Traffic Aware Beamformer Design for Integrated Access and Backhaul with Flexible TDD

Praneeth Jayasinghe, Antti Tölli, Jarkko Kaleva and Matti Latva-aho
Centre for Wireless Communications, University of Oulu, P.O. Box 4500, FIN-90014, Oulu, Finland.
{praneeth.laddu, antti.tolli, jarkko.kaleva, matti.latva-aho}@oulu.fi

Abstract—Integrated access and backhaul (IAB) networks consist of IAB-donor, IAB-nodes, and user-equipments (UEs). Both IAB-donor and IAB-node provide access to UEs while IAB-donor and IAB-nodes exchange UEs data via wireless in-band backhaul using the same frequency-time resources shared with access links. Multi-antenna beamformer techniques can be used to mitigate the complicated cross-link interference scenarios arising from IAB systems. In this paper, an iterative beamformer design with the weighted queue minimization (WQM) objective is proposed for the time-division-duplexed (TDD) based IAB system. In the considered TDD based IAB model, in a given timeslot, IAB-nodes and IAB-donor are assumed to be different uplink (UL)/downlink (DL) transmission modes to mitigate conventional half-duplex loss. Also, the beamformer design is carried out over two timeslots, considering both UL and DL transmission at each node. Specifically, user-specific UL/DL queues are introduced at the IAB-nodes to guarantee the BS to/from UE data delivery. The proposed beamformer solution is based on the iterative evaluation of Karush-Kuhn-Tucker (KKT) conditions of the optimization problem, which can practically be implemented in both centralized and decentralized manner. The numerical examples illustrate the superior system performance of the proposed method in comparison to the conventional half-duplex relaying system.

I. INTRODUCTION

To satisfy ever-growing traffic demand, small-cells should be densely deployed in the next generation of cellular systems [1]. However, connecting these small-cell base stations to the core network using optical fiber links/backhaul can be costly. Therefore, high-speed wireless backhaul comes into play as it is more cost-effective, flexible, and easier to deploy. Moreover, current developments of millimeter-wave (mmWave) communication have backed up the possibility to provide high-speed wireless backhaul to densified small-cell networks. Hence, the third-generation partnership project (3GPP) studies have focused on studying integrated access and backhaul (IAB) systems under the new radio (NR) study item on 3GPP specification [2], [3].

IAB or self-backhauling network consists of three components, IAB-nodes, IAB-donor, and user-equipments (UEs). The node that supports wireless access to UEs and wirelessly

backhauls the access traffic is known as IAB-node. IAB-donor provides UEs interface to the core network and wireless backhauling functionality to IAB-nodes. The IAB offers the flexibility to use the same wireless resources for access and backhaul data transmissions at the same time, in both uplink (DL) and downlink(DL) directions based on the traffic requirement. Therefore, IAB plays a vital role in future wireless communication networks by expanding coverage and improving the throughput with less transmit power requirements, with minimal planning and implementation cost.

Cooperative relaying has been extensively investigated, especially during the last two decades. These studies have shown that relays can be used to improve the reliability, throughput, and coverage of communication with lower power requirements [4]–[6]. Recently, there have been some studies on resource allocation/optimization on the self-backhaul system with the mmWave and massive multiple-input-multiple-output (MIMO) techniques [7], [8]. Most of these studies are either considered full-duplexed (FD) communication at the IAB-node or use of orthogonal time/frequency resources for UL/DL directions. Even though FD relaying facilitate both UL and DL at the same time, implementing them in practice is still not feasible due to production cost and network planning complexity [9]. Moreover, the decode-and-forward (DF) protocol is more suited at the IAB-node, so that IAB-node can decode and forward, access and backhaul traffic to respective UL and DL users. Consequently, it is essential to have user-specific queues for both UL and DL users at IAB-nodes to guarantee end-to-end data delivery.

In comparison to the traditional relay network, IAB allows uplink (UL) and downlink (DL) transmission at the same time. Furthermore, by employing flexible time-division-duplexing (TDD) based resource scheduling, frequency-time resources can be more efficiently managed in the network. However, it is challenging to handle complicated cross-link interference scenarios introduced in the IAB system. Moreover, MIMO systems provide more degrees of freedom to mitigate interference in complex interference-limited systems [10]–[12]. In [12], we have proposed iterative beamformer design to handle UL-DL and DL-UL cross-link interference in the dynamic TDD system by employing multiple antennas at each node. Also, interference coordination with wireless backhaul in the context of DF relaying has been considered in [11].

Motivated by the above concerns, we consider a flexible

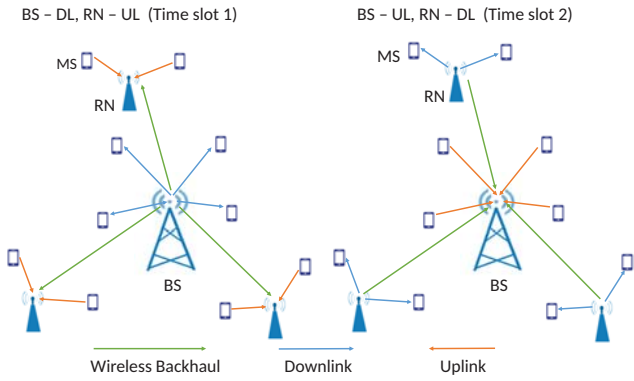BS – DL, RN – UL (Time slot 1)　　BS – UL, RN – DL (Time slot 2)

Fig. 1: Flexible duplexed IAB Network.

TDD based IAB network such that in a given timeslot, IAB-donor and IAB-nodes operate in different UL/DL modes. Then, we propose iterative transmit/receive beamformer design with the weighted queue minimization (WQM) objective by considering traffic dynamics at each node. We assume that IAB-nodes have user-specific UL/DL queues, in addition to queues at UEs and IAB-donor, to guarantee end-to-end data delivery. Due to the relaying function at each IAB node, both UL and DL timeslots are taken into account in the beamformer design. Furthermore, the beamformer design can be implemented in practice either in a centralized or a decentralized manner with additional control signaling to exchange information between nodes.

