Channel Estimation for Amplify-and-Forward Relaying

by

Berna Gedik

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2008

I hereby declare that I am the sole author of this thesis.
This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

Berna Gedik

I understand that my thesis may be made electronically available to the public.

Berna Gedik

Abstract

Cooperative diversity has been proposed as a powerful means to enhance the performance of high-rate communications over wireless fading channels. It realizes spatial diversity advantages in a distributed manner where two or more nodes share their antennas to mimic a virtual antenna array. Most of the current research on cooperative diversity assumes the availability of perfect channel state information (CSI) at the receiver. However, in practice, the fading channel coefficients need to be first estimated and then can be used in the detection process.

The first part of the thesis presents a comparative performance evaluation of so-called cascaded channel estimation (C-CE) and disintegrated channel estimation (D-CE) techniques proposed for amplify-and-forward (AaF) relaying. Our results demonstrate that C-CE is particularly attractive for low-complexity AaF relaying systems. C-CE is therefore our main focus in the later parts of this thesis.

In the second part of the thesis, we analyze the error rate performance of AaF relaying with partially-coherent and mismatched-coherent receivers. For both receiver types, we demonstrate the achievable diversity orders for cooperation protocols under consideration and quantify the impact of channel estimation through the derivation of pairwise error probability.

Finally, we propose an optimal power allocation scheme for performance optimization of AaF cooperative system in the presence of channel estimation training. Considering average received signal-to-noise ratio at the destination node as the objective function, we formulate an optimization problem for a single-relay scenario and determine how much power should be allocated to training and data transmission periods and how much training/data power should be allocated to broadcasting and relaying phases. Our results indicate that optimized power allocation scheme demonstrates significant performance gains over its equal power allocation competitor.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Professor Murat Uysal for his guidance throughout my research. His encouragement has been an incredible motivation for me to complete this thesis.

I would like to give my deepest thanks to my mother and my grandmother for all their endless love and support over these years.

Table of Contents

List of Figures

Abbreviations

AaF	Amplify-and-forward
APS	Average Power Scaling
AWGN	Additive white Gaussian noise
BER	Bit error rate
C-CE	Cascaded channel estimation
CSI	Channel state information
DaF	Decode-and-forward
D-CE	Disintegrated channel estimation
EGC	Equal gain combining
IPS	Instantaneous Power Scaling
LMMSE	Linear minimum mean square error
MRC	Maximal-ratio-combining
ML	Maximum-likelihood
NAF	Non-orthogonal Amplify-and-forward
OAF	Orthogonal Amplify-and-forward
PEP	Pairwise error probability
PSK	Phase shift keying
SNR	Signal-to-Noise Ratio
STBC	Space-time block coding
STTC	Space-time trellis coding

List of Tables

Notation:

Bold upper and lower letter denote matrices and column vectors, respectively.

$(.)^{\mathrm{T}}$ Transpose operation;

$(.)^{\mathrm{H}}$ Hermitian transpose operation;

$\mathrm{E}(.)$ Denotes expectation;

$\mathrm{var}(.)$ Denotes variance;

$\mathrm{trace}(.)$ Denotes trace of a matrix;

$\mathrm{det}(.)$ Denotes determinant of a matrix;

$\mathrm{diag}(.)$ Denotes diagonal of a matrix;

$\|.\|$ Denotes Frobenius norm;

$|.|$ Denotes absolute value;

$\mathbf{I}_N$ Denotes identity matrix of size N;

i Denotes $\sqrt{-1}$

Chapter 1

Introduction

The introduction of commercial wireless telephone systems dates back to the late 1940s when AT&T Labs implemented the first public mobile telephony in USA. These early systems allowed mobile users to connect each other within a restricted area and its performance was limited by low user capacity and poor speech quality. After many improvements, the first cellular radio networks emerged in late 1970s. With the introduction of the cellular concept, capacity and mobility of communication systems increased and further developments in semiconductor and microprocessor technology have given rise to low cost, high-speed communications with more compact user terminals.

A series of generations have been introduced in the technological history of wireless communications [1]. First-generation (1G) cellular systems appeared in the early 1980s. These systems provided only voice communications based on analog transmission techniques. As the number of cellular subscribers grew, the need for better system capacity and better transmission quality was perceived. Therefore, second-generation (2G) systems were developed using digital technologies to support better voice quality, enhanced security features for privacy, and larger capacity. The development of third-generation (3G) systems has been initiated by the need for higher data rates to support a wide range of high-speed applications such as multimedia transmission, internet access, and video telephony.

Cellular networks require a central base station to transmit one user's information to another. On the other hand, built on a decentralized architecture, wireless ad-hoc networks allow peer-to-peer communications without the support from a fixed infrastructure. In these networks, each node forwards data on behalf of the other nodes enabling

communication in a distributed manner. Such systems have a number of application areas such as home networking, sensor networks, redundant link in disaster recovery and relief efforts, and military operations in areas without a fixed infrastructure. The introduction of wireless ad-hoc networks to mobile phones would enable local communications between personal devices without the need of a base station. Unfortunately, as the number of users in such networks increases, maintaining a reliable communication becomes more challenging due to network coverage and interference problems; and for the present, there is no standard for ad-hoc applications in cellular communications [2].

In the current wireless market, there exist several wireless technologies (e.g. Wi-Fi, Wi-Max, Bluetooth) supporting different service types and data rates. The idea of next generation (4G) systems has been born with the vision of integrating all these technologies in a single standard. 4G technology aims to provide voice, data, and multimedia services to users anytime and anywhere at lower costs [3].

1.1 Diversity Techniques for Wireless Fading Channels

Wireless communication systems experience severe performance degradations due to unpredictable channel characteristics. The main impairments which significantly affect the integrity and thus efficiency of digital data transmission can be classified as long-term path loss/shadowing and short-term multipath fading. Wireless radio channels demonstrate multipath behavior because of reflections and refractions of the transmitted signal on the way from transmitter to receiver. Due to randomly delayed, reflected, and scattered signal components, significant attenuations in the received signal strength are observed in such environments. There are different statistical models to describe the envelope of multipath fading. For typical mobile wireless channels where there is no line of sight propagation, amplitude of the multipath fading follows a Rayleigh distribution. In terms of error rate performance, Rayleigh fading converts the exponential dependency of the bit-error probability on the signal-to-noise ratio (SNR) for the classical additive white

Gaussian noise (AWGN) channel into an approximately inverse linear one, resulting in a large SNR penalty.

A common technique to combat destructive effects of fading is to utilize two or more communication channels. If independent channels are provided to transmit multiple versions of the same signal, the probability that all replicas simultaneously experience deep fading is significantly reduced. This technique is called diversity. The four most common diversity schemes can be identified as time, frequency, space, and polarization diversity [4], [5].

In time diversity, replicas of the same signal are transmitted at different time instants separated by an interval longer than the coherence time of the channel. Although this technique performs very well in fast fading environments, it is not as effective in slow fading channels unless significant delays can be tolerated.

In frequency diversity, independence among diversity channels is ensured by sending the same signal over different carriers, whose frequencies are separated by more than the coherence bandwidth of the channel. The need for multiple frequencies makes this technique bandwidth inefficient.

Space diversity, also known as antenna diversity, utilizes more than one antenna at the transmitter and/or the receiver. To ensure independent fading channels, adjacent antennas must be spaced many carrier wavelengths apart. Although this technique is bandwidth efficient, it requires complex hardware circuitry which leads to higher power dissipation. Therefore, space diversity is usually not preferred at small mobile handsets.

In polarization diversity, diversity is achieved based on different propagation characteristics of the vertically and horizontally polarized electromagnetic waves. By using dual polarized antennas at the base station, the probability that the signals received on the two differently polarized antennas undergo a deep fading simultaneously is reduced. The advantage of polarization diversity over space diversity is the reduced cost for installation

and the smaller space needed in comparison to space diversity. On the other hand, the number of independently faded diversity branches is limited to two in polarization diversity.

1.2 Receive and Transmit Diversity

Space diversity is a practical way of enhancing the reliability of data transmission without requiring an excessive bandwidth expansion. Space diversity, in the form of multiple antenna deployment at the receive side (i.e., receive diversity), has been successfully used in uplink transmission (i.e., from mobile station to base station) of the cellular communication systems for many years. In receive diversity, using different diversity combining techniques, the signals on independent diversity branches associated with multiple receive antennas are processed together to obtain a single signal. The most common diversity combining techniques can be classified as selection, equal gain, and maximal ratio combining [4]. In selection combining (SC), only the received signal with the highest SNR is used. Therefore, SC requires the monitoring of each branch simultaneously and the ability to switch to the strongest one. Disregarding the information obtained from all branches except the one with highest SNR makes this technique's performance inferior. In equal gain combining (EGC), the signals on all diversity branches are combined linearly and weighting coefficients of the linear combination are computed by considering only channel phase information. Since the signals from all branches are taken into account, the performance of EGC is superior to that of SC. In maximal-ratio-combining (MRC), by using both phase and amplitude information, the SNR of linearly combined diversity branches is maximized. Although MRC achieves the best performance compared to the other two techniques, EGC can be preferred in practical systems because of its reduced complexity while achieving near-optimal MRC performance. A detailed treatment of diversity combining systems can be found in [4], [5] and the references therein.

Although the deployment of multiple receive antennas at the base stations yield high performance for uplink transmission, the use of more than one antenna at the mobile handsets is not possible due to size, power, and hardware complexity limitations. A popular technique to overcome this problem is transmit diversity which uses multiple transmit antennas at the base station for downlink transmission.

Transmit diversity techniques are classified as open loop or close loop based on whether they need channel state information (CSI) at the transmit side. Close loop transmit diversity requires CSI feedback from the receiver to the transmitter to compute weighting coefficients. Rapidly varying channel conditions and imperfections on the feedback link may cause outdated and erroneous channel observations at the transmitter. Unlike close loop schemes, open loop transmit diversity does not require CSI knowledge at the transmitter. Using linear processing to transmit information over multiple antennas enables full diversity with simple decoding techniques at the receiver. The schemes proposed by Wittneben [6], [7] are the earliest open loop transmit diversity techniques. Converting the operating frequency-flat fading channel intentionally into a frequency-selective channel, Wittneben's so-called delay diversity provides a diversity order that is equal to the number of transmit antennas [8]. The need to design channel codes for multiple-antenna systems has later resulted in the introduction of space-time trellis codes [9].

1.3 Space-Time Coding

Space- time trellis coding (STTC) is an open loop transmission technique which effectively provides diversity advantages and coding gain [9]. The receiver structure of this scheme relies on Maximum Likelihood Sequence Estimation via Viterbi Algorithm and its decoding complexity increases exponentially with the transmission rate for a fixed number of transmit antennas. Space-time block coding (STBC) [10], [11] has been proposed to extract full diversity gains with a much lower receiver complexity. STBCs of [10], [11] were inspired by Alamouti's earlier work in [12] which introduced a simple

open loop scheme that provides the same diversity order as MRC with two receive antennas. In Alamouti's code, symbols transmitted from the two antennas are mapped into space and time domains in a manner to ensure that transmissions from both antennas are orthogonal to each other. This allows receiver to use very simple linear processing techniques to perform decoding. More detailed overview of this topic can be found in [13], [14], [15].

1.4 Cooperative Diversity

Space-time coding has been acknowledged as one of the most efficient ways to improve link reliability and spectral efficiency, to the extent that it has been already incorporated into 3G wireless standards. The use of space-time coding in mobile terminals is not however practical due to size, power, and hardware limitations. The particular technique that allows single antenna mobiles to reap the benefits of space-time coding is introduced under the name of "cooperative diversity" or "user cooperation". In cooperative diversity, broadcast nature of the wireless medium allows cooperating nodes (named as "partners" or "relays") to transmit each other's message. Hence, a virtual antenna array is created to extract the spatial diversity advantages in a distributed manner.

The concept of cooperative diversity can be traced back to the work of Meulen [16] where he has analyzed a three terminal relay network and introduced a two hop communication system instead of point-to-point communication between source and destination. In [17], considering AWGN channels, Cover et al. have developed upper and lower bounds on capacity for a single-relay system. In [18] and [19], Sendonaris et al. have generalized the work of [17] for fading channels and shown that cooperating nodes in a wireless network provide higher data rates. Laneman et al. [20], [21] have proposed a cooperation protocol built upon a two phase transmission scheme and demonstrated that the receiver can achieve full diversity by user cooperation. In [20], authors have further studied so-called Amplify-and-Forward (AaF) and Decode-and-Forward (DaF) relaying

techniques. These techniques differ from each other in how the received signal is processed at the relay station before it is sent to the destination. In DaF relaying, the relay terminal decodes the source node's message and then re-encodes it before transmitting to the destination. On the other hand, in AaF relaying, the relay amplifies and retransmits its received noisy signal without decoding it. Although the noise of the partner is amplified in this scheme, the destination still receives independently faded versions of the signal. In [22], Laneman et al. have applied conventional space-time codes to user cooperation systems in a distributed manner. Their proposed use of STBC implements coding across the relay nodes assuming a scenario with more than one relay. Considering a single relay scenario, Nabar et al. [23] have also studied distributed STBC; however their setup realizes coding across the source and the relay nodes. Particularly, they consider the three following cooperation protocols[1]:

- **Transmit Diversity (TD) Protocol:** In the first transmission phase (broadcasting phase), the source terminal broadcasts to the relay and destination terminals. During the second transmission phase (relaying phase), both the relay and source terminals communicate with the destination terminal. TD Protocol in conjunction with AaF relaying is referred as "Non Orthogonal Amplify and Forward (NAF) protocol" in [25].

- **Receive Diversity (RD) Protocol:** The source terminal communicates with the relay and destination terminals in broadcasting phase. In the relaying phase, the source stops transmission and only the relay terminal communicates with the destination. This protocol is the same as proposed in [20] and also referred as "Orthogonal Amplify and Forward (OAF) protocol" in [25].

---

[1] These protocols are referred as Protocol I, Protocol II, and Protocol III in [23]. Here, we use the protocol names as proposed in [24].

- **Simplified Transmit Diversity (STD) Protocol:** This is similar to TD protocol except that the destination terminal does not receive from the source during the first transmission phase for reasons which are possibly imposed from the upper-layer networking protocols. For example, the destination terminal may be engaged in data transmission to another terminal during the first phase.

In RD protocol, the signal conveyed to the relay and destination terminals over the two phases is the same and RD protocol effectively realizes receive diversity in a distributed manner. On the other hand, both TD and STD protocols can potentially convey different signals to the relay and destination terminals. This makes possible the deployment of various conventional space-time codes (originally proposed for co-located antennas) in a distributed scenario.

1.5 Motivation and Thesis Outline

The current literature on cooperative diversity has been mainly limited to some idealistic assumptions such as the availability of perfect CSI at the receiver. However, in practical systems, fading channel coefficients need to be estimated for detection stage. The quality of channel estimates inevitably affects the overall performance of relay-assisted transmission and might become a performance limiting factor. In this thesis, our main focus will be the analysis and optimization of channel estimation schemes for cooperative communications. We consider AaF relaying throughout the work and consider a single-relay cooperative network unless otherwise noted.

1.5.1 Cascaded versus Disintegrated Channel Estimation

Channel estimation problem in the context of DaF relaying basically consists of individual estimation of source-to-relay and relay-to-destination channels. On the other hand, in AaF relaying, a cascaded channel from source-to-destination links needs to be estimated. Moreover, for this relaying mode, knowledge of source-to-relay CSI may also be required at the relay depending on the adopted scaling factor [26], [27].

In Chapter 2 of this thesis, we investigate the performance of RD protocol in AaF mode considering two different pilot-symbol-assisted channel estimation methods which we name as cascaded channel estimation (C-CE) and disintegrated channel estimation (D-CE). In C-CE, the cascaded channel consisting of source-to-relay and relay-to-destination links is estimated at the destination terminal. In D-CE, the estimation of cascaded channel is disintegrated into separate estimations of source-to-relay and relay-to-destination links which are carried out at the relay and destination terminals, respectively. D-CE method involves feed-forwarding a quantized version of the source-to-relay channel estimate to the destination terminal. Considering two different power scaling constraints of [27], we compare the performances of C-CE and D-CE by conducting extensive Monte-Carlo simulations. Our results demonstrate that cascaded channel estimator outperforms its competitor with small number of quantization bits. As the number of employed quantization bits increase, disintegrated channel estimator approaches to its competitor eventually outperforming it. However, this better performance comes at the cost of equipping the relay with channel estimator and vector quantizer. Therefore, C-CE is particularly attractive for low-complexity AaF relaying systems and will be our main focus in Chapter 3 and Chapter 4. The main results of this chapter have been already presented in [28].

1.5.2 Mismatched-Coherent and Partially-Coherent Detection for AaF Relaying

For a coherent wireless communication system, the fading channel coefficients need to be first accurately and efficiently estimated during the training period and then used in

the detection process. In the literature, coherent detection with imperfect channel estimation for AaF relaying has been first addressed by Mheidat and Uysal in [29], [30] and independently by Patel and Stuber in [31]. The main focus in [30] is actually the derivation of a non-coherent detector based on a maximum likelihood sequence estimator for distributed STBCs whereas the performance of a mismatched-coherent receiver (i.e., coherent detection with imperfect channel estimation) is studied as a benchmark. On the other hand, [31] considers a multi-hop relay scenario, derives a channel estimator tailored for cascaded Rayleigh fading channel, and further presents an approximate bit error rate performance analysis for a mismatched-coherent receiver.

In Chapter 3 of this thesis, we investigate the effects of channel estimation on the performance of both RD and TD protocols. Based on C-CE technique, we study mismatched-coherent and partially-coherent receivers at the destination terminal. In mismatched-coherent receiver, the channel coefficients are first estimated through pilot symbols based on a linear minimum-mean-square-error (LMMSE) estimation approach and then fed to a coherent maximum likelihood (ML) decoder as if the channel was perfectly known. In partially-coherent receiver, the estimates of channel phase information are obtained through a phase lock loop (PLL) while no effort is made for the estimation of channel amplitudes. For both receiver types, we analyze the achievable diversity orders for cooperative protocols under consideration and quantify the impact of channel estimation through the derivation of pairwise error probability. Our performance analysis reveals that a second order diversity order is obtained for the considered single-relay scenario indicating that full diversity is extracted. Furthermore, our performance results show that partially-coherent receiver presents a similar performance to mismatched-receiver for sufficiently large loop SNRs although channel amplitude is completely ignored. The main contributions of this chapter have been already presented in [32] and also submitted for publication [33].

1.5.3 Training Power Optimization for AaF Relaying

In the literature, there have been some research efforts on resource optimization for AaF relaying [34]-[36]. In [34], Wang et.al. have aimed to optimize the training and data powers with average received SNR as their objective function. Their transmission scenario is mainly limited to a multi-hop scenario where there is no direct link between the source and destination. On the other hand, in [35], Cui et.al. have considered RD protocol which allows direct link transmission and includes multi-hop scheme as a special case. The focus of [35] is to determine the optimum training pilot symbol sequence and relay precoding structure. Zhang and Gursoy [36] have also considered the RD protocol and formulated a training power allocation problem between relay and source nodes based on the maximization of a lower bound on the capacity. In their work, they assume that channel estimation is carried out both at the relay and destination nodes through D-CE technique. As earlier emphasized, this requires that the relay node should be equipped with a channel estimator. A feed-forward control channel should be deployed as well for the transmission of the quantized source-to-relay channel estimate.

In Chapter 4 of this thesis, we investigate the optimum power allocation scheme for an AaF relaying cooperative system with channel estimation. Similar to [35] and [36], we consider RD protocol, however we assume that the relay node is not equipped with a channel estimator to keep its complexity as low as possible (i.e., C-CE). Based on the maximization of the average received SNR at the destination node, we formulate an optimization problem and taking into account the effect of relay location, we attempt to answer the following questions:

1) How should the overall transmit power be shared between training and data transmission periods?

2) How should training power be allocated to broadcasting and relaying phases?

3) How should data power be allocated to broadcasting and relaying phases?

For the single-relay cooperative system under consideration, we derive the optimum power allocation rules to maximize the average received SNR of the destination node. Our simulation results demonstrate that optimized scheme significantly outperforms the original scheme with equal power allocation. The key results of this chapter have been submitted for publication [37].

Chapter 2

Cascaded versus Disintegrated Channel Estimation for AaF Relaying

Introduction

In this chapter, we investigate the performance of AaF relaying with two different pilot-symbol-assisted channel estimation methods. In the first estimation method, which we call "Cascaded Channel Estimation (C-CE)", the cascaded channel consisting of source-to-relay and relay-to-destination links is estimated at the destination terminal. In the so-called "Disintegrated Channel Estimation (D-CE)" method, the estimation of cascaded channel is disintegrated into separate estimations of source-to-relay and relay-to-destination links which are carried out at the relay and destination terminals, respectively.

2.1 Transmission Model

We consider a cooperative system with half-duplex nodes each of which is equipped with a single pair of transmit and receive antennas (Fig. 2.1). We assume a channel model which takes into account both long-term path loss and short-term fading. The path loss is proportional to d^{-a} where d is the propagation distance and a is the path loss coefficient. Typical values of a range from 1.6 to 6 [38]. Let d_{SD}, d_{SR}, and d_{RD} denote the distances of source-to-destination (S→D), source-to-relay (S→R), and relay-to-destination (R→D) links, respectively, and θ is the angle between lines S→R and R→D.
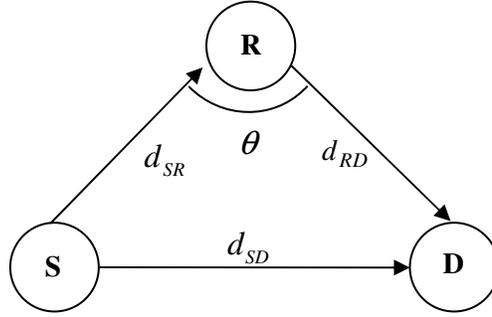
Fig.2.1: Relay-assisted transmission model.

Normalizing the path loss in S→D to be unity, the relative geometrical gain of S→R and R→D links are defined, respectively, as $G_{SR} = (d_{SD}/d_{SR})^a$ and $G_{RD} = (d_{SD}/d_{RD})^a$ [24]. They can be related to each other by $G_{SR}^{-2/a} + G_{RD}^{-2/a} - 2G_{SR}^{-1/a}G_{RD}^{-1/a}\cos\theta = 1$ through the cosine theorem. h_{SR}, h_{SD}, and h_{RD} denote the Rayleigh fading coefficients over S→R, S→D, and R→D links and are modeled as zero-mean complex Gaussian random variables with variance of $\sigma_{h_{SR}}^2$, $\sigma_{h_{SD}}^2$, and $\sigma_{h_{RD}}^2$ under a quasi-static channel assumption.

We consider RD protocol which is earlier described in Section 1.4. Without loss of gene-rality, we assume single-relay network for the brevity of the presentation in this section. In the first signaling interval, the source communicates with the relay and destination terminals. In the second signaling interval, the relay normalizes its received signal and forwards it to the destination. Let x denote the M-PSK (phase shift keying) modulation signal transmitted from the source in the first time slot with power E. Considering path-loss effects, the received signals at relay and destination terminals are given by

$$r_R = \sqrt{G_{SR}E}h_{SR}x + n_R, \tag{2.1}$$

$$r_{D1} = \sqrt{E}h_{SD}x + n_{D1}. \tag{2.2}$$

The relay terminal normalizes the received signal by a factor of $\sqrt{E[|r_R|^2]}$ and then re-transmits the normalized signal within the second time slot. The received signal model at the destination in the second time slot is therefore given by

$$r_{D2} = \sqrt{G_{RD}E}h_{RD}\frac{r_R}{\sqrt{E[|r_R|^2]}} + n_{D2}. \tag{2.3}$$

In (2.1)-(2.3), n_R, n_{D1}, and n_{D2} are the independent samples of zero-mean complex Gaussian random variables with variance $N_0/2$ per dimension and model the additive noise terms.

