Received 21 March 2021; revised 21 April 2021; accepted 5 May 2021. Date of publication 13 May 2021; date of current version 8 June 2021.

Digital Object Identifier 10.1109/OJCOMS.2021.3080239

Full-Duplex Amplify-and-Forward MIMO Relaying: Design and Performance Analysis Under Erroneous CSI and Hardware Impairments

OMID TAGHIZADEH [1] (Member, IEEE), SLAWOMIR STANCZAK [1] (Senior Member, IEEE), HIROKI IIMORI [2] (Graduate Student Member, IEEE), AND GIUSEPPE THADEU FREITAS DE ABREU [2] (Senior Member, IEEE)

[1] Network Information Theory Group, Technische Universitaät Berlin, 10587 Berlin, Germany

[2] Department of Electrical and Computer Engineering, Jacobs University Bremen, 28759 Bremen, Germany

CORRESPONDING AUTHOR: O. TAGHIZADEH (e-mail: taghizadehmotlagh@tu-berlin.de)

This work was supported by Deutsche Forschungsgemeinschaft (DFG), under Grant MA 1184/38-1 (DUPLINK). Part of this work has been presented in 2020 IEEE Global Communications Conference [58].

ABSTRACT Full-duplex amplify-and-forward multiple-input multiple-output relaying has been the focus of several recent studies, due to the potential for achieving a higher spectral efficiency and lower latency, together with inherent processing simplicity. However, when the impact of hardware distortions are considered, such relays suffer from a *distortion-amplification loop*, due to the inter-dependent nature of the relay transmit signal covariance and the residual self-interference covariance. The aforementioned behavior leads to a significant performance degradation for a system with a low or medium hardware accuracy. In this work, we analyze the relay transfer function as well as the mean squared-error performance of a full-duplex amplify-and-forward multiple-input multiple-output relaying communication, under the impact of collective sources of additive and multiplicative transmit and receive impairments. Building on the performed analysis, an optimization problem is devised to minimize the communication Mean Squared-Error and solved by employing the recently proposed penalty dual-decomposition framework. The proposed solution converges to a stationary point of the original problem via a sequence of convex quadratic programs, thereby enjoying an acceptable arithmetic complexity as the problem dimensions grow large. Numerical simulations verify the significance of the proposed distortion-aware design and analysis, compared to the common simplified approaches, as the hardware accuracy degrades.

INDEX TERMS Full-duplex, amplify-and-forward, multiple-input multiple-output relay, hardware impairments, penalty-dual-decomposition.

I. INTRODUCTION

FULL-DUPLEX (FD) relays have been the focus of several recent studies, due to their potential to achieve a higher level of spectral efficiency and a lower end-to-end latency, compared to their Half-Duplex (HD) counterparts. This is because an FD relay has the capability to transmit and receive the relayed signal at the same time and frequency, enabled by the recently-developed Self-Interference Cancellation (SIC) techniques, e.g., [1]–[4], which provide an adequate level of isolation between transmit (Tx) and receive (Rx) directions motivating a wide range of related applications, see, e.g., [5], [6]. A common idea of such SIC techniques is to attenuate the main interference paths in RF domain, i.e., prior to down-conversion, so that the remaining Self-Interference (SI) can be processed in the effective dynamic range of the analog-to-digital converter (ADC) and further attenuated in the baseband, i.e., digital domain. While the aforementioned SIC techniques are proved to be successful for specific scenarios, e.g., [3], it is easy to observe that the obtained

cancellation level may vary for different realistic conditions. This mainly includes *i)* aging and inaccuracy of the hardware components, e.g., ADC and digital-to-analog-converter (DAC) noise, power amplifier and oscillator phase noise in analog domain, as well as *ii)* inaccurate estimation of the remaining interference paths due to the limited channel coherence time. As a result, it is essential to take into account the aforementioned inaccuracies to obtain a design which remains efficient under realistic situations.

In this work we focus on the application of FD multiple-input multiple-output (MIMO) relays operating with an Amplify-and-Forward (AF) processing protocol. Compared to their FD-Decode-and-Forward (DF) counterpart, e.g., [7], [8], FD-Amplify-and-Forward (AF) relays are known for their processing simplicity and good performance, which have attracted much interest recently, e.g., [9]–[21]. In this regard, the FD-AF relays have been the focus of [22]–[27], where the relay is equipped with a single antenna. In the aforementioned works, the impact of linear system uncertainties in the digital domain have been studied in [24], where the hardware imperfections from analog domain components have been addressed in [22], [23], following the model in [7], [28].

While the aforementioned literature introduces the importance of an accurate transceiver modeling with respect to the effects of hardware impairments for an FD-AF relay, such works are not yet extended to MIMO relaying setups. This stems from the fact that in an FD-AF relay, the interdependent behavior of the transmit signal covariance from the relay and the residual SI covariance result in a distortion amplification loop effect, see Section II-E.3. The aforementioned effect results in a rather complicated mathematical description when the relay is equipped with multiple antennas. As a result, related studies resort to simplified models to reduce the consequent design complexity. In [9]–[17] a multiple-antenna FD-AF relay system is studied where perfect SIC is assumed, thanks to either the estimation and subtraction of the interference in the receiver [9]–[11], or to the spatial zero-forcing of the SI signal under the assumption that the number of transmit antennas exceeds the number of receive antennas at the relay [12]–[17]. For the scenarios where the number of transmit antennas is not higher than the receive antennas, a general framework is proposed in [29], [30], assuming fixed and known residual SI statistics, and in [18], [19], where perfect SIC[1] is assumed thanks to a combined analog/digital SIC scheme on the condition that the SI power does not exceed a certain threshold. In [31], an FD-AF relay has been studied for a single-stream communication, assuming perfect hardware at both ends. Furthermore, the impact of CSI error, resulting from noise and the impact of impairments is not considered. In [20], [21], [32] the residual SI signal is related to the transmit signal via a known and linear function, assuming a distortion-free hardware. To the best of the authors knowledge, the consideration of the known system uncertainties, i.e., Channel State Information (CSI) error and transmit/receive chain impairments, have not been addressed in the context of MIMO FD AF relays.

1. Residual SI is assumed to be buried in the thermal noise, following a known statistics.

A. CONTRIBUTION

In this work, we study a one-way MIMO FD-AF relay scheme, where the explicit impact of hardware distortions in the receiver and transmit chains are taken into account. Our goal is to analyze the instantaneous end-to-end performance under collective sources of impairments, and subsequently to improve the system performance via optimized linear transmit/receive strategies. The main contributions are as follows:

- Due to the joint consideration of hardware distortions in the receiver and transmitter chains, we observe an interdependent behavior of the relay transmit covariance and the residual SI covariance in an FD-AF relay, i.e., the distortion amplification loop. Note that this behavior may not be captured from prior works based on simplified residual SI models, e.g., [9]–[21], [29], [30], [32]. In the first step, we analyze the system operation under the impact of collective sources of additive and multiplicative impairments, i.e., CSI error and transmit/receive chain distortions, following the transceiver impairments characterization reported in [7], [28]. In this respect, the relay transfer function, which relates the relay transmit covariance to the received undistorted covariance, is analytically derived as a function of relay amplification, CSI error and impairments statistics. Please note that this is in contrast to [31] where the relay function is analyzed for a single stream communication, assuming a perfect CSI and accurate hardware at both ends. By employing the relay transfer function, the system performance is then derived in terms of the end-to-end Mean Squared-Error (MSE).
- Building on the obtained analysis, we propose linear transmit/amplification/receive strategies at the source, relay and destination, with the intention of minimizing MSE. The instantaneous CSI is utilized to control the impact of distortion, and to enhance the quality of the desired signal. This is in contrast to [9]–[21], [29], [30], [32] where the dependency of the distortion statistics to the intended transmit/receive signal is ignored. In this regard, an MSE minimization problem is formulated which shows an intractable mathematical structure. An iterative solution is proposed, following the recently developed Penalty Dual-Decomposition (PDD) framework for non-convex problems with non-linear equality constraints. The proposed solution converges to a stationary point of the original problem via a sequence of Convex Quadratic Program (CQP)s, thereby enjoying an acceptable arithmetic complexity as the problem dimensions grow large.

- The proposed PDD-based framework is then extended to different system setups and design metrics in order to be applicable in the diverse popular scenarios. This includes the extension of the proposed single-user design into a scenario with multiple MIMO users as well as the extension of the MSE minimization algorithm for rate maximization. Moreover, the consideration of per-antenna instantaneous power constraint, as an effective consideration for SIC in FD transceivers, have been integrated in the proposed framework.

Numerical simulations verify the significance of the proposed distortion-aware design and analysis, compared to the common simplified approaches, as the hardware accuracy degrades.

The rest of this paper is organized as follows. The system model is defined in Section II. The end-to-end MSE performance of the studied relaying system is then obtained in Section III. In Section IV the relay performance is optimized, employing the MSE analysis from Section III. The proposed design framework is then extended in Section V, covering a larger scenarios and design metrics. The numerical simulations are presented in Section VI. This paper is concluded in Section VII by summarizing the main findings.

B. MATHEMATICAL NOTATION

Throughout this paper, column vectors and matrices are denoted as lower-case and upper-case bold letters, respectively. The rank of a matrix, expectation, trace, transpose, conjugate, Hermitian transpose, determinant, Euclidean and Frobenius norms are denoted by $\text{rank}(\cdot)$, $\mathbb{E}(\cdot)$, $\text{tr}(\cdot)$, $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $|\cdot|$, $||\cdot||_2$, $|\cdot|_F$, respectively. The Kronecker product is denoted by $\otimes$. The identity matrix with dimension K is denoted as $\mathbf{I}_K$, the $\text{vec}(\cdot)$ operator stacks the elements of a matrix into a vector, and $(\cdot)^{-1}$ represents the inverse of a matrix. The sets of real, real and positive, complex, natural, and the set $\{1 \dots K\}$ are respectively denoted by , $^+$, $\mathbb{C}$, $\mathbb{N}$ and $\mathbb{F}_K$. $\mathcal{R}_i(\mathbf{X})$ returns the i-th row of the matrix $\mathbf{X}$. $\{a_i\}$ denotes the set of a_i, $\forall i$. The set of all positive semi-definite matrices is denoted by $\mathcal{H}$. $\perp$ represents statistical independence. $x^\star$ is the value of the variable x at optimality.

II. SYSTEM MODEL

We study a system where an HD source, equipped with N_s antennas, communicates with an HD destination node, equipped with M_d antennas, with the help of a one-way Full-Duplex (FD) relay. The relay is equipped with N_r (M_r) transmit (receive) antennas, and operates in AF mode. The channels between the source and the relay, between the relay and the destination, and between the source and the destination are denoted as $\mathbf{H}_\text{sr} \in \mathbb{C}^{M_\text{r} \times N_\text{s}}$, $\mathbf{H}_\text{rd} \in \mathbb{C}^{M_\text{d} \times N_\text{r}}$, and $\mathbf{H}_\text{sd} \in \mathbb{C}^{M_\text{d} \times N_\text{s}}$, respectively. The SI channel, which is the channel between the relay's transmit and receive ends is denoted as $\mathbf{H}_\text{rr} \in \mathbb{C}^{M_\text{r} \times N_\text{r}}$. All channels are assumed to follow the flat-fading model.

A. CSI ESTIMATION

An effective estimation method is presented in [7, Sec. III-A] for an FD relaying setup in the presence of hardware impairments.[2] However, the perfect CSI estimation is not practically feasible, due to the presence of thermal noise, hardware impairments, as well as the limited available resource which can be dedicated to channel estimation. In this work we consider the CSI inaccuracy as a correlated complex Gaussian additive error [7], [33], [34] as

$$\begin{aligned} &\mathbf{H}_\mathcal{X} = \tilde{\mathbf{H}}_\mathcal{X} + \Delta_\mathcal{X}, \\ &\Delta_\mathcal{X} \perp \tilde{\mathbf{H}}_\mathcal{X}, \quad \Delta_\mathcal{X} = \mathbf{C}_{\text{rx},\mathcal{X}}^{1/2} \tilde{\Delta}_\mathcal{X} \mathbf{C}_{\text{tx},\mathcal{X}}^{1/2}, \ \text{vec}\left(\tilde{\Delta}_\mathcal{X}\right) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}), \end{aligned} \tag{1}$$

where $\tilde{\mathbf{H}}_\mathcal{X}$ denotes the estimated channel with $\mathcal{X} \in \{\text{sr}, \text{rd}, \text{rr}, \text{sd}\}$, and $\mathbf{C}_{\text{rx},\mathcal{X}} \in \mathcal{H}$ ($\mathbf{C}_{\text{tx},\mathcal{X}} \in \mathcal{H}$) represent the receive (transmit) correlation matrices, which depend on the employed estimation method as well as the interference and impairments statistics. Please note that while the abovementioned correlation matrices represent correlation among the elements of the CSI error, the CSI error remains independent to the active signal components. This is since the CSI error at each active communication block is affected by the signal sources, e.g., data symbols, noise and distortion, during the training process but are independent to the active signal components during the communication phase. Please also see [7], [33], [34] for similar statistical assumptions on the CSI error.

B. SOURCE-TO-RELAY COMMUNICATION

The relay continuously receives and amplifies the received signal from the source, while estimating and subtracting the loopback SI signal from its own transmitter, see Fig. 1. The transmitted signal by the source and the received signal at the relay are respectively expressed as

$$\mathbf{x} = \underbrace{\mathbf{F}\mathbf{s}}_{\mathbf{u}_\text{s}} + \mathbf{e}_\text{tx,s}, \tag{2}$$

$$\mathbf{r}_\text{in} = \underbrace{\mathbf{H}_\text{sr}\mathbf{x} + \mathbf{H}_\text{rr}\mathbf{r}_\text{out} + \mathbf{n}_\text{r}}_{=:\mathbf{m}_\text{in}} + \mathbf{e}_\text{rx,r}, \tag{3}$$

where $\mathbf{x} \in \mathbb{C}^{N_\text{s}}$ and $\mathbf{s} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_d)$ are the transmitted signal and data symbols from the source and $\mathbf{F} \in \mathbb{C}^{N_\text{s} \times d}$ is the transmit precoder.

At the relay side, $\mathbf{r}_\text{in} \in \mathbb{C}^{M_\text{r}}$ and $\mathbf{r}_\text{out} \in \mathbb{C}^{N_\text{r}}$ respectively represent the received and transmitted signal at the relay and $\mathbf{n}_\text{r} \sim \mathcal{CN}(\mathbf{0}, \sigma_\text{nr}^2 \mathbf{I}_{M_\text{r}})$ represents the zero-mean additive white complex Gaussian (ZMAWCG) noise at the relay.

The *undistorted* transmitted signal at the source and the corresponding transmit hardware distortions, representing the combined effects of transmit chain impairments, e.g., limited DAC accuracy, oscillator phase noise, power-amplifier

2. A two-phase estimation is suggested to avoid interference; first, source transmits the pilot where relay is silent, thereby estimating the source-relay and source-destination channels, and then relay transmits pilot and source remain silent, hence estimating the SI and relay-destination channels.

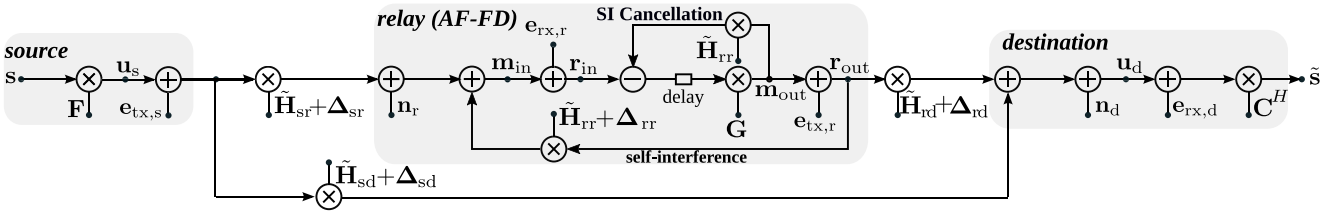


FIGURE 1. The signal model for the studied amplify-and-forward FD MIMO relay. The impact of hardware impairments form the transmitter and receiver chains, as well as the inaccurate CSI is observed at all nodes.

noise [7], are denoted as $\mathbf{u}_{\rm s} \in \mathbb{C}^{N_{\rm s}}$ and $\mathbf{e}_{\rm tx,s} \in \mathbb{C}^{N_{\rm s}}$, respectively. Similarly, the undistorted received signal at the relay and the receiver hardware distortions, representing the combined effects of receiver chain impairments, e.g., limited ADC accuracy, oscillator phase noise, low-noise-amplifier (LNA) distortion, are denoted as $\mathbf{m}_{\rm in} \in \mathbb{C}^{M_{\rm r}}$ and $\mathbf{e}_{\rm rx,r} \in \mathbb{C}^{M_{\rm r}}$, see Section II-D for more details. Please note that the impact of the aforementioned impairments are more pronounced in an FD transceiver, in comparison to an HD one, due to high strength of the SI path. The known, i.e., distortion-free, part of the SI signal is then suppressed in the receiver by utilizing the recently developed SIC techniques in analog and digital domains, e.g., [1], [3]. The remaining signal is then amplified to constitute the relay's output:

$$\mathbf{r}_{\rm out} = \mathbf{m}_{\rm out} + \mathbf{e}_{\rm tx,r}, \tag{4}$$

$$\mathbf{m}_{\rm out}(t) = \mathbf{G}\tilde{\mathbf{r}}_{\rm in}(t-\tau), \tag{5}$$

$$\begin{aligned}\tilde{\mathbf{r}}_{\rm in} &= \mathbf{r}_{\rm in} - \tilde{\mathbf{H}}_{\rm rr}\mathbf{m}_{\rm out} \\ &= \mathbf{H}_{\rm sr}\mathbf{x} + \mathbf{n}_{\rm r} + \underbrace{\mathbf{e}_{\rm rx,r} + \tilde{\mathbf{H}}_{\rm rr}\mathbf{e}_{\rm tx,r} + \Delta_{\rm rr}\mathbf{r}_{\rm out}}_{\text{residual SI}},\end{aligned} \tag{6}$$

where $\tilde{\mathbf{r}}_{\rm in} \in \mathbb{C}^{M_{\rm r}}$ and $\mathbf{G} \in \mathbb{C}^{N_{\rm r}\times M_{\rm r}}$ respectively represent the interference-suppressed version of the received signal and the relay amplification matrix, $t \in^{+}$ represents the time instance,[3] and $\tau \in^{+}$ is the relay processing delay, see Section II-E.2.