II. SYSTEM MODEL

As shown in Fig.1, we consider a flexible TDD based multi-user MIMO IAB system consisting of one IAB-donor (referred to as BS) and multiple DF IAB-nodes (i.e., RNs). The set of UEs served by the BS or RN i is denoted by $\mathcal{U}_i$. Here, for the simplicity of the notation, we use $i = 1$ for BS and $i \in \{2, 3, \ldots., N\} = \mathcal{B}_R$ for RNs. The total number of UEs in the system is K and, the number of UEs served by the BS or RN i is $K_i = |\mathcal{U}_i|$. Also, the serving BS/RN of the user k is denoted as i_k. Each UE k employs N_k antenna elements while each BS/RN i employs M_i antenna elements. The maximum number of spatial data streams allocated to UE $k \in \mathcal{U}_i$ is denoted by $L_k \leq \min(M_i, N_k)$. Also, the maximum number of spatial data streams between the BS and RS i is denoted as $\bar{L}_i \leq \min(M_i, M_1)$.

We consider two timeslots, as the data transmission from/to BS to/from UEs takes at least two timeslots via half-duplexed RNs. In the first timeslot, the BS is in DL mode while RNs are in UL mode. In the second timeslot, the BS is in UL mode, and RNs are in DL mode. Therefore, following transmissions and transmit precoders are applied during the first timeslot;

- Tx1 : The BS transmits data using the precoder $\mathbf{m}_{k,l}^{(\text{dl},1)} \in \mathbb{C}^{M_1}$ to DL UE $k \in \mathcal{U}_1$ via l^{th} spatial stream.
- Tx2 : Each UL UE $k \in \mathcal{U}_i$ transmits data to the RN using the precoder $\mathbf{m}_{k,l}^{(\text{ul},1)} \in \mathbb{C}^{N_k}$, via l^{th} spatial stream.
- Tx3 : The BS transmit backhaul data to each RN to serve DL users in the next time slot. Transmit precoder for RN i at BS, via l^{th} spatial stream is $\mathbf{v}_{i,l}^{(\text{dl})} \in \mathbb{C}^{M_1}$.

Due to the above transmissions during the first timeslot, the received signal $\mathbf{x}_k^{(\text{dl},1)} \in \mathbb{C}^{N_k}$ at DL user $k \in \mathcal{U}_1$ can be expressed as

$$\mathbf{x}_k^{(\text{dl},1)} = \mathbf{H}_{1,k}\Big(\sum_{j \in \mathcal{U}_1} \sum_{n=1}^{L_j} \mathbf{m}_{j,n}^{(\text{dl},1)} d_{j,n}^{(\text{dl},1)} + \sum_{i=2}^{N} \sum_{n=1}^{\bar{L}_i} \mathbf{v}_{i,n}^{(\text{dl})} d_{i,n}^{(\text{dl})} \Big)$$
$$+ \sum_{i=2}^{N} \sum_{j \in \mathcal{U}_i} \sum_{n=1}^{L_j} \tilde{\mathbf{H}}_{j,k} \mathbf{m}_{j,n}^{(\text{ul},1)} d_{j,n}^{(\text{ul},1)} + \mathbf{z}_k, \qquad (1)$$

where $\mathbf{H}_{i,k} \in \mathbb{C}^{N_k \times M_i}$ is the channel matrix between BS/RN i and UE k, $\tilde{\mathbf{H}}_{j,k} \in \mathbb{C}^{N_k \times N_j}$ is the UE-UE interference channel matrix between UE j and UE k. All transmit data symbols $d_{j,n}^{(\text{dl},1)}$, $d_{j,n}^{(\text{ul},1)}$ and $d_{i,n}^{(\text{dl})}$ ($\forall j, n, i$) are assumed to be independent and identically distributed (i.i.d.) with $\mathbb{E}\{|d_{j,n}^{(\text{dl},1)}|^2\} = 1$, $\mathbb{E}\{|d_{j,n}^{(\text{ul},1)}|^2\} = 1$ and $\mathbb{E}\{|d_{i,n}^{(\text{dl})}|^2\} = 1$. We assume complex white Gaussian noise vectors $\mathbf{z}_k \in \mathbb{C}^{N_k}$ with variance $\mathcal{N}_0$ per element. Similarly, the received signal $\mathbf{x}_i^{(\text{ul},1)} \in \mathbb{C}^{M_i}$ at RN i can be expressed as

$$\mathbf{x}_i^{(\text{ul},1)} = \hat{\mathbf{H}}_{1,i}\Big(\sum_{j \in \mathcal{U}_1} \sum_{n=1}^{L_j} \mathbf{m}_{j,n}^{(\text{dl},1)} d_{j,n}^{(\text{dl},1)} + \sum_{r=2}^{N} \sum_{n=1}^{\bar{L}_r} \mathbf{v}_{r,n}^{(\text{dl})} d_{r,n}^{(\text{dl})} \Big)$$
$$+ \sum_{r=2}^{N} \sum_{j \in \mathcal{U}_r} \sum_{n=1}^{L_j} \mathbf{H}_{i,j}^{\text{H}} \mathbf{m}_{j,n}^{(\text{ul},1)} d_{j,n}^{(\text{ul},1)} + \mathbf{z}_i, \qquad (2)$$

where $\hat{\mathbf{H}}_{1,i} \in \mathbb{C}^{M_i \times M_1}$ is the channel matrix between the BS and RN i. We assume complex white Gaussian noise vectors and $\mathbf{z}_i \in \mathbb{C}^{M_i}$ with variance $\mathcal{N}_0$ per element. Then, we employ following linear receivers at the receiver node to decode the received data.

- Rx1 : Each DL UE $k \in \mathcal{U}_1$ employs $\mathbf{u}_{k,l}^{(\text{dl},1)} \in \mathbb{C}^{N_k}$, then estimate data as $\hat{d}_{k,l}^{(\text{dl},1)} = (\mathbf{u}_{k,l}^{(\text{dl},1)})^{\text{H}} \mathbf{x}_k^{(\text{dl},1)}$.
- Rx2 : RN i employs $\mathbf{u}_{k,l}^{(\text{ul},1)} \in \mathbb{C}^{M_i}$ to decode the received data from UL UE $k \in \mathcal{U}_i$ as $\hat{d}_{k,l}^{(\text{ul},1)} = (\mathbf{u}_{k,l}^{(\text{ul},1)})^{\text{H}} \mathbf{x}_i^{(\text{ul},1)}$.
- Rx3 : RN i employs $\mathbf{w}_{i,l}^{(\text{dl})} \in \mathbb{C}^{M_i}$ to decode relayed data from BS via l^{th} spatial stream, then estimated data is $\hat{d}_{i,l}^{(\text{dl})} = (\mathbf{w}_{i,l}^{(\text{dl})})^{\text{H}} \mathbf{x}_i^{(\text{ul},1)}$.

