

Nonlinear MIMO Transceivers Improve Wireless-Powered and Self-Interference-Aided Relaying

Lei Zhang, Yunlong Cai, Minjian Zhao, Benoit Champagne, and Lajos Hanzo

Abstract—This paper investigates the design of robust nonlinear transceivers conceived for multiple-input multiple-output (MIMO) full-duplex (FD) wireless-powered relay (WPR) networks in the face of realistic imperfect channel state information (CSI). A novel self-energy recycling aided relaying protocol is employed, whereby the relay node benefits from energy harvesting (EH) gleaned from the self-interfering link in addition to its primary energy. The proposed nonlinear transceiver relies on a Tomlinson-Harashima (TH) precoder along with an amplify-and-forward (AF) relaying matrix and a linear receiver, where the TH precoder is composed of a feedback matrix and a source precoding matrix. Two different criteria are considered for the robust design of the nonlinear transceiver in the presence of channel estimation errors modeled by the Gaussian distribution. The first one aims for minimizing the mean-squared-error (MSE) at the destination subject to a transmit power constraint at the source and an EH constraint at the relay. The resultant optimization problem is converted to four subproblems and solved via an alternating optimization (AO) algorithm that iteratively updates the transceiver coefficients by sequentially addressing each subproblem, while keeping the other matrix variables fixed. Specifically, the optimal linear receiver matrix is derived in closed form; the AF relaying matrix is obtained via convex optimization; an iterative algorithm based on the constrained concave convex procedure (CCCP) is developed for optimizing the source's precoding matrix; finally, the feedback matrix of the TH precoder is derived with the aid of the Lagrangian multiplier method. The second design criterion aims for minimizing the transmit power at the source under both MSE and EH constraints. Similarly, an AO-based iterative algorithm is proposed for solving this problem. Our simulation results show that the robust design advocated is capable of alleviating the effects of CSI errors, hence improving the robustness of the system over that of the corresponding linear designs.

Index Terms—Transceiver design, Tomlinson-Harashima precoding, full-duplex, energy harvesting, MIMO relay

I. INTRODUCTION

At the current state-of-the-art wireless devices are usually powered by batteries, hence have a limited operating time.

The work of Y. Cai was supported in part by the National Natural Science Foundation of China under Grant 61471319 and the Fundamental Research Funds for the Central Universities. The work of M. Zhao was supported by the National Natural Science Foundation of China under Grant 91538103. (Corresponding authors: Yunlong Cai; Minjian Zhao.)

L. Zhang, Y. Cai, and M. Zhao are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: bestleileisara@zju.edu.cn; ylcai@zju.edu.cn; mjzhao@zju.edu.cn).

B. Champagne is with the Department of Electrical and Computer Engineering, McGill University, Montreal, Canada. (Email: benoit.champagne@mcgill.ca).

L. Hanzo is with the Department of Electronics and Computer Science, University of Southampton, Southampton, UK. (Email: lh@ecs.soton.ac.uk).

Hence, energy harvesting (EH) is considered as an attractive technique of prolonging the recharge period of energy-constrained wireless devices [1]. However, the traditional EH techniques rely on natural energy sources (solar, vibration, wind, etc.), which cannot be controlled and are not always available. Hence, EH from the ambient radio-frequency (RF) electromagnetic signals has emerged as a promising techniques of powering energy-constrained wireless networks. Consequently, RF-EH networks have found their way into wireless sensor networks, wireless charging systems and wireless body area networks [2]–[5]. In this emerging EH framework, the ambient RF signals are exploited as a precious source of energy for recharging ubiquitous wireless devices. As a result, the study of wirelessly powered communication networks (WPCN), in which wireless devices are powered by the RF energy harvested from wireless transmissions has drawn considerable attention [6]–[10]. More specifically, WPCN is a new networking paradigm where the battery of wireless communication devices can be remotely replenished by means of microwave wireless power transfer (WPT) technology. Compared to the conventional battery-powered networks, WPCN offers the potential to eliminate the need of manual recharging the batteries, which can efficiently reduce the operational cost and enhance the performance attained.

Furthermore, relay-aided transmissions may be invoked for enhancing the cell-edge coverage [11]. Motivated by the aforementioned benefits of WPCNs, the concept of wireless-powered relay (WPR) networks has emerged [12]–[14]. The existing contributions on WPR networks tend to use either time-switching based relaying (TSR) [12], [13] or power-splitting based relaying (PSR) protocols [13], [14]. As a further development, full-duplex (FD) transmissions have attracted substantial attention, since they can potentially double the spectral efficiency [15]. Inspired by the benefits of FD, the optimal TSR protocol was invoked in [16] for boosting the throughput of FD-WPR networks. Furthermore, a novel two-phase protocol was proposed for FD-WPR networks relying on self-interference (SI) recycling, in which the relay harvests energy not only from the source but also from its own SI link [17]. However, the existing body of research is strictly focused on WPR networks relying on single data streams, whilst the scenario of multiple-input multiple-output (MIMO) WPR networks supporting multiple streams has not as yet been studied.

When it comes to MIMO relay networks, the transceiver design has been extensively studied for enhancing the attainable performance. In particular, nonlinear transceivers based on

Tomlinson-Harashima precoding (THP) have attracted much interest as a benefit of their significant performance advantages over their linear counterparts [20]–[23]. In [24], the authors focused their attention on the transceiver design of a two-hop MIMO-aided relay network employing THP at the source. Given the limited processing capability of the relay node, THP is only used at the source. In [25], two methods are proposed for solving the design problem, including a non-iterative technique of finding closed-form solutions for the precoders and an iterative method of separately optimizing the precoders. Robust nonlinear transceiver designs were proposed in [26]–[29] for MIMO-aided relay networks relying on realistic imperfect channel state information (CSI). To the best of our knowledge, however, no contributions have studied robust nonlinear transceiver design of WPR networks.

Motivated by the above considerations, our new contributions are: 1) We propose robust nonlinear transceiver designs for FD-based MIMO WPR networks. Explicitly, the nonlinear THP-based transmitter is intrinsically amalgamated with a linear amplify-and-forward (AF) relaying matrix and a linear minimum-mean-squared-error (MMSE) receiver at the destination, where the THP is composed of a feedback matrix and a source precoding matrix. 2) A novel two-phase transmission protocol is used for the WPR network, in which the relay operates in FD mode relying on simultaneous information transmission and EH [17]. In particular, the relay consciously invokes energy recycling by harnessing a precious portion of the SI power in addition to the RF energy received from its primary source. 3) Within this framework, two compelling practical criteria are considered for the proposed robust transceiver design. The first criterion aims for minimizing the mean square error (MSE) at the destination subject to a realistic transmit power constraint at the source and to an EH constraint at the relay. The resultant non-convex optimization problem is decoupled into four subproblems and solved via a sophisticated alternating optimization (AO) algorithm that iteratively updates the transceiver coefficients by sequentially addressing each subproblem, while keeping the other matrix variables fixed. Specifically, given the THP feedback matrix, as well as the source precoding matrix and the AF relaying matrix, we derive the optimum linear receiver in a closed form. Given the feedback matrix, the source precoding matrix and the linear receiver, the AF relaying matrix can be found by convex optimization. We then transform the non-convex design subproblem of the source precoding matrix into a difference of convex (DC) functions programming [18] and propose an iterative solution scheme based on the constrained concave convex procedure (CCCP) [19]. Finally, by fixing the other three matrix variables, the feedback matrix is derived with the aid of the classic Lagrangian multiplier method. 4) The second criterion aims for minimizing the transmit power at the source under an MSE constraint at the destination and an EH constraint at the relay. Similar to the first condition, an AO-based iterative algorithm is developed for solving this problem sequentially, where an efficient initial feasibility search scheme is proposed. Since some of the search variables are not included in the objective function, we resort to a novel method to enlarge the feasible region so as to obtain the solution. 5) The convergence speed and computational complexity of

the proposed algorithms are also studied in detail. Simulation results are presented for demonstrating the efficiency of the proposed robust nonlinear transceiver designs.

The rest of this paper is organized as follows. Section II introduces the proposed MIMO WPR system model and channel error model, followed by a mathematical formulation of the joint transceiver design problems. In Section III, our AO-based algorithm is conceived for sequentially solving the constrained MSE minimization problem, while the transmit power minimization problem is addressed in Section IV. Simulation results and comparisons to a range of potent benchmark schemes are presented in Section V. Finally, conclusions are offered in Section VI.

Notation: Boldface lowercase and uppercase letters denote vectors and matrices, respectively. $\mathbb{E}[\cdot]$ stands for the statistical expectation. The operators $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$ and $\text{tr}(\cdot)$ denote the matrix transpose, conjugate, Hermitian transpose and trace, respectively. The Kronecker product of matrices is denoted by \otimes . $\text{vec}(\cdot)$ represents the matrix vectorization operator. $\|\cdot\|$ corresponds to the Euclidean norm of the vector. $\mathbf{A} \succeq \mathbf{0}$ indicates that \mathbf{A} is positive semi-definite. $\mathcal{R}(\cdot)$ denotes the real part of a complex-valued scalar. $\lfloor \cdot \rfloor$ represents the floor function, which returns the largest integer that is smaller than or equal to the argument. $\mathcal{CN}(\mathbf{m}, \mathbf{C})$ denotes a multivariate complex circular Gaussian distribution with mean \mathbf{m} and covariance matrix \mathbf{C} .

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

We consider a three-node AF MIMO WPR network, consisting of source-, relay- and destination- nodes, as illustrated in Fig. 1. It is assumed that there is no direct link between the source and the destination due to physical obstacles, which is a valid assumption in many real-world communication scenarios. Hence, an intermediate relay node is needed to assist in the transmission of the source information to the destination. The propagation channels between the source and relay as well as between the relay and destination are assumed to be frequency flat-fading. More specifically, the channel coefficients remain constant over a block's transmission time duration and vary independently and identically from one block to the other, with magnitude obeying the Rayleigh distribution. In this work, we concentrate on the study of transceiver optimization over a block fading channel environment. The source, relay and destination nodes are equipped with N_s , N_r and N_d antennas, respectively. Furthermore, a novel two-phase protocol is adopted for this WPR network [17]. As illustrated in Fig. 2, the information signal is transmitted from the source to the relay in the first phase of duration $T/2$. In the second phase, also of duration $T/2$, the relay operates in its FD mode with simultaneous energy harvesting and information transmission. It receives energy signals from the source, while simultaneously transmitting the information signals to the destination at the same time. The information transmission and EH aspects are further developed below.

1) *Information Transmission:* In the first phase, the source transmits a signal vector $\mathbf{s} = [s_1, \dots, s_{N_d}]^T$, where the entries are independent and identically distributed

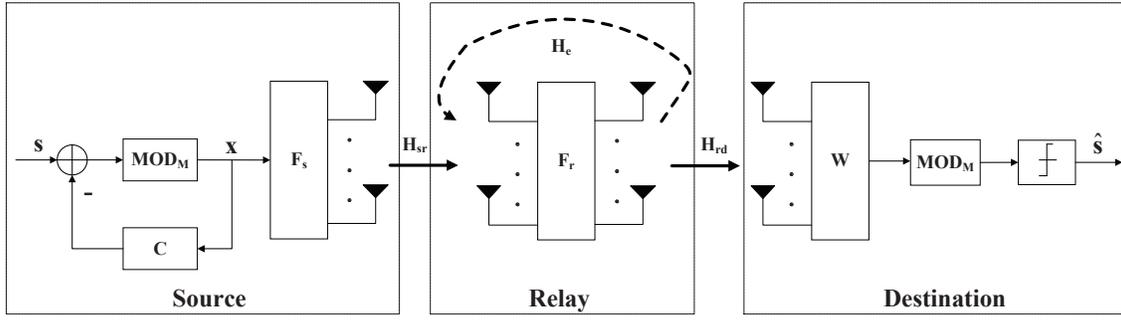


Fig. 1: MIMO WPR network with THP-based transceiver

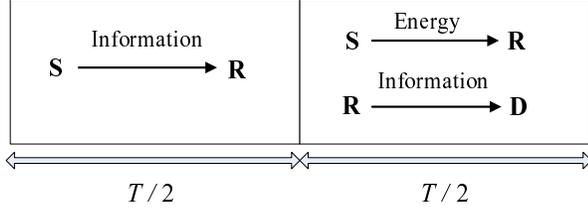


Fig. 2: The proposed two-phase protocol for MIMO WPR networks

(i.i.d.) random variables of an M -ary quadrature amplitude modulation (QAM) alphabet \mathcal{A} . Specifically, we chose $\mathcal{A} = \left\{ \pm 1\sqrt{\frac{3}{2(M-1)}}, \pm 3\sqrt{\frac{3}{2(M-1)}}, \dots, \pm (\sqrt{M}-1)\sqrt{\frac{3}{2(M-1)}} \right\}$ and let $E[s] = \mathbf{0}$ and $E[ss^H] = \mathbf{I}_{N_d}$, where the real and imaginary parts of s_k all belong to the set \mathcal{A} . Furthermore, to provide sufficient degrees of freedom for signal recovery, the number of the substreams N_d should satisfy $N_d \leq N_s$. To generate the THP signal, a modulo operator acts independently over the real and imaginary parts of its input signal components. This operation is formulated as

$$\text{MOD}_M(x) = x - 2\sqrt{\tau} \left\lfloor \frac{x + \sqrt{\tau}}{2\sqrt{\tau}} \right\rfloor = x + e, \quad (1)$$

where $\tau = \sqrt{\frac{3M}{2(M-1)}}$ and e is the residual error, such that $\text{MOD}_M(x)$ is constrained to be within the interval $[-\tau, \tau]$. The information signal $\mathbf{x} = [x_1, \dots, x_{N_d}]^T$ intended for the destination is recursively computed with the help of the modulo operation in (1) as

$$\begin{aligned} x_k &= \text{MOD}_M \left(s_k - \sum_{n=1}^{k-1} C_{k,n} x_n \right) \\ &= s_k - \sum_{n=1}^{k-1} C_{k,n} x_n + e_k, \quad k = 1, 2, \dots, N_d, \end{aligned} \quad (2)$$

where \mathbf{C} is a strictly lower triangular matrix with entries $C_{k,n}$, and e_k is the residual error for the k th entry.

