Worst-Case Robust Sum Rate Maximization for Full-Duplex Bi-Directional MIMO Systems Under Channel Knowledge Uncertainty

Omid Taghizadeh and Rudolf Mathar
Institute for Theoretical Information Technology, RWTH Aachen University, D-52074 Aachen, Germany
Email: {taghizadeh, mathar}@ti.rwth-aachen.de

***Abstract*—In this paper we address a worst-case weighted sum rate maximization problem for a full-duplex (FD) and point-to-point (P2P) system. The effects of channel-state information (CSI) error, as well as the signal distortion due to hardware impairments are jointly taken into account. Due to the intractable structure of the resulting problem, a weighted minimum mean squared error (WMMSE) method is applied to cast the rate maximization into a separately convex optimization problem, which can be iteratively solved with a guaranteed convergence. The provided rate maximization framework is also shown to provide a converging minimum mean squared error (MMSE) design as a special case. Moreover, a methodology to obtain the least favorable error matrices is proposed by casting the resulting non-convex quadratic optimization into a convex problem. The achievable guaranteed (worst-case) rate is then numerically studied, over different levels of CSI error intensity, transceiver accuracy, and available transmit power.**

I. Introduction

Full-Duplex transceivers are known for their capability to transmit and receive at the same time and frequency, and hence have the potential to enhance the spectral efficiency [1]. Nevertheless, such systems suffer from the inherent self-interference from their own transmitter. Recently, specialized cancellation techniques, e.g., [2]–[4], have demonstrated an adequate level of isolation between Tx and Rx directions to facilitate a FD communication and motivated a wide range of related studies, see, e.g., [1], [5], [6]. A common idea of such techniques is to subtract the dominant part of the self-interference signal, e.g., the direct interference path or near-end reflections, in the RF analog domain so that the remaining signal can be processed for further interference reduction in the baseband, i.e., digital domain. Nevertheless, such methods are still far from perfect in a realistic environment due to i) aging and inherent inaccuracy of the hardware (analog) components, as well as ii) inaccurate knowledge of the self-interference channel. In particular, a FD link is vulnerable to the CSI inaccuracy at the self-interference path in environments with a small channel coherence time, see [4, Subsection 3.4.1], [7, Subsection V.C]. A good example of such challenge is a high-speed vehicle that passes close to a FD device, and results in additional reflective self-interference paths[1]. This may render the self-interference cancellation module out of tune, and result in degradation of link quality as a result of the increased residual error.

In order to combat this effect, the transceiver may adapt its transmit/receive strategy to the expected nature of CSI inaccuracy, e.g., by directing the transmit beams away from the moving objects or operating in the directions with smaller impact of CSI error. Moreover, the accuracy of the transmit/receiver chain elements can be considered, e.g., by dedicating less task to the chains with noisier elements. In this regard, a widely used model for the operation of a multiple-antenna FD transceiver is proposed in [8], where CSI inaccuracy as well as the impact of hardware impairments are taken into account. A gradient-projection-based method is then proposed in the same work for maximizing the sum rate in a FD P2P setup. Afterwards, a convex optimization design framework is introduced in [9]–[11] by defining a price/threshold for the self-interference power, assuming the availability of perfect CSI and accurate transceiver operation. While this approach reduces the design computational complexity, it does not provide a reliable performance for a scenario with erroneous CSI, particularly regarding the self-interference path [12]. Consequently, the consideration of residual self-interference in a FD P2P system is further studied in [13]–[15] by maximizing the system sum rate, in [16] by minimizing the sum mean-squared-error (MSE), and in [17] for minimizing the total power consumption under given rate constraints.

The aforementioned works focus on optimizing the average system performance under CSI inaccuracy, where the first and second order statistics of CSI error is known. Nevertheless, they do not provide a guarantee for the instantaneous performance, i.e., they focus on the *average*, and not the *worst-case* performance optimization under CSI uncertainty. The consideration of the worst-case performance becomes significant for the scenarios where seamless and reliable connectivity is of high priority, e.g., remote medical and sensitive automotive control applications. In contrast, the worst-case rate maximization for a half-duplex (HD) communication setup has been studied, e.g., in [18], [19], following a deterministic representation of the CSI error region, and extended in the context of multiple antenna FD systems from the aspects of energy efficiency [20] and sum rate maximization in a FD cellular system [21] with perfect hardware assumptions[2]. Nevertheless, such studies have not been yet extended for sum rate maximization in a FD bi-directional setup.

In this paper we address a worst-case weighted sum rate

[1] Since the object is moving rapidly, the reflective paths can not be accurately estimated.

[2] In this model the channel matrices are not known, but located with a sufficiently high probability within a known feasible region, see [22]–[24] and references therein.

maximization problem for a FD and P2P system. In particular, the effects of CSI error, as well as the signal distortion due to hardware inaccuracies are jointly taken into account. In Section II the system model, and the intended optimization strategy is presented. Due to the intractable mathematical structure of the resulting problem, a WMMSE-based optimization method [25] is proposed in Section III, which results in an iterative optimization with a guaranteed convergence. The provided rate maximization framework is also shown to provide a converging MMSE-based design as a special case. Furthermore, a methodology to obtain the least favorable error matrices are then proposed by casting the resulting non-convex quadratic optimization into a convex optimization problem. The achievable (worst case) rate is then numerically evaluated in Section IV for different levels of CSI error, transceiver inaccuracy, and the available transmit power.

A. Mathematical Notation:

Throughout this paper, column vectors and matrices are denoted as lower-case and upper-case bold letters, respectively. Mathematical expectation, trace, inverse, determinant, transpose, conjugate and Hermitian transpose are denoted by $\mathbb{E}\{\cdot\}$, $\text{tr}(\cdot)$, $(\cdot)^{-1}$ $|\cdot|$, $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$, respectively. The Kronecker product is denoted by $\otimes$. The identity matrix with dimension K is denoted as $\mathbf{I}_K$ and $\text{vec}(\cdot)$ operator stacks the elements of a matrix into a vector. $\mathbf{0}_{m\times n}$ represents an all-zero matrix with size $m \times n$. $\|\cdot\|_2$ and $\|\cdot\|_F$ respectively represent the Euclidean and Frobenius norms. $\text{diag}(\cdot)$ returns a diagonal matrix by putting the off-diagonal elements to zero. $\lfloor \mathbf{A}_i \rfloor_{i=1,\ldots,K}$ denotes a tall matrix, obtained by stacking the matrices $\mathbf{A}_i$, $i = 1,\ldots,K$. $\mathcal{R}\{\mathbf{A}\}$ represents the range (column space) of the matrix $\mathbf{A}$.

