# Joint User Selection and Hybrid Precoder Design for Massive MIMO Systems

Hossein Vaezy and Steven D. Blostein, *Senior Member, IEEE*

***Abstract*—Massive multiple-input multiple-output (MIMO) systems are a cornerstone of modern wireless communication, enabling significant improvements in capacity and reliability. However, the joint optimization of user selection and hybrid precoder/decoder design remains challenging due to the complexity introduced by spatial correlation, noisy channel information, and the non-convex nature of the problem. This paper addresses these challenges by considering the downlink of multi-user massive MIMO systems. A noisy version of channel information with spatial correlation between antennas is assumed to be available at the transmitter, and an optimization problem is formulated for joint user selection and hybrid analog/digital precoder design. The total sum rate of the network is considered as a design metric that leads to non-convex and NP-hard mixed-integer optimization. To address the non-convexity, an iterative method is proposed which results in multiple simpler bounding and relaxed convex sub-problems with closed-form solutions for analog precoders/decoders, digital decoders, and user selection. As a by-product, the proposed algorithm also optimizes the number of selected users with perfect or imperfect channel state information (CSI). A generalized user selection metric is also derived for massive MIMO systems with multiple-antenna users under both perfect and imperfect CSI, and is further analyzed for specific scenarios such as ZF, MRT, block diagonalized precoders, and large-scale MIMO settings. Finally, the method is extended to finite-resolution phase shifters and assessed for Rayleigh fading channels. The simulation results show that the proposed method performs favorably compared to other recent joint user selection and precoder designs.**

***Index Terms*—Multiple antennas, hybrid precoding, user selection, imperfect channel state information.**

## I. Introduction

With growing demand for high data rates in next-generation networks, there is increasing interest in transmit and receive beamforming strategies for the cellular downlink that aim to increase the sum rate of wireless systems. In particular, spatial multiplexing at the transmitter can achieve high throughput. Emerging applications such as the internet of things (IoT) and massive machine-type communication (MMTC) need to support highly dense networks and limited resources of time, frequency spectrum, and antennas. Massive multiple-input multiple-output (MIMO) technology has emerged as a key mechanism to address this challenge, but conventional massive MIMO systems require a dedicated radio frequency (RF) chain for each antenna element, making fully digital precoding techniques costly and energy-consuming. To limit the number of RF chains, hybrid analog/digital beamforming has been proposed as a cost-effective alternative to support spatial multiplexing, whose potential has been demonstrated in many studies [1]. However, in overloaded networks, user selection is necessary, as limiting RF chains at the base station (BS) reduces the ability to simultaneously serve large numbers of user equipment (UE).

Multiple approaches to BS precoder design have been proposed for multiple-antenna systems to serve a given set of active users. The use of uplink-downlink duality has been proposed for joint precoder design and power allocation in multi-user MIMO (MU-MIMO) systems under perfect CSI [2]. Sum rate maximization with perfect CSI is an alternative approach to designing precoding vectors and user power allocation [3]. In [4], a reduced-WMMSE method is proposed, which achieves a fully digital precoder design with linear complexity. Sum rate maximization using a rate splitting multiple access (RSMA) technique has been exploited in [5] in satellite and aerial integrated networks. Beamforming design in reconfigurable intelligent surface (RIS) aided relay networks with a power minimization objective was investigated in [6].

Existing hybrid precoder methods can be categorized into joint or tandem designs where the first category jointly designs digital and analog precoders to approach fully digital performance, e.g., deep learning-aided rate maximization in single-user point-to-point networks was considered in [7]. Minimum mean squared error (MMSE), zero-forcing and matching pursuit are other criteria that have been extensively employed in the literature [8].

The second hybrid precoder design category in multi-user MIMO systems utilizes a tandem, i.e., two-stage approach. First, analog precoders are designed based on a performance metric, e.g., maximum array gain. The digital precoders are then designed to further improve sum rate or decrease interference among users. For instance, in [9], a large array gain was realized through DFT-based RF precoding/combining and the digital precoder was optimized by block diagonalization (BD)

Received 8 August 2024; revised 15 January 2025 and 8 April 2025; accepted 9 April 2025. Date of publication 21 April 2025; date of current version 12 June 2025. This work was supported in part by Huawei Technologies Canada and in part by the Natural Sciences and Engineering Research Council of Canada Discovery under Grant RGPIN-2019-06237. The associate editor coordinating the review of this article and approving it for publication was Mohammed Nabil El Korso. *(Corresponding author: Hossein Vaezy.)*

The authors are with the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: h.vaezy@gmail.com; steven.blostein@queensu.ca).

Digital Object Identifier 10.1109/TSP.2025.3562858

on the equivalent baseband channel. In [10], analog precoders are designed to match the channel matrix and digital precoders utilize BD. In [11], analog precoders are designed using a proposed stream selection method. Optimization based on zero forcing (ZF) was then used to design digital precoders. Fully digital precoders are obtained first by solving a semi-definite programming (SDP) problem in [12], and hybrid precoders are decomposed using a manifold optimization algorithm. In other tandem-type designs, a hybrid precoder/decoder for IoT systems via uplink-downlink duality was considered in [13], hybrid beamforming (BF) for MU-MIMO OFDM systems was recently proposed in [14], and energy-efficient hybrid BF in satellite-terrestrial integrated networks was addressed in [15].

Recently, hybrid precoder design was investigated using data-driven deep learning (DL) techniques in certain application-specific scenarios. An important advantage is the ability to avoid a costly CSI estimation process. In [16], channel prediction and power optimization was used to propagate CSI information over time. In [17], an unsupervised DL training method based on received signal strength information (RSSI) was proposed that avoids CSI. Data-driven DL was proposed in [18] based on implicit CSI, i.e., pilot design, channel feedback, and digital beamforming of air-to-ground channels modelled as an end-to-end neural network for massive MU-MIMO systems with orthogonal frequency division multiplexing (OFDM). In these techniques, non-linear optimization problems are reformulated as classification problems based on performance objectives. However, discretizing the general hybrid precoder design problem in a brute-force manner may be combinatorially infeasible.

The complementary problem of MU-MIMO user selection for a given precoder matrix has also been extensively investigated. For the single-antenna case, user selection assuming ZF precoding was considered in [19]. In [20], a method selects users with semi-orthogonal channels to mitigate inter-user interference. Other methods involve utilizing a sum-of-orthogonality metric as discussed in [21], as well as employing greedy user selection through the genetic algorithm and successive ZF [22]. Joint user selection with a fully digital precoder design for single-antenna users using generalized power iteration (GPI) was investigated in [23]. Sum rate maximization aimed at joint user scheduling and fully digital precoder design for single-antenna users was proposed in [24]. Heuristic user selection was combined with uplink-downlink duality in [25]. In [26], a method for joint user selection based on ZF-based precoder design was shown to lower complexity over that of dirty paper coding (DPC). Random user selection, CDF-based user selection combined with maximum channel gain selection metrics were also proposed in [27] to take into account fairness and multi-user diversity in the network.

User selection has been extended to multi-antenna users by employing the geometric angle between subspaces in [28]. Applying a similar subspace concept, the product of eigenvalues of effective channels was used in a greedy user selection method in [29] and chordal distance was applied for user selection in [30]. For multi-antenna receivers, a greedy user selection method based on BD was addressed in [31], [32] and a spatial modulation method based on singular value decomposition (SVD) was presented in [33]. Joint user selection and hybrid precoder design in MIMO-OFDM systems with perfect CSI was investigated in [34], [35].

A common precoding strategy in MU-MIMO systems is to nullify multi-user interference, which requires accurate CSI at the BS. However, in practice, this is infeasible due to pilot contamination, quantization error, and feedback limitations. This motivates the proposed investigation into the impact of imperfect channel information on performance, and the effects of a limited number of RF chains on user selection. The number of users in multi-antenna systems to fully nullify interference between users is upper bounded by $\frac{M^{RF}}{N^{RF}}$, where $M^{RF}$ and $N^{RF}$ represent numbers of RF chains at the transmitter and receiver, respectively. However, this bound assumes that perfect channel information is available and that there is no correlation between transmit and receive antennas. This motivates the proposed approach to joint user selection and hybrid precoding that incorporates channel estimation error while explicitly maximizing the total sum rate. The channel matrices model spatial correlation between antennas.

In MU-massive MIMO downlink systems, the problem of joint user selection and hybrid beamformer design that maximizes the sum rate is challenging even in the case of perfect CSI. Recent results that incorporate machine learning (ML) in antenna selection problems without considering hybrid beamformer design can be found in [36]. To the best of the authors' knowledge, ML and DL-based techniques have yet to be applied to joint user selection and hybrid precoder/decoder design for multi-antenna receivers.

The main difficulty arises from the coupling between beamforming matrices and user selection parameters, as well as constant-modulus constraints on the analog beamformers. The problem of sum-rate maximization in fully digital MU-MIMO downlink systems is also known to be NP-hard [5], [37]. The case of imperfect CSI poses further complications because an exact expression for the sum rate is not available. RIS-assisted high altitude networks with angular channel uncertainty for mmWave networks was investigated in [38]. Joint user selection and beamforming design for the case of single-antenna users was considered in [23] with imperfect CSI. While several studies have addressed aspects of joint user selection and hybrid precoding, they are often constrained by assumptions such as perfect CSI, single-antenna users, or uniform linear array beamforming strategies. To the best of the authors' knowledge, no prior work provides a comprehensive framework that dynamically optimizes user selection and hybrid precoder/decoder design for multi-antenna receivers, accounting for practical constraints such as imperfect CSI and finite-resolution phase shifters. Key features of previous approaches are summarized in Table I.

In this paper, a new optimization problem is formulated that integrates joint user selection and hybrid digital/analog precoding, aiming to explicitly maximize the sum-rate under a total transmit power constraint. The proposed approach incorporates

TABLE I
COMPARISON AMONG PREVIOUS APPROACHES

| | Antenna | Precoder Structure | CSI | UE Selection | Design |
|---|---|---|---|---|---|
| [2] | MIMO | Given-Fully digital | Perfect | × | sum rate |
| [19] | MISO | Given-Fully digital | Perfect | ✓ | ZF |
| [20], [21] | MIMO | Given-Fully digital | Perfect | ✓ | ZF |
| [22] | MIMO | Given-Fully digital | Perfect | ✓ | successive ZF |
| [39] | MISO | Fully digital | Perfect | ✓ | ZF |
| [40] | MISO | Fully digital | Perfect | ✓ | - |
| [25] | MISO | Fully digital | Perfect | ✓ | minimum weighted rate |
| [23] | MISO | Fully digital | Imperfect | ✓ | sum rate |
| [28] | MIMO | Fully digital | Perfect | ✓ | Greedy-subspace angles |
| [29] | MIMO | Fully digital | Perfect | ✓ | Greedy-product of eigenvalues |
| [41], [42] | MIMO | Fully digital | Perfect | ✓ | Greedy-BD |
| [1] | MISO | Hybrid | Perfect | × | sum rate |
| [43] | MISO | Hybrid | Perfect | × | ZF-MRT |
| [9],[10] | Massive MIMO | Hybrid | Perfect | × | BD |
| [11] | Massive MIMO | Hybrid | Perfect | × | ZF |
| [12] | Massive MIMO | Hybrid | Perfect | × | SRM |
| [44] | Massive MIMO | Hybrid | Perfect | × | SRM-MMSE |
| [33] | Massive MIMO | Hybrid | Perfect | ✓ | Greedy-SVD-GMD |
| [34], [35] | Massive MIMO | Hybrid | Perfect | ✓ | Greedy-BD |

imperfect CSI by leveraging the first- and second-order statistics of the error matrices at the BS. The method features a closed-form solution for analog precoders/decoders and user selection, combined with a convex optimization framework for digital precoder design. The main contributions of this work are as follows:

- Novel Joint Iterative Algorithm: An iterative algorithm is proposed that jointly optimizes user selection and hybrid beamformer design, addressing the challenges of channel estimation error at the BS. This ensures an efficient and tractable solution for sum-rate maximization. A systematic evaluation is conducted to investigate the impact of key design parameters, including the number of RF chains, channel estimation accuracy, phase resolution, and BS transmit power, under diverse scenarios.
- Closed-Form Solutions: The proposed method includes closed-form solutions for analog precoders/decoders and user selection in each iteration, significantly reducing computational complexity, enabling scalable precoder design, and introducing a generalized user selection metric for the cases of ideal perfect CSI and practical imperfect CSI.
- Adaptation to Hardware Constraints: The algorithm is adapted to account for practical limitations of finite phase resolution of analog phase shifters, with its effect on performance extensively evaluated through simulations with insights provided, highlighting its implications for the practicality and deployment of massive MIMO systems.
- Dynamic User Group Optimization: Unlike most existing methods, the proposed approach dynamically optimizes the user group size, determining the number of selected UEs to maximize system performance.
- Complexity and Convergence Analysis: The time complexity is determined by analysis of the number of operations involved in each step of the algorithm, including digital beamforming, analog beamforming, user selection, and the solution of convex sub-problems. The overall complexity takes into account the number of iterations required to achieve convergence. The convergence analysis of the iterative method reveals an incrementally increasing behavior of the objective function.

In the sequel, the system model and problem formulation for imperfect channel estimation are described in Sections II and III, respectively. Computational complexity, convergence, and extension to limited phase resolution are also included in Section III. Simulation results that evaluate system performance are presented in Section IV.

**Notation**: Bold lowercase letters are used for vectors and bold uppercase letters for matrices. $(\cdot)^T$ and $(\cdot)^H$ denote transpose and complex conjugate operations, respectively. $\mathbf{A}(m:n,:)$ refers to the submatrix consisting of the *m*-th to *n*-th rows of matrix $\mathbf{A}$. $\mathbf{A}(i,j)$ denotes the $(i,j)$ entries of matrix $\mathbf{A}$. $\mathbb{C}$ denotes the set of complex numbers. $\mathbf{I}_N$ represents the identity matrix in $\mathbb{C}^{N\times N}$. $\mathbf{0}_{N\times M}$ denotes the all-zero matrix of size $N\times M$. $tr(\cdot)$ and $|\cdot|$ are the trace and the determinant of a square matrix, respectively. The $vec(\cdot)$ operator represents column-wise vectorization, $\otimes$ denotes the Kronecker product, $\Re(\cdot)$ is the real part, and $(\cdot)^\star$ represents complex conjugate.