The *undistorted* transmit signal at the relay and the transmit distortion are denoted as $\mathbf{m}_{\rm out} \in \mathbb{C}^{N_{\rm r}}$ and $\mathbf{e}_{\rm tx,r} \in \mathbb{C}^{N_{\rm r}}$, respectively. In order to take into account the transmit power limitations we impose

$$\mathbb{E}\{\|\mathbf{r}_{\rm out}\|_2^2\} \le P_{\rm r,max},\ \mathbb{E}\{\|\mathbf{x}\|_2^2\} \le P_{\rm s,max}, \tag{7}$$

where $P_{\rm r,max}$ and $P_{\rm s,max}$ respectively represents the maximum transmit power from the relay and from the source.

C. RELAY-TO-DESTINATION COMMUNICATION

Consequently, the received signal at the destination can be expressed as

$$\mathbf{y} = \underbrace{\mathbf{H}_{\rm rd}\mathbf{r}_{\rm out} + \mathbf{H}_{\rm sd}\mathbf{x} + \mathbf{n}_{\rm d}}_{\mathbf{u}_{\rm d}} + \mathbf{e}_{\rm rx,d}, \tag{8}$$

$$\hat{\mathbf{s}}(t-\tau) = \mathbf{C}^H\mathbf{y}(t), \tag{9}$$

3. The argument indicating time instance, i.e., t, is dropped for simplicity for signals with a same time reference.

where $\mathbf{y} \in \mathbb{C}^{M_{\rm d}}$ ($\mathbf{u}_{\rm d} \in \mathbb{C}^{M_{\rm d}}$) is the received signal (undistorted received signal) at the destination, whereas $\mathbf{e}_{\rm rx,d} \in \mathbb{C}^{M_{\rm d}}$ and $\mathbf{n}_{\rm d} \sim \mathcal{CN}(\mathbf{0}, \sigma_{\rm nd}^2\mathbf{I}_{M_{\rm d}})$ denote the receiver hardware distortion and the ZMAWCG noise at the destination.

The linear receiver processing filter and the estimated received data symbols are denoted as $\mathbf{C} \in \mathbb{C}^{M_{\rm d}\times d}$ and $\hat{\mathbf{s}}$, respectively.

D. DISTORTION SIGNAL STATISTICS

The impact of hardware elements inaccuracy in each chain is modeled as additive distortion terms, following the transceiver impairments model proposed in [28] and widely used in the context of FD system design and performance analysis, e.g., [7], [8], [31], [35]–[40]. The proposed model in [28] is based on the following three observations. Firstly, the collective distortion signal in each transmit/receive chain can be approximated as an additive zero-mean Gaussian term [41]–[43]. Secondly, the variance of the distortion signal is proportional to the power of the intended transmit/received signal. And thirdly, the distortion signal is statistically independent to the intended transmit/receive signal at each chain, and for different chains, see [28, Sec. C], [31]. Please note that the statistical independence property can be concluded from the result of Bussgang's theorem [59], as well as from the fact that after proper signal processing any *known* correlation among the residual SI and the known signal elements is used to reduce the residual SI and hence eliminated. In the defined relaying system, the statistics of the distortion signals are characterized as

$$\begin{aligned}\mathbf{e}_{\rm tx,s} &\sim \mathcal{CN}\Big(\mathbf{0}, \kappa_s \mathrm{diag}\Big(\mathbb{E}\Big\{\mathbf{u}_s\mathbf{u}_s^H\Big\}\Big)\Big), \\ &\mathbf{e}_{\rm tx,s} \perp \mathbf{u}_s,\ \ \mathbf{e}_{\rm tx,s}(t) \perp \mathbf{e}_{\rm tx,s}(t'),\end{aligned} \tag{10a}$$

$$\begin{aligned}\mathbf{e}_{\rm rx,r} &\sim \mathcal{CN}\Big(\mathbf{0}, \beta_r \mathrm{diag}\Big(\mathbb{E}\Big\{\mathbf{m}_{\rm in}\mathbf{m}_{\rm in}^H\Big\}\Big)\Big), \\ &\mathbf{e}_{\rm rx,r} \perp \mathbf{m}_{\rm in},\ \ \mathbf{e}_{\rm rx,r}(t) \perp \mathbf{e}_{\rm rx,r}(t'),\end{aligned} \tag{10b}$$

$$\begin{aligned}\mathbf{e}_{\rm tx,r} &\sim \mathcal{CN}\Big(\mathbf{0}, \kappa_r \mathrm{diag}\Big(\mathbb{E}\Big\{\mathbf{m}_{\rm out}\mathbf{m}_{\rm out}^H\Big\}\Big)\Big), \\ &\mathbf{e}_{\rm tx,r} \perp \mathbf{m}_{\rm out},\ \ \mathbf{e}_{\rm tx,r}(t) \perp \mathbf{e}_{\rm tx,r}(t'),\end{aligned} \tag{10c}$$

$$\begin{aligned}\mathbf{e}_{\rm rx,d} &\sim \mathcal{CN}\Big(\mathbf{0}, \beta_d \mathrm{diag}\Big(\mathbb{E}\Big\{\mathbf{u}_{\rm d}\mathbf{u}_{\rm d}^H\Big\}\Big)\Big), \\ &\mathbf{e}_{\rm rx,d} \perp \mathbf{u}_{\rm d},\ \ \mathbf{e}_{\rm rx,d}(t) \perp \mathbf{e}_{\rm rx,d}(t'),\end{aligned} \tag{10d}$$

where $t \neq t'$ indicate the time instance and $\beta_{\mathcal{X}}, \kappa_{\mathcal{X}} \in^{+}$, $\mathcal{X} \in \{\rm s,r,d\}$ are the receive and transmit distortion coefficients.

It is worth mentioning that the values of $\kappa_{\mathcal{X}}$, $\beta_{\mathcal{X}}$ depend of the implemented SIC scheme, and reflect the quality of the

cancellation. For instance, in a massive MIMO transceiver implementation where the analog cancellation is less efficient and relies on baseband processing, e.g., [8], the quality of self-interference cancellation is dominated by quantization resolution and the values of κ, β [dB] are proportional to the number of quantization bits. For the implementation proposed in [3], the distortion coefficients depend on the number of the employed attenuation and phase shifter units in the RF cancellation module, for more discussions on the used distortion model please see [7], [8], [28], [31], [35] and the references therein.

E. REMARKS

1) DIRECT LINK

In this work, we assume that the direct link is weak and consider the source-destination path as a source of interference, similar to [7], [24], [31]. For the scenarios where the direct link is strong, it is shown in [44] that the receiver strategy can be gainfully updated as a RAKE receiver [45] to temporally align the desired signal in the direct and relay links.

2) PROCESSING DELAY

The relay output signals, i.e., $\mathbf{m}_{\text{out}}$ and $\mathbf{r}_{\text{out}}$, are generated from the received signals with a relay processing delay τ, see (4). This delay is assumed to be long enough, e.g., more than a symbol duration, such that the source signal is decorrelated, i.e., $\mathbf{s}(t) \perp \mathbf{s}(t-\tau)$ [44], [46]. The zero-mean and independent statistics of the samples from data signal, i.e., $\mathbf{s}(t)$ and $\mathbf{s}(t-\tau)$, as well as the noise and distortion signals, are basis for the analysis in the following section.

3) DISTORTION-AMPLIFICATION LOOP

As observed from (6) and (10a)-(10d), as the transmit signal covariance at the relay scales up, the statistics of transmit and receive distortion at the relay and the residual SI signal scale up consequently. The aforementioned distortions are later amplified in the relay process as part of the received signal at the relay, see (4), and further scale up the transmit signal covariance from the relay. The observed inter-dependence among the transmit signal covariance and the residual SI signal covariance, lead to a distortion-amplification-loop effect, which degrade the relay performance when the dynamic range is not high. In the following part, we analyze the performance of the defined relay-assisted communication, under the impact of hardware impairments. As the transmit power from the relay increases, the power of the error components increase in all receiver chains, see (3) in connection to (10b)-(10c). On the other hand, these errors are amplified in the relay process and further increase the relay transmit power, see (3) in connection to (6) and (4). The aforementioned effect causes a loop which signifies the problem of residual SI for the relays with AF process. In the following section, this impact is analytically studied and an optimization strategy is proposed in order to alleviate this effect.

III. ANALYSIS UNDER HARDWARE IMPAIRMENTS

In this part, we analyze the MSE performance of the relay-assisted communication system under the impact of impairments, as a function of the tunable parameters, i.e., $\mathbf{F}, \mathbf{G}, \mathbf{C}$. By incorporating (4) and (10a)-(10c) into (6), (5) we have

$$\mathbb{E}\{\mathbf{m}_{\text{in}}\} = 0, \tag{11}$$

$$\begin{aligned}\mathbb{E}\left\{\mathbf{m}_{\text{in}}\mathbf{m}_{\text{in}}^H\right\} &:= \mathbf{M}_{\text{in}} = \tilde{\mathbf{H}}_{\text{sr}}\left(\mathbf{F}\mathbf{F}^H + \kappa_{\text{s}}\text{diag}\left(\mathbf{F}\mathbf{F}^H\right)\right) \\ &\quad \times \tilde{\mathbf{H}}_{\text{sr}}^H + \sigma_{\text{nr}}^2\mathbf{I}_{M_{\text{r}}} \\ &\quad \underbrace{+\mathbf{C}_{\text{rx,sr}}\text{tr}\left(\mathbf{C}_{\text{tx,sr}}\left(\mathbf{F}\mathbf{F}^H + \kappa_{\text{s}}\text{diag}\left(\mathbf{F}\mathbf{F}^H\right)\right)\right)}_{=:\mathbf{M}_{\text{in},0}} \\ &\quad + \tilde{\mathbf{H}}_{\text{rr}}(\mathbf{M}_{\text{out}} + \kappa_{\text{r}}\text{diag}(\mathbf{M}_{\text{out}}))\tilde{\mathbf{H}}_{\text{rr}}^H \\ &\quad + \mathbf{C}_{\text{rx,rr}}\text{tr}\left(\mathbf{C}_{\text{tx,rr}}(\mathbf{M}_{\text{out}} + \kappa_{\text{r}}\text{diag}(\mathbf{M}_{\text{out}}))\right),\end{aligned} \tag{12}$$

and similarly,

$$\mathbb{E}\{\mathbf{m}_{\text{out}}\} = 0, \tag{13}$$

$$\mathbb{E}\left\{\mathbf{m}_{\text{out}}\mathbf{m}_{\text{out}}^H\right\} = \mathbf{G}\mathbb{E}\left\{\tilde{\mathbf{r}}_{\text{in}}\tilde{\mathbf{r}}_{\text{in}}^H\right\}\mathbf{G}^H =: \mathbf{M}_{\text{out}} \tag{14}$$

$$\begin{aligned}\mathbb{E}\left\{\tilde{\mathbf{r}}_{\text{in}}\tilde{\mathbf{r}}_{\text{in}}^H\right\} &= \mathbf{M}_{\text{in},0} + \tilde{\mathbf{H}}_{\text{rr}}\kappa_{\text{r}}\text{diag}(\mathbf{M}_{\text{out}})\tilde{\mathbf{H}}_{\text{rr}}^H \\ &\quad + \mathbf{C}_{\text{rx,rr}}\text{tr}\left(\mathbf{C}_{\text{tx,rr}}(\mathbf{M}_{\text{out}} + \kappa_{\text{r}}\text{diag}(\mathbf{M}_{\text{out}}))\right) \\ &\quad + \beta_{\text{r}}\text{diag}(\mathbf{M}_{\text{in}}),\end{aligned} \tag{15}$$

where $\mathbf{M}_{\text{in}} \in \mathcal{H}$ represents the undistorted receive signal covariance at the relay and (11), (12) hold as the noise, the desired signal at subsequent symbol durations and the distortion components are zero-mean and mutually independent. $\mathbf{M}_{\text{in},0}$ is the undistorted received signal covariance at the relay, when the relay transmission is turned off, i.e., $\mathbf{G} = \mathbf{0}$.

Similarly, $\mathbf{M}_{\text{out}} \in \mathcal{H}$ represents the covariance of the undistorted transmit signal from the relay, and holds the relation with $\mathbf{r}_{\text{out}}$ as $\mathbb{E}\{\mathbf{r}_{\text{out}}\mathbf{r}_{\text{out}}^H\} = \mathbf{M}_{\text{out}} + \kappa_{\text{r}}\text{diag}(\mathbf{M}_{\text{out}})$. Our goal is to obtain a closed-form expression of the relay transmit signal covariance, as a function of the $\mathbf{F}, \mathbf{G}$. By substituting (12) into (15) we have

$$\begin{aligned}\mathbb{E}\left\{\tilde{\mathbf{r}}_{\text{in}}\tilde{\mathbf{r}}_{\text{in}}^H\right\} &\approx \mathbf{M}_{\text{in},0} + \tilde{\mathbf{H}}_{\text{rr}}(\kappa_{\text{r}}\text{diag}(\mathbf{M}_{\text{out}}))\tilde{\mathbf{H}}_{\text{rr}}^H \\ &\quad + \mathbf{C}_{\text{rx,rr}}\text{tr}\left(\mathbf{C}_{\text{tx,rr}}(\mathbf{M}_{\text{out}} + \kappa_{\text{r}}\text{diag}(\mathbf{M}_{\text{out}}))\right) \\ &\quad + \beta_{\text{r}}\text{diag}\Big(\mathbf{M}_{\text{in},0} + \tilde{\mathbf{H}}_{\text{rr}}(\mathbf{M}_{\text{out}})\tilde{\mathbf{H}}_{\text{rr}}^H \\ &\qquad + \mathbf{C}_{\text{rx,rr}}\text{tr}\left(\mathbf{C}_{\text{tx,rr}}(\mathbf{M}_{\text{out}})\right)\Big) \\ &=: \mathcal{M}_1(\mathbf{F}, \mathbf{M}_{\text{out}})\end{aligned} \tag{16}$$

where the approximation (16) is obtained by considering $\kappa_{\text{r}}, \beta_{\text{r}} \ll 1$ and hence ignoring the expressions including higher order of the distortion coefficients [7], [31]. This subsequently leads to

$$\mathbf{M}_{\text{out}} = \mathbf{G}\mathcal{M}_1(\mathbf{F}, \mathbf{M}_{\text{out}})\mathbf{G}^H. \tag{17}$$

