



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Joint Power Allocation and Beamforming for In-band Full-duplex Multi-cell Multi-user Networks

Citation for published version:

Luo, H, Garg, N, Holme, M & Ratnarajah, T 2024, 'Joint Power Allocation and Beamforming for In-band Full-duplex Multi-cell Multi-user Networks', *IEEE Transactions on Vehicular Technology*, pp. 1-15.
<<http://10.1109/TVT.2024.3400460>>

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

IEEE Transactions on Vehicular Technology

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Joint Power Allocation and Beamforming for In-band Full-duplex Multi-cell Multi-user Networks

Haifeng Luo, *Member, IEEE*, Navneet Garg, *Member, IEEE*, Mark Holm, and Tharmalingam Ratnarajah, *Senior Member, IEEE*

Abstract—This paper investigates a robust joint power allocation and beamforming scheme for in-band full-duplex multi-cell multi-user (IBFD-MCMU) networks. A mean-squared error (MSE) minimization problem is formulated with constraints on the power budgets and residual self-interference (RSI) power. The problem is not convex, so we decompose it into two sub-problems: interference management beamforming and power allocation, and give closed-form solutions to the sub-problems. Then we propose an iterative algorithm to yield an overall solution. The computational complexity and convergence behavior of the algorithm are analyzed. Our method can enhance the analog self-interference cancellation (ASIC) depth provided by the precoder with less effect on the downlink communication than the existing null-space projection method, inspiring a low-cost but efficient IBFD transceiver design. It can achieve 42.9% of IBFD gain in terms of spectral efficiency with only antenna isolation, while this value increases to 60.9% with further digital self-interference cancellation (DSIC). Numerical results illustrate that our algorithm is robust to hardware impairments and channel uncertainty. With sufficient ASIC depth, our method reduces the computation time by at least 20% than the existing scheme due to its faster convergence speed at the cost of < 12.5% sum rate loss. The benefit is much more significant with single-antenna users that our algorithm saves at least 40% of the computation time at the cost of < 10% sum rate reduction.

Index Terms—analog self-interference cancellation, in-band full-duplex, joint power allocation and beamforming, multi-cell multi-user

I. INTRODUCTION

IN-band full-duplex (IBFD) is a promising candidate in beyond fifth-generation (5G) systems due to its ability to enhance spectral efficiency (SE) and reduce the end-to-end transmission delay compared to conventional half-duplex (HD) systems since the bandwidth is used simultaneously by uplink and downlink transmissions. However, IBFD radios introduce additional interferences, i.e., self-interference (SI) and co-channel interference (CCI), degrading the promising gain of IBFD, especially the significant SI due to the proximity of the transceiver [1]. In IBFD multi-cell multi-user (MCMU) networks, downlink users not only get interference from base stations as in traditional HD radios but also get interference from uplink users. Meanwhile, the uplink communication is interfered by downlink transmissions. The complex interference must be appropriately processed to achieve substantial system throughput.

It is revealed that, in multi-input and multi-output (MIMO) systems, the interference is caused by the non-zero inner product between the channel and transmitted signal matrices. Therefore, beamforming can be leveraged to cancel the interference effectively. The readers are referred to [2] for details of implementing interference cancellation by adjusting the weights on different antennas from the perspective of signal processing. Interference cancellation in IBFD-MCMU networks is challenging due to the contradiction between a large amount of interference and the limited degrees of freedom imposed by finite numbers of antennas. Nevertheless, an appropriate beamforming design could be leveraged to effectively utilize the limited degrees of freedom and maximize spectral efficiency [3]. Authors of [4] minimize the sum of mean-squared errors (MSE) of an IBFD multi-user network through beamforming, and authors of [5] maximize the sum rate of such a network by optimizing beamforming matrices, respectively. The results illustrate that appropriate beamforming schemes effectively manage the interference and improve the system capacity. IBFD-MCMU scenarios are further studied in [6] and a maximum weighted sum rate (MWSR) beamforming design with power constraints is formulated and solved by exploring the relationship between weighted minimum mean-squared error (WMMSE) and MWSR. The MWSR algorithm yields a joint power allocation and beamforming (JPABF) scheme that effectively manages the complex interference and maximizes the system throughput in IBFD-MCMU networks. Simulations demonstrate the spectral efficiency improvement of IBFD over HD with this design. However, existing studies usually assume self-interference cancellation (SIC) has been realized by other techniques rather than beamforming, which could be challenging in practice.

Indeed, various SIC techniques are proposed to suppress the SI in the propagation, analog, and digital domains. Studies demonstrate the effectiveness of these techniques by simulations and experiments as reported in the literature (see [1], [7] and references therein). However, enabling effective SIC in multi-input and multi-output (MIMO) systems is still challenging due to the exponentially increased cost and complexity with increasing antennas [8], [9]. For analog self-interference cancellation (ASIC), a total of $N_t N_r$ analog cancellers will be needed for an IBFD transceiver with N_t transmitting antennas and N_r receiving antennas. Although this number could be reduced to N_r by using auxiliary transmitters, it is still a large number, and additional transmitter chains are costly [10]. For nonlinear digital self-interference cancellation (DSIC), the parameters that need to be calculated become unacceptably

H. Luo, N. Garg, and T. Ratnarajah are with Institute for Imaging, Data and Communications (IDCOM), School of Engineering, The University of Edinburgh, UK, e-mails: {s1895225, navneet.garg, t.ratnarajah}@ed.ac.uk

M. Holm is with Huawei Technologies (Sweden) AB, Gothenburg, Sweden, e-mail: mark.holm@huawei.com

high with high-order MIMO systems. It increases almost by the cube of N_t with the nonlinearity order of 3 [11]. Although the principal component analysis is leveraged to reduce the parameters by 65% in [11], the computational complexity could still be unacceptable with massive MIMO (e.g., > 128 antennas). Therefore, it is desired to explore beamforming for SIC since it does not require additional processing circuits nor increases the parameters. Existing studies in this direction mainly use the null-space projection (NSP) [12]–[15] or choose orthogonal transmitting and receiving beamforming matrices [12]. The NSP-based methods steer the transmitted beams to the direction orthogonal to the SI channel, distorting the original precoders (i.e., transmitting beamformers¹) and introducing precoding errors. The trade-off between the precoding errors and SIC capability of the NSP-based methods is analyzed in [15]. It illustrates that achieving effective SIC may severely sacrifice the downlink capacity. Choosing orthogonal transmitting and receiving beamformers from the eigenvectors of the intended channel is a good approach to realizing both efficient SIC and effective downlink communications. However, it does not consider the limited dynamic range of practical receivers, which is an obstacle to all-digital SIC schemes [10]. So it is not applicable to practical systems.

In addition to the beamforming, an appropriate power allocation policy can also manage the interference since the strength of interference depends on the transmit power of associated nodes [16]. Power allocation could be implemented solely based on the channel strength in single-antenna systems as in [17]. However, in MIMO systems, power allocation and beamforming should be crossly designed for optimal performance since they will affect each other. Studies illustrate that an appropriate JPABF scheme can significantly improve the system capacity in MIMO systems. A sum rate maximization problem with respect to joint power allocation and beamforming is formulated for a single-cell 2-user network in [18]. It is solved using an iterative algorithm, and simulation results show that it achieves close-to-bound sum rate performance. In addition to the system throughput, a JPABF scheme can also be designed to maximize the energy efficiency [19], [20]. The main challenge of the JPABF design is that the joint optimization problem is non-convex, so it is difficult to derive closed-form solutions [19], [21]. Alternatively, it is solved by an iterative algorithm, causing high computational complexity.

The JPABF has not yet been intensively studied for IBFD-MCMU networks, where a large amount of CCI and significant SI pose even greater challenges. To the best knowledge of the authors, existing studies in this direction mainly use the MWSR algorithm proposed in [6]. The MWSR can indeed achieve significant spectral efficiency; however, there are two disadvantages: 1) it requires sufficient ASIC to be realized by other techniques rather than precoding, which is challenging in practical MIMO systems; 2) the overall computational complexity could be high due to the slow convergence speed. In this paper, we propose a joint power allocation and in-

terference management (JPAIM) algorithm for IBFD-MCMU networks, where the power allocation and beamforming are jointly optimized to minimize the sum of MSE of the network under constraints on the transmit power budget and residual self-interference (RSI). We do not assume a perfect SIC while we investigate the SIC capability of beamformers. Besides, we include practical imperfections (e.g., CSI errors, hardware impairments, and the limited dynamic range of receivers) to derive a robust solution. Compared to existing studies, our algorithm can achieve considerable IBFD gain with insufficient ASIC and reduce computational complexity. Our novel contributions can be summarised as follows.

- Robustness to channel uncertainty: we include the effects of imperfect CSI by introducing channel uncertainty and derive a robust design. Simulation results show that our algorithm is robust to imperfect CSI as large channel uncertainty has a smaller impact on our method than on existing methods.
- Beamforming cancellation: we exploit beamforming for SIC and reveal that precoders can be leveraged to suppress the SI in the propagation domain, preventing receiver saturation. Thus, we enhance the ASIC capability of precoders to remove or reduce the requirements of RF cancellers, which have high implementation complexity and energy consumption. Different from existing NSP-based methods [13]–[15], we realize it by adding a constraint on the received RSI power, yielding a joint design. Numerical results show the advantage of our design over existing methods that it can enhance the ASIC capability of precoders with less effect on the downlink communication. Thus, our method can enable a low-cost but efficient IBFD transceiver design.
- A new JPABF formulation: we use scalar coefficients to reflect the power allocation policy and formulate a minimization problem with respect to the coefficients and beamforming matrices. Since the formulated problem is not convex and may not be converted into a convex form by simple manipulations, we propose a two-stage method to obtain the solution. The sub-problems in each stage are convex, and closed-form solutions are given. Then we propose an iterative algorithm to obtain the overall solution as the optimized variables show interdependence on each other. We analyze the computational complexity and convergence behavior of the algorithm. Simulation results demonstrate that our algorithm significantly reduces the time complexity compared to the existing MWSR algorithm in [6] at the cost of acceptable sum rate loss due to the scalar power coefficients and decomposition. The benefit is significant with single-antenna users, which is meaningful for practical cellular network deployment.

The rest of the paper is organized as follows: In Section II, we give the system model of the IBFD-MCMU network with transceiver hardware impairments and channel uncertainty. Then, we formulate the joint optimization problem in Section III and solve it by decomposing it into two sub-problems. The solutions of the two sub-problems are derived in Section

¹In the MIMO context, transmitting beamforming can be considered a form of precoding, especially when it comes to optimizing the signal considering the spatial aspect of the channel. Therefore, in our study, we generally use precoding to refer to transmitting beamforming.

IV, and the iterative algorithm to obtain the overall solution is given and analyzed. In Section V, we evaluate the performance of the proposed design by simulations. Finally, conclusions are drawn in Section VI.

Notations: \mathbf{A} , \mathbf{a} , and a represent a matrix, a vector, and a scalar, respectively. $\text{tr}(\mathbf{A})$, $\text{Cov}(\mathbf{A})$, $|\mathbf{A}|$, $\|\mathbf{A}\|$, \mathbf{A}^H , \mathbf{A}^T , and \mathbf{A}^{-1} denote the trace, covariance matrix, determinant, 2-norm, Hermitian, transpose, and inverse of matrix \mathbf{A} . \mathbf{I}_k represents an identity matrix with k elements along its diagonal. $\mathcal{D}(\mathbf{A})$ denotes a diagonal matrix containing the elements along the diagonal of \mathbf{A} . $[\mathbf{A}]_{i,j}$ represents the element of the i^{th} row and j^{th} column of matrix \mathbf{A} . $\mathbf{A}(:,j)$ denotes the j^{th} column of matrix \mathbf{A} . $\mathbb{E}\{\cdot\}$ denotes the expectation operation. $\max\{\cdot\}$ and $\min\{\cdot\}$ denote the maximum and minimum element of the set. $\mathcal{R}\{\cdot\}$ represents the real part of complex numbers. $\mathcal{CN}(0, \sigma^2)$ denotes a complex normal distribution with zero mean and variance of σ^2 .

II. SYSTEM MODEL

We consider a G -cell network, where the base station in the g^{th} cell serves K_g^d downlink (DL) users and K_g^u uplink (UL) users. Assume all base stations (BS) have N_{bs} transmitting antennas and M_{bs} receiving antennas, while UL user equipment (UE) has N_{ue} transmitting antennas and DL UE has M_{ue} receiving antennas. Fig. 1 shows the interference between the nodes in this IBFD-MCMU network.

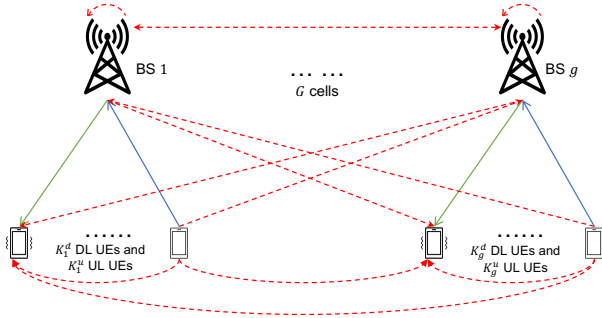


Fig. 1. Interference between nodes in an IBFD multi-cell multi-user network.

