m',(m,k),n}\}$. The above approach to precoder design has several advantages. It is very practical since acquiring full CSI for all $m', (m, k), n$ induces huge communication overhead due to CSI exchanges. Also, since the higher-order statistics can be costly to obtain, the second-order statistics often serve as the minimal statistics for various kinds of estimations.

In order to solve the nonconvex problem (P1), we need to find the expectation $\mathbb{E}_{\mathbf{H}}[r_{(m,k),n}]$. However, this expectation

makes the problem hard to analyse even for the single-user case [16]. Therefore, it is useful to replace the objective function by the analytically tractable approximation $\tilde{r}_{(m,k),n} \approx \mathbb{E}_{\mathbf{H}}[r_{(m,k),n}]$. Thanks to the random matrix theory, as will be shown in the next section, we can find the deterministic approximation $\tilde{r}_{(m,k),n}$ that allows analytical expressions for the ergodic rate [20–23, 29]. The approximation is appropriate for our problem as $\tilde{r}_{(m,k),n}$ is a function of second-order statistics. While we have asymptotic convergence $\tilde{r}_{(m,k),n} \rightarrow \mathbb{E}_{\mathbf{H}}[r_{(m,k),n}]$ when the number of BS antennas is large, our simulations show that the approximation is also accurate for MIMO-OFDM systems with very small number of antennas.

We now replace the ergodic rate $\mathbb{E}_{\mathbf{H}}[r_{(m,k),n}]$ with its deterministic approximation $\tilde{r}_{(m,k),n}$ and form the analytically tractable problem

$$\begin{aligned} \max_{\mathcal{F}} \quad & \sum_{m=1}^M \sum_{k=1}^K w_{(m,k)} \sum_{n=1}^N \tilde{r}_{(m,k),n} \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{n=1}^N \text{tr}(\mathbf{F}_{(m,k),n} \mathbf{F}_{(m,k),n}^H) = P_m, \quad \forall m. \end{aligned} \quad (\text{P2})$$

While analytically tractable, problem (P2) is still nonconvex. Since the global optimum to a nonconvex optimization problem is generally hard to obtain, we will develop an algorithm to find the local optimal solution to (P2). In the next sections, we will first derive the asymptotic approximation $\tilde{r}_{(m,k),n}$ and then propose an algorithm to obtain locally optimum precoder matrices.

3. Asymptotic Approximation of the Rate

In this section we derive the asymptotic approximation to the ergodic rate function. From (2) and (3), it follows that the rate function can be rewritten as

$$\begin{aligned} r_{(m,k),n} = & +\log \left| \mathbf{I} + \sum_{m'} \sum_{k'} \mathbf{H}_{m',(m,k),n} \right. \\ & \cdot \mathbf{F}_{(m',k'),n} \mathbf{F}_{(m',k'),n}^H \mathbf{H}_{m',(m,k),n}^H \left. - \log \left| \mathbf{I} \right. \right. \\ & \left. \left. + \sum_{(m',k') \neq (m,k)} \mathbf{H}_{m',(m,k),n} \mathbf{F}_{(m',k'),n} \mathbf{F}_{(m',k'),n}^H \mathbf{H}_{m',(m,k),n}^H \right. \right. \end{aligned} \quad (5)$$

To express the rate more compactly, define

$$\begin{aligned} \mathbf{A}_{(m,k),n} &= [\mathbf{H}_{1,(m,k),n} \bar{\mathbf{F}}_{1,n}, \dots, \mathbf{H}_{m,(m,k),n} \bar{\mathbf{F}}_{m,n}, \dots, \\ & \quad \mathbf{H}_{M,(m,k),n} \bar{\mathbf{F}}_{M,n}], \\ \mathbf{B}_{(m,k),n} &= [\mathbf{H}_{1,(m,k),n} \check{\mathbf{F}}_{1,n}, \dots, \mathbf{H}_{m,(m,k),n} \check{\mathbf{F}}_{m,k,n}, \dots, \\ & \quad \mathbf{H}_{M,(m,k),n} \bar{\mathbf{F}}_{M,n}], \end{aligned} \quad (6)$$

in which

$$\begin{aligned} \bar{\mathbf{F}}_{m',n} &= [\mathbf{F}_{(m',1),n}, \dots, \mathbf{F}_{(m',k),n}, \dots, \mathbf{F}_{(m',K),n}], \\ \check{\mathbf{F}}_{m,k,n} &= [\mathbf{F}_{(m,1),n}, \dots, \mathbf{F}_{(m,k-1),n}, \mathbf{F}_{(m,k+1),n}, \dots, \mathbf{F}_{(m,K),n}], \end{aligned} \quad (7)$$

so that the rate $r_{(m,k),n}$ is formed by the difference between two Shannon transforms as

$$\begin{aligned} r_{(m,k),n} = & \log \left| \mathbf{I} + \mathbf{A}_{(m,k),n} \mathbf{A}_{(m,k),n}^H \right| \\ & - \log \left| \mathbf{I} + \mathbf{B}_{(m,k),n} \mathbf{B}_{(m,k),n}^H \right|. \end{aligned} \quad (8)$$

The $N_U \times N_U KM$ matrix $\mathbf{A}_{(m,k),n}$ is associated with the total received signal while the $N_U \times N_U (KM - 1)$ matrix $\mathbf{B}_{(m,k),n}$ is associated with interference plus noise.

In order to approximate $\mathbb{E}_{\mathbf{H}}[r_{(m,k),n}]$, we need to find the asymptotic approximation to $\mathbb{E}_{\mathbf{X}}[\log |\mathbf{I} + \mathbf{X}\mathbf{X}^H|]$. According to the random matrix theory [21, 22, 40], the ergodic Shannon transform of $\mathbf{X}\mathbf{X}^H$ can be asymptotically approximated by a deterministic function that only depends on the second-order statistics of \mathbf{X} . Therefore, we only need the second-order statistics of $\mathbf{A}_{(m,k),n}$ and $\mathbf{B}_{(m,k),n}$ to characterize $\tilde{r}_{(m,k),n} \approx \mathbb{E}_{\mathbf{H}}[r_{(m,k),n}]$.

3.1. Second-Order Statistics of $\mathbf{A}_{(m,k),n}$ and $\mathbf{B}_{(m,k),n}$. We shall now calculate the $N_U^2 KM \times N_U^2 KM$ covariance matrix of $\mathbf{A}_{(m,k),n}$ defined as

$$\boldsymbol{\Sigma}_{(m,k),n}^{\mathbf{A}} = \mathbb{E}_{\mathbf{H}} \left[\text{vec} \{ \mathbf{A}_{(m,k),n} \} \text{vec}^H \{ \mathbf{A}_{(m,k),n} \} \right], \quad (9)$$

which is associated with the total received signal, and the $N_U^2 (KM - 1) \times N_U^2 (KM - 1)$ covariance matrix of $\mathbf{B}_{(m,k),n}$ defined as

$$\boldsymbol{\Sigma}_{(m,k),n}^{\mathbf{B}} = \mathbb{E}_{\mathbf{H}} \left[\text{vec} \{ \mathbf{B}_{(m,k),n} \} \text{vec}^H \{ \mathbf{B}_{(m,k),n} \} \right], \quad (10)$$

which is associated with interference plus noise. From the definition of $\mathbf{A}_{(m,k),n}$ and statistical independence of channel matrices across BSs, it follows that

$$\boldsymbol{\Sigma}_{(m,k),n}^{\mathbf{A}} = \begin{bmatrix} \boldsymbol{\Sigma}_{1,(m,k),n} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \cdots & \cdots & \mathbf{0} \\ \vdots & \vdots & \boldsymbol{\Sigma}_{m,(m,k),n} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & \boldsymbol{\Sigma}_{M,(m,k),n} \end{bmatrix}, \quad (11)$$

where $\boldsymbol{\Sigma}_{m',(m,k),n}$ is the $N_U^2 K \times N_U^2 K$ subcovariance matrix corresponding to BS m' and given by

$$\begin{aligned} \boldsymbol{\Sigma}_{m',(m,k),n} &= \mathbb{E}_{\mathbf{H}} \left[\text{vec} \{ \mathbf{H}_{m',(m,k),n} \bar{\mathbf{F}}_{m',n} \} \text{vec}^H \{ \mathbf{H}_{m',(m,k),n} \bar{\mathbf{F}}_{m',n} \} \right]. \end{aligned} \quad (12)$$

Similarly, from the definition of $\mathbf{B}_{(m,k),n}$ and statistical independence of channel matrices across BSs, it follows that

$$\Sigma_{(m,k),n}^{\mathbf{B}} = \begin{bmatrix} \Sigma_{1,(m,k),n} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \cdots & \cdots & \mathbf{0} \\ \vdots & \vdots & \check{\Sigma}_{m,(m,k),n} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & \Sigma_{M,(m,k),n} \end{bmatrix}, \quad (13)$$

where $\check{\Sigma}_{m,(m,k),n}$ is the $N_U^2(K-1) \times N_U^2(K-1)$ subcovariance matrix corresponding to BS m' and given by

$$\begin{aligned} & \check{\Sigma}_{m,(m,k),n} \\ &= \mathbb{E}_{\mathbf{H}} \left[\text{vec} \left\{ \mathbf{H}_{m,(m,k),n} \check{\mathbf{F}}_{m,k,n} \right\} \text{vec}^H \left\{ \mathbf{H}_{m,(m,k),n} \check{\mathbf{F}}_{m,k,n} \right\} \right]. \end{aligned} \quad (14)$$

Now, we only need to calculate the subcovariance matrices $\Sigma_{m',(m,k),n}$ and $\check{\Sigma}_{m,(m,k),n}$ to completely determine the covariance matrices $\Sigma_{(m,k),n}^{\mathbf{A}}$ and $\Sigma_{(m,k),n}^{\mathbf{B}}$. It is straightforward to show that

$$\begin{aligned} \Sigma_{m',(m,k),n} &= \left(\bar{\mathbf{F}}_{m',n}^T \otimes \mathbf{I} \right) \Omega_{m',(m,k),n} \left(\bar{\mathbf{F}}_{m',n}^* \otimes \mathbf{I} \right), \\ \check{\Sigma}_{m,(m,k),n} &= \left(\check{\mathbf{F}}_{m,k,n}^T \otimes \mathbf{I} \right) \Omega_{m,(m,k),n} \left(\check{\mathbf{F}}_{m,k,n}^* \otimes \mathbf{I} \right), \end{aligned} \quad (15)$$

where $\Omega_{m',(m,k),n}$ is defined in (4).

We remark here that, in contrast to the uplink transmission considered in [20] where each subcovariance matrix is $N_U N_B \times N_U N_B$, in the downlink transmission each subcovariance matrix is $N_U^2 K \times N_U^2 K$, so each dimension is approximately K times larger. The increased size of the covariance matrix is due to the fact that the downlink transmitter (BS) has K intended receivers (users), while the uplink transmitter (user) has only one intended receiver (BS). In addition to this, there are N covariance matrix groups in the MIMO-OFDM system, one for each subcarrier.

The deterministic approximation will depend on the eigenvalues of $\Sigma_{(m,k),n}^{\mathbf{A}}$ and $\Sigma_{(m,k),n}^{\mathbf{B}}$, so we shall define two eigenvalue matrices. Define the $N_U \times N_U K M$ eigenvalue matrix $\Lambda_{(m,k),n}^{\mathbf{A}}$ and $N_U \times N_U (K M - 1)$ eigenvalue matrix $\Lambda_{(m,k),n}^{\mathbf{B}}$, which are formed by stacking the eigenvalues of $\Sigma_{(m,k),n}^{\mathbf{A}}$ and $\Sigma_{(m,k),n}^{\mathbf{B}}$ in $N_U \times 1$ sized columns, respectively. Such a definition is required to apply the theorems from random matrix theory. Note that the first dimension of eigenvalue matrices is the receiver antenna number N_U . It is easy to see that

$$\begin{aligned} \Lambda_{(m,k),n}^{\mathbf{A}} &= [\Lambda_{1,(m,k),n}, \dots, \Lambda_{m,(m,k),n}, \dots, \Lambda_{M,(m,k),n}], \\ \Lambda_{(m,k),n}^{\mathbf{B}} &= [\Lambda_{1,(m,k),n}, \dots, \check{\Lambda}_{m,(m,k),n}, \dots, \Lambda_{M,(m,k),n}], \end{aligned} \quad (16)$$

where $N_U \times N_U K$ eigenvalue matrix $\Lambda_{m',(m,k),n}$ and $N_U \times N_U (K - 1)$ eigenvalue matrix $\check{\Lambda}_{m,(m,k),n}$ are formed by stacking the eigenvalues of $\Sigma_{m',(m,k),n}$ and $\check{\Sigma}_{m,(m,k),n}$ in $N_U \times 1$ sized columns, respectively.

3.2. Approximation of the Rate Function. In this subsection, we will derive the approximation to the rate based on the random matrix theory. We start by reviewing some concepts from the random matrix theory. For an $I \times J$ random matrix \mathbf{X} , the correlation function is defined as $R_{\mathbf{X}}(i, j; i', j') \triangleq \mathbb{E}[\mathbf{X}]_{i,j} [\mathbf{X}]_{i',j'}^*$. Define the kernel of the correlation function by a set of orthonormal eigenfunctions $\psi_{k,l}(i, j) \in \mathbb{C}$ satisfying $\sum_{i',j'} R_{\mathbf{X}}(i, j; i', j') \psi_{k,l}(i', j') = \lambda_{k,l} \psi_{k,l}(i, j)$ where $\lambda_{k,l}$ is the kernel eigenvalue [21]. We say the kernel is factorable if $\psi_{k,l}(i, j) = u_k(i) v_l(j)$ for some functions $u_k(i)$ and $v_l(j)$. Also, define $\Lambda^{\mathbf{X}} \in \mathbb{R}^{I \times J}$ as the eigenvalue matrix which is formed by stacking the eigenvalues of $\Sigma^{\mathbf{X}} = \mathbb{E}[\text{vec}[\mathbf{X}] \text{vec}^H[\mathbf{X}]]$ in $I \times 1$ sized columns. Matrices \mathbf{X} , $\Sigma^{\mathbf{X}}$, and $\Lambda^{\mathbf{X}}$ can represent $\mathbf{A}_{(m,k),n}$, $\Sigma_{(m,k),n}^{\mathbf{A}}$, and $\Lambda_{(m,k),n}^{\mathbf{A}}$ (or those of \mathbf{B}), respectively.

