THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Beamforming Design for In-Band Full-duplex Multi-cell Multi-user MIMO Networks with Global and Local CSI

OPEN ACCESS

Beamforming Design for In-Band Full-duplex Multi-cell Multi-user MIMO Networks with Global and Local CSI

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

Abstract—This paper investigates beamforming techniques for in-band full-duplex multi-cell multi-user (IBFD-MCMU) wireless networks considering hardware impairments (HWIs), channel uncertainty, and limited channel state information (CSI). For the global CSI scenario, we first enhance zero-forcing (ZF), maximum ratio transmission and combining (MRTC), and minimum mean-squared error (MMSE) beamforming to be compatible with multi-antenna users and IBFD base stations. Then, we investigate beamforming schemes for the local CSI assumption where only intra-cell channel knowledge is fully available at the base stations, and inter-cell channels are known statistically due to the limited training resources. With these limitations, two MMSE-based methods are proposed, which regard unknown CSI as random instances (i.e., eMMSE-RI) and noise (i.e., eMMSE-N). We explore the self-interference cancellation (SIC) capability of beamformers and evaluate the performance of these methods in 3GPP scenarios. Numerical results reveal that the enhanced MMSE beamforming can achieve the desired IBFD with low analog SIC depth, inspiring a low-cost IBFD transceiver design. The practical imperfections (e.g., transceiver HWIs, channel uncertainty, limited CSI) decrease the achievable sum rate but do not reduce the IBFD gain. With CSI limitations, eMMSE-RI outperforms eMMSE-N for microcells, where the probability of the presence of LOS is high, whereas eMMSE-N performs better when cell size is enlarged.

Index Terms—Beamforming; channel uncertainty; full-duplex; hardware impairments; local CSI

I. INTRODUCTION

THE next-generation wireless communication networks have stringent requirements on the transmission throughput and latency to support a broad range of emerging applications, e.g., unmanned vehicles and the internet of things [1]. In-band full-duplex (IBFD) is considered a promising candidate technique in beyond 5G communication systems due to its potential to enhance spectral efficiency and reduce end-to-end latency. In recent decades, many studies have successfully demonstrated the potential of full-duplex radios and their applications to emerging techniques [2]–[8]. However, applying the IBFD to multi-cell multi-user multi-input-multi-output (MCMU-MIMO) networks in practice is still challenging due to the additional interference compared to half-duplex (HD) radios, e.g., self-interference (SI) and co-channel interference (CCI), introduced by simultaneous transmission and reception.

With antenna arrays (e.g., MIMO or massive MIMO systems), beamforming is a promising technique to manage the interference [9] and improve the spectral efficiency [1]. The

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

zero-forcing (ZF), maximum ratio transmission and combining (MRTC), and minimum mean-squared error (MMSE) beamforming schemes are commonly used in wireless networks; however, they are not directly applicable for IBFD-MCMU networks due to SI and CCI. In [10], the performance of ZF and MRTC beamforming schemes are evaluated and compared to the optimal scheme for sum rate maximization in the IBFD single-cell multi-user network, demonstrating the IBFD gain and revealing that increasing the number of antennas at IBFD nodes can improve the IBFD gain. It is illustrated in [11] that a ZF/MRT processing can guarantee a balance between sum rate maximization and maintaining the fairness between uplink and downlink rate and improve the performance of the IBFD cloud radio access network. The performance of the ZF precoding and maximum ratio transmission (MRT) precoding is analyzed in MCMU networks in [12] and [13], respectively. Existing studies only perform beamforming at IBFD nodes or consider the single-cell scenario, so they are not compatible with our networks. In the first part of this paper, we enhance the ZF and MRTC beamforming schemes compatible with IBFD-MCMU networks with multi-antenna users.

In contrast, many studies have extended the MMSE beamforming to be compatible with IBFD-MCMU MIMO networks. Authors of [14] propose a beamforming scheme by minimizing the sum of mean-squared errors (MSE) of an IBFD multi-user network and show a significant performance improvement over HD. In [8], the IBFD-MCMU network is further studied, and a weighted sum rate (WSR) maximization beamforming design is proposed with the power constraints by exploiting the relationship between WSR and weighted MMSE. For massive MIMO systems, an MMSE-based hybrid beamforming scheme is proposed in [7] to achieve the IBFD gain with reduced radio frequency (RF) chains. These studies indeed illustrate the performance of MMSE-based beamforming schemes in IBFD-MCMU networks. However, they assume effective self-interference cancellation (SIC) has been realized in both the analog and digital domain by antenna isolation (see [15] and references therein), RF cancellation (see [5], [6] and references therein), and digital cancellation (see [16], [17] and references therein). Realizing the effective SIC by these techniques in antenna arrays significantly increases the complexity and cost of transceivers [6]. In contrast, many studies illustrate the potential of beamforming for SIC that the self-interference is eliminated by the null-space projection [18], zero-forcing beamforming [19], or orthogonal precoder and combiner at the IBFD nodes [20]. Thus, we enhance the

MMSE beamforming design to improve the SIC capability in the first part of the paper, saving the cost of IBFD transceivers.

The potential of MIMO systems strongly depends on the available instantaneous CSI at the base station (BS) [21]. Recently, deep learning has been exploited for CSI feedback with high performance and robustness (see [22] and references therein). In a traditional multi-cell network, the intra-cell CSI of the associated cell can be obtained by channel estimation and CSI feedback and shared between all BSs via a backhaul link [23]. However, acquiring the inter-cell CSI requires the users and BSs to send multiple orthogonal training pilots for channel estimation, increasing the computational overhead and occupying lots of time-frequency resources [24]. Thus, it is challenging and costly to have all the CSI globally (i.e., instantaneous CSI of all intra-cell and inter-cell channels) at the BS for beamforming matrices calculation in practice, particularly for large-scale antenna arrays. In contrast, the statistical CSI is stable and remains unchanged over a fairly long time [25], and can be easily obtained by long-term feedback or averaging over channel samples [26]. In [27], the statistical CSI is exploited to derive the optimal beamforming design in a single-cell multi-user network by maximizing a lower bound on the signal-to-leakage-plus-noise ratio. In [24], a mixed imperfect instantaneous and statistical CSI is considered in a single-cell multi-user MIMO network, and ZF-and MRT-based precoding schemes are proposed. The users utilizing statistical CSI are able to achieve a comparable sum rate with the users utilizing imperfect instantaneous CSI with the proposed methods. Thus, in the second part of the paper, we limit the CSI availability for practical implementation considerations, where the imperfect instantaneous CSI of intra-cell channels and statistical CSI of inter-cell channels are available. With the limited CSI knowledge, we propose two methods based on the MMSE beamforming.

The rest of the paper is organized as follows. In Section II, the transmitted and received signals, HWIs, and channel uncertainty are modeled. Enhanced beamforming designs with global CSI are derived in Section III, and two methods to cope with the local CSI are given in Section IV. Then, the performance of these beamforming schemes is evaluated by simulation results in Section V. Finally, conclusions are drawn in Section VI.

A. Our Contributions

In this paper, we consider practical imperfections, including HWIs, channel uncertainty, and limited CSI, to derive enhanced beamforming schemes for IBFD-MCMU MIMO networks. Our contributions can be summarised as follows.

- Enhanced beamforming methods: We enhance the conventional ZF, MRTC, and MMSE beamforming to be applicable to the system model under consideration and are named eZF, eMRTC, and eMMSE, respectively. The eZF and eMRTC can support IBFD-MCMU networks with multi-antenna users while existing studies only consider single-cell or single-antenna users. The novelty of eMMSE comes from its improvement of analog self-interference cancellation (ASIC) capability due to a different objective function compared to existing MMSE-

based beamforming for IBFD-MCMU networks. It is demonstrated that the eMMSE scheme can realize the desired IBFD gain with around 20dB less ASIC depth than existing studies, inspiring a low-cost IBFD transceiver design. The corresponding optimization problems are formulated and solved, followed by the construction of an iterative algorithm. The convergence behavior is demonstrated, and the computational complexity is analyzed. The eMRTC approaches the performance of eMMSE with extremely low interference, and eZF outperforms eMRTC with higher interference strength. The eMRTC requires effective SIC to realize the maximum IBFD gain as it cannot suppress the SI, whereas eZF and eMMSE require only 30dB of SIC depth since the resulting beamformers can suppress the residual SI. Among these three methods, the eMMSE achieves the highest spectral efficiency at the expense of the highest computational complexity of the three.

- Beamforming methods with limited CSI knowledge: Existing studies on beamforming schemes for IBFD-MCMU networks usually focus on performance optimization with global CSI assumption, while it is challenging to acquire the global CSI for practical implementations. We consider a feasible local CSI scenario, where the inter-cell channels are known statistically due to the limited resources for training pilots. We propose two methods. The first one regards the unknown CSI terms as random instances (i.e., eMMSE-RI), whereas the second one considers the unknown CSI terms as noise (i.e., eMMSE-N). It is concluded that treating unknown CSI as random instances gives better performance for microcells where LOS probability is high. On the other hand, the second method yields better performance with larger cells.

- Simulations with practical imperfections: We consider various practical imperfections in this paper, i.e., practical dynamic range-limited transceivers, channel uncertainty due to limited training resources, limited ASIC depth due to its high complexity, etc. We perform extensive simulations to compare the above five methods. Their achievable sum rates are plotted with varying SIC depths, transmitter HWIs, receiver HWIs, channel uncertainty, cell size, number of cells, number of users, number of antennas, and transmission power. In addition, four practical 3GPP scenarios (indoor hotspot, urban micro, urban macro, and rural macro) are considered for evaluation. We include nearly all imperfections considered by relevant studies and show their impact in multiple cases, which is meaningful for practical implementation.

B. Notations

$\mathbf{A}$, $\mathbf{a}$, and a represent a matrix, a vector, and a scalar, respectively. $tr(\mathbf{A})$, $Cov(\mathbf{A})$, $|\mathbf{A}|$, $\|\mathbf{A}\|_F$, $\mathbf{A}^H$, $\mathbf{A}^T$, and $\mathbf{A}^{-1}$ denote the trace, covariance matrix, determinant, Frobenius norm, Hermitian, transpose, and inverse of matrix $\mathbf{A}$. $\mathbf{I}$ represents an identity matrix. a^* represents the conjugate complex of a. $\mathcal{D}(\mathbf{A})$ denotes a diagonal matrix containing the elements along the diagonal of $\mathbf{A}$. $[\mathbf{A}]_{i,j}$ represents the element of the ith row and j^{th} column of matrix $\mathbf{A}$. $\mathbf{A}(:,j)$ denotes the jth column

TABLE I
SYMBOLS AND CORRESPONDING DESCRIPTIONS

Symbol	Description
$\mathbf{y}_\dagger$	received signal
$\tilde{\mathbf{y}}_\dagger$	interference and noise of corresponding user
$\mathbf{c}_\dagger$ $(\mathbf{e}_\dagger)$	transmitter (receiver) HWIs
$\mathbf{n}_\dagger$	additive white Gaussian noise (AWGN)
$\Delta_{\dagger,*}$	channel uncertainty
$\bar{\sigma}^2_{\dagger,*,k}$ $(= \bar{\sigma}^2_{\dagger,*})$	strength of channel uncertainty
$\bar{\sigma}^2_\dagger$	strength of unknown CSI regarded as noise
σ^2_t (σ^2_r)	transmitter (receiver) HWIs factor
σ^2_{ue} (σ^2_{bs})	AWGN strength of UEs (BSs)
$\varrho_{\dagger,*}$	pathloss between nodes $\dagger$ and $*$
$\kappa_{\dagger,*}$	Rician factor
$\mathbf{C}^2_\dagger$ $(\bar{\mathbf{C}}^2_\dagger)$	received signal covariance matrix with global (local) CSI
$R_\dagger$ $(\bar{R}_\dagger)$	achievable rate with global (local) CSI
P_{bs} (P_{ue})	transmission power of BSs (UEs)
$P_{s,\dagger}$	received power of signal of interest

of matrix $\mathbf{A}$. $\underline{\nu}_{1:d}[\mathbf{A}]$ and $\bar{\nu}_{1:d}[\mathbf{A}]$ represents the eigenvectors of $\mathbf{A}$ corresponding to d minimum or maximum eigenvalues, respectively. $\mathbb{E}\{\cdot\}$ denotes the mathematical expectation. $\mathcal{CN}(0, \sigma^2)$ denotes a complex normal distribution with zero mean and variance of σ^2. In addition, some important and similar symbols are listed in Table I to help the readers.

II. SYSTEM MODEL

We consider an IBFD-MCMU network, where G base stations (BS) operate in the in-band full-duplex mode, and the gth base station serves K^d_g downlink (DL) users and K^u_g uplink (UL) users operating in the half-duplex mode.