We consider Average Power Scaling (APS) and Instantaneous Power Scaling (IPS) constraints of [27] for the normalization of the received signal at the relay. In APS constraint, an average output power is maintained by performing the expectation operation in $\sqrt{\mathrm{E}[|r_R|^2]}$ with respect to both n_R and h_{SR}. On the other hand, in IPS power constraint, the expectation is carried only over n_R. Therefore, IPS constraint needs the estimate of h_{SR} to be available at the relay before the normalization of the received signal r_R.

Considering these different power scaling constraints, the received signal in (2.3) can be rewritten as

$$r_{D2} = \begin{cases} \sqrt{\dfrac{G_{SR}G_{RD}E^2}{G_{SR}E\sigma_{h_{SR}}^2 + N_0}}h_{SR}h_{RD}x + \sqrt{\dfrac{G_{RD}E}{G_{SR}E\sigma_{h_{SR}}^2 + N_0}}h_{RD}n_R + n_{D2}, & \text{for APS} \qquad (2.4) \\[4ex] \sqrt{\dfrac{G_{SR}G_{RD}E^2}{G_{SR}E\left|\hat{h}_{SR}\right|^2 + N_0}}h_{SR}h_{RD}x + \sqrt{\dfrac{G_{RD}E}{G_{SR}E\left|\hat{h}_{SR}\right|^2 + N_0}}h_{RD}n_R + n_{D2}, & \text{for IPS} \qquad (2.5) \end{cases}$$

where $\hat{h}_{SR}$ is the estimate of source-to- relay link.

At the receiver side, the received signals are first normalized and then, along with the channel estimates of underlying links are fed to a maximum likelihood (ML) decoder. In the following section, we elaborate on the channel estimation methods under consideration.

2.2 Channel Estimation

We consider two pilot-symbol-assisted approaches for channel estimation in AaF relaying. The first approach is to estimate the cascaded S→R→D channel at the destination terminal. The second one is to disintegrate the cascaded channel into individual links and separately estimate S→R and R→D links at the relay and destination terminals, respectively.

2.2.1 Cascaded Channel Estimation (C-CE)

In this channel estimation method, only the destination is equipped with a channel estimator. Therefore, the relay node cannot employ IPS power constraint which requires the knowledge of h_{SR} to compute the scaling term. For channel estimation purposes, prior to data transmission, the source terminal transmits a pilot symbol p with unit power. Let the received signals at destination terminal during training period be denoted as $r_{D1,p}$ and $r_{D2,p}$ which are simply obtained by replacing x in (2.2) and (2.4) with p. The destination processes received signals $r_{D1,p}$ and $r_{D2,p}$ and obtains channel estimates $\hat{h}_{SD}$ and $\hat{h}_{SRD}$ for S→D and cascaded S→R→D links through LMMSE (Linear Minimum Mean Squared Error Estimation) [39].

During the data transmission, to normalize the received signal r_{D2} in (2.4), the destination uses the scaling factor $\sqrt{1 + G_{RD}E\sigma_{hRD}^2 \big/ (G_{SR}E\sigma_{hSR}^2 + N_0)}$ [23] by replacing $|h_{RD}|^2$ with its variance σ_{hRD}^2 since it does not have access to individual estimate of h_{RD}. This normalization yields

$$\hat{r}_{D2} = \sqrt{B_1 E}\, h_{SR}h_{RD}x + \hat{n}_{D2}, \tag{2.6}$$

where $\hat{n}_{D2}$ is approximated as a zero-mean complex additive Gaussian noise with a variance of N_0 [31], [40] and B_1 is given by

$$B_1 = \frac{G_{SR}G_{RD}\, E/N_0}{1 + \sigma_{hSR}^2 G_{SR}\, E/N_0 + \sigma_{hRD}^2 G_{RD}E/N_0}. \tag{2.7}$$

Defining $\mathbf{r} = \begin{bmatrix} \hat{r}_{D2} & r_{D1} \end{bmatrix}^{\mathrm{T}}$, $\mathbf{n} = \begin{bmatrix} \hat{n}_{D2} & n_{D1} \end{bmatrix}^{\mathrm{T}}$, $\mathbf{h} = \begin{bmatrix} h_{SR}h_{RD} & h_{SD} \end{bmatrix}^{\mathrm{T}}$ and $\mathbf{X}_D = diag\left(\sqrt{B_1 E}, \sqrt{E}\right)x$, the received signals over two time slots can be rewritten in a compact matrix form as $\mathbf{r} = \mathbf{X}_D\mathbf{h} + \mathbf{n}$. The received signals along with channel estimate vector $\hat{\mathbf{h}} = [\ \hat{h}_{SRD} \quad \hat{h}_{SD}]$ are fed to a ML mismatched-coherent decoder [41] given by

$$\arg \min_{\mathbf{X}_D} \ \left\| \mathbf{r}\text{-}\mathbf{X}_D\hat{\mathbf{h}} \right\|^2. \tag{2.8}$$

2.2.2 Disintegrated Channel Estimation (D-CE)

In D-CE, both relay and destination nodes are equipped with an LMMSE channel estimator. Therefore, the relay can obtain an estimate of S→R channel, i.e. $\hat{h}_{SR}$, through pilot symbols. Before sending this estimate to the destination through a feed-forward control channel, the relay quantizes $\hat{h}_{SR}$ as $\hat{h}_{SR,b}$ where b indicates the number of quantization bits per dimension (assuming the deployment of a vector quantizer). Moreover, the relay also transmits a "clean" pilot symbol so that R→D link can be later estimated as $\hat{h}_{RD}$ at the destination.

In this channel estimation method, possessing $\hat{h}_{SR}$ information enables the relay to employ not only APS but also IPS power constraint. Considering APS and IPS power constraints, the destination normalizes r_{D2} in (2.4) and (2.5) either by $\sqrt{1 + G_{RD}E \mid \hat{h}_{RD} \mid^2 / (G_{SR}E\sigma_{hSR}^2 + N_0)}$ or $\sqrt{1 + G_{RD}E \mid \hat{h}_{RD} \mid^2 / (G_{SR}E \mid \hat{h}_{SR,b} \mid^2 + N_0)}$, yielding

$$\hat{r}_{D2} = \begin{cases} \sqrt{B_2 E} h_{SR} h_{RD} x + \hat{n}_{D2}, & \text{for APS} \\ \sqrt{B_3 E} h_{SR} h_{RD} x + \hat{n}_{D2}, & \text{for IPS} \end{cases}$$

$$(2.9)$$
$$(2.10)$$

where $\hat{n}_{D2}$ is approximated as a zero-mean complex additive Gaussian noise with a variance of N_0 [31], [40]. In (2.9) and (2.10), B_2 and B_3 are given by

$$B_2 = \frac{G_{SR}G_{RD} E/N_0}{1 + \sigma_{hSR}^2 G_{SR} E/N_0 + \left|\hat{h}_{RD}\right|^2 G_{RD}E/N_0}, \tag{2.11}$$

$$B_3 = \frac{G_{SR}G_{RD} E/N_0}{1 + \left|\hat{h}_{SR,b}\right|^2 G_{SR} E/N_0 + \left|\hat{h}_{RD}\right|^2 G_{RD} E/N_0} \cdot \frac{1 + \left|\hat{h}_{SR,b}\right|^2 G_{SR} E/N_0}{1 + \left|\hat{h}_{SR}\right|^2 G_{SR} E/N_0}. \tag{2.12}$$

Using the quantized S→R channel information, the destination terminal constructs the estimate for the cascaded link as $\hat{h}_{SRD} = \hat{h}_{SR,b}\hat{h}_{RD}$. The received signals over two time slots can be rewritten as $\mathbf{r} = \mathbf{X}_D\mathbf{h} + \mathbf{n}$. The received signals are then fed to the decoder given by (2.8) where the channel estimate vector is defined as $\hat{\mathbf{h}} = [\hat{h}_{SR,b}\hat{h}_{RD} \quad \hat{h}_{SD}]$.

2.3 Simulation Results

In this section, we present Monte Carlo simulations to demonstrate the error rate performance of AaF relaying with C-CE and D-CE methods under consideration. In our simulations, we consider $a = 2$, $\theta = \pi$, and 4-PSK modulation, and take the channel variances as $\sigma_{h_{SR}}^2 = \sigma_{h_{RD}}^2 = \sigma_{h_{SD}}^2 = 1$. For quantization and transmission of $\hat{h}_{SR}$ which is required for D-CE method, we employ a vector quantizer with Generalized Lloyd Algorithm [42] and assume an error-free and delay-free feedforward control channel.

In Figs. 2.2-2.5, we present the bit error rate (BER) performance of a single-relay AaF scheme for different relay locations. Specifically, we consider two scenarios where the relay is located close to the source ($G_{SR}/G_{RD} = 30 \, \text{dB}$) and close to the destination ($G_{SR}/G_{RD} = -30 \, \text{dB}$). In these figures, we illustrate the performance of AaF relaying with C-CE and D-CE methods along with APS and IPS power constraints assuming the availability of various number of quantization bits. The performances of genie-aided receivers for C-CE (i.e., perfect h_{SD} and $h_{SR}h_{RD}$ knowledge available at the destination) and D-CE (i.e., perfect h_{SD}, h_{SR}, and h_{RD} knowledge available at the destination) under APS and IPS constraints are also included as benchmarks. From the figures, we can make the following observations:

a. C-CE is superior to D-CE (either with APS or IPS) which experiences an error floor when the number of quantization bits b is less than 2 bits/dimension (see Figs 2.2-2.5). As b increases, quantization error experienced at the destination decreases and the performance of D-CE significantly improves eventually outperforming C-CE. At a target bit error rate of 10^{-3}, assuming $b = 5$ bits/dimension, D-CE with APS outperforms C-CE by 0.8 dB and 2.5dB, respectively, for $G_{SR}/G_{RD} = 30 \, \text{dB}$ and $G_{SR}/G_{RD} = -30 \, \text{dB}$, while the performance difference between D-CE with IPS and C-CE is 2.5dB for both relay locations.

b. For D-CE, $b = 5$ bits/dimension is sufficient to provide a nearly identical performance with *perfectly quantized* $\hat{h}_{SR}$ (which requires an infinite number of quantization bits).

c. For relay locations close to the source, D-CE with IPS outperforms D-CE with APS (see Figs 2.2 and 2.3). Specifically, at BER=10^{-3}, their performance difference is approximately 2dB for $b = 5$ bits/dimension. However, as b decreases, the performance of APS approaches to that of IPS and eventually, they show nearly identical performance.

d. When the relay is located close to the destination, the performance of D-CE with APS coincides with that of D-CE with IPS when the number of quantization bits is large (see Figs 2.4 and 2.5). As b decreases, APS becomes superior to IPS. At a target bit error rate of 2.10^{-2}, APS with $b = 1$ bits/dimension outperforms IPS by approximately 3.8dB. Our observations indicate that under the considered channel estimation method, IPS is advantageous only when the relay is close to source and large quantization bits are employed. When the relay is far away from the source and feedback is limited to a few bits, IPS has no gain over APS.

e. For *perfectly quantized* $\hat{h}_{SR}$, we can observe that DCE with APS achieves its best performance for $G_{SR}/G_{RD} = -30$ dB (see Figs 2.2 and 2.4).

f. For perfect quantization case, the performances of D-CE with APS and that of with IPS are observed to be the same when the relay is located close to the destination (See Figs 2.4 and 2.5). This can be explained by the fact that (2.11) and (2.12) simply reduce to $B_2 = B_3 = 1/\left|\hat{h}_{RD}\right|^2$ for this relay location (i.e., $G_{RD} \gg 1$, $G_{SR} \approx 1$). Moreover, by comparing Figs 2.3 and 2.5, it can be also observed that the performance of D-CE with IPS is identical for both $G_{SR}/G_{RD} = -30$ dB and $G_{SR}/G_{RD} = 30$ dB. Same conclusions can be driven by considering the corresponding genie receivers.
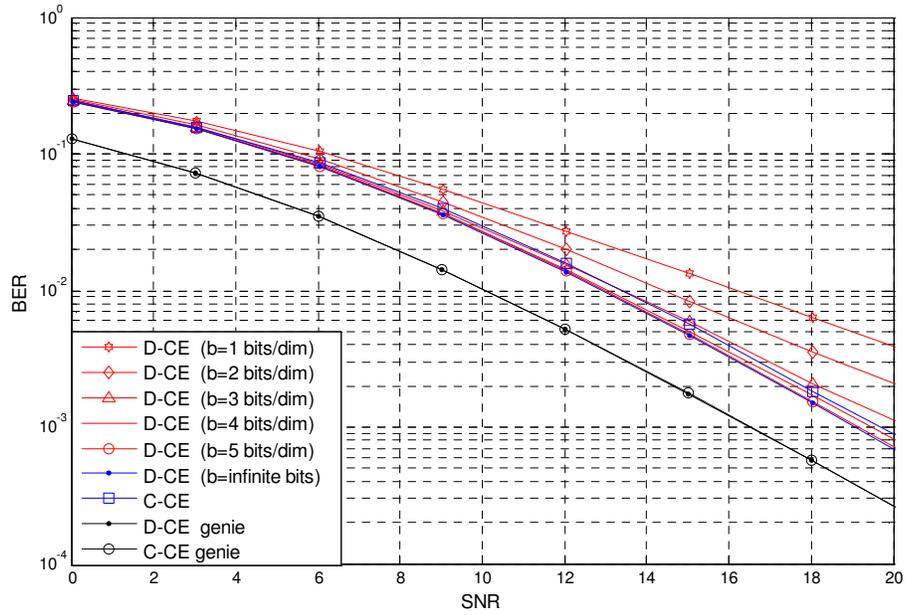
Fig.2.2: BER of C-CE and D-CE with APS in a single-relay AaF network for $G_{SR}/G_{RD} = 30\,\text{dB}$.
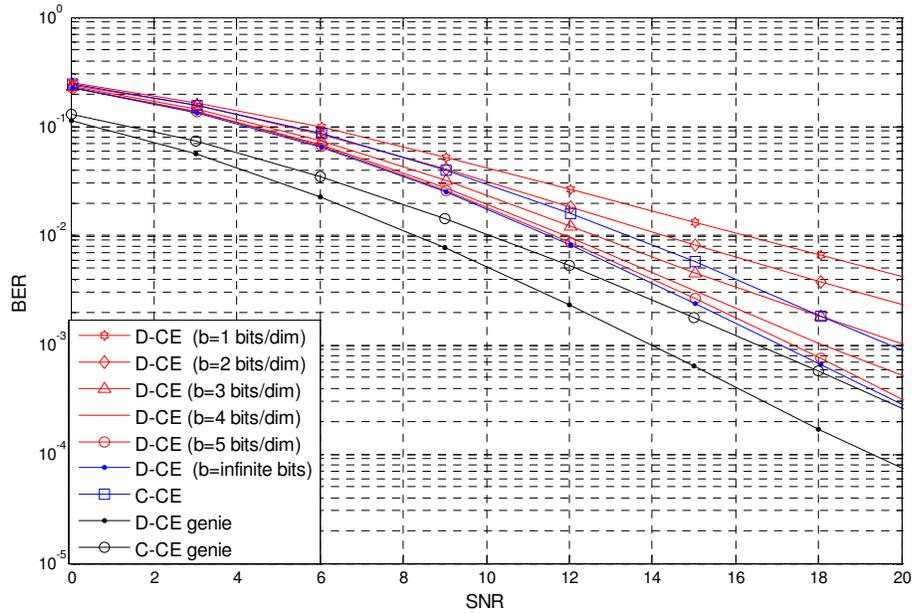
Fig.2.3: BER of C-CE and D-CE with IPS in a single-relay AaF network for $G_{SR}/G_{RD} = 30\,\text{dB}$.
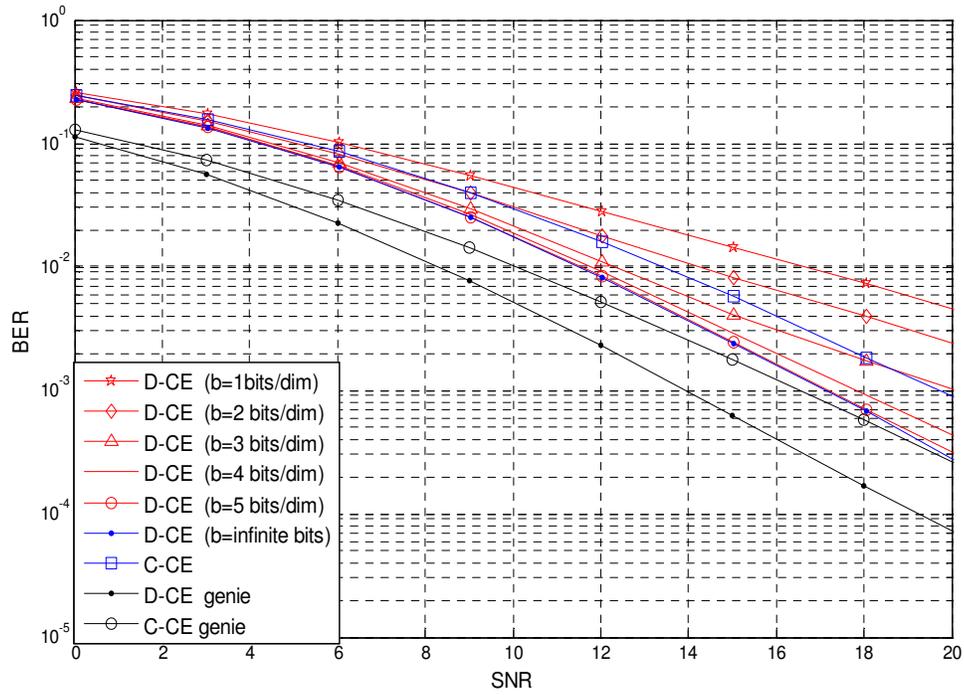
Fig.2.4: BER of C-CE and D-CE with APS in a single-relay AaF network for $G_{SR}/G_{RD} = -30\,\text{dB}$.
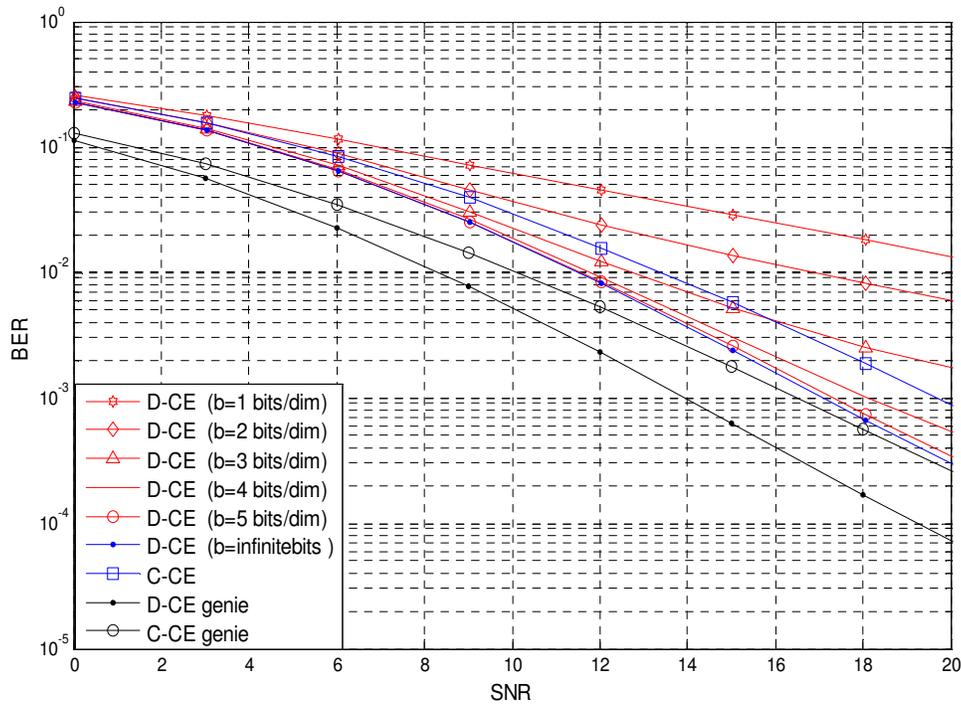
Fig.2.5: BER of C-CE and D-CE with IPS in a single-relay AaF network for $G_{SR}/G_{RD} = -30\,\text{dB}$.

In Figs. 2.6 and 2.7, we demonstrate the BER performance of AaF relaying with multiple relays. Specifically, we consider two and four relays which apply Alamouti space-time block coding (STBC) [12] and rotated quasi-orthogonal STBC [43] across their received signals before forwarding to the destination. From these figures, it is observed that for lower values of quantization bits (c.f., plots for b=1,2,3), severe error floors are experienced by D-CE due to quantization error accumulation at the destination, which is proportional to the number of relay nodes. However, as b increases, error floors are observed only in very high SNRs and performance of D-CE improves. For 2-relay case, D-CE with b=4 is able to outperform the C-CE. For 4-relay case, it requires b=5 quantization bits/dimension for D-CE to outperform its competitor.
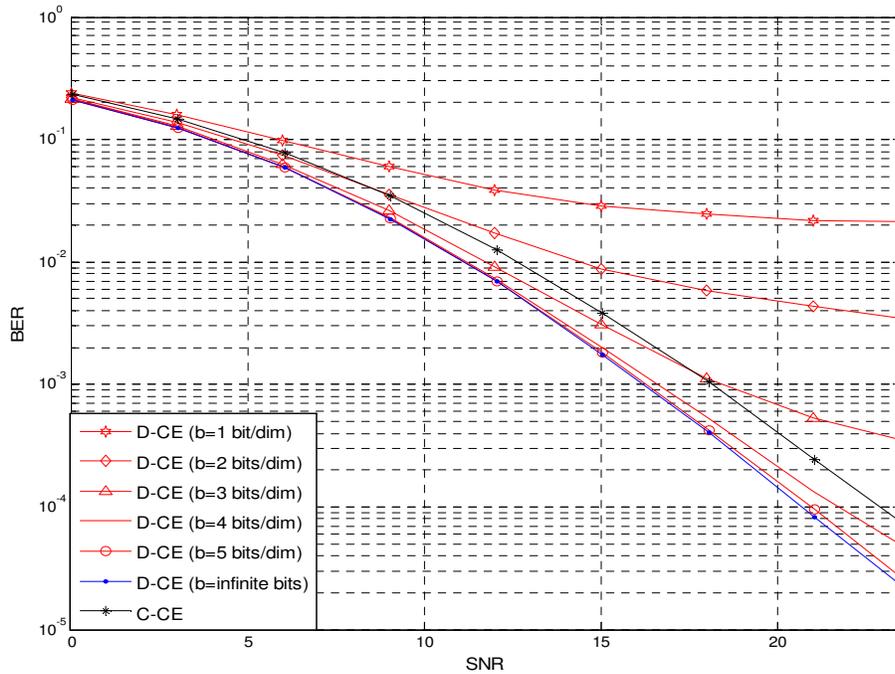
Fig. 2.6: BER of C-CE and D-CE in an AaF network with two relays. Alamouti STBC is applied $(G_{SR}/G_{RD} = 30\,\text{dB})$.
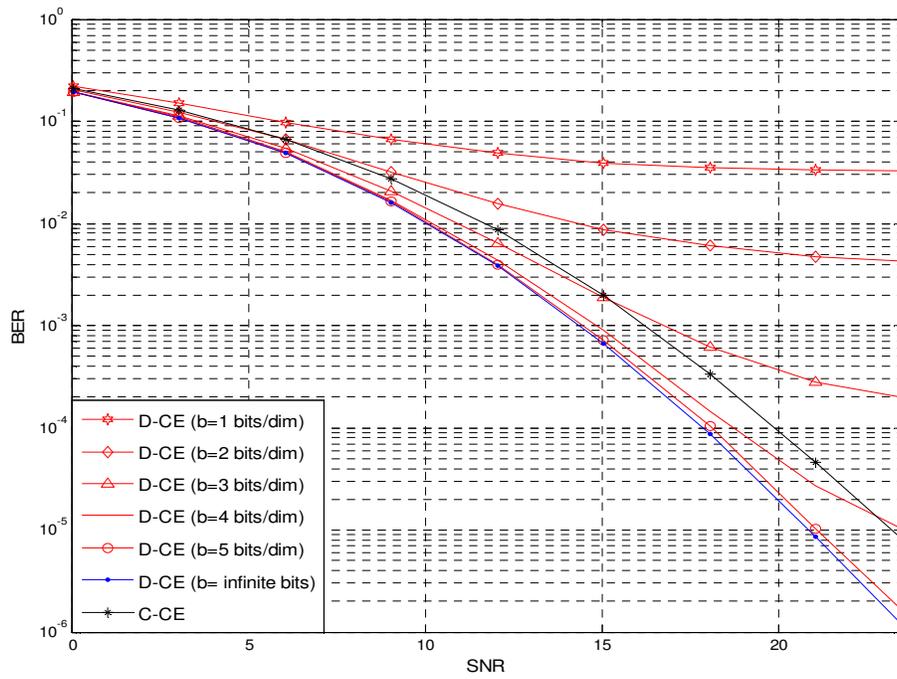
Fig. 2.7: BER of C-CE and D-CE in an AaF network with four relays. Rotated quasi-orthogonal STBC is applied ($G_{SR}/G_{RD} = 30\,\text{dB}$).

