國 立 交 通 大 學

電信工程研究所

碩 士 論 文

合作式放大傳遞多輸入多輸出中繼系統之
強健性 Tomlinson-Harashima 來源端與線性中繼端
前置編碼設計

Robust Tomlinson-Harashima Source and Linear Relay Precoders Design in Amplify-and-Forward MIMO Relay Systems

研 究 生：張閔堯

指導教授：吳文榕 博士

中 華 民 國 99 年 7 月

合作式放大傳遞多輸入多輸出中繼系統之強健性 Tomlinson-Harashima 來源端與線性中繼端前置編碼設計

Robust Tomlinson-Harashima Source and Linear Relay Precoders Design in Amplify-and-Forward MIMO Relay Systems

研 究 生：張閔堯　　Student：Min-Yao Chang

指導教授：吳文榕 博士　　Advisor：Dr. Wen-Rong Wu

國 立 交 通 大 學

電信工程研究所

碩 士 論 文

A Thesis

Submitted to Institute of Computer and Information Science

College of Electrical Engineering and Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer and Information Science

July 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年七月

合作式放大傳遞多輸入多輸出中繼系統
之強健性 Tomlinson-Harashima 來源端與線性中繼端前置編碼設計

研究生：張閔堯　　　　　　　　　　　　　　指導教授：吳文榕 博士

國立交通大學

電信工程研究所碩士論文

摘要

在現存的合作式放大傳遞多輸入多輸出中繼系統傳收機的設計，通常假設此系統可得到完美的通道狀態資訊(perfect CSI)。但在實際系統應用上，可能無法獲得完美的通道狀態資訊。基於非完美通道狀態資訊的考量下，強健性的傳收機設計在實際應用上是需要被考慮的。在本論文中，我們提出一種強健性的傳收機設計，此傳收機中的來源端使用 Tomlinson-Harashima 前置編碼 (THP)、線性中繼端前置編碼與最小均方錯誤(minimum mean-squared error)接收機。當兩個前置編碼的組合及非完美通道狀態資訊被考慮進來時，傳收機的設計變得相當困難。為了克服這個設計上的困難，我們提出一種前置編碼結構與設計方法，使得原本的傳收機設計可轉換為凹曲線最佳化問題，由此導出解析解。我們在 TH 前置編碼後串接一個單位前置編碼。這個額外的前置編碼的功能不僅可以簡化最佳化的問題而且可改善整個系統的效能表現。由於最佳化的問題是由多個前置編碼組成，我們使用最初分解(primal decomposition)將原本的最佳化問題分解成次要問題(subproblem)與主要問題(master problem)。依序解決次要問題與主要問題，原本由兩個前置編碼構成的問題，可簡化成設計單一中繼端前置編碼的問題。但是要解決主要問題仍然相當困難，因此我們對這個主要問題提出一個最低界線，並經由一些操作將此問題轉換成凹曲線最佳化的形式。透過 Karush-Kuhn-Tucker(KKT)條件可以推導出解析解。模擬的結果顯示在完美/非完美的通道狀態資訊環境下，所提出的強健性傳收機設計在效能上的表現比現存的線性傳收機設計好。

Robust Tomlinson-Harashima source and linear relay precoders design in amplify-and-forward MIMO relay systems

Student:Min-Yao Chang Advisor: Dr. Wen-Rong Wu

Institute of Communication Engineering
National Chiao Tung University

Abstract

The existing transceiver design in amplify-and-forward (AF) multiple-input multiple-output (MIMO) relay systems often assume the perfect channel state information (CSI) is available. In practice, the perfect CSI is not attainable and the robust design with considering imperfect CSI is applied. In this paper, we propose a robust transceiver design for the system with Tomlinson-Harashima (TH) precoder, a linear relay precoder and a minimum-mean-square (MMSE) receiver. Since two precoders and the imperfect CSI are involved, the robust design problem is difficult. To overcome the difficulty, we additionally cascade a unitary precoder after the TH precoder. The unitary precoder can not only simply the optimization problem but improve the performance of the system. With the precoders, we use the primal decomposition to divide the original optimization problem into a subproblem and a master problem. The subproblem can be solved and the two-precoder problem can be reduced to the problem composed of single relay precoder. However, the master problem is still difficult to solve. We then proposed a lower bound for the cost function and transfer the master problem to a convex optimization problem. A closed-form solution can then be obtained by Karush-Kuhn-Tucker (KKT) conditions. Simulations show the that the proposed transceiver design have better performance than existing linear transceiver with either perfect/imperfect CSI.

Acknowledgement

在完成本論文的過程中，感謝所有幫助過我的人。

感謝我的指導教授 - 吳文榕教授，感謝文榕老師在學業上的指導及提供專業的研究建議，使我明白學術研究的基礎需要正確的技巧與態度；在遇到未來規劃上的困難時，文榕老師也能適時分享經驗及想法，這些都令我受益匪淺。感謝我的父母親與家人，他們在精神上對我的支持給我很大的勇氣與毅力完成學業。感謝 720 實驗室的所有同仁，兩年的研究所期間從他們的身上學到很多東西，也共同經歷了多事情。感謝凡碩、鈞陶、兆元、弘道、俊芳、其翰、南鈞、光敏等學長們，每當我遇到難題，請教他們總是可以獲得許多學習的機會，也時常與他們討論社會上或其他領域上的議題，讓我從他們的人生閱歷中得到一些知識。特別感謝凡碩學長，由於他在學術及生活上的經驗分享，我學到做人處事應有積極的態度與熱誠，並瞭解人生有許多難題需要耐心面對，也有許多樂趣等待發掘。感謝勝富、沁寧、瑞慶、綺瑩等同學，感謝你們的陪伴並一起度過兩年的研究生活。感謝所有這兩年中認識的老師、朋友們，與你們的互動讓我在各方面成長很多。

感謝所有幫助過我的人，此為誌謝。

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In recent years, many works have been devoted to the study of cooperative communication due to its great potential to improve coverage, capacity and reliability of wireless link [1]-[3]. Due to shadowing, multipath fading, path losses observed in wireless channels, the link between a signal source and a destination may not be always reliable. In cooperative communication systems, relays are placed at some strong shadowing environments such that signals can be transmitted to the destination by a direct link and relay links. With the additional relay links, it provides the enhancement of diversity gain or capacity gain [1]-[18]. There are some relay strategies such as amplified-and-forward (AF), decode-and-forward (DF), and compress-and-forward(CF) [1],[2]. In AF, the relays receive signal from the source and retransmit it to the destination with signal amplification only. In DF, the relays decode the received signal, re-encode information bits, and then transmission the resultant signal to the destination. CF, being a compromise between AF and DF, estimates information bits, compress their information, and then transmit the modulated signal to the destination. In this thesis, we only consider the AF-based system since the AF strategy requires less implement complexity and smaller processing delay. Research in this subject has attracted a lot of attention.

MIMO systems have been widely studied in the literature since it can enhance spatial diversity or multiplexing gain in the rich scattering environments. It is known that a precoder can be used in a MIMO system to further enhance the performance. Linear or non-linear precoder designs in the point-to-point MIMO system have been extensively studied [7]-[9]. In cooperative systems, multiple antennas can be equipped at the source, the relays, and the destination, resulting MIMO relay systems. Similar to conventional MIMO systems, the precoding operation can be conducted in a MIMO relay system [12]-[19]. In this thesis, we only consider the transceiver design in AF MIMO systems.

For the AF MIMO relay system, the capacity bound of the systems have been derived in [12]. Apart from the capacity, the link quality is another criterion has been studied. In [15], a dual-hop single relay precoder in an AF MIMO system was designed for a minimum mean-square error (MMSE) receiver without considering direct link (source to destination). It has been shown in [15] that the joint design has a better performance than separate design scheme. In [16], a joint transceiver design for the multi-relay case has been discussed. In [18], a transceiver design has been proposed for a three-node AF MIMO system using MMSE criterion. This scheme takes both the direct and relay links into consideration, and uses a linear precoder at the source and another linear precoder at the relay. Recently, the joint source and relay precoders design for multiple transmission streams were studied in [17]-[20], where the source and relay precoders are jointly designed with the direct and relay links via MMSE [17], QR successive-interference-cancellation (SIC) [19], and MMSE-SIC [20], respectively.

As well known, the nonlinear transceivers are superior to linear transceivers. To obtain better performance for the precoded system, we then focus on a nonlinear source precoder design in this thesis. There exist several nonlinear MIMO transceivers, e.g. the system with a Tomlinson-Harashima precoder (THP), and that with a decision-feed-back equalizer (DFE). As well known, the DFE at the destination estimate and cancel the interferences and it may cause error propagation in low SNR environments. We consider a precoded AF MIMO relay system in which the nonlinear THP is used at the source, a linear precoder at the relay and a MMSE receiver at the destination [22]. The THP can pre-cancel the known interference at the source and will not induce error propagation. It is widely used in point-to-point or multiuser MIMO systems [21]. In the first part of the thesis, we consider a transceiver design in AF MIMO relay system. The design uses the MMSE criterion and take both the direct and relay links into consideration. Since the THP is involved, the cost function becomes a highly nonlinear and complicated function of the source precoder and relay precoder. The optimization problem becomes very difficult. To solve the problem, we first cascade a unitary precoder with the THP. The specially designed unitary precoder not can only simplify the optimization problem but also improve the MMSE performance. Then, we use the primary decomposition method, decomposing the problem into a master problem and a subproblem. With our formulation, the subproblem problem, designing the unitary precoder and feed-back matrix in the THP, can be degenerated to the system in [24], and the solution is readily obtained. In the master problem, the cost function becomes a function of the relay precoder only. With some precoder structure, we can translate the optimization problem from matrix-valued into a scalar-valued optimization problem, and use Karush-Kuhn-Tucker (KKT) conditions to obtain a closed-form solution for the source and relay precoders.

Most transceiver design in AF MIMO system assume that it knows perfect channel state information (CSI) of each link at each node [14]-[20]. In practice, the perfect CSI is not attainable due to channel estimation or quantization errors. For conventional MIMO systems, some works study the sensitivity of the MIMO precoder with respect to channel uncertainties [29], [30]. In [31]-[35], a robust design for the THP precoded point-to-point MIMO system has been studied. In [36], the design is extended to an AF MIMO relay system. In the design, the direct link is not taken into consideration.

In the second part of this thesis, we study a robust AF MIMO transceiver design with the THP. The optimization problem is similar to that of perfect CSI. The only difference is to consider the estimation errors as extra noise sources. Still, we use the primal decomposition method to decompose the optimization problem into subproblem and master problem. To ease the optimization, we then propose a method that can translate the master problem to a standard scalar-valued concave optimization. The key idea is to apply some approximations for the cost function such that the optimization in the master problem can be solved. We then propose a relay precoding structure in the optimization. Though the structure is suboptimal method, however, it can translate the master optimization problem to a standard scalar-valued concave optimization problem. Finally, similar to the perfect CSI case, we can obtained the close-form solution of the relay and source precoder by using KKT conditions.

The organization of the thesis is described as follows. In Chapter 2, we describe the proposed THP precoded AF MIMO relay system. In Chapter 3, we take the channel estimation error into consideration and propose a robust transceiver design. In Chapter 4, we present the simulation results and related discussions. Finally, we draw conclusions in Chapter 5.

Chapter 2

Joint MMSE Transceiver Desgin with Tomlinson-Harashima Source and Linear Relay Precoders

In this chapter, we consider the MMSE transceiver design with a nonlinear Tomlinson-Harashima precoder (THP) at the source, a linear precoder at the relay in AF MIMO relay systems. Here, we assume that perfect CSIs of all channels are known at the destination. In Section 2.1, we first give the system model, while in Section 2.2 we formulate the design problem under the MMSE criterion. It is found that the MSE is a complicated function of the source and relay precoders, and the optimization problem is non-convex. Thus, the problem is difficult to solve. In Section 2.3, we propose a method translating the two-precoder design problem into a single-relay problem. By using this method, the optimization problem can be formulated as a convex optimization problem, and the close-form solution can then be obtained.

2.1 System model

We consider a typical three-node AF MIMO relay system with a THP. The block diagram of the system is shown in Figure 2.1. The system includes a THP precoder cascaded with a unitary precoder $\mathbf{F}_S$ at the source, a linear precoder $\mathbf{F}_R$ at the relay, and a MMSE receiver $\mathbf{G}$ at the destination. Here, we define the number of antenna at the source, the relay and the destination as *N*, *R* and *M*, respectively. The MIMO channels are assumed to be flat fading. In this cooperative system, we use a half-duplex relay protocol, which means it require a two-phase transmission for a data packet. Let's start with the THP. The THP conducts a interference pre-cancelling operation characterized by a backward, strict low-triangular matrix $\mathbf{B}$ and a modulo operation $\mathrm{MOD}_m(\bullet)$. Let the input signal vector be $\mathbf{s} \in \mathbb{C}^{N\times 1}$; each element of $\mathbf{s} = [s_1, \cdots, s_N]^T$ is a symbol mapped to a square *m*-QAM constellation where. Each QAM symbol is drawn from the set $A = \left\{ s_I + js_Q \middle| s_I, s_Q \in \left\{ \pm 1, \ldots, \pm\sqrt{m} \right\} \right\}$. The feed-back operation conducted in THP may increase the transmit power, and it can be avoided by a modulo operation [21]. The modulo operation applied over both the real and image parts of the input $\mathbf{x}$ is expressed as:

$$\mathrm{MOD}_m(x) = x - 2\sqrt{m} \cdot \left\lfloor \frac{x+\sqrt{m}}{2\sqrt{m}} \right\rfloor \tag{2.1}$$

With $\mathbf{B}$ and the modulo operation, the elements of $\mathbf{x}$ can be expressed as:

$$\mathbf{x}_k = \mathbf{s}_k - \sum_{l=1}^{k-1} \mathbf{B}(k,l)\mathbf{x}_l + \mathbf{e}_k \tag{2.2}$$

where $\mathbf{x}_k$ is the kth element of vector $\mathbf{x}$, $\mathbf{x} \in \mathbb{C}^{N\times 1}$, $\mathbf{B}(k,l)$ is the (k,l) element of matrix $\mathbf{B}$, and $\mathbf{e} = [\mathbf{e}_1, \cdots, \mathbf{e}_N]^T$ denotes the errors caused by the modulo operation. Then, (2.2) can be rewritten with the following matrix form as:

$$\mathbf{x} = \mathbf{C}^{-1}\mathbf{v} \tag{2.3}$$

where $\mathbf{C} = \mathbf{B} + \mathbf{I}_N$ is a lower triangular matrix and $\mathbf{v} = \mathbf{s} + \mathbf{e}$ [21].

The transmission in the cooperative system has two-phase [1]. In the first phase, the THP precoded signal $\mathbf{x}$ is passed through the cascaded unitary precoder $\mathbf{F}_S$, and subsequently send to the relay and the destination simultaneously. As we will show, an appropriate design of the additional unitary precoder will improve the performance of the MIMO relay system.

In the second phase, the received signal at the relay is multiplied by the relay precoder $\mathbf{F}_R$ and the resultant signal is then transmitted to the destination. Therefore, the signal received at the destination after the two consecutive phases can be expressed as a vector form as [17]-[20], [22]:

$$\mathbf{y}_D := \underbrace{\begin{bmatrix} \mathbf{H}_{SD} \\ \mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR} \end{bmatrix}}_{:=\mathbf{H}} \mathbf{F}_S\mathbf{x} + \underbrace{\begin{bmatrix} \mathbf{n}_{D,1} \\ \mathbf{H}_{RD}\mathbf{F}_R\mathbf{n}_R + \mathbf{n}_{D,2} \end{bmatrix}}_{:=\mathbf{w}} \tag{2.4}$$

where $\mathbf{H}$ and $\mathbf{w}$ denote the equivalent channel matrix and the equivalent noise vector, respectively. In (2.4), $\mathbf{x} \in \mathbb{C}^{N\times 1}$ is the THP precoded signal vector (2.3); $\mathbf{y}_D \in \mathbb{C}^{2M\times 1}$ is the received signal vector at the destination; $\mathbf{H}_{SR} \in \mathbb{C}^{R\times N}$, $\mathbf{H}_{SD} \in \mathbb{C}^{M\times N}$ and $\mathbf{H}_{RD} \in \mathbb{C}^{M\times R}$ are the channel matrices of the source-to-relay, the source-to-destination, and the relay-to-destination channels, respectively. Note that these channel matrices are all assumed to be flat fading channel; $\mathbf{n}_{D,1} \in \mathbb{C}^{M\times 1}$, $\mathbf{n}_R \in \mathbb{C}^{R\times 1}$, and $\mathbf{n}_{D,2} \in \mathbb{C}^{M\times 1}$ are the received noise vectors at the destination, at the relay in the first-phase, and at the destination in the second-phase. Here, we assume that $N \le M$ such that sufficient degree of freedom for signal transmission can be assured.

Note that if $\mathbf{v}$ can be estimated at the destination, $\mathbf{s}$ can be reconstructed by the modulo operation. We define the mean-square-error of the estimation as:

$$J = E\left\{\left\|\mathbf{G}\mathbf{y}_D - \mathbf{v}\right\|^2\right\}. \tag{2.5}$$

By minimizing the MSE, we can derive the optimum $\mathbf{G}$. The signal elements of $\mathbf{s}$ is assumed to be independent each other and the variance is σ_s^2. It has been shown in [21] that each element of $\mathbf{x}$ is approximately i.i.d and distributed in the region $A=\left\{s_I + js_Q \middle| s_I, s_Q \in \left\{\pm 1, \ldots, \pm\sqrt{m}\right\}\right\}$ uniformly. The approximation error becomes small as the number of signal levels is large. Thus, it is valid only when m is sufficiently large. With the approximation, we have $E\left[\mathbf{x}\mathbf{x}^H\right] = \sigma_s^2 \mathbf{I}_N$, $E\left[\mathbf{v}\mathbf{v}^H\right] = \sigma_s^2 \mathbf{C}\mathbf{C}^H$. The modulo operation in the THP may cause a transmit power penalty, called precoding loss. According [21], precoding loss for a two-dimentional m-ary square constellations can be calculated as:

$$\gamma_p^2 = \frac{E\left[\left|x_k\right|^2\right]}{E\left[\left|s_k\right|^2\right]} = \frac{m}{m-1} \tag{2.6}$$

where x_k, s_k indicates the k-th element of signal vector $\mathbf{x}, \mathbf{s}$, respectively. We show the precoding loss for various m in Table 2.1. As we can see, the precoding loss is negligible and vanishes completely as m goes to infinity. It is recommended that at least $m \geq 16$ should be used.

$m =$	4	8	16	32	64
$10\log_{10}\left(\gamma_p^2\right)$ [dB]	1.25	0.58	0.28	0.14	0.07

Table 2.1: Precoding loss γ_p^2 (in dB) of Tomlinson-Harashima Precoding [21].

Taking the derivative the MSE with respect to $\mathbf{G}$ and setting the result to zero, we can obtain the optimum $\mathbf{G}$ as [24]:

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{G}^H} &= \frac{\partial}{\partial \mathbf{G}^H} E\left[\mathrm{Tr}\left\{\left(\mathbf{G}\mathbf{y}_D - \mathbf{v}\right)\left(\mathbf{G}\mathbf{y}_D - \mathbf{v}\right)^H\right\}\right] \\ &= \mathbf{G} \cdot E\left[\mathbf{y}_D \mathbf{y}_D^H\right] - E\left[\mathbf{v}\mathbf{y}_D^H\right] = 0 \end{aligned} \tag{2.7}$$

Then, the optimum $\mathbf{G}$, denoted by $\mathbf{G}_{opt}$, is given by

$$\mathbf{G}_{opt} = \sigma_s^2 \mathbf{C}\mathbf{F}_S^H \mathbf{H}^H \left(\sigma_s^2 \mathbf{H}\mathbf{F}_S \mathbf{F}_S^H \mathbf{H}^H + \mathbf{R}_w\right)^{-1} \tag{2.8}$$

where $\mathbf{R}_w = E\left[\mathbf{w}\mathbf{w}^H\right]$ is the covariance matrix of the equivalent noise. Note that the equivalent noise is colored. Denote the variance of the noise at the relay as $\sigma_{n,r}^2$, and that at the destination as $\sigma_{n,d}^2$. Substituting (2.8) into (2.5), we obtain the minimum MSE. To simplify the expression of the MSE, we consider the following leamma:

Lemma 2.1: Matrix inversion lemma [42]

$$\mathbf{A}^H\left(\mathbf{A}\mathbf{A}^H+\mathbf{I}\right)^{-1}\mathbf{A}=\mathbf{I}-\left(\mathbf{A}^H\mathbf{A}+\mathbf{I}\right)^{-1} \tag{2.9}$$

where $\mathbf{A},\mathbf{I}$ denote matrices with appropriate size. Using the lemma, we can rewrite the error matrix in (2.5) as

$$\begin{aligned}\mathbf{E} &= \mathbf{C}\left(\sigma_s^{-2}\mathbf{I}_N+\mathbf{F}_S^H\mathbf{H}^H\mathbf{R}_w^{-1}\mathbf{H}\mathbf{F}_S\right)^{-1}\mathbf{C}^H\\ &= \mathbf{C}\left(\sigma_s^{-2}\mathbf{I}_N+\mathbf{F}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{F}_S\right)^{-1}\mathbf{C}^H\end{aligned} \tag{2.10}$$

and (2.5) becomes

$$J_{\min} = \mathrm{Tr}\{\mathbf{E}\} \tag{2.11}$$

where

$$\begin{aligned}\tilde{\mathbf{H}} &= \mathbf{R}_w^{-1/2}\mathbf{H}\\ &=\begin{bmatrix}\sigma_{n,d}^{-1}\mathbf{H}_{SD}\\ \left(\sigma_{n,r}^2\mathbf{H}_{RD}\mathbf{F}_R\mathbf{F}_R^H\mathbf{H}_{RD}^H+\sigma_{n,d}^2\mathbf{I}_M\right)^{-1/2}\mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}\end{bmatrix}\end{aligned} \tag{2.12}$$

is defined as the equivalent channel matrix after noise whitening. Note that the MSE is contributed by both the relay link and direct links. If we ignore the direct link, the MSE will be reduced to that in [24]. From (2.10) and (2.11), we see that the achievable minimum MSE is a complicated function of $\mathbf{C}$, $\mathbf{F}_S$ and $\mathbf{F}_R$. In the next section, we formulate the precoders design problem using the signal model derived in this section.