Proposition 1. *Suppose the following technical conditions hold for the correlation function of $\mathbf{X} \in \mathbb{C}^{I \times J}$: (a) its kernel is factorable; (b) its eigenvalues multiplied by I are uniformly bounded. Then, based on [29, Theorem 2] and [23, Theorem 3], the random variable $(1/I) \log |\mathbf{I} + \mathbf{X} \mathbf{X}^H|$ will converge almost surely to its deterministic approximation $\tilde{f}_{\mathbf{X}}$ for J/I fixed as $I, J \rightarrow \infty$. We have*

$$\tilde{f}_{\mathbf{X}} \triangleq \frac{1}{I} \sum_{j=1}^J \log \left(1 + I \left[\Lambda^{\mathbf{X}^T \mathbf{u}} \right]_j \right) \quad (17)$$

$$+ \frac{1}{I} \sum_{i=1}^I \log \left(1 + J \left[\Lambda^{\mathbf{X} \mathbf{v}} \right]_i \right) - J \mathbf{u}^T \Lambda^{\mathbf{X}} \mathbf{v} \log e,$$

where $\Lambda^{\mathbf{X}}$ denotes the eigenvalue matrix and $\mathbf{u} \in \mathbb{R}^I$, $\mathbf{v} \in \mathbb{R}^J$ are solutions to the following fixed point equations:

$$\begin{aligned} [\mathbf{u}]_i^{-1} &= I \left(1 + J \left[\Lambda^{\mathbf{X} \mathbf{v}} \right]_i \right), \quad i = 1, \dots, I, \\ [\mathbf{v}]_j^{-1} &= J \left(1 + I \left[\Lambda^{\mathbf{X}^T \mathbf{u}} \right]_j \right), \quad j = 1, \dots, J. \end{aligned} \quad (18)$$

Now based on this approximation technique and the second-order statistics of $\mathbf{A}_{m',(m,k),n}$ and $\mathbf{B}_{m',(m,k),n}$ derived previously, we can obtain the asymptotic approximation to the ergodic rate in the following theorem.

Theorem 2. *Suppose the following conditions are satisfied for the correlation functions of $\mathbf{A}_{(m,k),n}$ and $\mathbf{B}_{(m,k),n}$: (a) they have factorable kernels; (b) their eigenvalues multiplied by N_U are uniformly bounded. Then for $N_B \geq N_U \rightarrow \infty$ we have*

$$\frac{r_{(m,k),n} - \tilde{r}_{(m,k),n}}{N_U} \rightarrow 0, \quad (19)$$

in which $\tilde{r}_{(m,k),n}$ is a deterministic function given by

$$\begin{aligned}
\tilde{r}_{(m,k),n} = & - \sum_{j=1}^{N_U KM} \log [\mathbf{v}_{(m,k),n}^{\mathbf{A}}]_j \\
& + \sum_{j=1}^{N_U(KM-1)} \log [\mathbf{v}_{(m,k),n}^{\mathbf{B}}]_j + \sum_{i=1}^{N_U} \left[\log \frac{[\mathbf{u}_{(m,k),n}^{\mathbf{B}}]_i}{[\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_i} \right. \\
& \left. + N_U \log e \left([\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_i - [\mathbf{u}_{(m,k),n}^{\mathbf{B}}]_i \right) \right] - N_U \\
& \cdot \log(N_U(KM-1)) + N_U KM \log \left(1 - \frac{1}{KM} \right),
\end{aligned} \tag{20}$$

where $\mathbf{u}_{(m,k),n}^{\mathbf{A}} \in \mathbb{R}^{N_U}$, $\mathbf{v}_{(m,k),n}^{\mathbf{A}} \in \mathbb{R}^{KM N_U}$, $\mathbf{u}_{(m,k),n}^{\mathbf{B}} \in \mathbb{R}^{N_U}$, and $\mathbf{v}_{(m,k),n}^{\mathbf{B}} \in \mathbb{R}^{(KM-1)N_U}$ are the solutions to the following fixed point equations:

$$[\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_i^{-1} = N_U \left(1 + KM N_U [\Lambda_{(m,k),n}^{\mathbf{A}} \mathbf{v}_{(m,k),n}^{\mathbf{A}}]_i \right), \tag{21}$$

$$[\mathbf{v}_{(m,k),n}^{\mathbf{A}}]_j^{-1} = N_U KM \left(1 + N_U [\Lambda_{(m,k),n}^{\mathbf{A}} \mathbf{u}_{(m,k),n}^{\mathbf{A}}]_j \right), \tag{22}$$

$$\begin{aligned}
[\mathbf{u}_{(m,k),n}^{\mathbf{B}}]_i^{-1} &= N_U \left(1 + N_U (KM-1) [\Lambda_{(m,k),n}^{\mathbf{B}} \mathbf{v}_{(m,k),n}^{\mathbf{B}}]_i \right), \\
[\mathbf{v}_{(m,k),n}^{\mathbf{B}}]_j^{-1} &= N_U (KM-1) \left(1 + N_U [\Lambda_{(m,k),n}^{\mathbf{B}} \mathbf{u}_{(m,k),n}^{\mathbf{B}}]_j \right),
\end{aligned} \tag{23}$$

where i and j range according to the vector sizes.

Proof. We will apply Proposition 1 to each term in (8). Consider the first term $\log |\mathbf{I} + \mathbf{A}_{(m,k),n} \mathbf{A}_{(m,k),n}^H|$ in which $\mathbf{A}_{(m,k),n} \in \mathbb{C}^{N_U \times N_U KM}$ must satisfy conditions (a) and (b) of Theorem 2 in accordance with the same conditions in Proposition 1. Since $N_U KM / N_U = KM$, the ratio of dimensions is always fixed as $N_B \geq N_U \rightarrow \infty$. Condition $N_B \geq N_U$ guarantees that the precoders have full column rank so (2) remains valid. Now, from Proposition 1, the first term in (8) divided by N_U is approximated as

$$\begin{aligned}
\tilde{f}_{\mathbf{A}_{(m,k),n}} &\triangleq \frac{1}{N_U} \sum_{j=1}^{N_U KM} \log \left(1 + N_U [\Lambda_{(m,k),n}^{\mathbf{A}} \mathbf{u}_{(m,k),n}^{\mathbf{A}}]_j \right) \\
&+ \frac{1}{N_U} \sum_{i=1}^{N_U} \log \left(1 + KM N_U [\Lambda_{(m,k),n}^{\mathbf{A}} \mathbf{v}_{(m,k),n}^{\mathbf{A}}]_i \right) \\
&- N_U KM \mathbf{u}_{(m,k),n}^{\mathbf{A}} \mathbf{u}_{(m,k),n}^{\mathbf{A}T} \Lambda_{(m,k),n}^{\mathbf{A}} \mathbf{v}_{(m,k),n}^{\mathbf{A}} \log e.
\end{aligned} \tag{24}$$

In the first term of approximation $\tilde{f}_{\mathbf{A}_{(m,k),n}}$, we can express $[\Lambda_{(m,k),n}^{\mathbf{A}} \mathbf{u}_{(m,k),n}^{\mathbf{A}}]_j$ in terms of $[\mathbf{v}_{(m,k),n}^{\mathbf{A}}]_j$ using (22). In the second term of approximation $\tilde{f}_{\mathbf{A}_{(m,k),n}}$, we can express $[\Lambda_{(m,k),n}^{\mathbf{A}} \mathbf{v}_{(m,k),n}^{\mathbf{A}}]_i$ in terms of $[\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_i$ using (21). For the third term of approximation $\tilde{f}_{\mathbf{A}_{(m,k),n}}$, we can write

$$\begin{aligned}
&\mathbf{u}_{(m,k),n}^{\mathbf{A}} \mathbf{u}_{(m,k),n}^{\mathbf{A}T} \Lambda_{(m,k),n}^{\mathbf{A}} \mathbf{v}_{(m,k),n}^{\mathbf{A}} \\
&= \sum_i \sum_j [\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_i [\Lambda_{(m,k),n}^{\mathbf{A}}]_{i,j} [\mathbf{v}_{(m,k),n}^{\mathbf{A}}]_j \\
&= \sum_i [\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_i \sum_j [\Lambda_{(m,k),n}^{\mathbf{A}}]_{i,j} [\mathbf{v}_{(m,k),n}^{\mathbf{A}}]_j \\
&= \sum_i [\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_i [\Lambda_{(m,k),n}^{\mathbf{A}} \mathbf{v}_{(m,k),n}^{\mathbf{A}}]_i
\end{aligned} \tag{25}$$

and express $[\Lambda_{(m,k),n}^{\mathbf{A}} \mathbf{v}_{(m,k),n}^{\mathbf{A}}]_i$ in terms of $[\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_i$ using (21). So all the terms in $\tilde{f}_{\mathbf{A}_{(m,k),n}}$ are expressed in terms of $[\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_i$, $[\mathbf{v}_{(m,k),n}^{\mathbf{A}}]_j$. In a similar manner, $\tilde{f}_{\mathbf{B}_{(m,k),n}}$ can be expressed in terms of $[\mathbf{u}_{(m,k),n}^{\mathbf{B}}]_i$, $[\mathbf{v}_{(m,k),n}^{\mathbf{B}}]_j$. Forming $(1/N_U) \tilde{r}_{(m,k),n} = \tilde{f}_{\mathbf{A}_{(m,k),n}} - \tilde{f}_{\mathbf{B}_{(m,k),n}}$ along with some mathematical manipulations completes the proof. \square

Some remarks are in order. It is shown in [23] that many channel models including the unitary-independent-unitary (UIU) model, Kronecker model, and independent nonidentically distributed (IND) model satisfy conditions (a) and (b) of Theorem 2. While the approximation is guaranteed to converge as $N_B \geq N_U \rightarrow \infty$, it is accurate even for very small number of antennas (e.g., $N_U = 3$, $N_B = 3$) as will be shown in the simulation section. The existence of positive solutions to the fixed point equations in Theorem 2 is proved in [21]. Computationally, the equations can be solved numerically by iteratively substituting the value on the right-hand side into the left-hand side. The convergence result suggests that the ergodic rate $\mathbb{E}_{\mathbf{H}}[r_{(m,k),n}]$ can be approximated by $\tilde{r}_{(m,k),n}$.

4. Weighted Sum Rate Maximization

In this section, we develop an algorithm to obtain a sub-optimal solution to problem (P2). We will often need to differentiate one matrix with respect to another matrix; therefore, to facilitate calculations, the following definition is presented [41].

Definition 3. Let $\mathbf{f}(\mathbf{x})$ be a function mapping $\mathbb{R}^n \rightarrow \mathbb{R}^m$, define its derivative $\mathcal{D}_{\mathbf{x}} \mathbf{f}$ as the $m \times n$ matrix with elements

$$[\mathcal{D}_{\mathbf{x}} \mathbf{f}]_{i,j} = \frac{\partial \mathbf{f}_i(\mathbf{x})}{\partial x_j}. \tag{26}$$

Furthermore, let $\mathbf{F}(\mathbf{X})$ be a function mapping $\mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{p \times q}$; define its derivative $\mathcal{D}_{\mathbf{X}}\mathbf{F}(\mathbf{X})$ as the $pq \times mn$ matrix

$$\mathcal{D}_{\mathbf{X}}\mathbf{F}(\mathbf{X}) = \mathcal{D}_{\text{vec}\{\mathbf{X}\}}\text{vec}\{\mathbf{F}(\mathbf{X})\}. \quad (27)$$

The most important property of this definition is that the chain rule holds, which enables the differentiation of complicated functions. Adopting this definition, we can now seek to solve the optimization problem.

Problem (P2) is nonconvex, so a global maximum is generally hard to obtain. Fortunately, the KKT conditions still serve as the necessary conditions for local optima, so we seek a suboptimal solution that satisfies the KKT conditions. To this end, we define

$$\mathbf{f}_{(m,k),n} = \begin{bmatrix} \text{vec}\{\text{Re}(\mathbf{F}_{(m,k),n})\} \\ \text{vec}\{\text{Im}(\mathbf{F}_{(m,k),n})\} \end{bmatrix} \in \mathbb{R}^{2N_U N_B \times 1}, \quad (28)$$

$$[\mathbf{f}_{(m,k),n}]_i = \begin{cases} \sqrt{P_m} \cos[\phi_{(m,k),n}]_i \prod_{(k',n',i') < (k,n,i)} \sin[\phi_{(m,k'),n'}]_{i'}, & (k,n,i) \neq (K,N,2N_B N_U) \\ \sqrt{P_m} \prod_{(k',n',i') < (K,N,2N_B N_U)} \sin[\phi_{(m,k'),n'}]_{i'}, & (k,n,i) = (K,N,2N_B N_U), \end{cases} \quad (30)$$

where $<$ denotes the lexicographical order of vectors. Under this parametrization, the power constraint is automatically satisfied since the variables are on the surface of a hypersphere with radius $\sqrt{P_m}$.

Since the local optima occur at the interior of the domain of angle parameters, the KKT condition for problem (P2) is simplified to

$$\sum_{m=1}^M \sum_{k=1}^K w_{(m,k)} \sum_{n=1}^N \mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n} = 0, \quad \forall m', \quad (31)$$

$\theta_{m'}$

$$\triangleq [\phi_{(m',1),1}^T, \dots, \phi_{(m',1),N}^T, \dots, \phi_{(m',K),1}^T, \dots, \phi_{(m',K),N}^T]^T,$$

where $\theta_{m'} \in \mathbb{R}^{(2N_U N_B KN - 1) \times 1}$ holds all the angle parameters corresponding to BS m' . The above expression simply states that the gradient vanishes at any optimum point. So finding a local optimum solution is equivalent to pursuing a set of points $\theta_{m'}^{\text{opt}}, \forall m'$ at which the gradient vanishes. In practice, the equation for setting the gradient to zero is highly nonlinear and it is impossible to solve it directly. However, starting from any initial point for $\theta_{m'}$, we can use a simple gradient search method to increase the objective function after each iteration and gradually approach a point where the gradient is zero. In order to do so, we need to derive the $1 \times (2N_U N_B KN - 1)$ vector $\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}$, which is the transpose of the gradient vector. The remaining materials of this section will thus be devoted to the derivation of $\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}$.