## II. SYSTEM MODEL

The downlink of a MU-Massive MIMO network is illustrated in Fig. 1, where a BS is equipped with *M* antennas and $M^{RF}$ RF chains that transmit to *K* UE receivers. Each UE has *N* antennas and $N^{RF}$ RF chains to support *d* data symbols, resulting in a total of *Kd* data symbols transmitted by the BS. Due to the limited number of RF chains, the number of transmitted symbols is restricted by $Kd \le M^{RF} \le M$ and $d \le N^{RF} \le N$ for BS and UEs, respectively. Due to the limited number of transmit RF chains, it is not possible to implement fully digital precoding, which requires one dedicated RF chain per antenna. Consequently, both the BS and UE employ a hybrid digital/analog

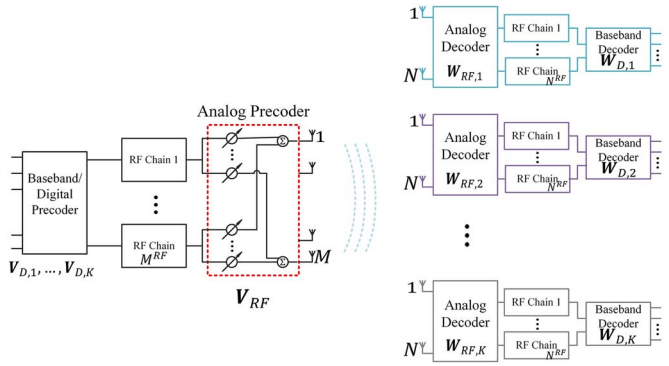


Fig. 1. Shown is the downlink channel hybrid precoder/decoder structure at the BS and at the UEs.

precoding structure. We denote $\mathcal{K} = \{1, \ ,2, \ldots, K\}$ as the set of all UEs.

For the fully digital deployment, the channel matrix between the BS and UE $k$ is denoted by $\mathbf{H}_k \in \mathbb{C}^{N\times M}$, while the precoding matrix and decoding matrix at UE $k$ are denoted by $\mathbf{V}_k \in \mathbb{C}^{M\times d}$ and $\mathbf{W}_k \in \mathbb{C}^{N\times d}$, respectively. To represent spatial correlation between antennas at the BS and UEs, the well-known Kronecker model is used. Specifically, $\mathbf{H}_k$ is given by $\mathbf{R}_{r,k}^{\frac{1}{2}}\hat{\mathbf{H}}_k\mathbf{R}_t^{\frac{1}{2}}$ [45], where $\hat{\mathbf{H}}_k \in \mathbb{C}^{N\times M}$ is an uncorrelated complex channel matrix whose entries are independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian random variables with zero mean and unit variance, i.e., $\hat{\mathbf{H}}_k \sim \mathcal{CN}(\mathbf{0}_{N\times M}, \mathbf{I}_M \otimes \mathbf{I}_N)$. $\mathbf{R}_t \in \mathbb{C}^{M\times M} \succeq \mathbf{0}$ and $\mathbf{R}_{r,k} \in \mathbb{C}^{N\times N} \succeq \mathbf{0}$ denote spatial correlation matrices between the BS transmit antennas and between receive antennas at the $k$th UE, respectively.

In the hybrid precoding structure, the BS first processes the data symbols intended for UEs, e.g., user $k$ at baseband as a digital precoder, $\mathbf{V}_{D,k} \in \mathbb{C}^{M^{RF}\times d}$, and then up converts the processed symbols to the carrier frequency via the $M^{RF}$ RF chains. The modified symbols are then processed by analog RF precoders, $\mathbf{V}_{RF} \in \mathbb{C}^{M\times M^{RF}}$ which are implemented using analog phase shifters with $\left|\mathbf{V}_{RF}(i,j)\right|^2 = \frac{1}{M}$. The baseband digital precoder $\mathbf{V}_{D,k}$, on the other hand, enables both amplitude and phase modification.

Denoting $\boldsymbol{s}_k \in \mathbb{C}^d$ as the symbol vector transmitted to UE $k$, which receives the signal,

$$\boldsymbol{y}_k = \sum_{\ell\in\mathcal{K}} x_\ell \mathbf{H}_k\mathbf{V}_\ell\boldsymbol{s}_\ell + \boldsymbol{n}_k, \qquad \forall k \in \mathcal{K}, \tag{1}$$

where $\boldsymbol{n}_k \in \mathbb{C}^N$ represents noise at UE $k$ and is assumed to be circularly symmetric complex Gaussian with covariance matrix $\mathrm{E}\{\boldsymbol{n}_k\boldsymbol{n}_k{}^H\} = \sigma^2\mathbf{I}_N$. In (1), $x_k$ is a user selection indicator function with value '1' if user $k \in \mathcal{K}$ is selected for transmission and '0' otherwise.

Receiver UE $k$ first uses analog decoder, $\mathbf{W}_{RF,k} \in \mathbb{C}^{N\times N^{RF}}$, implemented by phase shifters such that $|\mathbf{W}_{RF,k}(i,j)|^2 = \frac{1}{N}$, and then down converts the signals to baseband using $N^{RF}$ RF chains. Using digital combiner $\mathbf{W}_{D,k} \in \mathbb{C}^{N^{RF}\times d}$, the estimated symbols

$$\begin{aligned}\hat{\boldsymbol{s}}_k = {} & x_k\mathbf{W}_{D,k}^H\mathbf{W}_{RF,k}^H\mathbf{H}_k\mathbf{V}_{RF}\mathbf{V}_{D,k}\boldsymbol{s}_k \\ & + \sum_{\ell\in\mathcal{K}_{|k}} x_\ell\mathbf{W}_{D,k}^H\mathbf{W}_{RF,k}^H\mathbf{H}_k\mathbf{V}_\ell\boldsymbol{s}_\ell + \mathbf{W}_{D,k}^H\mathbf{W}_{RF,k}^H\boldsymbol{n}_k,\end{aligned} \tag{2}$$

where $\mathcal{K}_{|k}$ is the set of UEs except UE $k$. The first term in (2) corresponds to the desired signal while the second and last terms represent the interfering signals and noise, respectively.

By assuming Gaussian transmitted symbols, the rate for user $k$, $R_k$, is given as follows [46]:

$$R_k = \log_2\det\left(\mathbf{I}_d + x_k^2\mathbf{V}_k^H\mathbf{H}_k^H\mathbf{W}_k\mathbf{C}_k^{-1}\mathbf{W}_k^H\mathbf{H}_k\mathbf{V}_k\right), \tag{3}$$

where

$$\mathbf{C}_k = \sigma^2\mathbf{W}_k^H\mathbf{W}_k + \sum_{\ell\in\mathcal{K}_{|k}}^{K} x_\ell^2\mathbf{W}_k^H\mathbf{H}_k\mathbf{V}_\ell\mathbf{V}_\ell^H\mathbf{H}_k^H\mathbf{W}_k, \ \forall k \in \mathcal{K}.$$

By substituting $\mathbf{V}_k = \mathbf{V}_{RF}\mathbf{V}_{D,k}$ and $\mathbf{W}_k = \mathbf{W}_{RF,k}\mathbf{W}_{D,k}$, the rate expression for the hybrid structure are expressed in terms of digital and analog precoders/decoders defined earlier.

## III. Imperfect Channel Estimation

Precoding in the downlink and signal detection in the uplink require CSI at the BS. Although it is commonly assumed that the channel is perfectly known at both the BS and UEs, this is not realistic. In TDD massive MIMO systems, CSI acquisition relies on the assumption of channel reciprocity between the uplink and the downlink, which typically contains calibration errors. Additionally, limited channel coherence time may result in pilot contamination from neighboring cells. Similarly, in FDD systems, CSI is inherently noisy due to the feedback error, finite-resolution ADC, and quantization error [45].

Linear minimum mean squared error (LMMSE) channel estimation is commonly employed, which can provide high performance at low complexity. Following the channel estimation model for MU-MIMO systems in [47], the noisy channel matrix is modelled as follows:

$$\mathbf{H}_k = \mathbf{R}_{r,k}^{\frac{1}{2}}(\sqrt{1-\rho^2}\hat{\mathbf{H}}_k + \rho\boldsymbol{\Delta}_k)\mathbf{R}_t^{\frac{1}{2}}, \tag{4}$$

where $\hat{\mathbf{H}}_k$ and $\boldsymbol{\Delta}_k$ represent the estimated channel and channel estimation error for the $k$th UE, respectively, where the estimation accuracy parameter $0 < \rho \leq 1$, corresponds to partial channel knowledge at the BS. Here, $\rho = 0$ represents perfect CSI. The channel estimation error matrices are assumed to be uncorrelated with $\hat{\mathbf{H}}_k$ and elements of both $\hat{\mathbf{H}}_k$ and $\boldsymbol{\Delta}_k$ are assumed to be independent and identically distributed (i.i.d.) zero mean with circularly symmetric complex Gaussian distribution with unit variance. In general, $\rho$ may be system dependent, e.g., pilot sequence power, and for an analog feedback link, $\rho$ is a function of the number of channel uses per channel coefficient and power of feedback links [45]. For simplicity, $\rho$ is assumed here to be a constant parameter. However, specific scenarios can be modelled by expressing $\rho$ as a known

function of system parameters. For notational simplicity, define $\bar{\mathbf{H}}_k \triangleq \sqrt{1-\rho^2}\mathbf{R}_{r,k}^{\frac{1}{2}}\hat{\mathbf{H}}_k\mathbf{R}_t^{\frac{1}{2}}$ and $\bar{\mathbf{\Delta}}_k \triangleq \rho\mathbf{R}_{r,k}^{\frac{1}{2}}\mathbf{\Delta}_k\mathbf{R}_t^{\frac{1}{2}}$, then rewrite (4) as

$$\mathbf{H}_k = \bar{\mathbf{H}}_k + \bar{\mathbf{\Delta}}_k, \tag{5}$$

and the received signal at the $k$th UE can be expressed as

$$\boldsymbol{y}_k = \sum_{\ell\in\mathcal{K}} x_\ell(\bar{\mathbf{H}}_k + \bar{\mathbf{\Delta}}_k)\mathbf{V}_{RF}\mathbf{V}_{D,\ell}\boldsymbol{s}_\ell + \boldsymbol{n}_k. \tag{6}$$

Consequently, the estimated symbols at the $k$th UE for the imperfect channel become

$$\hat{\boldsymbol{s}}_k = \sum_{\ell\in\mathcal{K}} x_\ell \mathbf{W}_{D,k}^H\mathbf{W}_{RF,k}^H(\bar{\mathbf{H}}_k + \bar{\mathbf{\Delta}}_k)\mathbf{V}_\ell\boldsymbol{s}_\ell + \mathbf{W}_{D,k}^H\mathbf{W}_{RF,k}^H\boldsymbol{n}_k.$$

### A. Problem Formulation

The objective is to select a subset of UEs for simultaneous transmission and design hybrid precoders and decoders to maximize the total sum rate in the presence of channel estimation error with known error statistics. The objective function is therefore the sum rate averaged over the channel estimation errors, $\mathrm{E}_\Delta\{R_k\}$, which depends only on the first and second-order statistics of the channel estimation errors.

The technique used to incorporate error statistics uses the well-known approximation [48]:

$$\begin{aligned}\mathrm{E}_\Delta\{R_k\} = \bar{R}_k \approx \log\det\Bigg[&\mathbf{I}_d \\ &+ \mathrm{E}_\Delta\Big\{x_k^2\mathbf{W}_k^H(\bar{\mathbf{H}}_k + \bar{\mathbf{\Delta}}_k)\mathbf{V}_k\mathbf{V}_k^H(\bar{\mathbf{H}}_k + \bar{\mathbf{\Delta}}_k)^H\mathbf{W}_k\Big\} \\ &\mathrm{E}_\Delta\Big\{\mathbf{W}_k^H\big(\sigma^2\mathbf{I} \\ &+ \sum_{\ell\in\mathcal{K}_{|k}} x_\ell^2(\bar{\mathbf{H}}_k + \bar{\mathbf{\Delta}}_k)\mathbf{V}_\ell\mathbf{V}_\ell^H(\bar{\mathbf{H}}_k + \bar{\mathbf{\Delta}}_k)^H\big)\mathbf{W}_k\Big\}^{-1}\Bigg].\end{aligned} \tag{7}$$

To calculate the expectations in (7) the following is employed:

*Lemma 1:* The average rate for the $k$th UE under the channel error model (5) is given by:

$$\begin{aligned}\bar{R}_k \approx \log\det\Bigg[&\mathbf{I}_d \\ &+ \Big\{x_k^2\mathbf{W}_k^H(\bar{\mathbf{H}}_k\mathbf{V}_k\mathbf{V}_k^H\bar{\mathbf{H}}_k^H + \rho^2\|\mathbf{V}_k\|_F^2\mathbf{I}_N)\mathbf{W}_k\Big\} \\ &\Big\{\mathbf{W}_k^H\Big(\sigma^2\mathbf{I}_N \\ &+ \sum_{\ell\in\mathcal{K}_{|k}} x_\ell^2\big(\bar{\mathbf{H}}_k\mathbf{V}_\ell\mathbf{V}_\ell^H\bar{\mathbf{H}}_k^H + \rho^2\|\mathbf{V}_\ell\|_F^2\mathbf{I}_N\big)\Big)\mathbf{W}_k\Big\}^{-1}\Bigg].\end{aligned} \tag{8}$$

*Proof:* See Appendix A. ∎

The aim is to design digital/analog precoders, decoders, and select a subset of users to maximize the total sum rate subject to a total BS power constraint while accounting for imperfect CSI. This is expressed as the following optimization problem:

$$\max_{\{\mathbf{x},\mathbf{V}_{RF},\mathbf{V}_D,\mathbf{W}_{RF}\mathbf{W}_D\}} \quad \sum_{\ell\in\mathcal{K}} \alpha_\ell\bar{R}_\ell \tag{9}$$

$$\text{subject to} \quad \sum_{\ell\in\mathcal{K}} tr(\mathbf{V}_{RF}\mathbf{V}_{D,\ell}\mathbf{V}_{D,\ell}^H\mathbf{V}_{RF}^H) \le P_m, \tag{9a}$$

$$x_\ell \in \{0,1\},\ \forall\ell, \tag{9b}$$

$$|\mathbf{V}_{RF}(i,j)|^2 = \frac{1}{M},\ \forall i,j \tag{9c}$$

$$|\mathbf{W}_{RF,k}(i,j)|^2 = \frac{1}{N},\ \forall i,j,k. \tag{9d}$$

where $P_m$ is the total BS power constraint, and rate weight $\alpha_k \in [0,1]$ reflects user $k$'s priority. Also, the set of all user selection parameters, precoders, and decoders are denoted by $\mathbf{x}, \mathbf{V}_D, \mathbf{W}_D, \mathbf{W}_{RF}$, respectively.