Chapter 3

Performance Analysis of AaF Relaying with Imperfect Channel Estimation

Introduction

Our results in Chapter 2 demonstrate that C-CE is more advantageous when the quantization rate in D-CE is limited to a few bits. As the quantization rate increases, D-CE is able to outperform its competitor. However, it should be kept in mind that D-CE's better performance comes at the cost of equipping the relay with channel estimator and vector quantizer. Therefore, C-CE becomes particularly attractive for low-complexity AaF relaying and it will be our choice of channel estimation technique in this chapter.

The aim of the current chapter is to analyze the performance of AaF relaying and determine the available diversity in the presence of imperfect channel estimation. We consider both RD and TD protocols and assume two receiver architectures: In the mismatched-coherent receiver, which has been also used in Chapter 2, the complex fading channel coefficients (i.e., both phase and amplitude) are estimated based on LMMSE approach and fed to a coherent ML decoder as if the channels were perfectly known. In the partially-coherent receiver, channel amplitude is ignored and channel phase is estimated by a phase locked loop. For both receiver types, we derive pairwise error probability (PEP) expressions to determine the achievable diversity orders for cooperative protocols under consideration and to quantify the impact of channel estimation. Monte Carlo simulations are further conducted to provide a comprehensive performance comparison among the receiver structures under consideration.

3.1 Transmission Model

Our transmission model is based on a single-relay scenario where each of the half-duplex nodes is equipped with a single pair of transmit and receive antennas (Fig.2.1). We consider RD and TD cooperation protocols which are described in Section 1.4. As earlier emphasized, signal conveyed to the relay and destination terminals over the two time slots is same for RD protocol whereas TD protocol can convey different signals to the relay and destination terminals. Therefore, TD protocol can successfully deploy various conventional space-time codes in a distributed fashion.

3.1.1 TD Protocol

Although any conventional space-time code can, in principle, be used in conjunction with TD protocol, we consider the Golden code of [44], which has been recently shown to achieve optimum diversity-multiplexing tradeoff in the single relay AaF case [45]. Let $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^T$ denote M-PSK modulation signals with normalized power, i.e., $E[|x_i|^2] = 1$. Before transmission, the modulation signals are fed into a encoder given by

$$\mathbf{P} = \begin{bmatrix} \alpha & \alpha\Theta & 0 & 0 \\ 0 & 0 & i\overline{\alpha} & i\overline{\alpha}\overline{\Theta} \\ 0 & 0 & \alpha & \alpha\Theta \\ \overline{\alpha} & \overline{\alpha}\overline{\Theta} & 0 & 0 \end{bmatrix} \tag{3.1}$$

where $\Theta = \left(1 + \sqrt{5}\right)/2$, $\overline{\Theta} = \left(1 - \sqrt{5}\right)/2$, $\alpha = \left(1 + i - i\Theta\right)/\sqrt{5}$, $\overline{\alpha} = \left(1 + i - i\overline{\Theta}\right)/\sqrt{5}$. The output of the encoder is given by $\mathbf{c} = \mathbf{P}\mathbf{x} = \begin{bmatrix} c_1 & c_2 & c_3 & c_4 \end{bmatrix}^T$ where $c_1 = \alpha(x_1 + \Theta x_2)$, $c_2 = i\overline{\alpha}(x_3 + \overline{\Theta}x_4)$, $c_3 = \alpha(x_3 + \Theta x_4)$ and $c_4 = \overline{\alpha}(x_1 + \overline{\Theta}x_2)$. In the first signaling interval, the codeword c_1 is transmitted from the source with power E. Considering path-loss effects, the received signals at the relay and destination are given as

$$r_{R1} = \sqrt{G_{SR}E} h_{SR} c_1 + n_{R1} \ , \tag{3.2}$$

$$r_{D1} = \sqrt{E} h_{SD} c_1 + n_{D1} \ . \tag{3.3}$$

The relay terminal normalizes the received signal by a factor of $\sqrt{\mathrm{E}[|r_{R1}|^2]}$ to have average unit power and then re-transmits the normalized signal within the second time slot. The source terminal simultaneously transmits the codeword c_2. Therefore, the destination receives a superposition of the signals transmitted by the relay and source transmission as

$$r_{D2} = \sqrt{G_{RD}\frac{E}{2}}h_{RD}\frac{r_{R1}}{\sqrt{\mathrm{E}[|r_{R1}|^2]}} + \sqrt{\frac{E}{2}}h_{SD}c_2 + n \tag{3.4}$$

where scaling by $1/\sqrt{2}$ is included to ensure the total power consumption limited by E in a particular time slot. In (3.2)-(3.4), h_{SR}, h_{SD} and h_{RD} are modeled as zero-mean complex Gaussian random variables with variance of 0.5 per dimension. n_R, n_{D1} and n are the independent samples of zero-mean complex Gaussian noise with variance $N_0/2$ per dimension. Considering APS power constraint, we insert $\sqrt{\mathrm{E}[|r_{R1}|^2]} = \sqrt{G_{SR}E\kappa + N_0}$ with $\kappa = |\alpha|^2\left(1 + \Theta^2\right)$ in (3.4) and obtain

$$r_{D2} = \sqrt{\frac{0.5 G_{SR}G_{RD}E^2}{G_{SR}E\kappa + N_0}}h_{SR}h_{RD}c_1 + \sqrt{0.5E}h_{SD}c_2 + \tilde{n}_1 \tag{3.5}$$

where $\tilde{n}_1 = \sqrt{0.5 G_{RD}E/\left(G_{SR}E\kappa + N_0\right)}h_{RD}n_{R1} + n$. Due to the term involving $h_{RD}n_{R1}$, $\tilde{n}_1$ is of non-Gaussian nature which makes the analysis intractable. However, as in [31], [40], [52], we can treat it as Gaussian noise with the same average power. Therefore, $\tilde{n}_1$ is assumed to be complex Gaussian with statistics $\tilde{n}_1 \sim CN\left(0, N_0\left(1 + 0.5 G_{RD}E\sigma_{hRD}^2 \big/\left(G_{SR}E\kappa + N_0\right)\right)\right)$. The destination terminal normalizes r_{D2} by a factor of $\sqrt{1 + \left(0.5 G_{RD}E\right)\sigma_{hRD}^2 \big/\left(G_{SR}E\kappa + N_0\right)}$ yielding

$$r_{D2} = \sqrt{A_1 E}h_{SR}h_{RD}c_1 + \sqrt{A_2 E}h_{SD}c_2 + n_{D2} \tag{3.6}$$

where n_{D2} is assumed to be zero mean complex Gaussian random variable with a variance of N_0. In (3.6), A_1 and A_2 are defined as

$$A_1 = \frac{0.5 G_{SR}G_{RD}\,E/N_0}{1 + G_{SR}\kappa E/N_0 + 0.5 G_{RD}\sigma_{hRD}^2\,E/N_0} \quad, \tag{3.7}$$

$$A_2 = \frac{0.5\left(1 + G_{SR}\kappa E/N_0\right)}{1 + G_{SR}\kappa E/N_0 + 0.5 G_{RD}\sigma^2_{h_{RD}} E/N_0} \ . \tag{3.8}$$

The codewords c_3 and c_4 are transmitted from the source in the third and fourth time slots, respectively. The corresponding received signal models are obtained, similar to (3.3) and (3.6), as

$$r_{D3} = \sqrt{E}h_{SD}c_3 + n_{D3} \ , \tag{3.9}$$

$$r_{D4} = \sqrt{A_1 E}h_{SR}h_{RD}c_3 + \sqrt{A_2 E}h_{SD}c_4 + n_{D4} . \tag{3.10}$$

Defining $\quad \mathbf{r} = \begin{bmatrix} r_{D1} & r_{D2} & r_{D3} & r_{D4} \end{bmatrix}^{\mathrm{T}}, \quad \mathbf{n} = \begin{bmatrix} n_{D1} & n_{D2} & n_{D3} & n_{D4} \end{bmatrix}^{\mathrm{T}} \quad \mathbf{h} = \begin{bmatrix} h_{SR}h_{RD} & h_{SD} \end{bmatrix}^{\mathrm{T}}$ $= \begin{bmatrix} h_{SRD} & h_{SD} \end{bmatrix}^{\mathrm{T}}$, the received signals over four time slots can be rewritten in a compact matrix form as $\mathbf{r} = \mathbf{X}_D\mathbf{h} + \mathbf{n}$ where $\mathbf{X}_D$ is given by

$$\mathbf{X}_D = \begin{bmatrix} 0 & \sqrt{A_1 E}c_1 & 0 & \sqrt{A_1 E}c_3 \\ \sqrt{E}c_1 & \sqrt{A_2 E}c_2 & \sqrt{E}c_3 & \sqrt{A_2 E}c_4 \end{bmatrix}^{\mathrm{T}} . \tag{3.11}$$

3.1.2 RD Protocol

As earlier described in Chapter 2, RD protocol realizes receive diversity in a distributed fashion. Let x be the M-PSK signal transmitted by the source. The received signals at the destination are given by

$$r_{D1} = \sqrt{E}h_{SD}x + n_{D1} \tag{3.12}$$

$$r_{D2} = \sqrt{B_1 E}h_{SR}h_{RD}x + n_{D2} \tag{3.13}$$

where B_1 is defined as

$$B_1 = \frac{G_{SR}G_{RD} E/N_0}{1 + G_{SR} E/N_0 + \sigma^2_{h_{RD}}G_{RD} E/N_0} \ . \tag{3.14}$$

Defining $\mathbf{r} = \begin{bmatrix} r_{D1} & r_{D2} \end{bmatrix}^{\mathrm{T}}$ and $\mathbf{n} = \begin{bmatrix} n_{D1} & n_{D2} \end{bmatrix}^{\mathrm{T}}$, the received signals can be written in a matrix form as $\mathbf{r} = \mathbf{X}_D\mathbf{h} + \mathbf{n}$ where $\mathbf{X}_D$ is given by

$$\mathbf{X}_D = \begin{bmatrix} 0 & \sqrt{E} \\ \sqrt{B_1 E} & 0 \end{bmatrix} x . \tag{3.15}$$

3.2 Mismatched-Coherent and Partially-Coherent Receivers

In this section, we consider two different pilot-symbol-assisted receiver architectures: In the first receiver, the complex fading channel coefficients (i.e., both phase and amplitude) are estimated based on a LMMSE approach and fed to a ML mismatched-coherent receiver [41]. In the second receiver, channel amplitude is ignored. Only phase information of the channels is estimated by a PLL and these estimates are used in a partially-coherent receiver [46].

Let $\mathbf{X}_{lT}$, $l = 1,2,...N$, and $\mathbf{X}_{jD}$, $j = 1,2,...M$ denote the pilot and data matrices transmitted by the source terminal at transmission blocks l and j. Here, N and M denote the number of training and data transmission blocks, respectively. The length of data transmission block is equal to the codeword length which is 2 and 4 for RD and TD protocols, c.f., (3.11) and (3.15). The received signal is therefore given by

$$\mathbf{X}_{Tot} = [\underbrace{\mathbf{X}_{1T} \cdots \mathbf{X}_{NT}}_{\mathbf{X}_T} \quad \underbrace{\mathbf{X}_{1D} \cdots \mathbf{X}_{MD}}_{\mathbf{X}_D}]^{\mathrm{T}}. \tag{3.16}$$

3.2.1 Mismatched-coherent detection with LMMSE estimator

The LMMSE estimate of the channel matrix can be obtained as $\hat{\mathbf{h}} = \mathbf{B}\mathbf{r}_T$ where $\mathbf{r}_T = \mathbf{X}_T \mathbf{h} + \mathbf{n}_T$ is the received signal during the training period and $\mathbf{B}$ is a matrix obtained through the minimization of $\mathrm{E}[\|\mathbf{B}\mathbf{r}_T - \mathbf{h}\|^2]$. This minimization yields [39]

$$\mathbf{B} = \mathrm{E}(\mathbf{h}\mathbf{r}_T^{\mathrm{H}})\mathrm{E}(\mathbf{r}_T \mathbf{r}_T^{\mathrm{H}})^{-1}. \tag{3.17}$$

Using $\mathrm{E}(\mathbf{h}\mathbf{h}^{\mathrm{H}}) = \mathbf{I}_2$, the channel estimate $\hat{\mathbf{h}} = \mathbf{h} - \mathbf{e}$ is obtained as

$$\hat{\mathbf{h}} = \begin{bmatrix} \hat{h}_{SRD} & \hat{h}_{SD} \end{bmatrix}^{\mathrm{T}} = \mathbf{X}_T^{\mathrm{H}}(\mathbf{X}_T \mathbf{X}_T^{\mathrm{H}} + N_0 \mathbf{I}_2)^{-1} \mathbf{r}_T. \tag{3.18}$$

The covariance matrix of estimation errors $\mathbf{e} = \begin{bmatrix} e_{SRD} & e_{SD} \end{bmatrix}^T$ is given by

$$\mathbf{C}_{\mathbf{e}} = \mathrm{E}(\mathbf{h}\mathbf{h}^{\mathrm{H}}) - \mathbf{B}\mathrm{E}(\mathbf{r}_T \mathbf{r}_T^{\mathrm{H}})\mathbf{B}^{\mathrm{H}} = N_0(\mathbf{X}_T^{\mathrm{H}}\mathbf{X}_T + N_0 \mathbf{I}_2)^{-1} \tag{3.19}$$

The channel estimate $\hat{\mathbf{h}}$ is then used to minimize the following coherent ML metric

$$\arg \min_{\mathbf{X}_D} \|\mathbf{r} - \mathbf{X}_D \hat{\mathbf{h}}\|^2 \tag{3.20}$$

as if the channel was perfectly known.

3.2.2 Partially-coherent detection with PLL estimator

Let $h_{SRD} = |h_{SRD}| e^{j\phi_{SRD}}$ and $h_{SD} = |h_{SD}| e^{j\phi_{SD}}$ be the polar coordinates representations of the complex fading coefficients. ϕ_{SRD} and φ_{SD} are the phases introduced by S→R→D and S→D channels. We assume a first-order PLL at the destination terminal. Upon receiving the signal $\mathbf{r}_T$ during the training phase, PLL first compares the phases of the input signal and the locally generated oscillator output, then generates a control signal that is a function of the phase difference which is minimized to produce phase estimates $\hat{\varphi}_{SD}$ and $\hat{\varphi}_{SRD}$. The estimation errors are denoted as $\varepsilon_{SRD} = \phi_{SRD} - \hat{\phi}_{SRD}$ and $\varepsilon_{SD} = \phi_{SD} - \hat{\phi}_{SD}$ for S→R→D and S→D links, respectively. Their distribution can be well approximated by Tikhonov probability density function [47]

$$p_\varepsilon(\varepsilon) = \frac{\exp(\rho_L \cos\varepsilon)}{2\pi I_0(\rho_L)} \qquad -\pi \le \varepsilon \le \pi \qquad (3.21)$$

where we drop the index S→R→D and S→D for notational convenience. Here, $I_0(.)$ is the zeroth order modified Bessel function of the first kind [48] and ρ_L is the loop signal-to-noise ratio defined as [47]

$$\rho_L = \frac{\xi}{B_L T} \qquad (3.22)$$

where T denotes the symbol duration, B_L is the loop bandwidth, and ξ is the instantaneous received signal-to-noise-ratio [55]. If PLL is assumed to be in lock position, ε is sufficiently small; therefore, phase errors can be approximated as zero mean Gaussian random variables with variance $\sigma_\varepsilon^2 = 1/\rho_L$. These phase estimates are used in partially-coherent detection to minimize the metric

$$\arg\ \min_{\mathbf{X}_D} \|\mathbf{r} - \mathbf{X}_D \hat{\varphi}\|^2 \qquad (3.23)$$

where $\hat{\varphi} = \begin{bmatrix} e^{j\hat{\varphi}_{SRD}} & e^{j\hat{\varphi}_{SD}} \end{bmatrix}^{\mathrm{T}}$ and channel amplitudes are taken equal to one [46].

3.3 Diversity Gain Analysis

In this section, we investigate the achievable diversity order for the cooperative schemes under consideration through the derivation of PEP. PEP is the building block for the derivation of union bounds to the error probability. It is widely used in the literature to predict the attainable diversity order where the closed-form error probability expressions are unavailable. Let $P(\mathbf{X}_D \to \hat{\mathbf{X}}_D)$ denote PEP where the transmitted codeword vector and the erroneously-decoded codeword matrices are given by $\mathbf{X}_D$ and $\hat{\mathbf{X}}_D$, respectively. Following the derivation steps in Appendix A, PEP for TD protocol with Golden code and mismatched-coherent receiver can be found as

$$
\begin{aligned}
P(\mathbf{X}_D \to \hat{\mathbf{X}}_D) &< \frac{4 SNR_{eff}^{-2}/(\lambda_{\min}^2 \kappa_1 \kappa_2)}{\left(1-\sigma_{e_{SD}}^2\right)\left(1-\sigma_{e_{SR}}^2\right)\left(1-\sigma_{e_{RD}}^2\right)} \exp\left(\frac{2 SNR_{eff}^{-1}/(\lambda_{\min}\kappa_2)}{\left(1-\sigma_{e_{SR}}^2\right)\left(1-\sigma_{e_{RD}}^2\right)}\right) \\
&\quad \times \Gamma\left(0, \frac{2 SNR_{eff}^{-1}/(\lambda_{\min}\kappa_2)}{\left(1-\sigma_{e_{SR}}^2\right)\left(1-\sigma_{e_{RD}}^2\right)}\right)
\end{aligned}
\tag{3.24}
$$

where $SNR_{eff} = E/(2\Lambda)$. Here, Λ is trace of the covariance matrix for effective noise which contains both additive Gaussian channel noise and channel estimation error and is given by (A.5) of Appendix A. In the above, κ_1, κ_2 are defined by $\kappa_1 = \kappa + A_2 \kappa'$ and $\kappa_2 = A_1 \kappa$ where $\kappa = |\alpha|^2(1+\Theta^2)$ and $\kappa' = |\overline{\alpha}|^2(1+\overline{\Theta}^2)$. $\lambda_{\min}$ denotes the minimum value of $\lambda_1 + \lambda_3$ and $\lambda_2 + \lambda_4$ where $\lambda_1, \lambda_2, \lambda_3$ and λ_4 are the eigenvalues of $(\mathbf{x}-\hat{\mathbf{x}})(\mathbf{x}-\hat{\mathbf{x}})^{\mathrm{H}}$. $\sigma_{e_{SD}}^2$, $\sigma_{e_{SR}}^2$ and $\sigma_{e_{RD}}^2$ denote variances of S→D, S→R and R→D channel estimation errors and are, respectively, given by

$$
\sigma_{e_{SD}}^2 = (E/N_0)^{-1}/\left(N(2\kappa+\kappa')+(E/N_0)^{-1}\right),
\tag{3.25}
$$

$$
\sigma_{e_{SR}}^2 = (E/N_0)^{-1}/\left(2NG_{SR}\kappa+(E/N_0)^{-1}\right),
\tag{3.26}
$$

$$
\sigma_{e_{RD}}^2 = (E/N_0)^{-1}/\left(NG_{RD}\kappa+(E/N_0)^{-1}\right).
\tag{3.27}
$$

For large SNR_{eff} values, exponential term in (3.24) goes to zero and we can use the approximation $\lim_{t \to 0} \Gamma(0,t) \approx -\log(t)$ for the gamma term [50]. Then, $P(\mathbf{X}_D \to \hat{\mathbf{X}}_D)$ reduces to

$$P\left(\mathbf{X}_D \to \hat{\mathbf{X}}_D\right) < \frac{4SNR_{eff}^{-2}}{\lambda_{\min}^2 \kappa_1 \kappa_2 \left(1-\sigma_{e_{SD}}^2\right)\left(1-\sigma_{e_{SR}}^2\right)\left(1-\sigma_{e_{RD}}^2\right)} \log\left(\frac{SNR_{eff}\lambda_{\min}\kappa_2\left(1-\sigma_{e_{SR}}^2\right)\left(1-\sigma_{e_{RD}}^2\right)}{2}\right) \quad (3.28)$$

Under high SNR_{eff} assumption, $\log(SNR_{eff})$ term can be ignored with respect to the dominating term SNR_{eff}^{-2}. Thus, asymptotically, second order diversity is achieved, extracting the full diversity for the considered scenario with single relay. We observe that the presence of channel estimation errors does not affect the diversity order. The PEP expression for the perfect CSI case can be simply obtained when the estimation error variances become zero. Let $P_{genie}\left(\mathbf{X}_D \to \hat{\mathbf{X}}_D\right)$ denote the PEP for perfect CSI, then the performance degradation due to channel estimation is given by

$$\Xi = \frac{P\left(\mathbf{X}_D \to \hat{\mathbf{X}}_D\right)}{P_{genie}\left(\mathbf{X}_D \to \hat{\mathbf{X}}_D\right)}$$

$$= \frac{(\Lambda/N_0)^2}{\left(1-\sigma_{e_{SD}}^2\right)\left(1-\sigma_{e_{SR}}^2\right)\left(1-\sigma_{e_{RD}}^2\right)} \log\left(\frac{E\lambda_{\min}\kappa_2}{4}\left[\frac{\left(1-\sigma_{e_{SR}}^2\right)\left(1-\sigma_{e_{RD}}^2\right)}{\Lambda} - \frac{1}{N_0}\right]\right). \quad (3.29)$$

By increasing the number of pilot symbols or pilot symbol power, estimation error variances approach zero (i.e., $\sigma_{e_{SRD}}^2$, $\sigma_{e_{SD}}^2 \to 0$) leading to $\log \Xi = \log 1 = 0$.