A. Transmitted Signals

1) Downlink: Assume each base station has N_{bs} transmitting (TX) antennas and M_{bs} receiving (RX) antennas, and each downlink user has M_{ue} RX antennas; there are b_d data streams transmitted on each subcarrier. The transmitted signal from the gth BS on a single subcarrier can be written as (the subcarrier index is omitted throughout this paper for simplicity)

$$\mathbf{x}_g = \sum_{k=1}^{K^d_g} \left(\sqrt{\frac{P_{bs}}{K^d_g}} \mathbf{V}_{k^d_g} \mathbf{s}_{k^d_g} + \mathbf{c}_{k^d_g} \right) = \sqrt{\frac{P_{bs}}{K^d_g}} \mathbf{V}_g \mathbf{s}_g + \mathbf{c}_g, \quad (1)$$

where P_{bs} is the transmission power of the BS; $\mathbf{s}_{k^d_g} \in \mathbb{C}^{b_d \times 1}$ denotes the transmitted data streams with zero mean and $\mathbb{E}\left\{\mathbf{s}_{k^d_g}\mathbf{s}^H_{k^d_g}\right\} = \frac{1}{b_d}\mathbf{I}$; $\mathbf{V}_{k^d_g} \in \mathbb{C}^{N_{bs} \times b_d}$ is the associated precoding matrix, whose column is normalised such that $\mathbf{V}^H_{k^d_g}(:,j)\mathbf{V}_{k^d_g}(:,j) = 1$ for all j, yielding $tr\left(\mathbf{V}_{k^d_g}\mathbf{V}^H_{k^d_g}\right) = b_d$ and $\mathbb{E}\left\{\left\|\mathbf{V}_{k^d_g}\mathbf{s}_{k^d_g}\right\|^2_F\right\} = 1$. The transmitted signal is an accumulation of downlink payload, so the expression can be simplified with $\mathbf{s}_g = \left[\mathbf{s}^T_{1^d_g}, \ldots, \mathbf{s}^T_{K^d_g}\right]^T$ and $\mathbf{V}_g = \left[\mathbf{V}_{1^d_g}, \ldots, \mathbf{V}_{K^d_g}\right]$. The vector $\mathbf{c}_g \in \mathbb{C}^{N_{bs} \times 1}$ denotes the transmitter hardware impairments (HWIs), which is a natural consequence of imperfect digital-to-analog converters, oscillators, and power amplifiers. It is experimentally demonstrated that a circular complex Gaussian model can closely approximate the combined effects

of these non-ideal components (see [8], [28] and reference therein) as

$$\mathbf{c}_g \sim \mathcal{CN}\left(0, \sigma^2_t \frac{P_{bs}}{K^d_g} \mathbb{E}\left\{\mathcal{D}\left(\mathbf{V}_g \mathbf{s}_g \mathbf{s}^H_g \mathbf{V}^H_g\right)\right\}\right)$$
$$= \mathcal{CN}\left(0, \sigma^2_t \frac{P_{bs}}{K^d_g b_d} \mathcal{D}\left(\mathbf{V}_g \mathbf{V}^H_g\right)\right), \quad (2)$$

where σ^2_t describes the dynamic range of transmitters and high-quality transmitters have a small value of σ^2_t.

2) Uplink: Assume each uplink user equipment (UE) has N_{ue} TX antennas to send b_u data streams on a single subcarrier. The transmitted signal from the kth uplink user in the gth cell can be written as

$$\mathbf{x}_{k^u_g} = \sqrt{P_{ue}}\mathbf{V}_{k^u_g}\mathbf{s}_{k^u_g} + \mathbf{c}_{k^u_g}, \quad (3)$$

where P_{ue} is the transmission power of the UE for the uplink payload; $\mathbf{s}_{k^u_g} \in \mathbb{C}^{b_u \times 1}$ is the transmitted data streams with zero mean and $\mathbb{E}\left\{\mathbf{s}_{k^u_g}\mathbf{s}^H_{k^u_g}\right\} = \frac{1}{b_u}\mathbf{I}$; $\mathbf{V}_{k^u_g} \in \mathbb{C}^{N_{ue} \times b_u}$ is the associated precoding matrix, and the norm of each column of $\mathbf{V}_{k^u_g}$ is normalised to be 1, such that $\mathbf{V}^H_{k^u_g}(:,j)\mathbf{V}_{k^u_g}(:,j) = 1$ for all j, resulting $tr\left(\mathbf{V}_{k^u_g}\mathbf{V}^H_{k^u_g}\right) = b_u$ and $\mathbb{E}\left\{\left\|\mathbf{V}_{k^u_g}\mathbf{s}_{k^u_g}\right\|^2_F\right\} = 1$. The vector $\mathbf{c}_{k^u_g} \in \mathbb{C}^{N_{ue} \times 1}$ denotes the transmitter HWIs at this uplink UE, which can be modeled as

$$\mathbf{c}_{k^u_g} \sim \mathcal{CN}\left(0, \sigma^2_t P_{ue} \mathbb{E}\left\{\mathcal{D}\left(\mathbf{V}_{k^u_g}\mathbf{s}_{k^u_g}\mathbf{s}^H_{k^u_g}\mathbf{V}^H_{k^u_g}\right)\right\}\right)$$
$$= \mathcal{CN}\left(0, \sigma^2_t \frac{P_{ue}}{b_u}\mathcal{D}\left(\mathbf{V}_{k^u_g}\mathbf{V}^H_{k^u_g}\right)\right), \quad (4)$$

where σ^2_t is the HWIs factor that is assumed to be identical for the BSs and UEs.

3) Covariance matrices: The covariance matrices of the transmitted signals can be written as

$$\mathbf{T}_g = \mathbb{E}\left\{\mathbf{x}_g\mathbf{x}^H_g\right\} = \frac{P_{bs}}{K^d_g b_d}\left(\mathbf{V}_g\mathbf{V}^H_g + \sigma^2_t\mathcal{D}\left(\mathbf{V}_g\mathbf{V}^H_g\right)\right),$$

$$\mathbf{T}_{k^u_g} = \mathbb{E}\left\{\mathbf{x}_{k^u_g}\mathbf{x}^H_{k^u_g}\right\} = \frac{P_{ue}}{b_u}\left(\mathbf{V}_{k^u_g}\mathbf{V}^H_{k^u_g} + \sigma^2_t\mathcal{D}\left(\mathbf{V}_{k^u_g}\mathbf{V}^H_{k^u_g}\right)\right).$$

B. Channel Model

1) Wireless channels: In the multi-cell system, there are Rician and Rayleigh channels depending on the presence of line-of-sight (LOS) components, whose models are given as follows (each matrix $\mathbf{H}$ contains i.i.d zero-mean circularly symmetric complex Gaussian random variables of variance 0.5 per dimension).

- Rician channels: if there are LOS paths between the transmitter and receiver, the channel can be modeled by the Rician channel as [29]

$$\mathbf{H}_{r,t} = \sqrt{\varrho_{r,t}}\left(\sqrt{\frac{\kappa_{r,t}}{\kappa_{r,t}+1}}\mathbf{H}^{LOS}_{r,t} + \sqrt{\frac{1}{\kappa_{r,t}+1}}\mathbf{H}\right) \quad (5)$$

where $\varrho_{r,t}$ is the pathloss between the rth receiving node and the tth transmitting node (the receiving and transmitting nodes could be any BS or UE); $\kappa_{r,t}$ is the associated Rician factor; $\mathbf{H}^{LOS}_{r,t}$ denotes the LOS component such that $\mathbf{H}^{LOS}_{r,t} = \mathbf{a}^H_{M_r}(\theta_{\text{AOA},1})\mathbf{a}_{N_t}(\theta_{\text{AOD},1})$, where
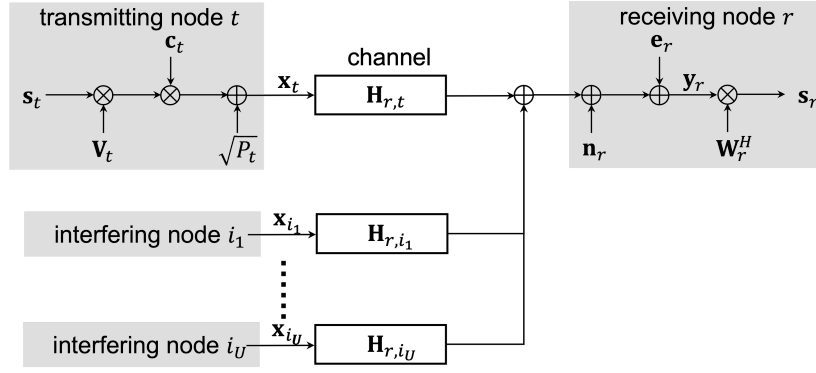
Fig. 1. Our model of the received IBFD-MCMU MIMO network with transceiver HWIs (the receiving node itself could be one of the interfering nodes if it operates in IBFD mode).

$\mathbf{a}_N(\theta) = \left[1, e^{j2\pi \frac{d}{\lambda_s}\sin\theta}, \cdots, e^{j2\pi \frac{d}{\lambda_s}(N-1)\sin\theta}\right]$ with θ, d, and λ_s representing the angle of departure (AoD) or angle of arrival (AoA), the distance between neighboring antennas, and signal wavelength, respectively.

- Rayleigh channels: if there are no LOS paths between the transmitter and receiver, the channel can be modeled by the Rayleigh channel, which has i.i.d. elements as

$$\mathbf{H}_{r,t} = \sqrt{\varrho_{r,t}}\,\mathbf{H}. \tag{6}$$

2) Channel uncertainty model: The channel uncertainty describes the difference between the actual CSI and the acquired CSI for calculations due to limited training resources [30]. It mainly comes from imperfect channel estimation, i.e., estimation errors caused by noise, interpolation, etc. Assume $\hat{\mathbf{H}}_{r,t}$ is the estimated channel between the rth receiver and the tth transmitter, while $\mathbf{H}_{r,t}$ is the corresponding actual wireless channel, they can be related as

$$\mathbf{H}_{r,t} = \hat{\mathbf{H}}_{r,t} + \Delta_{r,t} \tag{7}$$

where $\Delta_{r,t}$ is the channel uncertainty. The channel uncertainty can be described by Gaussian models [30], [31] with similar variances across each entry of matrices, such that $\Delta_{r,t}$ has i.i.d. elements as $[\Delta_{r,t}]_{i,j} \sim \mathcal{CN}\left(0, \tilde{\sigma}_{r,t}^2\right) \; \forall \; i,j$. We assume the variance of the channel uncertainty is identical for all subcarriers without loss of generality, while the channel uncertainties themselves are different on different subcarriers.

C. Received Signals

1) Downlink: The received signal at the kth downlink user in the gth cell can be denoted as

$$\mathbf{y}_{k_g^d} = \sum_{j=1}^G \mathbf{H}_{k_g^d,j}\mathbf{x}_j + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \mathbf{H}_{k_g^d,i_j^u}\mathbf{x}_{i_j^u} + \mathbf{n}_{k_g^d} + \mathbf{e}_{k_g^d}$$
$$= \sqrt{\frac{P_{bs}}{K_g^d}}\hat{\mathbf{H}}_{k_g^d,g}\mathbf{V}_{k_g^d}\mathbf{s}_{k_g^d} + \tilde{\mathbf{y}}_{k_g^d}, \tag{8}$$

where $\mathbf{H}_{k_g^d,j}$ represents the downlink channels from the jth BS to the DL user k_g^d, $\mathbf{H}_{k_g^d,i_j^u}$ is channel from the UL user i_j^u to the DL user k_g^d; $\mathbf{n}_{k_g^d}$ represents the additive white Gaussian noise (AWGN) distributed as $\mathbf{n}_{k_g^d} \sim \mathcal{CN}\left(\mathbf{0}, \sigma_{ue}^2\mathbf{I}\right)$; $\mathbf{e}_{k_g^d}$

denotes the receiver HWIs; and $\tilde{\mathbf{y}}_{k_g^d} = \Delta_{k_g^d,g}\mathbf{x}_g + (\hat{\mathbf{H}}_{k_g^d,g} + \Delta_{k_g^d,g})\mathbf{c}_{k_g^d} + \sum_{j=1}^G \sum_{i=1 (i,j)\neq(k,g)}^{K_j^d}(\hat{\mathbf{H}}_{k_g^d,j} + \Delta_{k_g^d,j})\mathbf{x}_{i_j^d} + \sum_{j=1}^G \sum_{i=1}^{K_j^u}(\hat{\mathbf{H}}_{k_g^d,i_j^u} + \Delta_{k_g^d,i_j^u})\mathbf{x}_{i_j^u} + \mathbf{n}_{k_g^d} + \mathbf{e}_{k_g^d}$ represent the interference and noise terms. The receiver HWIs are a natural consequence of imperfect analog-to-digital converters, oscillators, and low-noise amplifiers. Experimental measurements demonstrate that a circular complex Gaussian model can closely approximate the combined effects, so the receiver HWIs can be given as [8], [28]

$$\mathbf{e}_{k_g^d} \sim \mathcal{CN}\left(0, \sigma_r^2 \mathbb{E}\left\{\mathcal{D}\left((\mathbf{y}_{k_g^d} - \mathbf{e}_{k_g^d})(\mathbf{y}_{k_g^d} - \mathbf{e}_{k_g^d})^H\right)\right\}\right), \tag{9}$$

where σ_r^2 describes the dynamic range of receivers and high-quality receivers have a small value of σ_r^2.

2) Uplink: The received signal at the gth BS can be denoted as

$$\mathbf{y}_g = \sum_{j=1}^G \mathbf{H}_{g,j}\mathbf{x}_j + \sum_{j=1}^G \sum_{i=1}^{K_j^u} \mathbf{H}_{g,i_j^u}\mathbf{x}_{i_j^u} + \mathbf{n}_g + \mathbf{e}_g$$
$$= \sqrt{P_{ue}}\hat{\mathbf{H}}_{g,k_g^u}\mathbf{V}_{k_g^u}\mathbf{s}_{k_g^u} + \tilde{\mathbf{y}}_{k_g^u}, \tag{10}$$

where $\mathbf{H}_{g,j}$ represents the channels from the jth BS to the gth BS, specially, $\mathbf{H}_{g,g}$ is the self-interference (SI) channel at the gth BS; $\mathbf{H}_{g,i_j^u}$ denotes uplink channels from the UL user i_j^u to the gth BS; $\mathbf{n}_g$ is the AWGN distributed as $\mathbf{n}_g \sim \mathcal{CN}\left(\mathbf{0}, \sigma_{bs}^2\mathbf{I}\right)$; $\mathbf{e}_g$ denotes the receiver HWIs; and $\tilde{\mathbf{y}}_{k_g^u} = \Delta_{g,k_g^u}\mathbf{x}_{k_g^u} + (\hat{\mathbf{H}}_{g,k_g^u} + \Delta_{g,k_g^u})\mathbf{c}_{k_g^u} + \sum_{j=1}^G(\hat{\mathbf{H}}_{g,j} + \Delta_{g,j})\mathbf{x}_j + \sum_{j=1}^G \sum_{i=1 (i,j)\neq(k,g)}^{K_j^u}(\hat{\mathbf{H}}_{g,i_j^u} + \Delta_{g,i_j^u})\mathbf{x}_{i_j^u} + \mathbf{n}_g + \mathbf{e}_g$ represent the interference and noise terms. The receiver HWIs at base stations can be denoted as

$$\mathbf{e}_g \sim \mathcal{CN}\left(0, \sigma_r^2 \mathbb{E}\left\{\mathcal{D}\left((\mathbf{y}_g - \mathbf{e}_g)(\mathbf{y}_g - \mathbf{e}_g)^H\right)\right\}\right), \tag{11}$$

where σ_r^2 is the receiver HWIs factor that is assumed to be identical for both BSs and UEs.