The optimization problem (9) is non-convex due to the presence of integer and constant modulus variables, making the optimal solution hard to find and computationally challenging. Even if the set of users is fixed, their mutual interference results in a non-convex problem [47]. To address these challenges, the majorization-minimization (MM) technique, with its numerous associated advantages over other techniques due to its use of convexity arguments [48], is applied to the non-convex problem (9). In the MM technique, a surrogate function is constructed, known as a minorizer. An iterative method that reduces complexity is found that converges to a locally optimum point.

To apply the MM technique to Problem (9), a minorizer is required for the objective function. To proceed, first define $\mathbf{E}_k$ in (10), and define

$$\mathbf{U} \triangleq \begin{bmatrix}\mathbf{I}_{(N+d)} & \mathbf{0}_{(N+d)\times d}\end{bmatrix}^T \in \mathbb{R}^{(N+2d)\times(N+d)}. \tag{11}$$

Using the matrix inversion lemma [49], one can recast the average rate, i.e., $\bar{R}_k$ in (8) as

$$\bar{R}_k = \log\det\left[\mathbf{U}^T\mathbf{E}_k^{-1}\mathbf{U}\right],\ \forall k = 1,\ldots,K. \tag{12}$$

*Lemma 2:* Let $\mathbf{U}$ denote a full column-rank matrix. Then, the function $\log|\mathbf{U}^H\mathbf{X}^{-1}\mathbf{U}|$ is convex for any $\mathbf{X} \succ 0$.

*Proof:* See [50]. ∎

Therefore, using Lemma 2 and the first-order condition for convex functions [50], a minorizer at a given $\bar{\mathbf{E}}_k$ for $\bar{R}_k$ can be derived using its supporting hyperplane as follows:

$$\log|\mathbf{U}^T\mathbf{E}_k^{-1}\mathbf{U}| \ge \log|\mathbf{U}^T\bar{\mathbf{E}}_k^{-1}\mathbf{U}| - \mathrm{tr}[\mathbf{F}_k(\mathbf{E}_k - \bar{\mathbf{E}}_k)], \tag{13}$$

$$\mathbf{E}_k \triangleq \begin{bmatrix}\mathbf{I}_{N+d} & x_k\begin{bmatrix}\bar{\mathbf{H}}_k\mathbf{V}_k & \rho\|\mathbf{V}_k\|_F\mathbf{I}_N\end{bmatrix}^H\mathbf{W}_k \\ x_k\mathbf{W}_k^H\begin{bmatrix}\bar{\mathbf{H}}_k\mathbf{V}_k & \rho\|\mathbf{V}_k\|_F\mathbf{I}_N\end{bmatrix} & \mathbf{W}_k^H\Big(\sigma^2\mathbf{I}_N + \sum_{\ell\in\mathcal{K}} x_\ell^2(\bar{\mathbf{H}}_k\mathbf{V}_\ell\mathbf{V}_\ell^H\bar{\mathbf{H}}_k^H + \rho^2\|\mathbf{V}_\ell\|_F^2\mathbf{I}_N)\Big)\mathbf{W}_k\end{bmatrix} \in \mathbb{C}^{(N+2d)\times(N+2d)} \tag{10}$$

where $\bar{\mathbf{E}}_k$ is the value of $\mathbf{E}_k$ in the previous iteration and,

$$\mathbf{F}_k \triangleq \bar{\mathbf{E}}_k^{-1}\mathbf{U}(\mathbf{U}^T\bar{\mathbf{E}}_k^{-1}\mathbf{U})^{-1}\mathbf{U}^T\bar{\mathbf{E}}_k^{-H} \quad \in \mathbb{C}^{(N+2d)\times(N+2d)}. \tag{14}$$

Using the minorizer in (13) and applying the MM technique, after removing the constant terms, the optimization problem (9) is converted into the following minimization problem at each iteration (See Appendix B for more details):

$$\begin{aligned}\min_{\{x_k,\mathbf{V}_D,\mathbf{V}_{RF},\mathbf{W}_D,\mathbf{W}_{RF}\}} & \Bigg\{ \sum_{k\in\mathcal{K}} 2x_k\Re\Big[\mathrm{tr}\Big(\mathbf{F}_{12,k}(1:d,:)\mathbf{W}_k^H\bar{\mathbf{H}}_k\mathbf{V}_k\Big) \\ & + \rho\|\mathbf{V}_k\|_F\mathrm{tr}\Big(\mathbf{F}_{12,k}(d+1:N+d,:)\mathbf{W}_k^H\Big)\Big] \\ & + \mathrm{tr}\Big(\mathbf{W}_k\mathbf{F}_{22,k}(\sigma^2+\rho^2\sum_{\ell\in\mathcal{K}}x_\ell^2\|\mathbf{V}_\ell\|_F^2)\mathbf{W}_k^H\Big) \\ & + \mathrm{tr}\Big(\sum_{\ell\in\mathcal{K}}x_\ell^2\mathbf{F}_{22,k}\mathbf{W}_k^H\bar{\mathbf{H}}_k\mathbf{V}_\ell\mathbf{V}_\ell^H\bar{\mathbf{H}}_k^H\mathbf{W}_k\Big)\Bigg\}, \\ \text{subject to} \quad & \text{(9a)}, \ \text{(9b)}, \ \text{(9c)}, \ \text{(9d)}, \end{aligned} \tag{15}$$

where $\mathbf{W}_k = \mathbf{W}_{RF,k}\mathbf{W}_{D,k}$, $\mathbf{V}_k = \mathbf{V}_{RF}\mathbf{V}_{D,k}$ and $\mathbf{F}_{12,k}$ and $\mathbf{F}_{22,k}$ are sub-matrices given by

$$\mathbf{F}_k \triangleq \begin{bmatrix}\mathbf{F}_{11,k} & \mathbf{F}_{12,k}\\ \mathbf{F}_{21,k} & \mathbf{F}_{22,k}\end{bmatrix}, \ \forall k\in\mathcal{K}. \tag{16}$$

Due to the highly coupled precoder and decoder matrices, constant modulus constraints as well as integer variables, $x_\ell$, the transformed optimization problem (15) is still jointly non-convex for variables $\{x_k,\mathbf{V}_{D,k},\mathbf{V}_{RF},\mathbf{W}_{D,k},\mathbf{W}_{RF,k}\}_k$ and cannot be efficiently solved by conventional convex optimization methods. However, the optimization problem (15) is convex in the variables $x_k$ and $\mathbf{W}_{D,k}$ individually. To handle this intractable non-convex problem, iterative alternating optimization is employed by decomposing the optimization problem into multiple sub-problems. In the sequel, the optimization problems are solved in terms of each parameter individually while considering the other variables as fixed.

### B. Solving the Optimization Problem

*1) Optimizing Digital Precoder:* Given a set of fixed variables $\{x_k,\mathbf{V}_{RF},\mathbf{W}_{D,k},\mathbf{W}_{RF,k}\}_k$, the optimization problem with respect to the variables $\{\mathbf{V}_{D,k}\}_k$ is described as:

$$\begin{aligned}\min_{\mathbf{V}_D} \quad & \sum_{k\in\mathcal{K}}\Big\{2\Re[\boldsymbol{a}_{1,k}^H\boldsymbol{v}_{D,k}+a_{2,k}\|(\mathbf{I}_d\otimes\mathbf{V}_{RF})\boldsymbol{v}_{D,k}\|_2] \\ & + x_k^2\boldsymbol{v}_{D,k}^H(\mathbf{I}_d\otimes\boldsymbol{\Phi}_v)\boldsymbol{v}_{D,k}\Big\} \\ \text{subject to} \quad & \sum_{k\in\mathcal{K}}\|(\mathbf{I}_d\otimes\mathbf{V}_{RF})\boldsymbol{v}_{D,k}\|_2^2\le P_m, \end{aligned} \tag{17}$$

where

$$\begin{aligned}\boldsymbol{\Phi}_v &\triangleq \sum_{\ell\in\mathcal{K}}\Big\{\mathbf{V}_{RF}^H(\bar{\mathbf{H}}_\ell^H\mathbf{W}_\ell\mathbf{F}_{22,\ell}\mathbf{W}_\ell^H\bar{\mathbf{H}}_\ell \\ &\qquad + \rho^2\mathrm{tr}(\mathbf{F}_{22,\ell}\mathbf{W}_\ell^H\mathbf{W}_\ell))\mathbf{V}_{RF}\Big\} \\ \boldsymbol{a}_{1,k} &\triangleq x_k.\mathrm{vec}\big(\mathbf{V}_{RF}^H\bar{\mathbf{H}}_k^H\mathbf{W}_k\mathbf{F}_{12,k}(1:d,:)^H\big), \\ a_{2,k} &\triangleq \rho x_k.\mathrm{tr}\big(\mathbf{F}_{12,k}(d+1:N+d,:)\mathbf{W}_k^H\big). \end{aligned} \tag{18}$$

The problem described in (17) is convex if $a_{2,k}\ge 0$. For negative $a_{2,k}$, the objective function is the difference of two convex functions. To handle the non-convexity, a first-order condition for convex functions is used to bound $\|(\mathbf{I}_d\otimes\mathbf{V}_{RF})\boldsymbol{v}_{D,k}\|_2$ in each iteration by $\frac{\bar{\boldsymbol{v}}_{D,k}^H(\mathbf{I}_d\otimes\mathbf{V}_{RF}^H\mathbf{V}_{RF})}{\sqrt{\bar{\boldsymbol{v}}_{D,k}^H(\mathbf{I}_d\otimes\mathbf{V}_{RF}^H\mathbf{V}_{RF})\bar{\boldsymbol{v}}_{D,k}}}\boldsymbol{v}_{D,k}$. Here, $\bar{\boldsymbol{v}}_{D,k}$ is the vector $\boldsymbol{v}_{D,k}$ in the previous iteration and is treated as a fixed parameter in subsequent iterations. Problem (17) can be solved efficiently numerically, e.g., via the MATLAB CVX toolbox.

*2) Optimizing Digital Decoder:* Given the fixed variables $\{x,\mathbf{V}_D,\mathbf{V}_{RF},\mathbf{W}_{RF}\}$, the optimization problem for the digital decoders, $\{\mathbf{W}_{D,k}\}_k$, is formulated as

$$\min_{\{\boldsymbol{w}_{D,k}\}_{k\in\mathcal{K}}} \quad \sum_{k\in\mathcal{K}} 2\Re[\boldsymbol{a}_{w,k}^H\boldsymbol{w}_{D,k}]+\boldsymbol{w}_{D,k}^H\boldsymbol{\Phi}_{w,k}\boldsymbol{w}_{D,k}, \tag{19}$$

where $\boldsymbol{w}_{D,k}\triangleq\mathrm{vec}(\mathbf{W}_{D,k})$ and

$$\begin{aligned}\boldsymbol{a}_{w,k} &\triangleq x_k\Big\{\mathrm{vec}\big(\mathbf{W}_{RF,k}^H\bar{\mathbf{H}}_k\mathbf{V}_k\mathbf{F}_{12,k}(1:d,:)\big) \\ &\qquad + \rho\|\mathbf{V}_k\|_F\mathrm{vec}(\mathbf{W}_{RF,k}^H\mathbf{F}_{12,k}(d+1:N+d,:))\Big\}, \\ \boldsymbol{\Phi}_{w,k} &\triangleq \mathbf{F}_{22,k}^T\otimes\Big\{\mathbf{W}_{RF,k}^H\Big(\sum_{\ell\in\mathcal{K}}x_\ell^2(\bar{\mathbf{H}}_k\mathbf{V}_\ell\mathbf{V}_\ell^H\bar{\mathbf{H}}_k^H \\ &\qquad + \rho^2\|\mathbf{V}_\ell\|_F^2\mathbf{I}_N)+\sigma^2\mathbf{I}_N\Big)\mathbf{W}_{RF,k}\Big\}. \end{aligned} \tag{20}$$

This problem is an unconstrained quadratic program with a positive definite matrix $\boldsymbol{\Phi}_{w,k}$ which is convex. Hence, it is easily shown that the closed-form solution is given by [50]

$$\boldsymbol{w}_{D,k} = -\boldsymbol{\Phi}_{w,k}^{-1}\boldsymbol{a}_{w,k}, \ \forall k\in\mathcal{K}. \tag{21}$$

*3) Optimizing Analog Decoders:* Given a fixed set of variables $\{x_k,\mathbf{V}_{D,k},\mathbf{V}_{RF},\mathbf{W}_{D,k}\}_k$, the optimization problem corresponds to the problem for designing analog decoders can be written as

$$\min_{\{\mathbf{W}_{RF,k}\}_k}\sum_{k\in\mathcal{K}}2\Re\big[\mathrm{tr}(\mathbf{W}_{RF,k}^H\mathbf{T}_k)\big]+\mathrm{tr}\big(\mathbf{W}_{RF,k}^H\mathbf{R}_k\mathbf{W}_{RF,k}\mathbf{S}_k\big) \tag{22}$$

$$\text{subject to} \quad |\mathbf{W}_{RF,k}(i,j)|^2=\frac{1}{N}, \ \forall \ i,j,k. \tag{22a}$$

where

$$\begin{aligned}\mathbf{T}_k &\triangleq x_k\big(\bar{\mathbf{H}}_k\mathbf{V}_k\mathbf{F}_{12,k}(1:d,:) \\ &\qquad + \rho\|\mathbf{V}_k\|_F\mathbf{F}_{12,k}(d+1:N+d,:)\big)\mathbf{W}_{D,k}^H, \\ \mathbf{S}_k &\triangleq \mathbf{W}_{D,k}\mathbf{F}_{22,k}\mathbf{W}_{D,k}^H, \ \mathbf{R}_k\triangleq b_w\mathbf{I}_N+\bar{\mathbf{H}}_k\sum_{\ell\in\mathcal{K}}x_\ell^2\mathbf{V}_\ell\mathbf{V}_\ell^H\bar{\mathbf{H}}_k^H \\ b_w &\triangleq \sigma^2+\rho^2\sum_{\ell\in\mathcal{K}}x_\ell^2\|\mathbf{V}_\ell\|_F^2. \end{aligned} \tag{23}$$

To find the solution to the analog decoder design problem, an iterative algorithm is proposed that starts with an initial feasible $\mathbf{W}_{RF,k},\forall k\in\mathcal{K}$ and then sequentially updates each element by solving the optimization problem described next.