Similarly, following transmissions and transmit precoders are used in the second timeslot;

- Tx4 : Each UL UE $k \in \mathcal{U}_1$ transmit data using the precoder $\mathbf{m}_{k,l}^{(\text{ul},2)} \in \mathbb{C}^{N_k}$ via l^{th} spatial stream.
- Tx5 : Each RN transmit data using precoder $\mathbf{m}_{k,l}^{(\text{dl},2)} \in \mathbb{C}^{M_i}$ to DL UE $k \in \mathcal{U}_i$ via l^{th} spatial stream.
- Tx6 : Each RN i transmit received UL data, using precoder $\mathbf{v}_{i,l}^{(\text{ul})} \in \mathbb{C}^{M_i}$ to the BS via l^{th} spatial stream.

Then, corresponding to above transmissions, the received signal $\mathbf{x}_1^{(\text{ul},2)} \in \mathbb{C}^{M_1}$ at the BS can be expressed as

$$\mathbf{x}_1^{(\text{ul},2)} = \sum_{i=2}^{N} \hat{\mathbf{H}}_{1,i}^{\text{T}}\Big(\sum_{n=1}^{\bar{L}_i} \mathbf{v}_{i,n}^{(\text{ul})} d_{i,n}^{(\text{ul})} + \sum_{j \in \mathcal{U}_i} \sum_{n=1}^{L_j} \mathbf{m}_{j,n}^{(\text{dl},2)} d_{j,n}^{(\text{dl},2)} \Big)$$
$$+ \sum_{j \in \mathcal{U}_1} \sum_{n=1}^{L_j} \mathbf{H}_{1,j}^{\text{H}} \mathbf{m}_{j,n}^{(\text{ul},2)} d_{j,n}^{(\text{ul},2)} + \mathbf{z}_1, \qquad (3)$$

where $d_{j,n}^{(\mathrm{dl},2)}$, $d_{j,n}^{(\mathrm{ul},2)}$ and $d_{i,n}^{(\mathrm{ul})}$ ($\forall j, n, i$) are the transmit data symbols, which are i.i.d. with $\mathbb{E}\{|d_{j,n}^{(\mathrm{dl},2)}|^2\} = 1$, $\mathbb{E}\{|d_{j,n}^{(\mathrm{ul},2)}|^2\} = 1$ and $\mathbb{E}\{|d_{i,n}^{(\mathrm{ul})}|^2\} = 1$. Also, we consider $\mathbf{z}_1 \in \mathbb{C}^{M_1}$ to be complex white Gaussian noise vector with variance $\mathcal{N}_0$ per element. Similarly, the received signal $\mathbf{x}_k^{(\mathrm{dl},2)} \in \mathbb{C}^{N_k}$ at DL UE $k \in \mathcal{U}_i$ can be expressed as

$$
\begin{aligned}
\mathbf{x}_k^{(\mathrm{dl},2)} = \sum_{r=2}^{N} \mathbf{H}_{r,k} \Big(\sum_{n=1}^{\bar{L}_i} \mathbf{v}_{i,n}^{(\mathrm{ul})} d_{i,n}^{(\mathrm{ul})} + \sum_{j \in \mathcal{U}_r} \sum_{n=1}^{L_j} \mathbf{m}_{j,n}^{(\mathrm{dl},2)} d_{j,n}^{(\mathrm{dl},2)} \Big) \\
+ \sum_{j \in \mathcal{U}_1} \sum_{n=1}^{L_j} \tilde{\mathbf{H}}_{j,k}^{\mathrm{T}} \mathbf{m}_{j,n}^{(\mathrm{ul},2)} d_{j,n}^{(\mathrm{dl},2)} + \mathbf{z}_k, \quad (4)
\end{aligned}
$$

where $\mathbf{z}_k \in \mathbb{C}^{N_k}$ is complex white Gaussian noise vector with variance $\mathcal{N}_0$ per element. To decode each of the received data, we employ following linear receivers at the receiver node.

- Rx4 : The BS employs $\mathbf{u}_{k,l}^{(\mathrm{ul},2)} \in \mathbb{C}^{M_1}$ to decode data from UL UE $k \in \mathcal{U}_1$ as $\hat{d}_{k,l}^{(\mathrm{ul},2)} = (\mathbf{u}_{k,l}^{(\mathrm{ul},2)})^{\mathrm{H}} \mathbf{x}_1^{(\mathrm{ul},2)}$.
- Rx5 : Each DL UE $k \in \mathcal{U}_i$ employs $\mathbf{u}_{k,l}^{(\mathrm{dl},2)} \in \mathbb{C}^{N_k}$ to decode the received data as $\hat{d}_{k,l}^{(\mathrm{dl},2)} = (\mathbf{u}_{k,l}^{(\mathrm{dl},2)})^{\mathrm{H}} \mathbf{x}_k^{(\mathrm{dl},2)}$.
- Rx6 : The BS employs $\mathbf{w}_{i,l}^{(\mathrm{ul})} \in \mathbb{C}^{M_1}$ to decode the relaying data from RN i as $\hat{d}_{i,l}^{(\mathrm{ul})} = (\mathbf{w}_{i,l}^{(\mathrm{ul})})^{\mathrm{H}} \mathbf{x}_1^{(\mathrm{ul},2)}$.

