sensors

MDPI

Article

Robust Transceiver Design for IRS-Assisted Cascaded MIMO Communication Systems

Hossein Esmaeili [1,*], Ali Kariminezhad [2] and Aydin Sezgin [1]

[1] Institute of Digital Communication Systems, Ruhr University Bochum, 44801 Bochum, Germany
[2] e:fs TechHub GmbH, 85080 Gaimersheim, Germany
* Correspondence: hossein.esmaeili@rub.com

Abstract: Intelligent reconfigurable surfaces (IRSs) have gained much attention due to their passive behavior that can be a successor to relays in many applications. However, traditional relay systems might still be a perfect choice when reliability and throughput are the main concerns in a communication system. In this work, we use an IRS along with a decode-and-forward relay to provide a possible solution to address one of the main challenges of future wireless networks which is providing reliability. We investigate a robust transceiver design against the residual self-interference (RSI), which maximizes the throughput rate under self-interference channel uncertainty-bound constraints. The yielded problem turns out to be a non-convex optimization problem, where the non-convex objective is optimized over the cone of semidefinite matrices. We propose a novel mathematical method to find a lower bound on the performance of the IRS that can be used as a benchmark. Eventually, we show an important result in which, for the worst-case scenario, IRS can be helpful only if the number of IRS elements are at least as large as the size of the interference channel. Moreover, a novel method based on majorization theory and singular value decomposition (SVD) is proposed to find the best response of the transmitters and relay against worst-case RSI. Furthermore, we propose a multi-level water-filling algorithm to obtain a locally optimal solution iteratively. We show that our algorithm performs better that the state of the art in terms of time complexity as well as robustness. For instance, our numerical results show that the acheivable rate can be increased twofold and almost sixfold, respectively, for the case of small and large antenna array at transceivers.

Keywords: IRS; robust design; worst case; decode-and-forward; MIMO

check for updates

Citation: Esmaeili, H.; Kariminezhad, A.; Sezgin, A. Robust Transceiver Design for IRS-Assisted Cascaded MIMO Communication Systems. *Sensors* **2022**, *22*, 6587. https://doi.org/10.3390/s22176587

Academic Editors: Karim Seddik, Radwa Sultan and Hongliang Zhang

Received: 21 July 2022
Accepted: 19 August 2022
Published: 31 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Reliability and throughput are two of the most crucial requirements for the next generation of wireless networks. Optimally relaying the signal from a source to a destination can help enhance reliability and capacity of networks and is currently an active research area [1].

Another emerging candidate for relaying signals is reconfigurable intelligent surfaces (IRSs) [2]. An IRS is a device equipped with multiple passive reconfigurable reflectors that can reflect the colliding waves with an adjustable phase. One of the biggest advantages of IRSs is that they work in a real-time manner without consuming a noticeable amount of power [3]. However, the characteristics of the IRS (e.g., the lack of signal amplification and decode-and-forward processes) can potentially limit its functionality. As a result, in cases where reliability and throughput are of greater importance than power consumption, conventional relays might still be a better option than IRSs. For instance, authors in [4] show that a simple full-duplex relay can outperform an IRS in terms of throughput under certain conditions.

In this paper, we investigate the IRS-assisted MIMO full-duplex (FD) relay system that suffers channel uncertainties. It is also considered that the relays have practical issues such as self-interference (SI) as well as antenna and power limits. The combination of IRS and DF relay can potentially be advantageous. This is due to the fact that both the IRS and

the relay have their own limitations that can possibly be compensated for by exploiting each other. The objective of this paper is to maximize the achievable rate of the system by jointly optimizing the impact of the IRS as well as the covariance matrices of the source and the relay.

1.1. Related works

IRSs can be utilized in various ways to help the direct links enhance the performance of the system. In [5], authors utilized an IRS to maximize the weighted sum rate MISO system. Authors in [6] proposed a method to minimize the power consumption in a MISO system equipped with an IRS, and authors in [7] investigate the problem of energy efficiency in an IRS-assisted MISO downlink system. The problem of rate maximization in a MIMO system has been presented in [8], where the authors propose an iterative algorithm to find the best IRS pattern assuming the perfect channel state information (CSI) is given. Recently, the authors in [9] utilized IRS in a relay-aided network to minimize the power consumption and successfully showed that the combination of IRS and relay outperforms other scenarios. While the performance of IRS communication systems has been extensively studied, there is not much research that considers robust design when the perfect CSI is not available [10].

Recently, with the emergence of artificial intelligence, IRSs have shown a great potential to improve the existing protocols and technologies [11]. Authors in [12] investigate the benefit of employing an IRS equipped with a multi-task learning system on the transmit power and achievable throughput of aerial–terrestrial communications. In [13], authors use a reinforcement-learning-based approach to optimize the IRS reflection coefficients for buffer-aided relay selection.

One of the earliest studies of the robust transmission designs of IRS-assisted systems was undertaken in [14], where a bounded CSI error model is applied to a problem of power minimization in a MISO transmission system. There, by virtue of semidefinite programming (SDP), the authors turn the original problem into a sequence of convex sub-problems. The robust power minimization subject to the outage probability constraints under statistical cascaded channel error model is considered in [15], where the aim is to optimize the system under worst-case rate constraint. Authors in [16] have proposed a robust algorithm for mean squared error (MSE) minimization for a single user MISO system equipped with an IRS. Their method provides a closed form solution for each iteration. However, it can be used only for the case of a single user system and cannot be extended to more general cases where there are multiple users. Recently, a robust algorithm based on a penalty dual decomposition (PDD) technique is proposed in [10] for sum-rate maximization where they assumed that the channel estimation error follows a complex normal distribution.

Exploiting a relay to improve the communication throughput rate is a classic alternative for IRS in communication systems. However, utilizing a relay in a network raises some important questions to be answered. For instance, how should the relay process the received signal before dispatching it to the destination? Now, the relay can receive a signal from the source, process it and transmit it towards the destination in a successive manner. This type of relaying technique is known as half-duplex relaying. Alternatively, while receiving a signal at a certain time instant, the relay can simultaneously transmit the previously received signals. This technique is known as full-duplex relaying [17].

As a consequence of transmitting and receiving at a common resource unit, the relay is confronted with SI. Note that full-duplex relaying potentially increases the total throughput rate of the communication compared to the half-duplex counterpart only if the SI is handled properly at the relay input. By physically isolating the transmitter and receiver front ends of the relay, a significant portion of SI can be reduced [18]. Moreover, analog and/or digital signal processing at the relay input can be utilized to cancel a portion of SI [19–22]. This can be realized if the estimate of the SI can be obtained at the relay. These SI cancellation procedures can effectively mitigate the destructive impact of SI up to a certain level. Hence, the remaining portion, the so-called residual self-interference (RSI), is still present at the relay input. The distribution of the RSI is investigated in [23,24]. This RSI is mainly due

to the channel estimation uncertainties and transmitter noise. Therefore, the quality of channel estimation plays an important role for limiting RSI if the conventional modulation techniques are utilized.

The authors in [25] employ a superimposed signaling procedure (asymmetric modulation constellation) in the basic point-to-point FD communication for cancelling the SI and further retrieving the desired information contents without requiring channel estimates. They show that for the same average energy per transmission block, the bit error rate of their proposed method is better than that of conventional ones. The RSI evidently degrades the performance of the communication quality. To this end, the authors in [26] study the degrees-of-freedom (DoF), i.e., the slope of the rate curve at asymptotically high SNR and its relation to the performance of an FD cellular network in the presence of RSI. Moreover, the authors in [27,28] investigate the joint rate-energy and delivery time optimization of FD communication, respectively, when RSI is still present. Furthermore, the authors in [29] study the sum rate capacity of the FD channel with and without such degradation. In the presence of RSI, the authors in [30] study the capacity of a Gaussian two-hop FD relay.

Robust transceiver design against the worst-case RSI channel helps find the threshold for switching between HD and FD operating modes. This setup is commonly known as a hybrid relay system [31]. The authors in [32] investigate a robust design for multi-user full-duplex relaying with multi-antenna DF relay. In that work, the sources and destinations are equipped with single antennas. Moreover, the authors in [33] investigate a robust transceiver design for FD multi-user MIMO systems for maximizing the weighted sum-rate of the network. The robust design against worst-case RSI is investigated by authors in [34].

1.2. Contribution

Motivated by the above, in this work, we consider a DF multi-hop system with multiple antennas at the source, relay and destination along with an IRS to provide additional links. Then we try to maximize the throughput rate for the worst-case RSI scenario. To the best of our knowledge, this is the first time that the throughput rate maximization against the worst-case RSI is evaluated for IRS-assisted DF full-duplex relay in MIMO systems. First, we simplify the problem by finding an analytical lower bound for the performance of the IRS. Then, the optimization of maximum achievable rate of the DF full-duplex relaying is cast as a non-convex optimization problem. Thereafter, we propose a low complexity method to find the solution using majorization theory. We propose an efficient algorithm to solve this problem in polynomial time. Finally, the transmit signal covariances at the source and the relay are designed efficiently to improve robustness against worst-case RSI channel in a given uncertainty bound. Notice that once the covariances are known, one can easily find the precoders using conventional methods such as singular-value decomposition (SVD), etc. To the best of our knowledge, this is the first work that uses the IRS for RSI cancellation in MIMO full-duplex DF relay systems.

1.3. Organization

The rest of the paper is organized as follows. Section 2 outlines the system model and introduces its characteristics. The three different tasks for employing the IRS are also given in this section. In Section 3, the optimization problem belonging to the FD scenario is formulated, and its proper solution is presented. In addition, analytical bounds for the performance of the IRS are given and their corresponding proofs are provided. Section 4 provides the optimization problem for the HD scenario along with the solution. In Section 5, the effectiveness of the proposed algorithm is evaluated and verified by performing numerical simulations over various aspects. Finally, the paper is concluded in Section 6, and technical proofs of the theorems are given in the Appendices A–E.

2. System Model

We consider the communication from a source equipped with N_t antennas to a destination with N_r antennas. The reliable communication from the transmitter to the destination is assumed to be only feasible by means of a relay with K_t transmitter and K_r receiver

antennas at the output and input front ends, respectively. This means that the direct link from the transmitter to the destination and the link from the transmitter to the IRS and to the destination has a negligible impact on the throughput. This assumption is realistic for the scenarios where the path loss is high due to the high frequency ranges such as mmWave and Terahertz or due to far distances [35], as well as cases were there are objects that block the direct link between the source and the destination. An IRS consisting of M elements is established to either cancel the RSI or help enhance one of the transmitter–relay/relay–destination links. The overall system model can be found in Figure 1.

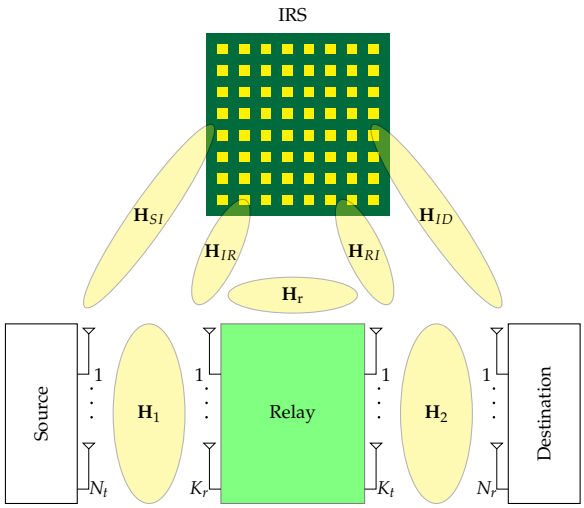

Figure 1. System model of an IRS assisted full-duplex relay. In our model both source and destination are equipped with N_t and N_r antennas, respectively. In addition, the relay is equipped with K_{t} transmitting and K_t receiving antennas, and the IRS has M passive elements.

In this paper, it is assumed that signal delivery over the transmitter–IRS–receiver link is not available. This is mainly due to the power attenuation and the power radiation pattern effects [36]. As the IRS is a passive device, it has some power attenuation in practice, which makes the reflected waves weaker than the received ones. In addition, due to the power radiation pattern, based on the angle of arrival and departure, both received and reflected waves are subject to attenuation, respectively. In our system, as the IRS is established in the vicinity of the relay and is faced towards it, both aforementioned effects cause the source–IRS–destination link to be extremely weaker than the source–relay–destination link.

Next, we present the achievable throughput rates for the HD and FD relaying. We start with the HD relay in which $\kappa = 0$. In the second case, IRS can be exploited to enhance the quality of the channel between the source and the relay. In such a case, the received signals at the relay and destination can be expressed as

$$\mathbf{y}_r = (\mathbf{H}_1 + \mathbf{H}_{IR}\boldsymbol{\Theta}\mathbf{H}_{SI})\mathbf{x}_s + \kappa\mathbf{H}_r\mathbf{x}_r + \mathbf{n}_t, \tag{1}$$

$$\mathbf{y}_{\mathrm{d}} = \mathbf{H}_2\mathbf{x}_r + \mathbf{n}_{\mathrm{d}}, \tag{2}$$

where $\mathbf{H}_{SI} \in \mathbb{C}^{N_{\mathrm{t}}M}$ is the channel from the source to the IRS. Finally, IRS can be used to help the channel from the relay to the destination. In this case, the received signals are going to be

$$\mathbf{y}_r = \mathbf{H}_1\mathbf{x}_s + \kappa\mathbf{H}_r\mathbf{x}_r + \mathbf{n}_t, \tag{3}$$

$$\mathbf{y}_{\mathrm{d}} = (\mathbf{H}_2 + \mathbf{H}_{ID}\boldsymbol{\Theta}\mathbf{H}_{RI})\mathbf{x}_r + \mathbf{n}_{\mathrm{d}}, \tag{4}$$

where $\mathbf{H}_{ID} \in \mathbb{C}^{M\times N_{\mathrm{r}}}$ is the channel from the source to the IRS.

In what follows, we find the achievable rate for three aforementioned cases and compare them to see under what conditions each of them should be applied. Notation and definitions are summarized in Table 1.

Table 1. Simulation parameters.

Notation	Definition	Notation	Definition
N_t	Number of antennas at the source	$\mathbf{\Theta}$	IRS phase profile
N_r	Number of antennas at the destination	x_s, x_r	Source and relay transmit signals
K_t	Number of antennas at the relay's transmitter	n_r, n_d	Additive noise at the relay and destination
K_r	Number of antennas at the relay's receiver	y_r, y_d	Received signals at the relay and destination, respectively
M	Number of IRS elements	κ	FD/HD mode indicator
$\mathbf{H}_r, \hat{\mathbf{H}}_r, \bar{\mathbf{H}}_r$	Self-interference actual channel, estimated channel and channel error, respectively	$R_{\text{sr}}^{\text{FD}}, R_{\text{rd}}^{\text{FD}}, R^{\text{FD}}$	Source–relay, relay–destination and the overall throughput in FD mode
$\mathbf{H}_1, \hat{\mathbf{H}}_1, \bar{\mathbf{H}}_1$	Source–Relay actual channel, estimated channel and channel error, respectively	$R_{\text{sr}}^{\text{HD}}, R_{\text{rd}}^{\text{HD}}, R^{\text{HD}}$	Source–relay, relay–destination and the overall throughput in HD mode
$\mathbf{H}_2, \hat{\mathbf{H}}_2, \bar{\mathbf{H}}_2$	Relay–Destination actual channel, estimated channel and channel error, respectively	$\mathbf{Q}_s, \mathbf{Q}_r$	Source and relay covarriance matrices
$\mathbf{H}_{SI}, \hat{\mathbf{H}}_{SI}, \bar{\mathbf{H}}_{SI}$	Source–IRS actual channel, estimated channel and channel error, respectively	$\mathbf{H}_{tot_1}$	$\mathbf{H}_{tot_1} = (\bar{\mathbf{H}}_r + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IR})$
$\mathbf{H}_{IR}, \hat{\mathbf{H}}_{IR}, \bar{\mathbf{H}}_{IR}$	IRS–Destination actual channel, estimated channel and channel error, respectively	$\hat{\mathbf{H}}_{tot_2}, \bar{\mathbf{H}}_{tot_2}$	$\hat{\mathbf{H}}_{tot_2} = (\hat{\mathbf{H}}_1 + \mathbf{H}_{RI}\mathbf{\Theta}\hat{\mathbf{H}}_{IR})$ $\bar{\mathbf{H}}_{tot_2} = (\bar{\mathbf{H}}_1 + \bar{\mathbf{H}}_{RI}\mathbf{\Theta}\bar{\mathbf{H}}_{IR})$
$\mathbf{H}_{RI}, \hat{\mathbf{H}}_{RI}, \bar{\mathbf{H}}_{RI}$	Relay–IRS actual channel, estimated channel and channel error, respectively	$\hat{\mathbf{H}}_{tot_3}, \bar{\mathbf{H}}_{tot_3}$	$\hat{\mathbf{H}}_{tot_3} = (\hat{\mathbf{H}}_2 + \mathbf{H}_{RI}\mathbf{\Theta}\hat{\mathbf{H}}_{IR})$ $\bar{\mathbf{H}}_{tot_3} = (\bar{\mathbf{H}}_2 + \bar{\mathbf{H}}_{RI}\mathbf{\Theta}\bar{\mathbf{H}}_{IR})$
$\mathbf{H}_{IR}, \hat{\mathbf{H}}_{IR}, \bar{\mathbf{H}}_{IR}$	IRS–Relay actual channel, estimated channel and channel error, respectively	$\lambda_i(X)$	i'th largest eigenvalue of matrix X
γ_{s_i}	$\gamma_{s_i} = \lambda_i(\mathbf{Q}_s)$	$\sigma_{1_i}^2$	$\sigma_{1_i}^2 = \lambda_i(\mathbf{H}_1\mathbf{H}_1^H)$
γ_{r_i}	$\gamma_{r_i} = \lambda_i(\mathbf{Q}_r)$	$\sigma_{r_i}^2$	$\sigma_{r_i}^2 = \lambda_i(\bar{\mathbf{H}}_{tot_1}\bar{\mathbf{H}}_{tot_1}^H)$
T_r	The RSI channel uncertainty bound	$\sigma_{2_i}^2$	$\sigma_{2_i}^2 = \lambda_i(\mathbf{H}_2\mathbf{H}_2^H)$
T_x	Channel estimation error bound $x \in \{1, 2, SI, IR, RI, ID\}$	T'	The remaining RSI channel uncertainty after considering the impact of the IRS

3. Achievable Rate (Full-Duplex Relay)

3.1. Overview

Suppose that the relay employs a DF strategy. In the full-duplex scenario, both source–relay and relay–destination links are active at the same time. As a result, the signals from the relay transmitter interfere with the receiving signal at the relay receiver. We assume that an estimate of the SI channel $\mathbf{H}_r$ is available at the relay denoted by $\hat{\mathbf{H}}_r$. Hence, the RSI represented by $\bar{\mathbf{H}}_r$ is given as

$$\bar{\mathbf{H}}_r = \mathbf{H}_r - \hat{\mathbf{H}}_r. \tag{5}$$

In the rest of the paper, we try to find an approach to deal with this RSI.