4.1. Derivation of $\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}$. It immediately follows by the chain rule that

which is the complex-real isomorphism of $\mathbf{F}_{(m,k),n}$. Problem (P2) then becomes

$$\begin{aligned} \max_{\tilde{\mathcal{F}}} \quad & \sum_{m=1}^M \sum_{k=1}^K w_{(m,k)} \sum_{n=1}^N \tilde{r}_{(m,k),n} \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{n=1}^N \mathbf{f}_{(m,k),n}^T \mathbf{f}_{(m,k),n} = P_m, \quad \forall m, \end{aligned} \quad (29)$$

where $\tilde{\mathcal{F}}$ is a set that has complex-real isomorphism with respect to \mathcal{F} .

To further simplify the problem, we can reformulate it as an unconstrained optimization. To do so, similar to [42], we make a change of variables from $\mathbf{f}_{(m,k),n}$ to the spherical coordinates $\phi_{(m,k),n}$ that belongs to $\mathbb{R}^{2N_B N_U}$ if $(k,n) \neq (K,N)$ and to $\mathbb{R}^{2N_B N_U - 1}$ if $(k,n) = (K,N)$. The transformation is described by

$$\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n} = \mathcal{D}_{\bar{\mathbf{f}}_{m',n}} \tilde{r}_{(m,k),n} \mathcal{D}_{\theta_{m'}} \bar{\mathbf{f}}_{m',n}, \quad (32)$$

$$\bar{\mathbf{f}}_{m',n} = [\mathbf{f}_{(m',1),n}^T, \mathbf{f}_{(m',2),n}^T, \dots, \mathbf{f}_{(m',K),n}^T]^T. \quad (33)$$

Now we need to find $\mathcal{D}_{\bar{\mathbf{f}}_{m',n}} \tilde{r}_{(m,k),n}$ and $\mathcal{D}_{\theta_{m'}} \bar{\mathbf{f}}_{m',n}$ in order to form the chain rule and obtain (33). We shall find derivatives with respect to the subvectors of $\bar{\mathbf{f}}_{m',n}$, that is, with respect to vectors $\mathbf{f}_{(m',k'),n}$. The components of $\mathcal{D}_{\theta_{m'}} \mathbf{f}_{(m',k'),n}$ can be obtained by differentiating (30) which yields the following:

$$\mathcal{D}_{\theta_{m'}} [\mathbf{f}_{(m',k'),n}]_i = [\mathbf{f}_{(m',k'),n}]_i \times \begin{cases} \begin{bmatrix} \cot[\phi_{(m',1),1}]_1 \\ \vdots \\ \cot[\phi_{(m',k'),n}]_{i-1} \\ -\tan[\phi_{(m',k'),n}]_i \\ \mathbf{0} \end{bmatrix}^T, & (k',n,i) \neq (K,N,2N_B^2) \\ \begin{bmatrix} \cot[\phi_{(m',1),1}]_1 \\ \vdots \\ \cot[\phi_{(m',k'),n}]_{i-2} \\ \cot[\phi_{(m',k'),n}]_{i-1} \end{bmatrix}^T, & \text{o.w.} \end{cases} \quad (34)$$

The components of $\mathcal{D}_{\mathbf{f}_{(m',k'),n}} \tilde{r}_{(m,k),n}$ can be found through the complex-real isomorphism as

$$\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{\mathbf{r}}_{(m,k),n} = 2 \begin{bmatrix} \text{Re}^T \left\{ \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{\mathbf{r}}_{(m,k),n} \right\} \\ -\text{Im}^T \left\{ \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{\mathbf{r}}_{(m,k),n} \right\} \end{bmatrix}^T. \quad (35)$$

Substituting (34) and (35) in (33) completes the derivation provided that $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{\mathbf{r}}_{(m,k),n}$ is available.

Calculation of $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{\mathbf{r}}_{(m,k),n}$ is more complicated since, according to Theorem 2, $\tilde{\mathbf{r}}_{(m,k),n}$ is a function of $\mathbf{u}_{(m,k),n}^{\mathbf{A}}$, $\mathbf{v}_{(m,k),n}^{\mathbf{A}}$, $\mathbf{u}_{(m,k),n}^{\mathbf{B}}$, and $\mathbf{v}_{(m,k),n}^{\mathbf{B}}$, which are in turn functions of $\Lambda_{(m,k),n}^{\mathbf{A}}$, $\Lambda_{(m,k),n}^{\mathbf{B}}$. The chain rule gives

$$\begin{aligned} & \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{\mathbf{r}}_{(m,k),n} \\ &= \mathcal{D}_{\mathbf{u}_{(m,k),n}^{\mathbf{A}}} \tilde{\mathbf{r}}_{(m,k),n} \mathcal{D}_{\Lambda_{(m,k),n}^{\mathbf{A}}} \mathbf{u}_{(m,k),n}^{\mathbf{A}} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^{\mathbf{A}} \\ &+ \mathcal{D}_{\mathbf{v}_{(m,k),n}^{\mathbf{A}}} \tilde{\mathbf{r}}_{(m,k),n} \mathcal{D}_{\Lambda_{(m,k),n}^{\mathbf{A}}} \mathbf{v}_{(m,k),n}^{\mathbf{A}} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^{\mathbf{A}} \quad (36) \\ &+ \mathcal{D}_{\mathbf{u}_{(m,k),n}^{\mathbf{B}}} \tilde{\mathbf{r}}_{(m,k),n} \mathcal{D}_{\Lambda_{(m,k),n}^{\mathbf{B}}} \mathbf{u}_{(m,k),n}^{\mathbf{B}} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^{\mathbf{B}} \\ &+ \mathcal{D}_{\mathbf{v}_{(m,k),n}^{\mathbf{B}}} \tilde{\mathbf{r}}_{(m,k),n} \mathcal{D}_{\Lambda_{(m,k),n}^{\mathbf{B}}} \mathbf{v}_{(m,k),n}^{\mathbf{B}} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^{\mathbf{B}}. \end{aligned}$$

There are three derivatives involved in each term of (36). In what follows, we shall calculate each of them.

4.1.1. First Chain of (36). The first chains can be obtained by differentiating $\tilde{\mathbf{r}}_{(m,k),n}$, as given in Theorem 2, which yields

$$\begin{aligned} \mathcal{D}_{[\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_i} \tilde{\mathbf{r}}_{(m,k),n} &= - \left([\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_i^{-1} - N_U \right) \log e, \\ \mathcal{D}_{[\mathbf{u}_{(m,k),n}^{\mathbf{B}}]_i} \tilde{\mathbf{r}}_{(m,k),n} &= \left([\mathbf{u}_{(m,k),n}^{\mathbf{B}}]_i^{-1} - N_U \right) \log e, \\ \mathcal{D}_{[\mathbf{v}_{(m,k),n}^{\mathbf{A}}]_i} \tilde{\mathbf{r}}_{(m,k),n} &= - [\mathbf{v}_{(m,k),n}^{\mathbf{A}}]_i^{-1} \log e, \\ \mathcal{D}_{[\mathbf{v}_{(m,k),n}^{\mathbf{B}}]_i} \tilde{\mathbf{r}}_{(m,k),n} &= [\mathbf{v}_{(m,k),n}^{\mathbf{B}}]_i^{-1} \log e. \end{aligned} \quad (37)$$

4.1.2. Second Chain of (36). The second chains can be obtained by differentiating the fixed point equations given in Theorem 2. It can be shown that

$$\begin{aligned} & \begin{bmatrix} \mathcal{D}_{[\Lambda_{(m,k),n}^{\mathbf{A}}]_{h,l}} \mathbf{u}_{(m,k),n}^{\mathbf{A}} \\ \mathcal{D}_{[\Lambda_{(m,k),n}^{\mathbf{A}}]_{h,l}} \mathbf{v}_{(m,k),n}^{\mathbf{A}} \end{bmatrix} = -KN_U M \\ & \times \begin{bmatrix} \mathbf{U}_{(m,k),n}^{\mathbf{A}} & KN_U^2 M \Lambda_{(m,k),n}^{\mathbf{A}} \\ KN_U^2 M \Lambda_{(m,k),n}^{\mathbf{A}} & \mathbf{V}_{(m,k),n}^{\mathbf{A}} \end{bmatrix}^{-1} \\ & \cdot \begin{bmatrix} [\mathbf{v}_{(m,k),n}^{\mathbf{A}}]_l \mathbf{e}_h^{N_U} \\ [\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_h \mathbf{e}_l^{KN_B M} \end{bmatrix}, \\ & \begin{bmatrix} \mathcal{D}_{[\Lambda_{(m,k),n}^{\mathbf{B}}]_{h,l}} \mathbf{u}_{(m,k),n}^{\mathbf{B}} \\ \mathcal{D}_{[\Lambda_{(m,k),n}^{\mathbf{B}}]_{h,l}} \mathbf{v}_{(m,k),n}^{\mathbf{B}} \end{bmatrix} = -(KM-1) N_U \end{aligned}$$

$$\begin{aligned} & \times \begin{bmatrix} \mathbf{U}_{(m,k),n}^{\mathbf{B}} & (KM-1) N_U^2 \Lambda_{(m,k),n}^{\mathbf{B}} \\ (KM-1) N_U^2 \Lambda_{(m,k),n}^{\mathbf{B}} & \mathbf{V}_{(m,k),n}^{\mathbf{B}} \end{bmatrix}^{-1} \\ & \times \begin{bmatrix} [\mathbf{v}_{(m,k),n}^{\mathbf{B}}]_l \mathbf{e}_h^{N_U} \\ [\mathbf{u}_{(m,k),n}^{\mathbf{B}}]_h \mathbf{e}_l^{(KM-1)N_B} \end{bmatrix}, \end{aligned} \quad (38)$$

where \mathbf{e}_l^P denotes the l th standard basis vector in \mathbb{R}^P and

$$\begin{aligned} \mathbf{U}_{(m,k),n}^{\mathbf{A}} &= \text{diag} \left\{ [\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_1^{-2}, \dots, [\mathbf{u}_{(m,k),n}^{\mathbf{A}}]_{N_U}^{-2} \right\}, \\ \mathbf{V}_{(m,k),n}^{\mathbf{A}} &= \text{diag} \left\{ [\mathbf{v}_{(m,k),n}^{\mathbf{A}}]_1^{-2}, \dots, [\mathbf{v}_{(m,k),n}^{\mathbf{A}}]_{KN_U M}^{-2} \right\}, \\ \mathbf{U}_{(m,k),n}^{\mathbf{B}} &= \text{diag} \left\{ [\mathbf{u}_{(m,k),n}^{\mathbf{B}}]_1^{-2}, \dots, [\mathbf{u}_{(m,k),n}^{\mathbf{B}}]_{N_U}^{-2} \right\}, \\ \mathbf{V}_{(m,k),n}^{\mathbf{B}} &= \text{diag} \left\{ [\mathbf{v}_{(m,k),n}^{\mathbf{B}}]_1^{-2}, \dots, [\mathbf{v}_{(m,k),n}^{\mathbf{B}}]_{(KM-1)N_U}^{-2} \right\}. \end{aligned} \quad (39)$$

4.1.3. Third Chain of (36). Using the chain rule together with (16), we have

$$\begin{aligned} & \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^{\mathbf{A}} \\ &= \begin{bmatrix} \mathbf{0} \\ \mathcal{D}_{\Sigma_{m',(m,k),n}} \Lambda_{m',(m,k),n} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Sigma_{m',(m,k),n} \\ \mathbf{0} \end{bmatrix}, \\ & \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^{\mathbf{B}} \\ &= \begin{cases} \begin{bmatrix} \mathbf{0} \\ \mathcal{D}_{\Sigma_{m',(m,k),n}} \Lambda_{m',(m,k),n} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Sigma_{m',(m,k),n} \\ \mathbf{0} \end{bmatrix}, & m' \neq m, \\ \begin{bmatrix} \mathbf{0} \\ \mathcal{D}_{\check{\Sigma}_{m,(m,k),n}} \check{\Lambda}_{m,(m,k),n} \mathcal{D}_{\mathbf{F}_{(m,k),n}} \check{\Sigma}_{m,(m,k),n} \\ \mathbf{0} \end{bmatrix}, & m' = m. \end{cases} \end{aligned} \quad (40)$$

We now need to calculate the terms in the above formulas. Note that the nonzero terms in the above matrices are equal to $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{m',(m,k),n}$ and $\mathcal{D}_{\mathbf{F}_{(m,k),n}} \check{\Lambda}_{m,(m,k),n}$ expressed through the chain rule.

Based on the results in [41], the derivative of the eigenvalues with respect to the matrix can be explicitly written as a function of the eigenvectors, so we obtain

$$\begin{aligned} & \mathcal{D}_{\Sigma_{m',(m,k),n}} \Lambda_{m',(m,k),n} \\ &= \begin{bmatrix} \text{vec}^T \left\{ \boldsymbol{\mu}_{m',(m,k),n}^1 * \boldsymbol{\mu}_{m',(m,k),n}^{1T} \right\} \\ \vdots \\ \text{vec}^T \left\{ \boldsymbol{\mu}_{m',(m,k),n}^{KN_U^2} * \boldsymbol{\mu}_{m',(m,k),n}^{KN_U^2 T} \right\} \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} & \mathcal{D}_{\tilde{\Sigma}_{m,(m,k),n}} \tilde{\Lambda}_{m,(m,k),n} \\ &= \begin{bmatrix} \text{vec}^T \left\{ \tilde{\boldsymbol{\mu}}_{m,(m,k),n}^1 * \tilde{\boldsymbol{\mu}}_{m,(m,k),n}^{1T} \right\} \\ \vdots \\ \text{vec}^T \left\{ \tilde{\boldsymbol{\mu}}_{m,(m,k),n}^{(K-1)N_U^2} * \tilde{\boldsymbol{\mu}}_{m,(m,k),n}^{(K-1)N_U^2 T} \right\} \end{bmatrix}, \end{aligned} \quad (41)$$

where $\tilde{\boldsymbol{\mu}}_{m,(m,k),n}^i$ and $\tilde{\boldsymbol{\mu}}_{m,(m,k),n}^i$ denote the i th eigenvector of $\Sigma_{m',(m,k),n}$ and $\tilde{\Sigma}_{m,(m,k),n}$, respectively.

