

Practical Hybrid Beamforming for Millimeter Wave Massive MIMO Full Duplex With Limited Dynamic Range

CHANDAN KUMAR SHEEMAR¹ (Student Member, IEEE),
CHRISTO KURISUMMOOTIL THOMAS² (Member, IEEE), AND DIRK SLOCK¹ (Fellow, IEEE)

¹Communication Systems Department, EURECOM, 06410 Sophia Antipolis, France

²Wireless Research and Development, Qualcomm Finland RFFE Oy, 02150 Espoo, Finland

CORRESPONDING AUTHOR: C. K. SHEEMAR (e-mail: sheemar@eurecom.fr)

ABSTRACT Full Duplex (FD) radio has emerged as a promising solution to increase the data rates by up to a factor of two via simultaneous transmission and reception in the same frequency band. This paper studies a novel hybrid beamforming (HYBF) design to maximize the weighted sum-rate (WSR) in a single-cell millimeter wave (mmWave) massive multiple-input-multiple-output (mMIMO) FD system. Motivated by practical considerations, we assume that the multi-antenna users and hybrid FD base station (BS) suffer from the limited dynamic range (LDR) noise due to non-ideal hardware and an impairment aware HYBF approach is adopted by integrating the traditional LDR noise model in the mmWave band. In contrast to the conventional HYBF schemes, our design also considers the joint sum-power and the practical per-antenna power constraints. A novel interference, self-interference (SI) and LDR noise aware optimal power allocation scheme for the uplink (UL) users and FD BS is also presented to satisfy the joint constraints. The maximum achievable gain of a multi-user (MU) mmWave FD system over a fully digital half duplex (HD) system with different LDR noise levels and numbers of the radio-frequency (RF) chains is investigated. Simulation results show that our design outperforms the HD system with only a few RF chains at any LDR noise level. The advantage of having amplitude control at the analog stage is also examined, and additional gain for the mmWave FD system becomes evident when the number of RF chains at the hybrid FD BS is small.

INDEX TERMS Millimeter wave, full duplex, hybrid beamforming, limited dynamic range, minorization-maximization.

I. INTRODUCTION

THE REVOLUTION in wireless communications has led to an exponential increase in the data rate requirements and number of users. The millimeter wave (mmWave) frequency band 30 – 300 GHz can accommodate the ever-increasing data demands and results to be a vital resource for future wireless communications [1]. It offers much wider bandwidths than the traditional cellular networks, and the available spectrum at such higher frequencies is 200 times greater [2]. Full Duplex (FD) communication in mmWave has the potential to further double the spectral efficiency by offering simultaneous transmission and reception in the same frequency band. Moreover, it can be beneficial for efficient

management of the vast mmWave spectrum, reducing end-to-end delays/latency, enabling advanced joint communication and sensing, and solving the hidden node problem [3]–[6].

Self-interference (SI), which can be 90 – 110 dB higher than the received signal [7], [8], is a key challenge to achieve an ideal FD operation. Given the tremendous amount of SI, signal reception is impossible without a proper SI cancellation scheme. Beamforming is a powerful tool for FD to mitigate the SI while serving multiple users and can lead to a significant performance gain compared to a half duplex (HD) system [9]–[18]. However, its gain in practical communication systems is restricted by the limited dynamic range (LDR) of the radio-frequency (RF) chains [12]. The signal

may suffer from LDR noise due to the distortions introduced by non-ideal power amplifiers (PAs), analog-to-digital-converters (ADCs), digital-to-analog-converters, mixers and low noise PAs. These impairments dictate the residual SI power which cannot be canceled and therefore establish the achievable gain for FD [12]. This adverse effect urges the requirement of impairment aware beamforming designs and investigating their performance in terms of the LDR noise levels such that correct conclusions on the achievable gain of FD could be drawn. Such an approach for the fully digital FD systems can be adopted with the well-established LDR noise model available in [10]–[18]. In general, impairment aware beamforming is more robust to distortions and can significantly outperform the naive schemes [19], [20], see, e.g., [20, Fig. 2].

The deployment of MU mmWave FD systems requires the FD base stations (BSs) to be equipped with a massive number of antennas to overcome the propagation challenges. Owing to the hardware cost, they will have to rely on a hybrid architecture consisting of only a few RF chains. Therefore, efficient hybrid beamforming (HYBF) schemes are required for such transceivers to manage the SI and interference jointly by performing large-dimensional phasor processing in the analog domain and lower-dimensional digital processing.

A. STATE-OF-THE-ART AND MOTIVATION

In [21]–[27], novel HYBF designs for a point-to-point mmWave massive MIMO (mMIMO) FD system are studied. HYBF schemes of mMIMO FD relays and integrated access and backhaul are presented in [28]–[30] and [31], respectively. HYBF designs with single antenna uplink (UL) and downlink (DL) users for a single-cell and a multi-cell mmWave FD system are proposed in [32] and [33], respectively. In [34], HYBF for mmWave mMIMO FD with only one UL and one DL multi-antenna user, under the receive LDR is proposed. In [35], HYBF for two fully connected mMIMO FD nodes that approaches SI-free sum-spectral efficiency is proposed. In [36], HYBF for a mmWave FD system equipped with analog SI cancellation stage is presented. In [37], HYBF to generalize the point-to-point mmWave mMIMO FD communication to the case of a K -pair links is presented. Frequency-selective HYBF for a wide-band mmWave FD system is studied in [38].

The literature on multi-antenna MU mmWave FD systems is limited only to the case of one UL and one DL user [34]–[36], [38]. In [34], the receive side LDR of FD BS is also considered, which is dominated by the quantization noise of the ADCs. However, LDR noise from the transmit side is ignored, which also affects the performance of FD systems significantly [39]. The effect of cross-interference generated from the UL user towards the DL user is also not considered in [34], which can have a major impact on the achievable performance. Cross-interference generated from the neighboring cells is well investigated in the dynamic time-division-duplexing networks [40]–[44], and it is more

harmful to the MU FD systems as it occurs in the same cell. For example, consider the case of a small cell, in which BSs and users are expected to operate with a similar amount of transmit power [44]. Suppose that one FD BS simultaneously serves one UL and one DL user and that both the users are located close to each other and sufficiently far from the BS. In such a case, cross-interference can become as severe as the SI and can completely drown the useful signal intended for the DL user if not considered in the beamforming design. In a MU scenario with multiple UL users located near the DL users, each DL user suffers from cross-interference, which is summed over all the UL users' transmit power, with each UL user transmitting with a similar amount of power as the BS. In such a case, cross-interference can become even more severe than the SI if not considered in the design.

B. MAIN CONTRIBUTIONS

We present a novel HYBF design to maximize the weighted sum-rate (WSR) in a single-cell mmWave mMIMO FD system, i.e., for multiple multi-antenna UL and DL users. The users are assumed to have a limited number of antennas and digital processing capability. The FD BS is assumed to have a massive number of antennas and hybrid processing capability. Our design is based on alternating optimization and relies on the mathematical tools offered by minorization-maximization [45]. The users and BS are assumed to be suffering from the LDR noise due to non-ideal hardware, modeled with the traditional LDR model [12] and by extending it to the case of a hybrid transceiver, respectively. Our work represents the first-ever impairment aware HYBF approach for mmWave FD and its analysis as a function of the LDR noise levels. Extension of the LDR noise model presented herein is applicable to any mmWave FD scenario.

In contrast to the conventional HYBF designs for mmWave FD, in this work, the beamformers are designed under the joint sum-power and the practical per-antenna power constraints. The sum-power constraint at each terminal is imposed by the regulations, which limits its total transmit power. In practice, each transmit antenna is equipped with its PA¹ [47] and the per-antenna power constraints arise due to power consumption limits imposed on the physical PAs [17], [47]–[50]. We also present a novel SI, interference, cross-interference and LDR noise aware optimal power allocation scheme to meet the joint constraints.

Compared to the digital part, optimization of the analog stage is more challenging as it must obey the unit-modulus constraint. Recently, new transceivers have started to emerge, which with the aid of amplitude modulators (AMs), also allow amplitude control for the analog stage [34], [51], [52]. Such transceivers alleviate the unit-modulus constraint but require additional hardware. Hence, we study both the unit-modulus and AMs cases and investigate when the amplitude control for mmWave FD could be advantageous. In practice,

1. The mMIMO systems are also expected to be deployed with one PA per-antenna to enable the deployment of very low-cost PAs [46].

as the analog beamformer and analog combiner can assume only finite values, a quantization constraint is also imposed on them during the optimization process. In our problem formulation, the WSR does not depend on the digital combiners, which are omitted in the design. They must be chosen as the minimum-mean-squared-error (MMSE) combiners after the convergence of the proposed algorithm. The number of digital combiners would be the sum of the number of UL and DL users. By omitting them, the HYBF design simplifies, and the per-iteration computational complexity reduces significantly.

Simulation results show that our design outperforms a fully digital HD system and can deal with the SI, MU interference and cross-interference with only a few RF chains. Results are reported with different LDR noise levels, and significant performance gain is observed at any level.

In summary, the contributions of our work are:

- Extension of the LDR noise model for the mmWave band.
- Introduction of the WSR maximization problem formulation for HYBF in a single-cell mmWave mMIMO FD system affected by the LDR noise.
- A novel SI, MU interference, cross-interference, LDR noise and practical per-antenna power constraints aware HYBF design.
- Investigation of the achievable WSR in a MU mmWave FD system as a function of the LDR noise.
- Optimal MU interference, SI, LDR noise and per-antenna power constraints aware power allocation scheme for the hybrid FD BS and UL users.

Paper Organization: The rest of the paper is organized as follows. Section II presents the system model, problem formulation and extends the LDR noise model. Sections III and IV present the minorization-maximization method and a novel HYBF design, respectively. Finally, Sections V and VI present the simulation results and conclusions, respectively.

Mathematical Notations: Boldface lower and upper case characters denote vectors and matrices, respectively. $\mathbb{E}\{\cdot\}$, $\text{Tr}\{\cdot\}$, $(\cdot)^H$, $(\cdot)^T$, \otimes , \mathbf{I} , \mathbf{D}_d and i denote expectation, trace, conjugate transpose, transpose, Kronecker product, identity matrix, d dominant vectors selection matrix and the imaginary unit, respectively. $\text{vec}(\mathbf{X})$ stacks the columns of \mathbf{X} into a vector \mathbf{x} and $\text{unvec}(\mathbf{x})$ reshapes \mathbf{x} into \mathbf{X} . $\angle \mathbf{X}$ and $\angle \mathbf{x}$ return the unit-modulus phasors of \mathbf{X} and the unit-modulus phasor of \mathbf{x} , respectively. $\text{Cov}(\cdot)$ and $\text{diag}(\cdot)$ denote the covariance and diagonal matrices, respectively. $\text{SVD}(\mathbf{X})$ returns the singular value decomposition (SVD) of \mathbf{X} . Element of \mathbf{X} at the m -th row and n -th column is denoted as $X(m, n)$. Vector of zeros of size M is denoted as $\mathbf{0}_{M \times 1}$. Operators $|\mathbf{X}|$ and $|x|$ return a matrix of moduli of \mathbf{X} and the modulus of scalar x , respectively.

II. SYSTEM MODEL

We consider a single-cell mmWave FD system consisting of one hybrid FD BS serving J DL and K UL fully digital

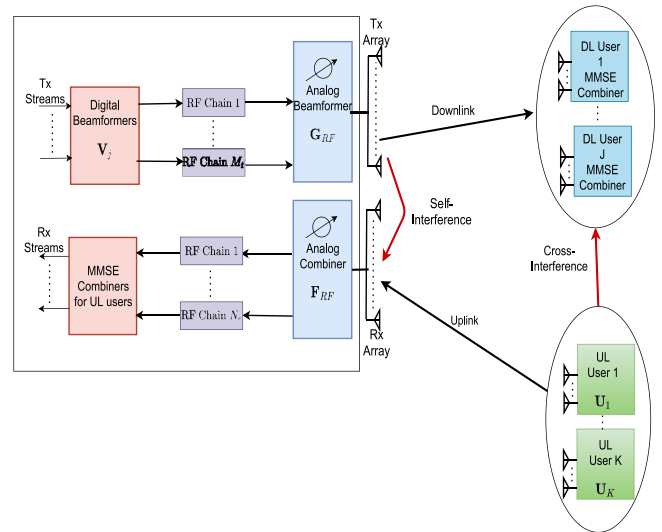


FIGURE 1. FD in mmWave with HYBF to serve multi-antenna users. Tx and Rx denote transmit and receive, respectively.

multi-antenna users, as shown in Fig. 1. We assume perfect channel state information (CSI).² The FD BS is assumed to have M_t transmit and N_r receive RF chains, and M_0 transmit and N_0 receive antennas. Let $\mathcal{U} = \{1, \dots, K\}$ and $\mathcal{D} = \{1, \dots, J\}$ denote the sets containing the indices of K UL and J DL users, respectively. Let M_k and N_j denote the number of transmit and receive antennas for k -th UL and j -th DL user, respectively. We consider a multi-stream approach and the number of data streams for k -th UL and j -th DL user are denoted as u_k and v_j , respectively. Let $\mathbf{U}_k \in \mathbb{C}^{M_k \times u_k}$ and $\mathbf{V}_j \in \mathbb{C}^{M_t \times v_j}$ denote the precoders for white unitary variance data streams $\mathbf{s}_k \in \mathbb{C}^{u_k \times 1}$ and $\mathbf{s}_j \in \mathbb{C}^{v_j \times 1}$, respectively. Let $\mathbf{G}_{RF} \in \mathbb{C}^{M_0 \times M_t}$ and $\mathbf{F}_{RF} \in \mathbb{C}^{N_0 \times N_r}$ denote the fully connected analog beamformer and combiner at the FD BS, respectively. Let $\mathcal{P} = \{1, e^{i2\pi/n_{ps}}, \dots, e^{i2\pi(n_{ps}-1)/n_{ps}}\}$ denote the set of n_{ps} possible discrete values that the phasors at the analog stage can assume on unit-circle.