III. PRECODER DESIGN

In this section, we present an iterative transmit/receive beamformer design with the WQM objective for the flexible TDD based IAB system. For the UEs served by RNs, it minimally takes two timeslots for the end-to-end data delivery. Hence, in the WQM objective, we consider traffic dynamics at each node during both timeslots. For a successful IAB communication, following UL/DL queues are required at each node;

- At the BS, DL user queues ($Q_k^{(\mathrm{dl})}$) are needed for both directly serving and relaying users, as all DL traffic passes through the BS.
- At each UE, UL user queues ($Q_k^{(\mathrm{ul})}$) are maintaining to send UL traffic to the BS. The UL user queues are updating based on the UL arrival process and the serving rate to BS or RN.
- At each RN, user-specific UL ($\bar{Q}_k^{(\mathrm{ul})}$) and DL ($\bar{Q}_k^{(\mathrm{dl})}$) queues are needed to store/relay access and backhaul data, hence guarantee end-to-end data delivery.

Note that the DL user queues at RN, fill-up when backhaul traffic arrives from the BS during the first timeslot, and flow-out when serving the DL users during the second timeslot. Similarly, the UL user queues at RN, fill-up due to the received UL user data during the first timeslot, and flow-out when relaying UL user data to the BS during the second timeslot. Hence, in the overall queue minimization perspective, it is crucial to consider the traffic dynamic at each node over two timeslots.

First, we can define the queue deviation metric $\Psi_k^{(\mathrm{dl})}$ for all DL user queues at the BS, after two timeslots as

$$
\Psi_k^{(\mathrm{dl})} = \begin{cases} Q_k^{(\mathrm{dl})} - \sum_{l=1}^{L_k} R_{k,l}^{(\mathrm{dl},1)} & k \in \mathcal{U}_1, \\ Q_k^{(\mathrm{dl})} - \sum_{l=1}^{L_i} a_k^{(\mathrm{dl})} R_{i,l}^{(\mathrm{dl})} & k \in \mathcal{U}_i, i \in \mathcal{B}_{\mathrm{R}}, \end{cases} \quad (5)
$$

where $R_{k,l}^{(a,s)}$ denotes the number of transmitted bits over the lth spatial stream to/from user k with $a = \{\mathrm{ul, dl}\}$ and $s = \{1, 2\}$. The backhaul rate over the lth spatial stream to/from BS to RN i is denoted as $R_{i,l}^{(a)}$. These backhaul streams are multiplexed with several users and, $\mu_k^{(a)} < 1$ is the multiplexed rate portion for UE k. Also, we assume rate multiplexing for each user in the backhaul is based on time-division-multiplexing (TDM). Moreover, the queue deviation metric $\Psi_k^{(\mathrm{ul})}$ for UL user queues at the UL UE, after two timeslots can be obtained as

$$
\Psi_k^{(\mathrm{ul})} = \begin{cases} Q_k^{(\mathrm{ul})} - \sum_{l=1}^{L_k} R_{k,l}^{(\mathrm{ul},1)} & k \in \mathcal{U}_i, i \in \mathcal{B}_{\mathrm{R}}, \\ Q_k^{(\mathrm{ul})} - \sum_{l=1}^{L_k} R_{k,l}^{(\mathrm{ul},2)} & k \in \mathcal{U}_1. \end{cases} \quad (6)
$$

Similarly, the queue deviation metric $\bar{\Psi}_k^{(\mathrm{dl})}$ and $\bar{\Psi}_k^{(\mathrm{ul})}$ for DL and UL user queues at the RN i_k, after two timeslots are given by

$$
\bar{\Psi}_k^{(\mathrm{dl})} = \bar{Q}_k^{(\mathrm{dl})} + \sum_{l=1}^{L_i} \mu_k^{(\mathrm{dl})} R_{i,l}^{(\mathrm{dl})} - \sum_{l=1}^{L_k} R_{k,l}^{(\mathrm{dl},2)}, \quad (7\mathrm{a})
$$

$$
\bar{\Psi}_k^{(\mathrm{ul})} = \bar{Q}_k^{(\mathrm{ul})} + \sum_{l=1}^{L_k} R_{k,l}^{(\mathrm{ul},1)} - \sum_{l=1}^{L_i} \mu_k^{(\mathrm{ul})} R_{i,l}^{(\mathrm{ul})}. \quad (7\mathrm{b})
$$

In order to simplify the notation, let $\tilde{\boldsymbol{\Psi}}^{(a)}$ and $\tilde{\bar{\boldsymbol{\Psi}}}^{(a)}$ denote vectors with elements $\tilde{\Psi}_k^{(a)} \triangleq \alpha_k^{1/q} \Psi_k^{(a)}$ and $\tilde{\bar{\Psi}}_k^{(a)} \triangleq \alpha_k^{1/q} \bar{\Psi}_k^{(a)}$, respectively. Here, α_k is the weighting factor that is used to prioritize users, based on their quality of service (QoS) requirements. Then, we can define weighted ℓ_q-norm queue minimization of the UL and DL users during two timeslots with sum transmit power constraints at the transmitters as

$$
\min_{\mathcal{M}, \mathcal{W}} \sum_{a \in \{\mathrm{ul, dl}\}} \|\tilde{\boldsymbol{\Psi}}^{(a)}\|_q + \|\tilde{\bar{\boldsymbol{\Psi}}}^{(a)}\|_q \quad (8\mathrm{a})
$$

$$
\text{s. t.} \quad \sum_{l=1}^{L_k} \|\mathbf{m}_{k,l}^{(\mathrm{ul},1)}\|^2 \leq P_k^{(\mathrm{ul})} \quad \forall k \in \mathcal{U}_i, i \neq 1, \quad (8\mathrm{b})
$$

$$
\sum_{k \in \mathcal{U}_1} \sum_{l=1}^{L_k} \|\mathbf{m}_{k,l}^{(\mathrm{dl},1)}\|^2 + \sum_{i=2}^{N} \sum_{l=1}^{\bar{L}_i} \|\mathbf{v}_{i,l}^{(\mathrm{dl})}\|^2 \leq P_1^{(\mathrm{dl})}, \quad (8\mathrm{c})
$$

$$
\sum_{l=1}^{L_k} \|\mathbf{m}_{k,l}^{(\mathrm{ul},2)}\|^2 \leq P_k^{(\mathrm{ul})} \quad \forall k \in \mathcal{U}_1, \quad (8\mathrm{d})
$$

$$
\sum_{k \in \mathcal{U}_i} \sum_{l=1}^{L_k} \|\mathbf{m}_{k,l}^{(\mathrm{dl},2)}\|^2 + \sum_{l=1}^{\bar{L}_i} \|\mathbf{v}_{i,l}^{(\mathrm{ul})}\|^2 \leq P_i^{(\mathrm{dl})} \quad \forall i \in \mathbb{B}_{\mathrm{R}}. \quad (8\mathrm{e})
$$

