

Robust Beamforming for IRS Aided MIMO Full Duplex Systems

Chandan Kumar Sheemar, Jorge Querol, Sourabh Solanki, Sumit Kumar, and Symeon Chatzinotas
SnT, University of Luxembourg, emails: {name.surname}@uni.lu

Abstract—In this paper, a novel robust beamforming for an intelligent reflecting surface (IRS) assisted FD system is presented. Since perfect channel state information (CSI) is often challenging to acquire in practice, we consider the case of imperfect CSI and adopt a statistically robust beamforming approach to maximize the ergodic weighted sum rate (WSR). We also analyze the achievable WSR of an IRS-assisted FD with imperfect CSI, for which the lower and the upper bounds are derived. The ergodic WSR maximization problem is tackled based on the expected Weighted Minimum Mean Squared Error (WMMSE), which is guaranteed to converge to a local optimum. The effectiveness of the proposed design is investigated with extensive simulation results. It is shown that our robust design achieves significant performance gain compared to the naive beamforming approaches and considerably outperforms the robust Half-Duplex (HD) system.

Index Terms—full duplex, intelligent reflecting surfaces, robust beamforming, ergodic weighted sum rate.

I. INTRODUCTION

FULL DUPLEX (FD) is a promising wireless transmission technology offering simultaneous transmission and reception in the same frequency band, which theoretically doubles the spectral efficiency [1], [2]. Beyond spectral efficiency, FD can be beneficial to improve security, enable advanced joint communication and sensing, and reduce end-to-end delays. However, FD systems suffer from Self-Interference (SI), which could be 90 – 110 dB higher than the received signal of interest, and it is a major challenge for achieving an ideal FD operation. Advanced SI Cancellation (SIC) techniques are pivotal to mitigate the SI power close to the noise floor and enable correct reception of the received signal. In parallel to FD, another emerging technology is the Intelligent Reflecting Surfaces (IRSs), which can provide a smart and reconfigurable wireless environment [3]–[5].

The IRS-aided FD (IRS-FD) systems hold great promise in supporting the forthcoming traffic demands while being highly spectrally and energy efficient [6]–[8]. In [3], the authors investigated the performance of the IRS-FD system by analyzing the outage and error probabilities. In [9], the authors present a novel beamforming design for an FD relay assisted with one IRS to maximize the minimum achievable rate with the max-min optimization. In [10], the authors studied the IRS-FD systems to improve the worst-case achievable security rate under imperfect channel state information (CSI). Furthermore, a deep neural network is also developed to reduce the computational burden of the proposed beamforming solution. Finally, in [11], the authors presented a joint beamforming design for weighted

sum rate (WSR) maximization in a single-cell multiple-input single-output (MISO) IRS-FD system. Note that the existing literature only provides the secrecy rate expression for IRS-FD systems under imperfect Channel State Information (CSI). However, there is currently no analytical expression available for the ergodic weighted sum rate (WSR) or a novel robust beamforming design considering imperfect CSI in this context.

Motivated by the aforementioned considerations, we address the challenge of maximizing the ergodic WSR for IRS-FD systems with multi-antenna users under imperfect CSI. Our objective is to enhance both the uplink (UL) and downlink (DL) sum rates by jointly optimizing the digital beamformers and the IRS phase response while accounting for imperfect CSI. It is important to note that incorporating CSI errors in the beamforming designs, with multi-antenna users, makes the problem of MIMO beamforming for IRS-FD systems highly complex. The CSI errors are modelled by using the Gaussian-Kronecker model and then an analytical approximation for the ergodic achievable WSR using this model is derived. Subsequently, we formulate the problem of maximizing the ergodic WSR, considering average sum-power and unit-modulus constraints for the IRS phase response under imperfect CSI. To solve this problem, we leverage the connection between the ergodic WSR and the expected weighted minimum mean squared error (EWMMSE), where the expectation is taken with respect to the CSI errors. Simulation results validate the accuracy of the analytically derived ergodic WSR expression. It is further shown that our proposed approach significantly outperforms the naive schemes in the presence of imperfect CSI.¹

II. SYSTEM MODEL

Let j and k denote the multi-antenna DL and UL user served by the MIMO FD base station (BS), respectively, and let M_k and N_j denote their number of transmit and receive antennas, respectively. The FD BS is assumed to have M_0 transmit and N_0 receive antennas. We consider a multi-stream approach, and the number of data streams for the UL user k and DL user j is denoted as u_k and v_j , respectively. Let $\mathbf{U}_k \in \mathbb{C}^{M_k \times u_k}$ and $\mathbf{V}_j \in \mathbb{C}^{M_j \times v_j}$ denote the precoders for white unitary-variance

¹*Notations:* Boldface lower and upper case characters denote vectors and matrices, respectively. $\mathbb{E}\{\cdot\}$, $\text{Tr}\{\cdot\}$, \mathbf{I} , \otimes , $\tilde{\mathbf{X}}$ denote expectation, trace, identity matrix, Kronecker product, and transmit covariance matrix, respectively. Finally, $\text{diag}(\mathbf{x})$ and \mathbf{x} denote a diagonal matrix with vector \mathbf{x} on its main diagonal and \mathbf{I} denotes identity matrix.

data streams $\mathbf{s}_k \in \mathbb{C}^{u_k \times 1}$ and $\mathbf{s}_j \in \mathbb{C}^{v_j \times 1}$, respectively. We assume that the considered FD system is aided with one IRS of size $R \times C$. Let $\boldsymbol{\theta} = [e^{i\theta_1}, \dots, e^{i\theta_{RC}}]$ denote the vector containing the phase-shift response of its RC elements, and let $\boldsymbol{\Theta} = \text{diag}(\boldsymbol{\theta})$ denote a diagonal matrix containing $\boldsymbol{\theta}$ on its main diagonal. Let \mathbf{n}_0 and \mathbf{n}_j denote the noise vectors at the FD BS and DL user j , respectively, which are modelled as $\mathbf{n}_0 = \mathcal{CN}(\mathbf{0}, \sigma_0^2 \mathbf{I})$ $\mathbf{n}_j = \mathcal{CN}(\mathbf{0}, \sigma_j^2 \mathbf{I})$, where σ_0^2 and σ_j^2 denote the noise variances at the FD node and the DL user j , respectively. The channel responses from the UL user k to the BS and from the BS to the DL user j are denoted with $\mathbf{H}_k \in \mathbb{C}^{N_0 \times M_k}$ and $\mathbf{H}_j \in \mathbb{C}^{N_j \times M_0}$, respectively. Let $\mathbf{H}_0 \in \mathbb{C}^{N_0 \times M_0}$ and $\mathbf{H}_{j,k} \in \mathbb{C}^{N_j \times M_k}$ denote the SI channel response for the FD BS and cross-interference channel response between the UL user k and the DL user j , respectively. The channel responses from the transmit antenna array of the FD BS to the IRS and from the IRS to the receive antenna array of the FD BS are denoted with $\mathbf{H}_{\theta,0} \in \mathbb{C}^{RC \times M_0}$ and $\mathbf{H}_{0,\theta} \in \mathbb{C}^{N_0 \times RC}$, respectively. Let $\mathbf{H}_{j,\theta} \in \mathbb{C}^{N_j \times RC}$ and $\mathbf{H}_{\theta,k} \in \mathbb{C}^{RC \times M_k}$ denote the channel responses from the IRS to the DL user j and from the UL user k to the IRS, respectively.