For HYBF with the unit-modulus constraint, we define the quantizer function $\mathcal{Q}_P(\cdot)$ to quantize the unit-modulus phasors of analog beamformer \mathbf{G}_{RF} and combiner \mathbf{F}_{RF} such that $\mathcal{Q}_P(\angle \mathbf{G}_{RF}(m, n)) \in \mathcal{P}$ and $\mathcal{Q}_P(\angle \mathbf{F}_{RF}(m, n)) \in \mathcal{P}$, $\forall m, n$. For HYBF with amplitude control, the phase part is still quantized with $\mathcal{Q}_P(\cdot)$ and belongs to \mathcal{P} . Let $\mathcal{A} = \{a_0, \dots, a_{A-1}\}$ denote the set of A possible values that the amplitudes can assume. Let $\mathcal{Q}_A(\cdot)$ denote the quantizer function to quantize the amplitudes of \mathbf{G}_{RF} and \mathbf{F}_{RF} such that $\mathcal{Q}_A(|\mathbf{G}_{RF}(m, n)|) \in \mathcal{A}$ and $\mathcal{Q}_A(|\mathbf{F}_{RF}(m, n)|) \in \mathcal{A}$, $\forall m, n$. A complex number $\mathbf{G}_{RF}(m, n)$ with amplitude in \mathcal{A} and phase part in \mathcal{P} can be written as $\mathbf{G}_{RF}(m, n) = \mathcal{Q}_A(|\mathbf{G}_{RF}(m, n)|)\mathcal{Q}_P(\angle \mathbf{G}_{RF}(m, n))$. The thermal noise vectors for FD BS and j -th DL user are denoted as $\mathbf{n}_0 \sim \mathcal{CN}(0, \sigma_0^2 \mathbf{I}_{N_0})$ and $\mathbf{n}_j \sim \mathcal{CN}(0, \sigma_j^2 \mathbf{I}_{N_j})$, respectively. Let \mathbf{c}_k and \mathbf{e}_j denote the LDR noise vectors

2. The CSI of the mmWave FD systems can be acquired similarly as in [53] for the mmWave HD system and it is part of the ongoing research [54].

for k -th UL and j -th DL user, respectively, which can be modeled as [12]

$$\mathbf{c}_k \sim \mathcal{CN}(\mathbf{0}_{M_k \times 1}, k_k \text{diag}(\mathbf{U}_k \mathbf{U}_k^H)), \quad (1)$$

$$\mathbf{e}_j \sim \mathcal{CN}(\mathbf{0}_{N_j \times 1}, \beta_j \text{diag}(\Phi_j)), \quad (2)$$

where $k_k \ll 1$, $\beta_j \ll 1$, $\Phi_j = \text{Cov}(\mathbf{r}_j)$ and \mathbf{r}_j denotes the undistorted signal received by j -th DL user. Let \mathbf{c}_0 and \mathbf{e}_0 denote the LDR noise vectors in transmission and reception for FD BS, respectively. We model them as

$$\mathbf{c}_0 \sim \mathcal{CN}\left(\mathbf{0}_{M_0 \times 1}, k_0 \text{diag}\left(\sum_{n \in \mathcal{D}} \mathbf{G}_{RF} \mathbf{V}_n \mathbf{V}_n^H \mathbf{G}_{RF}^H\right)\right), \quad (3)$$

$$\mathbf{e}_0 \sim \mathcal{CN}(\mathbf{0}_{N_r \times 1}, \beta_0 \text{diag}(\Phi_0)), \quad (4)$$

where $k_0 \ll 1$, $\beta_0 \ll 1$, $\Phi_0 = \text{Cov}(\mathbf{r}_0)$ and \mathbf{r}_0 denotes the undistorted signal received by FD BS after the analog combiner \mathbf{F}_{RF} . Note that (3) extends the transmit LDR noise model from [12] to the case of a hybrid transmitter. For the hybrid receiver at the mmWave FD BS, ADCs, the most dominant sources of receive LDR noise, are placed after the analog combiner \mathbf{F}_{RF} . Consequently, \mathbf{e}_0 in (4) considers the undistorted signal received after the analog combiner. We remark that the extension presented in (3)-(4) is slightly simplified. In practice, as some circuitry might be shared among multiple antennas, it can lead to some correlation.

Let \mathbf{y} and \mathbf{y}_j denote the signals received by the FD BS and j -th DL user, respectively, which can be written as

$$\begin{aligned} \mathbf{y} &= \mathbf{F}_{RF}^H \sum_{k \in \mathcal{U}} \mathbf{H}_k \mathbf{U}_k \mathbf{s}_k + \mathbf{F}_{RF}^H \sum_{k \in \mathcal{U}} \mathbf{H}_k \mathbf{U}_k \mathbf{c}_k + \mathbf{F}_{RF}^H \mathbf{n}_0 \\ &+ \mathbf{F}_{RF}^H \mathbf{H}_0 \sum_{j \in \mathcal{D}} \mathbf{G}_{RF} \mathbf{V}_j \mathbf{s}_j + \mathbf{F}_{RF}^H \mathbf{H}_0 \mathbf{c}_0 + \mathbf{e}_0, \end{aligned} \quad (5)$$

$$\begin{aligned} \mathbf{y}_j &= \mathbf{H}_j \sum_{n \in \mathcal{D}} \mathbf{G}_{RF} \mathbf{V}_n \mathbf{s}_n + \mathbf{H}_j \sum_{n \in \mathcal{D}} \mathbf{G}_{RF} \mathbf{V}_n \mathbf{c}_0 + \mathbf{e}_j + \mathbf{n}_j \\ &+ \sum_{k \in \mathcal{U}} \mathbf{H}_{j,k} \mathbf{U}_k \mathbf{s}_k + \sum_{k \in \mathcal{U}} \mathbf{H}_{j,k} \mathbf{c}_k. \end{aligned} \quad (6)$$

The matrices $\mathbf{H}_k \in \mathbb{C}^{N_0 \times M_k}$ and $\mathbf{H}_j \in \mathbb{C}^{N_j \times M_0}$ denote channel response from the k -th UL user to BS and from the BS to j -th DL user, respectively. The matrices $\mathbf{H}_0 \in \mathbb{C}^{N_0 \times M_0}$ and $\mathbf{H}_{j,k} \in \mathbb{C}^{N_j \times M_k}$ denote SI channel response for FD BS and cross-interference channel response between k -th UL and j -th DL users, respectively. At the mmWave, the channel response \mathbf{H}_k can be modeled as [23]

$$\mathbf{H}_k = \sqrt{\frac{M_k N_0}{N_c N_p}} \sum_{n_c=1}^{N_c} \sum_{n_p=1}^{N_p} \alpha_k^{(n_p, n_c)} \mathbf{a}_r(\phi_k^{n_p, n_c}) \mathbf{a}_t^T(\theta_k^{n_p, n_c}), \quad (7)$$

where N_c and N_p denote the number of clusters and number of rays ([23, Fig. 1]), respectively, and $\alpha_k^{(n_p, n_c)} \sim \mathcal{CN}(0, 1)$ denotes a complex Gaussian random variable whose amplitude and phase are Rayleigh and uniformly distributed, respectively. The vectors $\mathbf{a}_r(\phi_k^{n_p, n_c})$ and $\mathbf{a}_t^T(\theta_k^{n_p, n_c})$ denote the receive and transmit antenna array response with angle of arrival (AoA) $\phi_k^{n_p, n_c}$ and angle of departure (AoD) $\theta_k^{n_p, n_c}$,

TABLE 1. Notations.

M_t	Number of transmit RF chains for the BS
N_r	Number of receive RF chains for the FD BS
M_0	Number of transmit antennas for the BS
N_0	Number of receive antennas for the BS
M_k	Number of transmit antennas for UL user k
N_j	Number of receive antennas for DL user j
\mathbf{U}_k	Digital beamformer for UL user k
\mathbf{V}_j	Digital beamformer for DL user j
\mathbf{G}_{RF}	Analog beamformer for the FD BS
\mathbf{F}_{RF}	Analog combiner for the FD BS
\mathbf{c}_k	Transmit LDR noise from UL user k
\mathbf{c}_0	Transmit LDR noise from the FD BS
\mathbf{e}_0	Receive LDR noise at the FD BS
\mathbf{e}_j	Receive LDR noise at the DL user j
\mathbf{n}_0	Thermal noise at the FD BS
\mathbf{n}_j	Thermal noise at the DL user j
\mathbf{H}_0	SI channel
\mathbf{H}_k	Channel between the BS and UL user k
\mathbf{H}_j	Channel between the BS and DL user j
$\mathbf{H}_{j,k}$	Cross-interference channel between UL user k and DL user j
\preceq	Element-wise inequality

respectively. The channel matrices \mathbf{H}_j and $\mathbf{H}_{j,k}$ can be modeled similarly as in (7). The SI channel can be modeled as [23]

$$\mathbf{H}_0 = \sqrt{\frac{\kappa}{\kappa + 1}} \mathbf{H}_{LoS} + \sqrt{\frac{1}{\kappa + 1}} \mathbf{H}_{ref}, \quad (8)$$

where κ denotes the Rician factor, and the matrices \mathbf{H}_{LoS} and \mathbf{H}_{ref} denote the line-of-sight (LoS) and reflected contributions, respectively. The channel matrix \mathbf{H}_{ref} can be modeled as (7) and element of \mathbf{H}_{LoS} at the m -th row and n -th column can be modeled as [23]

$$\mathbf{H}_{LoS}(m, n) = \frac{\rho}{r_{m,n}} e^{-i2\pi \frac{r_{m,n}}{\lambda}}. \quad (9)$$

The scalar ρ denotes the power normalization constant to assure $\mathbb{E}(\|\mathbf{H}_{LoS}(m, n)\|_F^2) = M_0 N_0$ and λ denotes the wavelength. The scalar $r_{m,n}$ denotes distance between the m -th receive and n -th transmit antenna, which depends on the transmit and receive array geometry (9) [23]. The aforementioned notations are summarized in Table 1.

A. PROBLEM FORMULATION

Let \bar{k} and \bar{j} denote the indices in sets \mathcal{U} and \mathcal{D} without the elements k and j , respectively. The received (signal plus) interference and noise covariance matrices from UL user $k \in \mathcal{U}$ at the BS and by the DL user $j \in \mathcal{D}$ are denoted as (\mathbf{R}_k) $\mathbf{R}_{\bar{k}}$ and (\mathbf{R}_j) $\mathbf{R}_{\bar{j}}$, respectively. Let $\mathbf{T}_k, \forall k \in \mathcal{U}$, and $\mathbf{Q}_j, \forall j \in \mathcal{D}$, defined as

$$\mathbf{T}_k = \mathbf{U}_k \mathbf{U}_k^H, \quad (10a)$$

$$\mathbf{Q}_j = \mathbf{G}_{RF} \mathbf{V}_j \mathbf{V}_j^H \mathbf{G}_{RF}^H, \quad (10b)$$

denote the transmit covariance matrices from UL user $k \in \mathcal{U}$ and DL user $j \in \mathcal{D}$, respectively. By considering the distortions from non-ideal hardware with the extended LDR noise

model, cross-interference, interference and SI, the received covariance matrices at the BS after the analog combiner, i.e., \mathbf{R}_k and $\mathbf{R}_{\bar{k}}$, and at the DL user $j \in \mathcal{D}$, i.e., \mathbf{R}_j and $\mathbf{R}_{\bar{j}}$, can be written as (11), shown at the bottom of the page. In (11), \mathbf{S}_k and \mathbf{S}_j denote the useful received signal covariance matrices from k -th UL user at the FD BS and by j -th DL user, respectively. The undistorted received covariance matrices can be recovered from (11) as $\Phi_0 = \mathbf{R}_k$, with $\beta_0 = 0$, and $\Phi_j = \mathbf{R}_j$, with $\beta_j = 0$.

The WSR maximization problem with respect to the digital beamformers, analog beamformer and combiner with amplitudes in \mathcal{A} and phase part in \mathcal{P} , under the joint sum-power and per-antenna power constraints, can be stated as

$$\max_{\mathbf{U}, \mathbf{V}, \mathbf{G}_{RF}, \mathbf{F}_{RF}} \sum_{k \in \mathcal{U}} w_k \text{Indet}(\mathbf{R}_k^{-1} \mathbf{R}_k) + \sum_{j \in \mathcal{D}} w_j \text{Indet}(\mathbf{R}_j^{-1} \mathbf{R}_j) \quad (12a)$$

$$\text{s.t. } \text{diag}(\mathbf{U}_k \mathbf{U}_k^H) \leq \mathbf{\Lambda}_k, \quad \forall k \in \mathcal{U}, \quad (12b)$$

$$\text{diag}\left(\sum_{j \in \mathcal{D}} \mathbf{G}_{RF} \mathbf{V}_j \mathbf{V}_j^H \mathbf{G}_{RF}^H\right) \leq \mathbf{\Lambda}_0, \quad (12c)$$

$$\text{Tr}(\mathbf{U}_k \mathbf{U}_k^H) \leq \alpha_k, \quad \forall k \in \mathcal{U}, \quad (12d)$$

$$\text{Tr}\left(\sum_{j \in \mathcal{D}} \mathbf{G}_{RF} \mathbf{V}_j \mathbf{V}_j^H \mathbf{G}_{RF}^H\right) \leq \alpha_0. \quad (12e)$$

$$\angle \mathbf{G}_{RF}(m, n) \in \mathcal{P}, \text{ and } |\mathbf{G}_{RF}(m, n)| \in \mathcal{A}, \quad \forall m, n, \quad (12f)$$

$$\angle \mathbf{F}_{RF}(i, j) \in \mathcal{P}, \text{ and } |\mathbf{F}_{RF}(i, j)| \in \mathcal{A}, \quad \forall i, j. \quad (12g)$$

The scalars w_k and w_j denote rate weights for the UL user k and DL user j , respectively. The diagonal matrices $\mathbf{\Lambda}_k$ and $\mathbf{\Lambda}_0$ denote per-antenna power constraints for the k -th UL user and FD BS, respectively, and the scalars α_k and α_0 denote their sum-power constraint. The collections of digital UL and DL beamformers are denoted as \mathbf{U} and \mathbf{V} , respectively. For unit-modulus HYBF, the constraints in (12f) – (12g) on the amplitude part become unit-modulus.

Remark 1: Note that the rate achieved with (12) is not affected by the digital receivers if they are chosen as the MMSE combiners, see, e.g., (4) – (9) [55] for more details.

For WSR maximization, only the analog combiner has to be considered in the optimization problem as it affects the size of the received covariance matrices from UL users, i.e., the UL rate.

III. MINORIZATION-MAXIMIZATION

Problem (12) is non-concave in the transmit covariance matrices \mathbf{T}_k and \mathbf{Q}_j due to the interference terms and searching its globally optimum solution is very challenging. In this section, we present the minorization-maximization optimization method [45] for solving (12) to a local optimum.

