Differential Amplitude/Phase Space-Time Modulation

by

Juan Zi

B. Eng., Huazhong University of Science and Technology Hankou Branch, 1994

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
Master of Applied Science
in
THE FACULTY OF GRADUATE STUDIES
(Defense of Electrical and Computer Engineering)

We accept this thesis as conforming
to the required standard

The University of British Columbia
December 2003
© Juan Zi, 2003
Library Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Juan Z
Name of Author (please print) Dec. 10, 2003
Date

Title of Thesis: Differential Amplitude/Phase space-time modulation

Degree: M.A.Sc. Year: 2003
Abstract

In this thesis, a differential amplitude/phase space-time modulation (DAPSTM) is proposed for multiple transmit antenna wireless systems over flat Rayleigh fading channels. Two conventional noncoherent detection schemes, namely, simply heuristic (SH) differential detection (DD) and maximum likelihood (ML) DD are presented. Furthermore, two improved noncoherent detection schemes, multiple-symbol detection (MSD) and decision-feedback DD (DF-DD) with lower decoding complexity are derived. By taking the dependencies among the received symbols into account, MSD and DF-DD can reduce the error floor of ML-DD. The pairwise error probability (PEP) based on SH-DD, and an approximation of the bit error rate (BER) based on the union bound, are derived. Analytical considerations agree well with the simulation results.

Compared with the known differential unitary space-time modulation (DUSTM), DAPSTM can be said to generalize the diagonal structure from phase signals to a combination of phase signals and amplitude signals. This generalization potentially allows the spectral efficiency to be increased by carrying information, not only in phases, but also in amplitudes.

DAPSTM is not as power efficient as space-time codes with differential amplitude/phase shift keying (STC-DAPSK), which is based on Alamouti’s orthogonal space-time code (OSTC), when two transmit antennas are employed. However, DAPSTM allows easy implementation at the transmitter, due to the group property of its constellation under matrix
multiplication. DAPSTM can be employed for an arbitrary number of transmit antennas while keeping full diversity and full rate. It is also suitable for exploiting time diversity when only one transmit antenna is used in the system. In contrast, STC-DAPSK can only achieve full diversity and full rate for two transmit antennas and can not exploit time diversity, due to its nondiagonal structure.
# Contents

Abstract ii

List of Tables vii

List of Figures ix

Acknowledgments x

1 Introduction 1

1.1 Diversity ................................................................. 1

1.2 Motivation and Objectives ............................................. 2

1.3 Contributions .............................................................. 3

1.4 Thesis Outline ............................................................ 4

2 Background and Related Work 6

2.1 System Model ............................................................... 6

2.1.1 Space Diversity System Model ..................................... 6

2.1.2 Time Diversity System Model ...................................... 7

2.2 Channel Model ............................................................. 10

2.2.1 Flat Rayleigh Fading Channel Model .............................. 10

2.2.2 MIMO Channel Model ............................................... 12

2.3 Differential Modulation and Detection .............................. 14
3 Differential Amplitude and Phase Space-Time Modulation

3.1 DAPSTM ................................................. 27
3.2 Conventional Differential Detection .................................. 31
   3.2.1 Simple Heuristic Differential Detection ....................... 32
   3.2.2 Maximum Likelihood Differential Detection .................. 33
3.3 Multiple-Symbol Detection ....................................... 38
3.4 Decision-Feedback Differential Detection ........................ 41
3.5 Modifications for Time Diversity ................................. 43

4 Performance Analysis

4.1 Pairwise Error Probability ..................................... 44
4.2 Approximation for Bit Error Rate ................................ 48

5 Simulation Results

5.1 Numerical Results Compared with Simulations ....................... 49
5.2 Constellations Design ......................................... 52
5.3 Space Diversity .............................................. 55
5.4 Time Diversity ............................................... 61

6 Conclusions and Recommendations

6.1 Conclusions .................................................. 64
6.2 Recommendations for Future Work ................................. 65

Glossary .......................... 69
CONTENTS vi

Bibliography 70
List of Tables

2.1 DUSTM parameters for $N_T = 2$. .............................................. 21
2.2 DUSTM parameter for $N_T = 3$. .............................................. 22
5.1 DAPSTM parameters for $N_T = 2$ by minmax PEP. ......................... 54
5.2 DAPSTM parameters for $N_T = 2$ and $N_T = 3$. ............................ 55
List of Figures

2.1 Space diversity system model .............................................. 7
2.2 Time diversity system model ............................................. 8
2.3 $N_B \times N_I$ rectangular interleaver .................................. 9
2.4 Generic model for a MIMO channel ..................................... 12
2.5 Structure of the differential encoder for MDPSK modulation ......... 14
2.6 Block diagram of the conventional differential detector for MDPSK .. 15
2.7 Structure of the differential encoder for MDAPSK modulation .......... 16
2.8 Signal constellation for 16DAPSK ....................................... 18
2.9 Block diagram of the conventional differential detector for MDAPSK .. 19
2.10 Structure of the differential encoder for DUSTM ..................... 20
2.11 Structure of the differential encoder for STC-DAPSK ............... 23
2.12 STC-24APSK signal constellation ..................................... 24
3.1 Structure of the differential encoder for DAPSTM .................... 28
5.1 Comparison of simulation vs. numerical results of PEP, $N_T = 2$, $R = 1.5$ bits/(channel use) ($RP = 1$ bit/(channel use), $RA = 0.5$ bit/(channel use)), $B_fT = 0.001$, $\rho = 2.1$, $u_0 = 1$, $u_1 = 1$, $\theta_0 = 0$, and $\theta_1 = 1$ are valid ...................... 50
5.2 Comparison of simulation vs. numerical result of BER, $N_T = 2$, $R = 4$ bits/(channel use) ($RP = 3$ bits/(channel use), $RA = 1$ bit/(channel use)), $B_fT = 0.001$, $\rho = 1.4$, $u_0 = 1$, $u_1 = 15$, $\theta_0 = 0$, and $\theta_1 = 1$ are valid ...................... 51
5.3 BER vs. $10\log_{10}(E_b/N_0)$ for DAPSTM and DUSTM with $R = 3$ bits/(channel use) and $R = 4$ bits/(channel use). $N_T = 2$, $N_R = 1$ and $B_fT = 0.001$ are valid. DAPSTM is detected by SH-DD and ML-DD; DUSTM is only detected by ML-DD. ................................................................. 56

5.4 BER vs. $10\log_{10}(E_b/N_0)$ for MSD and DF-DD with (a) $R = 3$ bits/(channel use), $B_fT = 0.005$, (b) $R = 4$ bits/(channel use), $B_fT = 0.005$, (c) $R = 5$ bits/(channel use), $B_fT = 0.0035$, (d) $R = 6$ bits/(channel use), $B_fT = 0.0025$. $N_T = 2$, $N_R = 1$ is valid. ................................................................. 57

5.5 BER vs. $10\log_{10}(E_b/N_0)$ for DAPSTM with $N_T = 1$, $N_T = 2$, and $N_T = 3$. $N_R = 1$, $B_fT = 0.001$, $R = 4$ bits/(channel use) ($RA = 1$ bit/(channel use), $RP = 3$ bits/(channel use)), $\rho = 1.4$, $\tilde{u}_1 = 23$, $\tilde{P} = 3$ are used. ................................................................. 58

5.6 BER vs. $10\log_{10}(E_b/N_0)$ for DAPSTM and DUSTM with (a) $N_T = 2$, (b) $N_T = 3$. $N_R = 1$, $B_fT = 0.001$ are valid. ................................................................. 59

5.7 BER vs. $10\log_{10}(E_b/N_0)$ for DAPSTM vs. STC-DAPSK. $N_T = 2$, $N_R = 1$, and $B_fT = 0.001$ are valid. ................................................................. 61

5.8 BER vs. $10\log_{10}(E_b/N_0)$ for space diversity with $N_T = 2$ and time diversity with $N_B = 2$. (a) $R = 4$ bits/(channel use), (b) $R = 5$ bits/(channel use), (c) $R = 6$ bits/(channel use) are valid. ................................................................. 62

5.9 BER vs. $10\log_{10}(E_b/N_0)$ for time diversity with $N_B = 3$ and space diversity with $N_T = 3$. $R = 4$ bits/(channel use) is valid. ................................................................. 63
Acknowledgments

I would like to express my heartfelt gratitude to my supervisor, Dr. Robert Schober, for his continuous guidance and insight throughout this thesis. His enthusiasm and devotion have always inspired me during hard times. I would like to thank Dr. Victor C.M. Leung for his constant encouragement from the beginning of my graduate study at UBC. I would also like to thank all of my friends and fellow students, who have made my past years of study and work at UBC productive, cheerful and enjoyable. Last but not least, I would like to thank my family, for their never ending love and unconditional support.
Chapter 1

Introduction

1.1 Diversity

In wireless communication systems, fading [1] is a major obstacle to the efficient and reliable transmission of data. To mitigate the adverse effects of fading, it is desirable to provide the receiver with more than one replica of the transmitted signal to improve error-rate performance. This so-called diversity is based upon the observation that if the same information signal is received redundantly, over two or more independent fading channels ("diversity branches"), the probability that all the signals will fade simultaneously is reduced [2]. Space diversity and time diversity are two widely-applied diversity techniques.

Space diversity (antenna diversity) is an attractive technique for achieving a diversity advantage, since it does not incur any bandwidth expansion. The basic idea here is to employ two or more transmit antennas and/or receive antennas in order to receive uncorrelated signals. It is a practical technique for reducing the effects of fading in most scattering environments [2].

For time diversity, the same information-bearing signal is repeatedly transmitted in different time slots. If the time spacing equals or exceeds the coherence time of the channel, the multiple repetitions of the signal undergo nearly independent fading, thereby achieving
1.2 Motivation and Objectives

For single antenna wireless communication systems, differential encoding is employed for phase shift keying (PSK) when the phase shift introduced by the channel is unknown at the receiver [3]. In this so-called differential phase-shift keying (DPSK) scheme, the absolute phase of the received signal cannot be exploited, and the information is mapped to the phase difference between two successively transmitted symbols. To achieve a higher bandwidth efficiency, differential amplitude/phase-shift keying (DAPSK) modulation, see e.g. [4], is introduced. To make use of the diversity of multiple transmit antennas, differential unitary space-time modulation (DUSTM) [5], [6], and differential space-time block coding (DSTBC) [7], are proposed. These noncoherent transmission schemes avoid the need for channel estimation but often pay a price in performance. As a special case, diagonal signals are introduced for DUSTM, where at any given time only one antenna is active. The diagonal signal constellations form groups under matrix multiplication, allowing easy implementation at the transmitter while providing full diversity and full rate. DSTBC is more power efficient than DUSTM, but in general, it can only achieve full diversity and full rate with two transmit antennas. In [5]-[7], signal constellations consisting of unitary matrices are employed. DUSTM and DSTBC can be seen to be natural extensions of standard DPSK as used in single-antenna unknown-channel systems to multiple antenna transmission systems. To take advantage of the “amplitude”, space-time codes with DAPSK (STC-DAPSK) [8] are introduced to improve the bandwidth efficiency of DSTBC [7] for two transmit antenna systems.
1.3 Contributions

In this thesis, we propose a differential amplitude/phase space-time modulation (DAPSTM) scheme for multiple antenna wireless communication systems in flat Rayleigh fading channels. In our proposed DAPSTM scheme, diagonal signals are employed not only for phase but also for amplitude modulation. The total differential signal is the product of the amplitude signal and the phase signal. The DAPSTM signal keeps the group properties of DUSTM.

Two corresponding differential detection (DD) schemes, simple heuristic differential detection (SH-DD) and maximum likelihood DD (ML-DD) are derived for detection of DAPSTM. Two improved noncoherent receivers are also considered. The first one is the multiple-symbol detection (MSD) receiver. Because of the high computational complexity of MSD, a low-complexity decision-feedback DD (DF-DD) receiver is also derived. We also optimize the parameters in our signal design to achieve higher performance.

Performance analyses of the proposed DAPSTM for Rayleigh fading with uncorrelated diversity branches are presented. In particular, we calculate pairwise error probability (PEP) based on SH-DD. The analysis results agree well with the simulation results. Also, an approximation for the bit error rate (BER) is derived by the union bound.

Compared with DUSTM, the proposed constellations do not necessarily have a constant amplitude. In other words, we extend the differential phase modulation to combined differential amplitude and phase modulation in multiple-antenna systems. Since we take advantage of the amplitude information, DAPSTM can achieve a higher bandwidth efficiency than DUSTM.

DAPSTM is not as power efficient as the nondiagonal STC-DAPSK [8] when two trans-
mit antennas are employed. But DAPSTM can be easily extended to more than two transmit antennas, while keeping full diversity and the full transmission rate. For DAPSTM and fast fading, higher performance can be achieved if time diversity is exploited instead of space diversity. STC-DAPS K cannot be applied to exploit time diversity due to its nondiagonal structure.

Space and time diversity are closely related, since in both schemes a number of symbols are jointly modulated to obtain a multi-dimensional hypersymbol. All the features that work for space diversity also work for time diversity if signals are of a diagonal structure. But these two diversity techniques have different diversity characteristics when the channel varies slowly. Time diversity is less effective, because a very long interleaver is necessary to obtain sufficient diversity gain. On the other hand, space diversity requires multiple antennas at the transmitter, which increases implementation costs.

1.4 Thesis Outline

The outline of this thesis is as follows:

Chapter 2 reviews some basic concepts, such as the M-ary DPSK (MDPSK) and M-ary DAPS K (MDAPS K) modulation schemes, differential encoding, the employed channel models, DUSTM and STC-DAPS K, and conventional coherent and noncoherent receivers.

Chapter 3 discusses DAPSTM with ML-DD, MSD, and DF-DD receivers and introduces time diversity.

In chapter 4, we derive an exact, closed form expression for PEP and an approximation for BER.
Chapter 5 presents constellation designs and simulation results for space diversity and time diversity of DAPSTM, and compares the simulation results for space diversity with those for DUSTM and STC-DAPSK.