3.2. Mathematical Preliminaries

Considering a FD DF relay, the following rates are achievable [37],

$$R^{\mathrm{FD}} = \min(R_{\mathrm{sr}}^{\mathrm{FD}}, R_{\mathrm{rd}}^{\mathrm{FD}}), \tag{6}$$

in which, depending on how the IRS is applied to the system, the three following sets of rates are possible. First,

$$R_{\mathrm{sr}}^{\mathrm{FD}} = \log_2 \frac{\left|\sigma_t^2 \mathbf{I}_{K_r} + \hat{\mathbf{H}}_1 \mathbf{Q}_s \hat{\mathbf{H}}_1^H + \bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right|}{\left|\sigma_t^2 \mathbf{I}_{K_r} + \bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right|}, \tag{7}$$

$$R_{\mathrm{rd}}^{\mathrm{FD}} = \log_2 \frac{\left|\sigma_d^2 \mathbf{I}_N + \hat{\mathbf{H}}_2 \mathbf{Q}_r \hat{\mathbf{H}}_2^H + \bar{\mathbf{H}}_2 \mathbf{Q}_r \bar{\mathbf{H}}_2^H\right|}{\left|\sigma_d^2 \mathbf{I}_N + \bar{\mathbf{H}}_2 \mathbf{Q}_r \bar{\mathbf{H}}_2^H\right|}, \tag{8}$$

where $\mathbf{H}_{tot_1} = (\bar{\mathbf{H}}_r + \mathbf{H}_{RI} \boldsymbol{\Theta} \mathbf{H}_{IR})$ when the IRS is used to cancel the self interference. Second,

$$R_{\mathrm{sr}}^{\mathrm{FD}} = \log_2 \frac{\left|\sigma_t^2 \mathbf{I}_{K_r} + \hat{\mathbf{H}}_{tot_2} \mathbf{Q}_s \hat{\mathbf{H}}_{tot_2}^H + \bar{\mathbf{H}}_{tot_2} \mathbf{Q}_s \bar{\mathbf{H}}_{tot_2}^H + \bar{\mathbf{H}}_r \mathbf{Q}_r \bar{\mathbf{H}}_r^H\right|}{\left|\sigma_t^2 \mathbf{I}_{K_r} + \bar{\mathbf{H}}_{tot_2} \mathbf{Q}_s \bar{\mathbf{H}}_{tot_2}^H + \bar{\mathbf{H}}_r \mathbf{Q}_r \bar{\mathbf{H}}_r^H\right|}, \tag{9}$$

$$R_{\mathrm{rd}}^{\mathrm{FD}} = \log_2 \frac{\left|\sigma_d^2 \mathbf{I}_N + \hat{\mathbf{H}}_2 \mathbf{Q}_r \hat{\mathbf{H}}_2^H + \bar{\mathbf{H}}_2 \mathbf{Q}_r \bar{\mathbf{H}}_2^H\right|}{\left|\sigma_d^2 \mathbf{I}_N + \bar{\mathbf{H}}_2 \mathbf{Q}_r \bar{\mathbf{H}}_2^H\right|}, \tag{10}$$

where $\hat{\mathbf{H}}_{tot_2} = \left(\hat{\mathbf{H}}_1 + \hat{\mathbf{H}}_{RI} \boldsymbol{\Theta} \hat{\mathbf{H}}_{IR}\right)$ and $\bar{\mathbf{H}}_{tot_2} = (\bar{\mathbf{H}}_1 + \bar{\mathbf{H}}_{RI} \boldsymbol{\Theta} \bar{\mathbf{H}}_{IR})$ if the IRS is established to help the source–relay channel and finally

$$R_{\mathrm{sr}}^{\mathrm{FD}} = \log_2 \frac{\left|\sigma_t^2 \mathbf{I}_{K_r} + \hat{\mathbf{H}}_1 \mathbf{Q}_s \hat{\mathbf{H}}_1^H + \bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \bar{\mathbf{H}}_r \mathbf{Q}_r \bar{\mathbf{H}}_r^H\right|}{\left|\sigma_t^2 \mathbf{I}_{K_r} + \bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \bar{\mathbf{H}}_r \mathbf{Q}_r \bar{\mathbf{H}}_r^H\right|}, \tag{11}$$

$$R_{\mathrm{rd}}^{\mathrm{FD}} = \log_2 \frac{\left|\sigma_d^2 \mathbf{I}_N + \hat{\mathbf{H}}_{tot_3} \mathbf{Q}_r \hat{\mathbf{H}}_{tot_3}^H + \bar{\mathbf{H}}_{tot_3} \mathbf{Q}_r \bar{\mathbf{H}}_{tot_3}^H\right|}{\left|\sigma_d^2 \mathbf{I}_N + \bar{\mathbf{H}}_{tot_3} \mathbf{Q}_r \bar{\mathbf{H}}_{tot_3}^H\right|}, \tag{12}$$

where $\hat{\mathbf{H}}_{tot_3} = \left(\hat{\mathbf{H}}_2 + \hat{\mathbf{H}}_{RI} \boldsymbol{\Theta} \hat{\mathbf{H}}_{IR}\right)$ and $\bar{\mathbf{H}}_{tot_3} = (\bar{\mathbf{H}}_2 + \bar{\mathbf{H}}_{RI} \boldsymbol{\Theta} \bar{\mathbf{H}}_{IR})$ if the IRS is utilized to enhance the rate of the relay–destination channel. Notice that assuming that the RSI remains uncanceled, a robust transceiver against the worst-case RSI channel is required which is formulated as an optimization problem as follows

$$\max_{\mathbf{Q}_s, \mathbf{Q}_r, \boldsymbol{\Theta}} \min_{\bar{\mathbf{H}}_r} \quad \min\left(R_{\mathrm{sr}}^{\mathrm{FD}}, R_{\mathrm{rd}}^{\mathrm{FD}}\right) \tag{13}$$

$$\text{subject to} \quad \mathrm{Tr}(\mathbf{Q}_s) \leq P_s, \tag{13a}$$

$$\mathrm{Tr}(\mathbf{Q}_r) \leq P_r, \tag{13b}$$

$$\mathrm{Tr}(\bar{\mathbf{H}}_x \bar{\mathbf{H}}_x^H) \leq T_x,$$

$$x \in \{1, 2, r, RI, IR, ID, SI\} \tag{13c}$$

$$|\theta_m| \leq 1, \forall m \tag{13d}$$

in which the throughput rate with respect to the worst-case RSI channel is maximized. Two constraints, P_s and P_r, represent the transmit power budgets at the source and the relay, respectively. In constraint (13c), T_x represents the RSI or the channel estimation error bound corresponding to $\mathbf{H}_x$. Notice that $\mathrm{Tr}(\bar{\mathbf{H}}_x \bar{\mathbf{H}}_x^H)$ represents the sum of the squared singular values of $\mathbf{H}_x$. It should be noted that using a bounded matrix norm is the most common way for modeling the uncertainty of a matrix [38,39]. In practice, T_x can be found using stochastic methods when the distribution of the channel error is known. Otherwise, one may find it using a sample average approximation method. Finally, constraints (13d) are due to the unit modulus limitation of the IRS elements.

The problem (13) is non-convex and hard to solve. As a result, for each of the above-mentioned scenarios, we propose a simplified version of the optimization problem and try to solve it instead. Note that as we are interested in finding the throughput corresponding to the worst-case RSI, any simplification in the optimization problem should be in favor of the RSI and interference. First, we analyse the performance of the system when the IRS is helping the relay cancel the RSI. Consequently, we show that the problem (13) can be simplified to the following optimization problem

$$\begin{aligned} \max_{\mathbf{Q}_s,\mathbf{Q}_r} \min_{\mathbf{H}_{tot}} \quad & \min\left(R_{\text{sr}}^{\text{FD}}, R_{\text{rd}}^{\text{FD}}\right) && (14)\\ \text{subject to} \quad & \text{Tr}(\mathbf{Q}_s) \leq P_s, \\ & \text{Tr}(\mathbf{Q}_r) \leq P_r, \\ & \text{Tr}(\mathbf{H}_{tot}\mathbf{H}_{tot}^H) \leq T'(T_r, \mathbf{\Theta}). \\ & \text{Tr}(\bar{\mathbf{H}}_x\bar{\mathbf{H}}_x^H) \leq T_x, \quad x \in \{1,2\} && (14a) \end{aligned}$$

where

$$\begin{aligned} T'(T_r, \mathbf{\Theta}) = \min_{\mathbf{\Theta}} \max_{\bar{\mathbf{H}}_r} \quad & ||\bar{\mathbf{H}}_r + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IR}||_F^2 && (15)\\ \text{subject to} \quad & \text{Tr}(\bar{\mathbf{H}}_r\bar{\mathbf{H}}_r^H) \leq T_r, && (15a)\\ & ||Vec(\mathbf{\Theta})||_2^2 \leq 1, && (15b) \end{aligned}$$

and where $Vec(\cdot)$ denotes the vector of all non-zero elements of its input matrix. We can equivalently write T' as

$$\begin{aligned} T'(T_r, \mathbf{\Theta}) = \min_{\mathbf{\Theta}} \max_{\bar{\mathbf{H}}_r} \quad & ||Vec(\bar{\mathbf{H}}_r) + (\mathbf{H}_{IR} * \mathbf{H}_{RI}^T)Vec(\mathbf{\Theta})||_2^2 && (16)\\ \text{subject to} \quad & \text{Tr}(\bar{\mathbf{H}}_r\bar{\mathbf{H}}_r^H) \leq T_r, && (16a)\\ & ||Vec(\mathbf{\Theta})||_2^2 \leq 1, && (16b) \end{aligned}$$

where $*$ denotes a column-wise Khatri–Rao product defined as below

$$A * B = [A_1 \otimes B_1 | A_2 \otimes B_2 | \cdots | A_n \otimes B_n], \tag{17}$$

and where A_i is the i'th column of A, and $\otimes$ denotes the Kronecker product. See Appendix A for proof.

One can show that $T' \leq (\sqrt{T_r} - \sigma_{\min}(\mathbf{H}_{IR} * \mathbf{H}_{RI}^T))^2$. As mentioned before, problem (14) is a simplification of problem (13). This means every achievable rate which is inside the feasible set of (14) is also inside the feasible set of (13) (Notice that the reverse is not necessarily true, i.e., every achievable rate which is a feasible solution of (13) is not necessarily a feasible solution for (14) as well. However, as we look for achievable rates, we can still use this method). The reason is that in problem (13), the minimization over RSI happens only one time, and the worst-case RSI simultaneously tries to cancel the effect of the best configuration of IRS and the best covariance matrices. In (14), first, the RSI does its worst damage on the performance of the best IRS configuration and after that performs another optimization to bring the worst power allocation against the best covariance matrices (This will be clearer later on when the geometrical representation of the problem is given). In what follows, we provide our proposed ways to deal with optimization problems (14) and (15), respectively.

Theorem 1. *For the optimization problem* (14)*, one can show that* $T'(T, \Theta) \leq (\sqrt{T_r} - \sigma_{\min}(\mathbf{H}_{IR} * \mathbf{H}_{RI}^T))^2$.

Proof. We begin the proof with an intuitive example and then extend it to the more general case. Assume that $K_t = 1, K_r = 2$ and $M = 3$. Then we have

$$T' = \max_{\mathbf{\Theta}} \min_{\bar{\mathbf{h}}_r} \quad ||\bar{\mathbf{h}}_r + \mathbf{H}_{IR} diag(\mathbf{H}_{RI}^T) Vec(\mathbf{\Theta})||_2^2 \tag{18}$$

$$\text{subject to} \quad \bar{h}_{11}^2 + \bar{h}_{21}^2 \leq T_r, \tag{18a}$$

$$\theta_1^2 \leq 1,\ \theta_2^2 \leq 1,\ \theta_3^2 \leq 1. \tag{18b}$$

In addition, consider the following optimization problem

$$T'' = \max_{\mathbf{\Theta}} \min_{\bar{\mathbf{h}}_r} \quad ||\bar{\mathbf{h}}_r + \mathbf{H}_{IR} diag(\mathbf{H}_{RI}^T) Vec(\mathbf{\Theta})||_2^2 \tag{19}$$

$$\text{subject to} \quad \bar{h}_{11}^2 + \bar{h}_{21}^2 \leq T_r, \tag{19a}$$

$$\theta_1^2 + \theta_2^2 + \theta_3^2 \leq 1, \tag{19b}$$

Here, notice that $\mathbf{H}_{IR} diag(\mathbf{H}_{RI}^T)$ is a linear map from a three-dimensional into a two-dimensional space. One simple example of such a mapping can be found in Figure 2. Here, an example of mapping from a three-dimensional to a two-dimensional space is shown. The left shape shows the feasible set for the IRS with three elements in a real valued space. The cube belongs to the case of T', i.e., constraints $-1 \leq \theta_m \leq 1, \forall m$, while the sphere shows the constraint $\theta_1^2 + \theta_2^2 + \theta_3^2 \leq 1$ which belongs to T''. On the right, the feasible sets belonging to the two aforementioned regions after performing mapping f are presented as an example. It can be seen that the mapping of the first set of constraints (the hexagon) covers the whole area of that of the second one (the ellipse). One important key is, as mapping is a linear function, we have $\mathcal{A} \subset \mathcal{B} \rightarrow f(\mathcal{A}) \subset f(\mathcal{B})$, where $\mathcal{A}$ and $\mathcal{B}$ are two arbitrary sets and f is the mapping.

In general, as the number of IRS elements or the dimensions of $\bar{\mathbf{h}}_r$ increase, the mapping of the hypercube becomes more and more complicated and finding the optimal distance becomes more difficult. However, there is an upper bound for this distance. As shown in Figure 3, if instead of the cube, we limit the feasible set of IRS elements to the sphere inside the cube, i.e., replacing (18b) with (19b), the solution to the problem becomes $GE \geq GF$. It turns out that finding GE is very simple as by the definition we have $\sigma_{\min}(\mathbf{H}_{IR} diag(\mathbf{H}_{RI}^T)) = OE$, and also we know that $\sqrt{T_r} = GO$. Therefore, we can conclude $GE = \sqrt{T_r} - \sigma_{\min}(\mathbf{H}_{IR} diag(\mathbf{H}_{RI}^T))$. Finally, we use one last upper bound to make the original problem even easier to solve. Note that if instead of the ellipse, we consider the circle inscribed in it, we will have $\max_{\bar{\mathbf{h}}_r} \min_{\mathbf{\Theta}} ||\bar{\mathbf{h}}_r + (\mathbf{H}_{IR} * \mathbf{H}_{RI}^T)\theta||_2 = \sqrt{T_r} - \sigma_{\min}(\mathbf{H}_{IR} * \mathbf{H}_{RI}^T), \forall \bar{\mathbf{h}}_r$. As a result, we have

$$||\bar{\mathbf{h}}_{tot} \bar{\mathbf{h}}_{tot}^H||_2 \leq T', \quad \forall \bar{\mathbf{h}}_r, \tag{20}$$

where $T' = \left(\sqrt{T_r} - \sigma_{\min}(\mathbf{H}_{IR} * \mathbf{H}_{RI}^T)\right)^2$. It is worth mentioning that the geometrical representation for the optimization problem (13) is different because there, considering that the RSI wants to bring the worst representation against the IRS configuration and covariance matrices simultaneously, the RSI cannot freely span the whole circle. This is due to the fact that some regions in the circle might not be a good choice when it comes to RSI design against covariance matrices. However, if the best representation of RSI against the covariance matrices also provides the best RSI against the IRS configuration, the solution to (14) and (13) will be the same. Eventually, instead of optimization problem (13), one can solve optimization problem (14). The solution to the new problem is guaranteed to be achievable by the original problem as well.

Notice that one can readily extend this interpretation into the complex domain, as the constraint (19b) will still be a subset of constraints (18b). It remains to show one can generalize the geometrical proof for arbitrary large dimensions. This means that the channel dimensions and the number of IRS elements can be any natural numbers.

Interestingly, it is enough to show that the geometrical proof based on ℓ_2 norms and Euclidean distance exists for higher dimensions. This proof is given in Appendix A where it is shown $||\bar{\mathbf{H}}_r + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IR}||_F^2 = ||Vec(\bar{\mathbf{H}}_r) + (\mathbf{H}_{IR} * \mathbf{H}_{RI}^T)Vec(\mathbf{\Theta})||_2^2$. □

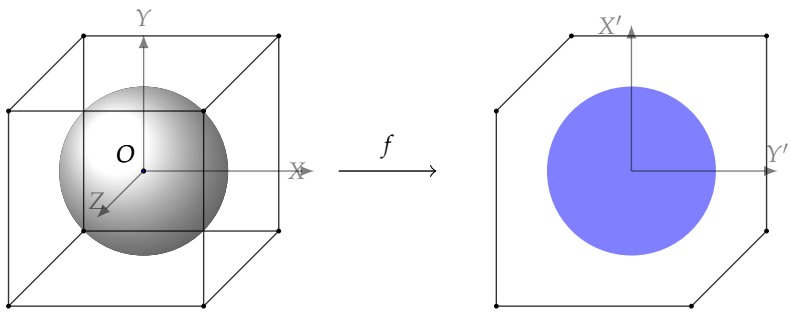

Figure 2. An example of mapping from three-dimensional to two-dimensional space. The hexagon and the circle are the mapping output of the cube and the sphere respectively.