PEP for RD protocol with mismatched coherent receiver is given by (see Appendix B for details of the derivation)

$$P\left(\mathbf{X}_D \to \hat{\mathbf{X}}_D\right) \leq \frac{4SNR_{eff}^{-2}/\left(\lambda^2 B_1\right)}{\left(1-\sigma_{e_{SD}}^2\right)\left(1-\sigma_{e_{SR}}^2\right)\left(1-\sigma_{e_{RD}}^2\right)} \exp\left(\frac{2SNR_{eff}^{-1}/\left(\lambda B_1\right)}{\left(1-\sigma_{e_{SR}}^2\right)\left(1-\sigma_{e_{RD}}^2\right)}\right) \Gamma\left(0, \frac{2SNR_{eff}^{-1}/\left(\lambda B_1\right)}{\left(1-\sigma_{e_{SR}}^2\right)\left(1-\sigma_{e_{RD}}^2\right)}\right) \quad (3.30)$$

where $\lambda = |x-\hat{x}|^2$ and $SNR_{eff} = E/(2\Lambda)$. Λ contains both channel estimation error and additive noise variances and is given by (A.18) of Appendix B. In (3.30), $\sigma_{e_{SD}}^2$, $\sigma_{e_{SR}}^2$ and $\sigma_{e_{RD}}^2$ denote variances of S→D, S→R and R→D channel estimation errors for this receiver type and are, respectively, given by

$$\sigma_{e_{SD}}^2 = \left(E/N_0\right)^{-1}/\left(N + \left(E/N_0\right)^{-1}\right), \quad (3.31)$$

$$\sigma_{e_{SR}}^2 = \left(E/N_0\right)^{-1}/\left(NG_{SR} + \left(E/N_0\right)^{-1}\right), \quad (3.32)$$

$$\sigma_{e_{RD}}^2 = \left(E/N_0\right)^{-1} \big/ \left(NG_{RD} + \left(E/N_0\right)^{-1}\right). \tag{3.33}$$

For large SNR_{eff} values, (3.30) reduces to

$$P\left(\mathbf{X}_D \to \hat{\mathbf{X}}_D\right) \le \frac{4SNR_{eff}^{-2} \big/ \left(\lambda^2 B_1\right)}{\left(1 - \sigma_{e_{SD}}^2\right)\left(1 - \sigma_{e_{SR}}^2\right)\left(1 - \sigma_{e_{RD}}^2\right)} \log\left(\frac{SNR_{eff}\,\lambda B_1\left(1 - \sigma_{e_{SR}}^2\right)\left(1 - \sigma_{e_{RD}}^2\right)}{2}\right). \tag{3.34}$$

Similar to TD protocol, a second order diversity is obtained. Performance degradation due to channel estimation is given by

$$\Xi = \frac{\left(\Lambda/N_0\right)^2}{\left(1 - \sigma_{e_{SD}}^2\right)\left(1 - \sigma_{e_{SR}}^2\right)\left(1 - \sigma_{e_{RD}}^2\right)} \log\left(\frac{E\lambda B_1}{4}\left[\frac{\left(1 - \sigma_{e_{SR}}^2\right)\left(1 - \sigma_{e_{RD}}^2\right)}{\Lambda} - \frac{1}{N_0}\right]\right) \tag{3.35}$$

As $\sigma_{e_{SRD}}^2$, $\sigma_{e_{SD}}^2 \to 0$, (3.34) reduces to perfect CSI case reported in [51] and $\log \Xi$ becomes 0.

Finally, for RD protocol with partially-coherent receiver, we obtain the PEP as (see Appendix C for details of the derivation)

$$P\left(\mathbf{X}_D \to \hat{\mathbf{X}}_D\right) < \frac{SNR_{eff}^{-2}}{B_1^2 \Delta_1^2} \sqrt{\frac{\left(\Delta_2 B_L T / (2\Delta_1 + 2B_1\Delta_1)\right)^2}{\Delta_1^2 + SNR_{eff}^{-1}}} \exp\left(\frac{\Delta_2 B_L T}{2} + \frac{SNR_{eff}^{-1}}{B_1^2 \Delta_1^2}\right)$$
$$\times K_1\left(\sqrt{\left(\frac{\Delta_2 B_L T}{2 + 2B_1}\right)^2 + \frac{\left(\Delta_2 B_L T / (2 + 2B_1)\right)^2 SNR_{eff}^{-1}}{\Delta_1^2}}\right)\Gamma\left(0, \frac{SNR_{eff}^{-1}}{B_1^2 \Delta_1^2}\right) \tag{3.36}$$

where $\quad \Delta_1 = \left(1 - \cos(\theta_\Delta)\right), \quad \Delta_2 = \left(1 - \cos(\theta_\Delta) - 2\sin(\theta_\Delta)\right), \quad \cos(\theta_\Delta) = \text{Re}\left\{x\hat{x}^*\right\}\quad$ and $SNR_{eff} = E / \left[2N_0\left(1 - \cos(\theta_\Delta)\right)\left(1 + B_1\right)\right]$. For large values of SNR_{eff}, we use the approximation $K_v(z) \approx 0.5\Gamma(v)(2/z)^v$, $z < 1$ [50] reducing (3.36) to

$$P\left(\mathbf{X}_D \to \hat{\mathbf{X}}_D\right) < \frac{SNR_{eff}^{-2}}{B_1^2 \Delta_1^4} \exp\left(\frac{\Delta_2 B_L T}{2}\right) \log\left(SNR_{eff} B_1^2 \Delta_1^2\right) \tag{3.37}$$

indicating a second order diversity. Performance degradation with respect to partially-coherent receiver having perfect channel phase knowledge is given by

$$\Xi = \exp\left(\frac{\Delta_2 B_L T}{2}\right). \tag{3.38}$$

As the loop signal-to-noise ratio (which is inversely proportional to $B_L T$) increases, estimation of the channel phases in PLL become error free and therefore the term in (3.38) approaches one.

3.4 Effect of Relay Location on the Quality of Channel Estimates

In this section, we investigate the effect of relay location on the quality of channel estimates. Let Σ_{SRD} denote mean square error (MSE) of S→R→D channel estimate. To minimize MSE with respect to relay location, we need to solve the following constrained optimization problem

$$\underset{G_{SR}, G_{RD}}{\text{minimize}} \ \Sigma_{SRD} \qquad (3.39)$$

$$\text{s.t.} \ \ G_{SR}^{-2/a} + G_{RD}^{-2/a} - 2G_{SR}^{-1/a}G_{RD}^{-1/a}\cos\theta = 1$$

where the constraint equation is obtained through law of cosines between relative gains G_{SR} and G_{RD} considering the relay geometry (c.f., Fig.2.1). For Golden-coded pilot symbols p_{1i} and p_{3i}, replacing $\mathbf{X}_T = [\sqrt{A_1 E}\, p_{11} \ \ \sqrt{A_1 E}\, p_{31}..... \ \ \sqrt{A_1 E}\, p_{1N} \ \ \sqrt{A_1 E}\, p_{3N}]^{\text{T}}$ in $N_0\left(\mathbf{X}_T^{\text{H}}\mathbf{X}_T + N_0\mathbf{I}_2\right)^{-1}$ we obtain Σ_{SRD} of TD protocol as

$$\Sigma_{SRD} = \frac{\left(E/N_0\right)^{-1}}{2A_1 N\kappa + \left(E/N_0\right)^{-1}} \qquad (3.40)$$

where A_1 is a function of G_{SR} and G_{RD}, c.f., (3.7). Assume a scenario with path loss coefficient $a = 2$ and $\theta = \pi$. Then we have $G_{RD} = \left(1 + 1/\sqrt{\eta}\right)^2$, $G_{SR} = \eta\left(1 + 1/\sqrt{\eta}\right)^2$ with $\eta = G_{SR}/G_{RD}$. Taking the derivative of (3.40) with respect to η and equating to zero, we have

$$\frac{\partial \Sigma_{SRD}}{\partial \eta} = \frac{\left(EN\kappa/N_0\right)^2 v^3 \left(-2/\sqrt{\eta} + v/\eta + v\chi_1/m_1\right)/(\eta m_1)}{\left(\left(EN\kappa v^2/N_0\right)^2/(\eta m_1) + 1\right)^2} = 0 \qquad (3.41)$$

where $v = \sqrt{\eta} + 1$, $m_1 = 1 + v^2 E(\kappa + 1/2\eta)/N_0$ and $\chi_1 = vE\left(\kappa + 1/2\eta - v/2\eta\sqrt{\eta}\right)/\left(\sqrt{\eta}N_0\right)$. The numeric solutions of (3.41) for various E/N_0 are provided in Table 3-1.

Table 3-1: Relay locations that minimize MSE (TD protocol with $N = 1$)

$\eta = G_{SR}/G_{RD}$ [dB]	$\theta = \pi$	$\theta = 3\pi/4$	$\theta = 2\pi/3$
$E/N_0 = 10$dB	-6.21dB	-7.55dB	-9.32dB
$E/N_0 = 25$dB	-6.021dB	-7.2dB	-8.74dB
$E/N_0 = 35$dB	-6.021dB	-7.19dB	-8.73dB
$E/N_0 = 100$dB	-6.020dB	-7.19dB	-8.73dB

It can be readily checked that Σ_{SRD} is a convex function of relay location η. Therefore, the results in Table 3-1 are in fact, global minimums of the optimization problem. We observe from the table that relay location which minimizes MSE remains nearly constant for a wide range of SNR values. Negative values of $\eta = G_{SR}/G_{RD}$ (in dB) indicate that quality of channel estimation improves when relay is closer to the destination. Similar observations can be made for other values of θ presented in Table 3-1. As θ decreases, optimum relay location comes even closer to destination.

For RD protocol, replacing $\mathbf{X}_T = [\sqrt{B_1 E}\, p_1 \quad ...... \quad \sqrt{B_1 E}\, p_N]^{\mathrm{T}}$ in $N_0 \left(\mathbf{X}_T^{\mathrm{H}} \mathbf{X}_T + N_0 \mathbf{I_2} \right)^{-1}$, Σ_{SRD} can be obtained as

$$\Sigma_{SRD} = \frac{(E/N_0)^{-1}}{B_1 N + (E/N_0)^{-1}} \tag{3.42}$$

where B_1 is a function of G_{SR} and G_{RD}, c.f. (3.14). Taking the derivative of (3.42) with respect to $\eta = G_{SR}/G_{RD}$, we have

$$\frac{\partial \Sigma_{SRD}}{\partial \eta} = \frac{(EN/N_0)^2 v^3 \left(-2/\sqrt{\eta} + v/\eta + v\chi_2/m_2 \right)/(\eta m_2)}{\left(\left(ENv^2/N_0 \right)^2 /(\eta m_2) + 1 \right)^2} = 0 \tag{3.43}$$

where $m_2 = 1 + v^2 E(1 + 1/\eta)/N_0$ and $\chi_2 = vE\left(1 + 1/\eta - v/\eta\sqrt{\eta}\right)/\left(\sqrt{\eta}N_0\right)$. Solving (3.43), we obtain $\eta = 0$ dB independent of E/N_0 and θ. Hence, minimum MSE is obtained when relay is at the mid-point between source and destination.

3.5 Simulation Results

In this section, we present the results of a Monte-Carlo simulation study to demonstrate the performance of RD and TD protocols with mismatched-coherent and partially-coherent receivers. In our simulations, we consider TD protocol with Golden code assuming 4-PSK modulation. This achieves a throughput of 2 bits/sec/Hz. To make a fair comparison, we consider RD relaying and non-cooperative direct transmission with 16-PSK and 4-PSK, respectively.

In Fig. 3.1, we present the BER performance of Golden-coded TD protocol with a mismatched-coherent receiver for a scenario in which the relay is located in the midway of source-to-destination link, i.e., $G_{SR}/G_{RD} = 0\text{dB}$. We assume $a = 2$ and $\theta = \pi$. The performance of genie-aided receiver (i.e., perfect CSI), non-cooperative direct transmission (i.e., no relaying), and TD protocol with Alamouti code are further included as benchmarks. It is observed that Golden-coded TD protocol with both perfect and imperfect channel estimation yields a diversity order of two confirming our PEP analysis. The mismatched-coherent detection results in a performance loss of approximately 1.5 dB at BER$=5.10^{-3}$ with respect to the genie bound. Although Alamouti-coded TD protocol extracts a diversity order of two as well, it is significantly outperformed by its Golden-coded counterpart. Specifically, at BER$= 5.10^{-3}$, we observe a performance difference of 4 dB between two codes. Our results further demonstrate that Golden code has a slightly better robustness than Alamouti in the presence of imperfect channel estimation.
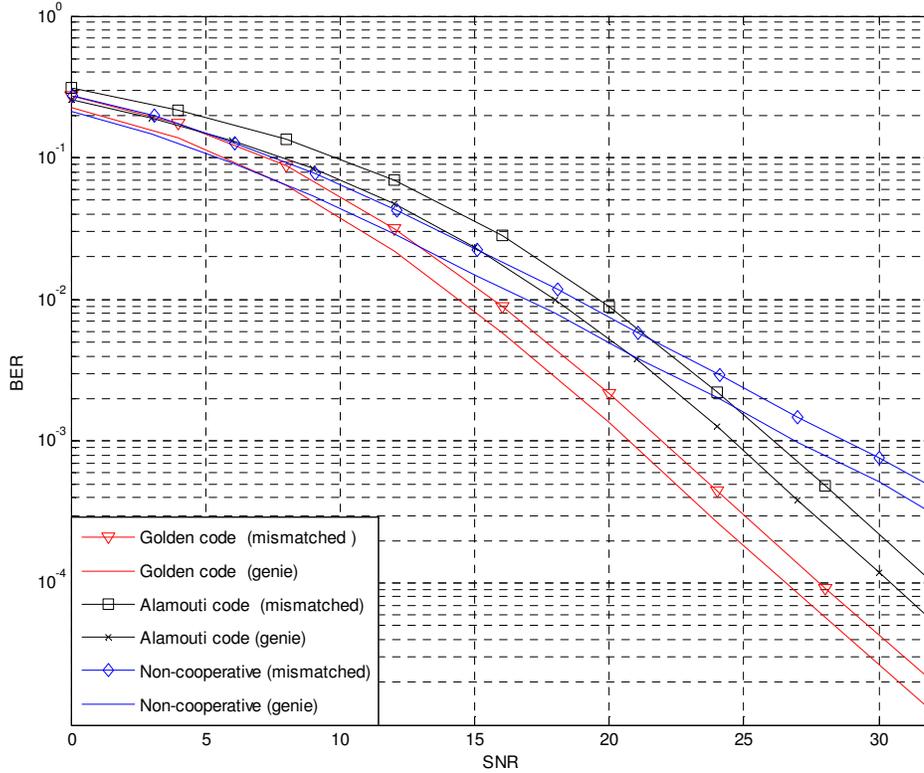
Fig. 3.1: BER of Golden-coded and Alamouti-coded TD protocol with mismatched-coherent and ge-
nie receivers ($G_{SR}/G_{RD} = 0$dB).

In Fig. 3.2, we demonstrate MSE and BER performance of TD protocol as a function of
relay location at SNR values of 10dB and 28dB. From Fig.3.2.a and 3.2.b, we observe that
the relay location which minimizes the estimation error takes place approximately at
$G_{SR}/G_{RD} = -6$dB. This confirms our earlier analytical derivations in Section 3.4. For low
SNR values, we observe from Fig.3.2.c that error rate performance of Golden code improves
slightly as relay continues to move away from the destination. However for high SNRs, a
better error rate performance is obtained when the relay is close to the destination (i.e., farth-
er than -6dB location). For such large negative values, relay is close to the destination and
in such a scenario, cooperative scheme mimics the behavior of a receive diversity scheme
with two co-located antennas. This demonstrates that nature of the cooperation protocol do-
minates the BER performance rather than the channel estimation quality.

36
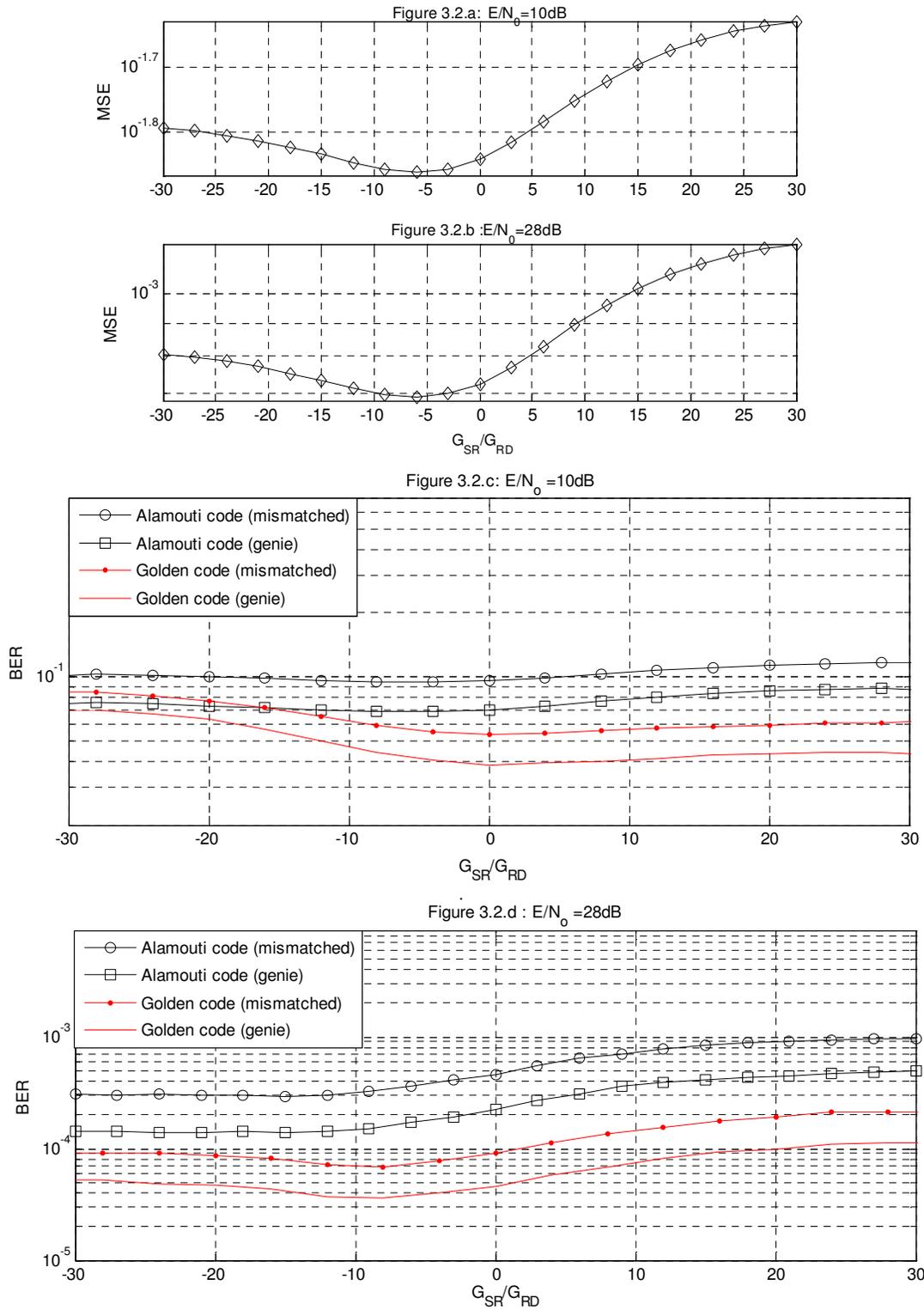
Fig. 3.2: MSE and BER of TD protocol with respect to relay location (E/N_0 = 10dB and 28 dB).
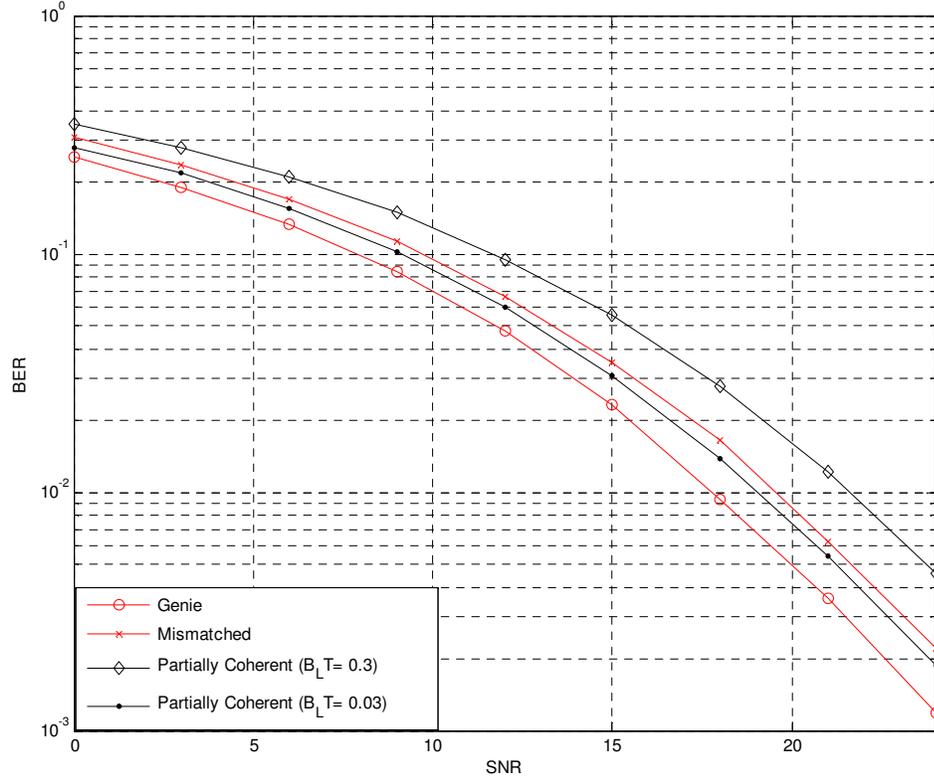
Fig. 3.3: BER of RD protocol with mismatched-coherent and partially-coherent receivers with relay location of $G_{SR}/G_{RD} = 0\text{dB}$.

In Fig. 3.3, we present BER performance of RD relaying with mismatched-coherent and partially-coherent receivers for $G_{SR}/G_{RD} = 0\text{dB}$ along with the genie bound. For partially-coherent detection with PLL-aided phase estimation, we consider two different $B_L T$ values. For $B_L T = 0.03$ and 0.3, performance degradations with respect to genie bound are, respectively, 1.1 dB and 4dB at BER=5.10^{-3}. This is an expected result as large $B_L T$ values result in inefficient phase estimation; whereas for small $B_L T$ values the estimation error variance tends to zero. It is interesting to note that partially-coherent detector with $B_L T = 0.03$ is able to slightly outperform the mismatched-coherent receiver although no effort is made for channel amplitude estimation. This points out that a reliable channel phase information is more essential in the detection process than the channel amplitude. Further comparison of Figs. 3.1

and 3.3 reveal that TD protocol with Alamouti code provides an identical performance to that of RD protocol. This observation has been earlier reported in [51] for perfect CSI case. Since Golden-coded TD protocol has a much superior performance over both RD and Alamouti-coded TD, it becomes the obvious choice for distributed implementation.