2.2 Problem formulation

For the MIMO relay system, two precoders are involved. Using (2.5), (2.8)-(2.12), we can formulate the precoders design problem as:

$$
\begin{aligned}
&\min_{\mathbf{C},\mathbf{F}_S,\mathbf{F}_R} \operatorname{Tr}\left\{\underbrace{\mathbf{C}\left(\sigma_s^{-2}\mathbf{I}_N + \mathbf{F}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{F}_S\right)^{-1}\mathbf{C}^H}_{:=\mathbf{E}(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R)}\right\} \\
&\text{s.t.} \\
&\mathbf{F}_S = \alpha\mathbf{U}_S, C_1, C_2
\end{aligned}
\tag{2.13}
$$

where

$$
\begin{aligned}
&\tilde{\mathbf{H}}^H\tilde{\mathbf{H}} = \sigma_{n,d}^{-2}\mathbf{H}_{SD}^H\mathbf{H}_{SD} + \\
&\mathbf{H}_{SR}^H\mathbf{F}_R^H\mathbf{H}_{RD}^H\left(\sigma_{n,r}^2\mathbf{H}_{RD}\mathbf{F}_R\mathbf{F}_R^H\mathbf{H}_{RD}^H + \sigma_{n,d}^2\mathbf{I}_M\right)^{-1}\mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}
\end{aligned}
\tag{2.14}
$$

and C_1, C_2 denote the power constraints at the source and the relay, respectively:

$$
\begin{aligned}
&C_1 : \operatorname{Tr}\left\{E\left[\mathbf{F}_S\mathbf{x}\mathbf{x}^H\mathbf{F}_S^H\right]\right\} = \sigma_s^2\operatorname{Tr}\left\{\mathbf{F}_S\mathbf{F}_S^H\right\} \le P_{S,T}, \\
&C_2 : \operatorname{Tr}\left\{\mathbf{F}_R\left(\sigma_{n,r}^2\mathbf{I}_R + \sigma_s^2\mathbf{H}_{SR}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}_{SR}^H\right)\mathbf{F}_R^H\right\} \le P_{R,T}.
\end{aligned}
\tag{2.15}
$$

Here we let $\mathbf{F}_S = \alpha\mathbf{U}_S$ in which α is a scalar and $\mathbf{U}_S$ is a unitary matrix. In next section, we show that the unitary structure can facilitate the derivation of the optimization problem and improve the performance of the MIMO rely system. From (2.13), it is apparent that both the cost function and the constraints are complicated function of $\mathbf{F}_S$ and $\mathbf{F}_R$. Yet, the problem is non-convex. Solving such a problem is a very difficult problem, if not impossible. In the next section, we propose a method to overcome the problems.

2.3 Joint source and relay precoders design

Since a direct solution for optimum $\mathbf{F}_S$ and $\mathbf{F}_R$ in (2.13) is difficult, we use the primal decomposition method [42] such that the problem can be translated into a subproblem and a master problem and $\mathbf{F}_S$ and $\mathbf{F}_R$ can be solved separately. In the subproblem, the relay precoder is assumed to be known and the source precoder is solved as a function of the relay precoder. Then, in the master problem, the relay precoder is solved. The problem now can be re-formulated as

$$\begin{aligned}
&\min_{\mathbf{C},\mathbf{F}_S,\mathbf{F}_R} \mathrm{Tr}\{\mathbf{E}\} = \min_{\mathbf{F}_R}\min_{\mathbf{C},\mathbf{F}_S} \mathrm{Tr}\left\{\mathbf{E}(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R)\right\} \\
&s.t. \\
&\mathbf{E} = \mathbf{C}\left(\sigma_s^{-2}\mathbf{I}_N + \mathbf{F}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{F}_S\right)^{-1}\mathbf{C}^H, \\
&\mathbf{F}_S = \alpha\mathbf{U}_S\ , \\
&C_1 : \sigma_s^2\mathrm{Tr}\left\{\mathbf{F}_S\mathbf{F}_S^H\right\} \le P_{S,T}, \\
&C_2 : \mathrm{Tr}\left\{\mathbf{F}_R\left(\sigma_{n,r}^2\mathbf{I}_R + \alpha^2\sigma_s^2\mathbf{H}_{SR}\mathbf{H}_{SR}^H\right)\mathbf{F}_R^H\right\} \le P_{R,T}.
\end{aligned} \tag{2.16}$$

In the subproblem, as mentioned, the optimum $\mathbf{C}$ and $\mathbf{F}_S$ are derived as a function of $\mathbf{F}_R$ by assuming $\mathbf{F}_R$ is given. Then the joint precoder design problem is reduced to the determination of $\mathbf{F}_R$ which is the master problem.

The unitary precoder $\mathbf{F}_S$ is included for two reasons: (i) It can simply the solutoin of the relay precoder. (ii) By a proper design of $\mathbf{U}_S$, the minimum MSE can have an amenable form, leading to a tractable optimization problem. Since $\mathbf{F}_S = \alpha\mathbf{U}_S$, the subproblem becomes the optimization of α, $\mathbf{U}_S$ and $\mathbf{C}$, given as

$$\begin{aligned}
&\min_{\mathbf{C}(\mathbf{F}_R),\alpha\mathbf{U}_S(\mathbf{F}_R)} \mathrm{Tr}\left(\mathbf{E}(\mathbf{C},\alpha\mathbf{U}_S,\mathbf{F}_R)\right) \\
&\mathrm{s.t.} \\
&\alpha\mathbf{U}_S, C_1', C_2 \\
&\mathbf{E} = \mathbf{C}\left(\sigma_s^{-2}\mathbf{I}_N + \alpha^2\mathbf{U}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{U}_S\right)^{-1}\mathbf{C}^H
\end{aligned} \tag{2.17}$$

where $C_1' : N\alpha^2\sigma_S^2 \le P_{S,T}$ is obtained from C_1 by setting $\mathbf{F}_S$ to be unitary. If we fix $\mathbf{U}_S$ and $\mathbf{C}$ in (2.17) , we can find that the trace of MSE matrix is a decreasing function of α. So under the transmission power constraint, we can have the optimum α as $\alpha_{opt} = \sqrt{\frac{P_{S,T}}{N\sigma_s^2}}$ to minimize the MSE. We substitute α_{opt} into constraint C_2 in (2.17) and find that C_2 is not a function of the source precoder. So, we can just consider it in the master problem. Thus the subproblem becomes:

$$\min_{\mathbf{C}(\mathbf{F}_R),\mathbf{U}_S(\mathbf{F}_R)} \mathrm{Tr}\left(\mathbf{C}\left(\sigma_s^{-2}\mathbf{I}_N + \frac{P_{S,T}}{N\sigma_s^2}\mathbf{U}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{U}_S^H\right)^{-1}\mathbf{C}^H\right) \tag{2.18}$$

With a known relay precoder, the problem (2.18) is similar to the THP design in the conventional MIMO system, and the optimum solution **C**, denote as $\mathbf{C}_{opt}$, has been solved as [24]

$$\mathbf{C}_{opt}=\mathbf{D}\mathbf{L}^{-1} \tag{2.19}$$

where

$$\mathbf{L}\mathbf{L}^{H}=\left(\sigma_s^{-2}\mathbf{I}_N+\frac{P_{S,T}}{N\sigma_s^2}\mathbf{U}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{U}_S\right)^{-1} \tag{2.20}$$

is the Cholesky factorization of $\left(\sigma_s^{-2}\mathbf{I}_N+\frac{P_{S,T}}{N\sigma_s^2}\mathbf{U}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{U}_S\right)^{-1}$ while $\mathbf{D}$ is a diagonal matrix that scales each element on the main diagonal of $\mathbf{C}$ to unity. (The proof is summarized in Appendix A).

Substituting (2.19), (2.20) into (2.18), we then have

$$\begin{aligned} J_{\min} &= \mathrm{Tr}\left\{\mathbf{C}\left(\sigma_s^{-2}\mathbf{I}_N+\frac{P_{S,T}}{N\sigma_s^2}\mathbf{U}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{U}_S\right)^{-1}\mathbf{C}^H\right\} \\ &= \sum_{k=1}^{N}\mathbf{L}(k,k)^2 \geq N\left(\prod_{k=1}^{N}\mathbf{L}(k,k)\right)^{2/N} \end{aligned} \tag{2.21}$$

The inequality in (2.21) is obtained from the arithmetic-mean-geometric-mean (AM-GM) inequality, and the equality is held when $\mathbf{L}(i,i)=\mathbf{L}(j,j),\forall\, i\neq j$. If $\mathbf{U}_S$ is designed properly, the bound in (2.21) can be achieved. For this purpose, we first decompose $\mathbf{U}_S$ as the form of

$$\mathbf{U}_S=\mathbf{V}_{\tilde{\mathbf{H}}}\mathbf{U}_S' \tag{2.22}$$

where $\mathbf{V}_{\tilde{\mathbf{H}}}\in\mathbb{C}^{N\times N}$ is the left singular matrices of $\tilde{\mathbf{H}}$, and $\mathbf{U}_S'\in\mathbb{C}^{N\times N}$ is an unitary matrix to be further specified. Note that this decomposition is always possible for any unitary matrix. Substituting (2.22) into (2.20), we can have (2.20) as

$$\mathbf{L}\mathbf{L}^{H}=\mathbf{U}_S'^{H}\underbrace{\left(\sigma_s^{-2}\mathbf{I}_N+\frac{P_{S,T}}{N\sigma_s^2}\Lambda\right)^{-1}}_{:=\tilde{\mathbf{D}}}\mathbf{U}_S' \tag{2.23}$$

where $\Lambda=diag\left\{\lambda_{\tilde{\mathbf{H}},1},\cdots,\lambda_{\tilde{\mathbf{H}},N}\right\}$, and $\lambda_{\tilde{\mathbf{H}},1},\cdots,\lambda_{\tilde{\mathbf{H}},N}$ are the eigenvalues of $\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}$. To obtain $\mathbf{U}_S'$, we apply geometric mean decomposition (GMD) [26] on $\tilde{\mathbf{D}}^{\frac{1}{2}}$ which can be expressed as

$$\tilde{\mathbf{D}}^{1/2}=\mathbf{Q}\mathbf{R}\mathbf{P}^H \tag{2.24}$$

$\mathbf{Q},\mathbf{P}$ are some unitary matrices, and $\mathbf{R}$ is a upper triangular matrix with equal diagonal elements. Letting $\mathbf{U}_S'=\mathbf{P}$ and substituting (2.24) in (2.23), we can have $\mathbf{L}=\mathbf{R}^H$. The lower bound in (2.21) is achieved since $\mathbf{L}(i,i)=\mathbf{L}(j,j), \forall\, i \neq j$. So, the optimum $\mathbf{F}_S$, denote as $\mathbf{F}_{S,opt}$, can be expressed as

$$\mathbf{F}_{S,opt} = \sqrt{\frac{P_{S,T}}{N\sigma_e^2}}\mathbf{V}_{\tilde{\mathbf{H}}}\mathbf{P} \tag{2.25}$$

From (2.23) and (2.24), we can have the resultant MSE

$$J_{\min} = \sum_{k=1}^{N}\mathbf{L}(k,k)^2 = N\prod_{k=1}^{N}\left(\frac{1}{\lambda_{\tilde{\mathbf{H}},k}\dfrac{P_{S,T}}{N\sigma_s^2}+\sigma_s^{-2}}\right)^{1/N} \tag{2.26}$$

Now, our problem becomes to minimize (2.26) in the master problem. Note that the equality of the right side of (2.26) is satisfied when the diagonal elements of $\mathbf{L}$ are all equal. With a propoer $\mathbf{F}_S$, we can not only minimize the MSE but also make the optimization problem more tractable. To proceed, let us consider the following equivalence:

$$\min_{\mathbf{F}_R} N\prod_{k=1}^{N}\left(\frac{1}{\lambda_{\tilde{\mathbf{H}},k}\dfrac{P_{S,T}}{N\sigma_s^2}+\sigma_s^{-2}}\right)^{1/N} = \max_{\mathbf{F}_R}\left(\sigma_s^{-2}\frac{P_{S,T}}{N}\right)^N \det\left(\left(\frac{N}{P_{S,T}}\mathbf{I}_N+\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\right)\right) \tag{2.27}$$

Note that $\left(\sigma_s^{-2}\dfrac{P_{S,T}}{N}\right)^N$ in (2.27) is a constant so we can reformulate the master problem as:

$$\begin{aligned}&\max_{\mathbf{F}_R}\ \det\left(\frac{N}{P_{S,T}}\mathbf{I}_N+\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\right)\\ &\text{s.t.}\\ &C_2:\mathrm{Tr}\left\{\mathbf{F}_R\left(\sigma_{n,r}^2\mathbf{I}_R+\frac{P_{S,T}}{N}\mathbf{H}_{SR}\mathbf{H}_{SR}^H\right)\mathbf{F}_R^H\right\}\leq P_{R,T}.\end{aligned} \tag{2.28}$$

The problem (2.28) is still difficult to solve because the cost function is a nonlinear function of $\mathbf{F}_R$ and the problem is not convex either. To solve the problem, we propose a relay precoder structure such that a closed-form solution can be solved. Directly solving (2.28) is not feasible. We then use a lemma describe below:

Lemma 2.2 [41]: Let $\mathbf{M} \in \mathbb{C}^{N\times N}$ be a positive definite matrix, then

$$\det(\mathbf{M}) \leq \prod_{i=1}^{N}\mathbf{M}(i,i) \tag{2.29}$$

where $\mathbf{M}(i,i)$ denotes the ith diagonal element of $\mathbf{M}$. Note that the equality in (2.29) holds when $\mathbf{M}$ is a diagonal matrix. If we let $\mathbf{M} = \tilde{\mathbf{H}}^H \tilde{\mathbf{H}}$, it turns out that when $\mathbf{M}$ is diagonal, the maximization of the cost function becomes possible. To have the diagonalization, we need another lemma shown below:

Lemma 2.3 [41]: Let $\mathbf{A} \in \mathbb{C}^{N \times N}$ be a positive matrix and $\mathbf{B} \in \mathbb{C}^{N \times N}$, then

$$\det(\mathbf{A} + \mathbf{B}) = \det(\mathbf{A}) \det\left(\mathbf{I}_N + \mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}\right) \tag{2.30}$$

Form (2.30), we let $\mathbf{B} = \mathbf{H}_{SR}^H \mathbf{F}_R^H \mathbf{H}_{RD}^H \left(\sigma_{n,r}^2 \mathbf{H}_{RD}\mathbf{F}_R\mathbf{F}_R^H\mathbf{H}_{RD}^H + \sigma_{n,d}^2 \mathbf{I}_M\right)^{-1} \mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}$ and $\mathbf{A} = \dfrac{N}{P_{S,T}}\mathbf{I}_N + \sigma_{n,d}^{-2}\mathbf{H}_{SD}^H\mathbf{H}_{SD}$ we have the following equivalence:

$$\begin{aligned}
&\arg\max_{\mathbf{F}_R} \det\left(\frac{N}{P_{S,T}}\mathbf{I}_N + \tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\right) \\
&= \arg\max_{\mathbf{F}_R} \det\left(\mathbf{I}_N + \mathbf{H}_{SR}'^H \mathbf{F}_R^H \mathbf{H}_{RD}^H \left(\sigma_{n,r}^2 \mathbf{H}_{RD}\mathbf{F}_R\mathbf{F}_R^H\mathbf{H}_{RD}^H + \sigma_{n,d}^2 \mathbf{I}_M\right)^{-1} \mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}'\right)
\end{aligned} \tag{2.31}$$

where $\mathbf{H}_{SR}' := \mathbf{H}_{SR}\left(N / P_{S,T}\mathbf{I}_N + \sigma_{n,d}^{-2}\mathbf{H}_{SD}^H\mathbf{H}_{SD}\right)^{-\frac{1}{2}}$. Note that $\det(\mathbf{A})$ is not the function of $\mathbf{F}_R$, so we can ignore it. The optimization problem can be rewritten as

$$\begin{aligned}
&\max_{\mathbf{F}_R} \ \det(\mathbf{M}') \\
&\text{s.t.} \\
&\mathbf{M}' = \left(\mathbf{I}_N + \mathbf{H}_{SR}'^H \mathbf{F}_R^H \mathbf{H}_{RD}^H\right. \\
&\left.\left(\sigma_{n,r}^2 \mathbf{H}_{RD}\mathbf{F}_R\mathbf{F}_R^H\mathbf{H}_{RD}^H + \sigma_{n,d}^2 \mathbf{I}_M\right)^{-1} \mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}'\right) \\
&C_2 : \sigma_{n,r}^2 \left\|\mathbf{F}_R\right\|_2^2 + \frac{P_{S,T}}{N}\left\|\mathbf{F}_R\mathbf{H}_{SR}\right\|_2^2 \le P_{R,T}.
\end{aligned} \tag{2.32}$$

Taking a close look at (2.32), we can see that there exists a precoder structure for the relay precoder $\mathbf{F}_R$ such that the diagonalization can be achieved. Consider the singular value decomposition (SVD) on $\mathbf{H}_{RD}$ and $\mathbf{H}_{SR}'$:

$$\mathbf{H}_{RD} = \mathbf{U}_{rd}\boldsymbol{\Sigma}_{rd}\mathbf{V}_{rd}^H \tag{2.33}$$

$$\mathbf{H}_{SR}' = \mathbf{U}_{sr}'\boldsymbol{\Sigma}_{sr}'\mathbf{V}_{sr}'^H \tag{2.34}$$

where $\mathbf{U}_{rd} \in \mathbb{C}^{M \times M}$ and $\mathbf{U}_{sr}' \in \mathbb{C}^{R \times R}$ are left singular matrices of $\mathbf{H}_{RD}$ and $\mathbf{H}_{SR}'$, respectively; $\boldsymbol{\Sigma}_{rd} \in \mathbb{R}^{M \times R}$ and $\boldsymbol{\Sigma}_{sr}' \in \mathbb{R}^{R \times N}$ are the diagonal singular value matrices of $\mathbf{H}_{RD}$ and $\mathbf{H}_{SR}'$, respectively; $\mathbf{V}_{rd}^H \in \mathbb{C}^{R \times R}$ and $\mathbf{V}_{sr}'^H \in \mathbb{C}^{N \times N}$ are the right singular matrices of $\mathbf{H}_{RD}$ and $\mathbf{H}_{SR}'$, respectively. If we let $\mathbf{F}_R$ have the structure as the following form, a full diagonalization of the matrix of the determinant in (2.32) can be achieved:

$$\mathbf{F}_{R,opt} = \mathbf{V}_{rd}\boldsymbol{\Sigma}_r\mathbf{U}'^H_{sr} \tag{2.35}$$

where $\boldsymbol{\Sigma}_r$ is a diagonal matrix with its ith diagonal element of $\sigma_{r,i}$. Let $\sigma_{rd,i}$ and $\sigma'_{sr,i}$ be the ith diagonal element of $\boldsymbol{\Sigma}_{rd}$ and $\boldsymbol{\Sigma}'_{sr}$, respectively. With the relay precoder structure, the optimization in the master problem can finally be translated into a scalar-value concave optimization problem. Substituting (2.33), (2.34), (2.35) into (2.32), we can rewrite (2.32) as:

$$\begin{aligned}
&\max_{p_{r,i},\ 1\le i\le N} \sum_{i=1}^{N} \ln\left(1+\frac{p_{r,i}\sigma^2_{n,d}\sigma^2_{rd,i}\sigma'^2_{sr,i}}{p_{r,i}\sigma^2_{n,r}\sigma^2_{rd,i}+\sigma^2_{n,d}}\right)\\
&s.t.\\
&\sum_{i=1}^{N} p_{r,i}\left(\frac{P_{S,T}}{N}\sigma'^2_{sr,i}\mathbf{D}'_{sr}(i,i)+\sigma^2_{n,r}\right)\le P_{R,T},\ p_{r,i}\ge 0,
\end{aligned} \tag{2.36}$$

where $p_{r,i}=\sigma^2_{r,i}$, $\mathbf{D}'_{sr}=\mathbf{V}'^H_{sr}\left(N/P_{S,T}\mathbf{I}_N+\mathbf{H}^H_{SD}\mathbf{H}_{SD}\right)\mathbf{V}'_{sr}$, and $\mathbf{D}'_{sr}(i,i)$ stands for the ith diagonal element of $\mathbf{D}'_{sr}$. It is apparent that the mast problem is now a scalar optimization problem. And, since $p_{r,i}\ge 0$, the cost function (2.36) is concave [42]. To solve this problem, we apply the KKT conditions given by [42] and find the solution for $p_{r,i}$, $i=1,\cdots,N$ as:

$$p_{r,i}=\left[\sqrt{\frac{\mu}{\sigma^2_{rd,i}\left(\frac{P_{S,T}}{N}\sigma'^2_{sr,i}\mathbf{D}'_{sr}(i,i)+\sigma^2_{n}\right)\left(\sigma^2_{n,r}\sigma^{-2}_{n,d}\sigma'^{-2}_{sr,i}+1\right)}+\frac{\dfrac{\frac{1}{4}\frac{\sigma^4_{n,d}}{\sigma^4_{n,r}}}{\sigma^4_{rd,i}\left(\frac{\sigma^2_{n,r}}{\sigma^2_{n,d}\sigma'^2_{sr,i}}+1\right)^2}}{}}-\frac{1+\frac{1}{2}\frac{\sigma^2_{n,d}\sigma'^2_{sr,i}}{\sigma^2_{n,r}}}{\sigma^2_{rd,i}\left(\frac{\sigma^2_{n,r}}{\sigma^2_{n,d}}+\sigma'^2_{sr,i}\right)}\right]^+, \tag{2.37}$$

where μ is chosen to satisfy the power constraint in (2.36). Substituting (2.37) into (2.35), we can obtain the optimum relay precoder, and then $\tilde{\mathbf{H}}$ in (2.12) can also be obtained. Finally, the unitary source precoder can be derived by substituting (2.24) into (2.25) , and the matrix $\mathbf{C}$ can be obtained by (2.19).