Chapter 6 summarizes the main contributions and conclusions of this thesis and gives recommendations for future work.

In this thesis, script English letters denote sets, bold upper case letters denote matrices, bold lower case letters denote vectors, and lower case letters denote scalars.
Chapter 2

Background and Related Work

2.1 System Model

2.1.1 Space Diversity System Model

Fig. 2.1 depicts the discrete-time equivalent baseband model of a single-user communication system with \( N_T \) transmit antennas and \( N_R \) receive antennas. The discrete message source first emits \( L \) binary bits \( b[i], b[i] \in \{0, 1\}, i = 0, 1, \ldots, L - 1 \), in each time interval \( N_T \). As usual, the bits at different time instances are assumed to be independent and identically distributed (i.i.d.). The mapper then modulates these \( L \) bits into a space-time matrix symbol \( V[k] \) of size \( N_T \times N_T \) according to some mapping rules. After that, the transmit matrix \( S[k] \) is obtained by differentially encoding \( V[k] \). The symbol \( s_m[N_T k + \kappa] \) in row \( \kappa \) and column \( m \) of the matrix \( S[k] \) is transmitted at time \( N_T k + \kappa, 0 \leq \kappa \leq N_T - 1 \), by the \( m \)th antenna, \( 0 \leq m \leq N_T - 1 \).

At the receiver, \( r_n[N_T k + \kappa] \) is received at time \( N_T k + \kappa, 0 \leq \kappa \leq N_T - 1 \) by the \( n \)th receive antenna, \( 0 \leq n \leq N_R - 1 \). The \( N_T N_R \) received signals constitute the matrix \( R[k] \) of size \( N_T \times N_R \), which is processed by a noncoherent receiver. This processing combines DD
2.1 System Model

Discrete Message Source \( h[i] \in \{0,1\} \) → Mapper → Differential Encoder → Channel → Noncoherent Receiver

Message Sink \( \hat{b}[i] \in \{0,1\} \) → DeMapper → Differential Detector → \( R[k] \)

Figure 2.1: Space diversity system model.

and demapping, and determines the estimates \( \hat{b}[i] \) of the transmitted binary symbols \( b[i] \).
Finally, the estimates \( \hat{b}[i] \) are passed to the message sink. A more detailed description of the different parts of Fig. 2.1 is given in the following sections.

2.1.2 Time Diversity System Model

The discrete-time equivalent complex-baseband model of the time diversity transmission system is sketched in Fig. 2.2. There are only one transmit antenna and \( N_R \) receive antennas in the system. To achieve \( N_B \) transmit diversity branches, an \( N_B \times N_f \) rectangular interleaver (cf. Fig. 2.3) is used to space the successive signals sufficiently far apart through the channel.
First, the $L \times N_I$ i.i.d. binary symbols $b[i], b[i] \in \{0,1\}, i = 0, 1, \ldots, N_I \times L - 1$, are mapped into $N_I$ diagonal matrices $V[k], 0 \leq k \leq N_I - 1$, of size $N_B \times N_B$. Each $V[k]$ is differentially encoded to obtain one transmission matrix $S[k]$. The main diagonal elements of $S[k], s_k[k], 0 \leq \kappa \leq N_B - 1, 0 \leq k \leq N_I - 1$ are written into the $k$th, $0 \leq k \leq N_I - 1$, row of the interleaver’s memory. After interleaving, a block of $N_B N_I$ symbols is read out from the interleaver, column by column, to form $N_I$ vectors $b[k]$ of size $N_B$. The $\kappa$th element of the vector $b[k], b_k[k], 0 \leq \kappa \leq N_B - 1, 0 \leq k \leq N_I - 1$, is transmitted at time $N_B k + \kappa, 0 \leq \kappa \leq N_B - 1, 0 \leq k \leq N_I - 1$. $b_k[k] = s_k[k'], 0 \leq \kappa' \leq N_B - 1, 0 \leq k' \leq N_I - 1$ and $N_B k + \kappa = \kappa' N_I + k'$. With this arrangement, the successive symbols of $S[k]$ are spaced $N_B$ symbols apart in $b[k]$, so that they encounter independent fading.
Each of the $N_R$ receive antennas collects the $N_I \times N_B$ received signals and feeds them in its own deinterleaver. Each deinterleaver performs the reverse operation on each block of $N_I \times N_B$ received symbols to recover their original sequence order and reads out $N_B$ symbols each time. At the output of the corresponding deinterleaver of the $n$th, $0 \leq n < N_R - 1$, receive antenna, the received signal at time $N_Bk + \kappa$, $0 \leq \kappa < N_B - 1$, $0 \leq k < N_I - 1$, is $r_n[N_Bk + \kappa]$,

$$r_n[N_Bk + \kappa] = s_\kappa[k]h_n[N_Ik + k] + n_n[N_Ik + k],$$  \hspace{1cm} (2.1)

where $0 \leq \kappa \leq N_B - 1$, $0 \leq k \leq N_I - 1$. The received $N_B \times N_R$ matrix $R[k]$, $0 \leq k \leq N_I - 1$, with elements $r_n[N_Bk + \kappa]$, $0 \leq \kappa \leq N_B - 1$, $0 \leq k \leq N_I - 1$, is processed by a noncoherent receiver to determine the estimates $\hat{v}[i]$ of the transmitted binary symbols $v[i]$, similar to space diversity. In particular, for given $k$ and $n$, the fading gains $h_n[N_Ik + k]$, $0 \leq \kappa \leq N_B - 1$, should be approximately uncorrelated to ensure full diversity for the jointly modulated symbols $s_\kappa[k]$, $0 \leq \kappa \leq N_B - 1$.

All concepts proposed for the space diversity system using diagonal signals can also be applied directly to the time diversity system. Therefore, in the following, we concentrate on the space diversity aspect, and discuss time diversity only when necessary. Note that in space diversity $N_T$ diversity branches are achieved by $N_T$ transmit antennas, and in time
diversity, where just one transmit antenna exists, $N_B$ diversity branches are achieved by the interleaver.

2.2 Channel Model

2.2.1 Flat Rayleigh Fading Channel Model

This thesis considers a stationary, slowly time-varying, frequency non-selective, Rayleigh fading channel along with additive white Gaussian noise (AWGN) distortion. We assume that the fading processes of the diversity branches are statistically independent. As usual, channel state and carrier phase are expected to be constant over a block of at least $N_T$ symbol periods, and perfect symbol synchronization is presumed.

The received signal at each receive antenna is a noisy superposition of independent, Rayleigh faded, transmitted signals. At time $N_Tk + \kappa$, $0 \leq \kappa \leq N_T - 1$, the equivalent complex-baseband received signal $r_n[N_Tk + \kappa]$ at the $n$th, $0 \leq n \leq N_R - 1$, receive antenna can be modelled as

$$r_n[N_Tk + \kappa] = \sum_{m=0}^{N_T-1} s_m[N_Tk + \kappa] h_{mn}[N_Tk + \kappa] + n_n[N_Tk + \kappa],$$

(2.2)

where the gain $h_{mn}[N_Tk + \kappa]$ models the zero mean complex flat fading channel from the $m$th transmit antenna to the $n$th receive antenna, and $n_n[N_Tk + \kappa]$ is an independent zero-mean complex white Gaussian noise process with two sided power spectral density $\sigma_n^2$.

All the fading processes are assumed to have identical statistical properties and are correlated in time. The fading autocorrelation function (ACF) $\varphi_{hh}[\cdot]$ is [9]
\[ \varphi_{hh}[\lambda] = \mathcal{E}\left\{ h_{mn}^*[N_T k + \kappa] h_{mn}[N_T k + \kappa + \lambda] \right\} \]
\[ = \sigma_f^2 \cdot J_0\left(2\pi B_f T \lambda\right), \]  

where \( \mathcal{E}\{ \cdot \} \) and \( (\cdot)^* \) denote expectation and complex conjugation, respectively, and the following definitions are used.

- \( \sigma_f^2 \) refers to the variance of the fading processing,

\[ \sigma_f^2 \triangleq \mathcal{E}\left\{ |h_{mn}[N_T k + \kappa]|^2 \right\}. \]  

For a simple representation, we can normalize \( h_{mn}[N_T k + \kappa] \) properly to ensure that the variance of the received noise-free signal is equal to 1. Since the transmitted symbols satisfy

\[ \mathcal{E}\left\{ \sum_{m=0}^{N_T-1} |s_m[N_T k + \kappa]|^2 \right\} = 1, \]  

we get, \( \sigma_f^2 = 1. \)

- \( J_0(\cdot) \) refers to the zeroth order Bessel function of the first kind.

\[ J_0(x) \triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ix\sin(\eta)} d\eta, \quad x \in \mathbb{R}. \]  

- \( B_f \) refers to one-sided bandwidth of the underlying continuous-time fading processes.

- \( T \) (scalar) refers to the symbol duration.

The white Gaussian noise processes of different receive antennas have equal variance

\[ \sigma_n^2 = \mathcal{E}\left\{ |n_n[N_T k + \kappa]|^2 \right\}. \]
2.2 Channel Model

The mean SNR per receive antenna is \( \text{SNR} = \frac{\sigma_f^2}{\sigma_n^2} = \frac{1}{\sigma_n^2} \).

Since the path gains are assumed to be constant during the transmission of one symbol matrix \( V[k] \) and vary from one block to the next, \( h_{mn}[N_Tk + \kappa] \) can be represented as \( h_{mn}[N_Tk] \). Hence, the \( r_n[N_Tk + \kappa] \) can also be obtained by

\[
grn[N_Tk + \kappa] = \sum_{m=0}^{N_T-1} s_m[N_Tk + \kappa] h_{mn}[N_Tk] + n_n[N_Tk + \kappa]. \tag{2.8}
\]

2.2.2 MIMO Channel Model

Fig. 2.4 represents the basic layout of a multiple-input multiple-output (MIMO) channel model with \( N_T \) transmit antennas and \( N_R \) receive antennas. The channel matrix \( H[k] \) of size \( N_T \times N_R \) consists of the fading gains for the different diversity branch pairs,
where \( h_{mn}[N_T k] \), \( 0 \leq m \leq N_T - 1 \), \( 0 \leq n \leq N_R - 1 \), is the complex fading gain between the \( m \)th transmit antenna and the \( n \)th receive antenna in \( k \)th \( N_T \) period.

The output is related to the input and the channel parameters by the following equation:

\[
R[k] = S[k]H[k] + N[k],
\]

(2.10)

where

- \( R[k] \) is an \( N_T \times N_R \) received signal matrix at the \( k \)th block with \((\kappa, n)\)th entry \( r_n[N_T k + \kappa] \) representing the received data at the \( n \)th, \( 0 \leq n \leq N_R - 1 \), receive antenna at time \( N_T k + \kappa \), \( 0 \leq \kappa \leq N_T - 1 \),

\[
R[k] = \begin{bmatrix}
    r_0[N_T k] & r_1[N_T k] & \cdots & r_{N_R-1}[N_T k] \\
    r_0[N_T k+1] & r_1[N_T k+1] & \cdots & r_{N_R-1}[N_T k+1] \\
    \vdots & \vdots & \ddots & \vdots \\
    r_0[N_T k+N_T-1] & r_1[N_T k+N_T-1] & \cdots & r_{N_R-1}[N_T k+N_T-1]
\end{bmatrix}
\]

(2.11)

- \( S[k] \) is an \( N_T \times N_T \) transmitted signal matrix at the \( k \)th block with \((\kappa, m)\)th entry \( s_m[N_T k + \kappa] \) representing the transmitted symbols from \( m \)th, \( 0 \leq m \leq N_T - 1 \), transmit antenna at time \( N_T k + \kappa \),

\[
S[k] = \begin{bmatrix}
    s_0[N_T k] & s_1[N_T k] & \cdots & s_{N_T-1}[N_T k] \\
    s_0[N_T k+1] & s_1[N_T k+1] & \cdots & s_{N_T-1}[N_T k+1] \\
    \vdots & \vdots & \ddots & \vdots \\
    s_0[N_T k+N_T-1] & s_1[N_T k+N_T-1] & \cdots & s_{N_T-1}[N_T k+N_T-1]
\end{bmatrix}
\]

(2.12)
2.3 Differential Modulation and Detection

Coherent modulation schemes require CSI at the receiver. However, in some situations, e.g., fast fading, reliable channel estimation is difficult. In these cases, differential modulation is usually used in wireless communication systems.

In this section, first, two differential modulation schemes with one transmit antenna and one receive antenna, namely, MDPSK and MDAPSK, are reviewed. Then, DSTM and STC-DAPSK are presented for more transmit antennas.

2.3.1 MDPSK

![Figure 2.5: Structure of the differential encoder for MDPSK modulation.](image)

MDPSK, the differential form of M-ary PSK (MPSK), is a very important modulation
scheme for single antenna wireless communication systems. The structure of the differential encoder for MDPSK is depicted in Fig. 2.5. In this scheme, notation \( v[k] \) and \( s[k] \) are used instead of \( V[k] \) and \( S[k] \), which only have one element in them, respectively. First, the information bits are mapped into the phase difference symbol \( v[k] \). The transmitted symbol at time \( k \) is generated from \( v[k] \) via differential encoding:

\[
s[k] = v[k] \cdot s[k - 1].
\] (2.14)

The phase difference symbols \( v[k] \) are drawn from an \( M \)-ary alphabet,

\[
\mathcal{A}_t(M) \triangleq \left\{ e^{j\frac{2\pi}{M} l} \mid l \in \{0, 1, \ldots, M-1\} \right\}.
\] (2.15)

In conventional differential detectors for MDPSK, cf., Fig. 2.6, the estimated transmitted MDPSK symbol \( \hat{v}[k] \) is determined from the decision variable [3]

\[
d[k] \triangleq r[k] \cdot r^*[k - 1].
\] (2.16)

![Figure 2.6: Block diagram of the conventional differential detector for MDPSK.](image)

For this, the complex plane is divided into \( M \) sectors, corresponding to the \( M \) possible values of \( v[k] \). The sector into which \( d[k] \) falls determines the value of \( \hat{v}[k] \). The decision also can be expressed as
\[ \hat{v}[k] = \arg\max_{\hat{v}[k]} \left\{ \Re\{ r[k] r^*[k-1] \hat{v}^*[k] \} \right\}, \]

where \( \Re\{ \cdot \} \) denotes the real part of a complex number and \( \hat{v}[k] \in \mathcal{A}_v(M) \) denotes a trial symbol.