By fixing all entries of $\mathbf{W}_{RF,k}$ except $\mathbf{W}_{RF,k}(i,j)$ the contribution of $\mathbf{W}_{RF,k}(i,j)$ in (22) is extracted (details in Appendix C). To iteratively solve the analog decoder design problem, an algorithm is proposed where each element of

$\mathbf{W}_{RF,k}$ is updated sequentially. To extract the contribution of $\mathbf{W}_{RF,k}(i,j)$ in (22), all entries of $\mathbf{W}_{RF,k}$ are fixed except for $\mathbf{W}_{RF,k}(i,j)$, and the resulting expression is denoted by

$$f_1(\mathbf{W}_{RF,k}(i,j)) = \Re\{\mathbf{W}_{RF,k}(i,j)\eta_{i,j,k}\}, \tag{24}$$

where

$$\eta_{i,j,k} = \mathbf{T}_k(i,j)^\star + \sum_{n=1}^{N^{RF}} \sum_{\substack{m=1 \\ (n,m)\neq(j,i)}}^{N} \mathbf{R}_k(m,i)\mathbf{S}_k(j,n)\mathbf{W}^\star_{RF,k}(m,n). \tag{25}$$

Further, assuming that $\mathbf{W}_{RF,k}(i,j) = \frac{1}{\sqrt{N}} e^{j\theta_{i,j,k}}$, $\forall i,\ j,\ k$ and substituting into (24), the optimization problem in terms of $\theta_{i,j,k}$ is formulated as follows:

$$\min_{\theta_{i,j,k}} \quad f_1(\mathbf{W}_{RF,k}(i,j)). \tag{26}$$

It is easily shown that there are two solutions,

$$\begin{cases} \theta^{(1)}_{i,j,k} = -\angle\eta_{i,j,k} \\ \theta^{(2)}_{i,j,k} = \pi - \angle\eta_{i,j,k} \end{cases}, \ \forall i,\ j,\ k. \tag{27}$$

By computing the objective function (22) for both phases,

$$\theta^\star_{i,j,k} = argmin_\theta \left( f_1\left(\frac{1}{\sqrt{N}} e^{j\theta^{(1)}_{i,j,k}}\right), f_1\left(\frac{1}{\sqrt{N}} e^{j\theta^{(2)}_{i,j,k}}\right)\right). \tag{28}$$

*4) Optimizing Analog Precoder:* Given a set of fixed variables $\{x_k, \mathbf{V}_{D,k}, \mathbf{W}_{D,k}, \mathbf{W}_{RF,k}\}_k$, the optimization problem (15) for analog precoder design can be formulated as

$$\min_{\mathbf{V}_{RF}} \ 2\Re\big[\mathrm{tr}(\mathbf{L}\mathbf{V}_{RF})\big] + \mathrm{tr}\big(\mathbf{V}^H_{RF}\mathbf{A}\mathbf{V}_{RF}\mathbf{D}\big) + \sum_{k\in\mathcal{K}} b_{v,k}\|\mathbf{V}_{RF}\mathbf{V}_{D,k}\|_F + \sum_{k\in\mathcal{K}} e_w x_k^2 \|\mathbf{V}_{RF}\mathbf{V}_{D,k}\|_F^2 \tag{29}$$

$$\text{subject to } (9a), \quad |\mathbf{V}_{RF}(i,j)|^2 = \frac{1}{M}, \ \forall i,j. \tag{29a}$$

where

$$\mathbf{L} \triangleq \sum_{k\in\mathcal{K}} x_k \mathbf{V}_{D,k}\mathbf{F}_{12,k}(1:d,:)\mathbf{W}_k^H\bar{\mathbf{H}}_k,$$
$$\mathbf{A} \triangleq \sum_{k\in\mathcal{K}} \bar{\mathbf{H}}_k^H\mathbf{W}_k\mathbf{F}_{22,k}\mathbf{W}_k^H\bar{\mathbf{H}}_k, \quad \mathbf{D} \triangleq \sum_{k\in\mathcal{K}} x_k^2\mathbf{V}_{D,k}\mathbf{V}_{D,k}^H$$
$$b_{v,k} \triangleq \rho x_k \mathrm{tr}\big(\mathbf{W}_k^H\mathbf{F}_{12,k}(d+1:d+N,:)\big),$$
$$e_w \triangleq \rho^2 \sum_{k\in\mathcal{K}} \mathrm{tr}\big(\mathbf{F}_{22,k}\mathbf{W}_k^H\mathbf{W}_k\big). \tag{30}$$

A limit on total power at the BS results in a constant modulus constrained optimization problem that cannot be solved using conventional methods. However, it has been shown that for large M, the RF precoder usually satisfies $\mathbf{V}^H_{RF}\mathbf{V}_{RF} \approx \mathbf{I}$ [1]. Therefore, neglecting the power constraint for the analog precoder design is possible. Furthermore, the third and fourth terms in equation (29) which contain $\mathbf{V}^H_{RF}\mathbf{V}_{RF}$ can be treated as constant terms and disregarded in the objective function. As a result, the optimization problem associated with the analog precoder can be formulated as follows:

$$\min_{\mathbf{V}_{RF}} \ 2\Re\big[\mathrm{tr}(\mathbf{L}\mathbf{V}_{RF})\big] + \mathrm{tr}\big(\mathbf{V}^H_{RF}\mathbf{A}\mathbf{V}_{RF}\mathbf{D}\big) \tag{31}$$

$$\text{subject to} \quad |\mathbf{V}_{RF}(i,j)|^2 = \frac{1}{M}, \ \forall i,j. \tag{31a}$$

To find the solution to the analog precoder design problem (31), an iterative algorithm is proposed that starts with an initial feasible $\mathbf{V}_{RF}$ then sequentially updates each element by solving the optimization problem described next. By fixing all entries of $\mathbf{V}_{RF}$ except $\mathbf{V}_{RF}(i,j)$, the contribution of $\mathbf{V}_{RF}(i,j)$ appears in the objective function as follows:

$$f_2(\mathbf{V}_{RF}(i,j)) = \Re\{\mathbf{V}_{RF}(i,j)\gamma_{i,j}\} \tag{32}$$

where

$$\gamma_{i,j} = \mathbf{L}(j,i) + \sum_{n=1}^{M^{RF}} \sum_{\substack{m=1 \\ (n,m)\neq(j,i)}}^{M} \mathbf{A}(m,i)\mathbf{D}(j,n)\mathbf{V}^\star_{RF}(m,n). \tag{33}$$

Finally, by assuming $\mathbf{V}_{RF}(i,j) = \frac{1}{\sqrt{M}} e^{j\phi_{i,j}}$, $\forall i,\ j$ the optimal problem w.r.t., $\phi_{i,j}$ is formulated as follows:

$$\min_{\phi_{i,j}} \quad f_2(\mathbf{V}_{RF}(i,j)) \tag{34}$$

which can be easily shown to have the following two solutions:

$$\begin{cases} \phi^{(1)}_{i,j} = -\angle\gamma_{i,j} \\ \phi^{(2)}_{i,j} = \pi - \angle\gamma_{i,j,k} \end{cases}, \ \forall i,\ j,\ k. \tag{35}$$

By computing the objective function in (32) for both phases,

$$\phi^\star_{i,j} = argmin_\phi \left( f_2\left(\frac{1}{\sqrt{M}} e^{j\phi^{(1)}_{i,j}}\right), f_2\left(\frac{1}{\sqrt{M}} e^{j\phi^{(2)}_{i,j}}\right)\right). \tag{36}$$

*5) Optimizing User Selection Parameters:* Given fixed variables $\{\mathbf{V}_{D,k}, \mathbf{V}_{RF}, \mathbf{W}_{D,k}, \mathbf{W}_{RF,k}\}_k$, based on (15), in every iteration of the algorithm, the following optimization problem needs to be solved in order to choose a subset of users:

$$\min_{\{x_k\}_k} \quad \sum_{k\in\mathcal{K}} \Phi_{x,k} x_k \tag{37}$$

$$\text{subject to} \quad x_k \in \{0,1\}, \ \forall k \in \mathcal{K}, \tag{37a}$$

where

$$\Phi_{x,k} \triangleq \Re[\mathrm{tr}\big(\mathbf{F}_{12,k}(1:d,:)\mathbf{W}_k^H\bar{\mathbf{H}}_k\mathbf{V}_k\big) + \rho\|\mathbf{V}_k\|_F\mathrm{tr}\big(\mathbf{F}_{12,k}(d+1:d+N,:)\mathbf{W}_k^H\big)] + \mathrm{tr}\big(\sum_{\ell\in\mathcal{K}} \mathbf{F}_{22,\ell}\mathbf{W}_\ell^H(\bar{\mathbf{H}}_\ell\mathbf{V}_k\mathbf{V}_k^H\bar{\mathbf{H}}_\ell^H + \rho^2\|\mathbf{V}_k\|_F^2\mathbf{I}_N)\mathbf{W}_\ell\big), \tag{38}$$

and the optimal solution for this problem is given by

$$x^\star_k = \begin{cases} 1 & : \Phi_{x,k} \leq 0, \\ 0 & : \Phi_{x,k} > 0. \end{cases} \tag{39}$$

**Algorithm 1:** Proposed Joint UE Selection and Hybrid Precoder/Decoder Design.

1 Initialize $\{x_k, \mathbf{V}_{D,k}, \mathbf{V}_{RF}, \mathbf{W}_{D,k}, \mathbf{W}_{RF,k}\}_k$ to satisfy the power and group size constraints.
2 $\ell \leftarrow 1$
3 Compute $\mathbf{E}_k$ and $\mathbf{F}_k$, $\forall k$ using (10) and (14) respectively.
4 **while** $|\sum_k R_k^{(\ell)} - \sum_k R_k^{(\ell-1)}| > \epsilon_{stop}$ **do**
5  **while** $\|\boldsymbol{V}_{RF}^{new} - \boldsymbol{V}_{RF}^{old}\|_F^2 > \epsilon_{stop}$ **do**
6  Compute $\gamma_{i,j}, \forall i,j$ using (33), then calculate $\phi_{i,j}^{\star} \forall i,j$ using (36).
7  **end**
8  **while** $\sum_k \|\boldsymbol{W}_{RF,k}^{new} - \boldsymbol{W}_{RF,k}^{old}\|_F^2 > \epsilon_{stop}$ **do**
9  Compute $\eta_{i,j,k}, \forall i,j,k$ using (25), then calculate $\theta_{i,j,k}^{\star} \forall i,j,k$ using (28).
10  **end**
11  Compute $\boldsymbol{a}_{w,k}, \boldsymbol{\Phi}_{w,k} \forall k$ using (20), then calculate $\boldsymbol{w}_{D,k}$ using (21).
12  Compute $\boldsymbol{a}_{1,k}, a_{2,k}\ \forall k, \boldsymbol{\Phi}_{RF}, \boldsymbol{\Phi}_v$ using (18), then solve (17) to obtain digital precoders.
13  Compute $\Phi_{x,k}, \forall k$ using (38), then obtain optimal $\{x_k\}_k$ using (39).
14  Update $\mathbf{E}_k$ and $\mathbf{F}_k$, $\forall k$ using (10) and (14) respectively.
15  $\ell \leftarrow \ell + 1$.
16 **end**

The proposed joint user selection and hybrid precoder/decoder design is summarized in Algorithm 1. The algorithm starts with feasible initial solution $\{x_k, \mathbf{V}_{D,k}, \mathbf{V}_{RF}, \mathbf{W}_{D,k}$ $\mathbf{W}_{RF,k}\}_{k\in\mathcal{K}}$. In the next step, $\mathbf{E}_k$ and $\mathbf{F}_k$, $\forall k \in \mathcal{K}$ are computed using (10) and (14), respectively. In step 6, the $(i,j)$th element of analog precoder matrix, $\mathbf{V}_{RF}(i,j)$ is found assuming all other elements are fixed. Similarly, in step 9, the $(i,j)$th element of analog decoder, $\mathbf{W}_{RF,k}(i,j)$ is computed using (28) for every user assuming all other elements are given. After computing $\boldsymbol{a}_{w,k}, \boldsymbol{\Phi}_{w,k} \forall k \in \mathcal{K}$ using (20), digital decoders can be obtained in closed form via (21). By computing $\boldsymbol{a}_{1,k}, a_{2,k}\ \forall k, \boldsymbol{\Phi}_{RF}, \boldsymbol{\Phi}_V$, digital precoders, $\boldsymbol{v}_{D,k}$ are calculated for every user using (17). User selection parameters $\{x_k\}_{k\in\mathcal{K}}$ are then computed in closed form using (39) based on $\Phi_{x,k}, \forall k \in \mathcal{K}$ in (38). By updating $\mathbf{E}_k$ and $\mathbf{F}_k$ for all UEs based on the optimized precoder/decoder and selected users, the next iteration proceeds by returning to step 5. The proposed algorithm sequentially updates precoder/decoder matrices and user selection parameters until the stopping criterion is satisfied.

*Remark 1:* The proposed method is in a general form which includes perfectly known CSI as well as imperfect CSI cases. In addition, by setting $M^{RF} = M, N^{RF} = N, \mathbf{W}_{RF,k} = \mathbf{I}_N, \forall k \in \mathcal{K}$, and $\mathbf{V}_{RF} = \mathbf{I}_M$, the proposed method for hybrid precoding/decoding reduces to that of fully digital precoding/decoding. Further, by specifying $x_k$ as 0 or 1 for all $k \in \mathcal{K}$, the proposed joint user selection and hybrid precoder/decoder reduces to precoder/decoder design, with user selection omitted in the proposed algorithm.