In Fig. 3.4, we provide MSE and BER performance of RD relaying as a function of relay location. Both mismatched-coherent and partially-coherent receivers are considered. We observe from Fig.3.4.a and Fig.3.4.b that RD protocol experiences the minimum estimation error when the relay is in the mid-point confirming our derivations in Section 3.4. Our results in Fig.3.4.c and Fig.3.4.d demonstrate that error rate performance improves as relay moves closer to the destination. This is similar to our earlier observations for Fig. 3.2. Specifically, for $G_{SR}/G_{RD} = -30$ dB and SNR=28dB, the performance degradations with respect to genie bound are, respectively, $1.5.10^{-4}$, $1.6.10^{-4}$, $5.6.10^{-4}$ dB for mismatched-coherent receiver, partially-coherent receivers with $B_L T = 0.03$ and $B_L T = 0.3$. For $G_{SR}/G_{RD} = 30$ dB, the performance degradations are $4.2.10^{-4}$, $1.2.10^{-4}$ and $5.4.10^{-4}$ respectively.
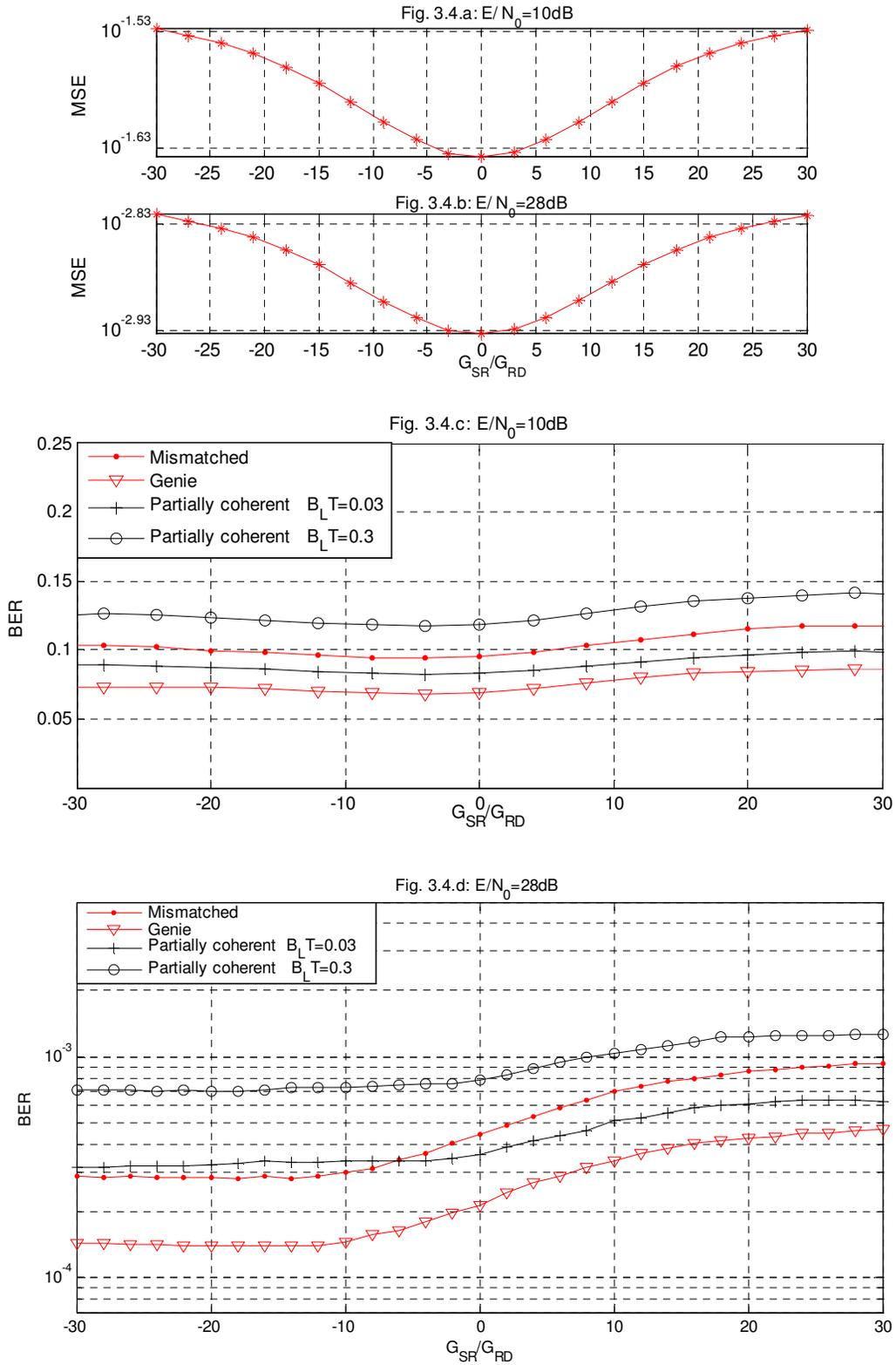
Fig. 3.4: MSE and BER of RD protocol with respect to relay location ($E/N_0 = 10$dB and 28 dB).

Chapter 4

Training Power Optimization for AaF Relaying

Introduction

In the previous chapter, we have shown that the diversity order is preserved even in the presence of channel estimation errors. However, further performance gains are possible through proper resource optimization. In this chapter, we aim to optimize the performance of AaF relaying through power allocation. Considering average received SNR at the destination node as the objective function, we formulate an optimization problem for a single-relay scenario with RD cooperation protocol to answer the following fundamental questions:

1) How should the overall transmit power be shared between training and data transmission periods?;

2) How should training power be allocated to broadcasting and relaying phases?;

3) How should data power be allocated to broadcasting and relaying phases?

4.1 Transmission Model

We consider a single-relay AaF cooperation system illustrated in Fig.2.1, where source, relay and destination terminals are equipped with single transmit and receive antennas. We assume mismatched-coherent detection with C-CE. Since we consider a quasi-static fading channel, the placement of pilot symbols is irrelevant in our optimization. In our work, we assume that training symbols precede the data symbols and consider a frame length of $T = T_t + T_d$ symbols where T_t and T_d denote the length of training and data symbols, re-

spectively. Let $2ET$ be the total available energy consumed during the transmission of T symbols, yielding an average power of E (per time slot). Over the two time slots (required to transmit a symbol for the cooperation protocol under consideration), $2E_t$ and $2E_d$ are, respectively, assigned for the transmission of a training and data symbol. Therefore, we have $2E_tT_t + 2E_dT_d = 2ET$. We further introduce a parameter $0 < \rho < 1$ to relate E_d and E_t as $E_d = \rho ET/T_d$ and $E_t = (1-\rho)ET/T_t$. This parameter will be later used in the optimization procedure to determine how much power should be allocated to data or training symbol transmission.

The transmission model under consideration builds upon RD cooperation protocol which conveys the same information to both relay and destination nodes over the two transmission phases. Let x_j denote M-PSK modulated symbol with normalized power $E[|x_j|^2]=1$ for $j = 1, 2, ... T_d$. The received signals at the relay and destination taking into account of the associated path losses are given by

$$r_{1,j} = \sqrt{2K_dG_{SR}E_d}\, h_{SR}x_j + n_{1,j}, \tag{4.1}$$

$$r_{2,j} = \sqrt{2K_dE_d}\, h_{SD}x_j + n_{2,j}. \tag{4.2}$$

where K_d is defined as another optimization parameter and controls the fraction of power reserved for the source terminal's use in the broadcasting phase. The relay node normalizes the received signal by a factor of $\sqrt{E[|r_{1,j}|^2]}$ to have average unit power and then forwards the scaled signal within the relaying phase with power $2(1-K_d)E_d$. The received signal at the destination is, therefore, given by

$$r_{3,j} = \sqrt{2(1-K_d)G_{RD}E_d}\, h_{RD}\, \frac{r_{1,j}}{\sqrt{E[|r_{1,j}|^2]}} + n_{3,j}. \tag{4.3}$$

In (4.1)-(4.3), $n_{1,j}$, $n_{2,j}$, and $n_{3,j}$ model the additive noise terms and are the independent samples of zero-mean complex Gaussian random variables with a variance of $N_0/2$ per dimension. h_{SR}, h_{SD}, and h_{RD} denote the Rayleigh fading coefficients over S→R, S→D, and R→D links and are modeled as zero-mean complex Gaussian fading coefficients with va-

riance of 0.5 per dimension. Replacing the normalization factor $\sqrt{E[|r_{1,j}|^2]}$ $=$ $\sqrt{2K_d G_{SR} E_d + N_0}$ in (4.3), we have

$$r_{3,j} = \sqrt{\frac{4K_d(1-K_d)G_{SR}G_{RD}E_d^2}{2K_d G_{SR} E_d + N_0}} h_{SR} h_{RD} x_j + \tilde{n} \tag{4.4}$$

where $\tilde{n} = \sqrt{[2(1-K_d)G_{RD}E_d]/(2K_d G_{SR} E_d + N_0)} h_{RD} n_{1,j} + n_{3,j}$. Due to the term involving $h_{RD} n_{1,j}$, $\tilde{n}$ is of non-Gaussian nature which makes the analysis intractable for most cases. As in Chapter 2 and 3, we treat this noise as Gaussian noise as the worst case scenario. After replacing $\tilde{n}$ with zero mean Gaussian noise with the same average power, the destination terminal normalizes the received signal by a factor of $\sqrt{1+2(1-K_d)G_{RD}E_d/(2K_d G_{SR} E_d + N_0)}$ which yields

$$r_{4,j} = \sqrt{B_d E_d} h_{SR} h_{RD} x_j + n_{4,j} \tag{4.5}$$

where $n_{4,j}$ is zero-mean complex Gaussian random variable with a variance of N_0 and B_d is given by

$$B_d = \frac{4K_d(1-K_d)G_{SR}G_{RD}E_d/N_0}{1+2K_d G_{SR} E_d/N_0 + 2(1-K_d)G_{RD}E_d/N_0} \quad . \tag{4.6}$$

Introducing $\mathbf{r}_{d,j} = [r_{2,j} \quad r_{4,j}]^T$ and $\mathbf{n}_{d,j} = [n_{2,j} \quad n_{4,j}]^T$, the received signals for the jth data block can be given as $\mathbf{r}_{d,j} = \mathbf{X}_{d,j}\mathbf{h} + \mathbf{n}_{d,j}$ where $\mathbf{h} = [h_{SR}h_{RD} \quad h_{SD}]^T$ and $\mathbf{X}_{d,j}$ is

$$\mathbf{X}_{d,j} = \begin{bmatrix} 0 & \sqrt{2K_d E_d} \\ \sqrt{B_d E_d} & 0 \end{bmatrix} x_j . \tag{4.7}$$

The received signals $\mathbf{r}_{d,j}$ during the data transmission period can be then stacked and written in a compact matrix form as $\mathbf{r} = \begin{bmatrix} \mathbf{r}_{d,1}, \mathbf{r}_{d,2}, ... \mathbf{r}_{d,T_d} \end{bmatrix}^T = \mathbf{X}_D \mathbf{h} + \mathbf{n}_D$ where $\mathbf{X}_D = [\mathbf{X}_{d,1} \ \mathbf{X}_{d,2} \ \cdots \ \mathbf{X}_{d,T_d}]^T$. $\mathbf{r}$ is fed to mismatched coherent decoder given by (3.20). Before performing the decoding, the receiver needs to know CSI. For this purpose, prior to data transmission, the source node transmits pilot symbols for channel estimation. The pilot-symbol matrix $\mathbf{X}_{t,i}$ for $i = 1, 2, ... T_t$ has a similar form of (4.7) and is given by

$$\mathbf{X}_{t,i} = \begin{bmatrix} 0 & \sqrt{2K_t E_t} \\ \sqrt{B_t E_t} & 0 \end{bmatrix} p \tag{4.8}$$

where p denotes the pilot symbol. Here, B_t is defined as

$$B_t = \frac{4K_t(1-K_t)G_{SR}G_{RD}\,E_t/N_0}{1+2K_tG_{SR}\,E_t/N_0 + 2(1-K_t)G_{RD}\,E_t/N_0} \;. \tag{4.9}$$

where we introduce another optimization parameter $0 < K_t < 1$ which controls the allocation of training power to broadcasting and relaying phases.

The received signals during the training period are given by $\mathbf{r}_T = \mathbf{X}_T \mathbf{h} + \mathbf{n}_T$ where $\mathbf{X}_T = [\mathbf{X}_{t,1}\ \mathbf{X}_{t,2}\ \cdots\ \mathbf{X}_{t,T_t}]^{\mathrm{T}}$. Based on an LMMSE estimator, channel estimate $\hat{\mathbf{h}} = \begin{bmatrix} \hat{h}_{SRD} & \hat{h}_{SD} \end{bmatrix}^{\mathrm{T}}$ is obtained from (3.18). The covariance matrix of estimation error $\mathbf{e} = \mathbf{h} - \hat{\mathbf{h}}$ can be found by using (3.19). The covariance matrix $\mathbf{C_e}$ has a diagonal structure and its diagonal elements $\sigma^2_{e_{SRD}}$ and $\sigma^2_{e_{SD}}$ are given by

$$\mathrm{var}(e_{SRD}) = \sigma^2_{e_{SRD}} = N_0/(B_t\,T_t E_t + N_0), \tag{4.10}$$

$$\mathrm{var}(e_{SD}) = \sigma^2_{e_{SD}} = N_0/(2K_t\,T_t E_t + N_0). \tag{4.11}$$

Then, MSE can be then obtained in terms of channel estimation error variances as

$$\Sigma_{MSE} = \mathrm{var}(e_{SRD}) + \mathrm{var}(e_{SD}) \;. \tag{4.12}$$

4.2 Optimization of K_d, K_t and ρ

In our model, we have introduced three optimization parameters K_d, K_t and ρ. ρ controls the power allocation to training and data periods. On the other hand, K_d and K_t respectively control the allocation of data and training powers between broadcasting and relaying phases. Here, we aim to optimize these parameters in order to maximize the average received SNR at the destination terminal. The average SNR at the destination is

$$SNR_{avg} = SNR_{S \to D} + SNR_{S \to R \to D} \tag{4.13}$$

$$= \frac{2K_d E_d \mathrm{var}(\hat{h}_{SD})}{2K_d E_d \mathrm{var}(e_{SD}) + N_0} + \frac{B_d E_d \mathrm{var}(\hat{h}_{SRD})}{B_d E_d \mathrm{var}(e_{SRD}) + N_0} \tag{4.14}$$

where $\text{var}(\hat{h}_{SD})$ and $\text{var}(\hat{h}_{SRD})$ denote the variances of $S{\rightarrow}D$ and $S{\rightarrow}R{\rightarrow}D$ channel estimates and are given by

$$\text{var}(\hat{h}_{SD}) = 2K_t T_t E_t \big/ (2K_t T_t E_t + N_0), \tag{4.15}$$

$$\text{var}(\hat{h}_{SRD}) = B_t T_t E_t \big/ (B_t T_t E_t + N_0). \tag{4.16}$$

After rewriting the $SNR_{S{\rightarrow}D}$ and $SNR_{S{\rightarrow}R{\rightarrow}D}$ expressions of (4.14) in common denominator, we can formulate our optimization problem as

$$\max_{\rho,K_t,K_d} \Phi(\rho,K_t,K_d) \quad \text{s.t.} \quad 0< \rho <1, \quad 0< K_d <1, \text{ and } \quad 0< K_t <1$$

where the objective function $\Phi(\rho,K_t,K_d)$ is

$$\Phi = \frac{2K_d \dfrac{E_d}{N_0} \text{var}(\hat{h}_{SD}) \left(B_d \dfrac{E_d}{N_0} \text{var}(e_{SRD})+1 \right) + B_d \dfrac{E_d}{N_0} \text{var}(\hat{h}_{SRD}) \left(2K_d \dfrac{E_d}{N_0} \text{var}(e_{SD})+1 \right)}{\left(2K_d \dfrac{E_d}{N_0} \text{var}(e_{SD})+1 \right) \left(B_d \dfrac{E_d}{N_0} \text{var}(e_{SRD})+1 \right)} \tag{4.17}$$

Inserting (4.10), (4.11), (4.15) and (4.16) in (4.17) and further replacing $E_d = \rho ET/T_d$ and $E_t = (1-\rho)ET/T_t$ in the resulting expression, we obtain

$$\Phi = \frac{\dfrac{4K_d K_t \rho(1-\rho)(ET/N_0)^2}{T_d} \left(B_d \dfrac{\rho ET}{T_d N_0} + B_t \dfrac{(1-\rho)ET}{N_0} +1 \right)}{\left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} +1 \right) \left(B_d \dfrac{\rho ET}{T_d N_0} + B_t \dfrac{(1-\rho)ET}{N_0} +1 \right)}$$
$$+ \frac{\dfrac{B_d B_t \rho(1-p)(ET/N_0)^2}{T_d} \left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} +1 \right)}{\left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} +1 \right) \left(B_d \dfrac{\rho ET}{T_d N_0} + B_t \dfrac{(1-\rho)ET}{N_0} +1 \right)} \tag{4.18}$$

where B_d and B_t, earlier defined by (4.6) and (4.9), are functions of optimization parameters and can be expanded as

$$B_d = \frac{4K_d(1-K_d)G_{SR}G_{RD}\dfrac{\rho ET}{T_d N_0}}{1+2K_d G_{SR}\dfrac{\rho ET}{T_d N_0}+2(1-K_d)G_{RD}\dfrac{\rho ET}{T_d N_0}}\ , \tag{4.19}$$

$$B_t = \frac{4K_t(1-K_t)G_{SR}G_{RD}\dfrac{(1-\rho)ET}{T_t N_0}}{1+2K_t G_{SR}\dfrac{(1-\rho)ET}{T_t N_0}+2(1-K_t)G_{RD}\dfrac{(1-\rho)ET}{T_t N_0}}\ . \tag{4.20}$$

To observe the dependence of the average SNR expression on optimization parameters more closely, we provide the three dimensional plot of (4.18) as a function of ρ and K_d for various T and T_t values in Fig.4.1. We assume $G_{SR}/G_{RD}=0\text{dB}$, $\theta=\pi$, $a=2$ and $K_d=K_t$ here. Fig. 4.1 demonstrates that (4.18) is a concave function of the optimization parameters, therefore, there is only one global maximum within the considered interval.
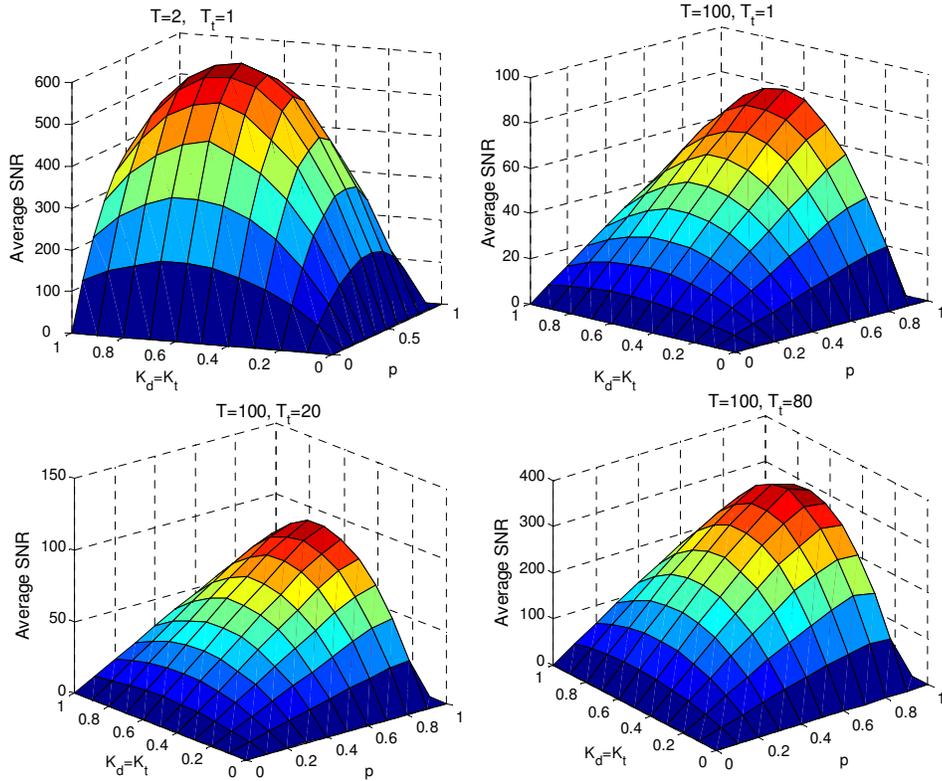
Fig.4.1: Average SNR with respect to ρ, K_d and K_t for $G_{SR}/G_{RD}=0\text{dB}$.

A general analytic solution for optimization parameters is very difficult, if not impossible. Therefore, in the following, we investigate certain scenarios where (4.18) simplifies and lets analytical derivation. As clearly seen from the dependence of (4.19) and (4.20) on relative geometrical gains, the optimization function is a function of the relay location. Next, we consider three representative scenarios based on the relay location.

4.2.1 Scenario I: Relay is Close to Destination

When the relay is located close to the destination node, we have $G_{SR}/G_{RD} \ll 1$ which lets us to write $G_{SR} \approx 1$ and $G_{RD} \gg 1$. Therefore, (4.19) and (4.20) simplify as $B_d \approx 2K_d G_{SR} \approx 2K_d$ and $B_t \approx 2K_t G_{SR} \approx 2K_t$. Inserting these in (4.18), the objective function is formulated as

$$
\Phi = \frac{4K_d K_t \dfrac{\rho(1-\rho)(ET/N_0)^2}{T_d}\left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} + 1\right)}{\left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} + 1\right)\left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} + 1\right)}
$$
$$
+ \frac{4K_d K_t \dfrac{\rho(1-\rho)(ET/N_0)^2}{T_d}\left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} + 1\right)}{\left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} + 1\right)\left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} + 1\right)}
$$

(4.21)

After some mathematical manipulations and simplifications, (4.21) becomes

$$
\Phi = \frac{8K_d K_t \dfrac{\rho(1-\rho)(ET/N_0)^2}{T_d}}{2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} + 1} \ .
$$

(4.22)

For a fixed value of ρ, it can be easily checked that (4.22) is an increasing function of K_d and K_t (i.e., both $d\Phi/dK_d$ and $d\Phi/dK_t$ are greater than zero). Therefore, it can be maximized by setting K_d and K_t as large as possible. Considering the range of K_d and K_t, we conclude that the objective function is maximized for $K_{d\_opt} \to 1$ and $K_{t\_opt} \to 1$. This

points out that majority of the training/data power should be allocated for the use of source node in the broadcasting phase. It would be interesting to note that for the relay location under consideration (i.e., $B_t \approx 2K_t G_{SR} \approx 2K_t$), MSE in (4.12) simplifies to $\Sigma_{MSE} = 2N_0/(2K_t T_t E_t + N_0)$. For a given training power E_t and training length T_t, the optimum K_t value to minimize the MSE expression is found as $K_{t\_opt} \to 1$ which coincides our result obtained through the maximization of average SNR.