Finally, differentiating (15) and using the results in [41], we obtain the formulas for $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Sigma_{m',(m,k),n}$ and $\mathcal{D}_{\mathbf{F}_{(m,k'),n}} \tilde{\Sigma}_{m,(m,k),n}$ given by

$$\begin{aligned} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Sigma_{m',(m,k),n} &= \left\{ \left[\left(\tilde{\mathbf{F}}_{m',n}^H \otimes \mathbf{I}_{N_U} \right) \left(\Omega_{m',(m,k),n}^{1/2} \right)^* \right] \right. \\ &\quad \left. \otimes \mathbf{I}_{KN_U^2} \right\} \times \mathbf{K}_{N_U N_B, KN_U^2} \left\{ \mathbf{e}_k^K \left[\left(\Omega_{m',(m,k),n}^{T/2} \right) \right. \right. \\ &\quad \cdot \left. \left. \left(\mathbf{I}_{N_B} \otimes \mathbf{K}_{N_U, N_U} \otimes \mathbf{I}_{N_U} \right) \left(\mathbf{I}_{N_U N_B} \otimes \text{vec} \left\{ \mathbf{I}_{N_U} \right\} \right) \right. \right. \\ &\quad \left. \left. \cdot \mathbf{K}_{N_B, N_U} \right] \right\}, \end{aligned} \quad (42)$$

$$\begin{aligned} \mathcal{D}_{\mathbf{F}_{(m,k'),n}} \tilde{\Sigma}_{m,(m,k),n} &= \left\{ \left[\left(\tilde{\mathbf{F}}_{m,k',n}^H \otimes \mathbf{I}_{N_U} \right) \left(\Omega_{m,(m,k),n}^{1/2} \right)^* \right] \right. \\ &\quad \left. \otimes \mathbf{I}_{(K-1)N_U^2} \right\} \times \mathbf{K}_{N_U N_B, (K-1)N_U^2} \left\{ \mathbf{u}_{k,k'} \right. \\ &\quad \left. \otimes \left[\left(\Omega_{m,(m,k),n}^{T/2} \right) \left(\mathbf{I}_{N_B} \otimes \mathbf{K}_{N_U, N_U} \otimes \mathbf{I}_{N_U} \right) \right. \right. \\ &\quad \left. \left. \cdot \left(\mathbf{I}_{N_U N_B} \otimes \text{vec} \left\{ \mathbf{I}_{N_U} \right\} \right) \mathbf{K}_{N_B, N_U} \right] \right\}, \end{aligned}$$

in which $\mathbf{K}_{k,l}$ is the commutation matrix defined in [41] satisfying $\mathbf{K}_{k,l} \text{vec}\{\mathbf{X}\} = \text{vec}\{\mathbf{X}^T\}$ for every matrix $\mathbf{X} \in \mathbb{C}^{k \times l}$, and

$$\mathbf{u}_{k,k'} = \begin{cases} \mathbf{e}_{k'}^{K-1} & k' < k, \\ \mathbf{e}_{k'-1}^{K-1} & k' > k, \\ \mathbf{0}_{K-1} & k' = k. \end{cases} \quad (43)$$

With all the chains derived, (36) is completely characterized which in turn enables an explicit expression for $\mathcal{D}_{\boldsymbol{\theta}_{m',\tilde{r}(m,k),n}}$ given by (33).

4.2. Local Optimum Solution. With the gradient $\mathcal{D}_{\boldsymbol{\theta}_{m',\tilde{r}(m,k),n}}$ computed, a gradient search method can be applied. We now propose Algorithm 1 to find a local optimum solution for problem (P2).

Algorithm 1 is not limited to multicell networks with partial cooperation, so with small modifications it can be used for precoder design based on statistical CSI for networked MIMO systems. In networked MIMO, in addition to the channel statistics, it is assumed that each BS has all the transmit data, so the whole system can be viewed as a MIMO single-cell with statistical CSI at the transmitter.

5. MIMO-OFDM Kronecker Channel Model

While the results obtained so far are valid for general correlation channel models, in this section, we will reduce the results to a more compact form with less computational complexity by considering the Kronecker channel model. The Kronecker model arises in practice when the immediate surrounding dominates the spatial correlation and the intermediate scattering clusters exist in a narrow angular range seen from the antennas [34, 39]. We shall start by describing the MIMO-OFDM channel and expressing it with the Kronecker structure.

5.1. Statistical Representation. A wideband MIMO channel is characterized by L channel taps $\tilde{\mathbf{H}}_0, \dots, \tilde{\mathbf{H}}_{L-1} \in \mathbb{C}^{I \times J}$. The channel matrix on the n th subcarrier is then given by $\mathbf{H}_n = \sum_{l=0}^{L-1} \tilde{\mathbf{H}}_l e^{-j2\pi n l / N} \in \mathbb{C}^{I \times J}$. Now the correlation among channel taps is given by the tap correlation matrix $\mathbf{Y} \in \mathbb{C}^{L \times L}$ that is defined as $[\mathbf{Y}]_{l,k} \triangleq \mathbb{E}\{\tilde{\mathbf{H}}_l \tilde{\mathbf{H}}_k^*\}$, $\forall i, j, i', j'$. On the other hand, the Kronecker model assumes that the correlation of transmitter side and the receiver side is separable, so for each channel tap $\tilde{\mathbf{H}}_l$, we have $\mathbb{E}\{\tilde{\mathbf{H}}_l \tilde{\mathbf{H}}_l^*\} = [\mathbf{R}]_{i,i'} [\mathbf{T}]_{j,j'}$ where $\mathbf{R} \in \mathbb{C}^{I \times I}$ and $\mathbf{T} \in \mathbb{C}^{J \times J}$ are receiver and transmitter correlation matrices, respectively. Based on the above definitions, it is easy to show that the channel correlation matrix on subcarrier n is given by

$$\begin{aligned} \Omega_n &= \mathbb{E} \left\{ \text{vec} \left\{ \mathbf{H}_n \right\} \text{vec}^H \left\{ \mathbf{H}_n \right\} \right\} = \omega_n^Y (\mathbf{T} \otimes \mathbf{R}) \in \mathbb{C}^{IJ \times IJ}, \\ \omega_n^Y &\triangleq \mathbf{w}_n^H \mathbf{Y} \mathbf{w}_n, \end{aligned} \quad (44)$$

$$\mathbf{w}_n = \left[e^{-j2\pi n(0)/N}, \dots, e^{-j2\pi n(L-1)/N} \right]^T.$$

Therefore, the channel correlation matrix Ω_n is characterized by the Kronecker product of transmitter and receiver correlation matrices \mathbf{T} and \mathbf{R} , multiplied by the quadratic form ω_n^Y which depends on the tap correlation matrix and the n th Fourier vector. Finally, the following statistical representation can be considered for the channel:

$$\mathbf{H}_n = \mathbf{R}^{1/2} \mathbf{H}_w \left[\left(\mathbf{Y}^{1/2} \mathbf{w}_n \right) \otimes \mathbf{T}^{1/2} \right], \quad (45)$$

where $\mathbf{H}_w \in \mathbb{C}^{I \times J L}$ is a white random matrix whose elements are uncorrelated with zero mean and unit variance. Through the properties of the Kronecker product, the above statistical representation yields $\Omega_n = \omega_n^Y (\mathbf{T} \otimes \mathbf{R})$. We shall use (45) to describe the MIMO-OFDM Kronecker channel model.

If the channel taps are uncorrelated, \mathbf{Y} is diagonal and ω_n^Y becomes independent of n ; hence, the channel correlation function Ω_n is the same over all subcarriers. However, whenever the channel taps are correlated, \mathbf{Y} is not diagonal and channel statistics are different on each subcarrier. Since the precoders depend on the channel statistics, when there is tap correlation, the MIMO-OFDM precoding matrices are frequency dependent which limits the system performance [17, 31–34]. But when there is no tap correlation, precoders are the same across all frequencies. We shall study the effect of tap correlation on the system sum rate in the simulation results section.

- (1) **Initialize:**
initialize $\boldsymbol{\theta}_{m'}(0)$, $\forall m'$ and $t \leftarrow 0$;

(2) **while** stopping criterion not met **do**

(3) compute $\mathbf{F}_{(m',k'),n}$, $\forall m', k', n$ using (28), (30);

(4) solve for $\mathbf{u}_{(m,k),n}^{\mathbf{A}}$, $\mathbf{v}_{(m,k),n}^{\mathbf{A}}$, $\mathbf{u}_{(m,k),n}^{\mathbf{B}}$, $\mathbf{v}_{(m,k),n}^{\mathbf{B}}$, $\forall m, k, n$, using Theorem 2;

(5) compute $\mathcal{D}_{\boldsymbol{\theta}_{m'}} \bar{\mathbf{r}}_{(m,k),n}$, $\forall m', m, k, n$ using (33);

(6) $\boldsymbol{\theta}_{m'}(t+1) \leftarrow \boldsymbol{\theta}_{m'}(t) + \alpha(t) \sum_{m,k,n} w_{(m,k)} \mathcal{D}_{\boldsymbol{\theta}_{m'}}^T \bar{\mathbf{r}}_{(m,k),n}$, $\forall m'$ and $t \leftarrow t+1$;

(7) **end while**

(8) **return** $\mathbf{F}_{(m',k'),n}$, $\forall m'$;

ALGORITHM 1: Gradient search solution to problem (P2).

5.2. *Eigendecomposition under Kronecker Model.* Based on (45), the downlink channel matrix between BS m' and user (m, k) can be expressed as

$$\mathbf{H}_{m', (m,k), n} = \mathbf{R}_{m', (m,k)}^{1/2} \mathbf{H}_w \left[\left(\mathbf{Y}_{m', (m,k)}^{1/2} \mathbf{w}_n \right) \otimes \mathbf{T}_{m', (m,k)}^{1/2} \right], \quad (46)$$

where $\mathbf{R}_{m', (m,k)} \in \mathbb{C}^{N_U \times N_U}$, $\mathbf{T}_{m', (m,k)} \in \mathbb{C}^{N_B \times N_B}$, and $\mathbf{Y}_{m', (m,k)} \in \mathbb{C}^{L \times L}$ are the corresponding receiver, transmitter, and tap correlation matrices, respectively, and $\mathbf{H}_w \in \mathbb{C}^{N_U \times N_B}$ is a white random matrix whose elements are uncorrelated with zero mean and unit variance. Obviously, we have $\boldsymbol{\Omega}_{m', (m,k), n} = \omega_n^{Y_{m', (m,k)}} (\mathbf{T}_{m', (m,k)} \otimes \mathbf{R}_{m', (m,k)})$ that along with (15) yields

$$\begin{aligned} \boldsymbol{\Sigma}_{m', (m,k), n} &= \omega_n^{Y_{m', (m,k)}} \left(\bar{\mathbf{F}}_{m', n}^T \mathbf{T}_{m', (m,k)} \bar{\mathbf{F}}_{m', n}^* \right) \otimes \mathbf{R}_{m', (m,k)}, \\ \check{\boldsymbol{\Sigma}}_{m, (m,k), n} &= \omega_n^{Y_{m, (m,k)}} \left(\check{\mathbf{F}}_{m, k, n}^T \mathbf{T}_{m, (m,k)} \check{\mathbf{F}}_{m, k, n}^* \right) \otimes \mathbf{R}_{m, (m,k)}. \end{aligned} \quad (47)$$

Observe that the above subcovariance matrices decompose into a separated form, so their eigenvalue matrices admit the following decomposition:

$$\begin{aligned} \boldsymbol{\Lambda}_{m', (m,k), n} &= \omega_n^{Y_{m', (m,k)}} \boldsymbol{\lambda}_{m', (m,k)}^{\mathbf{R}} \boldsymbol{\lambda}_{m', (m,k), n}^{\bar{\mathbf{T}}}^T, \\ \check{\boldsymbol{\Lambda}}_{m, (m,k), n} &= \omega_n^{Y_{m, (m,k)}} \boldsymbol{\lambda}_{m, (m,k)}^{\mathbf{R}} \boldsymbol{\lambda}_{m, (m,k), n}^{\check{\mathbf{T}}}^T, \end{aligned} \quad (48)$$

in which the vector $\boldsymbol{\lambda}_{m', (m,k)}^{\mathbf{R}} \in \mathbb{C}^{N_U \times 1}$ holds the eigenvalues of $\mathbf{R}_{m', (m,k)}$, and the vectors $\boldsymbol{\lambda}_{m', (m,k), n}^{\bar{\mathbf{T}}} \in \mathbb{C}^{KN_U \times 1}$ and $\boldsymbol{\lambda}_{m, (m,k), n}^{\check{\mathbf{T}}} \in \mathbb{C}^{(K-1)N_U \times 1}$ hold the eigenvalues of $\bar{\mathbf{F}}_{m', n}^T \mathbf{T}_{m', (m,k)} \bar{\mathbf{F}}_{m', n}^*$ and $\check{\mathbf{F}}_{m, k, n}^T \mathbf{T}_{m, (m,k)} \check{\mathbf{F}}_{m, k, n}^*$, respectively. Now the eigenvalue matrices $\boldsymbol{\Lambda}_{m', (m,k)}^{\mathbf{A}}$ and $\boldsymbol{\Lambda}_{m, (m,k)}^{\mathbf{B}}$ can then be found by inserting (48) into (16).

Computationally, we no longer need to compute the eigenvalues of very large matrices $\boldsymbol{\Sigma}_{m', (m,k), n} \in \mathbb{C}^{KN_U^2 \times KN_U^2}$ and $\check{\boldsymbol{\Sigma}}_{m, (m,k), n} \in \mathbb{C}^{(K-1)N_U \times (K-1)N_U}$, but instead, we need only compute the eigenvalues of the lower dimensional matrices $\mathbf{R}_{m', (m,k)} \in \mathbb{C}^{N_U \times N_U}$, $\bar{\mathbf{F}}_{m', n}^T \mathbf{T}_{m', (m,k)} \bar{\mathbf{F}}_{m', n}^* \in \mathbb{C}^{KN_U \times KN_U}$, and $\check{\mathbf{F}}_{m, k, n}^T \mathbf{T}_{m, (m,k)} \check{\mathbf{F}}_{m, k, n}^* \in \mathbb{C}^{(K-1)N_U \times (K-1)N_U}$. Using the above eigendecomposition, the computational complexity is reduced from $\mathcal{O}(K^3 N_U^6)$ to $\mathcal{O}(K^3 N_U^2)$.

