Robust Relay Design for Two-Way Multi-Antenna Relay Systems with Imperfect CSI

Chenyuan Wang, Xiaodai Dong, and Yi Shi

Abstract: **The paper investigates the problem of designing the multiple-antenna relay in a two-way relay network by taking into account the imperfect channel state information (CSI). The objective is to design the multiple-antenna relay based upon the CSI estimates, where the estimation errors are included to attain the robust design under the worst-case philosophy. In particular, the worst-case transmit power at the multiple-antenna relay is minimized while guaranteeing the worst-case quality of service requirements that the received signal-to-noise ratio (SNR) at both sources are above a prescribed threshold value. Since the worst-case received SNR expression is too complex for subsequent derivation and processing, its lower bound is explored instead by minimizing the numerator and maximizing the denominator of the worst-case SNR. The aforementioned problem is mathematically formulated and shown to be nonconvex. This motivates the pursuit of semidefinite relaxation coupled with a randomization technique to obtain computationally efficient high-quality approximate solutions. This paper has shown that the original optimization problem can be reformulated and then relaxed to a convex problem that can be solved by utilizing suitable randomization loop. Numerical results compare the proposed multiple-antenna relay with the existing non-robust method, and therefore validate its robustness against the channel uncertainty. Finally, the feasibility of the proposed design and the associated influencing factors are discussed by means of extensive Monte Carlo simulations.**

Index Terms: **Gaussian randomization, imperfect channel state information (CSI), semidefinite relaxation, two-way relay network, worst-case robust design.**

I. INTRODUCTION

Bidirectional relaying has gained a lot of interests recently for the benefit of higher spectral efficiency [1]–[3]. The two widely used relaying protocols in one-way relaying – amplify-and-forward (AF) and decode-and-forward (DF) – are naturally inherited by two-way relaying. Attracted by the benefits of lower complexity and easier implementation, AF protocol is more desirable for practical consideration compared to DF. Typically, AF two-way relaying consists of two sources transmitting information simultaneously to the relay in the first phase, and the relay amplifying the received signal and broadcasting in the second phase. The process of linear amplifying the sum signal received from both sources and then retransmitting the resulting signal is also referred to as analogy network coding (ANC) [3]–[5]. The essential of ANC relies on the observation that the collision at the relay in the first phase is totally harmless, and that the so-called self-interference can be removed from the received signals at the sources before data detection since both sources know their own transmitted signals.

Due to many practical issues, the channel state information (CSI) is usually imperfectly known at the transmitter and/or the receiver. As has been well recognized, optimal designs based on perfect CSI assumptions are very sensitive to channel errors [6]. If unaccounted for, CSI imprecision can lead to severe system performance degradation [7], [8]. Typically, there are two classes of models to characterize imperfect CSI: The stochastic model and the worst-case model. The stochastic model usually assumes the channel to be random with a known distribution, and the slowly-varying channel statistics, such as the mean or channel covariance, can be well estimated [9], [10]. System designs are typically based on optimizing average performance like the mean square error at the receiver. The worst-case model, on the other hand, assumes that a nominal value of the instantaneous channel is available, but lies in a bounded uncertainty region defined by some norm [11]. Accordingly, the system design is based on optimizing the worst-case performance (among all the possibilities in the uncertainty region), so that certain quality of service (QoS) can be guaranteed. An analysis of the penalty for using imperfect CSI would be of significant benefit to system designers.

In effect, developing optimal designs that are robust to CSI imprecision are not uncommon in the signal processing literature. Existing works in one-way relaying include, but not limited to [12]–[16]. In [12], Chalise *et al.* provide a robust design of multiple-input-multiple-output (MIMO) relay precoder taking into account the channel uncertainty for a system with multiple source-destination pairs assisted by a single multiple-antenna relay station, and further extend this work to multiple multi-antenna relays in [13]. In [14], the same system as in [12] is considered, and the multiple-antenna relay precoder as well as the destination filters are jointly designed to provide robustness to errors in CSI. Recent works on robust design for MIMO two-way relay systems include [9], [17], and [18]. In [17], the effects of channel estimation error on the transmission rate of MIMO two-way AF relay links are investigated and a lower bound is derived via the worst-case noise theorem. Gharavol, *et al.* in [18] address the joint source and relay optimization for a MIMO two-way relay channel with the objective of minimizing the worst-case sum mean squared error. This problem has also been studied in [9] but using the a stochastic approach. Other related works on worst-case robust design include multiple-antenna relay designs for downlink broadcast channel

Manuscript received January 2, 2013; approved for publication by Chung, Sae-Young, Division II Editor, December 1, 2013.

Chenyuan Wang is with QUALCOMM Inc., CA, USA. email: chenyuan@ece.uvic.ca.

Xiaodai Dong is with the Department of Electrical and Computer Engineering, University of Victoria, B.C. Canada. email: xdong@ece.uvic.ca.

Yi Shi is with Huawei Technologies, Beijing, China. email: wn.shiyi@huawei.com.

Digital object identifier 10.1109/JCN.2014.000008

[19], [20] and one-way single-user or multiuser MIMO system [15], [16].

Given the research significance of the robust design in the presence of imperfect CSI in two-way multiple-antenna relay network, we consider an AF two-way multiple-antenna relay network with a pair of communicating single-antenna sources and a single multi-antenna relay, and propose a robust design of the AF multiple-antenna relay precoder with imperfect CSI at both sources and relay. As in [11]–[16], [18]–[20], we have adopted the worst-case design and have taken into account both the imperfect self-interference cancellation and imperfect data detection due to CSI estimation errors. Our design target is to achieve the minimum system power consumption while fulfilling the instantaneous single-to-noise ratio (SNR) constraints at both sources. We formulate the worst-case received SNR at both sources, which is too complex for further analysis. So, we resort to its lower bound by minimizing the numerator while maximizing the denominator of the worst-case received SNR. The design optimization problem is then formulated but turns out to be nonconvex. By utilizing semidefinite relaxation (SDR) based approximation, a sub-optimal and robust solution to the original program can be obtained and further verified through numerical results. This design methodology can be applied to solve the non-robust design problem and other robust problems.

The remainder of the paper is organized as follows. Section II introduces the system model under consideration. Section III presents the channel uncertainty model as well as the optimization formulation with imperfect CSI. The robust design of the multiple-antenna relay in the presence of channel estimation error is proposed in Section IV, and the semidefinite relaxation as well as the randomization technique are exploited to solve the optimal multiple-antenna relay precoder. Numerical results are given in Section V, and Section VI concludes this paper.

Notation: Boldface uppercase/lowercase letters denote matrices/vectors. The superscripts $(\cdot)^T$ and $(\cdot)^H$ denote transpose and Hermitian transpose. tr$(\cdot)$, vec$(\cdot)$, $\|\cdot\|$, and rank$(\cdot)$ denote the trace, the vectorization, the Euclidean norm, and the rank operators, respectively. $\otimes$, $\odot$, $\mathcal{R}\{\}$, and $\mathcal{I}\{\}$ denote Kronecker product, elementwise product, the real part and the imaginary part, respectively. By $\mathbf{X} \succeq \mathbf{0}$ we denote that $\mathbf{X}$ is a Hermitian positive-semidefinite matrix. Finally, $\mathbf{I_N}$, $\mathbf{1}_{\mathbf{1}\times\mathbf{N}}$, and $\mathcal{C}^{\mathcal{N}\times\mathcal{N}}$ denote the $N\times N$ identity matrix, the $N\times 1$ all ones column vector, and the space of $N \times N$ matrices with complex entries, respectively.

II. SYSTEM MODEL

We consider a two-way AF relay-assisted system consisting of two single-antenna sources A and B exchanging information with the help of a single relay with N antennas, as shown in Fig. 1. There is no direct link between A and B, and both sources and the relay are operated under half-duplex mode which means they cannot transmit and receive at the same time. We assume that the two-way relay network (TWRN) is a time division duplexing (TDD) system, which means both the uplink and downlink channels occupy the same frequency slot, but are differentiated in a time duplex manner in information exchange. Moreover, the propagation channel is assumed to be constant and reciprocal between two consecutive time slots.

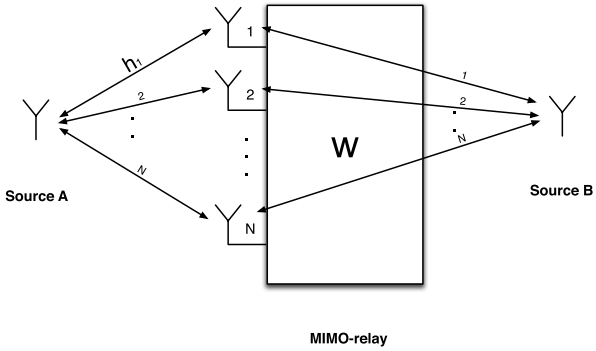

Fig. 1. System model for an AF two-way MIMO relay network.

A. Data Model

In the ANC-based system herein, we consider a two-phase cooperative strategy where the first phase involves the pair of sources transmitting simultaneously. At the multiple-antenna relay, the received sum signal is linearly processed and then broadcasted in the second phase. Since TDD is assumed, the signal received at the relay station in the first time slot is

$$\mathbf{r_R} = \mathbf{h x_A} + \mathbf{g x_B} + \mathbf{n_R} \tag{1}$$

where x_A and x_B denote the transmitted signals with unit average energy, i.e., $\mathbb{E}[|x_A|^2] = 1$ and $\mathbb{E}[|x_B|^2] = 1$, and unit transmit power is assumed at both sources A and B. Variables $\mathbf{h} = [\mathbf{h_1}, \cdots, \mathbf{h_N}]^\mathbf{T}$ and $\mathbf{g} = [\mathbf{g_1}, \cdots, \mathbf{g_N}]^\mathbf{T}$ are discrete baseband equivalent channel coefficient vectors for both sides of the link to the multiple-antenna relay, as indicated in Fig. 1. Vector $\mathbf{n_R} = [\mathbf{n_R^{(1)}}, \cdots, \mathbf{n_R^{(N)}}]^\mathbf{T}$ is circularly symmetric additive white Gaussian noise (AWGN) at the relay antennas. Each entry of $\mathbf{n_R}$ is independent and modeled as complex Gaussian random variable (CGRV) with zero-mean and the same variance N_R, i.e., $\mathbf{n_R} \sim \mathcal{CN}(\mathbf{0}, \mathbf{N_R I_N})$.

