IEEE Open Journal of the Communications Society

Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version XX Month, XXXX.

Digital Object Identifier 10.1109/OJCOMS.2022.1234567

Joint Beamforming and Combining Design for mmWave Integrated Access and Backhaul Networks

Zou Linfu[1],**Pan Zhiwen**[1],**Alaa Alameer Ahmad**[2], **Hayssam Dahrouj**[3], ***Senior Member, IEEE***, **Mohammed El-Hajjar**[4],***Senior Member, IEEE***

[1]Zou Linfu and Pan Zhiwen are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China and also with the Purple Mountain Laboratories, Nanjing 211111, China.
[2]Alaa Alameer Ahmad is with Cariad SE, Wolfsburg, Germany.
[3]Hayssam Dahrouj is with the Department of Electrical Engineering, University of Sharjah, Sharjah, UAE.
[4]Mohammed El-Hajjar is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO171BJ, U.K.

CORRESPONDING AUTHOR: Zou Linfu (e-mail: zoulinfu@seu.edu.cn).

This work is partially supported by the National Key Research and Development Project under Grant 2020YFB1806805, the Fundamental Research Funds for the Central Universities, 2242022K60001 and the China Scholarship Council (CSC, 202106090144).

ABSTRACT Integrated access and backhaul (IAB) networks operating in full-duplex (FD) mode at millimeter wave frequencies have been actively investigated in the context of future-generation communications networks. However, conventional analog cancellation techniques cannot adequately mitigate the self-interference resulting from the FD operation and the multi-user interference. Hence, in this paper, we consider a multi-cell, multi-user IAB network and jointly design the beamforming and combining matrices to maximize the networks weighted sum rate. Given the non-convex nature of the problem, we reformulate it using weighted minimum-mean-square-error (WMMSE) and extended fractional programming (FP) techniques followed by a block coordinate descent (BCD) approach. Extensive simulation results validate the superior performance of our proposed algorithms. Specifically, the WMMSE and FP methods can achieve 50 bits/sec/Hz higher than the benchmark scheme for a network employing three cells with two uplink and two downlink users per cell.

INDEX TERMS beamforming, full-duplex, multi-user, MIMO, integrated access and backhaul, millimeter wave.

I. Introduction

KEY technologies such as millimeter-wave (mmWave) communications, full-duplex (FD) transmissions, and integrated access and backhaul (IAB) networks have emerged as fundamental components of the next-generation wireless networks [1]. These technologies are essential to accommodate the exponential traffic growth and provide ubiquitous connectivity [2]. However, massive multiple-input and multiple-output (MIMO) systems are needed to enhance the coverage and compensate for the higher path loss experienced at mmWave [3]. Moreover, it is economically and physically impractical to establish traditional fiber backhaul connections for all small mmWave cells [1]. Hence, to address this challenge, the concept of IAB has been proposed as an effective alternative to fiber backhauling [1], where wireless backhauling is utilized as a benefit of the flexible and wide bandwidth available at mmWave. Along such lines, this paper addresses the problem of maximizing the sum-rate of a multi-cell multi-user IAB network by means of determining the beamforming and combining strategies of each link.

Consider a multi-cell multi-user network, where each multi-antenna IAB node aims at serving several users and communicating with adjacent IAB nodes simultaneously. The IAB nodes operating in the mmWave wideband operate in FD mode, which offers increased spectral efficiency (SE) and reduced communication delay compared with half-duplex (HD) transmission [4]. However, the simultaneous access and backhaul communications in an FD-IAB network give rise to self-interference (SI) at the receiver of the IAB node and the multi-user interference, when considering multi-cell multi-user networks. In FD-IAB networks, con-

ventional analog cancellation techniques are insufficient to mitigate the SI, as the SI power can exceed the signal of interest by more than 100 dB [5], [6]. This high SI power can surpass the hardware dynamic range and compromise the benefits of FD transmission. Thus, effective SI cancellation (SIC) techniques are required in antenna and digital domains, involving antenna isolation and digital beamforming.

In a typical IAB framework, many previous studies incorporate HD constraints for implementation simplicity [7], [8], where the IAB links must use the given radio resources orthogonally. However, a mmWave system with directional transmissions makes it possible to fully exploit the potential of the allocated radio resources by operating in FD mode [6], [9]. For example, [6], [10], [11] proposed various hybrid beamforming methods to fulfill the mmWave FD transmissions and further reduce the hardware complexity. Although these beamforming schemes were able to achieve array gain and mitigate SI in FD-aided mmWave communication, the backhaul link is not considered. Additionally, [12] proposed a distributed Stackelberg game-theoretic approach to address energy efficiency optimization under a RIS-aided and UAV-assisted IAB network.

To further enable the combination of IAB and mmWave FD communications, [13]–[16] presented beamforming and SI cancellation algorithms considering a single backhaul link. However, these studies primarily focused on scenarios with limited investigations into the multi-user multi-cell networks. Since different cells reuse the same frequency resources in multi-cell systems to mitigate spectrum scarcity, this leads to significant inter-cell interference [17], particularly in densely deployed IAB node configurations. On the other hand, the optimization problem in the IAB mmWave FD scenario is challenging to solve in closed form.

A novel fractional programming (FP) technique, which is widely used in power control [18] and beamforming [19], can tackle multiple ratio terms by transforming the original non-convex problem into a set of convex problems. However, joint precoding and combining design schemes utilizing FP sparsely exist in the multi-cell and multi-user IAB networks operating in mmWave. Consequently, in this paper, we propose to optimize the precoder and combiner design that maximizes the weighted sum rate (WSR) by extending the FP algorithm.

The contributions of this paper can be summarized as follows.

- We derive expressions for the received signals and relative sum rates for multi-cell multi-user IAB networks operating in mmWaves with base stations (BS) operating in FD mode, while considering MIMO systems at the BS and the user equipment.
- We formulate the joint precoding and combining matrices design to maximimze the weighted sum rate, whose complexity arises from the non-convex nature and strong inter-dependency among these matrices. Hence, the optimization problem for the joint multi-cell multi-user precoding and combining matrices design for the FD-IAB network is challenging due to the non-convexity and strong coupling between these matrices.
- We propose reformulating the original problem into an equivalent one to address these challenges using the weighted minimum-mean-square-error (WMMSE) and fractional programming (FP) methods tailored for IAB scenarios. Subsequently, a block coordinate descent (BCD) algorithm is developed to optimize the two matrices iteratively.
- We compare the performance of the FP and WMMSE algorithms against alternative methods such as separate-minimization self-interference (SMSI) [6], [13], the singular value decomposition (SVD), and the HD schemes. Our results demonstrate that the FP algorithm consistently outperforms the benchmark schemes across all simulated scenarios. Furthermore, the analysis considering imperfect CSI validates the robustness of the proposed algorithms in more practical scenarios.

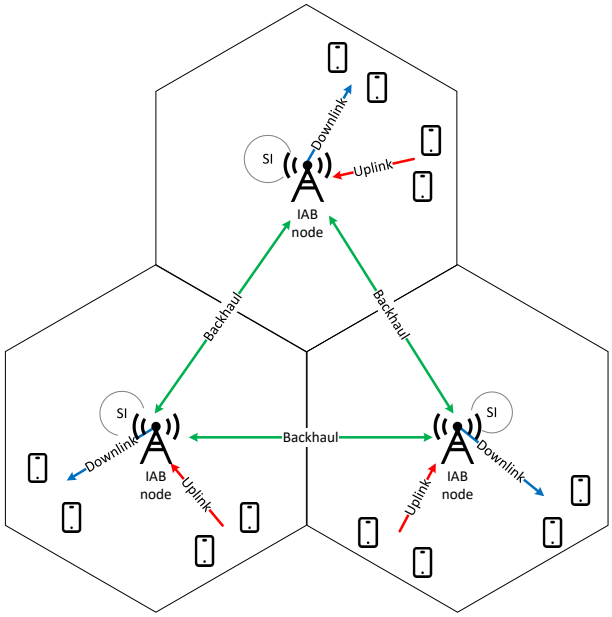

FIGURE 1. A multi-cell multi-user mmWave massive MIMO FD-IAB network architecture.

The rest of the paper is organized as follows. Section II describes the system model and objective function followed by Section III, where we present the design of beamforming and combining matrices. Section IV provides the numerical results obtained from simulations and finally, we provide our conclusions in Section V.

Notation: Bold lowercase $\mathbf{x}$ denotes column vectors, bold uppercase $\mathbf{X}$ denotes matrix, non-bold letters x, X denote scalar values. The Frobenius norm is denoted by $\|\mathbf{X}\|_F$, $\det(\mathbf{X})$ and $|\mathbf{X}|$ denote the determinant, $Tr(X)$ denotes the trace, $\mathbf{X}^H$ is the conjugate transpose, $\mathbf{X}^{-1}$ denotes the inverse of a square non-singular matrix. $\mathbb{C}^{M\times N}$ denotes the

set of $M \times N$ complex matrices. $\mathcal{CN}(0, \sigma^2)$ represents a random vector following the distribution of zero mean and σ^2 variance matrix.

II. System Model and Problem Formulation

We consider a multi-cell multi-user network, where mmWaves are used for IAB nodes, as shown in Fig. 1. IAB nodes operate in FD mode, while users are in HD mode. The network consists of L cells, each containing an IAB BS node serving K uplink and downlink users simultaneously. It is worth noting that the IAB nodes have the capability to communicate with each other through backhaul mmWave links [4]. Each IAB node is equipped with $M \gg 1$ transmit and receive antennas and each user equipment (UE) is equipped with $N \geq 1$ transmit or receive antennas, respectively. The stream of each link is set to be $N_s \leq N$. We aim to maximize the WSR by jointly optimizing the precoding and combining matrices of each link in the network.