5.3. *Derivation of $\mathcal{D}_{\boldsymbol{\theta}_{m'}} \bar{\mathbf{r}}_{(m,k),n}$ under Kronecker Model.* The procedure given in Section 4.1 can be applied here with great simplifications for $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \boldsymbol{\Lambda}_{(m,k),n}^{\mathbf{A}}$ and $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \boldsymbol{\Lambda}_{(m,k),n}^{\mathbf{B}}$ which constitute the third chain of (36) computed in Section 4.1.3. Note that the third chain was the most complicated component of the derivative. Considering (16), it suffices to find $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \boldsymbol{\Lambda}_{m', (m,k), n}$ and $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \check{\boldsymbol{\Lambda}}_{m, (m,k), n}$. It can be shown from (48) that

$$\begin{aligned} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \boldsymbol{\Lambda}_{m', (m,k), n} &= \omega_n^{Y_{m', (m,k)}} \left(\mathbf{I}_{KN_U} \otimes \boldsymbol{\lambda}_{m', (m,k)}^{\mathbf{R}} \right) \\ &\quad \times \begin{bmatrix} \text{vec}^T \left\{ \bar{\boldsymbol{\mu}}_{m', (m,k), n}^1 \bar{\boldsymbol{\mu}}_{m', (m,k), n}^{1T} \right\} \\ \vdots \\ \text{vec}^T \left\{ \bar{\boldsymbol{\mu}}_{m', (m,k), n}^{KN_U} \bar{\boldsymbol{\mu}}_{m', (m,k), n}^{KN_U T} \right\} \end{bmatrix} \\ &\quad \times \left(\bar{\mathbf{F}}_{m', n}^H \mathbf{T}_{m', (m,k)}^{1/2} \otimes \mathbf{I}_{KN_U} \right) \\ &\quad \times \mathbf{K}_{N_B, KN_U} \left(\mathbf{e}_{k'}^K \otimes \mathbf{I}_{N_U} \otimes \mathbf{T}_{m', (m,k)}^{1/2} \right), \end{aligned} \quad (49)$$

$$\begin{aligned} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \check{\boldsymbol{\Lambda}}_{m, (m,k), n} &= \omega_n^{Y_{m, (m,k)}} \left(\mathbf{I}_{KN_U} \otimes \boldsymbol{\lambda}_{m, (m,k)}^{\mathbf{R}} \right) \\ &\quad \times \begin{bmatrix} \text{vec}^T \left\{ \check{\boldsymbol{\mu}}_{m, (m,k), n}^1 \check{\boldsymbol{\mu}}_{m, (m,k), n}^{1T} \right\} \\ \vdots \\ \text{vec}^T \left\{ \check{\boldsymbol{\mu}}_{m, (m,k), n}^{(K-1)N_U} \check{\boldsymbol{\mu}}_{m, (m,k), n}^{(K-1)N_U T} \right\} \end{bmatrix} \\ &\quad \times \left(\check{\mathbf{F}}_{m, n}^H \mathbf{T}_{m, (m,k)}^{1/2} \otimes \mathbf{I}_{KN_U} \right) \\ &\quad \times \mathbf{K}_{N_B, (K-1)N_U} \left(\mathbf{u}_{k, k'} \otimes \mathbf{I}_{N_U} \otimes \mathbf{T}_{m, (m,k)}^{1/2} \right), \end{aligned}$$

where $\boldsymbol{\mu}_{m', (m,k), n}^i$ and $\check{\boldsymbol{\mu}}_{m, (m,k), n}^i$ denote the i th eigenvector of $\bar{\mathbf{F}}_{m', n}^T \mathbf{T}_{m', (m,k)} \bar{\mathbf{F}}_{m', n}^*$ and $\check{\mathbf{F}}_{m, k, n}^T \mathbf{T}_{m, (m,k)} \check{\mathbf{F}}_{m, k, n}^*$, respectively, and $\mathbf{u}_{k, k'}$ was defined in (43). Using the above equations, the computational complexity is reduced from $\mathcal{O}(KN_U^3 N_B)$ to $\mathcal{O}(KN_U N_B)$.

6. Intercarrier Interference

In this section, we allow ICI among OFDM subcarriers. The case without ICI is then a special case of this scenario. The ICI occurs when there is carrier frequency offset due to synchronization errors and Doppler shifts [1, 38, 43]. This leads to loss of orthogonality among subcarriers which introduces more interference to the system. To accommodate our method to this situation, we will approximate the rate under ICI and then extend the proposed algorithm to facilitate the precoder design with statistical CSI.

6.1. System Model and Problem Formulation under ICI. When there is ICI, the received signal at user (m, k) on subcarrier n can be modelled as

$$\mathbf{y}_{(m,k),n} = \sum_{m',k',n'} \sqrt{\rho_{n',n}} \mathbf{H}_{m',(m,k),n'} \mathbf{F}_{(m',k'),n'} \mathbf{x}_{(m',k'),n'} + \mathbf{z}_{(m,k),n} \quad (50)$$

where $|\rho_{n',n}| < 1$ models the power leaked from subcarrier n' to subcarrier n due to the ICI. Following [1], we model the ICI through the normalized frequency offset denoted by $0 < \epsilon < 1$, which is the ratio of the actual frequency offset to the intercarrier spacing. It is shown in [1] that $\rho_{n',n}$ relates to the ϵ by

$$\rho_{n',n} = \left(\frac{\sin(\pi(n' - n + \epsilon))}{N \sin((\pi/N)(n' - n + \epsilon))} \right)^2 \cdot \exp \left[j\pi \left(1 - \frac{1}{N} \right) \epsilon \right]. \quad (51)$$

Here at the receiver, in addition to the summations over m', k' that model intercell and intracell interference, the received signal over subcarrier n depends also on all other subcarriers through ICI and thus the summation over n' . Note that the formulation reduces to that of the non-ICI scenario when $\epsilon = 0$, or equivalently $\rho_{n',n} = \delta_{n',n}$. The achievable instantaneous data-rate under ICI for user (m, k) on subcarrier n is then given by

$$\begin{aligned} r_{(m,k),n}^{\text{ICI}} = & + \log \left| \mathbf{I} + \sum_{m',k',n'} |\rho_{n',n}| \right. \\ & \cdot \mathbf{H}_{m',(m,k),n} \mathbf{F}_{(m',k'),n} \mathbf{F}_{(m',k'),n}^H \mathbf{H}_{m',(m,k),n}^H \left. \right| - \log |\mathbf{I}| \\ & + \sum_{(m',k',n') \neq (m,k,n)} |\rho_{n',n}| \mathbf{H}_{m',(m,k),n} \\ & \cdot \mathbf{F}_{(m',k'),n} \mathbf{F}_{(m',k'),n}^H \mathbf{H}_{m',(m,k),n}^H \left. \right|. \end{aligned} \quad (52)$$

If we define

$$\begin{aligned} \tilde{\mathbf{A}}_{(m,k),n} \triangleq & \left[\sqrt{\rho_{1,n}} \mathbf{H}_{1,(m,k),1} \bar{\mathbf{F}}_{1,1}, \dots, \sqrt{\rho_{N,n}} \mathbf{H}_{1,(m,k),N} \bar{\mathbf{F}}_{1,N}, \right. \\ & \dots, \sqrt{\rho_{n,n}} \mathbf{H}_{m,(m,k),n} \bar{\mathbf{F}}_{m,n}, \dots, \sqrt{\rho_{1,n}} \mathbf{H}_{M,(m,k),1} \bar{\mathbf{F}}_{M,1}, \dots, \end{aligned}$$

$$\left. \sqrt{\rho_{N,n}} \mathbf{H}_{M,(m,k),N} \bar{\mathbf{F}}_{M,N} \right],$$

$$\begin{aligned} \tilde{\mathbf{B}}_{(m,k),n} \triangleq & \left[\sqrt{\rho_{1,n}} \mathbf{H}_{1,(m,k),1} \bar{\mathbf{F}}_{1,1}, \dots, \sqrt{\rho_{N,n}} \mathbf{H}_{1,(m,k),N} \bar{\mathbf{F}}_{1,N}, \right. \\ & \dots, \sqrt{\rho_{n,n}} \mathbf{H}_{m,(m,k),n} \bar{\mathbf{F}}_{m,n}, \dots, \sqrt{\rho_{1,n}} \mathbf{H}_{M,(m,k),1} \bar{\mathbf{F}}_{M,1}, \dots, \\ & \left. \sqrt{\rho_{N,n}} \mathbf{H}_{M,(m,k),N} \bar{\mathbf{F}}_{M,N} \right], \end{aligned} \quad (53)$$

then we can write

$$\begin{aligned} r_{(m,k),n}^{\text{ICI}} = & \log \left| \mathbf{I} + \tilde{\mathbf{A}}_{(m,k),n} \tilde{\mathbf{A}}_{(m,k),n}^H \right| \\ & - \log \left| \mathbf{I} + \tilde{\mathbf{B}}_{(m,k),n} \tilde{\mathbf{B}}_{(m,k),n}^H \right|. \end{aligned} \quad (54)$$

Based on this result, we can utilize the same random matrix method employed for the non-ICI scenario to derive the approximated rate $\tilde{r}_{(m,k),n}^{\text{ICI}} \approx \mathbb{E}_{\mathbf{H}} [r_{(m,k),n}^{\text{ICI}}]$.

The optimization problem we consider is similar to (P1) but using the rate under ICI, that is, $r_{(m,k),n}^{\text{ICI}}$. As before, we replace the rate $r_{(m,k),n}^{\text{ICI}}$ by its approximation $\tilde{r}_{(m,k),n}^{\text{ICI}}$ for tractability. Therefore, the following optimization problem is formed:

$$\begin{aligned} \max_{\mathcal{F}} \quad & \sum_{m=1}^M \sum_{k=1}^K w_{(m,k)} \sum_{n=1}^N \tilde{r}_{(m,k),n}^{\text{ICI}} \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{n=1}^N \text{tr}(\mathbf{F}_{(m,k),n} \mathbf{F}_{(m,k),n}^H) = P_m, \quad \forall m. \end{aligned} \quad (\text{P3})$$

6.2. Second-Order Statistics under ICI. Due to the ICI, in addition to per-carrier statistics, the BSs now share the additional information of cross-carrier correlation. Specifically, the cross-carrier covariance matrices,

$$\begin{aligned} \Omega_{m',(m,k),n_1,n_2} \\ \triangleq \mathbb{E}_{\mathbf{H}} \left[\text{vec} \left\{ \mathbf{H}_{m',(m,k),n_1} \right\} \text{vec}^H \left\{ \mathbf{H}_{m',(m,k),n_2} \right\} \right], \end{aligned} \quad (55)$$

are assumed to be available at all BSs. This additional information is important for the BSs to suppress the ICI. For Kronecker model, the transmitter, receiver, and tap correlation matrices are sufficient because we have

$$\Omega_{m',(m,k),n_1,n_2} = \omega_{n_1,n_2}^{\mathbf{Y}_{m',(m,k)}} (\mathbf{T}_{m',(m,k)} \otimes \mathbf{R}_{m',(m,k)}) \quad (56)$$

with $\omega_{n_1,n_2}^{\mathbf{Y}_{m',(m,k)}} = \mathbf{w}_{n_1}^H \mathbf{Y}_{m',(m,k)} \mathbf{w}_{n_2}$.

We will now find the second-order statistics of $\tilde{\mathbf{A}}_{(m,k),n}$ and $\tilde{\mathbf{B}}_{(m,k),n}$ which depend on $\Omega_{m',(m,k),n_1,n_2}$. It can be shown that

$$\begin{aligned} \Sigma_{(m,k),n}^{\tilde{\mathbf{A}}} \\ = \text{diag} \left\{ \Sigma_{1,(m,k),n}^{\text{ICI}}, \dots, \Sigma_{m,(m,k),n}^{\text{ICI}}, \dots, \Sigma_{M,(m,k),n}^{\text{ICI}} \right\} \\ \Sigma_{(m,k),n}^{\tilde{\mathbf{B}}} \\ = \text{diag} \left\{ \Sigma_{1,(m,k),n}^{\text{ICI}}, \dots, \Sigma_{m,(m,k),n}^{\text{ICI}}, \dots, \Sigma_{M,(m,k),n}^{\text{ICI}} \right\} \end{aligned} \quad (57)$$

where $\Sigma_{m', (m,k), n}^{\text{ICI}}$ and $\check{\Sigma}_{m, (m,k), n}^{\text{ICI}}$ are block matrices holding N^2 blocks with block (n_1, n_2) given by

$$\begin{aligned} \Sigma_{m', (m,k), n}^{\text{ICI}(n_1, n_2)} &= \sqrt{\rho_{n_1, n} \rho_{n_2, n}} \\ &\times \left(\bar{\mathbf{F}}_{m', n_1}^T \otimes \mathbf{I}_{N_U} \right) \mathbf{\Omega}_{m', (m,k), n_1, n_2} \left(\bar{\mathbf{F}}_{m', n_1}^T \otimes \mathbf{I}_{N_U} \right)^H, \end{aligned} \quad (58)$$

$$\check{\Sigma}_{m, (m,k), n}^{\text{ICI}(n_1, n_2)} = \sqrt{\rho_{n_1, n} \rho_{n_2, n}} \mathbf{U}_{m, k, n_1} \mathbf{\Omega}_{m, (m,k), n_1, n_2} \mathbf{U}_{m, k, n_2}^H,$$

with $\mathbf{U}_{m, k, \bar{n}} = \bar{\mathbf{F}}_{m, \bar{n}}^T \otimes \mathbf{I}_{N_U}$, $\forall \bar{n} \neq n$, and $\mathbf{U}_{m, k, \bar{n}} = \check{\mathbf{F}}_{m, k, n}^T \otimes \mathbf{I}_{N_U}$, $\forall \bar{n} \neq n$.

Finally, we define the eigenvalue matrices as

$$\begin{aligned} \Lambda_{(m,k), n}^{\bar{\mathbf{A}}} &= \left[\Lambda_{1, (m,k), n}^{\text{ICI}}, \dots, \Lambda_{m, (m,k), n}^{\text{ICI}}, \dots, \Lambda_{M, (m,k), n}^{\text{ICI}} \right], \\ \Lambda_{(m,k), n}^{\bar{\mathbf{B}}} &= \left[\Lambda_{1, (m,k), n}^{\text{ICI}}, \dots, \check{\Lambda}_{m, (m,k), n}^{\text{ICI}}, \dots, \Lambda_{M, (m,k), n}^{\text{ICI}} \right], \end{aligned} \quad (59)$$

where $N_U \times N_U KN$ eigenvalue matrix $\Lambda_{m', (m,k), n}^{\text{ICI}}$ and $N_U \times N_U (K-1)N$ eigenvalue matrix $\check{\Lambda}_{m, (m,k), n}^{\text{ICI}}$ are formed by stacking the eigenvalues of $\Sigma_{m', (m,k), n}^{\text{ICI}}$ and $\check{\Sigma}_{m, (m,k), n}^{\text{ICI}}$ in $N_U \times 1$ sized columns, respectively.