A. Imperfect CSI Modelling

Let $\Delta\mathbf{H}_k, \Delta\mathbf{H}_j, \Delta\mathbf{H}_0, \Delta\mathbf{H}_{j,k}, \Delta\mathbf{H}_{\theta,0}, \Delta\mathbf{H}_{0,\theta}, \Delta\mathbf{H}_{j,\theta}, \Delta\mathbf{H}_{\theta,k}$ denote the estimation errors for the channel responses $\mathbf{H}_k, \mathbf{H}_j, \mathbf{H}_0, \mathbf{H}_{j,k}, \mathbf{H}_{\theta,0}, \mathbf{H}_{0,\theta}, \mathbf{H}_{j,\theta}, \mathbf{H}_{\theta,k}$, respectively. The available CSI matrices can be written as a sum of the channel estimates and CSI errors as

$$\begin{aligned} \mathbf{H}_k &= \hat{\mathbf{H}}_k + \Delta\mathbf{H}_k, & \mathbf{H}_j &= \hat{\mathbf{H}}_j + \Delta\mathbf{H}_j, \\ \mathbf{H}_0 &= \hat{\mathbf{H}}_0 + \Delta\mathbf{H}_0, & \mathbf{H}_{j,k} &= \hat{\mathbf{H}}_{j,k} + \Delta\mathbf{H}_{j,k}, \\ \mathbf{H}_{\theta,0} &= \hat{\mathbf{H}}_{\theta,0} + \Delta\mathbf{H}_{\theta,0}, & \mathbf{H}_{0,\theta} &= \hat{\mathbf{H}}_{0,\theta} + \Delta\mathbf{H}_{0,\theta}, \\ \mathbf{H}_{j,\theta} &= \hat{\mathbf{H}}_{j,\theta} + \Delta\mathbf{H}_{j,\theta}, & \mathbf{H}_{\theta,k} &= \hat{\mathbf{H}}_{\theta,k} + \Delta\mathbf{H}_{\theta,k}, \end{aligned} \quad (1)$$

where the channel matrices of the form $\hat{\mathbf{X}}$ denote the channel estimates. To model the estimation errors, we adopt the Gaussian Kronecker model [12], which dictates that

$$\begin{aligned} \Delta\mathbf{H}_k &= \mathcal{CN}(\mathbf{0}, \mathbf{J}_k \otimes \mathbf{K}_k), & \Delta\mathbf{H}_{\theta,0} &= \mathcal{CN}(\mathbf{0}, \mathbf{J}_{\theta,0} \otimes \mathbf{K}_{\theta,0}), \\ \Delta\mathbf{H}_0 &= \mathcal{CN}(\mathbf{0}, \mathbf{J}_0 \otimes \mathbf{K}_0), & \Delta\mathbf{H}_{j,k} &= \mathcal{CN}(\mathbf{0}, \mathbf{J}_{j,k} \otimes \mathbf{K}_{j,k}), \\ \Delta\mathbf{H}_j &= \mathcal{CN}(\mathbf{0}, \mathbf{J}_j \otimes \mathbf{K}_j), & \Delta\mathbf{H}_{0,\theta} &= \mathcal{CN}(\mathbf{0}, \mathbf{J}_0 \otimes \mathbf{K}_\theta), \\ \Delta\mathbf{H}_{j,\theta} &= \mathcal{CN}(\mathbf{0}, \mathbf{J}_{j,\theta} \otimes \mathbf{K}_{j,\theta}), & \Delta\mathbf{H}_{\theta,k} &= \mathcal{CN}(\mathbf{0}, \mathbf{J}_{\theta,k} \otimes \mathbf{K}_{\theta,k}), \end{aligned} \quad (2)$$

where the matrices \mathbf{J} and \mathbf{K} are the covariance matrices seen from the transmitter and receiver for each link, respectively. The estimation errors are uncorrelated with the estimated channel matrices, and hence we can write

$$\begin{aligned} \mathbf{H}_k &= \mathcal{CN}(\hat{\mathbf{H}}_k, \mathbf{J}_k \otimes \mathbf{K}_k), & \mathbf{H}_{\theta,0} &= \mathcal{CN}(\hat{\mathbf{H}}_{\theta,0}, \mathbf{J}_{\theta,0} \otimes \mathbf{K}_{\theta,0}), \\ \mathbf{H}_0 &= \mathcal{CN}(\hat{\mathbf{H}}_0, \mathbf{J}_0 \otimes \mathbf{K}_0), & \mathbf{H}_{j,k} &= \mathcal{CN}(\hat{\mathbf{H}}_{j,k}, \mathbf{J}_{j,k} \otimes \mathbf{K}_{j,k}), \\ \mathbf{H}_j &= \mathcal{CN}(\hat{\mathbf{H}}_j, \mathbf{J}_j \otimes \mathbf{K}_j), & \mathbf{H}_{\theta,k} &= \mathcal{CN}(\hat{\mathbf{H}}_{\theta,k}, \mathbf{J}_{\theta,k} \otimes \mathbf{K}_{\theta,k}), \\ \mathbf{H}_{j,\theta} &= \mathcal{CN}(\hat{\mathbf{H}}_{j,\theta}, \mathbf{J}_{j,\theta} \otimes \mathbf{K}_{j,\theta}), & \mathbf{H}_{0,\theta} &= \mathcal{CN}(\hat{\mathbf{H}}_{0,\theta}, \mathbf{J}_{0,\theta} \otimes \mathbf{K}_{0,\theta}), \end{aligned} \quad (3)$$

Let $\bar{\mathbf{H}}_{0,k}, \bar{\mathbf{H}}_{0,0}, \bar{\mathbf{H}}_j, \bar{\mathbf{H}}_{j,k}$ denote the effective channel responses affected by the estimation errors defined as

$$\bar{\mathbf{H}}_k = (\hat{\mathbf{H}}_k + \Delta\mathbf{H}_k) + (\hat{\mathbf{H}}_{0,\theta} + \Delta\mathbf{H}_{0,\theta})\boldsymbol{\Theta}(\hat{\mathbf{H}}_{\theta,k} + \Delta\mathbf{H}_{\theta,k}) \quad (4a)$$

$$\bar{\mathbf{H}}_0 = (\hat{\mathbf{H}}_0 + \Delta\mathbf{H}_0) + (\hat{\mathbf{H}}_{0,\theta} + \Delta\mathbf{H}_{0,\theta})\boldsymbol{\Theta}(\hat{\mathbf{H}}_{\theta,0} + \Delta\mathbf{H}_{\theta,0}) \quad (4b)$$

$$\bar{\mathbf{H}}_j = (\hat{\mathbf{H}}_j + \Delta\mathbf{H}_j) + (\hat{\mathbf{H}}_{j,\theta} + \Delta\mathbf{H}_{j,\theta})\boldsymbol{\Theta}(\hat{\mathbf{H}}_{\theta,0} + \Delta\mathbf{H}_{\theta,0}) \quad (4c)$$

$$\bar{\mathbf{H}}_{j,k} = (\hat{\mathbf{H}}_{j,k} + \Delta\mathbf{H}_{j,k}) + (\hat{\mathbf{H}}_{j,\theta} + \Delta\mathbf{H}_{j,\theta})\boldsymbol{\Theta}(\hat{\mathbf{H}}_{\theta,k} + \Delta\mathbf{H}_{\theta,k}) \quad (4d)$$

Let \mathbf{y}_k and \mathbf{y}_j denote the signals received by the FD BS from UL user k and by the DL user j , respectively, given as

$$\mathbf{y}_k = \bar{\mathbf{H}}_k \mathbf{U}_k \mathbf{s}_k + \bar{\mathbf{H}}_0 \mathbf{V}_j \mathbf{s}_j + \mathbf{n}_0, \quad (5a)$$

$$\mathbf{y}_j = \bar{\mathbf{H}}_j \mathbf{V}_j \mathbf{s}_j + \bar{\mathbf{H}}_{j,k} \mathbf{U}_k \mathbf{s}_k + \mathbf{n}_j. \quad (5b)$$