We assume that both uplink (from source to relay) and downlink (from relay to source) CSI are available at the relay station to serve the precoder design. The relay can employ the pilots from both sources to estimate the uplink channels, and use them as the estimates in the downlink due to the channel reciprocity principle. In the second time slot, the relay linearly processes the received signal by multiplying it with the precoder matrix $\mathbf{W}$, and then forwards the following signal to the sources,

$$\mathbf{s_R} = \mathbf{W} \cdot \mathbf{r_R}. \tag{2}$$

Note that the transmit power at the multiple-antenna relay can be derived from (2) as

$$P_R = \text{tr}\left(\mathbb{E}[\mathbf{s_R s_R^H}]\right) = \text{tr}\left(\mathbf{W}\left(\mathbf{h h^H} + \mathbf{g g^H} + \mathbf{N_R I_N}\right)\mathbf{W^H}\right) \tag{3}$$

where $\mathbb{E}[\text{□}\text{□}^\mathbb{H}] = \text{□}\ (k \in \{A, B\})$ and $\mathbb{E}[\mathbf{n_R n_R^H}] = \mathbf{N_R I_N}$. The signals received by the two sources in scalar form are

$$y_A = \mathbf{h^T s_R} + \mathbf{n_A} = \underbrace{\mathbf{h^T W g x_B}}_{\text{Signal part}} + \underbrace{\mathbf{h^T W h x_A}}_{\text{Self-interference}} + \underbrace{\mathbf{h^T W n_R} + \mathbf{n_A}}_{\text{Noise part}}, \tag{4a}$$

$$y_B = \mathbf{g^T s_R} + \mathbf{n_B} = \underbrace{\mathbf{g^T W h x_A}}_{\text{Signal part}} + \underbrace{\mathbf{g^T W g x_B}}_{\text{Self-interference}} + \underbrace{\mathbf{g^T W n_R} + \mathbf{n_B}}_{\text{Noise part}} \tag{4b}$$

where n_A and n_B denote the AWGN with variance N_A and N_B, respectively. Let $\hat{\mathbf{h}}$ and $\hat{\mathbf{g}}$ denote the channel estimates. We assume the relay forwards the CSI estimates as well as the precoder matrix $\mathbf{W}$ to the sources through the feedback channels. The quantization error as well as the error in the feedback channels is not considered in this paper, so the CSI estimates used for the source-relay reciprocal channels are identical, and the sources have a copy of $\mathbf{W}$ for the further process of the received signals. Then the self-interference can be cancelled as $z_A = y_A - \hat{\mathbf{h}}^T \mathbf{W} \hat{\mathbf{h}} \mathbf{x_A}$ and $z_B = y_B - \hat{\mathbf{g}}^T \mathbf{W} \hat{\mathbf{g}} \mathbf{x_B}$. If perfect channel condition is assumed, self-interference can be totally removed since the sources know their own transmitted data. Nevertheless, when the CSI is imperfect, self-interference cannot be removed completely due to the fact that the channel estimates are not always equal to the actual channels in practice. To this end, we consider the case where imperfect self-interference cancellation leads to residual self-interference. We show later in this paper that residual self-interference can be ignored under some reasonable assumption.

For the convenience of future analysis, define $\mathbf{H_{DL}} = [\mathbf{g}\ \mathbf{h}]^{\mathbf{T}}$ and $\mathbf{H_{UL}} = [\mathbf{h}\ \mathbf{g}]$, and we can represent both (4a) and (4b) in an equivalent matrix form as

$$\mathbf{y} = \mathbf{H_{DL} W H_{UL} x} + \mathbf{H_{DL} W n_R} + \mathbf{n} \tag{5}$$

where $\mathbf{y} = [\mathbf{y_B}\ \mathbf{y_A}]^{\mathbf{T}}$, $\mathbf{x} = [\mathbf{x_A}\ \mathbf{x_B}]^{\mathbf{T}}$, and $\mathbf{n} = [\mathbf{n_B}\ \mathbf{n_A}]^{\mathbf{T}}$.

B. End-to-End SNR and Relay Power

In this paper, we only mathematically analyze the transmission from B to A, and similar method can be applied to that from A to B.

First, the following operation is defined to convert a column vector $\mathbf{f} = [\mathbf{f_1}, \cdots, \mathbf{f_N}]^{\mathbf{T}}$ with N entries to two new row vectors $\tilde{\mathbf{f}}_1$ and $\tilde{\mathbf{f}}_2$ with N^2 entries each as

$$\begin{aligned} \tilde{\mathbf{f}}_1 &= \mathbf{f^T} \left(\mathbf{I_N} \otimes \mathbf{1_{1\times N}}\right)_{\mathbf{N\times N^2}} \triangleq \mathbf{f^T D_1} = [\mathbf{f_1} \cdots \mathbf{f_1} \cdots \mathbf{f_N} \cdots \mathbf{f_N}], \\ \tilde{\mathbf{f}}_2 &= \mathbf{f^T} \left(\mathbf{1_{1\times N}} \otimes \mathbf{I_N}\right)_{\mathbf{N\times N^2}} \triangleq \mathbf{f^T D_2} = [\mathbf{f_1} \cdots \mathbf{f_N} \cdots \mathbf{f_1} \cdots \mathbf{f_N}]. \end{aligned} \tag{6}$$

So, the received signal at source A in (4a) can be rewritten as

$$\begin{aligned} y_A =& (\tilde{\mathbf{g}}_1 \odot \tilde{\mathbf{h}}_2)\mathbf{w_L x_B} + (\tilde{\mathbf{h}}_1 \odot \tilde{\mathbf{h}}_2)\mathbf{w_L x_A} \\ &+ (\tilde{\mathbf{n}}_{R,1} \odot \tilde{\mathbf{h}}_2)\mathbf{w_L} + \mathbf{n_A} \end{aligned} \tag{7}$$

where $\tilde{\mathbf{h}}_1 = \mathbf{h^T D_1}$, $\tilde{\mathbf{h}}_2 = \mathbf{h^T D_2}$, $\tilde{\mathbf{g}}_1 = \mathbf{g^T D_1}$, and $\mathbf{w_L} = \text{vec}(\mathbf{W})$. Note that the vec$(\mathbf{W})$ operation stacks the columns of $\mathbf{W}$ each by each, and finally converts the matrix $\mathbf{W}$ to a column vector as $[w_{11}\ w_{21} \cdots w_{N1}\ \cdots\ w_{1N}\ w_{2N} \cdots w_{NN}]^T$. After cancelling the self-interference, the signal at source A becomes

$$\begin{aligned} z_A =& \underbrace{(\tilde{\mathbf{g}}_1 \odot \tilde{\mathbf{h}}_2)\mathbf{w_L x_B}}_{\text{Signal part } r_A} + \underbrace{(\tilde{\mathbf{n}}_{R,1} \odot \tilde{\mathbf{h}}_2)\mathbf{w_L} + \mathbf{n_A}}_{\text{Noise part } t_n} \\ &+ \underbrace{\left((\tilde{\mathbf{h}}_1 \odot \tilde{\mathbf{h}}_2) - (\hat{\tilde{\mathbf{h}}}_1 \odot \hat{\tilde{\mathbf{h}}}_2)\right)\mathbf{w_L x_A}}_{\text{Interference part } t_s}. \end{aligned} \tag{8}$$

The received SNR at source A after self-interference cancellation is obtained as

$$\gamma_A = \frac{\mathbf{w_L^H} \left(\tilde{\mathbf{g}}_\mathbf{1} \odot \tilde{\mathbf{h}}_\mathbf{2}\right)^{\mathbf{H}} \left(\tilde{\mathbf{g}}_\mathbf{1} \odot \tilde{\mathbf{h}}_\mathbf{2}\right) \mathbf{w_L}}{t_n + t_s} \triangleq \frac{r_A}{s_A} \tag{9}$$

where

$$\begin{aligned} t_n =&\ \mathbf{w_L^H}(\mathbf{N^r} \odot (\tilde{\mathbf{h}}_\mathbf{2}^\mathbf{H} \tilde{\mathbf{h}}_\mathbf{2}))\mathbf{w_L} + \sigma_{\mathbf{n_A}}^\mathbf{2}, \\ t_s =&\ \mathbf{w_L^H}(\tilde{\mathbf{h}}_\mathbf{1} \odot \tilde{\mathbf{h}}_\mathbf{2})^\mathbf{H}(\tilde{\mathbf{h}}_\mathbf{1} \odot \tilde{\mathbf{h}}_\mathbf{2})\mathbf{w_L} \\ &+ \mathbf{w_L^H}(\hat{\tilde{\mathbf{h}}}_\mathbf{1} \odot \hat{\tilde{\mathbf{h}}}_\mathbf{2})^\mathbf{H}(\hat{\tilde{\mathbf{h}}}_\mathbf{1} \odot \hat{\tilde{\mathbf{h}}}_\mathbf{2})\mathbf{w_L} \\ &- \mathbf{w_L^H}(\tilde{\mathbf{h}}_\mathbf{1} \odot \tilde{\mathbf{h}}_\mathbf{2})^\mathbf{H}(\hat{\tilde{\mathbf{h}}}_\mathbf{1} \odot \hat{\tilde{\mathbf{h}}}_\mathbf{2})\mathbf{w_L} \\ &- \mathbf{w_L^H}(\hat{\tilde{\mathbf{h}}}_\mathbf{1} \odot \hat{\tilde{\mathbf{h}}}_\mathbf{2})^\mathbf{H}(\tilde{\mathbf{h}}_\mathbf{1} \odot \tilde{\mathbf{h}}_\mathbf{2})\mathbf{w_L}. \end{aligned} \tag{10}$$

Note that $\mathbf{N^r} = \mathbb{E}[(\tilde{\mathbf{n}}_{\mathbb{R},\mathbb{\mu}})^{\mathbb{H}} \tilde{\mathbf{n}}_{\mathbb{R},\mathbb{\mu}}] \in \mathcal{C}^{\mathcal{N}^\in \times \mathcal{N}^\in}$ which can also be expressed as $\mathbf{N^r} = \mathbf{N_R}(\mathbf{I_N} \otimes \mathbf{1_{N\times N}})$.

Also, by using some identical transformation the transmit power in (3) can be further rewritten in terms of $\mathbf{w_L}$ and $\mathbf{H_{UL}}$ as

$$\begin{aligned} P_R &= \text{tr}\left(\mathbf{W^H W}\left(\mathbf{H_{UL} H_{UL}^H} + \sigma_{\mathbf{n_R}}^\mathbf{2}\mathbf{I_N}\right)\right) \\ &= \text{vec}\left(\mathbf{W}\right)^H \left(\left(\mathbf{H_{UL} H_{UL}^H} + \sigma_{\mathbf{n_R}}^\mathbf{2}\mathbf{I_N}\right)^T \otimes \mathbf{I_N}\right) \text{vec}\left(\mathbf{W}\right) \\ &\triangleq \mathbf{w_L^H C w_L} \end{aligned} \tag{11}$$

where the identical transformation used in the derivation include tr$(\mathbf{AB})$ = tr$(\mathbf{BA})$, tr$(\mathbf{A^H B})$ = vec$(\mathbf{A})^\mathbf{H}$vec$(\mathbf{B})$, and vec$(\mathbf{AB}) = \left(\mathbf{B^T} \otimes \mathbf{I}\right)vec(\mathbf{A})$.

C. Non-Robust Design with Perfect CSI

In this subsection, we will present a quick review of the optimal multiple-antenna relay structure in the perfect CSI case. Since the SNR at the source has been derived, the power minimization problem under SNR constraints is then straightforward. Similar work has been done in [21], and an efficient algorithm is also provided to compute the optimal precoder matrix. But the design in [21] is based on the ideal situation where perfect CSI is available, which we will refer to as the non-robust method in the following. The non-robust precoder structure will provide us a baseline for our robust design, and some comparison will be given in Section V.

If the perfect channel is assumed, the self-interference can be completely removed and therefore t_s in (10) is equal to zero. Now, the optimal precoder design with the actual channel knowledge at the relay can be formulated as the optimization problem with the objective of minimizing the transmit power at the relay under the constraints of satisfying the prescribed SNR threshold value at both sources, i.e.,

$$\begin{aligned} \mathcal{P}^{\text{nrob}}:\ & \min_{\mathbf{w_L}} P_R \\ & \text{s.t. } \gamma_k \geq \gamma_{\text{th}},\ k \in \{A, B\}. \end{aligned} \tag{12}$$

It should be noted that $\mathcal{P}^{\text{nrob}}$ is not always feasible when the target SNR γ_{th} is set too high [22].