6.3. Rate Approximation under ICI. Using the second-order statistics derived above, and by extending Theorem 2, we present the rate approximation under ICI in the following corollary.

Corollary 4. For $N_B \geq N_U \rightarrow \infty$ we have

$$\frac{r_{(m,k), n}^{\text{ICI}} - \tilde{r}_{(m,k), n}^{\text{ICI}}}{N_U} \rightarrow 0, \quad (60)$$

in which $\tilde{r}_{(m,k), n}^{\text{ICI}}$ is a deterministic function given by

$$\begin{aligned} \tilde{r}_{(m,k), n}^{\text{ICI}} &= - \sum_{j=1}^{N_U KNM} \log \left[\mathbf{v}_{(m,k), n}^{\bar{\mathbf{A}}} \right]_j \\ &+ \sum_{j=1}^{N_U (KNM-1)} \log \left[\mathbf{v}_{(m,k), n}^{\bar{\mathbf{B}}} \right]_j + \sum_{i=1}^{N_U} \left[\log \frac{\left[\mathbf{u}_{(m,k), n}^{\bar{\mathbf{B}}} \right]_i}{\left[\mathbf{u}_{(m,k), n}^{\bar{\mathbf{A}}} \right]_i} \right. \\ &\left. + N_U \log e \left(\left[\mathbf{u}_{(m,k), n}^{\bar{\mathbf{A}}} \right]_i - \left[\mathbf{u}_{(m,k), n}^{\bar{\mathbf{B}}} \right]_i \right) \right] - N_U \\ &\cdot \log (N_U (KNM - 1)) + N_U KNM \log \left(1 - \frac{1}{KNM} \right), \end{aligned} \quad (61)$$

where $\mathbf{u}_{(m,k), n}^{\bar{\mathbf{A}}} \in \mathbb{R}^{N_U}$, $\mathbf{v}_{(m,k), n}^{\bar{\mathbf{A}}} \in \mathbb{R}^{KN_U MN}$, $\mathbf{u}_{(m,k), n}^{\bar{\mathbf{B}}} \in \mathbb{R}^{N_U}$, and $\mathbf{v}_{(m,k), n}^{\bar{\mathbf{B}}} \in \mathbb{R}^{(KNM-1)N_U}$ are the solutions to the following fixed point equations:

$$\begin{aligned} \left[\mathbf{u}_{(m,k), n}^{\bar{\mathbf{A}}} \right]_i^{-1} &= N_U \left(1 + KNMN_U \left[\Lambda_{(m,k), n}^{\bar{\mathbf{A}}} \mathbf{v}_{(m,k), n}^{\bar{\mathbf{A}}} \right]_i \right), \\ \left[\mathbf{v}_{(m,k), n}^{\bar{\mathbf{A}}} \right]_j^{-1} &= N_U KNM \left(1 + N_U \left[\Lambda_{(m,k), n}^{\bar{\mathbf{A}}} \mathbf{u}_{(m,k), n}^{\bar{\mathbf{A}}} \right]_j \right), \\ \left[\mathbf{u}_{(m,k), n}^{\bar{\mathbf{B}}} \right]_i^{-1} &= N_U \left(1 + N_U (KNM - 1) \left[\Lambda_{(m,k), n}^{\bar{\mathbf{B}}} \mathbf{v}_{(m,k), n}^{\bar{\mathbf{B}}} \right]_i \right), \\ \left[\mathbf{v}_{(m,k), n}^{\bar{\mathbf{B}}} \right]_j^{-1} &= N_U (KNM - 1) \left(1 + N_U \left[\Lambda_{(m,k), n}^{\bar{\mathbf{B}}} \mathbf{u}_{(m,k), n}^{\bar{\mathbf{B}}} \right]_j \right), \end{aligned} \quad (62)$$

where i and j range according to the vector sizes.

6.4. Weighted Sum Rate Optimization under ICI. After the change of variables to spherical coordinates in a similar manner to Section 4, the KKT conditions for (P3) are $\sum_{m=1}^M \sum_{k=1}^K \omega_{(m,k)} \sum_{n=1}^N \mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k), n}^{\text{ICI}} = 0$, $\forall m'$. Under ICI, the rate is a function of precoders over all subcarriers and not just n , so the gradient is modified as

$$\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k), n}^{\text{ICI}} = \sum_{n'=1}^N \mathcal{D}_{\bar{\mathbf{f}}_{m', n'}} \tilde{r}_{(m,k), n}^{\text{ICI}} \mathcal{D}_{\theta_{m'}} \bar{\mathbf{f}}_{m', n'}. \quad (63)$$

Here, due to the ICI, we need to sum over all subcarriers to calculate the gradient function. The above derivative can be obtained with the chain rule in a similar manner to Section 4; however, we need the derivatives with respect to all subcarriers and not just n , as is seen from (63). Another difficulty involved under ICI is the calculation of $\mathcal{D}_{\mathbf{F}_{(m', k'), n'}} \Lambda_{m', (m,k), n}^{\text{ICI}}$ and $\mathcal{D}_{\check{\mathbf{F}}_{(m', k'), n'}} \check{\Lambda}_{m, (m,k), n}^{\text{ICI}}$ which are rather different from the case without ICI. We shall only derive the above two terms under ICI because other terms are straightforward to find.

For general channels, it can be shown that

$$\begin{aligned} \mathcal{D}_{\mathbf{F}_{(m', k'), n'}} \Lambda_{m', (m,k), n}^{\text{ICI}} &= \begin{bmatrix} \text{vec}^T \left\{ \boldsymbol{\gamma}_{m', (m,k), n}^1 * \boldsymbol{\gamma}_{m', (m,k), n}^{1T} \right\} \\ \vdots \\ \text{vec}^T \left\{ \boldsymbol{\gamma}_{m', (m,k), n}^{KNN_U^2} * \boldsymbol{\gamma}_{m', (m,k), n}^{KNN_U^2 T} \right\} \end{bmatrix} \begin{bmatrix} \mathbf{Y}_{n', 1} \\ \vdots \\ \mathbf{Y}_{n', N} \end{bmatrix}, \\ \mathcal{D}_{\check{\mathbf{F}}_{(m', k'), n'}} \check{\Lambda}_{m, (m,k), n}^{\text{ICI}} &= \begin{bmatrix} \text{vec}^T \left\{ \check{\boldsymbol{\gamma}}_{m, (m,k), n}^1 * \check{\boldsymbol{\gamma}}_{m, (m,k), n}^{1T} \right\} \\ \vdots \\ \text{vec}^T \left\{ \check{\boldsymbol{\gamma}}_{m, (m,k), n}^{(KN-1)N_U^2} * \check{\boldsymbol{\gamma}}_{m, (m,k), n}^{(KN-1)N_U^2 T} \right\} \end{bmatrix} \begin{bmatrix} \check{\mathbf{Y}}_{n', 1} \\ \vdots \\ \check{\mathbf{Y}}_{n', N} \end{bmatrix}, \end{aligned} \quad (64)$$

where $\boldsymbol{\gamma}_{m', (m,k), n}^i$ and $\check{\boldsymbol{\gamma}}_{m, (m,k), n}^i$ are the i th eigenvectors of $\Sigma_{m', (m,k), n}^{\text{ICI}}$ and $\check{\Sigma}_{m, (m,k), n}^{\text{ICI}}$, respectively, and $\check{\mathbf{Y}}_{n', \bar{n}}$ are given by

$$\begin{aligned}
\mathbf{Y}_{n',\tilde{n}} &= \sqrt{\rho_{n',n}\rho_{\tilde{n},n}} \mathbf{K}_{KN_U^2,KN_U^2} \left[\mathbf{0} \left(\mathbf{F}_{m',\tilde{n}}^T \otimes \mathbf{I}_{N_U} \right)^* \boldsymbol{\Omega}_{m',(m,k),n',\tilde{n}}^T \otimes \mathbf{I}_{KN_U^2} \mathbf{0} \right]^T \left(\mathbf{I}_{N_B} \otimes \mathbf{K}_{N_U,KN_U} \right) \\
&\quad \times \left(\mathbf{I}_{KN_U N_B} \otimes \text{vec} \left\{ \mathbf{I}_{N_U} \right\} \right) \mathbf{K}_{N_B,KN_U} \left(\mathbf{e}_{k'}^K \otimes \mathbf{I}_{N_U N_B} \right), \\
\check{\mathbf{Y}}_{n',\tilde{n}} &= \sqrt{\rho_{n',n}\rho_{\tilde{n},n}} \times \begin{cases} \mathbf{K}_{(K-1)N_U^2,(K-1)N_U^2}, & \tilde{n} = n' \\ \mathbf{K}_{KN_U^2,(K-1)N_U^2}, & \tilde{n} \neq n' \end{cases} \times \left[\mathbf{0} \mathbf{U}_{\tilde{n}}^T \otimes \mathbf{I}_{(K-1)N_U^2} \mathbf{0} \right]^T \left(\mathbf{I}_{N_B} \otimes \mathbf{K}_{N_U,(K-1)N_U} \right) \\
&\quad \times \left(\mathbf{I}_{(K-1)N_U N_B} \otimes \text{vec} \left\{ \mathbf{I}_{N_U} \right\} \right) \mathbf{K}_{N_B,(K-1)N_U} \left[\mathbf{0} \mathbf{I}_{N_U N_B} \mathbf{0} \right]^T.
\end{aligned} \tag{65}$$

For Kronecker channels, it can be shown that $\mathcal{D}_{\mathbf{F}_{(m',k'),n'}} \boldsymbol{\Lambda}_{m',(m,k),n}^{\text{ICI}}$ and $\mathcal{D}_{\check{\mathbf{F}}_{(m,k),n'}} \check{\boldsymbol{\Lambda}}_{m,(m,k),n}^{\text{ICI}}$ are given by

$$\begin{aligned}
\mathcal{D}_{\mathbf{F}_{(m',k'),n'}} \boldsymbol{\Lambda}_{m',(m,k),n}^{\text{ICI}} &= \left(\mathbf{I}_{KN_U N} \otimes \boldsymbol{\lambda}_{m,(m',k')}^{\mathbf{R}} \right) \\
&\quad \cdot \begin{bmatrix} \text{vec}^T \left\{ \boldsymbol{\beta}_{m',(m,k),n}^1 * \boldsymbol{\beta}_{m',(m,k),n}^{1T} \right\} \\ \vdots \\ \text{vec}^T \left\{ \boldsymbol{\beta}_{m',(m,k),n}^{KN_U N} * \boldsymbol{\beta}_{m',(m,k),n}^{KN_U N T} \right\} \end{bmatrix} \left(\mathbf{W}_{m',(m,k),n} \right) \\
&\quad \otimes \mathbf{I}_{KN_U N} \times \mathbf{K}_{LN_B,KN_U N} \mathbf{e}_{(n'-1)K+k'}^{NK} \\
&\quad \otimes \left[\sqrt{\rho_{n',n}} \left(\mathbf{K}_{N_U,L} \otimes \mathbf{I}_{N_B} \right) \right. \\
&\quad \cdot \left(\text{vec} \left\{ \mathbf{Y}_{m',(m,k)}^{1/2} \right\}^T \mathbf{w}_n \otimes \mathbf{I}_{N_B N_U} \right) \\
&\quad \left. \cdot \left(\mathbf{I}_{N_U} \otimes \mathbf{T}_{m',(m,k)}^{1/2} \right) \right], \tag{66}
\end{aligned}$$

$$\begin{aligned}
\mathcal{D}_{\check{\mathbf{F}}_{(m,k),n'}} \check{\boldsymbol{\Lambda}}_{m,(m,k),n}^{\text{ICI}} &= \left(\mathbf{I}_{N_U(KN-1)} \otimes \boldsymbol{\lambda}_{m,(m,k)}^{\mathbf{R}} \right) \\
&\quad \cdot \begin{bmatrix} \text{vec}^T \left\{ \check{\boldsymbol{\beta}}_{m,(m,k),n}^1 * \check{\boldsymbol{\beta}}_{m,(m,k),n}^{1T} \right\} \\ \vdots \\ \text{vec}^T \left\{ \check{\boldsymbol{\beta}}_{m,(m,k),n}^{N_U(KN-1)} * \check{\boldsymbol{\beta}}_{m,(m,k),n}^{N_U(KN-1)T} \right\} \end{bmatrix} \left(\check{\mathbf{W}}_{m,(m,k),n} \right) \\
&\quad \otimes \mathbf{I}_{N_U(KN-1)} \times \mathbf{K}_{LN_B,N_U(KN-1)} \mathbf{v}_{k',n'} \\
&\quad \otimes \left[\sqrt{\rho_{n',n}} \left(\mathbf{K}_{N_U,L} \otimes \mathbf{I}_{N_B} \right) \right. \\
&\quad \cdot \left(\text{vec} \left\{ \mathbf{Y}_{m,(m,k)}^{1/2} \right\}^T \mathbf{w}_n \otimes \mathbf{I}_{N_B N_U} \right) \left(\mathbf{I}_{N_U} \otimes \mathbf{T}_{m,(m,k)}^{1/2} \right) \left. \right]
\end{aligned}$$

in which

$$\begin{aligned}
\mathbf{W}_{m',(m,k),n} &= \begin{bmatrix} \sqrt{\rho_{1,n}} \mathbf{w}_1^T \mathbf{Y}_{m',(m,k)}^{1/2} \otimes \left(\bar{\mathbf{F}}_{m',1} \mathbf{T}_{m',(m,k)}^{1/2} \right)^T \\ \vdots \\ \sqrt{\rho_{N,n}} \mathbf{w}_N^T \mathbf{Y}_{m',(m,k)}^{1/2} \otimes \left(\bar{\mathbf{F}}_{m',N} \mathbf{T}_{m',(m,k)}^{1/2} \right)^T \end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
\check{\mathbf{W}}_{m,(m,k),n} &= \begin{bmatrix} \sqrt{\rho_{1,n}} \mathbf{w}_1^T \mathbf{Y}_{m,(m,k)}^{1/2} \otimes \left(\bar{\mathbf{F}}_{m,1} \mathbf{T}_{m,(m,k)}^{1/2} \right)^T \\ \vdots \\ \sqrt{\rho_{n,n}} \mathbf{w}_n^T \mathbf{Y}_{m,(m,k)}^{1/2} \otimes \left(\check{\mathbf{F}}_{m,k,n} \mathbf{T}_{m,(m,k)}^{1/2} \right)^T \\ \vdots \\ \sqrt{\rho_{N,n}} \mathbf{w}_N^T \mathbf{Y}_{m,(m,k)}^{1/2} \otimes \left(\bar{\mathbf{F}}_{m,N} \mathbf{T}_{m,(m,k)}^{1/2} \right)^T \end{bmatrix}, \tag{67}
\end{aligned}$$

and $\boldsymbol{\beta}_{m',(m,k),n}^i$ and $\check{\boldsymbol{\beta}}_{m,(m,k),n}^i$ are the i th eigenvectors of $\mathbf{W}_{m',(m,k),n}$ and $\check{\mathbf{W}}_{m,(m,k),n}$, respectively, and

$$\begin{aligned}
\mathbf{v}_{k',n'} &= \begin{cases} \mathbf{e}_{(n'-1)K+k'-1}^{NK-1}, & (n-1)K+k < (n'-1)K+k', \\ \mathbf{e}_{(n'-1)K+k'}^{NK-1}, & (n-1)K+k > (n'-1)K+k', \\ \mathbf{0}_{NK-1}, & (n-1)K+k = (n'-1)K+k'. \end{cases} \tag{68}
\end{aligned}$$

The computational complexity is greatly reduced for the Kronecker channel model from $\mathcal{O}(K^2 N_U^4 N)$ to $\mathcal{O}(LN_B KN_U N)$.