A. Transmission Model

The received signal for the k-th downlink user from the i-th BS is expressed as

$$\begin{aligned}\mathbf{y}_{ik} =& \underbrace{\mathbf{W}_{ik}\mathbf{H}_{ii_k}\mathbf{F}_{ik}\mathbf{s}_{ik}}_{\text{downlink signal}} + \underbrace{\sum_{(j,m)\neq(i,k)}^{L}\sum^{K}\mathbf{W}_{ik}\mathbf{H}_{ji_k}\mathbf{F}_{jm}\mathbf{s}_{jm}}_{\text{downlink interference}} \\ &+ \underbrace{\sum_{j}^{L}\sum_{n}^{K}\mathbf{W}_{ik}\mathbf{H}_{j_n i_k}\mathbf{F}_{jn}\mathbf{s}_{jn}}_{\text{uplink interference}} + \mathbf{W}_{ik}\mathbf{n}_{ik} \\ &+ \underbrace{\sum_{j}^{L}\sum_{j'\neq j}^{L}\mathbf{W}_{ik}\mathbf{H}_{ji_k}\mathbf{F}_{jj'}\mathbf{s}_{jj'}}_{\text{backhaul interference}},\end{aligned} \tag{1}$$

where the downlink signal represents the desired signal from the corresponding IAB node, while the other three parts of (1) represent the interference from downlink, uplink, and backhaul links, respectively. Additionally, the different terms of (1) and subsequent equations denote the following:

- i, i', j, j' are used to denote the index of a cell.
- k, m are used to denote the index of a downlink user in a cell.
- l, n are used to denote the index of uplink user in a cell.
- $\mathbf{s}_{ik}$ is the $(N_s \times 1)$-element symbol vector transmitted to the kth user in the i-th cell satisfying $E[\mathbf{s}_{ik}\mathbf{s}_{ik}^H] = \mathbf{I}_d$.
- $\mathbf{F}_{ik}$ is the $\mathbb{C}^{M\times N_s}$ precoding matrix used by the i-th IAB node for transmitting its data vector $\mathbf{s}_{ik}$ to the k-th user.
- $\mathbf{W}_{ik}$ is the $\mathbb{C}^{N_s\times N}$ combining matrix for the k-th user in the i-th cell.
- $\mathbf{n}_{ik}$ is the noise vector whose elements are independent and identically distributed (i.i.d.) with distribution $\mathcal{CN}(0, \sigma^2)$.
- $\mathbf{H}_{ii_k}$ is the $\mathbb{C}^{N\times M}$ mmWave channel matrix spanning from the i-th IAB node to the k-th user of the i-th IAB node.
- $\mathbf{H}_{ii}$ represents the $\mathbb{C}^{M\times M}$ SI channel matrix between the transmit and receive antennas in i-th IAB node.
- $\mathbf{H}_{ji}$ represents the $\mathbb{C}^{M\times M}$ mmWave channel matrix spanning from the j-th IAB node to the i-th IAB node, where $j \neq i$.
- $\mathbf{H}_{j_n i_k}$ represents the $\mathbb{C}^{M\times N}$ mmWave channel matrix spanning from the n-th uplink user of the j-th IAB to the k-th downlink user of i-th IAB node.

The achievable rate of the k-th user in the i-th cell can be expressed as

$$R_{ik} = \log\left|\mathbf{I} + \mathbf{F}_{ik}^H\mathbf{H}_{ii_k}^H\mathbf{W}_{ik}^H\mathbf{J}_{ik}^{-1}\mathbf{W}_{ik}\mathbf{H}_{ii_k}\mathbf{F}_{ik}\right|, \tag{2}$$

where

$$\begin{aligned}\mathbf{J}_{ik} =& \sum_{(j,m)\neq(i,k)}^{L}\sum^{K}\mathbf{W}_{ik}\mathbf{H}_{ji_k}\mathbf{F}_{jm}(\mathbf{W}_{ik}\mathbf{H}_{ji_k}\mathbf{F}_{jm})^H \\ &+ \sum_{j}^{L}\sum_{n}^{K}\mathbf{W}_{ik}\mathbf{H}_{j_n i_k}\mathbf{F}_{jn}(\mathbf{W}_{ik}\mathbf{H}_{j_n i_k}\mathbf{F}_{jn})^H \\ &+ \sum_{j}^{L}\sum_{j'\neq j}^{L}\mathbf{W}_{ik}\mathbf{H}_{ji_k}\mathbf{F}_{jj'}(\mathbf{W}_{ik}\mathbf{H}_{ji_k}\mathbf{F}_{jj'})^H \\ &+ \sigma^2\mathbf{W}_{ik}\mathbf{W}_{ik}^H,\end{aligned} \tag{3}$$

represents the interference-plus-noise power of the k-th downlink user in the i-th cell.

Similarly, the received signal and the achievable rate for the n-th uplink user in the i-th cell can be represented as

$$\begin{aligned}\mathbf{y}_{in} =& \underbrace{\mathbf{W}_{in}\mathbf{H}_{i_n i}\mathbf{F}_{in}\mathbf{s}_{in}}_{\text{Uplink signal}} + \underbrace{\sum_{j}^{L}\sum_{k}^{K}\mathbf{W}_{in}\mathbf{H}_{ji}\mathbf{F}_{jk}\mathbf{s}_{jk}}_{\text{downlink interference}} \\ &+ \underbrace{\sum_{j}^{L}\sum_{j'\neq j}^{L}\mathbf{W}_{in}\mathbf{H}_{ji}\mathbf{F}_{jj'}\mathbf{s}_{jj'}}_{\text{backhaul interference}} + \mathbf{W}_{in}\mathbf{n}_{in} \\ &+ \underbrace{\sum_{(j,l)\neq(i,n)}^{L}\sum^{K}\mathbf{W}_{in}\mathbf{H}_{jli}\mathbf{F}_{jl}\mathbf{s}_{jl}}_{\text{uplink interference}},\end{aligned} \tag{4}$$

and

$$R_{in} = \log\left|\mathbf{I} + \mathbf{F}_{in}^H\mathbf{H}_{i_n i}^H\mathbf{W}_{in}^H\mathbf{J}_{in}^{-1}\mathbf{W}_{in}\mathbf{H}_{i_n i}\mathbf{F}_{in}\right|, \tag{5}$$

where

$$\begin{aligned}\mathbf{J}_{in} = &\sum_{(j,l)\neq(i,n)}^{L}\sum^{K}\mathbf{W}_{in}\mathbf{H}_{jli}\mathbf{F}_{jl}(\mathbf{W}_{in}\mathbf{H}_{jli}\mathbf{F}_{jl})^H \\ &+\sum_{j}^{L}\sum_{k}^{K}\mathbf{W}_{in}\mathbf{H}_{ji}\mathbf{F}_{jk}(\mathbf{W}_{in}\mathbf{H}_{ji}\mathbf{F}_{jk})^H \\ &+\sum_{j}^{L}\sum_{j'\neq j}^{L}\mathbf{W}_{in}\mathbf{H}_{ji}\mathbf{F}_{jj'}(\mathbf{W}_{in}\mathbf{H}_{ji}\mathbf{F}_{jj'})^H \\ &+\sigma^2\mathbf{W}_{in}\mathbf{W}_{in}^H,\end{aligned} \tag{6}$$

represents the interference-plus-noise power of the n-th uplink user in the i-th cell.

Furthermore, considering the backhaul link from the i-th IAB node to the j-th IAB node, the received signal and the achievable rate are given by

$$\begin{aligned}\mathbf{y}_{ij} = &\underbrace{\mathbf{W}_{ij}\mathbf{H}_{ij}\mathbf{F}_{ij}\mathbf{s}_{ij}}_{\text{backhaul}} + \underbrace{\sum_{i'}^{L}\sum_{k}^{K}\mathbf{W}_{ij}\mathbf{H}_{i'j}\mathbf{F}_{i'k}\mathbf{s}_{i'k}}_{\text{downlink interference}} \\ &+\underbrace{\sum_{i'}^{L}\sum_{n}^{K}\mathbf{W}_{ij}\mathbf{H}_{i'_nj}\mathbf{F}_{i'n}\mathbf{s}_{i'n}}_{\text{uplink interference}} + \mathbf{W}_{ij}\mathbf{n}_{ij} \\ &+\underbrace{\sum_{(i',j')\neq(i,j)}^{L}\sum_{j'\neq i'}^{L}\mathbf{W}_{ij}\mathbf{H}_{i'j}\mathbf{F}_{i'j'}\mathbf{s}_{i'j'}}_{\text{backhaul Interference}},\end{aligned} \tag{7}$$

and

$$R_{ij} = \log\left|\mathbf{I} + \mathbf{F}_{ij}^H\mathbf{H}_{ij}^H\mathbf{W}_{ij}^H\mathbf{J}_{ij}^{-1}\mathbf{W}_{ij}\mathbf{H}_{ij}\mathbf{F}_{ij}\right|, \tag{8}$$

where

$$\begin{aligned}\mathbf{J}_{ij} = &\sum_{(i',j')\neq(i,j)}^{L}\sum_{j'\neq i'}^{L}\mathbf{W}_{ij}\mathbf{H}_{i'j}\mathbf{F}_{i'j'}(\mathbf{W}_{ij}\mathbf{H}_{i'j}\mathbf{F}_{i'j'})^H \\ &+\sum_{i'}^{L}\sum_{n}^{K}\mathbf{W}_{ij}\mathbf{H}_{i'_nj}\mathbf{F}_{i'n}(\mathbf{W}_{ij}\mathbf{H}_{i'_nj}\mathbf{F}_{i'n})^H \\ &+\sum_{i'}^{L}\sum_{k}^{K}\mathbf{W}_{ij}\mathbf{H}_{i'j}\mathbf{F}_{i'k}(\mathbf{W}_{ij}\mathbf{H}_{i'j}\mathbf{F}_{i'k})^H \\ &+\sigma^2\mathbf{W}_{ij}\mathbf{W}_{ij}^H\end{aligned} \tag{9}$$

represents the interference-plus-noise power of the backhaul link from the i-th IAB node to the j-th IAB node.