Note: It should be noted that there can exist inter-carrier interference (ICI) due to time variations in high delay and Doppler spread environments. In [32], a piece-wise linear approximation is used to mitigate the ICI among pilot symbols. Without transceiver HWIs, the variance of interference after

mitigation is shown to be proportional to the product-to-average noise variance across subcarriers. With transceiver HWIs present, the coefficient of impairments adds to the noise variance, increasing the strength of ICI. This ICI term cannot be removed due to unknown HWIs coefficients. Thus, we can only consider its effect in design and analysis using an increased effective noise variance. Based on the model given in [32], it can be derived that ICI increases the effective noise variance by up to $(K-1)\varrho_{i,j}P_i\sigma_j^2(\sigma_r^2+\sigma_t^2+\sigma_r^2\sigma_t^2)$ with HWIs present, where K is the number of subcarriers, and i and j denote the transmitting and receiving nodes. Thus, we can use a larger transceiver HWIs factor to reflect the effects of ICI while designing and analyzing the system.

D. Achievable Rate

The achievable sum rate is used as the metric to evaluate the performance throughout this paper.

1) Downlink Rate: With the antenna array, receivers can perform combining to enhance the signal-to-interference-plus-noise ratio (SINR) for a higher achievable sum rate. The achievable downlink rate of the kth downlink user in the gth cell is given as [8]

$$R_{k_g^d} = \log_2\left|\mathbf{I}_{b_d} + \mathbf{U}_{k_g^d}^H\mathbf{C}_{k_g^d,S}\mathbf{U}_{k_g^d}\left(\mathbf{U}_{k_g^d}^H\mathbf{C}_{k_g^d,X}\mathbf{U}_{k_g^d}\right)^{-1}\right|,$$
(12)

where $\mathbf{U}_{k_g^d} \in \mathbb{C}^{M_{ue}\times b_d}$ is the combiner consisting of the combining coefficients, $\mathbf{C}_{k_g^d,S}$ denotes the covariance matrix of the signal of interest, and $\mathbf{C}_{k_g^d,X}$ represents the covariance matrix of the unknown part of the signal (i.e., interference and noise terms), which are given as

$$\mathbf{C}_{k_g^d,S} = Cov\left(\sqrt{\frac{P_{bs}}{K_g^d}}\hat{\mathbf{H}}_{k_g^d,g}\mathbf{V}_{k_g^d}\mathbf{s}_{k_g^d}\right)$$
$$= \frac{P_{bs}}{K_g^d b_d}\hat{\mathbf{H}}_{k_g^d,g}\mathcal{D}\left(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H\right)\hat{\mathbf{H}}_{k_g^d,g}^H.$$
(13)

Using the results in Appendix A, $\mathbf{C}_{k_g^d,X}$ can be written as

$$\mathbf{C}_{k_g^d,X} = Cov\left(\mathbf{y}_{k_g^d} - \sqrt{\frac{P_{bs}}{K_g^d}}\hat{\mathbf{H}}_{k_g^d,g}\mathbf{V}_{k_g^d}\mathbf{s}_{k_g^d}\right)$$
$$= \mathbf{C}_{k_g^d} + \sigma_r^2\cdot\mathcal{D}\left(\mathbf{C}_{k_g^d}\right) - \mathbf{C}_{k_g^d,S}.$$
(14)

2) Uplink Rate: Similarly, the achievable uplink rate of the kth uplink user in the gth cell can be given as

$$R_{k_g^u} = \log_2\left|\mathbf{I}_{b_u K_g^u} + \mathbf{U}_{k_g^u}^H\mathbf{C}_{k_g^u,S}\mathbf{U}_{k_g^u}\left(\mathbf{U}_{k_g^u}^H\mathbf{C}_{k_g^u,X}\mathbf{U}_{k_g^u}\right)^{-1}\right|,$$
(15)

where $\mathbf{U}_{k_g^u} \in \mathbb{C}^{M_{bs}\times b_u}$; $\mathbf{C}_{k_g^u,S}$ and $\mathbf{C}_{k_g^u,X}$ are given as

$$\mathbf{C}_{k_g^u,S} = Cov\left(\sqrt{P_{ue}}\hat{\mathbf{H}}_{g,k_g^u}\mathbf{V}_{k_g^u}\mathbf{s}_{k_g^u}\right)$$
$$= \frac{P_{ue}}{b_u}\hat{\mathbf{H}}_{g,k_g^u}\mathcal{D}\left(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H\right)\hat{\mathbf{H}}_{g,k_g^u}^H.$$
(16)

Using the results in Appendix A, $\mathbf{C}_{k_g^u,X}$ can be written as

$$\mathbf{C}_{k_g^u,X} = Cov\left(\mathbf{y}_g - \sqrt{P_{ue}}\hat{\mathbf{H}}_{g,k_g^u}\mathbf{V}_{k_g^u}\mathbf{s}_{k_g^u}\right)$$
$$= \mathbf{C}_g + \sigma_r^2\cdot\mathcal{D}\left(\mathbf{C}_g\right) - \mathbf{C}_{k_g^u,S}.$$
(17)

III. BEAMFORMING WITH GLOBAL CSI

In this section, we consider that imperfect instantaneous CSI of intra-cell and inter-cell channels (i.e., global CSI) is available for beamforming matrices calculation. We enhance the conventional zero-forcing (ZF), maximum ratio transmission and combining (MRTC), and minimum mean-squared error (MMSE) beamforming schemes to IBFD-MCMU MIMO networks.

A. Enhanced Zero-Forcing (eZF)

The objective of zero-forcing beamforming is to eliminate interference, which has stringent requirements on the number of transmitting and receiving antennas to leave null space, i.e., the number of antennas should be larger than the number of total interference. The stringent requirements may not be satisfied at each node in the IBFD-MCMU network. The users usually have very limited antennas, so they cannot perform the conventional zero-forcing precoding or combining, and base stations need a large number of antennas to eliminate the self-interference. Thus, the conventional ZF beamforming is not compatible with IBFD-MCMU networks. Based on the idea of ZF beamforming, we enhance it to be compatible with our networks by minimizing interference. At the base station, the precoder for downlink payload and the combiner for uplink payload should minimize the interference at associated users, which can be formulated as

$$\min_{\left\|\mathbf{V}_{k_g^d}\right\|_F^2=b_d}\left\|\mathbf{\Gamma}_{k_g^d}\mathbf{V}_{k_g^d}\right\|_F^2, \quad \min_{\left\|\mathbf{U}_{k_g^u}\right\|_F^2=b_u}\left\|\mathbf{U}_{k_g^u}^H\mathbf{\Pi}_{k_g^u}\right\|_F^2, \quad (18)$$

where $\mathbf{\Gamma}_{k_g^d} = \left[\mathbf{U}_{1_1^d}^H\hat{\mathbf{H}}_{1_1^d,g};\ldots;\mathbf{U}_{(k-1)_g^d}^H\hat{\mathbf{H}}_{(k-1)_g^d,g};\mathbf{U}_{(k+1)_g^d}^H\right.$

$\hat{\mathbf{H}}_{(k+1)_g^d,g};\ldots;\mathbf{U}_{K_G^d}^H\hat{\mathbf{H}}_{K_G^d,g};\mathbf{U}_{1_1^u}^H\hat{\mathbf{H}}_{1,g};\ldots;\mathbf{U}_{K_G^u}^H\hat{\mathbf{H}}_{G,g}\big];$

$\mathbf{\Pi}_{k_g^u} = \left[\hat{\mathbf{H}}_{g,1_1^u}\mathbf{V}_{1_1^u},\ldots,\hat{\mathbf{H}}_{g,(k-1)_g^u}\mathbf{V}_{(k-1)_g^u},\hat{\mathbf{H}}_{g,(k+1)_g^u}\right.$

$\mathbf{V}_{(k+1)_g^u},\ldots,\hat{\mathbf{H}}_{g,K_G^u}\mathbf{V}_{K_G^u},\hat{\mathbf{H}}_{g,1}\mathbf{V}_{1_1^d},\ldots,\hat{\mathbf{H}}_{g,G}\mathbf{V}_{K_G^d}\big].$ The solution can be given as [33]

$$\mathbf{V}_{k_g^d} = \underline{\nu}_{1:b_d}\left[\mathbf{\Gamma}_{k_g^d}^H\mathbf{\Gamma}_{k_g^d}\right], \quad (19)$$

$$\mathbf{U}_{k_g^u} = \underline{\nu}_{1:b_u}\left[\mathbf{\Pi}_{k_g^u}\mathbf{\Pi}_{k_g^u}^H\right], \quad (20)$$

where $\underline{\nu}_{1:b}[\mathbf{A}]$ denotes the b eigenvectors associated with the bth smallest eigenvalues of matrix $\mathbf{A}$. Similarly, the precoder for uplink payload and the combiner for downlink payload should minimize the interference at associated users, which can be described as

$$\min_{\left\|\mathbf{V}_{k_g^u}\right\|_F^2=b_u}\left\|\mathbf{\Gamma}_{k_g^u}\mathbf{V}_{k_g^u}\right\|_F^2, \quad \min_{\left\|\mathbf{U}_{k_g^d}\right\|_F^2=b_d}\left\|\mathbf{U}_{k_g^d}^H\mathbf{\Pi}_{k_g^d}\right\|_F^2, \quad (21)$$

where $\mathbf{\Gamma}_{k_g^u} = \left[\mathbf{U}_{1_1^u}^H\hat{\mathbf{H}}_{1,k_g^u};\ldots;\mathbf{U}_{(k-1)_g^u}^H\hat{\mathbf{H}}_{g,k_g^u};\mathbf{U}_{(k+1)_g^u}^H\right.$

$\hat{\mathbf{H}}_{g,k_g^u};\ldots;\mathbf{U}_{K_G^u}^H\hat{\mathbf{H}}_{G,k_g^u};\mathbf{U}_{1_1^d}^H\hat{\mathbf{H}}_{1_1^d,k_g^u};\ldots;\mathbf{U}_{K_G^d}^H\hat{\mathbf{H}}_{K_G^d,k_g^u}\big];$

$\mathbf{\Pi}_{k_g^d} = \left[\hat{\mathbf{H}}_{k_g^d,1}\mathbf{V}_{1_1^d},\ldots,\hat{\mathbf{H}}_{k_g^d,g}\mathbf{V}_{(k-1)_g^d},\hat{\mathbf{H}}_{k_g^d,g}\right.$

$$\mathbf{V}_{(k+1)_g^d}, \ldots, \hat{\mathbf{H}}_{k_g^d, G} \mathbf{V}_{K_G^d}, \hat{\mathbf{H}}_{k_g^d, 1_1^u} \mathbf{V}_{1_1^u}, \ldots, \hat{\mathbf{H}}_{k_g^d, K_G^u} \mathbf{V}_{K_G^u} \Big].$$

The solution can also be given as

$$\mathbf{V}_{k_g^u} = \underline{\nu}_{1:b_u} \left[\mathbf{\Gamma}_{k_g^u}^H \mathbf{\Gamma}_{k_g^u} \right], \tag{22}$$

$$\mathbf{U}_{k_g^d} = \underline{\nu}_{1:b_d} \left[\mathbf{\Pi}_{k_g^d} \mathbf{\Pi}_{k_g^d}^H \right]. \tag{23}$$

The loss function of the eZF beamforming design can be given by the sum of the objectives (i.e., interference strengths) of these minimization problems as

$$
\begin{aligned}
\mathcal{L}_{\text{eZF}} = &\sum_{g=1}^{G} \sum_{k=1}^{K_g^d} \left(\left\| \mathbf{\Gamma}_{k_g^d} \mathbf{V}_{k_g^d} \right\|_F^2 + \left\| \mathbf{U}_{k_g^d}^H \mathbf{\Pi}_{k_g^d} \right\|_F^2 \right) \\
&+ \sum_{g=1}^{G} \sum_{k=1}^{K_g^u} \left(\left\| \mathbf{\Gamma}_{k_g^u} \mathbf{V}_{k_g^u} \right\|_F^2 + \left\| \mathbf{U}_{k_g^u}^H \mathbf{\Pi}_{k_g^u} \right\|_F^2 \right).
\end{aligned}
\tag{24}
$$

The optimized variables should always reduce the loss as it reduces one of the objective functions.