After inserting K_{d_opt} and K_{t_opt} in (4.22), differentiating the resulting expression with respect to ρ and finally equating it to zero, optimum value of ρ is found as

$$\rho_{opt} = \begin{cases} 1/2, & \text{for } T_d = 1 \\ \dfrac{1 + 2K_{t\_opt} ET/N_0}{2K_{t\_opt} ET/N_0 - 2K_{d\_opt}\dfrac{ET}{T_d N_0}} \left(1 - \sqrt{1 - \dfrac{2K_{t\_opt} ET/N_0 - 2K_{d\_opt}\dfrac{ET}{T_d N_0}}{1 + 2K_{t\_opt} ET/N_0}} \right), & \text{for } T_d > 1 \end{cases} \quad (4.23)$$

Table 4-1 tabulates optimum values of ρ for a frame length of $T = 100$ assuming various number of pilot symbols (i.e., T_t). For $T_t = 1$, we observe that the optimum ρ value is ~ 0.9. As T_t increases, optimum ρ value, along with number of data symbols (i.e., $T_d = T - T_t$) decreases. This leads to higher $E_t T_t = (1 - \rho_{opt})ET$ values. From (4.10) and (4.11), we can readily check that channel estimation errors are inversely proportional to $E_t T_t$. Therefore, the increase in training interval would result in a better channel estimation quality (i.e., smaller MSE) and less bit errors at the decoder. However, this comes with a sacrifice in data throughput and channel capacity.

Table 4-2 tabulates optimum values of ρ for a training length of $T_t = 1$ assuming various frame lengths (i.e., T). Specifically for $T = 2$, we observe that ρ_{opt} is found as 0.5. Increase in T leads to larger ρ_{opt} values. Although $(1 - \rho_{opt})$ decreases for higher T values, $E_t T_t = (1 - \rho_{opt})ET$ increases, therefore results in a better quality channel estimation.

Table 4-1: Optimum ρ values for various T_t values when the relay is close to the destination

E/N_0	$T_t = 1$	$T_t = 5$	$T_t = 20$	$T_t = 50$	$T_t = 80$
5dB	0.9025	0.9009	0.8939	0.8719	0.8150
10dB	0.9067	0.9051	0.8975	0.8748	0.8164
15dB	0.9080	0.9063	0.8988	0.8756	0.8170
20dB	0.9082	0.9067	0.8992	0.8758	0.8171
25dB	0.9086	0.9068	0.8993	0.8760	0.8172

Table 4-2: Optimum ρ values for various T values when the relay is close to the destination

E/N_0	$T = 2$	$T = 50$	$T = 80$
5dB	0.5	0.8670	0.8921
10dB	0.5	0.8724	0.8967
15dB	0.5	0.8741	0.8981
20dB	0.5	0.8747	0.8986
25dB	0.5	0.8749	0.8987

4.2.2 Scenario II: Relay is in the Midway between Source and Destination

When the relay is equidistant from the source and the destination nodes, channel gain ratio becomes $G_{SR}/G_{RD} = 1$. Under the assumption of $G_{SR} = G_{RD}$, (4.19) and (4.20) simplify to $B_d \approx 2K_d(1-K_d)G_{SR}$ and $B_t \approx 2K_t(1-K_t)G_{SR}$. Inserting these in (4.18), we have

$$
\Phi = \frac{\dfrac{4K_dK_t\rho(1-\rho)(ET/N_0)^2}{T_d}\left(2K_d(1-K_d)G_{SR}\dfrac{\rho ET}{T_dN_0}+2K_t(1-K_t)G_{SR}\dfrac{(1-\rho)ET}{N_0}+1\right)}{\left(2K_d\dfrac{\rho ET}{T_dN_0}+2K_t\dfrac{(1-\rho)ET}{N_0}+1\right)\left(2K_d(1-K_d)G_{SR}\dfrac{\rho ET}{T_dN_0}+2K_t(1-K_t)G_{SR}\dfrac{(1-\rho)ET}{N_0}+1\right)}
$$
$$
+\frac{\dfrac{4K_tK_d(1-K_t)(1-K_d)G_{SR}^2\rho(1-\rho)(ET/N_0)^2}{T_d}\left(2K_d\dfrac{\rho ET}{T_dN_0}+2K_t\dfrac{(1-\rho)ET}{N_0}+1\right)}{\left(2K_d\dfrac{\rho ET}{T_dN_0}+2K_t\dfrac{(1-\rho)ET}{N_0}+1\right)\left(2K_d(1-K_d)G_{SR}\dfrac{\rho ET}{T_dN_0}+2K_t(1-K_t)G_{SR}\dfrac{(1-\rho)ET}{N_0}+1\right)}
$$

(4.24)

Now recall that for the previous scenarios we have found out that optimum K_t and K_d values are equal. Under the assumption of $K_t = K_d$, (4.24) reduces to

$$\Phi = \frac{\dfrac{4K_d^2\rho(1-\rho)(ET/N_0)^2}{T_d}\left(2K_d(1-K_d)G_{SR}\dfrac{\rho ET}{T_dN_0}+2K_d(1-K_d)G_{SR}\dfrac{(1-\rho)ET}{N_0}+1\right)}{\left(2K_d\dfrac{\rho ET}{T_dN_0}+2K_d\dfrac{(1-\rho)ET}{N_0}+1\right)\left(2K_d(1-K_d)G_{SR}\dfrac{\rho ET}{T_dN_0}+2K_d(1-K_d)G_{SR}\dfrac{(1-\rho)ET}{N_0}+1\right)}$$

$$+\frac{\dfrac{4K_d^2(1-K_d)^2G_{SR}^2\rho(1-\rho)(ET/N_0)^2}{T_d}\left(2K_d\dfrac{\rho ET}{T_dN_0}+2K_d\dfrac{(1-\rho)ET}{N_0}+1\right)}{\left(2K_d\dfrac{\rho ET}{T_dN_0}+2K_d\dfrac{(1-\rho)ET}{N_0}+1\right)\left(2K_d(1-K_d)G_{SR}\dfrac{\rho ET}{T_dN_0}+2K_d(1-K_d)G_{SR}\dfrac{(1-\rho)ET}{N_0}+1\right)}$$

$$(4.25)$$

Further imposing high SNR assumption and performing some mathematical manipulations, (4.25) takes the form of

$$\Phi = \frac{\dfrac{2K_d\rho(1-\rho)(ET/N_0)}{T_d}(1+(1-K_d)G_{SR})}{\left(\dfrac{\rho}{T_d}+(1-\rho)\right)}$$

$$(4.26)$$

For a fixed ρ, the optimum value of K_d value which maximizes (4.26) can be found by taking the derivative of (4.26) and solving for K_d. This yields

$$K_{d\_opt} = K_{t\_opt} = \frac{1+G_{SR}}{2G_{SR}}.$$

$$(4.27)$$

which takes values within the range of 0.5 and 1. After inserting (4.27) into (4.26), we take the derivative with respect to ρ and solve for ρ_{opt}. We obtain ρ_{opt} as

$$\rho_{opt} = \begin{cases} 1/2 & \text{for } T_d = 1 \\ \dfrac{T_d}{T_d-1}\left(1-\sqrt{\dfrac{1}{T_d}}\right) & \text{for } T_d > 1 \end{cases}$$

$$(4.28)$$

In Table 4-3, we have presented optimum values of ρ assuming various values for T_t. We assume $\theta = \pi$ and $a = 2$. For these particular values, G_{SR} can be solved from cosine theorem and, using (4.27), we find K_{d_opt} and K_{t_opt} as 0.625. Similar to two previous scenarios, we have $\rho_{opt} \simeq 0.9$ for $T_t = 1$. It is also observed that ρ_{opt} slightly decreases with increasing values of T_t.

Table 4-3: Optimum ρ for various T_t values when the relay is at the midpoint between the source and the destination nodes

E/N_0	$T_t = 1$	$T_t = 5$	$T_t = 20$	$T_t = 50$	$T_t = 80$
5dB	0.9087	0.9069	0.8994	0.8761	0.8173
10dB	0.9087	0.9069	0.8994	0.8761	0.8173
15dB	0.9087	0.9069	0.8994	0.8761	0.8173
20dB	0.9087	0.9069	0.8994	0.8761	0.8173
25dB	0.9087	0.9069	0.8994	0.8761	0.8173

4.2.3 Scenario III: Relay is Close to Source

When the relay is located near the source node, the relative gain ratio becomes $G_{SR}/G_{RD} \gg 1$ which leads to $G_{RD} \approx 1$ and $G_{SR} \gg 1$. Consequently, we have $B_d \approx 2G_{RD}(1 - K_d) \approx 2 - 2K_d$ and $B_t \approx 2G_{RD}(1 - K_t) \approx 2 - 2K_t$. Inserting these into (4.18), we have

$$
\begin{aligned}
\Phi = & \frac{4K_d K_t \dfrac{\rho(1-\rho)(ET/N_0)^2}{T_d}\left((2-2K_d)\dfrac{\rho ET}{T_d N_0} + (2-2K_t)\dfrac{(1-\rho)ET}{N_0} + 1\right)}{\left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} + 1\right)\left((2-2K_d)\dfrac{\rho ET}{T_d N_0} + (2-2K_t)\dfrac{(1-\rho)ET}{N_0} + 1\right)} \\[2em]
& + \frac{(2-2K_d)(2-2K_t)\dfrac{\rho(1-\rho)(ET/N_0)^2}{T_d}\left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} + 1\right)}{\left(2K_d \dfrac{\rho ET}{T_d N_0} + 2K_t \dfrac{(1-\rho)ET}{N_0} + 1\right)\left((2-2K_d)\dfrac{\rho ET}{T_d N_0} + (2-2K_t)\dfrac{(1-\rho)ET}{N_0} + 1\right)}
\end{aligned}
\tag{4.29}
$$

There is no easy way of solving (4.29) analytically for optimum values. After some mathe-matical manipulations, we obtain

$$\Phi = \frac{\dfrac{8K_d\left(1-K_d\right)\rho^2\left(1-\rho\right)\left(\dfrac{ET}{N_0}\right)^3}{T_d^2} + \dfrac{8K_t\left(1-K_t\right)\rho\left(1-\rho\right)^2\left(\dfrac{ET}{N_0}\right)^3}{T_d}}{\dfrac{4K_d\left(1-K_d\right)\rho^2\left(\dfrac{ET}{N_0}\right)^2}{T_d^2} + 4K_t\left(1-K_t\right)\left(1-\rho\right)^2\left(\dfrac{ET}{N_0}\right)^2 + \dfrac{4\rho\left(1-\rho\right)\left(\dfrac{ET}{N_0}\right)^2\left(K_d+K_t-2K_d K_t\right)}{T_d} + \dfrac{2\rho\left(\dfrac{ET}{N_0}\right)}{T_d} + 2\left(1-\rho\right)\left(\dfrac{ET}{N_0}\right)+1}$$

$$-\frac{\dfrac{4\rho\left(1-\rho\right)\left(\dfrac{ET}{N_0}\right)^2\left(K_t+K_d-2K_t K_d\right)}{T_d} - \dfrac{4\rho\left(1-\rho\right)\left(\dfrac{ET}{N_0}\right)^2}{T_d}}{\dfrac{4K_d\left(1-K_d\right)\rho^2\left(\dfrac{ET}{N_0}\right)^2}{T_d^2} + 4K_t\left(1-K_t\right)\left(1-\rho\right)^2\left(\dfrac{ET}{N_0}\right)^2 + \dfrac{4\rho\left(1-\rho\right)\left(\dfrac{ET}{N_0}\right)^2\left(K_d+K_t-2K_d K_t\right)}{T_d} + \dfrac{2\rho\left(\dfrac{ET}{N_0}\right)}{T_d} + 2\left(1-\rho\right)\left(\dfrac{ET}{N_0}\right)+1} \quad (4.30)$$

Similar to the previous scenario, an analytical solution from (4.30) is not possible. However, under the assumption of $K_t \approx K_d$, (4.30) reduces to

$$\Phi = \frac{\dfrac{8K_d\left(1-K_d\right)\rho\left(1-\rho\right)\left(ET/N_0\right)^3}{T_d}\left[\dfrac{\rho}{T_d}+\left(1-\rho\right)\right] + \dfrac{4\rho\left(1-\rho\right)\left(ET/N_0\right)^2}{T_d} - \dfrac{8K_d\left(1-K_d\right)\rho\left(1-\rho\right)\left(ET/N_0\right)^2}{T_d}}{\left(\dfrac{\rho\left(ET/N_0\right)}{T_d}+\left(1-\rho\right)\left(ET/N_0\right)\right)\left[4K_d\left(1-K_d\right)\left(ET/N_0\right)\left(\dfrac{\rho}{T_d}+\left(1-\rho\right)\right)+2\right]+1} \quad (4.31)$$

Further imposing high SNR assumption, we ignore the terms with $\left(ET/N_0\right)^2$ in numerator and second term in denominator of (4.31). After some simplifications, (4.31) yields

$$\Phi = \frac{8K_d\left(1-K_d\right)\dfrac{\rho\left(1-\rho\right)\left(ET/N_0\right)^2}{T_d}}{4K_d\left(1-K_d\right)\left(\dfrac{\rho\left(ET/N_0\right)}{T_d}+\left(1-\rho\right)\left(ET/N_0\right)\right)+2} \quad (4.32)$$

Eq. (4.32) is an increasing function of $K_d\left(1-K_d\right)$ for a fixed ρ value. Maximum value of $K_d\left(1-K_d\right)$ takes place at $K_{d\_opt}=K_{t\_opt}=0.5$. We also note that the same result can be ob-

tained through the minimization of MSE. Specifically, for $B_t \approx 2 - 2K_t$, MSE takes the form

of

$$\Sigma_{MSE} = \frac{N_0}{(2 - 2K_t)T_t E_t + N_0} + \frac{N_0}{2K_t T_t E_t + N_0} \tag{4.33}$$

The optimum value of K_t to minimize MSE can be readily found as 0.5.

After differentiating (4.32) and equating to zero, we obtain optimum ρ in terms of K_{d_opt} as

$$\rho_{opt} = \begin{cases} 1/2 & \text{for } T_d = 1 \\ \dfrac{1 + 2K_{d\_opt}\left(1 - K_{d\_opt}\right)\left(ET/N_0\right)}{2K_{d\_opt}\left(1 - K_{d\_opt}\right)\left(ET/N_0\right)\left(1 - \dfrac{1}{T_d}\right)}\left(1 - \sqrt{1 - \dfrac{2K_{d\_opt}\left(1 - K_{d\_opt}\right)\left(ET/N_0\right)\left(1 - 1/T_d\right)}{1 + 2K_{d\_opt}\left(1 - K_{d\_opt}\right)\left(ET/N_0\right)}}\right) & \text{for } T_d > 1 \end{cases} \tag{4.34}$$

Table 4-4 tabulates optimum values of ρ for a frame length of $T = 100$ assuming various number of pilot symbols. Specifically for $T_t = 1$, ρ_{opt} is found ~ 0.9. Similar to the previous scenario, we observe that optimum ρ value decreases as the training length increases. In contrast, K_{d_opt} and K_{t_opt} drop from 1 to 0.5. This indicates that the training power allocated to the relay node should increase as the relay moves closer to source node and under such scenarios the relay and source nodes should share the training power equally.

Table 4-4: Optimum ρ values for various T_t values when the relay is close to the source

E/N_0	$T_t = 1$	$T_t = 5$	$T_t = 20$	$T_t = 50$	$T_t = 80$
5dB	0.8867	0.8854	0.8796	0.8607	0.8086
10dB	0.9009	0.8994	0.8926	0.8709	0.8144
15dB	0.9061	0.9044	0.8972	0.8744	0.8163
20dB	0.9078	0.9061	0.8987	0.8755	0.8169
25dB	0.9084	0.9066	0.8992	0.8759	0.8171

4.3 Simulation Results

In this section, we present Monte Carlo simulations to demonstrate the performance of optimized scheme for different relay locations. In our simulations, we assume 4-PSK modulation, $\theta = \pi$, and $a = 2$. We consider a quasi-static channel with coherence time of $T_{coh} = 2T$ symbols.
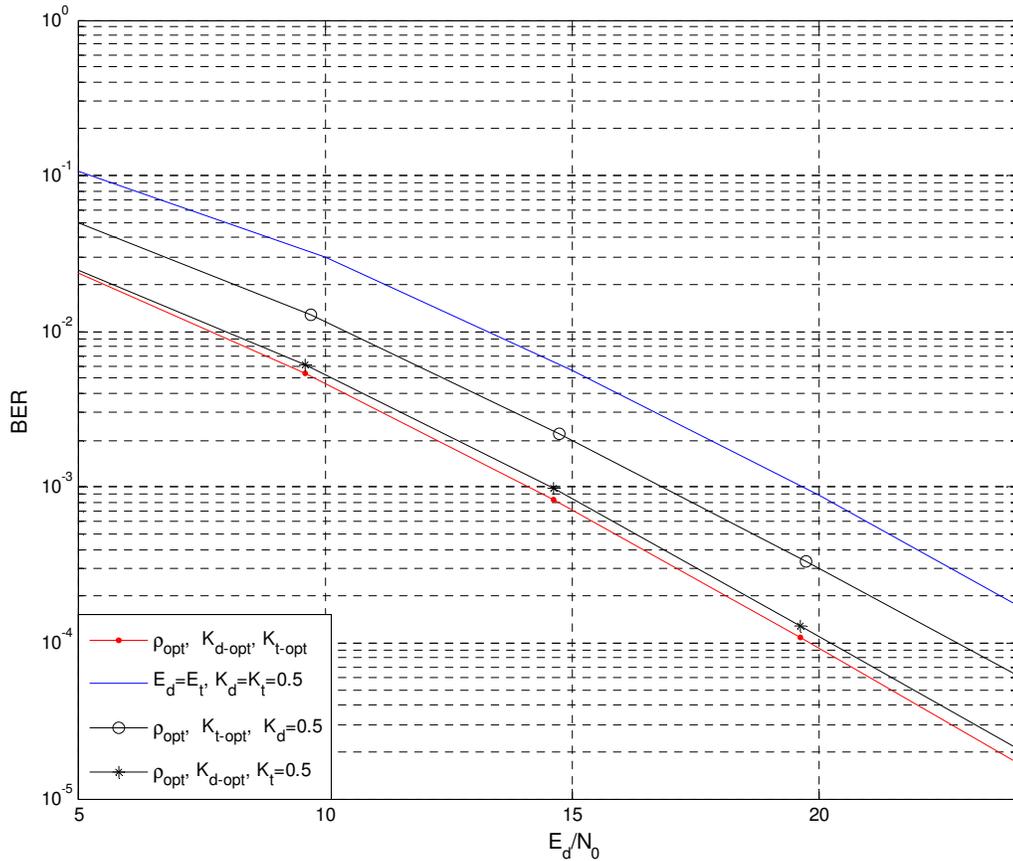
Fig.4.2: BER for optimum and equal power allocation ($G_{SR}/G_{RD} = -30\mathrm{dB}$)

In Fig. 4.2, we present the BER performance of optimum power allocation (i.e., K_{d_opt}, K_{t_opt} and ρ_{opt} in Table 4-1) along with equal power allocation (i.e., $E_d = E_t = E$ and $K_d = K_t = 0.5$) for $G_{SR}/G_{RD} = -30\mathrm{dB}$. We assume a transmission frame of $T = 100$ and deployment of a single pilot symbol, i.e., $T_t = 1$. Our results demonstrate that a performance

improvement of 5.5dB is obtained with respect to equal power allocation at a target BER of 10^{-3}. To test how K_t and K_d values individually affect the BER performance, we also consider the following two cases: In the first case, we set $K_t = 0.5$ and then optimize (4.21) with respect to ρ and K_d. In the second case, we set $K_d = 0.5$ and then optimize (4.21) for ρ and K_t. We include the corresponding performance of these cases in Fig. 4.2. The results demonstrate that the first case, i.e., $K_t = 0.5$, results in a minor degradation of 0.3dB with respect to optimum power allocation scheme while the second case, i.e., $K_d = 0.5$, results in a performance loss of 2.6dB. Therefore, we conclude that optimization of K_d plays a more significant role in performance improvement.
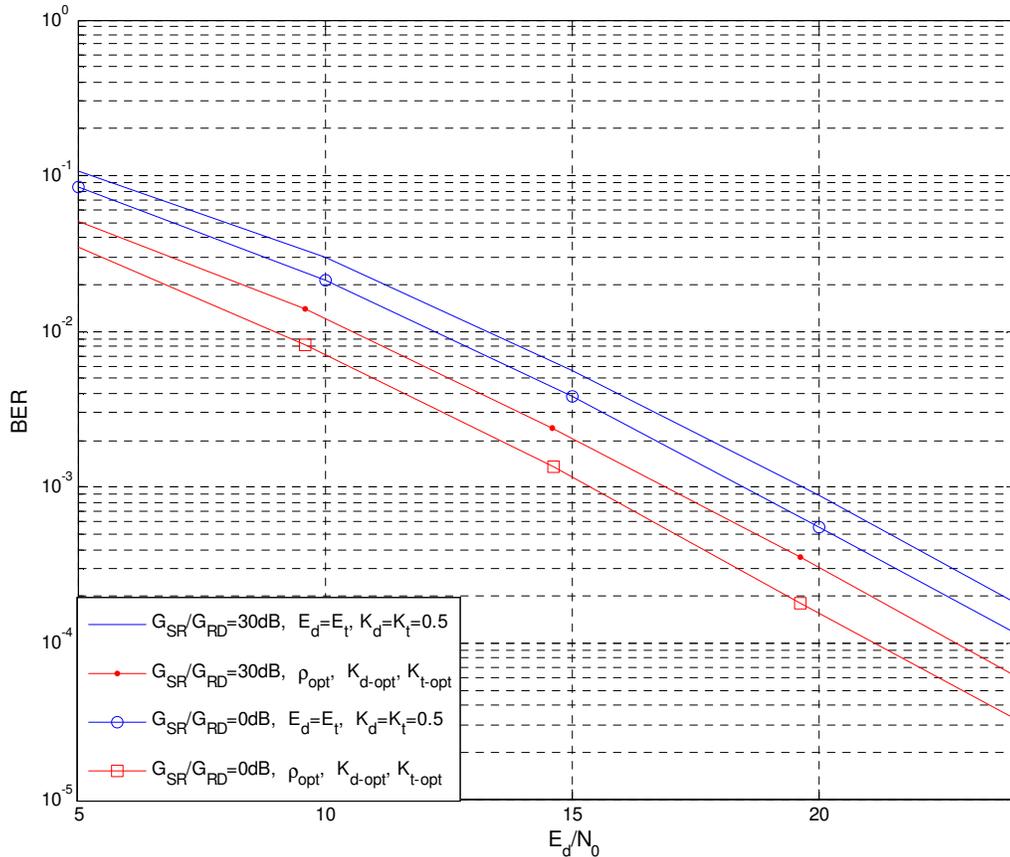
Fig. 4.3: BER for optimum and equal power allocation ($G_{SR}/G_{RD} = 30\text{dB}$ and $G_{SR}/G_{RD} = 0\text{dB}$)

In Fig. 4.3, we consider $G_{SR}/G_{RD} = 30\text{dB}$ (i.e., relay is located close to the source) and $G_{SR}/G_{RD} = 0\text{dB}$ (i.e., relay is at midway between the source and relay) and present BER for equal and optimized power allocation. At a target BER of 10^{-3}, we observe performance improvements of 2.7dB and 3.3dB for $G_{SR}/G_{RD} = 30\text{dB}$ and $G_{SR}/G_{RD} = 0\text{dB}$, respectively. We note that these gains are less than those observed for the case where the relay is located near the destination.
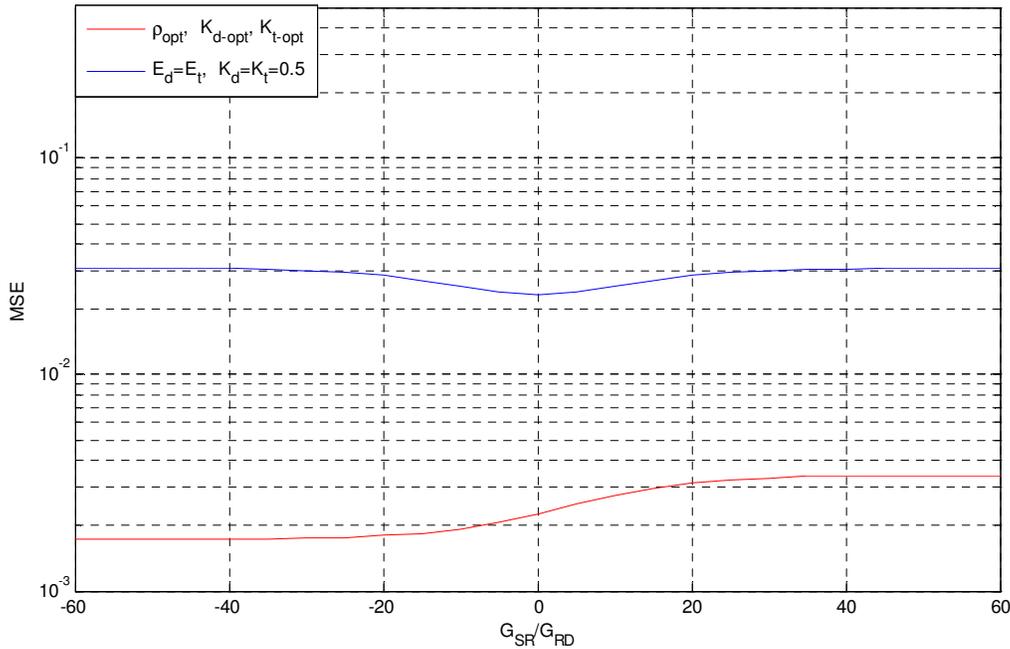
Fig.4.4: MSE with respect to relay location

In Fig. 4.4, we further investigate the impact of relay location on the MSE performance. We fix E/N_0 to 18dB and present MSE performance for equal and optimum power allocation as a function of relay location. As in Fig.3.4, it can be observed that MSE of equal power allocation becomes minimum for $G_{SR}/G_{RD} = 0\text{dB}$. On the other hand, MSE performance of optimum power allocation improves for negative G_{SR}/G_{RD} values. This indicates that larger optimization gains are available as the relay node moves close to the destination confirming our earlier observations.