B. Problem Formulation

We aim to maximize the ergodic WSR of the IRS-FD system under imperfect CSI given the CSI error statistics and the channel estimates. Let $\mathcal{R} = \mathcal{R}_k + \mathcal{R}_j$ denote the WSR of the system in the case of perfect CSI, with \mathcal{R}_k and \mathcal{R}_j denoting the weighted rate of the users k and j , respectively. The ergodic WSR of the system, i.e., the average WSR with respect to the CSI errors, can be written as $\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}[\mathcal{R}]$. However, such a problem is infeasible to be solved and can be tackled by applying Jensen's inequality, which allows moving the expectation operator insider as $\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}[\mathcal{R}] \geq \mathcal{R}(\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}})$ [13]. The optimization problem for ergodic WSR $\mathcal{R}(\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}})$, under the sum-power and the unit-modulus constraint can be written as

$$\max_{\mathbf{V}_j, \mathbf{U}_k, \boldsymbol{\Theta}} \mathcal{R}_k(\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}) + \mathcal{R}_j(\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}) \quad (6a)$$

$$\text{s.t. } \text{Tr}(\mathbf{U}_k \mathbf{U}_k^H) \leq \alpha_k, \text{ \& } \text{Tr}(\mathbf{V}_j \mathbf{V}_j^H) \leq \alpha_0, \quad (6b)$$

$$|\boldsymbol{\theta}(i)| = 1, \quad \forall i, \quad (6c)$$

where α_0 and α_k denote the sum-power constraint at the BS and UL user, respectively.

C. Ergodic WSR Analysis with Imperfect CSI

In the following, we derive the expression for the ergodic WSR $\mathcal{R}(\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}})$ in the presence of estimated channel responses and CSI errors for the IRS-FD system with multi-antenna UL and DL users. Let $\bar{\mathbf{U}}_k = \mathbf{U}_k \mathbf{U}_k^H$ and $\bar{\mathbf{V}}_j = \mathbf{V}_j \mathbf{V}_j^H$ denote the transmit covariance matrices for UL user k and DL user j , respectively. In the case of imperfect CSI, the received signal plus interference plus noise covariance matrices \mathbf{R}_k and \mathbf{R}_j , including the CSI errors and the IRS phase response $\boldsymbol{\Theta}$, are given as

$$\mathbf{R}_k = \bar{\mathbf{H}}_k \bar{\mathbf{U}}_k \bar{\mathbf{H}}_k^H + \bar{\mathbf{H}}_0 \bar{\mathbf{V}}_j \bar{\mathbf{H}}_0^H + \sigma_0^2 \mathbf{I}, \quad (7a)$$

$$\mathbf{R}_j = \bar{\mathbf{H}}_j \bar{\mathbf{V}}_j \bar{\mathbf{H}}_j^H + \bar{\mathbf{H}}_{j,k} \bar{\mathbf{U}}_k \bar{\mathbf{H}}_{j,k}^H + \sigma_j^2 \mathbf{I}. \quad (7b)$$

The interference plus noise covariance matrices can be obtained as $\mathbf{R}_{\bar{k}} = \mathbf{R}_k - \mathbf{S}_k$, $\mathbf{R}_{\bar{j}} = \mathbf{R}_j - \mathbf{S}_j$, with \mathbf{S}_k and \mathbf{S}_j denoting the useful received signal covariance part.

Given the statistical distribution of the CSI errors (2) and the channel estimates, the lower bound for the ergodic WSR $\mathcal{R}(\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}})$ of an IRS-FD system with multi-antenna users with imperfect CSI can be approximated as

$$\begin{aligned} \mathcal{R}(\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}) = & w_k \ln[\det(\mathbf{I} + \mathbf{U}_k^H (\hat{\mathbf{H}}_k + \hat{\mathbf{H}}_{0,\theta} \Theta \hat{\mathbf{H}}_{\theta,k})^H \Sigma_k^{-1} \\ & (\hat{\mathbf{H}}_k + \hat{\mathbf{H}}_{0,\theta} \Theta \hat{\mathbf{H}}_{\theta,k}) \mathbf{U}_k)] + w_j \ln[\det(\mathbf{I} + \mathbf{V}_j^H (\hat{\mathbf{H}}_j + \\ & \hat{\mathbf{H}}_{j,\theta} \Theta \hat{\mathbf{H}}_{\theta,0})^H \Sigma_j^{-1} (\hat{\mathbf{H}}_j + \hat{\mathbf{H}}_{j,\theta} \Theta \hat{\mathbf{H}}_{\theta,0}) \mathbf{V}_j)], \end{aligned} \quad (8)$$

where Σ_k and Σ_j are given as in (9). This result is derived based on the identities: For any $\mathbf{H} \sim \mathcal{CN}(\hat{\mathbf{H}}, \mathbf{J} \otimes \mathbf{K})$, there is $\mathbb{E}[\mathbf{H}\mathbf{X}\mathbf{H}^H] = \hat{\mathbf{H}}\mathbf{X}\hat{\mathbf{H}}^H + \text{Tr}(\mathbf{X}\mathbf{J}^T)\mathbf{K}$ and $\mathbb{E}[\mathbf{H}^H\mathbf{X}\mathbf{H}] = \hat{\mathbf{H}}^H\mathbf{X}\hat{\mathbf{H}} + \text{Tr}(\mathbf{K}\mathbf{X})\mathbf{J}^T$. Note that an upper bound for ergodic WSR can be easily derived from (9) in the presence of ideal CSI, i.e., by setting \mathbf{J} and \mathbf{I} equal to the identity.

III. ROBUST BEAMFORMING VIA EXPECTED WMMSE

The problem of ergodic WSR maximization (6) is non-convex due to interference and in the case of perfect CSI it can be solved with an equivalent problem formulation of EWMMSE, which exploits the relationship between the WSR and the expected error covariance matrices [13].

A. Digital Combining

We assume that the FD BS for UL user k and the DL user j apply the combiners \mathbf{F}_k and \mathbf{F}_j to estimate their data streams as $\hat{\mathbf{s}}_k = \mathbf{F}_k \mathbf{y}_k$ and $\hat{\mathbf{s}}_j = \mathbf{F}_j \mathbf{y}_j$. Let \mathbf{E}_k and \mathbf{E}_j denote the MSE error matrices for instantaneous CSI for UL user k and DL user j , respectively, which can be written as $\mathbf{E}_k = \mathbb{E}[\mathbf{e}_k \mathbf{e}_k^H]$, $\mathbf{E}_j = \mathbb{E}[\mathbf{e}_j \mathbf{e}_j^H]$ where $\mathbf{e}_k = \mathbf{F}_k \mathbf{y}_k - \hat{\mathbf{s}}_k$ and $\mathbf{e}_j = \mathbf{F}_j \mathbf{y}_j - \hat{\mathbf{s}}_j$ denote the error vectors. Let $\mathbf{Q}_k, \mathbf{T}_j, \mathbf{T}_0$ and $\mathbf{Q}_{j,k}$ denote the matrices defined as

$$\mathbf{Q}_k = \mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}(\bar{\mathbf{H}}_k \bar{\mathbf{U}}_k \bar{\mathbf{H}}_k^H), \quad \mathbf{T}_j = \mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}(\bar{\mathbf{H}}_j \bar{\mathbf{V}}_j \bar{\mathbf{H}}_j^H), \quad (10a)$$

$$\mathbf{T}_0 = \mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}(\bar{\mathbf{H}}_0 \bar{\mathbf{V}}_0 \bar{\mathbf{H}}_0^H), \quad \mathbf{Q}_{j,k} = \mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}(\bar{\mathbf{H}}_{j,k} \bar{\mathbf{U}}_k \bar{\mathbf{H}}_{j,k}^H). \quad (10b)$$