We can now propose Algorithm 2 to find a suboptimal solution to (P3). Note that this algorithm is computationally more cumbersome compared to Algorithm 1, since the number of components in $\mathcal{D}_{\boldsymbol{\theta}_{m'}} \check{\boldsymbol{\Gamma}}_{(m,k),n}^{\text{ICI}}$ is N times larger due to the ICI.

7. Simulation Results

In this section, we demonstrate the approximation accuracy and evaluate the performance of our algorithms. We consider a downlink multicell MIMO-OFDM system, where the number of cells and users in each cell is set to be 4, and the number of subcarriers is set to 8, that is, $M = 4$, $K = 4$, and $N = 8$. The number of antennas for users and BSs is set to 3 unless stated otherwise; that is, $N_B = 3$ and $N_U = 3$. The users are uniformly distributed in cells with 0.5 Km radius. Without loss of generality, the weighting for the sum rate maximization is uniform in the simulations, that is, $w_{(m,k)} = 1/NM$, $\forall m, k$. We assume the power constraints are the same

```

(1) Initialize:
    initialize  $\theta_{m'}(0)$ ,  $\forall m'$  and  $t \leftarrow 0$ ;
(2) while stopping criterion not met do
(3)   compute  $\mathbf{F}_{(m',k'),n}$ ,  $\forall m', k', n$  using (28), (30);
(4)   solve for  $\mathbf{u}_{(m,k),n}^{\mathbf{A}}$ ,  $\mathbf{v}_{(m,k),n}^{\mathbf{A}}$ ,  $\mathbf{u}_{(m,k),n}^{\mathbf{B}}$ ,  $\mathbf{v}_{(m,k),n}^{\mathbf{B}}$ ,  $\forall m, k, n$ , using Corollary 4;
(5)   compute  $\mathcal{D}_{\theta_{m'}}^T \tilde{\mathbf{r}}_{(m,k),n}^{\text{ICI}}$ ,  $\forall m', m, k, n$  using (63);
(6)    $\theta_{m'}(t+1) \leftarrow \theta_{m'}(t) + \alpha(t) \sum_{m,k,n} w_{(m,k)} \mathcal{D}_{\theta_{m'}}^T \tilde{\mathbf{r}}_{(m,k),n}^{\text{ICI}}$ ,  $\forall m'$  and  $t \leftarrow t+1$ ;
(7) end while
(8) return  $\mathbf{F}_{(m',k'),n}$ ,  $\forall m'$ ;

```

ALGORITHM 2: Gradient search solution to problem (P3).

for all BSs, that is, $P_m = P$, $\forall m$. Since the noise power was normalized to unity, the transmitter signal-to-noise ratio is $\text{SNR} = 10 \log_{10} P$.

We consider the Kronecker channel model introduced in Section 5 with correlation matrices

$$\begin{aligned}
 [\mathbf{T}_{m',(m,k)}]_{i,j} &= \left(\frac{\alpha}{d_{m',(m,k)}} \right)^2 \rho_{\mathbf{T}}^{|i-j|}, \\
 [\mathbf{R}_{m',(m,k)}]_{i,j} &= \left(\frac{\alpha}{d_{m',(m,k)}} \right)^2 \rho_{\mathbf{R}}^{|i-j|},
 \end{aligned} \tag{69}$$

in which $d_{m',(m,k)}$ denotes the distance between BS m' and user (m,k) , and $\rho_{\mathbf{T}}$, $\rho_{\mathbf{R}}$ are the spatial correlation factors as defined in [44]. The constant α is chosen so that $\mathbb{E}[(\alpha/d_{m',(m,k)})^2] = 1$. While not explicitly specified, we also choose $\rho_{\mathbf{T}}$, $\rho_{\mathbf{R}}$ to be random and uniformly distributed in the interval $[0.5, 1]$ for each user. The tap correlation matrix is chosen as

$$[\mathbf{Y}_{m',(m,k)}]_{i,j} = \sqrt{\gamma_i \gamma_j} \rho_{\mathbf{Y}}^{|i-j|}, \tag{70}$$

where $\rho_{\mathbf{Y}}$ is the tap correlation factor and $\sum_{i=1}^L \gamma_i = 1$ with $\gamma_i/\gamma_{i+1} = 3$ dB as defined in [34]. The number of channel taps is assumed to be 3 throughout our simulations; that is, $L = 3$. The ICI is modelled according to (51).

7.1. Approximation Accuracy. We compare the exact ergodic sum rate obtained by simulations and its approximation given by Theorem 2 or Corollary 4. The results are depicted in Figure 2, where it is seen that the approximations are extremely accurate. Note that although the approximations are asymptotic, exceptional accuracy is observed for our small size MIMO system ($N_B = 3$, $N_U = 3$); this fact justifies this approximation approach for practical antenna sizes. It is also seen that the approximation is accurate across a wide range of SNRs.

7.2. System Performance. Now we demonstrate the performance of our proposed method given by Algorithm 1, which is based only on statistical CSI at the BSs. For comparison, we shall also depict the results of two other methods. First, the results of the algorithm are based on perfect instantaneous CSI proposed in [12], which assumes the same system model

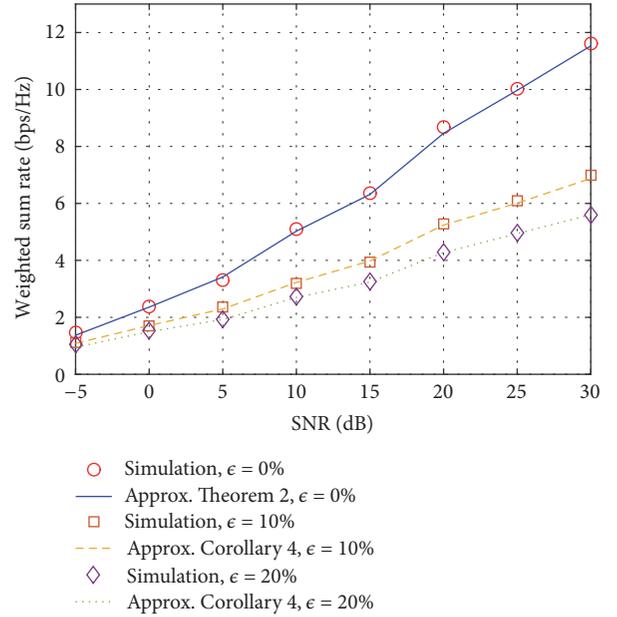


FIGURE 2: Approximation accuracy for the ergodic sum rate.

as this work but with full CSI at the BSs. Second, the results of networked MIMO are based on statistical CSI, where, in addition to the channel statistics, it is assumed that each BS has all the transmit data, so the whole system can be viewed as a MIMO super-cell with statistical CSI at the transmitter. The above two systems obviously require much heavier BS coordination, but as we shall see, our method competes with them. As mentioned before, our proposed algorithm is not limited to multicell networks with partial cooperation, so with small modifications it can be used for precoder design based on statistical CSI for networked MIMO systems.

Figure 3 depicts the results where it is observed that although our method uses statistical CSI with limited BS cooperation, the achievable weighted sum rate is comparable to that with full CSI and networked MIMO for a wide range of SNRs; note that statistical CSI incurs a much lower signaling overhead and requires less frequent update compared to full CSI. Here, we also see the results for different tap correlation factors. While tap correlations degrade system performance, using our method to incorporate the correlation information

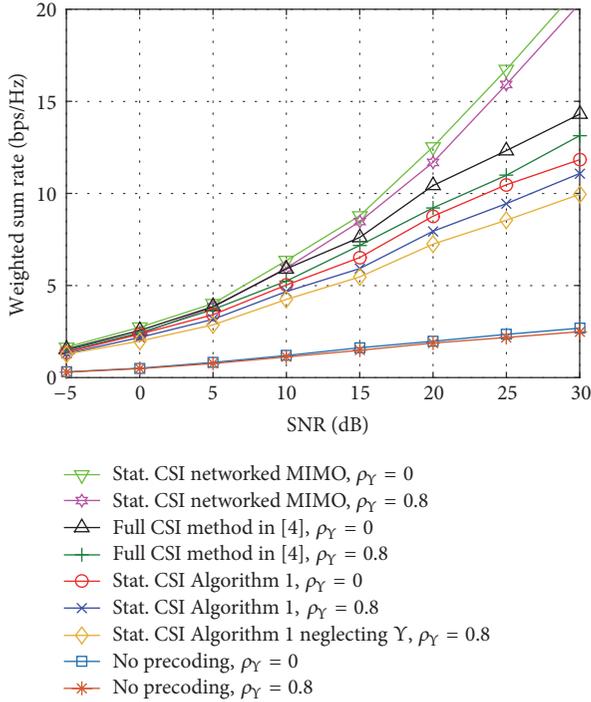


FIGURE 3: System performance under different conditions.

into precoder design alleviates the performance loss as seen in Figure 3. Moreover, it is observed that lack of knowledge about the tap correlations or neglecting them (assuming uncorrelated channel taps) leads to performance loss. When no precoders are employed (identity matrix precoding), there is a huge disadvantage, so precoder design is crucial in the MIMO-OFDM system.

7.2.1. Correlation Factors. In Figures 4 and 5, we investigate the effect of tap correlation and spatial correlation factors on the system performance. We see that both tap correlation and spatial correlation decrease the weighted sum rate. However, the spatial correlation shows more prominent effect on the sum rate. It is seen that as long as the correlation information is incorporated into precoder design via our proposed algorithm, the performance loss is not significant for a wide range of correlation degrees. It is noteworthy to mention that the performance loss becomes slightly larger at lower SNRs.

7.2.2. Number of Antennas. Figure 6 shows the sum rate as the number of transmitter antennas N_B increases; here the number of receiver antennas is fixed to $N_U = 2$. The weighted sum rate is seen to increase with N_B . As it is seen, the slope of performance increase for Algorithm 1 and that of the method in [12] based on full CSI is similar as N_B increases. This suggests that our proposed method is reliable for arbitrary large MIMO-OFDM systems.

7.3. System Performance under ICI. Now we present the system performance when there is ICI. Due to the cross

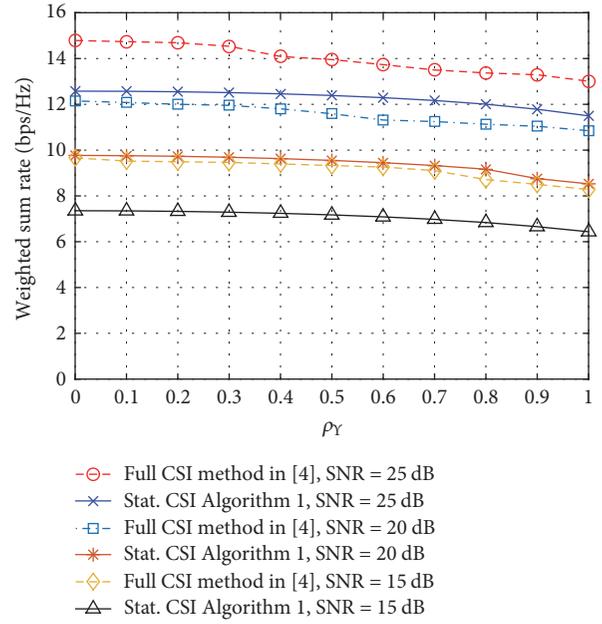


FIGURE 4: Impact of tap correlations.

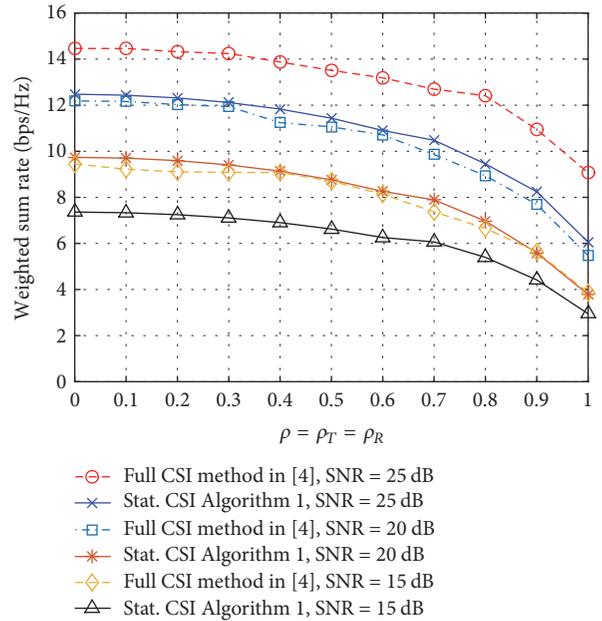


FIGURE 5: Impact of spatial correlations.

interference from other subcarriers, the sum rate decreases as is seen in Figure 7. It is seen that, for nonnegligible FO, there can be a serious decrease in the sum rate. However, smart precoder design implemented by Algorithm 2 can suppress ICI and achieve a reasonable sum rate. It is observed that the performance gain of our method over the nonprecoding scheme increases with SNR. Moreover, the advantages is more pronounced for smaller frequency offsets. However, we note that the gap between ICI-free system and system with ICI also increases with SNR.