II. System Model

We consider a MIMO and bi-directional communication between two FD transceivers, where both communication directions are accommodated on the same channel. Each communication direction is realized with N_i transmit and M_i receive antennas, where $i \in \mathcal{D}$, and $\mathcal{D} := \{1, 2\}$ represents the set of the communication directions, see Fig. 1. The desired channel in the communication direction i is denoted as $\mathbf{H}_{ii} \in \mathbb{C}^{M_i \times N_i}$ where the self-interference channel from i to j-th communication direction is denoted as $\mathbf{H}_{ji} \in \mathbb{C}^{M_j \times N_i}$. All channels are assumed to follow a flat-fading model, where the CSI is known erroneously. In this respect we follow the so-called deterministic model [22], where the error matrices are not known, but located within a known feasible error region. This is written as

$$\mathbf{H}_{ij} = \tilde{\mathbf{H}}_{ij} + \mathbf{\Delta}_{ij}, \;\; \|\mathbf{D}_{ij}\mathbf{\Delta}_{ij}\|_F \leq \zeta_{ij}, \;\; i,j \in \mathcal{D}, \quad (1)$$

where $\tilde{\mathbf{H}}_{ij}$ is the estimated channel matrix, $\mathbf{\Delta}_{ij}$ represents the channel estimation error, and $\mathbf{D}_{ij} \succeq 0$ and $\zeta_{ij} \geq 0$ jointly define a feasible ellipsoid region for $\mathbf{\Delta}_{ij}$. For further elaboration on the used error model please see [22]–[24]. The transmitted signal in the direction i is formulated as

$$\mathbf{x}_i = \underbrace{\mathbf{V}_i\mathbf{s}_i}_{=:\mathbf{v}_i} + \mathbf{e}_{\text{t},i}, \;\; \mathbb{E}\left\{\|\mathbf{x}_i\|_2^2\right\} \leq P_i, \quad (2)$$

where $\mathbf{s}_i \in \mathbb{C}^{d_i}, \mathbf{V}_i \in \mathbb{C}^{N_i \times d_i}, \mathbf{v}_i \in \mathbb{C}^{N_i}$ and $P_i \in \mathbb{R}$ respectively represent the vector of the data symbols, the transmit precoding matrix, the intended (undistorted) transmit signal, and the maximum affordable transmit power. The number of the transmited data streams in direction i is denoted as d_i, and $\mathbb{E}\{\mathbf{s}_i\mathbf{s}_i^H\} = \mathbf{I}_{d_i}$. Moreover, the inaccurate behavior of the transmit chain elements is modeled as an additional distortion term $\mathbf{e}_{\text{t},i}$ such that

$$\mathbf{e}_{\text{t},i} \sim \mathcal{CN}\Big(\mathbf{0}_{N_i\times 1}, \kappa_i \text{diag}\left(\mathbb{E}\{\mathbf{v}_i\mathbf{v}_i^H\}\right)\Big), \;\; \mathbf{e}_{\text{t},i} \perp \mathbf{v}_i, \quad (3)$$

where $\perp$ represents the statistical independence, and κ_i is the transmit distortion coefficient, see [8, Section II.C] for more elaboration. The received signal at the destination can be consequently written as

$$\mathbf{y}_i = \underbrace{\mathbf{H}_{ii}\mathbf{x}_i + \mathbf{H}_{ij}\mathbf{x}_j + \mathbf{n}_i}_{=:\mathbf{u}_i} + \mathbf{e}_{\text{r},i}, \quad (4)$$

where $\mathbf{n}_i$ is the additive thermal noise with variance $\sigma_{\text{n},i}^2$, $\mathbf{u}_i \in \mathbb{C}^{M_i}$ is the undistorted received signal, and the additive distortion term $\mathbf{e}_{\text{r},i}$ models the inaccuracies in the receive chains such that

$$\mathbf{e}_{\text{r},i} \sim \mathcal{CN}\Big(\mathbf{0}_{M_i\times 1}, \beta_i \text{diag}\left(\mathbb{E}\{\mathbf{u}_i\mathbf{u}_i^H\}\right)\Big), \;\; \mathbf{e}_{\text{r},i} \perp \mathbf{u}_i, \quad (5)$$

where β_i is the receiver chain distortion coefficient, see [8, Section II.D]. Please note that the distortion terms $\mathbf{e}_{\text{r},i}$ and $\mathbf{e}_{\text{t},i}$ model the combined effects of the transmit and receive chain inaccuracies, e.g., digital-to-analog and analog-to-digital converter error, power amplifier noise, oscillator phase noise and the automatic gain control noise at the respective chains. Hence, unlike the thermal noise components, the variance of the distortion terms are dependent on the power of the intended transmit/receive signal at each antenna and play an important role in a FD setup due to the strong self-interference path, see [8], [26] and the references therein. The *known* part of the self-interference signal, i.e, $\tilde{\mathbf{H}}_{ij}\mathbf{v}_j$, can be reduced at the receiver side using the self-interference cancellation techniques, e.g., [2], [3]. This is written as

$$\tilde{\mathbf{y}}_i = \mathbf{H}_i\mathbf{V}_i\mathbf{s}_i + \mathbf{w}_i, \quad (6)$$

where $\tilde{\mathbf{y}}_i$ is the received signal, after the self-interference cancellation, and $\mathbf{w}_i \in \mathbb{C}^{M_i}$ is the aggregate residual interference-plus-noise signal for the communication direction i

$$\mathbf{w}_i = \mathbf{H}_{ii}\mathbf{e}_{\text{t},i} + \mathbf{H}_{ij}\mathbf{e}_{\text{t},j} + \mathbf{e}_{\text{r},i} + \mathbf{\Delta}_{ij}\mathbf{V}_j\mathbf{s}_j + \mathbf{n}_i. \quad (7)$$

Finally, the estimated data vector is obtained at the receiver as

$$\tilde{\mathbf{s}}_i = \mathbf{U}_i{}^H\tilde{\mathbf{y}}_i, \quad (8)$$

where $\mathbf{U}_i \in \mathbb{C}^{M_i \times d_i}$ is the linear receive filter. The achievable communication rate in the direction i, i.e., the mutual information between $\mathbf{s}_i$ and $\tilde{\mathbf{s}}_i$, is hence written as

$$I_i = B\text{log}_2\left|\mathbf{I}_{M_i} + \mathbf{\Sigma}_i^{-1}\mathbf{H}_{ii}\mathbf{V}_i\mathbf{V}_i^H\mathbf{H}_{ii}^H\right|, \quad (9)$$

where B is the used bandwidth, I_i is the achievable rate, and $\mathbf{\Sigma}_i$ represents the combined residual interference-plus-noise covariance for the i-th communication direction. Please note that the above formulation holds for the Gaussian distribution of the desired, and residual interference signal terms, and otherwise can be viewed as an approximation.

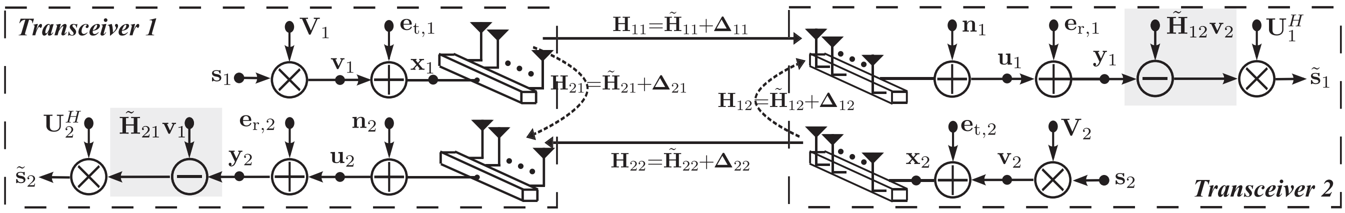

Figure 1. A full-duplex bi-directional system with multiple antennas. The communication quality suffers due to the additive white noise, i.e., $\mathbf{n}_i$, inaccuracies of transmit and receive chains, i.e., $\mathbf{e}_{\text{t},i}$ and $\mathbf{e}_{\text{r},i}$, as well as the CSI estimation error $\mathbf{\Delta}_{ij}$, $i,j \in \{1,2\}$.