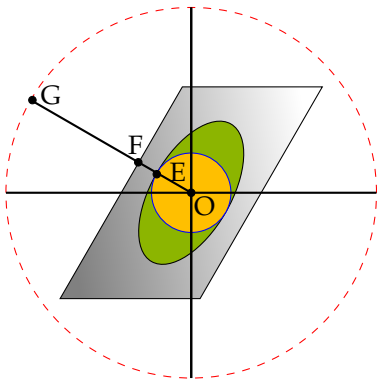

Figure 3. Geometrical representation of optimization problem (18).

Next we solve problem (14). Solving this problem is hard in general as it is non-convex. Hence, we first use the following lemma and theorem to solve it. There, it is shown that for every possible choice of $\mathbf{H}_1$ and $\mathbf{H}_2$, there exists at least one set of simultaneously diagonalizable matrices $\mathbf{H}_{tot}$, $\mathbf{Q}_s$ and $\mathbf{Q}_r$ that are the solutions to the problem (14).

Lemma 1. *For two positive semi-definite and positive definite matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *with eigenvalues* $\lambda_1(\mathbf{A}) \geq \lambda_2(\mathbf{A}) \geq ... \geq \lambda_N(\mathbf{A})$ *and* $\lambda_1(\mathbf{B}) \geq \lambda_2(\mathbf{B}) \geq ... \geq \lambda_N(\mathbf{B})$*, respectively, the following inequalities hold,*

$$\prod_{i=1}^{N}\left(1+\frac{\lambda_i(\mathbf{A})}{\lambda_i(\mathbf{B})}\right) \leq \left|\mathbf{I}+\mathbf{A}\mathbf{B}^{-1}\right| \leq \prod_{i=1}^{N}\left(1+\frac{\lambda_i(\mathbf{A})}{\lambda_{N+1-i}(\mathbf{B})}\right). \tag{21}$$

Proof. Consider Fiedler's inequality given by [40],

$$\prod_{i=1}^{N}(\lambda_i(\mathbf{A})+\lambda_i(\mathbf{B})) \leq \left|\mathbf{B}+\mathbf{A}\right| \leq \prod_{i=1}^{N}(\lambda_i(\mathbf{A})+\lambda_{N+1-i}(\mathbf{B})). \tag{22}$$

Furthermore, given $\mathbf{B}$ as a positive definite matrix, the following are true,

$$\left|\mathbf{B}\right| > 0, \tag{23}$$

$$\left|\mathbf{B}^{-1}\right| = \prod_{i=1}^{N}\frac{1}{\lambda_i(\mathbf{B})}. \tag{24}$$

Now, dividing the sides of (22) by $\left|\mathbf{B}\right|$, one can readily obtain (21). □

Note that in (21), the inequalities hold with equality if and only if $\mathbf{A}$ and $\mathbf{B}$ are diagonalizable over a common basis. Using the result of Lemma 1, $R_{\text{sr}}^{\text{FD}}$ can be lower-bounded as

$$\log_2 \frac{\left|\sigma_t^2 \mathbf{I}_{K_r} + \hat{\mathbf{H}}_1 \mathbf{Q}_s \hat{\mathbf{H}}_1^H + \bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right|}{\left|\sigma_t^2 \mathbf{I}_{K_r} + \bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right|} \geq \sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\lambda_i\left(\mathbf{H}_1 \mathbf{Q}_s {\mathbf{H}_1}^H\right)}{\lambda_i\left(\sigma_t^2 \mathbf{I}_{K_r} + \bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right)}\right). \tag{25}$$

In addition, it holds that $\lambda_i\left(\sigma_t^2 \mathbf{I}_{K_r} + \bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right) = \sigma_t^2 + \lambda_i\left(\bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right)$. Hence, we obtain

$$\log_2 \frac{\left|\sigma_t^2 \mathbf{I}_{K_r} + \hat{\mathbf{H}}_1 \mathbf{Q}_s \hat{\mathbf{H}}_1^H + \bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right|}{\left|\sigma_t^2 \mathbf{I}_{K_r} + \bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right|} \geq \sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\lambda_i\left(\mathbf{H}_1 \mathbf{Q}_s {\mathbf{H}_1}^H\right)}{\sigma_t^2 + \lambda_i\left(\bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right)}\right). \tag{26}$$

Note that the inequality holds with equality whenever $\mathbf{H}_1 \mathbf{Q}_s {\mathbf{H}_1}^H$ and $\bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H$ share a common basis. Next, we use the following inequality

$$\sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\lambda_i\left(\mathbf{H}_1 \mathbf{Q}_s {\mathbf{H}_1}^H\right)}{\sigma_t^2 + \lambda_i\left(\bar{\mathbf{H}}_1 \mathbf{Q}_s \bar{\mathbf{H}}_1^H + \mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right)}\right) \geq \sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\lambda_i\left(\mathbf{H}_1 \mathbf{Q}_s {\mathbf{H}_1}^H\right)}{\sigma_t^2 + T_1 P_s + \lambda_i\left(\mathbf{H}_{tot_1} \mathbf{Q}_r \mathbf{H}_{tot_1}^H\right)}\right). \tag{27}$$

The above inequality holds true since $\lambda_i(\bar{\mathbf{H}}_1 Q_s \bar{\mathbf{H}}_1^H) \leq T_1 P_s$.

Now, instead of completing the minimization over the left-hand side (LHS) of Equation (26), we can first minimize the right-hand side (RHS) of Equation (27) to find an achievable rate. Similarly, for $R_{\text{rd}}^{\text{FD}}$ we have

$$\log_2 \frac{\left|\sigma_d^2 \mathbf{I}_N + \hat{\mathbf{H}}_2 \mathbf{Q}_r \hat{\mathbf{H}}_2^H + \bar{\mathbf{H}}_2 \mathbf{Q}_r \bar{\mathbf{H}}_2^H\right|}{\left|\sigma_d^2 \mathbf{I}_N + \bar{\mathbf{H}}_2 \mathbf{Q}_r \bar{\mathbf{H}}_2^H\right|} \geq \sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\lambda_i\left(\mathbf{H}_2 \mathbf{Q}_r {\mathbf{H}_2}^H\right)}{\sigma_d^2 + T_2 P_r}\right). \tag{28}$$

Remark 1. *Having the equality $\mathbf{C} = \bar{\mathbf{H}}_r \mathbf{Q}_r \bar{\mathbf{H}}_r^H$, one can generally conclude the rule of multiplication is determinant, i.e., $\det(\mathbf{C}) = \det(\bar{\mathbf{H}}_r^H \bar{\mathbf{H}}_r) \det(\mathbf{Q}_r)$. Further, using the properties of determinants we can also conclude $\prod_{i=1}^N \lambda_i(\mathbf{C}) = \prod_{i=1}^N \left(\lambda_{\rho(i)}(\bar{\mathbf{H}}_r^H \bar{\mathbf{H}}_r) \lambda_i(\mathbf{Q}_r)\right)$ where $\rho(i)$ is a random permutation of i and indicates that there is no need for $\lambda_{\rho(i)}(\bar{\mathbf{H}}_r^H \bar{\mathbf{H}}_r)$ to be in decreasing order. However, one cannot generally conclude $\lambda_i(\mathbf{C}) = \lambda_{\rho(i)}(\bar{\mathbf{H}}_r^H \bar{\mathbf{H}}_r) \lambda_i(\mathbf{Q}_r), \forall i$, unless $\bar{\mathbf{H}}_r^H \bar{\mathbf{H}}_r$ and $\mathbf{Q}_r$ share common basis.*

As a result of Remark 1, in general, we cannot rewrite (26) in terms of $\lambda_i(\bar{\mathbf{H}}_r^H \bar{\mathbf{H}}_r)$, $\lambda_i(\mathbf{Q}_r)$, $\lambda_i(\mathbf{H}_1^H \mathbf{H}_1)$ and $\lambda_i(\mathbf{Q}_s)$. However, if we show that for every choice of $\mathbf{Q}_s$, there exists a matrix $\mathbf{Q}_s'$ with properties: (1) $\lambda_i(\mathbf{H}_1 \mathbf{Q}_s \mathbf{H}_1^H) = \lambda_i(\mathbf{H}_1 \mathbf{Q}_s' \mathbf{H}_1^H)$; (2) $\lambda_i(\mathbf{H}_1 \mathbf{Q}_s' \mathbf{H}_1^H) = \lambda_i(\mathbf{Q}_s') \lambda_i(\mathbf{H}_1^H \mathbf{H}_1)$ and (3) $\text{Tr}(\mathbf{Q}_s') \leq \text{Tr}(\mathbf{Q}_s)$; then we can use $\mathbf{Q}_s'$ instead and rewrite (26) in terms of $\lambda_i(\mathbf{H}_1^H \mathbf{H}_1)$ and $\lambda_i(\mathbf{Q}_s')$ to simplify the problem. The first property implies that both $\mathbf{Q}_s$ and $\mathbf{Q}_s'$ have the exact same impact on the capacity. Hence, if we find a $\mathbf{Q}_s$ which is the solution to the problem (14), its corresponding $\mathbf{Q}_s'$ will also be a solution. The second

property means, unlike $\mathbf{Q}_s$, $\mathbf{Q}'_s$ actually shares the common basis with $\mathbf{H}_1^H\mathbf{H}_1$. The last property implies that $\mathbf{Q}'_s$ is at least as good as $\mathbf{Q}_s$ in terms of power consumption. Observe that if we show for every feasible $\mathbf{Q}_s$ there exists at least one such $\mathbf{Q}'_s$, then we can solve the problem (14) in a much easier way. The reason is, in such a case, instead of searching for optimal $\mathbf{Q}_s$ over the whole feasible set, we can search for the optimal $\mathbf{Q}'_s$. Unlike $\mathbf{Q}_s$, finding $\mathbf{Q}'_s$ does not need a complete search over the whole feasible set since $\mathbf{Q}'_s$ shares a common basis with $\mathbf{H}_1^H\mathbf{H}_1$. Therefore, we can limit our search only to the portion of the feasible set in which the matrices have eigendirections identical to those of $\mathbf{H}_1^H\mathbf{H}_1$. Similarly, if we show for every choice of $\bar{\mathbf{H}}_r$, there exist at least one $\bar{\mathbf{H}}'_r$ for which we have three conditions $\lambda_i(\bar{\mathbf{H}}_r\mathbf{Q}_r\bar{\mathbf{H}}_r^H) = \lambda_i(\bar{\mathbf{H}}'_r\mathbf{Q}_r\bar{\mathbf{H}}_r'^H)$, $\lambda_i(\bar{\mathbf{H}}_r\mathbf{Q}_r\bar{\mathbf{H}}_r'^H) = \lambda_i(\mathbf{Q}_r)\lambda_i(\bar{\mathbf{H}}_1'^H\bar{\mathbf{H}}'_1)$ and $\mathrm{Tr}(\bar{\mathbf{H}}_1'^H\bar{\mathbf{H}}'_1) \leq \mathrm{Tr}(\bar{\mathbf{H}}_1^H\bar{\mathbf{H}}_1)$, we can simplify our search to finding $\bar{\mathbf{H}}'_r$ instead of $\bar{\mathbf{H}}_r$. In the next theorem, we show that such $\mathbf{Q}'_s$ and $\bar{\mathbf{H}}'_r$ exist.

Theorem 2. *For all matrices $\mathbf{Q}_s$ and $\mathbf{H}_1$, there exists at least one matrix $\mathbf{Q}'_s$ that satisfies the following conditions,*

$$\lambda_i(\mathbf{H}_1\mathbf{Q}_s\mathbf{H}_1^H) = \lambda_i(\mathbf{H}_1\mathbf{Q}'_s\mathbf{H}_1^H), \tag{29}$$

$$\lambda_i(\mathbf{H}_1\mathbf{Q}'_s\mathbf{H}_1^H) = \lambda_{\rho(i)}(\mathbf{Q}'_s)\lambda_i(\mathbf{H}_1^H\mathbf{H}_1), \tag{30}$$

$$Tr(\mathbf{Q}'_s) \leq Tr(\mathbf{Q}_s), \tag{31}$$

where $\rho(i)$ is a random permutation of i and indicates that there is no need for $\lambda_{\rho(i)}(\mathbf{Q}'_s)$ to be in decreasing order.

Proof. The proof is given in Appendix B. □

For the sake of simplicity, we use the following notions for the rest of the paper,

$$\gamma_{s_i} = \lambda_i(\mathbf{Q}_s), \tag{32}$$

$$\gamma_{r_i} = \lambda_i(\mathbf{Q}_r), \tag{33}$$

$$\sigma_{1_i}^2 = \lambda_i(\mathbf{H}_1\mathbf{H}_1^H), \tag{34}$$

$$\sigma_{r_i}^2 = \lambda_i(\bar{\mathbf{H}}_{tot_1}\bar{\mathbf{H}}_{tot_1}^H), \tag{35}$$

$$\sigma_{2_i}^2 = \lambda_i(\mathbf{H}_2\mathbf{H}_2^H). \tag{36}$$

Now, using Theorem 2 alongside Lemma 1, we infer that with no loss of generality, instead of optimising over matrices, one can complete the optimization over eigenvalues to find the optimal value for RSH of (27). Then we have

$$\max_{\gamma_s,\gamma_r} \min_{\sigma_r} \min\left(\sum_{i=1}^{\min(M,K_r)} \log\left(1+\frac{\sigma_{1_i}^2\gamma_{s_{\rho(i)}}}{\sigma_t^2+T_1P_s+\gamma_{r_i}\sigma_{r_{\rho(i)}}^2}\right), \sum_{i=1}^{\min(K_t,N)} \log\left(1+\frac{\sigma_{2_i}^2\gamma_{r_i}}{\sigma_d^2+T_2P_r}\right)\right) \tag{37}$$

$$\text{subject to} \quad \|\gamma_s\|_1 \leq P_s, \tag{37a}$$

$$\|\gamma_r\|_1 \leq P_r, \tag{37b}$$

$$\|\sigma_r^2\|_1 \leq T', \tag{37c}$$

$$\sigma_{1_i}^2\gamma_{s_{\rho(i)}} \geq \sigma_{1_{i+1}}^2\gamma_{s_{\rho(i+1)}}, \ \forall i \leq \min(M,K_r), \tag{37d}$$

$$\gamma_{r_i}\sigma_{r_{\rho(i)}}^2 \geq \gamma_{r_{i+1}}\sigma_{r_{\rho(i+1)}}^2, \ \forall i \leq \min(K_t,N). \tag{37e}$$

Note that the two additional constraints (37d) and (37e) need to be satisfied due to the conditions of Lemma 1 (i.e., eigenvalues have to be in decreasing order). Interestingly,

these two additional constraints are affine. The above optimization problem can further be simplified using the following lemma,

Lemma 2. *The objective function of the optimization problem* (37) *is optimized when the constraints* (37a) *and* (37c) *are satisfied with equality.*

Proof. Intuitively, as the objective function is an increasing and decreasing function of each element of γ_s and σ_r^2, respectively, at convergence, the constraints are met with equality. See Appendix C for the proof. □

3.3. Algorithm Description

In this subsection, our proposed algorithm is given. In short, it works as follows. First, based on the task of the IRS in the system, we compute the effect of IRS on the RSI, source–relay and/or relay–destination channel links. After that, we design the best signal design for the source and relays transmitters with the objective of maximizing the throughput. In the rest of this subsection, the detailed explanation of the algorithm is given. First, we need to solve the optimization problem (37). It can be readily shown that $R_{\text{rd}}^{\text{FD}}$ is a monotonically increasing function of P_r. Furthermore, one can show that $R_{\text{sr}}^{\text{FD}}$ is an increasing function with respect to P_s and a decreasing function with respect to T' and P_r (See Appendix D). Consequently, the worst-case RSI chooses a strategy to reduce the spectral efficiency, while the relay and the source cope with such strategy for improving the system robustness. That means, on one hand, the RSI hurts the stronger eigendirections of the received signal space more than the weaker ones. However, on the other hand, the source tries to cope with this strategy adaptively by smart eigen selection. This process clearly makes the optimization problem complicated at the source–relay hop. Unlike the source–relay hop, the resource allocation problem at the relay–receiver hop is rather easy. Since at the relay–receiver hop there is only one maximization, we can find the sum capacity simply by using the well-known water-filling algorithm.

Observe that although finding each $R_{\text{sr}}^{\text{FD}}$ and $R_{\text{rd}}^{\text{FD}}$ separately is a convex problem, the problem (37) as a whole is not convex. Therefore in this paper, we find the optimal $R_{\text{sr}}^{\text{FD}}$ by keeping $R_{\text{rd}}^{\text{FD}}$ fixed. Then we use the resulting $R_{\text{sr}}^{\text{FD}}$ to find optimal $R_{\text{rd}}^{\text{FD}}$ and again, using the new resulted $R_{\text{rd}}^{\text{FD}}$ to find optimal $R_{\text{sr}}^{\text{FD}}$. This iterative process repeats until the convergence. Our simulation showed that the algorithm has a very fast convergence and only in rare cases does it take more than 20 iterations for the algorithm to converge. This is mainly due to the fact that inequalities (37d) and (37e) restrict the eigenvalues to vary up to a certain limit, which in turn, makes the whole outputs more stable. Figure 4 depicts a typical histogram of iterations. As it can be seen, only less than 3% of cases did not converge until 50 iterations.