Here, $\mathcal{M}$ represent the set of all transmit precoders and $\mathcal{W}$ represent the set of all receive beamformers. Maximum transmit power at BS/RN i is denoted as $P_i^{(\mathrm{dl})}$ and maximum transmit power at UE k denoted as $P_k^{(\mathrm{ul})}$. The WQM problem in (8) is an NP-hard problem, but computationally efficient

$$\mathbf{M}_k^{(\text{dl},1)} = \mathbf{H}_{1,k}\Big(\sum_{j\in\mathcal{U}_1}\sum_{n=1}^{L_j}\mathbf{m}_{j,n}^{(\text{dl},1)}(\mathbf{m}_{j,n}^{(\text{dl},1)})^{\text{H}} + \sum_{r=2}^{N}\sum_{n=1}^{\bar{L}_r}\mathbf{v}_{r,n}^{(\text{dl})}(\mathbf{v}_{r,n}^{(\text{dl})})^{\text{H}}\Big)\mathbf{H}_{1,k}^{\text{H}} + \sum_{r=2}^{N}\sum_{j\in\mathcal{U}_r}\sum_{n=1}^{L_j}\tilde{\mathbf{H}}_{j,k}\mathbf{m}_{j,n}^{(\text{ul},1)}(\tilde{\mathbf{H}}_{j,k}\mathbf{m}_{j,n}^{(\text{ul},1)})^{\text{H}} + \mathcal{N}_0\mathbf{I}. \quad (10)$$

solutions can be found by iterative alternating optimization (AO) as detailed in [12]–[14]. The main steps are briefly reproduced here for clarity.

First, we employ MMSE receivers at each RX node. Then, we can represent individual rate terms using corresponding MSE terms [13]. Then, instantaneous rate for DL/UL user k and BS-RN backhaul via lth spatial stream can be expressed by $R_{k,l}^{(a,s)} = -\log_2(\epsilon_{k,l}^{(a,s)})$ and $R_{i,l}^{(a)} = -\log_2(\epsilon_{i,l}^{(a)})$, respectively. Here, $\epsilon_{k,l}^{(a,s)} = \mathbb{E}[|\hat{d}_{k,l}^{(a,s)} - \hat{d}_{k,l}^{(a,s)}|^2]$ is the user specific MSE for the DL/UL data detection and $\epsilon_{i,l}^{(a)} = \mathbb{E}[|\hat{d}_{i,l}^{(a)} - \hat{d}_{i,l}^{(a)}|^2]$ is the MSE for the data detection corresponding to backhual traffic. Then, the user-specific MSE value corresponding to Rx1 can be obtain as[1]

$$\epsilon_{k,l}^{(\text{dl},1)} = 1 - 2\Re((\mathbf{u}_{k,l}^{(\text{dl},1)})^{\text{H}}\mathbf{H}_{1,k}\mathbf{m}_{k,l}^{(\text{dl},1)}) + (\mathbf{u}_{k,l}^{(\text{dl},1)})^{\text{H}}\mathbf{M}_k^{(\text{dl},1)}\mathbf{u}_{k,l}^{(\text{dl},1)}, \quad (9)$$

where $\mathbf{M}_k^{(\text{dl},1)} = \mathbb{E}[\mathbf{x}_k^{(\text{dl},1)}(\mathbf{x}_k^{(\text{dl},1)})^{\text{H}}]$ is the received signal covariance matrix for DL UE k. The expression for $\mathbf{M}_k^{(\text{dl},1)}$ is given in (10) top of the page[2]. Then, the MMSE receiver corresponding to Rx1 is given by

$$\tilde{\mathbf{u}}_{k,l}^{(\text{dl},1)} = (\mathbf{M}_k^{(\text{dl},1)})^{-1}\mathbf{H}_{1,k}\mathbf{m}_{k,l}^{(\text{dl},1)}, \quad (11)$$

Then, by re-writing rate terms using the corresponding MSE terms and introducing auxiliary MSE constraints as in [14] to (8), we can construct an approximated optimization problem as

$$\min_{\mathcal{M},\mathcal{W},\mathcal{T}} \sum_{k\in\mathcal{U}_1}\Big(\alpha_k^{(\text{dl})}(Q_k^{(\text{dl})} - \sum_{l=1}^{L_k}K_0 t_{k,l}^{(\text{dl},1)})^q + \alpha_k^{(\text{ul})}(Q_k^{(\text{ul})}$$

$$-\sum_{l=1}^{L_k}K_0 t_{k,l}^{(\text{ul},2)})^q\Big) + \sum_{i=2}^{N}\sum_{k\in\mathcal{U}_i}\Big(\alpha_k^{(\text{ul})}(Q_k^{(\text{ul})} - \sum_{l=1}^{L_k}K_0 t_{k,l}^{(\text{ul},1)})^q$$

$$+\alpha_k^{(\text{dl})}(Q_k^{(\text{dl})} - \sum_{l=1}^{\bar{L}_i}\mu_k^{(\text{dl})}K_0 t_{i,l}^{(\text{dl})})^q$$

$$+\alpha_k^{(\text{dl})}(\bar{Q}_k^{(\text{dl})} + \sum_{l=1}^{\bar{L}_i}\mu_k^{(\text{dl})}K_0 t_{i,l}^{(\text{dl})} - \sum_{l=1}^{L_k}K_0 t_{k,l}^{(\text{dl},2)})^q$$

$$+\alpha_k^{(\text{ul})}(\bar{Q}_k^{(\text{ul})} + \sum_{l=1}^{L_k}K_0 t_{k,l}^{(\text{ul},1)} - \sum_{l=1}^{\bar{L}_i}\mu_k^{(\text{ul})}K_0 t_{i,l}^{(\text{ul})})^q\Big), \quad (12a)$$

$$\text{s. t. } \epsilon_{k,l}^{(a,s)} \le \beta^{-t_{k,l}^{(a,s)}} \quad \forall (k,l), a\in\{\text{dl,ul}\}, s\in\{1,2\}, \quad (12b)$$

$$\epsilon_{i,l}^{(a)} \le \beta^{-t_{i,l}^{(a)}} \quad \forall i\in\mathbb{B}_{\text{R}} \,\&\, \forall l, a\in\{\text{dl,ul}\}, \quad (12c)$$

$$(8b), (8c), (8d), (8e).$$

[1] Expressions for Rx2/Tx2 to Rx6/Tx6 are not presented due to the repetitive nature of the equations.