A. Optimization Problem

Our goal is to maximize the weighted sum rate of the system, corresponding to the worst-case feasible channel error matrix, see (1). This is written as

$$\max_{\bar{\mathbf{V}},\bar{\mathbf{U}}} \ \min_{\bar{\mathbf{\Delta}}} \ \sum_{i\in\mathcal{D}} \nu_i I_i \tag{10a}$$

$$\text{s.t.} \quad \mathbb{E}\left\{\|\mathbf{x}_i\|_2^2\right\} \le P_i, \ \ \forall i \in \mathcal{D}, \tag{10b}$$

$$\|\mathbf{D}_{ij}\mathbf{\Delta}_{ij}\|_F \le \zeta_{ij}, \ \ \forall i,j \in \mathcal{D}, \tag{10c}$$

where $\bar{\mathbf{X}} := \{\mathbf{X}_i, \ \forall i \in \mathcal{D}\}$, for $\mathbf{X} \in \{\mathbf{V},\mathbf{U}\}$, $\bar{\mathbf{\Delta}} := \{\mathbf{\Delta}_{ij}, \ \forall i,j \in \mathcal{D}\}$, and $\nu_i \in \mathbb{R}$ represents the weight of the communication rate.

III. WMMSE Method for Worst-Case Rate Maximization

With the application of $\mathbf{V}_i, \mathbf{U}_i$ as the transmit and receive linear filters, the MSE matrix for the communication direction i is written as

$$\begin{aligned}\mathbf{E}_i :&= \mathbb{E}\{(\tilde{\mathbf{s}}_i - \mathbf{s}_i)(\tilde{\mathbf{s}}_i - \mathbf{s}_i)^H\} \\ &= \left(\mathbf{U}_i{}^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}\right)\left(\mathbf{U}_i{}^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}\right)^H \\ &\quad + \mathbf{U}_i{}^H\mathbf{\Sigma}_i\mathbf{U}_i,\end{aligned} \tag{11}$$

where $\mathbf{\Sigma}_i$ is calculated as

$$\begin{aligned}\mathbf{\Sigma}_i :=& \mathbb{E}\left\{\mathbf{w}_i\mathbf{w}_i^H\right\} \\ =& \kappa_j\mathbf{H}_{ij}\text{diag}\left(\mathbf{V}_j\mathbf{V}_j^H\right)\mathbf{H}_{ij}^H + \beta_i\text{diag}\left(\mathbf{H}_{ij}\mathbf{V}_j\mathbf{V}_j^H\mathbf{H}_{ij}^H\right) \\ &+ \kappa_i\mathbf{H}_{ii}\text{diag}\left(\mathbf{V}_i\mathbf{V}_i^H\right)\mathbf{H}_{ii}^H + \beta_i\text{diag}\left(\mathbf{H}_{ii}\mathbf{V}_i\mathbf{V}_i^H\mathbf{H}_{ii}^H\right) \\ &+ \Delta_{ij}\mathbf{V}_j\mathbf{V}_j^H\Delta_{ij}^H + \sigma_{\text{n},i}^2\mathbf{I}_{M_i}.\end{aligned} \tag{12}$$

Note that the above expression holds, as $\kappa_i \ll 1$, $\beta_i \ll 1$, and the terms including higher orders of κ_i and β_i can be safely ignored. The MMSE receive linear filter can be calculated as

$$\mathbf{U}_i^{\text{mmse}} = \left(\mathbf{\Sigma}_i + \mathbf{H}_{ii}\mathbf{V}_i\mathbf{V}_i^H\mathbf{H}_{ii}^H\right)^{-1}\mathbf{H}_{ii}\mathbf{V}_i, \tag{13}$$

and the resulting MSE matrix is obtained as

$$\mathbf{E}_i^{\text{mmse}} = \left(\mathbf{I}_{d_i} + \mathbf{V}_i^H\mathbf{H}_{ii}\mathbf{\Sigma}_i^{-1}\mathbf{H}_{ii}\mathbf{V}_i\right)^{-1}. \tag{14}$$

By recalling (9), we observe the following useful connection to the rate function

$$I_i = -B\log_2\left|\mathbf{E}_i^{\text{mmse}}\right|, \tag{15}$$

which facilitates the decomposition of rate function via the following lemma, see also [27, Eq. (9)].

Lemma 1. *Let $\mathbf{E} \in \mathbb{C}^{d\times d}$ be a positive semi-definite matrix. The maximization of the term $-log|\mathbf{E}|$ is equivalent to the maximization*

$$\max_{\mathbf{E},\mathbf{W}} - tr\left(\mathbf{W}^H\mathbf{E}\mathbf{W}\right) + log\left|\mathbf{W}\mathbf{W}^H\right| + d, \tag{16}$$

where $\mathbf{W} \in \mathbb{C}^{d\times d}$ is a positive semi-definite matrix.

Proof: The proof is obtained following [18, Lemma 2], and decomposing the positive semi-definite matrix $\mathbf{S} \in \mathbb{C}^{d\times d}$ as $\mathbf{S} = \mathbf{W}\mathbf{W}^H$. ■

By recalling (11) and (9), and utilizing Lemma 1, the original optimization problem can be equivalently formulated as

$$\max_{\bar{\mathbf{V}}} \ \min_{\bar{\mathbf{\Delta}}} \ \max_{\bar{\mathbf{U}},\bar{\mathbf{W}}} \ \sum_{i\in\mathcal{D}} \nu_i\left(\log|\mathbf{W}_i\mathbf{W}_i^H| + d_i - \text{tr}\left(\mathbf{W}_i^H\mathbf{E}_i\mathbf{W}_i\right)\right) \tag{17a}$$

$$\text{s.t.} \ \ (10b), (10c) \tag{17b}$$

where $\bar{\mathbf{W}}$ is a set containing $\mathbf{W}_i \succeq 0$, $\forall i \in \mathcal{D}$. In order to cast the objective into a simpler form we calculate

$$\begin{aligned}\text{tr}\left(\mathbf{W}_i^H\mathbf{E}_i\mathbf{W}_i\right) =& \left\|\mathbf{W}_i^H\left(\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}\right)\right\|_F^2 \\ &+ \sum_{l\in\mathbb{F}_{N_j}} \kappa_j\left\|\mathbf{W}_i^H\mathbf{U}_i^H\mathbf{H}_{ij}\mathbf{\Gamma}_l\mathbf{V}_j\right\|_F^2 \\ &+ \sum_{l\in\mathbb{F}_{M_i}} \beta_i\left\|\mathbf{W}_i^H\mathbf{U}_i^H\mathbf{\Gamma}_l\mathbf{H}_{ij}\mathbf{V}_j\right\|_F^2 \\ &+ \sum_{l\in\mathbb{F}_{N_i}} \kappa_i\left\|\mathbf{W}_i^H\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{\Gamma}_l\mathbf{V}_i\right\|_F^2 \\ &+ \sum_{l\in\mathbb{F}_{M_i}} \beta_i\left\|\mathbf{W}_i^H\mathbf{U}_i^H\mathbf{\Gamma}_l\mathbf{H}_{ii}\mathbf{V}_i\right\|_F^2 \\ &+ \left\|\mathbf{W}_i^H\mathbf{U}_i^H\mathbf{\Delta}_{ij}\mathbf{V}_j\right\|_F^2 + \sigma_{\text{n},i}^2\left\|\mathbf{W}_i^H\mathbf{U}_i^H\right\|_F^2\end{aligned} \tag{18}$$