As far as the purpose of this paper is concerned, it is too optimistic to assume the precise knowledge of CSI at the relay,

and estimation errors should be taken into account to guarantee the robustness of multiple-antenna relay design. Hence, channel uncertainty ought to be incorporated in both objectives and the constraints in (12), and modeling the channel uncertainty along with the problem of designing the robust multiple-antenna relay is investigated in the following two sections.

III. CHANNEL UNCERTAINTY MODEL

In the aforementioned system, the CSI is estimated once at the multiple-antenna relay, and then forwarded to both sources through the feedback channels. We assume that the feedback channel is error-free, has zero delay, and that the quantization of the channel estimates at the relay is perfect. So, AWGN is the main source of CSI error, which is referred to as channel uncertainty in the following part of this paper. Note that the multiple-antenna relay here uses the uplink CSI estimates as that of downlink to design the precoder since the uplink and downlink is reciprocal to each other.

Let the estimates of the channel coefficients for both links to the multiple-antenna relay be $\hat{\mathbf{h}} = [\hat{h}_1, \cdots, \hat{h}_N]^T$ and $\hat{\mathbf{g}} = [\hat{g}_1, \cdots, \hat{g}_N]^T$. We assume that both entry pairs h_i, $\hat{h}_i$ and g_i, $\hat{g}_i$ can be modeled as jointly ergodic and stationary Gaussian processes. The relations between the actual channels and the channel estimates are given by

$$\mathbf{h} = \hat{\mathbf{h}} + \mathbf{e_h} \text{ and } \mathbf{g} = \hat{\mathbf{g}} + \mathbf{e_g} \tag{13}$$

where the channel estimation errors are $\mathbf{e_h} = [\mathbf{e_h^{(1)}} \mathbf{e_h^{(2)}} \cdots \mathbf{e_h^{(N)}}]^{\mathbf{T}}$ and $\mathbf{e_g} = [\mathbf{e_g^{(1)}} \mathbf{e_g^{(2)}} \cdots \mathbf{e_g^{(N)}}]^{\mathbf{T}}$. We consider the entries of $\mathbf{e_h}$ and $\mathbf{e_g}$ are independent and identically distributed (i.i.d.) following the distribution of $\mathcal{CN}(0, \sigma_{e_h}^2)$ and $\mathcal{CN}(0, \sigma_{e_g}^2)$, respectively [23].

We further assume that both $\mathbf{e_h}$ and $\mathbf{e_g}$ are norm-bounded as

$$\|\mathbf{e_h}\| \le \delta_h \text{ and } \|\mathbf{e_g}\| \le \delta_g \tag{14}$$

where $\|\cdot\|$ denotes the Euclidean norms of a vector. Equivalently, the uncertainty region for each channel can be specified as [15]

$$\mathcal{R}_{\mathbf{h}} = \{\zeta | \zeta = \hat{\mathbf{h}} + \mathbf{e}_h, \|\mathbf{e}_h\| \le \delta_h\}, \tag{15}$$

$$\mathcal{R}_{\mathbf{g}} = \{\zeta | \zeta = \hat{\mathbf{g}} + \mathbf{e}_g, \|\mathbf{e}_g\| \le \delta_g\}. \tag{16}$$

The relation between the error variance and an upper bound of the norm is also discussed in [15]. A more definite mathematical method is provided in [12] to determine $\delta_h(\delta_g)$ from $\sigma_{e_h}^2(\sigma_{e_g}^2)$ by means of numerical search, which will be summarized with δ_h and $\sigma_{e_h}^2$ as an example.

Let T be a random variable (r.v.) following the chi-square distribution with N degrees of freedom with the probability density function (PDF) denoted by $f_T(t, N)$. Define $\|\mathbf{e_h}\|^2 = \sigma_{e_h}^2 T$, $\Pr\{\sigma_{e_h}^2 T \le \delta_h^2\}$ is some predefined bounding probability and given by [15]

$$\begin{aligned} \Pr\{\sigma_{e_h}^2 T \le \delta_h^2\} &= \int_0^{\delta_h^2/\sigma_{e_h}^2} f_T(t, N) \mathrm{dt} \\ &= \frac{1}{\Gamma(N/2)} \gamma\left(\frac{N}{2}, \frac{\delta_h^2}{2\sigma_{e_h}^2}\right) \end{aligned} \tag{17}$$

where $\Gamma(\cdot)$ and $\gamma(\cdot)$ are the complete and lower complete Gamma functions, respectively. Moreover, we choose the upper bound δ to guarantee the bounding probability achieving a value of $1 - \exp(-c/\sigma_{e_h}^2)$, where c is a positive constant,. In this way, the achieved bounding probability scales with the channel estimation error variance. Therefore, given the estimation error variance, the upper bound of the norm of the channel uncertainty can be numerically determined.

Generally speaking, channel estimation error is much smaller compared to the actual channel, so it is reasonable to not take the second-order terms of $\mathbf{e_h}$ and $\mathbf{e_g}$ and their cross-product into account. By applying the operation in (6) to the channel estimate vectors, $\hat{\mathbf{h}}$ and $\hat{\mathbf{g}}$, as well as the estimation error vectors, $\mathbf{e_h}$ and $\mathbf{e_g}$, we can obtain

$$\tilde{\mathbf{h}}_j = \hat{\tilde{\mathbf{h}}}_{\mathbf{j}} + \tilde{\mathbf{e}}_{\mathbf{h,j}} \text{ and } \tilde{\mathbf{g}}_{\mathbf{j}} = \hat{\tilde{\mathbf{g}}}_{\mathbf{j}} + \tilde{\mathbf{e}}_{\mathbf{g,j}}, \ (\mathbf{j} = \mathbf{1}, \mathbf{2}). \tag{18}$$

Then, we substitute (18) into (9), the received SNR at source A under imperfect channel condition γ_A^{rob} in terms of channel estimates and estimation error can be determined. After ignoring the second-order terms and cross-product of channel estimation errors, both the numerator and denominator of γ_A^{rob} are given by

$$\begin{aligned} r_A^{\mathrm{rob}} &\approx \mathrm{tr}(\mathbf{w_L w_L^H a^H a}) \\ &\quad + \mathrm{tr}\Big(\mathbf{w_L w_L^H}(\mathbf{a^H}(\hat{\tilde{\mathbf{g}}}_{\mathbf{1}} \odot \tilde{\mathbf{e}}_{\mathbf{h,2}}) + (\hat{\tilde{\mathbf{g}}}_{\mathbf{1}} \odot \tilde{\mathbf{e}}_{\mathbf{h,2}})^{\mathbf{H}}\mathbf{a})\Big) \\ &\quad + \mathrm{tr}\Big(\mathbf{w_L w_L^H}(\mathbf{a^H}(\hat{\tilde{\mathbf{h}}}_{\mathbf{2}} \odot \tilde{\mathbf{e}}_{\mathbf{g},1}) + (\hat{\tilde{\mathbf{h}}}_{\mathbf{2}} \odot \tilde{\mathbf{e}}_{\mathbf{g},1})^{\mathbf{H}}\mathbf{a})\Big), \end{aligned} \tag{19a}$$

$$\begin{aligned} s_A^{\mathrm{rob}} &= t_n^{\mathrm{rob}} + t_s^{\mathrm{rob}} = t_n^{\mathrm{rob}} \\ &\approx \mathrm{tr}\Big(\mathbf{w_L w_L^H}(\mathbf{N^r} \odot (\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}^{\mathbf{H}} \hat{\tilde{\mathbf{h}}}_{\mathbf{2}}))\Big) \\ &\quad + \mathrm{tr}\Big(\mathbf{w_L w_L^H}(\mathbf{N^r} \odot (\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}^{\mathbf{H}} \tilde{\mathbf{e}}_{\mathbf{h,2}}))\Big) \\ &\quad + \mathrm{tr}\Big(\mathbf{w_L w_L^H}(\mathbf{N^r} \odot (\tilde{\mathbf{e}}_{\mathbf{h,2}}^{\mathbf{H}} \hat{\tilde{\mathbf{h}}}_{\mathbf{2}}))\Big), \end{aligned} \tag{19b}$$

respectively, where $\mathbf{a} = \hat{\tilde{\mathbf{g}}}_{\mathbf{1}} \odot \hat{\tilde{\mathbf{h}}}_{\mathbf{2}}$ and $t_s^{\mathrm{rob}} = 0$. It can be easily shown that the following relations hold true:

$$\mathbf{w_L w_L^H}\,(\mathbf{x} \odot \mathbf{y})\,(\mathbf{x} \odot \mathbf{z})^{\mathbf{H}} = \mathbf{z}\left((\mathbf{x}\mathbf{x})^{\mathbf{T}} \odot \mathbf{w_L w_L^H}\right)\mathbf{y}, \tag{20a}$$

$$\mathrm{tr}\left(\mathbf{w_L w_L^H}\left((\mathbf{x^H y}) \odot \mathbf{Z}\right)\right) = \mathbf{y}\left(\mathbf{Z^T} \odot \mathbf{w_L w_L^H}\right)\mathbf{x^H} \tag{20b}$$

where $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{C}^{\infty \times \mathcal{N}^{\in}}$ and $\mathbf{Z} \in \mathcal{C}^{\mathcal{N}^{\in} \times \mathcal{N}^{\in}}$. By using (20a) and (20b), and noting that $\mathcal{R}\{\mathbf{w_L^H Z w_L}\} = \mathcal{R}\{\mathbf{w_L^H Z^H w_L}\}$ and $\mathcal{I}\{\mathbf{w_L^H Z w_L}\} = -\mathcal{I}\{\mathbf{w_L^H Z^H w_L}\}$, r_A^{rob} and s_A^{rob} can be simplified to

$$\begin{aligned} r_A^{\mathrm{rob}} &\approx \mathrm{tr}(\mathbf{w_L w_L^H a^H a}) + 2\mathcal{R}\left\{\tilde{\mathbf{e}}_{\langle,\in}\left((\hat{\tilde{\mathbf{g}}}_{\mathbf{1}}^{\mathbf{H}} \hat{\tilde{\mathbf{g}}}_{\mathbf{1}})^{\mathbf{T}} \odot \mathbf{w_L w_L^H}\right)\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}^{\mathbf{H}}\right\} \\ &\quad + 2\mathcal{R}\left\{\tilde{\mathbf{e}}_{\},\infty}\left((\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}^{\mathbf{H}} \hat{\tilde{\mathbf{h}}}_{\mathbf{2}})^{\mathbf{T}} \odot \mathbf{w_L w_L^H}\right)\hat{\tilde{\mathbf{g}}}_{\mathbf{1}}^{\mathbf{H}}\right\}, \end{aligned} \tag{21a}$$

$$\begin{aligned} s_A^{\mathrm{rob}} &\approx \mathrm{tr}\Big(\mathbf{w_L w_L^H}(\mathbf{N^r} \odot (\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}^{\mathbf{H}} \hat{\tilde{\mathbf{h}}}_{\mathbf{2}}))\Big) \\ &\quad + 2\mathcal{R}\left\{\tilde{\mathbf{e}}_{\langle,\in}\left((\mathbf{N^r})^{\mathbf{T}} \odot \mathbf{w_L w_L^H}\right)\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}^{\mathbf{H}}\right\} + \sigma_{\mathbf{n_A}}^{\mathbf{2}}. \end{aligned} \tag{21b}$$

For the transmit power at the multiple-antenna relay, the channel uncertainty should also be included by incorporating $\mathbf{E_{UL}} =$

$[\mathbf{e_h}\ \mathbf{e_g}]$ into (11) as

$$\begin{aligned} P_R^{\text{rob}} &= \mathbf{w_L^H}\hat{\mathbf{C}}\mathbf{w_L} + \text{tr}\Big(\mathbf{W}\big(\mathbf{E_{UL}}\hat{\mathbf{H}}_{\mathbf{UL}}^{\mathbf{H}} + \hat{\mathbf{H}}_{\mathbf{UL}}\mathbf{E}_{\mathbf{UL}}^{\mathbf{H}}\big)\mathbf{W^H}\Big) \\ &\approx \mathbf{w_L^H}\hat{\mathbf{C}}\mathbf{w_L} + \mathbf{2}\mathcal{R}\Big\{\Big(\hat{\tilde{\mathbf{h}}}_{\mathbf{1}} \odot \text{vec}(\mathbf{W^H W})^{\mathbf{T}}\Big)\tilde{\mathbf{e}}_{\mathbf{h,2}}^{\mathbf{H}} \\ &\quad + \Big(\hat{\tilde{\mathbf{g}}}_{\mathbf{1}} \odot \text{vec}(\mathbf{W^H W})^{\mathbf{T}}\Big)\tilde{\mathbf{e}}_{\mathbf{g,2}}^{\mathbf{H}}\Big\} \end{aligned} \tag{22}$$

where $\hat{\mathbf{C}} = \left(\left(\hat{\mathbf{H}}_{\text{UL}}\hat{\mathbf{H}}_{\text{UL}}^{H} + \sigma_{n_R}^2\mathbf{I_N}\right)^T \otimes \mathbf{I_N}\right)$.