### 2.3.2 MDAPSK

Another more bandwidth-efficient modulation scheme for single antenna wireless communication systems is MDAPSK modulation [4]. MDAPSK, where both amplitude and phase are differentially encoded, may be viewed as a combination of differential amplitude-shift keying (DASK) and DPSK, which are independent of each other (cf. Fig. 2.7).

Figure 2.7: Structure of the differential encoder for MDAPSK modulation.

The MDAPSK symbol is given by \( s[k] = a[k] p[k] \), where the absolute amplitude symbol \( a[k] \) is a positive real number, and the absolute phase symbol has magnitude one. For convenience, an appropriate normalization is employed to ensure

\[ \mathcal{E}\left\{ |s[k]|^2 \right\} = \mathcal{E}\left\{ a^2[k] \right\} = 1, \]
The absolute amplitude at time $k$, which is taken from the alphabet

$$\mathcal{A}_a \triangleq \left\{ a_0, a_1, \cdots, a_{Y-1} \mid Y \in \mathbb{N}, Y \geq 2 \right\},$$

(2.19)

where $\mathbb{N}$ is the natural number set. $a[k]$ is differentially encoded by

$$a[k] = v_a[k] \cdot a[k-1],$$

(2.20)

where $v_a[k]$ is the amplitude difference symbols. The alphabet $\mathcal{A}_{v_a}$ of the amplitude difference symbols $v_a[k]$ depends on $\mathcal{A}_a$.

Similar to MDPSK, the absolute phase at time $k$ is also obtained from differentially encoding the phase difference symbols $v_p[k]$,

$$p[k] = v_p[k] \cdot p[k-1].$$

(2.21)

The absolute phase symbol is drawn from the alphabet,

$$\mathcal{A}_p(M/Y) \triangleq \left\{ e^{j\frac{2\pi l}{M} \theta} \mid l \in \{0, 1, \cdots, M/Y - 1\} \right\}.$$  

(2.22)

And the alphabet of the phase difference symbols $\mathcal{A}_{v_p}(M/Y)$ is identical to $\mathcal{A}_p(M/Y)$.

From another point of view, we can see the MDAPSK symbol $v[k]$ as the product of the amplitude difference symbol $v_a[k]$ and the phase difference symbol $v_p[k]$,

$$v[k] = v_a[k]v_p[k].$$

(2.23)

16DAPSK ($Y = 2$) modulation, which is also referred as 16-star quadrature-amplitude modulation (QAM) [4], is a popular example of MDAPSK modulation. The single constellation for 16DAPSK is shown in Fig. 2.8. It consists of 2DASK with absolute amplitude...
2.3 Differential Modulation and Detection

\{r_L, r_H\} and 8DPSK with

\[ v_p[k] \in \left\{ e^{j \frac{2\pi}{2^L}} | l \in \{0, 1, \ldots, 7\} \right\}. \quad (2.24) \]

The amplitude ratio \( \rho \triangleq r_H/r_L > 1 \) is the most significant parameter of 16DAPSK. The absolute amplitude can be expressed by \( \rho \) as

\[ r_L = \sqrt{\frac{2}{1 + \rho^2}}, \quad (2.25) \]

\[ r_H = \sqrt{\frac{2\rho^2}{1 + \rho^2}}, \quad (2.26) \]

where the normalization

\[ \mathcal{E}\left\{ a^2[k] \right\} = \frac{1}{2} (r_L^2 + r_H^2) = 1 \quad (2.27) \]

has been taken into account. One out of four bits is carried by 2DAPSK. This one bit is mapped to amplitude difference symbol \( v_a[k] \). The bit "0" causes no amplitude change.
(a[k] = a[k − 1]) and the bit “1” causes an amplitude change (a[k] ≠ a[k − 1]). The remaining 3 bits are carried by 8DPSK. They are Gray mapped onto the phase difference symbol $v_p[k]$, according to this rule: (000)$→e^{j\pi/4}$, (001)$→e^{j\pi/2}$, (010)$→e^{j3\pi/4}$, (101)$→e^{j\pi}$, (111)$→e^{-j3\pi/4}$, (100)$→e^{-j\pi/2}$, (100)$→e^{-j\pi/4}$.

In the first stage of the receiver, differential decoding involves as shown in Fig. 2.9. The decision variable

$$d[k] \triangleq \frac{r[k]}{r[k - 1]} \quad (2.28)$$

is generated. Amplitude and phase difference symbols are estimated separately in the receiver, since the amplitude and phase modulation are independent.

![Figure 2.9: Block diagram of the conventional differential detector for MDAPSK.](image)

In 16DAPSK, the phase decision rule is the same as that employed for conventional DD of 8DPSK. For the amplitude decision of 16DAPSK, the arithmetic mean is used for decision thresholds [4]:

$$\gamma_0 = \frac{1}{2(1 + 1/\rho)}, \quad (2.29)$$
$$\gamma_1 = \frac{1}{2(1 + \rho)}. \quad (2.30)$$

The above decision thresholds are not optimum. For a given channel and a given ring ratio, we can optimize $\gamma_0$ and $\gamma_1$ together with ring ratio $\rho$ by numerical evaluation of the
analytical expression for the BER provided in [10]. So, the estimated amplitude difference symbol becomes

\[
\Delta \hat{a}[k] = \begin{cases} 
\rho, & |d[k]| \geq \gamma_1 \\
1, & \gamma_0 \leq |d[k]| < \gamma_1 \\
1/\rho, & |d[k]| < \gamma_0
\end{cases}
\]  

(2.31)

2.3.3 Differential Unitary Space-Time Modulation

In the case of multiple transmit antennas, DUSTM is proposed by Hochwald et al. [5] and Hughes [6], independently. In DUSTM, cf. Fig. 2.10, the transmitted signal matrix at each time block is the product of the previous transmitted matrix and the current unitary data matrix:

\[
S[k] = V[k]S[k-1].
\]  

(2.32)

The transmitted matrix symbol

\[
S[k] = \text{diag}\{s_0[N_T k], s_1[N_T k + 1], \ldots, s_{N_T - 1}[N_T k + N_T - 1]\},
\]  

(2.33)

where \(\text{diag}\{ \cdot \}\) denotes a diagonal matrix. \(S[0] = I_{N_T}\), where \(I_{N_T}\) stands for \(N_T \times N_T\) identity matrix.
2.3 Differential Modulation and Detection

The constellations for unitary space-time modulated signals proposed in [5] and [6] form groups under matrix multiplication, thus simplifying the differential encoding process. In particular, the diagonal cyclic group constellations proposed in [5] are systematically designed for an arbitrary number of transmit antennas while providing full rate. The unitary data matrix

$$V[k] = \text{diag}\{e^{j\frac{2\pi}{L}u_0}, e^{j\frac{2\pi}{L}u_1}, \cdots, e^{j\frac{2\pi}{L}u_m}, \cdots, e^{j\frac{2\pi}{L}u_{N_T}}\}, 1 \leq m \leq N_T - 1, \quad (2.34)$$

where $0 \leq l[k] \leq L - 1$, $L = 2^{N_TR}$ with data rate $R$. For diagonal signals only one antenna is active at any time. The transmitted phase-shifted symbols $V[k]$ are potentially different for each antenna due to different values of $u_m$. The coefficients $u_0 = 1$, $u_m$, $1 \leq m \leq N_T - 1$ can be obtained by an exhaustive computer search to maximize the diversity product

$$\zeta = \min_{l[k] \in \{1,3,\cdots,L/2\}} \prod_{m=1}^{N_T} \sin\left(\pi u_m l[k]/L\right)^{1/N_T}, \quad (2.35)$$

where, $u_m \in \{1,3,\cdots,L/2\}$ and $l[k] \in \{1,3,\cdots,L/2 - 1\}$. The search results of $u_m$, $1 \leq m \leq N_T - 1$, of $N_T = 2$ and $N_T = 3$ are given in Tables 2.1 and 2.2, respectively. In order to yield a low BER, the distance between the two nearest neighbors $\bar{l}$, is determined using the method in [11] to generate a Gray mapping for a given signal constellation. The

Table 2.1: DUSTM parameters for $N_T = 2$.

<table>
<thead>
<tr>
<th>R=3</th>
<th>R=4</th>
<th>R=5</th>
<th>R=6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_1$</td>
<td>$\bar{l}$</td>
<td>$u_1$</td>
<td>$\bar{l}$</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>75 *</td>
<td>1 *</td>
</tr>
<tr>
<td>27 *</td>
<td>19 *</td>
<td>99</td>
<td>75</td>
</tr>
</tbody>
</table>
optimum value of $\bar{l}$ is also shown in Tables 2.1 and 2.2. The values with "*" are actually used for the simulations in Chapter 5.

Table 2.2: DUSTM parameter for $N_T = 3$.

<table>
<thead>
<tr>
<th></th>
<th>$R=3$</th>
<th></th>
<th></th>
<th>$R=4$</th>
<th></th>
<th></th>
<th>$R=5$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_1$</td>
<td>$u_2$</td>
<td>$\bar{l}$</td>
<td>$u_1$</td>
<td>$u_2$</td>
<td>$\bar{l}$</td>
<td>$u_1$</td>
<td>$u_2$</td>
<td>$\bar{l}$</td>
<td></td>
</tr>
<tr>
<td>119</td>
<td>221</td>
<td>185</td>
<td>1735</td>
<td>889</td>
<td>14</td>
<td>6921*</td>
<td>11375*</td>
<td>3*</td>
<td></td>
</tr>
<tr>
<td>139</td>
<td>157</td>
<td>75</td>
<td>1737*</td>
<td>961*</td>
<td>158*</td>
<td>6921*</td>
<td>11375*</td>
<td>3*</td>
<td></td>
</tr>
<tr>
<td>185</td>
<td>75</td>
<td>1</td>
<td>1737*</td>
<td>961*</td>
<td>158*</td>
<td>6921*</td>
<td>11375*</td>
<td>3*</td>
<td></td>
</tr>
</tbody>
</table>

In DD at the receiver, the channel $H[k]$ is unknown and assumed to be approximately constant during two adjacent blocks. The decision is based on the observation of $R[k - 1]$ and $R[k]$,

$$\bar{R}[k] \triangleq \begin{bmatrix} R[k - 1] \\ R[k] \end{bmatrix}.$$ (2.36)

In this case, $\bar{R}[k]$ can be expressed as

$$\bar{R}[k] = \bar{S}[k]H[k] + \bar{N}[k],$$ (2.37)

where

$$\bar{S}[k] \triangleq \begin{bmatrix} S[k - 1] \\ S[k] \end{bmatrix},$$ (2.38)

$$\bar{N}[k] \triangleq \begin{bmatrix} N[k - 1] \\ N[k] \end{bmatrix}.$$ (2.39)
2.3 Differential Modulation and Detection

It is shown in [11] that the probability density function (pdf) of $\mathbf{R}[k]$ conditioned on $\mathbf{S}[k]$, $p(\mathbf{R}[k]|\mathbf{S}[k])$, is independent of $\mathbf{S}[k-1]$. By maximizing $p(\mathbf{R}[k]|\mathbf{S}[k])$, the decoding metric of $\mathbf{V}[k]$ is

$$
\hat{\mathbf{V}}[k] = \arg\min_{\mathbf{V}[k]} \| \mathbf{R}[k] - \hat{\mathbf{V}}[k] \mathbf{R}[k-1] \|^2,
$$  \hspace{1cm} (2.40)

where $\| \cdot \|^2$ denotes the Frobenius norm. $\hat{\mathbf{V}}[k]$ is the estimated symbol, and $\mathbf{V}[k]$ is the trial symbol. The decoding metric can be either of quadratic form, or correlation form, or of minimum distance form, which are all equivalent.

2.3.4 Space-Time Codes with DAPSK

In DSTM and DSTBC, all the codes have the same norm. This implies that these space-time signal constellations are extensions of PSK, which is of limited bandwidth efficiency. To achieve a high bandwidth efficiency, Xia [8] proposes STC-DAPSK for Alamouti’s orthogonal space-time code (OSTC) [12] with two transmit antennas, cf. Fig. 2.11.

![Figure 2.11: Structure of the differential encoder for STC-DAPSK.](image-url)
For STC-($M_1 + M_2$)DAPSK, let $R_1 = \log_2 M_1$, $R_2 = \log_2 M_2$. $R_1 + R_2$ bits are mapped to one Alamouti codeword matrix [12],

\[
V_p[k] = \begin{bmatrix}
v_{p_0}[2k] & v_{p_1}[2k] \\
v_{p_1}[2k] & -v_{p_0}[2k]
\end{bmatrix}
\]  \hspace{1cm} (2.41)

where $v_{p_i}[2k] \in A_{M_i}$, $i \in \{0, 1\}$, is the $M_i$PSK signal,

\[
A_{M_i} = \left\{ e^{j \frac{2\pi l}{M_i}} | l \in \{0, 1, 2, \ldots, M_i - 1\} \right\}
\]  \hspace{1cm} (2.42)

The $(R_1 + R_2 + 1)$th bit is introduced to decide whether there is an amplitude change in the two successive transmit symbols. The bandwidth efficiency is $(R_1 + R_2 + 1)/2$ bits/(channel use).