$$= \boxed{\sum_{j\in\mathcal{D}} \left\|\mathbf{c}_{ij} + \mathbf{C}_{ij}\text{vec}\left(\mathbf{\Delta}_{ij}\right)\right\|_2^2}, \tag{19}$$

where

$$\mathbf{c}_{ii} := \begin{bmatrix} \text{vec}\left(\mathbf{W}_i^H\left(\mathbf{U}_i^H\tilde{\mathbf{H}}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}\right)\right) \\ \left\lfloor\sqrt{\kappa_i}\text{vec}\left(\mathbf{W}_i^H\mathbf{U}_i^H\tilde{\mathbf{H}}_{ii}\mathbf{\Gamma}_{N_i}^l\mathbf{V}_i\right)\right\rfloor_{l\in\mathbb{F}_{N_i}} \\ \left\lfloor\sqrt{\beta_i}\text{vec}\left(\mathbf{W}_i^H\mathbf{U}_i^H\mathbf{\Gamma}_{M_i}^l\tilde{\mathbf{H}}_{ii}\mathbf{V}_i\right)\right\rfloor_{l\in\mathbb{F}_{M_i}} \\ \sigma_{\text{n},i}\text{vec}\left(\mathbf{W}_i^H\mathbf{U}_i^H\right) \end{bmatrix}, \tag{20}$$

$$\mathbf{C}_{ii} := \begin{bmatrix} \mathbf{V}_i^T \otimes \left(\mathbf{W}_i^H \mathbf{U}_i^H\right) \\ \left\lfloor \sqrt{\kappa_i} \left(\mathbf{\Gamma}_{N_i}^l \mathbf{V}_i\right)^T \otimes \left(\mathbf{W}_i^H \mathbf{U}_i^H\right)\right\rfloor_{l \in \mathbb{F}_{N_i}} \\ \left\lfloor \sqrt{\beta_i} \mathbf{V}_i^T \otimes \left(\mathbf{W}_i^H \mathbf{U}_i^H \mathbf{\Gamma}_{M_i}^l\right)\right\rfloor_{l \in \mathbb{F}_{M_i}} \\ \mathbf{0}_{M_i d_i \times N_j M_i} \end{bmatrix}, \quad (21)$$

$$\mathbf{c}_{ij} \overset{i \neq j}{:=} \begin{bmatrix} \left\lfloor \sqrt{\kappa_j} \mathrm{vec}\left(\mathbf{W}_i^H \mathbf{U}_i^H \tilde{\mathbf{H}}_{ij} \mathbf{\Gamma}_{N_j}^l \mathbf{V}_j\right)\right\rfloor_{l \in \mathbb{F}_{N_j}} \\ \left\lfloor \sqrt{\beta_i} \mathrm{vec}\left(\mathbf{W}_i^H \mathbf{U}_i^H \mathbf{\Gamma}_{M_i}^l \tilde{\mathbf{H}}_{ij} \mathbf{V}_j\right)\right\rfloor_{l \in \mathbb{F}_{M_i}} \\ \mathbf{0}_{d_i d_j \times 1} \end{bmatrix}, \quad (22)$$

$$\mathbf{C}_{ij} \overset{i \neq j}{:=} \begin{bmatrix} \left\lfloor \sqrt{\kappa_j} \left(\mathbf{\Gamma}_{N_j}^l \mathbf{V}_j\right)^T \otimes \left(\mathbf{W}_i^H \mathbf{U}_i^H\right)\right\rfloor_{l \in \mathbb{F}_{N_j}} \\ \left\lfloor \sqrt{\beta_i} \mathbf{V}_j^T \otimes \left(\mathbf{W}_i^H \mathbf{U}_i^H \mathbf{\Gamma}_{M_i}^l\right)\right\rfloor_{l \in \mathbb{F}_{M_i}} \\ \mathbf{V}_j^T \otimes \left(\mathbf{W}_i^H \mathbf{U}_i^H\right) \end{bmatrix}, \quad (23)$$

such that $\mathbf{\Gamma}_K^l$ represents a square matrix with dimension K, with all zero elements except of the l-th diagonal element equal to 1, and $\mathbf{c}_{ij} \in \mathbb{C}^{\tilde{d}_{ij}}, \mathbf{C}_{ij} \in \mathbb{C}^{\tilde{d}_{ij} \times M_i N_j}$ where

$$\tilde{d}_{ii} = (1 + N_i + M_i)\, d_i^2 + M_i d_i,$$
$$\tilde{d}_{ij} = (1 + N_j + M_i)\, d_i d_j, \ \ \forall i, j \in \mathcal{D}, \ \ i \neq j.$$

In the above derivations, (18) is calculated by recalling (11) and (12) and the known matrix equality [28, Eq. (516)], and (20)-(23) are calculated via the application of [28, Eq. (496), (497)]. Unfortunately, the problem (17) is still intractable due to the joint inner and outer maximization. Hence, we focus on maximizing the lower bound of the objective following the max-min inequality, see [18, Eq. (12)]. This is written using the epigraph form as

$$\max_{\bar{\mathbf{V}}, \bar{\mathbf{U}}, \bar{\mathbf{W}}, \bar{\tau}} \ \min_{\bar{\mathbf{\Delta}}} \sum_{i \in \mathcal{D}} \nu_i \left(\log|\mathbf{W}_i \mathbf{W}_i^H| + d_i - \sum_{j \in \mathcal{D}} \tau_{ij} \right) \quad (24a)$$
$$\text{s.t.} \ \|\mathbf{c}_{ij} + \mathbf{C}_{ij} \mathrm{vec}\left(\mathbf{\Delta}_{ij}\right)\|_2^2 \leq \tau_{ij}, \ \forall i, j \in \mathcal{D}, \quad (24b)$$
$$(10c), (10b), \quad (24c)$$

where $\bar{\tau} := \{\tau_{ij}, \ \forall i, j \in \mathcal{D}\}$. By defining $\tilde{\mathbf{\Delta}}_{ij} := \mathbf{D}_{ij} \mathbf{\Delta}_{ij}$, $\tilde{\mathbf{D}}_{ij} := \mathbf{I}_{N_j} \otimes \mathbf{D}_{ij}^{-1}$, $\mathbf{b}_{ij} := \mathrm{vec}(\tilde{\mathbf{\Delta}}_{ij})$, and applying the Schur complement lemma, the constraint (24b) is equivalently written as

$$\begin{bmatrix} 0 & \mathbf{b}_{ij}^H \tilde{\mathbf{D}}_{ij}^H \mathbf{C}_{ij}^H \\ \mathbf{C}_{ij} \tilde{\mathbf{D}}_{ij} \mathbf{b}_{ij} & \mathbf{0}_{\tilde{d}_{ij} \times \tilde{d}_{ij}} \end{bmatrix} + \begin{bmatrix} \tau_{ij} & \mathbf{c}_{ij}^H \\ \mathbf{c}_{ij} & \mathbf{I}_{\tilde{d}_{ij}} \end{bmatrix} \succeq 0, \quad (25)$$

where $\tilde{d}_{ij}$ is equal to the size of $\mathbf{c}_{ij}$. Similarly, the constraint (10c) can be written as $\|\mathbf{b}_{ij}\|_2 \leq \zeta_{ij}$.