Since we adopt the EWMMSE approach, consider the expected MSE of the UL and DL user as

$$\begin{aligned} \mathbf{E}_k = & \mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}[\mathbf{E}_k] \\ = & \mathbf{F}_k \mathbf{Q}_k \mathbf{F}_k^H - \mathbf{F}_k (\hat{\mathbf{H}}_k + \hat{\mathbf{H}}_{0,\theta} \Theta \hat{\mathbf{H}}_{\theta,k}) \mathbf{U}_k + \mathbf{F}_k \mathbf{T}_0 \mathbf{F}_k^H \\ & + \sigma_0^2 \mathbf{F}_k \mathbf{F}_k^H - \mathbf{U}_k^H \hat{\mathbf{H}}_k^H \mathbf{F}_k^H - \mathbf{U}_k^H \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \hat{\mathbf{H}}_{0,\theta}^H \mathbf{F}_k^H - \mathbf{I}. \end{aligned} \quad (11a)$$

$$\begin{aligned} \mathbf{E}_j = & \mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}[\mathbf{E}_j] \\ = & \mathbf{F}_j \mathbf{T}_j \mathbf{F}_j^H - \mathbf{F}_j (\hat{\mathbf{H}}_j + \hat{\mathbf{H}}_{j,\theta} \Theta \hat{\mathbf{H}}_{\theta,0}) \mathbf{V}_j + \mathbf{F}_j \mathbf{Q}_{j,k} \mathbf{F}_j^H \\ & + \sigma_j^2 \mathbf{F}_j \mathbf{F}_j^H - \mathbf{V}_j^H \hat{\mathbf{H}}_j^H \mathbf{F}_j^H - \mathbf{V}_j^H \hat{\mathbf{H}}_{\theta,0}^H \Theta^H \hat{\mathbf{H}}_{j,\theta}^H \mathbf{F}_j^H - \mathbf{I}. \end{aligned} \quad (11b)$$

By minimizing the trace of the (11), the optimal EWMMSE combiners can be derived as

$$\mathbf{F}_k = \mathbf{U}_k^H (\hat{\mathbf{H}}_k^H + \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \hat{\mathbf{H}}_{0,\theta}^H) (\mathbf{Q}_k + \mathbf{T}_0 + \sigma_0^2 \mathbf{I})^{-1}, \quad (12a)$$

$$\mathbf{F}_j = \mathbf{V}_j^H (\hat{\mathbf{H}}_j^H + \hat{\mathbf{H}}_{\theta,0}^H \Theta^H \hat{\mathbf{H}}_{j,\theta}^H) (\mathbf{T}_j + \mathbf{Q}_{j,k} + \sigma_j^2 \mathbf{I})^{-1}. \quad (12b)$$

B. Active Digital Beamforming Under Imperfect CSI

Given the optimal combiners, the EWMMSE problem with respect to the digital beamformers under the average total sum-power constraint and given the IRS phase response Θ fixed, can be formally stated as

$$\min_{\mathbf{V}_j, \mathbf{U}_k} \text{Tr}(\mathbf{W}_k \mathbf{E}_k) + \text{Tr}(\mathbf{W}_j \mathbf{E}_j), \quad (13a)$$

$$\text{s.t. } \text{Tr}(\mathbf{U}_k \mathbf{U}_k^H) \leq \alpha_k, \quad \& \quad \text{Tr}(\mathbf{V}_j \mathbf{V}_j^H) \leq \alpha_0, \quad (13b)$$

where \mathbf{W}_i is a constant weight matrix associated with node i . The problem (13) and the WSR maximization problem (6) are equivalent if their gradient results to be the same, which can be assured if the weight matrices are chosen as

$$\mathbf{W}_k = \frac{w_k}{\ln 2} (\mathbf{E}_k)^{-1}, \quad \mathbf{W}_j = \frac{w_j}{\ln 2} (\mathbf{E}_j)^{-1}. \quad (14)$$

Such a result can be easily shown by following a similar proof provided in [14, Appendix A], but carried out for the case of average MSE error under the imperfect CSI case. To optimize the digital beamformers $\mathbf{V}_j, \mathbf{U}_k$ we take the partial derivative of the Lagrangian function of (13) with respect to their conjugate, which leads to the following optimal beamformers

$$\mathbf{U}_k = (\mathbf{X}_k + \lambda_k \mathbf{I})^{-1} (\hat{\mathbf{H}}_k^H + \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \hat{\mathbf{H}}_{0,\theta}^H) \mathbf{F}_k^H \mathbf{W}_k, \quad (15a)$$

$$\mathbf{V}_j = (\mathbf{X}_j + \lambda_0 \mathbf{I})^{-1} (\hat{\mathbf{H}}_j^H + \hat{\mathbf{H}}_{\theta,0}^H \Theta^H \hat{\mathbf{H}}_{j,\theta}^H) \mathbf{F}_j^H \mathbf{W}_j, \quad (15b)$$

where \mathbf{X}_k and \mathbf{X}_j are defined in (15a) and (15b), respectively, and the scalars λ_k and λ_0 denote the Lagrange multiplier for the uplink user k and the FD BS. The multipliers can be searched while performing power allocation for the users given the average total sum-power constraints. Namely, consider the singular value decomposition (SVD) of the matrices as $\mathbf{X}_k = \mathbf{A}_k \mathbf{\Lambda}_k \mathbf{B}_k$ and $\mathbf{X}_j = \mathbf{A}_j \mathbf{\Lambda}_j \mathbf{B}_j$, where \mathbf{A}_i and \mathbf{B}_i denote the left and right unitary matrices obtained with SVD and $\mathbf{\Lambda}_i$ denote the singular values. The average power constraints (13b), after some simplifications, can be written as

$$\text{Tr}(\mathbf{V}_j \mathbf{V}_j^H) = \frac{\sum_{i=1}^{M_0} \mathbf{S}_j(i, i)}{(\lambda_0 + \mathbf{\Lambda}_j(i, i))^2}, \quad (17a)$$

$$\text{Tr}(\mathbf{U}_k \mathbf{U}_k^H) = \frac{\sum_{i=1}^{M_k} \mathbf{S}_k(i, i)}{(\lambda_j + \mathbf{\Lambda}_k(i, i))^2}, \quad (17b)$$

where the matrices \mathbf{S}_k and \mathbf{S}_j are defined as

$$\mathbf{S}_j = \mathbf{B}_j (\hat{\mathbf{H}}_j + \hat{\mathbf{H}}_{j,\theta} \Theta \hat{\mathbf{H}}_{\theta,0})^H \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \mathbf{B}_j (\hat{\mathbf{H}}_j + \hat{\mathbf{H}}_{j,\theta} \Theta \hat{\mathbf{H}}_{\theta,0}) \mathbf{A}_j, \quad (18a)$$

$$\mathbf{S}_k = \mathbf{B}_k (\hat{\mathbf{H}}_k + \hat{\mathbf{H}}_{0,\theta} \Theta \hat{\mathbf{H}}_{\theta,k})^H \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k (\hat{\mathbf{H}}_k + \hat{\mathbf{H}}_{0,\theta} \Theta \hat{\mathbf{H}}_{\theta,k}) \mathbf{A}_k. \quad (18b)$$

The optimal Lagrange multipliers satisfying the average total power constraints can be searched with a linear search, and in this work, we adopt the Bisection method. If the obtained values of the multipliers are negative, then we replace them with zero. Note that the average power constraint is met based on the CSI estimates.