B. Enhanced Maximum Ratio Transmission and Combining (eMRTC)

The objective of the enhanced maximum ratio transmission and combining beamforming is to maximize the signal of interest, which can be described as

$$
\begin{aligned}
\max_{\mathbf{U}_{k_g^d}, \mathbf{V}_{k_g^d}, \mathbf{U}_{k_g^u}, \mathbf{V}_{k_g^u}} &\sum_{g=1}^{G} \sum_{k=1}^{K_g^d} \mathbb{E} \left\{ \left\| \mathbf{U}_{k_g^d}^H \mathbf{H}_{k_g^d, g} \mathbf{V}_{k_g^d} \mathbf{s}_{k_g^d} \right\|_F^2 \right\} \\
&+ \sum_{g=1}^{G} \sum_{k=1}^{K_g^u} \mathbb{E} \left\{ \left\| \mathbf{U}_{k_g^u}^H \mathbf{H}_{g, k_g^u} \mathbf{V}_{k_g^u} \mathbf{s}_{k_g^u} \right\|_F^2 \right\}. \\
\text{s.t.} \quad &\left\| \mathbf{V}_{k_g^d} \right\|_F^2 = \left\| \mathbf{U}_{k_g^d} \right\|_F^2 = b_d \\
&\left\| \mathbf{V}_{k_g^u} \right\|_F^2 = \left\| \mathbf{U}_{k_g^u} \right\|_F^2 = b_u
\end{aligned}
\tag{25}
$$

The objective function is convex and differentiable to a specific precoder or combiner. Fixing other beamforming matrices, the optimization problem with respect to a specific precoder or combiner has the same form as the maximization problem $\max_{\|\mathbf{X}\|_F^2 = b} \|\mathbf{AX}\|_F^2$. Thus, the solutions can be given as [33]

$$\mathbf{V}_{k_g^d} = \bar{\nu}_{1:b_d} \left[\mathbf{H}_{k_g^d, g}^H \mathbf{U}_{k_g^d} \mathbf{U}_{k_g^d}^H \mathbf{H}_{k_g^d, g} \right], \tag{26}$$

$$\mathbf{U}_{k_g^d} = \bar{\nu}_{1:b_d} \left[\mathbf{V}_{k_g^d}^H \mathbf{H}_{k_g^d, g}^H \mathbf{H}_{k_g^d, g} \mathbf{V}_{k_g^d} \right], \tag{27}$$

$$\mathbf{V}_{k_g^u} = \bar{\nu}_{1:b_u} \left[\mathbf{H}_{g, k_g^u}^H \mathbf{U}_{k_g^u} \mathbf{U}_{k_g^u}^H \mathbf{H}_{g, k_g^u} \right], \tag{28}$$

$$\mathbf{U}_{k_g^u} = \bar{\nu}_{1:b_u} \left[\mathbf{V}_{k_g^u}^H \mathbf{H}_{g, k_g^u}^H \mathbf{H}_{g, k_g^u} \mathbf{V}_{k_g^u} \right], \tag{29}$$

where $\bar{\nu}_{1:b}[\mathbf{A}]$ denotes the b eigenvectors associated with the bth largest eigenvalues of matrix $\mathbf{A}$. The loss function of eMRTC beamforming can be given by the reciprocal of the objective function as it is a maximization problem, denoted as

$$\mathcal{L}_{\text{eMRTC}} = \frac{1}{\sum_{g=1}^{G} \sum_{k=1}^{K_g^d} P_{s, k_g^d} + \sum_{g=1}^{G} \sum_{k=1}^{K_g^u} P_{s, k_g^u}}, \tag{30}$$

where the denominator is equal to the objective function with $P_{s, k_g^d} = \frac{1}{b_d} \left\| \mathbf{U}_{k_g^d}^H \mathbf{H}_{k_g^d, g} \mathbf{V}_{k_g^d} \right\|_F^2$ and $P_{s, k_g^u} =$

$\frac{1}{b_u} \left\| \mathbf{U}_{k_g^u}^H \mathbf{H}_{g, k_g^u} \mathbf{V}_{k_g^u} \right\|_F^2$. The optimized variables should reduce the loss as they maximize the objective.

C. Enhanced Minimum Mean-Squared Error (eMMSE)

The objective of the conventional MMSE beamforming scheme is to minimize the sum of mean-squared errors (MSE) of the network. Previous studies usually assume effective SIC has been realized in both analog and digital domains at IBFD nodes [7], [8], [14]. However, analog self-interference cancellation (ASIC) has high complexity and cost due to additional hardware. Thus, we would like to explore the precoders to help suppress the SI in the propagation domain to save the cost of ASIC. To this end, we enhance the conventional MMSE beamforming by minimizing the sum of MSE plus the received analog residual SI, which can be described as

$$
\begin{aligned}
\min_{\mathbf{V}_{k_g^d}, \mathbf{V}_{k_g^u}, \mathbf{U}_{k_g^d}, \mathbf{U}_{k_g^u}} &\sum_{g=1}^{G} \sum_{k=1}^{K_g^d} \mathbb{E} \left\{ \left\| \mathbf{s}_{k_g^d} - \mathbf{U}_{k_g^d}^H \mathbf{y}_{k_g^d} \right\|_F^2 \right\} \\
&+ \sum_{g=1}^{G} \mathbb{E} \left\{ \left\| \mathbf{s}_{u, g} - \mathbf{U}_g^H \mathbf{y}_g \right\|_F^2 \right\} \\
&+ \lambda \sum_{g=1}^{G} \mathbb{E} \left\{ \left\| \mathbf{H}_{g, g} \mathbf{x}_g \right\|_F^2 \right\}, \\
\text{s.t.} \quad &\left\| \mathbf{V}_{k_g^d} \right\|_F^2 = \left\| \mathbf{U}_{k_g^d} \right\|_F^2 = b_d \\
&\left\| \mathbf{V}_{k_g^u} \right\|_F^2 = \left\| \mathbf{U}_{k_g^u} \right\|_F^2 = b_u
\end{aligned}
\tag{31}
$$

where λ is the weight coefficient for ASIC. The detailed expression of the sum of MSE is given in Appendix B. According to the MSE expressions, we can see that the sum of the MSE is convex and differentiable to a single precoder or combiner, so we can differentiate the objective function with respect to a single precoder or combiner and set the derivatives to zero, yielding the solutions as

$$\mathbf{V}_{k_g^d} = \mathbb{N} \left[\left(\mathbf{\Omega}_g + \lambda \mathbf{H}_{rsi, g} \right)^{-1} \hat{\mathbf{H}}_{k_g^d, g}^H \mathbf{U}_{k_g^d} \right], \tag{32}$$

$$\mathbf{V}_{k_g^u} = \mathbb{N} \left[\mathbf{\Omega}_{k_g^u}^{-1} \hat{\mathbf{H}}_{g, k_g^u}^H \mathbf{U}_{k_g^u} \right], \tag{33}$$

$$\mathbf{U}_{k_g^d} = \mathbb{N} \left[\left(\mathbf{C}_{k_g^d} + \sigma_r^2 \mathcal{D}(\mathbf{C}_{k_g^d}) \right)^{-1} \hat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d} \right], \tag{34}$$

$$\mathbf{U}_{k_g^u} = \mathbb{N} \left[\left(\mathbf{C}_g + \sigma_r^2 \mathcal{D}(\mathbf{C}_g) \right)^{-1} \hat{\mathbf{H}}_{g, k_g^u} \mathbf{V}_{k_g^u} \right], \tag{35}$$

where $\mathbb{N}[\cdot]$ denote the normalization operation to satisfy the constraints; $\mathbf{H}_{rsi, g} = \hat{\mathbf{H}}_{g, g}^H \hat{\mathbf{H}}_{g, g} + \mathcal{D}\left(\hat{\mathbf{H}}_{g, g}^H \hat{\mathbf{H}}_{g, g} \right)$; $\mathbf{\Omega}_g$ and $\mathbf{\Omega}_{k_g^u}$ can be denoted using the function defined in Equation (36) as $\mathbf{\Omega}_g = \sum_{j=1}^{G} \sum_{i=1}^{K_j^d} \mathcal{F}\left(\hat{\mathbf{H}}_{i_j^d, g}^H, \mathbf{U}_{i_j^d} \right) + \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \mathcal{F}\left(\hat{\mathbf{H}}_{j, g}^H, \mathbf{U}_{i_j^u} \right)$ and $\mathbf{\Omega}_{k_g^u} = \sum_{j=1}^{G} \sum_{i=1}^{K_j^d} \mathcal{F}\left(\hat{\mathbf{H}}_{i_j^d, k_g^u}^H, \mathbf{U}_{i_j^d} \right) + \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \mathcal{F}\left(\hat{\mathbf{H}}_{j, k_g^u}^H, \mathbf{U}_{i_j^u} \right)$. λ strikes a balance between enhancing the ASIC and MSE minimization. It should be noted that when λ exceeds a bound, the precoder will be close to $\mathbf{V}_{k_g^d} = \mathbb{N} \left[\mathbf{H}_{rsi, g}^{-1} \hat{\mathbf{H}}_{k_g^d, g}^H \mathbf{U}_{k_g^d} \right]$ due to the normalization operation, maximizing the ASIC capability of precoders. According to the simulation results, we set $\lambda = 1000$ when

---

Algorithm 1 Iterative algorithm to obtain the beamformers

Input: CSI (i.e., $\hat{\mathbf{H}}_{\dagger,*}$ and $\tilde{\sigma}_{\dagger,*}$, $\forall \dagger, *$ (node)), and threshold ς_{min}

Output: $\mathbf{V}_{k_g^d}, \mathbf{U}_{k_g^d}, \mathbf{V}_{k_g^u}, \mathbf{U}_{k_g^u}$

1: Initialize precoders $\mathbf{V}_{k_g^d}$ and $\mathbf{V}_{k_g^u}, \forall k, g$ randomly.
2: **for** $t = 1, \ldots, \text{max\_iter}$ **do**
3:     Update decoders $\mathbf{U}_{k_g^d}$ and $\mathbf{U}_{k_g^d}, \forall k, g$ from associated solutions, i.e., from (23) and (20) for eZF, from (27) and (29) for eMRTC, and from (34) and (35) for eMMSE.
4:     Calculate the associated loss function at this iteration $\mathcal{L}_\star^{(t)}$, where $\star$ could represent "eZF", "eMRTC", and "eMMSE", then calculate the decreased amount of the loss function as $\mathcal{L}_- = \mathcal{L}_\star^{(t-1)} - \mathcal{L}_\star^{(t)}$.
5:     **if** $\mathcal{L}_- < \varsigma_{min}$ **then**
6:         break;
7:     **else**
8:         Update precoders $\mathbf{V}_{k_g^d}$ and $\mathbf{V}_{k_g^u}, \forall k, g$ from associated solutions, i.e., from (19) and (22) for eZF, from (26) and (28) for eMRTC, and from (32) and (33) for eMMSE.
9:     **end if**
10: **end for**

---

the ASIC depth is insufficient, while we set $\lambda = 0$ with sufficient ASIC. The required sufficient ASIC depth can be calculated according to the transmission power, the dynamic range of ADCs, and noise floor as in [6].

$$\mathcal{F}(\mathbf{Y}, \mathbf{X}) = \mathbf{Y}\left[\mathbf{X}\mathbf{X}^H + \mathcal{D}\left(\mathbf{X}\mathbf{X}^H\right)\right]\mathbf{Y}^H + \sigma_t^2 \mathcal{D}\left(\mathbf{Y}\left[\mathbf{X}\mathbf{X}^H + \mathcal{D}\left(\mathbf{X}\mathbf{X}^H\right)\right]\mathbf{Y}^H\right) \quad (36)$$

The loss function of the eMMSE beamforming can be given by the objective function of the minimization problem as

$$\mathcal{L}_{\text{eMMSE}} = \xi + \lambda \sum_{g=1}^{G} \mathbb{E}\left\{\|\mathbf{H}_{g,g}\mathbf{x}_g\|_F^2\right\}, \quad (37)$$

where ξ is the sum of MSE given in Appendix A. The optimized variables should reduce the loss.

D. Iterative Algorithm

The expressions of solutions to the eZF, eMRTC, and eMMSE beamformers show interdependence between precoders and combiners. Thus, we need to iteratively update one of the precoders or combiners with others fixed until it converges. The procedure is summarised as Algorithm 1.

1) Convergence Behaviour: At each iteration, the optimized beamformer can reduce the loss function of associated beamforming schemes. Thus, the loss function is guaranteed to be reduced with iterations until it converges to a local minimum. Fig. 2 shows the loss function variation of the three schemes with iterations, demonstrating their convergence behaviors. The results are obtained based on a random instance of a 2-cell-4-user network with settings given in Section V, while more cells and users and larger antenna arrays may need more iterations to converge, and vice versa.
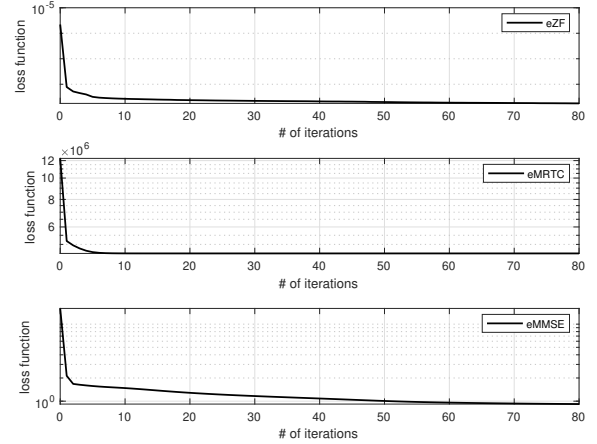

Fig. 2. The convergence behavior of the three beamforming schemes.