Notice that the optimum values for the transmission power on relay hop may not sum to P_r. The reason is that $R_{\text{sr}}^{\text{FD}}$ is a monotonically decreasing function of P_r and as we are interested in the $\min(R_{\text{sr}}^{\text{FD}}, R_{\text{rd}}^{\text{FD}})$, with $R_{\text{sr}}^{\text{FD}} < R_{\text{rd}}^{\text{FD}}$ we will have $\min(R_{\text{sr}}^{\text{FD}}, R_{\text{rd}}^{\text{FD}}) = R_{\text{sr}}^{\text{FD}}$. Therefore, it is in our interest to keep P_r as low as possible to increase $R_{\text{sr}}^{\text{FD}}$ as much as possible. Analogously, in the case of $R_{\text{sr}}^{\text{FD}} > R_{\text{rd}}^{\text{FD}}$ we have $\min(R_{\text{sr}}^{\text{FD}}, R_{\text{rd}}^{\text{FD}}) = R_{\text{rd}}^{\text{FD}}$ which can be increased by increasing the total power usage of relay's transmitter. As a result, the well-known bisection method can be used to find the optimal rate where we have $R_{\text{sr}}^{\text{FD}} = R_{\text{rd}}^{\text{FD}}$, unless the case $R_{\text{sr}}^{\text{FD}} \geq R_{\text{rd}}^{\text{FD}}$ happens even if the maximum allowed power is used at the relay transmitter. In such a case, the relay–destination link becomes the bottleneck.

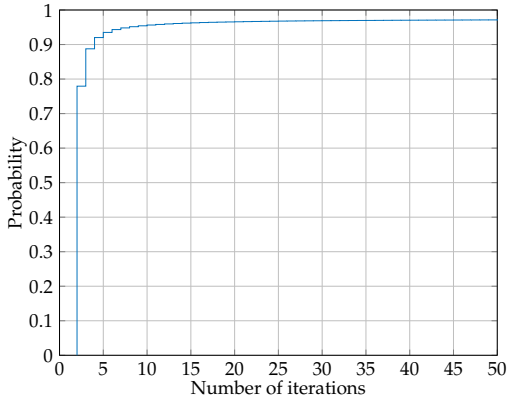

Figure 4. Cumulative distribution function (cdf) of iterations when $P_s = 5$ and $P_r = 1$ and $M = K_t = K_r = N = 10$. The maximum number of iterations is set to be 50. Cases that took more than 50 iterations to converge are considered to be divergent.

Now we focus on how to find R_{sr}^{FD}. In order to find the sum rate for the source–relay hop, we assume that we are already given $\gamma_r^\star$ which is the vector of relay input powers that maximizes the sum rate at the relay–destination hop. The next step is to complete the minimization over $\boldsymbol{\sigma}_r$ and the maximization over $\boldsymbol{\gamma}_s$. One approach to solve this problem is to solve it iteratively. With this method, first one finds the optimal $\boldsymbol{\gamma}_s$ by solving the maximization part of (37) under the assumption that the optimal $\boldsymbol{\sigma}_r$ is given, and then, having the optimal $\boldsymbol{\gamma}_s$, the minimization part of (37) can be solved efficiently. This process goes on until the convergence of $\boldsymbol{\gamma}_s$ and/or $\boldsymbol{\sigma}_r$. The maximization part is performed using the water-filling method. However, the additional conditions $\forall\, i \leq \min(M, K_r), \sigma_{1_i}^2 \gamma_{s_{\rho(i)}} \geq \sigma_{1_{i+1}}^2 \gamma_{s_{\rho(i+1)}}$ should be taken into account. For instance, if the optimal value for γ_{s_i} turns out to be equal to zero, then we should have $\gamma_{s_j} = 0$ for all $j > i$ irrespective of their SNR. Figure 5 depicts two different examples of multi-level water-filling algorithms. As it can be seen, first, a regular water-filling algorithm is considered where for each subchannel we have $\frac{\sigma_{1_i}^2}{1+\gamma_{r_i}\sigma_{r_{\rho(i)}}^2}$ as its channel gain. After finding the water-level in this way, we need to impose $\gamma_{s_{\rho(i+1)}} \leq \frac{\min_{1\leq i'\leq i}\{\sigma_{1_{i'}}^2 \gamma_{s_{\rho(i')}}\}}{\sigma_{1_{i+1}}^2}$. These additional restrictions act like caps on top of the water and create multilevel water-filling which can be interpreted as a cave. Figure 5a shows the case where these caps do not make any subchannel to have zero power. However, Figure 5b shows the case where subchannel $i = 13$ has to be zero as a result of the cap imposed by the additional constraints (37d). In this case, we have $\gamma_{s_{\rho(13)}} = 0$, and as a result $\min_{1\leq i'\leq 13}\{\sigma_{1_{i'}}^2 \gamma_{s_{\rho(i')}}\} = 0$. Thus, this condition forces all other subchannels (i.e., $i > 13$) to get no power. Algorithm 1 provides the detail of multilevel water-filling. For the minimization part, a Lagrangian multiplier is used. We have

$$L = \sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\sigma_{1_i}^2 \gamma_{s_{\rho(i)}}}{\sigma_t^2 + T_1 P_s + \gamma_{r_i}\sigma_{r_{\rho(i)}}^2}\right) + \lambda\left(\sum_{i=0}^{N} \sigma_{r_i}^2 - T_r\right). \tag{38}$$

Calculating $\frac{\partial L}{\partial \sigma_{r_i}^2} = 0$ we arrive at

$$\sigma_{r_i}^2 = \left[\frac{\sqrt{\left(\sigma_{1_i}^2 \gamma_{s_i}\right)^2 + \frac{4\sigma_{1_i}^2 \gamma_{s_i}\gamma_{r_i}}{\lambda}} - \sigma_{1_i}^2 \gamma_{s_i} - 2(\sigma_t^2 + T_1 P_s)}{2\gamma_{r_i}}\right]^+, \tag{39}$$

where λ is the water level.

Similarly to the maximization case, there are additional constraints $\gamma_{r_i}\sigma^2_{s_{\rho(i)}} \geq \gamma_{r_{i+1}}\sigma^2_{r_{i+1}}$ that must be considered during the minimization process. However, it can be shown that if the constraints $\gamma_{r_i} \geq \gamma_{r_{i+1}}$ and $\sigma^2_{1_i}\gamma_{s_{\rho(i)}} \geq \sigma^2_{1_{i+1}}\gamma_{s_{i+1}}$ are met, then the constraint $\gamma_{r_i}\sigma^2_{s_{\rho(i)}} \geq \gamma_{r_{i+1}}\sigma^2_{r_{i+1}}$ becomes redundant. Please refer to Appendix E for proof. The summary of the algorithm to find the achievable rate can be found in Algorithm 2. Next we deal with the optimization for the cases where IRS is utilized to help either the source–relay or relay–destination channels. In such cases, the optimization part over the covariance matrices remains the same as the abovementioned case. In addition, the optimization of the IRS elements can be performed using eigenvalue decomposition and the algorithm introduced in [8]. Notice that for the case in which IRS is assisting the source–relay link, the term T_1P_s in (27) should be replaced with $(T_1 + T_{SI}T_{IR})P_s$, and for the case where IRS helps the relay–destination link, the term T_2P_s in (28) should be replaced with $(T_2 + T_{RI}T_{ID})P_r$. The pseudo code for these scenarios is given in Algorithm 3.

Algorithm 1 The optimal γ_s

Find power allocation $\boldsymbol{P}^0$ using water-filling algorithm
while $|\boldsymbol{P}^{(q)} - \boldsymbol{P}^{(q-1)}|$ is large **do**
  Define temp $= 0$
  for i **do**
    Calculate $\text{cap}_i = \min_{1\leq i'\leq i-1}\{\sigma^2_{1_{i'}}\gamma_{s_{\rho(i')}}\}/\sigma^2_{1_i}$
    if $P_i > \text{cap}_i$ **then**
      $P_i = \text{cap}_i$
      $\text{temp} = \text{temp} + P_i - \text{cap}_i$
    end if
  end for
  $\boldsymbol{P} = \boldsymbol{P} + \frac{\text{temp}}{\text{number of channels}}$
end while

Algorithm 2 Robust Transceiver Design for FD scenario, the first case

Define $U = P_r$, $L = 0$, $\bar{P}_r^{(1)} = \frac{P_r}{2}$
while $|U - L|$ is large **do**
  Determine $\gamma_r = [\tau_r - \frac{1}{\sigma_2^2}]^+$, s.t. $\|\gamma_r\|_1 = \bar{P}_r$
  Define $\boldsymbol{\sigma}_r^{(0)^2} = \mathbf{0}$ and $\boldsymbol{\sigma}_r^{(1)^2} = \mathbf{1}$ and $q = 0$
  Set $T' = \left(\sqrt{T_r} - \sigma_{\min}(\mathbf{H}_{IR} * \mathbf{H}_{RI}^T)\right)^2$
  while $\|\boldsymbol{\sigma}_r^{(q)^2} - \boldsymbol{\sigma}_r^{(q-1)^2}\|_1$ is large **do**
    Obtain $\boldsymbol{\sigma}_r^{(q)^2}$, using Equation (39)
    Obtain $\gamma_s^{(q)}$, using Algorithm 1
    $q = q + 1$
  end while
  Calculate R_{sr} and R_{rd}
  if $R_{\text{sr}} > R_{\text{rd}}$ **then**
    $U = \bar{P}_r$
  else if $R_{\text{sr}} < R_{\text{rd}}$ **then**
    $L = \bar{P}_r$
  end if
  $\bar{P}_r = \frac{U+L}{2}$
end while

Algorithm 3 Robust Transceiver Design for FD scenario, the second and third case

Define $U = P_r$, $L = 0$, $\bar{P}_r^{(1)} = \frac{P_r}{2}$
Define $Case = 1$ if the IRS is being used help the source–relay link or $Case = 0$ if IRS is being used to help the relay–destination link
while $|U - L|$ is large **do**
    Determine $\gamma_r = [\tau_r - \frac{1}{\sigma_2^2}]^+$, s.t. $\|\gamma_r\|_1 = \bar{P}_r$
    Define $\sigma_r^{(0)^2} = \mathbf{0}$ and $\sigma_r^{(1)^2} = \mathbf{1}$ and $q = 0$
    if $Case = 1$ **then**
        $\bar{\mathbf{H}}_{tot_2} = (\bar{\mathbf{H}}_1 + \bar{\mathbf{H}}_{RI}\mathbf{\Theta}\bar{\mathbf{H}}_{IR})$
        Find the optimum $\mathbf{\Theta}$, using Algorithm 1 in [8]
    else if $Case = 0$ **then**
        $\bar{\mathbf{H}}_{tot_3} = (\bar{\mathbf{H}}_2 + \bar{\mathbf{H}}_{RI}\mathbf{\Theta}\bar{\mathbf{H}}_{IR})$
        Find the optimum $\mathbf{\Theta}$, using Algorithm 1 in [8]
    end if
    Obtain $\gamma_s^{(q)}$, using Algorithm 1
    $q = q + 1$
end while
Calculate R_{sr} and R_{rd}
if $R_{\text{sr}} > R_{\text{rd}}$ **then**
    $U = \bar{P}_r$
else if $R_{\text{sr}} < R_{\text{rd}}$ **then**
    $L = \bar{P}_r$
end if
$\bar{P}_r = \frac{U+L}{2}$

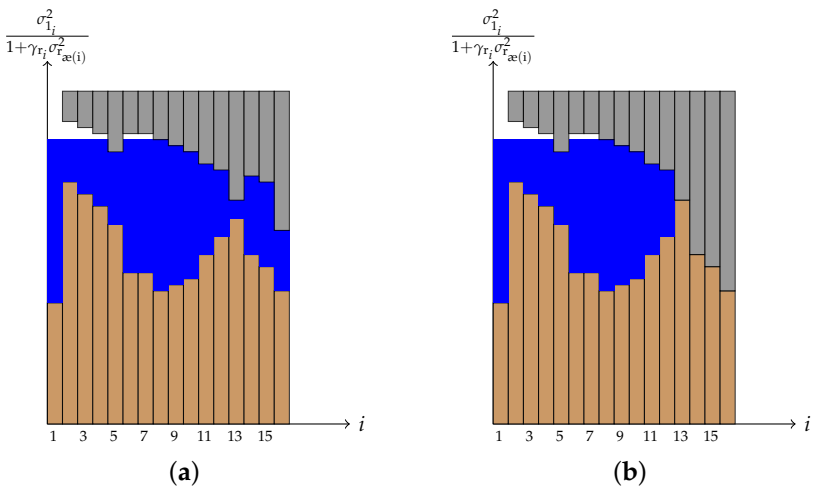

Figure 5. Examples of multilevel water-filling for two different cases. (**a**) No subchannel with optimum power equal to zero. Note that due to the power cap constraint (37d), water does not have the same level for all subchannels. (**b**) Subchannel $i = 13$ receives the optimum of zero as its input power. Notice that in this case, due to the additional power cap constraint (37d), all the remaining subchannels $i > 13$ also have zero power.

3.4. Discussion

In this part, we evaluate the various aspects of our method. First, we examine the complexity of our algorithm and compare it with the state of the art. Algorithms 1 and 2 are the main solutions provided in this paper. Algorithm 1 is a multi-level water-filling, and as a result, it has the complexity of $\mathcal{O}(I_w \min(N_t, K_r))$, where I_w is a constant that is independent of system parameters and is only related to the accuracy of the multi-level water-filling algorithm. Algorithm 2 requires the SVD for matrices $\mathbf{H}_1$, $\mathbf{H}_2$ and $(\mathbf{H}_{IR} * \mathbf{H}_{RI}^T)$, with the complexity $\mathcal{O}(N_t K_r \min(N_t, K_r))$, $\mathcal{O}(N_r K_t \min(N_r, K_t))$ and $\mathcal{O}(M K_r K_t \min(M, K_r K_t))$, re-

spectively. Furthermore, the Khatri–Rao multiplication $\mathbf{H}_{IR} * \mathbf{H}_{RI}^T$ is needed that has the complexity $\mathcal{O}(N_r K_t M)$. As a result, the overall complexity of our method is $\mathcal{O}(MK_rK_t \min(M, K_rK_t) + N_rK_tM + N_tK_r \min(N_t, K_r) + N_rK_t \min(N_r, K_t) + I_t(I_w \min(N_t, K_r)))$, where I_t is a constant independent of the system parameters. Interestingly, our method has a super linear complexity with respect to the number of IRS elements which is better than the state of the art works, e.g., [5,8]. This means that our algorithm in more energy efficient and suitable for latency sensitive applications. It should also be noted that our algorithm does not provide the optimal IRS pattern; instead, it provides analytical bounds for the performance of the IRS that can be used as a benchmark. In other words, our work provides a tool with which one can evaluate the efficiency of their robust design. A comparison between our method and previous works is summarized in Table 2

Table 2. Comparison of the proposed method with previous studies.

Method	Complexity	Robust Design *	Solution	System Model
This work	$\mathcal{O}(MK_rK_t \min(M, K_rK_t))$	Yes	Analytical bounds	Relay and IRS
Zhang et al. [8]	$\mathcal{O}(I_O M^3)$	No	Practical solution	Only IRS
Esmaeili et al. [34]	$\mathcal{O}(I_t(I_w \min(N_t, K_r)))$	Yes	Practical solution	Only relay
Obeed et al. [9]	$\mathcal{O}(I_\lambda I_O I_w M^3)$	No	Practical solution	Relay and IRS

* This means that the method is robust against channel uncertainties and/or RSI.

4. Achievable Rate (Half-Duplex Relay)

We consider a simple HD relay where the source and the relay transmit in two subsequent time instances. Notice that for the case of HD, IRS can be used to assist both the source–relay and the relay–destination channels as the signal is being sent over each of these channels in a different time slot. Therefore, the received signals at the relay and the destination can, respectively, be expressed as

$$\mathbf{y}_\mathrm{r} = (\hat{\mathbf{H}}_1 + \hat{\mathbf{H}}_{IR}\boldsymbol{\Theta}\hat{\mathbf{H}}_\mathrm{SI})\mathbf{x}_\mathrm{r} + (\bar{\mathbf{H}}_1 + \bar{\mathbf{H}}_{IR}\boldsymbol{\Theta}\bar{\mathbf{H}}_\mathrm{SI})\mathbf{x}_\mathrm{r} + \mathbf{n}_\mathrm{t}, \tag{40}$$

$$\mathbf{y}_\mathrm{d} = (\hat{\mathbf{H}}_2 + \hat{\mathbf{H}}_{ID}\boldsymbol{\Theta}\hat{\mathbf{H}}_{RI})\mathbf{x}_r + (\bar{\mathbf{H}}_2 + \bar{\mathbf{H}}_{ID}\boldsymbol{\Theta}\bar{\mathbf{H}}_{RI})\mathbf{x}_r + \mathbf{n}_\mathrm{d}. \tag{41}$$

Consequently, the achievable rates for the transmitter–relay and relay–destination links can be expressed as below

$$\begin{aligned} R_\mathrm{sr}^\mathrm{HD} = \log \Big| & \mathbf{I}_{K_r} + (\hat{\mathbf{H}}_1 + \hat{\mathbf{H}}_{IR}\boldsymbol{\Theta}\hat{\mathbf{H}}_\mathrm{SI})\mathbf{Q}_s(\hat{\mathbf{H}}_1 + \hat{\mathbf{H}}_{IR}\boldsymbol{\Theta}\hat{\mathbf{H}}_\mathrm{SI})^H \\ & \left(\sigma_t^2 \mathbf{I}_{K_r} + (\bar{\mathbf{H}}_1 + \bar{\mathbf{H}}_{IR}\boldsymbol{\Theta}\bar{\mathbf{H}}_\mathrm{SI})\mathbf{Q}_s(\bar{\mathbf{H}}_1 + \bar{\mathbf{H}}_{IR}\boldsymbol{\Theta}\bar{\mathbf{H}}_\mathrm{SI})^H\right)^{-1} \Big| \\ \geq & \sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\lambda_i\left(\mathbf{H}_1'\mathbf{Q}_s\mathbf{H}_1'^H\right)}{\sigma_t^2 + T_1 P_s}\right), \end{aligned} \tag{42}$$

$$\begin{aligned} R_\mathrm{rd}^\mathrm{HD} = \log \Big| & \mathbf{I}_N + (\hat{\mathbf{H}}_2 + \hat{\mathbf{H}}_{ID}\boldsymbol{\Theta}\hat{\mathbf{H}}_\mathrm{RI})\mathbf{Q}_r(\hat{\mathbf{H}}_2 + \hat{\mathbf{H}}_{ID}\boldsymbol{\Theta}\hat{\mathbf{H}}_\mathrm{RI})^H \\ & \left(\sigma_d^2 \mathbf{I}_N + (\bar{\mathbf{H}}_2 + \bar{\mathbf{H}}_{ID}\boldsymbol{\Theta}\bar{\mathbf{H}}_\mathrm{RI})\mathbf{Q}_r(\bar{\mathbf{H}}_2 + \bar{\mathbf{H}}_{ID}\boldsymbol{\Theta}\bar{\mathbf{H}}_\mathrm{RI})^H\right)^{-1} \Big| \\ \geq & \sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\lambda_i\left(\mathbf{H}_2'\mathbf{Q}_r\mathbf{H}_2'^H\right)}{\sigma_d^2 + T_2 P_r}\right). \end{aligned} \tag{43}$$

where $\mathbf{H}_1' = \hat{\mathbf{H}}_1 + \hat{\mathbf{H}}_{IR}\boldsymbol{\Theta}\hat{\mathbf{H}}_\mathrm{SI}$ and $\mathbf{H}_2' = \bar{\mathbf{H}}_2 + \bar{\mathbf{H}}_{ID}\boldsymbol{\Theta}\bar{\mathbf{H}}_\mathrm{RI}$. In addition, $R_\mathrm{sr}^\mathrm{HD}$ and $R_\mathrm{rd}^\mathrm{HD}$ are the achievable rates on the source–relay and relay–destination links, respectively. Using time sharing, the achievable rate between the source and destination nodes is given by

$$R^\mathrm{HD} = \min(\alpha R_\mathrm{sr}^\mathrm{HD}, (1-\alpha) R_\mathrm{rd}^\mathrm{HD}), \tag{44}$$

where α is the time-sharing parameter.