The mmWave channel matrix is given by [3]

$$\mathbf{H}_{ii_k} = \sqrt{\frac{NM}{CR_c}}\sum_{c=1}^{C}\sum_{r_c=1}^{R_c}\alpha_{c,r_c}\mathbf{a}_r\left(\phi_{c,r_c}\right)\mathbf{a}_t^T\left(\theta_{c,r_c}\right), \tag{10}$$

and $\mathbb{E}\left[\|\mathbf{H}_{ii_k}\|_F^2\right] = NM$, where C is the number of clusters, R_c is the number of rays per cluster, and ϕ_{c,r_c} and θ_{c,r_c} are the angles of arrival (AoA) and departure (AoD) of the (c,r_c)-th ray, respectively. Each ray has a complex path gain α_{c,r_c}. Also, $\mathbf{a}_r(\phi_{c,r_c})$ and $\mathbf{a}_t(\theta_{c,r_c})$ are the receive and transmit antenna array response vectors, respectively. The array response vector for a uniform linear array (ULA) can be represented as [20]

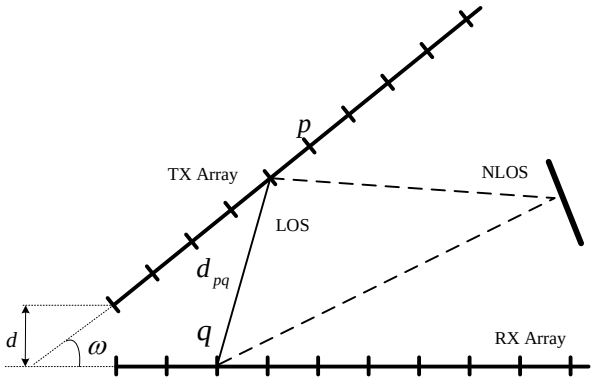

FIGURE 2. Relative position of the transmit and receive arrays in the FD-IAB nodes.

$$\mathbf{a}_r\left(\phi_{c,r_c}\right) = \left[1, \cdots, e^{j\frac{2\pi}{\lambda}(N_r-1)d\cos(\phi_{c,r_c})}\right]^T, \tag{11}$$

$$\mathbf{a}_t\left(\theta_{c,r_c}\right) = \left[1, \cdots, e^{j\frac{2\pi}{\lambda}(M_t-1)d\cos(\theta_{c,r_c})}\right]^T. \tag{12}$$

The SI channel matrix $\mathbf{H}_{ji}(j=i)$ in (4) at the IAB node is modeled as [6], [21]

$$\mathbf{H}_{ii} = \sqrt{\frac{\kappa}{\kappa+1}}\mathbf{H}_{\text{LOS}} + \sqrt{\frac{1}{\kappa+1}}\mathbf{H}_{\text{NLOS}}, \tag{14}$$

where κ represents the Rician factor, $\mathbf{H}_{\text{LOS}}$ represents the line-of-sight (LOS) parts derived from the geometry of the transceiver and $\mathbf{H}_{\text{NLoS}}$ follows the mmWave channel model (10). The entry at the (q,p)-th position of the LOS part can be denoted as [6], [21]

$$[\mathbf{H}_{\text{LOS}}]_{qp} = \frac{\rho}{d_{qp}}\mathrm{e}^{-j2\pi\frac{d_{qp}}{\lambda}}, \tag{15}$$

where ρ represents the power normalization factor to ensure $\mathbb{E}(||\mathbf{H}_{\text{LOS}}||_F^2) = M^2$, λ represents the wavelength and d_{qp} is the distance between the q-th receive and p-th transmit antenna, given by (13). The separation between the transmit and receive arrays of the FD transceiver is defined by distance d, while the transceiver inclination is determined by ω as depicted in Fig. 2.

In the following we assume that all links' channel state information (CSI) is perfectly known and that the optimization is performed at a central IAB node. In practical scenarios, the CSI in mmWave FD systems can be estimated using compressed sensing-based channel estimation algorithms as detailed in [22]. Furthermore, our evaluation encompasses an assessment of the effect of imperfect CSI on the performance of all employed algorithms. In a time-division duplexing (TDD)-assisted IAB system, channel reciprocity serves as a means to acquire CSI within the transceiver units [23]. In contrast, in a frequency-division duplexing (FDD) setup, CSI

$$d_{pq} = \sqrt{\left(\frac{d}{\tan(\omega)} + (q-1)\frac{\lambda}{2}\right)^2 + \left(\frac{d}{\sin(\omega)} + (p-1)\frac{\lambda}{2}\right)^2 - 2\left(\frac{d}{\tan(\omega)} + (q-1)\frac{\lambda}{2}\right)\left(\frac{d}{\sin(\omega)} + (p-1)\frac{\lambda}{2}\right)\cos(\omega)}. \tag{13}$$

acquisition at both the transmitter (TX) and the receiver (RX) can be achieved either through dedicated feedback channels or via the utilization of deep learning methodologies as discussed in [23]. Subsequently, the central node collects these acquired channels through feedback links [24].

B. Problem Formulation

We aim to maximize the weighted sum rate of all the uplink users, downlink users, and the backhaul links between the IAB nodes, by jointly optimizing the precoding $\mathbf{F}$ and combining matrices $\mathbf{W}$ for each link, while adhering to the power constraints imposed at each IAB node and UE. The optimization problem considered in this paper can be formulated as follows:

$$\max_{\mathbf{F},\mathbf{W}} \sum_{i,k} \beta_{ik} R_{ik} + \sum_{i,n} \beta_{in} R_{in} + \sum_{i,j,j\neq i} \beta_{ij} R_{ij} \tag{16a}$$

$$\text{s.t.} \sum_{k}^{K} \|\mathbf{F}_{ik}\|_F^2 + \sum_{j\neq i}^{L} \|\mathbf{F}_{ij}\|_F^2 \leq P_{i,\max},\ i = 1,\dots L, \tag{16b}$$

$$\|\mathbf{F}_{in}\|_F^2 \leq P_{in,\max},\ i = 1,\dots L, n = 1,\dots K, \tag{16c}$$

where β denotes the weighting factor representing the priority of the corresponding links, $P_{i,max}$ and $P_{in,max}$ refer to the transmit power budget at the IAB node side and the uplink user side, respectively.

Due to the coupling effect between the matrices $\mathbf{F}$ and $\mathbf{W}$ as well as multiple signal-to-interference-plus-noise ratio terms, this optimization problem is non-convex and challenging to solve. The following sections present two algorithms, WMMSE and FP, to solve the problem (16).

III. Proposed Beamforming and Combining design

To address the above mentioned problem (16), the WMMSE and FP techniques are employed to transform the original problem into a more tractable formulation. The WMMSE algorithm tackles this problem by utilizing the relation to the weighted minimum-mean-square-error problem. The FP algorithm reformulates it to a set of convex problems. Subsequently, the block coordinate descent (BCD) approach is utilized to solve the reformulated problem.

A. WMMSE

The weighted sum-rate maximization problem (16) is hard to solve for containing multiple ratio terms. However, there is equivalent relation between the WSR maximization problem and the weighted minimum-mean-square-error (WMMSE) problem [24]. Therefore, the optimal precoding and combining matrices can be obtained by exploiting this relationship. In particular, the optimal $\mathbf{W}$ is shown in (17), (18) and (19) .