Upon introducing $\mathbf{e} = [e_1, \dots, e_{N_d}]^T$, we can rewrite (2) as

$$\mathbf{x} = \mathbf{U}^{-1}\mathbf{v}, \quad (3)$$

where $\mathbf{U} = \mathbf{C} + \mathbf{I}$ is a lower triangular matrix with unity diagonal entries and $\mathbf{v} = \mathbf{s} + \mathbf{e}$ is the effective signal vector.

¹ Note that for a high-order M -QAM constellation, \mathbf{x} can be

¹The proposed nonlinear scheme can be treated as a general processing framework for precoding. Note that if the matrix \mathbf{U} is an identity matrix, the proposed nonlinear transceiver design can be reduced to the linear one.

assumed to have a covariance of $E[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_{N_d}$ [30]. Following THP, the resultant vector \mathbf{x} is passed through a linear precoding matrix $\mathbf{F}_s \in \mathbb{C}^{N_s \times N_d}$ prior to its transmission.

In the first phase, the signal $\mathbf{y}_{r,1}$ received at the relay is formulated as

$$\mathbf{y}_{r,1} = \mathbf{H}_{sr}\mathbf{F}_s\mathbf{x} + \mathbf{n}_{sr,1}, \quad (4)$$

where $\mathbf{H}_{sr} \in \mathbb{C}^{N_r \times N_s}$ denotes the flat-fading channel matrix of the link between the source and the relay, while $\mathbf{n}_{sr,1}$ is the additive white Gaussian noise (AWGN) vector at the relay, with a zero-mean and covariance matrix of $E[\mathbf{n}_{sr,1}\mathbf{n}_{sr,1}^H] = \sigma_{n_{sr}}^2 \mathbf{I}_{N_r}$.

During the second phase, the relay amplifies the received signal by a linear AF matrix $\mathbf{F}_r \in \mathbb{C}^{N_r \times N_r}$ and forwards it to the destination. The signal vector received at the destination is given by

$$\mathbf{y}_d = \mathbf{H}_{rd}\mathbf{F}_r\mathbf{y}_{r,1} + \mathbf{n}_{rd}, \quad (5)$$

where $\mathbf{H}_{rd} \in \mathbb{C}^{N_d \times N_r}$ is the channel matrix of the link between the relay and the destination, while \mathbf{n}_{rd} is the AWGN vector at the destination, with a zero-mean and covariance matrix of $E[\mathbf{n}_{rd}\mathbf{n}_{rd}^H] = \sigma_{n_{rd}}^2 \mathbf{I}_{N_d}$.

At the destination, a linear receiver represented by the matrix $\mathbf{W} \in \mathbb{C}^{N_d \times N_d}$ is employed for detecting the received signal. Specifically, the output of the linear receiver is expressed as

$$\hat{\mathbf{v}} = \mathbf{W}\mathbf{H}_{rd}\mathbf{F}_r\mathbf{H}_{sr}\mathbf{F}_s\mathbf{x} + \mathbf{W}\mathbf{H}_{rd}\mathbf{F}_r\mathbf{n}_{sr} + \mathbf{W}\mathbf{n}_{rd}, \quad (6)$$

while the final output is obtained as

$$\hat{\mathbf{s}} = \text{MOD}_M(\hat{\mathbf{v}}). \quad (7)$$

In general, transceiver optimization relies on the availability of accurate CSI. However, in practical systems, encountering channel estimation errors is inevitable and these should be taken into account in a robust design. Here, we consider the popular statistical error model to handle the channel uncertainties. This model has been well justified in the transceiver design literature [23], [28], [29] and references therein; it is especially suitable when the CSI errors are dominated by the channel estimation errors. Using the well-known Kronecker model [22], the channel matrices are expressed as

$$\mathbf{H}_{sr} = \bar{\mathbf{H}}_{sr} + \Delta\mathbf{H}_{sr}, \quad (8)$$

$$\mathbf{H}_{rd} = \bar{\mathbf{H}}_{rd} + \Delta\mathbf{H}_{rd}, \quad (9)$$

where $\bar{\mathbf{H}}_{sr}$ and $\bar{\mathbf{H}}_{rd}$ are the estimated channel matrices, while $\Delta\mathbf{H}_{sr}$ and $\Delta\mathbf{H}_{rd}$ are the corresponding channel estimation error matrices. Specifically, $\Delta\mathbf{H}_{sr}$ can be written as $\Delta\mathbf{H}_{sr} = \sum_{sr}^{1/2} \mathbf{H}_1 \Psi_{sr}^{T/2}$, while $\Delta\mathbf{H}_{rd}$ can be expressed as $\Delta\mathbf{H}_{rd} = \sum_{rd}^{1/2} \mathbf{H}_2 \Psi_{rd}^{T/2}$, where the entries of \mathbf{H}_1 and \mathbf{H}_2 are both i.i.d. with zero-mean and unit-variance circular complex Gaussian distribution [31]. Hence, $\Delta\mathbf{H}_{sr}$ and $\Delta\mathbf{H}_{rd}$ exhibit matrix-variate complex Gaussian distributions formulated as

$$\Delta\mathbf{H}_{sr} \sim \mathcal{CN}(\mathbf{0}_{N_r, N_s}, \Psi_{sr} \otimes \Sigma_{sr}), \quad (10)$$

$$\Delta\mathbf{H}_{rd} \sim \mathcal{CN}(\mathbf{0}_{N_d, N_r}, \Psi_{rd} \otimes \Sigma_{rd}), \quad (11)$$

where Ψ_{sr} and Σ_{sr} denote the normalized covariance matrices of the source-to-relay channel seen from the transmitter and receiver, respectively, and so do Ψ_{rd} and Σ_{rd} for the relay-to-destination channel. The proposed optimization is implemented at the source node. We assume that the system works in the time division duplex (TDD) mode [32]. The first-phase CSI (from the source to the relay) can be obtained at the source by using channel estimation and channel reciprocity. The second-phase CSI (from the relay to the destination) can be obtained by following the same approach at the relay; they can be conveyed to the source via control channels.

It is assumed that the magnitudes of the CSI errors are much smaller than those of the channel estimates themselves, hence both the third and higher order moments of the CSI errors of $\Delta\mathbf{H}_{sr}$ and $\Delta\mathbf{H}_{rd}$ can be neglected [33]. The MSE at the destination of the complete transceiver chain seen in Fig. 1 can be written as

$$\begin{aligned} \text{MSE}(\mathbf{W}, \mathbf{F}_s, \mathbf{F}_r, \mathbf{U}) &= \mathbb{E} \left[\|\mathbf{W}\mathbf{y}_d - \mathbf{v}\|^2 \right] \\ &= \text{tr}(\mathbf{W}\mathbf{M}\mathbf{W}^H) - \text{tr}(\mathbf{U}\mathbf{F}_s^H \bar{\mathbf{H}}_{sr}^H \mathbf{F}_r^H \bar{\mathbf{H}}_{rd}^H \mathbf{W}^H) \\ &\quad + \text{tr}(\mathbf{U}\mathbf{U}^H) - \text{tr}(\mathbf{W}\bar{\mathbf{H}}_{rd}^H \mathbf{F}_r^H \bar{\mathbf{H}}_{sr}^H \mathbf{F}_s \mathbf{U}^H), \end{aligned} \quad (12)$$

where we have:

$$\begin{aligned} \mathbf{M} &= \bar{\mathbf{H}}_{rd}^H \mathbf{F}_r (\bar{\mathbf{H}}_{sr}^H \mathbf{F}_s \mathbf{F}_s^H \bar{\mathbf{H}}_{sr}^H + \sigma_{n_{sr}}^2 \mathbf{I}) \mathbf{F}_r^H \bar{\mathbf{H}}_{rd}^H + \sigma_{n_{rd}}^2 \mathbf{I} \\ &\quad + \text{tr}(\mathbf{F}_r \bar{\mathbf{H}}_{sr}^H \mathbf{F}_s \mathbf{F}_s^H \bar{\mathbf{H}}_{sr}^H \mathbf{F}_r^H \Psi_{rd}^T) \Sigma_{rd} \\ &\quad + \text{tr}(\mathbf{F}_s \mathbf{F}_s^H \Psi_{sr}^T) \bar{\mathbf{H}}_{rd}^H \mathbf{F}_r \Sigma_{sr} \mathbf{F}_r^H \bar{\mathbf{H}}_{rd}^H \\ &\quad + \sigma_{n_{sr}}^2 \text{tr}(\mathbf{F}_r \mathbf{F}_r^H \Psi_{rd}^T) \Sigma_{rd}. \end{aligned} \quad (13)$$

Note that the expectation here is taken over all the distributions of $\Delta\mathbf{H}_{sr}$, $\Delta\mathbf{H}_{rd}$, \mathbf{n}_{sr} , \mathbf{n}_{rd} .

2) *Energy Harvesting*: Concurrently with the information transmission of the second phase, wireless energy is conveyed from the source to the relay for EH. In fact, we only concentrate on the energy itself, rather than the format of the energy signal. The transmit power of the energy signals at the source can be expressed as $\mathbb{E}[\|\mathbf{F}_s \mathbf{x}_{en}\|^2]$, where \mathbf{x}_{en} is the energy signal after the nonlinear THP operation. Since the power loss introduced by THP processing can be neglected [30], the conclusion reached in above discussion on information transmission still holds. Specifically, similar to (3), it can be observed that the covariance matrix of \mathbf{x}_{en} can still be approximated as $\mathbb{E}[\mathbf{x}_{en} \mathbf{x}_{en}^H] = \mathbf{I}_{N_d}$. By imposing a transmit power constraint at the source, the transmit power should satisfy: $\text{tr}(\mathbf{F}_s \mathbf{x}_{en} \mathbf{x}_{en}^H \mathbf{F}_s^H) = \text{tr}(\mathbf{F}_s \mathbf{F}_s^H) \leq P_s$, where P_s is the maximum transmit power available at the source.