Unfortunately, a direct expression of $\mathbf{M}_{\text{out}}$ in terms of $\mathbf{G}, \mathbf{F}$ can not be achieved from (17) in the current form. In order to facilitate further analysis we therefore resort to

the vectorized presentation of (17). By applying the well-known matrix equality $\mathrm{vec}(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3) = (\mathbf{A}_3^{\mathrm{T}} \otimes \mathbf{A}_1)\mathrm{vec}(\mathbf{A}_2)$, we have

$$\mathrm{vec}\Big(\mathbb{E}\Big\{\mathbf{r}_{\mathrm{out}}\mathbf{r}_{\mathrm{out}}^H\Big\}\Big) = \Big(\mathbf{I}_{N_{\mathrm{r}}^2} + \kappa_{\mathrm{r}}\mathbf{D}_{N_{\mathrm{r}}}\Big)\mathrm{vec}(\mathbf{M}_{\mathrm{out}}), \tag{18}$$

where $\mathbf{D}_M \in \{0,1\}^{M^2\times M^2}$ is a selection matrix with one or zero elements such that $\mathbf{D}_{N_{\mathrm{r}}}\mathrm{vec}(\mathbf{M}_{\mathrm{out}}) = \mathrm{vec}(\mathrm{diag}(\mathbf{M}_{\mathrm{out}}))$. Similarly from (17) we obtain

$$\begin{aligned}\mathrm{vec}(\mathbf{M}_{\mathrm{out}}) = &\Big(\mathbf{I}_{N_{\mathrm{r}}^2} - \big(\mathbf{G}^*\otimes\mathbf{G}\big)\mathbf{B}\Big)^{-1}\big(\mathbf{G}^*\otimes\mathbf{G}\big)\\ &\times\Big(\mathbf{I}_{M_{\mathrm{r}}^2} + \beta_{\mathrm{r}}\mathbf{D}_{M_{\mathrm{r}}}\Big)\mathrm{vec}\big(\mathbf{M}_{\mathrm{in},0}\big),\end{aligned} \tag{19}$$

where

$$\begin{aligned}\mathbf{B} = &\kappa_{\mathrm{r}}\Big(\tilde{\mathbf{H}}_{\mathrm{rr}}^*\otimes\tilde{\mathbf{H}}_{\mathrm{rr}}\Big)\mathbf{D}_{N_{\mathrm{r}}} + \beta_{\mathrm{r}}\mathbf{D}_{M_{\mathrm{r}}}\Big(\tilde{\mathbf{H}}_{\mathrm{rr}}^*\otimes\tilde{\mathbf{H}}_{\mathrm{rr}}\Big)\\ &+ \mathrm{vec}\big(\mathbf{C}_{\mathrm{rx,rr}}\big)\mathrm{vec}\big(\mathbf{C}_{\mathrm{tx,rr}}^*\big)^T\Big(\mathbf{I}_{N_{\mathrm{r}}^2} + \kappa_{\mathrm{r}}\mathbf{D}_{N_{\mathrm{r}}}\Big)\\ &+ \beta_{\mathrm{r}}\mathbf{D}_{M_{\mathrm{r}}}\mathrm{vec}\big(\mathbf{C}_{\mathrm{rx,rr}}\big)\mathrm{vec}\big(\mathbf{C}_{\mathrm{tx,rr}}^*\big)^T.\end{aligned} \tag{20}$$

The direct dependence of the relay transmit covariance matrix and $\mathbf{F}, \mathbf{G}$ can be hence obtained from (19) and (18) as

$$\mathrm{vec}\Big(\mathbf{r}_{\mathrm{out}}\mathbf{r}_{\mathrm{out}}^H\Big) = \Theta\Big(\mathbf{G}, \tilde{\mathbf{H}}_{\mathrm{rr}}, \mathbf{C}_{\mathrm{tx,rr}}, \mathbf{C}_{\mathrm{rx,rr}}, \kappa_{\mathrm{r}}, \beta_{\mathrm{r}}\Big)\mathrm{vec}\big(\mathbf{M}_{\mathrm{in},0}\big), \tag{21}$$

where

$$\begin{aligned}&\Theta\Big(\mathbf{G}, \tilde{\mathbf{H}}_{\mathrm{rr}}, \mathbf{C}_{\mathrm{tx,rr}}, \mathbf{C}_{\mathrm{rx,rr}}, \kappa_{\mathrm{r}}, \beta_{\mathrm{r}}\Big)\\ &= \Big(\mathbf{I}_{N_{\mathrm{r}}^2} + \kappa_{\mathrm{r}}\mathbf{D}_{N_{\mathrm{r}}}\Big)\times\Big(\mathbf{I}_{N_r^2} - \big(\mathbf{G}^*\otimes\mathbf{G}\big)\mathbf{B}\Big)^{-1}\\ &\quad\times\big(\mathbf{G}^*\otimes\mathbf{G}\big)\Big(\mathbf{I}_{M_{\mathrm{r}}^2} + \beta_{\mathrm{r}}\mathbf{D}_{M_{\mathrm{r}}}\Big),\end{aligned} \tag{22}$$

represents the transfer function of the relay; relating the distortion-less input signal statistics, i.e., $\mathbf{M}_{\mathrm{in},0}$ to the distorted transmit covariance, under the collective impact of transmit and receive chain impairments, CSI inaccuracy, and the strength of the SI channel.[4]

A. MEAN SQUARED ERROR (MSE)

Incorporating the collective sources of impairments, the transmit relay covariance is obtained from (21) as a function of the tunable parameters, i.e., $\mathbf{F}, \mathbf{G}$. Consequently, the communication MSE can be formulated as

$$\begin{aligned}\mathbf{E} &:= \mathbb{E}\Big\{(\mathbf{s}-\tilde{\mathbf{s}})(\mathbf{s}-\tilde{\mathbf{s}})^H\Big\} = \mathbb{E}\Big\{\Big(\mathbf{s}-\mathbf{C}^H\mathbf{y}\Big)\Big(\mathbf{s}-\mathbf{C}^H\mathbf{y}\Big)^H\Big\}\\ &= \mathbf{I}_d - \mathbf{C}^H\mathbb{E}\Big\{\mathbf{y}\mathbf{s}^H\Big\} - \mathbb{E}\Big\{\mathbf{s}\mathbf{y}^H\Big\}\mathbf{C} + \mathbf{C}^H\mathbb{E}\Big\{\mathbf{y}\mathbf{y}^H\Big\}\mathbf{C},\end{aligned} \tag{23}$$

where $\mathbf{E}$ denotes the MSE matrix. By recalling (8) and (21) we have

$$\mathbb{E}\Big\{\mathbf{y}(t)\mathbf{s}^H(t-\tau)\Big\} = \tilde{\mathbf{H}}_{\mathrm{rd}}\mathbf{G}\tilde{\mathbf{H}}_{\mathrm{sr}}\mathbf{F}, \tag{24}$$

4. It is observed that $\Theta(\mathbf{W}, \tilde{\mathbf{H}}_{\mathrm{rr}}, 0, 0, 0, 0) = \mathbf{G}^*\otimes\mathbf{G}$, which is similar to the known FD-AF relay operation with perfect CSI and accurate hardware, i.e., $\kappa_{\mathrm{r}}, \beta_{\mathrm{r}} = 0$.

$$\begin{aligned}\mathbb{E}\Big\{\mathbf{y}(t)\mathbf{y}^H(t)\Big\} \approx\ &\tilde{\mathbf{H}}_{\mathrm{sd}}\Big(\mathbf{F}\mathbf{F}^H + \kappa_{\mathrm{s}}\mathrm{diag}\Big(\mathbf{F}\mathbf{F}^H\Big)\Big)\tilde{\mathbf{H}}_{\mathrm{sd}}^H\\ &+ \mathbf{C}_{\mathrm{rx,sd}}\mathrm{tr}\Big(\mathbf{C}_{\mathrm{tx,sd}}\Big(\mathbf{F}\mathbf{F}^H + \kappa_{\mathrm{s}}\mathrm{diag}\Big(\mathbf{F}\mathbf{F}^H\Big)\Big)\Big)\\ &+ \tilde{\mathbf{H}}_{\mathrm{rd}}(\mathbf{M}_{\mathrm{out}} + \kappa_{\mathrm{s}}\mathrm{diag}(\mathbf{M}_{\mathrm{out}}))\tilde{\mathbf{H}}_{\mathrm{rd}}^H\\ &+ \mathbf{C}_{\mathrm{rx,rd}}\mathrm{tr}\big(\mathbf{C}_{\mathrm{tx,rd}}(\mathbf{M}_{\mathrm{out}} + \kappa_{\mathrm{r}}\mathrm{diag}(\mathbf{M}_{\mathrm{out}}))\big) + \sigma_{\mathrm{nd}}^2\mathbf{I}_{M_{\mathrm{d}}}\\ &+ \beta_{\mathrm{d}}\mathrm{diag}\Big(\tilde{\mathbf{H}}_{\mathrm{sd}}\Big(\mathbf{F}\mathbf{F}^H\Big)\tilde{\mathbf{H}}_{\mathrm{sd}}^H + \mathbf{C}_{\mathrm{rx,sd}}\mathrm{tr}\Big(\mathbf{C}_{\mathrm{tx,sd}}\Big(\mathbf{F}\mathbf{F}^H\Big)\Big)\\ &\qquad + \tilde{\mathbf{H}}_{\mathrm{rd}}(\mathbf{M}_{\mathrm{out}})\tilde{\mathbf{H}}_{\mathrm{rd}}^H + \mathbf{C}_{\mathrm{rx,rd}}\mathrm{tr}\big(\mathbf{C}_{\mathrm{tx,rd}}(\mathbf{M}_{\mathrm{out}})\big) + \sigma_{\mathrm{nd}}^2\mathbf{I}_{M_{\mathrm{d}}}\Big)\\ =:\ &\mathcal{M}_2(\mathbf{F}, \mathbf{M}_{\mathrm{out}}),\end{aligned} \tag{25}$$

where the approximation (25) is obtained, similar to that of (16), considering $\kappa_{\mathrm{s}}, \kappa_{\mathrm{r}}, \beta_{\mathrm{d}} \ll 1$ and hence ignoring the terms including higher order coefficients. Please note that similar to (11) and (12), the identities (24) and (25) are obtained recalling that all components of noise, additive transmit and receiver distortions, CSI error, and data symbols at subsequent symbol durations are zero-mean and mutually independent.

Consequently, we have

$$\begin{aligned}\mathbf{E}(\mathbf{F}, \mathbf{G}, \mathbf{C}) \approx\ &\mathbf{C}^H\mathcal{M}_2(\mathbf{F}, \mathbf{M}_{\mathrm{out}})\mathbf{C} + \mathbf{I}_d - \mathbf{C}^H\tilde{\mathbf{H}}_{\mathrm{rd}}\mathbf{G}\tilde{\mathbf{H}}_{\mathrm{sr}}\mathbf{F}\\ &- \mathbf{F}^H\tilde{\mathbf{H}}_{\mathrm{sr}}^H\mathbf{G}^H\tilde{\mathbf{H}}_{\mathrm{rd}}^H\mathbf{C},\end{aligned} \tag{26}$$

which presents the communication MSE as a function of the $\mathbf{F}, \mathbf{G}, \mathbf{C}$.

IV. JOINT DESIGN: A SEQUENTIAL QUADRATIC PROGRAMMING FRAMEWORK

In this part, we provide a framework for optimizing the defined relaying system, utilizing the expressions derived in Section III. In particular, the corresponding MSE minimization problem is formulated as

$$\min_{\mathbf{F},\mathbf{G},\mathbf{C}}\ \mathrm{tr}(\mathbf{E}(\mathbf{F}, \mathbf{G}, \mathbf{C})) \tag{27a}$$

$$\text{s.t.}\ \ \mathbb{E}\Big\{\|\mathbf{x}\|^2\Big\} \le P_{\mathrm{s,max}},\ \ \mathbb{E}\Big\{\|\mathbf{r}_{\mathrm{out}}\|^2\Big\} \le P_{\mathrm{r,max}}, \tag{27b}$$

where (27b) represents the power constraints at source and at the relay.

Unfortunately, the optimization problem (27) is not tractable in this form due to the non-linear and non-convex nature of the objective as well as the relay transmit power constraint with respect to $\mathbf{F}, \mathbf{G}$, see (25), (26) in connection to (21). In particular, the main challenge lies in the non-linear nature of (21), including the matrix inversion in the expanded variable space due to the vectorized representation of (17), which reflects the inter-dependent behavior of the transmit relay covariance and the residual SI. In order to obtain a tractable form of (27), in the first step, we add the undistorted relay transmit covariance $\mathbf{M}_{\mathrm{out}}$ as an auxiliary variable, i.e.,

$$\min_{\mathbf{F},\mathbf{G},\mathbf{C},\mathbf{M}_{\mathrm{out}}}\ \mathrm{tr}(\mathbf{E}(\mathbf{F}, \mathbf{G}, \mathbf{C}, \mathbf{M}_{\mathrm{out}})) \tag{28a}$$

$$\text{s.t.}\quad \|\mathbf{F}\|_F^2 \le \frac{P_{\mathrm{s,max}}}{(1+\kappa_{\mathrm{s}})},\ \ \mathrm{tr}(\mathbf{M}_{\mathrm{out}}) \le \frac{P_{\mathrm{r,max}}}{(1+\kappa_{\mathrm{r}})}, \tag{28b}$$

$$\mathbf{M}_{\mathrm{out}} = \mathbf{G}\mathcal{M}_1(\mathbf{F}, \mathbf{M}_{\mathrm{out}})\mathbf{G}^H, \tag{28c}$$

which is obtained by directly incorporating (17) as constraint (28c). Please note that the above reformulation eliminates the matrix inversion in (18), thereby resulting in a reduced numerical complexity. Furthermore, the resulting structure of the transmit relay power constraint as well as the objective are block-wise convex with respect the problem variables.

Nevertheless, the (28) is still intractable, due to the non-convex objective, as well as the tight nonlinear equality constraint (28c). Our goal is to obtain a tractable structure by employing the Penalty Dual Decomposition (PDD) framework, recently introduced in [47], [48] and used, e.g., in [49], [50], where the tight non-linear constraints can be resolved by employing the proposed updates on the augmented Lagrangian, see Section IV-A. In order to facilitate the updates on decoupled variable blocks, we employ similar trick as introduced in, e.g., [49, Sec. IV-C], by defining the auxiliary functions $\bar{\mathcal{M}}_1, \bar{\mathcal{M}}_2$ and $\bar{\mathcal{M}}_3$, respectively for the relay transmit covariance, received covariance at the destination, and the MSE metric in (29)-(31), shown at the bottom of the page, as well as the additional auxiliary variables $\mathbf{J}, \mathbf{L}_i, \tilde{\mathbf{F}}, \tilde{\mathbf{J}}, \mathbf{L}_i, \tilde{\mathbf{L}}_i, \forall i$. In this regard, the equivalent expressions

$$\mathcal{M}_1(\mathbf{F}, \mathbf{M}_{\text{out}}) = \bar{\mathcal{M}}_1\Big(\mathbf{F}, \tilde{\mathbf{F}}, \mathbf{J}, \tilde{\mathbf{J}}\Big),$$
$$\mathcal{M}_2(\mathbf{F}, \mathbf{M}_{\text{out}}) = \bar{\mathcal{M}}_2\Big(\mathbf{F}, \tilde{\mathbf{F}}, \mathbf{J}, \tilde{\mathbf{J}}\Big), \tag{32}$$
$$\text{s.t. } \mathbf{F} - \tilde{\mathbf{F}} = \mathbf{0}, \quad \mathbf{J} - \tilde{\mathbf{J}} = \mathbf{0}, \quad \mathbf{M}_{\text{out}} - \mathbf{J}\tilde{\mathbf{J}}^H = \mathbf{0}, \tag{33}$$

facilitate a decoupled and convex dependency of the problem metrics in each case to the variable blocks. By employing similar transformations on (28c) and (28a), problem (28) is equivalently reformulated as

$$\min_{\mathcal{B}_1, \mathcal{B}_2} \ \|\mathbf{L}_4\|_F^2 \tag{34a}$$
$$\text{s.t. } \mathcal{X} - \tilde{\mathcal{X}} = \mathbf{0}, \quad \forall \mathcal{X} \in \{\mathbf{F}, \mathbf{J}, \mathbf{L}_1, \ldots, \mathbf{L}_4\} \tag{34b}$$
$$\bar{\mathcal{M}}_i\Big(\mathbf{F}, \tilde{\mathbf{F}}, \mathbf{J}, \tilde{\mathbf{J}}\Big) - \mathbf{L}_i\tilde{\mathbf{L}}_i^H = \mathbf{0}, \quad \forall i \in \{1, 2\} \tag{34c}$$
$$\bar{\mathcal{M}}_3\Big(\mathbf{L}_3, \tilde{\mathbf{L}}_3, \mathbf{C}, \mathbf{L}_6\Big) - \mathbf{L}_4\tilde{\mathbf{L}}_4^H = \mathbf{0}, \quad \mathbf{C}^H\mathbf{L}_2 - \mathbf{L}_3 = \mathbf{0}, \tag{34d}$$
$$\mathbf{J} - \mathbf{G}\mathbf{L}_1 = \mathbf{0}, \quad \mathbf{L}_5 - \tilde{\mathbf{H}}_{\text{sr}}\mathbf{F} = \mathbf{0}, \quad \mathbf{L}_6 - \tilde{\mathbf{H}}_{\text{rd}}\mathbf{G}\mathbf{L}_5 = \mathbf{0}, \tag{34e}$$
$$\|\mathbf{F}\|_F^2 \le P_{\text{s,max}}/(1+\kappa_{\text{s}}), \quad \|\mathbf{J}\|_F^2 \le P_{\text{r,max}}/(1+\kappa_{\text{r}}), \tag{34f}$$

where

$$\mathcal{B}_1 := \Big\{\mathbf{F}, \mathbf{G}, \mathbf{C}, \mathbf{J}, \tilde{\mathbf{L}}_1, \ldots, \tilde{\mathbf{L}}_4\Big\}, \tag{35a}$$
$$\mathcal{B}_2 := \Big\{\tilde{\mathbf{F}}, \tilde{\mathbf{J}}, \mathbf{L}_1, \ldots, \mathbf{L}_6\Big\}. \tag{35b}$$

Note that problem (34) is still non-convex, due to the highly non-linear and coupled equality constraints. However, while the objective and inequality constraints in (34) are both de-coupled and convex, the equality constraints hold a block-affine structure over $\mathcal{B}_1$ and $\mathcal{B}_2$, when one of the variable blocks are considered as constant. Hence, it complies with the structure presented in [47], [48, Sec. II], and can be consequently solved as a sequence of CQPs with convergence towards a point satisfying Karush–Kuhn–Tucker (KKT) conditions. Please see Section (IV-B) for a detailed justification of the formulation in (34).