Upon introducing a set of auxiliary variables, we can obtain the optimal solution of $\mathbf{F}$ as follows:

$$\begin{aligned}
\mathbf{F}_{ik}^{\text{WMMSE}}(\mu_i) &= \beta_{ik} \left(\mathbf{A}_{ik} + \mu_i \mathbf{I}\right)^{-1} \mathbf{H}_{iik}^H \mathbf{W}_{ik}^H \mathbf{V}_{ik}, \\
\mathbf{F}_{in}^{\text{WMMSE}}(\mu_i) &= \beta_{in} \left(\mathbf{A}_{in} + \mu_{in} \mathbf{I}\right)^{-1} \mathbf{H}_{ini}^H \mathbf{W}_{in}^H \mathbf{V}_{in}, \\
\mathbf{F}_{ij}^{\text{WMMSE}}(\mu_i) &= \beta_{ij} \left(\mathbf{A}_{ij} + \mu_i \mathbf{I}\right)^{-1} \mathbf{H}_{ij}^H \mathbf{W}_{ij}^H \mathbf{V}_{ij},
\end{aligned} \tag{20}$$

where

$$\begin{aligned}
\mathbf{A}_{ik} = &\sum_{j}^{L}\sum_{m}^{K} \mathbf{H}_{ij_m}^H \mathbf{W}_{jm}^H \mathbf{V}_{jm} \mathbf{W}_{jm} \mathbf{H}_{ij_m} \\
&+ \sum_{j}^{L}\sum_{j'\neq j}^{L} \mathbf{H}_{ij'}^H \mathbf{W}_{jj'}^H \mathbf{V}_{jj'} \mathbf{W}_{jj'} \mathbf{H}_{ij'} \\
&+ \sum_{j}^{L}\sum_{n}^{K} \mathbf{H}_{ij}^H \mathbf{W}_{jn}^H \mathbf{V}_{jn} \mathbf{W}_{jn} \mathbf{H}_{ij},
\end{aligned} \tag{21}$$

$$\begin{aligned}
\mathbf{A}_{in} = &\sum_{j}^{L}\sum_{k}^{K} \mathbf{H}_{i_n i_k}^H \mathbf{W}_{jk}^H \mathbf{V}_{jk} \mathbf{W}_{jk} \mathbf{H}_{i_n j_k} \\
&+ \sum_{j}^{L}\sum_{j'\neq j}^{L} \mathbf{H}_{i_n j'}^H \mathbf{W}_{jj'}^H \mathbf{V}_{jj'} \mathbf{W}_{jj'} \mathbf{H}_{i_n j'} \\
&+ \sum_{j}^{L}\sum_{l}^{K} \mathbf{H}_{i_n j}^H \mathbf{W}_{jl}^H \mathbf{V}_{jl} \mathbf{W}_{jl} \mathbf{H}_{i_n j},
\end{aligned} \tag{22}$$

$$\begin{aligned}
\mathbf{A}_{ij} = &\sum_{i'}^{L}\sum_{k}^{K} \mathbf{H}_{ii'_k}^H \mathbf{W}_{i'k}^H \mathbf{V}_{i'k} \mathbf{W}_{i'k} \mathbf{H}_{ii'_k} \\
&+ \sum_{i'}^{L}\sum_{j'\neq i'}^{L} \mathbf{H}_{ij'}^H \mathbf{W}_{i'j'}^H \mathbf{V}_{i'j'} \mathbf{W}_{i'j'} \mathbf{H}_{ij'} \\
&+ \sum_{i'}^{L}\sum_{n}^{K} \mathbf{H}_{ii'}^H \mathbf{W}_{i'n}^H \mathbf{V}_{i'n} \mathbf{W}_{i'n} \mathbf{H}_{ii'},
\end{aligned} \tag{23}$$

$$\mathbf{V}_{ik} = \mathbf{E}_{ik}^{-1}, \quad \mathbf{V}_{in} = \mathbf{E}_{in}^{-1}, \text{ and } \mathbf{V}_{ij} = \mathbf{E}_{ij}^{-1}, \tag{24}$$

where $\mathbf{E}$ is the mean-square error (MSE) of the estimated signal.

The value of μ should be chosen to satisfy the power constraint

$$\begin{aligned}
\mu_i^\star &= \min\left\{\mu_i \geq 0 : \sum_{k}^{K} \|\mathbf{F}_{ik}\|_F^2 + \sum_{j\neq i}^{L} \|\mathbf{F}_{ji}\|_F^2 \leq P_{i,\max}\right\}, \\
\mu_{in}^\star &= \min\left\{\mu_{in} \geq 0 : \sum_{k}^{K} \|\mathbf{F}_{in}\|_F^2 \leq P_{in,\max}\right\}.
\end{aligned} \tag{25}$$

$$\mathbf{W}_{ik}^{\text{WMMSE}} = \mathbf{F}_{ik}^H \mathbf{H}_{ii_k}^H \left(\sigma^2 \mathbf{I} + \sum_{j}^{L} \sum_{m}^{K} \mathbf{H}_{ji_k} \mathbf{F}_{jm} \mathbf{F}_{jm}^H \mathbf{H}_{ji_k}^H + \sum_{j}^{L} \sum_{j' \neq j}^{L} \mathbf{H}_{ji_k} \mathbf{F}_{jj'} \mathbf{F}_{j'}^H \mathbf{H}_{ji_k}^H + \sum_{j}^{L} \sum_{n}^{k} \mathbf{H}_{j_n i_k} \mathbf{F}_{j_n} \mathbf{F}_{j_n}^H \mathbf{H}_{j_n i_k}^H \right)^{-1}. \tag{17}$$

$$\mathbf{W}_{in}^{\text{WMMSE}} = \mathbf{F}_{in}^H \mathbf{H}_{i_n i}^H \left(\sigma^2 \mathbf{I} + \sum_{j}^{L} \sum_{k}^{K} \mathbf{H}_{ji} \mathbf{F}_{jk} \mathbf{F}_{jk}^H \mathbf{H}_{ji}^H + \sum_{j}^{L} \sum_{j \neq j'}^{L} \mathbf{H}_{ji} \mathbf{F}_{jj'} \mathbf{F}_{jj'}^H \mathbf{H}_{ji}^H + \sum_{j}^{L} \sum_{l}^{K} \mathbf{H}_{jli} \mathbf{F}_{jl} \mathbf{F}_{jl}^H \mathbf{H}_{jli}^H \right)^{-1}. \tag{18}$$

$$\mathbf{W}_{ij}^{\text{WMMSE}} = \mathbf{F}_{ij}^H \mathbf{H}_{ij}^H \left(\sigma^2 \mathbf{I} + \sum_{i'}^{L} \sum_{k}^{K} \mathbf{H}_{i'j} \mathbf{F}_{i'k} \mathbf{F}_{i'k}^H \mathbf{H}_{i'j}^H + \sum_{i'}^{L} \sum_{j' \neq i'}^{L} \mathbf{H}_{i'j} \mathbf{F}_{i'j'} \mathbf{F}_{i'j}^H \mathbf{H}_{i'j'}^H + \sum_{i'}^{L} \sum_{n}^{K} \mathbf{H}_{i'nj} \mathbf{F}_{i'n} \mathbf{F}_{i'n}^H \mathbf{H}_{i'nj}^H \right)^{-1}. \tag{19}$$

Note that bisection search can efficiently determine the optimal μ [24].

The framework of the WMMSE algorithm to solve the problem (16) is summarized in Algorithm 1.

Algorithm 1 WMMSE strategy for beamforming and combining design in IAB networks

1) Initialize all the variables to feasible values.
2) **repeat**
3) Update $\mathbf{W}$ by (17), (18) and (19)
4) Update $\mathbf{V}$ by (24)
5) Update $\mathbf{F}$ by (20)
6) **until** Convergence

As stated above, the WMMSE method can solve the sum-rate maximization problem (16) by exploiting the equivalent relation with weighted mean-square-error minimization. However, the algorithm iterates to a local optimum while the majority of users will converge to zero and these users are then implicitly not scheduled, which is not only more computationally complex when users are enormous, but also has inferior performance [19], [24]. Inspired by the imperfection, we propose to take advantage of the extended FP method to jointly design the precoder and combiner achieving better performance in the following subsection.

B. Fractional Programming

Without an equivalence premise, the FP algorithm is able to deal with (16) by jointly optimizing the precoder and combiner matrices. Specifically, the main process of the FP is to decouple the denominator and numerator of the signal-to-interference-plus-noise ratio (SINR) terms in (16), which transforms the original non-convex problem to a set of convex ones [18]. This technique can achieve concavity over the optimization variables to transform the problem into a more tractable form, allowing for efficient optimization.

The sum of the weighted-logarithmic-matrix-ratios problem can be formulated as

$$\max_{\mathbf{x}} \sum_{m=1}^{M} \beta_m \log \left| \boldsymbol{I} + \sqrt{\mathbf{C}}_m^H(\mathbf{x}) \mathbf{B}_m^{-1}(\mathbf{x}) \sqrt{\mathbf{C}}_m(\mathbf{x}) \right|. \tag{26}$$

According to the Quadratic and Lagrangian Transform from fractional programming [19], it is equivalent to

$$\begin{aligned} \max_{\mathbf{x},\mathbf{\Gamma},\mathbf{Y}} \ & f^{\text{FP}}(\mathbf{x}, \mathbf{\Gamma}, \mathbf{Y}) \\ \text{s.t. } & \mathbf{x} \in \mathcal{X}, \\ & \mathbf{\Gamma}_m \in \mathbb{H}_+^{n \times n}, \\ & \mathbf{Y}_m \in \mathbb{C}^{n \times n}, \end{aligned} \tag{27}$$

where the new objective function f_{FP} is displayed in (28) and $\mathbf{\Gamma}$ and $\mathbf{Y}$ are auxiliary variables.

Notably, the new objective function exhibits linearity with respect to each $\sqrt{\mathbf{C}}_m(\mathbf{x})$ and $\mathbf{B}_m(\mathbf{x})$, while holding all other terms constant.