## IV. Special Cases of User Selection

Assuming that all UEs are scheduled and precoder/decoder matrices have been previously designed and remain fixed, $\mathbf{V}_n, \mathbf{W}_n \neq \mathbf{0}, \forall n$, one can select a subset of users based on (39). Deriving $\mathbf{F}_{22,k}$ and $\mathbf{F}_{12,k}$ from (14) and substituting into (38), result in the closed-form

$$\Phi_{x,k} = \mathrm{tr}\Bigg( -\mathbf{V}_k^H \mathbf{H}_k^H \mathbf{W}_k \mathbf{C}_k^{-1} \mathbf{W}_k^H \mathbf{H}_k \mathbf{V}_k + \sum_{\ell} \mathbf{F}_{22,\ell} \mathbf{W}_\ell^H (\mathbf{H}_\ell \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_\ell^H) \mathbf{W}_\ell \Bigg) + \rho^2 \|\mathbf{V}_k\|^2 . \mathrm{tr}\left( -\mathbf{W}_k\ \mathbf{C}_k^{-1} \mathbf{W}_k^H + \sum_{\ell} \mathbf{F}_{22,\ell} \mathbf{W}_\ell^H \mathbf{W}_\ell \right), \tag{40}$$

where $\mathbf{F}_{22,k}$ is easily computed by algebraic operations not included due to space limitations. It is obvious that the first and third terms in (40) are $\leq 0$ while the last term is $\geq 0$. In order to select one of the UEs, say UE $k$, $\Phi_{x,k}$ has to be negative which implies that positive terms should be zero. Due to the positive definiteness of $\mathbf{F}_{22,n},\ \forall n$, the second term could be made zero by selecting $\mathbf{H}_n \mathbf{V}_k = 0,\ \forall n \neq k$. The term $\rho^2 \|\mathbf{V}_k\|^2 \mathrm{tr}(\sum_\ell \mathbf{F}_{ff,\ell} \mathbf{W}_\ell^H \mathbf{W}_\ell)$ can be considered as a penalty term for imperfect CSI cases and is proportional to channel estimation quality as well as power allocated to UE $k$.

Let us investigate $\Phi_{x,k}$ for the following special cases:

### A. $N > 1$ and Perfect CSI, $\rho = 0$

For multi-antenna UEs, the UE selection is determined by $\sum_\ell \mathrm{tr}\big(\mathbf{F}_{22,\ell} \mathbf{W}_\ell^H (\mathbf{H}_{\ \ell} \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_\ell^H) \mathbf{W}_\ell\big)$.

### B. Perfect CSI, $\rho = 0$ and $N = 1$

$$\Phi_{x,k} = \frac{-|\mathbf{H}_k \mathbf{V}_k|^2}{\sigma^2 + \sum_n |\mathbf{H}_\ell \mathbf{V}_k|^2} + \sum_{\ell \neq k} \frac{|\mathbf{H}_\ell \mathbf{V}_k|^2 |\mathbf{H}_\ell \mathbf{V}_\ell|^2}{(\sigma^2 + \sum_{n\neq\ell} |\ \mathbf{H}_\ell \mathbf{V}_n|^2)(\sigma^2 + \sum_n |\mathbf{H}_\ell \mathbf{V}_n|^2)}. \tag{41}$$

*1) $M \to \infty$ With ZF Beamformer:* In this case, $|\mathbf{H}_\ell \mathbf{V}_k| = 0, \forall \ell \neq k$ and results in $\Phi_{x,k} = \frac{-|\mathbf{H}_k \mathbf{V}_k|^2}{\sigma^2 + |\mathbf{H}_k \mathbf{V}_k|^2} \leq 0$. This implies that for large numbers of antennas with ZF precoder UE selection is not required since $\Phi_{x,k} \leq 0, \forall k$.

*2) $M \to \infty$ With Maximum Ratio Transmission (MRT):* In this case, $\mathbf{V}_k = \mathbf{H}_k^H$ that leads to

$$\Phi_{x,k} = \frac{-|\mathbf{H}_k \mathbf{H}_k^H|^2}{\sigma^2 + \sum_n |\mathbf{H}_\ell \mathbf{H}_k^H|^2} + \sum_{\ell \neq k} \frac{|\mathbf{H}_\ell \mathbf{H}_k^H|^2 |\mathbf{H}_\ell \mathbf{H}_\ell^H|^2}{(\sigma^2 + \sum_n |\mathbf{H}_\ell \mathbf{H}_n^H|^2)(\sigma^2 + \sum_n |\mathbf{H}_\ell \mathbf{H}_n^H|^2)}. \tag{42}$$

From the weak law of large numbers as $M \to \infty$, the second terms converge to 0 and again $\Phi_{x,k} \leq 0$.

TABLE II
COMPUTATIONAL COMPLEXITY OF ALGORITHM 1

| | Complexity |
|---|---|
| Steps 5-7 | $\mathcal{O}\big(J_1(2M^2M^{RF^2}+KMM^{RF}N+KM^{RF^2}d+KMNd$ $+KM^{RF}Nd+KNN^{RF}d+KN^{RF}d^2+KNd^2)\big)$ |
| Steps 8-10 | $\mathcal{O}\big(J_2(2KMM^{RF}N+KMM^{RF}d+K^2M^2N+KMN^2+KM^{RF}N^{RF}d$ $+kMd^2+KNN^{RF}d+KN^{RF^2}d+KN^{RF}d^2+2KN^2N^{RF^2})\big)$ |
| Step 11 | $\mathcal{O}\big(K^2MNd+KMM^{RF}N^{RF}+KMd^2+KMN^{RF}N+K^2Nd^2+KM^{RF}N^{RF}d$ $+KM^{RF}d^2+KN^{RF}Nd+KNN^{RF^2}+2KN^{RF^2}d^2+KN^{RF^3}d^3\big)$ |
| Step 12 | $\mathcal{O}\big(KMM^{RF}d+KMNd+KM^{RF^3}d^3+3KNN^{RF}d+KN^2d+KNd^2+2KN^{RF}d^2+Kd^3\big)$ |
| Step 13 | $\mathcal{O}\big(K^2Md^2+K^2MNd+3K^2d^2N+KN^2d+2K^2d^3+KMM^{RF}d+KNN^{RF}d\big)$ |

### C. Limited Resolution Phase Shifters

In the previous subsection, infinite-resolution phase shifters were assumed in the hybrid structure. However, accurate phase control may be physically impractical, especially in large-scale antenna arrays where the number of phase shifters is proportional to the number of antennas. Let $b$ represent the number of bits of phase shifter resolution, and let $\mathcal{C}$ represent the set of $2^b$ realizable phases, i.e., $\mathcal{C}=\big\{1, e^{j\frac{j2\pi}{2^b}}, e^{j\frac{j4\pi}{2^b}}, \ldots, e^{j\frac{j2\pi(2^b-1)}{2^b}}\big\}$. For the finite-resolution scenario, the problem of maximizing the sum rate can be formulated as follows:

$$\max_{\{x_k,\mathbf{V}_{RF},\mathbf{V}_{D,k},\mathbf{W}_{RF,k}\mathbf{W}_{D,k}\}_k} \quad \sum_{\ell\in\mathcal{K}}\alpha_\ell\bar{R}_\ell \tag{43}$$

$$\text{subject to} \quad \sum_{\ell\in\mathcal{K}} tr(\mathbf{V}_{RF}\mathbf{V}_{D,\ell}\mathbf{V}_{D,\ell}^H\mathbf{V}_{RF}^H)\le P_m, \tag{43a}$$

$$x_\ell\in\{0,1\},\ \ \forall\ell\in\mathcal{K}, \tag{43b}$$

$$\sqrt{M}|\mathbf{V}_{RF}(i,j)|\in\mathcal{C},\ \ \forall i,j \tag{43c}$$

$$\sqrt{N}|\mathbf{W}_{RF,k}(i,j)|\in\mathcal{C},\ \ \forall i,j,k. \tag{43d}$$

Since the set of feasible RF precoder/decoders are finite, an exhaustive search over all feasible choices obtains an optimal solution. However, since the search space is exponential in phase shifter resolution, full search is impractical for systems with large numbers of antennas. An alternative approach is to modify Algorithm 1 to determine finite-resolution phase shifters sequentially by obtaining $\mathbf{V}_{RF}(i,j)$ and $\mathbf{W}_{RF,k}(i,j)$ in each iteration by minimizing (34) and (26), respectively, over set $\mathcal{C}$ using one-dimensional searches. Specifically, $\mathbf{V}_{RF}(n,m)=e^{j\frac{2\pi i^\star}{2^b}},\ \forall n,m$, and $i^\star=\arg\min_{0\le k\le 2^b-1}|\angle(\mathbf{V}_{RF}(m,n))-\frac{2\pi k}{2^b}|$.

Table III compares the complexity of hybrid precoder designs of existing precoding methods under perfect CSI and does not include user selection. Therefore, in the comparison, it is worth noting that the proposed algorithm has added complexity due to both its imperfect CSI assumption and user selection stage. Despite this extra complexity, the proposed method for joint precoder design and user selection has a lower complexity compared to previous methods proposed for precoder design only.

TABLE III
COMPUTATIONAL COMPLEXITY COMPARISON FOR GIVEN SET OF USERS

| Method | Complexity |
|---|---|
| WMMSE [44] | $\mathcal{O}\big(L_1L_2M^2M^{RF^2}\big)$ |
| BD-UCD [10] | $\mathcal{O}\big(MM^{RF^2}+KNN^{RF^2}\big)$ |
| BD-SS [11] | $\mathcal{O}\big(L_3M^3\big)$ |
| ITA [12] | $\mathcal{O}\big(L_4M^6\big)$ |
| DUD [13] | $\mathcal{O}\big(L_6K(L_5M^3+M^3)\big)$ |
| HBF-MISO [1] | $\mathcal{O}\big(L_7K^2M^2M^{RF^2}\big)$ |
| GPI-MISO [23] | $\mathcal{O}\big(L_8KM^3\big)$ |

### D. Computational Complexity

The time complexity is determined by thorough analysis of the number of operations involved in each algorithm step: user selection, digital precoder design, analog precoder design and solution of convex sub-problems. The overall complexity is calculated by taking into account the iterative nature of the algorithm and the number of iterations required to achieve convergence. Table II summarizes the computational complexity of the proposed method for each sub-problem, where $J_1$ denotes the number of inner iterations to compute $\mathbf{V}_{RF}$, while $J_2$ denotes the number of inner loop iterations to compute $\{\mathbf{W}_{RF,k}\}_k$. Steps 12 and 13 represent the complexity of the digital precoder optimization which involves quadratically constrained quadratic program (QCQP) optimization and user selection, respectively. In massive MIMO scenarios where the $M$ BS antennas greatly exceed the $N$ UE antennas, it suffices to consider terms containing $M$ and $M^{RF}$ to evaluate system complexity. Looking at Table II, the computation of the analog precoder matrix $\mathbf{V}_{RF}$ (Steps 5-7), which has dimension $M\times M^{RF}$, is the most computationally intensive task, requiring $\mathcal{O}(M^2(M^{RF})^2)$ operations per iteration, and dominates overall algorithm complexity.

### E. Convergence

To investigate the convergence of the proposed method, define the objective function in (9) as $f(\mathbf{V}_{RF}^{(\kappa)},\mathbf{V}_D^{(\kappa)},\mathbf{W}_{RF}^{(\kappa)},\mathbf{W}_D^{(\kappa)},\boldsymbol{x}^{(\kappa)})\triangleq\sum_{i\in\mathcal{K}}\log|\mathbf{U}^T\mathbf{E}_i^{(\kappa)^{-1}}\mathbf{U}|$ and $g(\mathbf{V}_{RF}^{(\kappa)},\mathbf{V}_D^{(\kappa)},\mathbf{W}_{RF}^{(\kappa)},\mathbf{W}_D^{(\kappa)},\boldsymbol{x}^{(\kappa)})\triangleq\sum_{i\in\mathcal{K}}\log|\mathbf{U}^T\bar{\mathbf{E}}_i^{(\kappa)^{-1}}\mathbf{U}|-\text{tr}[\mathbf{F}_i(\mathbf{E}_i-\bar{\mathbf{E}}_i^{(\kappa)})]$ as a lower bound for

$f(\cdot)$ where, superscript $(\kappa)$ denotes $\kappa$th iteration. The objective function is shown to be non-decreasing over each iteration.

*Lemma 3:* The proposed Algorithm 1 is non-decreasing in the objective function.