For the proposed relaying protocol aided by SI-energy recycling, the relay harvests the dedicated energy emanating from the source, and additionally recycles a portion of the energy used for information transmission via a SI-feedback loop represented by $\mathbf{H}_e \in \mathbb{C}^{N_r \times N_r}$. Accordingly, the signal received at the relay is given by

$$\begin{aligned} \mathbf{y}_{r,2} &= \mathbf{H}_{sr} \mathbf{F}_s \mathbf{x}_{en} + \mathbf{H}_e \mathbf{F}_r \mathbf{y}_{r,1} + \mathbf{n}_{sr,2} \\ &= \mathbf{H}_{sr} \mathbf{F}_s \mathbf{x}_{en} + \mathbf{H}_e \mathbf{F}_r \mathbf{H}_{sr} \mathbf{F}_s \mathbf{x} + \mathbf{H}_e \mathbf{F}_r \mathbf{n}_{sr,1} + \mathbf{n}_{sr,2}, \end{aligned} \quad (14)$$

where $\mathbf{n}_{sr,2}$ is an AWGN vector with zero-mean and covariance matrix of $\mathbb{E}[\mathbf{n}_{sr,2} \mathbf{n}_{sr,2}^H] = \sigma_{n_{sr,2}}^2 \mathbf{I}_{N_r}$. Thus the total energy harvested at the relay is expressed as

$$E_r = \frac{T}{2} \eta \text{tr}(\mathbb{E}[\mathbf{y}_{r,2} \mathbf{y}_{r,2}^H]), \quad (15)$$

where $0 < \eta < 1$ denotes the EH efficiency at the relay. In practice, the energy harvesting sub-system would have additional control circuits which consume power; however, this power consumption has been included in the conversion efficiency η . To ensure that the energy used for transmission at the relay does not exceed that being harvested, the constraint $\text{tr}(\mathbf{F}_r \mathbb{E}[\mathbf{y}_{r,1} \mathbf{y}_{r,1}^H] \mathbf{F}_r^H) T/2 \leq E_r$ must be satisfied. The EH constraint at the relay can therefore be formulated as

$$\text{tr}(\mathbf{F}_r \mathbb{E}[\mathbf{y}_{r,1} \mathbf{y}_{r,1}^H] \mathbf{F}_r^H) \leq \eta \text{tr}(\mathbb{E}[\mathbf{y}_{r,2} \mathbf{y}_{r,2}^H]). \quad (16)$$

B. Problem Formulation

Given the MSE expression (12) and the EH constraint (16), we can now formulate the proposed robust transceiver design in mathematical term. Our general goal is to jointly design the THP at the source, the linear AF matrix at the relay and the linear receiver at the destination, in order to minimize well-accepted performance metrics, such as the MSE and transmit power. Furthermore, we focus our attention on a robust design philosophy, where the statistical error model in (8)-(11) is adopted for describing the CSI mismatch. In light of the above system model equations, this goal translates into two different design problem formulations, defined as follows:

- Minimizing the total MSE at the destination, subject to a transmit power constraint at the source and the EH constraint at the relay, expressed as:

$$\min_{\mathbf{W}, \mathbf{F}_s, \mathbf{F}_r, \mathbf{U}} \text{MSE}(\mathbf{W}, \mathbf{F}_s, \mathbf{F}_r, \mathbf{U}) \quad (17a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{F}_s \mathbf{F}_s^H) \leq P_s \quad (17b)$$

$$\text{tr}(\mathbf{F}_r \mathbb{E}[\mathbf{y}_{r,1} \mathbf{y}_{r,1}^H] \mathbf{F}_r^H) \leq \eta \text{tr}(\mathbb{E}[\mathbf{y}_{r,2} \mathbf{y}_{r,2}^H]), \quad (17c)$$

- Minimizing the transmit power at the source, subject to the MSE constraint at the destination and the EH constraint at the relay, expressed as:

$$\min_{\mathbf{W}, \mathbf{F}_s, \mathbf{F}_r, \mathbf{U}} \text{tr}(\mathbf{F}_s \mathbf{F}_s^H) \quad (18a)$$

$$\text{s.t.} \quad \text{MSE}(\mathbf{W}, \mathbf{F}_s, \mathbf{F}_r, \mathbf{U}) \leq \gamma \quad (18b)$$

$$\text{tr}(\mathbf{F}_r \mathbb{E}[\mathbf{y}_{r,1} \mathbf{y}_{r,1}^H] \mathbf{F}_r^H) \leq \eta \text{tr}(\mathbb{E}[\mathbf{y}_{r,2} \mathbf{y}_{r,2}^H]), \quad (18c)$$

where γ denotes a common MSE target.

Both problems (17) and (18) are non-convex with multi-dimensional variables, making it quite challenging to simultaneously find the optimal \mathbf{W} , \mathbf{F}_s , \mathbf{F}_r and \mathbf{U} . To overcome this limitation and obtain tractable solutions, iterative algorithms inspired by the AO principle are developed for each one of the above problems in the following two sections.

III. ROBUST TRANSCIVER DESIGN BASED ON MSE MINIMIZATION

The MSE is a ubiquitous design criterion routinely used for evaluating the transmission reliability of communication systems. In this section, we address the robust transceiver design problem (17), where our aim is to minimize the MSE, while meeting both the transmit power and the EH constraints. An AO-based algorithm is developed for optimizing A) the source precoding matrix, B) the AF relaying matrix, C) the receiver matrix and D) the feedback matrix in an iterative manner by solving the resultant subproblems.

A. Optimization of the Receiver Matrix

Since \mathbf{W} is not involved in any constraints, the optimal linear receiver matrix \mathbf{W} of the destination can be derived by solving the unconstrained convex problem of (17a). Hence, given \mathbf{F}_r , \mathbf{F}_s and \mathbf{U} , the optimal Wiener filter is derived by solving $\frac{\partial}{\partial \mathbf{W}^*} \text{MSE}(\mathbf{W}, \mathbf{F}_s, \mathbf{F}_r, \mathbf{U}) = \mathbf{0}$, which results into

$$\mathbf{W} = \mathbf{U} \mathbf{F}_{sr}^H \bar{\mathbf{H}}_{sr}^H \mathbf{F}_r^H \bar{\mathbf{H}}_{rd}^H \mathbf{M}^{-1}. \quad (19)$$

B. Optimization of the AF Relaying Matrix

We next solve the sub-problem for the AF relaying matrix \mathbf{F}_r , while keeping \mathbf{F}_s , \mathbf{U} and \mathbf{W} fixed. Invoking the following identities from [34],

$$\text{tr}(\mathbf{A}^T \mathbf{B}) = (\text{vec}(\mathbf{A}))^T \text{vec}(\mathbf{B}), \quad (20)$$

$$\text{tr}(\mathbf{A}^T \mathbf{B} \mathbf{A} \mathbf{C}) = (\text{vec}(\mathbf{A}))^H (\mathbf{C}^T \otimes \mathbf{B}) \text{vec}(\mathbf{A}), \quad (21)$$

$$\text{vec}(\mathbf{A} \mathbf{B} \mathbf{C}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B}), \quad (22)$$

the MSE in (12) can be represented as a function of $\mathbf{f}_r \triangleq \text{vec}(\mathbf{F}_r)$ as follows,

$$\text{MSE}(\mathbf{F}_r) = \mathbf{f}_r^H \mathbf{A} \mathbf{f}_r - \mathbf{d}^H \mathbf{f}_r - \mathbf{f}_r^H \mathbf{d} + t_1, \quad (23)$$

where we have:

$$\begin{aligned} \mathbf{A} &= \left((\bar{\mathbf{H}}_{sr} \mathbf{F}_s \mathbf{F}_s^H \bar{\mathbf{H}}_{sr}^H)^T + \text{tr}(\mathbf{F}_s \mathbf{F}_s^H \Psi_{sr}^T) \Sigma_{sr}^T + \sigma_{n_{sr}}^2 \mathbf{I} \right) \\ &\quad \otimes (\bar{\mathbf{H}}_{rd}^H \mathbf{W}^H \mathbf{W} \bar{\mathbf{H}}_{rd}) + \text{tr}(\mathbf{W} \Sigma_{rd} \mathbf{W}^H) \\ &\quad \times \left((\bar{\mathbf{H}}_{sr} \mathbf{F}_s \mathbf{F}_s^H \bar{\mathbf{H}}_{sr}^H)^T + \sigma_{n_{sr}}^2 \mathbf{I} \right) \otimes \Psi_{rd}^T, \end{aligned} \quad (24)$$

$$\mathbf{d} = \text{vec}(\bar{\mathbf{H}}_{rd}^H \mathbf{W}^H \mathbf{U} \mathbf{F}_s^H \bar{\mathbf{H}}_{sr}^H), \quad (25)$$

$$t_1 = \text{tr}(\mathbf{U} \mathbf{U}^H) + \sigma_{n_{rd}}^2 \text{tr}(\mathbf{W} \mathbf{W}^H), \quad (26)$$

and t_1 does not depend on \mathbf{f}_r .

Note that the relay's EH constraint (17c) can be written as

$$\begin{aligned} &\text{tr}(\mathbf{F}_r \mathbf{E}[\mathbf{y}_{r,1} \mathbf{y}_{r,1}^H] \mathbf{F}_r^H) - \eta \text{tr}(\mathbf{H}_e \mathbf{F}_r \mathbf{E}[\mathbf{y}_{r,1} \mathbf{y}_{r,1}^H] \mathbf{F}_r^H \mathbf{H}_e^H) \\ &\leq \eta \text{tr}(\mathbf{H}_{sr} \mathbf{F}_s \mathbf{F}_s^H \mathbf{H}_{sr}^H + \sigma_{n_{sr}}^2 \mathbf{I}). \end{aligned} \quad (27)$$

Using the identities in (20)-(22), we have

$$\text{tr}(\mathbf{F}_r \mathbf{E}[\mathbf{y}_{r,1} \mathbf{y}_{r,1}^H] \mathbf{F}_r^H) = \mathbf{f}_r^H \mathbf{D}_1 \mathbf{f}_r, \quad (28)$$

$$\eta \text{tr}(\mathbf{H}_e \mathbf{F}_r \mathbf{E}[\mathbf{y}_{r,1} \mathbf{y}_{r,1}^H] \mathbf{F}_r^H \mathbf{H}_e^H) = \mathbf{f}_r^H \mathbf{D}_2 \mathbf{f}_r, \quad (29)$$

where

$$\begin{aligned} \mathbf{D}_1 &= \left((\bar{\mathbf{H}}_{sr} \mathbf{F}_s \mathbf{F}_s^H \bar{\mathbf{H}}_{sr}^H)^T + \sigma_{n_{sr}}^2 \mathbf{I} \right) \\ &\quad + \text{tr}(\mathbf{F}_s \mathbf{F}_s^H \Psi_{sr}^T) \Sigma_{sr}^T \otimes \mathbf{I}, \end{aligned} \quad (30)$$

$$\begin{aligned} \mathbf{D}_2 &= \left((\bar{\mathbf{H}}_{sr} \mathbf{F}_s \mathbf{F}_s^H \bar{\mathbf{H}}_{sr}^H)^T + \sigma_{n_{sr}}^2 \mathbf{I} \right) \\ &\quad + \text{tr}(\mathbf{F}_s \mathbf{F}_s^H \Psi_{sr}^T) \Sigma_{sr}^T \otimes (\eta \mathbf{H}_e^H \mathbf{H}_e). \end{aligned} \quad (31)$$

Thus (27) can be equivalently rewritten as

$$\mathbf{f}_r^H \tilde{\mathbf{D}} \mathbf{f}_r \leq \bar{P}_r, \quad (32)$$

where

$$\begin{aligned} \tilde{\mathbf{D}} &= \mathbf{D}_1 - \mathbf{D}_2 = \left((\bar{\mathbf{H}}_{sr} \mathbf{F}_s \mathbf{F}_s^H \bar{\mathbf{H}}_{sr}^H)^T + \sigma_{n_{sr}}^2 \mathbf{I} \right) \\ &\quad + \text{tr}(\mathbf{F}_s \mathbf{F}_s^H \Psi_{sr}^T) \Sigma_{sr}^T \otimes (\mathbf{I} - \eta \mathbf{H}_e^H \mathbf{H}_e), \end{aligned} \quad (33)$$

$$\begin{aligned} \bar{P}_r &= \eta \text{tr}(\bar{\mathbf{H}}_{sr} \mathbf{F}_s \mathbf{F}_s^H \bar{\mathbf{H}}_{sr}^H + \sigma_{n_{sr}}^2 \mathbf{I}) \\ &\quad + \eta \text{tr}(\mathbf{F}_s \mathbf{F}_s^H \Psi_{sr}^T) \text{tr}(\Sigma_{sr}). \end{aligned} \quad (34)$$

Since $(\mathbf{I} - \eta \mathbf{H}_e^H \mathbf{H}_e) \succeq \mathbf{0}$ holds in practice ², we can verify that $\tilde{\mathbf{D}} \succeq \mathbf{0}$.

From (23) and (32), the AF relaying matrix optimization problem can be formulated as

$$\min_{\mathbf{f}_r} (\mathbf{f}_r^H \mathbf{A} \mathbf{f}_r - \mathbf{d}^H \mathbf{f}_r - \mathbf{f}_r^H \mathbf{d} + t_1) \quad (35a)$$

$$\text{s.t. } \mathbf{f}_r^H \tilde{\mathbf{D}} \mathbf{f}_r \leq \bar{P}_r. \quad (35b)$$

The above is a quadratically constrained quadratic program (QCQP) [36], which can be equivalently transformed into standard second-order cone program (SOCP) [37] as follows,

$$\min_{p_1, \lambda_1, \mathbf{f}_r} p_1 \quad (36a)$$

$$\text{s.t. } \left\| \mathbf{A}^{1/2} \mathbf{f}_r - \mathbf{A}^{-1/2} \mathbf{d} \right\| \leq \lambda_1, \quad (36b)$$

$$\left\| \begin{bmatrix} 2\lambda_1 \\ p_1 + \mathbf{d}^H \mathbf{A}^{-1} \mathbf{d} - t_1 - 1 \end{bmatrix} \right\| \leq p_1 + \mathbf{d}^H \mathbf{A}^{-1} \mathbf{d} - t_1 + 1, \quad (36c)$$

$$\left\| \tilde{\mathbf{D}}^{1/2} \mathbf{f}_r \right\| \leq \sqrt{\bar{P}_r}, \quad (36d)$$

where p_1 and λ_1 are slack variables. Note that for the SOCP itself, the objective function (OF) is linear and its constraints are linear or second-order cones. Hence, the problem can be efficiently solved by using the convex programming toolbox CVX [38].