![Figure 2.12: STC-24APSK signal constellation.](image)

Specifically, in STC-24APSK, there are two independent 8PSK and 16PSK constellations and one 2ASK, with $\rho = r_H/r_L$, as shown in Fig. 2.12. $v_{p_0}[2k]$ and $v_{p_1}[2k]$ are picked up from the sets $A_8$ and $A_{16}$, respectively. In 2ASK, the $(4 + 3 + 1)$th bit is mapped to $v_a[k] = \rho^{\Delta a[k]}$, where $\Delta a[k] \in \{0, \pm 1\}$. By using differential encoding, cf. Fig. 2.10, we
have

\[ S[k] = a[k]P[k], \quad (2.43) \]
\[ a[k] = a[k-1] \cdot v_{a[k]}, \quad (2.44) \]
\[ P[k] = P[k-1] \cdot V_p[k], \quad (2.45) \]

where \( P[0] = I_{N_T}, a[0] = r_L \). In each two time steps, we first decide upon the amplitude ring, and then upon the codeword matrix \( P[k] \). In particular, if the \( r_H \) ring is selected, then the signal constellation is drawn from 16PSK. Otherwise, an 8PSK signal is used. In the next step, the amplitude may change and a different signal alphabet is used. On average, there are a total of \( 4 + 3 + 1 \) bits carried in 2 time slots. Hence, the bandwidth efficiency is \( 4 \text{ bits/(channel use)} \).

The single receive antenna case is considered in [8]. Here, the received signals in the \( 2N_T \) time intervals form the following \( 2 \times 1 \) vector:

\[ r[k] \triangleq \begin{bmatrix} r_{0[2k]} \\ r_{0[2k+1]} \end{bmatrix}. \quad (2.46) \]

Thus, we can get the receive signal

\[ r[k] = S[k]h[k] + n[k], \quad (2.47) \]

where

\[ h[k] \triangleq \begin{bmatrix} h_{0[2k]} \\ h_{1[2k]} \end{bmatrix}, \quad (2.48) \]
\[ n[k] \triangleq \begin{bmatrix} n_{0[2k]} \\ n_{0[2k+1]} \end{bmatrix}. \quad (2.49) \]
There are two steps in the differential decoding. The first step is to detect the value of the \((R_1 + R_2 + 1)\)th bit, “0” or “1”, by

\[
\Delta \hat{a}[k] = \arg \min_{\Delta \hat{a}[k] \in \{0, \pm 1\}} \left\| r[k] - \rho^{\Delta \hat{a}[k]} r[k - 1] \right\|.
\] (2.50)

If \(\Delta \hat{a}[k] = 0\), the \((R_1 + R_2 + 1)\)th bit is 0. Otherwise, it is equal to 1. After deciding on the bit to be carried by 2ASK, symbols \(v_{p_0}[2k]\), \(v_{p_1}[2k]\) are detected by

\[
\hat{v}_{p_0}[2k] = \arg \max_{\hat{v}_{p_0}[2k] \in A_{p_0}} \Re \left\{ \left( r_0^*[2k] r_0[2(k - 1)] + r_1[2k] r_1^*[2(k - 1)] \right) \hat{v}_{p_0}[2k] \right\}, \quad (2.51)
\]

\[
\hat{v}_{p_1}[2k] = \arg \max_{\hat{v}_{p_1}[2k] \in A_{p_1}} \Re \left\{ \left( r_0^*[2k] r_1[2(k - 1)] + r_1[2k] r_0^*[2(k - 1)] \right) \hat{v}_{p_1}[2k] \right\}. \quad (2.52)
\]

Since Alamouti’s OSTC is orthogonal, the decoding complexity of STC-DAPSK remains similar to that of a single antenna system.
Chapter 3

Differential Amplitude and Phase Space-Time Modulation

In this chapter, we propose a DAPSTM scheme based on diagonal signals, originally proposed for DUSTM [5], [6]. For DAPSTM, two differential detectors, SH-DD and ML-DD, are investigated. An improved noncoherent receiver, MSD for DAPSTM is also introduced. To reduce the computation complexity of the MSD, the decoding metric is simplified to DF-DD. In SH-DD, we assume the channel is unchanged in the two successive symbols. In ML-DD, MSD and DF-DD, the channel is assumed to be constant in one symbol interval, but the second order statistics of fading and noise are yet to be known.

3.1 DAPSTM

Fig. 3.1 shows the block diagram of the proposed DAPSTM transmission scheme. We consider a transmission scheme comprising $N_T$ transmit antennas and $N_R$ receive antennas. The transmitted signals are organized in the square matrix $S[k]$, $k = 0, 1, \cdots$, with element $s_m[N_Tk + \kappa]$ in row $\kappa$, $0 \leq \kappa \leq N_T - 1$, and column $m$, $0 \leq m \leq N_T - 1$. The transmitted signal matrix at $k^{th}$ time block
3.1 DAPSTM

\[ S[k] = A[k]P[k], \]  

(3.1)

consists of an amplitude matrix symbol \( A[k] \) and an unitary phase matrix symbol \( P[k] \).

The data rate \( R = R_A + R_P \), where \( R_A \) and \( R_P \) are the rate of the bits mapped to phase and amplitude symbols, respectively.

\[
V_P[k] \quad N_T T \quad V_{A[k]} \quad \times \quad S[k] \\
\quad \quad N_T T \quad V_p[k] \quad \times \quad P[k]
\]

Figure 3.1: Structure of the differential encoder for DAPSTM.

The phase matrix symbol \( P[k] \) is obtained through differential encoding from \( P[k - 1] \) and \( V_P[k] \):

\[
P[k] = V_P[k]P[k - 1].
\]  

(3.2)

For \( V_P[k] \), we restrict the phase signals to constellations whose elements form a group under matrix multiplication; that is, the possible values for \( V_P[k] \) and \( P[k] \) belong to a finite set with \( L_P = 2^{N_T R_P} \) elements [13]. In particular, we are interested in the diagonal constellations where the phase difference matrix symbols \( V_P[k] = V_{\Delta l_P[k]} \) are drawn from the set [5] [6],

\[
A = \left\{ V_{\Delta l_P} = \text{diag} \left\{ e^{j \frac{2\pi}{L_P} u_0}, e^{j \frac{2\pi}{L_P} u_1}, \ldots, e^{j \frac{2\pi}{L_P} u_{N_T-1}} \right\} \right\}^{\Delta l_P} \mid \Delta l_p \in \{0, 1 \ldots L_P-1\} \right\}.  
\]  

(3.3)
where \( j = \sqrt{-1} \) denotes the imaginary unit, the \( \Delta l p[k] \) refers to the information-carrying phase difference symbols and the coefficient \( u_m, 0 \leq u_m \leq LP - 1, 0 \leq m \leq N_T - 1 \) is a design parameter. Note that the matrix \( V_P[k] \) and \( P[k] \) are uniquely associated with the symbol \( \Delta l p[k] \) and \( lp[k] \) by

\[
V_P[k] = V_P^{\Delta l p[k]}, \tag{3.4}
\]

\[
P[k] = V_P^{lp[k]}, \tag{3.5}
\]

respectively, where \( V_P = V_P[1], lp[k] \) represents the absolute phase symbols. Through differential encoding, we have

\[
P[k] = V_P[k]P[k - 1]
= V_P^{\Delta l p[k]+lp[k-1]}. \tag{3.6}
\]

The amplitude matrix symbol \( A[k] \) is obtained by differentially encoding from \( V_A[k] \) and \( A[k - 1] \),

\[
A[k] = V_A[k]A[k - 1]. \tag{3.7}
\]

The possible values for \( V_A[k] \) belong to a finite set \( \mathcal{A}_V \) with \( LA = 2^{N_T R_A} \) elements [13]. The alphabet \( \mathcal{A}_A \) of the absolute amplitude matrix symbols \( A[k] \) depends on \( \mathcal{A}_V \). \( V_A[k] \) and \( A[k] \) are also diagonal matrices, and given by

\[
V_A[k] = \rho^{\Delta l a[k]} I_{N_T}, \tag{3.8}
\]

\[
A[k] = \text{diag}\left\{ \rho^{(\Delta l a[k]+la[k-1]+\theta_0)\text{mod} LA}, \rho^{(\Delta l a[k]+la[k-1]+\theta_1)\text{mod} LA}, \ldots, \right. \\
\rho^{(\Delta l a[k]+la[k-1]+\theta_{N_T-1})\text{mod} LA} \left. \} A[0]. \right. \tag{3.9}
\]
In the above equations, the following definition is used. \( A[0] = \varrho I_{N_T} \), with scalar \( \varrho > 0 \).

The information-carrying amplitude difference symbols \( \Delta a[k] \) belong to \( \{0, 1, \cdots, LA-1\} \).

The \( \lambda a[k] \) represents the absolute amplitude symbols. The design parameter \( \theta_m, 0 \leq \theta_m \leq N_T - 1 \), determines the initial amplitude value for each antenna. Another important design parameter \( \rho > 0 \) stands for the magnitude ratio between two amplitudes of the neighboring symbol matrices.

Since the \( A[k] \) and \( P[k] \) are diagonal matrices, we can define them as

\[
P[k] = \begin{bmatrix}
p_0[N_T k] & 0 & \cdots & 0 \\
0 & p_1[N_T k + 1] & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & p_{N_T-1}[N_T k + N_T - 1]
\end{bmatrix},
\]

(3.10)

\[
A[k] = \begin{bmatrix}
a_0[N_T k] & 0 & \cdots & 0 \\
0 & a_1[N_T k + 1] & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & a_{N_T-1}[N_T k + N_T - 1]
\end{bmatrix},
\]

(3.11)

Using Eq. (3.1), the transmitted signal matrix \( S[k] \) also becomes a diagonal matrix:

\[
S[k] = \begin{bmatrix}
s_0[N_T k] & 0 & \cdots & 0 \\
0 & s_1[N_T k + 1] & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & s_{N_T-1}[N_T k + N_T - 1]
\end{bmatrix},
\]

(3.12)

where

\[
s_m[N_T k + m] = e^{j \varphi_m \Delta p[k] + ip[k-1]} \cdot \varrho^0(\Delta la[k] + \lambda a[k-1] + \theta_m) \mod LA,
\]

(3.13)

with \( 0 \leq m \leq N_T - 1 \). \( s_m[N_T k + m] \) is transmitted by the \( m \)th antenna, \( 0 \leq m \leq N_T - 1 \) at time \( N_T k + m \). We transmit over \( N_T \) antennas in \( N_T \) time steps to achieve full rate. The
transmitted symbols are normalized to

$$\mathcal{E} \left\{ \sum_{m=0}^{N_T-1} |s_m[N_T k + m]|^2 \right\} = 1,$$

which means

$$\frac{q^2}{L_A} \left( (\rho^2)^0 + (\rho^2)^1 + (\rho^2)^0 + \cdots + (\rho^2)^{L_A-1} \right) = 1,$$

and hence we have,

$$\varrho = \frac{\sqrt{L_A}}{\sum_{m=0}^{L_A-1} (\rho^{2m})^{1/2}},$$

For given $N_T$ and data rate $R$, the design parameters $\rho$, $u_m$ and $\theta_m$ are optimized to achieve best performance, cf. Section 5.2.

### 3.2 Conventional Differential Detection

In conventional DD, two consecutively received signal matrices, $\mathbf{R}[k]$ and $\mathbf{R}[k-1]$, are stacked to form a matrix $\bar{\mathbf{R}}[k]$ with $2N_T$ rows:

$$\bar{\mathbf{R}}[k] = \begin{bmatrix} \mathbf{R}[k-1] \\ \mathbf{R}[k] \end{bmatrix}.$$  

By using Eqs. (2.9), (2.10), (2.13), (3.12), $\bar{\mathbf{R}}[k]$ can be expressed as

$$\bar{\mathbf{R}}[k] = \bar{\mathbf{S}}[k] \bar{\mathbf{H}}[k] + \bar{\mathbf{N}}[k],$$

with the definitions
3.2 Conventional Differential Detection

\[
\overline{S}[k] = \begin{bmatrix} S[k-1] & O_{NT} \\ O_{NT} & S[k] \end{bmatrix}, \tag{3.19}
\]

\[
\overline{H}[k] = \begin{bmatrix} H[k-1] \\ H[k] \end{bmatrix}, \tag{3.20}
\]

\[
\overline{N}[k] = \begin{bmatrix} N[k-1] \\ N[k] \end{bmatrix}. \tag{3.21}
\]

By substituting Eqs. (3.2), (3.7) into Eq. (3.1), we can get

\[
S[k] = V_A[k]A[k-1]V_P[k]P[k-1]. \tag{3.22}
\]

Since \( A[k-1] \) and \( V_P[k] \) commute, by defining \( V[k] = V_A[k]V_P[k] \), we can then write \( S[k] \) as

\[
S[k] = V_A[k]V_P[k]A[k-1]P[k-1],
\]

\[
= V[k]S[k-1], \tag{3.23}
\]

which is desirable for DD. In the following, we derive two conventional noncoherent decoding methods, namely SH-DD and ML-DD (in the sense of maximum likelihood over two consecutive received signal matrices).

### 3.2.1 Simple Heuristic Differential Detection

SH-DD is a suboptimum means of detecting DAPSTM, based on the minimum distance between the received signal matrix and the product of the current information conveying matrix \( V[k] \) and the previously received signal matrix. The estimated signal matrix is given by
3.2 Conventional Differential Detection

\[ \tilde{V}[k] = \arg\min_{\tilde{V}[k]} \left\| R[k] - \tilde{V}[k] R[k-1] \right\|^2. \] (3.24)

Trial matrix \( \tilde{V}[k] \) is an \( N_T \times N_T \) diagonal matrix,

\[
\tilde{V}[k] = \begin{bmatrix}
\tilde{v}_0[N_Tk] & 0 & \cdots & 0 \\
0 & \tilde{v}_1[N_Tk + 1] & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \tilde{v}_{N_T-1}[N_Tk + N_T - 1]
\end{bmatrix} \quad (3.25)
\]

where

\[ \tilde{v}_m[N_Tk + m] = e^{j \frac{2\pi m}{N_T} \Delta \hat{p}[k]} \cdot \rho \Delta \hat{a}[k], \] (3.26)

with \( 0 \leq m \leq N_T - 1 \). The decision rule can be written simply as

\[
(\Delta \hat{p}[k], \Delta \hat{a}[k]) = \arg \min_{\Delta \hat{p}[k], \Delta \hat{a}[k]} \left\{ \sum_{n=0}^{N_R-1} \sum_{m=0}^{N_T-1} \left| r_n[N_Tk + m] - \tilde{v}_m[N_Tk + m] r_n[N_T(k - 1) + m] \right|^2 \right\}, \quad (3.27)
\]

where \( \Delta \hat{a}[k] \) and \( \Delta \hat{p}[k] \) denote the estimated amplitude and phase difference symbols, respectively, while \( \Delta \hat{a}[k] \in \{0, 1, \cdots, LA - 1\} \) and \( \Delta \hat{p}[k] \in \{0, 1, \cdots, LP - 1\} \) present the corresponding trial symbols, respectively.