$$\begin{aligned}\Sigma_{\bar{k}} = & \text{Tr}(\widetilde{\mathbf{U}}_k \mathbf{J}_{\theta,k}^T) \mathbf{K}_k + \hat{\mathbf{H}}_{0,\theta} \Theta \text{Tr}(\widetilde{\mathbf{U}}_k \mathbf{J}_{\theta,k}^T) \mathbf{K}_{\theta,k} \Theta^H \hat{\mathbf{H}}_{0,\theta}^H + \text{Tr}(\Theta \hat{\mathbf{H}}_{\theta,k} \widetilde{\mathbf{U}}_k \hat{\mathbf{H}}_{\theta,k} \Theta^H \mathbf{J}_{0,\theta}^T) \mathbf{K}_{0,\theta} + \text{Tr}(\widetilde{\mathbf{U}}_k \mathbf{J}_{\theta,k}^T) \text{Tr}(\Theta \mathbf{K}_{\theta,k} \Theta^H \mathbf{J}_{0,\theta}^T) \mathbf{K}_{0,\theta} \\ & + \hat{\mathbf{H}}_0 \widetilde{\mathbf{V}}_j \hat{\mathbf{H}}_0^H + \hat{\mathbf{H}}_0 \widetilde{\mathbf{V}}_j \mathbf{H}_{\theta,0}^H \Theta^H \hat{\mathbf{H}}_{0,\theta} + \text{Tr}(\widetilde{\mathbf{V}}_j \mathbf{J}_0^T) \mathbf{K}_0^T + \hat{\mathbf{H}}_{0,\theta} \Theta \hat{\mathbf{H}}_{\theta,0} \widetilde{\mathbf{V}}_j \hat{\mathbf{H}}_0^H + \hat{\mathbf{H}}_{0,\theta} \Theta \mathbf{H}_{\theta,0} \widetilde{\mathbf{V}}_j \mathbf{H}_{\theta,0}^H \Theta^H \hat{\mathbf{H}}_{0,\theta}^H \\ & + \hat{\mathbf{H}}_{0,\theta} \Theta \text{Tr}(\widetilde{\mathbf{V}}_j \mathbf{J}_{\theta,0}^T) \mathbf{K}_{\theta,0} \Theta^H \hat{\mathbf{H}}_{0,\theta}^H + \text{Tr}(\Theta \hat{\mathbf{H}}_{\theta,0} \widetilde{\mathbf{V}}_j \hat{\mathbf{H}}_{\theta,0}^H \Theta^H \mathbf{J}_{\theta,0}^T) \mathbf{K}_{0,\theta} + \text{Tr}(\widetilde{\mathbf{V}}_j \mathbf{J}_{\theta,0}^T) \text{Tr}(\Theta \mathbf{K}_{\theta,0} \Theta^H \mathbf{J}_{0,\theta}^T) \mathbf{K}_{0,\theta} + \sigma_0^2 \mathbf{I},\end{aligned}\quad (9a)$$

$$\begin{aligned}\Sigma_{\bar{j}} = & \text{Tr}(\widetilde{\mathbf{V}}_j \mathbf{J}_j^T) \mathbf{K}_j + \hat{\mathbf{H}}_{j,\theta} \Theta \text{Tr}(\widetilde{\mathbf{V}}_j \mathbf{J}_{\theta,0}^T) \mathbf{K}_{\theta,0} \Theta^H \hat{\mathbf{H}}_{j,\theta}^H + \text{Tr}(\Theta \hat{\mathbf{H}}_{\theta,0} \widetilde{\mathbf{V}}_j \hat{\mathbf{H}}_{\theta,0}^H \Theta^H \mathbf{J}_{j,\theta}^T) \mathbf{K}_{j,\theta} + \text{Tr}(\widetilde{\mathbf{V}}_j \mathbf{J}_{\theta,0}^T) \text{Tr}(\Theta \mathbf{K}_{\theta,0} \Theta^H \mathbf{J}_{j,\theta}^T) \mathbf{K}_{j,\theta} \\ & + \hat{\mathbf{H}}_{j,k} \widetilde{\mathbf{U}}_k \hat{\mathbf{H}}_{j,k}^H + \hat{\mathbf{H}}_{j,k} \widetilde{\mathbf{U}}_k \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \hat{\mathbf{H}}_{j,\theta}^H + \text{Tr}(\widetilde{\mathbf{U}}_k \mathbf{J}_{j,k}^T) \mathbf{K}_{j,k} + \hat{\mathbf{H}}_{j,\theta} \Theta \mathbf{H}_{\theta,k} \widetilde{\mathbf{U}}_k \hat{\mathbf{H}}_{j,k}^H + \hat{\mathbf{H}}_{j,\theta} \Theta \mathbf{H}_{\theta,k} \widetilde{\mathbf{U}}_k \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \hat{\mathbf{H}}_{j,\theta}^H \\ & + \hat{\mathbf{H}}_{j,\theta} \Theta \text{Tr}(\widetilde{\mathbf{U}}_k \mathbf{J}_{\theta,k}^T) \mathbf{K}_{\theta,k} \Theta^H \hat{\mathbf{H}}_{j,\theta}^H + \text{Tr}(\Theta \hat{\mathbf{H}}_{\theta,k} \widetilde{\mathbf{U}}_k \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \mathbf{J}_{j,\theta}^T) \mathbf{K}_{j,\theta} + \text{Tr}(\widetilde{\mathbf{U}}_k \mathbf{J}_{\theta,k}^T) \text{Tr}(\Theta \mathbf{K}_{\theta,k} \Theta^H \mathbf{J}_{j,\theta}^T) \mathbf{K}_{j,\theta} + \sigma_j^2 \mathbf{I}.\end{aligned}\quad (9b)$$

$$\begin{aligned}\mathbf{X}_k = & (\hat{\mathbf{H}}_k^H + \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \hat{\mathbf{H}}_{0,\theta}^H) \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_k + \text{Tr}(\mathbf{K}_k \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k) \mathbf{J}_k^T + \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \hat{\mathbf{H}}_{0,\theta}^H \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_{0,\theta} \Theta \hat{\mathbf{H}}_{\theta,k} \\ & + \mathbf{H}_k^H \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_{0,\theta} \Theta \hat{\mathbf{H}}_{\theta,k} + \text{Tr}(\mathbf{K}_{\theta,k} \Theta^H \hat{\mathbf{H}}_{0,\theta}^H \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_{0,\theta} \Theta) \mathbf{J}_{\theta,k}^T + \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \text{Tr}(\mathbf{K}_{0,\theta} \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k) \mathbf{J}_{0,\theta}^T \Theta \hat{\mathbf{H}}_{\theta,k} \\ & + \hat{\mathbf{H}}_{j,k}^H \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_{j,\theta} \Theta \hat{\mathbf{H}}_{\theta,k} + (\hat{\mathbf{H}}_{j,k}^H + \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \hat{\mathbf{H}}_{j,\theta}^H) \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_{j,k} + \text{Tr}(\mathbf{K}_{j,k} \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j) \mathbf{J}_{j,k}^T \\ & + \text{Tr}(\mathbf{K}_{0,\theta} \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k) \text{Tr}(\mathbf{K}_{\theta,k} \Theta^H \mathbf{J}_{0,\theta}^T \Theta) \mathbf{J}_{\theta,k}^T + \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \hat{\mathbf{H}}_{j,\theta}^H \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_{j,\theta} \Theta \hat{\mathbf{H}}_{\theta,k} + \text{Tr}(\mathbf{K}_{\theta,k} \Theta^H \hat{\mathbf{H}}_{j,\theta}^H \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_{j,\theta} \Theta) \mathbf{J}_{\theta,k}^T \\ & + \hat{\mathbf{H}}_{\theta,k}^H \Theta^H \text{Tr}(\mathbf{K}_{j,\theta} \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j) \mathbf{J}_{j,\theta}^T \Theta \hat{\mathbf{H}}_{\theta,k} + \text{Tr}(\mathbf{K}_{j,\theta} \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j) \text{Tr}(\mathbf{K}_{\theta,k} \Theta^H \mathbf{J}_{j,\theta}^T \Theta) \mathbf{J}_{\theta,k}^T,\end{aligned}\quad (15a)$$