²In this contribution, \mathbf{H}_e represents the normalized SI loop channel matrix, which under the assumption of low SI power, introduces high attenuation. Since the source node can acquire accurate estimates of \mathbf{H}_e , the energy efficiency η ($0 < \eta < 1$) can be appropriately set to satisfy the requirement of positive semi-definiteness. The value of η should be less than $\max(1, 1/\gamma_{\max})$, where γ_{\max} denotes max eigenvalue of $\mathbf{H}_e^H \mathbf{H}_e$.

C. Optimization of the Source Precoding Matrix

Given \mathbf{U} , \mathbf{W} and \mathbf{F}_r , we now focus our attention on the optimization problem (17) with respect to the source precoding matrix \mathbf{F}_s . Introducing $\tilde{\mathbf{H}} = \mathbf{W}\tilde{\mathbf{H}}_{rd}\mathbf{F}_r\tilde{\mathbf{H}}_{sr}$, the MSE in (12) can be rewritten as a function of \mathbf{F}_s as follows,

$$\begin{aligned} \text{MSE}(\mathbf{F}_s) &= \text{tr}\left(\tilde{\mathbf{H}}\mathbf{F}_s\mathbf{F}_s^H\tilde{\mathbf{H}}^H\right) - \text{tr}\left(\tilde{\mathbf{H}}\mathbf{F}_s\mathbf{U}^H\right) \\ &\quad - \text{tr}\left(\mathbf{U}\mathbf{F}_s^H\tilde{\mathbf{H}}^H\right) + t_2, \end{aligned} \quad (37)$$

where

$$\begin{aligned} t_2 &= \sigma_{n_{sr}}^2 \text{tr}\left(\left(\mathbf{W}\tilde{\mathbf{H}}_{rd}\mathbf{F}_r\right)\left(\mathbf{W}\tilde{\mathbf{H}}_{rd}\mathbf{F}_r\right)^H\right) \\ &\quad + \sigma_{n_{sr}}^2 \text{tr}\left(\mathbf{F}_r\mathbf{F}_r^H\boldsymbol{\Psi}_{rd}^T\right) \text{tr}\left(\mathbf{W}\boldsymbol{\Sigma}_{rd}\mathbf{W}^H\right) \\ &\quad + \text{tr}\left(\mathbf{U}\mathbf{U}^H\right) + \sigma_{n_{rd}}^2 \text{tr}\left(\mathbf{W}\mathbf{W}^H\right), \end{aligned} \quad (38)$$

and t_2 does not depend on \mathbf{F}_s . Using the identities in (20)-(22), the MSE in (37) can be expressed as a function of $\mathbf{f}_s \triangleq \text{vec}(\mathbf{F}_s)$ as follows,

$$\text{MSE}(\mathbf{F}_s) = \mathbf{f}_s^H \mathbf{G} \mathbf{f}_s - \mathbf{Q}^H \mathbf{f}_s - \mathbf{f}_s^H \mathbf{Q} + t_2, \quad (39)$$

where we have:

$$\begin{aligned} \mathbf{G} &= \mathbf{I} \otimes \left(\tilde{\mathbf{H}}^H \tilde{\mathbf{H}} + \text{tr}\left(\mathbf{W}\boldsymbol{\Sigma}_{rd}\mathbf{W}^H\right) \tilde{\mathbf{H}}_{sr}^H \mathbf{F}_r^H \boldsymbol{\Psi}_{rd}^T \mathbf{F}_s \tilde{\mathbf{H}}_{sr} \right. \\ &\quad \left. + \text{tr}\left(\mathbf{W}\tilde{\mathbf{H}}_{rd}\mathbf{F}_r\boldsymbol{\Sigma}_{sr}\mathbf{F}_r^H\tilde{\mathbf{H}}_{rd}^H\mathbf{W}^H\right) \boldsymbol{\Psi}_{sr}^T \right), \end{aligned} \quad (40)$$

$$\mathbf{Q} = \text{vec}(\tilde{\mathbf{H}}^H \mathbf{U}). \quad (41)$$

Now considering the EH constraint (17c), we first note that

$$\text{tr}\left(\mathbf{F}_r\mathbf{H}_{sr}\mathbf{F}_s\mathbf{F}_s^H\mathbf{H}_{sr}^H\mathbf{F}_r^H\right) = \mathbf{f}_s^H \mathbf{E}_1 \mathbf{f}_s, \quad (42)$$

$$\eta \text{tr}\left(\mathbf{H}_e\mathbf{F}_r\mathbf{H}_{sr}\mathbf{F}_s\mathbf{F}_s^H\mathbf{H}_{sr}^H\mathbf{F}_r^H\mathbf{H}_e^H\right) = \mathbf{f}_s^H \mathbf{E}_2 \mathbf{f}_s, \quad (43)$$

$$\eta \text{tr}\left(\mathbf{H}_{sr}\mathbf{F}_s\mathbf{F}_s^H\mathbf{H}_{sr}^H\right) = \mathbf{f}_s^H \mathbf{E}_3 \mathbf{f}_s, \quad (44)$$

where

$$\mathbf{E}_1 = \mathbf{I} \otimes \left(\tilde{\mathbf{H}}_{sr}^H \mathbf{F}_r^H \mathbf{H}_{sr} \tilde{\mathbf{H}}_{sr} + \text{tr}\left(\mathbf{F}_r^H \mathbf{F}_r \boldsymbol{\Sigma}_{sr}\right) \boldsymbol{\Psi}_{sr}^T \right), \quad (45)$$

$$\begin{aligned} \mathbf{E}_2 &= \eta \mathbf{I} \otimes \left(\tilde{\mathbf{H}}_{sr}^H \mathbf{F}_r^H \mathbf{H}_e \mathbf{H}_{sr} \tilde{\mathbf{H}}_{sr} \right. \\ &\quad \left. + \text{tr}\left(\mathbf{F}_r^H \mathbf{H}_e \mathbf{H}_{sr} \boldsymbol{\Sigma}_{sr}\right) \boldsymbol{\Psi}_{sr}^T \right), \end{aligned} \quad (46)$$

$$\mathbf{E}_3 = \eta \mathbf{I} \otimes \left(\tilde{\mathbf{H}}_{sr}^H \tilde{\mathbf{H}}_{sr} + \text{tr}\left(\boldsymbol{\Sigma}_{sr}\right) \boldsymbol{\Psi}_{sr}^T \right). \quad (47)$$

With the help of (42)-(44), the EH constraint can be recast as

$$\mathbf{f}_s^H \mathbf{E}_1 \mathbf{f}_s - h(\mathbf{f}_s) \leq 0, \quad (48)$$

where

$$h(\mathbf{f}_s) = \mathbf{f}_s^H (\mathbf{E}_2 + \mathbf{E}_3) \mathbf{f}_s + t_3, \quad (49)$$

$$\begin{aligned} t_3 &= \eta \sigma_{n_{sr}}^2 \text{tr}\left(\mathbf{H}_e\mathbf{F}_r\mathbf{F}_r^H\mathbf{H}_e^H\right) + \eta \text{tr}\left(\sigma_{n_{sr}}^2 \mathbf{I}\right) \\ &\quad - \sigma_{n_{sr}}^2 \text{tr}\left(\mathbf{F}_r\mathbf{F}_r^H\right). \end{aligned} \quad (50)$$

We note that $\mathbf{f}_s^H \mathbf{E}_1 \mathbf{f}_s$ and $h(\mathbf{f}_s)$ in (48) are both convex functions of \mathbf{f}_s , which means that the EH constraint in (48) represents the difference of two convex functions [39]. Thus, the problem in (17) can be transformed into the following DC program [19]

$$\min_{\mathbf{f}_s} \mathbf{f}_s^H \mathbf{G} \mathbf{f}_s - \mathbf{Q}^H \mathbf{f}_s - \mathbf{f}_s^H \mathbf{Q} + t_2, \quad (51a)$$

$$\text{s.t. } \mathbf{f}_s^H \mathbf{f}_s \leq P_s, \quad (51b)$$

$$\mathbf{f}_s^H \mathbf{E}_1 \mathbf{f}_s - h(\mathbf{f}_s) \leq 0. \quad (51c)$$

In the following, we propose to employ the CCCP-based iterative algorithm of [18] for solving the DC program (51). The first-order Taylor expansion of $h(\mathbf{f}_s)$ around the current point $\mathbf{f}_s^{(i)}$ in the i th iteration is computed as [40]

$$\begin{aligned} \hat{h}\left(\mathbf{f}_s^{(i)}, \mathbf{f}_s\right) &= h\left(\mathbf{f}_s^{(i)}\right) + 2\mathcal{R}\left\{\nabla h\left(\mathbf{f}_s^{(i)}\right)^H\left(\mathbf{f}_s - \mathbf{f}_s^{(i)}\right)\right\} \\ &= 2\mathcal{R}\left\{\left(\mathbf{f}_s^{(i)}\right)^H\left(\mathbf{E}_2 + \mathbf{E}_3\right)\mathbf{f}_s\right\} \\ &\quad - \left(\mathbf{f}_s^{(i)}\right)^H\left(\mathbf{E}_2 + \mathbf{E}_3\right)\mathbf{f}_s^{(i)} + t_3, \end{aligned} \quad (52)$$

where $\nabla h(\mathbf{f}_s^{(i)})$ denotes the conjugate derivative of the function $h(\mathbf{f}_s)$ with respect to the complex vector \mathbf{f}_s , evaluated at $\mathbf{f}_s^{(i)}$.

Then, in the $(i+1)$ st iteration of the proposed CCCP-based iterative algorithm, we solve the following convex optimization problem:

$$\min_{\mathbf{f}_s} \left(\mathbf{f}_s^H \mathbf{G} \mathbf{f}_s - \mathbf{Q}^H \mathbf{f}_s - \mathbf{f}_s^H \mathbf{Q} + t_2 \right), \quad (53a)$$

$$\text{s.t. } \mathbf{f}_s^H \mathbf{f}_s \leq P_s, \quad (53b)$$

$$\mathbf{f}_s^H \mathbf{E}_1 \mathbf{f}_s - \hat{h}\left(\mathbf{f}_s^{(i)}, \mathbf{f}_s\right) \leq 0. \quad (53c)$$

As shown in Appendix A, problem (53) can be reformulated as the following SOCP which can be solved by standard techniques,

$$\min_{p_2, \lambda_2, \mathbf{f}_s} p_2, \quad (54a)$$

$$\text{s.t. } \left\| \mathbf{G}^{1/2} \mathbf{f}_s - \mathbf{G}^{-1/2} \mathbf{Q} \right\| \leq \lambda_2, \quad (54b)$$

$$\left\| \begin{bmatrix} 2\lambda_2 \\ p_2 + \mathbf{Q}^H \mathbf{G}^{-1} \mathbf{Q} - t_2 - 1 \end{bmatrix} \right\| \leq p_2 + \mathbf{Q}^H \mathbf{G}^{-1} \mathbf{Q} - t_2 + 1, \quad (54c)$$

$$\|\mathbf{f}_s\| \leq \sqrt{P_s}, \quad (54d)$$

$$\left\| \begin{bmatrix} 2\mathbf{E}_1^{1/2} \mathbf{f}_s \\ -\mathcal{R}\{\mathbf{g}^H \mathbf{f}_s\} - b - 1 \end{bmatrix} \right\| \leq -\mathcal{R}\{\mathbf{g}^H \mathbf{f}_s\} - b + 1. \quad (54e)$$

The CCCP-based iterative algorithm proposed for solving the DC program of (51) is summarized in Table I, where N_{\max}^a denotes the maximum number of iterations.

TABLE I: The CCCP-Based Iterative Algorithm to Solve (51)

1.	Initialize the algorithm with a feasible point $\mathbf{f}_s^{(0)}$. Set the iteration number $i = 0$.
2.	Repeat: – Solve problem (54) with the current feasible point $\mathbf{f}_s^{(i)}$. – Assign the solution to $\mathbf{f}_s^{(i+1)}$ and update $i := i + 1$.
3.	Until: Convergence or $i > N_{\max}^a$; obtain \mathbf{F}_s .

D. Optimization of Feedback Based THP Matrix

Finally, with fixed \mathbf{W} , \mathbf{F}_r and \mathbf{F}_s , we seek the specific THP feedback matrix \mathbf{U} , which minimizes

$$\begin{aligned} \text{MSE}(\mathbf{U}) &= \text{tr}\left(\mathbf{U}\mathbf{U}^H\right) - \text{tr}\left(\mathbf{U}\mathbf{F}_s^H\tilde{\mathbf{H}}_{sr}^H\mathbf{F}_r^H\tilde{\mathbf{H}}_{rd}^H\mathbf{W}^H\right) \\ &\quad - \text{tr}\left(\mathbf{W}\tilde{\mathbf{H}}_{rd}\mathbf{F}_r\tilde{\mathbf{H}}_{sr}\mathbf{F}_s\mathbf{U}^H\right) + t_4, \end{aligned} \quad (55)$$

where $t_4 = \text{tr}(\mathbf{W}\mathbf{M}\mathbf{W}^H)$ does not depend on \mathbf{U} .