IV. PROPOSED ROBUST MIMO RELAY DESIGN

Typically, robust techniques employ stochastic or worst-case approaches depending upon different CSI error models [24]. The statistics of the CSI error are known under the stochastic philosophy, while the channel error is specified in some uncertainty region in the worst-case method. Since the CSI errors are assumed to be inside the sets of $\mathcal{R}_\langle$ and $\mathcal{R}_\}$ in (16), the worst-case approach is exploited in this section to design the robust multiple-antenna relay precoder by optimizing the worst-case performance.

A. Problem Formulation with Imperfect CSI

The objective of the proposed robust multiple-antenna relay design is to minimize the worst-case transmit power while fulfilling the worst-case SNR constraints at both sources. The so-called worst-case is based upon the largest possible errors $\mathbf{e_h}$ and $\mathbf{e_g}$ which are norm-bounded by δ_h and δ_g, respectively. Hence, the worst-case SNR for each source is the minimum SNR above the predetermined threshold to guarantee the QoS in the presence of the largest possible errors. Similarly, the worst-case transmit power is the maximum power that the multiple-antenna relay needs to spend in forwarding the linearly processed signal under the largest possible CSI estimation errors.

Let P_R^{wc} denote the worst-case transmit power and γ_A^{wc} the worst-case SNR at source A. So, P_R^{wc} can be mathematically expressed as

$$P_R^{\text{wc}} = \max_{\|\mathbf{e_h}\|\le\delta_h, \|\mathbf{e_g}\|\le\delta_g} P_R^{\text{rob}}, \tag{23}$$

where P_R^{rob} is given by (22), and γ_A^{wc} can be written as

$$\gamma_A^{\text{wc}} = \min_{\|\mathbf{e_h}\|\le\delta_h, \|\mathbf{e_g}\|\le\delta_g} \gamma_A^{\text{rob}}. \tag{24}$$

Then, we modify both the objective of transmit power as well as the QoS conditions in (12) by replacing both with the worst-case transmit power of (23) and the worst-case received SNR of (24) to incorporate the robustness against unknown but norm-bounded channel estimation errors. The so-called robust design of multiple-antenna relay precoder under the worst-case philosophy becomes

$$\begin{aligned} \mathcal{P}^{rob}:\ &\min_{\mathbf{w_L}} \max_{\|\mathbf{e_h}\|\le\delta_h, \|\mathbf{e_g}\|\le\delta_g} P_R^{\text{rob}} \\ &\text{s.t.} \min_{\|\mathbf{e_h}\|\le\delta_h, \|\mathbf{e_g}\|\le\delta_g} \gamma_k^{\text{rob}} \ge \gamma_{\text{th}},\ k \in \{A, B\}. \end{aligned} \tag{25}$$

Note that the objective turns out to minimize the maximum multiple-antenna relay transmit power with respect to the largest possible CSI errors.

Unfortunately, the numerator and denominator of γ_A^{rob}, i.e., (21a) and (21b) are not independent, which complicates further solving (25) to obtain any tractable design. So, we strengthen the worst-case QoS constraints by replacing the worst-case received SNR of (24) with its lower bound $\tilde{\gamma}_A^{\text{wc}}$ as [25]

$$\gamma_A^{\text{wc}} \ge \frac{\min_{\|\mathbf{e_h}\|\le\delta_h, \|\mathbf{e_g}\|\le\delta_g} r_A^{\text{rob}}}{\max_{\|\mathbf{e_h}\|\le\delta_h} s_A^{\text{rob}}} \triangleq \tilde{\gamma}_A^{\text{wc}}. \tag{26}$$

To this end, the proposed optimization problem can be rewritten as

$$\begin{aligned} \mathcal{P}:\ &\min_{\mathbf{w_L}} P_R^{\text{wc}} \\ &\text{s.t. } \tilde{\gamma}_k^{\text{wc}} \ge \gamma_{\text{th}},\ k \in \{A, B\} \end{aligned} \tag{27}$$

where the worst-case SNR is substituted by its lower bound $\tilde{\gamma}_k^{\text{wc}}$.

Moreover, considering the numerator of $\tilde{\gamma}_A^{\text{wc}}$ in (26), it's a minimization problem with the estimation errors bounded by their Euclidean norms. Referring to (21a), since $\tilde{\mathbf{e}}_{h,2}$ and $\tilde{\mathbf{e}}_{g,1}$ are defined from (6) and equal to $\mathbf{e_h^T D_2}$ and $\mathbf{e_g^T D_1}$, respectively, we apply Cauchy-Schwarz inequality to the last two terms of r_A^{rob} and determine the minimum as

$$\begin{aligned} \min_{\|\mathbf{e_h}\|\le\delta_h, \|\mathbf{e_g}\|\le\delta_g} r_A^{\text{rob}} &= \text{tr}\left(\mathbf{w_L w_L^H a^H a}\right) - \\ &\quad 2\delta_h \left\|\mathbf{D_2}\left((\hat{\tilde{\mathbf{g}}}_{\mathbf{1}}^{\mathbf{H}}\hat{\tilde{\mathbf{g}}}_{\mathbf{1}})^{\mathbf{T}} \odot \mathbf{w_L w_L^H}\right)\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}^{\mathbf{H}}\right\| - \\ &\quad 2\delta_g \left\|\mathbf{D_1}\left((\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}^{\mathbf{H}}\hat{\tilde{\mathbf{h}}}_{\mathbf{2}})^{\mathbf{T}} \odot \mathbf{w_L w_L^H}\right)\hat{\tilde{\mathbf{g}}}_{\mathbf{1}}^{\mathbf{H}}\right\|. \end{aligned} \tag{28}$$

Similarly, the maximization problem of the denominator in (26) can also be solved by replacing $\tilde{\mathbf{e}}_{h,2}$ with $\mathbf{e_h^T D_2}$ and then applying Cauchy-Schwarz inequality to the second term. However, the solution is positive for maximization whereas negative for minimization. So, we get

$$\begin{aligned} \max_{\|\mathbf{e_h}\|\le\delta_h} s_A^{\text{rob}} &= \text{tr}\left(\mathbf{w_L w_L^H}\left(\mathbf{N^r} \odot \left(\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}^{\mathbf{H}}\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}\right)\right)\right) \\ &\quad + 2\delta_h \left\|\mathbf{D_2}\left((\mathbf{N^r})^{\mathbf{T}} \odot \mathbf{w_L w_L^H}\right)\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}^{\mathbf{H}}\right\| + \sigma_{n_A}^2. \end{aligned} \tag{29}$$

Applying (28) and (29) to (26), the robust SNR constraint at source A in (27) can be given in tractable form. For the problem of maximizing the robust transmit power at the multiple-antenna relay P_R^{rob} as given in (23), the worst-case transmit power can be derived as

$$\begin{aligned} \max_{\|\mathbf{e_h}\|\le\delta_h, \|\mathbf{e_g}\|\le\delta_g} P_R^{\text{rob}} &= 2\delta_h \left\|\mathbf{D_2}\left(\hat{\tilde{\mathbf{h}}}_{\mathbf{1}} \odot \text{vec}\left(\mathbf{W^H W}\right)^{\mathbf{T}}\right)\right\| \\ &\quad + 2\delta_g \left\|\mathbf{D_2}\left(\hat{\tilde{\mathbf{g}}}_{\mathbf{1}} \odot \text{vec}\left(\mathbf{W^H W}\right)^{\mathbf{T}}\right)\right\| \\ &\quad + \mathbf{w_L^H}\hat{\mathbf{C}}\mathbf{w_L} \\ &\triangleq 2\delta_h \left\|\mathbf{H_A}\text{vec}\left(\mathbf{w_L w_L^H}\right)\right\| \\ &\quad + 2\delta_g \left\|\mathbf{H_B}\text{vec}\left(\mathbf{w_L w_L^H}\right)\right\| \\ &\quad + \mathbf{w_L^H}\hat{\mathbf{C}}\mathbf{w_L} \end{aligned} \tag{30}$$

$$
\begin{aligned}
\mathcal{P}: \min_{\mathbf{w_L}} \ & \mathbf{w_L^H \hat{C} w_L} + \mathbf{2}\delta_{\mathbf{h}} \left\|\mathbf{H_A} \text{vec}\left(\mathbf{w_L w_L^H}\right)\right\| + \mathbf{2}\delta_{\mathbf{g}} \left\|\mathbf{H_B} \text{vec}\left(\mathbf{w_L w_L^H}\right)\right\| \\
\text{s.t. tr}\left(\mathbf{w_L w_L^H}\left(\mathbf{a^H a} - \gamma_{\mathbf{th}}\left(\mathbf{N^r} \odot \left(\hat{\mathbf{h}}_\mathbf{2}^\mathbf{H}\hat{\mathbf{h}}_\mathbf{2}\right)\right)\right)\right) & \geq 2\delta_h \left\|\mathbf{D_2}\left(\left(\hat{\mathbf{g}}_\mathbf{1}^\mathbf{H}\hat{\mathbf{g}}_\mathbf{1}\right)^\mathbf{T} \odot \mathbf{w_L w_L^H}\right)\hat{\mathbf{h}}_\mathbf{2}^\mathbf{H}\right\| \\
& + 2\delta_g \left\|\mathbf{D_1}\left(\left(\hat{\mathbf{h}}_\mathbf{2}^\mathbf{H}\hat{\mathbf{h}}_\mathbf{2}\right)^\mathbf{T} \odot \mathbf{w_L w_L^H}\right)\hat{\mathbf{g}}_\mathbf{1}^\mathbf{H}\right\| \\
& + 2\gamma_{\text{th}}\delta_h \left\|\mathbf{D_2}\left(\left(\mathbf{N^r}\right)^\mathbf{T} \odot \mathbf{w_L w_L^H}\right)\hat{\mathbf{h}}_\mathbf{2}^\mathbf{H}\right\| \\
& + \gamma_{\text{th}}\sigma_{n_A}^2, \\
\text{tr}\left(\mathbf{w_L w_L^H}\left(\mathbf{b^H b} - \gamma_{\mathbf{th}}\left(\mathbf{N^r} \odot \left(\hat{\mathbf{g}}_\mathbf{2}^\mathbf{H}\hat{\mathbf{g}}_\mathbf{2}\right)\right)\right)\right) & \geq 2\delta_g \left\|\mathbf{D_2}\left(\left(\hat{\mathbf{h}}_\mathbf{1}^\mathbf{H}\hat{\mathbf{h}}_\mathbf{1}\right)^\mathbf{T} \odot \mathbf{w_L w_L^H}\right)\hat{\mathbf{g}}_\mathbf{2}^\mathbf{H}\right\| \\
& + 2\delta_h \left\|\mathbf{D_1}\left(\left(\hat{\mathbf{g}}_\mathbf{2}^\mathbf{H}\hat{\mathbf{g}}_\mathbf{2}\right)^\mathbf{T} \odot \mathbf{w_L w_L^H}\right)\hat{\mathbf{h}}_\mathbf{1}^\mathbf{H}\right\| \\
& + 2\gamma_{\text{th}}\delta_g \left\|\mathbf{D_2}\left(\left(\mathbf{N^r}\right)^\mathbf{T} \odot \mathbf{w_L w_L^H}\right)\hat{\mathbf{g}}_\mathbf{2}^\mathbf{H}\right\| \\
& + \gamma_{\text{th}}\sigma_{n_B}^2 \\
\text{where } \mathbf{a} = \hat{\tilde{\mathbf{g}}}_\mathbf{1} \odot \hat{\tilde{\mathbf{h}}}_\mathbf{2} & \text{ and } \mathbf{b} = \hat{\tilde{\mathbf{h}}}_\mathbf{1} \odot \hat{\tilde{\mathbf{g}}}_\mathbf{2}.
\end{aligned} \tag{31}
$$