A. Transmitted Signals

Let $\mathbf{s}_{k_g^d} \in \mathbb{C}^{b_d \times 1}$ represent the data symbols intended to the k^{th} downlink user in the g^{th} cell with statistics $\mathbb{E}\{\mathbf{s}_{k_g^d} \mathbf{s}_{k_g^d}^H\} = \mathbf{I}_{b_d}$; $\mathbf{V}_{k_g^d} \in \mathbb{C}^{N_{bs} \times b_d}$ denotes the associated precoding matrix; and $\alpha_{k_g^d}$ is the power coefficient that reflects the transmit power allocated to the corresponding user. The signal transmitted at the g^{th} BS can be denoted as

$$\mathbf{x}_g = \sum_{k=1}^{K_g^d} \left(\alpha_{k_g^d} \mathbf{V}_{k_g^d} \mathbf{s}_{k_g^d} + \mathbf{c}_{k_g^d} \right), \quad (1)$$

where $\mathbf{c}_{k_g^d}$ represents the hardware impairments due to the limited dynamic range of practical transmitters. The limited dynamic range is a natural consequence of imperfect digital-to-analog converters (DACs), oscillators, and power amplifiers (PAs). Experimental measurements demonstrate that

the transmitter hardware impairments are independent of the transmitted signals and can be closely described by the circular complex Gaussian model as [22]

$$\mathbf{c}_g \sim \mathcal{CN}(\mathbf{0}, \kappa_{bs} \mathbb{E}\{\mathcal{D}(\mathbf{V}_g \mathbf{A}_g \mathbf{A}_g^H \mathbf{V}_g^H)\}), \quad (2)$$

where $\kappa_{bs} \ll 1$ characterizes the dynamic range of the transmitters at base stations. The base station transmits an accumulation of signals for DL users in the same cell. Let $\mathbf{s}_g = [\mathbf{s}_{1_g^d}^T, \mathbf{s}_{2_g^d}^T, \dots, \mathbf{s}_{K_g^d}^T]^T$, $\mathbf{A}_g = \mathcal{D}(\alpha_{1_g^d} \mathbf{I}_{b_d}, \alpha_{2_g^d} \mathbf{I}_{b_d}, \dots, \alpha_{K_g^d} \mathbf{I}_{b_d})$, and $\mathbf{V}_g = [\mathbf{V}_{1_g^d}, \mathbf{V}_{2_g^d}, \dots, \mathbf{V}_{K_g^d}]$, the accumulated transmitted signals can be denoted as $\tilde{\mathbf{x}}_g = \mathbf{V}_g \mathbf{A}_g \mathbf{s}_g$. The transmit power of the g^{th} base station is given as

$$P_g = \mathbb{E}\{\text{tr}(\tilde{\mathbf{x}}_g \tilde{\mathbf{x}}_g^H)\} = \text{tr}(\mathbb{E}\{\mathbf{V}_g \mathbf{A}_g \mathbf{s}_g \mathbf{s}_g^H \mathbf{A}_g^H \mathbf{V}_g^H\}) \\ = \sum_{k=1}^{K_g^d} \alpha_{k_g^d}^2 \text{tr}(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H). \quad (3)$$

Similarly, let $\mathbf{s}_{k_g^u} \in \mathbb{C}^{b_u \times 1}$ denote the uplink payload symbols with statistics $\mathbb{E}\{\mathbf{s}_{k_g^u} \mathbf{s}_{k_g^u}^H\} = \mathbf{I}_{b_u}$; $\mathbf{V}_{k_g^u}$ represents the associated precoding matrix; $\gamma_{k_g^u}$ represents the power coefficient that reflects the allocated transmit power; and $\kappa_{ue} \ll 1$ characterizes the dynamic range of the transmitters at the user equipment. The transmitted signal of the k^{th} uplink user in the g^{th} cell (i.e., k_g^u) can be denoted as

$$\mathbf{x}_{k_g^u} = \underbrace{\gamma_{k_g^u} \mathbf{V}_{k_g^u} \mathbf{s}_{k_g^u}}_{\tilde{\mathbf{x}}_{k_g^u}} + \mathbf{c}_{k_g^u}, \quad (4)$$

where $\mathbf{c}_{k_g^u}$ denotes the transmitter hardware impairments at this uplink user that can be described as

$$\mathbf{c}_{k_g^u} \sim \mathcal{CN}(\mathbf{0}, \kappa_{ue} \gamma_{k_g^u}^2 \mathbb{E}\{\mathcal{D}(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H)\}). \quad (5)$$

The transmit power of the k^{th} UE in the g^{th} cell is given as

$$P_{k_g^u} = \mathbb{E}\{\text{tr}(\tilde{\mathbf{x}}_{k_g^u} \tilde{\mathbf{x}}_{k_g^u}^H)\} = \gamma_{k_g^u}^2 \text{tr}(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H). \quad (6)$$

B. Received Signals

Let $\mathbf{H}_{B,A}$ represent the coefficients matrix of the wireless MIMO channel from node A to node B throughout this paper, the signal received by the g^{th} BS can be denoted as

$$\mathbf{y}_g = \underbrace{\mathbf{H}_{g,k_g^u} \mathbf{x}_{k_g^u} + \sum_{(i,j) \neq (k,g)} \sum_{i=1}^G \sum_{j=1}^{K_i^u} \mathbf{H}_{g,i_j^u} \mathbf{x}_{i_j^u} + \mathbf{H}_{g,g} \mathbf{x}_g + \sum_{j \neq g} \mathbf{H}_{g,j} \mathbf{x}_j}_{\tilde{\mathbf{y}}_g} \\ + \mathbf{n}_g + \mathbf{d}_g, \quad (7)$$

where the third and fourth terms at the right hand represent the additional interference due to IBFD operation; \mathbf{n}_g denotes the additive white Gaussian noise (AWGN) of the receiver such that $\mathbf{n}_g \sim \mathcal{CN}(\mathbf{0}, \sigma_g^2 \mathbf{I})$; and \mathbf{d}_g denotes the hardware impairments due to the limited dynamic range of practical receivers. The limited dynamic range is a natural consequence of imperfect low-noise amplifiers (LNAs), oscillators, and

analog-to-digital converters (ADCs). Experimental measurements demonstrate that the receiver hardware impairments are independent of the received signals and can be closely described by the circular complex Gaussian model as [22]

$$\mathbf{d}_g \sim \mathcal{CN}(\mathbf{0}, \beta_{bs} \mathbb{E}\{\mathcal{D}(\tilde{\mathbf{y}}_g \tilde{\mathbf{y}}_g^H)\}), \quad (8)$$

where $\beta_{bs} \ll 1$ characterizes the dynamic range of the receivers at base stations. Similarly, the signal received by the k^{th} downlink UE in the g^{th} cell can be denoted as

$$\begin{aligned} \mathbf{y}_{k_g^d} &= \mathbf{H}_{k_g^d, g} \mathbf{x}_{k_g^d} + \underbrace{\sum_{(i,j) \neq (k,g)} \sum_{i=1}^G \sum_{j=1}^{K_j^d} \mathbf{H}_{k_g^d, j} \mathbf{x}_{i_j^d}}_{\tilde{\mathbf{y}}_{k_g^d}} + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \mathbf{H}_{k_g^d, i_j^u} \mathbf{x}_{i_j^u} \\ &+ \mathbf{n}_{k_g^d} + \mathbf{d}_{k_g^d}, \end{aligned} \quad (9)$$

where the third term at the right hand represent the additional interference due to IBFD operation; $\mathbf{n}_{k_g^d}$ denotes the AWGN of the receiver such that $\mathbf{n}_{k_g^d} \sim \mathcal{CN}(\mathbf{0}, \sigma_{k_g^d}^2 \mathbf{I})$; and $\mathbf{d}_{k_g^d}$ denotes the hardware impairments of the receiver described as

$$\mathbf{d}_{k_g^d} \sim \mathcal{CN}(\mathbf{0}, \beta_{ue} \mathbb{E}\{\mathcal{D}(\tilde{\mathbf{y}}_{k_g^d} \tilde{\mathbf{y}}_{k_g^d}^H)\}), \quad (10)$$

where $\beta_{ue} \ll 1$ characterizes the dynamic range of the receivers at the user equipment.

Note: The values of κ_{bs} , κ_{ue} , β_{bs} , and β_{ue} are related to the measurable error vector magnitudes (EVMs) of corresponding RF transceivers. The HWIs model utilized is a verified model based on experiments [22] and has been adopted by many studies in the field of wireless communications (see [6] and references therein). Due to the power consumption problem in MIMO systems, the transceivers tend to use low-resolution DACs/ADCs [26], [30]. So we assume that the dynamic range of receivers is mainly limited by DACs/ADCs, which can be described by the additive quantization noise model (AQNM) given in [26].

C. Channel Uncertainty

Accurate channel estimation is challenging in practice due to limited training resources, resulting in channel uncertainty in the obtained channel state information (CSI). Let $\hat{\mathbf{H}}_{B,A}$ denote the estimate associated with the actual wireless channel $\mathbf{H}_{B,A}$, they are related as [24]

$$\mathbf{H}_{B,A} = \hat{\mathbf{H}}_{B,A} + \mathbf{\Delta}_{B,A}, \quad (11)$$

where $\mathbf{\Delta}_{B,A}$ denotes the channel uncertainty (i.e., estimation errors). In this paper, we adapt the stochastic error model, which describes the channel uncertainty as $\mathbf{\Delta}_{B,A} \sim \mathcal{CN}(\mathbf{0}, \hat{\sigma}_{B,A}^2 \mathbf{I})$ [6], [24]. Using the accessible estimated CSI with the statistical channel uncertainty, the received signals given in Equations (7) and (9) can be written as

$$\begin{aligned} \mathbf{y}_g &= \hat{\mathbf{H}}_{g, k_g^u} \mathbf{x}_{k_g^u} + \sum_{i,j \neq k, g} \sum_{i=1}^G \sum_{j=1}^{K_j^u} \hat{\mathbf{H}}_{g, i_j^u} \mathbf{x}_{i_j^u} + \mathbf{H}_{g, g} \mathbf{x}_g \\ &+ \sum_{j \neq g} \hat{\mathbf{H}}_{g, j} \mathbf{x}_j + \mathbf{n}_g + \mathbf{d}_g + \mathbf{e}_g, \end{aligned} \quad (12)$$

$$\begin{aligned} \mathbf{y}_{k_g^d} &= \hat{\mathbf{H}}_{k_g^d, g} \mathbf{x}_{k_g^d} + \sum_{i,j \neq k, g} \sum_{i=1}^G \sum_{j=1}^{K_j^d} \hat{\mathbf{H}}_{k_g^d, j} \mathbf{x}_{i_j^d} + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \hat{\mathbf{H}}_{k_g^d, i_j^u} \mathbf{x}_{i_j^u} \\ &+ \mathbf{n}_{k_g^d} + \mathbf{d}_{k_g^d} + \mathbf{e}_{k_g^d}, \end{aligned} \quad (13)$$

where \mathbf{e}_g and $\mathbf{e}_{k_g^d}$ represent the noise due to imperfect CSI with statistics that $\mathbf{e}_g \sim \mathcal{CN}(\mathbf{0}, \hat{\sigma}_g^2 \mathbf{I}_{M_{bs}})$ and $\mathbf{e}_{k_g^d} \sim \mathcal{CN}(\mathbf{0}, \hat{\sigma}_{k_g^d}^2 \mathbf{I}_{M_{ue}})$; $\hat{\sigma}_g^2$ and $\hat{\sigma}_{k_g^d}^2$ are given as

$$\hat{\sigma}_g^2 = \sum_{j=1}^G \sum_{i=1}^{K_j^u} \hat{\sigma}_{g, i_j^u}^2 \text{tr}(\mathbf{T}_{i_j^u}) + \sum_{j \neq g} \hat{\sigma}_{g, j}^2 \text{tr}(\mathbf{T}_j),$$

$$\hat{\sigma}_{k_g^d}^2 = \sum_{j=1}^G \sum_{k=1}^{K_j^d} \hat{\sigma}_{k_g^d, j}^2 \text{tr}(\mathbf{T}_{i_j^d}) + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \hat{\sigma}_{k_g^d, i_j^u}^2 \text{tr}(\mathbf{T}_{i_j^u}),$$

as detailed in Appendix A. We use the perfect CSI of SI channels since it can be estimated with high SINR [6].

III. PROBLEM FORMULATION

The power allocation and beamforming should have a common objective for the joint design. We chose the mean-squared error (MSE) minimization as the design objective, which is widely used in wireless networks [4], [8].

A. Mean-Squared Errors

Using the simplicities given in Appendix C, the MSE of the k^{th} downlink user and uplink user in the g^{th} cell can be given as Equations (14) and (15) on the top of next page.