3.2.2 Maximum Likelihood Differential Detection

Since \( h_{mn}[k] \), \( 0 \leq m \leq N_T - 1 \), \( 0 \leq n \leq N_R - 1 \), and \( n_n[k] \), \( 0 \leq n \leq N_R - 1 \) are zero-mean complex Gaussian random processes, the pdf of \( \tilde{R}[k] \) conditioned on \( \tilde{S}[k] \) is given by [11]

\[
p\left( \tilde{R}[k] | \tilde{S}[k] \right) = \frac{\exp \left( - \text{tr} \left( \tilde{R}^H[k] C_{\tilde{R}}^{-1}[k] \tilde{R}[k] \right) \right)}{\left( \pi^{N_{NT}} \det \left( C_{\tilde{R}[k]} \right) \right)^{N_R}}. \quad (3.28)
\]
Conventional Differential Detection

Here $N = 2$ stands for the observation window size, $[\cdot]^H$ denotes Hermitian transposition, and $\det\{\cdot\}$ and $\text{tr}\{\cdot\}$ denote the determinant and the trace of a matrix, respectively. $C_R[k]$ denotes the conditional covariance matrix,

$$C_R[k] = \mathcal{E}\left\{ \mathbf{R}[k] \mathbf{R}^H[k] | \mathcal{S}[k] \right\}. \quad (3.29)$$

Because of the mutual independence of $\bar{H}[k]$ and $\bar{N}[k]$, and the uncorrelatedness of $n_n[kNT + m]$ in space and time, $C_R[k]$ can be rewritten as

$$C_R[k] = \mathcal{S}[k] \mathcal{E}\left\{ \bar{H}[k] \bar{H}^H[k] \right\} \mathcal{S}^H[k] + \mathcal{E}\left\{ \bar{N}[k] \bar{N}^H[k] \right\}$$

$$= \mathcal{S}[k] C_H \mathcal{S}^H[k] + N_R \sigma_n^2 I_{2NT}, \quad (3.30)$$

with

$$C_H = \mathcal{E}\left\{ \bar{H}[k] \bar{H}^H[k] \right\}. \quad (3.31)$$

Due to the spatial uncorrelatedness of $h_{mn}[\cdot]$, $C_H$ can be rewritten as

$$C_H = N_R \left( C_h \otimes I_{2NT} \right), \quad (3.32)$$

where $\otimes$ denotes the Kronecker product [14] and $C_h$ is defined as

$$C_h \triangleq \begin{bmatrix} \varphi_{hh}[0] & \varphi_{hh}[NT] \\ \varphi_{hh}[-NT] & \varphi_{hh}[0] \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \varphi_{hh}[NT] \\ \varphi_{hh}[NT] & 1 \end{bmatrix}. \quad (3.33)$$

$\varphi_{hh}[\cdot]$ is given by Eq. (2.3). By substituting $\mathcal{S}[k]$ with Eq. (3.1), and taking into account that $\mathcal{S}[k]$, $\bar{P}[k]$, and $\bar{A}[k]$ are diagonal matrices, we obtain
3.2 Conventional Differential Detection

\[ C_R[k] = \bar{P}[k]A[k]C_H A^H[k]\bar{P}^H[k] + N_R\sigma_n^2 I_{2NT} \]
\[ = \bar{P}[k]\left( A[k]C_H A^H[k] + N_R\sigma_n^2 I_{2NT}\right)\bar{P}^H[k] \]
\[ = \bar{P}[k]M[k]\bar{P}^H[k], \quad (3.34) \]

where

\[ \bar{P}[k] = \begin{bmatrix} P[k-1] & O_{NT} \\ O_{NT} & P[k] \end{bmatrix}, \quad (3.35) \]

\[ \bar{A}[k] = \begin{bmatrix} A[k-1] & O_{NT} \\ O_{NT} & A[k] \end{bmatrix}, \quad (3.36) \]

\[ M[k] = \bar{A}[k]C_H A^H[k] + N_R\sigma_n^2 I_{2NT} \]
\[ = N_R \begin{bmatrix} A^2[k-1] + \sigma^2 I_{NT} & \phi_{hh}[N_T]A[k]A[k-1] \\ \phi_{hh}[N_T]A[k]A[k-1] & A^2[k] + \sigma^2 I_{NT} \end{bmatrix} \quad (3.37) \]

By using the matrix formula

\[ \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det A \det (D - CA^{-1}B), \quad (3.38) \]

the determinant of \( M[k] \) is

\[ \det \{ M[k] \} = N_R^{2NT} \prod_{m=0}^{NT-1} \left\{ \prod_{\xi=0}^{1} \left\{ \sigma_n^2 + a^2[N_T(k-\xi) + m] \right\} \right. \]
\[ \left. -\phi_{hh}^2[N_T] \prod_{\xi=0}^{1} a^2[N_T(k-\xi) + m] \right\} \]
\[ = N_R^{2NT} \prod_{m=0}^{NT-1} \left\{ \prod_{\xi=0}^{1} \left\{ \sigma_n^2 + (\rho^{2\Delta t\alpha[K]})^\xi a^2[N_T(k-1) + m] \right\} \right. \]
\[ \left. -\phi_{hh}^2[N_T] \prod_{n=0}^{1} \left( \rho^{2\Delta t\alpha[K]} \right)^\xi a^2[N_T(k-1) + m] \right\} \]}
3.2 Conventional Differential Detection

\[ N_T^{2N_{R}} \prod_{m=0}^{N_{T}-1} \left\{ \prod_{\xi=0}^{1} \left( \sigma_n^2 + \rho^2(\xi \Delta \alpha[k]+\alpha[k-1]+\theta_m) \right) \mod L_A \right\} \]

\[ -\varphi_{kh}[NT] \prod_{\xi=0}^{1} \rho^2(\xi \Delta \alpha[k]+\alpha[k-1]+\theta_m) \mod L_A \right\} \]

Now, the determinant of \( C_R[k] \) can be calculated to

\[ \det\{C_R[k]\} = \det\{\hat{P}[k]\}\det\{M[k]\}\det\{\hat{P}^H[k]\} = \det\{M[k]\}. \quad (3.40) \]

From Eqs. (3.39), (3.40), we observe that \( \det\{C_R[k]\} \) depends on \( \Delta \alpha[k] \) and the previous absolute amplitude symbol \( \alpha[k-1] \). We also obtain

\[ -C_R^{-1}[k] = \hat{P}[k]\left\{ -M[k]\right\}^{-1}\hat{P}^H[k] \]

\[ = \hat{P}[k]\hat{T}[k]\hat{P}^H[k], \quad (3.41) \]

with the definition

\[ \hat{T}[k] \triangleq -M^{-1}[k] \]

\[
\begin{bmatrix}
t_{00} & 0 & \cdots & 0 & t_{0N_T} & 0 & \cdots & 0 \\
0 & t_{11} & \cdots & 0 & 0 & t_{1N_T} & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & t_{N_T-1N_T-1} & 0 & \cdots & 0 & t_{N_T-12N_T-1} \\
t_{N_T0} & 0 & \cdots & 0 & t_{N_TN_T} & 0 & \cdots & 0 \\
0 & t_{N_T+11} & \cdots & 0 & 0 & t_{N_T+1N_T} & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & t_{2N_T-1N_T-1} & 0 & \cdots & 0 & t_{2N_T-12N_T-1}
\end{bmatrix}
\]

\[ (3.42) \]
3.2 Conventional Differential Detection

where

\[
t_{m,n_1} = \frac{-\left(\sigma_n^2 + \rho^2(\Delta a[k]+la[k-1]+\theta_m)\mod LA\right)}{N_R \vartheta},
\]

\[
t_{(N_T+m)(N_T+m)} = \frac{-\left(\sigma_n^2 + \rho^2(2\Delta a[k]+2la[k-1]+2\theta_m)\mod LA\right)}{N_R \vartheta},
\]

\[
t_{m(N_Tn_1)} = \frac{\varphi_{hh}[N_T]p^2(\Delta a[k]+2la[k-1]+2\theta_m)\mod LA}{N_R \vartheta},
\]

with \( \vartheta = \prod_{\xi=0}^{1} \left\{ \sigma_n^2 + \rho^2(\xi\Delta a[k]+la[k-1]+\theta_m)\mod LA\right\} - \varphi_{hh}^2[N_T] \prod_{\xi=0}^{1} \rho^2(\xi\Delta a[k]+la[k-1]+\theta_m)\mod LA \).

Notice that \( t_{m(N_T+m)} = t_{(N_T+m)m} \), for \( 0 \leq m \leq N_T - 1 \).

Also we have the relation [11]

\[
p_m[N_T(k-1) + m]p_m^*[N_Tk + m] = \exp \left( j \frac{2\pi}{LP} u_m \Delta \eta[k] \right).
\]

So we can further simplify the numerator of Eq. (3.28)

\[
-\text{tr} \left( \bar{R}^H[k]C^{-1}_R[k]\bar{R}[k] \right) = -\text{tr} \left( \bar{R}^H[k]\bar{P}[k]N_T\bar{P}^H[k]\bar{R}[k] \right)
\]

\[
= \sum_{m=0}^{N_T-1} \sum_{n=0}^{N_R-1} \sum_{\xi=0}^{1} \left\{ t_{\xi(N_T+m)(\xi(N_T+m)} r_n[N_T(k-1+\xi) + m] \right\}^2
\]

\[
+ 2\Re \left\{ \sum_{m=0}^{N_T-1} \sum_{n=0}^{N_R-1} t_{m(N_T+m)} e^{j \frac{2\pi}{LP} u_m \Delta \eta[k]} \cdot r_n^*[N_Tk + m]r_n[N_T(k-1) + m] \right\}.
\]

Maximizing \( p(\bar{R}[k]|\bar{S}[k]) \) is equivalent to maximizing \( \ln(p(\bar{R}[k]|\bar{S}[k])) \), where \( \ln(\cdot) \) refers to the natural logarithm. Therefore, we can get the decision rule
\[
\left(\Delta \hat{a}[k], \Delta \tilde{p}[k]\right) = \arg \max_{\Delta \hat{a}[k], \Delta \tilde{p}[k]} \left\{ \prod_{m=0}^{N_T-1} \left\{ \prod_{\xi=0}^{N_R-1} \left\{ \sigma_n^2 + \rho^2(\xi \Delta \hat{a}[k] + la[k-1] + \theta_m) \mod LA \right\} - \varphi^2[N_T] \prod_{\xi=0}^{N_R-1} \rho^2(\xi \Delta \hat{a}[k] + la[k-1] + \theta_m) \mod LA \right\} \right. \\
+ \sum_{m=0}^{N_T-1} \sum_{n=0}^{N_R-1} \sum_{\xi=0}^{N_R-1} \left\{ t_{\xi N_T+m}(\xi N_T+m) \left| r_{n[N_T(k-1+\xi)+m]} \right|^2 \right\} \\
+ 2\Re \left\{ \sum_{m=0}^{N_T-1} \sum_{n=0}^{N_R-1} t_{m(N_T+m)} e^{\left( j \frac{2\pi}{N_T} u_m \Delta \tilde{p}[k] \right)} \cdot r_{n[N_T k + m]}^* r_{n[N_T(k-1)+m]} \right\} \right. \\
\left. \right\}, \quad (3.48)
\]

where \( \Delta \hat{a}[k] \in \{0, 1, \ldots, LA - 1\} \) and \( \Delta \tilde{p}[k] \in \{0, 1, \ldots, LP - 1\} \). This decision rule depends on \( la[k-1] \), which is estimated from the previous estimated amplitude difference symbol as

\[
\hat{a}[k-1] = \left( \sum_{\chi=0}^{k-1} \Delta \hat{a}[\chi] \right) \mod LA \\
= \Delta \hat{a}[k-1] + \hat{a}[k-2], \quad (3.49)
\]

with \( \Delta \hat{a}[0] = 0 \).