Lemma 2. *Generalized Petersen's sign-definiteness lemma: Let $\mathbf{Y} = \mathbf{Y}^H$, and $\mathbf{X}, \mathbf{P}, \mathbf{Q}$ are arbitrary matrices with complex valued elements. Then we have*

$$\mathbf{Y} \succeq \mathbf{P}^H \mathbf{X} \mathbf{Q} + \mathbf{Q}^H \mathbf{X}^H \mathbf{P}, \ \ \forall \mathbf{X} \ : \ \|\mathbf{X}\|_F \leq \zeta, \quad (26)$$

if and only if

$$\exists \lambda \geq 0, \ \begin{bmatrix} \mathbf{Y} - \lambda \mathbf{Q}^H \mathbf{Q} & -\zeta \mathbf{P}^H \\ -\zeta \mathbf{P} & \lambda \mathbf{I} \end{bmatrix} \succeq 0. \quad (27)$$

Proof: See [29, Proposition 2], [30]. ■

By choosing the matrices in Lemma 2 such that $\mathbf{X} = \mathbf{b}_{ij}$, $\mathbf{Q} = \left[-1, \ \mathbf{0}_{1 \times \tilde{d}_{ij}},\right]$ and

$$\mathbf{Y} = \begin{bmatrix} \tau_{ij} & \mathbf{c}_{ij}^H \\ \mathbf{c}_{ij} & \mathbf{I}_{\tilde{d}_{ij}} \end{bmatrix}, \mathbf{P} = \left[\ \mathbf{0}_{M_i N_j \times 1}, \ \ \tilde{\mathbf{D}}_{ij}^H \mathbf{C}_{ij}^H \ \right], \quad (28)$$

the optimization problem in (24) is equivalently written as

$$\max_{\bar{\mathbf{V}}, \bar{\mathbf{U}}, \bar{\mathbf{W}}, \bar{\tau}, \bar{\lambda}} \sum_{i \in \mathcal{D}} \nu_i \left(2\log|\mathbf{W}_i| + d_i - \sum_{j \in \mathcal{D}} \tau_{ij} \right) \quad (29a)$$

$$\text{s.t.} \ \begin{bmatrix} \tau_{ij} - \lambda_{ij} & \mathbf{c}_{ij}^H & \mathbf{0}_{1 \times M_i N_j} \\ \mathbf{c}_{ij} & \mathbf{I}_{\tilde{d}_{ij}} & -\zeta_{ij} \mathbf{C}_{ij} \tilde{\mathbf{D}}_{ij} \\ \mathbf{0}_{M_i N_j \times 1} & -\zeta_{ij} \tilde{\mathbf{D}}_{ij}^H \mathbf{C}_{ij}^H & \lambda_{ij} \mathbf{I}_{M_i N_j} \end{bmatrix} \succeq 0, \quad (29b)$$

$$\begin{bmatrix} P_i & \mathbf{m}^H \\ \mathbf{m} & \mathbf{I}_{\sum_{i \in \mathcal{D}} N_i d_i} \end{bmatrix} \succeq 0, \ \forall i, j \in \mathcal{D}, \quad (29c)$$

where $\bar{\lambda} := \{\lambda_{ij}, \ \forall i, j \in \mathcal{D}\}$, $\mathbf{m} := \lfloor \mathrm{vec}\left(\mathbf{V}_i\right) \rfloor_{i \in \mathcal{D}}$, and the constraint (29c) is the schur complement representation of the power constraint (10b).

Note that (29) is not a jointly convex optimization problem. Nevertheless, it is separately convex over the optimization variables, see (20)-(23). This facilitates an iterative optimization over separated variable sets, where in each iteration a convex sub-problem is solved, see Algorithm 1. In this regard, the maximization over $\bar{\mathbf{V}}, \bar{\mathbf{U}}$ can be separately cast as a general SDP, where the optimization over $\bar{\mathbf{W}}$ can be efficiently implemented using MAX-DET algorithm [31]. Moreover, due to the monotonic increase of the objective in each optimization iteration the algorithm convergences to a stationary point, see also [18, Section III] for arguments regarding convergence and optimization steps for a problem with similar variable separation.

A. Worst-Case MMSE Optimization as a Special Case

The proposed optimization framework provides a system sum rate maximization design, where the rate functions are re-structured using the WMMSE method. Alternatively, an optimization problem for obtaining the optimal MMSE system operation can be written as

$$\min_{\bar{\mathbf{V}}, \bar{\mathbf{U}}} \ \max_{\bar{\mathbf{\Delta}}} \ \sum_{i \in \mathcal{D}} \mu_i \mathrm{tr}\left(\mathbf{E}_i\right), \ \ \text{s.t.} \ (10b), (10c), \quad (30)$$

where the maximization over $\bar{\mathbf{\Delta}}$ represents the worst-case channel estimation error in the MMSE sense, and $\mu_i \in \mathbb{R}$ represents the price of the estimation MSE at the communication direction i. As it can be observed from (24), the defined MMSE optimization (30) can be interpreted as the special case of (24) by choosing $\nu_i = \mu_i$, and setting $\mathbf{W}_i = \mathbf{I}_{d_i}$ in the objective and in the definitions of $\mathbf{c}_{ij}$ and $\mathbf{C}_{ij}, \forall i, j \in \mathcal{D}$. As a result, a similar iterative optimization can be employed where at each step a convex sub-problem is solved over $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$. Note that unlike the rate maximization case, where the max-min inequality is employed to construct a lower bound of the rate function, the provided framework acts on the exact value of the MMSE value as the objective. Moreover, the utilization of the MAX-DET algorithm is not necessary and all sub-problems are presented as an SDP, due to the elimination of $\bar{\mathbf{W}}$ from the optimization variables set.

B. Worst-Case Channel Error Matrices

The obtained design from (29) is intended to maximize the worst-case weighted system sum rate, over the feasible error region. On the other hand it is beneficial to obtain the least favorable channel error matrices, as it provides guidelines for the future channel estimation strategy, e.g., to reduce the radius of the error feasible regions in the most destructive directions. Moreover, it is a necessary step for cutting-set based methods [32] that aim to reduce the design complexity by iteratively identifying the most destructive error matrices and incorporating them into the future design steps. In the current setup, the worst-case channel error matrices are identified by minimizing the rate function (10) within their feasible region. Due to the intractable structure of such a rate minimization problem, we focus on the modified objective in (24) where the worst case error matrices can be obtained via the inner minimization as

$$\min_{\mathbf{\Delta}} \quad -\sum_{i\in\mathcal{D}} \nu_i \mathrm{tr}\left(\mathbf{W}_i^H \mathbf{E}_i \mathbf{W}_i\right), \tag{31a}$$

$$\text{s.t.} \quad \|\mathbf{D}_{ij}\mathbf{\Delta}_{ij}\| \leq \zeta_{ij}, \quad \forall i,j \in \mathcal{D}. \tag{31b}$$