A. PDD METHOD

In this subsection, we briefly review the employed PDD method, as a general framework for solving coupled equality-constraint non-convex optimization problems. Please note that as opposed to the widely used alternating direction method of multipliers (ADMM) method, the PDD method deals with the problems where the equality constraints need not to be linear and de-coupled. Consider the optimization problem

$$\min_{\mathbf{x}\in\mathbb{X}} \ f(\mathbf{x}) \tag{36a}$$
$$\text{s.t. } \mathbf{h}(\mathbf{x}) = \mathbf{0}, \tag{36b}$$
$$\mathbf{g}_i(\mathbf{x}_i) \le \mathbf{0}, \quad \forall i \in \{1, \ldots, q\}, \tag{36c}$$

where $f(\mathbf{x}) \in \mathbb{R}$ is a scalar and continuously differentiate function, $\mathbf{h}(\mathbf{x}) \in \mathbb{R}^p$ represents a vector of p continuously differentiable but potentially non-convex and coupled functions, and $\mathbf{g}_i(\mathbf{x}_i) \in \mathbb{R}^{q_i}$ represents the vector of differentiable functions over the variable block $\mathbf{x}_i$, such that $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_q)$.

Note that if the coupling constraint $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ does not exist for problem (36a), the classic Block Coordinate Descent (BCD) methods can be applied to obtain a KKT

$$\begin{aligned}\bar{\mathcal{M}}_1\Big(\mathbf{F}, \tilde{\mathbf{F}}, \mathbf{J}, \tilde{\mathbf{J}}\Big) = {} & \tilde{\mathbf{H}}_{\text{sr}}\Big(\mathbf{F}\tilde{\mathbf{F}}^H + \kappa_{\text{s}}\text{diag}\Big(\mathbf{F}\tilde{\mathbf{F}}^H\Big)\Big)\tilde{\mathbf{H}}_{\text{sr}}^H + \mathbf{C}_{\text{rx,sr}}\text{tr}\Big(\mathbf{C}_{\text{tx,sr}}\Big(\mathbf{F}\tilde{\mathbf{F}}^H + \kappa_{\text{s}}\text{diag}\Big(\mathbf{F}\tilde{\mathbf{F}}^H\Big)\Big)\Big) \\ & + \tilde{\mathbf{H}}_{\text{rr}}\Big(\kappa_{\text{r}}\text{diag}\Big(\mathbf{J}\tilde{\mathbf{J}}^H\Big)\Big)\tilde{\mathbf{H}}_{\text{rr}}^H + \mathbf{C}_{\text{rx,rr}}\text{tr}\Big(\mathbf{C}_{\text{tx,rr}}\Big(\mathbf{J}\tilde{\mathbf{J}}^H + \kappa_{\text{r}}\text{diag}\Big(\mathbf{J}\tilde{\mathbf{J}}^H\Big)\Big)\Big) + \beta_{\text{r}}\text{diag}\Big(\tilde{\mathbf{H}}_{\text{sr}}\Big(\mathbf{F}\tilde{\mathbf{F}}^H\Big)\tilde{\mathbf{H}}_{\text{sr}}^H \\ & + \mathbf{C}_{\text{rx,sr}}\text{tr}\Big(\mathbf{C}_{\text{tx,sr}}\Big(\mathbf{F}\tilde{\mathbf{F}}^H\Big)\Big) + \tilde{\mathbf{H}}_{\text{rr}}\Big(\mathbf{J}\tilde{\mathbf{J}}^H\Big)\tilde{\mathbf{H}}_{\text{rr}}^H + \mathbf{C}_{\text{rx,rr}}\text{tr}\Big(\mathbf{C}_{\text{tx,rr}}\Big(\mathbf{J}\tilde{\mathbf{J}}^H\Big)\Big) + \sigma_{\text{nr}}^2\mathbf{I}_{M_{\text{r}}}\Big) + \sigma_{\text{nr}}^2\mathbf{I}_{M_{\text{r}}}\end{aligned} \tag{29}$$
$$\begin{aligned}\bar{\mathcal{M}}_2\Big(\mathbf{F}, \tilde{\mathbf{F}}, \mathbf{J}, \tilde{\mathbf{J}}\Big) = {} & \tilde{\mathbf{H}}_{\text{sd}}\Big(\mathbf{F}\tilde{\mathbf{F}}^H + \kappa_{\text{s}}\text{diag}\Big(\mathbf{F}\tilde{\mathbf{F}}^H\Big)\Big)\tilde{\mathbf{H}}_{\text{sd}}^H + \mathbf{C}_{\text{rx,sd}}\text{tr}\Big(\mathbf{C}_{\text{tx,sd}}\Big(\mathbf{F}\tilde{\mathbf{F}}^H + \kappa_{\text{s}}\text{diag}\Big(\mathbf{F}\tilde{\mathbf{F}}^H\Big)\Big)\Big) \\ & + \tilde{\mathbf{H}}_{\text{rd}}\Big(\mathbf{J}\tilde{\mathbf{J}}^H + \kappa_{\text{r}}\text{diag}\Big(\mathbf{J}\tilde{\mathbf{J}}^H\Big)\Big)\tilde{\mathbf{H}}_{\text{rd}}^H + \mathbf{C}_{\text{rx,rd}}\text{tr}\Big(\mathbf{C}_{\text{tx,rd}}\Big(\mathbf{J}\tilde{\mathbf{J}}^H + \kappa_{\text{r}}\text{diag}\Big(\mathbf{J}\tilde{\mathbf{J}}^H\Big)\Big)\Big) + \sigma_{\text{nd}}^2\mathbf{I}_{M_{\text{d}}}\end{aligned} \tag{30}$$
$$\bar{\mathcal{M}}_3\Big(\mathbf{X}, \tilde{\mathbf{X}}, \mathbf{C}, \mathbf{Y}\Big) = \mathbf{X}_1\tilde{\mathbf{X}}_1^H - \mathbf{C}^H\mathbf{Y} - \mathbf{Y}^H\mathbf{C} + \mathbf{I} \tag{31}$$

solution by decomposing the problem (36a) into a sequence of small-scale problems. In order to resolve the coupled equality constraint, the PDD method aims for minimizing the corresponding Augmented Lagrangian (AL) function:

$$\min_{\mathbf{x},\mathbf{g}_i(\mathbf{x}_i)\leq\mathbf{0}} \; f(\mathbf{x})+\boldsymbol{\lambda}^T\mathbf{h}(\mathbf{x})+\frac{1}{2\rho}\|\mathbf{h}(\mathbf{x})\|^2 = f(\mathbf{x})+\frac{1}{2\rho}\|\mathbf{h}(\mathbf{x})+\rho\boldsymbol{\lambda}\|^2, \tag{37a}$$

where $\boldsymbol{\lambda}$ is the dual variable, and ρ is the constraint violation penalty.

Note that as $1/\rho\to\infty$, the problem (37a) yields an identical solution as (36a). However, as obseved in the penalty methods for constrained optimization problems, a large penalty parameter leads to a significantly slower algorithm convergence, as the direction of movement will be dominated by the terms representing constraints violation [47], [48]. In order to resolve this issue, the PDD method employs a dual loop structure, where in the inner loop the AL is minimized over $\mathbf{x}$, whereas the variables $\boldsymbol{\lambda}$, ρ are updated in the outer loop, until a given stability criteria is met. In particular, the minimization over the variable blocks $\mathbf{x}_i$ is done following the Block Successive Upper-bound Minimization (BSUM) algorithm [51], when $f(\mathbf{x})$ is a convex function over each block $\mathbf{x}_k$. At each outer iteration, depending on the constraint violation level, either the penalty parameter ρ (when violation level is high) or the dual variables $\boldsymbol{\lambda}$ are updated. The algorithm is proven to converge to a solution that satisfies KKT optimality conditions. For a detailed convergence proof and a summary of the general PDD framework please see [47, Sec. III].

B. PDD-BASED MSE MINIMIZATION

The core of the PDD method is to construct the AL function corresponding to the studied optimization problem, where the tight inequality constraints are transformed as an integral part of the objective. It is worth mentioning that in order to utilize the specific structure of the PDD method, the equivalent optimization problem in (34) is defined such that, firstly, all of the tight equality constraints are seperately affine with respect to the variable blocks $\mathcal{B}_1,\mathcal{B}_2$. Secondly, the variables appearing in the inequality constraints are decoupled. And third, all of the main search variables $\mathbf{F},\mathbf{G},\mathbf{C}$ are updated within the same block. Please note that the first property facilitates the implementation of the BSUM algorithm [51], as a standard CQP during each block update, which enables a low-complexity computation. Furthermore, the second and third property facilitates the convergence towards a KKT point of the original problem. This is of high significance, considering that the BCD-type methods for the problems with coupled constraints do not usually achieve first-order optimality conditions [51]. The AL-minimization problem, corresponding to (34), is hence formulated as

$$\begin{aligned}\min_{\mathcal{B}_1,\mathcal{B}_2}\quad & \mathbf{AL}(\mathcal{B}_1,\mathcal{B}_2,\{\boldsymbol{\lambda}_{\mathcal{X}}\},\{\boldsymbol{\lambda}_i\},\rho)\\ \text{s.t.}\quad & \|\mathbf{F}\|_F^2\leq P_{\text{s,max}}/(1+\kappa_{\text{s}}),\quad \|\mathbf{J}\|_F^2\leq P_{\text{r,max}}/(1+\kappa_{\text{r}}),\end{aligned} \tag{38}$$

where

$$\begin{aligned}\mathbf{AL}(\mathcal{B}_1,\mathcal{B}_2,\{\boldsymbol{\lambda}_{\mathcal{X}}\},\{\boldsymbol{\lambda}_i\},\rho) &\coloneqq \|\mathbf{L}_4\|_F^2\\ &+\frac{1}{2\rho}\Bigg(\sum_{\mathcal{X}}\Big\|\mathcal{X}-\tilde{\mathcal{X}}^H+\rho\boldsymbol{\lambda}_{\mathcal{X}}\Big\|_F^2\\ &+\sum_{i\in\{1,2\}}\Big\|\bar{\mathcal{M}}_i\Big(\mathbf{F},\tilde{\mathbf{F}},\mathbf{J},\tilde{\mathbf{J}}\Big)-\mathbf{L}_i\tilde{\mathbf{L}}_i^H+\rho\boldsymbol{\lambda}_i\Big\|_F^2\\ &+\|\mathbf{J}-\mathbf{G}\mathbf{L}_1+\rho\boldsymbol{\lambda}_3\|_F^2+\Big\|\mathbf{C}^H\mathbf{L}_2-\mathbf{L}_3+\rho\boldsymbol{\lambda}_4\Big\|_F^2\\ &+\Big\|\mathbf{L}_5-\tilde{\mathbf{H}}_{\text{sr}}\mathbf{F}+\rho\boldsymbol{\lambda}_5\Big\|_F^2+\Big\|\mathbf{L}_6-\tilde{\mathbf{H}}_{\text{rd}}\mathbf{G}\mathbf{L}_5+\rho\boldsymbol{\lambda}_7\Big\|_F^2\\ &+\Big\|\bar{\mathcal{M}}_3\Big(\mathbf{L}_3,\tilde{\mathbf{L}}_3,\mathbf{C},\mathbf{L}_6\Big)-\mathbf{L}_4\tilde{\mathbf{L}}_4^H+\rho\boldsymbol{\lambda}_6\Big\|_F^2\Bigg),\\ \mathcal{X}&\in\{\mathbf{F},\mathbf{J},\mathbf{L}_1,\ldots,\mathbf{L}_4\}.\end{aligned} \tag{39}$$

It can be observed that **AL** holds a convex quadratic form over each of the variable blocks $\mathcal{B}_1,\mathcal{B}_2$ in each case when the other variable block is kept constant. Furthermore, the problem (38) becomes equivalent to (27), once the penalty variables are chosen large enough. In the following we present the detailed steps for the variable updates, following the PDD method.

1) INNER LOOP: VARIABLE UPDATE $\{\mathcal{B}_1,\mathcal{B}_2\}$

The minimization of **AL** over the variable blocks $\mathcal{B}_1,\mathcal{B}_2$ can be done as a CQP, in each case when the other variable block is fixed. In particular, the minimization over $\mathcal{B}_1$ is done as a quadratic-constraint CQP, whereas the minimization over $\mathcal{B}_2$ is an unconstrained CQP. The sequence of the aforementioned updates lead to a monotonic decrease and converge to a stationary point of the **AL** function, please see [51, Sec. IV].

2) OUTER LOOP: VARIABLE UPDATE: $\{\boldsymbol{\lambda}_{\mathcal{X}}\},\{\boldsymbol{\lambda}_I\},\rho$

In the standard penalty method [52], the elimination of the penalty violations may lead to significantly large penalty coefficients, which lead to inefficient numerical convergence and ill-conditioned sub-problems. This is since the penalty imposed on the square of the constraint violations vanish quadratically as the violation becomes small. This calls for a significantly larger penalty coefficients as the violations become small. In order to resolve this issue, PDD method employs a hybrid approach; the penalty coefficient ρ is updated when the violations are large, whereas the dual variables $\{\boldsymbol{\lambda}_{\mathcal{X}}\},\{\boldsymbol{\lambda}_i\}$ are updated when the violation is small.

Similar to [49], the decision on the magnitude of constraint violation is made by evaluating a total violation function, defined as $\zeta(\mathcal{B}_1,\mathcal{B}_2)$ in (40), shown at the bottom of the next page, representing the total constraint violation such that the problems (27) and (34) become equivalent as $\zeta(\mathcal{B}_1,\mathcal{B}_2)\to 0$. Assuming a given threshold ζ_0, the variable update in the outer loop is defined as

$$\rho^{m+1}=c_\rho\rho^m, \tag{41}$$

when $\zeta(\mathcal{B}_1^m, \mathcal{B}_2^m) \geq \zeta_0$, and as

$$\boldsymbol{\lambda}_{\mathcal{X}}^{m+1} = \boldsymbol{\lambda}_{\mathcal{X}}^{m} + \frac{1}{\rho^m}\Big(\mathcal{X} - \tilde{\mathcal{X}}\Big), \tag{42}$$

$$\boldsymbol{\lambda}_{i}^{m+1} = \boldsymbol{\lambda}_{i}^{m} + \frac{1}{\rho^m}\Big(\bar{\mathcal{M}}_i\Big(\mathbf{F}, \tilde{\mathbf{F}}, \mathbf{J}, \tilde{\mathbf{J}}\Big) - \mathbf{L}_i\tilde{\mathbf{L}}_i^{H}\Big), \tag{43}$$

$$\boldsymbol{\lambda}_{3}^{m+1} = \boldsymbol{\lambda}_{3}^{m} + \frac{1}{\rho^m}\Big(\bar{\mathcal{M}}_3\Big(\mathbf{L}_3, \tilde{\mathbf{L}}_3, \mathbf{C}, \mathbf{L}_6\Big) - \mathbf{L}_4\tilde{\mathbf{L}}_4^{H}\Big), \tag{44}$$

$$\boldsymbol{\lambda}_{4}^{m+1} = \boldsymbol{\lambda}_{4}^{m} + \frac{1}{\rho^m}(\mathbf{J} - \mathbf{G}\mathbf{L}_1), \tag{45}$$

$$\boldsymbol{\lambda}_{5}^{m+1} = \boldsymbol{\lambda}_{5}^{m} + \frac{1}{\rho^m}\Big(\mathbf{L}_5 - \tilde{\mathbf{H}}_{\text{sr}}\mathbf{F}\Big), \tag{46}$$

$$\boldsymbol{\lambda}_{6}^{m+1} = \boldsymbol{\lambda}_{6}^{m} + \frac{1}{\rho^m}\Big(\mathbf{L}_6 - \tilde{\mathbf{H}}_{\text{rd}}\mathbf{G}\mathbf{L}_5\Big), \tag{47}$$

when $\zeta(\mathcal{B}_1^m, \mathcal{B}_2^m) < \zeta_0$. Please note that the updates in (42)-(47) follow from the method of incrementing the penalty parameters in the direction of the violation, please see [49, eq. (49)], [47]. In the above expressions, m indicates the algorithm iterations in the outer loop, and $0 < c_\rho < 1$ is a constant that characterizes the growth of the penalty parameter.

The iterations of the outer loop continue until $\zeta(\mathcal{B}_1^m, \mathcal{B}_2^m) \geq \zeta_{\text{th}}$, where $\zeta_{\text{th}} \ll 1$ indicates the tolerable numerical threshold for the total constraint violation. The detailed procedure of the PDD-based MSE minimization method is summarized in Algorithm 1.