*Proof:* From iteration $\kappa$ to $\kappa+1$,

$$f(\mathbf{V}_{RF}^{(\kappa)}, \mathbf{V}_{D}^{(\kappa)}, \mathbf{W}_{RF}^{(\kappa)}, \mathbf{W}_{D}^{(\kappa)}, \boldsymbol{x}^{(\kappa)}) =_a \quad (44)$$
$$g(\mathbf{V}_{RF}^{(\kappa)}, \mathbf{V}_{D}^{(\kappa)}, \mathbf{W}_{RF}^{(\kappa)}, \mathbf{W}_{D}^{(\kappa)}, \boldsymbol{x}^{(\kappa)}) \leq_b \quad (44a)$$
$$g(\mathbf{V}_{RF}^{(\kappa)}, \mathbf{V}_{D}^{(\kappa)}, \mathbf{W}_{RF}^{(\kappa)}, \mathbf{W}_{D}^{(\kappa)}, \boldsymbol{x}^{(\kappa+1)}) \leq_c \quad (44b)$$
$$g(\mathbf{V}_{RF}^{(\kappa)}, \mathbf{V}_{D}^{(\kappa)}, \mathbf{W}_{RF}^{(\kappa)}, \mathbf{W}_{D}^{(\kappa+1)}, \boldsymbol{x}^{(\kappa+1)}) \leq_d \quad (44c)$$
$$g(\mathbf{V}_{RF}^{(\kappa)}, \mathbf{V}_{D}^{(\kappa)}, \mathbf{W}_{RF}^{(\kappa+1)}, \mathbf{W}_{D}^{(\kappa+1)}, \boldsymbol{x}^{(\kappa+1)}) \leq_e \quad (44d)$$
$$g(\mathbf{V}_{RF}^{(\kappa)}, \mathbf{V}_{D}^{(\kappa+1)}, \mathbf{W}_{RF}^{(\kappa+1)}, \mathbf{W}_{D}^{(\kappa+1)}, \boldsymbol{x}^{(\kappa+1)}) \leq_f \quad (44e)$$
$$g(\mathbf{V}_{RF}^{(\kappa+1)}, \mathbf{V}_{D}^{(\kappa+1)}, \mathbf{W}_{RF}^{(\kappa+1)}, \mathbf{W}_{D}^{(\kappa+1)}, \boldsymbol{x}^{(\kappa+1)}) =_a \quad (44f)$$
$$f(\mathbf{V}_{RF}^{(\kappa+1)}, \mathbf{V}_{D}^{(\kappa+1)}, \mathbf{W}_{RF}^{(\kappa+1)}, \mathbf{W}_{D}^{(\kappa+1)}, \boldsymbol{x}^{(\kappa+1)}). \quad (44g)$$

Equality $(a)$ follows since the objective function, $f(\cdot)$, at the previous optimal point is equal to its lower bound, $g(\cdot)$, at that point; inequality $(b)$ is implied from the update of variable $x$, i.e., $x^{(\kappa+1)}$. Similar arguments are employed for inequalities $(c)$, $(d)$, $(e)$, and $(f)$. Therefore, Algorithm 1 is non-decreasing in the objective function from iteration $\kappa$ to $\kappa+1$ and converges to a local optimum point. ■

## V. Simulation Results

Numerical simulations illustrate the performance of the proposed method for various MU-MIMO scenarios. The proposed method for joint user selection and hybrid precoder/decoder design is referred to as hybrid beamforming (HBF). Joint user selection with the fully digital precoder/decoder is termed fully digital beamforming (FDBF). The numbers of transmit antennas, RF chains at the BS, receive antennas, and RF chains at UEs are represented by $M$, $M^{RF}$, $N$, and $N^{RF}$, respectively. The number of UEs is denoted by $K$, and each user receives $d$ data streams of symbols. The number of symbol streams transmitted to every UE is assumed to be equal to the number of user RF chains. In the proposed method $\epsilon_{stop} = 10^{-3}$ is used as the stopping criterion for convergence.

First, the convergence of the proposed method for perfect CSI, i.e., $\rho = 0$ and imperfect CSI are illustrated in Fig. 2. As shown, the proposed FDBF and HBF have monotonically increasing sum rates. The proposed HBF has a digital/analog decoder, digital/analog decoder, and user selection variables, requiring a larger number of parameters than FDBF. As a result, convergence of the proposed HBF is usually slower than that of the proposed FDBF. Unlike the hybrid case, the existence of a dedicated RF chain for each antenna in the fully digital implementation results in an appreciable gap between their sum rate curves at the converged points. The cases of perfect CSI, i.e., $\rho = 0$ and imperfect CSI with $\rho = 0.2$ and $\rho = 0.4$ are also shown for both proposed HBF and FDBF, where there is noticeable performance degradation for decreased channel estimation quality, i.e., for $\rho > 0$. Due to the problem's (i) non-convexity, (ii) coupling between optimization parameters in the objective function and constraints, and (iii) multi-variable nature, there is no guarantee for a global solution.

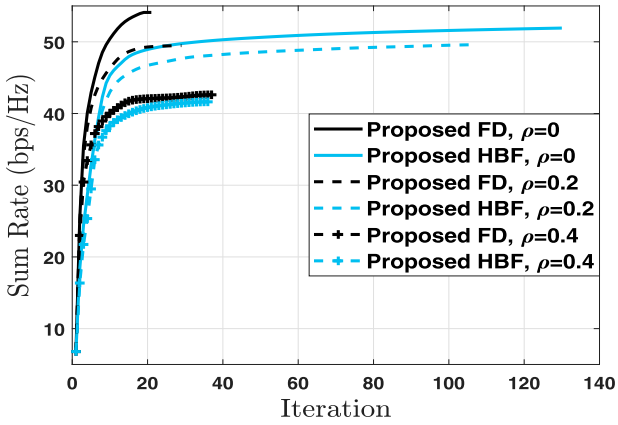


Fig. 2. Convergence of the proposed algorithm is investigated for $M = 32$, $N = 4$, $M^{RF} = 16$, $N^{RF} = 2$, $K = 40$, $P_m = 0dB$.

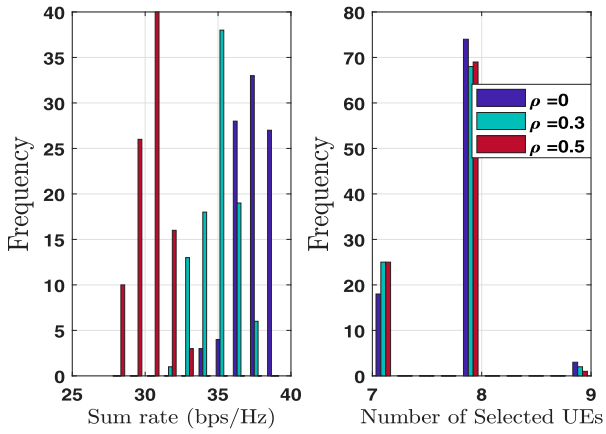


Fig. 3. Histogram for various random initialization points for $M = 64$, $N = 4$, $M^{RF} = 16$, $N^{RF} = 2$, $K = 32$, $K_{max} = 8$, $P_m = 0dB$.

Fig. 3 shows that the solution obtained by the proposed algorithm is independent of the order of optimization of the parameters. The sum rate frequency of occurrence is histogrammed for different initialization points for 100 random and feasible initial variables. To investigate the effect of channel estimation error, simulations are performed for $\rho = 0, 0.3, 0.5$. As expected, decreasing channel estimation quality results in sum rate loss. The dominant peak for each value of $\rho$ indicates low sensitivity to initialization. To investigate the effects of user selection variables on performance, the number of selected UEs is shown on the right-hand side of Fig. 3. As shown, in most scenarios, the number of selected UEs is equal to the maximum allowed, $K_{max} = \frac{M^{RF}}{N^{RF}}$, in the system model to within a small degree of rounding error.

Fig. 4 compares the proposed method to the well-known hybrid beamforming [52] (HyBF), hybrid block diagonalization [9] (HBD2), limited phase hybrid block diagonalization method (HBD1) [51] with perfect channel estimation, that is, $\rho = 0$.

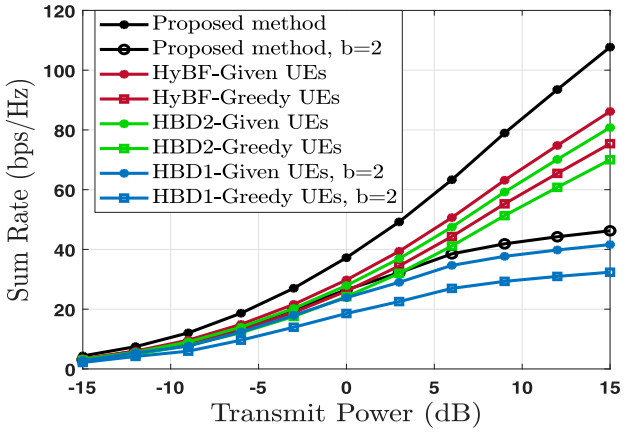


Fig. 4. Comparison of proposed method, HBD1 [51], HBD2 [9], and HBF [52] for $M = 64$, $N = 4$, $M^{RF} = 16$,$N^{RF} = 2$, $K = 20$.

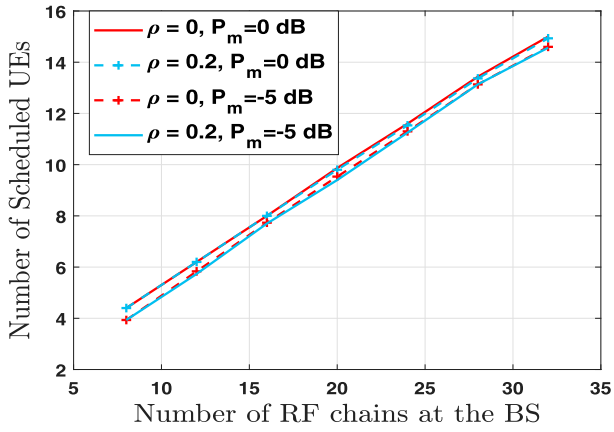


Fig. 5. Effect of number of RF chains at the BS on number of scheduled users for $M = 64$, $N = 4$, $N^{RF} = 2$, $K = 20$.

To ensure fair comparisons of precoder performance, simulations are performed for the same set of UEs selected by the proposed method (labelled "Given"). The above methods are also evaluated using a greedy user selection method (labelled "Greedy") that incrementally adds UEs to a group by choosing the user with the largest sum-rate increase. Fig. 4 shows that as the maximum BS transmit power increases from $-15$ dB to 15 dB, after a slow initial increase, a nearly linear increase in sum rate occurs at higher transmit powers. The figure also illustrates that the proposed method substantially outperforms HyBF and HBD2 with greedy user selection. In addition, using HyBF, HBD1, and HBD2 in combination with the proposed set of selected UEs outperforms greedy user selection. Finally, even with limited phase resolution, the proposed method outperforms HBD1 with 2-bit phase resolution for both the set of predetermined UEs and greedily selected UEs. It is important to note that sum rates of the methods may not increase monotonically with BS transmit power for the case of limited-resolution phase shifters since phase-shifter quantization is not incorporated into the CSI uncertainty model parameterized by $\rho$.

The results in Fig. 4 indicate that the proposed precoding method in isolation outperforms the precoding used for the other methods. It must be pointed out that the proposed joint user selection/precoding performance would exceed that shown in the comparison in Fig. 4. However, the intention of Fig. 4 is to systematically compare precoding performance in isolation. Reiterating, the other methods did not formulate joint precoding and user selection optimization and instead employed greedy user selection based on channel metrics not connected to performance.

To investigate the impact of limited RF chains on the number of served UEs, the number of selected UEs is plotted as a function of $M^{RF}$ in Fig. 5 for various transmit power and channel quality scenarios. The results show an almost linear increase in the number of selected UEs with $M^{RF}$. Moreover, we observe that increasing the transmit power at the BS leads to an average increase in the number of selected UEs.

To compare the proposed optimization with the optimal exhaustive search method, the following scenario is used with

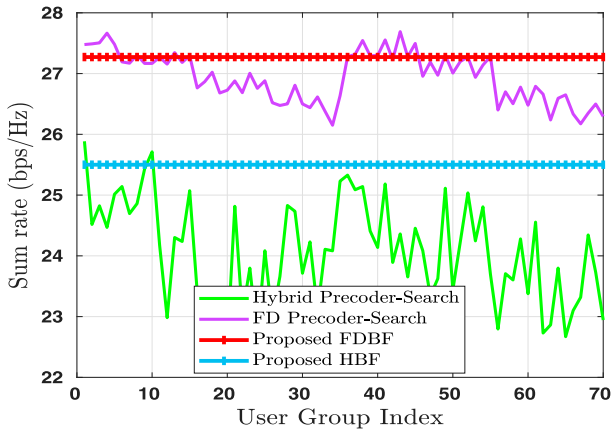


Fig. 6. Comparison between hybrid and fully digital precoders for $M = 32$, $N = 4$, $M^{RF} = 16$, $N^{RF} = 2$, $K = 4$, $\rho^2 = 0$ for each of the 70 possible 4-user groups chosen out of the 8-user pool.

$K = 8, M = 32, N = 4, M^{RF} = 16, N^{RF} = 2$ with $SNR = 10dB$. First, assuming all users are scheduled, a set of hybrid analog-digital precoders is designed. Then, by selecting 4 users out of an 8-user pool, sum rates of fully digital and hybrid analog-digital systems are compared to those of each of the $\binom{8}{4} = 70$ exhaustively searched user selections. The sum rate performance gap, as illustrated in Fig. 6 is observed to be small: the horizontal lines indicating joint user selection/precoding either exceed or are close to that of the fixed user selections. For hybrid precoding, (blue horizontal line) the proposed joint user selection/precoding outperforms 68 out 70 possible 4-user group selections. For fully digital precoding (red horizontal line), the proposed joint user selection/precoding outperforms 63 out 70 possible user selections, and for the remaining 7 exhaustively searched user selections, the joint user selection sum rates are within 2-3% of the global optimum. Overall, the proposed approach has the advantages of (i) avoiding expensive exponential UE group selection search, (ii) avoiding perfect CSI requirement, and (iii) jointly optimizing user selection and hybrid precoder/decoders.

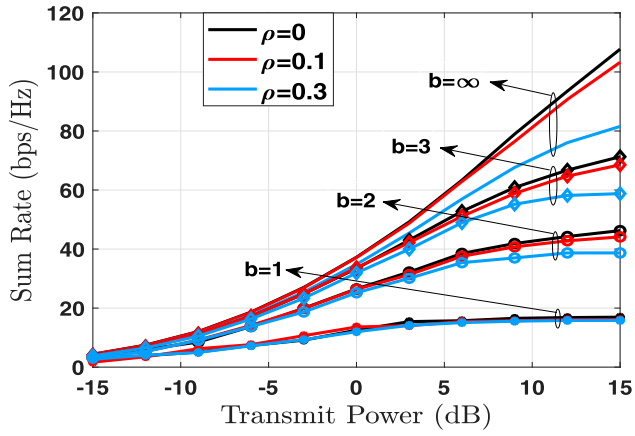


Fig. 7. Sum rate as a function of transmit power is shown for $M = 64$, $M^{RF} = 16$, $N = 4$, $N^{RF} = 2$, $K = 16$.