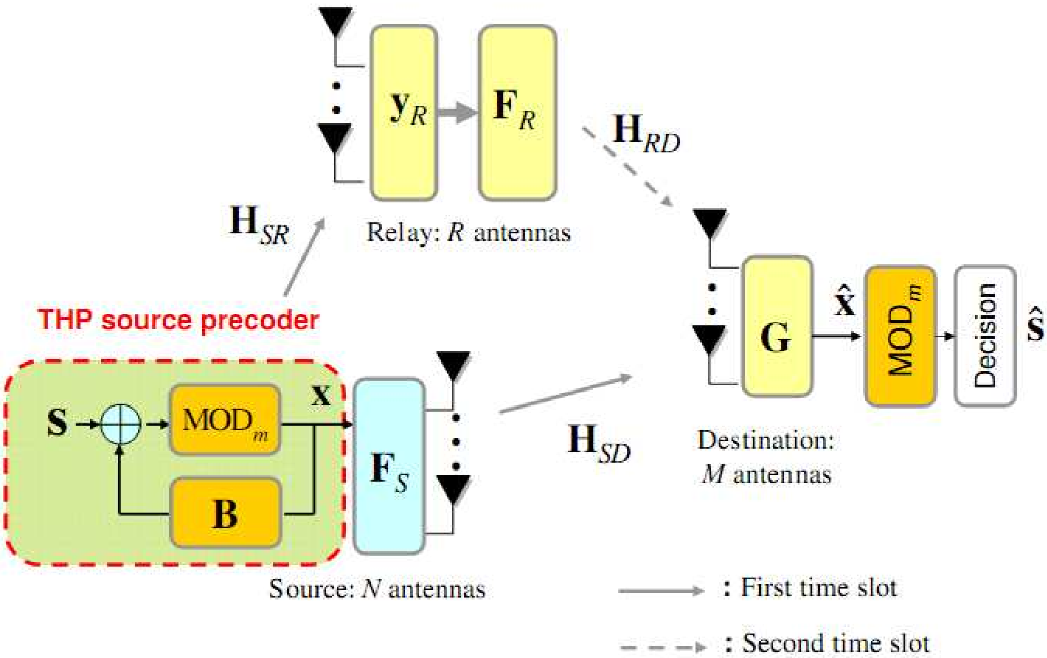

Figure 2.1: THP source and linear relay precoded AF MIMO relay system with MMSE receiver.

Chapter 3

Robust Joint MMSE Transceiver Desgin with Tomlinson-Harashima Source and Linear Relay Precoders

As well known, the performance of transceiver design relies on the accuracy of channel state information (CSI). In the literature, most transceiver designs assume perfect CSI. Our design in Chapter 2 also assumes that the destination has perfect CSIs of the three links. However, for real-world implementation, perfect CSIs are usually not attainable due to channel estimation or quantization error. The performance of the transceiver designed with imperfect CSIs may be degraded seriously. In this chapter, we consider a robust nonlinear transceiver design in which the THP, the linear relay and the MMSE receiver are used at the source, the relay and the destination, respectively. The imperfect CSIs from both the relay and direct links are incorporated into design where channel estimation errors are modeled as Gaussian random variables. The design procedure is similar to that we have used in Chapter 2. The main idea is to use the primal decomposition and some approximations such that a close-form solution of the optimization problem can be obtained. In Section 3.1, we build the system model taking the channel uncertainty into consideration. In Section 3.2, we formulate the designing problem under the MMSE criterion. In Section 3.3, we propose a new approach to solve the problem in closed-form.

3.1 System model

We consider a three-node AF MIMO relay precoded system which has been presented in Chapter 2, as shown in Figure 2.1. With the two-phase transmission protocol, the signal from the source is transmitted to the relay and the destination simultaneously in the first phase. Then the received signal at the relay is multiplied by the relay precoder and send to the destination in the second phase. For simplicity, we let the signal notations are similar to those in Chapter 2, (2.1)-(2.4). In general, the actual channel matrix can be modeled as a summation of an estimated channel and an error matrix [36]. Thus, we have:

$$\mathbf{H}_{SR} = \widehat{\mathbf{H}}_{SR} + \Delta\mathbf{H}_{SR} \tag{3.1}$$

$$\mathbf{H}_{RD} = \widehat{\mathbf{H}}_{RD} + \Delta\mathbf{H}_{RD} \tag{3.2}$$

$$\mathbf{H}_{SD} = \widehat{\mathbf{H}}_{SD} + \Delta\mathbf{H}_{SD} \tag{3.3}$$

where $\mathbf{H}_{SR} \in \mathbb{C}^{R\times N}$, $\mathbf{H}_{SD} \in \mathbb{C}^{M\times N}$ and $\mathbf{H}_{RD} \in \mathbb{C}^{M\times R}$ are the actual channel matrices of the source-to-relay, the source-to-destination, and the relay-to-destination links, respectively; $\widehat{\mathbf{H}}_{SR}$, $\widehat{\mathbf{H}}_{RD}$, and $\widehat{\mathbf{H}}_{SD}$ are the estimated channel matrices of $\mathbf{H}_{SR}$, $\mathbf{H}_{RD}$, and $\mathbf{H}_{SD}$, respectively. $\Delta\mathbf{H}_{SR}$, $\Delta\mathbf{H}_{RD}$, and $\Delta\mathbf{H}_{SD}$ are the corresponding channel errors matrices in which all the elements are assumed to be zero mean Gaussian random variables. An estimation channel matrix can be further decomposed into a product of three matrices. For example, $\Delta\mathbf{H}_{SR}$ can be express as:

$$\Delta\mathbf{H}_{SR} = \mathbf{\Sigma}_{SR}^{1/2}\mathbf{H}_{\text{i.i.d.}}\mathbf{\Psi}_{SR}^{1/2} \tag{3.4}$$

where the elements of $\mathbf{H}_{\text{i.i.d.}}$ are independent and identically distributed (i.i.d) Gaussian random variables with zero-mean and unit variance; $\mathbf{\Sigma}_{SR} \in \mathbb{C}^{R\times R}$ and $\mathbf{\Psi}_{SR} \in \mathbb{C}^{N\times N}$ are the row and column covariance matrices of $\Delta\mathbf{H}_{SR}$, respectively [38].

From (3.4), it is clear that $\text{vec}(\Delta\mathbf{H}_{SR}) \sim CN\left(\mathbf{0}_{NR\times 1}, \mathbf{\Sigma}_{SR} \otimes \mathbf{\Psi}_{SR}^{T}\right)$, where $CN(\mathbf{m},\mathbf{C})$ denotes a complex Gaussian random vector with mean $\mathbf{m}$ and covariance $\mathbf{C}$ [44]. Similarly, we can set the distribution of channel estimation error $\Delta\mathbf{H}_{RD}$ and $\Delta\mathbf{H}_{SD}$ as $\text{vec}(\Delta\mathbf{H}_{RD})$ $\sim CN\left(\mathbf{0}_{RM\times 1}, \mathbf{\Sigma}_{RD} \otimes \mathbf{\Psi}_{RD}^{T}\right)$ and $\text{vec}(\Delta\mathbf{H}_{SD}) \sim CN\left(\mathbf{0}_{NM\times 1}, \mathbf{\Sigma}_{SD} \otimes \mathbf{\Psi}_{SD}^{T}\right)$. It is noteworthy that the expression of $\mathbf{\Sigma}_{SR}$, $\mathbf{\Psi}_{SR}$, $\mathbf{\Sigma}_{RD}$, $\mathbf{\Psi}_{RD}$, $\mathbf{\Sigma}_{SD}$, $\mathbf{\Psi}_{SD}$ depend on specific channel estimation algorithms. For example, $\mathbf{\Psi}_{SR} = \mathbf{R}_{T,SR}$ and $\mathbf{\Sigma}_{SR} = \sigma_{e,sr}^{2}\mathbf{R}_{R,SR}$ if we use the estimation method proposed in [29]. $\mathbf{R}_{T,SR}$ and $\mathbf{R}_{R,SR}$ are the transmit and receive antenna correlation matrices; $\sigma_{e,sr}^{2}$ is the source-relay link channel estimation error variance. And note that the other two matrices have the similar structures. Here, we assume all the channels are time-invariant and all second-order statistics - $\mathbf{\Sigma}_{SR}$, $\mathbf{\Psi}_{SR}$, $\mathbf{\Sigma}_{RD}$, $\mathbf{\Psi}_{RD}$, $\mathbf{\Sigma}_{SD}$, $\mathbf{\Psi}_{SD}$ are known as a prior.

At the destination, we can have a single received vector for the two-phase transmission:

$$\mathbf{y}_{D} := \underbrace{\begin{bmatrix} \mathbf{H}_{SD} \\ \mathbf{H}_{RD}\mathbf{F}_{R}\mathbf{H}_{SR} \end{bmatrix}}_{:=\mathbf{H}}\mathbf{F}_{S}\mathbf{x} + \underbrace{\begin{bmatrix} \mathbf{n}_{D,1} \\ \mathbf{H}_{RD}\mathbf{F}_{R}\mathbf{n}_{R} + \mathbf{n}_{D,2} \end{bmatrix}}_{:=\mathbf{w}} \tag{3.5}$$

We use the MMSE receiver, that takes both the noise and the channel estimation error into account at the destination, to recover the transmitted signal. Define the MSE as:

$$
\begin{aligned}
J\left(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R,\mathbf{G}\right) &= E_{\Delta,\mathbf{w}}\left\{\left\|\mathbf{G}\mathbf{y}_D-\mathbf{v}\right\|^2\right\} \\
&= E_{\Delta,\mathbf{w}}\left\{\mathrm{Tr}\left\{\left(\left(\mathbf{GHF}_S-\mathbf{C}\right)\mathbf{x}+\mathbf{Gw}\right)\left(\left(\mathbf{GHF}_S-\mathbf{C}\right)\mathbf{x}+\mathbf{Gw}\right)^H\right\}\right\}
\end{aligned}
\tag{3.6}
$$

where $\mathbf{G}$ represents the equalization matrix and the subscript Δ, $\mathbf{w}$ denote that the expectation is taken over both the channel estimation error and noise. With the same assumption in Chapter 2, we can have $E\left[\mathbf{x}\mathbf{x}^H\right]=\sigma_s^2\mathbf{I}_N$ and $E\left[\mathbf{v}\mathbf{v}^H\right]=\sigma_s^2\mathbf{C}\mathbf{C}^H$. Thus, we can rewrite the MSE in (3.6) as:

$$
\begin{aligned}
&J\left(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R,\mathbf{G}\right)= \\
&\mathrm{Tr}\left\{\sigma_s^2\mathbf{G}\begin{bmatrix} \mathbf{T}_{SD} & \hat{\mathbf{H}}_{SD}\mathbf{F}_S\mathbf{F}_S^H\hat{\mathbf{H}}_{SR}^H\mathbf{F}_R^H\hat{\mathbf{H}}_{RD}^H \\ \hat{\mathbf{H}}_{RD}\mathbf{F}_R\hat{\mathbf{H}}_{SR}\mathbf{F}_S\mathbf{F}_S^H\hat{\mathbf{H}}_{SD}^H & \mathrm{Tr}\left(\mathbf{F}_R\mathbf{T}_{SR}\mathbf{F}_R^H\boldsymbol{\Psi}_{RD}\right)\boldsymbol{\Sigma}_{RD}+\hat{\mathbf{H}}_{RD}\mathbf{F}_R\mathbf{T}_{SR}\mathbf{F}_R^H\hat{\mathbf{H}}_{RD}^H \end{bmatrix}\mathbf{G}^H\right\} \\
&\quad-\mathrm{Tr}\left\{\sigma_s^2\mathbf{G}\hat{\mathbf{H}}\mathbf{F}_S\mathbf{C}^H\right\}-\mathrm{Tr}\left\{\sigma_s^2\mathbf{C}\mathbf{F}_S^H\hat{\mathbf{H}}^H\mathbf{G}^H\right\}+\mathrm{Tr}\left\{\sigma_s^2\mathbf{C}\mathbf{C}^H\right\} \\
&\quad+\mathrm{Tr}\left\{\mathbf{G}\begin{bmatrix} \sigma_{n,d}^2\mathbf{I}_M & \mathbf{0} \\ \mathbf{0} & \sigma_{n,r}^2\mathrm{Tr}\left(\mathbf{F}_R\mathbf{F}_R^H\boldsymbol{\Psi}_{RD}\right)\boldsymbol{\Sigma}_{RD}+\sigma_{n,r}^2\hat{\mathbf{H}}_{RD}\mathbf{F}_R\mathbf{F}_R^H\hat{\mathbf{H}}_{RD}^H+\sigma_{n,d}^2\mathbf{I}_M \end{bmatrix}\mathbf{G}^H\right\} \\
&=\mathrm{Tr}\left\{\mathbf{G}\left(\sigma_s^2\hat{\mathbf{H}}\mathbf{F}_S\mathbf{F}_S^H\hat{\mathbf{H}}^H+\hat{\mathbf{R}}_w+\Delta\mathbf{err}\right)\mathbf{G}^H\right\}-\mathrm{Tr}\left\{\sigma_s^2\mathbf{G}\hat{\mathbf{H}}\mathbf{F}_S\mathbf{C}^H\right\}- \\
&\mathrm{Tr}\left\{\sigma_s^2\mathbf{C}\mathbf{F}_S^H\hat{\mathbf{H}}^H\mathbf{G}^H\right\}+\mathrm{Tr}\left\{\sigma_s^2\mathbf{C}\mathbf{C}^H\right\}
\end{aligned}
\tag{3.7}
$$

where

$$
\hat{\mathbf{H}}=\begin{bmatrix} \hat{\mathbf{H}}_{SD} \\ \hat{\mathbf{H}}_{RD}\mathbf{F}_R\hat{\mathbf{H}}_{SR} \end{bmatrix}
\tag{3.8}
$$

$$
\Delta\mathbf{err}=\begin{bmatrix} \sigma_s^2\mathrm{Tr}\left(\mathbf{F}_S\mathbf{F}_S^H\boldsymbol{\Psi}_{SD}\right)\boldsymbol{\Sigma}_{SD} & \mathbf{0} \\ \mathbf{0} & \begin{matrix}\mathrm{Tr}\left(\mathbf{F}_R\left(\sigma_s^2\mathbf{T}_{SR}+\sigma_{n,r}^2\mathbf{I}_R\right)\mathbf{F}_R^H\boldsymbol{\Psi}_{RD}\right)\boldsymbol{\Sigma}_{RD}+ \\ \sigma_s^2\hat{\mathbf{H}}_{RD}\mathbf{F}_R\left(\mathrm{Tr}\left(\mathbf{F}_S\mathbf{F}_S^H\boldsymbol{\Psi}_{SR}\right)\boldsymbol{\Sigma}_{SR}\right)\mathbf{F}_R^H\hat{\mathbf{H}}_{RD}^H\end{matrix} \end{bmatrix}
\tag{3.9}
$$

$$
\hat{\mathbf{R}}_w=E\left[\begin{bmatrix} \mathbf{n}_{D,1} \\ \hat{\mathbf{H}}_{RD}\mathbf{F}_R\mathbf{n}_R+\mathbf{n}_{D,2} \end{bmatrix}\begin{bmatrix} \mathbf{n}_{D,1}^H & \mathbf{n}_R^H\mathbf{F}_R^H\hat{\mathbf{H}}_{RD}^H+\mathbf{n}_{D,2}^H \end{bmatrix}\right]
\tag{3.10}
$$

$$
\mathbf{T}_{SD}=\mathrm{Tr}\left(\mathbf{F}_S\mathbf{F}_S^H\boldsymbol{\Psi}_{SD}\right)\boldsymbol{\Sigma}_{SD}+\hat{\mathbf{H}}_{SD}\mathbf{F}_S\mathbf{F}_S^H\hat{\mathbf{H}}_{SD}^H
\tag{3.11}
$$

$$
\mathbf{T}_{SR}=\mathrm{Tr}\left(\mathbf{F}_S\mathbf{F}_S^H\boldsymbol{\Psi}_{SR}\right)\boldsymbol{\Sigma}_{SR}+\hat{\mathbf{H}}_{SR}\mathbf{F}_S\mathbf{F}_S^H\hat{\mathbf{H}}_{SR}^H
\tag{3.12}
$$

The derivation of the MSE in (3.7) can be found in Appendix B.

Since the cost function is convex in $\mathbf{G}$, we can find the minimum MSE in (3.7) by taking the gradient of J with respect to $\mathbf{G}$ and set the result to zero. Note that we keep $\mathbf{F}_S$, $\mathbf{F}_R$, and $\mathbf{C}$ fixed in the operation. Then we can obtain the optimum equalization matrix $\mathbf{G}$, denote as $\mathbf{G}_{opt}$, as

$$\mathbf{G}_{opt} = \sigma_s^2 \mathbf{C}\mathbf{F}_S^H \widehat{\mathbf{H}}^H \left(\sigma_s^2 \widehat{\mathbf{H}}\mathbf{F}_S\mathbf{F}_S^H \widehat{\mathbf{H}}^H + \widehat{\mathbf{R}}_w + \Delta\mathbf{err} \right)^{-1} \tag{3.13}$$

Substituting (3.13) to (3.7), we then have the minimum MSE as

$$\begin{aligned} &J\left(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R,\mathbf{G}_{opt}\right) \triangleq J_{\min}\left(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R\right) \\ &= \sigma_s^2 \mathrm{Tr}\left\{\mathbf{C}\mathbf{C}^H\right\} - \sigma_s^2 \mathrm{Tr}\left\{ \mathbf{C}\mathbf{F}_S^H \widehat{\mathbf{H}}^H \left(\widehat{\mathbf{H}}\mathbf{F}_S\mathbf{F}_S^H \widehat{\mathbf{H}}^H + \sigma_s^{-2}\left(\widehat{\mathbf{R}}_w + \Delta\mathbf{err}\right)\right)^{-1} \widehat{\mathbf{H}}\mathbf{F}_S\mathbf{C}^H \right\}. \end{aligned} \tag{3.14}$$

By using matrix inversion lemma $\mathbf{A}^H\left(\mathbf{A}\mathbf{A}^H + \mathbf{I}\right)^{-1}\mathbf{A} = \mathbf{I} - \left(\mathbf{A}^H\mathbf{A} + \mathbf{I}\right)^{-1}$ [42], (3.14) can be further rewritten as:

$$\begin{aligned} &J_{\min}\left(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R\right) \\ &= \sigma_s^2 \mathrm{Tr}\left\{\mathbf{C}\mathbf{C}^H\right\} - \sigma_s^2 \mathrm{Tr}\left\{ \mathbf{C}\mathbf{F}_S^H \widehat{\mathbf{H}}^H \left(\widehat{\mathbf{H}}\mathbf{F}_S\mathbf{F}_S^H \widehat{\mathbf{H}}^H + \underbrace{\sigma_s^{-2}\left(\widehat{\mathbf{R}}_w + \Delta\mathbf{err}\right)}_{\mathbf{R}_\Delta} \right)^{-1} \widehat{\mathbf{H}}\mathbf{F}_S\mathbf{C}^H \right\} \\ &= \sigma_s^2 \mathrm{Tr}\left\{\mathbf{C}\mathbf{C}^H\right\} - \sigma_s^2 \mathrm{Tr}\left\{ \mathbf{C}\underbrace{\mathbf{F}_S^H \widehat{\mathbf{H}}^H \mathbf{R}_\Delta^{-\frac{H}{2}}}_{\mathbf{A}^H} \left(\underbrace{\mathbf{R}_\Delta^{-\frac{1}{2}} \widehat{\mathbf{H}}\mathbf{F}_S}_{\mathbf{A}} \underbrace{\mathbf{F}_S^H \widehat{\mathbf{H}}^H \mathbf{R}_\Delta^{-\frac{H}{2}}}_{\mathbf{A}^H} + \mathbf{I}_{2M} \right)^{-1} \underbrace{\mathbf{R}_\Delta^{-\frac{1}{2}} \widehat{\mathbf{H}}\mathbf{F}_S}_{\mathbf{A}} \mathbf{C}^H \right\} \\ &= \sigma_s^2 \mathrm{Tr}\left\{ \mathbf{C}\left(\mathbf{F}_S^H \widehat{\mathbf{H}}^H \mathbf{R}_\Delta^{-1} \widehat{\mathbf{H}}\mathbf{F}_S + \mathbf{I}_N \right)^{-1} \mathbf{C}^H \right\} \\ &= \sigma_s^2 \mathrm{Tr}\left\{ \mathbf{E}\left(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R\right) \right\} \end{aligned} \tag{3.15}$$

where

$$\begin{aligned} \mathbf{R}_\Delta &= \sigma_s^{-2}\left(\widehat{\mathbf{R}}_w + \Delta\mathbf{err}\right) \\ &= \sigma_s^{-2}\begin{bmatrix} \mathbf{R}_{\Delta,1,1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{\Delta,2,2} \end{bmatrix} \end{aligned} \tag{3.16}$$

and

$$\mathbf{R}_{\Delta,1,1} = \sigma_{n,d}^2 \mathbf{I}_M + \sigma_s^2 \mathrm{Tr}\left(\mathbf{F}_S\mathbf{F}_S^H \mathbf{\Psi}_{SD}\right)\mathbf{\Sigma}_{SD} \tag{3.17}$$

$$\begin{aligned} \mathbf{R}_{\Delta,2,2} = &\mathrm{Tr}\left(\mathbf{F}_R\left(\sigma_s^2\left(\mathrm{Tr}\left(\mathbf{F}_S\mathbf{F}_S^H\mathbf{\Psi}_{SR}\right)\mathbf{\Sigma}_{SR} + \widehat{\mathbf{H}}_{SR}\mathbf{F}_S\mathbf{F}_S^H\widehat{\mathbf{H}}_{SR}^H \right) + \sigma_{n,r}^2\mathbf{I}_R \right)\mathbf{F}_R^H\mathbf{\Psi}_{RD} \right)\mathbf{\Sigma}_{RD} \\ &+ \sigma_s^2 \widehat{\mathbf{H}}_{RD}\mathbf{F}_R\left(\mathrm{Tr}\left(\mathbf{F}_S\mathbf{F}_S^H\mathbf{\Psi}_{SR}\right)\mathbf{\Sigma}_{SR} \right)\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H + \sigma_{n,r}^2 \widehat{\mathbf{H}}_{RD}\mathbf{F}_R\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H + \sigma_{n,d}^2\mathbf{I}_M. \end{aligned} \tag{3.18}$$

From (3.16)-(3.18), we can see that the MSE in (3.15) is a complicate and nonlinear function of $\mathbf{F}_S$ and $\mathbf{F}_R$.