Algorithm 1 1 PDD-Based MSE Minimization Algorithm for (27). $\epsilon, \zeta_{\text{th}}$ Are the Stability Threshold

initialize $\mathcal{B}_1^{0,0}, \mathcal{B}_2^{0,0} \in \mathbb{X}$, $\rho^0 > 0$, $0 < c_\rho < 1$, $\boldsymbol{\lambda}^0 = \mathbf{0}$, $k = 0$, $m = 0$;
Repeat ▷ Outer loop
  if $\zeta\Big(\mathcal{B}_1^{k,m}, \mathcal{B}_2^{k,m}\Big) \geq \zeta_{\text{th}}$: **then**
    Update: $\rho^{k+1} \leftarrow$ (41)
  else
    Update: $\Big\{\boldsymbol{\lambda}_{\mathcal{X}}^{k+1}\Big\}, \Big\{\boldsymbol{\lambda}_{i}^{k+1}\Big\} \leftarrow$ (42)
  end if
  $k \leftarrow k + 1$
  Repeat ▷ Inner loop
    $\mathcal{B}_1^{k,m+1} \leftarrow \underset{\mathcal{B}_1}{\text{argmin}}\ \mathbf{AL}\Big(\mathcal{B}_1, \mathcal{B}_2^{k,m}, \{\boldsymbol{\lambda}_{\mathcal{X}}\}, \Big\{\boldsymbol{\lambda}_{i}^{k}\Big\}, \rho^k\Big)$,
    $\mathcal{B}_2^{k,m+1} \leftarrow \underset{\mathcal{B}_2}{\text{argmin}}\ \mathbf{AL}\Big(\mathcal{B}_1^{k,m}, \mathcal{B}_2, \Big\{\boldsymbol{\lambda}_{\mathcal{X}}^{k}\Big\}, \Big\{\boldsymbol{\lambda}_{i}^{k}\Big\}, \rho^k\Big)$,
    $\text{AL}^{k,m} \leftarrow \mathbf{AL}\Big(\mathcal{B}_1^{k,m}, \mathcal{B}_2^{k,m}, \Big\{\boldsymbol{\lambda}_{\mathcal{X}}^{k}\Big\}, \Big\{\boldsymbol{\lambda}_{i}^{k}\Big\}, \rho^k\Big)$
    $m \leftarrow m + 1$
  Until $\text{AL}^{k,m} - \text{AL}^{k,m-1} \geq \epsilon$
Until $\zeta\Big(\mathcal{B}_1^{k,m}, \mathcal{B}_2^{k,m}\Big) \geq \zeta_{\text{th}}$
Return $\mathcal{B}_1^{k,m}, \mathcal{B}_2^{k,m}$

C. CONVERGENCE

The problem (34) is a special case of the general problem studied in [47], [48], with differentiable objective and inequality constraints. In this regard, the sequence of the dual-loop PDD updates lead to a solution satisfying KKT conditions when iterates of the inner loop converge to a stationary point of the AL function [47, Sec. III-B]. In particular, since at each block-update a convex sub-problem is solved to the optimality we have

$$\text{AL}^{k,m} \geq \text{AL}^{k,m+1} \geq \cdots \geq 0, \quad \forall k, \tag{48}$$

which ensures eventual convergence of the inner iterations. Furthermore, due to *i)* separate convexity of the AL function over the variable blocks, *ii)* de-coupledness and joint convexity of the inequality constraints (38) and *iii)* smoothness of the both objective and the constratins in (38), the updates of the inner iterations comply with the requirements specified in [51, Th. 2-b] in the context of the BSUM algorithm, and hence the converging point will be a solution satisfying KKT conditions for the problem (38). The convergence behavior of the inner iterations, as well as the required iterations until elimination of constraint violations in the outer loop are also numerically evaluated in Section VI-A.

D. COMPUTATIONAL COMPLEXITY

In this part, we analyze the arithmetic complexity associated with the Algorithm 1 in relation to the system dimensions. It is observed that the arithmetic complexity of the algorithm is dominated by the inner-loop computations of the CQP sub-problems. The canonical form of a real-valued conic QCP is expressed as [53]

$$\min_{\mathbf{x}}\ \mathbf{b}^T\mathbf{x}, \quad \text{s.t.}\ \ \|\mathbf{x}\|_2 \leq d_0, \quad \|\mathbf{A}_m\mathbf{x} + \mathbf{b}_m\|_2 \leq \mathbf{c}_m^T\mathbf{x} + d_m,$$
$$\forall m \in \{1, \ldots, \tilde{M}\}, \quad \mathbf{x} \in^{\tilde{N}}, \quad \mathbf{b}_m \in^{L_m}, \tag{49}$$

where $\tilde{N}, \tilde{M}$, and L_m respectively describe the dimention of the real-valued variable space, number of constraints and the constraint dimensions. The arithmetic complexity of obtaining an ϵ-solution to the defined problem, i.e., the convergence to the ϵ-distance vicinity of the optimum is upper-bounded by

$$\mathcal{O}(1)\Big(1 + \tilde{M}\Big)^{\frac{1}{2}}\tilde{N}\left(\tilde{N}^2 + \tilde{M} + \sum_{m=1}^{\tilde{M}} l_m^2\right)\text{digit}(\epsilon), \tag{50}$$

where $\text{digit}(\epsilon)$ is a constant obtained from [53, Sec. 4.1.2], and affected by the required solution precision.

$$\zeta(\mathcal{B}_1, \mathcal{B}_2) = \sum_{\mathcal{X}}\Big\|\mathcal{X} - \tilde{\mathcal{X}}^H\Big\|_F^2 + \sum_{i\in\{1,2\}}\Big\|\bar{\mathcal{M}}_i\Big(\mathbf{F}, \tilde{\mathbf{F}}, \mathbf{J}, \tilde{\mathbf{J}}\Big) - \mathbf{L}_i\tilde{\mathbf{L}}_i^{H}\Big\|_F^2 + \|\mathbf{J} - \mathbf{G}\mathbf{L}_1\|_F^2 + \Big\|\mathbf{C}^H\mathbf{L}_2 - \mathbf{L}_3\Big\|_F^2$$
$$+ \Big\|\mathbf{L}_5 - \tilde{\mathbf{H}}_{\text{sr}}\mathbf{F}\Big\|_F^2 + \Big\|\bar{\mathcal{M}}_3\Big(\mathbf{L}_3, \tilde{\mathbf{L}}_3, \mathbf{C}, \mathbf{L}_6\Big) - \mathbf{L}_4\tilde{\mathbf{L}}_4^{H}\Big\|_F^2 + \Big\|\mathbf{L}_6 - \tilde{\mathbf{H}}_{\text{rd}}\mathbf{G}\mathbf{L}_5\Big\|_F^2, \quad \mathcal{X} \in \{\mathbf{F}, \mathbf{J}, \mathbf{L}_1, \ldots, \mathbf{L}_4\} \tag{40}$$

1) COMPLEXITY OF $\mathcal{B}_1$ UPDATE, I.E., $\mathcal{C}(\mathcal{B}_2)$

As observed from (38) the update of $\mathcal{B}_1$ lead to a standard conic QCP. By transforming the problem into its canonical form[5] (49), the size of the variable space is given as $\tilde{N} = 4d(N_\mathrm{s}+M_\mathrm{d}) + 2(N_\mathrm{r}M_\mathrm{r} + N_\mathrm{r}^2 + M_\mathrm{r}^2 + M_\mathrm{d}^2 + d^2 + d(2M_\mathrm{d} + M_\mathrm{r}))$. Furthermore, the number of separate constraints and the respective dimensions are calculated as $\tilde{M} = 3$ where $l_1 = 2dN_\mathrm{s}$, $l_2 = 2N_\mathrm{r}^2$, and $l_3 = 2d(N_\mathrm{s}+M_\mathrm{d}+d) + 2(N_\mathrm{r}^2 + M_\mathrm{r}^2 + M_\mathrm{d}^2)$.

2) COMPLEXITY OF $\mathcal{B}_2$ UPDATE, I.E., $\mathcal{C}(\mathcal{B}_1)$

Similarly, the update of $\mathcal{B}_1$ lead to a standard conical QCP such that we have $\tilde{N} = 2d(2M_\mathrm{d} + M_\mathrm{r} + N_\mathrm{s} + d) + 2(N_\mathrm{r}^2 + M_\mathrm{r}^2 + M_\mathrm{d}^2)$, $M = 1$ and $l_1 = 2dN_\mathrm{s}$, $l_2 = 2N_\mathrm{r}^2$ and $l_3 = 2d(N_\mathrm{s}+M_\mathrm{d}+d) + 2(N_\mathrm{r}^2 + M_\mathrm{r}^2 + M_\mathrm{d}^2)$.

The total arithmetic complexity of Algorithm 1 until convergence can be hence expressed as $I_\mathrm{in}(\mathcal{C}(\mathcal{B}_1) + \mathcal{C}(\mathcal{B}_2))$, where I_in represent the total number of computed inner-iterations until convergence.

E. INDEPENDENT OPTIMAL RECEIVER STRATEGY

As explained in Section I, the previous studies in the literature employ simplified assumptions on the hardware and self-interference cancellation accuracy, in order to facilitate a low complexity design. In this regard, the studies [12]–[19] employ an approach where the impact of inaccurate hardware behavior is modeled with the combination of, firstly, imposing an additional artifical power constraint in the received self-interference power, thereby implicitly modeling the limited dynamic range of the FD transceiver, and secondly, considering an additional noise variance at the receiver, imitating the impact of residual self-interference. Nevertheless, the aforementioned studies neglect the impact of the elaborated distortion amplification phenomena, due to the simplified assumptions. Moreover, this approach results in a mismatch in the evaluation of the relay transmit covariance, which is an essential part for designing the receiver equilizer filter at the destination, i.e., $\mathbf{C}$. In this regard, when employing the low complexity solutions offered in [12]–[19], we propose a receiver operation where the receiver signal-plus-interference-plus-noise-plus-distortion covariance, i.e., $\mathbb{E}\{\mathbf{y}\mathbf{y}^H\}$ can be estimated directly by sensing the collective received signal at the destination. In this case, when the low-dimensional effective equivalent channel between source and the destination, i.e.,

$$\mathbb{E}\Big\{\mathbf{y}(t)\mathbf{x}^H(t-\tau)\Big\} = \tilde{\mathbf{H}}_\mathrm{rd}\mathbf{G}\tilde{\mathbf{H}}_\mathrm{sr} =: \mathbf{H}_\mathrm{eq}, \tag{51}$$

is communicated to the destination, the optimal receiver strategy is directly obtained as $\mathbf{C}^\star = (\mathbb{E}\{\mathbf{y}\mathbf{y}^H\})^{-1} \times \mathbf{H}_\mathrm{eq}\mathbf{F}$. As we will observe, the proposed improvement leads to a notable performance gain for the aforementioned simplified design strategies, when hardware accuracy degrades. Please note that the aforementioned extension is not applicable to the relay receiver, as the received relay covariance is not directly measurable in the digital domain, and can be only accurately measured after the application of self-interference cancellation.

5. This includes reformulating the AL objective as epigraph form and also transforming the variables/constraints into the equivalent real-valued ones.

V. FRAMEWORK EXTENSION

In the previous section, we have developed an impairments-aware MSE minimization method for a single-user FD AF relaying system, following the PDD framework, where the design parameters can be obtained via a sequence of CQPs. Nevertheless, it is also of high interest to enable the proposed framework for different design metrics or system setups in order to be applied in diverse scenarios. In this section, we elaborate on how the proposed framework can be extended to the other popular scenarios and design metrics.

A. RATE MAXIMIZATION

Recalling (24) and (25), which express the desired source-destination signal dependency as well as the total received signal covariance at the destination, the achievable rate is given as

$$R(\mathbf{F},\mathbf{G}) = W\log_2\Big|\mathbf{I}_d + \mathbf{F}^H\mathbf{H}_\mathrm{eq}^H\Big(\Gamma^{-1}\Big)\mathbf{H}_\mathrm{eq}\mathbf{F}\Big|, \tag{52}$$

where W is the bandwidth, $R(\mathbf{F},\mathbf{G})$ is the achievable rate as a function of transmit precoder and the relay amplification matrix, and

$$\Gamma = \mathcal{M}_2(\mathbf{F},\mathbf{M}_\mathrm{out}) - \mathbf{H}_\mathrm{eq}\mathbf{F}\mathbf{F}^H\mathbf{H}_\mathrm{eq}^H \tag{53}$$

is the interference-plus-noise covariance at the destination.[6] The optimization problem for maximizing the achievable rate is hence formulated as

$$\max_{\mathbf{F},\mathbf{G}} \; R(\mathbf{F},\mathbf{G}), \quad \text{s.t. (27b)}, \tag{54}$$

which shows an intractable mathematical structure due to the non-convex objective, as well as the non-convex and coupled constraints. The following lemmas facilitate the application of the developed PDD-based solution also for the problem (54).

Lemma 1 (WMMSE Lemma [54], [55]): The achievable source-destination communication rate can be equivalently expressed as

$$R(\mathbf{F},\mathbf{G}) = -W\log_2|\mathbf{E}_\mathrm{mmse}(\mathbf{F},\mathbf{G})|, \tag{55}$$

where $\mathbf{E}_\mathrm{mmse} \coloneqq \mathbf{E}(\mathbf{F},\mathbf{G},\mathbf{C}^\star)$ represents the MMSE matrix, such that $\mathbf{C}^\star = \underset{\mathbf{C}}{\operatorname{argmin}}\, \mathbf{E}(\mathbf{F},\mathbf{G},\mathbf{C})$.

Proof: See [54, Lemma 2], [55, eq. (9)]. ∎

Lemma 2: Let $\mathbf{X} \in \mathbb{C}^{d\times d}$ be a positive definite matrix. The maximization of the term $-\log|\mathbf{X}|$ is equivalent to the maximization

$$\max_{\mathbf{X},\mathbf{S}} \; \log|\mathbf{S}| - \mathrm{tr}(\mathbf{S}\mathbf{X}) + d, \tag{56}$$

6. Please note that the expressions (52), (53) hold true assuming a Gaussian distribution of the transmit codewords and distortion terms as well as a sufficiently long coding block-length, and otherwise can be viewed as an approximation.

where $\mathbf{S} \in \mathcal{H}$ and we have $\mathbf{S}^\star = \mathbf{E}^{-1}$ at the optimality.

Proof: See [54, Lemma 2], [40]. ■

Utilizing the introduced lemmas and following the same steps as studied in Section IV, an equivalent optimization problem for rate maximization is formulated as

$$\min_{\mathcal{B}_1,\mathcal{B}_2,\mathbf{S}} \quad \mathrm{tr}\Big(\mathbf{L}_4^H \mathbf{S} \mathbf{L}_4\Big) - \log|\mathbf{S}| - d \tag{57a}$$

$$\text{s.t.} \quad \text{(34b)-(34f)}. \tag{57b}$$

Please note that the optimization problem (57) can be solved following the same procedure as for (34). In particular, while each of the updates over $\mathcal{B}_1$ and $\mathcal{B}_2$ remains as a CQP, the update of $\mathbf{S}$ can be calculated in closed-form from Lemma 2 as $\mathbf{S}^\star = (\mathbf{L}_4\mathbf{L}_4^H)^{-1}$, which leads to a negligible additional computational cost. Consequently, similar outer and inner iterations to the Algorithm 1 will be followed until convergence.

B. MULTI-USER RELAYING

A common FD relaying setup consists of a single source, e.g., a base station, which is communicating with multiple users in the downlink with the help of an FD relay. In order to extend the signal model in Section II to a multi-user scenario, it is sufficient to consider $\mathbf{s} = \lfloor \mathbf{s}_l \rfloor_{l\in\mathcal{L}}$, where $\mathbf{s}_l \in \mathbb{C}^d,\ l \in \mathcal{L}$ is the vector of transmit data symbols for the l-th user, $\mathcal{L}$ being the index set of all users. Similarly, $\mathbf{H}_{\mathrm{sd},l}$ and $\mathbf{H}_{\mathrm{rd},l}$ denote the source-destination and relay-destination channels corresponding to l-th user. The receiver linear filter and the total received signal covariance corresponding to the l-th user is denoted as $\mathcal{M}_{2,l}(\mathbf{F}, \mathbf{M}_{\mathrm{out}})$ and $\mathbf{C}_l$, respectively. In particular, the function $\mathcal{M}_{2,l}(\mathbf{F}, \mathbf{M}_{\mathrm{out}})$ is obtained similar to $\mathcal{M}_2(\mathbf{F}, \mathbf{M}_{\mathrm{out}})$ but by replacing the channels $\mathbf{H}_{\mathrm{sd},l}$ and $\mathbf{H}_{\mathrm{rd},l}$ for the respective user.