The WSR maximization problem (12) will be reformulated at each iteration as a concave reformulation with its minorizer, using the difference-of-convex (DC) programming [56] in terms of the variable to be updated, while the other variables will be kept fixed. To proceed, note that the WSR in (12) can be written with the weighted-rate (WR) of user $k \in \mathcal{U}$, user $j \in \mathcal{D}$, WSRs for \bar{k} and \bar{j} as

$$\text{WSR} = \underbrace{\text{WR}_k^{UL} + \text{WR}_{\bar{k}}^{UL}}_{\triangleq \text{WSR}^{UL}} + \underbrace{\text{WR}_j^{DL} + \text{WR}_{\bar{j}}^{DL}}_{\triangleq \text{WSR}^{DL}}, \quad (13)$$

where WSR^{UL} and WSR^{DL} denote the WSR in UL and DL, respectively. Considering the dependence of the transmit covariance matrices, only WR_k^{UL} is concave in \mathbf{T}_k , meanwhile $\text{WR}_{\bar{k}}^{UL}$ and WR^{DL} are non-concave in \mathbf{T}_k , when $\mathbf{T}_{\bar{k}}$ and \mathbf{Q}_j , $\forall j \in \mathcal{D}$, are fixed. Similarly, only WR_j^{DL} is concave in \mathbf{Q}_j and non-concave in $\text{WR}_{\bar{j}}^{DL}$ and WR_k^{UL} , when $\mathbf{Q}_{\bar{j}}$ and \mathbf{T}_k , $\forall k \in \mathcal{U}$, are fixed. Since a linear function is simultaneously convex and concave, DC programming introduces the first order Taylor series expansion of WR_k^{UL} and WR^{DL} in \mathbf{T}_k , around $\hat{\mathbf{T}}_k$ (i.e., around all \mathbf{T}_k), and of $\text{WR}_{\bar{j}}^{DL}$ and WR^{UL} in \mathbf{Q}_j , around $\hat{\mathbf{Q}}_j$ (i.e., around all \mathbf{Q}_j). Let $\hat{\mathcal{T}}$ and $\hat{\mathcal{Q}}$ denote the set containing all such $\hat{\mathbf{T}}_k$ and $\hat{\mathbf{Q}}_j$, respectively. Let $\hat{\mathbf{R}}_k(\hat{\mathcal{T}}, \hat{\mathcal{Q}})$, $\hat{\mathbf{R}}_{\bar{k}}(\hat{\mathcal{T}}, \hat{\mathcal{Q}})$, $\hat{\mathbf{R}}_j(\hat{\mathcal{T}}, \hat{\mathcal{Q}})$, and $\hat{\mathbf{R}}_{\bar{j}}(\hat{\mathcal{T}}, \hat{\mathcal{Q}})$ denote the covariance matrices \mathbf{R}_k , $\mathbf{R}_{\bar{k}}$, \mathbf{R}_j and $\mathbf{R}_{\bar{j}}$ as a function of $\hat{\mathcal{T}}$ and $\hat{\mathcal{Q}}$, respectively. The linearized tangent expressions for each communication link by computing the gradients

$$\hat{\mathbf{A}}_k = -\left. \frac{\partial \text{WSR}_k^{UL}}{\partial \mathbf{T}_k} \right|_{\hat{\mathcal{T}}, \hat{\mathcal{Q}}}, \quad \hat{\mathbf{B}}_k = -\left. \frac{\partial \text{WSR}^{DL}}{\partial \mathbf{T}_k} \right|_{\hat{\mathcal{T}}, \hat{\mathcal{Q}}}, \quad (14a)$$

$$\begin{aligned} \mathbf{R}_k &= \underbrace{\mathbf{F}_{RF}^H \mathbf{H}_k \mathbf{T}_k \mathbf{H}_k^H \mathbf{F}_{RF}}_{\triangleq \mathbf{S}_k} + \sum_{\substack{i \in \mathcal{U} \\ i \neq k}} \mathbf{F}_{RF}^H \mathbf{H}_i \mathbf{T}_i \mathbf{H}_i^H \mathbf{F}_{RF} + \sum_{i \in \mathcal{U}} k_i \mathbf{F}_{RF}^H \mathbf{H}_i \text{diag}(\mathbf{T}_i) \mathbf{H}_i^H \mathbf{F}_{RF} + \sigma_0^2 \mathbf{I}_{N_0} + \beta_0 \text{diag}(\Phi_0) \\ &+ \mathbf{F}_{RF}^H \mathbf{H}_0 \left(\sum_{n \in \mathcal{D}} \mathbf{Q}_n + k_0 \text{diag} \left(\sum_{n \in \mathcal{D}} \mathbf{Q}_n \right) \right) \mathbf{H}_0^H \mathbf{F}_{RF} \end{aligned} \quad (11a)$$

$$\mathbf{R}_j = \underbrace{\mathbf{H}_j \mathbf{Q}_j \mathbf{H}_j^H}_{\triangleq \mathbf{S}_j} + \mathbf{H}_j \sum_{\substack{n \in \mathcal{D} \\ n \neq j}} \mathbf{Q}_n \mathbf{H}_j^H + k_0 \mathbf{H}_j \text{diag} \left(\sum_{n \in \mathcal{D}} \mathbf{Q}_n \right) \mathbf{H}_j^H + \sigma_j^2 \mathbf{I}_{N_j} + \sum_{i \in \mathcal{U}} \mathbf{H}_{j,i} (\mathbf{T}_i + k_i \text{diag}(\mathbf{T}_i)) \mathbf{H}_{j,i}^H + \beta_j \text{diag}(\Phi_j), \quad (11b)$$

$$\mathbf{R}_{\bar{k}} = \mathbf{R}_k - \mathbf{S}_k, \quad \mathbf{R}_{\bar{j}} = \mathbf{R}_j - \mathbf{S}_j \quad (11c)$$

$$\hat{\mathbf{C}}_j = -\frac{\partial \text{WSR}_j^{DL}}{\partial \mathbf{Q}_j} \Big|_{\hat{\mathbf{T}}, \hat{\mathbf{Q}}}, \quad \hat{\mathbf{D}}_j = -\frac{\partial \text{WSR}_j^{UL}}{\partial \mathbf{Q}_j} \Big|_{\hat{\mathbf{T}}, \hat{\mathbf{Q}}}, \quad (14b)$$

with respect to the transmit covariance matrices \mathbf{T}_k and \mathbf{Q}_j can be written as

$$\underline{\text{WSR}}_k^{UL}(\mathbf{T}_k, \hat{\mathbf{T}}, \hat{\mathbf{Q}}) = \text{WSR}_k^{UL}(\hat{\mathbf{T}}, \hat{\mathbf{Q}}) - \text{Tr}((\mathbf{T}_k - \hat{\mathbf{T}}_k) \hat{\mathbf{A}}_k), \quad (15a)$$

$$\underline{\text{WSR}}_k^{DL}(\mathbf{T}_k, \hat{\mathbf{T}}, \hat{\mathbf{Q}}) = \text{WSR}_k^{DL}(\hat{\mathbf{T}}, \hat{\mathbf{Q}}) - \text{Tr}((\mathbf{T}_k - \hat{\mathbf{T}}_k) \hat{\mathbf{B}}_k), \quad (15b)$$

$$\underline{\text{WSR}}_j^{DL}(\mathbf{Q}_j, \hat{\mathbf{Q}}, \hat{\mathbf{T}}) = \text{WSR}_j^{DL}(\hat{\mathbf{T}}, \hat{\mathbf{Q}}) - \text{Tr}((\mathbf{Q}_j - \hat{\mathbf{Q}}_j) \hat{\mathbf{C}}_j), \quad (15c)$$

$$\underline{\text{WSR}}_j^{UL}(\mathbf{Q}_j, \hat{\mathbf{Q}}, \hat{\mathbf{T}}) = \text{WSR}_j^{UL}(\hat{\mathbf{T}}, \hat{\mathbf{Q}}) - \text{Tr}((\mathbf{Q}_j - \hat{\mathbf{Q}}_j) \hat{\mathbf{D}}_j). \quad (15d)$$

We remark that the tangent expressions (15a)-(15d) constitute a touching lower bound for WSR_k^{UL} , WSR_j^{DL} , WSR_k^{DL} and WSR_j^{UL} , respectively. Hence, the DC programming approach is also a minorization-maximization approach, regardless of the restatement of the transmit covariance matrices \mathbf{T}_k and \mathbf{Q}_j as a function of the beamformers.

Theorem 1: The gradients $\hat{\mathbf{A}}_k$ and $\hat{\mathbf{B}}_k$ which linearize WSR_k^{UL} and WSR_k^{DL} , respectively, with respect to \mathbf{T}_k , $\forall k \in \mathcal{U}$, and the gradients $\hat{\mathbf{C}}_j$ and $\hat{\mathbf{D}}_j$ which linearize WSR_j^{DL} and WSR_j^{UL} , respectively, with respect to \mathbf{Q}_j , $\forall j \in \mathcal{D}$, with the first order Taylor series expansion are given in (16), shown at the bottom of the page.

Proof: Please see Appendix A. ■

A. CONCAVE REFORMULATION

In this section, we simplify the non-concave WSR maximization problem (12). By using the gradients (16), (12) can be reformulated as (17), given at the bottom of the page.

Lemma 1: The WSR maximization problem (12) for a single-cell mmWave FD system with multi-antenna users reformulated at each iteration with its first-order Taylor series expansion as in (17) is a concave reformulation for each link.

Proof: The optimization problem (12) restated as in (17) for each link is made of a concave part, i.e., $\log(\cdot)$, and a linear part, i.e., $\text{Tr}(\cdot)$. Since a linear function is simultaneously concave and non-concave, (17) results to be concave for each link. ■

$$\begin{aligned} \hat{\mathbf{A}}_k = \sum_{i \in \mathcal{U}, i \neq k} w_i & \left(\mathbf{H}_k^H \mathbf{F}_{RF} \left[\hat{\mathbf{R}}_i(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_i(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \beta_0 \text{diag} \left(\hat{\mathbf{R}}_i(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_i(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} \right) \right] \mathbf{F}_{RF}^H \mathbf{H}_k \right. \\ & \left. - k_i \text{diag} \left(\mathbf{H}_k^H \mathbf{F}_{RF} \left(\hat{\mathbf{R}}_i(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_i(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} \right) \mathbf{F}_{RF}^H \mathbf{H}_k \right) \right) \end{aligned} \quad (16a)$$

$$\begin{aligned} \hat{\mathbf{B}}_k = \sum_{l \in \mathcal{D}} w_l & \left(\mathbf{H}_{l,k}^H \left[\hat{\mathbf{R}}_l(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_l(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \beta_j \text{diag} \left(\hat{\mathbf{R}}_l(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_l(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} \right) \right] \mathbf{H}_{l,k} \right. \\ & \left. - k_k \text{diag} \left(\mathbf{H}_{l,k}^H \left(\hat{\mathbf{R}}_l(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_l(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} \right) \mathbf{H}_{l,k} \right) \right) \end{aligned} \quad (16b)$$

$$\begin{aligned} \hat{\mathbf{C}}_j = \sum_{n \in \mathcal{D}, n \neq j} w_n & \left(\mathbf{H}_n^H \left[\hat{\mathbf{R}}_n(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_n(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \beta_n \text{diag} \left(\hat{\mathbf{R}}_n(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_n(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} \right) \right] \mathbf{H}_n \right. \\ & \left. - k_0 \text{diag} \left(\mathbf{H}_n^H \left(\hat{\mathbf{R}}_n(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_n(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} \right) \mathbf{H}_n \right) \right) \end{aligned} \quad (16c)$$

$$\begin{aligned} \hat{\mathbf{D}}_j = \sum_{m \in \mathcal{U}} w_m & \left(\mathbf{H}_0^H \mathbf{F}_{RF} \left[\hat{\mathbf{R}}_m(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_m(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \beta_0 \text{diag} \left(\hat{\mathbf{R}}_m(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_m(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} \right) \right] \mathbf{F}_{RF}^H \mathbf{H}_0 \right. \\ & \left. - k_0 \text{diag} \left(\mathbf{H}_0^H \mathbf{F}_{RF} \left(\hat{\mathbf{R}}_m(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} - \hat{\mathbf{R}}_m(\hat{\mathbf{T}}, \hat{\mathbf{Q}})^{-1} \right) \mathbf{F}_{RF}^H \mathbf{H}_0 \right) \right) \end{aligned} \quad (16d)$$

$$\begin{aligned} \max_{\substack{\mathbf{U}, \mathbf{V} \\ \mathbf{G}_{RF}, \mathbf{F}_{RF}}} & \sum_{k \in \mathcal{U}} \left[w_k \text{Indet} \left(\mathbf{I} + \mathbf{U}_k^H \mathbf{H}_k^H \mathbf{F}_{RF} \mathbf{R}_k^{-1} \mathbf{F}_{RF}^H \mathbf{H}_k \mathbf{U}_k \right) - \text{Tr} \left(\mathbf{U}_k^H \left(\hat{\mathbf{A}}_k + \hat{\mathbf{B}}_k \right) \mathbf{U}_k \right) \right] \\ & + \sum_{j \in \mathcal{D}} \left[w_j \text{Indet} \left(\mathbf{I} + \mathbf{V}_j^H \mathbf{G}_{RF}^H \mathbf{H}_j^H \mathbf{R}_j^{-1} \mathbf{H}_j \mathbf{G}_{RF} \mathbf{V}_j \right) - \text{Tr} \left(\mathbf{V}_j^H \mathbf{G}_{RF}^H \left(\hat{\mathbf{C}}_j + \hat{\mathbf{D}}_j \right) \mathbf{G}_{RF} \mathbf{V}_j \right) \right] \\ \text{s.t.} & \quad (12b) - (12g) \end{aligned} \quad (17)$$

Remark 2: The problem (12) and its reformulated version (17) have the same KarushKuhnTucker (KKT) conditions and therefore any sub-optimal (optimal) solution of (17) is also sub-optimal (optimal) for (12).

Let $\Psi_0 = \text{diag}([\psi_1, \dots, \psi_{M_0}])$ and $\Psi_k = \text{diag}([\psi_{k,1}, \dots, \psi_{k,M_k}])$, denote diagonal matrices containing the Lagrange multipliers associated with per-antenna power constraints for the FD BS and UL user k , respectively. Let l_0 and l_1, \dots, l_K denote the Lagrange multipliers associated with the sum-power constraint for FD BS and K UL users, respectively. Let Ψ denote the collection of Lagrange multipliers associated with the per-antenna power constraints, i.e., Ψ_0 and $\Psi_k, \forall k \in \mathcal{U}$. Let L denote the collection of Lagrange multipliers associated with the sum-power constraints. Augmenting the linearized WSR maximization problem (17) with the sum-power and practical per-antenna power constraints, yields the Lagrangian (18), given at the bottom of the page. In (18), unconstrained analog beamformer and combiner are assumed and their constraints will be incorporated later.

IV. HYBRID BEAMFORMING AND COMBINING

This section presents a novel HYBF design for a MU mmWave mMIMO FD system based on alternating optimization. In the following, optimization of the digital beamformers, analog beamformer and analog combiner is presented into separate sub-sections. We will assume the other variables to be fixed during the alternating optimization process while updating one variable. Information of the other variables updated during previous iterations will be captured in the gradients.

A. DIGITAL BEAMFORMING

To optimize the digital beamformers, we take the derivative of (18) with respect to the conjugate of U_k and V_j , which leads to the following KKT conditions

$$\begin{aligned} & \mathbf{H}_k^H \mathbf{F}_{RF} \mathbf{R}_k^{-1} \mathbf{F}_{RF}^H \mathbf{H}_k U_k \left(\mathbf{I} + \mathbf{U}_k^H \mathbf{H}_k^H \mathbf{F}_{RF} \mathbf{R}_k^{-1} \mathbf{F}_{RF}^H \mathbf{H}_k U_k \right)^{-1} \\ & - \left(\hat{\mathbf{A}}_k + \hat{\mathbf{B}}_k + \Psi_k + l_k \mathbf{I} \right) U_k = 0, \end{aligned} \quad (19a)$$

$$\begin{aligned} & \mathbf{G}_{RF}^H \mathbf{H}_j^H \mathbf{R}_j^{-1} \mathbf{H}_j \mathbf{G}_{RF} V_j \left(\mathbf{I} + \mathbf{V}_j^H \mathbf{G}_{RF}^H \mathbf{H}_j^H \mathbf{R}_j^{-1} \mathbf{H}_j \mathbf{G}_{RF} V_j \right)^{-1} \\ & - \mathbf{G}_{RF}^H \left(\hat{\mathbf{C}}_j + \hat{\mathbf{D}}_j + \Psi_0 + l_0 \mathbf{I} \right) \mathbf{G}_{RF} V_j = 0. \end{aligned} \quad (19b)$$

Given (19)-(19b), the digital beamformers can be optimized based on the result stated in the following.

Theorem 2: Digital beamformers U_k and V_j , fixed the other variables, can be optimized as the generalized dominant eigenvector solution of the pair of the following matrices

$$U_k = D_{u_k} \left(\mathbf{H}_k^H \mathbf{F}_{RF} \mathbf{R}_k^{-1} \mathbf{F}_{RF}^H \mathbf{H}_k, \hat{\mathbf{A}}_k + \hat{\mathbf{B}}_k + \Psi_k + l_k \mathbf{I} \right) \quad (20a)$$

$$V_j = D_{v_j} \left(\mathbf{G}_{RF}^H \mathbf{H}_j^H \mathbf{R}_j^{-1} \mathbf{H}_j \mathbf{G}_{RF}, \mathbf{G}_{RF}^H \left(\hat{\mathbf{C}}_j + \hat{\mathbf{D}}_j + \Psi_0 + l_0 \mathbf{I} \right) \mathbf{G}_{RF} \right), \quad (20b)$$

where $D_d(X)$ selects d generalized dominant eigenvectors from matrix X .