Table 1. Pesudocode for constructing $\mathbf{H_A}$ and $\mathbf{H_B}$.

Step 1:	Initialization: $l = 0, \tilde{\mathbf{H}}_A = \tilde{\mathbf{H}}_B = \text{zeros}(N^2, N^4)$.
Step 2:	For $j = 1 : N$
Step 3:	For $i = 1 : N$
Step 4:	$l = l + 1$;
	$p = [(i-1)N^3 + (j-1)N + 1 : N^2$
	$+1 : iN^3 + jN]$;
	$\tilde{\mathbf{H}}_A(l,p) = \hat{\mathbf{h}}_1(l)$, $\tilde{\mathbf{H}}_B(l,p) = \hat{\mathbf{g}}_1(l)$
Step 5:	End Step 3.
Step 6:	End Step 2.
Step 7:	$\mathbf{H_A} = \mathbf{D_2}\tilde{\mathbf{H}}_\mathbf{A}$, $\mathbf{H_B} = \mathbf{D_2}\tilde{\mathbf{H}}_\mathbf{B}$ where
	$\mathbf{D_2} = \mathbf{1_{1\times N}} \otimes \mathbf{I_N}$ as provided in (6).

where $\mathbf{H_A}, \mathbf{H_B} \in \mathcal{C}^{\mathcal{N}\times\mathcal{N}^\triangle}$ can be constructed by using the pseudocode presented in Table 1.

The same approach can be applied to the transmission from A to B to derive the robust SNR constraint at source B. Finally, the proposed optimization problem in (27) can be specified as

B. Semidefinite Relaxation based Approximation

By observing the optimization problem of (31), both the objective and constraints consist of the second-order terms of $\mathbf{w_L}$ inside the norms and accordingly are nonconvex. However, after some tricky manipulation step by step, this problem can be solved, and an approximate solution to $\mathcal{P}$ can be obtained. We refer to this approach as the two-stage SDR-based approximation [22].

Define $\tilde{\mathbf{W}} = \mathbf{w_L w_L^H} \in \mathcal{C}^{\mathcal{N}\mathcal{N}\times\mathcal{N}\mathcal{N}}$, then $\tilde{\mathbf{W}} \succeq 0$ and $\text{rank}(\tilde{\mathbf{W}}) = 1$. It can be easily shown that $\mathbf{w_L^H X w_L} = \text{tr}(\mathbf{w_L^H X w_L}) = \text{tr}(\mathbf{X w_L w_L^H}) = \text{tr}(\mathbf{X}\tilde{\mathbf{W}})$. Now, we change the optimization variable in (31) from $\mathbf{w_L}$ to $\tilde{\mathbf{W}}$, and rewrite the optimization problem in terms of $\tilde{\mathbf{W}}$ as

$$
\begin{aligned}
\mathcal{P}_2: \min_{\tilde{\mathbf{W}}} \ & \text{tr}(\hat{\mathbf{C}}\tilde{\mathbf{W}}) + 2\delta_h \left\|\mathbf{H_A}\text{vec}(\tilde{\mathbf{W}})\right\| + 2\delta_g \left\|\mathbf{H_B}\text{vec}(\tilde{\mathbf{W}})\right\| \\
\text{s.t. } & \text{tr}(\mathbf{A}\tilde{\mathbf{W}}) \geq \mathbf{2}\delta_\mathbf{h} \left\|\mathbf{D_2}\left((\hat{\mathbf{g}}_\mathbf{1}^\mathbf{H}\hat{\mathbf{g}}_\mathbf{1})^\mathbf{T} \odot \tilde{\mathbf{W}}\right)\hat{\mathbf{h}}_\mathbf{2}^\mathbf{H}\right\| \\
& \quad + 2\gamma_{\text{th}}\delta_h \left\|\mathbf{D_2}\left((\mathbf{N^r})^\mathbf{T} \odot \tilde{\mathbf{W}}\right)\hat{\mathbf{h}}_\mathbf{2}^\mathbf{H}\right\| \\
& \quad + \gamma_{\text{th}}\sigma_{n_A}^2, \\
& \text{tr}(\mathbf{B}\tilde{\mathbf{W}}) \geq \mathbf{2}\delta_\mathbf{g} \left\|\mathbf{D_2}\left((\hat{\mathbf{h}}_\mathbf{1}^\mathbf{H}\hat{\mathbf{h}}_\mathbf{1})^\mathbf{T} \odot \tilde{\mathbf{W}}\right)\hat{\mathbf{g}}_\mathbf{2}^\mathbf{H}\right\| \\
& \quad + 2\delta_h \left\|\mathbf{D_1}\left((\hat{\mathbf{g}}_\mathbf{2}^\mathbf{H}\hat{\mathbf{g}}_\mathbf{2})^\mathbf{T} \odot \tilde{\mathbf{W}}\right)\hat{\mathbf{h}}_\mathbf{1}^\mathbf{H}\right\| \\
& \quad + 2\gamma_{th}\delta_g \left\|\mathbf{D_2}\left((\mathbf{N^r})^\mathbf{T} \odot \tilde{\mathbf{W}}\right)\hat{\mathbf{g}}_\mathbf{2}^\mathbf{H}\right\| \\
& \quad + \gamma_{th}\sigma_{n_B}^2, \\
& \tilde{\mathbf{W}} \succeq 0, \ \text{rank}(\tilde{\mathbf{W}}) = 1
\end{aligned} \tag{32}
$$

where $\mathbf{A} = \mathbf{a^H a} - \gamma_{\text{th}}\left(\mathbf{N^r} \odot \left(\hat{\mathbf{h}}_\mathbf{2}^\mathbf{H}\hat{\mathbf{h}}_\mathbf{2}\right)\right)$ and $\mathbf{B} = \mathbf{b^H b} - \gamma_{\mathbf{th}}\left(\mathbf{N^r} \odot \left(\hat{\mathbf{g}}_\mathbf{2}^\mathbf{H}\hat{\mathbf{g}}_\mathbf{2}\right)\right)$.

The problem $\mathcal{P}_2$ is still nonconvex due to the last $\tilde{\mathbf{W}}$ rank-one constraint in (32). After removing this nonconvex constraint, we obtain a relaxation counterpart of (32) denoted by $\mathcal{P}_3$ as shown below.

$$
\begin{aligned}
\mathcal{P}_3: \min_{\tilde{\mathbf{W}}} \ & \text{tr}(\hat{\mathbf{C}}\tilde{\mathbf{W}}) + 2\delta_h \left\|\mathbf{H_A}\text{vec}(\tilde{\mathbf{W}})\right\| + 2\delta_g \left\|\mathbf{H_B}\text{vec}(\tilde{\mathbf{W}})\right\| \\
\text{s.t. } & \text{tr}(\mathbf{A}\tilde{\mathbf{W}}) \geq \mathbf{2}\delta_\mathbf{h} \left\|\mathbf{D_2}\left((\hat{\mathbf{g}}_\mathbf{1}^\mathbf{H}\hat{\mathbf{g}}_\mathbf{1})^\mathbf{T} \odot \tilde{\mathbf{W}}\right)\hat{\mathbf{h}}_\mathbf{2}^\mathbf{H}\right\| \\
& \quad + 2\delta_g \left\|\mathbf{D_1}\left((\hat{\mathbf{h}}_\mathbf{2}^\mathbf{H}\hat{\mathbf{h}}_\mathbf{2})^\mathbf{T} \odot \tilde{\mathbf{W}}\right)\hat{\mathbf{g}}_\mathbf{1}^\mathbf{H}\right\| \\
& \quad + 2\gamma_{\text{th}}\delta_h \left\|\mathbf{D_2}\left((\mathbf{N^r})^\mathbf{T} \odot \tilde{\mathbf{W}}\right)\hat{\mathbf{h}}_\mathbf{2}^\mathbf{H}\right\| \\
& \quad + \gamma_{\text{th}}\sigma_{n_A}^2,
\end{aligned}
$$

$$\begin{aligned}\mathrm{tr}(\mathbf{B}\tilde{\mathbf{W}}) \geq\ & \mathbf{2}\delta_{\mathbf{g}}\left\|\mathbf{D_2}\left((\hat{\mathbf{h}}_{\mathbf{1}}^{\mathbf{H}}\hat{\mathbf{h}}_{\mathbf{1}})^{\mathbf{T}}\odot\tilde{\mathbf{W}}\right)\hat{\mathbf{g}}_{\mathbf{2}}^{\mathbf{H}}\right\| \\ & + 2\delta_h\left\|\mathbf{D_1}\left((\hat{\mathbf{g}}_{\mathbf{2}}^{\mathbf{H}}\hat{\mathbf{g}}_{\mathbf{2}})^{\mathbf{T}}\odot\tilde{\mathbf{W}}\right)\hat{\mathbf{h}}_{\mathbf{1}}^{\mathbf{H}}\right\| \\ & + 2\gamma_{\mathrm{th}}\delta_g\left\|\mathbf{D_2}\left((\mathbf{N}^{\mathbf{r}})^{\mathbf{T}}\odot\tilde{\mathbf{W}}\right)\hat{\mathbf{g}}_{\mathbf{2}}^{\mathbf{H}}\right\| \\ & + \gamma_{\mathrm{th}}\sigma_{n_B}^2, \\ \tilde{\mathbf{W}} \succeq\ & 0\end{aligned} \tag{33}$$

which is a convex optimization problem. Note that by introducing a new variable, $\mathcal{P}_3$ can be easily reformulated as a standard semidefinite programming (SDP) problem consisting of linear objective and second-order cone and positive semidefinite constraints. The convex problem of (33) can be efficiently solved with Matlab by using the YALMIP toolbox [26].