$$\begin{aligned}\mathbf{X}_j = & (\hat{\mathbf{H}}_j^H + \hat{\mathbf{H}}_{\theta,0}^H \Theta^H \hat{\mathbf{H}}_{j,\theta}^H) \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_j + \text{Tr}(\mathbf{K}_j \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j) \mathbf{J}_j^T + \hat{\mathbf{H}}_{\theta,0}^H \Theta^H \hat{\mathbf{H}}_{j,\theta}^H \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_{j,\theta} \Theta \hat{\mathbf{H}}_{\theta,0} + \hat{\mathbf{H}}_j^H \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_{j,\theta} \Theta \hat{\mathbf{H}}_{\theta,0} \\ & + \text{Tr}(\mathbf{K}_{\theta,0} \Theta^H \hat{\mathbf{H}}_{j,\theta}^H \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_{j,\theta} \Theta) \mathbf{J}_{\theta,0}^T + \hat{\mathbf{H}}_{\theta,0}^H \Theta^H \text{Tr}(\mathbf{K}_{j,\theta} \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j) \mathbf{J}_{j,\theta}^T \Theta \hat{\mathbf{H}}_{\theta,0} + \mathbf{H}_0^H \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_{0,\theta} \Theta \hat{\mathbf{H}}_{\theta,0} \\ & + (\hat{\mathbf{H}}_0^H + \hat{\mathbf{H}}_{\theta,0}^H \Theta^H \hat{\mathbf{H}}_{0,\theta}^H) \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_0 + \text{Tr}(\mathbf{K}_0 \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k) \mathbf{J}_0^T + \text{Tr}(\mathbf{K}_{j,\theta} \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j) \text{Tr}(\mathbf{K}_{\theta,0} \Theta^H \mathbf{J}_{j,\theta}^T \Theta) \mathbf{J}_{\theta,0}^T \\ & + \hat{\mathbf{H}}_{\theta,0}^H \Theta^H \hat{\mathbf{H}}_{0,\theta}^H \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_{0,\theta} \Theta \hat{\mathbf{H}}_{\theta,0} + \text{Tr}(\mathbf{K}_{\theta,0} \Theta^H \hat{\mathbf{H}}_{0,\theta}^H \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_{0,\theta} \Theta) \mathbf{J}_{\theta,0}^T + \hat{\mathbf{H}}_{\theta,0}^H \Theta^H \text{Tr}(\mathbf{K}_{0,\theta} \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k) \mathbf{J}_{0,\theta}^T \Theta \hat{\mathbf{H}}_{\theta,0} \\ & + \text{Tr}(\mathbf{K}_{0,\theta} \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k) \text{Tr}(\mathbf{K}_{\theta,0} \Theta^H \mathbf{J}_{0,\theta}^T \Theta) \mathbf{J}_{\theta,0}^T.\end{aligned}\quad (15b)$$

C. Passive Beamforming Under Imperfect CSI

In this section, we consider optimizing the phase response of the IRS to jointly assist the UL and DL channels for rate enhancement. Let \mathbf{S} , \mathbf{T} and \mathbf{Z} denote the matrices independent of the IRS phase response, given in Appendix A. The EWMMSE optimization problem for the IRS phase response Θ under imperfect CSI, given the matrices \mathbf{S} , \mathbf{T} and \mathbf{Z} , can be formally stated as

$$\min_{\Theta} \text{Tr}(\Theta^H \mathbf{Z} \Theta \mathbf{T}) + \text{Tr}(\Theta^H \mathbf{S}^H) + \text{Tr}(\Theta \mathbf{S}) + c, \quad (19a)$$

$$\text{s.t. } |\theta(i)| = 1, \quad \forall i, \quad (19b)$$

where the scalar c denotes the constant terms, independent of Θ . In (19a), the problem is stated with respect to the matrix Θ . However, we wish to maximize only the diagonal response of such matrix to maximize the ergodic WSR or minimize the expected MSE, because the off-diagonal elements result to be zero. Therefore, we first consider restating the problem (19a) with respect to θ , being a vector made of the diagonal elements of Θ . For doing so, we write the trace terms as

$$\text{Tr}(\Theta^H \mathbf{Z} \Theta \mathbf{T}) = \theta^H \Sigma \theta, \quad \text{where } \Sigma = \mathbf{Z} \odot \mathbf{T}^T, \quad (20a)$$

$$\text{Tr}(\Theta^H \mathbf{S}^H) = \mathbf{s}^H \theta^*, \quad \text{Tr}(\Theta \mathbf{S}) = \mathbf{s}^T \theta, \quad (20b)$$

where \mathbf{s} denotes the vector containing only the diagonal elements of the matrix \mathbf{S} . Problem (19a) can be restated with respect to the vector θ as

$$\min_{\theta} \theta^H \Sigma \theta + \mathbf{s}_r^H \theta^* + \mathbf{s}^T \theta, \quad (21a)$$

$$\text{s.t. } |\theta(i)| = 1, \quad \forall i. \quad (21b)$$

Problem (21) is still very challenging as it is a non-convex problem due to the unit-modulus constraint. To solve it we adopt the majorization-maximization method [15]. Such method aims to solve a more difficult problem by constructing a series of more tractable problems stated with the upper bound. Let $\mathcal{R}(\theta^{(n)})$ denote the function evaluating (21) at the n -th iteration for computed θ . Let $\mathcal{R}_u(\theta|\theta^{(n)})$ denote an upper bound constructed at n -th iteration for \mathcal{R} . According to [15], for the problem of the form (21), an upper bound can be constructed as

$$\mathcal{R}_u(\theta|\theta^{(n)}) = 2\text{Re}\{\mathbf{s}^H \mathbf{q}^{(n)}\} + c_u, \quad (22)$$

where c_u denote constant terms in the upper bound independent of θ and $\mathbf{q}^{(n)}$ is given by

$$\mathbf{q}^{(n)} = (\lambda^{\max} \mathbf{I} - \Sigma) \theta^{(n)} - \mathbf{s}^*, \quad (23)$$

with λ^{\max} denoting the maximum eigenvalues of Σ . Based on the result above, the hard non-convex optimization problem (21a) simplifies to a series of the following problem

$$\min_{\theta} 2\text{Re}\{\mathbf{s}^H \mathbf{q}^{(n)}\}, \quad (24a)$$

$$\text{s.t. } |\theta(i)| = 1, \quad \forall i, \quad (24b)$$

which need to be solved iteratively until convergence for each update of θ . By solving (24a), we get the optimal solution of θ at $n+1$ -th iteration, given the θ at the n -th iteration as

$$\theta^{(n+1)} = e^{i\angle \mathbf{q}^{(n)}}. \quad (25)$$

Formally, the optimization algorithm to optimize the IRS phase response is given in Algorithm 1 and the joint optimization procedure is provided in Algorithm 2.

Algorithm 1 Optimization of IRS Phase Response

Initialize: θ^0 , iteration index $n = 1$, accuracy ϵ .

Evaluate: $\mathcal{R}(\theta^0)$.

Repeat until convergence

Calculate $\mathbf{q}_i^{(n)}$ with (23).

Update $\theta_i^{(n+1)}$ with (25).

if $|\mathcal{R}^{(n+1)} - \mathcal{R}(\theta_i^{(n)})|/\mathcal{R}(\theta_i^{(n+1)}) \leq \epsilon$

Stop and return $\phi_i^{(n+1)}$.

Algorithm 2 Robust Beamforming for MIMO IRS-FD system

Initialize the iteration index n , accuracy ϵ , beamformers and combiners.

Repeat until convergence

for i , where $i = k$ or $i = j$

Update \mathbf{F}_i with (12).

Update \mathbf{W}_i with (14).

if $i=k \rightarrow$ update \mathbf{U}_k with (16a).

else \rightarrow update \mathbf{V}_j with (16b).

Update Θ with Algorithm 2.

if convergence condition is satisfied

Stop and return the optimized variables.

Note that each update of the beamformers leads to a monotonic decrease in the EWMSE which assures the convergence of the proposed methods. A more formal proof can be presented by following a similar strategy as [16]. However, due to space limitations, we omit it here, which will be available in the extended version of the paper.