Proof: Please see Appendix B. ■

The generalized dominant eigenvector solution provides the optimized beamforming directions but not power [56]. To include the optimal stream power allocation, we normalize the columns of digital beamformers to unit-norm. This operation preserves the optimized beamforming directions and allows to design the optimal power allocation scheme.

B. ANALOG BEAMFORMING

This section presents a novel approach to design the analog beamformer for hybrid FD BS in a MU scenario to maximize the WSR. The structure of the fully connected analog beamformer \mathbf{G}_{RF} is shown in Figure 2. Assuming the remaining variables to be fixed, we first consider the optimization of unconstrained analog beamformer \mathbf{G}_{RF} as

$$\begin{aligned} & \max_{\mathbf{G}_{RF}} \sum_{j \in \mathcal{D}} \left[w_j \ln \det \left(\mathbf{I} + \mathbf{V}_j^H \mathbf{G}_{RF}^H \mathbf{H}_j^H \mathbf{R}_j^{-1} \mathbf{H}_j \mathbf{G}_{RF} V_j \right) \right. \\ & \left. - \text{Tr} \left(\mathbf{V}_j^H \mathbf{G}_{RF}^H \left(\hat{\mathbf{C}}_j + \hat{\mathbf{D}}_j + l_0 \mathbf{I} + \Psi_0 \right) \mathbf{G}_{RF} V_j \right) \right]. \end{aligned} \quad (21)$$

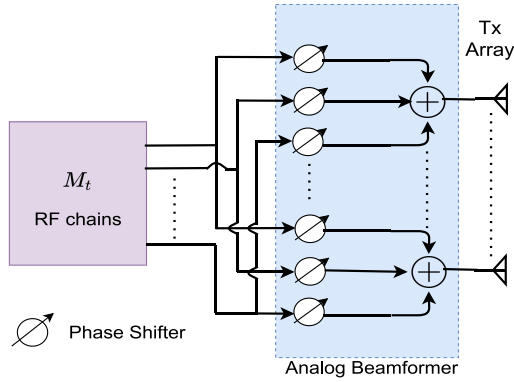
Note that from (17) only the terms shown in (21) depend on the analog combiner \mathbf{G}_{RF} and information about other variables is captured in gradients $\hat{\mathbf{C}}_j$ and $\hat{\mathbf{D}}_j$. To solve (21), we take its derivative with respect to the conjugate of \mathbf{G}_{RF} , which yields the following KKT condition

$$\begin{aligned} & \mathbf{H}_j^H \mathbf{R}_j^{-1} \mathbf{H}_j \mathbf{G}_{RF} V_j V_j^H \left(\mathbf{I} + \mathbf{V}_j V_j^H \mathbf{G}_{RF}^H \mathbf{H}_j^H \mathbf{R}_j^{-1} \mathbf{H}_j \mathbf{G}_{RF} \right)^{-1} \\ & - \left(\hat{\mathbf{C}}_j + \hat{\mathbf{D}}_j + \Psi_0 + l_0 \mathbf{I} \right) \mathbf{G}_{RF} V_j V_j^H = 0. \end{aligned} \quad (22)$$

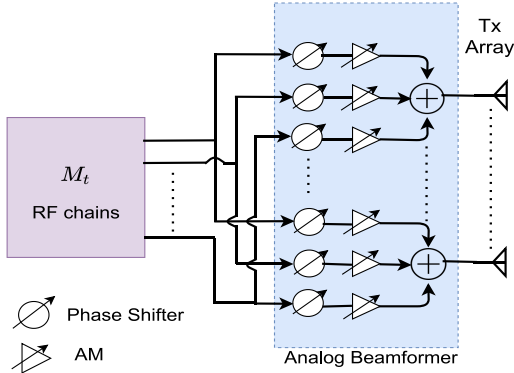
Given (22), the analog beamformer \mathbf{G}_{RF} for mmWave FD BS can be optimized as stated in the following.

Theorem 3: The vectorized unconstrained analog beamformer $\text{vec}(\mathbf{G}_{RF})$ can be optimized as one generalized

$$\begin{aligned} \mathcal{L}(\mathbf{U}, \mathbf{V}, \mathbf{G}_{RF}, \mathbf{F}_{RF}, \Psi, L) &= \sum_{l=0}^K l_l \alpha_l + \text{Tr}(\Psi_0 \Lambda_0) + \sum_{u \in \mathcal{U}} \text{Tr}(\Psi_u \Lambda_u) \\ &+ \sum_{k \in \mathcal{U}} \left[w_k \ln \det \left(\mathbf{I} + \mathbf{U}_k^H \mathbf{H}_k^H \mathbf{F}_{RF} \mathbf{R}_k^{-1} \mathbf{F}_{RF}^H \mathbf{H}_k U_k \right) - \text{Tr} \left(\mathbf{U}_k^H \left(\hat{\mathbf{A}}_k + \hat{\mathbf{B}}_k + l_k \mathbf{I} + \Psi_k \right) U_k \right) \right] \\ &+ \sum_{j \in \mathcal{D}} \left[w_j \ln \det \left(\mathbf{I} + \mathbf{V}_j^H \mathbf{G}_{RF}^H \mathbf{H}_j^H \mathbf{R}_j^{-1} \mathbf{H}_j \mathbf{G}_{RF} V_j \right) - \text{Tr} \left(\mathbf{V}_j^H \mathbf{G}_{RF}^H \left(\hat{\mathbf{C}}_j + \hat{\mathbf{D}}_j + l_0 \mathbf{I} + \Psi_0 \right) \mathbf{G}_{RF} V_j \right) \right] \end{aligned} \quad (18)$$



(a) Analog beamformer with unit-modulus phase shifters.



(b) Analog beamformer with amplitude modulators.

FIGURE 2. Analog beamformer with unit-modulus phase shifters (a) or additionally with amplitude control (b).

dominant eigenvector solution of the pair of the following matrices

$$\text{vec}(\mathbf{G}_{RF}) = \mathbf{D}_1 \left(\sum_{j \in \mathcal{D}} \left(\mathbf{V}_j \mathbf{V}_j^H \left(\mathbf{I} + \mathbf{V}_j \mathbf{V}_j^H \mathbf{G}_{RF}^H \mathbf{H}_j^H \mathbf{R}_j^{-1} \mathbf{H}_j \mathbf{G}_{RF} \right)^{-1} \right)^T \otimes \mathbf{H}_j^H \mathbf{R}_j^{-1} \mathbf{H}_j, \sum_{j \in \mathcal{D}} \left(\mathbf{V}_j \mathbf{V}_j^H \right)^T \otimes \left(\hat{\mathbf{C}}_j + \hat{\mathbf{D}}_j + \boldsymbol{\Psi}_0 + l_0 \mathbf{I} \right) \right), \quad (23)$$

where $\mathbf{D}_1(\mathbf{X})$ selects the first generalized dominant eigenvector from matrix \mathbf{X} .

Proof: Please see Appendix B. ■

Note that Theorem 3 provides the optimized vectorized unconstrained analog beamformer \mathbf{G}_{RF} and we need to reshape it with $\text{unvec}(\text{vec}(\mathbf{G}_{RF}))$. To satisfy the unit-modulus and quantization constraints, we do $\mathbf{G}_{RF}(m, n) = \mathbb{Q}_P(\angle \mathbf{G}_{RF}(m, n))$, $\forall m, n$. For HYBF with AMs, the columns are first scaled to be unit-norm and the quantization constraint is satisfied as $\mathbf{G}_{RF}(m, n) = \mathbb{Q}_A(|\mathbf{G}_{RF}(m, n)|) \mathbb{Q}_P(\angle \mathbf{G}_{RF}(m, n))$, $\forall m, n$.

C. ANALOG COMBINING

This section presents a novel approach to design the analog combiner \mathbf{F}_{RF} for mmWave FD BS to serve multiple UL users. Its design is more straightforward than the analog beamformer. Note that the trace terms appearing

in (17) have the objective to make beamformers' update aware of the interference generated towards other links. However, \mathbf{F}_{RF} being a combiner, does not generate any interference and therefore does not appear in the trace terms of (17). Consequently, to optimize \mathbf{F}_{RF} , we can solve the optimization problem (12) instead of using its minorized version (17). By considering the unconstrained analog combiner \mathbf{F}_{RF} , from (12) we have

$$\max_{\mathbf{F}_{RF}} \sum_{k \in \mathcal{U}} w_k \text{Indet}(\mathbf{R}_k^{-1} \mathbf{R}_k). \quad (24)$$

To solve (24), \mathbf{F}_{RF} has to combine the signal received at the antenna level of hybrid FD BS but \mathbf{R}_k and $\mathbf{R}_{\bar{k}}$ represent the received covariance matrices after analog combining. Let (\mathbf{R}_k^{ant}) $\mathbf{R}_{\bar{k}}^{ant}$ denote the (signal-plus) interference and noise covariance matrix received at the antennas of FD BS, which can be obtained from (\mathbf{R}_k) $\mathbf{R}_{\bar{k}}$ given in (11) by omitting \mathbf{F}_{RF} . After analog combining, we can recover \mathbf{R}_k and $\mathbf{R}_{\bar{k}}$ as $\mathbf{R}_k = \mathbf{F}_{RF}^H \mathbf{R}_k^{ant} \mathbf{F}_{RF}$ and $\mathbf{R}_{\bar{k}} = \mathbf{F}_{RF}^H \mathbf{R}_{\bar{k}}^{ant} \mathbf{F}_{RF}$, respectively, $\forall k \in \mathcal{U}$. Problem (24) can be restated as a function of \mathbf{R}_k^{ant} and $\mathbf{R}_{\bar{k}}^{ant}$ as

$$\max_{\mathbf{F}_{RF}} \sum_{k \in \mathcal{U}} \left[w_k \text{Indet}(\mathbf{F}_{RF}^H \mathbf{R}_k^{ant} \mathbf{F}_{RF}) - w_k \text{Indet}(\mathbf{F}_{RF}^H \mathbf{R}_{\bar{k}}^{ant} \mathbf{F}_{RF}) \right]. \quad (25)$$

In (17), the trace term was only linear, which made the restated optimization problem concave for each link. In (25), all the terms are fully concave. To optimize \mathbf{F}_{RF} , we take the derivative with respect to the conjugate of \mathbf{F}_{RF} , which yields the following KKT condition

$$\sum_{k \in \mathcal{U}} w_k \mathbf{R}_k^{ant} \mathbf{F}_{RF} \left(\mathbf{F}_{RF}^H \mathbf{R}_k^{ant} \mathbf{F}_{RF} \right)^{-1} - \sum_{k \in \mathcal{U}} w_k \mathbf{R}_{\bar{k}}^{ant} \mathbf{F}_{RF} \left(\mathbf{F}_{RF}^H \mathbf{R}_{\bar{k}}^{ant} \mathbf{F}_{RF} \right)^{-1} = 0. \quad (26)$$

It is immediate from (26) that the unconstrained analog combiner can be optimized as the generalized dominant eigenvector solution of the pair of sum of the received covariance matrices at the antenna level from all the K UL users, i.e.,

$$\mathbf{F}_{RF} \rightarrow \mathbf{D}_{Nr} \left(\sum_{k \in \mathcal{U}} w_k \mathbf{R}_k^{ant}, \sum_{k \in \mathcal{U}} w_k \mathbf{R}_{\bar{k}}^{ant} \right). \quad (27)$$

To satisfy the unit-modulus and quantization constraints for \mathbf{F}_{RF} , we do $\mathbf{F}_{RF}(m, n) = \mathbb{Q}_P(\angle \mathbf{F}_{RF}(m, n)) \in \mathcal{P}$, $\forall m, n$. If AMs are available, the columns are scaled to be unit-norm and quantization constraint is satisfied as $\mathbf{F}_{RF}(m, n) = \mathbb{Q}_A(|\mathbf{F}_{RF}(m, n)|) \mathbb{Q}_P(\angle \mathbf{F}_{RF}(m, n))$, $\forall m, n$.

D. OPTIMAL POWER ALLOCATION

Given the normalized digital beamformers and analog beamformer, optimal power allocation can be included while searching for the Lagrange multipliers satisfying the joint sum-power and practical per-antenna power constraints.

Let $\Sigma_k^{(1)}$ and $\Sigma_k^{(2)}$, $\forall k \in \mathcal{U}$ and $\Sigma_j^{(1)}$ and $\Sigma_j^{(2)}$, $\forall j \in \mathcal{D}$, be defined as

$$U_k^H H_k^H F_{RF} R_{\bar{k}}^{-1} F_{RF}^H H_k U_k = \Sigma_k^{(1)}, \quad (28a)$$

$$U_k^H (\hat{A}_k + \hat{B}_k + \Psi_k + l_k I) U_k = \Sigma_k^{(2)}, \quad (28b)$$

$$V_j^H G_{RF}^H H_j^H R_j^{-1} H_j G_{RF} V_j = \Sigma_j^{(1)}, \quad (28c)$$

$$V_j^H G_{RF}^H (\hat{C}_j + \hat{D}_j + \Psi_0 + l_0 I) G_{RF} V_j = \Sigma_j^{(2)}. \quad (28d)$$

Given (28), the optimal stream power allocation can be included based on the result stated in the following.

Lemma 2: Optimal power allocation for the hybrid FD BS and multi-antenna UL users can be obtained by multiplying $\Sigma_j^{(1)}$ and $\Sigma_j^{(2)}$ with the diagonal power matrix P_j , $\forall j \in \mathcal{D}$ and $\Sigma_k^{(1)}$ and $\Sigma_k^{(2)}$ with the diagonal power matrix P_k , $\forall k \in \mathcal{U}$, respectively.