B. Beamforming for Self-interference Cancellation

We consider all-digital beamforming throughout this paper. The received RSI at the g^{th} base station is given as

$$\mathbf{y}_{g, si} = \mathbf{H}_{g, g} \sum_{k=1}^{K_g^d} \left(\alpha_{k_g^d} \mathbf{V}_{k_g^d} \mathbf{s}_{k_g^d} + \mathbf{c}_{k_g^d} \right), \quad (16)$$

which gives the power of the RSI as $\text{tr}(\mathbf{y}_{g, si} \mathbf{y}_{g, si}^H)$. The RSI may exceed the dynamic range of receivers without active ASIC. Equation (16) indicates that the precoder can be leveraged to steer the beams to the desired direction in the propagation domain to suppress the SI as

$$\begin{aligned} \min_{\mathbf{V}_{k_g^d}} \mathbb{E} \left\{ \|\mathbf{y}_{g, si}\|^2 \right\} &= \min_{\mathbf{V}_{k_g^d}} \alpha_{k_g^d}^2 \left(\text{tr} \left(\mathbf{H}_{g, g} \mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \mathbf{H}_{g, g}^H \right) \right. \\ &\quad \left. + \kappa_{bs} \text{tr} \left(\mathbf{H}_{g, g} \mathcal{D} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right) \mathbf{H}_{g, g} \right) \right). \end{aligned} \quad (17)$$

It should be noted that the precoder suppresses the analog SI before the receiver to prevent receiver saturation. Thus the precoder-based cancellation is regarded as an ASIC technique, although it is implemented by digital signal processing. The receiving beamformers can also minimize the RSI as

$$\min_{\mathbf{U}_g} \mathbb{E} \left\{ \|\mathbf{U}_g^H \mathbf{y}_{g, si}\|^2 \right\}. \quad (18)$$

$$\begin{aligned}
\varepsilon_{k_g^d}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) &= \mathbb{E} \left\{ \left\| \mathbf{s}_{k_g^d} - \mathbf{U}_{k_g^d}^H \mathbf{y}_{k_g^d} \right\|^2 \right\} = \text{tr} \left(\mathbf{U}_{k_g^d}^H \left(\sum_{j=1}^G \sum_{i=1}^{K_j^d} \hat{\mathbf{H}}_{k_g^d, j} \mathbf{T}_{i_j^d} \hat{\mathbf{H}}_{k_g^d, j}^H + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \hat{\mathbf{H}}_{k_g^d, i_j^u} \mathbf{T}_{i_j^u} \hat{\mathbf{H}}_{k_g^d, i_j^u}^H \right) \mathbf{U}_{k_g^d} \right) \\
&+ \sigma_r^2 \text{tr} \left(\mathbf{U}_{k_g^d}^H \mathcal{D} \left(\sum_{j=1}^G \sum_{i=1}^{K_j^d} \hat{\mathbf{H}}_{k_g^d, j} \mathbf{T}_{i_j^d} \hat{\mathbf{H}}_{k_g^d, j}^H + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \hat{\mathbf{H}}_{k_g^d, i_j^u} \mathbf{T}_{i_j^u} \hat{\mathbf{H}}_{k_g^d, i_j^u}^H \right) \mathbf{U}_{k_g^d} \right) \\
&+ \sigma_{k_g^d}^2 \text{tr} \left(\mathbf{U}_{k_g^d}^H \mathbf{U}_{k_g^d} \right) - 2\alpha_{k_g^d} \mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^d}^H \hat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d} \right) \right\} + b_d,
\end{aligned} \tag{14}$$

$$\begin{aligned}
\varepsilon_{k_g^u}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) &= \mathbb{E} \left\{ \left\| \mathbf{s}_{k_g^u} - \mathbf{U}_{k_g^u}^H \mathbf{y}_{k_g^u} \right\|^2 \right\} = \text{tr} \left(\mathbf{U}_{k_g^u}^H \left(\sum_{j=1}^G \sum_{i=1}^{K_j^u} \hat{\mathbf{H}}_{g, i_j^u} \mathbf{T}_{i_j^u} \hat{\mathbf{H}}_{g, i_j^u}^H + \sum_{j=1}^G \hat{\mathbf{H}}_{g, j} \mathbf{T}_j \hat{\mathbf{H}}_{g, j}^H \right) \mathbf{U}_{k_g^u} \right) \\
&+ \sigma_r^2 \text{tr} \left(\mathbf{U}_{k_g^u}^H \mathcal{D} \left(\sum_{j=1}^G \sum_{i=1}^{K_j^u} \hat{\mathbf{H}}_{g, i_j^u} \mathbf{T}_{i_j^u} \hat{\mathbf{H}}_{g, i_j^u}^H + \sum_{j=1}^G \hat{\mathbf{H}}_{g, j} \mathbf{T}_j \hat{\mathbf{H}}_{g, j}^H \right) \mathbf{U}_{k_g^u} \right) \\
&+ \sigma_g^2 \text{tr} \left(\mathbf{U}_{k_g^u}^H \mathbf{U}_{k_g^u} \right) - 2\gamma_{k_g^u} \mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^u}^H \hat{\mathbf{H}}_{g, k_g^u} \mathbf{V}_{k_g^u} \right) \right\} + b_u.
\end{aligned} \tag{15}$$

This process is taken after the ADCs of receivers, so it is regarded as a DSIC technique and does not help prevent receiver saturation. As previous studies illustrate, sufficient ASIC is critical to any effective SIC scheme for practical implementation, while active ASCI has high complexity and cost [?], [25]. Therefore, it is desired to enhance the ASIC capability of precoders to implement a low-cost but effective IBFD transceiver. The RSI can be eliminated if $\mathbf{V}_{k_g^d}$ spans of vectors from the null-space of $\mathbf{H}_{g, g} \mathbf{H}_{g, g}^H + \kappa_{bs} \mathcal{D}(\mathbf{H}_{g, g} \mathbf{H}_{g, g}^H)$. Existing NSP-based methods follow this idea [12]–[15], but they did not consider the downlink communication while doing the projection nor the transmitter distortion. Thus, the NSP may significantly reduce the downlink rate, and its performance is limited to transmitter distortions. Choosing orthogonal transmitting and receiving beamforming matrices as [12] can minimize $\mathbb{E} \left\{ \left\| \mathbf{U}_g^H \mathbf{y}_{g, \text{si}} \right\|^2 \right\}$, but it cannot minimize the RSI power before ADCs, i.e., $\mathbb{E} \left\{ \left\| \mathbf{y}_{g, \text{si}} \right\|^2 \right\}$. Thus, its performance is compromised by the limited dynamic range. Let l_g denote the pathloss of the effective SI channel at the g^{th} base station, which includes the effects of other ASIC techniques (e.g., antenna isolation or RF cancellers), the ASIC capability of precoders (i.e., the ASIC depth provided by precoders) can be described in dB as

$$\text{ASIC}_{g, \text{pre}} = 10 \times \log_{10} \frac{l_g \mathbb{E} \left\{ \text{tr} \left(\mathbf{x}_g \mathbf{x}_g^H \right) \right\}}{\mathbb{E} \left\{ \text{tr} \left(\mathbf{y}_{g, \text{si}} \mathbf{y}_{g, \text{si}}^H \right) \right\}}, \tag{19}$$

where the numerator describes the received power without precoding, and the denominator is the RSI power with precoding; \mathbf{x}_g and $\mathbf{y}_{g, \text{si}}$ are given in (1) and (16), respectively.

C. Constrained Optimization Problem

Minimizing the sum of MSE by optimizing precoders will include the effects of combiners, as the MSE expressions

suggest. As a result, precoders may not provide a sufficient ASIC depth to prevent receiver saturation if combiners have effectively suppressed the RSI and distortions caused by it. However, the degrees of freedom of combiners occupied by SIC and uncorrelated distortions due to receiver saturation will naturally compromise the uplink communication. Thus, the existing beamforming schemes (e.g., MWSR in [6], and other schemes in [5], [8], [20]) for IBFD networks will require sufficient ASIC realized by other techniques to guarantee their performance. In contrast, we would like the precoder to provide a sufficient ASIC depth to reduce the RSI power, which can be done by adding the constraint on the RSI power, formulating the problem as

$$\begin{aligned}
(P.1) \quad & \min_{\{\alpha, \gamma, \mathbf{V}, \mathbf{U}\}_{\forall k, g}} \sum_{g=1}^G \sum_{k=1}^{K_g^d} \varepsilon_{k_g^d}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) \\
& + \sum_{g=1}^G \sum_{k=1}^{K_g^u} \varepsilon_{k_g^u}(\alpha, \gamma, \mathbf{V}, \mathbf{U})
\end{aligned} \tag{20}$$

$$\text{s.t. } (C.1) \quad \sum_{k=1}^{K_g^d} \alpha_{k_g^d}^2 \text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right) \leq P_{bs}, \quad \forall g \tag{21}$$

$$(C.2) \quad \gamma_{k_g^u}^2 \text{tr} \left(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H \right) \leq P_{ue}, \quad \forall k, g \tag{22}$$

$$(C.3) \quad \epsilon_{\text{rsi}, g}(\mathbf{V}) \leq \bar{\epsilon}_{\text{rsi}, g}, \tag{23}$$

where $\epsilon_{\text{rsi}, g}(\mathbf{V}) = \mathbb{E} \left\{ \text{tr} \left(\mathbf{y}_{g, \text{si}} \mathbf{y}_{g, \text{si}}^H \right) \right\}$ represents the received analog RSI power; $\bar{\epsilon}_{\text{rsi}, g}$ is the tolerable RSI power that could be chosen according to the dynamic range of receivers. We want to suppress the SI solely by precoding schemes instead of reducing the transmit power of the downlink payload. Thus, the RSI is a function of only precoders, as denoted in the formulation. According to the weighted sum rate and WMMSE relationship in [6], minimizing the sum MSE under

the RSI power constraint (i.e., problem (P.1)) can maximize the sum rate of the network with a constrained RSI power (i.e., enhanced ASIC depth to prevent receiver saturation).

Lemma 1. *The optimization problem (P.1) is equivalent to the problem (P.2) shown below with appropriate ν_g .*

$$(P.2) \quad \min_{\{\alpha, \gamma, \mathbf{V}, \mathbf{U}\}_{\forall k, g}} \sum_{g=1}^G \sum_{k=1}^{K_g^d} \varepsilon_{k_g^d}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) \\ + \sum_{g=1}^G \sum_{k=1}^{K_g^u} \varepsilon_{k_g^u}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) + \sum_{g=1}^G \nu_g \epsilon_{rsi, g}(\mathbf{V}) \quad (24)$$

$$s.t. \quad (C.1) \quad \sum_{k=1}^{K_g^d} \alpha_{k_g^d}^2 \text{tr}(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H) \leq P_{bs}, \quad \forall g \quad (25)$$

$$(C.2) \quad \gamma_{k_g^u}^2 \text{tr}(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H) \leq P_{ue}, \quad \forall k, g, \quad (26)$$

Proof. See Appendix B. \square

According to Lemma 1, the optimal power allocation and beamforming schemes can be obtained from solving problem (P.2). Still, an appropriate value of ν_g needs to be chosen at first to formulate the problem. As stated in Appendix B, ν_g should be equal to $\varpi_{rsi, g}$ to make the problems equivalent, while $\varpi_{rsi, g}$ depends on the chosen tolerable RSI power $\bar{\epsilon}_{rsi, g}$. It is challenging to mathematically derive the value of ν_g from given $\bar{\epsilon}_{rsi, g}$ since the multiple variables are entangled. However, we can directly give the value of ν_g , and it will reflect a specific tolerable RSI power. The resulting solutions will maximize the sum rate under the corresponding RSI power constraint. We choose ν_g according to its properties. As (24) suggests, ν_g can adjust the preference of the precoder: the objective function will be dominated by the sum of MSE when ν_g is relatively small, and conversely, the objective function will be dominated by the RSI when ν_g is rather large. As a result, precoders will tend to minimize the sum of MSE or RSI accordingly. Therefore, we can set a relatively large ν_g to help suppress SI before the receiver if there is not sufficient ASIC depth and set a relatively small ν_g to minimize the precoding errors if there is sufficient ASIC depth. According to experiments, we can set the value of ν_g based on the realized ASIC depth l_g as $\nu_g = 10^{-l_g/5}$.

From the MSE expressions, we can see that the main difficulty of the minimization problem (P.2) is that the optimized variables are entangled, leading to a non-convex problem. The non-convex constrained optimization problem may not be converted to a convex form with simple manipulations. Thus, we propose a two-stage approach to solve it, where the original joint optimization problem is decomposed into two convex sub-problems: interference management through beamforming and power allocation. The two sub-problems are relatively easy to be solved through standard solutions, but the decomposition leads to a suboptimal solution to the joint optimization problem.

IV. JOINT POWER ALLOCATION AND INTERFERENCE MANAGEMENT ALGORITHM

In the interference management stage, we will fix the power coefficients to form the first sub-problem regarding beamforming matrix optimization. However, the objective function in (24) is still not jointly convex to the transmitting and receiving beamforming matrices (i.e., precoders and combiners). Thus, we further divide this sub-problem into combiner updating and precoder updating stages.

A. Receiving Beamforming

With fixed precoding matrices and power coefficients, the combining matrices are optimized to minimize the sum of MSE. The analog RSI power is not affected by the combiners, so it is removed from the objective function, and the combiner optimization problem is formulated as

$$(S.1.1) \quad \min_{\{\mathbf{U}\}_{\forall k, g}} \sum_{g=1}^G \sum_{k=1}^{K_g^d} \varepsilon_{k_g^d}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) \\ + \sum_{g=1}^G \sum_{k=1}^{K_g^u} \varepsilon_{k_g^u}(\alpha, \gamma, \mathbf{V}, \mathbf{U}). \quad (27)$$

The objective function (i.e., the sum of MSE) is convex and differentiable to the combining matrices with other variables fixed, and there is no constraint associated with combining matrices. Thus, we can differentiate the objective function with respect to $\mathbf{U}_{k_g^d}$ and $\mathbf{U}_{k_g^u}$, and set the derivatives to zero, then the optimal combining matrices for problem (S.1.1) in (27) are given as

$$\mathbf{U}_{k_g^d}^* = \alpha_{k_g^d} \mathbf{C}_{k_g^d}^{-1} \hat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d}, \quad (28)$$

$$\mathbf{U}_{k_g^u}^* = \gamma_{k_g^u} \mathbf{C}_g^{-1} \hat{\mathbf{H}}_{g, k_g^u} \mathbf{V}_{k_g^u}, \quad (29)$$

where $\mathbf{C}_{k_g^d}$ and \mathbf{C}_g are covariance matrices of the received signals at the downlink user k_g^d and the g^{th} base station given in Equations (64) and (65) in Appendix C.