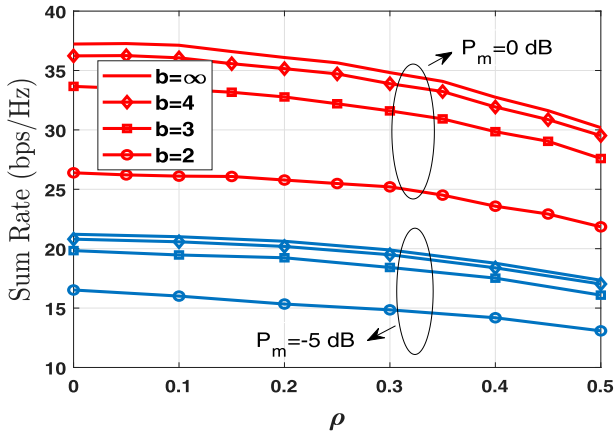


Fig. 9. Sum rate as a function of $\rho$ is shown for $M = 64$, $M^{RF} = 16$, $N = 4$, $N^{RF} = 2$, $K = 16$.

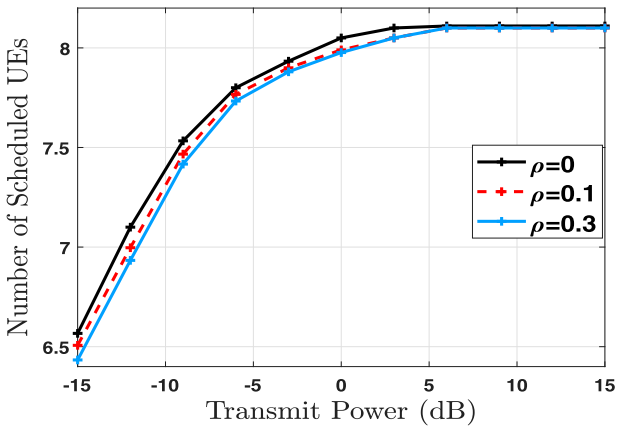


Fig. 8. Average number of selected UEs is shown for $M = 64$, $M^{RF} = 16$, $N = 4$, $N^{RF} = 2$, $K = 16$.

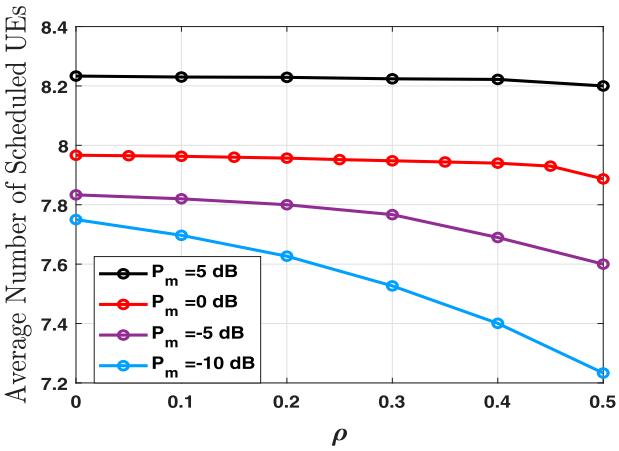


Fig. 10. Average number of selected UEs as a function of $\rho$ is shown for $M = 64$, $M^{RF} = 16$, $N = 4$, $N^{RF} = 2$, $K = 16$.

Fig. 7 investigates the impact of channel estimation error on the proposed method under different transmit power budgets and phase shifter resolutions. It is evident that if the phase shifter has unlimited resolution ($b = \infty$), the sum rate is a linear function of transmit power for high values of $P_m$. However, under limited phase shifter resolution, the sum rate degrades and saturates at 2-bit resolution. Thus, sum rate is more sensitive to phase resolution at higher transmit power. The effect of channel estimation error on sum rate also becomes more pronounced at higher transmit powers, leading to a wider gap between the perfect and imperfect CSI scenarios. As described by Eq. (8), increasing UE transmit power, represented by $|\mathbf{V}_\ell|_F^2$, amplifies the impact of channel estimation error, which adds to the receiver's noise power term $|\sum_{\ell \neq k}^{K} x_\ell \mathbf{V}_\ell|_F^2 \rho^2$.

Fig. 8 illustrates the impact of channel estimation error on the number of scheduled UEs, which is bounded by $\frac{M^{RF}}{N^{RF}}$. The number of selected UEs decreases with larger values $\rho$, i.e., lower channel estimation quality. It is also observed that with increasing transmit power, the number of selected UEs initially increases and then saturates at $\frac{M^{RF}}{N^{RF}}$.

In Fig. 9, sum rate versus channel estimation quality is investigated for various phase shifter resolutions, as expressed in bits, $b$. By decreasing channel estimation quality, i.e., increasing $\rho$, performance loss becomes more significant. In addition, increasing phase resolution, $b$, leads to an increase in sum rate. For the case of $b = 4$ bits, the sum rate is nearly that of unlimited phase resolution.

Fig. 10 demonstrates how channel estimation affects the number of selected UEs for different transmit power levels. At lower transmit power, the number of UEs is more dependent on channel estimation quality. For the case of $\rho = 0.5$, by reducing the transmit power to $-10dB$, the number of selected UEs decreases to 7.2. Clearly, a countermeasure against the effect of channel estimation error is to boost transmit power.

Fig. 11 presents the sum rate for single-antenna UEs and compares it with the hybrid beamforming method in [1] (HBF-SA) in a perfect channel estimation scenario. The performance gap between the proposed method with $\rho = 0$ and HBF-SA arises due to antenna spatial correlation in zero-forcing (ZF) digital precoders used in HBF-SA. The sum rate also saturates

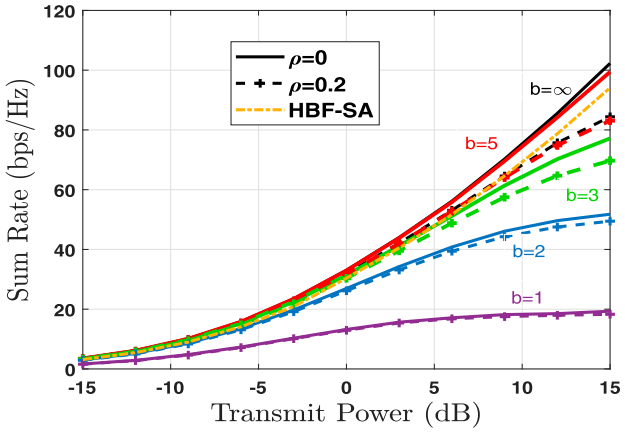


Fig. 11. Sum rate as a function transmit power is shown for $M = 64$, $M^{RF} = 16$, $N = N^{RF} = 1$, $K = 32$.

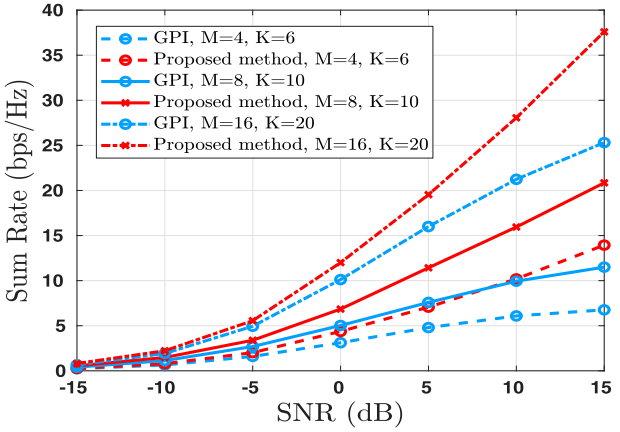


Fig. 12. Comparison of proposed with GPI method [23] in a single-antenna user and fully digital precoding scenario for $M = 4, 8, 16$, and $K = 6, 10, 20$.

for low-resolution phase shifters, while 5-bit resolution has only small degradation for single-antenna users.

Fig. 12 depicts the total sum rate as a function of *SNR* for a fully digital precoding scenario with single antenna users in comparison to the GPI method [23]. In the GPI method, user selection relies on the precoder and the power allocated to each user. It can be observed from Fig. 12, the differences in performance become more pronounced in high *SNR* scenarios, with a saturation effect in the achievable sum-rate.

## VI. Conclusion and Future Work

A framework for jointly optimizing user group selection, group size, hybrid structure precoders and decoders in the presence of channel estimation error is proposed based on bounding by convex functions. Subsets of parameters are determined in an alternating fashion. Numerical results reveal that local extrema are reached after a few iterations. As expected, the fully digital precoders/decoders outperform their hybrid versions but at the cost of more RF chains. Performance is also observed to degrade gracefully as channel estimation accuracy decreases.

Owing to the high complexity and overhead required for channel estimation combined with recent advances in machine learning techniques, possible future work would be to investigate the application of ML data-driven approaches with off-line training to reduce both run-time optimization complexity and CSI overhead.

## Appendix A

Let $\boldsymbol{\Xi}$ denote a matrix in $\mathbb{C}^{M\times N}$ of zero mean circularly symmetric complex Gaussian random variables with variance $\rho^2$ and let $\boldsymbol{A} \in \mathbb{C}^{M\times M}$ be an arbitrary deterministic matrix.

Let $\boldsymbol{A} = \begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_M\end{bmatrix} \in \mathbb{C}^{M\times M}$, and $\boldsymbol{\Xi} = \begin{bmatrix}\boldsymbol{\xi}_1^H \ ; \ \boldsymbol{\xi}_2^H \ ; \ \dots \ ; \ \boldsymbol{\xi}_N^H\end{bmatrix} \in \mathbb{C}^{N\times M}$.

Using the independence of elements of $\boldsymbol{\Xi}$, we obtain

$$E\{\boldsymbol{\Xi}\boldsymbol{A}\boldsymbol{\Xi}^H\} = E\begin{bmatrix}\sum_{i=1}^M E\{\xi_{1i}\boldsymbol{\xi}_1^H\}\boldsymbol{a}_i & & 0 \\ & \ddots & \\ 0 & & \xi_{Ni}\boldsymbol{\xi}_N^H\}\boldsymbol{a}_i\end{bmatrix}$$
$$= \begin{bmatrix}\rho^2\sum_{i=1}^M a_{ii} & & 0 \\ & \ddots & \\ 0 & & \rho^2\sum_{i=1}^M a_{ii}\end{bmatrix} = \rho^2\mathrm{tr}(\boldsymbol{A})\mathbf{I}_N, \tag{45}$$

where $\xi_{ji}$ is the $i$th entry of the vector $\boldsymbol{\xi}_j$ and $a_{ii}$ is the $i$th diagonal element of matrix $\boldsymbol{A}$. Hence,

$$\begin{aligned}&\mathrm{E}_\Delta\Big\{x_k^2\mathbf{W}_k^H(\bar{\mathbf{H}}_k + \bar{\boldsymbol{\Delta}}_k)\mathbf{V}_k\mathbf{V}_k^H(\bar{\mathbf{H}}_k + \bar{\boldsymbol{\Delta}}_k)^H\mathbf{W}_k\Big\} \\ &= x_k^2\mathbf{W}_k^H\Big(\bar{\mathbf{H}}_k\mathbf{V}_k\mathbf{V}_k^H\bar{\mathbf{H}}_k^H + \bar{\mathbf{H}}_k\mathbf{V}_k\mathbf{V}_k^H\mathrm{E}\{\bar{\boldsymbol{\Delta}}_k^H\} \\ &\quad + \mathrm{E}\{\bar{\boldsymbol{\Delta}}_k\}\mathbf{V}_k\mathbf{V}_k^H\bar{\mathbf{H}}_k^H + \mathrm{E}\{\bar{\boldsymbol{\Delta}}_k\mathbf{V}_k\mathbf{V}_k^H\bar{\boldsymbol{\Delta}}_k^H\}\Big)\mathbf{W}_k \\ &= x_k^2\mathbf{W}_k^H\Big(\bar{\mathbf{H}}_k\mathbf{V}_k\mathbf{V}_k^H\bar{\mathbf{H}}_k^H + \rho^2\|\mathbf{V}_k\mathbf{V}_k^H\|_F^2\mathbf{I}_N\Big)\mathbf{W}_k.\end{aligned} \tag{46}$$

The second term in (7) can be obtained in a similar way.

## Appendix B

Using (10) and partitioning matrix $\mathbf{F}_k$,

$$\mathrm{tr}(\mathbf{F}_k\mathbf{E}_k) = \mathrm{tr}\left(\begin{bmatrix}\mathbf{F}_{11,k} & \mathbf{F}_{12,k} \\ \mathbf{F}_{21,k} & \mathbf{F}_{22,k}\end{bmatrix}\mathbf{E}_k\right) \tag{47}$$

$$\begin{aligned}&= \mathrm{tr}(\mathbf{F}_{11,k}) + x_k\mathrm{tr}(\mathbf{F}_{12,k}[\mathbf{W}_k^H\bar{\mathbf{H}}_k\mathbf{V}_k \quad \rho\|\mathbf{V}_k\|_F]) \\ &\quad + \mathrm{tr}([\mathbf{W}_k^H\bar{\mathbf{H}}_k\mathbf{V}_k \quad \rho\|\mathbf{V}_k\|_F]^H\mathbf{F}_{21,k}) \\ &\quad + \mathrm{tr}(\mathbf{F}_{22,k}\mathbf{W}_k^H\Big(\sigma^2\mathbf{I}_N + \sum_{\ell=1}^K x_\ell^2(\bar{\mathbf{H}}_k\mathbf{V}_\ell\mathbf{V}_\ell^H\bar{\mathbf{H}}_k^H \\ &\quad + \rho^2\|\mathbf{V}_\ell\|_F^2\mathbf{I}_N)\Big)\mathbf{W}_k) \\ &= \mathrm{tr}(\mathbf{F}_{11,k}) + 2x_k\Re\Big[\mathrm{tr}\Big(\mathbf{F}_{12,k}(1:d,:)\mathbf{W}_k^H\bar{\mathbf{H}}_k\mathbf{V}_k\Big) \\ &\quad + \rho\|\mathbf{V}_k\|_F\mathrm{tr}\Big(\mathbf{F}_{12,k}(d+1:N+d,:)\mathbf{W}_k^H\Big)\Big] \\ &\quad + \mathrm{tr}\Big(\mathbf{W}_k\mathbf{F}_{22,k}(\sigma^2\mathbf{I}_N + \rho^2\sum_{\ell=1}^K x_\ell^2\|\mathbf{V}_\ell\|_F^2)\mathbf{W}_k^H\Big) \\ &\quad + \mathrm{tr}\Big(\sum_{\ell=1}^K x_\ell^2\mathbf{F}_{22,k}\mathbf{W}_k^H\bar{\mathbf{H}}_k\mathbf{V}_\ell\mathbf{V}_\ell^H\bar{\mathbf{H}}_k^H\mathbf{W}_k\Big).\end{aligned} \tag{48}$$