3.2 Problem formulation

In Section 3.1, we construct the system model using the MMSE criterion. The objective is to find $\mathbf{C}$, $\mathbf{F}_S$ and $\mathbf{F}_R$ so that the MSE (3.15) can be minimized. We now can formulate the joint precoder design problem as below:

$$\begin{aligned}
&\min_{\mathbf{C},\mathbf{F}_S,\mathbf{F}_R} \sigma_s^2 \mathrm{Tr}\left\{\mathbf{C}\left(\mathbf{F}_S^H \widehat{\mathbf{H}}^H \mathbf{R}_\Delta^{-1} \widehat{\mathbf{H}}\mathbf{F}_S + \mathbf{I}_N\right)^{-1} \mathbf{C}^H\right\} \\
&\text{s.t} \\
&C_1, C_2
\end{aligned} \tag{3.19}$$

where

$$\begin{aligned}
&C_1 : \sigma_s^2 \mathrm{Tr}\left\{\mathbf{F}_S \mathbf{F}_S^H\right\} \le P_{S,T}, \\
&C_2 : E\left[\mathrm{Tr}\left\{\mathbf{F}_R\left(\sigma_{n,r}^2 \mathbf{I}_R + \sigma_s^2 \mathbf{H}_{SR}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}_{SR}^H\right)\mathbf{F}_R^H\right\}\right] = \\
&\sigma_{n,r}^2 \mathrm{Tr}\left\{\mathbf{F}_R\mathbf{F}_R^H\right\} + \sigma_s^2 \mathrm{Tr}\left\{\mathbf{F}_R\left(\mathrm{Tr}\left(\mathbf{F}_S\mathbf{F}_S^H\mathbf{\Psi}_{SR}\right)\mathbf{\Sigma}_{SR} + \widehat{\mathbf{H}}_{SR}\mathbf{F}_S\mathbf{F}_S^H\widehat{\mathbf{H}}_{SR}^H\right)\mathbf{F}_R^H\right\} \le P_{R,T},
\end{aligned} \tag{3.20}$$

In (3.20), C_1 and C_2 stand for the transmission power constraints at the source and the relay, respectively, $\mathbf{C}$ is a lower triangular matrix with unit diagonal, and $\mathbf{R}_\Delta$ is the matrix specified in (3.16).

From (3.20), we see that the cost function and constraints are functions of $\mathbf{C}$, $\mathbf{F}_S$ and $\mathbf{F}_R$. As we can also see, the functions are complicated and the problem is not a convex optimization problem. Since imperfect CSIs of all links are involved, the problem becomes much more difficult than that in Chapter 2. In the next section, we propose a new approach to solve this problem.

3.3 Robust joint source and relay precoder design

Similar to that in Chapter 2, we also cascade an unitary precoder $\mathbf{F}_S$ after the THP. The precoder $\mathbf{F}_S$ can not only facilitate the optimization but improve the BER performance. The cost function is a nonlinear function of $\mathbf{F}_S$, $\mathbf{F}_R$ and $\mathbf{C}$. It is then difficult to find the optimum precoders simultaneously. We use the primal decomposition method so that the optimization problem can be translated into a subproblem and a master problem [42]. The procedure is first to split unknown variables into two group and the variables in the first group are treated as known constants. Then, the variables in second group are solved as the functions of the functions of the variables in first group (the subproblem), and the cost function is reduce to a function of the variables in the first group. Finally, the variables in first group are solved (the master problem). In this case, the subproblem is to find the optimal $\mathbf{C}$ and $\mathbf{F}_S$ by letting $\mathbf{F}_R$ be fixed, and the master problem is optimized for $\mathbf{F}_R$. For convenience, we rewrite the problem in (3.19) as:

$$\begin{aligned}
&\min_{\mathbf{C},\mathbf{F}_S,\mathbf{F}_R} \operatorname{Tr}\left\{\mathbf{E}\left(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R\right)\right\} = \min_{\mathbf{F}_R}\min_{\mathbf{C},\mathbf{F}_S} \operatorname{Tr}\left\{\mathbf{E}\left(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R\right)\right\}\\
&\text{s.t.}\\
&\mathbf{F}_S=\alpha_S\mathbf{U}_S,\\
&C_1' : \sigma_s^2\alpha_S^2 N \le P_{S,T},\\
&C_2 : Tr\left\{\mathbf{F}_R\left(\sigma_{n,r}^2\mathbf{I}_R + \sigma_s^2\alpha_S^2\widehat{\mathbf{H}}_{SR}\widehat{\mathbf{H}}_{SR}^H + \sigma_s^2\alpha_S^2\operatorname{Tr}\left(\boldsymbol{\Psi}_{SR}\right)\boldsymbol{\Sigma}_{SR}\right)\mathbf{F}_R^H\right\}\\
&= \operatorname{Tr}\left\{\mathbf{F}_R\left(\sigma_{n,r}^2\mathbf{I}_R + \sigma_s^2\alpha_S^2\underbrace{\left(\widehat{\mathbf{H}}_{SR}\widehat{\mathbf{H}}_{SR}^H + \operatorname{Tr}\left(\boldsymbol{\Psi}_{SR}\right)\boldsymbol{\Sigma}_{SR}\right)}_{:=\widehat{\mathbf{H}}'_{SR}\widehat{\mathbf{H}}'^H_{SR}}\right)\mathbf{F}_R^H\right\} \le P_{R,T}
\end{aligned} \tag{3.21}$$

where $\mathbf{E}\left(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R\right) = \mathbf{C}\left(\mathbf{F}_S^H\widehat{\mathbf{H}}^H\mathbf{R}_\Delta^{-1}\widehat{\mathbf{H}}\mathbf{F}_S + \mathbf{I}_N\right)^{-1}\mathbf{C}^H$ and $\widehat{\mathbf{H}}'_{SR} := \left(\widehat{\mathbf{H}}_{SR}\widehat{\mathbf{H}}_{SR}^H + \operatorname{Tr}\left(\boldsymbol{\Psi}_{SR}\right)\boldsymbol{\Sigma}_{SR}\right)^{1/2}$. We let $\mathbf{F}_S$ have the form as that shown in (3.21), where α_S is a scalar and $\mathbf{U}_S \in \mathbb{C}^{N\times N}$ is a unitary matrix. The constraint C_1' is obtained by substituting $\mathbf{F}_S = \alpha\mathbf{U}_S$ into C_1 in (3.20). In the subproblem, we can observe that the transmitted power constraint at the relay, C_2, is not the function of $\mathbf{C}$ or $\mathbf{F}_S$, so we can move it to the master problem. Let $\mathbf{F}_S = \alpha\mathbf{U}_S$, the cost function and the constraints can then become functions of $\mathbf{C}$, α and $\mathbf{U}_S$. Then the subproblem can be rewritten as:

$$
\begin{aligned}
&\min_{\mathbf{C}(\mathbf{F}_R),\alpha\mathbf{U}_S(\mathbf{F}_R)} \operatorname{Tr}\left(\mathbf{E}\left(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R\right)\right) \\
&\text{s.t.} \\
&\mathbf{E}\left(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R\right) = \mathbf{C}\left(\mathbf{F}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{F}_S + \mathbf{I}_N\right)^{-1}\mathbf{C}^H \ , \\
&\mathbf{F}_S = \alpha_S\mathbf{U}_S, \\
&C_1' : \sigma_s^2\alpha_S^2 N \le P_{S,T}, \\
&C_2 : \operatorname{Tr}\left\{\mathbf{F}_R\left(\sigma_{n,r}^2\mathbf{I}_R + \sigma_s^2\alpha_S^2\widehat{\mathbf{H}}'_{SR}\widehat{\mathbf{H}}'^H_{SR}\right)\mathbf{F}_R^H\right\} \le P_{R,T}
\end{aligned}
\tag{3.22}
$$

where $\mathbf{E}$ is now a function of $\mathbf{C}$, α_S, $\mathbf{U}_S$, and

$$
\begin{aligned}
\tilde{\mathbf{H}}^H\tilde{\mathbf{H}} &:= \widehat{\mathbf{H}}^H\mathbf{R}_\Delta^{-1}\widehat{\mathbf{H}} \\
&= \sigma_s^2\left[\widehat{\mathbf{H}}_{SD}^H \quad \widehat{\mathbf{H}}_{SR}^H\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H\right]\begin{bmatrix}\left(\sigma_{n,d}^2\mathbf{I}_M + \sigma_s^2\alpha_s^2\operatorname{Tr}(\boldsymbol{\Psi}_{SD})\Sigma_{SD}\right)^{-1} & \mathbf{0} \\ \mathbf{0} & (\Delta\mathbf{A}+\mathbf{A})^{-1}\end{bmatrix}\begin{bmatrix}\widehat{\mathbf{H}}_{SD} \\ \widehat{\mathbf{H}}_{RD}\mathbf{F}_R\widehat{\mathbf{H}}_{SR}\end{bmatrix} \\
&= \sigma_s^2\left(\sigma_{n,d}^{-2}\widehat{\mathbf{H}}_{SD}^H\left(\sigma_{n,d}^2\mathbf{I}_M + \sigma_s^2\alpha_s^2\operatorname{Tr}(\boldsymbol{\Psi}_{SD})\Sigma_{SD}\right)^{-1}\widehat{\mathbf{H}}_{SD} + \widehat{\mathbf{H}}_{SR}^H\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H(\Delta\mathbf{A}+\mathbf{A})^{-1}\widehat{\mathbf{H}}_{RD}\mathbf{F}_R\widehat{\mathbf{H}}_{SR}\right)
\end{aligned}
\tag{3.23}
$$

with

$$
\begin{aligned}
\Delta\mathbf{A} = \operatorname{Tr}&\left(\mathbf{F}_R\left(\sigma_s^2\alpha_S^2\underbrace{\left(\operatorname{Tr}(\boldsymbol{\Psi}_{SR})\Sigma_{SR} + \widehat{\mathbf{H}}_{SR}\widehat{\mathbf{H}}_{SR}^H\right)}_{=\widehat{\mathbf{H}}'_{SR}\widehat{\mathbf{H}}'^H_{SR}} + \sigma_{n,r}^2\mathbf{I}_R\right)\mathbf{F}_R^H\boldsymbol{\Psi}_{RD}\right)\Sigma_{RD} \\
&+\sigma_s^2\alpha_S^2\operatorname{Tr}(\boldsymbol{\Psi}_{SR})\widehat{\mathbf{H}}_{RD}\mathbf{F}_R\Sigma_{SR}\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H
\end{aligned}
\tag{3.24}
$$

$$
\mathbf{A} = \sigma_{n,r}^2\widehat{\mathbf{H}}_{RD}\mathbf{F}_R\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H + \sigma_{n,d}^2\mathbf{I}_M \tag{3.25}
$$

We first optimize α_S by treating $\mathbf{C}$, $\mathbf{F}_R$ and $\mathbf{U}_S$ as known entities. It is simple to see that the cost function in (3.22) is monotonically decreasing with α_S (see the proof in Appendix C). So, the optimum value can be found as

$$
\alpha_{S,opt} = \sqrt{\frac{P_{S,T}}{N\sigma_s^2}} \tag{3.26}
$$

This corresponds to the largest α_S under the constraint C_1' in (3.22). Substituting (3.26) into (3.22). The subproblem can be re-written into a conventional point-to-point THP MIMO problem [24] given by

$$
\min_{\mathbf{C}(\mathbf{F}_R),\mathbf{U}_S(\mathbf{F}_R)} \operatorname{Tr}\left\{\mathbf{C}\left(\frac{P_{S,T}}{N\sigma_s^2}\mathbf{U}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{U}_S + \mathbf{I}_N\right)^{-1}\mathbf{C}^H\right\} \tag{3.27}
$$

The optimum solution of $\mathbf{C}$ has been derived in [24] (see Appendix A), denote as $\mathbf{C}_{opt}$, as :

$$\mathbf{C}_{opt} = \mathbf{D}\mathbf{L}^{-1} \tag{3.28}$$

and

$$\mathbf{L}\mathbf{L}^{H} = \left(\frac{P_{S,T}}{N\sigma_s^2}\mathbf{U}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{U}_S + \mathbf{I}_N\right)^{-1} \tag{3.29}$$

is the Cholesky factorization of $\left(\frac{P_{S,T}}{N\sigma_s^2}\mathbf{U}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{U}_S + \mathbf{I}_N\right)^{-1}$; $\mathbf{L}$ is a lower triangular matrix with real diagonal elements; $\mathbf{D}$ is a diagonal matrix scaling the diagonal elements of $\mathbf{C}$ to unity, that is $\mathbf{D} = \operatorname{diag}\left(\mathbf{L}(k,k)\right), k = 1,\ldots,N$. Substituting (3.28) and (3.29) into (3.27), we can have the cost function as:

$$\begin{aligned}&\operatorname{Tr}\left\{\mathbf{C}\left(\frac{P_{S,T}}{N\sigma_s^2}\mathbf{U}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{U}_S + \mathbf{I}_N\right)^{-1}\mathbf{C}^H\right\}\\ &= \sum_{k=1}^{N}\mathbf{L}(k,k)^2 \geq N\left(\prod_{k=1}^{N}\mathbf{L}(k,k)^2\right)^{1/N}\end{aligned} \tag{3.30}$$

which is a function of $\mathbf{U}_S$. Note that $\mathbf{L}(k,k)$ means the kth diagonal element of $\mathbf{L}$. The inequality in (3.30) is the arithmetic-geometric inequality (AGI) and the equality is held when $\mathbf{L}(i,i) = \mathbf{L}(j,j),\ \forall i \neq j$.

The next step is to find $\mathbf{U}_S$ so that the bound in (3.30) can be achieved. First, we decompose $\mathbf{U}_S$ as

$$\mathbf{U}_S = \mathbf{V}_{\tilde{\mathbf{H}}}\mathbf{U}'_S \tag{3.31}$$

where $\mathbf{V}_{\tilde{\mathbf{H}}} \in \mathbb{C}^{N\times N}$ is the right singular matrices of $\tilde{\mathbf{H}}$ and $\mathbf{U}'_S \in \mathbb{C}^{N\times N}$ is a unitary matrix to be determined later. Note that the decomposition can always be conducted for a unitary matrix. Substituting (3.31) in (3.30), we have:

$$\mathbf{L}\mathbf{L}^{H} = \left(\frac{P_{S,T}}{N\sigma_s^2}\mathbf{U}_S^H\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{U}_S + \mathbf{I}_N\right)^{-1} = \mathbf{U}'^H_S\underbrace{\left(\frac{P_{S,T}}{N\sigma_s^2}\Lambda + \mathbf{I}_N\right)^{-1}}_{:=\tilde{\mathbf{D}}}\mathbf{U}'_S \tag{3.32}$$

where $\Lambda = \operatorname{diag}\left\{\lambda_{\tilde{\mathbf{H}},1},\cdots,\lambda_{\tilde{\mathbf{H}},N}\right\}$, $\lambda_{\tilde{\mathbf{H}},k}$ are the eigenvalues of $\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}$. To achieve the equality of the AGI, we apply GMD [26] method on $\tilde{\mathbf{D}}^{1/2}$. Let $\tilde{\mathbf{D}} = \tilde{\mathbf{D}}^{1/2}\tilde{\mathbf{D}}^{1/2}$ and $\tilde{\mathbf{D}}^{1/2}$ is the square-root matrix of $\tilde{\mathbf{D}}$. Then we have:

$$\tilde{\mathbf{D}}^{1/2} = \mathbf{QRP}^{H} \tag{3.33}$$

where $\mathbf{Q}$ and $\mathbf{P}$ are unitary matrices, and $\mathbf{R}$ is an upper triangular matrix with equal and real diagonal elements. Letting $\mathbf{U}'_S = \mathbf{P}$ and substituting (3.33) into (3.32), we can have

$$\mathbf{LL}^{H} = \mathbf{R}^{H}\mathbf{R} \tag{3.34}$$

which indicates that $\mathbf{L} = \mathbf{R}^{H}$, and the diagonal elements of $\mathbf{L}$ are real and all equal. Then the equality of AGI in (3.30) is held so that the lower bound is achieved. Therefore, the optimum $\mathbf{F}_S$, denoted as $\mathbf{F}_{S,opt}$, can be expressed as

$$\mathbf{F}_{S,opt} = \sqrt{\frac{P_{S,T}}{N\sigma_s^2}}\mathbf{V}_{\tilde{\mathbf{H}}}\mathbf{P} \tag{3.35}$$

By substituting (3.28) and (3.35) into (3.15), the result MSE can be expressed as:

$$\begin{aligned} J_{\min} &= \sigma_s^2 \sum_{k=1}^{N} \mathbf{L}(k,k)^2 = \sigma_s^2 N \left(\prod_{k=1}^{N} \mathbf{L}(k,k)^2 \right)^{\frac{1}{N}} \\ &= \sigma_s^2 N \prod_{k=1}^{N} \left(\frac{P_{S,T}}{N\sigma_s^2} \lambda_{\tilde{\mathbf{H}},k} + 1 \right)^{-1/N} \end{aligned} \tag{3.36}$$

Now, the problem becomes the minimization of (3.36), which is master problem. It is seen that the cost function now is a function of $\mathbf{F}_R$.

To proceed, we consider the following equivalence:

$$\prod_{k=1}^{N} \left(\frac{P_{S,T}}{N\sigma_s^2} \lambda_{\tilde{\mathbf{H}},k} + 1 \right) = \det\left(\left(\mathbf{I}_N + \frac{P_{S,T}}{N\sigma_s^2} \tilde{\mathbf{H}}^{H}\tilde{\mathbf{H}} \right) \right) \tag{3.37}$$

Substituting (3.37) into (3.36) , we can reformulate the problem as

$$\begin{aligned} &\max_{\mathbf{F}_R} \ \det\left(\mathbf{I}_N + \frac{P_{S,T}}{N\sigma_s^2} \tilde{\mathbf{H}}^{H}\tilde{\mathbf{H}} \right) \\ &\text{s.t.} \\ &C_2 : \mathrm{Tr}\left\{ \mathbf{F}_R \left(\sigma_{n,r}^2 \mathbf{I}_R + \frac{P_{S,T}}{N} \hat{\mathbf{H}}'_{SR} \hat{\mathbf{H}}'^{H}_{SR} \right) \mathbf{F}_R^{H} \right\} \le P_{R,T}. \end{aligned} \tag{3.38}$$

Since the structure of $\tilde{\mathbf{H}}^{H}\tilde{\mathbf{H}}$ is still complicate and difficult to deal with, we find another equivalent form as follows:

$$\frac{P_{S,T}}{N\sigma_s^2} \tilde{\mathbf{H}}^{H}\tilde{\mathbf{H}} = \hat{\mathbf{H}}''^{H}_{SR} \mathbf{F}_R^{H} \hat{\mathbf{H}}_{RD}^{H} (\Delta\mathbf{A} + \mathbf{A})^{-1} \hat{\mathbf{H}}_{RD} \mathbf{F}_R \hat{\mathbf{H}}''_{SR} \tag{3.39}$$

where

$$\hat{\mathbf{H}}''_{SR} = \sqrt{\frac{P_{S,T}}{N}}\hat{\mathbf{H}}_{SR}\left(\mathbf{I}_N + \frac{P_{S,T}}{N\sigma_{n,d}^2}\hat{\mathbf{H}}_{SD}^H\left(\sigma_{n,d}^2\mathbf{I}_M + \frac{P_{S,T}}{N}\mathrm{Tr}\left(\mathbf{\Psi}_{SD}\right)\mathbf{\Sigma}_{SD}\right)^{-1}\hat{\mathbf{H}}_{SD}\right)^{-1/2} \tag{3.40}$$

Then (3.38) can be rewritten as

$$\begin{aligned} &\max_{\mathbf{F}_R}\ \det\left(\mathbf{I}_N + \hat{\mathbf{H}}''^H_{SR}\mathbf{F}_R^H\hat{\mathbf{H}}_{RD}^H(\Delta\mathbf{A}+\mathbf{A})^{-1}\hat{\mathbf{H}}_{RD}\mathbf{F}_R\hat{\mathbf{H}}''_{SR}\right) \\ &\text{s.t.}\quad C_2. \end{aligned} \tag{3.41}$$

Taking a close look at (3.41), we can find that $(\Delta\mathbf{A}+\mathbf{A})$ is a complicated function of $\mathbf{F}_R$. The master problem is difficult to solve, even using the numerical method [42]. To overcome the problem, we propose maximizing a lower bound of (3.41) instead of trying to maximize it directly. In this manner, we can have an amenable form and the optimization problem in the master optimization can be made easier. For this purpose, we consider the following property.