Note that \mathbf{U} must be a lower trigular matrix with unity diagonal elements. This additional constraint can be expressed as

$$\mathbf{S}_i (\mathbf{U} - \mathbf{I}) \mathbf{e}_i = \mathbf{0}, \quad i = 1, 2, \dots, N_d, \quad (56)$$

where $\mathbf{S}_i = [\mathbf{I}_{i \times i}, \mathbf{0}_{i \times (N_d - i)}]$ is the selection matrix, which extracts the first i elements of a N_d -dimensional vector, while \mathbf{e}_i corresponds to the i th column of \mathbf{I}_{N_d} . By using the Lagrangian multiplier method, the optimal THP feedback matrix \mathbf{U} is derived as

$$\begin{aligned} \mathbf{U} = & \mathbf{W}\bar{\mathbf{H}}_{rd}\mathbf{F}_r\bar{\mathbf{H}}_{sr}\mathbf{F}_s \\ & - \sum_{i=1}^{N_d} \mathbf{S}_i^T (\mathbf{S}_i \mathbf{S}_i^T)^{-1} \mathbf{S}_i (\mathbf{W}\bar{\mathbf{H}}_{rd}\mathbf{F}_r\bar{\mathbf{H}}_{sr}\mathbf{F}_s - \mathbf{I}) \mathbf{e}_i \mathbf{e}_i^T. \end{aligned} \quad (57)$$

The detailed derivation is provided in Appendix B.

E. The Complete AO-based Algorithm

To sum up, the original non-convex problem (17) is first decoupled into four subproblems, where each of them only involves one subset of the variables, while the others are fixed. Subsequently, the proposed AO-based nonlinear transceiver design relies on sequentially updating the four matrices \mathbf{F}_r , \mathbf{F}_s , \mathbf{U} and \mathbf{W} in a cyclic, alternating manner. Specifically, given the THP feedback matrix \mathbf{U} , the source precoding matrix \mathbf{F}_s and the AF relaying matrix \mathbf{F}_r , we derive the optimum linear receiver \mathbf{W} in a closed-form. Given \mathbf{U} , \mathbf{F}_s and \mathbf{W} , \mathbf{F}_r can be found by convex optimization. Given \mathbf{U} , \mathbf{F}_r and \mathbf{W} , we transform the non-convex design subproblem of \mathbf{F}_s into a DC program and propose an iterative scheme based on the CCCP to solve the problem. Finally, by fixing the other three matrix variables, \mathbf{U} is derived with the aid of the classic Lagrangian multiplier method. The internal SOCP optimizations needed to implement this AO algorithm are performed by applying the readily available CVX software package. This process is terminated when the desired accuracy is attained or the affordable number of iterations is reached. The complete procedure is summarized in Table II, where N_{\max}^b is the maximum number of iterations and ε denotes the accuracy to be achieved before terminating the algorithm.

TABLE II: Proposed Robust AO Algorithm

- | | |
|-----|---|
| 1. | Initialization: $\mathbf{U}^{(0)} = \mathbf{I}$, the value of $\mathbf{F}_r^{(0)}$ should satisfy (35b), the value of $\mathbf{F}_s^{(0)}$ should satisfy (51b). Then, calculate $\mathbf{W}^{(0)}$ with fixed $\mathbf{U}^{(0)}$, $\mathbf{F}_s^{(0)}$ and $\mathbf{F}_r^{(0)}$, as shown in (19). Set $j = 0$. |
| 2. | Repeat: |
| 2.1 | Update $\mathbf{F}_r^{(j+1)}$ by solving (36) for fixed $\mathbf{W}^{(j)}$, $\mathbf{U}^{(j)}$ and $\mathbf{F}_s^{(j)}$. |
| 2.2 | Update $\mathbf{F}_s^{(j+1)}$ by solving (54) for fixed $\mathbf{F}_r^{(j+1)}$, $\mathbf{W}^{(j)}$ and $\mathbf{U}^{(j)}$, employ the CCCP-based algorithm in Table I. |
| 2.3 | Update $\mathbf{U}^{(j+1)}$ with fixed $\mathbf{F}_s^{(j+1)}$, $\mathbf{F}_r^{(j+1)}$ and $\mathbf{W}^{(j)}$, as shown in (57). |
| 2.4 | Update $\mathbf{W}^{(j+1)}$ with fixed $\mathbf{U}^{(j+1)}$, $\mathbf{F}_s^{(j+1)}$ and $\mathbf{F}_r^{(j+1)}$, as shown in (19). |
| 3. | Until: $\text{MSE}^{(j)} - \text{MSE}^{(j+1)} \leq \varepsilon$ or $j > N_{\max}^b$. Otherwise, set $j := j + 1$ and go to Step 2. |

F. Convergence Analysis and Computational Complexity

1) *Convergence Analysis:* Let us now focus our attention on the global convergence of the proposed AO-based algorithm. As shown in Table II, we update \mathbf{F}_r , \mathbf{F}_s , \mathbf{U} as well as \mathbf{W} sequentially and iteratively. Note that (19) and (57) represent closed-form solutions, while the optimization subproblem (35) is convex and can be formulated as an SOCP problem. As a result, we can readily show that the solutions to these three subproblems are optimal. For the non-convex subproblem (51), the proposed CCCP-based algorithm iteratively approximates the original non-convex feasible set around the current point by a convex set and then solves the resultant convex problem in each iteration. We next introduce the following lemma, which will be used in our analysis:

Lemma 1: The iterations of the CCCP-based algorithm in Table I produce a non-increasing sequence of OF values.

Proof: Please refer to Appendix C. \blacksquare

The importance of the above lemma lies in the fact that, although we are unable to find the optimal solution to the subproblem (51), the monotonicity of the CCCP-based algorithm with respect to \mathbf{F}_s can always be guaranteed. Hence, iteratively updating \mathbf{F}_r , \mathbf{F}_s , \mathbf{U} and \mathbf{W} as proposed may either decrease or maintain the OF value of problem (17), but cannot increase it. Hence, we arrive at a monotonically non-increasing sequence of OF values, as the number of iterations j increases. Specifically, we have

$$\begin{aligned} & \text{MSE} \left(\mathbf{F}_r^{(j+1)}, \mathbf{F}_s^{(j+1)}, \mathbf{U}^{(j+1)}, \mathbf{W}^{(j+1)} \right) \\ & \leq \text{MSE} \left(\mathbf{F}_r^{(j+1)}, \mathbf{F}_s^{(j+1)}, \mathbf{U}^{(j+1)}, \mathbf{W}^{(j)} \right) \\ & \leq \text{MSE} \left(\mathbf{F}_r^{(j+1)}, \mathbf{F}_s^{(j+1)}, \mathbf{U}^{(j)}, \mathbf{W}^{(j)} \right) \\ & \leq \text{MSE} \left(\mathbf{F}_r^{(j+1)}, \mathbf{F}_s^{(j)}, \mathbf{U}^{(j)}, \mathbf{W}^{(j)} \right) \\ & \leq \text{MSE} \left(\mathbf{F}_r^{(j)}, \mathbf{F}_s^{(j)}, \mathbf{U}^{(j)}, \mathbf{W}^{(j)} \right). \end{aligned} \quad (58)$$

Moreover, the sequence of values of the problem (17) is lower-bounded by zero, since the MSE is positive. We can therefore conclude that the proposed AO-based algorithm is always convergent. It is observed that the proposed AO-based algorithm converges to a local optimum of the original problems. The proof is similar to that of Theorem 2 in [42], and we therefore omit the details.

2) *Computational Complexity:* Let us now analyze the computational complexity of the proposed algorithm. Here, we apply the basic concepts of complexity analysis as used in [43]. The complexity of the proposed algorithm is dominated by solving the problem (36) N_{\max}^b times and solving problem (54) $N_{\max}^a N_{\max}^b$ times. Problem (36) involves 3 SOC constraints, including 2 SOCs of dimension $(N_r^2 + 1)$ and 1 SOC of dimension 3. The number of variables is on the order of $m_1 = O(N_r^2)$. Thus, the complexity of the SOCP problem can be expressed as $N_{\max}^b O[m_1(N_r^4 + m_1^2)]$. Similarly, problem (54) involves 4 SOC constraints, including 2 SOCs of dimension $(N_s^2 + 1)$, 1 SOC of dimension $(N_s^2 + 2)$ and 1 SOC of dimension 3. The number of variables is on the order of $m_2 = O(N_s^2)$. The complexity of the CCCP-based algorithm invoked for solving problem (54) is given by $N_{\max}^a N_{\max}^b O[m_2(N_s^4 + m_2^2)]$. Thus, the overall complexity

of the proposed algorithm can be written as

$$N_{\max}^b O[m_1(N_r^4 + m_1^2)] + N_{\max}^a N_{\max}^b O[m_2(N_s^4 + m_2^2)]. \quad (59)$$

Due to the complex nature of the underlying optimization problem and the use of sophisticated computational tools for its solution (i.e., CVX), indeed the proposed algorithm is not deemed to impose a low complexity. However, we would like to emphasize that the complexity of our algorithm is manageable using existing computing technology (see also Section V). While we focus on the formulation and theoretical aspect of the optimization in this manuscript, the search of low-complexity transceiver design algorithm remains an open avenue for our future work.

IV. ROBUST TRANSCIEVER DESIGN BASED ON TRANSMIT POWER MINIMIZATION

In this section, we focus on the robust transceiver design problem (18), where the aim to minimize the transmit power at the source under the MSE and EH constraints. This power minimization problem serves as a complement to the MSE minimization problem.

A. Joint Transceiver Optimization

For fixed \mathbf{F}_r , \mathbf{F}_s and \mathbf{U} , we first optimize the receiver matrix \mathbf{W} . Since \mathbf{W} is not included in the objective function of problem (18) and only involved in the MSE constraint, we seek to minimize the MSE for ensuring that the constraint is satisfied with some slack assuming a feasible starting point. Hence, we can tolerate an increase in MSE for accommodating change in other matrix variables, when we move on to the next parameter \mathbf{F}_r . We can obtain the same solution as the Wiener filter in (19).

We then optimize the AF relaying matrix \mathbf{F}_r by fixing \mathbf{F}_s , \mathbf{U} and \mathbf{W} . Similar to the derivation in (23), the MSE constraint of (18b) can also be represented as a function of \mathbf{f}_r as follows

$$\mathbf{f}_r^H \mathbf{A} \mathbf{f}_r - \mathbf{d}^H \mathbf{f}_r - \mathbf{f}_r^H \mathbf{d} + t_1 \leq \gamma. \quad (60)$$

The EH constraint in (18c) has the same expression as in (32). Note that the OF of the problem (18) does not include \mathbf{f}_r . Hence, directly solving problem (18) with respect to \mathbf{f}_r is not applicable. Instead, here we aim for achieving a non-increasing transmit power at the source in the next iteration. Hence, we can reformulate the following optimization problem

$$\min_{p_1, \mathbf{f}_r} p_1, \quad (61a)$$

$$\text{s.t. } \frac{\mathbf{f}_r^H \mathbf{A} \mathbf{f}_r - \mathbf{d}^H \mathbf{f}_r - \mathbf{f}_r^H \mathbf{d} + t_1}{\gamma} \leq p_1, \quad (61b)$$

$$\frac{\mathbf{f}_r^H \tilde{\mathbf{D}} \mathbf{f}_r}{\bar{P}_r} \leq p_1, \quad (61c)$$

where p_1 is a slack variable. We can observe that once problem (61) is solved, p_1 must be less than 1. This implies that the feasible region of problem (18) with respect to \mathbf{F}_s , \mathbf{U} and \mathbf{W} in the next iteration would become larger, which often results in a reduced transmit power.

Given \mathbf{U} , \mathbf{W} and \mathbf{F}_r , we address the optimization problem (18) with respect to the source precoding matrix \mathbf{F}_s . Similarly, the optimization problem can be expressed as

$$\min_{p_2, \mathbf{f}_s} p_2, \quad (62a)$$

$$\text{s.t. } \mathbf{f}_s^H \mathbf{G} \mathbf{f}_s - \mathbf{Q}^H \mathbf{f}_s - \mathbf{f}_s^H \mathbf{Q} + t_2 \leq \gamma, \quad (62b)$$

$$\mathbf{f}_s^H \mathbf{f}_s \leq p_2, \quad (62c)$$

$$\mathbf{f}_s^H \mathbf{E}_1 \mathbf{f}_s - h(\mathbf{f}_s) \leq 0, \quad (62d)$$

where p_2 is a slack variable. Note that the EH constraint in (62d) is a DC program. To solve this problem, we employ a CCCP-based iterative algorithm similar to that discussed in the Subsection III-C. We approximate the convex function $h(\mathbf{f}_s)$ in the i th iteration by its first-order Taylor expansion denoted as $\hat{h}(\mathbf{f}_s^{(i)}, \mathbf{f}_s)$. Then, in the $(i+1)$ st iteration of the proposed CCCP-based algorithm, we have the following convex optimization problem

$$\min_{p_2, \mathbf{f}_s} p_2, \quad (63a)$$

$$\text{s.t. } \mathbf{f}_s^H \mathbf{G} \mathbf{f}_s - \mathbf{Q}^H \mathbf{f}_s - \mathbf{f}_s^H \mathbf{Q} + t_2 \leq \gamma, \quad (63b)$$

$$\mathbf{f}_s^H \mathbf{f}_s \leq p_2, \quad (63c)$$

$$\mathbf{f}_s^H \mathbf{E}_1 \mathbf{f}_s - \hat{h}(\mathbf{f}_s^{(i)}, \mathbf{f}_s) \leq 0. \quad (63d)$$

Similar to Subsection III-C, the problem (63) can be converted into a standard SOCP. The details of the transformation are omitted here for space-economy.