[2] Other covariance/weighted covariance matrices are not presented in the paper due to the repetitive nature of the equations.

Here, $\mathcal{T}$ represents the set of the newly introduce auxiliary variables $t_{k,l}^{(a,s)}/t_{i,l}^{(a)}$. Also, $K_0 = \log_2(\beta)$ and β is a predefined constant to adjust the approximation function such that $\beta > 0$ [14]. By introducing these MSE constraints, the objective becomes a convex function of auxiliary variables $t_{k,l}^{(a,s)}/t_{i,l}^{(a)}$. However, constraints in (12b) and (12c) are still non-convex, and that non-convexity can be handled iteratively by using the first-order Taylor series approximation [14]. For example, (12b) can be approximated as,

$$\beta^{-t_{k,l}^{(a,s)}} = -K_1 t_{k,l}^{(a,s)} + K_2, \quad (13)$$

where $K_1 = \beta^{-\bar{t}_{k,l}^{(a,s)}}\log(\beta)$, $K_2 = \beta^{-\bar{t}_{k,l}^{(a,s)}} + \bar{t}_{k,l}^{(a,s)}K_1$ and $\bar{t}_{k,l}^{(a,s)}$ is the point of approximation. For (12c), same approximation is applied as in (13). Then, by substituting approximated expressions in (13) to (12b) and (12c), the optimization problem in (12) can be efficiently solved using the KKT optimality conditions [13].

A. Alternating optimization Method

Here, we present the iterative AO method using the KKT optimality conditions as follows; We begin by fixing transmit precoders and solving for the receive beamformers and other variables (auxiliary and dual). First, we calculate the MMSE receivers using (11), then corresponding MSE values are obtained from (9). Then, by using the complementary slackness of (12b) and (12c), we can update auxiliary variables $t_{k,l}^{(a,s)}/t_{i,l}^{(a)}$ as

$$t_{k,l}^{(a,s)} = \bar{t}_{k,l}^{(a,s)} + \frac{1}{\log(\beta)}\Big(1 - \epsilon_{k,l}^{(a,s)}\beta^{\bar{t}_{k,l}^{(a,s)}}\Big), \quad (14)$$

where $\bar{t}_{k,l}^{(a,s)}$ denotes $t_{k,l}^{(a,s)}$ from the previous iteration. This is corresponding to sub-gradient update of dual variable $t_{k,l}^{(a,s)}$ with the step size $\frac{1}{\log(\beta)}$. Hence for the faster convergence, we can experiment with the step size as in [14]. Next, dual variables $\omega_{k,l}^{(a,s)}/\omega_{i,l}^{(a)}$ corresponding to (12b) and (12c) are obtained as

$$\omega_{k,l}^{(a,s)} = (1-\rho)\bar{\omega}_{k,l}^{(a,s)} + \rho\frac{qK_0}{K_1}\Upsilon, \quad (15)$$

where $\bar{\omega}_{k,l}^{(a,s)}$ denotes fixed $\omega_{k,l}^{(a,s)}$ from the previous iteration. Here, $\rho \in (0,1)$ controls the rate of convergence and is used to prevent over-allocation. For Rx1 Υ is given by[1]

$$\Upsilon = \alpha_k^{(\text{dl})}[(Q_k^{(\text{dl})} - \sum_{l=1}^{L_k}K_0 t_{k,l}^{(\text{dl},1)})^{q-1}]^+, \quad (16)$$

where $[x]^+ \triangleq \max\{x,0\}$. Note that to simplify the solution, multiplexing factors $\mu_k^{(a)}$ are calculated based on the queue

$$\mathbf{\Phi}_1^{(dl,1)} = \sum_{k \in \mathcal{U}_1} \sum_{l=1}^{L_j} \omega_{k,l}^{(dl,1)} \mathbf{H}_{1,k}^H \mathbf{u}_{k,l}^{(dl,1)} (\mathbf{H}_{1,k}^H \mathbf{u}_{k,l}^{(dl,1)})^H + \sum_{i=2}^{N} \hat{\mathbf{H}}_{1,i}^H (\sum_{l=1}^{\bar{L}_i} \omega_{i,l}^{(dl)} \mathbf{w}_{i,l}^{(dl)} (\mathbf{w}_{i,l}^{(dl)})^H + \sum_{k \in \mathcal{U}_i} \sum_{l=1}^{L_j} \omega_{k,l}^{(ul,1)} \mathbf{u}_{k,l}^{(ul,1)} (\mathbf{u}_{k,l}^{(ul,1)})^H) \hat{\mathbf{H}}_{1,i}. \quad (19)$$

state of the users. We consider following method to multi-plexing rate for each user via backhaul;

$$\mu_k^{(a)} = \frac{Q_k^{(a)}}{\sum_{k=1}^{|\mathcal{U}_i|} Q_k^{(a)}}, \quad i \in \mathbb{B}_{\mathbf{R}}. \quad (17)$$

Next, we fix the MMSE receivers and solve for the trans-mit precoders. The transmit precoders can be derived from the first-order optimality conditions of (12). Hence, transmit precoders for Tx1 transmitter type can be obtained as[1]

$$\mathbf{m}_{k,l}^{(dl,1)} = \left(\mathbf{\Phi}_1^{(dl,1)} + \nu_1^{(dl,1)}\mathbf{I}\right)^{-1} \omega_{k,l}^{(dl,1)} \mathbf{H}_{1,k}^H \mathbf{u}_{k,l}^{(dl,1)}, \quad (18)$$

where $\mathbf{\Phi}_1^{(dl,1)}$ is the weighted transmit covariance matrix[2] and expression for $\mathbf{\Phi}_1^{(dl,1)}$ is obtained as in (19) top of the page. Also, $\nu_1^{(dl,1)}$, is the dual variable corresponding to power constraint in (8c). Hence, the transmit beamformers can efficiently solved from (18), by bisection search over the dual variables to satisfy the power constraint.