Note that similar to the arguments in Subsection III-A, the above formulation also represents the special case of the MMSE design criteria, by setting $\nu_i = \mu_i$, and $\mathbf{W}_i = \mathbf{I}_{d_i}, i \in \mathcal{D}$. Due to the uncoupled nature of the error feasible set, and the value of the objective function over $\mathbf{\Delta}_{ij},\ i,j \in \mathcal{D}$, following (19), the above minimization can be decomposed as

$$\min_{\mathbf{b}_{ij}} \quad -\mathbf{b}_{ij}^H\tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{C}_{ij}\tilde{\mathbf{D}}_{ij}\mathbf{b}_{ij} - 2\mathrm{Re}\{\mathbf{b}_{ij}^H\tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{c}_{ij}\} - \mathbf{c}_{ij}^H\mathbf{c}_{ij} \tag{32a}$$

$$\text{s.t.} \quad \mathbf{b}_{ij}^H\mathbf{b}_{ij} \leq \zeta_{ij}^2, \tag{32b}$$

where $\mathrm{Re}\{\cdot\}$ represents the real part of a complex value. Note that the objective in (32a) is a non convex function and can not be minimized using the usual numerical solvers in the current form. Following the zero duality gap results for the non-convex quadratic problems [33], [34], we focus on the dual function of (32). The corresponding Lagrangian function to (32) is hence constructed as

$$\mathcal{L}\left(\mathbf{b}_{ij},\rho_{ij}\right) = \mathbf{b}_{ij}^H\mathbf{A}_{ij}\mathbf{b}_{ij} - 2\mathrm{Re}\{\mathbf{b}_{ij}^H\tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{c}_{ij}\} - \mathbf{c}_{ij}^H\mathbf{c}_{ij} - \rho_{ij}\zeta_{ij}^2, \tag{33}$$

where ρ_{ij} is the dual variable and $\mathbf{A}_{ij} := \rho_{ij}\mathbf{I} - \tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{C}_{ij}\tilde{\mathbf{D}}_{ij}$. Consequently, the value of the dual function is obtained as

$$\mathbf{g}\left(\rho_{ij}\right) = -\mathbf{c}_{ij}^H\mathbf{C}_{ij}\tilde{\mathbf{D}}_{ij}\mathbf{A}_{ij}^{-1}\tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{c}_{ij} - \mathbf{c}_{ij}^H\mathbf{c}_{ij} - \rho_{ij}\zeta_{ij}^2, \tag{34}$$

if $\mathbf{A}_{ij} \succeq 0$, and $\tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{c}_{ij} \in \mathcal{R}\{\mathbf{A}_{ij}\}$, and otherwise is unbounded from below[3]. By applying the Schur complement lemma, the maximization of the dual function is written using the epigraph form as

$$\max_{\rho_{ij}\geq 0,\ \phi_{ij}} \quad \phi_{ij} \tag{35a}$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{A}_{ij} & -\tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{c}_{ij} \\ -\mathbf{c}_{ij}^H\mathbf{C}_{ij}\tilde{\mathbf{D}}_{ij} & -\mathbf{c}_{ij}^H\mathbf{c}_{ij} - \rho_{ij}\zeta_{ij}^2 - \phi_{ij} \end{bmatrix} \succeq 0, \tag{35b}$$

where $\phi_{ij} \in \mathbb{R}$ is an auxiliary variable[4]. By plugging the obtained dual variable ρ_{ij} into (33), and considering the fact that $-\tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{C}_{ij}\tilde{\mathbf{D}}_{ij}+\rho_{ij}^\star\mathbf{I} \succeq 0$ as a result of (35), the optimal value of $\mathbf{b}_{ij}$ is obtained from (33) as

$$\mathbf{b}_{ij}^\star = \left(-\tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{C}_{ij}\tilde{\mathbf{D}}_{ij} + \rho_{ij}^\star\mathbf{I}\right)^{-1}\tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{c}_{ij}, \tag{36}$$

where $(\cdot)^\star$ represents the optimality and the worst case $\mathbf{\Delta}_{ij}$ is consequently calculated via $\mathrm{vec}(\mathbf{\Delta}_{ij}) = \tilde{\mathbf{D}}_{ij}\mathbf{b}_{ij}^\star$.

[3]If one of the aforementioned conditions is not satisfied, an infinitely large value of $\mathbf{b}_{ij}$ can be chosen in the negative direction of $\mathbf{A}_{ij}$, if $\mathbf{A}_{ij}$ is not positive semi-definite, or in the direction $\tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{c}_{ij}$ in the complementary null-space of $\mathbf{A}_{ij}$.

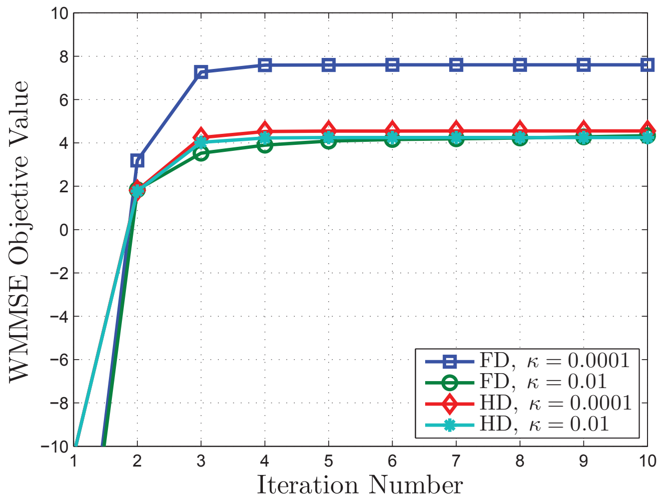

Figure 2. Average convergence behavior of the proposed iterative method. The proposed method converges to a stable point within fewer than 10 optimization iterations.

Algorithm 1 Iterative semi-definite-programming (SDP) framework for worst-case sum rate maximization in a bidirectional FD system

1: $\ell \leftarrow 0$ (set iteration number to zero)
2: $\mathbf{V}_i, \mathbf{U}_i \leftarrow \mathbf{0},\ \mathbf{W}_i \leftarrow \mathbf{I}_{d_i},\ i \in \mathcal{D}$ (initialization)
3: **repeat**
4: $\quad\ell \leftarrow \ell + 1$
5: $\quad\bar{\mathbf{V}}, \bar{\tau}, \bar{\lambda} \leftarrow$ solve SDP (29), with fixed $\bar{\mathbf{U}}, \bar{\mathbf{W}}$
6: $\quad\bar{\mathbf{U}}, \bar{\tau}, \bar{\lambda} \leftarrow$ solve SDP (29), with fixed $\bar{\mathbf{V}}, \bar{\mathbf{W}}$
7: $\quad\bar{\mathbf{W}}, \bar{\tau}, \bar{\lambda} \leftarrow$ solve MAX-DET (29), with fixed $\bar{\mathbf{U}}, \bar{\mathbf{V}}$
8: **until** a stable point, or maximum number of ℓ reached
9: **return** $\left(\bar{\mathbf{W}}, \bar{\mathbf{U}}, \bar{\mathbf{V}}\right)$

IV. Simulation Results

In this section we numerically evaluate the resulting worst-case sum rate of a FD P2P system, using the proposed design. In this respect, we assume that $\mathbf{H}_{ii}$ follows an uncorrelated Reyligh distribution, with variance 0.01 for each element and

[4]Note that the semi-definite presentation in (35b) automatically satisfies $\mathbf{A}_{ij} \succeq 0$, and $\tilde{\mathbf{D}}_{ij}^H\mathbf{C}_{ij}^H\mathbf{c}_{ij} \in \mathcal{R}\{\mathbf{A}_{ij}\}$.