B. Transmitting Beamforming

With fixed combining matrices and power coefficients, the sub-problem regarding precoder optimization is formulated as

$$(S.1.2) \quad \min_{\{\mathbf{V}\}_{\forall k, g}} \sum_{k=1}^{K_g^d} \varepsilon_{k_g^d}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) + \sum_{g=1}^G \nu_g \epsilon_{rsi, g}(\mathbf{V}) \\ + \sum_{g=1}^G \sum_{k=1}^{K_g^u} \varepsilon_{k_g^u}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) \quad (30)$$

$$s.t. \quad (C.1) \quad \sum_{k=1}^{K_g^d} \alpha_{k_g^d}^2 \text{tr}(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H) \leq P_{bs}, \quad \forall g \quad (31)$$

$$(C.2) \quad \gamma_{k_g^u}^2 \text{tr}(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H) \leq P_{ue}, \quad \forall k, g. \quad (32)$$

To solve the constrained optimization problem, we need to augment the objective function with a weighted sum of the constraint functions [27], yielding the Lagrange function as Equation (33), where $\varpi_g \geq 0$ and $\varpi_{k_g^u} \geq 0$ are Lagrange multipliers associated with constraints (31) and (32). Differ-

$$\begin{aligned} \mathcal{L}(\mathbf{V}, \nu, \varpi) = & \sum_{g=1}^G \sum_{k=1}^{K_g^d} \varepsilon_{k_g^d}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) + \sum_{g=1}^G \sum_{k=1}^{K_g^u} \varepsilon_{k_g^u}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) + \sum_{g=1}^G \varpi_g \left(\sum_{k=1}^{K_g^d} \alpha_{k_g^d}^2 \text{tr}(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H) - P_{bs} \right) \\ & + \sum_{g=1}^G \sum_{k=1}^{K_g^u} \varpi_{k_g^u} \left(\gamma_{k_g^u}^2 \text{tr}(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H) - P_{ue} \right) + \sum_{g=1}^G \nu_g \sum_{k=1}^{K_g^d} \alpha_{k_g^d}^2 \text{tr}(\mathbf{H}_{g,g} \mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \mathbf{H}_{g,g}^H + \kappa_{bs} \mathbf{H}_{g,g} \mathcal{D}(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H) \mathbf{H}_{g,g}^H) \end{aligned} \quad (33)$$

entiate the Lagrange function (33) with respect to $\mathbf{V}_{k_g^d}$ and $\mathbf{V}_{k_g^u}$, and set the derivatives to zero, the optimal precoding matrices for the sub-problem (S.1.2) are given as

$$\begin{aligned} \mathbf{V}_{k_g^d}^* = & \frac{1}{\alpha_{k_g^d}} \left(\boldsymbol{\Omega}_g + \nu_g (\mathbf{H}_{g,g}^H \mathbf{H}_{g,g} + \kappa_{bs} \mathcal{D}(\mathbf{H}_{g,g}^H \mathbf{H}_{g,g})) \right. \\ & \left. + \varpi_g^* \mathbf{I} \right)^{-1} \mathbf{H}_{k_g^d, g}^H \mathbf{U}_{k_g^d}, \end{aligned} \quad (34)$$

$$\mathbf{V}_{k_g^u}^* = \frac{1}{\gamma_{k_g^d}} \left(\boldsymbol{\Omega}_{k_g^u} + \varpi_{k_g^u}^* \mathbf{I} \right)^{-1} \mathbf{H}_{g, k_g^u}^H \mathbf{U}_{k_g^u}. \quad (35)$$

where $\boldsymbol{\Omega}_{k_g^d}$ and $\boldsymbol{\Omega}_{k_g^u}$ represent the interference-plus-distortions matrices for associated users, which can be denoted using the function $\mathcal{F}_1(\cdot)$ defined in Appendix E as

$$\begin{aligned} \boldsymbol{\Omega}_g = & \sum_{j=1}^G \sum_{i=1}^{K_j^d} \mathcal{F}_1(\hat{\mathbf{H}}_{i_g^d, g}^H, \mathbf{U}_{i_g^d}) + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \mathcal{F}_1(\hat{\mathbf{H}}_{j, g}^H, \mathbf{U}_{i_j^u}), \\ \boldsymbol{\Omega}_{k_g^u} = & \sum_{j=1}^G \sum_{i=1}^{K_j^d} \mathcal{F}_1(\hat{\mathbf{H}}_{i_g^d, k_g^u}^H, \mathbf{U}_{i_g^d}) + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \mathcal{F}_1(\hat{\mathbf{H}}_{j, k_g^u}^H, \mathbf{U}_{i_j^u}). \end{aligned}$$

Then, we need to find $\varpi_g^* \geq 0$ and $\varpi_{k_g^u}^* \geq 0$ satisfying power constraints (31) and (32), which could be found through Bisection searching. However, calculating the precoder and its norm with different values of ϖ_g and $\varpi_{k_g^u}$ for each iteration during the Bisection searching will yield high computational complexity. To reduce the complexity, we use singular value decomposition (SVD) to convert the expressions of the transmit power into a scalar form. Performing SVD to $\boldsymbol{\Omega}_g + \nu (\mathbf{H}_{g,g}^H \mathbf{H}_{g,g} + \kappa_{bs} \mathcal{D}(\mathbf{H}_{g,g}^H \mathbf{H}_{g,g}))$, which is Hermitian, we have $\boldsymbol{\Omega}_g + \nu (\mathbf{H}_{g,g}^H \mathbf{H}_{g,g} + \kappa_{bs} \mathcal{D}(\mathbf{H}_{g,g}^H \mathbf{H}_{g,g})) = \mathbf{Q}_{k_g^d} \mathbf{D}_{k_g^d} \mathbf{Q}_{k_g^d}^H$, where $\mathbf{Q}_{k_g^d} \mathbf{Q}_{k_g^d}^H = \mathbf{I}$. Then, Equation (34) can be written as

$$\begin{aligned} \mathbf{V}_{k_g^d}^* = & \frac{1}{\alpha_{k_g^d}} \left(\mathbf{Q}_{k_g^d} \mathbf{D}_{k_g^d} \mathbf{Q}_{k_g^d}^H + \mathbf{Q}_{k_g^d} \varpi_g^* \mathbf{Q}_{k_g^d}^H \right)^{-1} \mathbf{H}_{k_g^d, g}^H \mathbf{U}_{k_g^d} \\ = & \frac{1}{\alpha_{k_g^d}} \mathbf{Q}_{k_g^d} \left(\mathbf{D}_{k_g^d} + \varpi_g^* \mathbf{I} \right)^{-1} \mathbf{Q}_{k_g^d}^H \mathbf{H}_{k_g^d, g}^H \mathbf{U}_{k_g^d}. \end{aligned}$$

Let $\mathbf{G}_{k_g^d} = \mathbf{Q}_{k_g^d}^H \mathbf{H}_{k_g^d, g}^H \mathbf{U}_{k_g^d} \mathbf{U}_{k_g^d}^H \mathbf{H}_{k_g^d, g} \mathbf{Q}_{k_g^d}$, the power constraint in Equation (31) can be written as

$$\sum_{k=1}^{K_g^d} \alpha_{k_g^d}^2 \text{tr}(\mathbf{V}_{k_g^d}^* (\mathbf{V}_{k_g^d}^*)^H) = \sum_{k=1}^{K_g^d} \sum_{n=1}^{N_{bs}} \frac{[\mathbf{G}_{k_g^d}]_{n,n}}{([\mathbf{D}_{k_g^d}]_{n,n} + \varpi_g)} \leq P_{bs}$$

Similarly, performing SVD to $\boldsymbol{\Omega}_{k_g^u}$ such that $\boldsymbol{\Omega}_{k_g^u} = \mathbf{Q}_{k_g^u} \mathbf{D}_{k_g^u} \mathbf{Q}_{k_g^u}^H$, and let $\mathbf{G}_{k_g^u} = \mathbf{Q}_{k_g^u}^H \mathbf{H}_{g, k_g^u}^H \mathbf{U}_{k_g^u} \mathbf{U}_{k_g^u}^H \mathbf{H}_{g, k_g^u} \mathbf{Q}_{k_g^u}$. The power constraint in Equation (32) can be written as

$$\sum_{n=1}^{N_{ue}} \frac{[\mathbf{G}_{k_g^u}]_{n,n}}{([\mathbf{D}_{k_g^u}]_{n,n} + \varpi_{k_g^u})} \leq P_{ue}. \quad (36)$$

Since $\mathbf{G}_{k_g^d}$ and $\mathbf{D}_{k_g^d}$ (or $\mathbf{G}_{k_g^u}$ and $\mathbf{D}_{k_g^u}$) are fixed while updating the precoding matrices, we only need to calculate their values once before the Bisection searching, then only scalar calculations are required during the process of Bisection searching.

C. Power Allocation

In this section, we will fix the beamforming matrices and derive the solutions to the power coefficients. The sub-problem with regard to power allocation policy optimization is formulated as

$$(S.2) \quad \min_{\{\alpha, \gamma\}_{\forall k, g}} \sum_{k=1}^{K_g^d} \varepsilon_{k_g^d}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) + \sum_{g=1}^G \sum_{k=1}^{K_g^u} \varepsilon_{k_g^u}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) \quad (37)$$

$$\text{s.t. (C.1)} \quad \sum_{k=1}^{K_g^d} \alpha_{k_g^d}^2 \text{tr}(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H) \leq P_{bs}, \quad \forall g \quad (38)$$

$$(C.2) \quad \gamma_{k_g^u}^2 \text{tr}(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H) \leq P_{ue}, \quad \forall k, g. \quad (39)$$

Similarly, we augment the objective function with a weighted sum of the constraint functions, yielding the Lagrange function

$$\begin{aligned} \mathcal{L}(\alpha, \gamma, \lambda) = & \sum_{g=1}^G \sum_{k=1}^{K_g^d} \varepsilon_{k_g^d}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) + \sum_{g=1}^G \sum_{k=1}^{K_g^u} \varepsilon_{k_g^u}(\alpha, \gamma, \mathbf{V}, \mathbf{U}) \\ & + \sum_{g=1}^G \lambda_g \left(\sum_{k=1}^{K_g^d} \alpha_{k_g^d}^2 \text{tr}(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H) - P_{bs} \right) \\ & + \sum_{g=1}^G \sum_{k=1}^{K_g^u} \lambda_{k_g^u} \left(\gamma_{k_g^u}^2 \text{tr}(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H) - P_{ue} \right), \end{aligned} \quad (40)$$

where $\lambda_g \geq 0$ and $\lambda_{k_g^u} \geq 0$ are Lagrange multipliers associated with constraints (38)-(39). The Lagrange function is convex to the power coefficients. Differentiate the Lagrange

function (40) with respect to $\alpha_{k_g^d}$ and $\gamma_{k_g^u}$ and set the derivatives to zero, the optimal power coefficients for the sub-problem (S.2) are given as

$$\alpha_{k_g^d}^* = \frac{\mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^d}^H \hat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d} \right) \right\}}{\chi_{k_g^d} + \lambda_g^* \text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right)}, \quad (41)$$

$$\gamma_{k_g^u}^* = \frac{\mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^u}^H \hat{\mathbf{H}}_{g, k_g^u} \mathbf{V}_{k_g^u} \right) \right\}}{\chi_{k_g^u} + \lambda_{k_g^u}^* \text{tr} \left(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H \right)}, \quad (42)$$

where $\chi_{k_g^d}$ and $\chi_{k_g^u}$ can be denoted using the function $\mathcal{F}_2(\cdot)$ defined in Appendix E as

$$\chi_{k_g^d} = \sum_{j=1}^G \sum_{i=1}^{K_j^d} \mathcal{F}_2 \left(\mathbf{U}_{i_j^d}, \hat{\mathbf{H}}_{i_j^d, g}, \mathbf{V}_{k_g^d} \right) + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \mathcal{F}_2 \left(\mathbf{U}_{i_j^u}, \hat{\mathbf{H}}_{j, g}, \mathbf{V}_{k_g^d} \right), \quad (43)$$

$$\chi_{k_g^u} = \sum_{j=1}^G \sum_{i=1}^{K_j^d} \mathcal{F}_2 \left(\mathbf{U}_{i_j^d}, \hat{\mathbf{H}}_{i_j^d, k_g^u}, \mathbf{V}_{k_g^u} \right) + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \mathcal{F}_2 \left(\mathbf{U}_{i_j^u}, \hat{\mathbf{H}}_{j, k_g^u}, \mathbf{V}_{k_g^u} \right). \quad (44)$$