IV. NUMERICAL RESULTS

In this section, we present extensive simulation results to evaluate the performance of the proposed robust beamforming design. We consider the FD BS and the IRS to be centered in the positions (0m,0m,0m) and (20m,10m,0m), respectively, in the three-dimensional coordinates. The UL and the DL user are assumed to be randomly located in circles of radius $r = 8$ m centered in the positions (20m,0m,30m) and (30m,0m,20m), respectively. We assume that the FD BS and the users are equipped with the uniform linear array (ULA), placed at the distance of half-wavelength. The IRS of $10 \times 10 = 100$ elements is assumed to assist the MIMO FD communication unless otherwise stated. The MIMO FD BS is assumed to have $M_0 = 15$ transmit and $N_0 = 8$ receive antennas. The UL user k and the DL user j is assumed to have $M_k = 5$ transmit and 5 receive antennas, respectively, and served with $d_k = d_j = 2$ data streams. The large scale path loss is modelled as [15] and the SI channel is modelled as a Rician fading [17] with Rician factor 1. The direct links between the users and the FD BS are also modelled with a Rician fading channel model with a Rician factor of 1. The channels involving the IRS are modelled according to Rayleigh fading [18]. We define the transmit signal-to-noise (SNR) of our system as $SNR = \frac{\alpha_0}{\sigma_j^2} = \frac{\alpha_k}{\sigma_0^2}$, where α_0 and α_k are the average total transmit power at the FD BS and the multi-antenna UL user k . We assume that the CSI errors to be i.i.d zero-mean circularly symmetric complex Gaussian distribution with the same variance σ_{csi}^2 , and therefore, the error covariance matrices set chosen as $\mathbf{J}_i = \mathbf{I}$ and $\mathbf{K}_j = \sigma_{csi}^2 \mathbf{I}$, $\forall i, j$. Since the variance of the CSI errors

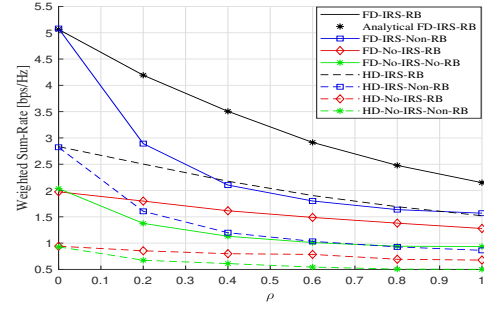


Fig. 1. Average WSR as a function of ρ with SNR= 30 dB.

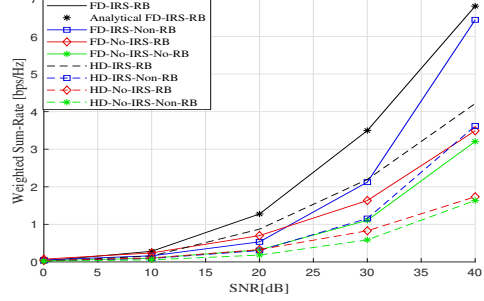


Fig. 2. Average WSR as a function of SNR with $\rho = 0.4$.

strictly depend on the average transmit power (used also to estimate the channels) and the noise variance, we assume that σ_{csi}^2 decays as $\mathcal{O}(SNR^{-\alpha})$, for some constant α , satisfying the CSI error decay rate inversely proportional to the SNR. As the SNR represents the transmit SNR, note that $SNR \rightarrow \infty$ is equivalent to $\alpha_k = \alpha_0 \rightarrow \infty$, which enhances the CSI quality and reduces the CSI error variance as $\sigma_{csi}^2 \rightarrow 0$. The CSI errors variance is set as $\sigma_{csi}^2 = \rho/SNR^\alpha$, where ρ denotes a scale factor and $\alpha \in [0, 1]$ determines the quality of the CSI. Maintaining the quality of the CSI with $\alpha = 1$ could be exhausting in terms of resources required to acquire the CSI and therefore we set $\alpha = 0.6$.

We label our proposed design as a *FD-IRS-RB*. For comparison, we define the following benchmark schemes: 1) *FD-IRS-Non-RB*: FD system assisted with IRS and non-robust beamforming, i.e., CSI is treated as perfect; 2) *FD-No-IRS-RB*: FD system with no IRS and robust beamforming; 3) *FD-No-IRS-Non-RB*: FD system with no IRS and non-robust beamforming; 4) *HD-IRS-RB*: HD system assisted with IRS and robust beamforming; 5) *HD-IRS-Non-RB*: HD system assisted with IRS and non-robust beamforming; 6) *HD-No-IRS-RB*: HD system with no IRS and robust beamforming; 7) *HD-No-IRS-Non-RB*: HD system with no IRS and non-robust beamforming. We also compare our approximation proposed for the analytical ergodic WSR derived above, which is labelled as *Analytical FD-IRS-RB* and compare it with the ergodic rate achieved with the EWMSE robust beamforming approach.

Fig. 1 shows the performance of the proposed robust joint beamforming design as a function of the scale factor ρ dictating the CSI error variance, by means of Monte Carlo simulations at SNR= 30 dB as a function of ρ . It is shown that the proposed robust design achieves significant performance gain compared

to the naive FD-IRS-Non-RB scheme, which does not account for the CSI errors. Moreover, we can also see that the achieved ergodic WSR accurately matches the analytical result derived above. It is to be noted that as the CSI error variance gets extremely large, the proposed statistically robust beamforming design preserves significant robustness against the uncertainties in the CSI quality.

Fig. 2 shows the ergodic WSR of the proposed design as a function of the transmit SNR with $\rho = 0.4$. We can see that our robust design achieves significant gain in the presence of CSI errors compared to the other schemes. Theoretically, the FD systems offer a twofold gain in the WSR compared to the HD systems. However, Fig. 2 for the scheme FD-IRS-Non-RB, at low SNR, i.e., high CSI error variance, shows that the IRS-FD system may achieve no gain compared to the IRS-aided HD system at any SNR level if non-robust beamforming approach is adopted. This is due to the fact that the residual SI can become extremely large in the presence of large CSI errors and can significantly degrade the performance. However, we can observe that as the CSI error variance decreases, the performance of the non-robust beamforming scheme tends toward the performance of the robust beamforming schemes, which tends towards the two-fold improvement in the WSR due to quasi-ideal simultaneous transmission and reception.

V. CONCLUSIONS

In this paper, a novel and statistically robust beamforming design for an IRS-FD system under imperfect CSI is presented. A closed-form approximation for the ergodic WSR is first derived given the statistical distribution of the errors. Then the ergodic WSR maximization problem is handled based on the EWMMSE method leading to two layers of sub-problems which are solved iteratively based on alternating optimization. Simulation results show that imperfect CSI can potentially limit the performance of the IRS-FD systems, and adopting robust beamforming leads to significant improvement in the performance gains.

APPENDIX

$$\begin{aligned} \mathbf{S} = & \hat{\mathbf{H}}_{0,\theta}^H \mathbf{F}_k^H (\mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_k \tilde{\mathbf{U}}_k \hat{\mathbf{H}}_{\theta,k}^H + \mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_0 \tilde{\mathbf{V}}_j \hat{\mathbf{H}}_{\theta,0}^H \\ & + \mathbf{W}_k \mathbf{U}_k^H \hat{\mathbf{H}}_{\theta,k}^H) + \hat{\mathbf{H}}_{j,\theta}^H \mathbf{F}_j^H (\mathbf{W}_j \mathbf{F}_j \mathbf{H}_j \tilde{\mathbf{V}}_j \hat{\mathbf{H}}_{\theta,0}^H \\ & + \mathbf{W}_j \mathbf{F}_j \mathbf{H}_{j,k} \tilde{\mathbf{U}}_k \hat{\mathbf{H}}_{\theta,k}^H - \mathbf{W}_j \mathbf{V}_j^H \hat{\mathbf{H}}_{\theta,0}^H), \end{aligned} \quad (26)$$