Proof: The beamformers U_k and V_k , are computed as the generalized dominant eigenvectors, which make the matrices $\Sigma_k^{(1)}$, $\Sigma_k^{(2)}$, $\forall k$ and $\Sigma_j^{(1)}$, $\Sigma_j^{(2)}$, $\forall j$ diagonal at each iteration. Multiplying any generalized dominant eigenvector solution matrix with a diagonal matrix still yields a generalized dominant eigenvector solution. Therefore, multiplying $\Sigma_k^{(1)}$, $\Sigma_k^{(2)}$ with P_k , $\forall k \in \mathcal{U}$ and $\Sigma_j^{(1)}$, $\Sigma_j^{(2)}$ with P_j , $\forall j \in \mathcal{D}$ still preserves the validity of optimized beamforming directions. ■

Given the optimized beamformers and fixed Lagrange multipliers, by using the result stated in Lemma 2, stream power allocation optimization problems for UL and DL users can be formally stated as

$$\max_{P_k} \cdot \left[w_k \text{Indet}(\mathbf{I} + \Sigma_k^{(1)} P_k) - \text{Tr}(\Sigma_k^{(2)} P_k), \quad \forall k \in \mathcal{U}, \right] \quad (29a)$$

$$\max_{P_j} \cdot \left[w_j \text{Indet}(\mathbf{I} + \Sigma_j^{(1)} P_j) - \text{Tr}(\Sigma_j^{(2)} P_j), \quad \forall j \in \mathcal{D} \right]. \quad (29b)$$

Solving (29) leads to the following optimal power allocation scheme

$$P_k = \left(w_k (U_k^H (\hat{A}_k + \hat{B}_k + \Psi_k + l_k I) U_k)^{-1} - (U_k^H H_k^H F_{RF} R_{\bar{k}}^{-1} F_{RF}^H H_k U_k)^{-1} \right)^+, \quad (30a)$$

$$P_j = \left(w_j (V_j^H G_{RF}^H (\hat{C}_j + \hat{D}_j + \Psi_0 + l_0 I) G_{RF} V_j)^{-1} - (V_j^H G_{RF}^H H_j^H R_j^{-1} H_j G_{RF} V_j)^{-1} \right)^+, \quad (30b)$$

where $(X)^+ = \max\{\mathbf{0}, X\}$. We remark that the proposed power allocation scheme is interference, SI, cross-interference and LDR noise aware as it takes into account their effect in the gradients, which are updated at each iteration. Fixed the beamformers, we can search for multipliers satisfying the joint constraints while doing water-filling for powers. To do so, consider the dependence of

Lagrangian (18) on multipliers and powers as

$$\begin{aligned} \mathcal{L}(\Psi, \mathbf{L}, \mathbf{P}) = & \sum_{l=0}^K l_l p_l + \text{Tr}(\Psi_0 \Lambda_0) + \sum_{u \in \mathcal{U}} \text{Tr}(\Psi_u \Lambda_u) \\ & + \sum_{k \in \mathcal{U}} \left[w_k \text{Indet}(\mathbf{I} + \Sigma_k^{(1)} P_k) - \text{Tr}(\Sigma_k^{(2)} P_k) \right] \\ & + \sum_{j \in \mathcal{D}} \left[w_j \text{Indet}(\mathbf{I} + \Sigma_j^{(1)} P_j) - \text{Tr}(\Sigma_j^{(2)} P_j) \right], \end{aligned} \quad (31)$$

where \mathbf{P} is the set of stream powers in UL and DL. The multipliers in Ψ and \mathbf{L} should be such that the Lagrange dual function (31) is finite and the values of multipliers should be strictly positive. Formally, Lagrange multipliers' search problem can be stated as

$$\begin{aligned} \min_{\Psi, \mathbf{L}} \max_{\mathbf{P}} \quad & \mathcal{L}(\Psi, \mathbf{L}, \mathbf{P}), \\ \text{s.t.} \quad & \Psi, \mathbf{L} \geq 0. \end{aligned} \quad (32)$$

The dual function $\max_{\mathbf{P}} \mathcal{L}(\Psi, \mathbf{L}, \mathbf{P})$ is the pointwise supremum of a family of functions of Ψ, \mathbf{L} , it is convex [57] and the globally optimal values for Ψ and \mathbf{L} can be obtained by using any of the numerous convex optimization techniques. In this work, we adopt the Bisection algorithm to search the multipliers. Let $\mathcal{M}_0 = \{\lambda_0, \psi_1, \dots, \psi_{M_0}\}$ and $\mathcal{M}_k = \{\lambda_k, \psi_{k,1}, \dots, \psi_{k,M_k}\}$ denote the sets containing Lagrange multipliers associated with the sum-power and practical per-antenna power constraints for FD BS and UL user $k \in \mathcal{U}$, respectively. Let $\underline{\mu}_i$ and $\bar{\mu}_i$ denote the lower and upper bound for the search range of multiplier μ_i , where $\mu_i \in \mathcal{M}_0$ or $\mu_i \in \mathcal{M}_k$. While searching multipliers and performing water-filling for powers, the UL and DL power matrices become non-diagonal. Therefore, we consider the SVD of power matrices to shape them back as diagonal. Namely, let P_i denote the power matrix for user i , where $i \in \mathcal{U}$ or $i \in \mathcal{D}$. When P_i becomes non-diagonal, we consider its SVD as

$$[U_{P_i}, D_{P_i}, V_{P_i}] = \text{SVD}(P_i). \quad (33)$$

where U_{P_i}, D_{P_i} and V_{P_i} are the left unitary, diagonal and right unitary matrices, respectively, obtained with the SVD decomposition, and we set $P_i = D_{P_i}$ to obtain diagonal power matrices.

For unit-modulus HYBF, the complete alternating optimization based procedure to maximize the WSR based on minorization-maximization is formally stated in Algorithm 1. For HYBF with AMs, the steps $\angle G_{RF}$ and $\angle F_{RF}$ must be omitted and amplitudes of the analog beamformer and combiner must be quantized with $\mathbb{Q}_A(\cdot)$. Once the proposed algorithm converges, all the combiners can be chosen as the MMSE combiners, which will not affect the WSR achieved with Algorithm 1 (4) – (9) [55].

Algorithm 1 Practical Hybrid Beamforming Design**Given:** The CSI and rate weights.**Initialize:** $\mathbf{G}_{RF}, \mathbf{V}_j, \forall j \in \mathcal{D}$ and $\mathbf{U}_k, \forall k \in \mathcal{U}$.**Set:** $\mu_i = 0$ and $\bar{\mu}_i = \mu_{i_{max}} \forall i \in \mathcal{M}_0$ or $\forall i \in \mathcal{M}_k, \forall k \in \mathcal{U}$ **Repeat until convergence** Compute \mathbf{G}_{RF} (23), $\text{unvec}(\mathbf{G}_{RF})$ and $\mathbf{G}_{RF} = \angle \mathbf{G}_{RF}$. Compute \mathbf{F}_{RF} with (27), and do $\mathbf{F}_{RF} = \angle \mathbf{F}_{RF}$. **for:** $j = 1:J$ Compute $\hat{\mathbf{C}}_j, \hat{\mathbf{D}}_j$ with (16) Compute \mathbf{V}_j with (20b) and normalize it **end** **Set:** $\mu_0 = 0$ and $\bar{\mu}_0 = \mu_{i_{max}} \forall i \in \mathcal{M}_0$ **for:** $\forall \mu_0 \in \mathcal{M}_0$ **Repeat until convergence** set $\mu_0 = (\mu_0 + \bar{\mu}_0)/2$ Compute \mathbf{P}_j with (30b) $\forall j$ **if** constraint for μ_0 is violated set $\mu_0 = \mu_0$, **else** $\bar{\mu}_0 = \mu_0$ $[\mathbf{U}_{P_j}, \mathbf{D}_{P_j}, \mathbf{V}_{P_j}] = \text{SVD}(\mathbf{P}_j), \forall j$ Set $\mathbf{P}_j = \mathbf{D}_{P_j}$ and $\mathbf{Q}_j = \mathbf{G}_{RF} \mathbf{V}_j \mathbf{P}_j \mathbf{V}_j^H \mathbf{G}_{RF}^H, \forall j$ **for:** $k = 1:K$ Compute $\hat{\mathbf{A}}_k, \hat{\mathbf{B}}_k$ with (16) Compute \mathbf{U}_k with (20a) and normalize it **Set:** $\mu_k = 0$ and $\bar{\mu}_k = \mu_{l_{max}}$ **for:** $\forall \mu_k \in \mathcal{M}_k$ **Repeat until convergence** set $\mu_k = (\mu_k + \bar{\mu}_k)/2$ Compute \mathbf{P}_k with (30a). **if** constraint for μ_0 is violated set $\mu_k = \mu_k$ **else** $\bar{\mu}_k = \mu_k$ $[\mathbf{U}_{P_k}, \mathbf{D}_{P_k}, \mathbf{V}_{P_k}] = \text{SVD}(\mathbf{P}_k)$ Set $\mathbf{P}_k = \mathbf{D}_{P_k}$ and $\mathbf{T}_k = \mathbf{U}_k \mathbf{P}_k \mathbf{U}_k^H$

Repeat

 Quantize $\angle \mathbf{G}_{RF}$ and $\angle \mathbf{F}_{RF}$ ($|\mathbf{G}_{RF}|$ and $|\mathbf{F}_{RF}|$ with AMs)**E. CONVERGENCE**

In our context, the ingredients required to prove the convergence are minorization [45], alternating or cyclic optimization [45], Lagrange dual function [57], saddle-point interpretation [57] and KKT conditions [57]. For the WSR cost function (12), we construct its minorizer as in (15a), (15b), (15c), (15d), which restates the WSR maximization as a concave problem (17) for each link. The minorizer is a touching lower bound for the original WSR problem (12), so we can write

$$\text{WSR} \geq \underline{\text{WSR}} = \underline{\text{WR}}_k^{UL} + \underline{\text{WSR}}_k^{UL} + \underline{\text{WR}}_j^{DL} + \underline{\text{WSR}}_j^{DL}. \quad (34)$$

The minorizer, which is concave in \mathbf{T}_k and \mathbf{Q}_j , still has the same gradient of the original WSR and hence the KKT conditions are not affected. Reparameterizing \mathbf{T}_k or \mathbf{Q}_j in terms of $\mathbf{U}_k, \forall k \in \mathcal{U}$ and \mathbf{G}_{RF} or $\mathbf{V}_j, \forall j \in \mathcal{D}$, respectively, as in (10) with the optimal power matrices and

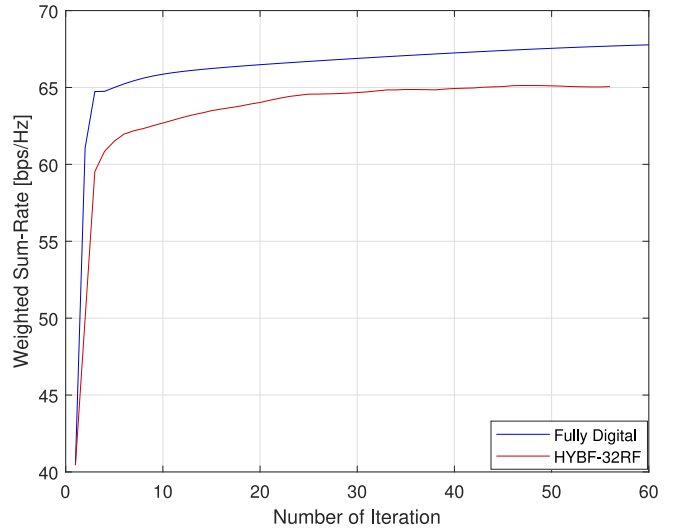


FIGURE 3. Typical convergence behavior of the proposed HYBF for mmWave mMIMO FD system.

adding the power constraints to the minorizer, we get the Lagrangian (18). Every alternating update of \mathcal{L} for $\mathbf{V}_j, \mathbf{G}_{RF}, \mathbf{U}_k, \forall j \in \mathcal{D}, \forall k \in \mathcal{U}$ or for $\mathbf{P}, \mathbf{\Lambda}, \mathbf{\Psi}$ leads to an increase of the WSR, ensuring convergence. For the KKT conditions, at the convergence point, the gradients of \mathcal{L} for $\mathbf{V}_j, \mathbf{G}_{RF}, \mathbf{U}_j$ or \mathbf{P} correspond to the gradients of Lagrangian (12), i.e., for the original WSR problem. For fixed analog and digital beamformers, \mathcal{L} is concave in \mathbf{P} , hence we have a strong duality for the saddle point, i.e.,

$$\max_{\mathbf{P}} \min_{\mathbf{L}, \mathbf{\Psi}} \mathcal{L}(\mathbf{L}, \mathbf{\Psi}, \mathbf{P}). \quad (35)$$

Let \mathbf{X}^* and x^* denote the optimal solution for matrix \mathbf{X} or scalar x at the convergence, respectively. When Algorithm 1 converges, solution of the following optimization problem

$$\min_{\mathbf{\Lambda}, \mathbf{\Psi}} \mathcal{L}(\mathbf{V}^*, \mathbf{G}^*, \mathbf{U}^*, \mathbf{P}^*, \mathbf{L}, \mathbf{\Psi}) \quad (36)$$

satisfies the KKT conditions for powers in \mathbf{P} and the complementary slackness conditions

$$r_0^* \left(\alpha_0 - \sum_{j \in \mathcal{D}} \text{Tr}(\mathbf{G}_{RF}^* \mathbf{V}_j^* \mathbf{P}_j^* \mathbf{V}_j^{*H} \mathbf{G}_{RF}^{*H}) \right) = 0, \quad (37a)$$

$$\text{Tr} \left(\mathbf{\Psi}_0^* \left(\mathbf{P}_0 - \sum_{j \in \mathcal{D}} \text{Tr}(\mathbf{G}_{RF}^* \mathbf{V}_j^* \mathbf{P}_j^* \mathbf{V}_j^{*H} \mathbf{G}_{RF}^{*H}) \right) \right) = 0, \quad (37b)$$

$$r_k^* \left(\alpha_k - \text{Tr}(\mathbf{U}_k^* \mathbf{P}_k^* \mathbf{U}_k^{*H}) \right) = 0, \quad (37c)$$

$$\text{Tr} \left(\mathbf{\Psi}_k^* \left(\mathbf{P}_k - \text{Tr}(\mathbf{U}_k^* \mathbf{P}_k^* \mathbf{U}_k^{*H}) \right) \right) = 0, \quad (37d)$$

where all the individual factors in the products are non-negative, and for per-antenna power constraints $\mathbf{\Psi}_0^*$ and $\mathbf{\Psi}_k^*$, the sum of non-negative terms being zero implies all terms result to be zero.

Remark 3: The unit-modulus HYBF scheme converges to a local optimum where $\angle \mathbf{G}_{RF}(m, n), \angle \mathbf{F}_{RF}(m, n) \in \mathcal{P}$ with $|\mathbf{G}_{RF}(m, n)|, |\mathbf{F}_{RF}(m, n)| = 1, \forall m, n$. Unconstrained HYBF with AMs converges to a different local optimum, where $\angle \mathbf{G}_{RF}(m, n), \angle \mathbf{F}_{RF}(m, n) \in \mathcal{P}$ and $|\mathbf{G}_{RF}(m, n)|, |\mathbf{F}_{RF}(m, n)| \in \mathcal{A}, \forall m, n$. Due to quantization, \mathbf{G}_{RF} and \mathbf{F}_{RF} obtained with Algorithm 1 tend to lose their optimality and consequently achieve less WSR compared to their infinite resolution case. For unit-modulus HYBF, the loss in WSR depends only on the resolution of phases. For HYBF with AMs, the loss in WSR depends on the resolution of both amplitudes and phases.

F. COMPLEXITY ANALYSIS

In this section, we analyze the per-iteration computational complexity of Algorithm 1, assuming that the dimensions of antennas get large. Its one iteration consists in updating K and J digital beamformers for the UL and DL user, respectively, and one analog beamformer and combiner for the FD BS. One dominant generalized eigenvector computation to update analog beamformer \mathbf{G}_{RF} from a matrix of size $M_t M_0 \times M_t M_0$ in (23), is $\mathcal{O}(M_0^2 M_t^2)$. To update the gradients $\hat{\mathbf{A}}_k$ and $\hat{\mathbf{B}}_k$ for one UL user, the complexity is given by $\mathcal{O}((K-1)N_r^3)$ and $\mathcal{O}(JN_j^3)$, respectively. For the gradient $\hat{\mathbf{C}}_j$ and $\hat{\mathbf{D}}_j$, required to update the beamformer of j -th DL user, computational complexity is $\mathcal{O}((J-1)N_j^3)$ and $\mathcal{O}(KN_r^3)$, respectively. Updating the beamformers of k -th UL and j -th DL users as the generalized dominant eigenvectors adds additional complexity of $\mathcal{O}(u_k M_k^2)$ and $\mathcal{O}(v_j N_j^2)$, respectively. The Lagrange multipliers' update associated with the per-antenna power constraints for FD BS and UL users is linear in the number of antennas M_0 or M_k , respectively. However, as we jointly perform the multipliers' search and power allocation, it adds $\mathcal{O}(v_i^3)$, where $i \in \mathcal{D}$ or $i \in \mathcal{U}$, which can be ignored. Updating the analog combiner \mathbf{F}_{RF} for FD BS is $\mathcal{O}(N_r N_0^2)$. Under the assumption that the dimensions of antennas get large, the per-iteration complexity is $\approx \mathcal{O}(K^2 N_r^3 + K J N_j^3 + J^2 N_j^3 + J K N_r^3 + M_0^2 M_t^2 + N_r N_0^2)$ which depends on the number of UL and DL users served by the mmWave FD BS.