The optimization problem $\mathcal{P}_3$ subsumes the problem $\mathcal{P}_2$, since the feasible region of $\mathcal{P}_2$ is a subset of that of $\mathcal{P}_3$. In general, the optimal solution to $\mathcal{P}_3$ is of rank r with $r > 1$ rather than rank-one, which makes $\mathcal{P}_2$ infeasible. The Gaussian randomization technique [12], [22], [27] can be adopted to construct a feasible rank-one solution to $\mathcal{P}_2$ from the optimal solution of relaxed SDP problem $\mathcal{P}_3$. So the two-stage SDR-based approximation method used in this paper consists of solving the relaxation problem $\mathcal{P}_3$ in the first stage and then applying the randomization technique to the optimal solution of $\mathcal{P}_3$ in the second stage. This method will lead to an approximate solution to $\mathcal{P}_2$. In the following, the Gaussian randomization method is presented.

Let $\tilde{\mathbf{W}}^{\mathrm{opt}}$ denote the optimal solution to $\mathcal{P}_3$ in (33). The idea behind the Gaussian randomization is to generate a large number of candidate vectors representing the multiple-antenna relay precoder matrix from $\tilde{\mathbf{W}}^{\mathrm{opt}}$ and choose the one that can be scaled to guarantee the SNR constraints of $\mathcal{P}_2$ in (32) at the minimum transmit power cost. Initially, the eigenvalue decomposition of the optimal matrix $\tilde{\mathbf{W}}^{\mathrm{opt}}$ is calculated as $\tilde{\mathbf{W}}^{\mathrm{opt}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^{\mathbf{H}}$. Let $\mathbf{v_L}$ be a column vector of N^2 zero-mean, unit-variance complex circularly symmetric uncorrelated Gaussian r.v.'s, i.e., $\mathbf{v_L} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I_{N^2}})$. Then, the candidate vector $\mathbf{w}_{\mathbf{L}}^{\mathrm{c}}$ is constructed as $\mathbf{w}_{\mathbf{L}}^{\mathrm{c}} = \mathbf{U}\boldsymbol{\Sigma}^{\mathbf{1/2}}\mathbf{v_L}$, and the multiple-antenna relay precoder matrix can herein be obtained through the reverse vectorization operation as $\mathbf{W}^{\mathrm{c}} = \mathrm{vec}^{-\mathbf{1}}(\mathbf{w}_{\mathbf{L}}^{\mathrm{c}})$. This ensures that $\mathbb{E}[\mathbf{w}_{\mathbf{L}}^{\mathrm{c}}(\mathbf{w}_{\mathbf{L}}^{\mathrm{c}})^{\mathbf{H}}] = \tilde{\mathbf{W}}^{\mathrm{opt}}$, denoted by $\mathbf{w}_{\mathbf{L}}^{\mathrm{c}} \sim \mathcal{CN}(\mathbf{0}, \tilde{\mathbf{W}}^{\mathrm{opt}})$.

Since $\mathbf{w}_{\mathbf{L}}^{\mathrm{c}}$ depends on the particular realization of $\mathbf{v_L}$, the constraints of the original problem $\mathcal{P}$ in (31) might be violated or over-satisfied. Accordingly we can seek a positive boost or reduction factor to scale $\mathbf{w}_{\mathbf{L}}^{\mathrm{c}}$ to the minimum length that is necessary to satisfy the constraints. Denote the scale factor as $\sqrt{\lambda}$, and based upon the problem $\mathcal{P}_2$ in (32) define

$$\begin{aligned}\alpha &= \mathrm{tr}\left(\hat{\mathbf{C}}\tilde{\mathbf{W}}^{\mathrm{c}}\right) + 2\delta_h\left\|\mathbf{H_A}\mathrm{vec}(\tilde{\mathbf{W}}^{\mathrm{c}})\right\| + \mathbf{2}\delta_{\mathbf{g}}\left\|\mathbf{H_B}\mathrm{vec}(\tilde{\mathbf{W}}^{\mathrm{c}})\right\|, \\ \beta_A &= \mathrm{tr}\left(\mathbf{A}\tilde{\mathbf{W}}^{\mathrm{c}}\right) - 2\delta_h\left\|\mathbf{D_2}\big((\hat{\mathbf{g}}_{\mathbf{1}}^{\mathbf{H}}\hat{\mathbf{g}}_{\mathbf{1}})^{\mathbf{T}}\odot\tilde{\mathbf{W}}^{\mathrm{c}}\big)\hat{\mathbf{h}}_{\mathbf{2}}^{\mathbf{H}}\right\| \\ &\quad - 2\delta_g\left\|\mathbf{D_1}\big((\hat{\mathbf{h}}_{\mathbf{2}}^{\mathbf{H}}\hat{\mathbf{h}}_{\mathbf{2}})^{\mathbf{T}}\odot\tilde{\mathbf{W}}^{\mathrm{c}}\big)\hat{\mathbf{g}}_{\mathbf{1}}^{\mathbf{H}}\right\| \\ &\quad - 2\gamma_{\mathrm{th}}\delta_h\left\|\mathbf{D_2}\big((\mathbf{N}^{\mathbf{r}})^{\mathbf{T}}\odot\tilde{\mathbf{W}}^{\mathrm{c}}\big)\hat{\tilde{\mathbf{h}}}_{\mathbf{2}}^{\mathbf{H}}\right\|, \\ \beta_B &= \mathrm{tr}(\mathbf{B}\tilde{\mathbf{W}}^{\mathrm{c}}) - \mathbf{2}\delta_{\mathbf{g}}\left\|\mathbf{D_2}\big((\hat{\mathbf{h}}_{\mathbf{1}}^{\mathbf{H}}\hat{\mathbf{h}}_{\mathbf{1}})^{\mathbf{T}}\odot\tilde{\mathbf{W}}^{\mathrm{c}}\big)\hat{\mathbf{g}}_{\mathbf{2}}^{\mathbf{H}}\right\| \\ &\quad - 2\delta_h\left\|\mathbf{D_1}\big((\hat{\mathbf{g}}_{\mathbf{2}}^{\mathbf{H}}\hat{\mathbf{g}}_{\mathbf{2}})^{\mathbf{T}}\odot\tilde{\mathbf{W}}^{\mathrm{c}}\big)\hat{\mathbf{h}}_{\mathbf{1}}^{\mathbf{H}}\right\| \\ &\quad - 2\gamma_{\mathrm{th}}\delta_g\left\|\mathbf{D_2}\big((\mathbf{N}^{\mathbf{r}})^{\mathbf{T}}\odot\tilde{\mathbf{W}}^{\mathrm{c}}\big)\hat{\mathbf{g}}_{\mathbf{2}}^{\mathbf{H}}\right\|\end{aligned} \tag{34}$$

where $\tilde{\mathbf{W}}^{\mathrm{c}} = \mathbf{w}_{\mathbf{L}}^{\mathrm{c}}(\mathbf{w}_{\mathbf{L}}^{\mathrm{c}})^{\mathbf{H}}$, and hence $\tilde{\mathbf{W}}^{\mathrm{c}} \succeq 0, \mathrm{rank}(\tilde{\mathbf{W}}) = 1$. Then, it turns out that the following problem can be resorted to converting the candidate of multiple-antenna relay precoder to the candidate solution to $\mathcal{P}_2$.

$$\mathcal{Q}: \min_{\lambda\geq 0} \lambda\alpha \text{ s.t. } \lambda\beta_k \geq \gamma_{\mathrm{th}}\sigma_{n_k}^2, k \in \{A, B\}. \tag{35}$$

Problem $\mathcal{Q}$ is a linear programming (LP) with a single variable λ and linear inequality constraints. For a feasible instance of the LP problem $\mathcal{Q}$, it is obvious that β_k should be positive. So, for those $\mathbf{w}_{\mathbf{L}}^{\mathrm{c}}$ candidates that make β_k positive, the scaling factor can be easily solved as

$$\lambda = \max\left(\frac{\gamma_{\mathrm{th}}\sigma_{n_A}^2}{\beta_A}, \frac{\gamma_{\mathrm{th}}\sigma_{n_B}^2}{\beta_B}\right). \tag{36}$$

Therefore, the two-stage SDR-based approximation algorithm for generating an approximate solution to the original problem $\mathcal{P}$ can be summarized as

- **Relaxation**: Solve the relaxed equivalent SDP problem $\mathcal{P}_3$ and obtain the optimal solution $\tilde{\mathbf{W}}^{\mathrm{opt}}$.
- **Randomization**: Check the rank of $\tilde{\mathbf{W}}^{\mathrm{opt}}$.

– If $\tilde{\mathbf{W}}^{\mathrm{opt}} = 1$, then use its principal eigenvector as the optimal solution to problem $\mathcal{P}$.

– Otherwise, generate a candidate $\mathbf{w}_{\mathbf{L}}^{\mathrm{c}}$ by using the aforementioned Gaussian randomization method. Calculate $\beta_k, k \in \{A, B\}$ in (34), and if negative discard the corresponding candidate $\mathbf{w}_{\mathbf{L}}^{\mathrm{c}}$. Otherwise, determine and record the scaling factor λ from (36), as well as the associated objective value in (35) and the candidate vector. Repeat a large enough number of the randomization procedures. In the end, choose the candidate and scaling factor with the minimum objective value, denoted by $\mathbf{w}_{\mathbf{L}}^{\circ}$ and λ°, respectively.

The approximated optimal solution to problem $\mathcal{P}$ can be given as $\mathbf{w}_{\mathbf{L}}^{\mathrm{opt}} = \sqrt{\lambda^{\circ}}\mathbf{w}_{\mathbf{L}}^{\circ}$. The robust multiple-antenna relay precoder matrix can herein be determined through the reverse vectorization operation as $\mathbf{W}^{\mathrm{opt}} = \mathrm{vec}^{-\mathbf{1}}(\mathbf{w}_{\mathbf{L}}^{\mathrm{opt}})$.

V. NUMERICAL RESULTS

This section will first provide a numerical example to show the robustness capabilities of the proposed design and compare its performance with the non-robust approach in (12). In this example, we take $N = 3$, and consider the standard i.i.d. Rayleigh fading model – the elements of each channel vector between the source and relay are independent and identically distributed (i.i.d.) CGRV with zero-mean and unit variance. For all simulations, the channel estimates are given as $\hat{\mathbf{h}} = [0.6282 - 0.8111i \quad -2.0819 + 1.0171i \quad 0.9689 - 1.2102i]^T$ and $\hat{\mathbf{g}} = [-0.7558 - 0.5724i \quad 0.2299 - 0.5338i \quad -0.0723 - 0.1707i]^T$, and the noise power at the multiple-antenna relay and sources

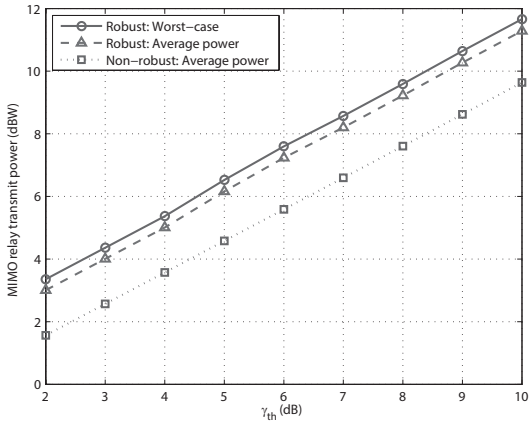

Fig. 2. Transmit power at the MIMO relay versus received SNR threshold $\gamma_{\rm th}$ for the robust ($1-\eta = 0.12$) and non-robust methods. σ_e^2 is fixed at 0.002.