Given fixed \mathbf{W} , \mathbf{F}_r and \mathbf{F}_s , the THP feedback matrix \mathbf{U} can be obtained by using a similar approach to that of (57).

B. Initial Feasibility Search Algorithm

If problem (18) is initialized with an infeasible point, the algorithm may fail at the first iteration. However, the task of finding a feasible point of a non-convex optimization problem is non-deterministic polynomial hard (NP-hard) [44] in general. Thus, the study of an efficient initial feasibility search algorithm is of substantial importance.

Inspired by the algorithm in [44] and the *phase I* approach of [36], we formulate the feasibility problem as the following convex program

$$\min_{\mathbf{W}, \mathbf{F}_s, \mathbf{F}_r, \mathbf{U}} \tau \quad (64a)$$

$$\text{s.t. } \text{MSE}(\mathbf{W}, \mathbf{F}_s, \mathbf{F}_r, \mathbf{U}) \leq \tau, \quad (64b)$$

$$\text{tr}(\mathbf{F}_r \mathbf{E}[\mathbf{y}_{r,1} \mathbf{y}_{r,1}^H] \mathbf{F}_r^H) \leq \eta \text{tr}(\mathbf{E}[\mathbf{y}_{r,2} \mathbf{y}_{r,2}^H]), \quad (64c)$$

where τ is a slack parameter, which can be regarded as an abstract measure for the violation of the constraint (18b). Instead of minimizing the transmit power at the source as in problem (18), we minimize the OF value τ . The algorithm continues until convergence is achieved or until the maximum number of iterations is reached. If $\tau^{(n+1)}$ achieves the MSE target γ in the $(n+1)$ st iteration, the procedure finds a feasible initial point. If no feasible point can be found using the procedure, we have to adjust γ to an appropriate value.

Note that (64) can be reformulated as

$$\min_{\mathbf{W}, \mathbf{F}_s, \mathbf{F}_r, \mathbf{U}} \text{MSE}(\mathbf{W}, \mathbf{F}_s, \mathbf{F}_r, \mathbf{U}) \quad (65a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{F}_s \mathbf{F}_s^H) \leq P_s^\infty, \quad (65b)$$

$$\text{tr}(\mathbf{F}_r \mathbf{E}[\mathbf{y}_{r,1} \mathbf{y}_{r,1}^H] \mathbf{F}_r^H) \leq \eta \text{tr}(\mathbf{E}[\mathbf{y}_{r,2} \mathbf{y}_{r,2}^H]), \quad (65c)$$

where $P_s^\infty \rightarrow \infty$, which is equivalent to removing the transmit power constraint at the source. Using the connection between the two problems, we summarize the proposed feasibility search algorithm in Table III.

TABLE III: Initial Feasibility Search Algorithm

1.	Set the iteration number $n = 0$.
2.	Repeat: – Solve one cycle of the problem (64) and compute the current objective value by $\tau^{(n+1)}$. – Verify if $\tau^{(n+1)} \leq \gamma$, and if so, stop the algorithm. – Update $n := n + 1$.
3.	Until: $ \tau^{(n)} - \tau^{(n-1)} \leq \varepsilon$.

C. Convergence Analysis and Computational Complexity

Note that the solution of subproblem (61) would enlarge the feasible region of the problem. Consequently the transmit power is reduced for problem (18) with respect to \mathbf{F}_s in the next iteration. Similar to the previous analysis in Subsection III-F, by iteratively updating \mathbf{F}_r , \mathbf{F}_s , \mathbf{U} and \mathbf{W} , the proposed procedure produces a monotonically non-increasing sequence of transmit power values. Thus, the monotonic convergence of the proposed algorithm is ensured.

Proceeding as in Subsection III-F, the overall complexity of the resultant AO-based algorithm for the constrained power minimization problem can be expressed as

$$N_{\max}^b O[m_1(N_r^4 + m_1^2)] + N_{\max}^a N_{\max}^b O[m_2(N_s^4 + m_2^2)], \quad (66)$$

but the details are omitted owing to lack of space.

V. SIMULATION RESULTS

In this section, we present simulation results for validating the performance of the proposed nonlinear transceiver designs. We consider the AF MIMO WPR network illustrated in Fig. 1 in conjunction with $N_s = N_r = N_d = 4$. For each simulation run, the true and estimated source-to-relay and relay-to-destination channels have Rayleigh flat-fading distributions obtained according to the model equations of Section II-A. Specifically, by using the exponential model [45], the covariance matrices of the channel estimation error matrices $\Delta \mathbf{H}_{sr}$ and $\Delta \mathbf{H}_{rd}$ in (10)-(11) can be expressed as

$$\begin{aligned} \mathbf{\Psi}_{sr} = \mathbf{\Psi}_{rd} &= \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 \\ \alpha & 1 & \alpha & \alpha^2 \\ \alpha^2 & \alpha & 1 & \alpha \\ \alpha^3 & \alpha^2 & \alpha & 1 \end{bmatrix} \\ \mathbf{\Sigma}_{sr} = \mathbf{\Sigma}_{rd} &= \sigma_e^2 \begin{bmatrix} 1 & \beta & \beta^2 & \beta^3 \\ \beta & 1 & \beta & \beta^2 \\ \beta^2 & \beta & 1 & \beta \\ \beta^3 & \beta^2 & \beta & 1 \end{bmatrix}, \end{aligned} \quad (67)$$

where α and β denote correlation coefficients, and σ_e^2 is the estimation error variance. The estimated channel matrices, $\bar{\mathbf{H}}_{sr}$ and $\bar{\mathbf{H}}_{rd}$, are generated according to the following distribution-
s,

$$\bar{\mathbf{H}}_{sr} \sim \mathcal{CN}\left(\mathbf{0}_{N_r, N_s}, \frac{(1 - \sigma_e^2)}{\sigma_e^2} \mathbf{\Psi}_{sr} \otimes \mathbf{\Sigma}_{sr}\right) \quad (68)$$

$$\bar{\mathbf{H}}_{rd} \sim \mathcal{CN}\left(\mathbf{0}_{N_d, N_r}, \frac{(1 - \sigma_e^2)}{\sigma_e^2} \mathbf{\Psi}_{rd} \otimes \mathbf{\Sigma}_{rd}\right). \quad (69)$$

It follows from (8) and (9) that the source-to-relay and relay-to-destination channels obey complex Gaussian distributions. In the special case of $\alpha = \beta = 0$, the entries of these matrices have unit variance. It is also assumed that the relay's SI loop channel matrix \mathbf{H}_e in Fig. 1 has i.i.d. complex Gaussian entries with zero-mean. In the simulations, we consider data transmission with 16-QAM modulation. The signal-to-noise ratio (SNR) for the source-to-relay channel is defined as $\text{SNR}_{sr} = P_s / |N_r \sigma_{n_{sr}}^2|$. The maximum transmit power of the source is set to be $P_s = 30$ dBm (or 1 W). The EH efficiency at the relay is assumed to be $\eta = 0.4$. Regarding the stopping criteria in Tables II and III, here we set $\epsilon = 10^{-4}$ and $N_{\max}^b = 20$. All the results are averaged over 10^3 independent channel realizations. Based on the algorithms proposed in this paper, we consider the following benchmark approaches for comparison:

- **Nonlinear Design:** The proposed nonlinear transceiver design consisting of a THP at the source along with an AF relaying matrix and a linear MMSE receiver at the destination.
- **Linear Design:** The corresponding linear transceiver designs with THP replaced by a linear precoder at the source.
- **Perfect Design:** The proposed transceiver designs under the idealized simplifying assumption of having perfect CSI.
- **Robust Design:** The proposed transceiver designs, which take the CSI estimation errors into consideration.
- **Non-robust Design:** Simplified versions of the proposed transceiver designs based on estimated CSI only (that is, knowledge of CSI errors is not used).

A. Results for the MSE Minimization Problem of Section III

In this subsection, we evaluate the performance of the proposed robust nonlinear transceiver design algorithm for MIMO FD-WPR networks, as conceived for the constrained MSE minimization problem of Section III. First, to illustrate the convergence behavior of the proposed robust nonlinear design algorithm, the MSE value of Eq. (12) as a function of the iteration index is shown in Fig. 3 for SNR_{sr} values of 5, 15 and 25 dB. Here we set $\alpha = \beta = 0$ and $\sigma_e^2 = 0.005$. It can be observed in Fig. 3 that the proposed algorithm converges within a reasonable number of iterations for all the SNR_{sr} values considered. Moreover, the convergence rate is faster for lower SNR_{sr} values.

The BER performance of the proposed transceiver design (after convergence of the iterative algorithm) versus SNR_{sr}

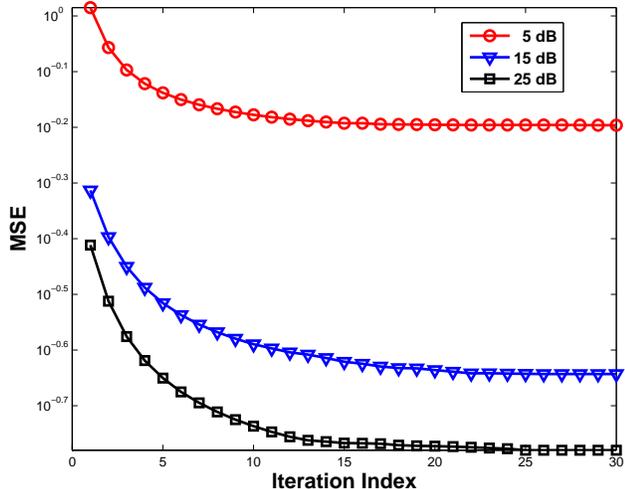


Fig. 3: Convergence behavior of the proposed robust transceiver design based on MSE minimization for different SNR_{sr} .

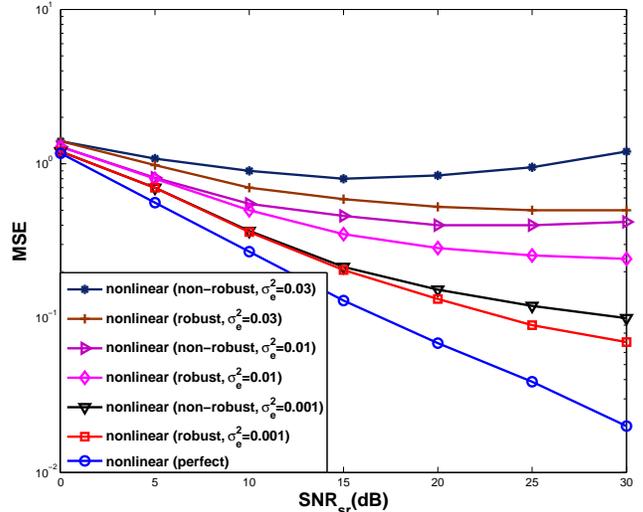


Fig. 5: MSE performance comparison of the proposed transceiver designs versus SNR_{sr} .

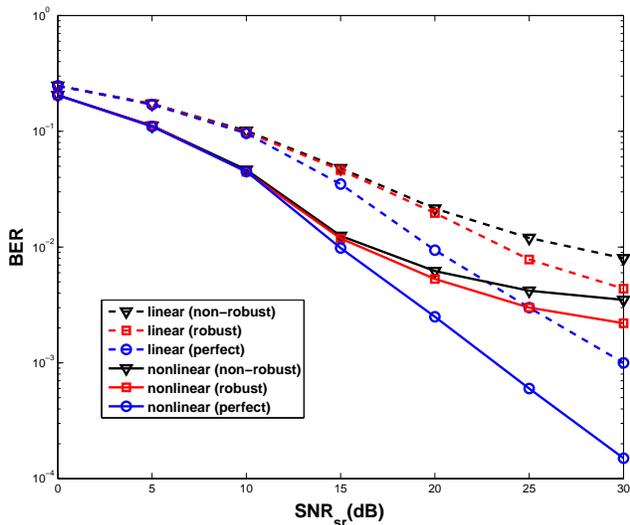


Fig. 4: BER performance comparison of the proposed transceiver designs versus SNR_{sr} .

is shown in Fig. 4. For the case of $\alpha = \beta = 0$ and $\sigma_e^2 = 0.001$, we observe that the BER performance of the nonlinear transceiver design is always superior to that of the linear one. The proposed nonlinear transceiver design with perfect CSI can lead to a 5 dB SNR gain compared with the corresponding linear design at a BER level of 10^{-2} . Naturally, the proposed algorithm with perfect CSI achieves the best BER performance. However, the performance of the proposed robust design, which explicitly take channel errors into consideration, is much better than that of the non-robust design based on estimated channels only. This shows the ability of our proposed design algorithm to take proactive advantage of the CSI errors.