V. SIMULATION RESULTS

This section presents simulation results to evaluate the performance of the proposed HYBF scheme. For comparison, we define the following benchmark schemes:

- a) A *Fully digital HD* scheme with LDR noise, serving the UL and DL users with time division duplexing. Being HD, it is neither affected by the SI nor by the cross-interference.
- b) A *Fully digital FD* scheme with LDR noise. This scheme sets an upper bound for the maximum achievable gain by a hybrid FD system.

Hereafter, HYBF designs with the unit-modulus constraint and with AMs are denoted as HYBF-UM and HYBF-AMs, respectively. We define the signal-to-noise-ratio (SNR) for the mmWave mMIMO FD system as

$$\text{SNR} = \alpha_0 / \sigma_0^2, \quad (38)$$

TABLE 2. Simulation parameters to simulate the MU mmWave FD system.

Simulation Parameters		
UL and DL users	K, J	2
Data streams	v_j, u_k	2
Antennas for the BS	M_0, N_0	100, 50
Clusters and Paths	N_c, N_p	3, 3
RF chains (BS)	$M_t = N_r$	8, 10, 16 or 32
User antennas	$M_k = N_j$	5
Rician Factor	κ	1
Tx and Rx array response	$\mathbf{a}_r, \mathbf{a}_t$	ULA, ULA
Angles	$\phi_k, \phi_j, \theta_k, \theta_j$	$\mathcal{U} \sim [-30^\circ, 30^\circ]$
Rate weights	w_k, w_j	1
Uniform Quantizer	$\mathbb{Q}_P(\cdot), \mathbb{Q}_A(\cdot)$	8, 3 bits
Angle between Tx and Rx array (BS)	Θ	90°
Antenna array separation (BS)	D	20 cm
Per-antenna power constraint	$\mathbf{\Lambda}_k, \mathbf{\Lambda}_0$	$\alpha_k / M_k \mathbf{I}, \alpha_0 / M_0 \mathbf{I}$

where the scalars α_0 and σ_0^2 denote the total transmit power and thermal noise variance for FD BS, respectively. We set the thermal noise level for DL users to be $\sigma_0^2 = \sigma_j^2, \forall j$, and the transmit power for UL users as $\alpha_0 = \alpha_k, \forall k$. We consider the total transmit power normalized to 1 and choose the noise variance based on desired SNR. To compare the gain of a FD system over a HD system, we define the additional gain in percentage as

$$\text{Gain} = \frac{WSR_{FD} - WSR_{HD}}{WSR_{HD}} \times 100 [\%], \quad (39)$$

where WSR_{FD} and WSR_{HD} denote the WSR of a FD and HD system, respectively. To evaluate the performance, we set the per-antenna power constraints for FD BS and UL users as the total transmit power divided by the number of antennas, i.e., $\alpha_0 / M_0 \mathbf{I}$ and $\alpha_k / M_k \mathbf{I}, \forall k$. The BS and users are assumed to be equipped with a uniform linear array (ULA) with antennas separated by half-wavelength. The transmit and receive antenna array at the BS are assumed to be placed $D = 20$ cm apart, with the relative angle $\Theta = 90^\circ$, and $r_{m,n}$ is modeled as (9) [23]. The Rician factor κ for the SI channel is set to be 1. We assume that the FD BS has $M_0 = 100$ transmit and $N_0 = 50$ receive antennas. It serves two UL and two DL users with $M_k = N_j = 5$ antennas and with 2 data streams for each user. The phases for both designs are quantized in the interval $[0, 2\pi]$ with an 8-bit uniform quantizer $\mathbb{Q}_P(\cdot)$. For HYBF with AMs, the amplitudes are uniformly quantized with a 3-bit uniform quantizer $\mathbb{Q}_A(\cdot)$ in the interval $[0, a_{max}]$, where $a_{max} = \max\{\max\{|\mathbf{G}_{RF}|\}, \max\{|\mathbf{F}_{RF}|\}\}$ is the maximum of the maximum modulus of \mathbf{G}_{RF} or \mathbf{F}_{RF} . We assume the same LDR noise level for the users and FD BS, i.e., $k_0 = \beta_0 = \kappa_k = \beta_j$. The rate weights for the UL and DL users are set to be 1. Aforementioned simulation parameters are summarized in Table 2. The digital beamformers are initialized as the dominant eigenvectors of the channel covariance matrices of the intended users.

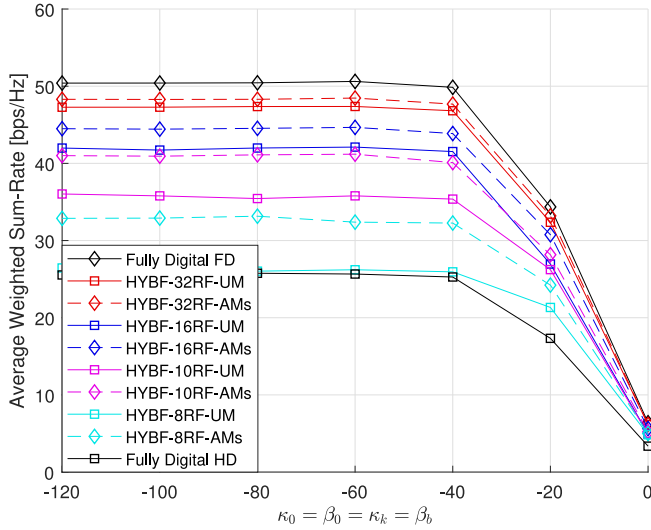


FIGURE 4. Average WSR as a function of the LDR noise with SNR = 0 dB.

Analog beamformer and combiner are initialized as the dominant eigenvectors of the sum of channel covariance matrices across all the UL and DL users, respectively. Note that as we assume perfect CSI, the SI can be canceled with HYBF only up to the LDR noise level, which represents the residual SI.

Figure 4 shows the achieved average WSR with the proposed HYBF designs as a function of the LDR noise with SNR = 0 dB. The fully digital FD scheme achieves an additional gain of $\sim 97\%$ over a fully digital HD scheme. The impact of different LDR noise levels on the maximum achievable WSR for a mmWave FD system with different number of RF chains is also shown. For $k_0 \leq -40$ dB, HYBF-UM and HYBF-AMs achieve an additional gain of $\sim 85\%$, 64% , 42% , 3% and $\sim 89\%$, 74% , 60% , 28% with 32, 16, 10, 8 RF chains, respectively. We can see that as the LDR noise variance increases, achievable WSR for both the hybrid FD and fully digital HD system degrades severely. Figure 5 shows the achieved average WSR as a function of the LDR noise with SNR = 40dB. For $k_0 \leq -80$ dB, HYBF-UM and HYBF-AMs achieve an additional gain of $\sim 65\%$, 55% , 41% , 15% and $\sim 67\%$, 62% , 55% , 26% with 32, 16, 10, 8 RF chains, respectively, and increasing the LDR noise variance degrades the achieved average WSR. By comparing Figure 4 with Figure 5, we can see that at low SNR, HYBF-UM with only 8 RF chains performs close to the fully digital HD scheme. As the SNR increases to 40 dB, HYBF-UM with 8 RF achieves an additional gain of $\sim 15\%$. HYBF-AMs with only 8 RF chains outperforms the fully digital HD scheme for all the SNR levels. Figures 4–5 also show that HYBF-AMs with 10 RF chains achieves similar average WSR as the HYBF-UM with 16 RF chains. It is interesting to observe that increasing the SNR from 0 dB to 40 dB decreases the thermal noise variance and the LDR noise variance dominates the noise floor already with $k_0 = -80$ dB at SNR= 40 dB. For SNR= 0 dB, the LDR noise variance dominates only for $k_0 > -40$ dB. From

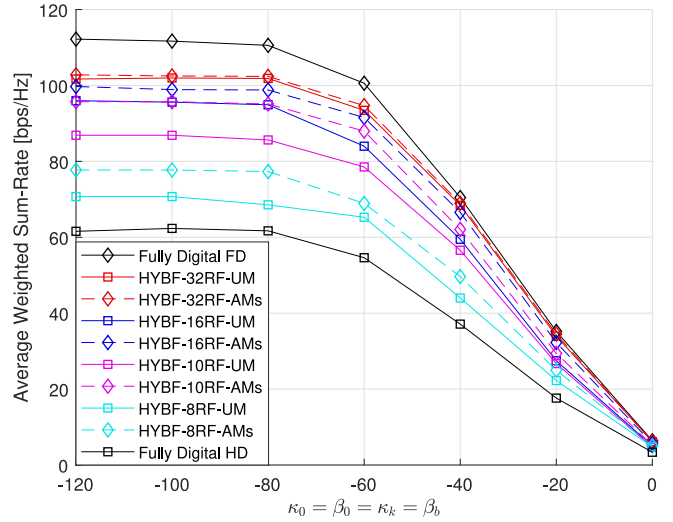


FIGURE 5. Average WSR as a function of the LDR noise with SNR = 40 dB.

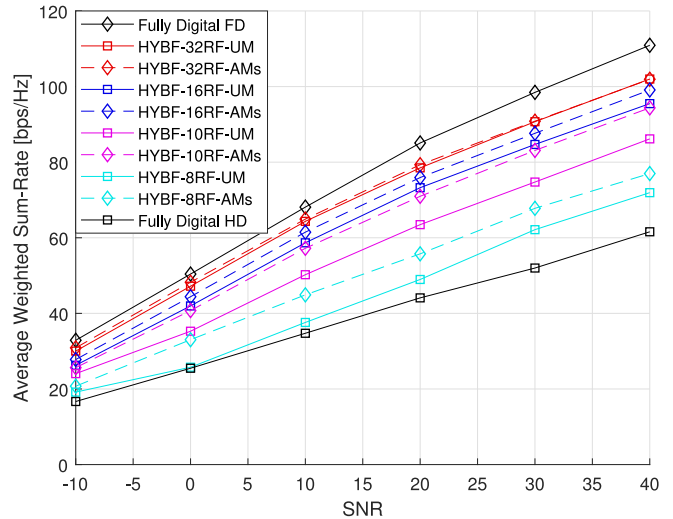


FIGURE 6. Average WSR as a function of the SNR with LDR noise $k_0 = -80$ dB.

this observation, we can conclude that hardware with a low LDR noise is required to benefit from a high SNR in the mmWave FD systems.

Figure 6 shows the average WSR with a low LDR noise level $k_0 = -80$ dB with 32, 16, 10 and 8 RF chains as a function of the SNR. Both the proposed designs perform very close to the fully digital FD scheme with 32 RF chains. HYBF-UM and HYBF-AMs outperform the fully digital HD scheme with only 8 RF chains at high SNR and at any SNR level, respectively. It is evident the advantage of AMs, which add additional gain for all the SNR levels when the number of RF chains at the FD BS is small. With a high number of RF chains, digital beamforming has enough amplitude manipulation liberty to manage the interference and adding AMs does not bring further improvement. Figure 7 shows the average WSR achieved with a moderate LDR noise level $k_0 = -60$ dB. We can see that for a low SNR, the achieved

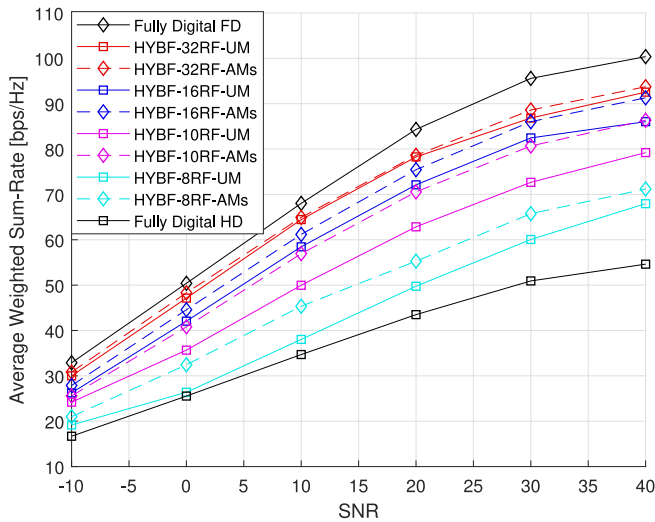


FIGURE 7. Average WSR as a function of the SNR with LDR noise $k_0 = -60$ dB.

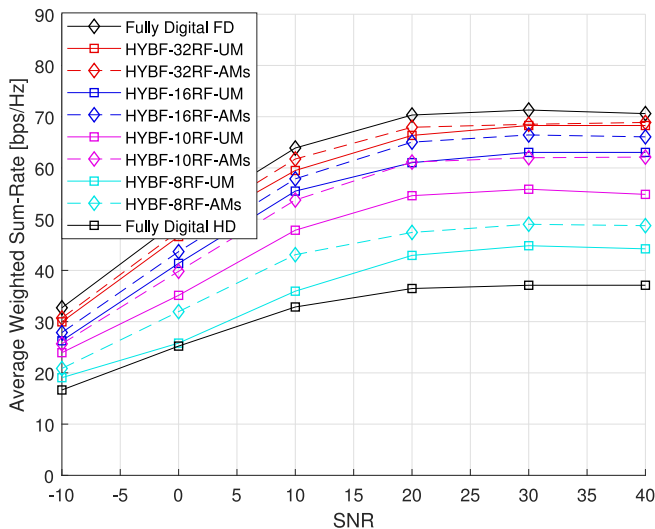


FIGURE 8. Average WSR as function of the SNR with LDR noise $k_0 = -40$ dB.

average WSR results to be similar as reported in Figure 6. At high SNR, the LDR noise variance starts dominating, which leads to less achieved average WSR compared to the case of Figure 6. Figure 8 shows the achieved WSR as a function of the SNR with a very large LDR noise variance of $k_0 = -40$ dB. By comparing the results reported in Figure 8 and Figures 6–7, we can see that the LDR noise variance dominates for most of the considered SNR range. For a very low SNR, the achieved WSR is similar as reported in Figures 6–7. However, as the SNR increases, it does not map into higher WSR. It is clear that the maximum achievable WSR with $k_0 = -40$ dB saturates already at SNR= 20 dB for both the HD and FD systems. Further improvement in the SNR does not dictate into higher WSR. When the LDR noise variance dominates, it acts as a ceiling to the effective received-signal-to-LDR-plus-thermal-noise-ratio (RSLTR). The transmit and receive LDR noise variance

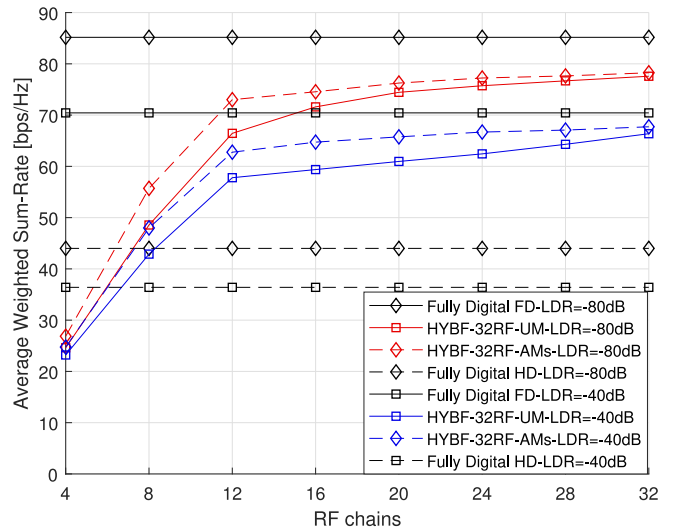


FIGURE 9. Average WSR as a function of the RF chains with LDR noise $k_0 = -80$ dB and $k_0 = -40$ dB at SNR = 20 dB.

is proportional to the total transmit power per-antenna and received power per RF chain after the analog combining, respectively. When the LDR noise variance is large, the thermal noise variance has a negligible effect on the effective RSLTR. Consequently, a decrease in the thermal noise variance (increasing SNR) does not dictate a better WSR.