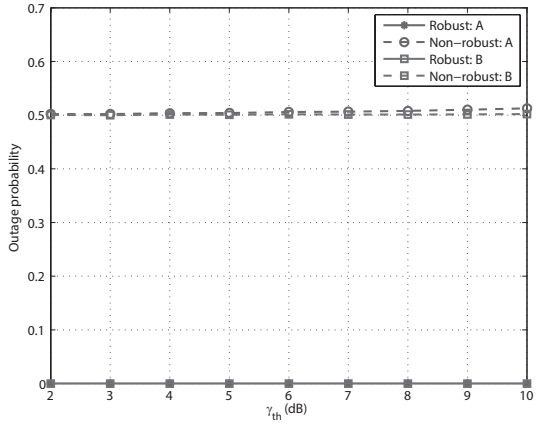

Fig. 3. Outage probability of SNR versus received SNR at both sources for the robust ($1-\eta = 0.12$) and non-robust methods. σ_e^2 is fixed at 0.002. The outage probability for the non-robust method is zero through all the SNR range.

are fixed at -20 and 0 dBW, respectively[1]. Furthermore, we assume that the error variances are identical across all links, i.e., $\sigma_{e_{h_i}}^2 = \sigma_{e_{g_i}}^2 = \sigma_e^2$ or $\delta_h = \delta_g = \delta$. Through Figs. 2 to 5, we take the error variance $\sigma_e^2 = 0.002$, and the upper norm-bound δ_h can be calculated by using a numerical search based on (17). The optimization problem $\mathcal{P}_3$ in (33) is solved using the YALMIP toolbox[2]. If the solution matrix $\tilde{\mathbf{W}}^{\rm opt}$ turns out being rank one, the associated principal component solves the original problem $\mathcal{P}$. Otherwise, the Gaussian randomization loop in Section IV-B is carried out. The computational complexity of this scheme is dominated by solving the relaxed equivalent SDP problem $\mathcal{P}_3$, which can be efficiently solved using the interior point method at a complexity cost that is at most $O((N^4+2)^{3.5})$ and usually much less [27]. With the obtained optimal $\mathbf{w_L}$, the worst-case transmit power at the multiple-antenna relay is calculated based on (30). Moreover, after substituting $\mathbf{H_{UL}} = \hat{\mathbf{H}}_{\mathbf{UL}} + \mathbf{E_{UL}}$ into (11) and then averaging over 1,000 independent realizations of estimation errors $\mathbf{e_h}$ and $\mathbf{e_g}$, the average transmit power of the robust method is obtained. For the comparison purpose, the non-robust method in (12) is implemented to obtain the optimal solution denoted by $\mathbf{w_L^{nrob}}$ without taking into account the channel estimation error. Similarly, the transmit power for the non-robust method with $\mathbf{w_L^{nrob}}$ in (11) is averaged by considering 1,000 random realizations of $\mathbf{e_h}$ and $\mathbf{e_g}$.

Fig. 2 illustrates the transmit power at the multiple-antenna relay versus SNR threshold $\gamma_{\rm th}$ for both robust and non-robust methods. Denote $\eta = \Pr\{\sigma_{e_h}^2 T \le \delta_h^2\}$, and $1-\eta$ is fixed at 0.12 in this figure which corresponds to c of 0.0042 and δ of 0.108. As shown in Fig. 2, the transmit power increases for all cases with increasing $\gamma_{\rm th}$. It can be intuitively reasoned that the multiple-antenna relay requires more power in order to guarantee higher QoS requirement at the sources. For the robust method, the worst-case transmit power is slightly larger than the average power, which manifests itself in the fact that (30) has two more terms comparing to (11). Moreover, the average transmit power curve of the non-robust method lies below that of the robust method with a power gap of around -2 dBW, since more power is needed to combat channel uncetainty. Based on the powers obtained in Fig. 2, we plot the outage probability for each source in Fig. 3 as a function of the SNR threshold. It should be noted that the outage probability here is not defined to indicate the fading of the channel itself by convention, but refers to the probability that the received SNR[3] at the sources is below the threshold for the given channel estimates due to the CSI errors[4]. It can be observed from Fig. 3 that both sources have almost the same outage probability resulting from the symmetric channel condition. Furthermore, the outage probability for the non-robust method is affected deeply by the channel estimation errors, whereas the proposed robust method efficiently combats the effect of estimation errors as the outage is zero through all the SNR range.

Next, Figs. 4 and 5 investigate the outage probability for the robust and non-robust methods under the same amount of average transmit power. Taking the average transmit power of the non-robust method in Fig. 2 as a baseline, the appropriate β or δ in the robust method is determined with the help of numerical search for each SNR threshold to achieve almost the same average transmit power at the relay as in the non-robust method. The resulting average transmit power is plotted in Fig. 4. As shown in Fig. 5, the outage probability of each source for the robust method is still zero, fully exhibiting the robustness of the proposed multiple-antenna relay design against the influence of estimation errors. So, we can conclude that the robust method far outperforms the non-robust method in terms of outage probability with the same or similar average transmit powers at the

[1]In practice, the relay is implemented as a fixed station, and employs expensive hardware to effectively reduce the receiving noise. On the other hand, the sources are usually end users like cellphone or portable wireless device, and limited to the battery power and chip cost. Hence, we assume much higher noise power at the sources in comparison to the relay.

[2]In all of our simulations, YALMIP takes seconds to solve $\mathcal{P}_3$ if a feasible solution exists.

[3]The actual received SNR in the simulations is calculated based on (8) by using the solved optimum relay precoder form either non-robust (12) or robust design (25) and the real channels from (13).

[4]The outage probability is averaged over 1,000 independent trials, and in each trial 1,000 outage events are captured by randomly generating the estimation errors.

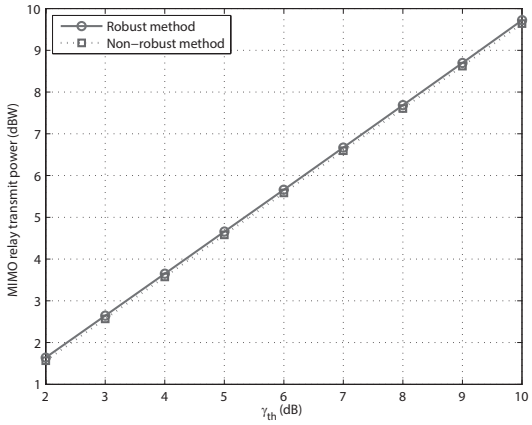

Fig. 4. Transmit power at the MIMO relay versus received SNR threshold $\gamma_{\rm th}$ for the robust ($1-\eta$ is adjusted for each $\gamma_{\rm th}$) and non-robust methods. σ_e^2 is fixed at 0.002.

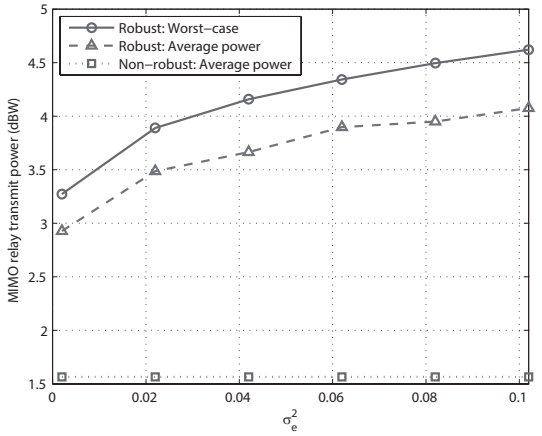

Fig. 6. Transmit power at the multiple-antenna relay versus error variance σ_e^2 for the robust and non-robust methods. $\gamma_{\rm th}$ is fixed at 2 dB.

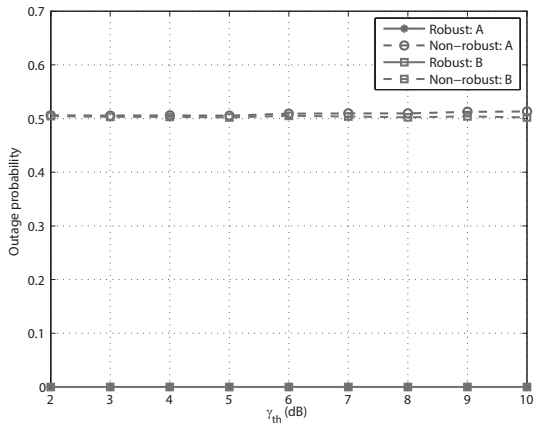

Fig. 5. Outage probability of SNR at both sources for the robust ($1-\eta$) is adjusted for each $\gamma_{\rm th}$) and non-robust methods. σ_e^2 is fixed at 0.002. The outage probability for the non-robust method is zero through all the SNR range.

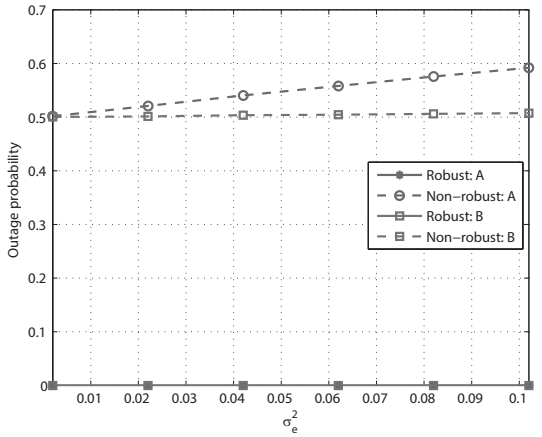

Fig. 7. Outage probability of SNR versus error variance σ_e^2 at both sources for the robust and non-robust methods. $\gamma_{\rm th}$ is fixed at 2 dB. The outage probability for the non-robust method is zero through all the error variance range.

multiple-antenna relay.

In Figs. 6 and 7, we plot the transmit power at the multiple-antenna relay and the outage probability of SNR for both sources versus the error variance σ_e^2, respectively. We take $c = 0.0042$ and $\gamma_{\rm th} = 2$ dB for both figures. Since the non-robust method is based upon the perfect channels, the CSI errors don't impact the design, and the average transmit power at the multiple-antenna relay remains almost the same across different error variances as shown in Fig. 6. But for the robust method, the increasing value of σ_e^2 implies worse channel condition, and more transmit power is needed to provide the robustness. As in previous cases, the outage probability of SNR versus σ_e^2 in the given range remains zero for the robust method whereas more than 50% is observed for the non-robust method. Hence, the proposed robust method can serve the QoS requirement for both sources but at the cost of a larger transmit power at the multiple-antenna relay.

Besides the performance improvement, the feasibility issue is also of interest. Since problem $\mathcal{P}_3$ is a relaxation of problem $\mathcal{P}_2$, $\mathcal{P}_3$ is feasible as long as $\mathcal{P}_2$ is. But the converse is generally not true. In order to establish the feasibility of the proposed two-stage SDR-based approximation, the first is that $\mathcal{P}_3$ is feasible, and once a feasible solution to $\mathcal{P}_3$ is found the randomization loop should yield at least one feasible solution, i.e., $\beta_k, k \in \{A, B\}$ in (34) should be positive. It has been verified in the simulations that if the randomization loop can yield at least one feasible solution, each randomization procedure of the loop can produce a feasible solution. After a large number of randomization procedures have generated a predetermined number[5] of feasible solutions, the best feasible solution can be selected as the approximated optimum solution to the original problem $\mathcal{P}$. Otherwise, the randomization loop will fail to return any feasible solution. Note that the randomization loop is another factor contributing to the complexity of the SDR-based approximation algorithm. As long as the randomization loop can yield at least one feasible solution, the randomization loop will run a prede-

[5] In our simulations, we choose the number to be 1,000.