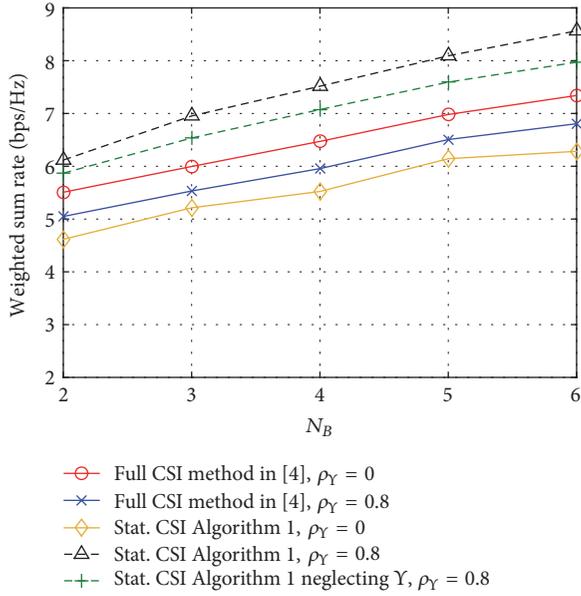


FIGURE 6: Impact of number of antennas.

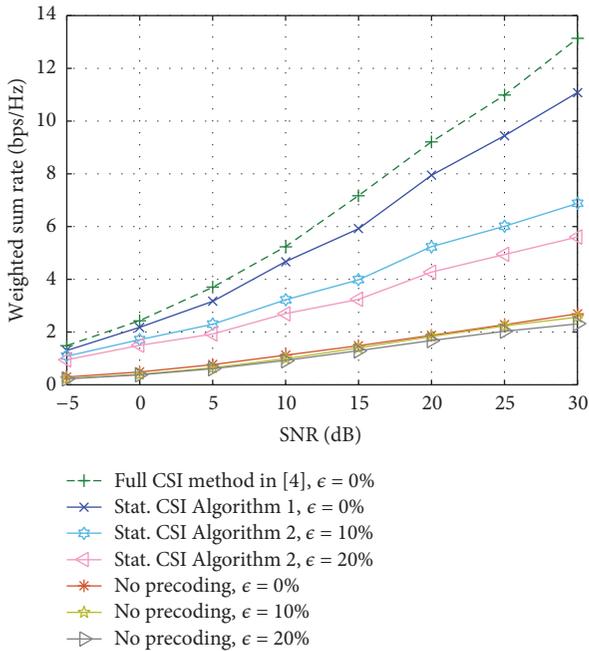


FIGURE 7: System performance under ICI.

In Figure 8, we plot the weighted sum rate versus the ICI intensity factor for SNR = 15 dB. We see that the system sum rate is highly sensitive to ICI for small ϵ .

7.4. Convergence Rate. Now we compare the convergence rates for the Algorithms 1 and 2 for SNR = 15 dB and $\epsilon = 0\%$, 1%, 10%, and 20%. The initial precoders are chosen to be the identity matrix. Under various degrees of ICI, we see from Figure 9 that Algorithm 1 always converges faster

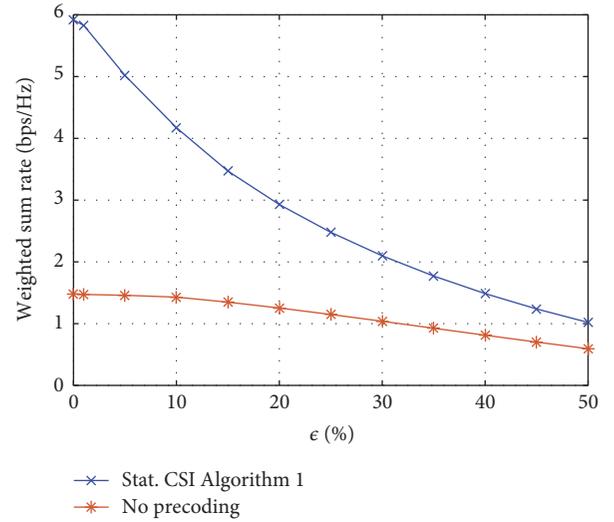


FIGURE 8: Impact of ICI intensity.

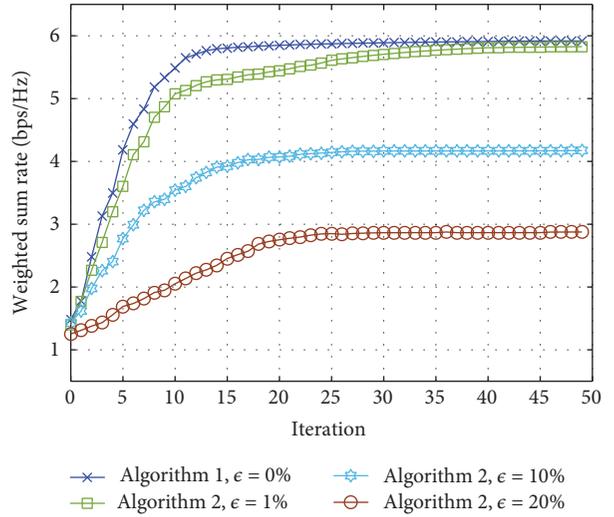


FIGURE 9: Convergence rate of Algorithms 1 and 2.

than Algorithm 2. When there is ICI, the convergence rate of Algorithm 2 is similar for various degrees of ICI.

8. Conclusions

We investigated linear precoding for downlink multicell MIMO-OFDM systems based on statistical CSI. The main contribution of this work was applying the already established results of random matrix theory to the MIMO-OFDM scenario in order to study the impact of frequency selectivity, tap correlations, and ICI on the statistical precoder design and system performance. The asymptotic approximations to the ergodic rates in ICI and ICI-free scenarios were derived, based on which, we formulated two nonconvex sum rate maximization problems and proposed locally optimum gradient based solutions to them. Simulation results showed that while spatial correlations, tap correlations, and ICI decrease

the system sum rate, our method alleviates this performance loss by incorporating the correlation information and ICI intensity information into the precoder design.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

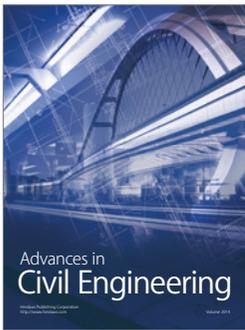
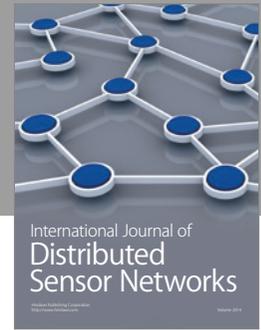
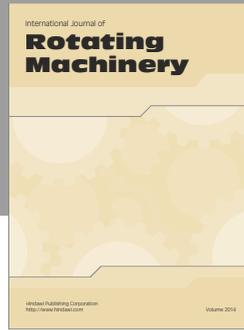
Acknowledgments

This work was supported in part by the Leading Talents Program of Guangdong Province under Grant 00201510, the Basic Research Program of Shenzhen under Grant JCYJ20151117161854942, and the Shenzhen Peacock Program under Grant KQTD2015071715073798.

References

- [1] D. N. Dao and C. Tellambura, "Intercarrier interference self-cancellation space-frequency codes for MIMO-OFDM," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 5, pp. 1729–1738, 2005.
- [2] J. M. García-Loygorri, L. de Haro, J. M. Riera, L. Cuéllar, and C. Rodríguez, "Keyholes in MIMO-OFDM: Train-to-Wayside Communications in Railway Tunnels," *Wireless Communications and Mobile Computing*, vol. 2017, pp. 1–7, 2017.
- [3] Y. Jin and X.-G. Xia, "A robust precoder design based on channel statistics for MIMO-OFDM systems with insufficient cyclic prefix," *IEEE Transactions on Communications*, vol. 62, no. 4, pp. 1249–1257, 2014.
- [4] C. Xu, J. Zhang, M. Liu, and C. Yin, "Pilot design for sparse channel estimation in large-scale MIMO-OFDM system," *International Journal of Antennas and Propagation*, vol. 2016, Article ID 6142574, 8 pages, 2016.
- [5] P. Xia, S. Zhou, and G. B. Giannakis, "Adaptive MIMO-OFDM based on partial channel state information," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 202–213, 2004.
- [6] R. El Chall, F. Nouvel, M. Héléard, and M. Liu, "Performance and complexity evaluation of iterative receiver for coded MIMO-OFDM systems," *Mobile Information Systems*, vol. 2016, Article ID 7642590, 22 pages, 2016.
- [7] J. Abdoli, M. Jia, and J. Ma, "Filtered OFDM: A new waveform for future wireless systems," in *Proceedings of the 16th IEEE International Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2015*, pp. 66–70, Sweden, July 2015.
- [8] P. N. Rani and C. S. Rani, "UFMC: The 5G modulation technique," in *Proceedings of the 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICIC)*, pp. 1–3, Chennai, December 2016.
- [9] M. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Communications Magazine*, vol. 13, no. 4, pp. 56–61, 2006.
- [10] H. Huang, M. Trivellato, A. Hottinen, M. Shafi, P. J. Smith, and R. Valenzuela, "Increasing downlink cellular throughput with limited network MIMO coordination," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 2983–2989, 2009.
- [11] W. Wu, K. Wang, W. Zeng, Z. Ding, and C. Xiao, "Cooperative multi-cell MIMO downlink precoding with finite-alphabet inputs," *IEEE Transactions on Communications*, vol. 63, no. 3, pp. 766–779, 2015.
- [12] K. Wang, X. Wang, W. Xu, and X. Zhang, "Coordinated linear precoding in downlink multicell MIMO-OFDM networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4264–4277, 2012.
- [13] L. Sanguinetti, R. Couillet, and M. Debbah, "Large system analysis of base station cooperation for power minimization," *IEEE Transactions on Wireless Communications*, vol. 99, p. 1, 2016.
- [14] D. J. Love, R. W. Heath Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1341–1365, 2008.
- [15] E. Bjornson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4298–4310, 2010.
- [16] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, 2003.
- [17] E. Yoon, J. Hansen, and A. Paulraj, "Space-frequency precoding for an OFDM based system exploiting spatial and path correlation," in *Proceedings of the GLOBECOM'04 - IEEE Global Telecommunications Conference*, pp. 436–440, December 2004.
- [18] C. Artigue and P. Loubaton, "On the precoder design of flat fading MIMO systems equipped with MMSE receivers: a large-system approach," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 57, no. 7, pp. 4138–4155, 2011.
- [19] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1165–1175, 2004.
- [20] H. Wu, X. Gao, X. Wang, and X. You, "Sum-rate-optimal precoding for multi-cell large-scale MIMO uplink based on statistical CSI," *IEEE Transactions on Communications*, vol. 63, no. 8, pp. 2924–2935, 2015.
- [21] A. M. Tulino and S. Verdú, "Random Matrix Theory and Wireless Communications," *Foundations and Trends™ in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.
- [22] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*, Cambridge University Press, Cambridge, UK, 2011.
- [23] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of antenna correlation on the capacity of multi-antenna channels," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 51, no. 7, pp. 2491–2509, 2005.
- [24] Z. Bai and J. W. Silverstein, *Spectral analysis of large dimensional random matrices*, Springer Series in Statistics, Springer, New York, Second edition, 2010.
- [25] R. Couillet, M. Debbah, and J. . Silverstein, "A deterministic equivalent for the analysis of correlated MIMO multiple access channels," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 57, no. 6, pp. 3493–3514, 2011.
- [26] J. Zhang, C.-K. Wen, S. Jin, X. Gao, and K. Wong, "On capacity of large-scale MIMO multiple access channels with distributed sets of correlated antennas," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 133–148, 2013.
- [27] F. Dupuy and P. Loubaton, "On the capacity achieving covariance matrix for frequency selective MIMO channels using the asymptotic approach," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 57, no. 9, pp. 5737–5753, 2011.

- [28] S. Chatterjee, "A generalization of the Lindeberg principle," *Annals of Probability*, vol. 34, no. 6, pp. 2061–2076, 2006.
- [29] H. Wu, L. Wang, X. Wang, and X. You, "Asymptotic and non-asymptotic analysis of uplink sum rate for relay-assisted MIMO cellular systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 6, pp. 1348–1360, 2014.
- [30] Z. Liu, Y. Xin, and G. B. Giannakis, "Space-time-frequency coded OFDM over frequency-selective fading channels," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2465–2476, 2002.
- [31] C. Xiao and Y. R. Zheng, "On the ergodic capacity of MIMO triply selective rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2272–2279, 2008.
- [32] V. K. V. Gottumukkala and H. Minn, "Ergodic capacity analysis of MISO/SIMO-OFDM with arbitrary antenna and channel tap correlation," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 3062–3068, 2013.
- [33] J. Pang, J. Li, and L. Zhao, "Ergodic capacity formula of MIMO-OFDM systems under correlated frequency selective Rayleigh fading channels," *Journal of Electronics (China)*, vol. 24, no. 5, pp. 588–592, 2007.
- [34] E. Yoon, J. Hansen, and A. Paulraj, "Space-frequency precoding with space-tap correlation information at the transmitter," *IEEE Transactions on Communications*, vol. 55, no. 9, pp. 1702–1711, 2007.
- [35] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and wiener phase noise," *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 191–193, 1995.
- [36] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2451–2464, 2002.
- [37] S. Diggavi, N. Al-Dhahir, and A. Stamoulis, "Intercarrier interference in MIMO OFDM," in *Proceedings of the 2002 International Conference on Communications (ICC 2002)*, pp. 485–489, usa, May 2002.
- [38] P. H. Moose, "Technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Transactions on Communications*, vol. 42, no. 10, pp. 2908–2914, 1994.
- [39] J. P. Kermaol, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, 2002.
- [40] V. L. Girko, "Random determinants," in *Probability theory. Mathematical statistics. Theoretical cybernetics, Vol. 24 (Russian)*, Itogi Nauki i Tekhniki, pp. 3–51, i, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1986.
- [41] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: techniques and key results," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, part 1, pp. 2740–2746, 2007.
- [42] X. Wang, V. Krishnamurthy, and J. Wang, "Stochastic gradient algorithms for design of minimum error-rate linear dispersion codes in MIMO wireless systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1242–1255, 2006.
- [43] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, 2008.
- [44] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-time Wireless Communication*, Cambridge University Press, 2003.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