To find the Lagrange multiplier $\lambda_{k_g^u}$, we can differentiate the Lagrange function (40) with respect to $\lambda_{k_g^u}$ and set the derivative to zero as

$$\begin{aligned} & \gamma_{k_g^u}^2 \text{tr} \left(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H \right) - P_{ue} \\ & = \left(\frac{\mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^u}^H \hat{\mathbf{H}}_{g, k_g^u} \mathbf{V}_{k_g^u} \right) \right\}}{\chi_{k_g^u} + \lambda_{k_g^u}^* \text{tr} \left(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H \right)} \right)^2 \text{tr} \left(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H \right) - P_{ue} = 0, \end{aligned}$$

yielding the optimal Lagrange multiplier as

$$\lambda_{k_g^u}^* = \max \left[0, -\frac{\chi_{k_g^u}}{\text{tr} \left(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H \right)} + \frac{\mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^u}^H \hat{\mathbf{H}}_{g, k_g^u} \mathbf{V}_{k_g^u} \right) \right\}}{\sqrt{\text{tr} \left(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H \right) P_{ue}}} \right],$$

where $\max[0, x]$ guarantees $\lambda_{k_g^u} \geq 0$ to strictly satisfy the constraint (39). For λ_g , it can be obtained from the positive

$$\text{root of } \sum_{k=1}^{K_g^d} \left(\frac{\mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^d}^H \hat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d} \right) \right\}}{\chi_{k_g^d} + \lambda_g \text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right)} \right)^2 - \frac{P_{bs}}{\text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right)} = 0$$

according to the derivative, or $\lambda_g^* = 0$ if the positive root does not exist. However, it is difficult to derive the close-form expression of its root since the variable is in the denominator. Alternatively, λ_g^* can be obtained using Bisection search within the search range of $[0, \bar{\lambda}_g]$, where $\bar{\lambda}_g$ is the upper bound of λ_g . The power constraints at the BS can be written as Equation (45), which gives Equation (46). Equation (46) is in the form of $a\lambda_g^2 + 2b\lambda_g + c \leq 0$, where $a = \sum_{k=1}^{K_g^d} \text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right)$, $b = \sum_{k=1}^{K_g^d} \chi_{k_g^d}$, and $c =$

Algorithm 1 Iterative JPAIM Algorithm

- 1: Initialize the power allocation coefficients as $\alpha_{k_g^d} = \sqrt{\frac{P_{bs}}{b_d K_g^d}}$ and $\gamma_{k_g^u} = \sqrt{\frac{P_{ue}}{b_u}} \forall k, g$.
 - 2: Randomly initializes the precoding matrices and normalize.
 - 3: **repeat**
 - 4: Update the combining matrices $\left\{ \mathbf{U}_{k_g^d}, \mathbf{U}_{k_g^u} \right\}_{\forall k, g}$ as (28) and (29).
 - 5: Calculate and record the current sum of MSE as $\varepsilon^{(t)}$
 - 6: Update the precoding matrices $\left\{ \mathbf{V}_{k_g^d}, \mathbf{V}_{k_g^u} \right\}_{\forall k, g}$ as (34) and (35).
 - 7: Update the power allocation coefficients $\left\{ \alpha_{k_g^d}, \gamma_{k_g^u} \right\}_{\forall k, g}$ as (41) and (42).
 - 8: **until** $\varepsilon^{(t)} - \varepsilon^{(t-1)} < \iota$ or $t > t_m$, where ι is the threshold and t_m is the iteration limit.
-

$$\sum_{k=1}^{K_g^d} \frac{\chi_{k_g^d}^2}{\text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right)} - \frac{1}{P_{bs}} \sum_{k=1}^{K_g^d} \mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^d}^H \hat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d} \right) \right\}^2.$$

Thus, we can derive the upper bound as

$$\lambda_g \leq -\frac{b}{a} + \sqrt{\frac{b^2}{a^2} - \frac{c}{a}} = \bar{\lambda}_g.$$

D. Iterative Processing

The solutions of beamforming matrices and power coefficients in Equations (28), (29), (34), (35), (41), and (42) show inter-dependence on each other. Thus, we need an iterative algorithm to obtain the overall solution by updating them until they converge. We first initialize power coefficients to meet the power budgets of each node and randomly initialize beamforming matrices. Then, we continuously update the precoding matrices, power coefficients, and combining matrices in order. The iterative procedure stops if the loss function, i.e., the objective function in (24), does not decrease with iterations (the decreasing amount is less than the threshold) or the iteration time exceeds the limitation. This processing is summarized as Algorithm 1. Note that, since the joint optimization problem is not jointly convex to the power coefficients and beamforming matrices, the algorithm can only converge to the local minimal closest to the initial point. Thus, the performance strongly depends on the initial point, which is similar to existing studies [6], [19].

1) *Convergence Behaviour*: This algorithm can keep reducing the objective function of the joint optimization problem (i.e., loss function) and converge to a local minimum. Let $\theta(\alpha, \beta, \mathbf{V}, \mathbf{U})$ denote the loss function. Within an iteration, one of the optimized variables is updated with others fixed, and the Lagrange function is convex to the target variable under this condition. Thus, we always obtain the optimal value of the target variable with other variables fixed, reducing the loss function at each step as:

$$\theta \left(\alpha, \gamma, \mathbf{V}, \mathbf{U}^{(t+1)} \right) \leq \theta \left(\alpha, \gamma, \mathbf{V}, \mathbf{U}^{(t)} \right) \quad (47)$$

$$\theta \left(\alpha, \gamma, \mathbf{V}^{(t+1)}, \mathbf{U} \right) \leq \theta \left(\alpha, \gamma, \mathbf{V}^{(t)}, \mathbf{U} \right) \quad (48)$$

$$\begin{aligned}
 0 &= P_{bs} - \sum_{k=1}^{K_g^d} \left(\frac{\mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^d}^H \hat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d} \right) \right\}}{\chi_{k_g^d} + \lambda_g \text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right)} \right)^2 \text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right) \leq P_{bs} - \frac{\sum_{k=1}^{K_g^d} \mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^d}^H \hat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d} \right) \right\}^2 \text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right)}{\sum_{k=1}^{K_g^d} \left(\chi_{k_g^d} + \lambda_g \text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right) \right)^2} \\
 &= P_{bs} - \frac{\sum_{k=1}^{K_g^d} \mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^d}^H \hat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d} \right) \right\}^2 \text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right)}{\lambda_g^2 \sum_{k=1}^{K_g^d} \left(\text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right) \right)^2 + 2\lambda_g \sum_{k=1}^{K_g^d} \text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right) \chi_{k_g^d} + \sum_{k=1}^{K_g^d} \chi_{k_g^d}^2}
 \end{aligned} \quad (45)$$

$$\lambda_g^2 \sum_{k=1}^{K_g^d} \text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right) + 2\lambda_g \sum_{k=1}^{K_g^d} \chi_{k_g^d} + \sum_{k=1}^{K_g^d} \frac{\chi_{k_g^d}^2}{\text{tr} \left(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \right)} - \frac{1}{P_{bs}} \sum_{k=1}^{K_g^d} \mathcal{R} \left\{ \text{tr} \left(\mathbf{U}_{k_g^d}^H \hat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d} \right) \right\}^2 \leq 0 \quad (46)$$

$$\theta \left(\alpha^{(t+1)}, \gamma^{(t+1)}, \mathbf{V}, \mathbf{U} \right) \leq \theta \left(\alpha^{(t)}, \gamma^{(t)}, \mathbf{V}, \mathbf{U} \right) \quad (49)$$

After the t^{th} iteration, we have $\alpha^{(t)}, \gamma^{(t)}, \mathbf{V}^{(t)}, \mathbf{U}^{(t)}$, and then update the variables to $\mathbf{U}^{(t+1)}, \mathbf{V}^{(t+1)}, \alpha^{(t+1)}, \gamma^{(t+1)}$ in order. According to (47)-(49), the loss function is guaranteed to be reduced at each iteration until it converges to a local optimal.

2) *Computational Complexity*: We measure the computational complexity by accounting for the required multiplication operations. The required multiplications of some basic operations are given in Table I, where \mathbf{A}_1 is of size $a \times b$ and \mathbf{A}_2 is of size $b \times c$. According to this, the total number

TABLE I
COMPLEXITY OF OPERATIONS

Operation	Number of multiplications
$\mathbf{A}_1 \mathbf{A}_2$	$a \times b \times c$
Eigen decomposition (\mathbf{A}_1)	$\mathcal{O}(a^3)$
Matrix inverse (\mathbf{A}_1)	$\mathcal{O}(a^3)$
$\mathcal{F}_1(\mathbf{A}_1, \mathbf{A}_2)$	$2a^2b + ab^2 + b^2c + 2abc + 3a^2$

of multiplications of calculating beamforming matrices (taking $\mathbf{V}_{k_g^d}$ as an instance) can be given as

$$\begin{aligned}
 M_v &= \underbrace{GK[3A_b^3 + A_b^2(2A_u + 3N_s + 6)] + A_b(A_u^2 + 2A_u N_s)}_{\Omega_g} \\
 &\quad + \underbrace{A_u^2 N_s}_{\Omega_g} + \underbrace{A_b^3}_{\mathbf{H}_{g,g}^H \mathbf{H}_{g,g}} + \underbrace{A_b^2 A_u + A_b A_u N_s}_{\text{rest multiplication}},
 \end{aligned}$$

plus an Eigen decomposition and an inverse to a matrix of size $A_b \times A_b$, where the calculation process is divided into calculating Ω_g , $\mathbf{H}_{g,g}^H \mathbf{H}_{g,g}$, and rest matrix multiplication, and we assume $K_g^d = K_g^u = K, \forall g, N_{bs} = M_{bs} = A_b, N_{ue} = M_{ue} = A_u$, and $b_d = b_u = N_s$ for simplicity. Usually, we will have $A_b \gg A_u \geq N_s$. To update the corresponding power coefficient $\alpha_{k_g^d}$, the complexity mainly comes from calculating $\chi_{k_g^d}$ and the rest matrix multiplication. Thus, the total number of multiplication operations required to compute $\alpha_{k_g^d}$ is

$$M_\alpha = \underbrace{GK[2A_b^3 + A_b^2(A_u + 5N_s + 2)] + A_b(A_u^2 + 4A_u N_s)}_{\chi_{k_g^d}}$$

$$\begin{aligned}
 &\underbrace{+ N_s^2 + A_u^2 N_s + 2A_u N_s + 2N_s^2}_{\chi_{k_g^d}} \\
 &\quad + \underbrace{A_b^2 N_s + A_b(A_u N_s + N_s^2)}_{\text{rest multiplication}}.
 \end{aligned}$$

Therefore, the number of multiplication operations required for one iteration of the JPAIM algorithm is given as $M_v + M_\alpha$ plus an Eigen decomposition and inverse operation to a matrix of size $A_b \times A_b$, which is in the order of GKA_b^3 such that

$$C_{\text{JPAIM}} = \mathcal{O}(GKA_b^3). \quad (50)$$

The computational complexity is in the same order as the MWSR algorithm proposed in [6]. The complexity analysis of the MWSR algorithm is not detailed in this paper since it is out of the scope, but it can be analyzed similarly based on the solution given in [6]. Although the complexities of the two algorithms are in the same order, JPAIM needs $M_\alpha + A_b^3$ more multiplications, which almost double the time complexity of a single iteration. Therefore, the JPAIM algorithm needs at least half the convergence time to reduce the overall complexity.

V. SIMULATION RESULTS

We consider OFDM modulation with 15kHz subcarrier spacing and 10 MHz bandwidth centered at 2.5GHz, and other parameters are given in Table II. The system parameters are chosen based on 3GPP specifications [28], [29] and our preliminary work [3], [6]. The ASIC depth realized by any other methods (e.g., antenna isolation as in [31] or RF cancellers as in [9]) will be directly reflected by the pathloss of the SI channel regardless of the implementation methods. We assume the pathloss without any ASIC is 0dB due to the proximity of IBFD transceivers [1]. Therefore, x dB of realized ASIC by other techniques means $l_g = x$ dB $\forall g$ in simulations, where l_g is defined in Section III-B. We do not presume a realized ASIC depth but evaluate the algorithms' performance at different ASIC depths. In the later simulations, we consider the ideal case, i.e., perfect ASIC with $l_g = 120$ dB, to demonstrate the characteristics of the algorithm itself, but this does not mean that our algorithm requires such perfect ASIC.