In Fig. 4.5, we study the effect of training length T_t . We illustrate BER and MSE performances as well as data and training powers with respect to T_t. We assume $G_{SR}/G_{RD} = -30\text{dB}$, frame length of $T = 100$, and a E/N_0 value of 15dB. From Fig.4.5.a, we observe that BER performance of equal power allocation scheme first improves as T_t increases, then saturates after $T_t = 18$. Increasing T_t leads to better channel estimates in general as observed from the decrease in MSE (c.f., Fig. 4.5.c). However, after a certain number of pilots, decoding performance at the receiver is mainly determined by E_d , not by channel estimation quality. For $G_{SR}/G_{RD} = 30\text{dB}$ and $G_{SR}/G_{RD} = 0\text{dB}$, we can also make similar observations.
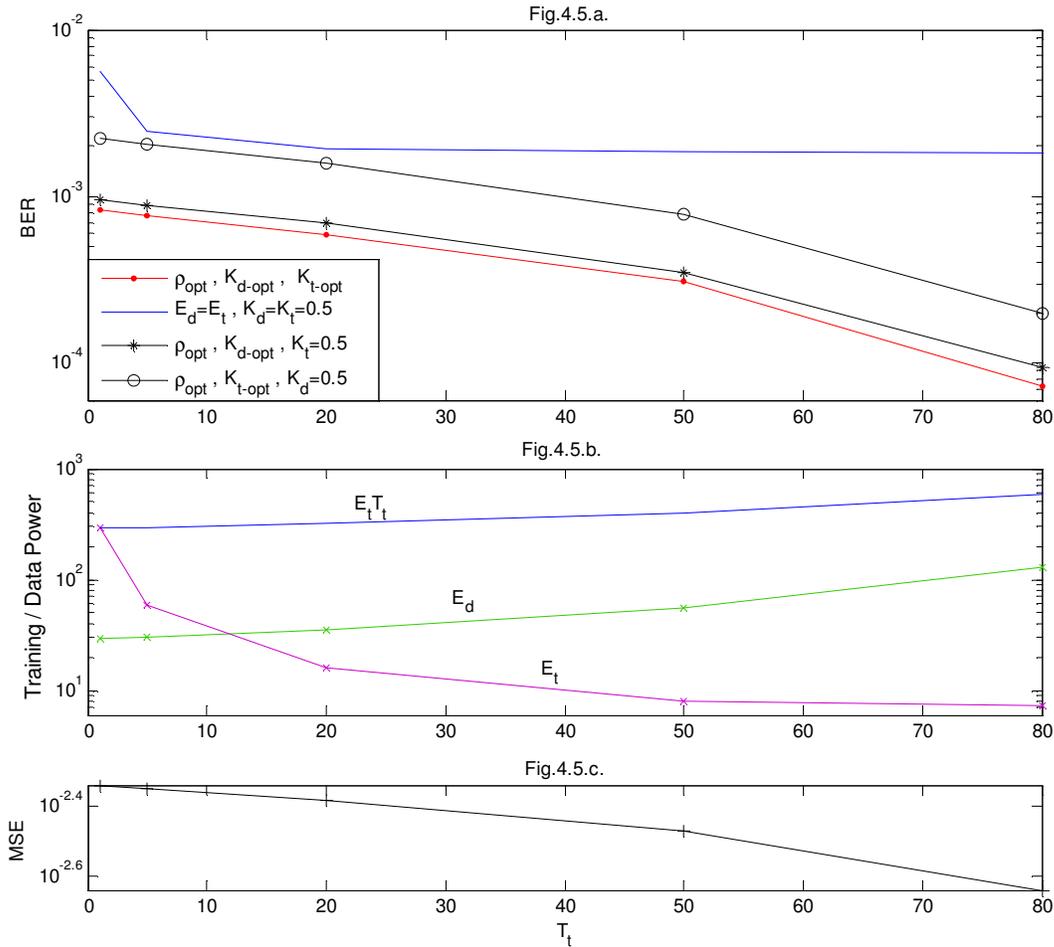
Fig.4.5: BER, training/data power, and MSE versus T_t

57

In Fig. 4.6, we study the effect of frame length of T. We illustrate BER, MSE, E_t and E_d in Figs.4.6.a-c with respect to T. We assume $G_{SR}/G_{RD} = -30\text{dB}$, deployment of a single pilot $(T_t = 1)$ and $E/N_0 = 15\text{dB}$. From Figs. 4.6.b and 4.6.c, we see that the increase in E_t results in smaller channel estimation errors, therefore, leading to less bit errors at the decoder. However, further increase in training power does not significantly affect error rate performance because after a certain training power is allocated, decoding performance is mainly determined by E_d. These findings are in line with Fig. 4.5 where we have made similar observations in terms of training length.
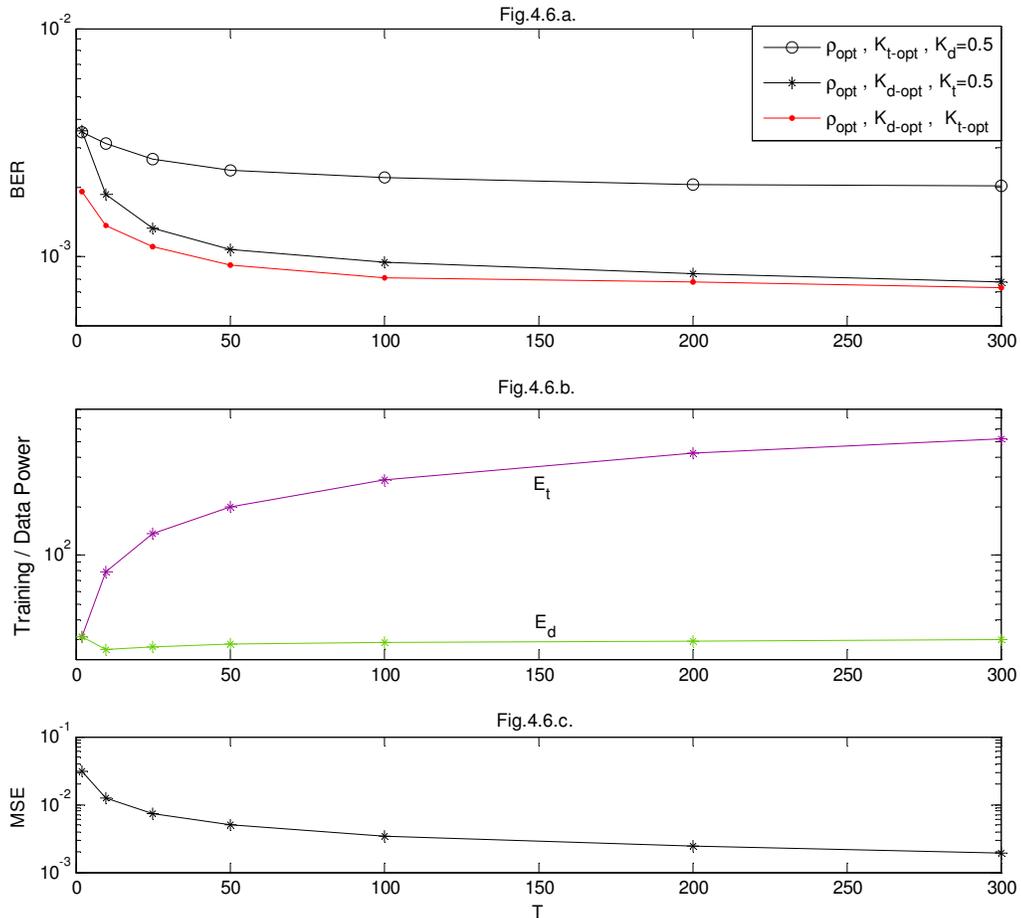
Fig.4.6.a: Fig.4.5: BER, training/data power, and MSE versus T.

Appendix

Appendix A - Derivation of (3.24)

In this appendix, we present the derivation of PEP for Golden coded TD protocol with mismatched-coherent detector. Replacing the channel estimate $\hat{\mathbf{h}} = \mathbf{h} - \mathbf{e}$ in the received signal vector $\mathbf{r} = \mathbf{X}_D \mathbf{h} + \mathbf{n}$, we have $\mathbf{r} = \mathbf{X}_D \hat{\mathbf{h}} + \bar{\mathbf{n}}$ where we define the effective noise term $\bar{\mathbf{n}} = \mathbf{X_D}\mathbf{e} + \mathbf{n}$. Under the assumption of Gaussian channel estimation errors, an upper PEP bound for transmitted codeword matrix $\mathbf{X}_D$ and erroneously decoded codeword matrix $\hat{\mathbf{X}}_D$ is given by [30], [56]

$$P\left(\mathbf{X}_D \rightarrow \hat{\mathbf{X}}_D \middle| \hat{\mathbf{h}}\right) \leq \exp\left(-\frac{\left\|(\mathbf{X}_D - \hat{\mathbf{X}}_D)\hat{\mathbf{h}}\right\|^2}{4\text{trace}\{\text{E}(\bar{\mathbf{n}}\bar{\mathbf{n}}^H)\}}\right) \tag{A.1}$$

where $\mathbf{X}_D$ is earlier defined by (3.11). For simplifying the ensuing derivation, we rewrite $\mathbf{r} = \mathbf{X}_D\hat{\mathbf{h}} + \mathbf{X}_D\mathbf{e} + \mathbf{n}$ as $\mathbf{r} = \hat{\mathbf{H}}\mathbf{x} + \bar{\mathbf{n}}$ where $\bar{\mathbf{n}} = \boldsymbol{\delta}\mathbf{x} + \mathbf{n}$. $\hat{\mathbf{H}}$ and $\boldsymbol{\delta}$ are, respectively, given by

$$\hat{\mathbf{H}} = \begin{bmatrix} \alpha\sqrt{E}\,\hat{h}_{SD} & \alpha\Theta\sqrt{E}\hat{h}_{SD} & 0 & 0 \\ \alpha\sqrt{A_1 E}\,\hat{h}_{SRD} & \alpha\Theta\sqrt{A_1 E}\hat{h}_{SRD} & i\bar{\alpha}\sqrt{A_2 E}\hat{h}_{SD} & i\bar{\alpha}\Theta\sqrt{A_2 E}\hat{h}_{SD} \\ 0 & 0 & \alpha\sqrt{E}\,\hat{h}_{SD} & \alpha\Theta\sqrt{E}\hat{h}_{SD} \\ \bar{\alpha}\sqrt{A_2 E}\hat{h}_{SD} & \bar{\alpha}\bar{\Theta}\sqrt{A_2 E}\hat{h}_{SD} & \alpha\sqrt{A_1 E}\,\hat{h}_{SRD} & \alpha\Theta\sqrt{A_1 E}\hat{h}_{SRD} \end{bmatrix}, \tag{A.2}$$

$$\boldsymbol{\delta} = \begin{bmatrix} \alpha\sqrt{E}\,e_{SD} & \alpha\Theta\sqrt{E}e_{SD} & 0 & 0 \\ \alpha\sqrt{A_1 E}\,e_{SRD} & \alpha\Theta\sqrt{A_1 E}e_{SRD} & i\bar{\alpha}\sqrt{A_2 E}e_{SD} & i\bar{\alpha}\bar{\Theta}\sqrt{A_2 E}e_{SD} \\ 0 & 0 & \alpha\sqrt{E}\,e_{SD} & \alpha\Theta\sqrt{E}e_{SD} \\ \bar{\alpha}\sqrt{A_2 E}e_{SD} & \bar{\alpha}\bar{\Theta}\sqrt{A_2 E}e_{SD} & \alpha\sqrt{A_1 E}\,e_{SRD} & \alpha\Theta\sqrt{A_1 E}e_{SRD} \end{bmatrix}. \tag{A.3}$$

Replacing (A.2) and (A.3) in (A.1), we have

$$P\left(\mathbf{x} \rightarrow \hat{\mathbf{x}} \middle| \hat{\mathbf{H}}\right) \leq \exp\left(-\frac{\left\|\hat{\mathbf{H}}(\mathbf{x} - \hat{\mathbf{x}})\right\|^2}{4\Lambda}\right) = \exp\left(-\frac{1}{4\Lambda}\text{trace}\left\{\hat{\mathbf{H}}\mathbf{A}\hat{\mathbf{H}}^H\right\}\right). \tag{A.4}$$

where we define $\mathbf{A} = (\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^H$ and

$$\Lambda = \text{trace}\left\{\text{E}\left(\bar{\mathbf{n}}\bar{\mathbf{n}}^H\right)\right\} = 4N_0 + 2\sigma_{eSD}^2\kappa_1 E + 2\sigma_{eSRD}^2\kappa_2 E. \tag{A.5}$$

$\sigma_{e_{SD}}^2$ and $\sigma_{e_{SRD}}^2$ denote the variances of S $\rightarrow$D and S$\rightarrow$R$\rightarrow$D link estimation errors and are given by

$$\sigma_{e_{SD}}^2 = (E/N_0)^{-1}\big/\big(N(2\kappa+\kappa')+(E/N_0)^{-1}\big), \tag{A.6}$$

$$\sigma_{e_{SRD}}^2 = (E/N_0)^{-1}\big/\big(2N\kappa_2+(E/N_0)^{-1}\big). \tag{A.7}$$

Since $\mathbf{A}$ in (A.4) is Hermitian and non-negative definite, it can be decomposed into $\mathbf{A} = \mathbf{UDU}^H$ where $\mathbf{U}$ is unitary matrix and $\mathbf{D} = \text{diag}\{\lambda_1,\lambda_2,\lambda_3,\lambda_4\}$ is a diagonal matrix having real-valued eigenvalues of $\mathbf{A}$. Replacing $\mathbf{A}$ with $\mathbf{UDU}^H$, (A.4) becomes

$$P\big(\mathbf{x} \rightarrow \hat{\mathbf{x}}\,|\,\hat{\mathbf{H}}\big) \leq \exp\left(-\frac{1}{4\Lambda}\text{trace}\Big\{\underbrace{\hat{\mathbf{H}}\mathbf{U}}_{\hat{\mathbf{H}}'}\mathbf{D}\underbrace{\mathbf{U}^H\hat{\mathbf{H}}^H}_{\hat{\mathbf{H}}'^H}\Big\}\right).$$

Noting that multiplication by a unitary matrix does not change the statistics of $\hat{\mathbf{H}}$, we have

$$P\big(\mathbf{x} \rightarrow \hat{\mathbf{x}}\,|\,\hat{\mathbf{h}}\big) \leq \exp\left(-E\frac{|\alpha|^2\big(\overline{\lambda}_1+\Theta^2\overline{\lambda}_2\big)\big|\hat{h}_{SD}\big|^2 +|\overline{\alpha}|^2 A_2\big(\overline{\lambda}_1+\overline{\Theta}^2\overline{\lambda}_2\big)\big|\hat{h}_{SD}\big|^2 +|\alpha|^2 A_1\big(\overline{\lambda}_1+\Theta^2\overline{\lambda}_2\big)\big|\hat{h}_{SRD}\big|^2}{4\Lambda}\right) \tag{A.8}$$

where $\overline{\lambda}_1 = \lambda_1+\lambda_3$ and $\overline{\lambda}_2 = \lambda_2+\lambda_4$. Defining $\lambda_{\min} = \min\{\overline{\lambda}_1,\overline{\lambda}_2\}$, the unconditional PEP can be found as

$$P\big(\mathbf{x} \rightarrow \hat{\mathbf{x}}\big) < \Psi_{|\hat{h}_{SD}|^2}\left(-\frac{E\kappa_1\lambda_{\min}}{4\Lambda}\right)\Psi_{|\hat{h}_{SRD}|^2}\left(-\frac{E\kappa_2\lambda_{\min}}{4\Lambda}\right) \tag{A.9}$$

where $\Psi_{|\hat{h}_{SD}|^2}$ and $\Psi_{|\hat{h}_{SRD}|^2}$ are the moment generating functions (MGFs) of $\big|\hat{h}_{SD}\big|^2$ and $\big|\hat{h}_{SRD}\big|^2$ respectively. $\big|\hat{h}_{SD}\big|^2$ is central chi-squared distributed with second degree of freedom. Hence its MGF can be readily found as [46]

$$\Psi_{|\hat{h}_{SD}|^2}(-s) = 1\big/\big(1+s\sigma_{\hat{h}_{SD}}^2\big) \tag{A.10}$$

where the variance of S $\rightarrow$D channel estimate is $\sigma_{\hat{h}_{SD}}^2 = 1-\sigma_{e_{SD}}^2$. For the S$\rightarrow$R$\rightarrow$D channel, $\hat{h}_{SRD}$ is the estimate of the product of two Gaussian terms, i.e., $h_{SRD} = h_{SR}h_{RD}$. Unfortunately, the exact distribution function of $\hat{h}_{SRD}$ is unknown. Here, we follow a similar approach to [31] where the estimate of cascaded channels is modeled as the product of estimates of individual channels: Assume S$\rightarrow$R and R$\rightarrow$D channels are estimated individually as

$h_{SR} = \hat{h}_{SR} + e_{SR}$ and $h_{RD} = \hat{h}_{RD} + e_{RD}$ where the estimation errors are modeled as zero-mean complex Gaussian random variables with variances given as

$$\sigma_{e_{SR}}^2 = (E/N_0)^{-1} / \left(2NG_{SR}\kappa + (E/N_0)^{-1}\right), \tag{A.11}$$

$$\sigma_{e_{RD}}^2 = (E/N_0)^{-1} / \left(NG_{RD}\kappa + (E/N_0)^{-1}\right). \tag{A.12}$$

The variance of $\hat{h}_{SR}\hat{h}_{RD}$ (i.e., the product of individual estimates) is then found as

$$\text{var}(\hat{h}_{SR}\,\hat{h}_{RD}) = \left(1 - \sigma_{e_{SR}}^2\right)\left(1 - \sigma_{e_{RD}}^2\right) = 1 - \sigma_{e_{SR}}^2 - \sigma_{e_{RD}}^2 + \sigma_{e_{SR}}^2\sigma_{e_{RD}}^2 \;. \tag{A.13}$$

On the other hand, the variance of $\hat{h}_{SRD}$ (i.e., estimate of the cascaded channel) is given by

$$\text{var}(\hat{h}_{SRD}) = 1 - \sigma_{e_{SRD}}^2 = 2N\kappa_2 / \left(2N\kappa_2 + (E/N_0)^{-1}\right)$$

Asymptotic relative efficiency of two estimates is defined as [49]

$$\frac{\text{var}(\hat{h}_{SR}\,\hat{h}_{RD})}{\text{var}(\hat{h}_{SRD})} = \frac{1 - \sigma_{e_{SR}}^2 - \sigma_{e_{RD}}^2 + \sigma_{e_{SR}}^2\sigma_{e_{RD}}^2}{1 - \sigma_{e_{SRD}}^2} \;. \tag{A.14}$$

For high SNR and sufficiently large pilot numbers, the relative efficiencies of two estimators become the same, i.e.,

$$\lim_{\substack{E/N_0 \to \infty \\ N \to \infty}} \frac{\text{var}(\hat{h}_{SR}\,\hat{h}_{RD})}{\text{var}(\hat{h}_{SRD})} = 1 \tag{A.15}$$

indicating that the statistics of two estimates converge to each other and, ultimately, two estimators perform equivalently. Under this assumption, we have $\Psi_{|\hat{h}_{SRD}|^2}(-s) \approx \Psi_{|\hat{h}_{SR}\hat{h}_{RD}|^2}(-s)$ which is given by

$$\Psi_{|\hat{h}_{SR}\hat{h}_{RD}|^2}(-s) = \frac{1}{s\sigma_{\hat{h}_{SR}}^2\sigma_{\hat{h}_{RD}}^2} \exp\left(\frac{1}{s\sigma_{\hat{h}_{SR}}^2\sigma_{\hat{h}_{RD}}^2}\right)\Gamma\left(0, \frac{1}{s\sigma_{\hat{h}_{SR}}^2\sigma_{\hat{h}_{RD}}^2}\right). \tag{A.16}$$

Replacing $\Psi_{|\hat{h}_{SD}|^2}$ and $\Psi_{|\hat{h}_{SR}\hat{h}_{RD}|^2}$ in (A.9), we obtain

$$P(\mathbf{x} \to \hat{\mathbf{x}}) < \frac{4SNR_{eff}^{-2}/\left(\lambda_{\min}^2\kappa_1\kappa_2\right)}{\left(1 - \sigma_{e_{SD}}^2\right)\left(1 - \sigma_{e_{SR}}^2\right)\left(1 - \sigma_{e_{RD}}^2\right)} \exp\left(\frac{2SNR_{eff}^{-1}/\left(\lambda_{\min}\kappa_2\right)}{\left(1 - \sigma_{e_{SR}}^2\right)\left(1 - \sigma_{e_{RD}}^2\right)}\right)\Gamma\left(0, \frac{2SNR_{eff}^{-1}/\left(\lambda_{\min}\kappa_2\right)}{\left(1 - \sigma_{e_{SR}}^2\right)\left(1 - \sigma_{e_{RD}}^2\right)}\right) \tag{A.17}$$

which yields (3.24).