A multi-user sum-MSE minimization problem can be expressed as

$$\min_{\mathbf{F},\mathbf{G},\{\mathbf{C}_l\}} \ \sum_{l\in\mathcal{L}} \mathrm{tr}(\mathbf{E}_l(\mathbf{F},\mathbf{G},\mathbf{C}_l)) \ \text{ s.t. } (27\text{b}). \tag{58}$$

By utilizing similar steps as in Section IV, the above problem is equivalently formulated as

$$\min_{\mathcal{B}_1,\mathcal{B}_2} \ \sum_{l\in\mathcal{L}} \big\|\mathbf{L}_{4,l}\big\|_F^2 \tag{59a}$$

$$\text{s.t.} \ \ \mathcal{X} - \tilde{\mathcal{X}} = \mathbf{0}, \ \ \forall \mathcal{X} \in \Big\{\mathbf{F}, \mathbf{J}, \mathbf{L}_1, \Big\{\mathbf{L}_{2,l}\Big\} \ldots, \Big\{\mathbf{L}_{4,l}\Big\}\Big\} \tag{59b}$$

$$\bar{\mathcal{M}}_1\Big(\mathbf{F}, \tilde{\mathbf{F}}, \mathbf{J}, \tilde{\mathbf{J}}\Big) - \mathbf{L}_1\tilde{\mathbf{L}_1}^H = \mathbf{0}, \tag{59c}$$

$$\bar{\mathcal{M}}_{2,l}\Big(\mathbf{F}, \tilde{\mathbf{F}}, \mathbf{J}, \tilde{\mathbf{J}}\Big) - \mathbf{L}_{2,l}\tilde{\mathbf{L}_{2,l}}^H = \mathbf{0}, \ \ \forall l \in \mathcal{L}, \tag{59d}$$

$$\bar{\mathcal{M}}_{3,l}\Big(\mathbf{L}_{3,l}, \tilde{\mathbf{L}}_{3,l}, \mathbf{C}_l, \mathbf{L}_{6,l}\Big) - \mathbf{L}_{4,l}\tilde{\mathbf{L}}_{4,l}^H = \mathbf{0},$$

$$\mathbf{C}_l^H\mathbf{L}_{2,l} - \mathbf{L}_{3,l} = \mathbf{0}, \ \ \forall l \in \mathcal{L}, \tag{59e}$$

$$\mathbf{J} - \mathbf{G}\mathbf{L}_1 = \mathbf{0}, \ \ \mathbf{L}_5 - \tilde{\mathbf{H}}_{\mathrm{sr}}\mathbf{F} = \mathbf{0},$$

$$\mathbf{L}_{6,l} - \tilde{\mathbf{H}}_{\mathrm{rd},l}\mathbf{G}\mathbf{L}_5 = \mathbf{0}, \ \ \forall l \in \mathcal{L}, \tag{59f}$$

$$\|\mathbf{F}\|_F^2 \le P_{\mathrm{s,max}}/(1+\kappa_{\mathrm{s}}), \ \ \|\mathbf{J}\|_F^2 \le P_{\mathrm{r,max}}/(1+\kappa_{\mathrm{r}}), \tag{59g}$$

where

$$\mathcal{B}_1 \coloneqq \Big\{\mathbf{F}, \mathbf{G}, \{\mathbf{C}_l\}, \mathbf{J}, \tilde{\mathbf{L}}_1, \Big\{\tilde{\mathbf{L}}_{2,l}\Big\}, \ldots, \Big\{\tilde{\mathbf{L}}_{4,l}\Big\}\Big\}, \tag{60a}$$

$$\mathcal{B}_2 \coloneqq \Big\{\tilde{\mathbf{F}}, \tilde{\mathbf{J}}, \mathbf{L}_1, \big\{\mathbf{L}_{2,l}\big\}, \ldots, \big\{\mathbf{L}_{6,l}\big\}\Big\}. \tag{60b}$$

It is observed that the above problem holds a similar structure as in (34) and hence can be solved employing the proposed PDD-based updates.

C. PROTECTION AGAINST INSTANTANEOUS CHAIN SATURATION

The characterization of the hardware impairments have been conducted in [7], [28] via a statistical model, as elaborated in Section II-D. Nevertheless, in particular to the FD transceivers with a strong self-interference channel, it is also of interest to protect the chains from an instantaneous saturation; when an instantaneously high transmit or receive power lead to a destructive distortion of the desired signal and SIC [18], [19]. Please note that the constraints on power consumption are usually imposed on the average signal power, e.g., (27b). Nevertheless, this does not capture the instantaneous fluctuations of the signal which is related to the average signal power via the Peack-to-Average Power Ratio (PAPR) and depends on various parameters of the transmission strategy, e.g., transmit codeword and waveform design [56].

The imposition of the per-chain instantaneous signal strength at the relay can be hence expressed as

$$\mathbb{E}\Big\{|\mathcal{R}_l(\mathbf{r}_{\mathrm{in}})|^2\Big\} \le P_{\mathrm{I,rx}}/\omega_{\mathrm{rx}}, \ \ \forall l \in \{1\ldots M_{\mathrm{r}}\}, \tag{61a}$$

$$\mathbb{E}\Big\{|\mathcal{R}_l(\mathbf{r}_{\mathrm{out}})|^2\Big\} \le P_{\mathrm{I,tx}}/\omega_{\mathrm{tx}}, \ \ \forall l \in \{1\ldots N_{\mathrm{r}}\}, \tag{61b}$$

where $\mathcal{R}_l(\mathbf{X})$ returns the l-th row of the matrix $\mathbf{X}$, $P_{\mathrm{I,tx}}$ ($P_{\mathrm{I,rx}}$) denotes the maximum instantaneous per-chain transmit (receive) power, and ω_{tx} (ω_{rx}) is the transmit (receive) PAPR. In the transformed formulation in (34), the non-convex constraints (61a) are equivalently expressed as

$$\|\mathcal{R}_l(\mathbf{J})\|_2^2 \le P_{\mathrm{I,tx}}/(\omega_{\mathrm{rx}}(1+\kappa_{\mathrm{r}})), \ \ \forall l \in \{1\ldots M_{\mathrm{r}}\}, \tag{62}$$

$$\mathrm{diag}\Big(\mathbf{M}_{\mathrm{in},0} + \tilde{\mathbf{H}}_{\mathrm{rr}}\Big(\mathbf{J}\mathbf{J}^H + \kappa_{\mathrm{r}}\mathrm{diag}\Big(\mathbf{J}\mathbf{J}^H\Big)\Big)\tilde{\mathbf{H}}_{\mathrm{rr}}^H + \mathbf{C}_{\mathrm{rx,rr}}\mathrm{tr}\Big(\mathbf{C}_{\mathrm{tx,rr}}\Big(\mathbf{J}\mathbf{J}^H + \kappa_{\mathrm{r}}\mathrm{diag}\Big(\mathbf{J}\mathbf{J}^H\Big)\Big)\Big)\Big) \prec \mathbf{I}_{M_{\mathrm{r}}}P_{\mathrm{I,rx}}/(\omega_{\mathrm{tx}}(1+\beta_{\mathrm{r}})), \tag{63}$$

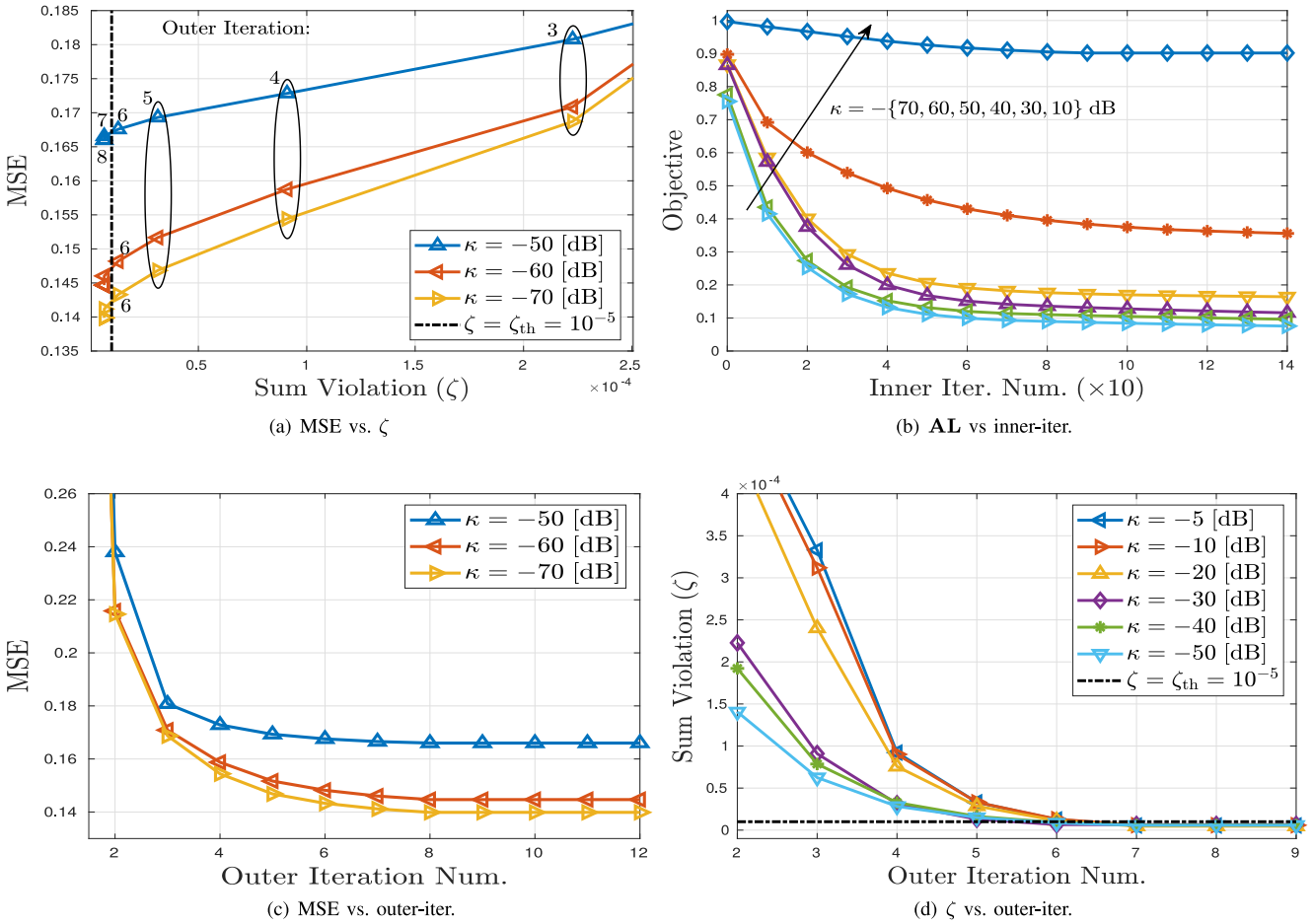
where (62), (63) represent the union of multiple jointly convex quadratic constraints over $\mathbf{J}$ and $\mathbf{F}$, and hence represent a jointly convex set. The corresponding MSE minimization problem, with the consideration of instantaneous per-chain power constraint is consequently obtained from (34) as

$$\min_{\mathcal{B}_1,\mathcal{B}_2} \ \ \|\mathbf{L}_4\|_F^2 \tag{64a}$$

$$\text{s.t.} \ \ \text{(34b)-(34f)}, \tag{64b}$$

$$\text{(62), (63)}. \tag{64c}$$

Please note that while (63) presents a coupled inequality constraint over $\mathbf{F}$ and $\mathbf{J}$, it is a jointly convex and decoupled



(a) MSE vs. ζ

(b) **AL** vs inner-iter.

(c) MSE vs. outer-iter.

(d) ζ vs. outer-iter.

FIGURE 2. **Convergence behaviour of Algorithm 1 in the outer and inner loop. (a) depicts MSE vs. ζ at different outer iteration number. In (b) the objective function is depicted at instances of the inner iterations for different levels of κ. Similarly, (c) and (d) depict MSE value vs. the number of outer iterations for different values of κ.**

constraint over the variable block $\mathcal{B}_1$, which simultaneously includes $\mathbf{F}, \mathbf{J}$. As a result, the above problem holds a similar structure to (34) and can be solved via the proposed PDD-based method, see Section IV-B.

VI. SIMULATION RESULTS

In this section we evaluate the behavior of the studied MIMO FD-AF relaying setup via numerical simulations, under the collective impact of impairments. In particular, we evaluate the proposed PDD-based method in Section IV, under the impact of hardware inaccuracies, in comparison with the available relevant methods in the literature. We assume that $\mathbf{H}_{\text{sr}}, \mathbf{H}_{\text{rd}}$ and $\mathbf{H}_{\text{sd}}$ follow an uncorrelated Rayleigh flat-fading model, where ρ_{sr}, ρ_{rd} and $\rho_{\text{sd}} \in^{+}$ represent the path loss. For the self-interference channel, we follow the characterization reported in [57]. In this respect we have $\mathbf{H}_{\text{rr}} \sim \mathcal{CN}(\sqrt{\frac{\rho_{\text{rr}} K_R}{1+K_R}}\mathbf{H}_0, \frac{1}{1+K_R}\mathbf{I}_{M_{\text{t}}} \otimes \mathbf{I}_{M_{\text{r}}})$ where ρ_{rr} represents the self-interference channel strength, $\mathbf{H}_0$ is a deterministic term[7] and K_R is the Rician coefficient. Unless explicitly stated, the following parameters are used as the default setup: $N := N_{\text{s}} = N_{\text{r}} = 4$, $M := M_{\text{r}} = M_{\text{d}} = 4$, $\rho_{\text{si}} = 0$ [dB], $\rho_{\text{sr}} = \rho_{\text{rd}} = -30$ [dB], $\rho_{\text{sd}} = -40$ [dB], $\sigma_{\text{n}}^2 = \sigma_{\text{nr}}^2 = \sigma_{\text{nd}}^2 = -40$ [dBm], $P_{\max} := P_{\text{s,max}} = P_{\text{r,max}} = 0$ [dBm], $\kappa = \kappa_{\text{s}} = \kappa_{\text{r}} = \beta_{\text{r}} = \beta_{\text{d}} = -40$ [dB], $K_R = 10$, $d = 2$. We follow the transmission protocol suggested in [7, Sec. II-B] for channel estimation, where $\Delta_{\mathcal{X}} = \mathbf{C}_{\text{rx},\mathcal{X}}^{1/2}\tilde{\Delta}_{\mathcal{X}}$ and

$$\mathbf{C}_{\text{rx},\mathcal{X}} = \frac{1}{2T}\left(\frac{\sigma_{\text{n}}^2}{P_{\max}}\mathbf{I}_M + \frac{2\kappa}{N}\tilde{\mathbf{H}}_{\mathcal{X}}\tilde{\mathbf{H}}_{\mathcal{X}}^H + \frac{2\kappa}{N}\text{diag}\left(\tilde{\mathbf{H}}_{\mathcal{X}}\tilde{\mathbf{H}}_{\mathcal{X}}^H\right)\right), \quad \mathcal{X} \in \{\text{sr}, \text{rr}, \text{sd}, \text{rd}\}, \tag{65}$$

where T is the number of symbol transmissions dedicated for CSI estimation during a channel coherence time interval. Please note that the chosen parameter setup corresponds to the signal-to-(noise or distortion or interference) level of 10 dB, which will be then modified to study a wide range, in each case by changing the corresponding parameters.

Please note that (65) relates the CSI error statistics to the instantaneous channel, statistics of noise, transmit and receive distortions and the number of transmissions dedicated for CSI estimation. For each channel realization the resulting performance in terms of the communication MSE

7. Please note that this part corresponds to the mutual coupling or direct Tx-Rx interference path, which can be viewed as a line-of-sight component. For simplicity, we choose $\mathbf{H}_0$ as a matrix of all-1 elements.

is evaluated by employing different design strategies and for various system parameters. The overall system performance is then evaluated via Monte-Carlo simulations by employing 100 channel realizations.[8]

A. ALGORITHM ANALYSIS

The proposed PDD-based algorithm (Algorithm 1) is a dual-loop procedure such that the inner loop is dedicated to the minimization of the AL function, whereas the outer loop is responsible for controlling the penalty/dual parameters and thereby pushing the constraint violations to zero. In this respect, the convergence of both sequences are essential, in order to ensure that the obtained solution is feasible and satisfies the first order optimality conditions [47, Sec. III]. In Figs. 2-c and 2-d, the resulting MSE and constraint violation function, i.e., $\zeta(\mathcal{B}_1, \mathcal{B}_2)$ are depicted for the iterations of the outer loop. The joint MSE-ζ convergence of the outer loop is also depicted in Fig. 2-a. It is observed that the outer loop converges in about 8 iterations, leading to the feasible and stable solution, with the numerical tolerance margin of $\zeta_{\text{th}} = 10^{-5}$. On the other hand, the convergence of the inner loop is depicted in Fig. 2-b, where the value of the AL function is depicted over the iterations of the inner loop, corresponding to the first outer iteration. As expected, a monotonic decrease of the AL is observed, leading to an eventual convergence. Please note that the value of the AL function reaches a smaller value compared to the actual MSE value as observed from Fig. 2-b. This is since when the penalty parameters are not large enough, as they are small in the first outer iteration and consequently grow to push the violations to zero, the AL function takes advantage of the tolerated violation margin and moves towards minimizing the AL value. Such mismatch will be later compensated during the iterations of the outer loop by adjusting the penalty/dual parameters to enforce violations to zero, see Fig. 2-d.