Property: The utility function in (3.38) is lower bounded by

$$\begin{aligned} &\det\left(\mathbf{I}_N + \hat{\mathbf{H}}''^H_{SR}\mathbf{F}_R^H\hat{\mathbf{H}}_{RD}^H(\Delta\mathbf{A}+\mathbf{A})^{-1}\hat{\mathbf{H}}_{RD}\mathbf{F}_R\hat{\mathbf{H}}''_{SR}\right) \\ &\geq \det\left(\mathbf{I}_N + \hat{\mathbf{H}}''^H_{SR}\mathbf{F}_R^H\hat{\mathbf{H}}_{RD}^H(\Delta\mathbf{A}'+\mathbf{A})^{-1}\hat{\mathbf{H}}_{RD}\mathbf{F}_R\hat{\mathbf{H}}''_{SR}\right) \end{aligned} \tag{3.42}$$

where

$$\begin{aligned} \Delta\mathbf{A} = &\,\mathrm{Tr}\left(\mathbf{F}_R\left(\sigma_s^2\alpha_S^2\underbrace{\left(\mathrm{Tr}\left(\mathbf{\Psi}_{SR}\right)\mathbf{\Sigma}_{SR} + \hat{\mathbf{H}}_{SR}\hat{\mathbf{H}}_{SR}^H\right)}_{=\hat{\mathbf{H}}'_{SR}\hat{\mathbf{H}}'^H_{SR}} + \sigma_{n,r}^2\mathbf{I}_R\right)\mathbf{F}_R^H\mathbf{\Psi}_{RD}\right)\mathbf{\Sigma}_{RD} \\ &+\sigma_s^2\alpha_S^2\mathrm{Tr}\left(\mathbf{\Psi}_{SR}\right)\hat{\mathbf{H}}_{RD}\mathbf{F}_R\mathbf{\Sigma}_{SR}\mathbf{F}_R^H\hat{\mathbf{H}}_{RD}^H \end{aligned} \tag{3.43}$$

$$\Delta\mathbf{A}' = P_{R,T}\lambda_{\max}\left(\mathbf{\Psi}_{RD}\right)\lambda_{\max}\left(\mathbf{\Sigma}_{RD}\right)\mathbf{I}_M + \frac{P_{S,T}}{N}\mathrm{Tr}\left(\mathbf{\Psi}_{SR}\right)\lambda_{\max}(\mathbf{\Sigma}_{SR})\hat{\mathbf{H}}_{RD}\mathbf{F}_R\mathbf{F}_R^H\hat{\mathbf{H}}_{RD}^H \tag{3.44}$$

the equality of (3.42) is held when $\mathbf{\Psi}_{RD} = \beta_{RD}\mathbf{I}_R$, $\mathbf{\Sigma}_{RD} = \gamma_{RD}\mathbf{I}_{M\times M}$, $\mathbf{\Psi}_{SR} = \beta_{SR}\mathbf{I}_N$ and $\mathbf{\Sigma}_{SR} = \gamma_{SR}\mathbf{I}_{R\times R}$ with some scalars β_{RD}, β_{SR}, γ_{RD} and γ_{SR}. As we will see, the optimization with the lower bound is much easer. The derivation of the lower bound in (3.42) can be found in Appendix D.

Using the lower bound of the utility function, we can reformulate the master optimization problem as:

$$
\begin{aligned}
&\max_{\mathbf{F}_R}\ \det\left(\mathbf{I}_N+\widehat{\mathbf{H}}''^{H}_{SR}\mathbf{F}_R^H\widehat{\mathbf{H}}^H_{RD}\left(\Delta\mathbf{A}'+\mathbf{A}\right)^{-1}\widehat{\mathbf{H}}_{RD}\mathbf{F}_R\widehat{\mathbf{H}}''_{SR}\right)\\
&\text{s.t.}\\
&\left(\Delta\mathbf{A}'+\mathbf{A}\right)=\alpha\mathbf{I}_M+\beta\widehat{\mathbf{H}}_{RD}\mathbf{F}_R\mathbf{F}_R^H\widehat{\mathbf{H}}^H_{RD},\\
&C_2:\operatorname{tr}\left\{\mathbf{F}_R\left(\sigma^2_{n,r}\mathbf{I}_R+\frac{P_{S,T}}{N}\widehat{\mathbf{H}}'_{SR}\widehat{\mathbf{H}}'^{H}_{SR}\right)\mathbf{F}_R^H\right\}\le P_{R,T}.
\end{aligned}
\tag{3.45}
$$

where

$$\alpha=P_{R,T}\lambda_{\max}\left(\Psi_{RD}\right)\lambda_{\max}\left(\Sigma_{RD}\right)+\sigma^2_{n,d} \tag{3.46}$$

$$\beta=\frac{P_{S,T}}{N}\operatorname{Tr}\left(\Psi_{SR}\right)\lambda_{\max}\left(\Sigma_{SR}\right)+\sigma^2_{n,r} \tag{3.47}$$

As we will see in latter development, $\Delta\mathbf{A}'$ in (3.45) is easier to handle compared to $\Delta\mathbf{A}$. Although the function in (3.45) is simplified, it is still a complicated function of $\mathbf{F}_R$. We now let $\mathbf{F}_R$ have a specific structure such that the master optimization problem can be easier to solve. This relay precoder structure can transfer the matrix-valued optimization problem in (3.45) to a scalar-valued problem, though it is suboptimal.

Similar the procedure in Chapter 2, we first use Lemma 2.2 and Lemma 2.3. Consider the following singular-value-decomposition (SVD):

$$\widehat{\mathbf{H}}_{RD}=\widehat{\mathbf{U}}_{rd}\widehat{\Sigma}_{rd}\widehat{\mathbf{V}}^H_{rd} \tag{3.48}$$

$$\widehat{\mathbf{H}}''_{SR}=\mathbf{U}''_{sr}\Sigma''_{sr}\mathbf{V}''^{H}_{sr} \tag{3.49}$$

where $\widehat{\mathbf{U}}_{rd}\in\mathbb{C}^{M\times M}$ and $\mathbf{U}''_{sr}\in\mathbb{C}^{R\times R}$ are left singular matrices of $\widehat{\mathbf{H}}_{RD}$ and $\widehat{\mathbf{H}}''_{SR}$, respectively; $\Sigma_{rd}\in\mathbb{R}^{M\times R}$ and $\Sigma''_{sr}\in\mathbb{R}^{R\times N}$ are the diagonal singular value matrices of $\mathbf{H}_{RD}$ and $\widehat{\mathbf{H}}''_{SR}$, respectively; $\mathbf{V}^H_{rd}\in\mathbb{C}^{R\times R}$ and $\mathbf{V}''^{H}_{sr}\in\mathbb{C}^{N\times N}$ are the right singular matrices of $\mathbf{H}_{RD}$ and $\widehat{\mathbf{H}}''_{SR}$, respectively. Substitute (3.48) and (3.49) into (3.45), we can rewrite (3.45) as

$$
\begin{aligned}
&\det\left(\mathbf{M}'\right)\\
&\text{s.t.}\quad C_2\\
&\mathbf{M}'=\mathbf{I}_N+\Sigma''_{sr}\mathbf{U}''_{sr}\mathbf{F}_R^H\widehat{\mathbf{V}}^H_{rd}\widehat{\Sigma}_{rd}\left(\alpha\mathbf{I}_M+\beta\widehat{\Sigma}_{rd}\widehat{\mathbf{V}}^H_{rd}\mathbf{F}_R\mathbf{F}_R^H\widehat{\mathbf{V}}_{rd}\widehat{\Sigma}_{rd}\right)^{-1}\widehat{\Sigma}_{rd}\widehat{\mathbf{V}}_{rd}\mathbf{F}_R\mathbf{U}''^{H}_{sr}\Sigma''_{sr}
\end{aligned}
\tag{3.50}
$$

From Lemma 2.2, we see that if $\mathbf{M}'$ is diagonalized, the utility function in (3.50) can be maximized. Note here that the result is not held when the power constraint in (3.50) is included. However, we use it to obtain a suboptimal solution. Let $\mathbf{F}_R$ have a structure shown below:

$$\mathbf{F}_{R,opt} = \widehat{\mathbf{V}}_{rd}\boldsymbol{\Sigma}_r\mathbf{U}''^{H}_{sr} \tag{3.51}$$

where Σ_r is a diagonal matrix with ith diagonal element $\sigma_{r,i}$, $i=1,\cdots,\kappa$, $\kappa=\min\{N,R\}$ whose value will be determined. The general structure for the relay precoder should have the form of $\left\{\mathbf{F}_R \middle| \mathbf{F}_R = \mathbf{U}_r\boldsymbol{\Sigma}_r\mathbf{V}_r^H,\ \mathbf{F}_R \in C_2\right\}$ with $\mathbf{U}_r$ is a $M\times M$ unitary matrix and $\mathbf{V}_r$ is a $R\times R$ unitary matrix. However, the optimum $\mathbf{F}_R$ in (3.50) is very difficult to find. Here, we only consider a specific feasible set of $\mathbf{F}_R$ in (3.51) simplifying the optimization problem.

The next step is to transfer the matrix-valued problem to the scalar-valued one with the precoder structure of $\mathbf{F}_R$ in (3.51). Let $\sigma_{rd,i}$ and $\sigma''_{sr,i}$ be the ith diagonal element of $\widehat{\boldsymbol{\Sigma}}_{rd}$ and $\boldsymbol{\Sigma}''_{sr}$, respectively. Substituting (3.51) into (3.50) and taking *ln* operation on the utility function, we can rewrite (3.50) as:

$$\begin{aligned}
&\max_{p_{r,i},1\le i\le\kappa}\sum_{i=1}^{\kappa}\ln\left(1+\frac{p_{r,i}\sigma''^2_{sr,i}\sigma^2_{rd,i}}{\alpha+\beta p_{r,i}\sigma^2_{rd,i}}\right)\\
&\text{s.t.}\\
&\sum_{i=1}^{\kappa}p_{r,i}\left(\sigma^2_{n,r}+\mathbf{D}_{sr}(i,i)\right)\le P_{R,T},\\
&p_{r,i}\ge 0,
\end{aligned} \tag{3.52}$$

where $p_{r,i}=\sigma^2_{r,i}$ and $\mathbf{D}_{sr}=\frac{P_{S,T}}{N}\mathbf{U}''^{H}_{sr}\left(\widehat{\mathbf{H}}_{SR}\widehat{\mathbf{H}}^{H}_{SR}+\mathrm{Tr}\left(\boldsymbol{\Psi}_{SR}\right)\boldsymbol{\Sigma}_{SR}\right)\mathbf{U}''_{sr}$ with $\mathbf{D}_{sr}(i,i)$ being the ith diagonal element of $\mathbf{D}_{sr}$. Now, the utility function and the constraints are all functions of scalars. Since the utility function and the constraints are all concave for $p_{r,i}\ge 0$ [42], we see that (3.52) is a standard concave optimization problem. As a result, we can find the optimum solution of $p_{r,i}$, $i=1,\cdots,\kappa$ by Karush-Kuhn-Tucker (KKT) conditions. The solution is given by:

$$p_{r,i}=\left[\sqrt{\frac{\mu}{\left(\sigma^2_{n,r}+\mathbf{D}_{sr}(i,i)\right)\sigma^2_{rd,i}\cdot\frac{\beta}{\alpha}\left(\frac{\beta}{\sigma''^2_{sr,i}}+1\right)}+\frac{1/4}{\sigma^4_{rd,i}\frac{\beta^2}{\alpha^2}\left(\frac{\beta}{\sigma''^2_{sr,i}}+1\right)^2}}-\frac{\alpha+\frac{\alpha\sigma''^2_{sr,i}}{2\beta}}{\sigma^2_{rd,i}\left(\beta+\sigma''^2_{sr,i}\right)}\right]^+ \tag{3.53}$$

where μ is chosen to satisfy the power constraint in (3.52). The detail derivation can be found in Appendix E. Substituting (3.53) into (3.51), we can have the optimum relay precoder $\mathbf{F}_R$. After the optimum $\mathbf{F}_R$ is found, we can obtain $\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}$ by (3.23). Subsequently, $\mathbf{F}_S$ and $\mathbf{C}$ can also be obtained by (3.35) and (3.28) with the same procedure described in Chapter 2.

In this chapter, we joint design the robust transceiver in an AF MIMO relay system in which a THP is used at the source, a linear precoder at the relay, and an MMSE receiver at the destination. Since the channel uncertainty has been taken into consideration, we can expect that the design will outperform that in Chapter 2. The price we pay is a more complicated design. The computational complexity of the non-robust/robust designs includes SVD, GMD, and matrix inversion operations, are mentioned in Chapter 2 and Chapter 3. The overall computational complexity and steps of the non-robust/robust designs, measured by FLOPs, are summarized in Table 3.1 and Table 3.2.

Table 3.1: Computational complexity of THP source and linear relay precoders (MMSE receiver)

Step	Operation	FLOPs
1	$\mathbf{H}'_{SR}$	$O\left(N^2(R+M)\right)$
2	SVD $\mathbf{H}_{RD}=\mathbf{U}_{rd}\boldsymbol{\Sigma}_{rd}\mathbf{V}_{rd}^H$	$O\left(MR^2+R^3\right)$
3	SVD $\mathbf{H}'_{SR}=\mathbf{U}'_{sr}\boldsymbol{\Sigma}'_{sr}\mathbf{V}'^H_{sr}$	$O\left(RN^2+N^3\right)$
4	$\boldsymbol{\Sigma}_r$	$O\left(\kappa I_r\right)$
5	$\mathbf{F}_R$	$O\left(R^2\right)$
6	$\tilde{\mathbf{H}}=\mathbf{R}_n^{-\frac{1}{2}}\mathbf{H}$	$O\left(M^2N\right)$
7	SVD $\tilde{\mathbf{H}}$	$O\left(MN^2+N^3\right)$
8	GMD $\tilde{\mathbf{D}}^{\frac{1}{2}}=\mathbf{QRP}^H$ $\mathbf{L}=\mathbf{R}^H$	$O\left(N^3\right)$
9	$\mathbf{C}_{opt}$	$O\left(N^2\right)$
10	$\mathbf{F}_{S,opt}$	$O\left(N^3\right)$
I_r denotes the iteration number of the water-filling process for solving $\sigma_{r,i}$		

Table 3.2: Computational complexity of robust THP source and linear relay precoders (MMSE receiver)

step	operation	FLOPs
1	$\hat{\mathbf{H}}''_{SR}$	$O\left(NM^2+(R+M)N^2\right)$
2	SVD $\hat{\mathbf{H}}_{RD}$	$O\left(MR^2+R^3\right)$
3	SVD $\hat{\mathbf{H}}''_{SR}$	$O\left(RN^2+N^3\right)$
4	$\boldsymbol{\Sigma}_r$	$O\left(\kappa I_r\right)$
5	$\mathbf{F}_R$	$O\left(R^2\right)$
6	$\mathbf{R}_\Delta$	$O\left(N^3+RN^2+R^3+MR^2\right)$
7	$\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}$	$O\left(M^3N\right)$
8	SVD $\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}$	$O\left(N^3\right)$
9	GMD $\tilde{\mathbf{D}}^{1/2}=\mathbf{QRP}^H$ $\mathbf{L}=\mathbf{R}^H$	$O\left(N^3\right)$
10	$\mathbf{C}_{opt}$	$O\left(N^2\right)$
11	$\mathbf{F}_{S,opt}$	$O\left(N^3\right)$
I_r denotes the iteration number of the water-filling process for solving $\sigma_{r,i}$		

Chapter 4

Simulation results and discussions

4.1 Simulation Setup

In this section, we describe our simulation environment. We consider an AF MIMO relay system with *N*, *R* and *M* antennas at the source, the relay and the destination, respectively. We let $N = R = M = 4$. The widely used exponential model [29] is chosen for the generation of the channel estimation error covariance matrices, which can be represented by

$$\boldsymbol{\Psi}_{SR} = \boldsymbol{\Psi}_{RD} = \boldsymbol{\Psi}_{SD} = \begin{bmatrix} 1 & \delta & \delta^2 & \delta^3 \\ \delta & 1 & \delta & \delta^2 \\ \delta^2 & \delta & 1 & \delta \\ \delta^3 & \delta^2 & \delta & 1 \end{bmatrix} \tag{4.1}$$

and

$$\boldsymbol{\Sigma}_{SR} = \boldsymbol{\Sigma}_{RD} = \boldsymbol{\Sigma}_{SD} = \sigma_e^2 \begin{bmatrix} 1 & \gamma & \gamma^2 & \gamma^3 \\ \gamma & 1 & \gamma & \gamma^2 \\ \gamma^2 & \gamma & 1 & \gamma \\ \gamma^3 & \gamma^2 & \gamma & 1 \end{bmatrix} \tag{4.2}$$

where δ and γ are the correlation coefficients of the row and column covariance matrices, σ_e^2 denotes the estimation error variance. The resultant covariance matrices can be obtained from the channel estimation method proposed in [29]. The estimate channels, $\widehat{\mathbf{H}}_{SR}$, $\widehat{\mathbf{H}}_{RD}$ and $\widehat{\mathbf{H}}_{SD}$, are generated base on the following distributions

$$vec\left(\widehat{\mathbf{H}}_{SR}\right) \sim CN\left(\mathbf{0}_{NR\times 1}, \frac{1-\sigma_e^2}{\sigma_e^2}\boldsymbol{\Sigma}_{SR} \otimes \boldsymbol{\Psi}_{SR}^T\right) \tag{4.3}$$

$$vec\left(\widehat{\mathbf{H}}_{RD}\right) \sim CN\left(\mathbf{0}_{MR\times 1}, \frac{1-\sigma_e^2}{\sigma_e^2}\boldsymbol{\Sigma}_{RD} \otimes \boldsymbol{\Psi}_{RD}^T\right) \tag{4.4}$$

$$vec\left(\widehat{\mathbf{H}}_{SD}\right) \sim CN\left(\mathbf{0}_{MN\times 1}, \frac{1-\sigma_e^2}{\sigma_e^2}\boldsymbol{\Sigma}_{SD} \otimes \boldsymbol{\Psi}_{SD}^T\right) \tag{4.5}$$

So that the relationships of the actual and estimation channels can be expressed as:

$\mathbf{H}_{SR} = \widehat{\mathbf{H}}_{SR} + \Delta\mathbf{H}_{SR}$, $\mathbf{H}_{RD} = \widehat{\mathbf{H}}_{RD} + \Delta\mathbf{H}_{RD}$ and $\mathbf{H}_{SD} = \widehat{\mathbf{H}}_{SD} + \Delta\mathbf{H}_{SD}$. Note that $\mathbf{H}_{SR}$, $\mathbf{H}_{RD}$ and $\mathbf{H}_{SD}$ have a unit variance for each element. Let SNR_{sr}, SNR_{sd} and SNR_{rd} denote the received signal-to-noise ratio (SNR) at each relay antenna in the first phase, that at each destination antenna in the first phase, and that at each destination antenna in the second phase, respectively. As defined in Chapter 2 and Chapter 3, the definition of SNR of each link can be represented as:

$$\begin{aligned} \mathrm{SNR}_{sr} &= P_{S,T} \big/ \mathrm{tr}\left(\mathbf{R}_{nR}\right) \\ \mathrm{SNR}_{rd} &= P_{R,T} \big/ \mathrm{tr}\left(\mathbf{R}_{nD,2}\right) \\ \mathrm{SNR}_{sd} &= P_{S,T} \big/ \mathrm{tr}\left(\mathbf{R}_{nD,1}\right) \end{aligned} \tag{4.6}$$

where $P_{S,T}$ and $P_{R,T}$ are the total power constrained at the source and the relay, $\mathbf{R}_{nR}$ is the covariance matrix of the noise vector at the relay antennas, and $\mathbf{R}_{nD,1}$ and $\mathbf{R}_{nD,2}$ are the covariance matrices of the noise vector at the destination antennas for first phase and second phase transmission, respectively.

Without of generality, we use the 16-QAM scheme for each transmission symbol stream. The data symbols are assumed to be independently transmitted form the four antennas with the same power. We regard the channel estimation error as the channel uncertainty, generated based on (3.4). We assume that the destination has the perfect knowledge of channel statistics such that the precoders can be calculated there.