TABLE II
 SIMULATION SETTING

Parameters	Values
Number of antennas	$N_{bs} = M_{bs} = 16, N_{ue} = M_{ue} = 2$ if not specified
Cells	hexagon cells with 200m of inter-site distance and 10m of a minimum base station to users distance
LOS probability and pathloss	models for UMi scenarios in [28]
Channel matrices	$\mathbf{H}_{B,A} = \sqrt{l_{B,A}} \mathbf{N}$ if the LOS probability ≥ 0.5 $\mathbf{H}_{B,A} = \sqrt{l_{B,A}} \left(\sqrt{\frac{k_{B,A}}{k_{B,A}+1}} \mathbf{I} + \sqrt{\frac{1}{k_{B,A}+1}} \mathbf{N} \right)$ if the LOS probability < 0.5 (\mathbf{N} has elements independently and identically distributed as $\mathcal{CN}(0,1)$)
Transmit power budgets	24dBm for BSs and 23dBm for UEs
Thermal noise density	-174dBm/Hz
Noise figure	13dB for BSs and 9dB for UEs
Number of effective bits of DACs/ADCs	$b = 12$ if not specified
Channel uncertainty factor	$\hat{\sigma}_{B,A}^2 = \varrho \ \mathbf{H}_{B,A}\ _F^2 \forall A, B$, where $\varrho = -120$ dB

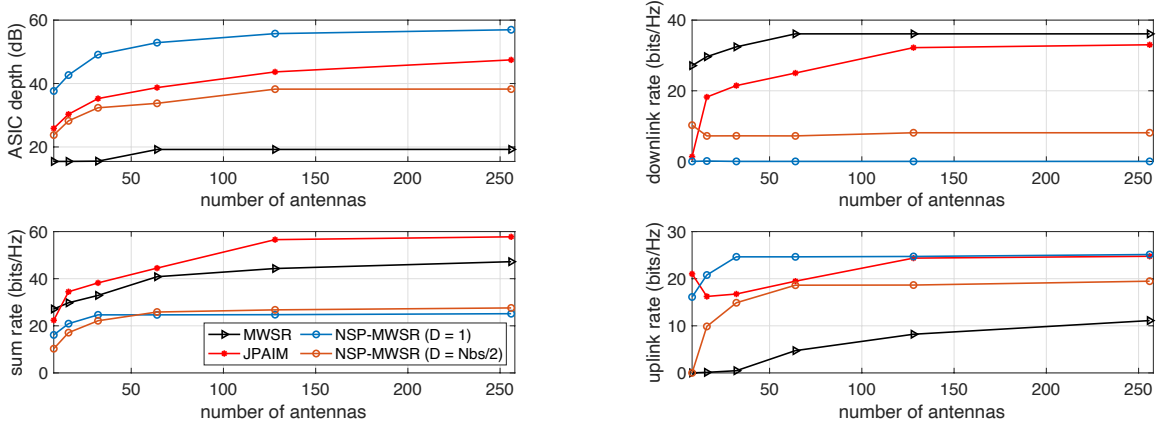


Fig. 2. The ASIC capability and achievable spectral efficiency versus the number of transmitting and receiving antennas of the FD node.

Since the throughput maximization for MCMU networks is not convex and the performance depends on the initial point, we run 1,000 times Monte Carlo simulations for each scenario and evaluate the performance via the averaged values. The realization is randomly generated at each step of the loop and the initial values of the optimized variable are also randomly generated. For a fair comparison, the simulation settings, e.g., user location, channel matrices, initial points, etc., remain identical for the two algorithms in each realization in 1,000 iterations of the Monte Carlo simulations, and the two algorithms will be performed separately.

A. SIC Capability

Fig. 2 shows the achievable ASIC depth and SE versus the number of transmitting and receiving antennas of IBFD transceivers using different beamforming schemes. In order to demonstrate the ASIC capability of precoders, we consider the scenario that no ASIC is realized by other techniques since precoders will tend to suppress other significant interference rather than SI if it has been efficiently suppressed. In this case, the MWSR precoder optimizes the interference management for the downlink payload regardless of SI, resulting in a large received SI power to saturate the IBFD receiver and significantly degrade the uplink rate. More receiving antennas compensate for the uplink rate reduction due to more degrees of freedom of combiners, managing the interference better for

the uplink payload. Combining the null-space projection [15] with MWSR beamforming as explained in Appendix D can enhance its ASIC capability. The uplink rate is increased at the cost of the downlink rate degradation. The trade-off can be made by adjusting the dimensions of the subspace chosen for projection, i.e., D (defined in Appendix D), which also determines the ASIC capability. The disadvantage of NSP-MWSR is that the downlink rate decreases significantly with increasing D . In contrast, our proposed JPAIM beamforming has less impact on the downlink rate with enhanced ASIC capability. It can achieve higher SE than other schemes and simultaneously support effective downlink and uplink communications with ≥ 16 antennas at the IBFD transceiver. It can offer > 30 dB of ASIC depth with ≥ 16 transmitting antennas at the IBFD node and can be improved by more antennas. When the antenna array is small (e.g., 8-antenna arrays), the power allocation tends to reduce the transmit power of the downlink payload to maximize the sum rate by maximizing the uplink rate. With enlarging antenna arrays (e.g., 16-antenna arrays), it is able to provide effective ASIC to enable efficient IBFD communications. However, it introduces interference for the uplink payload so that the uplink rate goes down at first.

B. Convergence Behaviour

Fig. 3 shows the convergence behavior of the JPAIM algorithm compared to the MWSR algorithm proposed in [6].

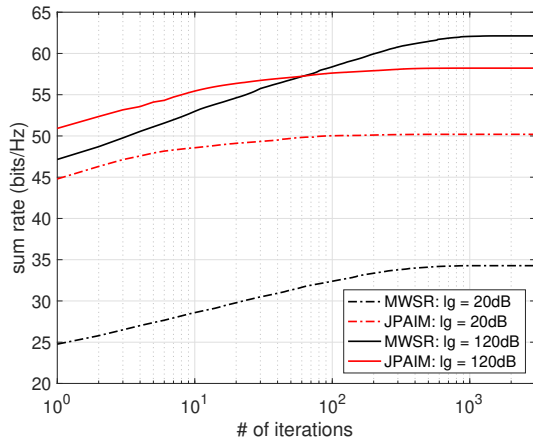


Fig. 3. Convergence behavior (x-axis is in the log scale).

Since the joint power allocation and beamforming optimization problem is not jointly convex, the convergence behavior and achievable sum rate depend on the initial points as the algorithms converge to the local minimum closest to the initial point. Therefore, we ran 1,000 times Monte Carlo simulations and calculated the average value to show their performance. It shows that the JPAIM algorithm converges faster than the MWSR algorithm under both conditions (i.e., insufficient and sufficient ASIC realized). The JPAIM algorithm reduces the iteration time to converge by at least half compared to the MWSR Algorithm, which meets the condition analyzed in Section IV-D2 to reduce the overall time complexity. Besides, both algorithms converge faster with insufficient ASIC realized.

C. Improvement with Insufficient ASIC

In this section, we evaluate the performance of the JPAIM algorithm in terms of achievable SE (i.e., sum rate), as the most attractive factor of IBFD is its potential to double the SE of its HD counterpart. Fig. 4 shows the achievable SE of the JPAIM algorithm versus realized ASIC depths compared to the MWSR algorithm. When the ASIC depth realized by other techniques is insufficient (i.e., < 60 dB), the proposed JPAIM achieves higher SE than MWSR. This is attractive as other ASIC techniques have incredibly high complexity and cost in MIMO systems, as we stated in Section I. Active antenna isolation, which has feasible implementation complexity, can achieve around 30 dB of ASIC in practice [31]. In contrast, deeper ASIC needs to be realized by the much more complex RF cancellation, which is almost practically prohibited in MIMO systems. In this case, the JPAIM algorithm can improve the SE by 42.9% in IBFD compared to HD, while the MWSR algorithm cannot. This value can be improved to 60.9% with further digital processing if the received SI has already been within the dynamic range of receivers. In contrast, the MWSR must require RF cancellers to achieve the IBFD gain due to its limited ASIC capability, as illustrated in Fig. 2.

D. Time Efficiency

Results in Fig. 4 suggest that MWSR can achieve higher SE than JPAIM with sufficient ASIC applied. This is reasonable

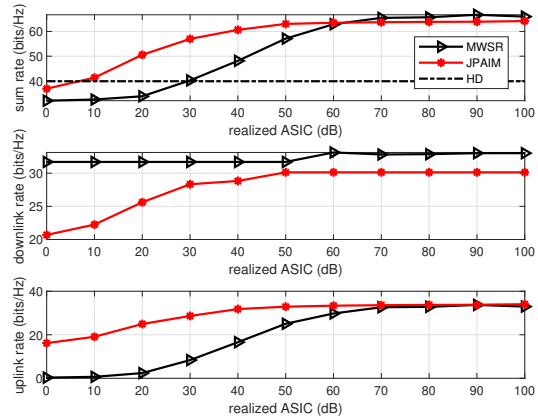
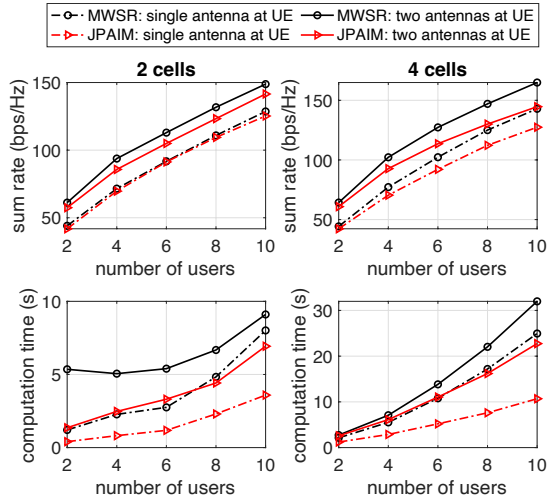


Fig. 4. Achievable spectral efficiency vs realized ASIC depth (i.e., l_g).

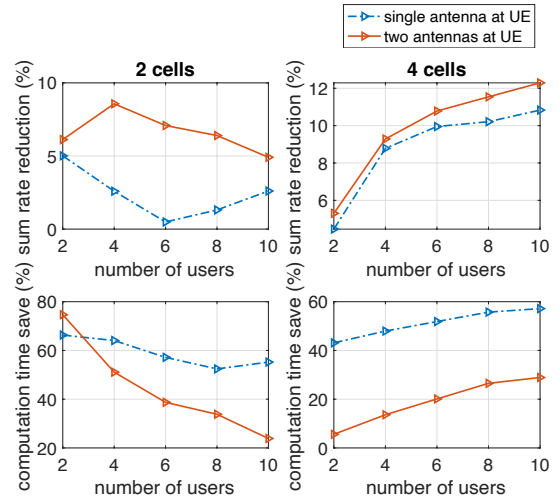
since MWSR yields a joint power allocation and beamforming solution while JPAIM decomposes it into two sub-problems to solve. However, the decomposition also brings benefits in terms of time efficiency (i.e., $= \frac{\text{sum rate}}{\text{computation time}}$). In order to evaluate the performance of the proposed algorithm in terms of time efficiency, we compare the SE and time complexity of the two algorithms in networks of different scales. We use Monte Carlo simulations to compare 2-cell and 4-cell networks with 2-10 users.

Fig. 5(a) shows the performance difference between the two algorithms in terms of sum rate and average computation time of a realization with different network sizes, i.e., different numbers of cells, users, and antennas at users. Fig. 5(b) shows the corresponding sum rate reduction and computation save percentages of JPAIM compared to MWSR, and Fig. 6 shows their time efficiency. The results show that JPAIM achieves a sum rate close to the one achieved by MWSR, but it takes much less computation time. The sum rate loss generally increases with increasing numbers of cells and users, and it is more significant with multi-antenna users. For single-antenna users, JPAIM saves at least 40% of the computation time at the cost of $< 10\%$ sum rate reduction for most cases. Although the benefits are compromised with two-antenna users, the computation time saved is always higher than the sum rate reduction.

Fig. 6 also illustrates that JPAIM achieves higher time efficiency than MWSR under all conditions. This indicates that JPAIM has a lower time complexity in practical implementations than MWSR at the cost of an acceptable sum rate loss. The benefit of JPAIM is significant with single-antenna users, which is meaningful for practical cellular network deployment since many UE devices still use a single antenna. This can be understood since our JPAIM algorithm uses a coarse power allocation scheme, while MWSR uses a refined power allocation. Therefore, MWSR can achieve higher SE than JPAIM. However, refined power allocation will drastically increase the computational complexity, making it less time-efficient than our proposed algorithm. When the interference is not complicated (i.e., with fewer cells and users), the refined power allocation is not so important, while a coarse power



(a) Sum rate and computation time versus the number of users.



(b) Sum rate reduction and computation time save percentages.

Fig. 5. Performance comparison of JPAIM and MWSR in terms of SE and time complexity.

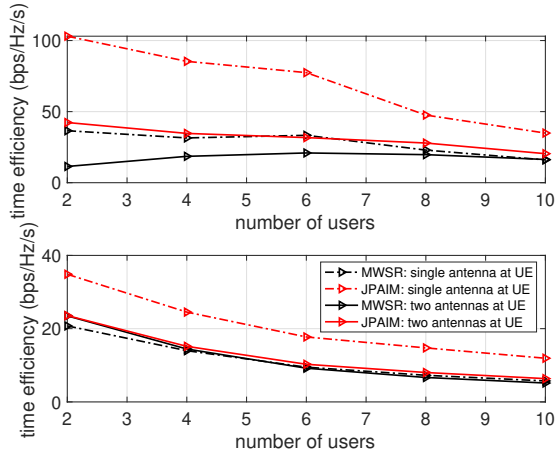


Fig. 6. Time efficiency of the proposed JPAIM algorithm.

allocation can greatly reduce the time complexity due to its fast convergence characteristics, so JPAIM has a much higher time efficiency than MWSR. When the interference becomes complicated (i.e., in enlarging networks), a refined power allocation will significantly improve network capacity, while a coarse power allocation can reduce the complexity relatively less (as shown in Fig. 5). In addition, our algorithm demonstrates a significant time efficiency improvement for single-antenna users, while the improvement is less pronounced for multi-antenna users. The reason is that multi-antenna users can perform beamforming to increase channel capacity, which requires user terminals to also perform iterative calculations, thereby significantly increasing the computational complexity. This is not necessary for the JPAIM algorithm based on coarse power allocation when single antenna users are considered, while the refined power allocation-based MWSR algorithm still requires iterative calculations for power allocation of the data streams. Therefore, the time complexity benefit of JPAIM compared to MWSR is greatly reduced when there are

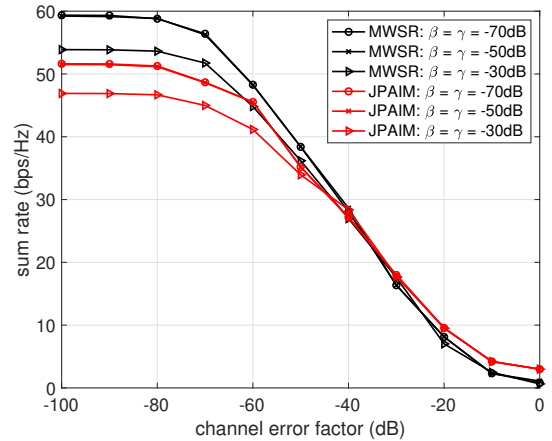


Fig. 7. Robustness to channel uncertainty and transceiver HWIs.

multiple-antenna users (see Fig. 5(b)), resulting in the overall time efficiency is not obvious.