### 3.3 Multiple-Symbol Detection

The observation window of ML-DD can be generalized from 2 to \( N, \ N > 2, \) consecutive received signal matrices, which leads to multiple-symbol detection (MSD) [15]. MSD makes a joint decision on \( N-1 \) symbols based on \( N \) received matrices. These \( N \) consecutive received matrices are stacked to form a matrix \( \hat{R}[k] \) with \( NN_T \) rows to obtain
3.3 Multiple-Symbol Detection

\[
\mathbf{R}[k] = \begin{bmatrix}
\mathbf{R}[k - N + 1] \\
\mathbf{R}[k - N + 2] \\
\vdots \\
\mathbf{R}[k]
\end{bmatrix},
\] (3.50)

For this, \( \mathbf{R}[k] \) can be expressed as

\[
\mathbf{R}[k] = \mathbf{S}[k] \mathbf{H}[k] + \mathbf{N}[k],
\] (3.51)

with the definitions

\[
\mathbf{S}[k] = \begin{bmatrix}
\mathbf{S}[k - N + 1] \\
\mathbf{O}_{NT} \\
\vdots \\
\mathbf{O}_{NT} \\
\mathbf{S}[k - N + 2] \\
\vdots \\
\mathbf{O}_{NT} \\
\vdots \\
\mathbf{O}_{NT} \\
\mathbf{S}[k - N + 2] \\
\vdots \\
\mathbf{O}_{NT} \\
\mathbf{S}[k]
\end{bmatrix},
\] (3.52)

\[
\mathbf{H}[k] = \begin{bmatrix}
\mathbf{H}[k - N + 1] \\
\mathbf{H}[k - N + 2] \\
\vdots \\
\mathbf{H}[k]
\end{bmatrix},
\] (3.53)

\[
\mathbf{N}[k] = \begin{bmatrix}
\mathbf{N}[k - N + 1] \\
\mathbf{N}[k - N + 2] \\
\vdots \\
\mathbf{N}[k]
\end{bmatrix},
\] (3.54)

The amplitude matrix becomes

\[
\mathbf{A}[k] = \begin{bmatrix}
\mathbf{A}[k - N + 1] \\
\mathbf{O}_{NT} \\
\vdots \\
\mathbf{O}_{NT} \\
\mathbf{A}[k - N + 2] \\
\vdots \\
\mathbf{O}_{NT} \\
\vdots \\
\mathbf{O}_{NT} \\
\vdots \\
\mathbf{O}_{NT} \\
\mathbf{A}[k]
\end{bmatrix},
\] (3.55)

with the elements of \( \mathbf{A}[k - \xi] \) given by
3.3 Multiple-Symbol Detection

\[ a[N_T(k - \xi) + m] = \rho \left( \sum_{\nu = \xi}^{N-2} \Delta a[k-\nu] + a[k-N+1] + \theta_m \right) \mod L_A, \quad 0 \leq \xi \leq N - 2, \quad (3.56) \]

The channel covariance matrix

\[ C_H = N_R \left( C_h \otimes I_{NN_T} \right), \quad (3.57) \]

with

\[ C_h = \begin{bmatrix}
\varphi_{hh}[0] & \varphi_{hh}[N_T] & \cdots & \varphi_{hh}[(N - 1)N_T] \\
\varphi_{hh}[-N_T] & \varphi_{hh}[0] & \cdots & \varphi_{hh}[(N - 2)N_T] \\
\vdots & \vdots & \ddots & \vdots \\
\varphi_{hh}[-(N - 1)N_T] & \cdots & \cdots & \varphi_{hh}[0]
\end{bmatrix}. \quad (3.58) \]

The matrix \( T[k] \) can be represented as

\[ T[k] = \left\{ - \left( \bar{A}[k] C_H \bar{A}[k]^H + N_R \sigma_n^2 I_{N_N T} \right) \right\}^{-1} \]

\[ = \begin{bmatrix}
t_{00} & \cdots & 0 & \cdots & t_{0(N-1)N_T} & \cdots & 0 \\
0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & t_{N_T-1N_T-1} & \cdots & 0 & 0 & t_{N_T-1N_N T-1} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
t_{(N-1)N_T} & \cdots & 0 & \cdots & t_{(N-1)N_T(N-1)N_T} & \cdots & 0 \\
0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & t_{N_N T-1N_T-1} & \cdots & 0 & 0 & t_{N_N T-1N_N T-1}
\end{bmatrix}. \quad (3.59) \]

Also we have the relation [11]

\[ p_m[N_T(k - \xi_1) + m] p_m^*[N_T(k - \xi_2) + m] = \prod_{\nu = \xi_1 + 1}^{\xi_2} \exp \left( j \frac{2\pi}{LP} u_m \Delta \nu [k - \nu] \right). \quad (3.60) \]
3.4 Decision-Feedback Differential Detection

When $H[k]$ is unknown, the pdf of $\bar{R}[k]$ conditioned on $\bar{S}[k]$ is given by Eq. (3.28) with $N > 2$. Following the same step as ML-DD for $N = 2$, we get the decision rule of MSD

$$
\left( (\Delta \hat{a}[k], \Delta \hat{p}[k]), \cdots, (\Delta \hat{a}[k-N+2], \Delta \hat{p}[k-N+2]) \right)
= \arg \max_{(\Delta \hat{a}[k], \Delta \hat{p}[k]) \cdots (\Delta \hat{a}[k-N+2], \Delta \hat{p}[k-N+2])}
\left\{ \left| \det \{ \hat{A}[k] \mathbf{C}_H \hat{A}^H[k] + N_R \sigma_n^2 \mathbf{I}_{NN_T} \} \right|^2 + \sum_{m=0}^{N_T-1} \sum_{n=0}^{N-1} \sum_{\xi=0}^{N-1} \sum_{(\xi N_T+m) (\xi N_T+m)} \left| r_n[N_T(k+1+\xi)+m] \right|^2 \right\}
+ 2\Re \left\{ \sum_{m=0}^{N_T-1} \sum_{n=0}^{N-1} \sum_{\xi_1=0}^{N-1} \sum_{\xi_2=\xi_1+1} \sum_{(\xi_1 N_T+m) (\xi_2 N_T+m)} \prod_{\nu=\xi_1}^{\xi_2} e^{j \frac{2\pi}{N} m \Delta \hat{p}[k-\nu] r_n[N_T(k-\xi_1)+m] r_n[N_T(k-\xi_2)+m]} \right\},
$$

(3.61)

where $\{\Delta \hat{a}[k], \cdots, \Delta \hat{a}[k-N+2]\} \in \{0, 1, \cdots, LA-1\}$ and $\{\Delta \hat{p}[k], \cdots, \Delta \hat{p}[k-N+2]\} \in \{0, 1, \cdots, LP - 1\}$. This decision rule depends on $la[k-N+1]$, which is estimated from the previous estimated amplitude symbols in the same way as for ML-DD. MSD takes the dependencies among the received symbols into account and can reduce the irreducible error floor of ML-DD. Unfortunately, the above MSD decision rule requires the calculation of $2^{N_T R(N-1)}/(N_T R(N-1))$ metrics per bit decision, i.e., its computational complexity is exponential in the number of transmit antennas $N_T$, data rate $R$, and observation window size $N$. An alternative scheme with similar performance to MSD, but with a complexity almost independent of $N$, is subsequently derived.

3.4 Decision-Feedback Differential Detection

As mentioned in the previous section, we can reduce computational complexity by introducing decision-feedback [11], i.e., the $\Delta la[k-\nu]$ and $\Delta lp[k-\nu]$ are replaced by the previously
decided symbols $\Delta \hat{a}[k - \nu]$ and $\Delta \hat{p}[k - \nu]$, $1 \leq \nu \leq N - 2$, in Eq. (3.61). Therefore the decision is made only on $\Delta \hat{a}[k]$ and $\Delta \hat{p}[k]$.

\[
(\Delta \hat{a}[k], \Delta \hat{p}[k]) = \arg \max_{\Delta \hat{a}[k], \Delta \hat{p}[k]} \left\{ \det \left\{ \tilde{A}[k]C_H \tilde{A}^H[k] + N_R \sigma^2 I_{NT} \right\} \right. \\
\left. + \sum_{m=0}^{N_T-1} \sum_{n=0}^{N_R-1} \sum_{\xi=0}^{N-1} t(\xi T + m) (\xi T + m) r_n[N_T(k - 1 + \xi) + m] \right. \\
\left. + 2 \Re \left\{ \sum_{m=0}^{N_T-1} \sum_{n=0}^{N_R-1} e^{j \frac{2 \pi}{L \phi} u_m \Delta \hat{p}[k]} \cdot r_n^* [N_T k + m] r_{\text{ref}, n}[N_T(k - 1) + m] \right\} \right\}, \tag{3.62}
\]

with the reference phase signal

\[
r_{\text{ref}, n}[N_T(k - 1) + m] \\
= \sum_{\xi_1=0}^{N-1} \sum_{\xi_2=\xi_1+1}^{N-1} t(\xi_1 T + m) (\xi_2 T + m) \prod_{\nu=\xi_1+1}^{\xi_2} \exp \left( j \frac{2 \pi}{L \phi} u_m \Delta \hat{p}[k - \nu] \right) \cdot r_n[N_T(k - \xi_2) + m]. \tag{3.63}
\]

The $(N_T k + m, N_T k + m)$th element of $\tilde{A}[k]$ is

\[
\tilde{a}[N_T k + m] = \rho \left( \sum_{\nu=\xi}^{N-2} \hat{a}[k - \nu] + \hat{a}[k - N + 1] + \theta_m + \Delta \hat{a}[k] \right) \mod LA, \tag{3.64}
\]

and the $(N_T(k - \xi) + m, N_T(k - \xi) + m)$th element of $\tilde{A}[k]$ is

\[
\tilde{a}[N_T(k - \xi) + m] = \rho \left( \sum_{\nu=\xi}^{N-2} \hat{a}[k - \nu] + \hat{a}[k - N + 1] + \theta_m \right) \mod LA, \tag{3.65}
\]

with $1 \leq \xi \leq N - 1$. The trial symbols $\Delta \hat{a}[k] \in \{0, 1, \ldots, LA - 1\}$ and $\Delta \hat{p}[k] \in \{0, 1, \ldots, LP - 1\}$. Now, only $2^{NT R}/(NT R)$ metrics per bit decision have to be calculated,
i.e., the computational complexity is only exponential in $N_T$ and $R$. For the special case $N = 2$, both the MSD (cf. Eq. (3.61)) and the DF-DD (cf. Eq. (3.62)) decision rules are identical to the decision rule for ML-DD given in Eq. (3.48).

### 3.5 Modifications for Time Diversity

For DAPSTM to exploit space diversity, each transmit antenna is active only in every $N_T$-th symbol interval, therefore; the effective fading bandwidth relevant for the receiver is $N_TB_fT$, instead of $B_fT$, which has a negative influence on receiver performance [16]. In time diversity, the effective fading bandwidth is $B_fT$, independent of $N_B$. Time diversity leads to a lower error floor. On the other hand, since an interleaver is used for time diversity, a transmission delay of $N_B \times N_I$ is introduced, whereas only a delay of $N_T$ symbol periods is introduced in space diversity.
Chapter 4

Performance Analysis

In this chapter, the performance of DAPSTM is analyzed for flat Rayleigh fading channels without spatial correlation. First, the PEP is evaluated based on SH-DD, since the calculation of PEP for MI-DD is too involved, if not impossible. An approximation for BER is obtained from the weighted PEPs.

4.1 Pairwise Error Probability

The PEP $P_e(\alpha, \beta)$ is the probability of detecting $\hat{V}[k] = V_\beta[k]$ ($\Delta l_\beta[k] = \Delta l p_\beta[k]$, $\Delta l a[k] = \Delta l a_\beta[k]$), when $V[k] = V_\alpha[k]$ ($\Delta l p[k] = \Delta l p_\alpha[k]$, $\Delta l a[k] = \Delta l a_\alpha[k]$), $V_\alpha[k] \neq V_\beta[k]$, is transmitted. Here $\Delta l a_\alpha[k], \Delta l a_\beta[k] \in \{0, 1, \cdots LA-1\}$ and $\Delta l p_\alpha[k], \Delta l p_\beta[k] \in \{0, 1, \cdots LP-1\}$. Using Eq. (3.27), it becomes a straightforward task to show that the PEP of DAPSTM with SH-DD can be expressed as

$$P_e(\alpha, \beta) = \text{Pr}\left\{ \| R[k] - V_\beta[k] R[k-1] \|^2 < \| R[k] - V_\alpha[k] R[k-1] \|^2 \right\}$$
4.1 Pairwise Error Probability

\[
\Pr \left\{ \sum_{n=0}^{N_R-1} \sum_{m=0}^{N_T-1} \left| r_n[N_Tk + m] - v_{\beta_m}[N_Tk + m]r_n[N_T(k-1) + m] \right|^2 \right. \\
\left. < \sum_{n=0}^{N_R-1} \sum_{m=0}^{N_T-1} \left| r_n[N_Tk + m] - v_{\alpha_m}[N_Tk + m]r_n[N_T(k-1) + m] \right|^2 \right\}
\]

\[
= \Pr \left\{ \sum_{n=0}^{N_R-1} \sum_{m=0}^{N_T-1} \left\{ 2\Re \left\{ r_n^*[N_Tk + m] (v_{\beta_m}[N_Tk + m] - v_{\alpha_m}[N_Tk + m]) r_n[N_T(k-1) + m] \right\} \\
+ \left| r_n[N_T(k-1) + m] \right|^2 \left( \left| v_{\beta_m}[N_Tk + m] \right|^2 - \left| v_{\alpha_m}[N_Tk + m] \right|^2 \right) \right\} < 0 \right\}
\]

\[
= \Pr \{ \Delta(\alpha, \beta) < 0 \}. \tag{4.1}
\]

where \( v_m[N_Tk + m] = e^{j \frac{2\pi}{T} n m \Delta[k]} \cdot \rho \Delta[k] \). Here \( \Delta(\alpha, \beta) \) is defined as

\[
\Delta(\alpha, \beta) \triangleq \sum_{n=0}^{N_R-1} \sum_{m=0}^{N_T-1} \left\{ x_{mn}^*[k] C_m y_{mn}[k-1] + x_{mn}[k] C_m^* y_{mn}^*[k-1] \\
+ y_{mn}^*[k-1] y_{mn}[k-1] B_m \right\}, \tag{4.2}
\]

\[
B_m \triangleq \left| v_{\beta_m}[N_Tk + m] \right|^2 - \left| v_{\alpha_m}[N_Tk + m] \right|^2, \tag{4.3}
\]

\[
C_m \triangleq v_{\beta_m}[N_Tk + m] - v_{\alpha_m}[N_Tk + m], \tag{4.4}
\]

\[
x_{mn}[k] \triangleq r_n[N_Tk + m], \tag{4.5}
\]

\[
y_{mn}[k-1] \triangleq r_n[N_T(k-1) + m]. \tag{4.6}
\]