An iterative approach is proposed to solve the joint beamforming and combining problem in (16). The proposed methodology involves first reformulating the problem using Quadratic and Lagrangian Transform. To do this, specify the variable $\mathbf{x}$ in (28) as the variables $\mathbf{F}$ and $\mathbf{W}$. In this way, the equivalence of the problem (16) with the following reformulated expression is established

$$\begin{aligned} \max_{\mathbf{F},\mathbf{W},\mathbf{\Gamma},\mathbf{Y}} \ & f^{\text{FP}}(\mathbf{F}, \mathbf{W}, \mathbf{\Gamma}, \mathbf{Y}) \\ \text{s.t.} \ & \text{(16b)}, \\ & \text{(16c)}, \\ & \mathbf{\Gamma} \in \mathbb{H}_+^{n \times n}, \\ & \mathbf{Y} \in \mathbb{C}^{n \times n}, \end{aligned} \tag{32}$$

where

$$f^{\text{FP}}(\mathbf{F}, \mathbf{W}, \mathbf{\Gamma}, \mathbf{Y}) = R_{dl}^{\text{FP}} + R_{ul}^{\text{FP}} + R_{bl}^{\text{FP}}. \tag{33}$$

R_{dl}^{FP}, R_{ul}^{FP} and R_{bl}^{FP} express the downlink, uplink and backhaul parts and can be obtained by (29), (30) and (31). Then, we can optimize the variables within the new problem formulation presented in (32) using an iterative approach. Initially, with $\mathbf{F}$ and $\mathbf{W}$ being held constant, the auxiliary variables $\mathbf{\Gamma}$ and $\mathbf{Y}$ can be determined optimally as

$$\begin{aligned} \mathbf{\Gamma}_{ik} &= \mathbf{F}_{ik}^H \mathbf{H}_{ii_k}^H \mathbf{W}_{ik}^H \mathbf{J}_{ik}^{-1} \mathbf{W}_{ik} \mathbf{H}_{ii_k} \mathbf{F}_{ik}, \\ \mathbf{\Gamma}_{in} &= \mathbf{F}_{in}^H \mathbf{H}_{i_n i}^H \mathbf{W}_{in}^H \mathbf{J}_{in}^{-1} \mathbf{W}_{in} \mathbf{H}_{i_n i} \mathbf{F}_{in}, \\ \mathbf{\Gamma}_{ij} &= \mathbf{F}_{ij}^H \mathbf{H}_{ij}^H \mathbf{W}_{ij}^H \mathbf{J}_{ij}^{-1} \mathbf{W}_{ij} \mathbf{H}_{ij} \mathbf{F}_{ij}, \end{aligned} \tag{34}$$

$$f^{\text{FP}}(\mathbf{x},\mathbf{\Gamma},\mathbf{Y}) = \sum_{m=1}^{M}\left(\log|\mathbf{I}+\mathbf{\Gamma}_m| - \text{tr}\left(\mathbf{\Gamma}_m\right) + \text{tr}\left(\left(\mathbf{I}+\mathbf{\Gamma}_m\right)\left(2\sqrt{\mathbf{C}}_m^H(\mathbf{x})\mathbf{Y}_m - \mathbf{Y}_m^H\left(\mathbf{C}_m(\mathbf{x})+\mathbf{B}_m(\mathbf{x})\right)\mathbf{Y}_m\right)\right)\right). \quad (28)$$

$$R_{dl}^{\text{FP}} = \sum_{i}^{L}\sum_{k}^{K}\Big(\beta_{ik}\log|\mathbf{I}+\mathbf{\Gamma}_{ik}| - \beta_{ik}\,\text{tr}(\mathbf{\Gamma}_{ik}) + \text{tr}\Big(\left(\mathbf{I}+\mathbf{\Gamma}_{ik}\right)\Big(2\sqrt{\beta_{ik}}\mathbf{W}_{ik}\mathbf{H}_{ii_k}\mathbf{F}_{ik}\mathbf{Y}_{ik}^H - \mathbf{Y}_{ik}^H\left(\mathbf{J}_{ik}+\mathbf{W}_{ik}\mathbf{H}_{ii_k}\mathbf{F}_{ik}\mathbf{F}_{ik}^H\mathbf{H}_{ii_k}^H\mathbf{W}_{ik}^H\right)\mathbf{Y}_{ik}\Big)\Big)\Big). \quad (29)$$

$$R_{ul}^{\text{FP}} = \sum_{i}^{L}\sum_{n}^{K}\Big(\beta_{in}\log|\mathbf{I}+\mathbf{\Gamma}_{in}| - \beta_{in}\,\text{tr}(\mathbf{\Gamma}_{in}) + \text{tr}\Big(\left(\mathbf{I}+\mathbf{\Gamma}_{in}\right)\Big(2\sqrt{\beta_{in}}\mathbf{W}_{in}\mathbf{H}_{i_ni}\mathbf{F}_{in}\mathbf{Y}_{in}^H - \mathbf{Y}_{in}^H\left(\mathbf{J}_{in}+\mathbf{W}_{in}\mathbf{H}_{i_ni}\mathbf{F}_{in}\mathbf{F}_{in}^H\mathbf{H}_{i_ni}^H\mathbf{W}_{in}^H\right)\mathbf{Y}_{in}\Big)\Big)\Big). \quad (30)$$

$$R_{bl}^{\text{FP}} = \sum_{i}^{L}\sum_{j\neq i}^{L}\Big(\beta_{ij}\log|\mathbf{I}+\mathbf{\Gamma}_{ij}| - \beta_{ij}\,\text{tr}(\mathbf{\Gamma}_{ij}) + \text{tr}\Big(\left(\mathbf{I}+\mathbf{\Gamma}_{ij}\right)\Big(2\sqrt{\beta_{ij}}\mathbf{W}_{ij}\mathbf{H}_{ij}\mathbf{F}_{ij}\mathbf{Y}_{ij}^H - \mathbf{Y}_{ij}^H\left(\mathbf{J}_{ij}+\mathbf{W}_{ij}\mathbf{H}_{ij}\mathbf{F}_{ij}\mathbf{F}_{ij}^H\mathbf{H}_{ij}^H\mathbf{W}_{ij}^H\right)\mathbf{Y}_{ij}\Big)\Big)\Big). \quad (31)$$

and

$$\begin{aligned}\mathbf{Y}_{ik} &= \left(\mathbf{W}_{ik}\mathbf{H}_{ii_k}\mathbf{F}_{ik}(\mathbf{W}_{ik}\mathbf{H}_{ii_k}\mathbf{F}_{ik})^H + \mathbf{J}_{ik}\right)^{-1}\\ &\quad\cdot\mathbf{W}_{ik}\mathbf{H}_{ii_k}\mathbf{F}_{ik},\\ \mathbf{Y}_{in} &= \left(\mathbf{W}_{in}\mathbf{H}_{i_ni}\mathbf{F}_{in}(\mathbf{W}_{in}\mathbf{H}_{i_ni}\mathbf{F}_{in})^H + \mathbf{J}_{in}\right)^{-1}\\ &\quad\cdot\mathbf{W}_{in}\mathbf{H}_{i_ni}\mathbf{F}_{in},\\ \mathbf{Y}_{ij} &= \left(\mathbf{W}_{ij}\mathbf{H}_{ij}\mathbf{F}_{ij}(\mathbf{W}_{ij}\mathbf{H}_{ij}\mathbf{F}_{ij})^H + \mathbf{J}_{ij}\right)^{-1}\\ &\quad\cdot\mathbf{W}_{ij}\mathbf{H}_{ij}\mathbf{F}_{ij}.\end{aligned} \quad (35)$$

This is then followed by optimizing the beamforming matrices $\mathbf{F}$ and the combining matrices $\mathbf{W}$. An important observation from (28) and (32) is that each link's beamformer or combiner can be determined optimally with $\mathbf{\Gamma}$ and $\mathbf{Y}$ held fixed. By computing the derivatives of (33) with respect to $\mathbf{F}$ and $\mathbf{W}$, the combiner can be computed using (36), (37), and (38). Similarly, the beamforming matrices can be obtained using (39), (40), and (41). Here, the Lagrangian multiplier μ for the power constraint plays a crucial role and can be optimally determined through bisection search using (25). Algorithm 2 summarizes the overall approach of applying FP for the beamforming design in the IAB network.

Algorithm 2 FP strategy for beamforming and combining design in IAB networks

1) Initialize all the variables to feasible values.
2) **repeat**
3) Update $\mathbf{Y}$ by (35)
4) Update $\mathbf{\Gamma}$ by (34)
5) Update $\mathbf{W}$ by (36), (37) and (38),
6) Update $\mathbf{F}$ by (39), (40) and (41)
7) **until** Convergence

The non-decreasing convergence of the FP algorithm can be proved by considering the following chain of reasoning going from iteration i to $i+1$:

$$\begin{aligned}&f^{\text{FP}}\left(\mathbf{F}^{(i)},\mathbf{W}^{(i)},\mathbf{\Gamma}^{(i)},\mathbf{Y}^{(i)}\right)\\ \leq& f^{\text{FP}}\left(\mathbf{F}^{(i)},\mathbf{W}^{(i)},\mathbf{\Gamma}^{(i+1)},\mathbf{Y}^{(i)}\right)\\ \leq& f^{\text{FP}}\left(\mathbf{F}^{(i)},\mathbf{W}^{(i)},\mathbf{\Gamma}^{(i+1)},\mathbf{Y}^{(i+1)}\right)\\ \leq& f^{\text{FP}}\left(\mathbf{F}^{(i)},\mathbf{W}^{(i+1)},\mathbf{\Gamma}^{(i+1)},\mathbf{Y}^{(i+1)}\right)\\ \leq& f^{\text{FP}}\left(\mathbf{F}^{(i+1)},\mathbf{W}^{(i+1)},\mathbf{\Gamma}^{(i+1)},\mathbf{Y}^{(i+1)}\right),\end{aligned} \quad (42)$$

which follows from the fact that the update of $\mathbf{\Gamma}, \mathbf{Y}, \mathbf{W}$, and $\mathbf{F}$ maximizes f^{FP}, when all other variables are fixed.