4.2 Simulation results and discussions

In this section, simulation results are reported demonstrating the effectiveness of the proposed scheme. In the first set of simulations, we let $\mathrm{SNR}_{sr} = 30$ dB, $\mathrm{SNR}_{sd} = 15$ dB and SNR_{rd} be varied. We also let $\delta = \gamma = 0$ and $\sigma_e^2 = 0$, which means perfect CSIs are available. Figure 4.1 and Figure 4.2 compare the MSE and BER performances for (a) an un-precoded system (U-U) (b) the relay precoded system [15] (U-L) (c) the linear source and linear relay precoded system [18] (L-L) (d) the TH source and linear relay precoded system (TH-L) and (e) the proposed robust TH source and linear relay precoded system (TH-L robust). All the systems considered, (a)-(e), use the MMSE receiver. The notation "SNR (dB)" in the figures is the average signal-to-noise ratio of the relay-to-destination link in dB scale. For fairly comparison, we include the direct link in (c) although the original relay precoded scheme in [15] only considers the relay link. As shown in Figure 4.1 and Figure 4.2, we can see that the relay precoded system outperforms the un-precoded system. However, it is inferior to the linear source and linear relay precoded system. Since the nonlinear TH source and linear relay precoded system are considered in (d) and (e), both systems outperform the other systems. We can observe that the performance of the proposed robust TH source and linear relay precoded system (e) is the same with the non-robust design system (d). This indicates that the proposed robust precoded system is a generalization of the non-robust precoded system. When perfect CSIs are available, the proposed robust system is degenerated to the non-robust one.

In the second set of simulations, we let $\sigma_e^2 = 0.003$ and the other parameters be the same as those in the first set of simulations. Figure 4.3 and Figure 4.4 show the MSE and BER performances for the imperfect CSI case. From the figures, we can also observe that the un-precoded system is still inferior to the precoded systems. The nonlinear source precoded systems are superior to the linear ones. Since the TH-L-robust take channel uncertainties into consideration, it outperforms TH-L. Note that the performance of the non-robust precoded system slightly is degraded at the high SNR region. This is because when SNR is high, the interference caused by channel uncertainties offset the noise effect and dominate the overall performance.

In the third set of simulations, we compare the performance of the proposed non-linear source and linear relay robust system and the existed linear robust relay precoded system (U-L-robust) [36]. Figure 4.5 and Figure 4.6 show the MSE and BER performances for $\delta = \gamma = 0$ and $\sigma_e^2 = 0$ ($\sigma_e^2 = 0.003$). As we can see, both U-L-robust and TH-L-robust are degenerated to U-L and TH-L, respectively in this case. In imperfect CSIs environments, the performance of U-L-robust is superior to that of U-L. However, since it only considers a relay precoder, its performance is still inferior to the proposed TH-L-robust.

In the fourth set of simulations, we show the MSE performance of the proposed method with various correlation covariance parameters. Here, we let $\gamma = 0$, $\sigma_e^2 = 0.002$ and δ be varied. Also let $\mathrm{SNR}_{sr} = 30$ dB, $\mathrm{SNR}_{sd} = 15$ dB and SNR_{rd} be varied. Figure 4.7 shows the simulation result. From Figure 4.7, we can see that the performance of the proposed method is increased as the decrease of δ. As δ becomes small, the covariance matrices Σ_{SR}, Σ_{RD} and Σ_{SD} will approach to $\sigma_e^2\mathbf{I}$. This matches our expectation that the performance of the proposed method depends on the quality of channel estimation.

In the last set of simulations, we evaluate the MSE performance of TH-L and TH-L-robust under the scenario that $\delta = \gamma = 0$ and σ_e^2 is varied. The results are shown in

Figure 4.8. The SNR of each link in the system is the same as the previous case. As we can see from Figure 4.8, the performance of TH-L and TH-robust both are significantly degraded when σ_e^2 becomes large. The performance gap between TH-L and TH-L-robust increases when the CSI uncertainty increases.

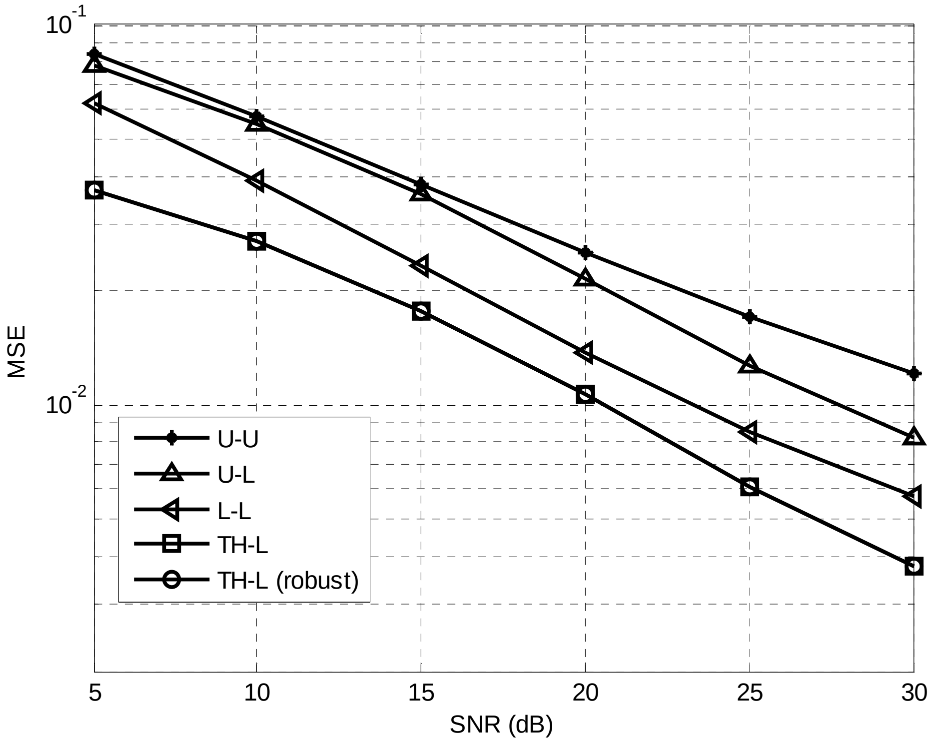

Figure 4.1: MSE performance comparison for existing precoded systems and proposed robust/non-robust TH source and linear relay precoded system. (All with MMSE receiver.)

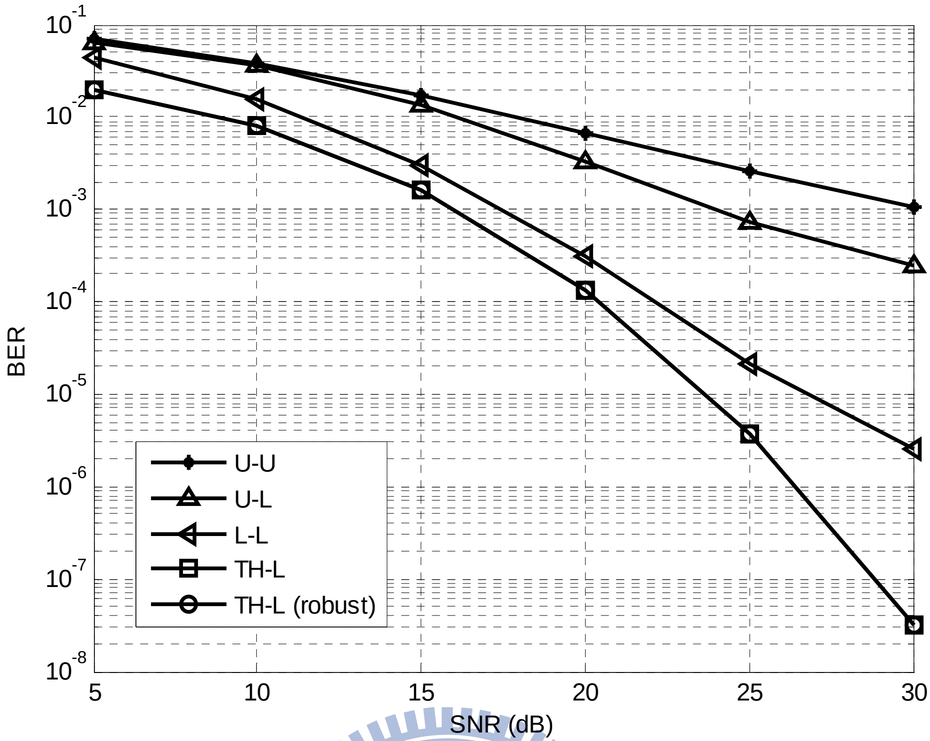

Figure 4.2: BER performance comparison for existing precoded systems and proposed robust/non-robust TH source and linear relay precoded system. (All with MMSE receiver.)

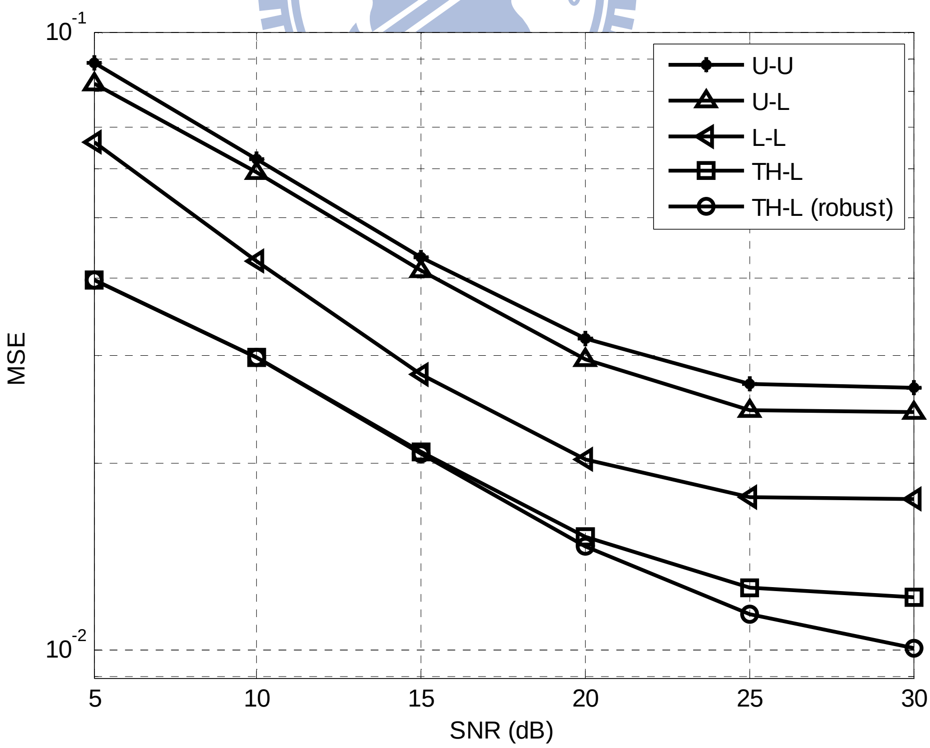

Figure 4.3: MSE performance comparison for existing precoded systems and proposed robust/non-robust TH source and linear relay precoded system. ($\delta = \gamma = 0 , \sigma_e^2 = 0.003$)

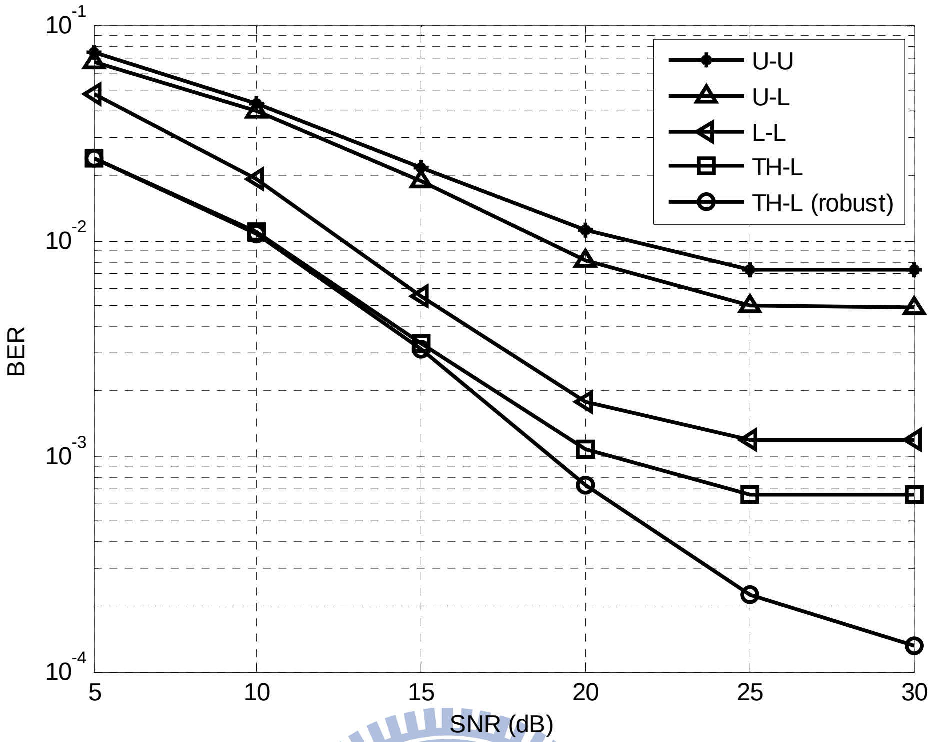

Figure 4.4: BER performance comparison for existing precoded systems and proposed robust/non-robust TH source and linear relay precoded system. ($\delta=\gamma=0,\sigma_e^2=0.003$)

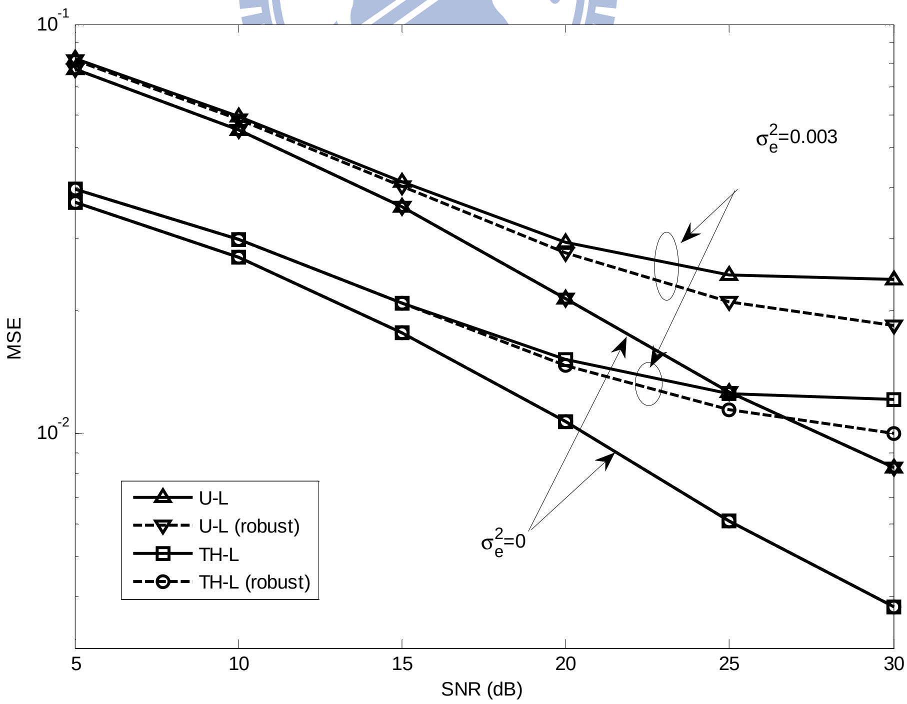

Figure 4.5: MSE performance comparison for existing robust/non-robust relay precoded systems and proposed robust/non-robust TH source and linear relay precoded system. ($\delta=\gamma=0,\sigma_e^2=0/\sigma_e^2=0.003$)

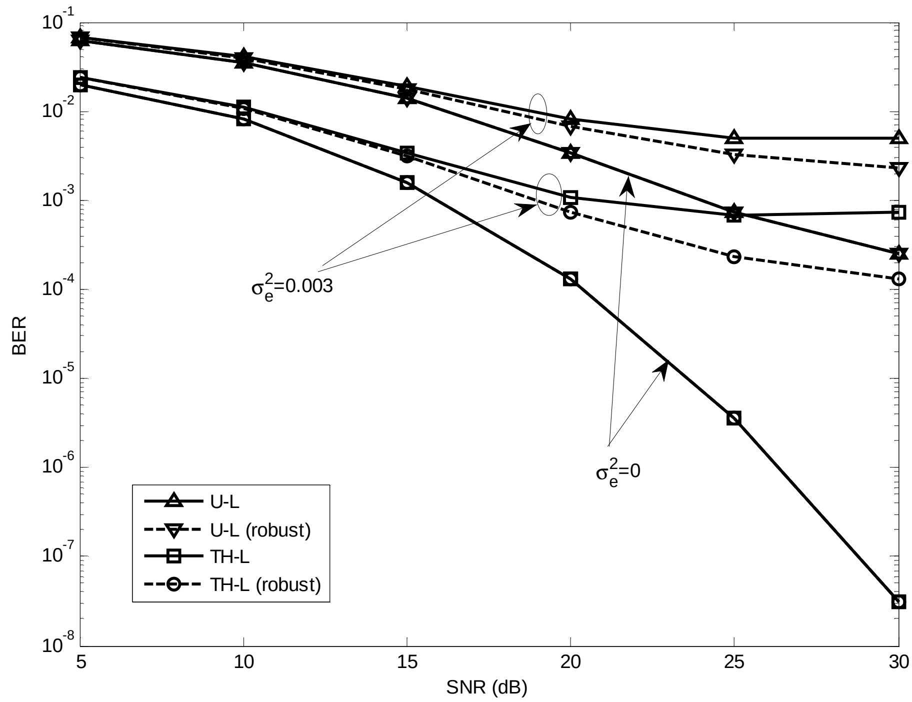

Figure 4.6: BER performance comparison for existing robust/non-robust relay precoded systems and proposed robust/non-robust TH source and linear relay precoded system. ($\delta = \gamma = 0$, $\sigma_e^2 = 0 / \sigma_e^2 = 0.003$)

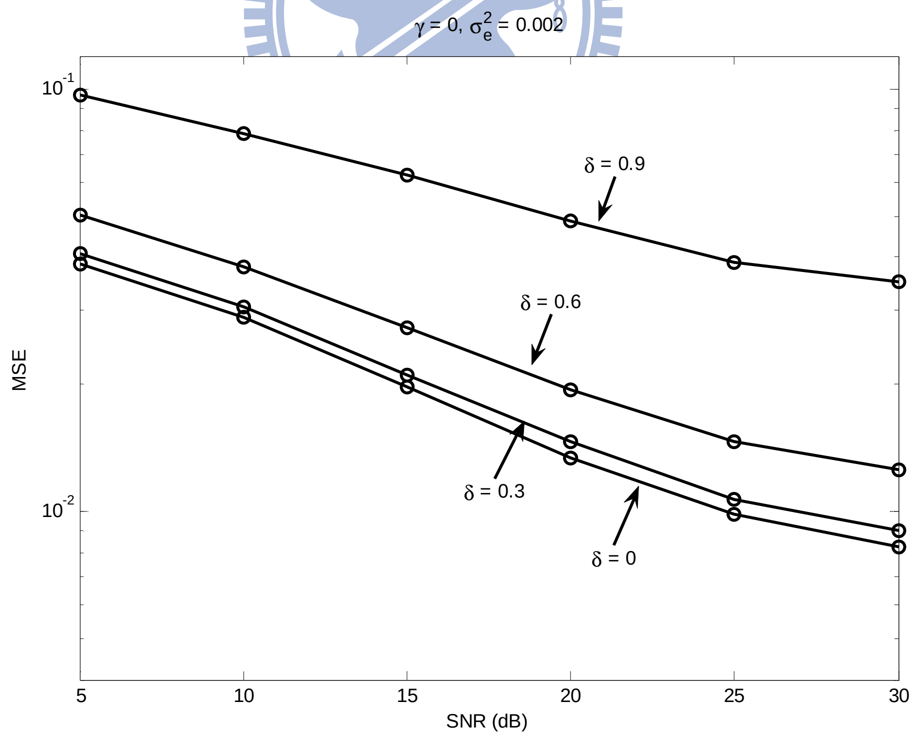

Figure 4.7: MSE comparison for proposed robust precoded system with different δ . ($\gamma = 0$, $\sigma_e^2 = 0.002$)

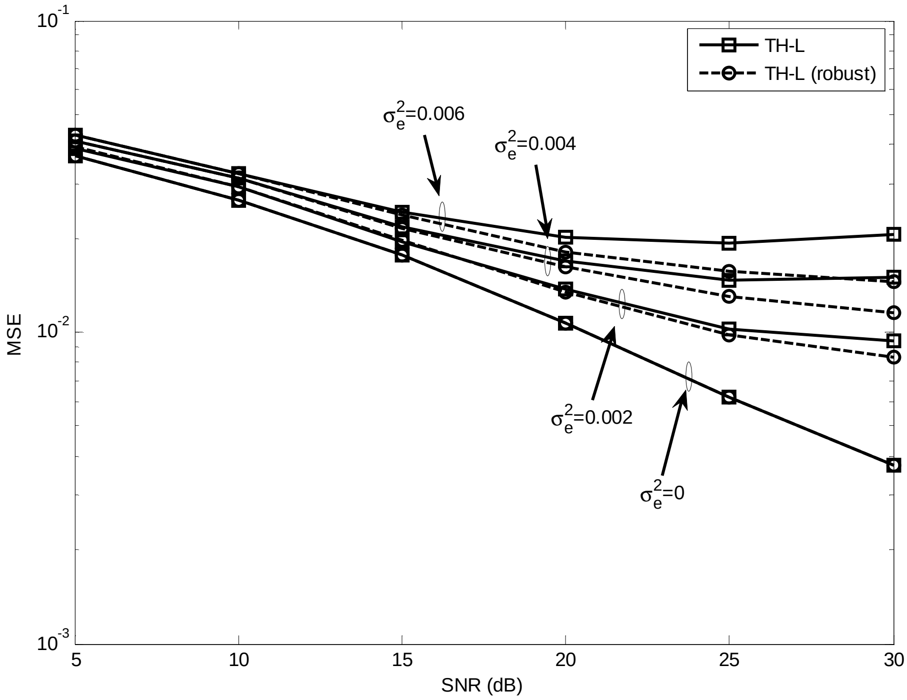

Figure 4.8: MSE comparison for proposed robust precoded system with different σ_e^2. ($\delta = \gamma = 0$)

Chapter 5
Conclusions

Many transceiver designs in three-node MIMO AF relay systems only consider the relay precoder. Also, the direct link is often ignored. As a result, the resource provided by the channel is not fully explored. This motivates us to study the transceiver designs such that both the source and relay pecoders are considered and both and the direct and relay links are taken into account. In this thesis, we consider a three-node MIMO AF relay system with a THP at the source, a linear precoder at the relay, and a MMSE receiver at the receiver. We employ the primal decomposition method, transferring the problem into a subproblem and master problem, to solve the design problem. Using a special precoder structure, we are able to obtain a suboptimum solution in closed-form. In real-world applications, however, perfect CSIs, required for the precoders design, may not be available. So, we step forward to consider a robust design. We take the CSI uncertainties into consideration and formulate the optimization problem. Similar to the non-robust design, we use the primal decomposition approach, transferring the problem into a subproblem and master problem. To facilitate the development, we derive a lower bound for the utility function of the optimization. It is shown that the tightness of the bound depends on the channel correlation between the transmit and receive antennas. If no correlation exists, the bound will be equal to the utility function itself. Finally, using the same precoder structure as that in the non-robust design, a closed-form solution is obtained. Simulation results shows that the proposed robust precoded system outperform existing un-precoded and relay precoded systems no matter perfect CSIs are available or not. In concluding the thesis, we suggest some possible topics for future research.