Finally, we repeat above precoder/decoder optimization until the convergence of the objective function. Above beamformer design can be summarized as in Algorithm 1.

Algorithm 1 Iterative Beamformer Design

1: Initializing feasible transmit beamformers $\mathbf{m}_{k,l}^{(a,s)}/\mathbf{v}_{i,l}^{(a)}$.
2: Calculate multiplexing factors $\mu_k^{(a)}$ to each user via back-haul from (17).
3: **repeat**
4:    Estimate MMSE receivers $\mathbf{u}_{k,l}^{(a,s)}/\mathbf{w}_{i,l}^{(a)}$ from (11).
5:    Calculate MSE values $\epsilon_{k,l}^{(a,s)}/\epsilon_{i,l}^{(a)}$ from (9) .
6:    Calculate auxiliary variables $t_{k,l}^{(a,s)}/t_{i,l}^{(a)}$ from (14).
7:    Calculate dual variables $\omega_{k,l}^{(a,s)}/\omega_{i,l}^{(a)}$ using (15) and (16).
8:    Estimate transmit precoders $\mathbf{m}_{k,l}^{(a,s)}/\mathbf{v}_{i,l}^{(a)}$ from (18).
9: **until** convergence.

B. Practical Implementation

It is essential to look at the feasibility of implementing the proposed iterative precoder/decoder designs in practice. Inter-estingly, the proposed beamformer design can be implemented in practice, either in a centralized or a decentralized manner.

In the centralized design, the central processing unit (CPU) performs all the steps stated in Algorithm 1. To do that, the CPU collects the following information; 1) All the BS-UE, RN-UE, BS-RN, and UE-UE fading channels. 2) All the UL/DL queue states and power constraints from each node. Once all the information is collected, the CPU can execute iterative transmit and receive beamformer calculations until the objective or beamformers are converged. Then, the optimized beamformers are distributed to the relevant nodes before

the data transmission. However, the centralized approach is challenging in dense networks due to the increment of over-head and complexity in the information collection process. Specifically, user-specific cross-user channel estimation and reporting these CSI to CPU is a tedious task. On the other hand, the proposed beamformer design can be implemented in a decentralized manner by estimating node-specific beam-formers at the corresponding nodes, which is a more flexible and scalable solution.

With the development of the 5G NR study, lots of flexibility is provided to work on application-specific frame structures. Specifically, with the minislot concept, faster OTA information exchange is possible in either direction. By carefully following the 5G NR frame structure and Numerology, we in [12] studied decentralized beamformer design for dynamic TDD system, by using bi-directional training to exchange interme-diate beamformers between coordinated nodes. Accordingly, we can employ a similar OTA bi-directional signaling strategy to implement the proposed iterative beamformer design decen-tralized manner. In the decentralized design, transmitter nodes are initializing/estimating transmit precoders and multiplexing factors while receiver nodes are estimating MMSE receivers, MSE values, auxiliary variable, and user-specific weights. To do that, precoded pilot sequences are used to exchange initial/intermediate beamformers and user-specific weights it-eratively between coordinated nodes in both forward and back-ward directions. Each node estimates their precoder/decoder based on the received forward/backward training sequences and, the estimated precoder/decoder is used for precoding for the next iteration forward/backward training.

IV. NUMERICAL EXAMPLES

In the simulation model, we consider $N = 5$ (a BS and 4 RNs) with 4 UEs ($K_i = 4$) associated to each BS/RN ($K = 20$). The number of antennas at the BS is $M_1 = 20$, at each RN is $M_i = 8$ and at each UE is $N_k = 2$. The distances between the BS and RNs are considered to be 200m and the UEs are randomly placed within 50m from the serving BS/RN. The path loss exponent is fixed to 3.67. The power constraint for BS is normalized to $P_1 = 10$ and, power constraints at RNs and UEs are $P_i = P_1/(N-1)$ and $p_k = P_1/K$, respectively. Noise power ($\mathcal{N}_0$) is obtained assuming the cell edge (100m from the BS) SNR for BS transmission to be 20 dB. We consider uncorrelated fading, which is modeled using Clarke's channel model. We consider Poisson arrival process to generate the traffic in the network, where $\lambda_k^{(dl)}(\tau) \sim \mathbf{Pois}(A_k^{(dl)})$ is the generated traffic for DL user k in time instance τ. Similarly, $\lambda_k^{(ul)}(\tau) \sim \mathbf{Pois}(A_k^{(ul)})$ defines the generated traffic at UL user k at time instance τ. Here, $A_k^{(dl)} = \mathbb{E}_\tau\{\lambda_k^{(dl)}\}$ and $A_k^{(ul)} = \mathbb{E}_\tau\{\lambda_k^{(ul)}\}$ are the average number of packet arrivals in bits for the corresponding
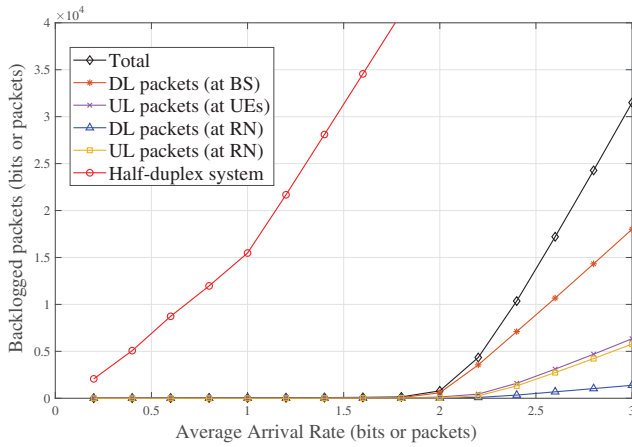
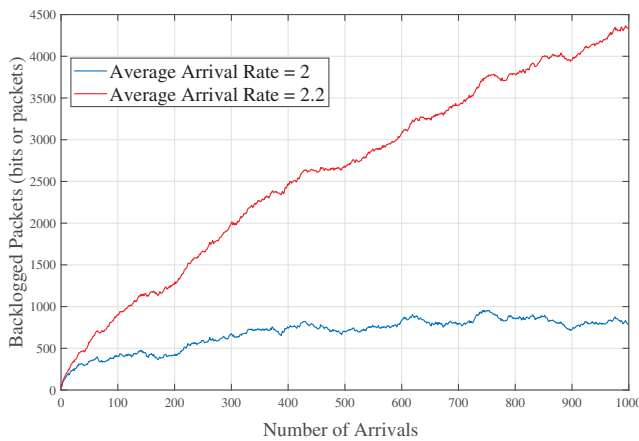Fig. 2: Total backlogged packets after 1000 traffic arrivals.