The FP algorithm is a general strategy to jointly design the precoder and combiner matrices, which first reformulates the original non-convex problem to a set of convex ones using Quadratic and Lagrangian Transform, then iterates to update the precoder and combiner matrices with closed-form expression. Simulation results in the next part show that it outperforms the benchmark, including WMMSE method.

C. Complexity Analysis

Assume that $M \gg N > N_s$, thereby establishing the context for the subsequent complexity analysis. The per-iteration complexity of the FP and WMMSE algorithms is found to be $\mathcal{O}(L^2K^2M^3)$ [19], [24]. It is worth noting that the FP algorithm exhibits slightly higher complexity due to incorporating additional auxiliary variables than the WMMSE algorithms. However, as elaborated in the following section, this increased complexity can be viewed as a trade-off for the superior performance offered by the FP approach.

$$\mathbf{W}_{ik}^{\text{FP}} = \mathbf{Y}_{ik}^{-1}\mathbf{F}_{ik}^{H}\mathbf{H}_{ii_k}^{H}\left(\sigma^2\mathbf{I} + \sum_{j}^{L}\sum_{m}^{K}\mathbf{H}_{ji_k}\mathbf{F}_{jm}\mathbf{F}_{jm}^{H}\mathbf{H}_{ji_k}^{H} + \sum_{j}^{L}\sum_{j'\neq j}^{L}\mathbf{H}_{ji_k}\mathbf{F}_{jj'}\mathbf{F}_{j'}^{H}\mathbf{H}_{ji_k}^{H} + \sum_{j}^{L}\sum_{n}^{k}\mathbf{H}_{j_n i_k}\mathbf{F}_{j_n}\mathbf{F}_{j_n}^{H}\mathbf{H}_{j_n i_k}^{H}\right)^{-1}. \tag{36}$$

$$\mathbf{W}_{in}^{\text{FP}} = \mathbf{Y}_{in}^{-1}\mathbf{F}_{in}^{H}\mathbf{H}_{i_n i}^{H}\left(\sigma^2\mathbf{I} + \sum_{j}^{L}\sum_{k}^{K}\mathbf{H}_{ji}\mathbf{F}_{jk}\mathbf{F}_{jk}^{H}\mathbf{H}_{ji}^{H} + \sum_{j}^{L}\sum_{j\neq j'}^{L}\mathbf{H}_{ji}\mathbf{F}_{jj'}\mathbf{F}_{jj'}^{H}\mathbf{H}_{ji}^{H} + \sum_{j}^{L}\sum_{l}^{K}\mathbf{H}_{jli}\mathbf{F}_{jl}\mathbf{F}_{jl}^{H}\mathbf{H}_{jli}^{H}\right)^{-1}. \tag{37}$$

$$\mathbf{W}_{ij}^{\text{FP}} = \mathbf{Y}_{ij}^{-1}\mathbf{F}_{ij}^{H}\mathbf{H}_{ij}^{H}\left(\sigma^2\mathbf{I} + \sum_{i'}^{L}\sum_{k}^{K}\mathbf{H}_{i'j}\mathbf{F}_{i'k}\mathbf{F}_{i'k}^{H}\mathbf{H}_{i'j}^{H} + \sum_{i'}^{L}\sum_{j'\neq i'}^{L}\mathbf{H}_{i'j}\mathbf{F}_{i'j'}\mathbf{F}_{i'j}^{H}\mathbf{H}_{i'j'}^{H} + \sum_{i'}^{L}\sum_{n}^{K}\mathbf{H}_{i'nj}\mathbf{F}_{i'n}\mathbf{F}_{i'n}^{H}\mathbf{H}_{i'_n j}^{H}\right)^{-1}. \tag{38}$$

$$\mathbf{F}_{ik}^{\text{FP}} = \left(\mu_i\mathbf{I} + \sum_{j}^{L}\sum_{m}^{K}\mathbf{H}_{ij_m}^{H}\mathbf{W}_{jm}^{H}\mathbf{Y}_{jm}(\mathbf{I}+\mathbf{\Gamma}_{jm})\mathbf{Y}_{jm}^{H}\mathbf{W}_{jm}\mathbf{H}_{ij_m} + \sum_{j}^{L}\sum_{j'\neq j}^{L}\mathbf{H}_{ij'}^{H}\mathbf{W}_{jj'}^{H}\mathbf{Y}_{jj'}(\mathbf{I}+\mathbf{\Gamma}_{jj'})\mathbf{Y}_{jj'}^{H}\mathbf{W}_{jj'}\mathbf{H}_{ij'} + \sum_{j}^{L}\sum_{n}^{K}\mathbf{H}_{ij}^{H}\mathbf{W}_{jn}^{H}\mathbf{Y}_{jn}(\mathbf{I}+\mathbf{\Gamma}_{jn})\mathbf{Y}_{jn}^{H}\mathbf{W}_{jn}\mathbf{H}_{ij}\right)^{-1} \cdot \sqrt{\beta_{ik}}\mathbf{H}_{ii_k}^{H}\mathbf{W}_{ik}^{H}\mathbf{Y}_{ik}(\mathbf{I}+\mathbf{\Gamma}_{ik}). \tag{39}$$

$$\mathbf{F}_{in}^{\text{FP}} = \left(\mu_{in}\mathbf{I} + \sum_{j}^{L}\sum_{k}^{K}\mathbf{H}_{i_n j_k}^{H}\mathbf{W}_{jk}^{H}\mathbf{Y}_{jk}(\mathbf{I}+\mathbf{\Gamma}_{jk})\mathbf{Y}_{jk}^{H}\mathbf{W}_{jk}\mathbf{H}_{i_n j_k} + \sum_{j}^{L}\sum_{j'\neq j}^{L}\mathbf{H}_{i_n j'}^{H}\mathbf{W}_{jj'}^{H}\mathbf{Y}_{jj'}(\mathbf{I}+\mathbf{\Gamma}_{jj'})\mathbf{Y}_{jj'}^{H}\mathbf{W}_{jj'}\mathbf{H}_{i_n j'} + \sum_{j}^{L}\sum_{l}^{K}\mathbf{H}_{i_n j}^{H}\mathbf{W}_{jl}^{H}\mathbf{Y}_{jl}(\mathbf{I}+\mathbf{\Gamma}_{jl})\mathbf{Y}_{jl}^{H}\mathbf{W}_{jl}\mathbf{H}_{i_n j}\right)^{-1} \cdot \sqrt{\beta_{in}}\mathbf{H}_{i_n i}^{H}\mathbf{W}_{in}^{H}\mathbf{Y}_{in}(\mathbf{I}+\mathbf{\Gamma}_{in}). \tag{40}$$

$$\mathbf{F}_{ij}^{\text{FP}} = \left(\mu_i\mathbf{I} + \sum_{i'}^{L}\sum_{k}^{K}\mathbf{H}_{ii'_k}^{H}\mathbf{W}_{i'k}^{H}\mathbf{Y}_{i'k}(\mathbf{I}+\mathbf{\Gamma}_{i'k})\mathbf{Y}_{i'k}^{H}\mathbf{W}_{i'k}\mathbf{H}_{ii'_k} + \sum_{i'}^{L}\sum_{j'\neq i'}^{L}\mathbf{H}_{ij'}^{H}\mathbf{W}_{i'j'}^{H}\mathbf{Y}_{i'j'}(\mathbf{I}+\mathbf{\Gamma}_{i'j'})\mathbf{Y}_{i'j'}^{H}\mathbf{W}_{i'j'}\mathbf{H}_{ij'} + \sum_{i'}^{L}\sum_{n}^{K}\mathbf{H}_{ii'}^{H}\mathbf{W}_{i'n}^{H}\mathbf{Y}_{i'n}(\mathbf{I}+\mathbf{\Gamma}_{i'n})\mathbf{Y}_{i'n}^{H}\mathbf{W}_{i'n}\mathbf{H}_{ii'}\right)^{-1} \cdot \sqrt{\beta_{ij}}\mathbf{H}_{ij}^{H}\mathbf{W}_{ij}^{H}\mathbf{Y}_{ij}(\mathbf{I}+\mathbf{\Gamma}_{ij}). \tag{41}$$

For hardware complexity, the residual SI power after the analog circuit domain must remain below the dynamic range threshold of the analog-to-digital converter (ADC) [25] to ensure adequate suppression of SI in the digital domain. Furthermore, collecting comprehensive CSI at the central node represents a practical trade-off in real-world deployments. The acquisition of full CSI significantly demands valuable resources as detailed in [24], [26].

IV. Simulation Results

In this section, the performance of the WSR in an IAB network shown in Fig. 1 is analyzed. The IAB node is placed in the center of each cell, while the uplink and downlink users are randomly distributed. For the mmWave channel, the number of clusters and rays per cluster is set to be 6 and 8, respectively [19]. The distance-dependent path loss experienced in mmWave is modeled by $128.1+37.6\log 10(D)+\tau$ [18], where D is the distance between the transceivers in kilometers and τ is the path loss factor (measured in dB). Table 1 provides the relevant parameters used in simulations, where equal weights are assigned to all links to maximize the sum rate and the interference-to-noise ratio (INR) is used to measure the influence of the SI. For each case, we perform 100 Monte Carlo trails using MATLAB to obtain the desired results, while considering random distribution of users within each cell and all communication links adhere to mmWave path loss, while meeting the power constraint[1]. The precoder matrix, denoted as $\mathbf{F}$, is initialized using complex-Gaussian random variables that adhere to predefined power constraints. Simultaneously, the combiner matrix $\mathbf{W}$ is initialized with complex-Gaussian random variables and normalized to maintain a unit magnitude.