1. In this thesis, we consider a three-node AF MIMO system with direct link included. In the design procedure, the estimation of CSIs is less addressed. The design of training sequences or pilots for effective channel estimation is important in real-world applications.
2. In the precoded system, the estimated CSIs or computed precoders have to be fed back to the source and relay node. How to design efficient feedback systems deserves further studies.
3. In this thesis, we only study a typical three-node MIMO relay system with one

source node, one relay node and one destination node. In a general relay system, there may have multiple source, relay and destination nodes. The precoders design in such a system is challenge and need for further study.

4. The relay precoder in our consideration is linear. How to design a nonlinear source and nonlinear relay precoders is still an open problem.

Appendix

Appendix A : Optimal feed-back matrix B [37]

We try to find the optimum $\mathbf{C}$ to minimize $\mathrm{tr}(\mathbf{E})$, described in (2.17)-(2.20), which is a function of the precoding matrices. This can be formulated to minimize $\mathrm{tr}(\mathbf{E}) = \mathrm{tr}(\mathbf{CMC}^H)$ subject to $\mathbf{C}$ being a unit diagonal lower triangular matrix. Using Cholesky factorization, we have

$$\mathbf{M} = \mathbf{LL}^H \tag{A.1}$$

where $\mathbf{L}$ is a real diagonal lower triangular matrix. Write the cost function as $\mathrm{tr}(\mathbf{CMC}^H) = \mathrm{tr}(\mathbf{CLL}^H\mathbf{C}^H) = \|\mathbf{CL}\|_F^2$ and note that $\mathbf{CL}$ is a positive definite lower triangular matrix. Let $\lambda_1 \geq \cdots \geq \lambda_K$ and $\sigma_1 \geq \cdots \geq \sigma_K$ denote the eigenvalues and singular values of $\mathbf{CL}$. Then we can obtain the following lower bound by applying Weyl's inequality [40]:

$$\begin{aligned} \|\mathbf{CL}\|_F^2 &= \sum_{i=1}^{K} \sigma_i^2 \geq \sum_{i=1}^{K} \lambda_i^2 = \sum_{i=1}^{K} [\mathbf{CL}]_{ii}^2 \\ &= \sum_{i=1}^{K} \mathbf{L}_{ii}^2 \end{aligned} \tag{A.2}$$

We find that the term in the right hand side of (A.2) is a bound of $\|\mathbf{CL}\|_F^2$ and it is dependent on $\mathbf{C}$. The equality of (A.2) is held when $\mathbf{CL}$ is normal. If a lower triangular matrix is normal, it must be a diagonal matrix [40]. Therefore, we can choose a $\mathbf{C}$ to achieve the bound in (A.2). So, we have:

$$\mathbf{C} = \mathrm{diag}\{\mathbf{L}_{11}, \ldots, \mathbf{L}_{KK}\} \cdot \mathbf{L}^{-1} \tag{A.3}$$

Using the optimum $\mathbf{C}$, we can rewrite the MSE matrix $\mathbf{E}$ as:

$$\mathbf{E} = \mathrm{diag}(\mathbf{L}_{11}^2, \ldots, \mathbf{L}_{KK}^2) \tag{A.4}$$

So, the optimum $\mathbf{C}$ will yield a diagonal MSE matrix, and it also minimizes the trace of the MSE matrix.

Appendix B: MSE in (3.7)

In this section, we derive (3.7). To proceed, we start from the MSE in (3.6). As assumed, the elements in the precoded signal $\mathbf{x}_k$'s are statistically independent and have zero-mean and a same variance σ_s^2. Then we can have $E\left[\mathbf{x}\mathbf{x}^H\right]=\sigma_s^2\mathbf{I}_N$ and $E\left[\mathbf{v}\mathbf{v}^H\right]=\sigma_s^2\mathbf{C}\mathbf{C}^H$. Note that the independent assumption is valid only for the large QAM size (e.g. $m \geq 16$) [21]. Therefore, we rewrite (3.6) as

$$
\begin{aligned}
J(\mathbf{C},\mathbf{F}_S,\mathbf{F}_R,\mathbf{G}) &= E_\Delta\left\{\mathrm{Tr}\left\{\sigma_s^2\mathbf{G}\mathbf{H}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}^H\mathbf{G}^H\right\}\right\} - E_\Delta\left\{\mathrm{Tr}\left\{\sigma_s^2\mathbf{G}\mathbf{H}\mathbf{F}_S\mathbf{C}^H\right\}\right\} \\
&\quad - E_\Delta\left\{\mathrm{Tr}\left\{\sigma_s^2\mathbf{C}\mathbf{F}_S^H\mathbf{H}^H\mathbf{G}^H\right\}\right\} + \mathrm{Tr}\left\{\sigma_s^2\mathbf{C}\mathbf{C}^H\right\} + E_\Delta\left\{\mathrm{Tr}\left\{\mathbf{G}\mathbf{R}_w\mathbf{G}^H\right\}\right\} \\
&= \mathrm{Tr}\left\{E_\Delta\left\{\sigma_s^2\mathbf{G}\mathbf{H}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}^H\mathbf{G}^H\right\}\right\} - \mathrm{Tr}\left\{E_\Delta\left\{\sigma_s^2\mathbf{G}\mathbf{H}\mathbf{F}_S\mathbf{C}^H\right\}\right\} \\
&\quad - \mathrm{Tr}\left\{E_\Delta\left\{\sigma_s^2\mathbf{C}\mathbf{F}_S^H\mathbf{H}^H\mathbf{G}^H\right\}\right\} + \mathrm{Tr}\left\{\sigma_s^2\mathbf{C}\mathbf{C}^H\right\} + \mathrm{Tr}\left\{E_\Delta\left\{\mathbf{G}\mathbf{R}_w\mathbf{G}^H\right\}\right\},
\end{aligned} \tag{B.1}
$$

where $\mathbf{R}_w = E\left[\mathbf{w}\mathbf{w}^H\right]$. We consider the first term of (B.1) :

$$
\begin{aligned}
&\mathrm{Tr}\left\{E_\Delta\left\{\sigma_s^2\mathbf{G}\mathbf{H}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}^H\mathbf{G}^H\right\}\right\} = \mathrm{Tr}\left\{\sigma_s^2\mathbf{G}E_\Delta\left\{\mathbf{H}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}^H\right\}\mathbf{G}^H\right\} \\
&= \sigma_s^2\mathbf{G}\begin{bmatrix} E_\Delta\left\{\mathbf{H}_{SD}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}_{SD}^H\right\} & E_\Delta\left\{\mathbf{H}_{SD}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}_{SR}^H\mathbf{F}_R^H\mathbf{H}_{RD}^H\right\} \\ E_\Delta\left\{\mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}_{SD}^H\right\} & E_\Delta\left\{\mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}_{SR}^H\mathbf{F}_R^H\mathbf{H}_{RD}^H\right\}\end{bmatrix}\mathbf{G}^H
\end{aligned} \tag{B.2}
$$

Since $\Delta\mathbf{H}_{SD}$, $\Delta\mathbf{H}_{SR}$ and $\Delta\mathbf{H}_{RD}$ are matrix-variate complex Gaussian random variables with zero mean and all independent, we can use a useful property as described below:

Property [44]**.** Let $\mathbf{X} \sim N_{p,n}(\mathbf{M},\Sigma\otimes\Psi)$ and $\Sigma=(\sigma_{ij})$, $\Psi=(\psi_{ij})$, then

$$
E(\mathbf{X}\mathbf{A}\mathbf{X}') = \mathrm{tr}(\mathbf{A}'\Psi)\Sigma + \mathbf{M}\mathbf{A}\mathbf{M}' \tag{B.3}
$$

This property has been proved in [44]**.** Thus, we can have the first diagonal term of (B.2) as

$$
\begin{aligned}
&E_\Delta\left\{\mathbf{H}_{SD}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}_{SD}^H\right\} = E_\Delta\left\{\left(\widehat{\mathbf{H}}_{SD}+\Delta\mathbf{H}_{SD}\right)\mathbf{F}_S\mathbf{F}_S^H\left(\widehat{\mathbf{H}}_{SD}+\Delta\mathbf{H}_{SD}\right)^H\right\} \\
&= Tr\left\{\mathbf{F}_S\mathbf{F}_S^H\Psi_{SD}\right\}\Sigma_{SD} + \widehat{\mathbf{H}}_{SD}\mathbf{F}_S\mathbf{F}_S^H\widehat{\mathbf{H}}_{SD}^H := \mathbf{T}_{SD}
\end{aligned} \tag{B.4}
$$

For the second diagonal term in (B.2), we have

$$
E_\Delta\left\{\mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}_{SR}^H\mathbf{F}_R^H\mathbf{H}_{RD}^H\right\} = E_\Delta\left\{\mathbf{H}_{RD}\mathbf{F}_R E_\Delta\left\{\mathbf{H}_{SR}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}_{SR}^H\right\}\mathbf{F}_R^H\mathbf{H}_{RD}^H\right\} \tag{B.5}
$$

$$
= E_\Delta\left\{\mathbf{H}_{RD}\mathbf{F}_R\mathbf{T}_{SR}\mathbf{F}_R^H\mathbf{H}_{RD}^H\right\} \tag{B.6}
$$

$$
= \mathrm{Tr}\left\{\Psi_{RD}\right\}\Sigma_{RD} + \widehat{\mathbf{H}}_{RD}\mathbf{F}_R\mathbf{T}_{SR}\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H \tag{B.7}
$$

where the equality in (B.5) is due to the assumption that $\Delta\mathbf{H}_{SR}$ and $\Delta\mathbf{H}_{RD}$ are independent.

The matrix $\mathbf{T}_{SR}$ has a similar form as that of $\mathbf{T}_{SD}$ in (B.4) and it can be expressed as

$$\mathbf{T}_{SR} := E_{\Delta}\left\{\mathbf{H}_{SR}\mathbf{F}_{S}\mathbf{F}_{S}^{H}\mathbf{H}_{SR}^{H}\right\} = \mathrm{Tr}\left\{\mathbf{F}_{S}\mathbf{F}_{S}^{H}\mathbf{\Psi}_{SR}\right\}\Sigma_{SR} + \widehat{\mathbf{H}}_{SR}\mathbf{F}_{S}\mathbf{F}_{S}^{H}\widehat{\mathbf{H}}_{SR}^{H} \tag{B.8}$$

The equality (B.7) is obtained by using the property outlined above. For the off-diagonal matrices in (B.2), we have:

$$E_{\Delta}\left\{\mathbf{H}_{SD}\mathbf{F}_{S}\mathbf{F}_{S}^{H}\mathbf{H}_{SR}^{H}\mathbf{F}_{R}^{H}\mathbf{H}_{RD}^{H}\right\} = \widehat{\mathbf{H}}_{SD}\mathbf{F}_{S}\mathbf{F}_{S}^{H}\widehat{\mathbf{H}}_{SR}^{H}\mathbf{F}_{R}^{H}\widehat{\mathbf{H}}_{RD}^{H} \tag{B.9}$$

$$E_{\Delta}\left\{\mathbf{H}_{RD}\mathbf{F}_{R}\mathbf{H}_{SR}\mathbf{F}_{S}\mathbf{F}_{S}^{H}\mathbf{H}_{SD}^{H}\right\} = \widehat{\mathbf{H}}_{RD}\mathbf{F}_{R}\widehat{\mathbf{H}}_{SR}\mathbf{F}_{S}\mathbf{F}_{S}^{H}\widehat{\mathbf{H}}_{SD}^{H} \tag{B.10}$$

For the second and third term of (B.1), it is clear that

$$\mathrm{Tr}\left\{E_{\Delta}\left\{\sigma_{s}^{2}\mathbf{G}\mathbf{H}\mathbf{F}_{S}\mathbf{C}^{H}\right\}\right\} = \mathrm{Tr}\left\{\sigma_{s}^{2}\mathbf{G}\widehat{\mathbf{H}}\mathbf{F}_{S}\mathbf{C}^{H}\right\} \tag{B.11}$$

$$\mathrm{Tr}\left\{E_{\Delta}\left\{\sigma_{s}^{2}\mathbf{C}\mathbf{F}_{S}^{H}\mathbf{H}^{H}\mathbf{G}^{H}\right\}\right\} = \sigma_{s}^{2}Tr\left\{\mathbf{C}\mathbf{F}_{S}^{H}\widehat{\mathbf{H}}^{H}\mathbf{G}^{H}\right\} \tag{B.12}$$

For the last term of (B.1), we have

$$\begin{aligned}\bar{\mathbf{R}}_{w} &= E_{\Delta}\left\{\begin{bmatrix}\mathbf{n}_{D,1}\\ \left(\widehat{\mathbf{H}}_{RD}+\Delta\mathbf{H}_{RD}\right)\mathbf{F}_{R}\mathbf{n}_{R}+\mathbf{n}_{D,2}\end{bmatrix}\begin{bmatrix}\mathbf{n}_{D,1}^{H} & \left(\left(\widehat{\mathbf{H}}_{RD}+\Delta\mathbf{H}_{RD}\right)\mathbf{F}_{R}\mathbf{n}_{R}+\mathbf{n}_{D,2}\right)^{H}\end{bmatrix}\right\}\\ &= \begin{bmatrix}\sigma_{n,d}^{2}\mathbf{I}_{M} & \mathbf{0}\\ \mathbf{0} & \sigma_{n,r}^{2}\mathrm{Tr}\left(\mathbf{F}_{R}\mathbf{F}_{R}^{H}\mathbf{\Psi}_{RD}\right)\Sigma_{RD}+\sigma_{n,r}^{2}\widehat{\mathbf{H}}_{RD}\mathbf{F}_{R}\mathbf{F}_{R}^{H}\widehat{\mathbf{H}}_{RD}^{H}+\sigma_{n,d}^{2}\mathbf{I}_{M}\end{bmatrix}\end{aligned} \tag{B.13}$$

Finally, substituting (B.2)-(B.13) into (B.1), we can obtain (3.7) after some simplifications.

Appendix C: To prove the cost function (3.22) is monotonically decreasing in $\alpha_S \geq 0$.

Let's rewrite the MSE matrix **E** in (3.22) as a function of α_S,

$$\begin{aligned}\mathbf{E}\left(\alpha_{S}\right) &= \sigma_{s}^{2}\mathbf{C}\left(\alpha_{S}^{2}\mathbf{U}_{S}^{H}\tilde{\mathbf{H}}^{H}\tilde{\mathbf{H}}\mathbf{U}_{S}+\mathbf{I}_{N}\right)^{-1}\mathbf{C}^{H}\\ &= \sigma_{s}^{2}\mathbf{C}\Big(\sigma_{s}^{2}\mathbf{U}_{S}^{H}\Big(\sigma_{n,d}^{-2}\widehat{\mathbf{H}}_{SD}^{H}\left(\alpha_{S}^{-2}\sigma_{n,d}^{2}\mathbf{I}_{M}+\sigma_{s}^{2}\mathrm{Tr}\left(\mathbf{\Psi}_{SD}\right)\Sigma_{SD}\right)^{-1}\widehat{\mathbf{H}}_{SD}+\\ &\qquad \widehat{\mathbf{H}}_{SR}^{H}\mathbf{F}_{R}^{H}\widehat{\mathbf{H}}_{RD}^{H}\left(\alpha_{S}^{-2}\Delta\mathbf{A}+\alpha_{S}^{-2}\mathbf{A}\right)^{-1}\widehat{\mathbf{H}}_{RD}\mathbf{F}_{R}\widehat{\mathbf{H}}_{SR}\Big)\mathbf{U}_{S}+\mathbf{I}_{N}\Big)^{-1}\mathbf{C}^{H},\end{aligned} \tag{C.1}$$

where $\Delta\mathbf{A}$ and $\mathbf{A}$ are those defined in (3.24) and (3.25). We know that $tr\left\{\mathbf{E}\left(\alpha_{S}\right)\right\}$ is monotonically decreasing in $\alpha_S \geq 0$ which implies $\mathbf{E}\left(\alpha_{S,1}\right) \prec \mathbf{E}\left(\alpha_{S,2}\right)$ for any $\alpha_{S,1} \geq \alpha_{S,2}$,

$\alpha_{S,1}$ and $\alpha_{S,2}$ are real numbers. Our purpose is to check if $\mathbf{E}(\alpha_{S,1}) \prec \mathbf{E}(\alpha_{S,2})$ for any $\alpha_{S,1} \ge \alpha_{S,2}$ or not. First, we start the proof with the following lemmas.

Lemma C.1 [43]: For any two Hermitian matrices, $\mathbf{P}_1$ and $\mathbf{P}_2$, if $\mathbf{P}_1 \ge \mathbf{P}_2$, then $\mathbf{X}^H\mathbf{P}_1\mathbf{X} \ge \mathbf{X}^H\mathbf{P}_2\mathbf{X}$ with an arbitrary matrix $\mathbf{X}$ where $\mathbf{P}_1 \ge \mathbf{P}_2$ indicates that $\mathbf{P}_1 - \mathbf{P}_2$ is a semi-positive definite matrix.

Lemma C.2 [43]: For any two Hermitian matrices, $\mathbf{P}_1$ and $\mathbf{P}_2$, $\mathbf{P}_1 \ge \mathbf{P}_2$ if and only if $\mathbf{P}_1^{-1} \le \mathbf{P}_2^{-1}$.