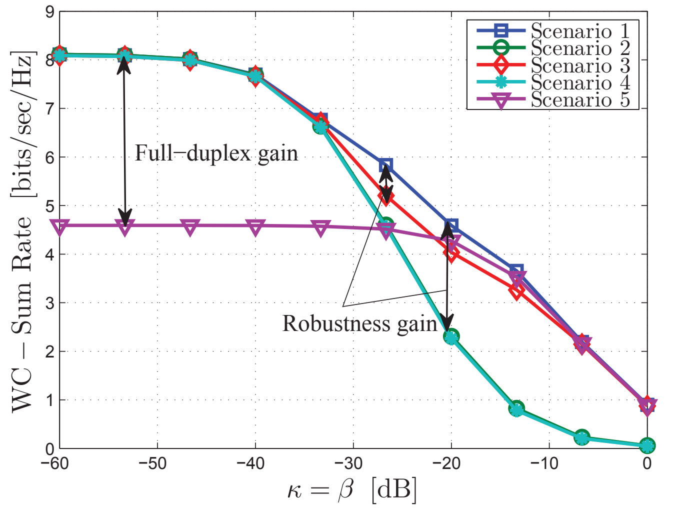

Figure 3. Guaranteed (worst-case) sum rate vs. transceiver inaccuracy ($\kappa = \kappa_i = \beta_j, \; \forall i, j$). Achievable worst case rate as well as the FD system advantage suffers as the transceiver inaccuracy increases.

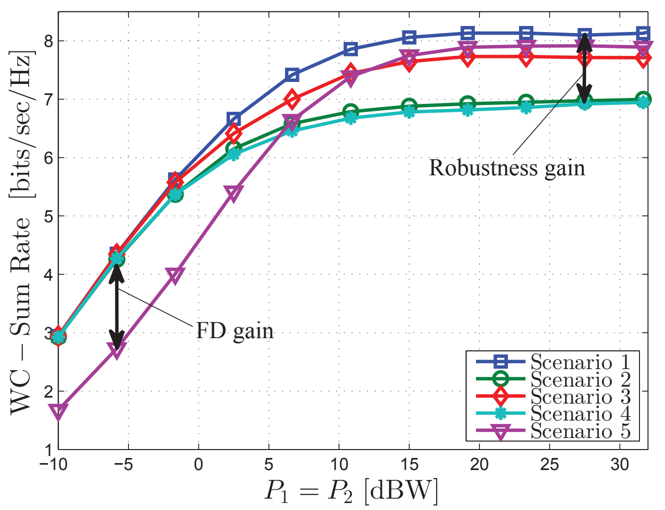

Figure 5. Guaranteed (worst-case) sum rate vs. maximum transmit power ($P = P_i, \; \forall i$). Achievable rate is dominated by the effects of CSI error, and transceiver inaccuracy for high power regime.

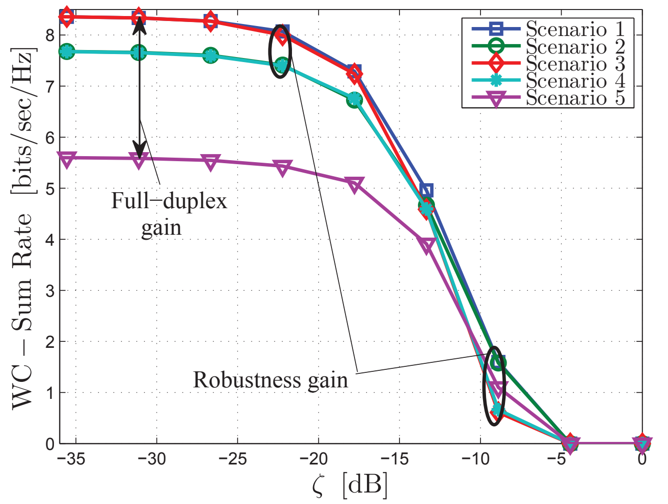

Figure 4. Guaranteed (worst-case) sum rate vs. feasible CSI error radius ($\zeta = \zeta_{ij}, \; \forall i, j$). Achievable rate suffers as the CSI error region expands.

$\mathbf{H}_{ij} \sim \mathcal{CN}\left(\sqrt{\frac{K_R}{1+K_R}}\mathbf{H}_0, \frac{1}{1+K_R}\mathbf{I}_{N_j} \otimes \mathbf{I}_{M_i}\right)$, $i \neq j$. $\mathbf{H}_0$ is matrix with all elements equal to 1 and K_R is the Rician coefficient. The results are then averaged over 100 channel realizations. Unless otherwise is stated we use the following values to define our default setup: $N_j = M_i = 4$, $K_R = 1$, $P_i = 1$ [Watt], $\kappa_i = \beta_j = 0.001$, $\sigma_{\mathrm{n},i}^2 = 0.001$ [Watt], $\mathbf{D}_{ij} = \mathbf{I}$, $\nu_i = 1$, $\zeta_{ij} = -15$ [dB], $\forall i, j \in \mathcal{D}$.

In Fig. 2 - Fig. 5 the worst-case (WC) achievable sum rate is numerically evaluated. In this respect, 'Scenario 1' represents the proposed design in Section III. 'Scenario 2', 'Scenario 3', and 'Scenario 4' represent a simplified sum-rate maximizing approach, by (respectively) ignoring the impact of distortion signals, i.e., assuming $\kappa = \beta = 0$, the impact of CSI inaccuracy, i.e., assuming $\zeta = 0$, and jointly ignoring the impacts of chain imperfections and CSI errors. Scenario 2 is adopted from the proposed design in [17], where Scenario 4 is previously studied in [10], [11]. Finally, 'Scenario 5' represents the achievable WC sum rate for the HD setup [18].

In Fig. 2 the average convergence behavior of the proposed iterative method is depicted. As it is observed, the convergence is obtained within 3-10 optimization iterations, which verifies the efficiency of the proposed iterative algorithm in terms of the required computational effort. Moreover, it is observed that such convergence behavior remains relatively constant for different levels of transceiver accuracy.

In Fig. 3 the obtained guaranteed (worst-case) system sum rate is evaluated for different values of transceiver inaccuracy. It is observed that while the achievable performance in all scenarios degrade as $\kappa = \beta$ increases, the performance of the FD system is more sensitive to the distortion intensity, compared to the HD system, due to the strong self-interference path. The gain of the proposed design ('Scenario 1') is particularly observable as the chain inaccuracy increases.

In Fig. 4 the obtained guaranteed (worst-case) system sum rate is evaluated for different values of channel error intensity. Similar to Fig. 3, it is observed that the proposed robust design provides a margin of gain as the CSI error radius increases.