## APPENDIX C

To compute the contribution of an element in (22), we note:

$$
\begin{aligned}
&2\Re\big[\mathrm{tr}(\mathbf{W}_{RF,k}^H\mathbf{T}_k)\big]+\mathrm{tr}\big(\mathbf{W}_{RF,k}^H\mathbf{R}_k\mathbf{W}_{RF,k}\mathbf{S}_k\big)\\
&\quad=2\Re\left[\sum_n\sum_m\mathbf{W}_{RF,k}^\star(m,n)\mathbf{T}_k(m,n)\right]\\
&\quad+\sum_n\sum_r\sum_m\sum_\ell\ \mathbf{W}_{RF,k}^\star(m,n)\mathbf{R}_k(m,r)\mathbf{W}_{RF,k}(r,\ell)\mathbf{S}_k(\ell,n)\\
&\quad=2\Re[\mathbf{W}_{RF,k}(i,j)\mathbf{T}_k^\star(i,j)]\\
&\quad+\sum_{\substack{(n,m)\\\neq(j,i)}}\mathbf{W}_{RF,k}^\star(m,n)\mathbf{R}_k(m,i)\mathbf{W}_{RF,k}(i,j)\mathbf{S}_k(j,n)\\
&\quad+\sum_{\substack{(r,\ell,)\\\neq(i,j)}}\mathbf{W}_{RF,k}^\star(i,j)\mathbf{R}_k(i,r)\mathbf{W}_{RF,k}(r,\ell)\mathbf{S}_k(\ell,j)+\mathbf{C},
\end{aligned}
\tag{49}
$$

where C comprises terms that are not a function of $\mathbf{W}_{RF,k}(i,j)$. Due to Hermitian symmetry of $\mathbf{R}_k$ and $\mathbf{S}_k$, (49) becomes

$$
\begin{aligned}
&2\Re\big[\mathbf{W}_{RF,k}(i,j)\mathbf{T}_k^\star(i,j)\big]\\
&\quad+2\Re\left[\sum_{\substack{(n,m)\\\neq(j,i)}}\mathbf{W}_{RF,k}^\star(m,n)\mathbf{R}_k(m,i)\mathbf{W}_{RF,k}(i,j)\mathbf{S}_k(j,n)\right]+\mathbf{C}\\
&=2\Re\Bigg[\mathbf{W}_{RF,k}(i,j)\big(\mathbf{T}_k^\star(i,j)\\
&\quad+\sum_{\substack{(n,m)\\\neq(j,i)}}\mathbf{W}_{RF,k}^\star(m,n)\mathbf{R}_k(m,i)\mathbf{S}_k(j,n)\big)\Bigg]+\mathbf{C}.
\end{aligned}
\tag{50}
$$

## REFERENCES


[1] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.

[2] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[3] L.-N. Tran, M. Juntti, M. Bengtsson, and B. Ottersten, "Weighted sum rate maximization for MIMO broadcast channels using dirty paper coding and zero-forcing methods," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2362–2373, Jun. 2013.

[4] X. Zhao, S. Lu, Q. Shi, and Z.-Q. Luo, "Rethinking WMMSE: Can its complexity scale linearly with the number of BS antennas?" *IEEE Trans. Signal Process.*, vol. 71, pp. 433–446, 2023.

[5] Z. Lin, M. Lin, T. de Cola, J.-B. Wang, W.-P. Zhu, and J. Cheng, "Supporting IoT with rate-splitting multiple access in satellite and aerial-integrated networks," *IEEE Internet Things J.*, vol. 8, no. 14, pp. 11123–11134, Jul. 2021.

[6] Z. Lin et al., "Refracting RIS-aided hybrid satellite-terrestrial relay networks: Joint beamforming design and optimization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 4, pp. 3717–3724, Aug. 2022.

[7] B. Shi, F. Liu, and R. Du, "JCNN-based DD hybrid precoding algorithm with TPSAS structure for FD-MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 13, no. 2, pp. 357–361, Feb. 2024.

[8] T. Lin, J. Cong, Y. Zhu, J. Zhang, and K. Ben Letaief, "Hybrid beamforming for millimeter wave systems using the MMSE criterion," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3693–3708, May 2019.

[9] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 201–211, Jan. 2016.

[10] M. Yuan, H. Wang, and Y. Sun, "BD-UCD-based nonlinear hybrid precoding for millimeter wave massive multiuser MIMO systems," *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 1010–1014, Mar. 2021.

[11] J.-Y. Lin, H.-J. Su, C.-C. Hong, and Y. Takano, "Low complexity hybrid precoder design for mmWave multi-user MIMO systems: A non-iterative approach," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC)*, 2019, pp. 1–7.

[12] J. P. Gonzalez-Coma, O. Fresnedo, and L. Castedo, "Rank constrained precoding for the downlink of mmWave massive MIMO hybrid systems," *IEEE Access*, vol. 9, pp. 28459–28470, 2021.

[13] J. Fang, H. Liu, C. Xing, S. Xie, S. Gong, and J. An, "Hybrid multi-antenna transceiver optimizations for IoT systems via downlink-uplink duality," *IEEE Internet Things J.*, vol. 11, no. 5, pp. 8156–8169, Mar. 2024.

[14] M. Hui, X. Zhao, T. Lin, and Y. Zhu, "Hybrid beamforming for utility maximization in multiuser broadband millimeter wave systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 16042–16057, Dec. 2023.

[15] Z. Lin, M. Lin, B. Champagne, W.-P. Zhu, and N. Al-Dhahir, "Secrecy-energy efficient hybrid beamforming for satellite-terrestrial integrated networks," *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 6345–6360, Sep. 2021.

[16] J. Zhang, G. Zheng, Y. Zhang, I. Krikidis, and K.-K. Wong, "Deep learning based predictive beamforming design," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 8122–8127, Jun. 2023.

[17] H. Hojatian, J. Nadal, J.-F. Frigon, and F. Leduc-Primeau, "Unsupervised deep learning for massive MIMO hybrid beamforming," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7086–7099, Nov. 2021.

[18] Z. Gao et al., "Data-driven deep learning based hybrid beamforming for aerial massive MIMO-OFDM systems with implicit CSI," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 2894–2913, Oct. 2022.

[19] S. Huang, H. Yin, J. Wu, and V. C. M. Leung, "User selection for multiuser MIMO downlink with zero-forcing beamforming," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3084–3097, Sep. 2013.

[20] Y. Zhang, B. Golkar, E. S. Sousa, and Q. Zhang, "Efficient user selection for downlink zero-forcing based multiuser MIMO systems," *IEEE VTC Fall*, 2011, pp. 1–5.

[21] A. Bayesteh and A. K. Khandani, "On the user selection for MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1086–1107, Mar. 2008.

[22] R. C. Elliott, S. Sigdel, W. A. Krzymien, M. Al-Shalash, and A. C. K. Soong, "Genetic and greedy user scheduling for multiuser MIMO systems with successive zero-forcing," *IEEE Global Commun. Workshops*, 2009, pp. 1–6.

[23] J. Choi, N. Lee, S.-N. Hong, and G. Caire, "Joint user selection, power allocation, and precoding design with imperfect CSIT for multi-cell MU-MIMO downlink systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 162–176, Jan. 2020.

[24] S. He, Z. An, J. Zhu, M. Zhang, Y. Huang, and Y. Zhang, "Cross-layer optimization: Joint user scheduling and beamforming design with QoS support in joint transmission networks," *IEEE Trans. Commun.*, vol. 71, no. 2, pp. 792–807, Feb. 2023.

[25] B. Song, Y.-H. Lin, and R. L. Cruz, "Weighted max-min fair beamforming, power control, and scheduling for a MISO downlink," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 464–469, Feb. 2008.

[26] G. Dimic and n N. Sidiropoulos, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.

[27] M. Wang, T. Samarasinghe, and J. S. Evans, "Optimizing user selection schemes in vector broadcast channels," *IEEE Trans. Commun.*, vol. 63, no. 2, pp. 565–577, Feb. 2015.

[28] X. Yi and E. K. S. Au, "User scheduling for heterogeneous multiuser MIMO systems: A subspace viewpoint," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 4004–4013, Oct. 2011.

[29] S. Nam, J. Kim, and Y. Han, "A user selection algorithm using angle between subspaces for downlink MU-MIMO systems," *IEEE Trans. Commun.*, vol. 62, no. 2, pp. 616–624, Feb. 2014.

[30] K. Ko and J. Lee, "Multiuser MIMO user selection based on chordal distance," *IEEE Trans. Commun.*, vol. 60, no. 3, pp. 649–654, Mar. 2012.

[31] L.-N. Tran, M. Bengtsson, and B. Ottersten, "Iterative precoder design and user scheduling for block-diagonalized systems," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3726–3739, Jul. 2012.

[32] M. Min, Y.-S. Jeon, and G.-H. Im, "On achievable rate of user selection for MIMO broadcast channels with limited feedback," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 122–135, Jan. 2017.

[33] T. Y. Elganimi and A. A. Aturki, "Joint user selection and GMD-based hybrid beamforming for generalized spatial modulation aided

millimeter-wave massive MIMO systems," in *Proc. IEEE 3rd Int. Conf. Inf. Commun. Signal Process. (ICICSP),* 2020, pp. 364–369.
[34] N. Nonaka, K. Muraoka, S. Suyama, and Y. Okumura, "Two-step user selection algorithm for multi-user massive MIMO with hybrid beamforming," in *Proc. 22nd Int. Symp. Wireless Pers. Multimedia Commun. (WPMC),* 2019, pp. 1–5.
[35] H. Miyazaki, S. Suyama, T. Okuyama, J. Mashino, and Y. Okumura, "User selection and rank adaptation for multi-user massive MIMO with hybrid beamforming," in *Proc. IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, 2017, pp. 1–5.
[36] T. X. Vu et al., "Machine learning-enabled joint antenna selection and precoding design: From offline complexity to online performance," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3710–3722, Jun. 2021.
[37] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.
[38] K. An et al., "Exploiting multi-layer refracting RIS-assisted receiver for HAP-SWIPT networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 12638–12657, Oct. 2024.
[39] S. Huang, H. Yin, H. Li, and V. C. M. Leung, "Decremental user selection for large-scale multi-user MIMO downlink with zero-forcing beamforming," *IEEE Wireless Commun. Lett.*, vol. 1, no. 5, pp. 480–483, Oct. 2012.
[40] M. Bavand and S. D. Blostein, "User selection and multiuser widely linear precoding for one-dimensional signalling," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11642–11653, Dec. 2018.
[41] X. Zhang and J. Lee, "Low complexity MIMO scheduling with channel decomposition using capacity upper bound," *IEEE Trans. Commun.*, vol. 56, no. 6, pp. 871–876, Jun. 2008.
[42] Z. Shen, R. Chen, J. Andrews, R. Heath, and B. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3658–3663, Sep. 2006.
[43] Y. Zhao, W. Xu, J. Xu, S. Jin, K. Wang, and M.-S. Alouini, "Analog versus hybrid precoding for multiuser massive MIMO with quantized CSI feedback," *IEEE Commun. Lett.*, vol. 24, no. 10, pp. 2319–2323, Oct. 2020.
[44] S. Gong, C. Xing, V. K. N. Lau, S. Chen, and L. Hanzo, "Majorization-minimization aided hybrid transceivers for MIMO interference channels," *IEEE Trans. Signal Process.*, vol. 68, pp. 4903–4918, 2020.
[45] W. A. Al-Hussaibi and F. H. Ali, "A closed-form approximation of correlated multiuser MIMO ergodic capacity with antenna selection and imperfect channel estimation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 5515–5519, Jun. 2018.
[46] S. S. Christensen et al., "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.
[47] B. Nosrat-Makouei, J. G. Andrews, and R. W. Heath, "MIMO interference alignment over correlated channels with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2783–2794, Jun. 2011.
[48] Q. Zhang, S. Jin, M. McKay, D. Morales-Jimenez, and H. Zhu, "Power allocation schemes for multicell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 5941–5955, Nov. 2015.
[49] C. D. Meyer, "*Matrix Analysis and Applied Linear Algebra*," *Siam*, vol. 712000.
[50] S. P. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, MA, USA: Cambridge Univ. Press, 2004,
[51] X. Li, Y. Huang, W. Heng, and J. Wu, "Hybrid block diagonalization for multiuser MIMO system with domestic switch network," in *Proc. IEEE WCNC*, 2021, pp. 1–6.
[52] X. Wu, D. Liu, and F. Yin, "Hybrid beamforming for multi-user massive MIMO systems," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 3879–3891, Sep. 2018.

**Hossein Vaezy** was born in Yazd, Iran, in 1988. He received the Ph.D. degree in electrical engineering from Isfahan University of Technology, Isfahan, Iran, in 2019. From 2018 to 2019, he was a Research Associate with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada. From 2021 to 2023, he was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, Canada. Since 2023, he has been working as a Senior R&D Engineer with Synopsys Inc., Canada. His research interests include communication theory, wireless communications, and machine learning, with an emphasis on mmWave and MIMO technologies for 5G and 6G networks.

**Steven D. Blostein** (Senior Member, IEEE) received the B.S. degree in electrical engineering from Cornell University, Ithaca, NY, USA, in 1983, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois, Urbana-Champaign, IL, USA, in 1985 and 1988, respectively. Since 1988, he has been on the Faculty with the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, Canada, and currently holds the position of Professor. From 2004 to 2009, he was the Department Head. He has also been a Consultant to industry and government in the areas of image compression, target tracking, radar imaging, and wireless communications. His research interests include wireless communications systems, including detection and estimation, signal processing, energy efficiency, MIMO, dynamic access, and dense deployments. He was the Chair of IEEE Kingston Section, the Chair of the Biennial Symposium on Communications, the Publications Chair IEEE ICASSP, an Associate Editor of IEEE TRANSACTIONS ON IMAGE PROCESSING and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and served on numerous Technical Program Committees for the IEEE Communications Society conferences. He is a registered Professional Engineer in Ontario.