B. PERFORMANCE COMPARISON

In this part we evaluate the performance of the proposed design for various system parameters, in comparison with the available designs in the literature.

1) AVAILABLE DESIGN APPROACHES

We divide the relevant available literature on the FD-AF relaying design into three main approaches. Firstly, as considered in [9], [10], the self-interference cancellation is purely relegated to the relay receiver end, via a combined time domain analog/digital cancellation techniques. The aforementioned approach considers no limitation on the self-interference process, i.e., assuming a perfect hardware, CSI, and self-interference cancellation.[9] Secondly, the SIC is purely done via transmit beamforming at the null space of the relay receive antennas, e.g., [12]–[17], hence imposing a zero interference power constraint for transmit beamforming design, i.e., $P_{\text{intf}} \leq 0$, where P_{intf} is the received self-interference power at the FD transceiver. Finally, as a generalization of the aforementioned extreme approaches, a combined transmit beamforming and analog/digital cancellation at the receiver is considered in [18], [19], [49]. In the aforementioned case it is assumed that the received self-interference power should not exceed a certain threshold, i.e., $P_{\text{intf}} \leq P_{\text{th}}$, where at the same time, the relay receiver noise covariance is increased with $\sigma_{\text{si}}^2 \mathbf{I}_M$, which imitates the impact of residual self-interference. In all of the aforementioned cases, due to the simplified assumptions on the hardware and residual self-interference statistics, the impact of the discussed distortion-amplification loop is not taken into account.

In view of the aforementioned literature, we employ the following comparison benchmarks.

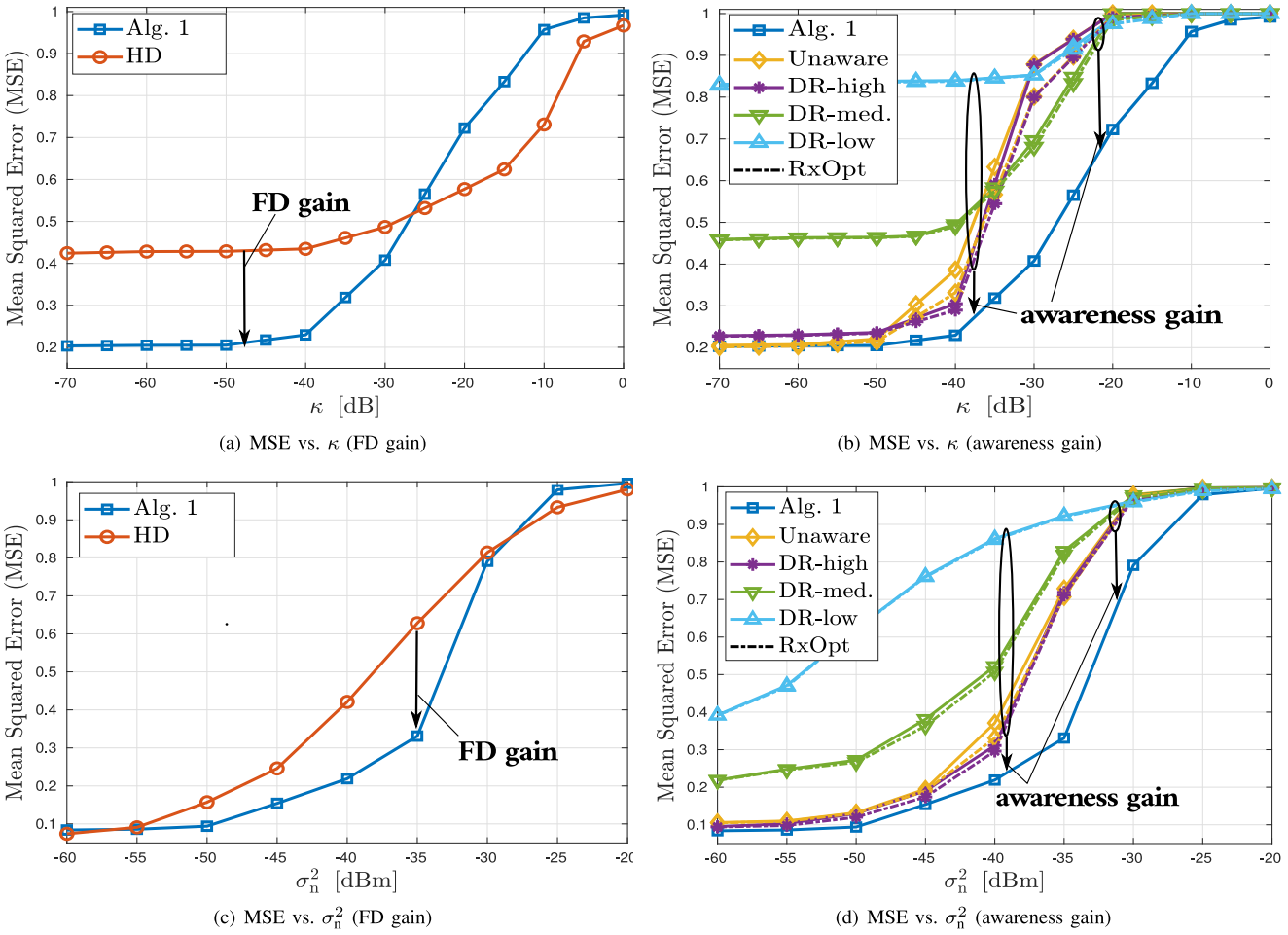
- Algorithm 1: which indicates the proposed Algorithm 1, where the impact of hardware and CSI inaccuracies are incorporated in the design process,
- HD: corresponding HD system performance, employing equal system and communication parameters as the defined FD setup,
- Unaware [9], [10]: a design algorithm assuming a perfect hardware and self-interference cancellation, where all sources of impairments are ignored,
- DR-high/med./low [12]–[19], [49]: a design algorithm assuming $P_{\text{th}} = \{10^2, 10^4, 10^6\} \times \sigma_{\text{n}}^2$ and $\sigma_{\text{si}}^2 = \sigma_{\text{n}}^2/\{10, 1, 0.1\}$, corresponding to an implicit assumption for a transceiver with a high, medium, and low hardware and self-interference cancellation accuracy.
- RxOpt [18], [19]: a design algorithm where the previous benchmarks are implemented with an optimum receiver strategy at the destination, following the proposed extension in Section IV-E.

2) NUMERICAL EVALUATION

In Fig. 3-a and Fig. 3-b the average communication MSE is depicted for different levels of hardware accuracy, i.e., κ. As expected, a higher impairments level results in a larger MSE for all system and design strategies. Moreover, it is observed that the HD system falls short at a system with an adequately large dynamic range, while offering a better performance as the system dynamic range degrades, see Fig. 3-a. This is due to the strong self-interference channel which significantly deteriorates the performance in an FD relaying setup as κ increases. Nevertheless, it is observed from Fig. 3-b that the application of an impairments-aware design, i.e., Algorithm 1, leads to a significantly lower MSE, compared to the designs with simplified modeling. In particular, the design assuming a perfect hardware that ignores all sources of impairments, i.e., 'Unaware', leads to a severely larger MSE compared to Algorithm 1 for a large κ region, while coincides with the optimum performance as the dynamic

8. For each channel realization, the system performance is evaluated for all of the reported system parameters in order to obtain a reliable and smooth comparison.

9. This approach is equivalent to ignoring the impact of SIC in the design of transmit and receive strategies, as it has been usual in the earlier literature.



(a) MSE vs. κ (FD gain)

(b) MSE vs. κ (awareness gain)

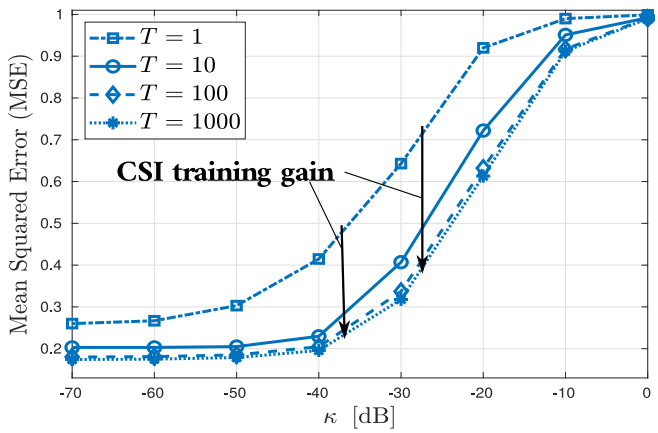
(c) MSE vs. σ_n^2 (FD gain)

(d) MSE vs. σ_n^2 (awareness gain)

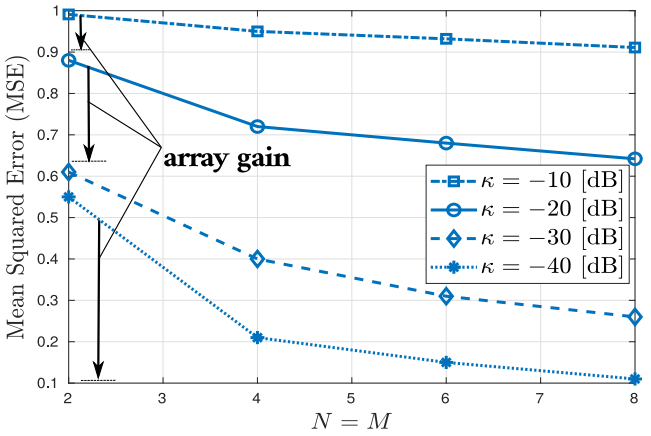
FIGURE 3. **Performance comparison of different benchmarks under various system conditions.**

range is large. Moreover, it is observed that the simplified design strategies, i.e., DR-high/med./low, when accompanied with the proposed improvement in Section IV-E, lead to a close performance to Algorithm 1, when applied at the *dedicated parameter region*. For instance, 'DR-low, RxOpt' reaches close to the optimum performance, when applied at the region corresponding to large values of κ. Similarly, this observation is also true for the algorithms 'DR-med' and 'DR-high' when applied together with 'RxOpt' for a medium and low values of κ. As expected, the obtained improvement via the proposed amendment in Section IV-E becomes notable for a medium and low range of hardware accuracy, while it vanishes as κ grows very large or very small.

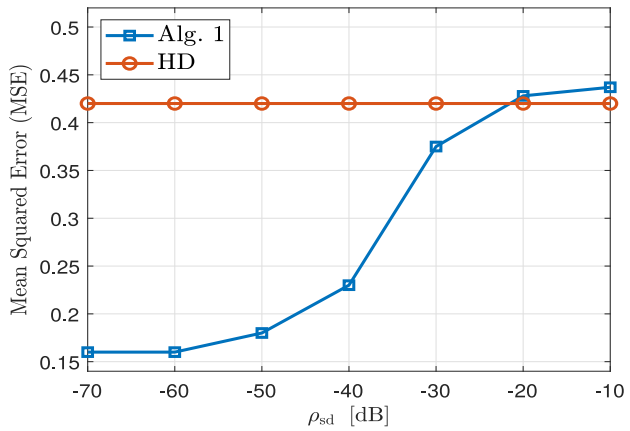
In Fig. 3-c and Fig. 3-d the average communication MSE is depicted for different levels of thermal noise, i.e., σ_n^2. Similar to the impact of κ, a higher level of thermal noise results in a larger MSE for all system and design strategies. It is observed from Fig. 3-c that the HD setup outperforms the FD setup at both very high and very low signal-to-noise ratio (SNR) regions, whereas the FD system shows better performance at the intermediate to high SNR conditions. This is since, similar to the observation from Fig. 3-a, the FD system makes use of the simultaneous transmission and reception and outperforms the HD setup, when the impact of hardware impairments is not dominant. Nevertheless, for the very high SNR region, the system performance will be dominated by the impact of hardware impairments, which limits the performance of the FD system in comparison to the HD setup. On the other hand, in contrast to the similar studies which ignore or simplify the channel estimation process in the perofmance evaluation, i.e., assuming a perfect CSI or a fixed CSI error statistics which is not impacted by noise [18], [19], [49], it is observed that the FD system performance reaches below that of the HD system for the high levels of thermal noise. This is since, a higher level of thermal noise also degrades the CSI accuracy at the self-interference channel, which deteriorates the performance of the FD system in comparison to the HD setup. Similar to Fig. 3-b, it is observed that the application of an impairments-aware design is essential for the studied FD system, also as the thermal noise variance, and hence the intensity of CSI error at the self-interference channel, increases.



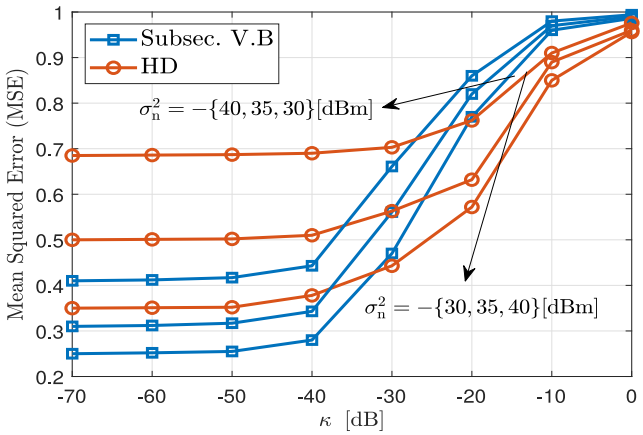
(a) MSE vs. T (CSI training gain)



(b) MSE vs. $M = N$ (array gain)



(c) MSE vs. ρ_{sd} (direct inrerf.)



(d) MSE vs. κ

FIGURE 4. Performance comparison of different scenarios, the impact of CSI training (a), number of antennas indicating the array gain (b), the impact of direct source-destination interference (c) and the multi-user relaying scenario (d).

A more clear impact of the CSI estimation accuracy is depicted in Fig. 4-a, where the resulting communication MSE is evaluated for different values of T, i.e., the transmission period dedicated for CSI training. Please note that the choice of T should be made with the consideration of the channel coherence time as well as the available overall communication resources. As expected, it is observed that the resulting MSE is reduced as T increases. Moreover, it is observed that the system performance greatly benefits from a higher T for $1 \leq T \leq 10$, specially in a medium range of κ when CSI accuracy is a limiting factor, but almost saturates when $T \geq 100$.

In Fig. 4-b, the average communication MSE is depicted for different array dimensions, i.e., $N \coloneqq N_{\mathrm{s}} = N_{\mathrm{r}}$, $M \coloneqq M_{\mathrm{r}} = M_{\mathrm{d}}$, and for different impairments level. It is observed that as the array dimension increases, the system can take advantage of the increased spatial degree of freedom to reduce MSE. Nevertheless, due to the almost non-correlated nature of the impairments, the array gain is not significant for an FD relaying system with a high level of impairments, e.g., $\kappa = -10$ [dB], also due to the deteriorating impact of self-interference and CSI inaccuracy in a larger antenna setup. Nevertheless, an increase in the number of antennas leads to a significantly better MSE performance for a system with a more accurate hardware, as observed by Fig. 4-b.

The impact of direct source-destination interference is depicted in Fig. 4-c. Please note that the direct interference, in addition to the self-interference at the relay, is emerging inherently for the FD relaying setup, due to the coexistence of the relay reception and transmission. As expected, the communication performance degrades as the interference intensity increases for the FD relaying channel, and goes below that of HD relaying. Please note that in an HD system the direct interference does not exist, where the source-relay and the relay-destination transmissions are separated at different resources. As a result, the intensity of the direct interference channel does not impact the performance of HD relaying system.

In Fig. 4-d, the performance of the multi-user relaying scenario is evaluated, following the proposed design extension in Section V-B under different levels of hardware distortions and thermal noise variance. Please note that the extension in

Section V-B presents the same PDD-based procedure as in Algorithm 1, when the MSE is considered at the receiver for a group of multiple-antenna users. In particular, we consider the scenario where $|\mathcal{L}| = 2$, $M_{\mathrm{d},l} = 2$, $\forall l \in \mathcal{L}$, where $M_{\mathrm{d},l}$ is the number of antennas at each receiver. The parameters regarding the transmit power budget at the source and the relay nodes, as well as the experienced path loss are defined similar to defined default values for the single-user evaluations in Fig. 3. As expected, it is observed that a higher thermal noise or hardware distortion intensity degrade the performance of the both FD and HD systems. Please note that in contrast to the single user MIMO case, due to the separated receiver processing, the multi-user scenario leads to a naturally higher value of error due to the additional inter-user interference. Similar to Fig. 3, it is observed that while for a system with a high chain resolution, i.e., low κ value, the FD system leads to a significant gain in comparison to the HD benchmark, the FD gain rapidly disappears as the hardware accuracy degrades. Moreover, it is observed that the value of κ for which the FD relaying scenario surpasses the performance of an HD system shifts slightly to a smaller value as the thermal noise variance becomes smaller. This is since, for a system with a higher noise floor, a higher level of hardware distortions can be tolerated, hence leading to a performance advantage for an FD relaying system. Similar to Fig. 3, it can be concluded that the use of the studied multiuser AF FD relaying is not motivated for a relay with a high hardware distortions due to the impact of distortion-amplification loop, as well as the deteriorating impact of the direct source-destination interference.