$$\begin{aligned} \mathbf{Z} = & \hat{\mathbf{H}}_{0,\theta}^H \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \mathbf{H}_{0,\theta} + \hat{\mathbf{H}}_{0,\theta}^H \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \mathbf{H}_{0,\theta} \\ & + \text{Tr}(\mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \mathbf{J}_{0,\theta}^T) \mathbf{K}_{0,\theta} + \text{Tr}(\mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \mathbf{J}_{0,\theta}^T) \mathbf{K}_{0,\theta} \\ & + \text{Tr}(\mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \mathbf{J}_{0,\theta}^T) \mathbf{K}_{0,\theta} + \hat{\mathbf{H}}_{0,\theta}^H \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_{0,\theta} \\ & + \hat{\mathbf{H}}_{0,\theta}^H \mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \hat{\mathbf{H}}_{0,\theta} + \text{Tr}(\mathbf{F}_k^H \mathbf{W}_k \mathbf{F}_k \mathbf{J}_{0,\theta}^T) \mathbf{K}_{0,\theta} \\ & + \hat{\mathbf{H}}_{j,\theta}^H \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_{j,\theta} + \hat{\mathbf{H}}_{j,\theta}^H \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_{j,\theta} \\ & + \text{Tr}(\mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \mathbf{J}_{j,\theta}^T) \mathbf{K}_{j,\theta} + \text{Tr}(\mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \mathbf{J}_{j,\theta}^T) \mathbf{K}_{j,\theta} \\ & + \hat{\mathbf{H}}_{j,\theta}^H \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_{j,\theta} + \hat{\mathbf{H}}_{j,\theta}^H \mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \hat{\mathbf{H}}_{j,\theta} \\ & + \text{Tr}(\mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \mathbf{J}_{j,\theta}^T) \mathbf{K}_{j,\theta} + \text{Tr}(\mathbf{F}_j^H \mathbf{W}_j \mathbf{F}_j \mathbf{J}_{j,\theta}^T) \mathbf{K}_{j,\theta}, \end{aligned} \quad (27)$$

$$\begin{aligned} \mathbf{T} = & \hat{\mathbf{H}}_{\theta,k} \tilde{\mathbf{U}}_k \hat{\mathbf{H}}_{\theta,k}^H + \text{Tr}(\tilde{\mathbf{U}}_k \mathbf{J}_{\theta,k}^T) \mathbf{K}_{\theta,k} + \hat{\mathbf{H}}_{\theta,k} \tilde{\mathbf{U}}_k \hat{\mathbf{H}}_{\theta,k}^H \\ & + \text{Tr}(\tilde{\mathbf{U}}_k \mathbf{J}_{\theta,k}^T) \mathbf{K}_{\theta,k} + \hat{\mathbf{H}}_{\theta,0} \tilde{\mathbf{V}}_j \hat{\mathbf{H}}_{\theta,0}^H + \text{Tr}(\tilde{\mathbf{V}}_j \mathbf{J}_{\theta,0}^T) \mathbf{K}_{\theta,0} \\ & + \hat{\mathbf{H}}_{\theta,0} \tilde{\mathbf{V}}_j \hat{\mathbf{H}}_{\theta,0}^H + \text{Tr}(\tilde{\mathbf{V}}_j \mathbf{J}_{\theta,0}^T) \mathbf{K}_{\theta,0} + \hat{\mathbf{H}}_{\theta,0} \tilde{\mathbf{V}}_j \hat{\mathbf{H}}_{\theta,0}^H \\ & + \text{Tr}(\tilde{\mathbf{V}}_j \mathbf{J}_{\theta,0}^T) \mathbf{K}_{\theta,0} + \hat{\mathbf{H}}_{\theta,k} \tilde{\mathbf{U}}_k \hat{\mathbf{H}}_{\theta,k}^H + \text{Tr}(\tilde{\mathbf{U}}_k \mathbf{J}_{\theta,k}^T) \mathbf{K}_{\theta,k} \\ & + \hat{\mathbf{H}}_{\theta,k} \tilde{\mathbf{U}}_k \hat{\mathbf{H}}_{\theta,k}^H + \text{Tr}(\tilde{\mathbf{U}}_k \mathbf{J}_{\theta,k}^T) \mathbf{K}_{\theta,k}. \end{aligned} \quad (28)$$

ACKNOWLEDGEMENT

This work is supported by the Luxembourg National Fund (FNR)-RISOTTI—the Reconfigurable Intelligent Surfaces for Smart Cities under Project FNR/C20/IS/14773976/RISOTTI.

REFERENCES

- [1] C. K. Sheemar, C. K. Thomas, and D. Slock, "Practical hybrid beamforming for millimeter wave massive MIMO full duplex with limited dynamic range," *IEEE OJ-COMS*, vol. 3, pp. 127–143, 2022.
- [2] C. K. Sheemar and D. Slock, "Beamforming for bidirectional MIMO full duplex under the joint sum power and per antenna power constraints," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 4800–4804.
- [3] P. K. Sharma and P. Garg, "Intelligent reflecting surfaces to achieve the full-duplex wireless communication," *IEEE Communications Letters*, vol. 25, no. 2, pp. 622–626, 2020.
- [4] C. K. Sheemar and D. Slock, "Hybrid beamforming and combining for millimeter wave full duplex massive MIMO interference channel," in *2021 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2021, pp. 1–6.
- [5] C. K. Sheemar, "Hybrid beamforming techniques for massive mimo full duplex radio systems," Ph.D. dissertation, Ph. D. dissertation, EURECOM, 2022.
- [6] C. K. Sheemar, G. C. Alexandropoulos, D. Slock, J. Querol, and S. Chatzinotas, "Full-duplex-enabled joint communications and sensing with reconfigurable intelligent surfaces," *arXiv preprint arXiv:2306.10865*, 2023.
- [7] C. K. Sheemar, S. Tomasin, D. Slock, and S. Chatzinotas, "Near-field intelligent reflecting surfaces for millimeter wave MIMO full duplex," *arXiv preprint arXiv:2211.10700*, 2022.
- [8] C. K. Sheemar and D. Slock, "Receiver design and agc optimization with self interference induced saturation," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 5595–5599.
- [9] Z. Abdullah, G. Chen, S. Lambbotharan, and J. A. Chambers, "Optimization of intelligent reflecting surface assisted full-duplex relay networks," *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 363–367, 2020.
- [10] Y. Ge and J. Fan, "Robust secure beamforming for intelligent reflecting surface assisted full-duplex MISO systems," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 253–264, 2021.
- [11] M. A. Saeidi, M. J. Emadi, H. Masoumi, M. R. Mili, D. W. K. Ng, and I. Krikidis, "Weighted sum-rate maximization for multi-IRS-assisted full-duplex systems with hardware impairments," *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 2, pp. 466–481, 2021.
- [12] Y. Rong, "Robust design for linear non-regenerative MIMO relays with imperfect channel state information," *IEEE transactions on signal processing*, vol. 59, no. 5, pp. 2455–2460, 2011.
- [13] F. Negro, I. Ghauri, and D. T. Slock, "Sum rate maximization in the noisy MIMO interfering broadcast channel with partial CSIT via the expected weighted MSE," in *IEEE ISWCS*, 2012, pp. 576–580.
- [14] A. C. Cirik, R. Wang, Y. Hua, and M. Latva-aho, "Weighted sum-rate maximization for full-duplex MIMO interference channels," *IEEE Transactions on communications*, vol. 63, no. 3, pp. 801–815, 2015.
- [15] C. Pan, H. Ren, K. Wang, W. Xu, M. El-kashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5218–5233, 2020.
- [16] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4792–4799, 2008.
- [17] C. K. Sheemar and D. Slock, "Hybrid beamforming for bidirectional massive MIMO full duplex under practical considerations," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*. IEEE, 2021, pp. 1–5.
- [18] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate optimization for intelligent reflecting surface enhanced wireless networks," *arXiv preprint arXiv:1905.07920*, 2019.