E. Robustness to Channel Uncertainty and Transceiver HWIs

In our study, the criteria for robustness are defined based on the algorithm's ability to maintain consistent performance levels under varying channel conditions and hardware impairments. That is, the algorithm can be considered robust if its performance does not degrade dramatically with increasing channel uncertainty and hardware impairments regardless of the highest sum rate. Fig. 7 shows the achievable sum rate variation of the JPAIM algorithm compared to the MWSR algorithm against the channel error factor ϱ . It demonstrates the robustness of JPAIM to channel uncertainty. It shows that significant channel uncertainty has less impact on JPAIM than MWSR. The benefits come from the fact that the effects of channel uncertainty are included in our algorithm as detailed in Appendix A. In addition, it is also robust to the transceiver hardware impairments that reducing the dynamic range of

transceivers from 70dB to 50dB has no noticeable effect on the achievable sum rate. Although a smaller dynamic range (e.g., 30dB) of transceivers reduces the achievable sum rate, it is acceptable as the reduction is less than 10%. The robustness to transceiver HWIs is similar for MSER and JPAIM since they both consider the HWIs in their algorithm designs.

VI. CONCLUSIONS

In this paper, we have formulated and solved an MSE minimization problem regarding power allocation and beamforming with hardware limits. We have revealed that downlink precoders have to sufficiently suppress SI before the receiver (i.e., without combiners); otherwise, the uplink communication rate will be naturally reduced due to the limited dynamic range of receivers. To this end, we have enhanced the ASIC capability of precoders by adding a constraint on the received RSI power and have demonstrated the advantages of our method over existing NSP-based methods. The formulated problem has been converted to an MSE plus RSI minimization problem and then decomposed into two sub-problems to be solved. The closed-form solutions to the sub-problems are derived, and the overall solution is obtained by an iterative algorithm. It has been demonstrated that our JPAIM algorithm could achieve 42.9% of the IBFD gain in terms of spectral efficiency with 3GPP specifications. Then, we evaluated its performance under ideal SIC conditions by comparing it to the existing MWSR algorithm [6], showing it takes much less computation time due to its faster convergence speed at the cost of acceptable sum rate loss. The benefit of our proposed JPAIM algorithm is significant with single-antenna users that it saves at least 40% of the computation time at the cost of < 10% sum rate reduction. In addition, we have proved that it is robust to channel uncertainty and transceiver hardware impairments due to the robust design.

ACKNOWLEDGMENTS

The work was supported in part by the research grant from Huawei Technologies (Sweden) AB.

APPENDIX A

EFFECTS OF CHANNEL UNCERTAINTY

In this appendix, we give the statistics of the errors caused by the channel uncertainty at the base station and user equipment. The errors caused by the channel uncertainty at the g^{th} BS and the downlink user k_g^d can be written as

$$\mathbf{e}_g = \Delta_{g,k_g^u} \mathbf{x}_{k_g^u} + \sum_{i,j \neq k,g} \Delta_{g,i_j^u} \mathbf{x}_{i_j^u} + \sum_{j \neq g} \Delta_{g,j} \mathbf{x}_j, \quad (51)$$

$$\mathbf{e}_{k_g^d} = \Delta_{k_g^d,g} \mathbf{x}_{k_g^d} + \sum_{i,j \neq k,g} \Delta_{k_g^d,j} \mathbf{x}_{i_j^d} + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \Delta_{k_g^d,i_j^u} \mathbf{x}_{i_j^u}. \quad (52)$$

For any two specific nodes A and B , we can calculate the covariance of the associated errors as

$$Cov(\Delta_{B,A} \mathbf{x}_A) = \mathbb{E} \{ \Delta_{B,A} \mathbf{x}_A \mathbf{x}_A^H \Delta_{B,A}^H \} \quad (53)$$

$$= \mathbb{E}_{\Delta_{B,A}} \{ \Delta_{B,A} \mathbf{T}_A \Delta_{B,A}^H \} \stackrel{(a)}{=} \tilde{\sigma}_{B,A}^2 tr(\mathbf{T}_A) \mathbf{I},$$

where (a) comes from following derivation:

$$\begin{aligned} & \mathbb{E}_{\Delta_{B,A}} \left\{ \left[\Delta_{B,A} \mathbf{T}_A \Delta_{B,A}^H \right]_{m,n} \right\} \\ &= \mathbb{E}_{\Delta} \left\{ \sum_{k=1}^{N_{\Delta}} \sum_{j=1}^{N_{\Delta}} \Delta_{mk} T_{kj} \Delta_{nj}^* \right\} \\ &= \sum_{k=1}^{N_{\Delta}} \sum_{j=1}^{N_{\Delta}} \mathbb{E}_{\Delta} \{ \Delta_{mk} T_{kj} \Delta_{mk}^* \} \delta_{mn} \delta_{kj} \\ &\stackrel{(c)}{=} \sum_{k=1}^{N_{\Delta}} \mathbb{E}_{\Delta} \{ |\Delta_{mk}|^2 T_{kk} \delta_{mn} \} = \tilde{\sigma}_{B,A}^2 tr(\mathbf{T}_A) \delta_{mn}, \end{aligned} \quad (54)$$

where N_{Δ} denotes the number of columns of matrix $\Delta_{B,A}$, T_{kj} denotes the element at the k^{th} row and j^{th} column of matrix \mathbf{T}_A , and Δ_{mk} denotes the element at the m^{th} row and k^{th} column of matrix $\Delta_{B,A}$. δ represents the correlation coefficient, and $\delta_{kj} = 0, \forall k \neq j$, $\delta_{kk} = 1, \forall k$, $\delta_{mn} = 0, \forall m \neq n$, and $\delta_{mm} = 1, \forall m$. Thus, the covariance matrices of the errors are given as

$$Cov(\mathbf{e}_g) = \underbrace{\sum_{j=1}^G \sum_{i=1}^{K_j^u} \tilde{\sigma}_{g,i_j^u}^2 tr(\mathbf{T}_{i_j^u}) + \sum_{j \neq g} \tilde{\sigma}_{g,j}^2 tr(\mathbf{T}_j)}_{\tilde{\sigma}_g^2} \mathbf{I}_{M_{bs}} \quad (55)$$

$$\begin{aligned} Cov(\mathbf{e}_{k_g^d}) &= \underbrace{\sum_{j=1}^G \sum_{k=1}^{K_j^d} \tilde{\sigma}_{k_g^d,j}^2 tr(\mathbf{T}_{i_j^d}) + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \tilde{\sigma}_{k_g^d,i_j^u}^2 tr(\mathbf{T}_{i_j^u})}_{\tilde{\sigma}_{k_g^d}^2} \mathbf{I}_{M_{ue}}. \end{aligned} \quad (56)$$

APPENDIX B PROOF OF LEMMA

The only difference between the two formulations ($P.1$) and ($P.2$) is that the RSI power-related terms appear in different places (one in the constraint condition and one in the objective function). Since the RSI power is only related to the downlink precoders, this difference does not affect the solution to other variables. The Lagrange function of the two problems concerning the downlink precoding matrices $\mathbf{V}_{k_g^d}$ can be given as

$$\mathcal{L}_{P_1} = \Sigma_L(\mathbf{V}) + \sum_{g=1}^G \nu_g \epsilon_{rsi,g}(\mathbf{V}), \quad (57)$$

$$\mathcal{L}_{P_2} = \Sigma_L(\mathbf{V}) + \sum_{g=1}^G \varpi_{rsi,g} (\epsilon_{rsi,g}(\mathbf{V}) - \bar{\epsilon}_{rsi,g}), \quad (58)$$

where $\varpi_{rsi,g}$ is the Lagrange multiplier; $\Sigma_L(\mathbf{V})$ denotes the objective and power constraint-related terms given as

$$\begin{aligned} \Sigma_L(\mathbf{V}) &= \sum_{g=1}^G \sum_{k=1}^{K_g^d} \varepsilon_{k_g^d}(\mathbf{V}) + \sum_{g=1}^G \sum_{k=1}^{K_g^u} \varepsilon_{k_g^u}(\mathbf{V}) \\ &+ \sum_{g=1}^G \varpi_g \left(\sum_{k=1}^{K_g^d} \alpha_{k_g^d}^2 \text{tr}(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H) - P_{bs} \right). \end{aligned} \quad (59)$$

The optimal solutions to $\mathbf{V}_{k_g^d}$ are obtained by deriving the Lagrange function (57) or (58) with respect to $\mathbf{V}_{k_g^d}$ and set the derivatives to zero, which will be identical if $\nu_g = \varpi_{\text{rsi},g}$. Thus, optimization problems (P.1) and (P.2) share the same solutions (i.e., (P.1) and (P.2) are equivalent) with $\nu_g = \varpi_{\text{rsi},g}$. It should be noted that the optimal value of the Lagrange multiplier $\varpi_{\text{rsi},g}$ depends on the tolerable RSI power we set.

APPENDIX C SIMPLICITIES FOR MSE EXPRESSIONS

In this appendix, we give some simplicities as

$$\text{tr} \left(\mathbb{E} \left\{ \mathbf{s}_{k_g^d} \mathbf{s}_{k_g^d}^H \right\} \right) = \text{tr}(\mathbf{I}_{b_d}) = b_d; \quad (60)$$

$$\text{tr} \left(\mathbb{E} \left\{ \mathbf{s}_{k_g^u} \mathbf{s}_{k_g^u}^H \right\} \right) = \text{tr}(\mathbf{I}_{b_u}) = b_u; \quad (61)$$

$$\begin{aligned} \mathbf{T}_g &= \mathbb{E} \left\{ \mathbf{x}_g \mathbf{x}_g^H \right\} = \mathbb{E} \left\{ (\mathbf{V}_g \mathbf{A}_g \mathbf{s}_g + \mathbf{c}_g) (\mathbf{V}_g \mathbf{A}_g \mathbf{s}_g + \mathbf{c}_g)^H \right\} \\ &= \mathbf{V}_g \mathbf{A}_g \mathbf{A}_g^H \mathbf{V}_g^H + \sigma_t^2 \mathcal{D}(\mathbf{V}_g \mathbf{A}_g \mathbf{A}_g^H \mathbf{V}_g^H) \\ &= \underbrace{\sum_{k=1}^{K_g^d} \alpha_{k_g^d}^2 (\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H + \sigma_t^2 \mathcal{D}(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H))}_{\mathbf{T}_{k_g^d}}; \end{aligned} \quad (62)$$

$$\begin{aligned} \mathbf{T}_{k_g^u} &= \mathbb{E} \left\{ \mathbf{x}_{k_g^u} \mathbf{x}_{k_g^u}^H \right\} \\ &= \mathbb{E} \left\{ \left(\gamma_{k_g^u} \mathbf{V}_{k_g^u} \mathbf{s}_{k_g^u} + \mathbf{c}_{k_g^u} \right) \left(\gamma_{k_g^u} \mathbf{V}_{k_g^u} \mathbf{s}_{k_g^u} + \mathbf{c}_{k_g^u} \right)^H \right\} \\ &= \gamma_{k_g^u}^2 \mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H + \sigma_t^2 \gamma_{k_g^u}^2 \mathcal{D}(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H) \\ &= \gamma_{k_g^u}^2 \left(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H + \sigma_t^2 \mathcal{D}(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H) \right); \end{aligned} \quad (63)$$

$$\begin{aligned} \mathbf{C}_{k_g^d} &= \mathbb{E} \left\{ \mathbf{y}_{k_g^d} \mathbf{y}_{k_g^d}^H \right\} \\ &= \sum_{j=1}^G \sum_{i=1}^{K_j^d} \hat{\mathbf{H}}_{k_g^d,j} \mathbf{T}_{i_j^d} \hat{\mathbf{H}}_{k_g^d,j}^H + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \hat{\mathbf{H}}_{k_g^d,i_j^u} \mathbf{T}_{i_j^u} \hat{\mathbf{H}}_{k_g^d,i_j^u}^H \\ &+ \sigma_r^2 \mathcal{D} \left(\sum_{j=1}^G \sum_{i=1}^{K_j^d} \hat{\mathbf{H}}_{k_g^d,j} \mathbf{T}_{i_j^d} \hat{\mathbf{H}}_{k_g^d,j}^H \right) \\ &+ \left(\sum_{j=1}^G \sum_{i=1}^{K_j^u} \hat{\mathbf{H}}_{k_g^d,i_j^u} \mathbf{T}_{i_j^u} \hat{\mathbf{H}}_{k_g^d,i_j^u}^H \right) + (\sigma_{k_g^d}^2 + \tilde{\sigma}_{k_g^d}^2) \mathbf{I}_{M_{ue}}. \end{aligned} \quad (64)$$

$$\begin{aligned} \mathbf{C}_g &= \mathbb{E} \left\{ \mathbf{y}_g \mathbf{y}_g^H \right\} \\ &= \sum_{j=1}^G \sum_{i=1}^{K_j^d} \hat{\mathbf{H}}_{g,i_j^d} \mathbf{T}_{i_j^d} \hat{\mathbf{H}}_{g,i_j^d}^H + \sum_{j=1}^G \hat{\mathbf{H}}_{g,j} \mathbf{T}_j \hat{\mathbf{H}}_{g,j}^H \end{aligned}$$

$$\begin{aligned} &+ \sigma_r^2 \mathcal{D} \left(\sum_{j=1}^G \sum_{i=1}^{K_j^u} \hat{\mathbf{H}}_{g,i_j^u} \mathbf{T}_{i_j^u} \hat{\mathbf{H}}_{g,i_j^u}^H \right) \\ &+ \sum_{j=1}^G \hat{\mathbf{H}}_{g,j} \mathbf{T}_j \hat{\mathbf{H}}_{g,j}^H \Big) + (\sigma_g^2 + \tilde{\sigma}_g^2) \mathbf{I}_{M_{bs}}. \end{aligned} \quad (65)$$