Next, we compare the MSE performance for the proposed nonlinear transceiver design under a correlated channel sce-

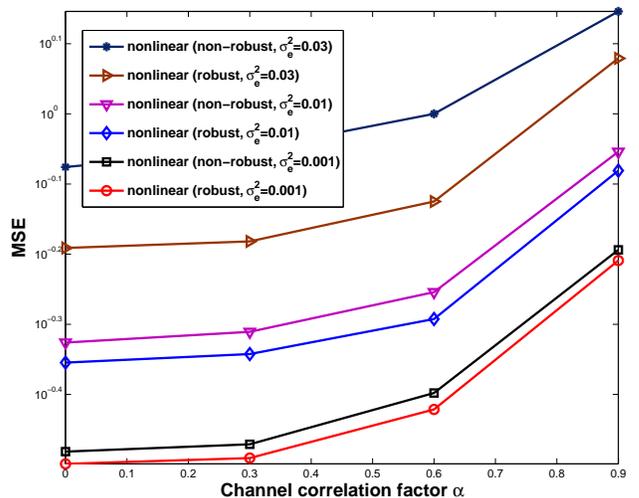


Fig. 6: MSE performance comparison of the proposed transceiver designs versus channel correlation α .

nario associated with $\alpha = \beta = 0.5$. As shown in Fig. 5, smaller estimation errors lead to a better MSE performance for both the robust and non-robust designs. The performance of the proposed robust algorithm is always better than that of the non-robust one. On the other hand, the performance of the non-robust algorithm degrades in the high-SNR region. Furthermore, the performance gap between the proposed robust design and non-robust one becomes larger, as the CSI estimation error σ_e^2 increases.

We further study the effects of channel correlation on the MSE performance of the proposed nonlinear transceiver design in Fig. 6. Here we let $\beta = 0.45$ and $\text{SNR}_{sr} = 10$ dB, whilst varying α . It can be seen that smaller correlation coefficients lead to a better performance. When the value of α increases, the performance of both the robust and non-

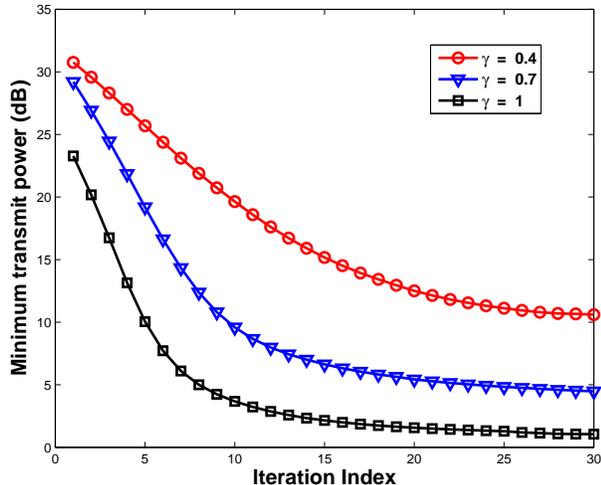


Fig. 7: Convergence behavior of the proposed transceiver design with different MSE targets γ .

robust algorithms is degraded. Naturally, the performance of the proposed robust algorithm is always superior to that of the non-robust one. Furthermore, the performance gain of the proposed robust design over the non-robust one becomes smaller, as α increases. This is reasonable, because a higher channel correlation leads to a reduced spatial multiplexing gain.

B. Results for the Transmit Power Minimization Problem of Section IV

Here, we illustrate the performance of the proposed non-linear transceiver design for the transmit power minimization problem of Section IV. In Fig. 7, we present the transmit power versus the number of iterations when the MSE target γ is set to 0.4, 0.7 and 1, while $\alpha = \beta = 0$ and $\sigma_e^2 = 0.005$. As shown, the proposed robust nonlinear algorithm is capable of converging within about 25 iterations for different MSE requirements. Moreover, the convergence rate is faster for higher MSE target values. When the MSE requirement becomes stricter, a higher transmit power is required for meeting the transmission reliability requirement.

Fig. 8 shows the transmit power at the source versus the MSE target γ . Here, we set $\alpha = \beta = 0.5$, whilst the value of γ is varied. It can be seen that the minimum transmit power required for both the nonlinear and linear algorithms decreases upon increasing γ . Meanwhile, the perfect schemes require a lower transmit power compared to the robust designs for the same γ . This is because the source has to allocate extra power for counteracting the effects of channel uncertainties in order to satisfy the MSE requirement. Furthermore, when compared to the corresponding linear scheme, the proposed nonlinear design requires a reduced transmit power, as expected.

Finally, we present the feasibility rate comparison versus the MSE target γ in Fig. 9, where we set $\alpha = \beta = 0.5$ for the robust transceiver design. In the simulations, a transceiver design algorithm is considered infeasible for a given MSE target if CVX reports an infeasible status, which indicate that

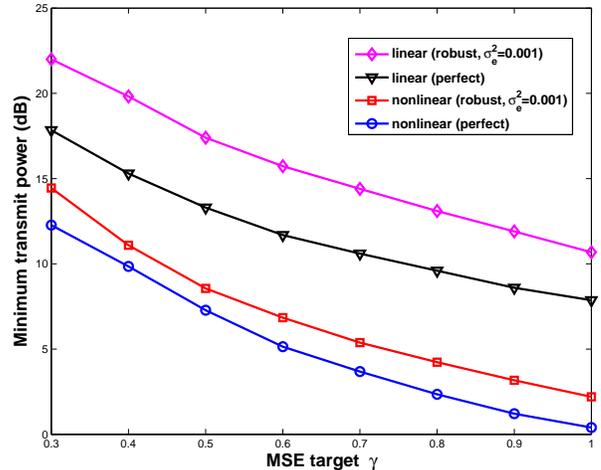


Fig. 8: Minimum transmit power of proposed transceiver design versus MSE target γ .

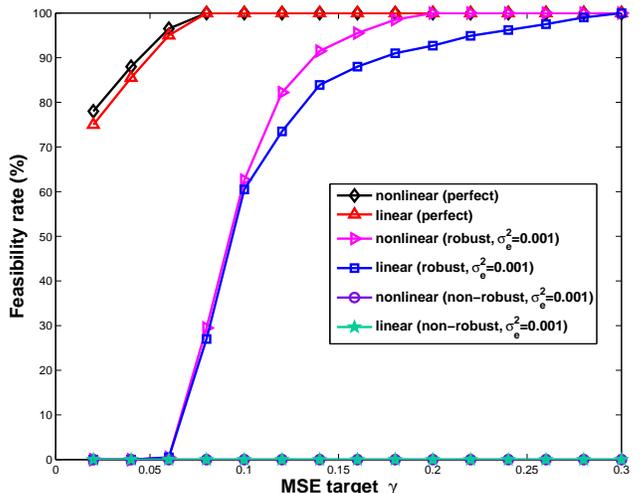


Fig. 9: Comparison of feasibility rate versus MSE target γ for the different algorithms.

the underlying constraints cannot be satisfied. The feasibility rate is formally as the probability of feasibility, but estimated as a ratio based on the given number of simulation runs. It is observed that the feasibility rate of the linear algorithm is inferior to that of the nonlinear algorithm. The transceiver design relying on the idealized simplifying assumption of perfect CSI achieves the highest feasibility rate. The non-robust algorithms fail to satisfy both the MSE and EH constraints most of the time.

VI. CONCLUSIONS

In this paper, we have conceived a robust nonlinear transceiver design for MIMO FD-WPR networks based on a novel SI-energy recycling relaying protocol. Two robust criteria have been considered for the nonlinear transceiver design, namely, MSE minimization under transmit power and EH

constraints, and transmit power minimization under MSE and EH constraints. Since the optimization problems associated with both design criteria are non-convex, we proposed AO-based algorithms for successively optimizing the transceiver coefficients corresponding to the source THP matrix, the AF relaying matrix, the receiver matrix and the feedback matrix. Each subproblem can be solved by employing the SOCP method, the CCCP-based algorithm, the Lagrangian multiplier method or closed-form solutions. Our simulation results have demonstrated the improved robustness of the proposed non-linear transceiver designs against CSI uncertainties.

VII. ACKNOWLEDGEMENT

The authors would like to thank the editor and the anonymous reviewers for their insightful comments and suggestions that helped to improve the quality of the paper.

APPENDIX A

TRANSFORMATION OF (53) INTO A STANDARD SOCP

We will show that problem (53) can be further transformed into a standard SOCP. By introducing

$$\mathbf{g} = -2\tilde{\mathbf{E}}\mathbf{f}_s^{(i)}, \quad (70)$$

$$b = \left(\mathbf{f}_s^{(i)}\right)^H \tilde{\mathbf{E}}\mathbf{f}_s^{(i)} + \sigma_{n_{sr}}^2 \text{tr}(\mathbf{F}_r \mathbf{F}_r^H) - \eta \sigma_{n_{sr}}^2 \text{tr}(\mathbf{H}_e \mathbf{F}_r \mathbf{F}_r^H \mathbf{H}_e^H) - \eta \text{tr}(\sigma_{n_{sr}}^2 \mathbf{I}), \quad (71)$$

(53c) can be rewritten as

$$\mathbf{f}_s^H \mathbf{E}_1 \mathbf{f}_s + \mathcal{R}\{\mathbf{g}^H \mathbf{f}_s\} + b \leq 0, \quad (72)$$

which can be converted into a second-order cone (SOC) constraint [44]

$$\left\| \begin{bmatrix} 2\mathbf{E}_1^{1/2} \mathbf{f}_s \\ -\mathcal{R}\{\mathbf{g}^H \mathbf{f}_s\} - b - 1 \end{bmatrix} \right\| \leq -\mathcal{R}\{\mathbf{g}^H \mathbf{f}_s\} - b + 1. \quad (73)$$

Similarly, we can convert (53a) into an SOC constraint. By introducing the slack variables p_2 and λ_2 , (53a) becomes equivalent to

$$\min_{p_2, \lambda_2, \mathbf{f}_s} p_2, \quad (74a)$$

$$\text{s.t.} \quad \left\| \mathbf{G}^{1/2} \mathbf{f}_s - \mathbf{G}^{-1/2} \mathbf{Q} \right\| \leq \lambda_2, \quad (74b)$$

$$\lambda_2 - \mathbf{Q}^H \mathbf{G}^{-1} \mathbf{Q} + t_2 \leq p_2. \quad (74c)$$

To tackle (74c), we use the following identity,

$$x^2 \leq yz \Leftrightarrow \left\| \begin{bmatrix} 2x \\ y - z \end{bmatrix} \right\| \leq y + z, \quad (75)$$

where $y \geq 0$, $z \geq 0$. Thus, we can apply (75) to transform (74c) into

$$\left\| \begin{bmatrix} 2\lambda_2 \\ p_2 + \mathbf{Q}^H \mathbf{G}^{-1} \mathbf{Q} - t_2 - 1 \end{bmatrix} \right\| \leq p_2 + \mathbf{Q}^H \mathbf{G}^{-1} \mathbf{Q} - t_2 + 1. \quad (76)$$

Substituting (74c) with (76), we can observe that (53a) is now in the form of a standard SOCP.