Figure 9 shows the achievable performance of HYBF-UM and HYBF-AMs as a function of the RF chains with SNR= 20 dB, in comparison with the benchmark schemes, with very high and very small LDR noise levels. In particular, with very high LDR noise $k_k = -40$ dB and 8 RF chains, HYBF-UM and HYBF-AMs perform close to the fully HD system, and an increase in the number of RF chains improves the performance, which tends towards the achieved WSR by a fully digital FD system with LDR noise level $k_k = -40$ dB. Similar behavior can be observed for the case of low LDR noise $k_k = -80$ dB. Both the proposed schemes achieve higher WSR with the same number of RF chains in the latter case. We can also see that AMs add additional gain with a low number of RF chains, and as the number of RF chains increase, the gap in the achievable WSR with HYBF-AMs and HYBF-UM closes. In particular, with 32 RF chains, the difference in the WSR with or without AMs becomes negligible.

From the results reported in Figures 4–9, we can conclude that the proposed HYBF schemes achieve significant performance improvement, in terms of average WSR, compared to a fully digital HD system. LDR noise plays a key role in determining the maximum achievable WSR for both the FD and HD systems. Figures 4–5 shows how an increase in the LDR noise variance degrades the average WSR at low and high SNR levels. Figures 6–7 shows that with a large to moderate dynamic range, the LDR noise degrades the performance only at very high SNR. Figure 8 shows the achieved WSR as a function of a very large LDR noise variance. In that case, it is observed that the WSR saturates

at SNR= 20 dB and further improvement in the SNR does not dictate higher WSR. From Figure 9, it is clear how the number of RF chains at the mmWave FD BS affects the achievable WSR with different LDR noise levels and with or without the AMs.

VI. CONCLUSION

This paper has presented a novel HYBF design to maximize the WSR in a single-cell mmWave FD system with multi-antenna users and suffering from LDR noise. The beamformers were designed under the joint sum-power and the practical per-antenna power constraints. Simulation results showed that the MU mmWave FD systems can outperform the fully digital HD system with only a few RF chains. The advantage of having amplitude control at the analog processing stage is also investigated, and the benefit resulted to be evident with a small number of RF chains. Achievable average WSR with different levels of the LDR noise variance is also investigated, and the proposed HYBF designs outperformed the fully digital HD system at any LDR noise level.

APPENDIX A GRADIENT DERIVATION

The proof of Theorem 1 is based on the result derived in the following.

Lemma 3: Let $Y = AXB + a A \text{diag}(X + Q)B + b \text{diag}(CXD + E) + F$. The derivative of $\text{Indet}(Y)$ with respect to X is given by

$$\frac{\partial \text{Indet} Y}{\partial X} = A^H Y^{-H} B^H + a \text{diag}(A^H Y^{-H} B^H) + b C^H \text{diag}(Y^{-H}) D^H. \quad (40)$$

Proof: By substituting $\phi = \text{Indet}(Y)$, we can write

$$\partial \phi = Y^{-H} : dY = \text{Tr}(Y^{-1} dY), \quad (41)$$

where operator $:$ denotes the Frobenius inner product, i.e., $G_{RF} : H = \text{Tr}(G_{RF}^H H)$. Its derivative with respect to X can be written as

$$\frac{\partial \phi}{\partial X} = Y^{-H} : \left[\frac{d}{\partial X} (AXB + a A \text{diag}(X)B + b \text{diag}(CXD + E) + F) \right], \quad (42)$$

where the last term results to be zero as independent from X . Substituting the Frobenius product with the trace operator, using its cyclic shift and separating terms, yields

$$\begin{aligned} \frac{\partial \phi}{\partial X} &= \underbrace{\frac{\partial \text{Tr}(BY^{-1}AX)}{\partial X}}_I + a \underbrace{\frac{\partial \text{Tr}(BY^{-1}A \text{diag}(X))}{\partial X}}_II \\ &+ b \underbrace{\frac{\partial \text{Tr}(Y^{-1} \text{diag}(CXD))}{\partial X}}_III + b \frac{\partial \text{Tr}(Y^{-1} \text{diag}(E))}{\partial X}, \end{aligned} \quad (43)$$

where the last term being independent of X is also zero. To proof the aforementioned result, we proof the derivatives of I , II and III separately. Firstly, for I , by using $:$ and doing some simple algebraic manipulations leads to

$$\frac{\partial \text{Tr}(BY^{-1}AX)}{\partial X} = A^H Y^{-H} B^H : \partial X = A^H Y^{-H} B^H. \quad (44)$$

To obtain the derivative of II , we first define $\text{diag}(X) = Z$. The diagonal of X can be written as $\text{diag}(X) = I \circ X$ where \circ denotes the Hadamard product. By writing II with $:$ and expressing the diagonal term as a function of \circ , and using the commutative property of the Hadamard product leads to the following result

$$\begin{aligned} a \frac{\partial \text{Tr}(BY^{-1}AZ)}{\partial Z} &= a A^H Y^{-H} B^H : \partial Z, \\ &= a A^H Y^{-H} B^H : I \circ \partial X, \\ &= a A^H Y^{-H} B^H \circ I : \partial X, \\ &= a \text{diag}(A^H Y^{-H} B^H). \end{aligned} \quad (45)$$

To compute the derivative of III , we first define $\text{diag}(CXD) = W$. By using a similar approach as in (45), we get

$$\begin{aligned} b \frac{\partial \text{Tr}(Y^{-1}W)}{\partial W} &= b Y^{-H} : \partial W, \\ &= b Y^{-H} : I \circ C \partial X D, \\ &= b Y^{-H} \circ I : C \partial X D, \\ &= b \text{diag}(Y^{-H}) : C \partial X D, \\ &= b C^H \text{diag}(Y^{-1})^H D^H. \end{aligned} \quad (46)$$

Combining the result from each term concludes the proof for Lemma (40). \blacksquare

To prove Theorem 1, note that the covariance matrices in (11) has a special (Hermitian) structure, i.e., $B = A^H$ and $D = C^H$. Therefore, the result of Lemma 3 for this particular case is given in the following.

Lemma 4: Let $Y = AXB + a A \text{diag}(X + Q)B + b \text{diag}(CXD + E) + F$, where the size of matrices involved is such that the product is valid. Let $B = A^H$ and $D = C^H$ and the derivative of $\text{Indet}(Y)$ is given by

$$\begin{aligned} \frac{\partial \text{Indet} Y}{\partial X} &= A^H Y^{-H} A + a \text{diag}(A^H Y^{-H} A) \\ &+ b C^H \text{diag}(Y^{-H}) C. \end{aligned} \quad (47)$$

Proof: The result follows directly by relying on the result given in Lemma 3 by substituting $B = A^H$ and $D = C^H$. \blacksquare

Proof (Theorem 1): To prove the gradients to linearize the WSR with respect to T_k and Q_j , we proceed by simplifying

the WSR as

$$\text{WSR} = \sum_{k \in \mathcal{U}} \left[w_k \text{Indet}(\mathbf{R}_k) - w_k \text{Indet}(\mathbf{R}_{\bar{k}}) \right] + \sum_{j \in \mathcal{D}} \left[w_j \text{Indet}(\mathbf{R}_j) - w_j \text{Indet}(\mathbf{R}_{\bar{j}}) \right]. \quad (48)$$

The WSR_k^{UL} and WSR^{DL} should be linearized for \mathbf{T}_k and WSR_j^{DL} and WSR^{UL} for \mathbf{Q}_j . Note from (11) that \mathbf{T}_k appears in $\text{WSR}_{\bar{k}}^{UL}$ and WSR^{DL} with the structure $\mathbf{Y} = \mathbf{A}\mathbf{X}\mathbf{A}^H + a\mathbf{A} \text{diag}(\mathbf{X} + \mathbf{Q})\mathbf{A}^H + b \text{diag}(\mathbf{C}\mathbf{X}\mathbf{C}^H + \mathbf{E}) + \mathbf{F}$, where the scalars a and b are due to the LDR noise model, \mathbf{A} and \mathbf{C} are the interfering channels, \mathbf{F} and \mathbf{E} contain the noise contributions from other transmit covariance matrices but independent from \mathbf{T}_k . The same structure holds also for the DL covariance matrices $\mathbf{Q}_j, \forall j \in \mathcal{D}$. By applying the result from Lemma 4 with $\mathbf{Y} = \mathbf{R}_k$ or $\mathbf{Y} = \mathbf{R}_{\bar{k}}$ repetitively $K-1$ time for linearizing $\text{WSR}_{\bar{k}}$ with respect to \mathbf{T}_k yield the gradient \mathbf{A}_k . Similarly, by considering $\mathbf{Y} = \mathbf{R}_j$ or $\mathbf{Y} = \mathbf{R}_{\bar{j}}, \forall j \in \mathcal{D}$ and applying the result from Lemma 4 yield the gradient \mathbf{B}_k .

The same reasoning holds also for \mathbf{Q}_j , which leads to the gradients $\hat{\mathbf{C}}_j$ and $\hat{\mathbf{D}}_j$ by applying the result provided in Lemma 4 for WSR_j^{DL} $J-1$ times and for WSR^{UL} K times, respectively, $\forall j \in \mathcal{D}$.

APPENDIX B PROOF OF THEOREM 3

The dominant generalized eigenvector solution maximizes the reformulated concave WSR maximization problem

$$\text{WSR} = \sum_{k \in \mathcal{U}} \left[w_k \text{Indet}(\mathbf{I} + \mathbf{U}_k^H \mathbf{H}_k^H \mathbf{F}_{RF} \mathbf{R}_{\bar{k}}^{-1} \mathbf{F}_{RF}^H \mathbf{H}_k \mathbf{U}_k) - \text{Tr}(\mathbf{U}_k^H (\hat{\mathbf{A}}_k + \hat{\mathbf{B}}_k + l_k \mathbf{I} + \mathbf{\Psi}_k) \mathbf{U}_k) \right] + \sum_{j \in \mathcal{D}} \left[w_j \text{Indet}(\mathbf{I} + \mathbf{V}_j^H \mathbf{G}_{RF}^H \mathbf{H}_j^H \mathbf{R}_{\bar{j}}^{-1} \mathbf{H}_j \mathbf{G}_{RF} \mathbf{V}_j) - \text{Tr}(\mathbf{V}_j^H \mathbf{G}_{RF}^H (\hat{\mathbf{C}}_j + \hat{\mathbf{D}}_j + l_0 \mathbf{I} + \mathbf{\Psi}_0) \mathbf{G}_{RF} \mathbf{V}_j) \right]. \quad (49)$$

To prove Theorem 3 for solving (49), we first consider the UL digital beamforming solution by keeping the analog beamformer and the digital DL beamformers fixed. We proceed by considering user $k \in \mathcal{U}$ for which we wish to compute the WSR maximizing digital UL beamformer. The same proof will be valid $\forall k \in \mathcal{U}$. The proof relies on simplifying

$$\max_{\mathbf{U}_k} \left[w_k \text{Indet}(\mathbf{I} + \mathbf{U}_k^H \mathbf{H}_k^H \mathbf{F}_{RF} \mathbf{R}_{\bar{k}}^{-1} \mathbf{F}_{RF}^H \mathbf{H}_k \mathbf{U}_k) - \text{Tr}(\mathbf{U}_k^H (\hat{\mathbf{A}}_k + \hat{\mathbf{B}}_k + l_k \mathbf{I} + \mathbf{\Psi}_k) \mathbf{U}_k) \right] \quad (50)$$

until the Hadamard's inequality applies as in [56, Proposition 1] or [58, Th. 1]. The Cholesky

decomposition of the matrix $(\hat{\mathbf{A}}_k + \hat{\mathbf{B}}_k + l_k + \mathbf{\Psi}_k)$ is given as $\mathbf{L}_k \mathbf{L}_k^H$ where \mathbf{L}_k is the lower triangular Cholesky factor. By defining $\tilde{\mathbf{U}}_k = \mathbf{L}_k^H \mathbf{U}_k$, (50) reduces to

$$\max_{\mathbf{U}_k} \left[w_k \text{Indet}(\mathbf{I} + \tilde{\mathbf{U}}_k^H \mathbf{L}_k^{-1} \mathbf{H}_k^H \mathbf{F}_{RF} \mathbf{R}_{\bar{k}}^{-1} \mathbf{F}_{RF}^H \mathbf{H}_k \mathbf{L}_k^{-H} \tilde{\mathbf{U}}_k) - \text{Tr}(\tilde{\mathbf{U}}_k^H \tilde{\mathbf{U}}_k) \right]. \quad (51)$$

Let $\mathbf{E}_k \mathbf{D}_k \mathbf{E}_k^H$ be the eigen-decomposition of $\mathbf{L}_k^{-1} \mathbf{H}_k^H \mathbf{R}_{\bar{k}}^{-1} \mathbf{H}_k \mathbf{L}_k^{-H}$, where \mathbf{E}_k and \mathbf{D}_k are the unitary and diagonal matrices, respectively. Let $\mathbf{O}_k = \mathbf{E}_k^H \tilde{\mathbf{U}}_k \tilde{\mathbf{U}}_k^H \mathbf{E}_k$ and (51) can be expressed as

$$\max_{\mathbf{O}_k} [w_k \text{Indet}(\mathbf{I} + \mathbf{O}_k \mathbf{D}_k) - \text{Tr}(\mathbf{O}_k)]. \quad (52)$$

By Hadamard's inequality [59, p. 233], it can be seen that the optimal \mathbf{O}_k must be diagonal. Therefore, $\mathbf{U}_k = \mathbf{L}_k^{-H} \mathbf{E}_k \mathbf{O}_k^{\frac{1}{2}}$ and thereby

$$\mathbf{H}_k^H \mathbf{F}_{RF} \mathbf{R}_{\bar{k}}^{-1} \mathbf{F}_{RF}^H \mathbf{H}_k \mathbf{U}_k = \mathbf{L}_k \mathbf{L}_k^H \mathbf{L}_k^{-H} \mathbf{E}_k \mathbf{O}_k^{\frac{1}{2}} \mathbf{D}_k = (\hat{\mathbf{A}}_k + \hat{\mathbf{B}}_k + l_k + \mathbf{\Psi}_k) \mathbf{U}_k \mathbf{D}_k, \quad (53)$$

from which we select u_k dominant eigenvectors, which concludes the proof for the UL beamformer for user $k \in \mathcal{U}$. For the digital DL beamformers the proof follow similarly by considering the following optimization problem $\forall j$

$$\max_{\mathbf{V}_j} \left[w_j \text{Indet}(\mathbf{I} + \mathbf{V}_j^H \mathbf{G}_{RF}^H \mathbf{H}_j^H \mathbf{R}_{\bar{j}}^{-1} \mathbf{H}_j \mathbf{G}_{RF} \mathbf{V}_j) - \text{Tr}(\mathbf{V}_j^H \mathbf{G}_{RF}^H (\hat{\mathbf{C}}_j + \hat{\mathbf{D}}_j + l_0 + \mathbf{\Psi}_0) \mathbf{G}_{RF} \mathbf{V}_j) \right]. \quad (54)$$

and simplifying it until the Hadamard's inequality applies to yield a similar result as expressed in (53).