VII. CONCLUSION

The impact of hardware inaccuracies is of particular importance for an FD transceiver, due to the high strength of the self-interference channel. In particular, for an FD-AF relaying system, such impact is more pronounced due to the *inter-dependency* of the relay transmit covariance, as well as the residual self-interference covariance, which results in a *distortion-amplification loop* effect. In this work, we have analyzed and optimized the MSE performance of a MIMO FD-AF relaying system under the impact of collective sources of hardware impairments. The proposed optimization framework converges to a stationary point by solving a sequence of convex quadratic programs, thereby enjoying a favorable arithmetic complexity as problem dimensions increase. Numerical simulations show the high significance of a distortion-aware design and system modeling, specially when hardware dynamic range is not high.

REFERENCES

[1] D. Bharadia and S. Katti, "Full duplex MIMO radios," in *Proc. 11th USENIX Conf. Netw. Syst. Design Implement.*, Berkeley, CA, USA, 2014, pp. 359–372.

[2] Y. Hua, P. Liang, Y. Ma, A. C. Cirik, and Q. Gao, "A method for broadband full-duplex MIMO radio," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 793–796, Dec. 2011.

[3] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. ACM SIGCOMM*, Aug. 2013, pp. 375–386.

[4] A. K. Khandani, "Two-way (true full-duplex) wireless," in *Proc. 13th Can. Workshop Inf. Theory (CWIT)*, Sep. 2013, pp. 33–38.

[5] S. Hong *et al.*, "Applications of self-interference cancellation in 5G and beyond," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 114–121, Feb. 2014.

[6] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[7] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1541–1553, Sep. 2012.

[8] X. Xia, D. Zhang, K. Xu, W. Ma, and Y. Xu, "Hardware impairments aware transceiver for full-duplex massive MIMO relaying," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6565–6580, Dec. 2015.

[9] K. Lee, H. M. Kwon, M. Jo, H. Park, and Y. H. Lee, "MMSE-based optimal design of full-duplex relay system," in *Proc. IEEE Veh. Technol. Conf. (VTC Fall)*, Sep. 2012, pp. 1–5.

[10] O. Somekh, O. Simeone, H. V. Poor, and S. Shamai, "Cellular systems with full-duplex amplify-and-forward relaying and cooperative base-stations," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 16–20.

[11] Z. Wen, X. Liu, N. C. Beaulieu, R. Wang, and S. Wang, "Joint source and relay beamforming design for full-duplex MIMO AF relay SWIPT systems," *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 320–323, Feb. 2016.

[12] H. A. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P. J. Smith, "Low-complexity end-to-end performance optimization in MIMO full-duplex relay systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 913–927, Feb. 2014.

[13] D. Choi and D. Park, "Effective self interference cancellation in full duplex relay systems," *Electron. Lett.*, vol. 48, no. 2, pp. 129–130, Jan. 2012.

[14] B. Chun and H. Park, "A spatial-domain joint-nulling method of self-interference in full-duplex relays," *IEEE Commun. Lett.*, vol. 16, no. 4, pp. 436–438, Apr. 2012.

[15] C. Y. Shang, P. J. Smith, G. K. Woodward, and H. A. Suraweera, "Linear transceivers for full duplex MIMO relays," in *Proc. Aust. Commun. Theory Workshop (AusCTW)*, Feb. 2014, pp. 11–16.

[16] U. Ugurlu, T. Riihonen, and R. Wichman, "Optimized in-band full-duplex MIMO relay under single-stream transmission," *IEEE Trans. Veh. Technol.*, vol. 65, no. 2, pp. 155–168, Jan. 2016.

[17] Q. Shi, M. Hong, X. Gao, E. Song, Y. Cai, and W. Xu, "Joint source-relay design for full-duplex MIMO AF relay systems," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6118–6131, Dec. 2016.

[18] O. Taghizadeh, J. Zhang, and M. Haardt, "Transmit beamforming aided amplify-and-forward MIMO full-duplex relaying with limited dynamic range," *Signal Process.*, vol. 127, pp. 266–281, Oct. 2016.

[19] Y. Y. Kang, B.-J. Kwak, and J. H. Cho, "An optimal full-duplex AF relay for joint analog and digital domain self-interference cancellation," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2758–2772, Aug. 2014.

[20] H. Shen, W. Xu, and C. Zhao, "Transceiver optimization for full-duplex massive MIMO AF relaying with direct link," *IEEE Access*, vol. 4, pp. 8857–8864, 2016.

[21] T. Guo and B. Wang, "Joint transceiver beamforming design for end-to-end optimization in full-duplex MIMO relay system with self-interference," *IEEE Commun. Lett.*, vol. 20, no. 9, pp. 1733–1736, Sep. 2016.

[22] O. Taghizadeh and R. Mathar, "Cooperative strategies for distributed full-duplex relay networks with limited dynamic range," in *Proc. IEEE Int. Conf. Wireless Space Extreme Environ. (WiSEE)*, Noordwijk, The Netherlands, 2014, pp. 1–7.

[23] O. Taghizadeh, M. Rothe, A. C. Cirik, and R. Mathar, "Distortion-loop analysis for full-duplex amplify-and-forward relaying in cooperative multicast scenarios," in *Proc. 9th Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, Cairns, QLD, Australia, Dec. 2015, pp. 107–115.

[24] T. Riihonen, S. Werner, and R. Wichman, "Optimized gain control for single-frequency relaying with loop interference," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2801–2806, Jun. 2009.

[25] L. J. Rodriguez, N. H. Tran, and T. Le-Ngoc, "Optimal power allocation and capacity of full-duplex AF relaying under residual self-interference," *IEEE Wireless Commun. Lett.*, vol. 3, no. 2, pp. 233–236, Apr. 2014.

[26] C. Dang, L. J. Rodríguez, N. H. Tran, S. Shelly, and S. Sastry, "Secrecy capacity of the full-duplex AF relay wire-tap channel under residual self-interference," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2015, pp. 99–104.

[27] X. Cheng, B. Yu, X. Cheng, and L. Yang, "Two-way full-duplex amplify-and-forward relaying," in *Proc. IEEE Military Commun. Conf. (MILCOM)*, Nov. 2013, pp. 1–6.

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

[29] C.-T. Lin, F.-S. Tseng, and W.-R. Wu, "MMSE transceiver design for full-duplex MIMO relay systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 6849–6861, Aug. 2017.

[30] C.-T. Lin and W.-R. Wu, "Linear transceiver design for full-duplex MIMO relay systems: A non-iterative approach," *IEEE Wireless Commun. Lett.*, vol. 6, no. 4, pp. 518–521, Aug. 2017.

[31] O. Taghizadeh, A. C. Cirik, and R. Mathar, "Hardware impairments aware transceiver design for full-duplex amplify-and-forward MIMO relaying," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1644–1659, Mar. 2018.

[32] Z. Wen, S. Wang, X. Liu, and J. Zou, "Joint relay–user beamforming design in a full-duplex two-way relay channel," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2874–2879, Mar. 2017.

[33] A. C. Cirik, Y. Rong, and Y. Hua, "Achievable rates of full-duplex MIMO radios in fast fading channels with imperfect channel estimation," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3874–3886, Aug. 2014.

[34] S. Zarei, W. H. Gerstacker, J. Aulin, and R. Schober, "Multi-cell massive MIMO systems with hardware impairments: Uplink-downlink duality and downlink precoding," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5115–5130, Aug. 2017.

[35] E. Antonio-Rodriguez, R. Lopez-Valcarce, T. Riihonen, S. Werner, and R. Wichman, "SINR optimization in wideband full-duplex MIMO relays under limited dynamic range," in *Proc. IEEE 8th Sensor Array Multichannel Signal Process. Workshop (SAM)*, Jun. 2014, pp. 177–180.

[36] B. Zhong, D. Zhang, Z. Zhang, Z. Pan, K. Long, and A. Vasilakos, "Opportunistic full-duplex relay selection for decode-and-forward cooperative networks over Rayleigh fading channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 5717–5722.

[37] E. Antonio-Rodriguez, R. Lopez-Valcarce, T. Riihonen, S. Werner, and R. Wichman, "Subspace-constrained SINR optimization in MIMO full-duplex relays under limited dynamic range," in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2015, pp. 281–285.

[38] H. Iimori, G. T. F. de Abreu, and G. C. Alexandropoulos, "MIMO beamforming schemes for hybrid SIC FD radios with imperfect hardware and CSI," *IEEE Trans. Wireless Commun.*, vol. 18, no. 10, pp. 4816–4830, Oct. 2019.

[39] O. Taghizadeh, P. Neuhaus, R. Mathar, and G. Fettweis, "Secrecy energy efficiency of MIMOME wiretap channels with full-duplex jamming," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5588–5603, Aug. 2019.

[40] O. Taghizadeh, V. Radhakrishnan, A. C. Cirik, R. Mathar, and L. Lampe, "Hardware impairments aware transceiver design for bidirectional full-duplex MIMO OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7450–7464, Aug. 2018.

[41] G. Santella and F. Mazzenga, "A hybrid analytical-simulation procedure for performance evaluation in M-QAM-OFDM schemes in presence of nonlinear distortions," *IEEE Trans. Veh. Technol.*, vol. 47, no. 1, pp. 142–151, Feb. 1998.

[42] W. Namgoong, "Modeling and analysis of nonlinearities and mismatches in AC-coupled direct-conversion receiver," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 163–173, Jan. 2005.

[43] H. Suzuki, T. V. A. Tran, I. B. Collings, G. Daniels, and M. Hedley, "Transmitter noise effect on the performance of a MIMO-OFDM hardware implementation achieving improved coverage," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 6, pp. 867–876, Aug. 2008.

[44] D. P. M. Osorio, E. E. B. Olivo, H. Alves, J. C. S. S. Filho, and M. Latva-aho, "Exploiting the direct link in full-duplex amplify-and-forward relaying networks," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1766–1770, Oct. 2015.

[45] G. E. Bottomley, T. Ottosson, and Y. P. E. Wang, "A generalized RAKE receiver for interference suppression," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 8, pp. 1536–1545, Aug. 2000.

[46] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3074–3085, Sep. 2011.

[47] Q. Shi and M. Hong, "Penalty dual decomposition method for nonsmooth nonconvex optimization—Part I: Algorithms and convergence analysis," in *IEEE Trans. Signal Process.*, vol. 68, no. 1, pp. 4108–4122, Jun. 2020, doi: 10.1109/TSP.2020.3001906.

[48] Q. Shi, M. Hong, X. Fu, and T.-H. Chang, "Penalty dual decomposition method for nonsmooth nonconvex optimization—Part II: Applications," *IEEE Trans. Signal Process.*, vol. 68, no. 1, pp. 4242–4257, Jun. 2020, doi: 10.1109/TSP.2020.3001397.

[49] Y. Cai, Y. Xu, Q. Shi, B. Champagne, and L. Hanzo, "Robust joint hybrid transceiver design for millimeter wave full-duplex MIMO relay systems," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1199–1215, Feb. 2019.

[50] Q. Shi and M. Hong, "Spectral efficiency optimization for millimeter wave multiuser MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 455–468, Jun. 2018.

[51] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, 2013.

[52] D. P. Bertsekas, *Nonlinear Programming*. Belmont, CA, USA: Athena Sci., 1999.

[53] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, vol. 2. Philadelphia, PA, USA: SIAM, 2001,

[54] J. Jose, N. Prasad, M. Khojastepour, and S. Rangarajan, "On robust weighted-sum rate maximization in MIMO interference networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2011, pp. 1–6.

[55] 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 Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[56] R. Zayani, H. Shaiek, and D. Roviras, "PAPR-aware massive MIMO-OFDM downlink," *IEEE Access*, vol. 7, pp. 25474–25484, 2019.

[57] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

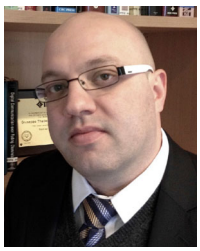
[58] O. Taghizadeh, S. Stanczak, H. Iimori, and G. Abreu, "Full-duplex AF MIMO Relaying: Impairments aware design and performance analysis," in *Proc. IEEE Global Commun. Conf. (Globecom SPC)*, Dec. 2020, pp. 1–6.

[59] J. J. Bussgang, *Crosscorrelation Functions of Amplitude-Distorted Gaussian Signals*. Cambridge, MA, USA: MIT, 1952.

OMID TAGHIZADEH (Member, IEEE) received the M.Sc. degree (with Distinction) in communications and signal processing from the Ilmenau University of Technology, Ilmenau, Germany, and the Ph.D. degree (with Distinction) from RWTH Aachen University, Aachen, Germany. He is currently a Research Associate with the Network Information Theory Group, Technical University of Berlin. His research interests include full-duplex wireless systems, MIMO communications, machine learning, optimization, and resource allocation in wireless networks.

SLAWOMIR STANCZAK (Senior Member, IEEE) received the Electrical Engineering degree with specialization in control theory from the Wroclaw University of Technology, Wroclaw, Poland, and the Technical University of Berlin (TU Berlin), Berlin, Germany, and the Dipl.-Ing. degree and the Dr.-Ing. degree *(summa cum laude)* in electrical engineering from TU Berlin, in 1998 and 2003, respectively, and the Habilitation degree (*venialegendi*) in 2006. Since 2015, he has been a Full Professor for network information theory with TU Berlin and the Head of the Wireless Communications and Networks Department, Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Berlin. He has been involved in research and development activities in wireless communications since 1997. In 2004 and 2007, he was a Visiting Professor with RWTH Aachen University and he was a Visiting Scientist with Stanford University, Stanford, CA, USA, in 2008. He has coauthored two books and more than 200 peer-reviewed journal articles and conference papers in the area of information theory, wireless communications, signal processing, and machine learning. He was a recipient of the Research Fellowships from the German Research Foundation and the Best Paper Award from the German Communication Engineering Society in 2014. He was a Co-Chair of the 14th International Workshop on Signal Processing Advances in Wireless Communications in 2013. From 2009 to 2011, he was an Associate Editor of the *European Transactions for Telecommunications* (information theory) and an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2012 to 2015. Since February 2018, he has been the Chairman of the ITU-T Focus Group on machine learning for future networks including 5G.

HIROKI IIMORI (Graduate Student Member, IEEE) received the B.Eng. degree in electrical and electronic engineering and the M.Eng. degree (Hons.) in advanced electrical, electronic and computer systems from Ritsumeikan University, Kyoto, Japan, in 2017 and 2019, respectively. He is currently pursuing the Ph.D. degree with a focus on electrical and computer engineering with Jacobs University Bremen, Germany. He was a Visiting Scholar with the Electrical and Computer Engineering Department, University of Toronto, Toronto, ON, Canada. His research interests lie in the field of optimization theory, wireless communications, and signal processing. He is a recipient of the YKK Doctoral Fellowship awarded by the Yoshida Scholarship Foundation, Japan.



GIUSEPPE THADEU FREITAS DE ABREU (Senior Member, IEEE) received the B.Eng. degree in electrical engineering and the specialization Latu Sensu degree in telecommunications engineering from the Universidade Federal da Bahia, Salvador, Bahia, Brazil, in 1996 and 1997, respectively, and the M.Eng. and D.Eng. degrees in physics, electrical and computer engineering from the Yokohama National University, Japan, in March 2001 and March 2004, respectively. He was a Postdoctoral Fellow and a later Adjunct Professor (Docent) on Statistical Signal Processing and Communications Theory with the Department of Electrical and Information Engineering, University of Oulu, Finland, from 2004 to 2006 and from 2006 to 2011, respectively. Since 2011, he has been a Professor of Electrical Engineering with Jacobs University Bremen, Germany. From April 2015 to August 2018, he also simultaneously held a Full Professorship with the Department of Computer and Electrical Engineering, Ritsumeikan University, Japan. His research interest spans a wide range of topics within communications and signal processing, including communications theory, estimation theory, statistical modeling, wireless localization, cognitive radio, wireless security, MIMO systems, ultra-wideband and millimeter wave communications, full-duplex and cognitive radio, compressive sensing, energy harvesting networks, random networks, and connected vehicles networks. He was the co-recipient of best paper awards at several international conferences. He received the Uenohara Award by Tokyo University in 2000 for his Master's Thesis work. He was also awarded the prestigious JSPS, the Heiwa Nakajima, and the NICT Fellowships in 2010, 2013, and 2015, respectively. He served as an Associate Editor of TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2009 to 2014, TRANSACTIONS ON COMMUNICATIONS from 2014 to 2017. He serves currently as an Executive Editor of TRANSACTIONS ON WIRELESS COMMUNICATIONS.