The WSR performance versus the number of iterations, number of antennas, and transmit power are analyzed for the proposed WMMSE and FP based beamforming optimization techniques with the following benchmark techniques. First, the SVD method is used to decompose the channel matrix into simpler components, and by selecting N_s nonzero singular values, the signal transmission is determined. Additionally, we consider the beamforming and combining matrix design approaches proposed in the works of [6], [13],

[1]The authors would like to acknowledge the use of the IRIDIS High Performance Computing Facility and associated support services at the University of Southampton during the course of this research.

TABLE 1. Simulation Parameters

Parameters	Values	Parameters	Values
f	28 GHz	bandwidth	10 MHz
K	8	L	8
d	10λ	τ	8 dB
M	128	N	4
N_s	1,2	κ	30 dB
INR	50 dB	ω	$\pi/6$
cell radius	400 m	noise	-169 dBm/Hz
$P_{in,max}$	20 dBm	$P_{i,max}$	43 dBm

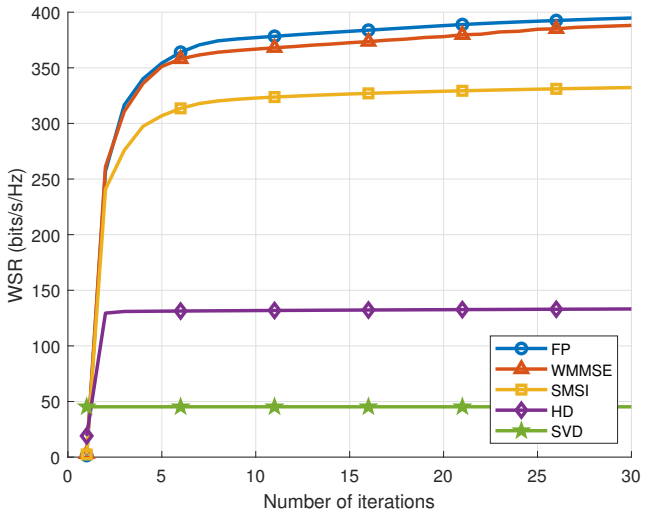

FIGURE 3. The convergence versus iteration step of all optimization methods for a network employing 3 cells with 2 uplink users and 2 downlink users per cell.

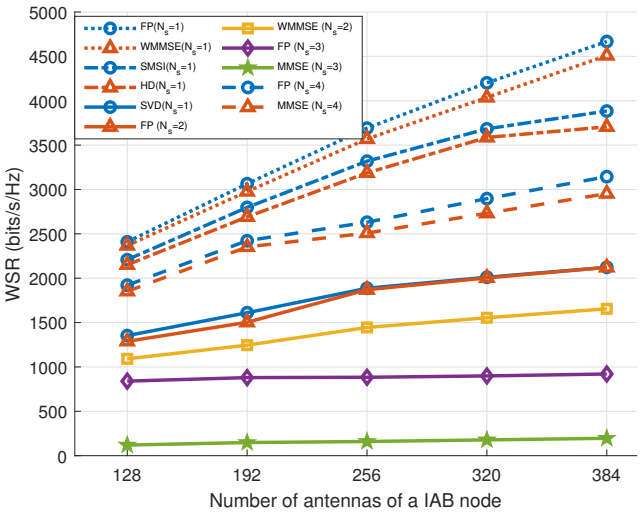

FIGURE 4. Comparison of the WSR versus the number of IAB node antennas for a variable number of streams.

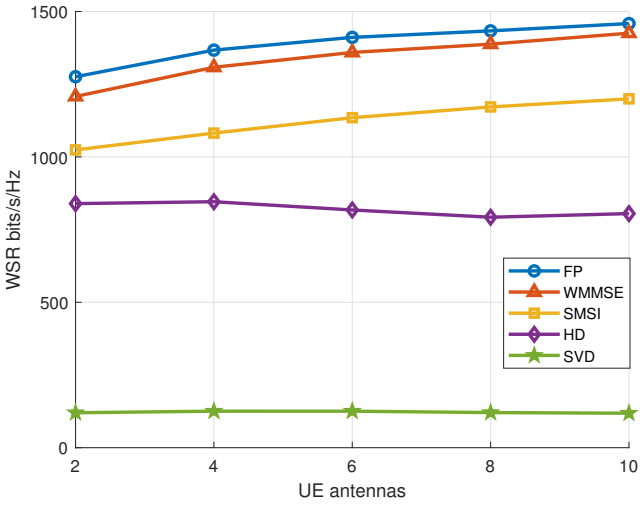

FIGURE 5. Comparison of the WSR of all methods versus the number of UE antennas with $N_s = 1$.

as a benchmark scheme. These methods are designed to separately minimize the self-interference, referred as SMSI, in the beamforming and combining domains, while taking no consideration of other interference terms. These methods are extended and adapted to the multi-cell multi-user network and used as benchmarks for comparative analysis. Finally, the performance of an HD mode scheme is investigated, where the IAB nodes activate the uplink transmission in one-time slot and the downlink transmission in another time slot. In this case, the average rate for the uplink and downlink links is calculated and referred to as the HD sum rate, where the beamforming and combining matrices are optimized using the FP method.

Fig. 3 illustrates the WSR versus the number of iterations. It can be observed from Fig. 3 that both the FP and WMMSE algorithms have a similar convergence speed and value that outperform all other methods, which confirms the practical benefits of our proposed techniques. Specifically, these methods achieve a 15% higher throughput, equivalent to 50 bits/s/Hz, than the SMSI scheme. In addition, they show a 115% higher throughput, reaching 120 bits/s/Hz, compared to the HD strategy in a network employing three cells with two uplink and downlink users per cell. Notably, these throughput gains exceed the 81.5% gain observed in the prototype [25]. Such results highlight the effectiveness of the WMMSE and FP schemes in significantly improving the network's throughput and practical deployment. The superior performance shows that joint optimization of the precoder and combiner can achieve higher gain than the separate optimization scheme.

Fig. 4 and 5 show the effect of the number of IAB nodes and UE antennas, when considering a network with 8 cells, 8 uplink and 8 downlink users per cell. The FD-IAB aided schemes show significant performance gains compared with the HD scheme when the number of antennas increases, as shown in the solid line of the two figures. The SVD-based approach exhibits poor performance due to its inability to entirely mitigate interference, rendering it susceptible to

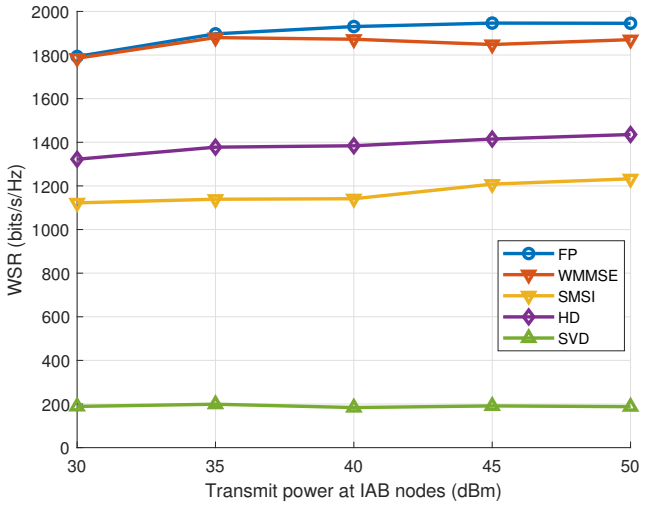

FIGURE 6. Comparison of the WSR of all methods versus the transmit power with $N_s = 2$.

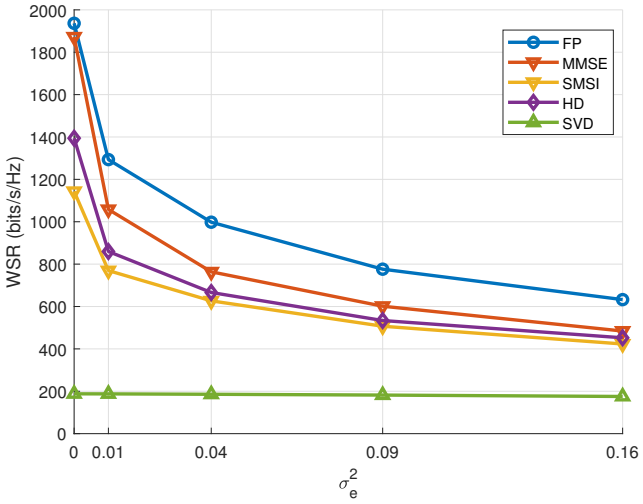

FIGURE 7. Comparison of the WSR of all methods versus imperfect CSI error variance variance of σ_e^2.