Using vector notation, \( \Delta(\alpha, \beta) \) can be written as quadratic form

\[
\Delta(\alpha, \beta) = x^H C^H y + y^H C^H x + y^H B y
\]

\[
= g^H F g, \tag{4.7}
\]

with
4.1 Pairwise Error Probability

\[ g \triangleq [x^T \ y^T]^T, \quad (4.8) \]
\[ F \triangleq \begin{bmatrix} O_{N_T N_R} & C^H \\ C & B \end{bmatrix}, \quad (4.9) \]
\[ x \triangleq \begin{bmatrix} r_0[N_T k] & r_0[N_T k + 1] & \cdots & r_0[N_T k + N_T - 1] & r_1[N_T k] \\
& \cdots & \cdots & \cdots & \cdots \\
& & & & r_{N_R - 1}[N_T k + N_T - 1] \end{bmatrix}^T, \quad (4.10) \]
\[ y \triangleq \begin{bmatrix} r_0[N_T (k - 1)] & r_0[N_T (k - 1) + 1] & \cdots & r_0[N_T (k - 1) + N_T - 1] \\
& \cdots & \cdots & \cdots & \cdots \\
& & & & r_{N_R - 1}[N_T (k - 1) + N_T - 1] \end{bmatrix}^T, \quad (4.11) \]
\[ B \triangleq I_{N_R} \otimes \text{diag}\{B_0, \ B_1, \ \cdots, \ B_{N_T - 1}\}, \quad (4.12) \]
\[ C \triangleq I_{N_R} \otimes \text{diag}\{C_0, \ C_1, \ \cdots, \ C_{N_T - 1}\}. \quad (4.13) \]

Since \( \Delta(\alpha, \beta) \) is a quadratic form of Gaussian random variables, the two-sided Laplace transform \( \Phi_{\Delta(\alpha, \beta)}(s) \) of its pdf can be expressed as [16]

\[
\Phi_{\Delta(\alpha, \beta)}(s) = \mathcal{E}\left\{e^{-s\Delta(\alpha, \beta)}\right\} = \int_{-\infty}^{\infty} p_{\Delta(\alpha, \beta)}(x) e^{-s\Delta(\alpha, \beta)} d\Delta(\alpha, \beta)
\]
\[
= \exp \left( -s\bar{g}^H \left( F^{-1} + s\Psi_{gg} \right)^{-1} \bar{g} \right) / \det \left( I_{2N_T N_R} + s\Psi_{gg} F \right), \quad (4.14) \]

where the definitions

\[ \bar{g} \triangleq \mathcal{E}\{g\}, \quad (4.15) \]
\[ \Psi_{gg} \triangleq \mathcal{E}\{(g - \bar{g})(g - \bar{g})^H\}, \quad (4.16) \]

are used. Since vector \( g \) has zero mean, \( \bar{g} = 0_{2N_T N_R \times 1} \), where \( 0_{2N_T N_R \times 1} \) is a \( 2N_T N_R \times 1 \) all 0 vector. \( \Phi_{\Delta(\alpha, \beta)}(s) \) can be simplified to

\[
\Phi_{\Delta(\alpha, \beta)}(s) = \frac{1}{\det \left( I_{2N_T N_R} + s\Psi_{gg} F \right)}. \quad (4.17) \]
4.1 Pairwise Error Probability

We can calculate $P_e(\alpha, \beta)$ from

$$P_e(\alpha, \beta) = \frac{1}{2\pi j} \int_{-\infty}^{\infty} \Phi_{\Delta(\alpha,\beta)}(s) \frac{ds}{s}$$

(4.18)

for $0 < \gamma < \Re\{s_1\}$, where $s_1$ refers to the pole of $\Phi_{\Delta(\alpha,\beta)}(s)$ which has the minimum positive real part.

The integral in Eq. (4.18) can be calculated in closed form using the residues theorem [16]

$$P_e(\alpha, \beta) = -\sum_{\text{RH poles}} \text{Residues}(\Phi_{\Delta(\alpha,\beta)}(s)/s),$$

(4.19)

where the summation is taken over all residues corresponding to poles located in the right-hand (RH) side of the complex $s$-plane. Eq. (4.19) constitutes a very general method of calculating $P_e(\alpha, \beta)$. We can easily get $P_e(\alpha, \beta)$ when the poles are simple. But it is very difficult to calculate the residues if $\Phi_{\Delta}(s)/s$ has multiple poles or essential singularities. To avoid this problem, some approximations, such as the Chernoff bound [17] can be used to simplify the calculation. The bound takes the simple form

$$P_e(\alpha, \beta) \leq \min_{0 < \gamma < s_1} \Phi_{\Delta(\alpha,\beta)}(\gamma).$$

(4.20)

The Chernoff bound is a general technique for upperbounding the PEP. But in most cases, it fails to provide a tight enough bound for error probability. Therefore, we favor a numerical calculation approach, which is based on Gauss-Chebyshev quadrature rules [17]:

$$P_e(\alpha, \beta) = \frac{1}{N_G} \sum_{k=1}^{N_G/2} \left( \Re\{\Phi_{\Delta(\alpha,\beta)}(\gamma + j\gamma\tau_k)\} + \tau_k \Im\{\Phi_{\Delta(\alpha,\beta)}(\gamma + j\gamma\tau_k)\} \right) + E_{N_G},$$

(4.21)
where $N_G$ is even, and $0 < \gamma < \Re\{s_1\}$ and $\tau_k \triangleq \tan((2k - 1)\pi/(2N_G))$ are valid. The error $E_{N_G}$ vanishes for $N_G \to \infty$. In practice, relatively small values for $N_G$ can be used if $\gamma$ is chosen properly [17]. For the numerical results in this section, $N_G = 128$ and $\gamma = \Re\{s_1\}$ are adopted.

### 4.2 Approximation for Bit Error Rate

The exact calculation of the BER is quite involved. A simple approximation of the BER is given by the union bound [18]

$$P_b = \frac{1}{2N_b} \sum_{\alpha=1}^{2N_b} \sum_{\beta \neq \alpha}^{2N_b} \frac{n(\alpha, \beta)}{N_b} P_e(\alpha, \beta), \quad (4.22)$$

where $N_b = 2^{N_T(R_A+R_P)}$, and $n(\alpha, \beta)$ denotes the number of bit errors if $V_\alpha$ is transmitted and $V_\beta$ is detected. $P_b$ is an upper bound for the achievable BER for space-time coded transmission. However, for moderate-to-large signal-to-noise ratios (SNR), the union bound becomes tight and is a very good approximation for the achievable BER.
Chapter 5
Simulation Results

In this chapter, simulation and numerical results are presented and discussed. We have simulated a system as described in Chapter 3, i.e., independent flat Rayleigh fading channels are assumed between any pair of transmit and receive antennas. We restrict ourselves to consider only one receive antenna \( N_R = 1 \), since the main focus of this work is on transmit (modulation) diversity. In the simulation, \( E_b/N_0 = 1/(\sigma_n^2 R) \) is used as the channel SNR, where \( E_b \) is the total energy per bit used in the transmission. For the numerical evaluation, the results given in Chapter 4 are used.

5.1 Numerical Results Compared with Simulations

First, we compare the simulation and numerical results of PEP. For simulation SH-DD is applied, and for the numerical expression of PEP Eq. (4.21) is used. Fig. 5.1 shows the BER vs. \( 10\log_{10}(E_b/N_0) \) for \( N_T = 2 \). \( R = 1.5 \) bits/(channel use) (\( RP = 1 \) bit/(channel use), \( RA = 0.5 \) bit/(channel use)), \( B_f T = 0.001 \), \( \rho = 2.1 \), \( u_0 = 1 \), \( u_1 = 1 \), \( \theta_0 = 0 \), and \( \theta_1 = 1 \) are adopted. So, \( \Delta l a \in \{0,1\} \) and \( \Delta l p \in \{0,1,2,3\} \). \( V_\alpha \to V_\beta \) denotes the probability of detecting \( V_\beta \), when \( V_\alpha \) is transmitted. Two particular pairs of PEP are
considered. The first one is the PEP between $V_{13}(\Delta l_a = 1, \Delta l_p = 3)$ and $V_{03}(\Delta l_a = 0, \Delta l_p = 3)$, where only an amplitude change occurs, and the other is the PEP between $V_{13}(\Delta l_a = 1, \Delta l_p = 3)$ and $V_{10}(\Delta l_a = 1, \Delta l_p = 0)$, where only a phase change occurs. We observe that PEP($V_{13}, V_{10}$) = PEP($V_{10}, V_{13}$), and in this case, the numerical results of the PEP match well with simulations. When the amplitude change happens, both numerical and simulation results of PEP are not symmetric, i.e., PEP($V_{13}, V_{03}$) ≠ PEP($V_{03}, V_{13}$) in either numerical or simulation results. For numerical results, we assume the previous absolute amplitude symbols are known at the receiver, and no error propagation
5.1 Numerical Results Compared with Simulations

is considered, while error propagation actually occurs in the simulation. The numerical result for $\text{PEP}(V_{13}, V_{03})$ is lower than the simulation result. These two results match remarkably well when we cancel the effect of the error propagation in simulation.

![Figure 5.2: Comparison of simulation vs. numerical result of BER. $N_T = 2$, $R = 4$ bits/(channel use) ($RP = 3$ bits/(channel use), $RA = 1$ bit/(channel use)), $B_fT = 0.001$, $\rho = 1.4$, $u_0 = 1$, $u_1 = 15$, $\theta_0 = 0$, and $\theta_1 = 1$ are valid.](image)

Fig. 5.2 shows BER vs. $10\log_{10}(E_b/N_0)$ for simulation and numerical evaluation using Eq. (4.22). Again, SH-DD is assumed. $N_T = 2$, $R = 4$ bits/(channel use) ($RP = 3$ bits/(channel use), $RA = 1$ bit/(channel use)) and $B_fT = 0.001$ are adopted. $\rho = 1.4$, $u_0 = 1$, $u_1 = 15$, $\theta_0 = 0$, and $\theta_1 = 1$ optimized in Section 5.2 are valid. To make this comparison conclusive, we assume all the decisions of the previous absolute amplitude symbols are correct, i.e., no error propagation occurs in the simulation. It can be observed
5.2 Constellations Design

As mentioned in Section 3.1, the BER performance depends on three parameters in our code designs, magnitude ratio $\rho$, $\rho > 0$, the initial amplitude value for each antenna $\theta_m$, $0 \leq m \leq N_T - 1$, and coefficients $u_m$, $0 \leq m \leq N_T - 1$, for the phase signal. It is necessary to find the optimum combination of these three parameters to get the best performance.

For a very simple case, $N_T = 2$, $R = 1.5$ bits/(channel use) ($RP = 1$ bit/(channel use), $RA = 0.5$ bit/(channel use)), $\theta_0$ and $\theta_1$ can be taken from the set $\{0, 1\}$, and $u_0$, $u_1$ is drawn from the set $\{0, 1, 2, 3\}$. We find the optimum values, $u_0 = 1$, $u_1 = 1$ according to the search method in [5], $\rho = 2.1$ and $\theta_0 \neq \theta_1$ ($\theta_0 = 0, \theta_1 = 1$ or $\theta_0 = 1, \theta_1 = 0$) by simulation. In more complicated cases, such as with multiple antennas or high data rates, there are many possible combinations of $\rho$, $\theta_m$, $u_m$. To optimize the parameters $\rho$, $\theta_m$, $u_m$, we need to calculate and compare the BER for different combinations of parameters.

However, it is too difficult to take into account all the PEPs, which are used to calculate BER. Consequently, we turn to a suboptimum way, in which we use the maximum PEP as the metric. The PEP (cf. Eq. (4.1)) is derived for SH-DD, and the parameters resulting from the optimization also give good performance for ML-DD. Our goal is to find $\rho$, $\theta_m$, and $u_m$ satisfying

$$\left\{ \rho, \theta_1, \cdots, \theta_{N_T-1}, u_0, u_1, \cdots, u_{N_T-1} \right\} = \arg \min_{\rho > 0} \left\{ \max_{\alpha, \beta, \alpha \neq \beta} P_e(\alpha, \beta) \right\}. \quad (5.1)$$

where (cf. Eq. (4.21))
5.2 Constellations Design

\[ P_e(\alpha, \beta) = \frac{1}{N_G} \sum_{k=1}^{N_G/2} \left( \Re\{\Phi_{\Delta(\alpha, \beta)}(\gamma + j\gamma \tau_k)\} + \tau_k \Im\{\Phi_{\Delta(\alpha, \beta)}(\gamma + j\gamma \tau_k)\} \right). \] (5.2)

Because each antenna is statistically equivalent to the others, we may impose the ordering \( u_0 \leq u_1 \leq \cdots \leq u_{N_T-1}, \theta_0 \leq \theta_1 \leq \cdots \leq \theta_{N_T-1}. \)

It is impossible to get an explicit solution for the procedure described above; therefore, we resort to exhaustive computer searches. The search space can be reduced using the following rules:

1. \( u_0 = 1, u_m \in \{3, 5, \ldots, LP/2 - 1\}, \) where \( 1 \leq m \leq N_T - 1 \) [5].

2. We restrict \( p > 1. \) If \( p < 1, \) it is effectively equivalent to another \( p > 1. \)

3. The phase difference symbol \( \Delta lp[k] \) may be taken from the set \( \{1, 2, \ldots, LP/2 - 1\} \) [5], and the amplitude difference symbol \( \Delta la[k] \in \{0, 1, 2, \ldots, LA - 1\}. \)

4. We assume \( \theta_0 = 0, \theta_{N_T-1} = LA - 1 \) and try to arrange the rest of the \( \theta_1, \theta_2, \ldots, \theta_{N_T-2} \) equally spaced between 0 and \( LA - 1 \) to maximize the distance between any two elements in a diagonal amplitude matrix symbol. For the special case of \( N_T = 2, RA = 1 \text{ bit/(channel use)}, \) we have \( \theta_0 = 0, \theta_1 = 3. \) For the case of \( N_T = 3, RA = 1 \text{ bit/(channel use)}, \) we assign \( \theta_0 = 0, \theta_1 = 3 \) and \( \theta_2 = 7. \)

Using the proposed procedure, we find the optimum values of the parameters for a given \( E_b/N_0 \) and \( R = 3 \text{ bits/(channel use)}, R = 4 \text{ bits/(channel use)} \) when \( N_T = 2, \) cf. Table 5.1. In order to yield a low BER for a given signal constellation, a kind of Gray mapping is desirable for the phase signal \( \Delta lp[k] \) and the amplitude signal \( \Delta la[k], \) respectively. For this, the nearest neighbors of the phase signal for each matrix symbol are first determined, according to the minmax value of PEP (cf. Eq. (5.1)), which can be calculated by Eq. (4.21).
5.2 Constellations Design

Table 5.1: DAPSTM parameters for \( N_T = 2 \) by minmax PEP.