APPENDIX B

DERIVATION OF FEEDBACK MATRIX \mathbf{U}

First, by ignoring the transmit power and EH constraints, the Lagrangian function is established as

$$\begin{aligned} L(\mathbf{U}) &= \mathbb{E} \left[\|\mathbf{W}\mathbf{Y}_d - \mathbf{v}\|^2 \right] + \sum_{i=1}^{N_d} \mu_i \mathbf{S}_i (\mathbf{U} - \mathbf{I}) \mathbf{e}_i \\ &= \text{tr}(\mathbf{W}\mathbf{M}\mathbf{W}^H) - \text{tr}(\mathbf{U}\mathbf{F}_s^H \bar{\mathbf{H}}_{sr}^H \mathbf{F}_r^H \bar{\mathbf{H}}_{rd}^H \mathbf{W}^H) \\ &\quad + \text{tr}(\mathbf{U}\mathbf{U}^H) - \text{tr}(\mathbf{W}\bar{\mathbf{H}}_{rd} \mathbf{F}_r \bar{\mathbf{H}}_{sr} \mathbf{F}_s \mathbf{U}^H) \\ &\quad + \sum_{i=1}^{N_d} \mu_i \mathbf{S}_i (\mathbf{U} - \mathbf{I}) \mathbf{e}_i, \end{aligned} \quad (77)$$

where μ_i is a Lagrange multiplier. By solving $\partial L(\mathbf{U})/\partial \mathbf{U}^* = \mathbf{0}$, we can obtain

$$\mathbf{U} = \mathbf{W}\bar{\mathbf{H}}_{rd} \mathbf{F}_r \bar{\mathbf{H}}_{sr} \mathbf{F}_s - \sum_{i=1}^{N_d} \mu_i \mathbf{S}_i^T \mathbf{e}_i^T. \quad (78)$$

Substituting \mathbf{U} into the strict lower triangular constraint (55), we have

$$\mu_i = (\mathbf{S}_i \mathbf{S}_i^T)^{-1} \mathbf{S}_i (\mathbf{W}\bar{\mathbf{H}}_{rd} \mathbf{F}_r \bar{\mathbf{H}}_{sr} \mathbf{F}_s - \mathbf{I}) \mathbf{e}_i. \quad (79)$$

Finally, substituting (79) into (78), we obtain the feedback matrix shown in (57).

APPENDIX C

THE PROOF OF LEMMA 1

To proceed, we first denote problem (53) by $\mathcal{P}(\mathbf{f}_s^{(i)})$, and let $\mathbb{S}(\mathbf{f}_s^{(i)})$ and $\mathbb{C}_{\leq}(\mathbf{f}_s^{(i)})$ denote the solution set and constraint set of problem $\mathcal{P}(\mathbf{f}_s^{(i)})$, respectively. Let us now define $R(\mathbf{f}_s) \triangleq \mathbf{f}_s^H \mathbf{E}_1 \mathbf{f}_s - h(\mathbf{f}_s)$ and $\bar{R}(\mathbf{f}_s, \mathbf{f}_s^{(i)}) \triangleq \mathbf{f}_s^H \mathbf{E}_1 \mathbf{f}_s - \bar{h}(\mathbf{f}_s, \mathbf{f}_s^{(i)})$, where we have $h(\mathbf{f}_s) = \mathbf{f}_s^H (\mathbf{E}_2 + \mathbf{E}_3) \mathbf{f}_s + t_3$, and $\bar{h}(\mathbf{f}_s, \mathbf{f}_s^{(i)}) = (\mathbf{f}_s^{(i)})^H (\mathbf{E}_2 + \mathbf{E}_3) \mathbf{f}_s^{(i)} + t_3 + (\mathbf{f}_s^{(i)})^H (\mathbf{E}_2 + \mathbf{E}_3) (\mathbf{f}_s - \mathbf{f}_s^{(i)}) + (\mathbf{f}_s - \mathbf{f}_s^{(i)})^H (\mathbf{E}_2 + \mathbf{E}_3) \mathbf{f}_s^{(i)}$. It follows that $\bar{R}(\mathbf{f}_s, \mathbf{f}_s^{(i)}) = R(\mathbf{f}_s)$. We then complete the proof in two steps.

In the first step, we show that each $\mathbf{f}_s^{(i)}$, $i = 1, 2, \dots$ is feasible to problem (51). It suffices to show that $\mathbf{f}_s^{(i+1)}$ is a feasible solution of problem (51), provided that $\mathbf{f}_s^{(i)}$ is feasible. Assuming that $\mathbf{f}_s^{(i)}$ is feasible to problem (51), we have $\bar{R}(\mathbf{f}_s^{(i)}, \mathbf{f}_s^{(i)}) = R(\mathbf{f}_s^{(i)}) \geq 0$. It follows that there must exist an $\mathbf{f}_s^{(i+1)}$ that is feasible to problem $\mathcal{P}(\mathbf{f}_s^{(i)})$, implying that $\mathbf{f}_s^{(i+1)}$ is such that $\bar{R}(\mathbf{f}_s^{(i+1)}, \mathbf{f}_s^{(i)}) \geq 0$. Furthermore, since $h(\mathbf{f}_s)$ is a convex function in \mathbf{f}_s , we have $h(\mathbf{f}_s) \geq \bar{h}(\mathbf{f}_s, \mathbf{f}_s^{(i)})$ for any \mathbf{f}_s and $\mathbf{f}_s^{(i)}$. It follows that

$$R(\mathbf{f}_s^{(i+1)}) \leq \bar{R}(\mathbf{f}_s^{(i+1)}, \mathbf{f}_s^{(i)}) \leq 0. \quad (80)$$

This implies that $\mathbf{f}_s^{(i+1)}$ is feasible to problem (51). This completes the first step.

In the second step, we show that the OF value sequence $\{P(\mathbf{f}_s^{(i)})\}$ monotonically decreases as the iteration index i increases. According to the first step, $\mathbf{f}_s^{(i)}$ is a feasible solution to $\mathcal{P}(\mathbf{f}_s^{(i)})$. Moreover, in the $(i+1)$ st iteration, $\mathbf{f}_s^{(i+1)}$ is the optimal solution of $\mathcal{P}(\mathbf{f}_s^{(i)})$. Thus, we have $P(\mathbf{f}_s^{(i+1)}) \leq P(\mathbf{f}_s^{(i)})$, which implies the monotonic convergence of $\{P(\mathbf{f}_s^{(i)})\}$, since $P(\mathbf{f}_s)$ is lower-bounded. This completes the proof.

REFERENCES

- [1] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys & Tuts.*, vol. 17, no. 2, pp. 757–789, May 2015.
- [2] H. J. Visser and R. J. Vullers, "RF energy harvesting and transport for wireless sensor network applications: Principles and requirements," *Proc. IEEE*, vol. 101, no. 6, pp. 1410–1423, Jun. 2013.
- [3] X. Zhang, H. Jiang, L. Zhang, C. Zhang, Z. Wang, and X. Chen, "An energy-efficient ASIC for wireless body sensor networks in medical applications," *IEEE Trans. Biomed. Circuits Syst.*, vol. 4, no. 1, pp. 11–18, Feb. 2010.
- [4] X. Lu, D. Niyato, P. Wang, D. I. Kim, and Z. Han, "Wireless charger networking for mobile devices: Fundamentals, standards, and applications," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 126–135, Apr. 2015.
- [5] D. W. K. Ng, E. S. Lo, and R. Schober, "Wireless information and power transfer: Energy efficiency optimization in OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6352–6370, Dec. 2013.
- [6] K. Huang and V. K. Lau, "Enabling wireless power transfer in cellular networks: Architecture, modeling and deployment," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 902–912, Feb. 2014.
- [7] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 418–428, Jan. 2014.
- [8] L. Liu, R. Zhang, and K.-C. Chua, "Multi-antenna wireless powered communication with energy beamforming," *IEEE Trans. Commun.*, vol. 62, no. 12, pp. 4349–4361, Dec. 2014.
- [9] S. Bi, C. Ho, and R. Zhang, "Wireless powered communication: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 117–125, Apr. 2015.
- [10] Q. Liu, J. Wu, P. Xia, S. Zhao, and L. Hanzo, "Distributed laser charging for mobile wireless power transfer - Will it work?," to appear in *IEEE Trans. Veh. Technol.*, 2016.
- [11] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [12] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Wireless-powered relays in cooperative communications: Time-switching relaying protocols and throughput analysis," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1607–1622, May 2015.
- [13] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3622–3636, Jul. 2013.
- [14] Z. Ding, I. Krikidis, B. Sharif, and H. V. Poor, "Wireless information and power transfer in cooperative networks with spatially random relays," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4440–4453, Aug. 2014.
- [15] 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.
- [16] C. Zhong, H. A. Suraweera, G. Zheng, I. Krikidis, and Z. Zhang, "Wireless information and power transfer with full duplex relaying," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3447–3461, Oct. 2014.
- [17] Y. Zeng and R. Zhang, "Full-duplex wireless-powered relay with self-energy recycling," *IEEE Wireless Commun. Lett.*, vol. 4, no. 2, pp. 201–204, Apr. 2015.
- [18] Y. Cheng and M. Pesavento, "Joint optimization of source power allocation and distributed relay beamforming in multiuser peer-to-peer relay networks," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2962–2973, Jun. 2012.
- [19] G. R. Lanckriet and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," *Neural Inf. Process. Syst.*, 2009, pp. 1759–1767.
- [20] W. Guan and H. Luo, "Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Commun. Lett.*, vol. 12, no. 7, pp. 517–519, Jul. 2008.
- [21] F. Tseng and W. Wu, "Linear MMSE transceiver design in amplify-and-forward MIMO relay systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 754–765, Feb. 2010.
- [22] C. Xing, S. Ma, and Y. Wu, "Robust joint design of linear relay precoder and destination equalizer for dual-hop amplify-and-forward MIMO relay systems," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2273–2283, Apr. 2010.
- [23] C. Xing and S. Ma and Z. Fei and Y. Wu and H. V. Poor, "A general robust linear transceiver design for multi-hop amplify-and-forward MIMO relaying systems," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1196–1209, Mar. 2013.
- [24] A. Millar, S. Weiss, and R. Stewart, "Tomlinson Harashima precoding design for non-regenerative MIMO relay networks," in *Proc. IEEE Veh. Technol. Conf.*, May 2011, pp. 1–5.
- [25] F. Tseng, M. Chang, and W. Wu, "Joint Tomlinson–Harashima source and linear relay precoder design in amplify-and-forward MIMO relay systems via MMSE criterion," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1687–1698, May 2011.
- [26] A. Millar, S. Weiss, and R. Stewart, "Robust transceiver design for MIMO relay systems with Tomlinson Harashima Precoding," in *Proc. IEEE EUSIPCO*, Aug. 2012, pp. 1374–1378.
- [27] C. Xing, S. Ma, F. Gao, and Y. Wu, "Robust transceiver with Tomlinson-Harashima precoding for amplify-and-forward MIMO relaying systems," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1370–1382, Sep. 2012.
- [28] F. Tseng, M. Chang, and W. Wu, "Robust Tomlinson-Harashima source and linear relay precoders design in amplify-and-forward MIMO relay systems," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1124–1137, Apr. 2012.
- [29] L. Zhang, Y. Cai, R. de Lamare, and M. Zhao, "Robust multibranch Tomlinson-Harashima precoding design in amplify-and-forward MIMO relay systems," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3476–3490, Oct. 2014.
- [30] M. B. Shenouda and T. N. Davidson, "A framework for designing MIMO systems with decision feedback equalization or Tomlinson-Harashima precoding," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 401–411, Feb. 2008.
- [31] A. Gupta and D. Nagar, *Matrix Variate Distributions*. London, U.K.: Chapman & Hall/CRC, 2000.
- [32] Y. Zeng and R. Zhang, "Optimized training design for wireless energy transfer," *IEEE Trans. Commun.*, vol. 63, no. 2, pp. 536–550, Feb. 2015.
- [33] J. Liu, F. Gao, and Z. Qiu, "Robust transceiver design for downlink multiuser MIMO AF relay systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2218–2231, Apr. 2015.
- [34] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. 25, no. 9, pp. 772–781, Sep. 1978.
- [35] Y. Che, J. Xu, L. Duan, and R. Zhang, "Multiantenna wireless powered communication with cochannel energy and information transfer," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2266–2269, 2015.
- [36] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. press, 2004.
- [37] O. Mehanna, K. Huang, B. Gopalakrishnan, A. Konar and N. Sidiropoulos, "Feasible point pursuit and successive approximation of non-convex QCQPs," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 804–808, Jul. 2015.
- [38] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," Sep. 2013. [Online]. Available: <http://cvxr.com/cvx>.
- [39] R. Horst and N. V. Thoai, "DC programming: Overview," *J. Optim. Theory Appl.*, vol. 103, no. 1, pp. 1–43, Oct. 1999.
- [40] A. J. Smola, S. Vishwanathan, and T. Hofmann, "Kernel methods for missing variables," in *Proc. 10th. Int. Workshop Artif. Intell. Stat.*, Mar. 2005, pp. 325–332.
- [41] Q. Shi, W. Xu, J. Wu, E. Song, and Y. Wang, "Secure beamforming for MIMO broadcasting with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2841–2853, May 2015.
- [42] S. Gong, C. Xing, N. Yang, Y.-C. Wu, and Z. Fei, "Energy efficient transmission in multi-user MIMO relay channels with perfect and imperfect channel state information," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3885–3898, Jun. 2017.
- [43] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.
- [44] J. Yang, B. Champagne, Y. Zou, and L. Hanzo, "Joint optimization of transceiver matrices for MIMO-aided multiuser AF relay networks: Improving the QoS in the presence of CSI errors," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1434–1451, Mar. 2015.
- [45] M. Ding and S. Blostein, "MIMO minimum total MSE transceiver design with imperfect CSI at both ends," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1141–1150, Mar. 2009.