2) Computational Complexity: We measure the computational complexity by accounting for the required multiplication operations. We assume $K_j^d = K_j^u = K, \forall j$, $N_{bs} = M_{bs} = A_b$, $N_{ue} = M_{ue} = A_u$, and $b_d = b_u = N_s$ for simplicity. Usually, we will have $A_b \gg A_u \geq N_s$. The required multiplication operations of calculating the matrix multiplication of $\mathbf{A}_1\mathbf{A}_2$ is $a \times b \times c$, where $\mathbf{A}_1$ is of size $a \times b$ and $\mathbf{A}_2$ is of size $b \times c$. Taking $\mathbf{V}_{k_g^d}$ as an instance, calculating the stack matrix $\mathbf{\Gamma}_{k_g^d}$ and $\mathbf{\Gamma}_{k_g^d}^H \mathbf{\Gamma}_{k_g^d}$ for eZF needs a total of $(3KG-1)A_b^2 N_s + (KG-1)A_b A_u N_s$ multiplications; calculating $\mathbf{H}_{k_g^d,g}^H \mathbf{U}_{k_g^d} \mathbf{U}_{k_g^d}^H \mathbf{H}_{k_g^d,g}$ for eMRTC requires only $A_b^2 N_s + A_b A_u N_s$ multiplications, while the Eigen decomposition has the same complexity (i.e., $\mathcal{O}\left(A_b^3\right)$). Thus, eZF has higher computational complexity than eMRTC with even $G = 1$ and $K = 1$, and the difference increases significantly with increasing G and K. For eMMSE, calculating $\mathbf{\Omega}_g + \lambda \mathbf{H}_{rsi,g}$ needs a total of $KG(2A_b^3 + A_b^2 A_u + A_u^2 A_b + (A_b^2 + A_u^2)N_s + 2A_b^2) + A_b^3$ multiplications, and the inverse operation has the computational complexity of $\mathcal{O}\left(A_b^3\right)$, the rest matrix multiplication requires $A_b^2 A_u + A_b A_u N_s$ multiplications. Thus, the eMMSE beamforming has a much higher computational complexity than eZF and eMRTC in MCMU networks since it is in the order of GKA_b^3. The computational complexity of the three schemes can be given as

$$C_{\text{eZF}} = \mathcal{O}\left(A_b^3 + GKA_b^2 N_s\right), \quad (38)$$

$$C_{\text{eMRTC}} = \mathcal{O}\left(A_b^3\right), \quad (39)$$

$$C_{\text{eMMSE}} = \mathcal{O}\left(GKA_b^3\right). \quad (40)$$

It should be noted that although eMMSE, eZF, and eMRTC have the computational complexity of the same order with $G = K = 1$, eMMSE needs more multiplication, as analyzed above. The difference is more significant with large antenna arrays and a number of cells and users.

IV. MMSE BEAMFORMING WITH LOCAL CSI

In this section, we consider a multi-cell interfering broadcast network in which only local CSI (i.e., mixed imperfect instantaneous and statistical CSI) is available. We assume that

imperfect instantaneous CSI of intra-cell channels (i.e., channels between the BS and UEs within the same cell, channels between the UEs and UEs within the same cell, and self-interference channel) is available, while the CSI of inter-cell channels (i.e., channels between the BSs and UEs, channels between UEs and UEs in different cells) is known statistically that only the channel covariance matrices are known to save the cost of time-frequency resources for training pilots. The imperfect instantaneous CSI of intra-cell channels of the gth cell is available at all base stations due to the capacity of the backhaul link. We only extend the eMMSE beamforming to cope with the local CSI scenario since it achieves the best performance with global CSI, and eZF cannot utilize second-order statistics. It should be noted that eMRTC can be applied directly to local CSI scenarios since its solutions do not rely on inter-cell CSI.

The most straightforward approach is to omit these terms with inter-cell channels. However, it will degenerate the solutions into single-cell solutions and degrade the performance in multi-cell networks. We propose two methods in this paper. In the first one, we assume random dummy instances for unknown channel matrices, while we average out dummy matrices in covariance matrices (i.e., the terms of unknown inter-cell interference are regarded as noise with known strength) in the second approach. Then, we derive two variants of eMMSE beamforming with local CSI as follows.

A. With Unknown CSI as Random Instances: eMMSE-RI

This idea takes advantage of the fact that the channel matrix associated with the LOS component is mainly determined by the pathloss only if the distance between the transmitting and receiving antennas is much further than the size of the uniform linear array. In this case, we can view the antenna arrays at the transmitting and receiving nodes as two points, such that the AoD and AoA could be viewed as 0 degrees, yielding each element of the LOS component to be 1 such that $\left[\mathbf{H}_{r,t}^{LOS}\right]_{m,n} = 1 \; \forall m,n$. Thus, we can ignore the minor phase difference between the elements of the LOS component without loss of generality. On the other hand, the LOS component usually dominates the channel in practice if it presents. Fig. 3 shows the pathloss of the LOS and NLOS components versus the distance between the communication nodes, where the pathloss model follows the 3GPP specifications in [34]. It shows that the LOS path has > 10dB less pathloss than the NLOS path and the LOS probability is high if the distance is short. Thus, it can potentially regenerate these unknown CSI terms with trivial errors for microcells using the pathloss of associated channels only. With the pathloss of the unknown channels, one can generate a random instance of the unknown channel matrix according to the wireless channel models given in Equations (5) and (6). Let $\tilde{H}$ denote a random realization that consists of i.i.d. complex Gaussian elements such that $[\tilde{\mathbf{H}}]_{i,j} \sim \mathcal{CN}(0,1) \;\; \forall \; i,j$, representing a dummy variable. With this variable, the unknown channel matrices can be given as follows.

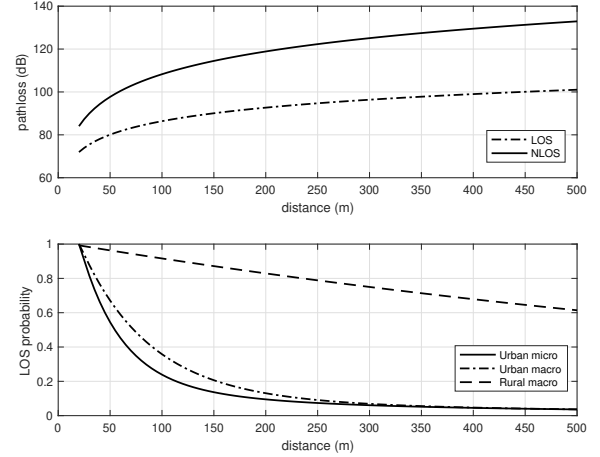- Rician channels: we can generate a random instance of

Fig. 3. The pathloss of the LOS and NLOS components and the LOS probability versus the distance between communication nodes calculated from the formula in 3GPP document [34].

the channel matrix as

$$\hat{\mathbf{H}}_{\dagger,*} = \sqrt{\varrho_{\dagger,*}} \left(\sqrt{\frac{\kappa}{1+\kappa}} \tilde{\mathbf{H}}^{LOS} + \sqrt{\frac{1}{1+\kappa}} \tilde{\mathbf{H}} \right) \quad (41)$$

where $\tilde{\mathbf{H}}^{LOS}$ is simply given by a matrix with each element equal to 1 (i.e., $\left[\tilde{\mathbf{H}}^{LOS}\right]_{m,n} = 1, \forall \; m,n$), and the channel uncertainty can be given as

$$\boldsymbol{\Delta}_{\dagger,*} = \mathbf{H}_{\dagger,*} - \hat{\mathbf{H}}_{\dagger,*} = \sqrt{\frac{1}{1+\kappa}\varrho_{\dagger,*}}\mathbf{H} - \sqrt{\frac{1}{1+\kappa}\varrho_{\dagger,*}}\tilde{\mathbf{H}} \quad (42)$$

which leads to the statistics as $\mathbb{E}\{\boldsymbol{\Delta}_{\dagger,*}\} = \mathbf{0}$ and $Var(\boldsymbol{\Delta}_{\dagger,*}) = \frac{2}{1+\kappa}\varrho_{\dagger,*}\mathbf{I}$.

- Rayleigh channels: the random instances of Rayleigh channel matrices can be given as

$$\hat{\mathbf{H}}_{\dagger,*} = \sqrt{\varrho_{\dagger,*}}\tilde{\mathbf{H}} \quad (43)$$

which gives the channel uncertainty as

$$\boldsymbol{\Delta}_{\dagger,*} = \mathbf{H}_{\dagger,*} - \hat{\mathbf{H}}_{\dagger,*} = \sqrt{\varrho_{\dagger,*}}\left(\mathbf{H} - \tilde{\mathbf{H}}\right) \quad (44)$$

which leads to the statistics as $\mathbb{E}\{\boldsymbol{\Delta}_{\dagger,*}\} = \mathbf{0}$ and $Var(\boldsymbol{\Delta}_{\dagger,*}) = 2\varrho_{\dagger,*}\mathbf{I}$.

This case can be regarded as a special case of the global CSI scenario, where the inter-cell channel matrices are replaced by these random instances. Thus, the procedure of the eMMSE-RI beamforming with local CSI regarded as random insurances is similar to Algorithm 1, where we have an additional step to generate the random instances and replace the inter-cell channel matrices by the random instances. Also, the channel uncertainty of inter-cell channels is different from the global CSI scenario, which is given as

$$\begin{cases} \tilde{\sigma}_{\dagger,*}^2 = 2\varrho_{\dagger,*}, \text{if } \kappa = 0 \\ \tilde{\sigma}_{\dagger,*}^2 = \frac{2}{1+\kappa}\varrho_{\dagger,*}, \text{if } \kappa \neq 0 \end{cases}, \forall \; \dagger, * \quad (45)$$

Note that the estimated inter-cell channels are regarded as random instances with channel uncertainty given in Equation (45) here. Therefore, the rate expression can be computed as

in Section II-D with inter-cell channel matrices replaced by the random instances and corresponding channel uncertainty.

B. With Unknown CSI as Noise: eMMSE-N

Indeed, the LOS component can be regenerated with trivial errors, as we explained above, but the LOS probability decreases significantly with increasing distance. Thus, we can speculate that in the case of large cells, there may be few LOS components present in inter-cell channels. In this case, regarding unknown CSI as random instances may further decrease the capacity since the random NLOS components could double the channel errors as (45) suggests. We proposed an alternative approach for processing the unknown CSI terms in this case, where they are regarded as noise with known strength. The beamformers minimize the MSE with interferences with known strength but unknown directions. This is a more general approach as it does not depend on the channel conditions. Considering unknown CSI terms as noise leads to different received signal covariance matrices, which are given as

$$\bar{\mathbf{C}}_{k_g^d} = \hat{\mathbf{H}}_{k_g^d,g} \mathbf{T}_g \hat{\mathbf{H}}_{k_g^d,g}^H + \left(\bar{\sigma}_{k_g^d}^2 + \sigma_{ue}^2 \right) \mathbf{I} \tag{46}$$

$$\bar{\mathbf{C}}_g = \hat{\mathbf{H}}_{g,g} \mathbf{T}_g \hat{\mathbf{H}}_{g,g}^H + \sum_{i=1}^{K_j^u} \hat{\mathbf{H}}_{g,i_g^u} \mathbf{T}_{i_g^u} \hat{\mathbf{H}}_{g,i_g^u}^H + \left(\bar{\sigma}_g^2 + \sigma_{ue}^2 \right) \mathbf{I} \tag{47}$$

Since we have the statistic knowledge of the unknown CSI terms, we have $\mathbb{E}\left\{ \mathbf{H}_{\dagger,*} \mathbf{T}_* \mathbf{H}_{\dagger,*}^H \right\} = \mathbb{E}_{\mathbf{H}}\left\{ \mathbf{H}_{\dagger,*} \mathbf{H}_{\dagger,*}^H \right\} \cdot tr\left(\mathbf{T}_* \right) = \varrho_{\dagger,*} \cdot tr\left(\mathbf{T}_* \right)$. Therefore, $\bar{\sigma}_{k_g^d}^2$ and $\bar{\sigma}_g^2$ are given as

$$\bar{\sigma}_{k_g^d}^2 = \tilde{\sigma}_{k_g^d,g}^2 tr\left(\mathbf{T}_g \right) + \sum_{j \neq g}^{G} \varrho_{k_g^d,j} tr\left(\mathbf{T}_j \right)$$
$$+ \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \varrho_{k_g^d,i_j^u} tr\left(\mathbf{T}_{i_j^u} \right) \tag{48}$$
$$= \left(1 + \sigma_t^2 \right) \left(\tilde{\sigma}_{k_g^d,g}^2 P_{bs} + \sum_{j \neq g}^{G} \varrho_{k_g^d,j} P_{bs} + \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \varrho_{k_g^d,i_j^u} P_{ue} \right)$$

$$\bar{\sigma}_g^2 = \tilde{\sigma}_{g,g}^2 tr\left(\mathbf{T}_g \right) + \sum_{i=1}^{K_j^u} \tilde{\sigma}_{g,i_g^u} tr\left(\mathbf{T}_{i_g^u} \right)$$
$$+ \sum_{j \neq g}^{G} \varrho_{g,j} tr\left(\mathbf{T}_j \right) + \sum_{j \neq g}^{G} \sum_{i=1}^{K_j^u} \varrho_{g,i_j^u} tr\left(\mathbf{T}_{i_j^u} \right)$$
$$= \left(1 + \sigma_t^2 \right) \left(\tilde{\sigma}_{g,g}^2 P_{bs} + \sum_{i=1}^{K_j^u} \tilde{\sigma}_{g,i_g^u} P_{ue} \right. \tag{49}$$
$$\left. + \sum_{j \neq g}^{G} \varrho_{g,j} P_{bs} + \sum_{j \neq g}^{G} \sum_{i=1}^{K_j^u} \varrho_{g,i_j^u} P_{ue} \right)$$

Replacing the covariance matrices $\mathbf{C}_g$ and $\mathbf{C}_{k_g^d}$ in the MSE expression (55) in Appendix B with $\bar{\mathbf{C}}_g$ and $\bar{\mathbf{C}}_{k_g^d}$, we get the sum of MSE with local CSI. Take derivation of the sum of MSE with respect to precoders and combiners and set the derivatives to zero, the solutions are given similarly as (32) -

(35) with the $\mathbf{\Omega}_g$, $\mathbf{\Omega}_{k_g^u}$, $\mathbf{C}_{k_g^d}$, and $\mathbf{C}_g$ replaced by $\bar{\mathbf{\Omega}}_g$, $\bar{\mathbf{\Omega}}_{k_g^u}$, $\bar{\mathbf{C}}_{k_g^d}$, and $\bar{\mathbf{C}}_g$, respectively, where $\bar{\mathbf{\Omega}}_g$ and $\bar{\mathbf{\Omega}}_{k_g^u}$ can be denoted using the functions defined in (36) as

$$\bar{\mathbf{\Omega}}_g = \sum_{i=1}^{K_g^d} \mathcal{F}\left(\hat{\mathbf{H}}_{i_g^d,g}^H, \mathbf{U}_{i_g^d} \right) + \sum_{i=1}^{K_g^u} \mathcal{F}\left(\hat{\mathbf{H}}_{g,g}^H, \mathbf{U}_{i_g^u} \right) \tag{50}$$

$$\bar{\mathbf{\Omega}}_{k_g^u} = \sum_{i=1}^{K_g^d} \mathcal{F}\left(\hat{\mathbf{H}}_{i_g^d,k_g^u}^H, \mathbf{U}_{i_g^d} \right) + \sum_{i=1}^{K_g^u} \mathcal{F}\left(\hat{\mathbf{H}}_{g,k_g^u}^H, \mathbf{U}_{i_g^u} \right) \tag{51}$$