Table 2. Simulation results with various transmit antennas $N = 3$ and $N = 6$ versus received SNR threshold based on 1,000 different independent channel realizations. The noise variance at the MIMO relay N_r is fixed at 0.1.

γ_{th}	$N = 3$				$N = 6$			
	Feas. $\mathcal{P}_3$		Feas. appr.		Feas. $\mathcal{P}_3$		Feas. appr.	
	#	%	#	%	#	%	#	%
2	997	99.7	981	98.4	1000	100.0	1000	100.0
4	988	98.8	958	97.0	1000	100.0	1000	100.0
6	964	96.4	914	94.8	1000	100.0	999	99.9
8	911	91.1	799	87.7	1000	100.0	999	99.9
10	784	78.4	590	75.3	998	99.8	983	98.5
12	532	53.2	324	60.9	980	98.0	900	91.8
14	203	20.3	96	47.3	863	86.3	651	75.4

Table 3. Simulation results with various noise variance $N_r = 0.1$ and $N_r = 0.01$ versus received SNR threshold based on 1,000 different independent channel realizations. The number of transmit antennas at MIMO relay N is fixed at 3.

γ_{th}	$N_r = 1$				$N_r = 0.01$			
	Feas. $\mathcal{P}_3$		Feas. appr.		Feas. $\mathcal{P}_3$		Feas. appr.	
	#	%	#	%	#	%	#	%
2	533	53.3	329	61.7	1000	100.0	999	99.9
4	204	20.4	96	47.1	999	99.9	998	99.9
6	38	3.8	14	36.8	999	99.9	997	99.8
8	3	0.3	1	33.3	999	99.9	996	99.7
10	-	-	-	-	998	99.8	991	99.3
12	-	-	-	-	997	99.7	980	98.3
14	-	-	-	-	987	98.7	957	97.0

termined number of times and each run will generate a feasible solution. So the complexity of the randomization loop is $O(1)$ under the condition that the feasible solution exists[6]. Hence, the feasibility issue is of more significance than the algorithm complexity.

Feasibility depends on various factors: the number of transmit antennas, the channel characteristics, the noise variance, and finally the received SNR threshold, which is investigated in the simulations by exploiting 1,000 different channel realizations. For each simulation run, the feasibility of $\mathcal{P}_3$ is first verified. If $\mathcal{P}_3$ is feasible, the randomization loop will be carried out to see if any feasible $\mathbf{w_L}$ to the original problem $\mathcal{P}$ can be yielded. Table 2 compares different number of transmit antennas, i.e., $N = 3$ and $N = 6$ with fixed noise variance of 0.1 at the multiple-antenna relay, while Table 3 fixes the number of transmit antennas at $N = 3$ and compares the different noise variance of 1 and 0.01. For each of these configurations, the same received SNR targets are requested for both sources ranging from 2 dB to 14 dB to investigate the impact of received SNR threshold. The other simulation parameters are in accordance with the ones used previously. The number of 1,000 simulations and corresponding percentage for which $\mathbf{P_3}$ is feasible are listed in the columns "Feas. $\mathcal{P}_3$", in which "#" denotes the number and "%" represents the percentage. Columns "Feas. appr." report the number of problem instances for which, once a feasible solution to $\mathcal{P}_3$ is found, the randomization loop yields at least one feasible solution, as well as the corresponding percentage. In all configurations considered, the higher the received SNR threshold, the less likely that $\mathcal{P}_3$ is feasible and that the randomization loop yields a feasible solution to $\mathcal{P}$. It can be observed in Table 2 that increasing the number of transmit antennas at the multiple-antenna relay increases the feasibility for both relaxation and randomization. This observation is expected since exploring more antennas can better combat the channel uncertainty by providing higher received SNR for both sources. Finally, $\mathcal{P}_3$ is getting more difficult to solve for higher noise variance at the multiple-antenna relay as illustrated in Table 3.

[6] In our simulations, it usually takes around one second to generate 1,000 candidates based on 2.4 GHz Intel Core 2 Duo processor. So plus the time solving the relaxed optimization problem $\mathcal{P}_3$, the total processing time to yield an optimal approximate solution to the original problem $\mathcal{P}$ is in the time order of seconds.

VI. CONCLUSIONS

In this paper, we have addressed the robust multiple-antenna relay design problem in TWRN and provided the robust multiple-antenna relay design based on the channel estimates. The channel estimation error has been explicitly taken into account in the design, and the worst-case philosophy has been adopted to include the robustness. Based upon the design criteria of QoS, we seek to minimize the worst-case transmit power at multiple-antenna relay while guaranteeing the worst-case received SNR above a prescribed threshold at both sources. The formulation turns out nonconvex, but has been transformed into a convex optimization problem that can be efficiently solved by means of SDR and randomization technique. The simulation shows that transmit power expectedly increases with higher QoS requirement. The extra power is used to combat channel uncertainty. Nevertheless, the proposed system triumphs at limiting power outage probability, as it has led to almost zero outage probability through an extensive SNR range at similar range of transmit power. From the numerical results, it can also be concluded that feasibility decreases with the increasing SNR threshold or noise variance levels. However, the problem can be alleviated by increasing the number of transmit antennas.

REFERENCES

[1] C. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Stat. Prob.*, vol. 1, pp. 611–644, Univ. California Press, 1961.

[2] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 379–389, Feb. 2007.

[3] F. Gao, C. Schnurr, I. Bjelakovic, and H. Boche, "Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, pp. 3024–3033, Oct. 2009.

[4] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," *SIGCOMM Comput. Commun. Rev.*, vol. 37, pp. 397–408, Aug. 2007.

[5] S. J. Kim, P. Mitran, and V. Tarokh, "Performance bounds for bidirectional coded cooperation protocols," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5235–5241, Nov. 2008.

[6] S. Zhou and G. Giannakis, "How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels?," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1285–1294, July 2004.

[7] J. Choi, "Performance analysis for transmit antenna diversity with/without channel information," *IEEE Trans. Veh. Technol.*, vol. 51, pp. 101–113, Jan. 2002.

[8] A. Pascual-Iserte, A. Perez-Neira, and M. Lagunas, "On power allocation strategies for maximum signal to noise and interference ratio in an OFDM-MIMO system," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 808–820, May 2004.
[9] J. Zou, H. Luo, M. Tao, and R. Wang, "Joint source and relay optimization for non-regenerative MIMO two-way relay systems with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 3305–3315, Sept. 2012.
[10] C. Wang, S. Ma, and Y.-C. 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, pp. 2273–2283, Apr. 2010.
[11] J. Wang and D. Palomar, "Robust MMSE precoding in MIMO channels with pre-fixed receivers," *IEEE Trans. Signal Process.*, vol. 58, pp. 5802–5818, Nov. 2010.
[12] B. Chalise and L. Vandendorpe, "MIMO relay design for multipoint-to-multipoint communications with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 57, pp. 2785–2796, July 2009.
[13] B. Chalise, "Optimization of MIMO relays for multipoint-to-multipoint communications: Nonrobust and robust designs," *IEEE Trans. Signal Process.*, vol. 58, pp. 6355–6368, Dec. 2010.
[14] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in MIMO-relay networks with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, pp. 5473–5484, Nov. 2011.
[15] A. Pascual-Iserte, D. Palomar, A. Perez-Neira, and M. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 54, pp. 346–360, Dec. 2006.
[16] G. Zheng, K.-K. Wong, and T.-S. Ng, "Robust linear MIMO in the downlink: A worst-case optimization with ellipsoidal uncertainty regions," *EURASIP J. Adv. in Signal Process.*, pp. 1–16, 2008.
[17] A. Panah and R. Heath, "MIMO two-way amplify-and-forward relaying with imperfect receiver CSI," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 4377–4387, Nov. 2010.
[18] E. Gharavol and E. Larsson, "Robust joint optimization of MIMO two-way relay channels with imperfect CSI," in *Annual Allerton Conf. on Commun., Control, and Comput.*, pp. 1657–1664, Sept. 2011.
[19] M. Botros and T. Davidson, "Convex conic formulations of robust downlink precoder designs with quality of service constraints," *IEEE J. Sel. Topics Signal Process.*, vol. 1, pp. 714–724, Dec. 2007.
[20] M. Payaro, A. Pascual-Iserte, and M. Lagunas, "Robust power allocation designs for multiuser and multi-antenna downlink communication systems through convex optimization," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 1390–1401, Sept. 2007.
[21] R. Zhang, Y.-C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, pp. 699–712, June 2009.
[22] E. Karipidis, N. Sidiropoulos, and Z.-Q. Luo, "Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups," *IEEE Trans. Signal Process.*, vol. 56, pp. 1268–1279, Mar. 2008.
[23] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203–2214, 2006.
[24] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Mar. 2004.
[25] M. Biguesh, S. Shahbazpanahi, and A. B. Gershman, "Robust downlink power control in wireless cellular systems," *EURASIP J. Wireless Commun. Netw.*, pp. 261–272, Dec. 2004.
[26] J. Lofberg, "Yalmip: A toolbox for modeling and optimization in matlab," in *Proc. CACSD Conf.*, Taipei, Taiwan, Sept. 2004.
[27] N. D. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, pp. 2239–2251, June 2006.

Chenyuan Wang received her B.A.Sc. degree from Beijing University of Posts and Telecommunications, China in 2009, and the M.A.Sc. degree from the University of Victoria, BC, Canada, in 2012. Now she is working full time in QUALCOMM Inc., CA, USA. Her research interests are in cooperative diversity, bidirectional communications, and MIMO wireless systems.

Xiaodai Dong (S'97-M'00-SM'09) received her B.Sc. degree in Information and Control Engineering from Xi'an Jiaotong University, China in 1992, her M.Sc. degree in Electrical Engineering from National University of Singapore in 1995 and her Ph.D. degree in Electrical and Computer Engineering from Queen's University, Kingston, ON, Canada in 2000. Since January 2005 she has been with the University of Victoria, Victoria, Canada, where she is now a Professor and Canada Research Chair (Tier II) at the Department of Electrical and Computer Engineering. Between 2002 and 2004, she was an Assistant Professor at the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. From 1999 to 2002, she was with Nortel Networks, Ottawa, ON, Canada and worked on the base transceiver design of the third-generation (3G) mobile communication systems.

Dr. Dong is an Editor for *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Vehicular Technology* and *Journal of Communications and Networks*. Her research interests include communication theory, radio propagation, ultra-wideband radio, machine to machine communications, wireless security, smart grid, and signal processing for communication applications.

Yi Shi (S'06–M'10) received the B.E. degree in Telecommunications from Beijing University of Posts and Telecommunications, Beijing, China, in 2006. He received the Ph.D. Degree in Electronic and Computer Engineering from The Hong Kong University of Science and Technology, Hong Kong, in August 2010. From 2010 to 2012, he worked as a Post-Doctoral Fellow in the Department of Electrical and Computer Engineering, University of Victoria. Dr. Shi is currently an Engineer in Huawei Technologies, Beijing, China. His work involves the 3GPP standardization of the next-generation wireless mobile communication systems. His research interests include cooperative communications, resource allocation and scheduling using optimization and game theory.