To start with, we consider the cumbersome part in (C.1)

$$\begin{aligned}\alpha_S^{-2}\Delta\mathbf{A} = &\operatorname{Tr}\left(\mathbf{F}_R\left(\sigma_s^2\underbrace{\left(\operatorname{Tr}\left(\boldsymbol{\Psi}_{SR}\right)\boldsymbol{\Sigma}_{SR} + \widehat{\mathbf{H}}_{SR}\widehat{\mathbf{H}}_{SR}^H\right)}_{=\widehat{\mathbf{H}}'_{SR}\widehat{\mathbf{H}}'^H_{SR}} + \alpha_S^{-2}\sigma_{n,r}^2\mathbf{I}_R\right)\mathbf{F}_R^H\boldsymbol{\Psi}_{RD}\right)\boldsymbol{\Sigma}_{RD} \\ &+\sigma_s^2\operatorname{Tr}\left(\boldsymbol{\Psi}_{SR}\right)\widehat{\mathbf{H}}_{RD}\mathbf{F}_R\boldsymbol{\Sigma}_{SR}\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H\end{aligned} \tag{C.2}$$

Since $\Delta\mathbf{A}$ and $\mathbf{A}$ are Hermitian matrices and $\Delta\mathbf{A}$ is a function of α_S. By (C.2), it is easy to observe that

$$\alpha_S^{-2}\Delta\mathbf{A}\Big|_{\alpha_S=\alpha_{S,1}} \le \alpha_S^{-2}\Delta\mathbf{A}\Big|_{\alpha_S=\alpha_{S,2}}, \text{ if } \alpha_{S,1} \ge \alpha_{S,2} \tag{C.3}$$

By Lemma C.2, we have

$$\left(\alpha_S^{-2}\Delta\mathbf{A} + \alpha_S^{-2}\mathbf{A}\right)^{-1}\Big|_{\alpha_S=\alpha_{S,1}} \ge \left(\alpha_S^{-2}\Delta\mathbf{A} + \alpha_S^{-2}\mathbf{A}\right)^{-1}\Big|_{\alpha_S=\alpha_{S,2}} \text{ if } \alpha_{S,1} \ge \alpha_{S,2} \tag{C.4}$$

By Lemma C.1, we have

$$\begin{aligned}&\left(\widehat{\mathbf{H}}_{SR}^H\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H\left(\alpha_S^{-2}\Delta\mathbf{A} + \alpha_S^{-2}\mathbf{A}\right)^{-1}\widehat{\mathbf{H}}_{RD}\mathbf{F}_R\widehat{\mathbf{H}}_{SR}\right)\Big|_{\alpha_S=\alpha_{S,1}} \\ &\ge \left(\underbrace{\widehat{\mathbf{H}}_{SR}^H\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H\left(\alpha_S^{-2}\Delta\mathbf{A} + \alpha_S^{-2}\mathbf{A}\right)^{-1}\widehat{\mathbf{H}}_{RD}\mathbf{F}_R\widehat{\mathbf{H}}_{SR}}_{:=\mathbf{M}}\right)\Big|_{\alpha_S=\alpha_{S,2}}, \text{ if } \alpha_{S,1} \ge \alpha_{S,2}\end{aligned} \tag{C.5}$$

Similarly,

$$\begin{aligned}&\left(\sigma_{n,d}^{-2}\widehat{\mathbf{H}}_{SD}^H\left(\alpha_S^{-2}\sigma_{n,d}^2\mathbf{I}_M + \sigma_s^2\operatorname{Tr}\left(\boldsymbol{\Psi}_{SD}\right)\boldsymbol{\Sigma}_{SD}\right)^{-1}\widehat{\mathbf{H}}_{SD}\right)\Big|_{\alpha_S=\alpha_{S,1}} \ge \\ &\left(\underbrace{\sigma_{n,d}^{-2}\widehat{\mathbf{H}}_{SD}^H\left(\alpha_S^{-2}\sigma_{n,d}^2\mathbf{I}_M + \sigma_s^2\operatorname{Tr}\left(\boldsymbol{\Psi}_{SD}\right)\boldsymbol{\Sigma}_{SD}\right)^{-1}\widehat{\mathbf{H}}_{SD}}_{:=\mathbf{N}}\right)\Big|_{\alpha_S=\alpha_{S,1}}, \text{ if } \alpha_{S,1} \ge \alpha_{S,2}\end{aligned} \tag{C.6}$$

Summing $\mathbf{M}$ and $\mathbf{N}$ in (C.5) and (C.6), and using Lemma C.1, we can have

$$\begin{aligned}&\left(\sigma_s^2\mathbf{C}\left(\sigma_s^2\mathbf{U}_S^H(\mathbf{M}+\mathbf{N})\mathbf{U}_S+\mathbf{I}_N\right)^{-1}\mathbf{C}^H\right)\Big|_{\alpha_S=\alpha_{S,1}}\\ &\le\left(\sigma_s^2\mathbf{C}\left(\sigma_s^2\mathbf{U}_S^H(\mathbf{M}+\mathbf{N})\mathbf{U}_S+\mathbf{I}_N\right)^{-1}\mathbf{C}^H\right)\Big|_{\alpha_S=\alpha_{S,2}}\end{aligned}\quad\text{, if }\ \alpha_{S,1}\ge\alpha_{S,2}\tag{C.7}$$

Therefore, we have $\mathbf{E}(\alpha_{S,1})\le\mathbf{E}(\alpha_{S,2})$ for any $\alpha_{S,1}\ge\alpha_{S,2}$ which implies that $tr\{\mathbf{E}(\alpha_{S,1})\}\le tr\{\mathbf{E}(\alpha_{S,2})\}$ for any $\alpha_{S,1}\ge\alpha_{S,2}$, e.g., $tr\{\mathbf{E}(\alpha_S)\}$ is monotonically decreasing with α_S.

Appendix D: To prove the property in (3.42)

In this section, we derive the lower bound in (3.42). We start the proof by using Lemma C.1-Lemma C.2 in Appendix C and the following lemmas.

Lemma D.1 [41]: For any two semi-positive definite matrices, $\mathbf{P}_1$ and $\mathbf{P}_2$, we have $\mathrm{Tr}(\mathbf{P}_1\mathbf{P}_2)\le\mathrm{Tr}(\mathbf{P}_1)\lambda_{\max}(\mathbf{P}_2)$, and the equality is satisfied if $\mathbf{P}_2=\lambda_{\max}\mathbf{I}$.

Lemma D.2 [41]: For any semi-positive definite matrix $\mathbf{P}$, we have $\lambda_{\max}(\mathbf{P})\mathbf{I}\ge\mathbf{P}$.

Lemma D.3: For any two positive matrices $\mathbf{P}_1$, $\mathbf{P}_2$, and semi-positive definite matrix $\mathbf{Y}$, we have if $\mathbf{P}_1\ge\mathbf{P}_2$, then $\det\left(\mathbf{Y}+\mathbf{X}\mathbf{P}_1^{-1}\mathbf{X}^H\right)\le\det\left(\mathbf{Y}+\mathbf{X}\mathbf{P}_2^{-1}\mathbf{X}^H\right)$.

Proof: This is obviously because if $\mathbf{P}_1\ge\mathbf{P}_2$, then $\det\left(\mathbf{Y}+\mathbf{X}\mathbf{P}_1^{-1}\mathbf{X}^H\right)\le\det\left(\mathbf{Y}+\mathbf{X}\mathbf{P}_2^{-1}\mathbf{X}^H\right)$ by Lemma C.1 and C.2.

We now consider $\Delta\mathbf{A}$ in (3.43). By Lemma D.1 and Lemma D.2, we have

$$\begin{aligned}&\mathrm{Tr}\left(\mathbf{F}_R\left(\frac{P_{S,T}}{N}\widehat{\mathbf{H}}'_{SR}\widehat{\mathbf{H}}'^H_{SR}+\sigma_{n,r}^2\mathbf{I}_R\right)\mathbf{F}_R^H\boldsymbol{\Psi}_{RD}\right)\boldsymbol{\Sigma}_{RD}+\frac{P_{S,T}}{N}\mathrm{Tr}(\boldsymbol{\Psi}_{SR})\widehat{\mathbf{H}}_{RD}\mathbf{F}_R\boldsymbol{\Sigma}_{SR}\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H\\ &\le\underbrace{\mathrm{Tr}\left(\mathbf{F}_R\left(\frac{P_{S,T}}{N}\widehat{\mathbf{H}}'_{SR}\widehat{\mathbf{H}}'^H_{SR}+\sigma_{n,r}^2\mathbf{I}_R\right)\mathbf{F}_R^H\right)}_{\le P_{R,T}}\lambda_{\max}(\boldsymbol{\Psi}_{RD})\lambda_{\max}(\boldsymbol{\Sigma}_{RD})\mathbf{I}_M\\ &+\frac{P_{S,T}}{N}\mathrm{Tr}(\boldsymbol{\Psi}_{SR})\lambda_{\max}(\boldsymbol{\Sigma}_{SR})\widehat{\mathbf{H}}_{RD}\mathbf{F}_R\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H\\ &\le P_{R,T}\lambda_{\max}(\boldsymbol{\Psi}_{RD})\lambda_{\max}(\boldsymbol{\Sigma}_{RD})\mathbf{I}_M+\frac{P_{S,T}}{N}\mathrm{Tr}(\boldsymbol{\Psi}_{SR})\lambda_{\max}(\boldsymbol{\Sigma}_{SR})\widehat{\mathbf{H}}_{RD}\mathbf{F}_R\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H.\end{aligned}\tag{D.1}$$

The last inequality of (D.1) is due to power constraint at the relay. Note that the equality is held if $\boldsymbol{\Psi}_{RD}=\beta_{RD}\mathbf{I}_R$, $\boldsymbol{\Sigma}_{RD}=\gamma_{RD}\mathbf{I}_{M\times M}$, $\boldsymbol{\Psi}_{SR}=\beta_{SR}\mathbf{I}_N$ and $\boldsymbol{\Sigma}_{SR}=\gamma_{SR}\mathbf{I}_{R\times R}$, which means the channel of each antenna pair is uncorrelated (transmit or receive antenna correlation matrix is an identity matrix). Then, from (D.1) and (3.40), we have

$$(\Delta\mathbf{A}+\mathbf{A}) \le (\Delta\mathbf{A}'+\mathbf{A}) \tag{D.2}$$

Now, let $\mathbf{Y}=\mathbf{I}_N+\frac{P_{S,T}}{N\sigma_{n,d}^2}\widehat{\mathbf{H}}_{SD}^H\left(\sigma_{n,d}^2\mathbf{I}_M+\frac{P_{S,T}}{N}\mathrm{Tr}(\mathbf{\Psi}_{SD})\mathbf{\Sigma}_{SD}\right)^{-1}\widehat{\mathbf{H}}_{SD}$, $\mathbf{B}=\sqrt{\frac{P_{S,T}}{N}}\widehat{\mathbf{H}}_{SR}^H\mathbf{F}_R^H\widehat{\mathbf{H}}_{RD}^H$, $\mathbf{P}_1=\Delta\mathbf{A}'+\mathbf{A}$ and $\mathbf{P}_2=\Delta\mathbf{A}+\mathbf{A}$ in Lemma D.1, and we can obtain the desired result.

Appendix E: Derivation of optimum solution in (3.53)

To solve the optimization problem in (3.53), we first consider the corresponding Lagrangian function

$$L=\sum_{i=1}^{K}\ln\left(1+\frac{p_{r,i}\sigma''^2_{sr,i}\sigma^2_{rd,i}}{a+bp_{r,i}\sigma^2_{rd,i}}\right)+\lambda\left[\sum_{i=1}^{\kappa}p_{r,i}\left(\sigma^2_{n,r}+\mathbf{D}_{sr}(i,i)\right)-P_{R,T}\right]-\sum_{i=1}^{K}v_{r,i}p_{r,i} \tag{E.1}$$

where $\lambda\ge 0$, $v_{r,i}\ge 0$ with $i=1,\cdots,\kappa$. By the KKT conditions (for all i), we have

$$\frac{\partial L}{\partial p_k}=-\frac{\dfrac{\alpha\sigma''^2_{sr,i}\sigma^2_{rd,i}}{\left(\alpha+\beta p_{r,i}\sigma^2_{rd,i}\right)^2}}{1+\dfrac{p_{r,i}\sigma''^2_{sr,i}\sigma^2_{rd,i}}{\alpha+\beta p_{r,i}\sigma^2_{rd,i}}}+\lambda\left[\sigma^2_{n,r}+\mathbf{D}_{sr}(i,i)\right]-v_{r,i}=0 \tag{E.2}$$

$$v_{r,i}p_{r,i}=0 \tag{E.3}$$

$$\lambda\left[\sum_{i=1}^{\kappa}p_{r,i}\left(\sigma^2_{n,r}+\mathbf{D}_{sr}(i,i)\right)-P_{R,T}\right]=0 \tag{E.4}$$

$$\lambda, v_{r,i}, p_{r,i}\ge 0 \tag{E.5}$$

Substituting (E.2) into (E.3) with the consideration of $p_{r,i}>0$, we have $v_{r,i}=0$, and thus

$$\frac{\dfrac{\alpha\sigma''^2_{sr,i}\sigma^2_{rd,i}}{\left(\alpha+\beta p_{r,i}\sigma^2_{rd,i}\right)^2}}{1+\dfrac{p_{r,i}\sigma''^2_{sr,i}\sigma^2_{rd,i}}{\alpha+\beta p_{r,i}\sigma^2_{rd,i}}}=\frac{1}{p^2_{r,i}\underbrace{\sigma^2_{rd,i}\cdot\frac{\beta}{\alpha}\left(\frac{\beta}{\sigma''^2_{sr,i}}+1\right)}_{:=A_i}+2p_{r,i}\underbrace{\left(\frac{\beta}{\sigma''^2_{sr,i}}+\frac{1}{2}\right)}_{:=B_i}+\underbrace{\frac{\alpha}{\sigma''^2_{sr,i}\sigma^2_{rd,i}}}_{:=C_i}} \tag{E.6}$$

$$=\lambda\left[\sigma^2_{n,r}+\mathbf{D}_{sr}(i,i)\right]$$

Then, we have

$$p_{r,i} = \sqrt{\frac{1}{\lambda\left(\sigma_{n,r}^2 + \mathbf{D}_{sr}(i,i)\right)A_i} + \frac{B_i^2}{A_i^2} - \frac{C_i}{A_i}} - \frac{B_i}{A_i} \tag{E.7}$$

where

$$\frac{B_i^2 - A_i C_i}{A_i^2} = \frac{1/4}{\sigma_{rd,i}^4 \dfrac{\beta^2}{\alpha^2}\left(\dfrac{\beta}{\sigma''^2_{sr,i}} + 1\right)^2} \tag{E.8}$$

and

$$\frac{B_i}{A_i} = \frac{\alpha + \dfrac{\alpha\sigma''^2_{sr,i}}{2b}}{\sigma_{rd,i}^2\left(\beta + \sigma''^2_{sr,i}\right)} \tag{E.9}$$

After some manipulations and using the constraint in (E.4), the optimum $p_{r,i}$ can be expressed as

$$p_{r,i} = \left[\sqrt{\frac{\mu}{\left(\sigma_{n,r}^2 + \mathbf{D}_{sr}(i,i)\right)\sigma_{rd,i}^2 \cdot \dfrac{\beta}{\alpha}\left(\dfrac{\beta}{\sigma''^2_{sr,i}} + 1\right)} + \frac{1/4}{\sigma_{rd,i}^4 \dfrac{\beta^2}{\alpha^2}\left(\dfrac{\beta}{\sigma''^2_{sr,i}} + 1\right)^2}} - \frac{\alpha + \dfrac{\alpha\sigma''^2_{sr,i}}{2\beta}}{\sigma_{rd,i}^2\left(\beta + \sigma''^2_{sr,i}\right)}\right]^+ \tag{E.10}$$

where $\mu := 1/\lambda$ is chosen to satisfy the constraint in (3.52).

References

[1] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.

[2] A. Scaglione, D. L. Goeckel, and J. N. Laneman, “Cooperative communications in mobile Ad Hoc networks,” *IEEE Signal Processing magazine*, vol. 23, no.5, pp.18-29, Sep. 2006.

[3] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, “Fading relay channels: Performance limits and space-time signal design,”*IEEE Trans. Selected Areas in Communications*, vol. 22, no. 6, pp. 1099-1109, Aug. 2004.

[4] I. E. Telatar, ”Capacity of multiantenna Gaussian channel,” *Eur. Trans. Telecommun.*, vol. 6, pp. 311-335, 1998.

[5] A. Goldsmith, S.A. Jafar, N. Jindal, and S. Vishwanath, ”Capacity limits of MIMO channels,” *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 684-702, June 2003.

[6] A. Goldsmith, *Wireless Communication*, Cambridge University Press 2005.

[7] H. Sampath, P. Stoica, and A. Paulraj, “Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion,” *IEEE Trans. Communications*, vol. 49, no. 12, pp. 2198-2206, Dec. 2001.

[8] D. P. Palomar, J. M. Cioffi, and M. A. Laguna, “Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization,” *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2381- 2401, Sept. 2003.

[9] D. P. Palomar and S. Barbarossa, “Designing MIMO communication systems: Constellation choice and linear transceiver design,” *IEEE Trans. Signal Processing*, vol. 53, no. 10, pp. 3804-3818, Oct. 2005.

[10] S. Yang and J. C. Belfiore, “Optimal space-time codes for the MIMO amplify-and-forward cooperative channel,” *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 647-663, Feb. 2007.

[11] G. Scutari and S. Barbarossa, “Distributed space-time coding for regenerative relay networks,” *IEEE Trans. Wireless Communications*, vol. 4, no. 5, pp. 2387-2399, Sep. 2005.

[12] B. Wang, J. Zhang, and A. Host-Madsen, “On the capacity of MIMO relay channels,” *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29-43, Jan. 2005.

[13] O. Munoz-Medina, J. Vidal, and A. Agustin, “Linear transceiver design in non-regenerative relays with channel state information,” *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2593-2604, June 2007.

[14] X. Tang and Y. Hua, “Optimal design of non-regenerative MIMO wireless relays,” *IEEE Trans. Wireless Communications*, vol. 6, no. 4, pp. 1398-1407, April 2007.

[15] W. Guan and H. Luo, “Joint MMSE transceiver design in non-regenerative MIMO relay systems,” *IEEE Trans. Communications Letters*, vol. 12, no. 7, pp. 517-519, July 2008.

[16] A. S. Behbahani, R. Merched, and A. M. Eltawil, “Optimization of a MIMO relay network,” *IEEE Trans. Signal Processing*, vol. 56, no. 10, pp. 5062-5073, Oct. 2008.

[17] F. S. Tseng, W. R. Wu and J. Y. Wu, “Joint source/relay precoder design in non-regenerative cooperative

systems using an MMSE criterion," *IEEE Trans. Wireless Communications*, vol. 8, no. 10, pp. 4928-4933, Oct. 2009.

[18] F. S. Tseng and W. R. Wu, "Linear MMSE transceiver design in amplify-and-forward MIMO relay systems," *IEEE Trans. Vehicular Technology*, vol. 59, no. 2, pp. 754-765, Feb. 2010.

[19] F. S. Tseng and W. R. Wu, "Joint source/relay precoders design in amplify-and-forward relay systems: A geometric mean decomposition approach," in *Proc. IEEE ICASSP 2009*.

[20] F. S. Tseng and W. R. Wu, "MMSE-SIC transceiver design in amplify-and-forward MIMO relay systems," in *Proc. IEEE VTC-Spring 2010*.

[21] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*. Wiley-IEEE Press, July 2002.

[22] F. S. Tseng, M. Y. Chang, and W. R. Wu, "Joint MMSE transceiver design in amplify-and-forward MIMO relay systems with Tomlinson-Harashima source precoding," in *Proc. IEEE PIMRC 2010*.

[23] R. F. H. Fischer, C. Windpassinger, A.Lampe, and J. H. Huber, "Space time transmission using Tomlinson-Harashima precoding," in *Proc. 4th ITC Conf. on Source and Channel Coding*, Berlin, 2002.

[24] A. A. D'Amico and M. Morelli, "Joint Tx-Rx MMSE design for MIMO multicarrier systems with Tomlinson-Harashima precoding," *IEEE Trans. Wireless Communications*, vol. 7, no. 8, pp. 3118-3127, Aug. 2008.

[25] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 2700-2712, Aug. 2005.

[26] Y. Jiang, J. Li, and W. W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Trans. Signal Processing*, vol. 53, no. 10, pp. 3791-3803, Oct. 2005.

[27] J. K. Zhang, A. Kavcic, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive cancellation detection," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 154-172, Jan. 2005.

[28] Y. Jiang, J. Li, and W. W. Hager, "Uniform channel decomposition for MIMO Communications," *IEEE Trans. Signal Processing*, vol. 53, no. 11, pp. 4283-4294, Nov. 2005.

[29] L. Musavian, M. R. Nakhi, M. Dohler, and A. H. Aghvami, "Effect of channel uncertainty on the mutual information of MIMO fading channels," *IEEE Trans. Vehicular Technology*, vol. 56, no. 5, pp. 2798-2806, Sept. 2007.

[30] D. Tsipouridou and A. P. Liavas, "On the sensitivity of the MIMO Tomlinson-Harashima precoder with respect to channel uncertainties," *IEEE Trans. Signal Processing*, vol. 58, no. 4, pp. 2261-2272, April 2010.

[31] M. Ding and S. D. Blostein, "MIMO minimum total MSE transceiver design with imperfect CSI at both ends," *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 1141V1150, March 2009.

[32] H. K. Bizaki and A. Falahati, "Tomlinson-Harashima precoding with imperfect channel state information," Communication IET, vol. 2, pp. 151~158, Jan. 2008.

[33] J. Duplicy and L. Vandendorpe, "Robust MMSE Precoding for the MIMO complex Gaussian broadcast channel," in *Proc. IEEE ICASSP 2007*.

[34] Y. Guo, B. C. Levy, "Worst-case MSE precoder design for imperfectly known MIMO communications channels," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 2918-2903, Aug. 2005.

[35] J. W. Huang, E. K. S. Au, and V. K. N. Lau, "Precoder design for space-time coded MIMO systems with

imperfect channel state information," *IEEE Trans. Wireless Communications*, vol. 7, no. 6, pp. 1997-1981, June 2008.

[36] C. Xing, S. Ma, and Y. C. Wu, "Robust joint design of linear relay precoder and destination equalizer for dual-hop amplify-and-forward MIMO relay systems," *IEEE Trans. Signal Processing*, vol. 58, no.4, pp. 2273-2283, April 2010.

[37] M. B. Shenouda and T. N. Davidson, "A framework for designing MIMO systems with decision feedback equalization or Tomlinson-Harashima Precoding," *IEEE journal on Selected Areas in Communications*, vol. 26, no. 2, pp. 401-411, February 2008.

[38] E. G. Larsson and P. Stoica, Space-Time Block Coding for Wireless Communications. Cambridge University Press, 2003.

[39] A. T. James, "Distributions of matrix variants and latent roots derived from normal samples," *Ann. Math. Statisitcs*, vol. 35, pp. 475-501, 1964.

[40] H. Weyl, "Inequalities between the two kinds of eigenvalues of a linear transformation," *Proc. Nat. Acad. Sci*., vol. 35, pp. 408-411, July 1949.

[41] D.S. Bernstein, *Matrix Mathematics*, Princeton University Press 2005.

[42] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press 2004.

[43] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[44] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*, 2000.