Since the estimated inter-cell channels are considered statistically with noise here, the achievable rate will be computed differently from the expressions in Section II-D. In this case, we will have

$$\bar{R}_{k_g^d} = \log_2 \left| \mathbf{I}_{b_d} + \mathbf{U}_{k_g^d}^H \mathbf{C}_{k_g^d,S} \mathbf{U}_{k_g^d} \left(\mathbf{U}_{k_g^d}^H \bar{\mathbf{C}}_{k_g^d,X} \mathbf{U}_{k_g^d} \right)^{-1} \right|,$$
$$\bar{R}_{k_g^u} = \log_2 \left| \mathbf{I}_{b_u} + \mathbf{U}_{k_g^u}^H \mathbf{C}_{k_g^u,S} \mathbf{U}_{k_g^u} \left(\mathbf{U}_{k_g^u}^H \bar{\mathbf{C}}_{k_g^u,X} \mathbf{U}_{k_g^u} \right)^{-1} \right|, \tag{52}$$

where $\bar{\mathbf{C}}_{k_g^d,X}$ and $\bar{\mathbf{C}}_{g,X}$ are given as

$$\bar{\mathbf{C}}_{k_g^d,X} = \bar{\mathbf{C}}_{k_g^d} - \mathbf{C}_{k_g^d,S} + \sigma_r^2 \mathcal{D}\left(\bar{\mathbf{C}}_{k_g^d} \right), \tag{53}$$
$$\bar{\mathbf{C}}_{g,X} = \bar{\mathbf{C}}_g - \mathbf{C}_{g,S} + \sigma_r^2 \mathcal{D}\left(\bar{\mathbf{C}}_g \right). \tag{54}$$

The iterative procedure for this case is the same as the Algorithm 1.

V. NUMERICAL RESULTS

Our simulations follow the 3GPP specifications, where the LOS probability and pathloss are calculated according to models given in Table 7.4.2-1, and Table 7.4.1-1 in [34]. We consider 10MHz bandwidth centered at 2.5GHz. We consider each cell has the same number of downlink and uplink users in simulations, i.e., $K_g^d = K_g^u = K \ \forall g$. The base station has 16 transmit and receive antennas (i.e., $N_{bs} = M_{bs} = 16$), while the users have 2 transmit and receive antennas (i.e., $N_{ue} = M_{ue} = 2$). The thermal noise density is -174dBm/Hz, and the noise figure is 13dB and 9dB for the base station and users, respectively. The transmission power is set to be 23dBm for each user, i.e., 23dBm at UEs and $(23 + 10\log_{10} K)$dBm at BSs. In order to show the necessity of ASIC, we limit σ_r^2 not less than the quantization noise strength caused by 12-bit ADCs at receivers in simulations. The Rician factor of the self-interference channel is set as 20dB, while the Rician factor of other channels is calculated according to the pathloss of the LOS and NLOS components calculated from [34]. The network is constructed based on 3GPP specifications, e.g., urban-micro (UMi), urban-macro (UMa), rural-macro(RMa), and indoor hotspot (InH) scenarios defined in [34]. The parameters of the four scenarios are given in Table II. We consider a common UMi scenario with 200m of inter-site distance (i.e., the distance between two neighboring base stations) if not specified.

TABLE II
3GPP SCENARIOS

Parameters		UMi - street canyon	UMa	RMa	InH
Cell layout		Hexagonal grid, ISD = 200m	Hexagonal grid, ISD = 500m	Hexagonal grid, ISD = 1732m	120m×50m×3m, ISD = 20m
BS antenna height h_{BS}		10m	25m	35m	3m (ceiling)
UE location	Outdoor/indoor	20% outdoor 80% indoor	20% outdoor 80% indoor	50% indoor 50% in car	Indoor
	LOS/NLOS	LOS and NLOS	LOS and NLOS	LOS and NLOS	LOS and NLOS
	Height h_{UE}	1.5-22.5m	1.5-22.5m	1.5m	1m
Minimum BS-to-UE distance		10m	35m	35m	0m

A. Self-interference Cancellation

We start with a simple scenario with 2 cells, and each BS serves 2 users as Fig. 4, where the downlink users and uplink users are $\frac{4}{10}r$ and $\frac{6}{10}r$ away from the BS (r is the radius of the cell), respectively. The CCI is stronger in scenario 2 than in scenario 1 as the uplink and downlink users are closer. In order to show the effect of self-interference cancellation (SIC), we use scenario 1 in simulations to minimize the effect of CCI.

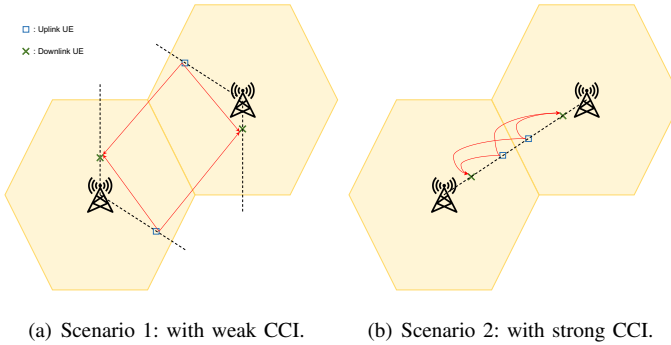(a) Scenario 1: with weak CCI. (b) Scenario 2: with strong CCI.

Fig. 4. A 2-cell IBFD network with 2 uplink UEs and 2 downlink UEs with CCI highlighted.

Fig. 5(a) shows the achievable rate of IBFD radios through different beamforming schemes versus the SIC depth compared to HD. The achievable rate of HD radios is achieved using the eMMSE beamforming with global CSI and eMMSE-N with local CSI as they achieve the highest rate. We assume an infinite dynamic range of receivers here, where the SIC in the analog and digital domain contribute the same, and we set $\lambda = 0$ for eMMSE since ASIC is not essential here. The results demonstrate the ability of eMMSE (or MMSE) and eZF beamformers to suppress the self-interference that full IBFD gain is obtained with only ≥ 30dB of SIC since the residual self-interference can be canceled by the beamformers. Deeper SIC (i.e., > 30dB) does not contribute to the ergodic sum rate with global CSI, but > 80dB SIC can still improve the system throughput with local CSI since the eMMSE beamformers can more effectively suppress the unknown inter-cell interference when the SI is not so significant. In contrast, the eMRTC beamformers cannot suppress the self-interference so that it realizes the full IBFD gain with > 140dB of SIC.

Then we consider practical receivers, but we ignore the nonlinearity of low noise amplifiers here. Fig. 5(b) shows the achievable sum rate versus ASIC depth, where a total of 140dB of SIC is assumed to be realized. It illustrates the importance of ASIC that the desired IBFD gain cannot be realized without sufficient ASIC even if a total of 140dB of SIC is applied due to the significant quantization noise. The eMMSE beamforming can improve the uplink rate when ASIC is insufficient compared to the conventional MMSE, but it will decrease the achievable downlink rate. The effective ASIC depth is related to the transmission power, noise level, and effective bits of ADCs, which can be calculated as in [6]. The advantage of our proposed eMMSE beamforming over the MMSE is that it can reduce the requirement for the ASIC depth to obtain the desired IBFD gain (i.e., from 80dB to 60dB). 20dB lower requirement on the ASIC depth can significantly reduce the complexity and cost of RF cancellers [6]. Besides, digital SIC is not needed with global CSI since the ASIC depth has already exceeded the required total SIC depth of 30dB as in Fig. 5(a), inspiring a low-complexity IBFD transceiver design, while it is still needed to make up for the 40db gap to achieve the maximum IBFD gain with local CSI as Fig. 5(a) shows the maximum IBFD gain is achieved with ≥ 100dB of SIC with local CSI.

B. Transceiver HWIs and Channel Uncertainty

In this section, we study the effects of transceiver HWIs and channel uncertainty on the IBFD gain. We assume effective SIC (i.e., ≥ 80dB of ASIC and ≥ 140dB) is achieved at each IBFD base station. Fig. 6 shows the achievable sum rate variation with increasing transceiver HWIs and channel uncertainty. The transceiver HWIs and channel uncertainty decrease the achievable sum rate in general. However, they do not decrease the IBFD gain with global CSI that 98.6%, 99.5%, and 95.2% higher sum rate of the HD radios are achieved by applying eMMSE beamforming with 0dB of transmitter and receiver HWIs factor and 0dB of channel uncertainty factor with global CSI, respectively. The eMRTC beamforming shows stronger robustness to the channel uncertainty than eZF and eMMSE that 0dB of channel uncertainty factor introduces 47.9% of sum rate degradation while this value increases to 55.1% and 71.1% for eMMSE and eZF, respectively. This is understandable as eMRTC use much less CSI, and eMMSE includes the effects of channel uncertainty. Furthermore, the performance difference between eMMSE with global CSI and eMMSE-N with local CSI decreases with increasing practical imperfections since their impact will outweigh the impact of the inter-cell interference. For local CSI scenarios, large channel uncertainty will reduce the IBFD gain that only 47.9% higher sum rate of HD is achieved with 0dB of channel uncertainty factor, while the improvement amount can be
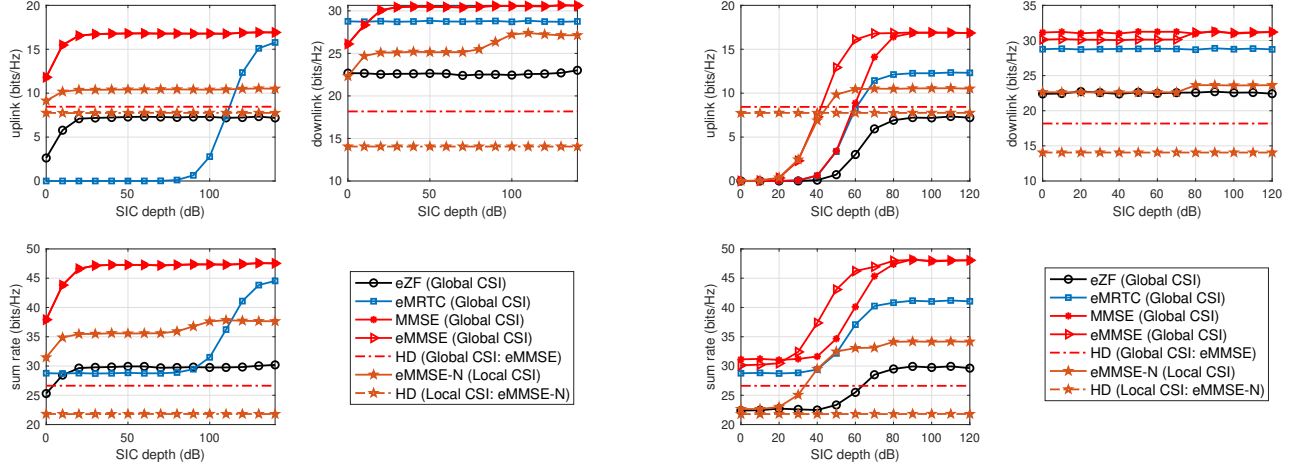
(a) Achievable sum rate versus SIC depth (infinite dynamic range of receivers). (b) Achievable sum rate versus ASIC depth (140 dB of total SIC is assumed).

Fig. 5. Achievable sum rate with different self-interference conditions.
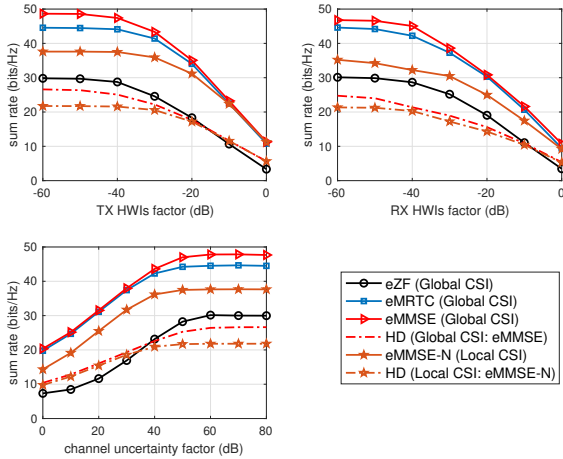

Fig. 6. Sum rate variation with increasing transceiver HWIs and channel uncertainty.

72.7% with accurate CSI ($\geq$ 60dB of channel uncertainty factor).