Fig. 3: Total backlogged packets variation with the traffic arrivals.

UL/DL users. Then, the total number of queued packets in each UL/DL queue at $(\tau + 1)$th time instant is given by $Q_k^{(a)}(\tau + 1) = \left[Q_k^{(a)}(\tau) - R_k^{(a)}(\tau) \right]^+ + \lambda_k^{(a)}(\tau)$, where $R_k^{(a)}$ is the transmission rate to/from user k. Also, the user priority weights assumed to be 1.

Fig. 2 illustrates the average total of backlogged packets after 1000 traffic arrivals (which is equivalent to 2000 times-lots) in the system queues for a given traffic arrival rate. As the reference case, we consider a conventional half-duplex relaying system without simultaneous UL/DL transmission. In the reference system, we assume 50% of the time it is in DL mode, and rest of the time it is in UL mode. We can observe that flexible TDD based IAB system always perform much better in comparison to the reference case in all traffic arrival rates. The IAB system becomes unstable after the arrival rate is 2.2 (linear growth as a function of arrival rate), while the half-duplex case gets saturated for lower arrival rates. Also, we can observe that RNs have lower queues when compared to UEs and BS. Hence, the RNs can potentially have smaller buffer sizes without causing a negative impact on system performance. Fig. 3 illustrates the total backlogged packets in the IAB system with the number of packet arrivals. For the arrival rate is 2, the system is in a stable region

where backlogged packets fluctuate around 700 bits/packets. However, when the average arrival rate is 2.2, the IAB system starts to be unstable and grows the backlogged packets with the number of arrivals.

V. CONCLUSIONS

In this paper, a flexible TDD based IAB system was investigated with complex interference conditions due to simultaneous UL/DL traffic, in-band access, and backhaul traffic. Multiantenna beamforming techniques were applied to mitigate the interference with the WQM objective. The original NP-hard optimization problem was solved to get computationally efficient solutions iteratively using the AO method. In numerical examples, a flexible TDD based IAB system is shown superior performance with the proposed beamformer design in comparison to the conventional half-duplex relaying system. IAB-nodes can potentially have smaller buffer sizes without causing a negative impact on system performance.

REFERENCES

[1] Cisco Systems, "Cisco Visual Networking Index: Forecast and Trends, 2017–2022," [Online Oct. 2019], https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.pdf.

[2] 3GPP, " Study on Integrated Access and Backhaul for NR TR 38.874," Tech. Rep., 3rd Generation Partnership Project 3GPP, www.3gpp.org, 2018.

[3] 3GPP, " Study on security aspects of Integrated Access and Backhaul (IAB) for Next Radio (NR) TR 33.824," Tech. Rep., 3rd Generation Partnership Project 3GPP, www.3gpp.org, 2019.

[4] J. N. Laneman, D. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans on Information theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[5] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. on Information Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[6] P. Jayasinghe, L. K S. Jayasinghe, M. Juntti, and M. Latva-Aho, "Performance analysis of optimal beamforming in fixed-gain AF MIMO relaying over asymmetric fading channels," *IEEE Trans. on commun.*, vol. 62, no. 4, pp. 1201–1217, Feb. 2014.

[7] L. Sanguinetti, A. L Moustakas, and M. Debbah, "Interference management in 5G reverse TDD HetNets with wireless backhaul: A large system analysis," *IEEE journal on Sel. Areas in Commun.*, vol. 33, no. 6, pp. 1187–1200, Mar. 2015.

[8] B. Li, D. Zhu, and P. Liang, "Small cell in-band wireless backhaul in massive MIMO systems: A cooperation of next-generation techniques," *IEEE Trans. on Wireless Commun.*, vol. 14, no. 12, pp. 7057–7069, Aug. 2015.

[9] R. Gupta and S. Kalyanasundaram, "Resource allocation for self-backhauled networks with half-duplex small cells," in *IEEE Int. Conf. on Commun. Workshops (ICC Workshops)*. IEEE, May. 2017, pp. 198–204.

[10] J. G. Andrews, W. Choi, and R. W. Heath, "Overcoming interference in spatial multiplexing MIMO cellular networks," *IEEE Wireless Commun.*, vol. 14, no. 6, pp. 95–104, 2007.

[11] J. Kaleva, A. Tölli, and M. Venkatraman, G.and Juntti, "Downlink precoder design for coordinated regenerative multi-user relaying," *IEEE trans. on signal processing*, vol. 61, no. 5, pp. 1215–1229, Mar. 2013.

[12] P. Jayasinghe, A. Tölli, J. Kaleva, and M. Latva-aho, "Bi-directional beamformer training for dynamic TDD networks," *IEEE Trans on Signal Processing*, vol. 66, no. 23, pp. 6252–6267, 2018.

[13] G. Venkatraman, A. Tölli, M. Juntti, and L. N. Tran, "Traffic aware resource allocation schemes for multi-cell MIMO-OFDM systems," *IEEE Trans. Signal Processing*, vol. 64, no. 11, pp. 2730–2745, Jun. 2016.

[14] J. Kaleva, A. Tölli, and M. Juntti, "Decentralized sum rate maximization with QoS constraints for interfering broadcast channel via successive convex approximation," *IEEE Trans. Signal Processing*, vol. 64, no. 11, pp. 2788–2802, Jun. 2016.