APPENDIX D NSP-BASED METHOD

In this appendix, we will explain how to apply the null-space projection to enhance the ASIC capability of precoders. The processing consists of two steps: 1) obtain the desired precoder; 2) project the precoder to the null-space of the SI channel. The first step could be done using any off-the-shelf precoding techniques, and we use the MWSR beamforming here for consistency, which can be obtained as in [6]. Let $\mathbf{V}_{k_g^d}^\dagger$ denote the desired precoder obtained by the MWSR algorithm, then we project the desired precoder as $\Gamma_{g,D} \Gamma_{g,D}^H \mathbf{V}_{k_g^d}^\dagger$ [15], where $\Gamma_{g,D}$ is the chosen subspace that spans of the eigenvectors associated with the D smallest eigenvalues of $\mathbf{H}_{g,g} \mathbf{H}_{g,g}^H + \kappa_{bs} \mathcal{D}(\mathbf{H}_{g,g} \mathbf{H}_{g,g}^H)$.

APPENDIX E FUNCTIONS FOR EXPLICIT EXPRESSIONS

In this appendix, we will define two functions for explicit expressions.

$$\begin{aligned} \mathcal{F}_1(\mathbf{Y}, \mathbf{X}) &= \mathbf{Y} \mathbf{X} \mathbf{X}^H \mathbf{Y}^H + \sigma_t^2 \mathcal{D}(\mathbf{Y} \mathbf{X} \mathbf{X}^H \mathbf{Y}^H) \\ &+ \sigma_r^2 \mathbf{Y} \mathcal{D}(\mathbf{X} \mathbf{X}^H) \mathbf{Y}^H + \sigma_r^2 \sigma_t^2 \mathcal{D}(\mathbf{Y} \mathcal{D}(\mathbf{X} \mathbf{X}^H) \mathbf{Y}^H) \end{aligned} \quad (66)$$

$$\begin{aligned} \mathcal{F}_2(\mathbf{Z}, \mathbf{Y}, \mathbf{X}) &= \text{tr}(\mathbf{Z}^H \mathbf{Y} (\mathbf{X} \mathbf{X}^H + \sigma_t^2 \mathcal{D}(\mathbf{X} \mathbf{X}^H)) \mathbf{Y}^H \mathbf{Z}) \\ &+ \sigma_r^2 \mathbf{Z}^H \mathcal{D}(\mathbf{Y} (\mathbf{X} \mathbf{X}^H + \sigma_t^2 \mathcal{D}(\mathbf{X} \mathbf{X}^H)) \mathbf{Y}^H) \mathbf{Z} \end{aligned} \quad (67)$$

REFERENCES

- [1] K.E. Kolodziej, B.T. Perry, and J.S. Herd, "In-band full-duplex technology: Techniques and systems survey," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 3025-3041, 2019.
- [2] M. Wang, F. Gao, S. Jin, and H. Lin, "An overview of enhanced massive MIMO with array signal processing techniques," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 886-901, 2019.
- [3] H. Luo, N. Garg, and T. Ratnarajah, "Beamforming Design for In-Band Full-Duplex Multi-Cell Multi-User MIMO Networks With Global and Local CSI," *IEEE Trans. Veh. Technol.*, vol. 72, no. 8, pp. 10218-10233, Aug. 2023.
- [4] A.C. Cirik, R. Wang, Y. Rong, and Y. Hua, "MSE-based transceiver designs for full-duplex MIMO cognitive radios," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2056-2070, 2015.
- [5] J. Kim, W. Choi, and H. Park, "Beamforming for full-duplex multiuser MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2423-2432, 2016.
- [6] P. Aquilina, A.C. Cirik, and T. Ratnarajah, "Weighted sum rate maximization in full-duplex multi-user multi-cell MIMO networks," *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1590-1608, 2017.
- [7] X. Han, X. Su, S. Fu, Y. Gu, Z. Wu, X. Li, and M. Zhao, "RF self-interference cancellation by using photonic technology," *Chin. Opt. Lett.*, vol. 19, no. 7, pp. 073901, 2021.
- [8] 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 Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4183-4196, Jun. 2022.
- [9] H. Luo, A. Bishnu, and T. Ratnarajah, "Design and Analysis of In-Band Full-Duplex Private 5G Networks Using FR2 Band," *IEEE Access*, vol. 9, pp. 166886-166905, Dec. 2021.

- [10] H. Luo, M. Holm, T. Ratnarajah, "On the Performance of Active Analog Self-Interference Cancellation Techniques for Beyond 5G Systems," *China Commun.*, vol. 10, pp. 158-168, Oct. 2021.
- [11] D. Korpi, L. Anttila, and M. Valkama, "Nonlinear self-interference cancellation in MIMO full-duplex transceivers under crosstalk," *EURASIP J. Wirel. Commun. Netw.* no. 1, pp. 24, Dec. 2017.
- [12] I.P. Roberts, J. G. Andrews, and S. Vishwanath, "Hybrid beamforming for millimeter wave full-duplex under limited receive dynamic range," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 7758-7772, 2021.
- [13] M.M. Fadoul, and C.Y. Leow, "Joint Nullspace Projection-Based Interference Mitigation for Full-Duplex Relay-Assisted Multicell Networks," *IEEE Syst. J.*, vol. 14, no. 2, pp. 2392-2399, 2020.
- [14] M.A. Ahmed, C.C. Tsimenidis, and A.F. Al Rawi, "Performance analysis of full-duplex-MRC-MIMO with self-interference cancellation using nullspace-projection," *IEEE Trans. Signal Process.*, vol. 64, no. 12, pp. 3093-3105, 2016.
- [15] X. Huang, A.T. Le, and Y.J. Guo, "Transmit beamforming for communication and self-interference cancellation in full duplex MIMO systems: A trade-off analysis," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3760-3769, 2021.
- [16] C. Masouros, M. Sellathurai, and T. Rantajaraj, "Maximizing Energy-Efficiency in the Vector Precoded MU-MISO Downlink by Selective Perturbation," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4974-4984, Sep. 2014.
- [17] K. Singh, K. Wang, S. Biswas, Z. Ding, F.A. Khan, and T. Ratnarajah, "Resource optimization in full duplex non-orthogonal multiple access systems," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4312-4325, 2019.
- [18] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X.G. Xia, "Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2961-2974, 2018.
- [19] X. Yu, T. Teng, X. Dang, S.H. Leung, and F. Xu, "Joint Power Allocation and Beamforming for Energy-Efficient Design in Multiuser Distributed MIMO Systems," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4128-4143, 2021.
- [20] A.C. Cirik, S. Biswas, S. Vuppala, and T. Ratnarajah, "Beamforming design for full-duplex MIMO interference channels—QoS and energy-efficiency considerations," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4635-4651, 2016.
- [21] Z. Sheng, H.D. Tuan, T.Q. Duong, and H.V. Poor, "Joint power allocation and beamforming for energy-efficient two-way multi-relay communications," *IEEE Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6660-6671, 2017.
- [22] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-Duplex MIMO Relaying: Achievable Rates Under Limited Dynamic Range," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1541-1553, Sep. 2012.
- [23] S. Ahmed, T. Ratnarajah, M. Sellathurai, C.F.N. Cowan, "Iterative Receiver for MIMO-OFDM and their Convergence Behavior," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 461-468, Jan. 2009.
- [24] M.M. Zhao, Q. Wu, M.J. Zhao, and R. Zhang, "Exploiting amplitude control in intelligent reflecting surface aided wireless communication with imperfect CSI," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4216-4231, 2021.
- [25] H. Luo, N. Garg and T. Ratnarajah, "A Channel Frequency Response-Based Secret Key Generation Scheme in In-Band Full-Duplex MIMO-OFDM Systems," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 9, pp. 2951-2965, Sept. 2023.
- [26] J. Choi, J. Park, and N. Lee, "Energy efficiency maximization precoding for quantized massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 21, no. 9, pp. 6803-6817, 2022.
- [27] S. Boyd, S.P. Boyd, and L. Vandenberghe, "Convex optimization," *Cambridge university press*, Mar. 2004.
- [28] *5G; Study on channel model for frequencies from 0.5 to 100 GHz (Release 16)*, document TR 38.901, 3rd Generation Partnership Project, Sophia Antipolis Cedex, France, 3GPP, Jan. 2020.
- [29] *5G; NR; Base Station (BS) radio transmission and reception (Release 15)*, document TR 38.104, 3rd Generation Partnership Project, Sophia Antipolis Cedex, France, 3GPP, May. 2019.
- [30] J. Dai, J. Liu, J. Wang, J. Zhao, C. Cheng, and J.Y. Wang, "Achievable rates for full-duplex massive MIMO systems with low-resolution ADCs/DACs," *IEEE Access*, vol. 7, pp. 24343-24353, 2019.
- [31] D. Lee, and B.W. Min, "Demonstration of Self-Interference Antenna Suppression and RF Cancellation for Full Duplex MIMO Communications," in *Proc. IEEE Wireless Commun. Netw. Conf.*, pp. 1-4, Apr. 2020.

- [32] P. Aquilina, and T. Ratnarajah, "Linear interference alignment in full-duplex MIMO networks with imperfect CSI," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5226-5243, 2017.



Haifeng Luo received the bachelor's degree from the Civil Aviation University of China in 2018 and received the M.S. degree and Ph.D. degree from the University of Edinburgh in 2019 and 2023, respectively. His research interests include signal processing for beyond 5G and 6G wireless networks, full-duplex radios, machine learning-aided communications, and integrated sensing and communications.



Navneet Garg (Member, IEEE) received the B.Tech. degree in electronics and communication engineering from College of Science & Engineering, Jhansi, India, in 2010 and the M.Tech. degree in digital communications from ABV-Indian Institute of Information Technology and Management, Gwalior, in 2012, and the Ph.D. degree from the Department of Electrical Engineering, IIT Kanpur, India, in June 2018. From July 2018 to January 2019, he was a visitor with The University of Edinburgh. From February 2019 to 2020, he was a Research Associate with Heriot-Watt University. Since February 2020, he has been a Research Associate with The University of Edinburgh. His research interests include wireless communications, signal processing, optimization, and machine learning.



Mark Holm received his B.Sc. (hons) in laser physics and optoelectronics from the University of Strathclyde in 1997, before studying for his Ph.D. in physics from the University of Strathclyde, which he received in 2001. He currently works as a technical lead and hardware system architect for Huawei Technologies (Sweden) AB with interests in microwave radio, phased array antennas, full-duplex radio systems, and photonic radios. In the past, he was the microwave lead on AESA Radar systems, senior engineer responsible for GaAs pHEMT modeling, and also a laser and package design engineer for SFP/XENPACK Fibre modules. He is published in the fields of laser design and GaAs device modeling.



Tharmalingam Ratnarajah (Senior Member, IEEE) is a Professor of Digital communications and Signal Processing at the Institute for Digital Communications, University of Edinburgh, Edinburgh, UK. His research interests include signal processing and information-theoretic aspects of beyond 5G cellular networks, full-duplex radio, mmWave communications, random matrix theory, big data analytics and machine learning for wireless networks, statistical and array signal processing, physical-layer secrecy and interference alignment. He has published over 450 Peer-reviewed papers in these areas and holds four US patents. He was the lead coordinator of the European Union (EU) projects HARP (4.6M€) in the area of highly distributed MIMO and ADEL (3.7M€) in the area of licensed shared access. He was also the coordinator of the European Union Future and Emerging Technologies project CROWN (3.4M€) in the area of cognitive radio networks and HIATUS (3.6M€) in the area of interference alignment. Prof Ratnarajah was an associate editor of *IEEE Transactions on Signal Processing*, 2015-2017, and Technical co-chair of the 17th IEEE International Workshop on Signal Processing Advances in Wireless Communications, Edinburgh, UK, 3-6, July 2016. Prof. Ratnarajah is a Fellow of the Higher Education Academy (FHEA).