The proof for analog beamformer \mathbf{G}_{RF} does not apply directly as the KKT condition have the form $\mathbf{A}_1 \mathbf{G}_{RF} \mathbf{A}_2 = \mathbf{B}_1 \mathbf{G}_{RF} \mathbf{B}_2$, which are not resolvable. To solve it for the analog beamformer \mathbf{G}_{RF} , we apply the result $\text{vec}(\mathbf{A}\mathbf{X}\mathbf{B}) = \mathbf{B}^T \otimes \text{Avec}(\mathbf{X})$ [60], which allows to rewrite (22) as

$$\sum_{j \in \mathcal{D}} w_j \left(\left(\mathbf{V}_j \mathbf{V}_j^H (\mathbf{I} + \mathbf{V}_j \mathbf{V}_j^H \mathbf{G}_{RF}^H \mathbf{H}_j^H \mathbf{R}_{\bar{j}}^{-1} \mathbf{H}_j \mathbf{G}_{RF})^{-1} \right)^T \otimes \mathbf{H}_j^H \mathbf{R}_{\bar{j}}^{-1} \mathbf{H}_j \right) \text{vec}(\mathbf{G}_{RF}) - \sum_{j \in \mathcal{D}} \left((\mathbf{V}_j \mathbf{V}_j^H)^T \otimes (\hat{\mathbf{C}}_j + \hat{\mathbf{D}}_j + \mathbf{\Psi}_0 + l_0 \mathbf{I}) \right) \text{vec}(\mathbf{G}_{RF}) = 0. \quad (55)$$

The WSR maximizing analog beamformer can alternatively be derived as follows (which allows the proof for the digital beamformers to be applicable directly). First we apply a noise whitening procedure using the noise plus interference covariance matrix $\mathbf{R}_j^{1/2}$ on the received signal. Further, we can rewrite the whitened signal as follows

$$\tilde{\mathbf{y}}_j = \left((\mathbf{s}_{jd}^T \mathbf{V}_j^T) \otimes \mathbf{R}_j^{-1/2} \mathbf{H}_j \right) \text{vec}(\mathbf{G}_{RF}) + \tilde{\mathbf{n}}_j, \quad (56)$$

where $\tilde{\mathbf{y}}_j = \mathbf{R}_j^{-1/2} \mathbf{y}_j$ and $\tilde{\mathbf{n}}_j$ represents the whitened noise plus interference signal. We can write the resulting WSR optimization problem, after the approximation to concave form and some algebraic manipulations on the linearized term, as

$$\max_{\mathbf{G}_{RF}} \sum_{j \in \mathcal{D}} \left[w_j \text{Indet} \left(\mathbf{I} + \text{vec}(\mathbf{G}_{RF})^H \left((\mathbf{V}_j \mathbf{V}_j^H)^T \otimes \mathbf{H}_j^H \mathbf{R}_j^{-1} \mathbf{H}_j \right) \text{vec}(\mathbf{G}_{RF}) \right) - \text{Tr} \left(\text{vec}(\mathbf{G}_{RF})^H \left(\mathbf{V}_j \mathbf{V}_j^H \otimes (\hat{\mathbf{C}}_j + \hat{\mathbf{D}}_j) + \boldsymbol{\Psi}_0 + l_0 \mathbf{I} \right) \text{vec}(\mathbf{G}_{RF}) \right) \right]. \quad (57)$$

Taking the derivative of (57) for the conjugate of \mathbf{G}_{RF} leads to the same generalized eigenvector solution as in (23). Note that this alternative representation has the same form as (50), which is resolvable for the vectorized version of the analog beamformer \mathbf{G}_{RF} . Therefore, the proof for the UL and DL digital beamformers can now be applied directly on the vectorized analog beamformer $\text{vec}(\mathbf{G}_{RF})$, which is summed over all the DL users served by the mmWave FD BS.

ACKNOWLEDGMENT

EURECOMs research is partially supported by its industrial members: ORANGE, BMW, Symantec, SAP, Monaco Telecom, iABG, and by the projects MASS-START (French FUI), DUPLEX (French ANR), SPOTLIGHT (EU ITN) and the Qualcomm Fab5G project.

REFERENCES

- [1] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.
- [2] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.
- [3] S. Liu, L. Fu, and W. Xie, "Hidden-node problem in full-duplex enabled CSMA networks," *IEEE Trans. Mobile Comput.*, vol. 19, no. 2, pp. 347–361, Feb. 2020.
- [4] H. Alves, T. Riihonen, and H. A. Suraweera, *Full-Duplex Communications for Future Wireless Networks*. Springer, 2020.
- [5] M. T. Kabir and C. Masouros, "A scalable energy vs. latency trade-off in full-duplex mobile edge computing systems," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5848–5861, Aug. 2019.
- [6] C. B. Barneto, S. D. Liyanarachchi, M. Heino, T. Riihonen, and M. Valkama, "Full duplex radio/radar technology: The enabler for advanced joint communication and sensing," *IEEE Wireless Commun.*, vol. 28, no. 1, pp. 82–88, Feb. 2021.
- [7] M. Gan *et al.*, "Full duplex for next generation of 802.11," in *Proc. IEEE 30th Int. Symp. Personal Indoor Mobile Radio Commun. (PIMRC Workshops)*, Sep. 2019, pp. 1–6.
- [8] P. Rosson, C. K. Sheemar, N. Valecha, and D. Slock, "Towards massive MIMO in-band full duplex radio," in *Proc. IEEE 16th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2019, pp. 69–74.
- [9] S. Huberman and T. Le-Ngoc, "MIMO full-duplex precoding: A joint beamforming and self-interference cancellation structure," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2205–2217, Apr. 2015.
- [10] P. Aquilina, A. C. Cirik, and T. Ratnarajah, "Weighted sum rate maximization in full-duplex multi-user multi-cell MIMO networks," *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1590–1608, Apr. 2017.
- [11] T. Riihonen and R. Wichman, "Analog and digital self-interference cancellation in full-duplex MIMO-OFDM transceivers with limited resolution in A/D conversion," in *Proc. IEEE 46th Asilomar Conf. Signals Syst. Comput. (ASILOMAR)*, Nov. 2012, pp. 45–49.
- [12] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3702–3713, Jul. 2012.
- [13] O. Taghizadeh, J. Zhang, and M. Haardt, "Transmit beamforming aided amplify-and-forward MIMO full-duplex relaying with limited dynamic range," *Signal Process.*, vol. 127, pp. 266–281, Oct. 2016.
- [14] A. C. Cirik, S. Biswas, S. Vuppala, and T. Ratnarajah, "Beamforming design for full-duplex MIMO interference channels-QoS and energy-efficiency considerations," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4635–4651, Nov. 2016.
- [15] E. Antonio-Rodríguez, R. López-Valcarce, T. Riihonen, S. Werner, and R. Wichman, "SINR optimization in wideband full-duplex MIMO relays under limited dynamic range," in *Proc. IEEE 8th Sens. Array Multichannel Signal Process. Workshop (SAM)*, Jun. 2014, pp. 177–180.
- [16] S. Biswas, K. Singh, O. Taghizadeh, and T. Ratnarajah, "Design and analysis of FD MIMO cellular systems in coexistence with MIMO radar," *IEEE Trans. Wireless Commun.*, vol. 19, no. 7, pp. 4727–4743, Jul. 2020.
- [17] A. C. Cirik, O. Taghizadeh, L. Lampe, R. Mathar, and Y. Hua, "Linear transceiver design for full-duplex multi-cell MIMO systems," *IEEE Access*, vol. 4, pp. 4678–4689, 2016.
- [18] O. Taghizadeh, V. Radhakrishnan, A. C. Cirik, R. Mathar, and L. Lampe, "Hardware impairments aware transceiver design for bidirectional full-duplex MIMO OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7450–7464, Aug. 2018.
- [19] T. Schenk, *RF Imperfections in High-Rate Wireless Systems: Impact and Digital Compensation*. Springer, Jan. 2008.
- [20] S. R. Aghdam, S. Jacobsson, and T. Eriksson, "Distortion-aware linear precoding for millimeter-wave multiuser MISO downlink," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2019, pp. 1–6.
- [21] R. López-Valcarce and N. González-Prelcic, "Analog beamforming for full-duplex millimeter wave communication," in *Proc. IEEE 16th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2019, pp. 687–691.
- [22] C. K. Sheemar and D. T. Slock, "Hybrid beamforming for bidirectional massive MIMO full duplex under practical considerations," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC) Spring*, Apr. 2021, pp. 1–5.
- [23] K. Satyanarayana, M. El-Hajjar, P.-H. Kuo, A. Mourad, and L. Hanzo, "Hybrid beamforming design for full-duplex millimeter wave communication," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1394–1404, Feb. 2019.
- [24] C. K. Thomas, C. K. Sheemar, and D. Slock, "Multi-stage/hybrid BF under limited dynamic range for OFDM FD backhaul with MIMO SI nulling," in *Proc. IEEE 16th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2019, pp. 96–101.
- [25] E. Balti, N. Mensi, and S. Yan, "A modified zero-forcing max-power design for hybrid beamforming full-duplex systems," 2020, *arXiv:2003.00147*.
- [26] S. Huang, Y. Ye, and M. Xiao, "Learning based hybrid beamforming design for full-duplex millimeter wave systems," *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 1, pp. 120–132, Mar. 2021.
- [27] J. Palacios, J. Rodriguez-Fernandez, and N. González-Prelcic, "Hybrid precoding and combining for full-duplex millimeter wave communication," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.
- [28] H. Abbas and K. Hamdi, "Full duplex relay in millimeter wave backhaul links," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2016, pp. 1–6.
- [29] Y. Cai, K. Xu, A. Liu, M. Zhao, B. Champagne, and L. Hanzo, "Two-timescale hybrid analog-digital beamforming for mmwave full-duplex MIMO multiple-relay aided systems," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 2086–2103, Sep. 2020.
- [30] S. Han, Y. Zhang, W. Meng, C. Li, and Z. Zhang, "Full-duplex relay-assisted macrocell with millimeter wave backhauls: Framework and prospects," *IEEE Netw.*, vol. 33, no. 5, pp. 190–197, Sep./Oct. 2019.
- [31] C. K. Sheemar and D. Slock, "Massive MIMO mmWave full duplex relay for IAB with limited dynamic range," in *Proc. IEEE 11th IFIP Int. Conf. New Technol. Mobility Security (NTMS)*, Apr. 2021, pp. 1–5.
- [32] J. M. B. da Silva, A. Sabharwal, G. Fodor, and C. Fischione, "1-bit phase shifters for large-antenna full-duplex mmWave communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6916–6931, Oct. 2020.

- [33] M.-M. Zhao, Y. Cai, M.-J. Zhao, Y. Xu, and L. Hanzo, "Robust joint hybrid analog-digital transceiver design for full-duplex mmWave multicell systems," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4788–4802, Aug. 2020.
- [34] I. P. Roberts, J. G. Andrews, and S. Vishwanath, "Hybrid beamforming for millimeter wave full-duplex under limited receive dynamic range," 2020, *arXiv:2012.11647*.
- [35] I. P. Roberts and S. Vishwanath, "Beamforming cancellation design for millimeter-wave full-duplex," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.
- [36] I. P. Roberts, H. B. Jain, and S. Vishwanath, "Equipping millimeter-wave full-duplex with analog self-interference cancellation," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2020, pp. 1–6.
- [37] C. K. Sheemar and D. Slock, "Hybrid beamforming and combining for millimeter wave full duplex massive MIMO interference channel," 2021, *arXiv:2108.00465*.
- [38] I. P. Roberts, H. B. Jain, and S. Vishwanath, "Frequency-selective beamforming cancellation design for millimeter-wave full-duplex," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.
- [39] D. Korpi, T. Riihonen, V. Syrjälä, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014.
- [40] C. K. Sheemar, L. Badia, and S. Tomasin, "Game-theoretic mode scheduling for dynamic TDD in 5G systems," *IEEE Commun. Lett.*, vol. 25, no. 7, pp. 2425–2429, Jul. 2021.
- [41] J. M. B. da Silva, G. Wikström, R. K. Mungara, and C. Fischione, "Full duplex and dynamic TDD: Pushing the limits of spectrum reuse in multi-cell communications," *IEEE Wireless Commun.*, vol. 28, no. 1, pp. 44–50, Feb. 2021.
- [42] H. Kim, J. Kim, and D. Hong, "Dynamic TDD systems for 5G and beyond: A survey of cross-link interference mitigation," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2315–2348, 4th Quart., 2020.
- [43] E. de Olivindo Cavalcante, G. Fodor, Y. C. Silva, and W. C. Freitas, "Bidirectional sum-power minimization beamforming in dynamic TDD MIMO networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 9988–10002, Oct. 2019.
- [44] S. Guo, X. Hou, and H. Wang, "Dynamic TDD and interference management towards 5G," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–6.
- [45] P. Stoica and Y. Selen, "Cyclic minimizers, majorization techniques, and the expectation-maximization algorithm: A refresher," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 112–114, Jan. 2004.
- [46] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [47] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.
- [48] T. Lan and W. Yu, "Input optimization for multi-antenna broadcast channels with per-antenna power constraints," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2004, pp. 420–424.
- [49] R. Chaluvari, S. S. Nair, and S. Bhashyam, "Optimal multi-antenna transmission with multiple power constraints," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3382–3394, Jul. 2019.
- [50] C. K. Sheemar and D. T. Slock, "Beamforming for bidirectional MIMO full duplex under the joint sum power and per antenna power constraints," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 4800–4804.
- [51] M. R. Castellanos *et al.*, "Hybrid multi-user precoding with amplitude and phase control," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.
- [52] M. Majidzadeh, J. Kaleva, N. Tervo, H. Pennanen, A. Tölli, and M. Latva-Aho, "Rate maximization for partially connected hybrid beamforming in single-user MIMO systems," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.
- [53] L. Chen, A. Liu, and X. Yuan, "Structured turbo compressed sensing for massive MIMO channel estimation using a Markov prior," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4635–4639, May 2018.
- [54] I. P. Roberts, J. G. Andrews, H. B. Jain, and S. Vishwanath, "Millimeter-wave full duplex radios: New challenges and techniques," *IEEE Wireless Commun.*, vol. 28, no. 1, pp. 36–43, Feb. 2021.
- [55] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted mmse for mimo-bc beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.
- [56] S.-J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3117–3131, May 2011.
- [57] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [58] D. Hoang and R. A. Iltis, "Noncooperative eigencoding for MIMO ad-hoc networks," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 865–869, Feb. 2008.
- [59] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 1991.
- [60] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus With Applications in Statistics and Econometrics*. Hoboken, NJ, USA: Wiley, 2019.