In Fig. 5 the impact of the maximum transmit power is observed on the achievable (worst-case) system sum rate. It is observed that for a low power system, where the system performance is dominated by the effect of noise, a significant gain is obtained via the proposed FD setup. Nevertheless, the gain obtained by the proposed robust design is marginal. Conversely, in a high power regime where the system performance is dominated by the effect of distortion signals and CSI inaccuracy, the application of a robust design strategy becomes crucial.

V. Conclusion

While the application of bi-directional FD communication paradigm presents a potential for improving the spectral efficiency, such gains may be limited due to the imperfect self-interference cancellation. In this regard, the impact of inaccurate CSI estimation, as well as the common hardware inaccuracies play the dominant role. In this work we have presented a multi-convex optimization framework, for a worst-case sum rate maximization problem. The presented method provides a higher margin of robustness, compared to the previously presented designs in terms of worst-case performance. In particular, the observed gain increases as the CSI error region expands, or as the transceiver accuracy decreases.

Acknowledgment

The authors would like to thank M.Sc. Vimal Radhakrishnan, and Dr. Ali Cagatay Cirik for constructive discussions during this work.

References

[1] S. Hong, J. Brand, J. Choi, M. Jain, J. Mehlman, S. Katti, and P. Levis, "Applications of self-interference cancellation in 5G and beyond," *IEEE Communications Magazine*, February 2014.

[2] D. Bharadia and S. Katti, "Full duplex MIMO radios," in *Proceedings of the 11th USENIX Conference on Networked Systems Design and Implementation*, ser. NSDI'14, Berkeley, CA, USA, 2014, pp. 359–372.

[3] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proceedings of the ACM SIGCOMM*, Aug. 2013.

[4] M. Jain, J. I. Choi, T. Kim, D. Bharadia, K. Srinivasan, S. Seth, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proceedings of 17th Annual International Conference on Mobile Computing and Networking (MobiCom)*, Las Vegas, NV, Sep. 2011.

[5] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 2017–2046, Fourthquarter 2015.

[6] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *Selected Areas in Communications, IEEE Journal on*, vol. 32, no. 9, pp. 1637–1652, Sept 2014.

[7] V. Aggarwal, M. Duarte, A. Sabharwal, and N. Shankaranarayanan, "Full-or half-duplex? a capacity analysis with bounded radio resources," in *Information Theory Workshop (ITW), 2012 IEEE*. IEEE, 2012, pp. 207–211.

[8] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3702–3713, July 2012.

[9] S. Huberman and T. Le-Ngoc, "Self-interference-threshold-based MIMO full-duplex precoding," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 8, pp. 3803–3807, Aug 2015.

[10] ——, "Self-interference pricing-based MIMO full-duplex precoding," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 549–552, Dec 2014.

[11] J. Zhang, O. Taghizadeh, J. Luo, and M. Haardt, "Full duplex wireless communications with partial interference cancellation," *Proceedings of the 46th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA*, Nov. 2012.

[12] J. Zhang, O. Taghizadeh, and M. Haardt, "Robust transmit beamforming design for full-duplex point-to-point MIMO systems," *Proceedings of the Tenth International Symposium on Wireless Communication Systems (ISWCS)*, Aug. 2013.

[13] A. C. Cirik, R. Wang, and Y. Hua, "Weighted-sum-rate maximization for bi-directional full-duplex mimo systems," in *2013 Asilomar Conference on Signals, Systems and Computers*, Nov 2013, pp. 1632–1636.

[14] A. C. Cirik, J. Zhang, M. Haardt, and Y. Hua, "Sum-rate maximization for bi-directional full-duplex MIMO systems under multiple linear constraints," in *Proc. 15th Int. Workshop Signal Processing Advances in Wireless Communications (SPAWC 2014)*, Toronto, Canada, Jun. 2014.

[15] J. Zhang, O. Taghizadeh, and M. Haardt, "Transmit strategies for full-duplex point-to-point systems with residual self-interference," *in Proceedings of International ITG Workshop on Smart Antennas (WSA), Stuttgart, Germany*, March 2013.

[16] A. C. Cirik, R. Wang, Y. Rong, and Y. Hua, "Mse based transceiver designs for bi-directional full-duplex mimo systems," in *2014 IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2014, pp. 384–388.

[17] A. C. Cirik, O. Taghizadeh, L. Lampe, R. Mathar, and Y. Hua, "Sum-power minimization under rate constraints in full-duplex MIMO systems," *VTC 2016, Montreal, Canada*, 2016.

[18] J. Jose, N. Prasad, M. Khojastepour, and S. Rangarajan, "On robust weighted-sum rate maximization in MIMO interference networks," in *2011 IEEE International Conference on Communications (ICC)*, June 2011, pp. 1–6.

[19] M. F. Hanif, L. N. Tran, A. Tlli, M. Juntti, and S. Glisic, "Efficient solutions for weighted sum rate maximization in multicellular networks with channel uncertainties," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5659–5674, Nov 2013.

[20] A. Cirik, S. Biswas, S. Vuppala, and T. Ratnarajah, "Beamforming design for full-duplex mimo interference channels: Qos and energy-efficiency considerations," *IEEE Transactions on Communications*, vol. PP, no. 99, pp. 1–1, 2016.

[21] A. C. Cirik, S. Biswas, S. Vuppala, and T. Ratnarajah, "Robust transceiver design for full duplex multiuser MIMO systems," *IEEE Wireless Communications Letters*, vol. 5, no. 3, pp. 260–263, June 2016.

[22] J. Wang and D. P. Palomar, "Worst-case robust MIMO transmission with imperfect channel knowledge," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3086–3100, 2009.

[23] R. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Transactions on Signal Processing*, vol. 53, May 2005.

[24] A. Pascual, D. P. Palomar, A. I. Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with partial channel state information based on convex optimization," *IEEE Transactions on Signal Processing*, vol. 54, Jan. 2006.

[25] Q. Shi, M. Razaviyayn, Z. Q. Luo, and C. He, "An iteratively weighted mmse approach to distributed sum-utility maximization for a mimo interfering broadcast channel," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, Sept 2011.

[26] B. Day, A. Margetts, D. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE Journal on Selected Areas in Communications*, Sep. 2012.

[27] S. S. Christensen, R. Agarwal, E. D. 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, December 2008.

[28] S. Petersen and M. Pedersen, *Matrix Cookbook*, 2008.

[29] Y. C. Eldar and N. Merhav, "A competitive minimax approach to robust estimation of random parameters," *IEEE Transactions on Signal Processing*, vol. 52, no. 7, pp. 1931–1946, July 2004.

[30] M. V. Khlebnikov and P. S. Shcherbakov, "Petersens lemma on matrix uncertainty and its generalizations," *Automation and Remote Control*, vol. 69, no. 11, pp. 1932–1945, 2008.

[31] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM Journal on matrix analysis and applications*, vol. 19, no. 2, pp. 499–533, 1998.

[32] A. Mutapcic and S. Boyd, "Cutting-set methods for robust convex optimization with pessimizing oracles," *Optimization Methods & Software*, vol. 24, no. 3, pp. 381–406, 2009.

[33] X. Zheng, X. Sun, D. Li, and Y. Xu, "On zero duality gap in nonconvex quadratic programming problems," *Journal of Global Optimization*, vol. 52, no. 2, pp. 229–242, 2012.

[34] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.