Note that in half-duplex relaying, the source and relay transmissions are conducted in separate channel uses. Hence, the transmit covariance matrices $\mathbf{Q}_s \in \mathbb{H}^{N_t \times N_t}$ and $\mathbf{Q}_r \in \mathbb{H}^{K_t \times K_t}$ are optimized by maximizing the achievable rate from the source to the destination. Here, the convex cone of Hermitian positive semidefinite matrices of dimensions $N_t \times N_t$ and $K_t \times K_t$ are represented by $\mathbb{H}^{N_t \times N_t}$ and $\mathbb{H}^{K_t \times K_t}$, respectively. Importantly, for maximizing this achievable rate, the time-sharing parameter, i.e., α, needs to be optimized alongside the system parameters, e.g., power allocation. Readily, optimal α occurs at $\alpha R_{\text{sr}}^{\text{HD}} = (1-\alpha) R_{\text{rd}}^{\text{HD}}$. Therefore, the achievable rate becomes as follows

$$R^{\text{HD}} = \frac{R_{\text{sr}}^{\text{HD}} R_{\text{rd}}^{\text{HD}}}{R_{\text{sr}}^{\text{HD}} + R_{\text{rd}}^{\text{HD}}}. \tag{45}$$

Notice that as the objective function of the above optimization problem is a monotonically increasing function of both $R_{\text{sr}}^{\text{HD}}$ and $R_{\text{rd}}^{\text{HD}}$, the problem can be simplified to maximizing $R_{\text{sr}}^{\text{HD}}$ and $R_{\text{rd}}^{\text{HD}}$ separately.

Next, we provide the solution to the rate optimization problem when IRS is assisting the source–relay link. We have

$$\max_{\mathbf{Q}_s} \quad \min_{\bar{\mathbf{H}}_{SI}, \bar{\mathbf{H}}_1, \bar{\mathbf{H}}_{\text{IR}}} \quad R_{\text{sr}}^{\text{HD}} \tag{46}$$

$$\text{subject to} \quad \text{Tr}(\mathbf{Q}_s) \leq P_s, \tag{46a}$$

$$\text{Tr}(\bar{\mathbf{H}}_1 \bar{\mathbf{H}}_1^H) \leq T_1, \tag{46b}$$

$$\text{Tr}(\bar{\mathbf{H}}_{SI} \bar{\mathbf{H}}_{SI}^H) \leq T_{SI}, \tag{46c}$$

$$\text{Tr}(\bar{\mathbf{H}}_{IR} \bar{\mathbf{H}}_{IR}^H) \leq T_{IR}, \tag{46d}$$

The above optimization problem follows the same approach applied for the optimization of the relay–destination link in the FD scenario. As a result, the same method could be applied to find it. In other words, the well-known water-filling algorithm can be used to find the optimal covariance matrices along with the algorithm introduced in [8] to find the best IRS pattern. This process continues iteratively until it finally converges. The solution to R_{rd}^{HD} is the same as well, and the same procedure can be applied to find $\mathbf{Q}_r$. The overall procedure of finding the solution for the HD mode is summarized in Algorithm 4.

Algorithm 4 Robust Transceiver Design for HD scenario

1: $\bar{\mathbf{H}}_{tot_2} = (\bar{\mathbf{H}}_1 + \bar{\mathbf{H}}_{RI} \boldsymbol{\Theta} \bar{\mathbf{H}}_{IR})$
2: Find the optimum $\boldsymbol{\Theta}$, using Algorithm 1 in [8]
3: $\bar{\mathbf{H}}_{tot_3} = (\bar{\mathbf{H}}_2 + \bar{\mathbf{H}}_{RI} \boldsymbol{\Theta} \bar{\mathbf{H}}_{IR})$
4: Find the optimum $\boldsymbol{\Theta}$, using Algorithm 1 in [8]
5: Find $R_{\text{sr}}^{\text{HD}}$, using Equation (42)
6: Find $R_{\text{rd}}^{\text{HD}}$, using Equation (43)
7: Find R^{HD}, using Equation (45)

5. Numerical Results

We assume the transmit power budgets at the source and at the relay are $P_s = 5$ and $P_r = 1$, respectively. Moreover, the AWGN spectral density is assumed to be -175 dBm and the bandwidth is $BW = 180$ MHz. In this section, we investigate the performance of IRS-assisted full-duplex relaying with RSI channel uncertainty bound T_r, i.e., $\mathbf{Tr}(\bar{\mathbf{H}}_r \bar{\mathbf{H}}_r^H) \leq T_r$. We consider all the channels to follow the Rician distribution with the factor $\epsilon = 0.1$ and the specificaiton given in Table 3. We also assume $T_x = 0.001, x \in \{1, 2, SI, IR, RI, ID\}$. We perform Monte Carlo simulations with $L = 10^3$ realizations from random channels and noise vectors. Hence, the average worst-case throughput rate is defined as the average of worst-case rates for L randomization, i.e., $R_{\text{av}} = \frac{1}{L} \sum_{l=1}^{L} R_l$. Notice that for each set of

realizations, we solve the robust transceiver design as is elaborated in Algorithm 2. We run different sets of simulations as described in the following subsections.

Table 3. Simulation parameters.

Parameters	Values
Transmitter location	$(0\text{ m}, 0\text{ m})$
IRS location	$(4000\text{ m}, 20\text{ m})$
Relay location	$(4000\text{ m}, 0\text{ m})$
Receiver location	$(8000\text{ m}, 0\text{ m})$
Path-loss	$32.6 + 36.7\log(d)$
Transmission bandwidth B	180 Mb

5.1. Antenna Array Increment with No IRS

In this part, first we assume that there is no IRS installed. Then we evaluate the performance of the system using different strategies. Thereafter, we examine how installing an IRS can help increase the throughput. We consider two cases where the source, relay and destination are equipped with (a) a small antenna array, and (b) a large antenna arrays. In order to see the impact of IRS, we first assume that there is no IRS installed. For these cases, we have

(a) $N_t = 4, K_r + K_\mathrm{t} = 10, N_r = 4,$
(b) $N_t = 10, K_r + K_\mathrm{t} = 24, N_r = 10.$

These cases are considered to highlight the performance of full-duplex DF relaying as a function of the number of antennas with the worst-case RSI. Interestingly, as the number of antennas at the source, relay and destination increase, full-duplex relaying achieves a higher throughput rate even with strong RSI. This can be seen by comparing rates from Figure 6a to those from Figure 6b.

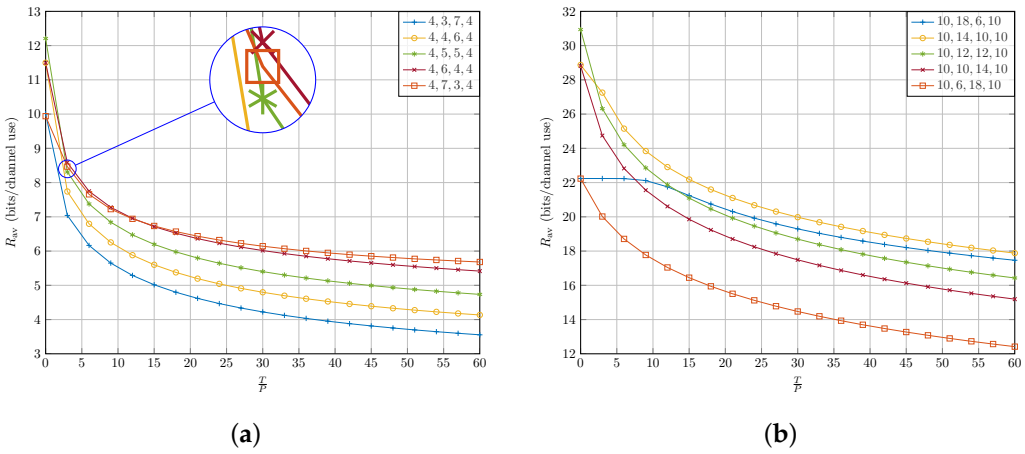

Figure 6. Sum rate throughput of the FD mode as a function of normalized interference $\frac{T}{P}$ with different number of antennas at the source, relay and destination. The transmit power budget at the source and the relay are assumed to be equal, i.e., $P_s = 5$ and $P_\mathrm{r} = 1$. (**a**) $\{N_t, K_r + K_\mathrm{t}, N_r\} = \{4, 10, 4\}$. (**b**) $\{N_t, K_r + K_\mathrm{t}, N_r\} = \{10, 24, 10\}$.

Furthermore, notice that the worst-case RSI casts strong interference on the strong streams from the source to the destination. With very low RSI power $T_r \to 0$, full-duplex almost doubles the throughput rate compared to the half-duplex counterpart. This can be seen in Figure 6, where the curves have their intercept point with the vertical axis.

However, as T_r increases, the efficiency of full-duplex operation drops. It is worth noting that at low RSI power the DoF plays the most important role to have a higher sum rate. For instance, consider Figure 6a in which the cases FD = {4,5,5,4}, FD = {4,4,6,4} and FD = {4,6,4,4} have $\mathrm{DoF_{total}} = 4$ – $\mathrm{DoF_{total}}$ is the minimum of the DoF of source–relay and relay–destination channels, i.e., $\mathrm{DoF_{total}} = \min(\mathrm{DoF_{sr}}, \mathrm{DoF_{rd}})$ – while the cases FD = {4,7,3,4} and FD = {4,3,7,4} have $\mathrm{DoF_{total}} = 3$. At $T_r = 0$, there is a noticeable gap between the first three cases and the last two, while the difference of the first three cases from each other is small. The big gap is due to the difference in $\mathrm{DoF_{total}}$, and the small one is due to the difference in SNR. Similarly, in Figure 6b, the three cases FD = {10,12,12,10}, FD = {10,10,14,10} and FD = {10,10,14,10} with $\mathrm{DoF_{total}} = 10$ have higher rates than the two cases FD = {10,6,18,10} and FD = {10,18,6,10} with $\mathrm{DoF_{total}} = 6$.

Finally, it can be seen in both Figure 6a,b that at $T_r = 0$ there is no difference between cases that have the same $\mathrm{DoF_{total}}$ but different $\mathrm{DoF_{sr}}$ and $\mathrm{DoF_{rd}}$. As it can be seen in both Figure 6a,b, for cases with $K_t > K_r$ the sum rate drops quickly as RSI increases. In fact, the more relative antennas at the relay transmitter compared to its receiver, the faster the sum rate drops with the rise in RSI. To understand this behaviour of the system better, again, consider case {10,18,6,10} and also suppose $T_r \to \infty$. As discussed before, we have $\mathrm{DoF_{sr}} = 10$ and $\mathrm{DoF_{rd}} = 6$. Moreover, we have $\mathrm{DoF_I} = 6$ for the interference channel ($\bar{\mathbf{H}}_\mathrm{r}$). Unlike the case with no interference, in this case the bottleneck is no longer the relay–destination link. This is due to the fact that interference can act to the detriment of some six of the source–relay subchannels. As we have $\mathrm{DoF_I} = 6$, interference can choose at most six independent subchannels, and as we assumed $T_r \to \infty$, for those subchannels we obtain SINR $\to 0$. Therefore, no information can be conveyed from those links, and the bottleneck becomes the source–relay link with 4 usable subchannels. It can be seen in Figure 6a,b that as T_r increases, the cases with the same sum-rate at $T_r = 0$ start to diverge because of the different characteristics of the interference they experience. We explain the effect of interference in the following subsection in more detail.

5.2. The Impact of IRS

In this part, we evaluate the impact of IRS on the throughput rate when it is used to perform different tasks. Figure 7 shows the throughput for three different scenarios, namely, when IRS is used to help the transmitter–relay link, when it is applied to cancel RSI and when the IRS job is to help the relay–destination link. Then the results are compared with two cases where the system is working in HD with IRS and the case where the system is working in FD with no IRS. It is also assumed that $\frac{T_r}{(\mathbf{H}_r\mathbf{H}_r^H)} = 75\%$, i.e., the system works at high RSI range. As it can be seen, the highest performance is achieved when the IRS is utilized to deal with the RSI. As a result, for the rest of the paper we use the IRS for this purpose.

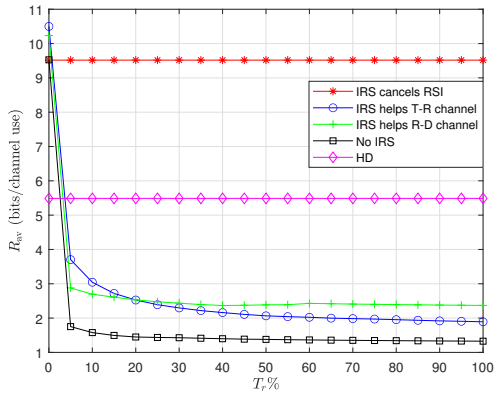

Figure 7. Comparison of five different scenarios: HD with IRS, FD with no IRS, FD with IRS as RSI cancelator; FD with IRS to help transmitter–relay link; FD with IRS to help relay–destination link. We considered the case where $\{N_t, K_t, K_r, N_r, M\} = \{4, 5, 5, 4, 100\}$.

Figure 8 shows the impact of IRS on the throughput. For Figure 8a, we considered the case $\{N_t, K_t, K_r, N_r\} = \{4, 5, 5, 4\}$, and for Figure 8b, we considered $\{N_t, K_t, K_r, N_r\} = \{10, 12, 12, 10\}$. As shown in the figure, the number of IRS elements has a great impact on RSI cancellation to the extend that having an IRS with $M = 100$ and $M = 300$ can cancel interference of $\frac{T_r}{\mathrm{Tr}(\mathbf{H}_r\mathbf{H}_r^H)} = 0.75$ for $\{N_t, K_t, K_r, N_r\} = \{4, 5, 5, 4\}$, and $\{N_t, K_t, K_r, N_r\} = \{10, 12, 12, 10\}$, respectively. Further, it can also be seen in the figure that having IRS with 20 and 100 elements for the small and large antenna array cases, respectively, is not helpful at all. This is mainly due to the fact that unlike the average case, for the case of worst-case scenario, the number of IRS elements should be at least as large as the dimension of $\bar{\mathbf{H}}_r$. Otherwise, the IRS feasible set cannot span into all dimensions of $\bar{\mathbf{H}}_r$. Therefore, there is always at least one representation for $\bar{\mathbf{H}}_r$ in which IRS cannot perform any RSI cancellation. In addition, comparing two figures Figure 8a,b one can conclude that, when the dimension of $\bar{\mathbf{H}}_r$ increases, the effort that IRS has to make in order to cancel RSI remarkably increases which is consistent with the previous statement.

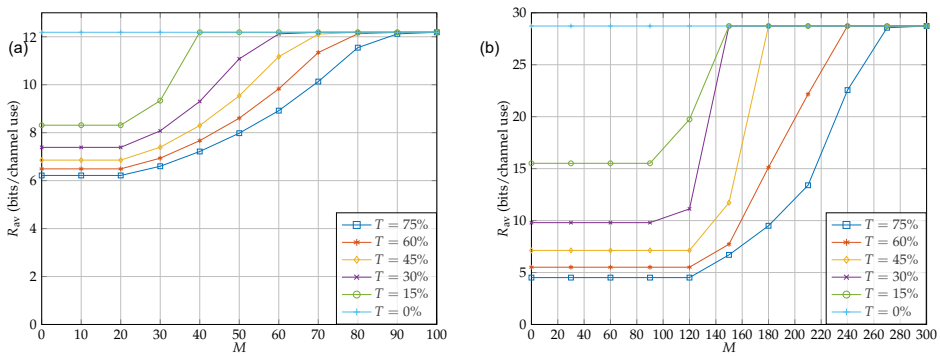

Figure 8. Sum rate throughput as a function of IRS elements. The transmit power budget at the source and the relay are assumed to be equal, i.e., $P_s = 5$ and $P_r = 1$: (**a**) $\{N_t, K_t, K_r, N_r\} = \{4, 5, 5, 4\}$; (**b**) $\{N_t, K_t, K_r, N_r\} = \{10, 12, 12, 10\}$.