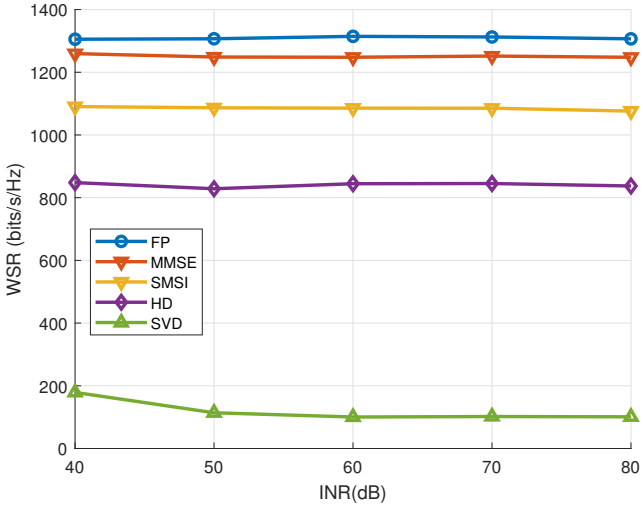

FIGURE 8. Comparison of the WSR of all methods versus INR.

potential disruptions from noise and interference within the data [13]. The number of streams transmitted per link is set to $N_s = 1$ in the above simulations. Additionally, the considered network can support $N_s > 1$ streams per user. Therefore, our study also demonstrates the WSR performance considering multiple streams. Particularly noteworthy is the considerable performance advantage exhibited by the FP scheme, especially in scenarios supporting more streams. This is due to the joint design of the precoding and combining matrices, which enables improved signal transmission and reception. These results highlight the significant benefits of using more antennas and enabling multiple streams, particularly with the FP scheme, to improve the system's performance.

Fig. 6 shows the impact of the transmit power at the IAB node. It can be observed from Fig. 6 that the FP algorithm outperforms all other methods, and the performance improvement becomes even more significant with the increased transmit power. In contrast, the other algorithms gain a slight sum rate, emphasizing the trade-off between the power and the system's performance.

Additionally, we analyze the effect of imperfect channel knowledge on the performance of the considered system. We model CSI imperfections as $h = h_p + \Delta h$ [27], where h represents the imperfect CSI, h_p is the perfect CSI, Δh measures the channel error assumed to follow Gaussian distribution with zero mean and variance of σ_e^2. Fig. 7 shows the effect of imperfect CSI on the performance of all considered algorithms. Explicitly, the FP algorithm shows superior performance compared to the benchmark schemes, which validates the practical implication and robustness of our proposed algorithm in mitigating the effects of imperfect CSI, indicating its potential for real-world applicability and reliability.

Furthermore, we analyze the WSR as a function of the self-interference-to-noise ratio (INR) in Figure 8, where we show that the FP and WMMSE algorithms consistently outperform the SMSI and SVD methods as the INR increases. This highlights the effectiveness of our proposed algorithms in reducing self-interference and emphasizes their robustness, showing their ability to maintain superior performance across varying INR levels.

V. Conclusion

This paper presents a joint precoder and combiner design for multi-cell, multi-user mmWave IAB networks using massive MIMO systems in full-duplex mode. We focus on mitigating residual self-interference and multi-user interference by maximizing the WSR. To deal with the inherent non-convexity of this optimization challenge, we employ two-step approaches: WMMSE and FP. These methods effectively optimize precoding and combining matrices. Simulation results demonstrate significant performance improvements in weighted sum rate maximization, highlighting the relevance

of our methods for addressing this complex optimization problem in multi-cell, multi-user mmWave IAB networks. Future research can consider efficient channel estimation techniques and practical prototype deployment.

REFERENCES

[1] M. Polese, M. Giordani, T. Zugno, A. Roy, S. Goyal, D. Castor, and M. Zorzi, "Integrated Access and Backhaul in 5G mmWave Networks: Potential and Challenges," *IEEE Communications Magazine*, vol. 58, no. 3, pp. 62–68, Mar. 2020.

[2] J. Zhang, N. Garg, M. Holm, and T. Ratnarajah, "Design of Full Duplex Millimeter-Wave Integrated Access and Backhaul Networks," *IEEE Wireless Communications*, vol. 28, no. 1, pp. 60–67, Feb. 2021.

[3] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-Wave Communications: Physical Channel Models, Design Considerations, Antenna Constructions, and Link-Budget," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 870–913, 2018.

[4] 3GPP, "NR; Study on Integrated Access and Backhaul," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 36.874, Dec. 2018.

[5] B. Lee, J.-B. Lim, C. Lim, B. Kim, and J.-y. Seol, "Reflected Self-Interference Channel Measurement for mmWave Beamformed Full-Duplex System," in *2015 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2015, pp. 1–6.

[6] K. Satyanarayana, M. El-Hajjar, P.-H. Kuo, A. Mourad, and L. Hanzo, "Hybrid Beamforming Design for Full-Duplex Millimeter Wave Communication," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1394–1404, Feb. 2019.

[7] A. Khodmi, S. B. Rejeb, N. Agoulmine, and Z. Choukair, "A Joint Power Allocation and User Association Based on Non-Cooperative Game Theory in an Heterogeneous Ultra-Dense Network," *IEEE Access*, vol. 7, pp. 111 790–111 800, 2019.

[8] Z. Guo, Y. Niu, S. Mao, R. He, N. Wang, Z. Zhong, and B. Ai, "Joint Design of Access and Backhaul in Densely Deployed MmWave Small Cells," *IEEE Transactions on Vehicular Technology*, pp. 1–18, 2023.

[9] M. Cudak, A. Ghosh, A. Ghosh, and J. Andrews, "Integrated Access and Backhaul: A Key Enabler for 5G Millimeter-Wave Deployments," *IEEE Communications Magazine*, vol. 59, no. 4, pp. 88–94, Apr. 2021.

[10] A. Koc and T. Le-Ngoc, "Intelligent Non-Orthogonal Beamforming With Large Self-Interference Cancellation Capability for Full-Duplex Multiuser Massive MIMO Systems," *IEEE Access*, vol. 10, pp. 51 771–51 791, 2022.

[11] C. K. Sheemar, C. K. Thomas, and D. Slock, "Practical Hybrid Beamforming for Millimeter Wave Massive MIMO Full Duplex With Limited Dynamic Range," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 127–143, 2022.

[12] M. Diamanti, P. Charatsaris, E. E. Tsiropoulou, and S. Papavassiliou, "The Prospect of Reconfigurable Intelligent Surfaces in Integrated Access and Backhaul Networks," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 2, pp. 859–872, Jun. 2022.

[13] E. Balti, C. Dick, and B. L. Evans, "Low Complexity Hybrid Beamforming for mmWave Full-Duplex Integrated Access and Backhaul," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, Dec. 2022, pp. 1606–1611.

[14] J. Zhang, H. Luo, N. Garg, A. Bishnu, M. Holm, and T. Ratnarajah, "Design and Analysis of Wideband In-Band-Full- Duplex FR2-IAB Networks," *IEEE Transactions on Wireless Communications*, vol. 21, no. 6, pp. 4183–4196, Jun. 2022.

[15] D. Yu, Y. Liu, and H. Zhang, "Energy-Efficient Beamforming Design for User-Centric Full-Duplex Wireless Backhaul Networks," in *2021 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2021, pp. 1–6.

[16] B. Yin, N. Wang, Y. Fan, X. Sun, D. He, and W. Liu, "Evaluation of TDM-based Integrated Access and Backhaul Schemes for 5G and Beyond at mmWave Band," in *2020 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*, Oct. 2020, pp. 348–353.

[17] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-Cell MIMO Cooperative Networks: A New Look at Interference," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[18] K. Shen and W. Yu, "Fractional Programming for Communication Systems—Part I: Power Control and Beamforming," *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2616–2630, May 2018.

[19] K. Shen, W. Yu, L. Zhao, and D. P. Palomar, "Optimization of MIMO Device-to-Device Networks via Matrix Fractional Programming: A Minorization–Maximization Approach," *IEEE/ACM Transactions on Networking*, vol. 27, no. 5, pp. 2164–2177, Oct. 2019.

[20] H. Liu, S. Lu, M. El-Hajjar, and L.-L. Yang, "Machine Learning Assisted Adaptive Index Modulation for mmWave Communications," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1425–1441, 2020.

[21] K. Satyanarayana, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, "Multi-User Full Duplex Transceiver Design for mmWave Systems Using Learning-Aided Channel Prediction," *IEEE Access*, vol. 7, pp. 66 068–66 083, 2019.

[22] Y. Cai, Y. Xu, Q. Shi, B. Champagne, and L. Hanzo, "Robust Joint Hybrid Transceiver Design for Millimeter Wave Full-Duplex MIMO Relay Systems," *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 1199–1215, Feb. 2019.

[23] Z. Linfu, P. Zhiwen, J. Huilin, L. Nan, and Y. Xiaohu, "Deep Learning Based Downlink Channel Covariance Estimation for FDD Massive MIMO Systems," *IEEE Communications Letters*, vol. 25, no. 7, pp. 2275–2279, Jul. 2021.

[24] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[25] G. Y. Suk, S.-M. Kim, J. Kwak, S. Hur, E. Kim, and C.-B. Chae, "Full Duplex Integrated Access and Backhaul for 5G NR: Analyses and Prototype Measurements," *IEEE Wireless Communications*, vol. 29, no. 4, pp. 40–46, Aug. 2022.

[26] K. Shen and W. Yu, "Fractional Programming for Communication Systems—Part II: Uplink Scheduling via Matching," *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2631–2644, May 2018.

[27] S. Katla, L. Xiang, Y. Zhang, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, "Deep Learning Assisted Detection for Index Modulation Aided mmWave Systems," *IEEE Access*, vol. 8, pp. 202 738–202 754, 2020.