C. Co-channel Interference

In this section, we investigate the effects of CCI on IBFD gain. We assume 80dB of ASIC, 120dB of SIC, -120dB of transceiver HWIs factor, and 120dB of channel uncertainty factor to minimize their effects. Fig. 7 shows the achievable sum rate of different beamforming schemes versus the inter-site distance (ISD) compared to the HD radios, where Fig.7(a) and Fig.7(b) show the results for 2-cell and 7-cell networks. We consider Scenario 1 and Scenario 2 in Fig. 4 for 2-cell networks, which have weak and strong CCI, respectively. For 7-cell networks, we also consider two scenarios: 1) weaker CCI: one downlink user and one uplink user per cell; 1) stronger CCI: two downlink users and two uplink users per cell. It should be noted that both the received signal of interest

and interference decrease with increasing ISD due to the long propagation distance. With global CSI, the sum rate of eMMSE decreases with enlarging cells since the interference could be effectively suppressed even if it is significant (i.e., with short ISD and more cells) so that the decreased signal of interest decreases the sum rate. However, the IBFD gain increases with enlarging cells such that 80% higher sum rate of HD can be achieved with ISD= 400m, while the IBFD operation may even decrease the spectral efficiency with the 7-cell-4-user network due to the strong CCI. The eZF beamforming has a similar behavior as eMMSE with 2-cell networks. However, with 7 cells, it improves the achievable sum rate with enlarging cells at first (i.e., ISD$\leq$ 200m). It suggests that eZF cannot effectively suppress interference if it is too complex. The achievable sum rate of the eMRTC beamforming increases with increasing ISD at first since it improves the received power of the signal of interest to compensate for its increasing pathloss, but the sum rate decreases with larger cells. In scenario 1, the eMRTC beamforming approaches the performance of eMMSE when ISD$>$ 350m, while it cannot achieve this in Scenario 2 due to the stronger CCI in this scenario. The eMRTC beamforming could approach the performance of eMMSE even with 7 cells as long as there is only a single uplink and downlink user within a cell. Furthermore, it shows that losing the instantaneous CSI knowledge of inter-cell channels significantly degrades the system capacity for microcells (e.g., ISD$<$ 100m). The impact of losing this knowledge decreases with enlarging cells that only 14.7% of performance degradation is yielded with proper processing when ISD= 400m in Scenario 2, respectively. The performance degradation due to local CSI is higher in Scenario 2 and 7-cell-4-user networks since the inter-cell interference is stronger than in Scenario 1 and 7-cell-2-user networks. As introduced before, eMMSE-RI takes advantage of the properties of Rician channels, so it outperforms eMMSE-N with microcells, where the probability of having LOS paths between users and BSs is high. The LOS probability decreases with

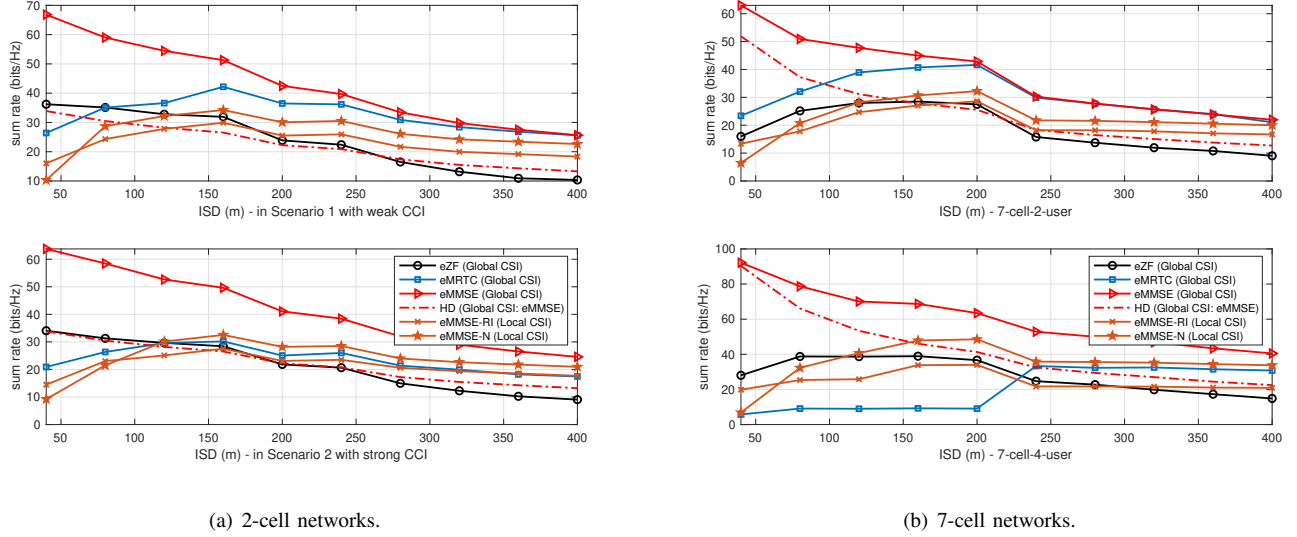(a) 2-cell networks.

(b) 7-cell networks.

Fig. 7. Achievable sum rate variation with increasing inter-site distance (ISD).
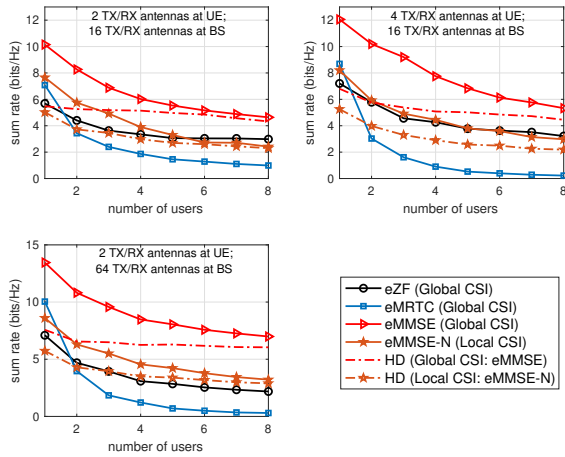

Fig. 8. Averaged achievable sum rate per user versus the number of users per cell.

enlarging cells, as suggested by Fig. 3. Therefore, eMMSE-N outperforms eMMSE-RI with large cells, and the performance gap increases with the increasing number of cells and users that eMMSE-N achieves 61.3% higher sum rate than eMMSE-RI with a 7-cell-4-user network when ID= 400m while it is only around 18.4% higher in Scenario 2.

Fig. 8 shows the average achievable sum rate per user versus the number of downlink and uplink users per cell with different antenna array configurations. We use 2 cells here, and users are randomly located. It can be seen that the IBFD gain decreases significantly with an increasing number of users due to the increased CCI. However, this can be compensated by more antennas at BSs or UEs that the sum rate improvement over HD increases from 7.2% with 2 antennas at UEs and 16 antennas at BSs to 19.8% with 4 antennas at UEs and 16 antennas at BSs and 17.6% with 2 antennas at UEs and

64 antennas at BSs. It also shows that eMRTC outperforms eZF when there are only one downlink user and one uplink user, while eZF outperforms eMRTC with more users. This can be understood as the eZF minimizes interference, so it performs better with stronger interference. Increasing the number of antennas at users and base stations can improve the performance of eMMSE and eZF as it provides more degrees of freedom to manage interference. The eMRTC beamforming benefits from more antennas at BSs, but more antennas at users may degrade its performance. The eMMSE-N beamforming with local CSI has similar behavior to the eMMSE beamforming with global CSI, although local CSI decreases the achievable rate.

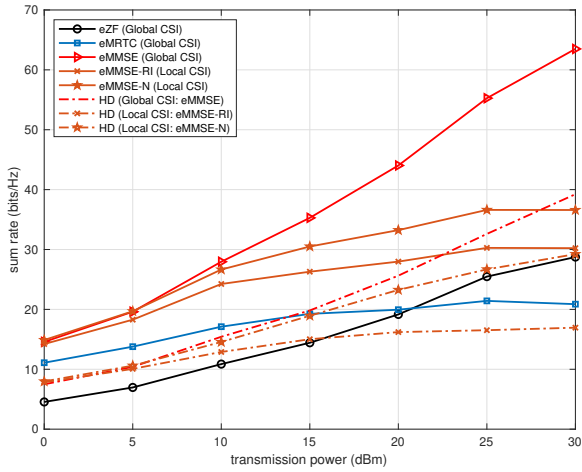### D. Performance Evaluation Based on 3GPP Specifications

In this section, we will evaluate the performance of these beamforming schemes under 3GPP-specified indoor hotspot, urban-micro, urban-macro, and rural-macro scenarios detailed in Table II. We consider a 2-cell network in simulations, where each base station serves 2 uplink users and 2 downlink users. The performance with more cells and users can be inferred from the above conclusions. Fig. 9 shows the achievable sum rate of various beamforming schemes versus the transmission power under the four scenarios. It shows that enhancing the transmission power can increase the system capacity. However, IBFD gain decreases with increasing transmission power. For instance, the sum rate improvement of IBFD over HD decreases from 66.3% with 0dBm of transmission power to 47.9% with 30dBm of transmission power under UMi scenarios. As for the eMRTC beamforming scheme, it can achieve considerable performance and outperforms eZF with low transmission power except in indoor hotspot scenarios, but enhancing the transmission power does not have significant improvements. In indoor hotspot scenarios, enhancing the transmission power does not significantly increase the achievable sum rate of beamforming schemes except eMMSE with
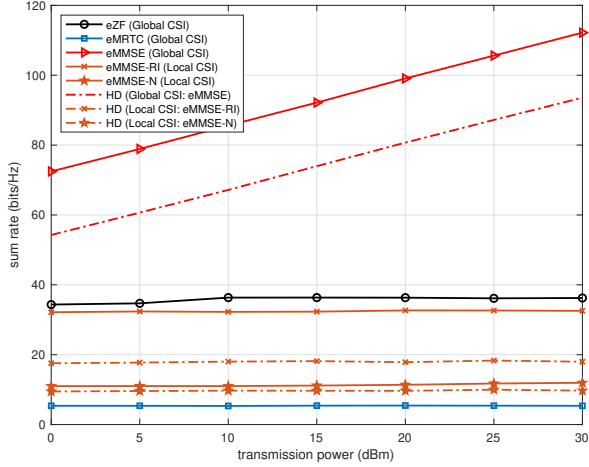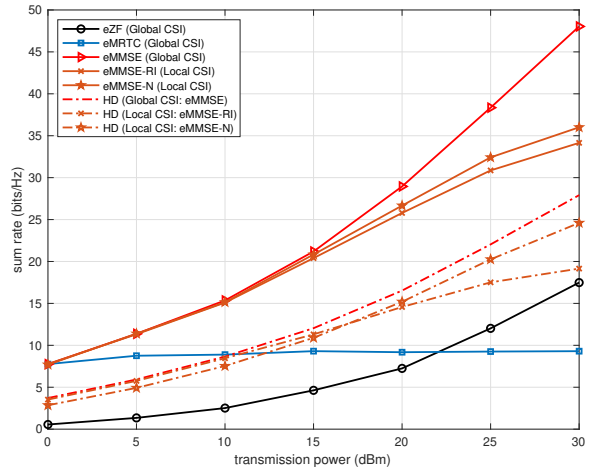
(a) Indoor hotspot scenario.
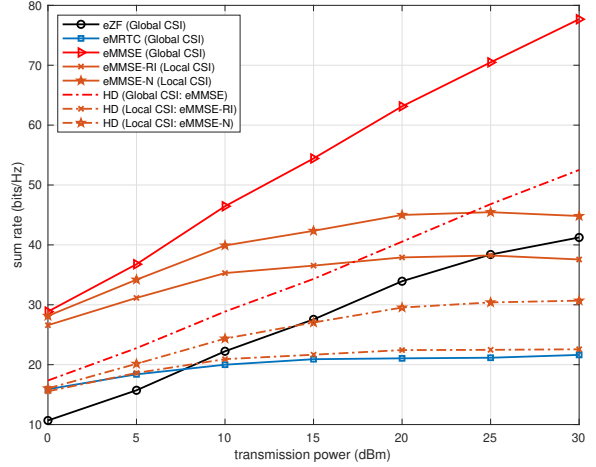
(b) Urban-micro scenario.

(c) Urban-macro scenario.

(d) Rural-macro scenario.

Fig. 9. Performance evaluation under four 3GPP-specified scenarios.

global CSI. However, with proper processing (i.e., eMMSE-RI), the IBFD operation can improve the spectral efficiency by 82.5%, which is higher than the global CSI scenario, although the local CSI seriously reduces the system capacity compared to global CSI. In the rest three scenarios, eMMSE-N performs better than eMMSE-RI. In rural-macro scenarios, having only local CSI does not introduce significant performance degradation with transmission power ≤ 15dBm since the inter-cell interference is not significant due to the large cell size.

VI. CONCLUSION

In this paper, we have studied beamforming schemes for IBFD-MCMU networks with practical imperfections. We revealed that the ASIC is critical to realizing the desired IBFD gain, while the residual self-interference can be effectively suppressed by the eZF and eMMSE beamformers without digital SIC for global CSI scenarios. With local CSI, the

digital SIC is still needed to maximize the IBFD gain. Our proposed eMMSE beamforming can reduce the requirement on the minimum effective ASIC depth. The transceiver HWIs decrease the achievable sum rate but do not decrease the IBFD gain, while large channel uncertainty reduces the IBFD gain with local CSI. The eMRTC beamforming has considerable performance with trivial interference (e.g., large cells and single-user network) that it could approach the performance of eMMSE. In contrast, the eZF beamforming outperforms eMRTC with increasing interference (e.g., microcells and more users). The eMMSE beamforming benefits from large-scale antenna arrays at both BSs and UEs, while eMRTC benefits only from more antennas at BSs. With local CSI, the unknown CSI can be regarded as random instances for microcells, while it should be regarded as noise with enlarging cells for better performance. In 3GPP-specified MCMU networks, the eMMSE beamforming could achieve a much higher sum

$$\xi = \sum_{g=1}^{G}\sum_{k=1}^{K_g^d}\left[1 - 2\frac{\sqrt{P_d}}{b_d}\Re\left\{tr\left(\mathbf{U}_{k_g^d}^H\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d}\right)\right\} + tr\left(\mathbf{U}_{k_g^d}^H\mathbf{C}_{k_g^d}\mathbf{U}_{k_g^d}\right)\right) + \sigma_r^2 tr\left(\mathbf{U}_{k_g^d}^H\mathcal{D}\left(\mathbf{C}_{k_g^d}\right)\mathbf{U}_{k_g^d}\right]$$

$$+ \sum_{g=1}^{G}\sum_{k=1}^{K_g^u}\left[1 - 2\frac{\sqrt{P_u}}{b_u}\Re\left\{tr\left(\mathbf{U}_{k_g^u}^H\mathbf{H}_{g,k_g^u}\mathbf{V}_{k_g^u}\right)\right\} + tr\left(\mathbf{U}_{k_g^u}^H\mathbf{C}_g\mathbf{U}_{k_g^u}\right) + \sigma_r^2 tr\left(\mathbf{U}_{k_g^u}^H\mathcal{D}\left(\mathbf{C}_g\right)\mathbf{U}_{k_g^u}\right)\right] \quad (55)$$

---

rate than other schemes, and it can be further improved with higher transmission power. In indoor hotspot scenarios, having local CSI will significantly reduce the system throughput due to the substantial inter-cell interference in this scenario, but considerable spectral efficiency (i.e., 32 bits/Hz) is achievable with proper processing (i.e., eMMSE-RI). In other scenarios, the effect of local CSI is mitigated by the enlarged cell size but will be exacerbated by enhancing transmission power.