5.3. Relay Tx/Rx Antenna allocation

Suppose that the relay has $K_t + K_r = 8$ in total. Furthermore, following cases in which the number of antenna at the source and destination are $\{N_t, N_r\} = \{4, 4\}$. The question is, from eight antennas at the relay, how many should be used for reception for the robust design? Figure 9 shows the sum rate as a function of K_r for different values of T where there is no IRS and there is an IRS with $M = 60$ elements, respectively. As it can be seen, by using more antennas for reception than for transmission, i.e., $K_r > K_t$, at the relay, the throughput rate is maximized. This is due to the fact that increasing the signal-to-noise ratio (SNR) of the source–relay streams enhances the overall throughput rate more than increasing the number of antennas for transmission in order to enhance the DoF of the relay–destination link. Furthermore, notice that in this setup the overall DoF from the source to destination is limited by the DoF of the source–relay link, i.e., the bottleneck is in the first hop.

By comparing two scenarios, we see that having an IRS not only improves the rates in all cases, but also it may change the best antenna allocation. For instance, for the case of $T = 15\%$, it is best to have six antennas at the relay receiver and four4 antennas at the relay transmitter. However, after establishing the IRS, the best antenna allocation changes to five antennas at each end. For instance, the results show that although the $\mathrm{DoF}_{\mathrm{total}}$ for both $\{N_t, K_t, K_r, N_r\} = \{4, 3, 5, 4\}$ and $\{N_t, K_t, K_r, N_r\} = \{4, 5, 3, 4\}$ is three, the sum rate capacity of the latter is much better than that of the former at high interference. This is because of the fact when $\mathrm{DoF}_{\mathrm{sr}} > \mathrm{DoF}_{\mathrm{rd}}$, the source–relay link enjoys $\mathrm{DoF}_{\mathrm{sr}} - \mathrm{DoF}_{\mathrm{rd}}$ subchannels with no interference. Therefore, the source can manage to obtain a higher sum rate by choosing its power allocation wisely. However, in the case of $\mathrm{DoF}_{\mathrm{sr}} \leq \mathrm{DoF}_{\mathrm{rd}}$, no matter how well the power allocation is performed, all sub channels suffer from interference at the source–relay end.

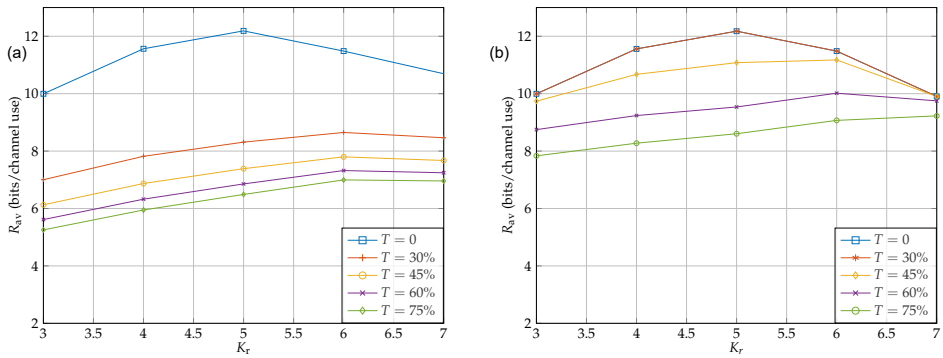

Figure 9. Sum rate throughput as a function of relay receiver antennas K_r with and without RSI. The transmit power budget at the source and the relay are assumed to be equal, i.e., $P_s = 5$ and $P_r = 1$: (**a**) $\{N_t, K_t + K_r, N_r\} = \{4, 10, 4\}$ with no IRS; (**b**) $\{N_t, K_t + K_r\ N_r\} = \{4, 10, 4\}$ with M = 60.

5.4. Full-Duplex vs. Half-Duplex

In this subsection, we determine the thresholds where the HD relaying outperforms the FD relaying. This threshold provides a mode-switching threshold in hybrid HD/FD relay systems. As it can be seen in Figure 10, for each case of K_r, there is a maximum value of $\frac{T}{P}$ above which the HD mode outperforms the FD mode in terms of sum rate maximization. Furthermore, Figure 10 shows the threshold for different IRS configurations. For this part, we continued with the case of $N_t = 4, K_r = 5, K_t = 5, N_r = 4$. As it can be seen, by increasing the number of antennas, the threshold occurs at higher RSI. This is in fact a direct result of obtaining better performance by having more antennas at the relay's receiver. It is worth noting that the IRS has a great impact on the performance of FD relaying. For instance, by having an IRS consisting of only 60 elements, the FD mode outperforms the HD mode in almost all cases.

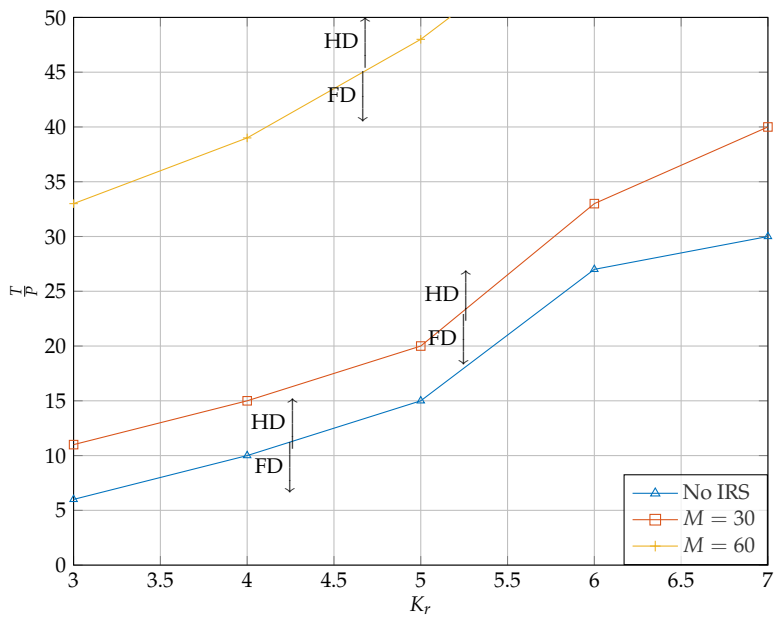

Figure 10. Thresholds for different K_r and M. The region above each curve indicates values of $\frac{T}{P}$ for which HD outperforms FD. In contrast, points below the curve belong to cases where FD performs better than HD.

6. Conclusions

In this paper, we investigated a multi-antenna source communicating with a multi-antenna destination through a multi-antenna relay. The relay is assumed to exploit a decode-and-forward (DF) strategy. An IRS is installed to help the relay cope with the RSI.

The transceivers are designed in order to be robust against the worst-case residual self-interference (RSI). To this end, the worst-case achievable throughput rate is maximized. This optimization problem turns out to be a non-convex problem. Assuming that the degrees-of-freedom (DoF) of the source–relay link is less than the DoF of the relay–destination link, we determined the left and right matrices of the singular vectors of the worst-case RSI channel. Then, the problem is simplified to the optimal power allocation at the transmitters, which guarantees robustness against the worst-case RSI singular values. This simplified problem is still non-convex. Based on the intuitions for optimal power allocation at the source and relay, we proposed an efficient algorithm to capture a stationary point. Our proposed method showed a significant improvement in robustness. More precisely, we showed that in the case of high uncertainty, using our method can lead to at least 100% worst-case throughput improvement for the case of few antenna arrays and up to 500% for the case of large antenna arrays at transceivers. Furthermore, we confirmed that there is a direct relation between the performance of the system and the number of IRS elements. The simulations show that having the IRS with as low as 90 and 300 elements can completely remove the RSI for our system configuration. Finally, we showed that when there is no RSI, the impact of the relay can be fully harnessed where the number of antennas are equal at the relay transmitter and receiver. Therefore, employing the IRS to deal with the RSI can lead to the best performance of the relay.

Author Contributions: Conceptualization, H.E.; methodology, A.S. and H.E.; software, H.E.; validation, H.E., A.K. and A.S.; formal analysis, H.E; investigation, H.E, A.K. and A.S.; resources, H.E, A.K. and A.S.; data curation, H.E., A.K., A.S.; writing—original, draft preparation, H.E.; writing—review and editing, H.E.; visualization, H.E., A.K. and A.S.; supervision, A.S.; project administration, A.S.; funding acquisition, A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported in part by the German Research Foundation, Deutsche Forschungsgemeinschaft (DFG), Germany, under the project grant number SE1967/16.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data can be provided by the authors H.E. upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Equation (16)

To show the proof, we use the Bra–Ket notation. We have

$$\mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IR} = \left(\sum_{r_1=1}^{K_r}\sum_{r_2=1}^{M} h_{RI_{r_1,r_2}}\,|r_1\rangle\langle r_2|\right)\left(\sum_{r_3=1}^{M}\theta_{r_3,r_3}\,|r_3\rangle\langle r_3|\right)\left(\sum_{r_4=1}^{M}\sum_{r_5=1}^{K_t} h_{IR_{r_4,r_5}}\,|r_4\rangle\langle r_5|\right) \tag{A1}$$

$$= \sum_{r_1=1}^{K_r}\sum_{r_2=1}^{M}\sum_{r_3=1}^{M}\sum_{r_4=1}^{M}\sum_{r_5=1}^{K_t} h_{RI_{r_1,r_2}}\theta_{r_3,r_3}h_{IR_{r_4,r_5}}|r_1\rangle\langle r_2|r_3\rangle\langle r_3|r_4\rangle\langle r_5| \tag{A2}$$

$$= \sum_{r_1=1}^{K_r}\sum_{r_3=1}^{M}\sum_{r_5=1}^{K_t} h_{RI_{r_1,r_3}}\theta_{r_3,r_3}h_{IR_{r_3,r_5}}|r_1\rangle\langle r_5|\,. \tag{A3}$$

Now, after turning the matrix representation into the vector representation we obtain

$$Vec(\mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IR}) = \sum_{r_1=1}^{K_r}\sum_{r_3=1}^{M}\sum_{r_5=1}^{K_t} h_{RI_{r_1,r_3}}\theta_{r_3,r_3}h_{IR_{r_3,r_5}}|r_1,r_5\rangle \tag{A4}$$

$$= \sum_{r_1=1}^{K_r}\sum_{r_3=1}^{M}\sum_{r_4=1}^{M}\sum_{r_5=1}^{K_t} h_{RI_{r_1,r_4}}\theta_{r_3,r_3}h_{IR_{r_4,r_5}}|r_1,r_5\rangle\,\langle r_4|r_3\rangle \tag{A5}$$

$$= \sum_{r_1=1}^{K_r} \sum_{r_4=1}^{M} \sum_{r_5=1}^{K_t} h_{RIr_1,r_4} h_{IRr_4,r_5} |r_1, r_5\rangle \langle r_4| \sum_{r_3=1}^{M} \theta_{r_3,r_3} |r_3\rangle \tag{A6}$$

$$= (\mathbf{H}_{IR} * \mathbf{H}_{RI}^T) Vec(\mathbf{\Theta}). \tag{A7}$$

As a result, we have, $||\bar{\mathbf{H}}_r + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IR}||_F^2 = ||Vec(\bar{\mathbf{H}}_r) + (\mathbf{H}_{IR} * \mathbf{H}_{RI}^T) Vec(\mathbf{\Theta})||_2^2$, and the proof is complete.

Appendix B. Proof of Theorem 1

Before stating the proof, first we introduce the following definitions.

Definition A1. *For a vector $\boldsymbol{a}$, we denote vector $\boldsymbol{a}^{\downarrow}$ which has the same components as $\boldsymbol{a}$ except that they are sorted in a decreasing order.*

Definition A2. *Vector $\boldsymbol{a}$ is said to be majorized by vector $\boldsymbol{b}$ and denoted by $\boldsymbol{a} \prec \boldsymbol{b}$ if:*

$$\sum_{i=1}^{K} a_i^{\downarrow} \leq \sum_{i=1}^{K} b_i^{\downarrow}, \tag{A8}$$

$$\sum_{i=1}^{N} a_i^{\downarrow} = \sum_{i=1}^{N} b_i^{\downarrow}, \tag{A9}$$

where $a_i^{\downarrow}$ is the i'th component of $\boldsymbol{a}^{\downarrow}$, N_r is the number of vector components and $K \leq N$. If the last equality does not hold, $\boldsymbol{a}$ is said to be weakly majorized by $\boldsymbol{b}$ and denoted by $\boldsymbol{a} \prec_w \boldsymbol{b}$.

Definition A3. *Vector $\boldsymbol{a}$ is said to be multiplicatively majorized by vector $\boldsymbol{b}$ and denoted by $\boldsymbol{a} \prec_\times \boldsymbol{b}$ if:*

$$\prod_{i=1}^{K} a_i^{\downarrow} \leq \prod_{i=1}^{K} b_i^{\downarrow}, \tag{A10}$$

$$\prod_{i=1}^{N} a_i^{\downarrow} = \prod_{i=1}^{N} b_i^{\downarrow}. \tag{A11}$$

In addition, it is easy to check

$$\boldsymbol{a} \prec_\times \boldsymbol{b} \quad \Leftrightarrow \quad \log(\boldsymbol{a}) \prec \log(\boldsymbol{b}). \tag{A12}$$

To begin with, we know that for $n \times m$ matrix $\boldsymbol{A}$ and $m \times n$ matrix $\boldsymbol{B}$ we have $\lambda_i(\boldsymbol{AB}) = \lambda_i(\boldsymbol{BA})$, $\forall i \in \{1, \cdots, \min(m, n)\}$. In addition, the only difference between eigenvalues of $\boldsymbol{BA}$ and $\boldsymbol{AB}$ are the number of eigenvalues 0. Thus, non-zero eigenvalues of $\bar{\mathbf{H}}_r\mathbf{Q}_r\bar{\mathbf{H}}_r^H$ and $\mathbf{Q}_r\bar{\mathbf{H}}_r^H\bar{\mathbf{H}}_r$ and also $\mathbf{H}_1\mathbf{Q}_s\mathbf{H}_1^H$ and $\mathbf{Q}_s\mathbf{H}_1^H\mathbf{H}_1$ are equal, respectively. Notice that all $\mathbf{Q}_s$, $\mathbf{H}_1^H\mathbf{H}_1$, $\mathbf{Q}_r$ and $\bar{\mathbf{H}}_r^H\bar{\mathbf{H}}_r$ are square matrices. For $\mathbf{H}_1^H\mathbf{H}_1$ and $\bar{\mathbf{H}}_r^H\bar{\mathbf{H}}_r$ we define $\lambda_i(\mathbf{H}_1^H\mathbf{H}_1) = \sigma_i^2(\mathbf{H}_1)$ and $\lambda_i(\bar{\mathbf{H}}_r^H\bar{\mathbf{H}}_r) = \sigma_i^2(\bar{\mathbf{H}}_r)$, respectively.

As discussed in Remark 1, the equality $\lambda_i(\mathbf{Q}_s\mathbf{H}_1^H\mathbf{H}_1) = \lambda_{\rho(i)}(\mathbf{Q}_s)\lambda_i(\mathbf{H}_1^H\mathbf{H}_1)$ does not hold in general. However, using the definition of determinant one can arrive at the following equality

$$\prod_{i=1}^{\min(M,K_r)} \lambda_i(\mathbf{Q}_s\mathbf{H}_1^H\mathbf{H}_1) = \prod_{i=1}^{\min(M,K_r)} \lambda_i(\mathbf{Q}_s)\sigma_i^2(\mathbf{H}_1). \tag{A13}$$

Now, we define vector $\boldsymbol{\lambda}(\mathbf{Q}_s')$ and set its components to be $\lambda_{\rho(i)}(\mathbf{Q}_s') = \frac{\lambda_i(\mathbf{Q}_s\mathbf{H}_1^H\mathbf{H}_1)}{\sigma_i^2(\mathbf{H}_1)}$. By defining $\lambda_{\rho(i)}(\mathbf{Q}_s')$ instead of $\lambda_i(\mathbf{Q}_s')$, we emphasize that the elements of $\boldsymbol{\lambda}(\mathbf{Q}_s')$ are

not necessarily in decreasing order. Then, we construct the matrix $\mathbf{Q}_s'$ having the same eigenvectors as those of $\mathbf{H}_1^H\mathbf{H}_1$ and the eigenvalues $\lambda_{\rho(i)}(\mathbf{Q}_s')$. One can check that for each i we have $\lambda_i(\mathbf{Q}_s'\mathbf{H}_1^H\mathbf{H}_1) = \lambda_i(\mathbf{Q}_s\mathbf{H}_1^H\mathbf{H}_1)$. In addition, by the definition of $\lambda_{\rho(i)}(\mathbf{Q}_s')$ we have

$$\lambda_{\rho(i)}(\mathbf{Q}_s') = \frac{\lambda_i(\mathbf{Q}_s\mathbf{H}_1^H\mathbf{H}_1)}{\sigma_i^2(\mathbf{H}_1)}, \tag{A14}$$

$$\Rightarrow \log\Big(\lambda_{\rho(i)}(\mathbf{Q}_s')\Big) = \log\Big(\lambda_i(\mathbf{Q}_s\mathbf{H}_1^H\mathbf{H}_1)\Big) - \log\Big(\sigma_i^2(\mathbf{H}_1)\Big), \tag{A15}$$

$$\Rightarrow \log\left(\boldsymbol{\lambda}(\mathbf{Q}_s')\right) = \log\Big(\boldsymbol{\lambda}(\mathbf{Q}_s\mathbf{H}_1^H\mathbf{H}_1)\Big) - \log\Big(\boldsymbol{\sigma}^2(\mathbf{H}_1)\Big). \tag{A16}$$

Lemma A1. *Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be semidefinite Hermitian matrices with $\lambda_{\min(m,n)}(\boldsymbol{AB}) > 0$. Then*

$$\log(\boldsymbol{\lambda}(\boldsymbol{AB})) - \log(\boldsymbol{\lambda}(\boldsymbol{B})) \prec \log(\boldsymbol{\lambda}(\boldsymbol{A})). \tag{A17}$$

Proof. The proof is given in [41] (H.1,e). □

Using the above lemma, we can conclude

$$\log\big(\boldsymbol{\lambda}(\mathbf{Q}_s')\big) \prec \log(\boldsymbol{\lambda}(\mathbf{Q}_s)). \tag{A18}$$

Then, immediately we can conclude

$$\boldsymbol{\lambda}(\mathbf{Q}_s') \prec_\times \boldsymbol{\lambda}(\mathbf{Q}_s). \tag{A19}$$

Remark A1. *It is worth mentioning that, depending on channel realizations, the optimal $\mathbf{Q}_s$ might contain some zero eigenvalues. In such cases, we can simply ignore the zeros and construct matrix $\mathbf{Q}_s'$ with dimension $(n-k)\times(n-k)$. Similarly, in the cases where $\bar{\mathbf{H}}_r^H\bar{\mathbf{H}}_r$ has some zero eigenvalues, we can do the same and proceed to constitute $\bar{\mathbf{H}}_s'$ using only nonzero eigenvalues of $\bar{\mathbf{H}}_r^H\bar{\mathbf{H}}_r$ and add the zeros back to the result again at the end.*

Finally, we use the following lemma to show that $\bar{\mathbf{H}}_\mathrm{r}'$ and $\mathbf{Q}_\mathrm{s}'$ are in the feasible set.