Appendix B - Derivation of (3.30)

In this appendix, we present the derivation of PEP for RD protocol with mismatched-coherent detector. The Chernoff bound on the PEP is given by (3.44) where $\mathbf{X}_D$ is defined by (3.15) and the effective noise term $\bar{\mathbf{n}} = \mathbf{n} + \mathbf{X}_D \mathbf{e}$ is assumed to be Gaussian with

$$\Lambda = \text{trace}\left(\mathbb{E}\left\{\bar{\mathbf{n}}\bar{\mathbf{n}}^H\right\}\right) = 2N_0 + B_1 \sigma_{eSRD}^2 E + \sigma_{eSD}^2 E. \tag{A.18}$$

The unconditional PEP can be found as

$$P(x \rightarrow \hat{x}) \leq \Psi_{|\hat{h}_{SD}|^2}\left(-\frac{E\lambda}{4\Lambda}\right)\Psi_{|\hat{h}_{SRD}|^2}\left(-\frac{B_1 E \lambda}{4\Lambda}\right) \tag{A.19}$$

where $\lambda = |x - \hat{x}|^2$. $\Psi_{|\hat{h}_{SD}|^2}$ and $\Psi_{|\hat{h}_{SRD}|^2}$ are already given by (A.10) and (A.16), respectively. Replacing them in (A.19), we have

$$P(x_1 \rightarrow \hat{x}_1) \leq \frac{4 SNR_{eff}^{-2}\big/\left(\lambda^2 B_1\right)}{\left(1 - \sigma_{eSD}^2\right)\left(1 - \sigma_{eSR}^2\right)\left(1 - \sigma_{eRD}^2\right)} \exp\left(\frac{2 SNR_{eff}^{-1}\big/\left(\lambda B_1\right)}{\left(1 - \sigma_{eSR}^2\right)\left(1 - \sigma_{eRD}^2\right)}\right)\Gamma\left(0, \frac{2 SNR_{eff}^{-1}\big/\left(\lambda B_1\right)}{\left(1 - \sigma_{eSR}^2\right)\left(1 - \sigma_{eRD}^2\right)}\right) \tag{A.20}$$

which yields (3.30).

Appendix C – Derivation of (3.36)

In this appendix, we present the derivation of PEP for RD protocol with partially-coherent detector. For this case, PEP is given by

$$P(\mathbf{X}_D \to \hat{\mathbf{X}}_D | \hat{\varphi}, |\mathbf{h}|) = P\left(\|\mathbf{r} - \mathbf{X}_D \hat{\varphi}\|^2 > \|\mathbf{r} - \hat{\mathbf{X}}_D \hat{\varphi}\|^2 \right) \tag{A.21}$$

$$= P\left(\|\mathbf{r}\|^2 - 2\,\mathrm{Re}\{\mathbf{r}^H(\mathbf{X}_D \hat{\varphi})\} + \|\mathbf{X}_D \hat{\varphi}\|^2 > \|\mathbf{r}\|^2 - 2\,\mathrm{Re}\{\mathbf{r}^H(\hat{\mathbf{X}}_D \hat{\varphi})\} + \|\hat{\mathbf{X}}_D \hat{\varphi}\|^2 \right) \tag{A.22}$$

Recall that $\mathbf{X}_D$ and $\hat{\mathbf{X}}_D$ consist of M-PSK modulated transmitted symbols with unit power. Thus, we have $\|\mathbf{X}_D \hat{\varphi}\|^2 = \|\hat{\mathbf{X}}_D \hat{\varphi}\|^2$ which lets us rewrite (A.22) as

$$P(\mathbf{X}_D \to \hat{\mathbf{X}}_D | \hat{\varphi}, |\mathbf{h}|) = P\left(2\,\mathrm{Re}\{\mathbf{r}^H(\hat{\mathbf{X}}_D \hat{\varphi} - \mathbf{X}_D \hat{\varphi})\} > 0 \right) \tag{A.23}$$

$$= P\left(2\,\mathrm{Re}\{\mathbf{h}^H \mathbf{X}_D^H(\hat{\mathbf{X}}_D - \mathbf{X}_D)\hat{\varphi} + \mathbf{n}^H(\hat{\mathbf{X}}_D - \mathbf{X}_D)\hat{\varphi}\} > 0 \right) . \tag{A.24}$$

where (A.24) is simply obtained from (A.23) replacing $\mathbf{r} = \mathbf{X}_D \mathbf{h} + \mathbf{n}$. Defining the phase difference between the transmitted symbol and incorrect decision as $\theta_\Delta = \theta_x - \theta_{\hat{x}}$, the variance of the effective noise $n' = 2\,\mathrm{Re}\{\mathbf{n}^H(\hat{\mathbf{X}}_D - \mathbf{X}_D)\hat{\varphi}\}$ can be obtained as [53]

$$\mathrm{var}(n') = 2N_0 E(1 + B_1)|\hat{x} - x|^2 = 4N_0 E(1 - \cos(\theta_\Delta))(1 + B_1) \tag{A.25}$$

where $|\hat{x} - x|^2 = |x|^2 + |\hat{x}|^2 - 2\,\mathrm{Re}\{\hat{x}^* x\} = 2(1 - \cos(\theta_\Delta))$. Using $\varepsilon_{SRD} = \phi_{SRD} - \hat{\phi}_{SRD}$ and $\varepsilon_{SD} = \phi_{SD} - \hat{\phi}_{SD}$, (A.24) can be rewritten as

$$P(\mathbf{X}_D \to \hat{\mathbf{X}}_D | \varepsilon, |\mathbf{h}|) = P\left(2\,\mathrm{Re}\left\{ \left(|h_{SD}| E e^{-j\varepsilon_{SD}} + |h_{SRD}| B_1 E e^{-j\varepsilon_{SRD}} \right) x^*(\hat{x} - x) \right\} + n' > 0 \right) \tag{A.26}$$

$$= P\left(n' > 2|h_{SD}| E(\cos(\varepsilon_{SD}) - \cos(\varepsilon_{SD} + \theta_\Delta)) + 2|h_{SRD}| B_1 E(\cos(\varepsilon_{SRD}) - \cos(\varepsilon_{SRD} + \theta_\Delta)) \right)$$

$$= Q\left(\frac{2|h_{SD}| E(\cos(\varepsilon_{SD}) - \cos(\varepsilon_{SD} + \theta_\Delta)) + 2|h_{SRD}| B_1 E(\cos(\varepsilon_{SRD}) - \cos(\varepsilon_{SRD} + \theta_\Delta))}{\sqrt{\mathrm{var}(n')}} \right) \tag{A.27}$$

Unconditional PEP can be found by averaging (A.27) with respect to $|\mathbf{h}|$ and ε. However, this averaging can be much complicated due to cross terms [46]. In order to simplify proceeding derivation steps, we further upper bound (A.27) as

$$P(x \to \hat{x} \mid \varepsilon, |\mathbf{h}|)$$
$$< Q\left(\sqrt{\frac{|h_{SD}|^2 E\left(\cos(\varepsilon_{SD}) - \cos(\varepsilon_{SD} + \theta_\Delta)\right)^2 + |h_{SRD}|^2 B_1^2 E\left(\cos(\varepsilon_{SRD}) - \cos(\varepsilon_{SRD} + \theta_\Delta)\right)^2}{N_0\left(1 - \cos(\theta_\Delta)\right)\left(1 + B_1\right)}} \right) \quad (A.28)$$

ignoring the cross terms resulting from squaring the numerator of the Q function in (A.27). Under the assumption that phase estimate errors are sufficiently small for high PLL loop gain, the expectation of (A.28) with respect to ε can be approximated by replacing the terms $\cos(\varepsilon_{SD}) - \cos(\varepsilon_{SD} + \theta_\Delta)$ and $\cos(\varepsilon_{SRD}) - \cos(\varepsilon_{SRD} + \theta_\Delta)$ with their expected values [54], i.e., $\cos(\varepsilon_{SD}) - \cos(\varepsilon_{SD} + \theta_\Delta) \approx (1 - \cos(\theta_\Delta)) \mathrm{E}\{\cos(\varepsilon_{SD})\} + \sin(\theta_\Delta) \mathrm{E}\{\sin(\varepsilon_{SD})\}$ where

$$\mathrm{E}\{\sin(\varepsilon_{SD})\} \cong \mathrm{E}\{\varepsilon_{SD}\} = 1/\rho_{SD} , \quad (A.29)$$

$$\mathrm{E}\{\cos(\varepsilon_{SD})\} = \int_{-\pi}^{\pi} \cos(\varepsilon'_{SD}) \frac{e^{\rho_{SD}\cos(\varepsilon'_{SD})}}{2\pi \mathrm{I}_0(\rho_{SD})} d\varepsilon'_{SD} = I_1(\rho_{SD})/I_0(\rho_{SD}) \quad . \quad (A.30)$$

Here, $I_1(.)$ and $I_0(.)$ are the first and zero order modified Bessel functions of the first kind and $\rho_{SD} = |h_{SD}|^2 E/N_0 B_L T$. Similarly, for S→R→D link phase error, we obtain $\cos(\varepsilon_{SRD}) - \cos(\varepsilon_{SRD} + \theta_\Delta) \approx (1 - \cos(\theta_\Delta)) I_1(\rho_{SRD})/I_0(\rho_{SRD}) + \sin(\theta_\Delta)/\rho_{SRD}$ where $\rho_{SRD} = |h_{SR}|^2 |h_{RD}|^2 B_1 E/N_0 B_L T$. Replacing these approximations in (A.28) and applying the Chernoff bound, we have

$$P(x \to \hat{x} \mid |\mathbf{h}|) < \exp\left(-\frac{|h_{SD}|^2 E\left(\Delta_1 I_1(\rho_{SD})/I_0(\rho_{SD}) + \sin(\theta_\Delta)/\rho_{SD}\right)^2}{\Omega} \right)$$
$$\times \exp\left(-\frac{|h_{SRD}|^2 B_1^2 E\left(\Delta_1 I_1(\rho_{SRD})/I_0(\rho_{SRD}) + \sin(\theta_\Delta)/\rho_{SRD}\right)^2}{\Omega} \right) \quad (A.31)$$

where $\Delta_1 = 1 - \cos(\theta_\Delta)$ and $\Omega = 2N_0\Delta_1(1 + B_1)$. To find the unconditional PEP, we still need to take an expectation with respect to $|\mathbf{h}|$. For high loop SNRs, the ratio of two Bessel functions in (A.30) can be approximated as [50]

$$I_1(\rho_{SD})/I_0(\rho_{SD}) \approx 1 - 1/2\rho_{SD} \text{ and } I_1(\rho_{SRD})/I_0(\rho_{SRD}) \approx 1 - 1/2\rho_{SRD} . \quad (A.32)$$

Then, (A.31) simplifies to

$$P(x \rightarrow \hat{x}) < \underbrace{\text{E}_{|h_{SD}|^2} \left\{ \exp\left(-\frac{E|h_{SD}|^2 \left(\Delta_1 - (1 - \cos(\theta_\Delta) - 2\sin(\theta_\Delta))/2\rho_{SD} \right)^2}{\Omega} \right) \right\}}_{E_1}$$

$$\times \underbrace{\text{E}_{|h_{SR}h_{RD}|^2} \left\{ \exp\left(-\frac{E|h_{SRD}|^2 B_1^2 \left(\Delta_1 - (1 - \cos(\theta_\Delta) - 2\sin(\theta_\Delta))/2\rho_{SRD} \right)^2}{\Omega} \right) \right\}}_{E_2}$$

(A.33)

Noting $|h_{SD}|^2$ is central chi-squared distributed with second degree of freedom, E_1 can be calculated as

$$E_1 = \exp\left(\frac{\Delta_1 \Delta_2}{\Omega/N_0 B_L T} \right) \int_0^\infty \exp\left(-\alpha\left(\frac{\Delta_1^2 E}{\Omega} + 1 \right) - \frac{1}{4\alpha}\left(\frac{\Delta_2^2 (N_0 B_L T)^2}{E\Omega} \right) \right) d\alpha$$

(A.34)

where $\Delta_2 = 1 - \cos(\theta_\Delta) - 2\sin(\theta_\Delta)$.

Using the integral result $\int_0^\infty \exp(-\beta/4x - tx)dx = \sqrt{\beta/t}\,K_1(\sqrt{\beta t})$ [48], (A.34) yields

$$E_1 = \exp\left(\frac{\Delta_1 \Delta_2}{\Omega/N_0 B_L T} \right) \sqrt{\frac{\Delta_2^2 (N_0 B_L T/E)^2}{\Delta_1^2 + \Omega/E}} K_1\left(\sqrt{\Delta_2^2 \left(\Delta_1^2 \left(\frac{N_0 B_L T}{\Omega} \right)^2 + \frac{(N_0 B_L T)^2}{E\Omega} \right)} \right)$$

(A.35)

where K_1 is the first order modified Bessel function of second kind. Noting $|h_{SR}|^2$ and $|h_{RD}|^2$ are chi-squared distributed, we have E_2 as

$$E_2 = \exp\left(\frac{B_1 \Delta_1 \Delta_2}{\Omega/N_0 B_L T} \right)$$
$$\times \int_0^\infty \sqrt{\frac{\Delta_2^2 (N_0 B_L T/\alpha_2 B_1 E)^2}{\Delta_1^2 + \Omega/\alpha_2 B_1^2 E}} \, K_1\left(\sqrt{\Delta_2^2 B_1^2 \left(\Delta_1^2 \left(\frac{N_0 B_L T}{\Omega} \right)^2 + \frac{(N_0 B_L T)^2}{\alpha_2 B_1^2 E\Omega} \right)} \right) \exp(-\alpha_2)\, d\alpha_2$$

(A.36)

There is, unfortunately, no closed form solution of the above integral. However, using the approximation $K_\nu(z) \approx 0.5\Gamma(\nu)(2/z)^\nu$, $z < 1$ [50] which gives satisfactory results for small $B_L T$ values, we can readily solve (A.36) as

$$E_2 = \frac{\Omega}{B_1^2 E\Delta_1^2} \exp\left(\frac{B_1 \Delta_1 \Delta_2}{\Omega/N_0 B_L T} + \frac{\Omega}{B_1^2 E\Delta_1^2} \right) \Gamma\left(0, \frac{\Omega}{B_1^2 E\Delta_1^2} \right).$$

(A.37)

Replacing (A.35) and (A.37) in (A.33), we have

$$P(x \rightarrow \hat{x}) < \frac{SNR_{eff}^{-2}}{B_1^2 \Delta_1^2} \sqrt{\frac{(\Delta_2 B_L T / (2\Delta_1 + 2B_1 \Delta_1))^2}{\Delta_1^2 + SNR_{eff}^{-1}}} \exp\left(\frac{\Delta_2 B_L T}{2} + \frac{SNR_{eff}^{-1}}{B_1^2 \Delta_1^2} \right)$$

$$\times K_1\left(\sqrt{\left(\frac{\Delta_2 B_L T}{2 + 2B_1} \right)^2 + \frac{(\Delta_2 B_L T / (2 + 2B_1))^2 SNR_{eff}^{-1}}{\Delta_1^2}} \right) \Gamma\left(0, \frac{SNR_{eff}^{-1}}{B_1^2 \Delta_1^2} \right)$$

(A.38)

which yields (3.36).

References

[1] J.E Padgett, C. G. Gunther, and T. Hattori, "Overview of wireless personal communications," *IEEE Commun. Mag.*, vol.33,   pp. 28 – 41, Jan. 1995.

[2] C. K. Toh, *Ad-hoc Mobile Wireless Networks: Protocols and Systems*, Prentice Hall, 2002

[3] S. Y. Hui and K. H. Yeung, "Challenges in the Migration to 4G Mobile Systems," *IEEE Commun. Mag.*, vol. 41, pp. 54 - 59, Dec. 2003.

[4] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, 2000.

[5] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.

[6] A. Wittneben, "Base station modulation diversity for digital SIMULCAST," in *Proc. of IEEE Vehicular Technology Conf. (VTC),* pp. 848-853, May 1991.

[7] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *Proc. IEEE International Conf. Commun. (ICC)*, pp. 1630-1634, May 1993.

[8] J. H. Winters, "Diversity gain of transmit diversity in wireless systems with Rayleigh fading," *IEEE Trans. Veh. Tech.*, vol. 47, pp. 119-123, February 1998.

[9] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rates wireless communications: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744-765, 1998.

[10] V. Tarokh, H. Jafarkhani, and A.R. Calderbank, "Space-time block coding from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456-1467, 1999.

[11] V. Tarokh, H. J. Jafarkhani and A. R. Calderbank, "Space-time block coding for wireless communications: Performance results," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 451-460, March 1999.

[12] S. M. Alamouti, "Simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1451-1458, 1998.

[13] B. Vucetic and J. Yuan, *Space-Time Coding*, John Wiley & Sons, May 2003.

[14] A. Paulraj, D. Gore and R. Nabar, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, May 2003.

[15] H. Jafarkhani, *Space-Time Coding: Theory and Practice*, Cambridge University Press, September 2005.

[16] E. V. D. Meulen, "Three-terminal communication channels," *Advances in Applied Probability*, vol. 3, pp. 120-154, 1971.

[17] T. M. Cover and A. El Gamal. "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572-584, Sept. 1979.

[18] A. Sendonaris, E. Erkip, and B. Aazhang, "User Cooperation Diversity – Part I: System Description," *IEEE Trans. Commun.*, vol. 51, pp. 1927-1938, Nov. 2003.

[19] A. Sendonaris, E. Erkip, and B. Aazhang, "User Cooperation Diversity – Part II: Implementation, Aspects and Performance Analysis," *IEEE Trans. Commun.*, vol. 51, pp. 1939-1948, Nov. 2003.

[20] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Trans. Inform. Theory,* vol.50, no.12, pp. 3062-3080, Dec. 2004.

[21] J. N. Laneman, G. W. Wornell, and D. N. C. Tse, "An Efficient Protocol for Realizing Cooperative Diversity in Wireless Networks," in *Proc. IEEE International Symposium Inform. Theory (ISIT)*, , June 2001.

[22] J. N. Laneman and G. W. Wornell, "Distributed Space-Time Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415-2525, Oct. 2003.

[23] R. U. Nabar, H. Boelcskei, and F. W. Kneubhueler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun.*, vol 22, pp. 1099-1109, Aug. 2004.

[24] H. Ochiai, P. Mitran, and V. Tarokh, "Variable-Rate two-Phase Collaborative Communication Protocols for Wireless Networks", *IEEE Trans. Inform. Theory,* vol. 52, no. 9, pp. 4299-4313, September 2006.

[25] K. Azarian, H. E. Gamal, and P. Schniter, "On the Achievable Diversity-Multiplexing Tradeoff in Half-duplex Cooperative channels", *IEEE Trans. Inform. Theory,* vol. 51, no. 12, pp. 4152 - 4172, Dec. 2005.

[26] J.N. Laneman, "Cooperative diversity in wireless networks: algorithms and architectures", Ph.D dissertation, Massachusetts Institute of Technology, Aug. 2002.

[27] H. Mheidat and M. Uysal, "Impact of Receive Diversity on the Performance of Amplify-and-Forward Relaying under APS and IPS Power Constraints", *IEEE Commun. Letters*, vol. 10, no. 6, pp. 468-470, June 2006.

[28] B. Gedik and M. Uysal, "Two Channel Estimation Methods for Amplify-and-Forward Relay Networks", accepted for presentation in *IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, May 2008.

[29] M. Uysal and H. Mheidat, "Maximum-Likelihood Detection for Distributed Space-Time Block Coding", in *Proc. of IEEE Vehicular Technology Conf. (VTC)*, Sept 2004.

[30] H. Mheidat and M. Uysal, "Non-Coherent and Mismatched-Coherent Receivers for Distributed STBCs with Amplify-and-Forward Relaying", *IEEE Trans. Wireless Commun.*, vol. 6, no. 11, pp. 4060-4070, Nov. 2007.

[31] C. S. Patel and G.L. Stuber, "Channel Estimation for Amplify and Forward Relay Based Cooperation Diversity Systems", *IEEE Trans. Wireless Commun.,* vol. 6, no. 6, pp . 2348-2356, June 2007.

[32] B. Gedik and M. Uysal, "Amplify-and-Forward Relaying with Mismatched-Coherent and Partially-Coherent Receivers", *IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2008.

[33] Berna Gedik and Murat Uysal, "Impact of Imperfect Channel Estimation on the Performance of Amplify-and-Forward Relaying", under review in *IEEE Trans. Wireless Commun.*

[34] B. Wang, J. Zhang, and L. Zheng, "Achievable Rates and Scaling Laws of Power-Constrained Wireless Sensory Relay Networks", *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp 4084- 4104, Sept 2006.

[35] T. Cui, F. Gao, and A. Nallanathan, "Optimal Training Design for Channel estimation in Amplify and Forward Relay Networks", in *Proc. of Globecom*, Nov. 2007.

[36] J. Zhang and M.C. Gursoy, "Achievable Rates and Optimal Resource Allocation for Imperfectly Known Relay Channels", in *Proc. of 45 th. Annual Allerton Conf. on Commun.*, Sept. 2007.

[37] B. Gedik and M. Uysal, "How Much Training Power Do We Need for Amplify-and-Forward Relay Channels?", under review in *IEEE Trans. Signal Proc.*

[38] J. W. Mark and W. Zhuang, *Wireless Communications and Networking*, Prentice Hall, New Jersey, 2003.

[39] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, New Jersey, 1993.

[40] D. Chen and J. N. Laneman, "Cooperative Diversity for Wireless Fading Channels without Channel state Information", in *Proc. Asilomar Conf. Signals*, Systems and Computers, Nov. 2004.

[41] G. Taricco and E. Biglieri, "Space-Time Decoding With Imperfect Channel Estimation", *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1874-1888, July 2005.

[42] Y. Linde, A. Buzo, and R. M. Gray, "An Algorithm for Vector Quantizer Design", *IEEE Trans. Commun.,* vol. 28, no.11, pp. 84-95, Jan. 1980.

[43] W. Su and X. Xia, "Signal Constellations for Quasi-Orthogonal Space-Time Block Codes with Full Diversity", *IEEE Trans. Inform. Theory*, vol. 50, no.10, pp. 2331–2347, Oct. 2003.

[44] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The Golden Code: A 2 x 2 Full-Rate Space-Time Code with Non-Vanishing Determinants", *IEEE Trans. Inform. Theory*, vol. 51, no 4, April 2005.

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

[46] M. K. Simon and M.-S Alouini, *Digital Communications over Fading Channels*, Wiley, New York, 2000.

[47] A. J. Viterbi, *Principles of Coherent Communication*, McGraw-Hill, New York, 1966.

[48] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, 5th ed, San Diago,1994.

[49] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, Wiley, New York, 1980.

[50] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover Pub., New York, 1964.

[51] M. M. Fareed and M. Uysal, "BER-Optimized Power Allocation for Fading Relay Channels", accepted for publication in *IEEE Trans. Wireless Commun.*

[52] B. Hassibi and B.M. Hochwald, "How much training is needed in multiple antenna wireless links", *IEEE Trans. Inform. Theory*, vol. 49, no 4, pp. 951-963, April 2003.

[53] E. Zehavi and G. Kaplan, "Phase noise Effects on M-ary PSK Trellis Codes", *IEEE Trans. Commun.*, vol. 39, no. 3, pp. 373-379, March 1991.

[54] T. Eng and L. B. Milstein, "Partially coherent DS-SS performance in frequency selective multipath Fading", *IEEE Trans. Wireless Comm.,* vol. 45, no. l, pp 110-118, Jan. 1997.

[55] W. J. Weber, "Performance of Phase-Locked Loops in the presence of fading communication channels", *IEEE Trans. Commun.*, vol. 24, no. 5, pp. 487-499, May 1976.

[56] K. Ahmed, C. Tepedelenlioglu, and A. Spanias, "Performance of precoded OFDM with channel estimation error," *IEEE Trans. Signal Proc.*, vol. 54, no. 3, pp. 1165 -1171, March 2006.