Future work could consider converting such a fully-digital beamforming scheme to hybrid beamforming for massive MIMO systems and do power allocation to improve spectral efficiency.

APPENDIX A

Here, we derive and write the expressions for the covariance matrices of received signals, excluding receiver HWIs. Receiver HWI will be written separately in terms of the following derived matrices. The covariance matrix of the received signals at the k_g^d-th downlink user can be computed as

$$\mathbf{C}_{k_g^d} = \mathbb{E}\left\{(\mathbf{y}_{k_g^d} - \mathbf{e}_{k_g^d})(\mathbf{y}_{k_g^d} - \mathbf{e}_{k_g^d})^H\right\}$$

$$= \mathbb{E}\left\{\sum_{j=1}^{G}\mathbf{H}_{k_g^d,j}\mathbf{T}_j\mathbf{H}_{k_g^d,j}^H + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u}\mathbf{H}_{k_g^d,i_j^u}\mathbf{T}_{i_j^u}\mathbf{H}_{k_g^d,i_j^u}^H\right\} + \sigma_{ue}^2\mathbf{I}$$

$$\overset{(a)}{=} \sum_{j=1}^{G}\left(\hat{\mathbf{H}}_{k_g^d,j}\mathbf{T}_j\hat{\mathbf{H}}_{k_g^d,j}^H + \tilde{\sigma}_{k_g^d,j}^2 tr\left(\mathbf{T}_j\right)\mathbf{I}\right)$$

$$+ \sum_{j=1}^{G}\sum_{i=1}^{K_j^u}\left(\hat{\mathbf{H}}_{k_g^d,i_j^u}\mathbf{T}_{i_j^u}\hat{\mathbf{H}}_{k_g^d,i_j^u}^H + +\tilde{\sigma}_{k_g^d,i_j^u}^2 tr\left(\mathbf{T}_{i_j^u}\right)\mathbf{I}\right) + \sigma_{ue}^2\mathbf{I}$$

$$\overset{(b)}{=} \sum_{j=1}^{G}\hat{\mathbf{H}}_{k_g^d,j}\mathbf{T}_j\hat{\mathbf{H}}_{k_g^d,j}^H + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u}\hat{\mathbf{H}}_{k_g^d,i_j^u}\mathbf{T}_{i_j^u}\hat{\mathbf{H}}_{k_g^d,i_j^u}^H$$

$$+ \left(\bar{\sigma}_{k_g^d}^2 + \sigma_{ue}^2\right)\mathbf{I}$$

where in (a), the channel uncertainty model $\mathbf{H} = \hat{\mathbf{H}} + \boldsymbol{\Delta}$ is inserted, such that $\mathbb{E}_\Delta\left\{\mathbf{HTH}^H\right\} = \mathbb{E}_\Delta\left\{\left(\hat{\mathbf{H}} + \boldsymbol{\Delta}\right)\mathbf{T}\left(\hat{\mathbf{H}} + \boldsymbol{\Delta}\right)^H\right\} = \hat{\mathbf{H}}\mathbf{T}\hat{\mathbf{H}}^H + \mathbb{E}_\Delta\left\{\boldsymbol{\Delta}\mathbf{T}\boldsymbol{\Delta}^H\right\}$, and following simplification is used

$$\mathbb{E}_\Delta\left\{\left[\boldsymbol{\Delta}\mathbf{T}\boldsymbol{\Delta}^H\right]_{m,n}\right\}$$

$$= \mathbb{E}_\Delta\left\{\sum_{k=1}^{N_\Delta}\sum_{j=1}^{N_\Delta}\Delta_{mk}T_{kj}\Delta_{nj}^*\right\}$$

$$= \sum_{k=1}^{N_\Delta}\sum_{j=1}^{N_\Delta}\mathbb{E}_\Delta\left\{\Delta_{mk}T_{kj}\Delta_{mk}^*\right\}\delta_{mn}\delta_{kj}$$

$$\overset{(c)}{=} \sum_{k=1}^{N_\Delta}\mathbb{E}_\Delta\left|\Delta_{mk}\right|^2 T_{kk}\delta_{mn} = \sigma_\Delta^2 tr\left(\mathbf{T}\right)\delta_{mn}$$

where in (c), we assume $\delta_{kj} = 0, \forall\, k \neq j$ and $\delta_{kk} = 1, \forall\, k$. This also suggests that $\mathbb{E}\left\{\boldsymbol{\Delta}\mathbf{T}\boldsymbol{\Delta}^H\right\}$ only has values on its diagonal since $\delta_{mn} = 0, \forall\, m \neq n$, such that $\mathbb{E}\left\{\boldsymbol{\Delta}\mathbf{T}\boldsymbol{\Delta}^H\right\} = \sigma_\Delta^2 tr\left(\mathbf{T}\right)\mathbf{I}$. In (b), we simplify the equation through the substitution as

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

$$= \sum_{j=1}^{G}\tilde{\sigma}_{k_g^d,j}^2 P_{bs} + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u}\tilde{\sigma}_{k_g^d,i_j^u}^2 P_{ue}. \quad (56)$$

Thus, we can derive according to (9)

$$Cov\left(\mathbf{y}_{k_g^d}\right) = \mathbf{C}_{k_g^d} + \sigma_r^2\mathcal{D}\left(\mathbf{C}_{k_g^d}\right). \quad (57)$$

Similarly, the covariance matrices for the signal received by the gth BS can be written as

$$\mathbf{C}_g = \mathbb{E}\left\{(\mathbf{y}_g - \mathbf{e}_g)(\mathbf{y}_g - \mathbf{e}_g)^H\right\}$$

$$= \sum_{j=1}^{G}\hat{\mathbf{H}}_{g,j}\mathbf{T}_j\hat{\mathbf{H}}_{g,j}^H + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u}\hat{\mathbf{H}}_{g,i_j^u}\mathbf{T}_{i_j^u}\hat{\mathbf{H}}_{g,i_j^u}^H$$

$$+ \left(\tilde{\sigma}_g^2 + \sigma_{bs}^2\right)\mathbf{I} \quad (58)$$

$$Cov\left(\mathbf{y}_g\right) = \mathbf{C}_g + \sigma_r^2\mathcal{D}\left(\mathbf{C}_g\right), \quad (59)$$

where $\bar{\sigma}_g^2 = \sum_{j=1}^{G}\tilde{\sigma}_{g,j}^2 P_{bs} + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u}\tilde{\sigma}_{g,i_j^u}^2 P_{ue}$.

APPENDIX B

In this appendix, we give detailed expressions for the sum of MSE. The MSE of a single downlink or uplink user is given as

$$\mathbb{E}\left\{\left\|\mathbf{s}_{k_g^d} - \mathbf{U}_{k_g^d}^H\mathbf{y}_{k_g^d}\right\|_F^2\right\}$$

$$= 1 - 2\frac{\sqrt{P_d}}{b_d}\Re\left\{tr\left(\mathbf{U}_{k_g^d}^H\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d}\right)\right\} + tr\left(\mathbf{U}_{k_g^d}^H\mathbf{C}_{k_g^d}\mathbf{U}_{k_g^d}\right)$$

$$+ \sigma_r^2 tr\left(\mathbf{U}_{k_g^d}^H\mathcal{D}\left(\mathbf{C}_{k_g^d}\right)\mathbf{U}_{k_g^d}\right),$$

$$\mathbb{E}\left\{\left\|\mathbf{s}_{k_g^u} - \mathbf{U}_{k_g^u}^H \mathbf{y}_g\right\|_F^2\right\}$$

$$= 1 - 2\frac{\sqrt{P_u}}{b_u}\Re\left\{tr\left(\mathbf{U}_{k_g^u}^H \mathbf{H}_{g,k_g^u} \mathbf{V}_{k_g^u}\right)\right\} + tr\left(\mathbf{U}_{k_g^u}^H \mathbf{C}_g \mathbf{U}_{k_g^u}\right)$$

$$+ \sigma_r^2 tr\left(\mathbf{U}_{k_g^u}^H \mathcal{D}\left(\mathbf{C}_g\right) \mathbf{U}_{k_g^u}\right).$$

Thus, we can give the sum of MSE as Equation (55).

REFERENCES

[1] 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.

[2] A. Bishnu, M. Holm, and T. Ratnarajah, "Performance Evaluation of Full-Duplex Enabled 5G IAB Network for Multi-Cell and Multi-User Environment for FR2 Band," *IEEE Access*, vol. 9, pp. 72269-72283, May 2021.

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

[4] T. Zhang, S. Biswas, and T. Ratnarajah, "An analysis on wireless edge caching in in-band full-duplex FR2-IAB networks," *IEEE Access*, vol. 8, pp. 164987-165002, 2020.

[5] H. Luo, M. Holm, T. Ratnarajah, "Wideband Active Analog Self- Interference Cancellation for 5G and Beyond Full-Duplex Systems," *In Proc. Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, California, pp. 868-872, Nov. 2020.

[6] H. Luo, M. Holm, and T. Ratnarajah, "On the performance of active analog self-interference cancellation techniques for beyond 5G systems," *China Commun.*, vol. 18, no. 10, pp. 158-168, 2021.

[7] 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. Wirel. Commun.*, vol. 21, no. 6, pp. 4183-4196, Jun. 2022.

[8] 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.

[9] 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.

[10] M. Mohammadi, H. A. Suraweera, Y. Cao, I. Krikidis and C. Tellambura, "Full-Duplex Radio for Uplink/Downlink Wireless Access With Spatially Random Nodes," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 5250-5266, Dec. 2015.

[11] M. Mohammadi, H. A. Suraweera and C. Tellambura, "Uplink/Downlink Rate Analysis and Impact of Power Allocation for Full-Duplex Cloud-RANs," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 9, pp. 5774-5788, Sep. 2018.

[12] R. Shafin and L. Liu, "Multi-Cell Multi-User Massive FD-MIMO: Downlink Precoding and Throughput Analysis," *IEEE Trans. Wirel. Commun.*, vol. 18, no. 1, pp. 487-502, Jan. 2019.

[13] J. Ye, A. Kammoun, and M.S. Alouini, "Sum-Rate Analysis of a Multi-cell Multi-user MISO System under Double Scattering Channels," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 332-349, 2022.

[14] 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.

[15] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 2, pp. 680-694, 2014.

[16] K. Komatsu, Y. Miyaji, and H. Uehara, "Iterative nonlinear self-interference cancellation for in-band full-duplex wireless communications under mixer imbalance and amplifier nonlinearity," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 7, pp. 4424-4438, 2020.

[17] Y. Kurzo, A.T. Kristensen, A. Burg, A. Balatsoukas-Stimming, "Hardware implementation of neural self-interference cancellation," *IEEE Trans. Emerg. Sel. Topics Circuits Syst.*, vol. 10, no. 2, pp. 204-216, 2020.

[18] 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.

[19] E. Balti, and N. Mensi, "Zero-forcing max-power beamforming for hybrid mmwave full-duplex mimo systems," *In Proc. lnt. Conf. Adv. Syst. Emerg. Technol.*, pp. 344-349, 2020.

[20] I.P. Roberts, J. G. Andrews, and S. Vishwanath, "Hybrid beamforming for millimeter wave full-duplex under limited receive dynamic range," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 12, pp. 7758-7772, 2021.

[21] J. Guo, C.K. Wen, and S. Jin, "Deep learning-based CSI feedback for beamforming in single-and multi-cell massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1872-1884, 2020.

[22] H. Ye, F. Gao, J. Qian, H. Wang, and G.Y. Li, "Deep learning-based denoise network for CSI feedback in FDD massive MIMO systems," *IEEE Commun. Lett.*, vol. 24, no. 8, pp. 1742-1746, 2020.

[23] X. Chen, and H.H. Chen, "Physical layer security in multi-cell MISO downlinks with incomplete CSI—A unified secrecy performance analysis," *IEEE Trans. Signal Process.*, vol. 62, no. 23, pp. 6286-6297, 2014.

[24] S. Qiu, D. Chen, D. Qu, K. Luo, and T. Jiang, "Downlink precoding with mixed statistical and imperfect instantaneous CSI for massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3028-3041, 2017.

[25] I. Viering, H. Hofstetter, and W. Utschick, "Spatial long-term variations in urban, rural and indoor environments," *In Proc. COST*, vol. 273, Sep. 2002.

[26] X. Mestre, "Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates," *IEEE Trans. Inf. Theory.*, vol. 54, no. 11, pp. 5113-5129, 2008.

[27] X. Li, S. Jin, X. Gao, and R.W. Heath, "Three-dimensional beamforming for large-scale FD-MIMO systems exploiting statistical channel state information," *IEEE Trans. Veh. Technol.*, vol. 65, no. 11, pp. 8992-9005, 2016.

[28] 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.

[29] Y. Han, W. Tang, S. Jin, C.K. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8238-8242, 2019.

[30] 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.

[31] J. Li, Q. Lv, P. Zhu, D. Wang, J. Wang, and X. You, "Network-Assisted Full-Duplex Distributed Massive MIMO Systems With Beamforming Training Based CSI Estimation," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 4, pp. 2190-2204, 2021.

[32] Y. Mostofi, and D.C. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. Wirel. Commun.*, vol. 4, no. 2, pp. 765-774, 2005.

[33] S. Boyd, S.P. Boyd, and L. Vandenberghe, "Convex optimization," *Cambridge university press*, Mar. 2004.

[34] *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, 2020.