Lemma A2. *For two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, if $\boldsymbol{a} \prec_\times \boldsymbol{b}$, then $\boldsymbol{a} \prec_w \boldsymbol{b}$ follows.*

Proof. The proof is given in [41] (5.A.2.b). □

Exploiting the above lemma, one concludes

$$\boldsymbol{\lambda}(\mathbf{Q}_s') \prec_\times \boldsymbol{\lambda}(\mathbf{Q}_s) \;\Rightarrow\; \boldsymbol{\lambda}(\mathbf{Q}_s') \prec_w \boldsymbol{\lambda}(\mathbf{Q}_s), \tag{A20}$$

which consequently results in

$$\sum_{i=1}^{N}\lambda_i(\mathbf{Q}_s') \le \sum_{i=1}^{N}\lambda_i(\mathbf{Q}_s) \Rightarrow \mathrm{Tr}(\mathbf{Q}_s') \le \mathrm{Tr}(\mathbf{Q}_s). \tag{A21}$$

Therefore, there exists $\mathbf{Q}_s'$ and $\bar{\mathbf{H}}_r'$ fulfilling (29)–(31), which satisfy

$$\sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\lambda_i(\mathbf{H}_1\mathbf{Q}_s\mathbf{H}_1^H)}{1+\lambda_i(\bar{\mathbf{H}}_r\mathbf{Q}_r\bar{\mathbf{H}}_r^H)}\right) = \tag{A22}$$

$$\sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\lambda_{\rho(i)}(\mathbf{Q}_s')\sigma_i^2(\mathbf{H}_1)}{1+\lambda_i(\mathbf{Q}_r)\sigma_{\rho(i)}^2(\bar{\mathbf{H}}_\mathrm{r}')}\right). \tag{A23}$$

Appendix C. Proof of Proposition 2

In this section, we prove that the problem

$$\max_{\gamma_s,\gamma_r} \quad \min_{\sigma_r} \quad \sum_{i=1}^{\min(M,K_r)} \log_2\left(1+\frac{\sigma_{1_i}^2\gamma_{s_{\rho(i)}}}{1+\gamma_{r_i}\sigma_{s_{\rho(i)}}^2}\right) \tag{A24}$$

$$\text{s.t.} \quad \|\boldsymbol{\gamma}_s\|_1 \leq P_s, \tag{A24a}$$

$$\|\boldsymbol{\gamma}_r\|_1 \leq P_r, \tag{A24b}$$

$$\|\boldsymbol{\sigma}_r^2\|_1 \leq T, \tag{A24c}$$

$$\sigma_{1_i}^2\gamma_{s_{\rho(i)}} \geq \sigma_{1_{i+1}}^2\gamma_{s_{\rho(i+1)}}, \ \forall i \leq \min(M,K_r), \tag{A24d}$$

$$\gamma_{r_i}\sigma_{r_{\rho(i)}}^2 \geq \gamma_{r_{i+1}}\sigma_{r_{\rho(i+1)}}^2, \ \forall i \leq \min(K_\mathrm{t},N). \tag{A24e}$$

can be further simplified to

$$\max_{\gamma_s,\gamma_r} \quad \min_{\sigma_r} \quad \sum_{i=1}^{\min(M,K_r)} \log_2\left(1+\frac{\sigma_{1_i}^2\gamma_{s_{\rho(i)}}}{1+\gamma_{r_i}\sigma_{s_{\rho(i)}}^2}\right) \tag{A25}$$

$$\text{s.t.} \quad \|\boldsymbol{\gamma}_s\|_1 = P_s, \tag{A25a}$$

$$\|\boldsymbol{\gamma}_r\|_1 \leq P_r, \tag{A25b}$$

$$\|\boldsymbol{\sigma}_r^2\|_1 = T. \tag{A25c}$$

$$\sigma_{1_i}^2\gamma_{s_{\rho(i)}} \geq \sigma_{1_{i+1}}^2\gamma_{s_{\rho(i+1)}}, \ \forall i \leq \min(M,K_r), \tag{A25d}$$

$$\gamma_{r_i}\sigma_{r_{\rho(i)}}^2 \geq \gamma_{r_{i+1}}\sigma_{r_{\rho(i+1)}}^2, \ \forall i \leq \min(K_\mathrm{t},N). \tag{A25e}$$

The proof is by contradiction. Starting with the minimization, assume that the optimal vector $\boldsymbol{\sigma}^{\star 2}_r$, for which we have $R_{\mathrm{sr}}^{\mathrm{FD}}(\boldsymbol{\sigma}^{\star 2}_r) \leq R_{\mathrm{sr}}^{\mathrm{FD}}(\boldsymbol{\sigma}_r^2)$, does not sum to T and thus, we have $\|\boldsymbol{\sigma}^{\star 2}_r\|_1 < T$. Then there exists $\varepsilon > 0$ for which we have $\|\boldsymbol{\sigma}^{\star 2}_r\|_1 + \varepsilon = T$. Now define

$$\varepsilon_i = \frac{\varepsilon\frac{\sigma_{1_i}^2}{\gamma_{r_i}}}{\sum_j \frac{\sigma_{1_j}^2}{\gamma_{r_j}}}. \tag{A26}$$

Note that we have

$$\sum_i \varepsilon_i = \varepsilon, \quad \varepsilon_i \geq 0. \tag{A27}$$

In addition, as we have $\varepsilon > 0$, there is at least one ε_i which is strictly greater than zero, i.e., $\varepsilon_i > 0$. Now define

$$\sigma'^2_{r_{\rho(i)}} = \sigma^{\star 2}_{r_{\rho(i)}} + \varepsilon_i. \tag{A28}$$

One can check that $\sum_i \sigma'^2_{r_{\rho(i)}} = T$ and $\forall i \leq \min(K_\mathrm{t},N) \Rightarrow \gamma_{r_i}\sigma'^2_{r_{\rho(i)}} \geq \gamma_{r_{i+1}}\sigma'^2_{r_{\rho(i+1)}}$. As a result, $\sigma'^2_{r_{\rho(i)}}$ meets the constraints and could be a feasible solution. Note that as $\gamma_s^\star$ is the optimal source power allocation based on all other parameters, by changing $\boldsymbol{\sigma}^{\star 2}_r$ to $\boldsymbol{\sigma}'^2_r$, $\gamma_s^\star$ might also change. However, we created each $\sigma'^2_{r_{\rho(i)}}$ in a special way to avoid this change. To show this, first notice that we have

$$\gamma^\star_{s_{\rho(i)}} = \left[\lambda - \frac{1+\gamma_{r_i}\sigma^{\star 2}_{r_{\rho(i)}}}{\sigma_{1_i}^2}\right]^+, \tag{A29}$$

where λ is water level and can be found based on power constraints. Substituting the new power allocation for interference, we obtain new power allocation for input power as follows

$$\gamma_{s_{\rho(i)}} = \left[\lambda - \frac{1+\gamma_{r_i}\sigma'^2_{r_{\rho(i)}}}{\sigma^2_{1_i}}\right]^+ = \left[\lambda - \frac{1+\gamma_{r_i}(\sigma^{\star 2}_{r_{\rho(i)}}+\varepsilon_i)}{\sigma^2_{1_i}}\right]^+ \tag{A30}$$

$$= \left[\lambda - \frac{1+\gamma_{r_i}\sigma^{\star 2}_{r_{\rho(i)}}}{\sigma^2_{1_i}} + \frac{\varepsilon}{\sum_{j=1}^N \frac{\sigma^2_{1_j}}{\gamma_{r_j}}}\right]^+ \overset{(a)}{=} \left[\lambda' - \frac{1+\gamma_{r_i}\sigma^{\star 2}_{r_{\rho(i)}}}{\sigma^2_{1_i}}\right]^+ = \gamma^\star_{s_{\rho(i)}}, \tag{A31}$$

where (a) comes from the fact that $\varepsilon/\sum_{j=1}^N \frac{\sigma^2_{1_j}}{\gamma_{r_j}}$ is a constant independent of i. So we can define $\lambda' = \lambda + \varepsilon/\sum_{j=1}^N \frac{\sigma^2_{1_j}}{\gamma_{r_j}}$. This shows, for σ'^2_r, all the optimal variables and parameters remain the same as those of $\sigma^{\star 2}_r$. Now we compare $R^{\text{FD}}_{\text{sr}}$ for both cases. First, notice that we have $\forall i, \varepsilon_i \geq 0$ and among them there is at least one index i', for which we have $\varepsilon_{i'} > 0$. This means $\forall i, \sigma'^2_{r_{\rho(i)}} \geq \sigma^{\star 2}_{r_{\rho(i)}}$ and $\sigma'^2_{r_{\rho(i')}} > \sigma^{\star 2}_{r_{\rho(i')}}$. Now, notice that $f_i(x) = \log_2\left(1 + \frac{\sigma^2_{1_i}\gamma^\star_{s_{\rho(i)}}}{1+\gamma_{r_i}x}\right)$ is a monotonically decreasing function of x. Thus, we have $f_i(\sigma'^2_{r_{\rho(i)}}) \leq f_i(\sigma^{\star 2}_{r_{\rho(i)}})$ and $f_{i'}(\sigma'^2_{r_{\rho(i')}}) < f_{i'}(\sigma^{\star 2}_{r_{\rho(i')}})$. Adding all above inequalities, we obtain

$$\sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\sigma^2_{1_i}\gamma^\star_{s_{\rho(i)}}}{1+\gamma_{r_i}\sigma'^2_{r_{\rho(i)}}}\right) < \sum_{i=1}^{\min(M,K_r)} \log_2\left(1 + \frac{\sigma^2_{1_i}\gamma^\star_{s_{\rho(i)}}}{1+\gamma_{r_i}\sigma^{\star 2}_{r_{\rho(i)}}}\right). \tag{A32}$$

The above equation indicates $R^{\text{FD}}_{\text{sr}}(\boldsymbol{\sigma}^{\star 2}_r) > R^{\text{FD}}_{\text{sr}}(\boldsymbol{\sigma}^2_r)$ which contradicts the first assumption $R^{\text{FD}}_{\text{sr}}(\boldsymbol{\sigma}^{\star 2}_r) \leq R^{\text{FD}}_{\text{sr}}(\boldsymbol{\sigma}^2_r)$. This completes the proof of the minimization part.

For the maximization part, the general idea is the same. Again, the proof is by contradiction. We assume the optimal vector $\boldsymbol{\gamma}^\star_s$, for which we have $R^{\text{FD}}_{\text{sr}}(\boldsymbol{\gamma}^\star_s) \geq R^{\text{FD}}_{\text{sr}}(\boldsymbol{\gamma}_s)$, does not sum to P_s. Therefore, we have $\|\boldsymbol{\gamma}_s\|_1 < P_s$. Then there exists $\varepsilon > 0$ for which we have $\|\boldsymbol{\gamma}_s\|_1 + \varepsilon = P_s$. Now we define

$$\varepsilon_i = \frac{\varepsilon}{\eta}\left(\frac{1+\sigma^{\star 2}_{r_i}\gamma_{r_i}}{\sigma^2_{1_i}} + \gamma^\star_{s_i}\right), \tag{A33}$$

where, $\eta = \sum_i \left(\frac{1+\sigma^{\star 2}_{r_i}\gamma_{r_i}}{\sigma^2_{1_i}} + \gamma^\star_{s_i}\right)$. Now we define the new source power allocation as below

$$\gamma'_{s_{\rho(i)}} = \gamma^\star_{s_{\rho(i)}} + \varepsilon_i. \tag{A34}$$

One can check that $\sum_i \gamma'_{s_{\rho(i)}} = P_s$ and $\sigma^2_{1_i}\gamma'_{s_{\rho(i)}} \geq \sigma^2_{1_{i+1}}\gamma'_{s_{\rho(i+1)}}$. Thus, the new source power allocation is in the feasible set. Now the remaining is to make sure the new allocation does not change the corresponding $\boldsymbol{\sigma}^2_r$. Using Lagrangian multiplier, we have

$$L = \sum_i \log_2\left(1 + \frac{\sigma^2_{1_i}\gamma'_{s_{\rho(i)}}}{1+\gamma_{r_i}\sigma^2_{s_{\rho(i)}}}\right) + \lambda\left(\sum_{i=0}^N \sigma^2_{r_i} - T\right), \tag{A35}$$

$$= \sum_i \log_2\left(1 + \frac{\sigma^2_{1_i}(\gamma^\star_{s_{\rho(i)}}+\varepsilon_i)}{1+\gamma_{r_i}\sigma^2_{s_{\rho(i)}}}\right) + \lambda\left(\sum_{i=0}^N \sigma^2_{r_i} - T\right), \tag{A36}$$

$$= \sum_i \log_2\left((1+\frac{\varepsilon}{\eta})\left(1 + \frac{\sigma^2_{1_i}\gamma^\star_{s_{\rho(i)}}}{1+\gamma_{r_i}\sigma^2_{s_{\rho(i)}}}\right)\right) + \lambda\left(\sum_{i=0}^N \sigma^2_{r_i} - T\right), \tag{A37}$$

$$= \sum_i \log_2(1+\frac{\varepsilon}{\eta}) + \sum_i \log_2\left(1+\frac{\sigma_{1_i}^2 \gamma_{s_{\rho(i)}}^\star}{1+\gamma_{r_i}\sigma_{s_{\rho(i)}}^2}\right) + \lambda\left(\sum_{i=0}^{N}\sigma_{r_i}^2 - T\right). \tag{A38}$$

Now notice that as $\sum_i \log_2(1+\frac{\varepsilon}{\eta})$ is a constant, we have $\frac{\partial \sum_i \log_2(1+\frac{\varepsilon}{\eta})}{\partial \sigma_{r_i}^2} = 0$ and $\frac{\partial \sum_i \log_2(1+\frac{\varepsilon}{\eta})}{\partial \lambda} = 0$. As a result, the optimum interference allocation for γ_r' is the same as that of $\gamma_r^\star$. Similarly to the case of minimization, here we have $\sum_i \varepsilon_i = \varepsilon$. In addition, we have $\varepsilon_i \geq 0$ and there exists at least one i' for which we have $\varepsilon_{i'} > 0$. Finally, as $f_i(x) = \log\left(1+\frac{\sigma_{1_i}^2 x}{1+P\sigma_2^\star{}_{r_{\rho(i)}}}\gamma_{r_i}\right)$ is a monotonically increasing function of x, we conclude $R_{\mathrm{sr}}^{\mathrm{FD}}(\gamma_s^\star) < R_{\mathrm{sr}}^{\mathrm{FD}}(\gamma'_s)$ which contradicts the first assumption of $\gamma_s^\star$ being the optimal source power allocation, and the proof is complete.

Appendix D

First, we show $R_{\mathrm{sr}}^{\mathrm{FD}}$ is a decreasing function of T and an increasing function of P_s. It is sufficient to show $\frac{dR_{\mathrm{sr}}^{\mathrm{FD}}}{dP_s} \geq 0$ and $\frac{dR_{\mathrm{sr}}^{\mathrm{FD}}}{dT} \leq 0$. We have

$$\frac{dR_{\mathrm{sr}}^{\mathrm{FD}}}{dP_{\mathrm{s}}} = \frac{\sum_i \frac{\partial R_{\mathrm{sr}}^{\mathrm{FD}}}{\partial \gamma_{s_{\rho(i)}}} d\gamma_{s_{\rho(i)}}}{\sum_i \frac{\partial P_{\mathrm{s}}}{\partial \gamma_{s_{\rho(i)}}} d\gamma_{s_{\rho(i)}}} = \frac{\sum_i \frac{\sigma_{1_i}^2 d\gamma_{s_{\rho(i)}}}{1+\sigma_{r_{\rho(i)}}^2\gamma_{r_i}+\sigma_{1_i}^2\gamma_{s_{\rho(i)}}}}{\sum_i d\gamma_{s_{\rho(i)}}} \geq \frac{\sum_i \phi_1 d\gamma_{s_{\rho(i)}}}{\sum_i d\gamma_{s_{\rho(i)}}} = \phi_1 > 0, \tag{A39}$$

$$\frac{dR_{\mathrm{sr}}^{\mathrm{FD}}}{dT} = \frac{\sum_i \frac{\partial R_{\mathrm{sr}}^{\mathrm{FD}}}{\partial \sigma_{r_{\rho(i)}}^2} d\sigma_{r_{\rho(i)}}^2}{\sum_i \frac{\partial T}{\partial \sigma_{r_{\rho(i)}}^2} d\sigma_{r_{\rho(i)}}^2} \tag{A40}$$

$$= \frac{\sum_i \frac{-\sigma_{1_i}^2 \gamma_{s_{\rho(i)}}\gamma_{r_i}}{\left(1+\sigma_{r_{\rho(i)}}^2\gamma_{r_i}\right)\left(1+\sigma_{1_i}^2\gamma_{s_{\rho(i)}}+\sigma_{r_i}^2\gamma_{r_i}\right)} d\sigma_{r_{\rho(i)}}^2}{\sum_i d\sigma_{r_{\rho(i)}}^2} \leq \frac{\sum_i -\phi_2 d\sigma_{r_{\rho(i)}}^2}{\sum_i d\sigma_{r_{\rho(i)}}^2} = -\phi_2 \leq 0, \tag{A41}$$

where

$$\phi_1 \doteq \min_i \left\{\frac{\sigma_{1_i}^2}{1+\sigma_{r_{\rho(i)}}^2\gamma_{r_i}+\sigma_{1_i}^2\gamma_{s_{\rho(i)}}}\right\} \tag{A42}$$

and

$$\phi_2 \doteq \min_i \left\{\frac{-\sigma_{1_i}^2\gamma_{s_{\rho(i)}}\gamma_{r_i}}{\left(1+\sigma_{r_{\rho(i)}}^2\gamma_{r_i}\right)\left(1+\sigma_{1_i}^2\gamma_{s_{\rho(i)}}+\sigma_{r_{\rho(i)}}^2\gamma_{r_i}\right)}\right\} \tag{A43}$$

respectively.

Next, we show $g(P_r) = R_{\mathrm{sr}}^{\mathrm{FD}}(P_r) - R_{\mathrm{rd}}^{\mathrm{FD}}(P_r)$ is a monotonically decreasing function of P_r. It is sufficient to show $\frac{dR_{\mathrm{sr}}^{\mathrm{FD}}}{dP_r} \leq 0$ and $\frac{dR_{\mathrm{rd}}^{\mathrm{FD}}}{dP_r} > 0$. We have

$$dR_{\mathrm{rd}}^{\mathrm{FD}} = \sum_i \frac{\partial R_{\mathrm{rd}}^{\mathrm{FD}}}{\partial \gamma_{r_i}} d\gamma_{r_i} = \sum_i \frac{\sigma_{2_i}^2}{1+\sigma_{2_i}^2\gamma_{r_i}} d\gamma_{r_i} \tag{A44}$$

$$dR_{\mathrm{sr}}^{\mathrm{FD}} = \sum_i \frac{\partial R_{\mathrm{sr}}^{\mathrm{FD}}}{\partial \gamma_{r_i}} d\gamma_{r_i} = \sum_i \frac{-\sigma_{1_i}^2\gamma_{s_{\rho(i)}}\sigma_{r_{\rho(i)}}^2}{\left(1+\sigma_{r_{\rho(i)}}^2\gamma_{r_i}\right)\left(1+\sigma_{1_i}^2\gamma_{s_{\rho(i)}}+\sigma_{r_{\rho(i)}}^2\gamma_{r_i}\right)} d\gamma_{r_i} \tag{A45}$$

$$dP_{\mathrm{r}} = \sum_i \frac{\partial P_{\mathrm{r}}}{\partial \gamma_{r_i}} d\gamma_{r_i} = \sum_i d\gamma_{r_i}. \tag{A46}$$

Now we define

$$\psi_1 \doteq \min_i \left\{ \frac{\sigma_{2_i}^2}{1+\sigma_{2_i}^2 \gamma_{r_i}} \right\} \tag{A47}$$

$$\psi_2 \doteq \min_i \left\{ \frac{\sigma_{1_i}^2 \gamma_{s_{\rho(i)}} \sigma_{r_{\rho(i)}}^2}{\left(1+\sigma_{r_{\rho(i)}}^2 \gamma_{r_i}\right)\left(1+\sigma_{1_i}^2 \gamma_{s_{\rho(i)}} + \sigma_{r_{\rho(i)}}^2 \gamma_{r_i}\right)} \right\}. \tag{A48}$$

It is obvious that $\psi_1 > 0$ and $\psi_2 \geq 0$. Now we have

$$\frac{dR_{\text{sr}}^{\text{FD}}}{dP_r} = \frac{\sum_i \frac{-\sigma_{1_i}^2 \gamma_{s_{\rho(i)}} \sigma_{r_{\rho(i)}}^2}{\left(1+\sigma_{r_{\rho(i)}}^2 \gamma_{r_i}\right)\left(1+\sigma_{1_i}^2 \gamma_{s_{\rho(i)}} + \sigma_{r_{\rho(i)}}^2 \gamma_{r_i}\right)} d\gamma_{r_i}}{\sum_i d\gamma_{r_i}} \leq \frac{\sum_i -\psi_2 d\gamma_{r_i}}{\sum_i d\gamma_{r_i}} = -\psi_2 \leq 0, \tag{A49}$$

and

$$\frac{dR_{\text{rd}}^{\text{FD}}}{dP_r} = \frac{\sum_i \frac{\sigma_{2_i}^2}{1+\sigma_{2_i}^2 \gamma_{r_i}} d\gamma_{r_i}}{\sum_i d\gamma_{r_i}} \geq \frac{\sum_i \psi_1 d\gamma_{r_i}}{\sum_i d\gamma_{r_i}} = \psi_1 > 0. \tag{A50}$$

Finally, one can conclude

$$\frac{dg}{dP_{\text{r}}} = \frac{dR_{\text{sr}}^{\text{FD}}}{dP_r} - \frac{dR_{\text{rd}}^{\text{FD}}}{dP_r} \leq -\psi_2 - \psi_1 < 0. \tag{A51}$$

Appendix E

Here we show that if $\gamma_{r_i} \geq \gamma_{r_{i+1}}$ and $\sigma_{1_i}^2 \gamma_{s_{\rho(i)}} \geq \sigma_{1_{i+1}}^2 \gamma_{s_{\rho(i+1)}}$ then $\gamma_{r_i} \sigma_{s_{\rho(i)}}^2 \geq \gamma_{r_{i+1}} \sigma_{r_{\rho(i+1)}}^2$. First we define $f(x,y) = \sqrt{x^2+axy} - x - b,\ x \geq 0,\ y \geq 0$ in which a and b are positive constants. Now we have,

$$\frac{\partial f}{\partial y} = \frac{ax}{2\sqrt{x^2+axy}} \geq 0. \tag{A52}$$

In addition, for $\frac{\partial f}{\partial x}$ we have

$$\frac{\partial f}{\partial x} = \frac{2x+ay}{2\sqrt{x^2+axy}} - 1 \geq 0. \tag{A53}$$

One can check that for positive values $x,\ y$ and a, we always have $\frac{2x+ay}{2\sqrt{x^2+axy}} \geq 1$. As a result, f is an increasing function of both x and y. The rest of the proof is as follows

$$\sigma_{r_i}^2 \gamma_{r_i} = \left[\frac{\sqrt{\left(\sigma_{1_i}^2 \gamma_{s_{\rho(i)}}\right)^2 + \frac{4\sigma_{1_i}^2 \gamma_{s_{\rho(i)}} \gamma_{r_i}}{\lambda}} - \sigma_{1_i}^2 \gamma_{s_{\rho(i)}} - 2(\sigma_t^2 + T_1 P_s)}{2} \right]^+ \tag{A54}$$

$$\overset{(a)}{\geq} \left[\frac{\sqrt{\left(\sigma_{1_i}^2 \gamma_{s_{\rho(i)}}\right)^2 + \frac{4\sigma_{1_i}^2 \gamma_{s_{\rho(i)}} \gamma_{r_{i+1}}}{\lambda}} - \sigma_{1_i}^2 \gamma_{s_{\rho(i)}} - 2(\sigma_t^2 + T_1 P_s)}{2} \right]^+ \tag{A55}$$

$$\overset{(b)}{\geq} \left[\frac{\sqrt{\left(\sigma_{1_{i+1}}^2 \gamma_{s_{\rho(i+1)}}\right)^2 + \frac{4\sigma_{1_{i+1}}^2 \gamma_{s_{\rho(i+1)}} \gamma_{r_{i+1}}}{\lambda}}}{2} \right.$$

$$- \frac{-\sigma_{1_{i+1}}^2 \gamma_{s_{\rho(i+1)}} - 2(\sigma_t^2 + T_1 P_s)}{2} \Bigg]^+ \tag{A56}$$

$$= \sigma_{r_{i+1}}^2 \gamma_{r_{i+1}}, \tag{A57}$$

in which (a) holds because $\gamma_{r_i} \geq \gamma_{r_{i+1}}$ and (b) holds because $\sigma_{1_i}^2 \gamma_{s_{\rho(i)}} \geq \sigma_{1_{i+1}}^2 \gamma_{s_{\rho(i+1)}}$.

References

1. Kariminezhad, A.; Sezgin, A.; Pesavento, M. Power efficiency of improper signaling in MIMO full-duplex relaying for K-user interference networks. In Proceedings of the 2017 IEEE International Conference on Communications (ICC), Paris, France, 21–25 May 2017.
2. Wu, Q.; Zhang, R. Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network. *IEEE Commun. Mag.* **2019**, *58*, 106–112. [CrossRef]
3. Di Renzo, M.; Zappone, A.; Debbah, M.; Alouini, M.S.; Yuen, C.; De Rosny, J.; Tretyakov, S. Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 2450–2525. [CrossRef]
4. Bazrafkan, A.; Poposka, M.; Hadzi-Velkov, Z.; Zlatanov, N. A simple single RF-chain multi-antenna full-duplex relay can outperform an intelligent reflecting surface. *arXiv* **2021**, arXiv:2104.07980.
5. Guo, H.; Liang, Y.C.; Chen, J.; Larsson, E.G. Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 3064–3076. [CrossRef]
6. Wu, Q.; Zhang, R. Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 5394–5409. [CrossRef]
7. Huang, C.; Zappone, A.; Alex Ropoulos, G.C.; Debbah, M.; Yuen, C. Reconfigurable intelligent surfaces for energy efficiency in wireless communication. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 4157–4170. [CrossRef]
8. Zhang, S.; Zhang, R. Capacity characterization for intelligent reflecting surface aided MIMO communication. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 1823–1838. [CrossRef]
9. Obeed, M.; Chaaban, A. Joint beamforming design for multiuser MISO downlink aided by a reconfigurable intelligent surface and a relay. *IEEE Trans. Wirel. Commun.* **2022**. [CrossRef]
10. Omid, Y.; Shahabi, S.M.; Pan, C.; Deng, Y.; Nallanathan, A. Low-complexity robust beamforming design for IRS-aided MISO systems with imperfect channels. *IEEE Commun. Lett.* **2021**, *25*, 1697–1701. [CrossRef]
11. Sejan, M.A.S.; Rahman, M.H.; Shin, B.S.; Oh, J.H.; You, Y.H.; Song, H.K. Machine Learning for Intelligent-Reflecting-Surface-Based Wireless Communication towards 6G: A Review. *Sensors* **2022**, *22*, 5405. [CrossRef]
12. Cao, X.; Yang, B.; Huang, C.; Yuen, C.; Di Renzo, M.; Niyato, D.; Han, Z. Reconfigurable intelligent surface-assisted aerial-terrestrial communications via multi-task learning. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 3035–3050. [CrossRef]
13. Huang, C.; Chen, G.; Wong, K.K. Multi-agent reinforcement learning-based buffer-aided relay selection in IRS-assisted secure cooperative networks. *IEEE Trans. Inf. Forensics Secur.* **2021**, *16*, 4101–4112. [CrossRef]
14. Zhou, G.; Pan, C.; Ren, H.; Wang, K.; Di Renzo, M.; Nallanathan, A. Robust beamforming design for intelligent reflecting surface aided MISO communication systems. *IEEE Wirel. Commun. Lett.* **2020**, *9*, 1658–1662. [CrossRef]
15. Zhou, G.; Pan, C.; Ren, H.; Wang, K.; Nallanathan, A. A framework of robust transmission design for IRS-aided MISO communications with imperfect cascaded channels. *IEEE Trans. Signal Process.* **2020**, *68*, 5092–5106. [CrossRef]
16. Zhang, J.; Zhang, Y.; Zhong, C.; Zhang, Z. Robust design for intelligent reflecting surfaces assisted MISO systems. *IEEE Commun. Lett.* **2020**, *24*, 2353–2357. [CrossRef]
17. Bliss, D.W.; Parker, P.A.; Margetts, A.R. Simultaneous transmission and reception for improved wireless network performance. In Proceedings of the 2007 IEEE/SP 14th Workshop on Statistical Signal Processing, Madison, WI, USA, 26–29 August 2007.
18. Everett, E. Achaleshwar Sahai, and Ashutosh Sabharwal. Passive self-interference suppression for full-duplex infrastructure nodes. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 680–694. [CrossRef]
19. Ahmed, E.; Eltawil, A.M. All-digital self-interference cancellation technique for full-duplex systems. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 3519–3532. [CrossRef]
20. Vogt, H.; Enzner, G.; Sezgin, A. State-space adaptive nonlinear self-interference cancellation for full-duplex communication. *IEEE Trans. Signal Process.* **2019**, *67*, 2810–2825. [CrossRef]
21. Lee, H.; Kim, D.; Kim, S.; Kim, J.; Lee, C.; Hong, D. Rotated precoder based self-interference cancellation in full-duplex communication. In Proceedings of the 2014 International Conference on Electronics, Information and Communications (ICEIC), Kota Kinabalu, Malaysia, 15–18 January 2014.
22. Soriano-Irigaray, F.J.; Fernez-Prat, J.S.; Lopez-Martinez, F.J.; Martos-Naya, E.; Cobos-Morales, O.; Entrambasaguas, J.T. Adaptive self-interference cancellation for full duplex radio: Analytical model and experimental validation. *IEEE Access* **2018**, *6*, 65018–65026. [CrossRef]

23. Irio, L.; Oliveira, R.; Oliveira, L. Characterization of the residual self-interference power in full-duplex wireless systems. In Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, 27–30 May 2018.
24. Irio, L.; Oliveira, R. Distribution of the residual self-interference power in in-band full-duplex wireless systems. *IEEE Access* **2019**, *7*, 57516–57526. [CrossRef]
25. Koohian, A.; Mehrpouyan, H.; Nasir, A.A.; Durrani, S.; Blostein, S.D. Residual self-interference cancellation and data detection in full-duplex communication systems. In Proceedings of the 2017 IEEE International Conference on Communications (ICC), Paris, France, 21–25 May 2017.
26. Chae, S.H.; Lee, K. Degrees of freedom of full-duplex cellular networks: Effect of self-interference. *IEEE Trans. Commun.* **2017**, *65*, 507–4518. [CrossRef]
27. Kariminezhad, A.; Sezgin, A. Heterogeneous multi-tier networks: Improper signaling for joint rate-energy optimization. *IEEE Trans. Wirel. Commun.* **2018**, *18*, 680–694. [CrossRef]
28. Kariminezhad, A.; Gherekhloo, S.; Sezgin, A. Full-duplex vs. half-duplex: Delivery-time optimization in cellular downlink. In Proceedings of the European Wireless 2017, 23th European Wireless Conference, Dresden, Germany, 17–19 May 2017; VDE: Merianstraße, Germany, 2017.
29. Herath, S.P.; Le-Ngoc, T. Sum-rate performance and impact of self-interference cancellation on full-duplex wireless systems. In Proceedings of the 2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), London, UK, 8–11 September 2013.
30. Zlatanov, N.; Sippel, E.; Jamali, V.; Schober, R. Capacity of the Gaussian two-hop full-duplex relay channel with residual self-interference. *IEEE Trans. Commun.* **2017**, *65*, 1005–1021. [CrossRef]
31. Riihonen, T.; Werner, S.; Wichman, R. Hybrid full-duplex/half-duplex relaying with transmit power adaptation. *IEEE Trans. Wirel. Commun.* **2011**, *10*, 3074–3085. [CrossRef]
32. Taghizadeh, O.; Mathar, R. Robust multi-user decode-and-forward relaying with full-duplex operation. In Proceedings of the 11th International Symposium on Wireless Communications Systems (ISWCS), Barcelona, Spain, 26–29 August 2014.
33. Cirik, A.C.; Biswas, S.; Vuppala, S.; Ratnarajah, T. Robust transceiver design for full duplex multiuser MIMO systems. *IEEE Wirel. Commun. Lett.* **2016**, *5*, 260–263. [CrossRef]
34. Esmaeili, H.; Kariminezhad, A.; Sezgin, A. Robust transceiver design for full-duplex decode-and-forward relay-assisted MIMO systems. In Proceedings of the 54th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 1–5 November 2020.
35. Han, C.; Akyildiz, I.F. Distance-aware bandwidth-adaptive resource allocation for wireless systems in the terahertz band. *IEEE Trans. Terahertz Sci. Technol.* **2016**, *6*, 541–553. [CrossRef]
36. Tang, W.; Chen, M.Z.; Chen, X.; Dai, J.Y.; Han, Y.; Di Renzo, M.; Zeng, Y.; Jin, S.; Cheng, Q.; Cui, T.J. Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement. *IEEE Trans. Wirel. Commun.* **2020**, *20*, 421–439. [CrossRef]
37. Wang, B.; Zhang, J.; Host-Madsen, A. On the capacity of MIMO relay channels. *IEEE Trans. Inf. Theory* **2005**, *51*, 29–43. [CrossRef]
38. Wang, J.; Bengtsson, M.; Ottersten, B.; Palomar, D.P. Robust MIMO precoding for several classes of channel uncertainty. *IEEE Trans. Signal Process.* **2013**, *61*, 3056–3070. [CrossRef]
39. Shen, H.; Wang, J.; Levy, B.C.; Zhao, C. Robust optimization for amplify-and-forward MIMO relaying from a worst-case perspective. *IEEE Trans. Signal Process.* **2013**, *61*, 5458–5471. [CrossRef]
40. Fiedler, M. Bounds for the determinant of the sum of hermitian matrices. *Proc. Am. Math. Soc.* **1971**, *30*, 27–31. [CrossRef]
41. Marshall, A.W.; Olkin, I.; Arnold, B.C. *Inequalities: Theory of Majorization and Its Applications*; Academic Press: New York, NY, USA, 1979; Volume 143.