<table>
<thead>
<tr>
<th>( E_b/N_0) (dB)</th>
<th>( R = 3 ) (( RP = 2 ), ( RA = 1 ))</th>
<th>( R = 4 ) (( RP = 3 ), ( RA = 1 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \rho )</td>
<td>( u_1 )</td>
</tr>
<tr>
<td>20</td>
<td>1.5</td>
<td>5</td>
</tr>
<tr>
<td>30</td>
<td>1.5</td>
<td>3</td>
</tr>
<tr>
<td>200</td>
<td>1.9</td>
<td>7</td>
</tr>
</tbody>
</table>

Assume each phase signal has two nearest neighbors with the same distance \( \tilde{l}_p \), and \( \tilde{l}_p \) for all the phase signals are identical; that is the nearest neighbor symbols for any given \( \Delta l_p[k] \), \( 0 \leq \Delta l_p[k] \leq LP - 1 \), can be expressed as \( (\Delta l_p[k] \pm \tilde{l}_p) \mod LP \), and \( \tilde{l}_p \) is unique. In this case, a Gray mapping may be constructed by assigning bit patterns to the phase signal in such a way that \( \Delta l_p[k] \) and its neighboring signals differ by only one bit. Otherwise, there may be constellations for which no Gray mapping exists. We managed to find \( \tilde{l}_p \) for \( N_T = 2 \) and \( R = 3,4 \) bits/(channel use) as given in Table 5.1. However, following the same procedure as for \( \Delta l_p[k] \), we could not express the nearest neighbors of \( \Delta l_a[k] \) as \( (\Delta l_a[k] \pm \tilde{l}_a) \mod LA \) with a unique distance \( \tilde{l}_a \). Here, we simply use Gray mapping with \( \tilde{l}_a = 1 \) for the amplitude signal mapping. Simulation results show that for \( N_T = 3, RA = 1 \) bit/(channel use) this Gray mapping for the amplitude signal yields the best performance. Further, while \( N_T = 2 \) and \( RA = 1 \) bit/(channel use), different mapping methods yield a similar performance.

Notice that the differential transmission schemes proposed in [5] and [6] appear as special cases of our scheme for \( RA = 0 \). Therefore, it appears to be reasonable to use optimum values of \( u_m, \tilde{l}_p \) for DUSTM (cf. Table 2.1, Table 2.2) as the suboptimum values in our signal design. This then provides an alternative means to the parameter optimization method described above. By this means, we may lose a little performance quality, but the computational complexity is significantly reduced. Again, we assign \( \theta_0 = 0 \) and \( \theta_1 = 3 \) for
5.3 Space Diversity

the case of $N_T = 2$ and $RA = 1$ bit/(channel use). Further we assign $\theta_0 = 0$, $\theta_1 = 3$, and $\theta_2 = 7$ in the case of $N_T = 3$ and $RA = 1$ bit/(channel use). Once $\theta_m$, $u_m$ and $\tilde{l}_p$ are obtained, the magnitude ratio $\rho$, $\rho > 1$, can be found by

$$\rho = \arg\min_{\rho > 1} \left\{ \max_{\alpha, \beta, \alpha \neq \beta} P_e(\alpha, \beta) \right\}.$$  \hspace{1cm} (5.3)

Table 5.2 lists the optimized parameters using the two described methods. For $N_T = 2$, $R = 3, 4$ bits/(channel use), the parameters are obtained by reduced-space-exhaustive-searching when $E_b/N_0 = 20$ dB. For the remaining cases, we borrow the optimal parameters $u_m$ and $l_a$ from DUSTM and then find $\rho$, according to Eq. (5.3). We use the parameters in Table 5.2 for our simulations and analytical calculations in the following.

Table 5.2: DAPSTM parameters for $N_T = 2$ and $N_T = 3$.

<table>
<thead>
<tr>
<th>$N_T=2$</th>
<th>$N_T=3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R=3$ ($RP = 2$, $RA = 1$)</td>
<td>$\rho$</td>
</tr>
<tr>
<td>1.5</td>
<td>5</td>
</tr>
<tr>
<td>$R=4$ ($RP = 3$, $RA = 1$)</td>
<td>1.4</td>
</tr>
<tr>
<td>$R=5$ ($RP = 4$, $RA = 1$)</td>
<td>1.2</td>
</tr>
<tr>
<td>$R=6$ ($RP = 5$, $RA = 1$)</td>
<td>1.2</td>
</tr>
</tbody>
</table>

5.3 Space Diversity

In this section, we simulate the system with $B_jT = 0.001$, without making special claims. As mentioned in the last section, the parameters for DAPSTM are adopted from Table 5.2. For DUSTM, the parameters with "*" in Tables 2.1 and 2.2 are drawn for simulation.

Fig. 5.3 shows the simulation results of DAPSTM and DUSTM with $N_T = 2$ for $R = 3$
bits/(channel use) and $R = 4$ bits/(channel use). DAPSTM is detected by SH-DD and ML-DD, while DUSTM is only detected by ML-DD. At the rate of $R = 3$ bits/(channel use), the performance of DAPSTM with SH-DD is worse than that of DUSTM, while DAPSTM with ML-DD performs slightly better than DUSTM. When $R = 4$ bits/(channel use), DAPSTM with both SH-DD and ML-DD yields better performance than DUSTM. DAPSTM with ML-DD outperforms DUSTM by 3.3 dB at a BER = $10^{-3}$. For DAPSTM, ML-DD always outperforms SH-DD; therefore, we use ML-DD in our simulation for DAPSTM.

Figure 5.3: BER vs. $10\log_{10}(E_b/N_0)$ for DAPSTM and DUSTM with $R = 3$ bits/(channel use) and $R = 4$ bits/(channel use). $N_T = 2$, $N_R = 1$ and $B/T = 0.001$ are valid. DAPSTM is detected by SH-DD and ML-DD; DUSTM is only detected by ML-DD.

Fig. 5.4 shows the simulation results of MSD and DF-DD for DAPSTM with $N_T = 2$
at four different data rates ($R = 3, 4, 5, 6$ bits/(channel use)). The results of MSD and DF-DD ($N = 3$) are compared with those of ML-DD ($N = 2$) at different fading velocities, including (a) $B_f T = 0.005$, (b) $B_f T = 0.005$, (c) $B_f T = 0.0035$, and (d) $B_f T = 0.0025$. It is shown in Fig. 5.4(a) that DAPSTM detected by DF-DD and MSD performs almost equally well and can take full advantage of the enhanced diversity provided by multiple transmit antennas even for fast fading. All the figures in Fig. 5.4 show that DF-DD with observation window $N = 3$ can yield a significant gain over conventional DD at high $E_b/N_0$. Also in the $E_b/N_0$ range of interest an error floor can be avoided, even though both yield almost the same performance at lower $E_b/N_0$s, $E_b/N_0 \leq 30$ dB. It is worth mentioning
that the computational complexity of MSD is much higher than that of DF-DD. These observations are also in accordance with the results for DUSTM, cf.[11].

In order to illustrate the diversity gain of DAPSTM for different numbers of transmit antennas, Fig. 5.5 shows the performance for $R = 4$ bits/(channel use) ($RA = 1$ bit/(channel use), $RP = 3$ bits/(channel use)) with $N_T = 1$, $N_T = 2$, and $N_T = 3$. We observe that if $E_b/N_0$ is low, DAPSTM with three transmit antennas yields the worst performance of the three schemes, while a single antenna has the best performance. With increasing $E_b/N_0$, DAPSTM with multiple antennas starts to benefit from the diversity gain. The performance

![Figure 5.5: BER vs. $10\log_{10}(E_b/N_0)$ for DAPSTM with $N_T = 1$, $N_T = 2$, and $N_T = 3$. $N_R = 1$, $B_f T = 0.001$, $R = 4$ bits/(channel use) ($RA = 1$ bit/(channel use), $RP = 3$ bits/(channel use)), $\rho = 1.4$, $u_1 = 23$, $\bar{p} = 3$ are used.](image-url)
for two transmit antennas surpasses that for the single antenna at $E_b/N_0 = 18$ dB, while DAPSTM with three antennas yields the best performance at $E_b/N_0 = 23$ dB. Obviously, at high $E_b/N_0$ BER decreases as the number of transmit antenna increases. DAPSTM is especially effective at high $E_b/N_0$. This is consistent with the results in [5], where unitary space-time signals are especially effective at high $E_b/N_0$s.

Fig. 5.6 compares the performance of DAPSTM and DUSTM with various rates for $N_T = 2, 3$. It is observed that DAPSTM outperforms DUSTM when $R \geq 3$ bits/(channel use). The gap between DAPSTM and DUSTM increases with increasing the data rate and

![Figure 5.6: BER vs. $10\log(E_b/N_0)$ for DAPSTM and DUSTM with (a) $N_T = 2$, (b) $N_T = 3$. $N_R = 1$, $B_fT = 0.001$ are valid.](image)

the number of transmit antennas. Specifically for $N_T = 2$, we compare the two schemes
from $R = 4$ bits/(channel use) to $R = 6$ bits/(channel use). At a BER = $10^{-2}$, DAPSTM outperforms DUSTM by 3 dB for $R = 4$ bits/(channel use) to 10.4 dB for $R = 6$ bits/(channel use). For $N_T = 3$, when the data rate increases from $R = 4$ bits/(channel use) to $R = 5$ bits/(channel use), DAPSTM only degrades slightly, while DUSTM degrades by over 10 dB at a BER = $10^{-3}$. Since all signals in DUSTM are restricted to diagonal unitary matrices, with increasing data rates, the minimum distance between two arbitrary signals becomes smaller, which leads to poorer performance. With DAPSTM we take advantage of amplitude so that the signals are not restricted to unitary matrices. Its minimum signal distance, which determines the BER, is larger than that of DUSTM, given a fixed data rate. This is the major difference between our scheme and DUSTM, which allows us to increase the spectral efficiency by carrying information, not only on the phase, but also on the amplitude of a data matrix. The advantage of DAPSTM over DUSTM is more pronounced at higher data rates.

Fig. 5.7 compares the performance of DAPSTM and STC-DAPSK with $N_T = 2$. For 96STC-DAPSK with $R = 6$ bits/(channel use), two independent 64PSK and 32PSK constellations are used and the amplitude ratio $\rho = 1.2$. For 48STC-DAPSK with $R = 5$ bits/(channel use), two independent 32PSK and 16PSK constellations are used and $\rho = 1.4$. For STC-DAPSK with $R = 4$ bits/(channel use), two independent 16PSK and 8PSK constellations are adopted and $\rho = 1.5$. Since our proposed DAPSM scheme is based on a diagonal signal, it is not as power efficient as a nondiagonal signal constellation. We can see that the performance of DAPSTM is worse than STC-DAPSK when $N_T = 2$. However, the gap between DAPSTM and STC-DAPSK becomes smaller with increasing the data rate. At BER = $10^{-3}$, the gap decreases from 2.0 dB for $R = 4$ bits/(channel use) to 1.4 dB for $R = 6$ bits/(channel use). We expect that the gap to be narrower and for DAPSTM to surpass STC-DAPSK at a certain higher data rate. Since STC-DAPSK extends Alamouti’s OSTC [12] from phase modulation to combined differential amplitude and phase modula-
tion, we conjecture that STC-DAPSK cannot achieve full diversity and full rate when more than two transmit antennas are employed. DAPSTM can achieve full diversity and full rate for an arbitrary number of transmit antennas. As shown subsequently, DAPSTM can be used to exploit time diversity when only one transmit antenna is employed. STC-DAPSK, however, cannot due to its nondiagonal structure.

Figure 5.7: BER vs. $10 \log_{10}(E_b/N_0)$ for DAPSTM vs. STC-DAPSK. $N_T = 2$, $N_R = 1$, and $B_f T = 0.001$ are valid.

5.4 Time Diversity

In this section, interleaver length $N_I = 250$ is adopted to achieve an uncorrelated diversity branch in time diversity (TD) and ML-DD is applied for all the simulations. In Fig. 5.8,
two different diversity schemes with $N_T = 2$ for space diversity (SD), and $N_T = 1$ and $N_B = 2$ for TD are compared for independent diversity branches. Fig.5.8, (a), (b), and (c) shows BER vs. $10\log_{10}(E_b/N_0)$ for $R = 4, 5, 6$ bits/(channel use), with different fading velocities, respectively. Clearly the major difference between SD and TD is the error floor. TD leads to a lower error floor than SD. For $R = 4$ bits/(channel), $B_fT = 0.005$, SD yields an error floor of $BER = 7.7 \times 10^{-4}$ at $E_b/N_0 = 40$ dB, while TD significantly decreases the error floor to $7 \times 10^{-5}$. The same applies to $R = 5$ and 6 bits/(channel use). The effective fading bandwidth in TD is $B_fT$, which is lower than $N_TB_fT$ in SD [16]. Since the error floor increases with increasing fading bandwidth, a better performance can be achieved if

![Figure 5.8: BER vs. $10\log_{10}(E_b/N_0)$ for space diversity with $N_T = 2$ and time diversity with $N_B = 2$. (a) $R = 4$ bits/(channel use), (b) $R = 5$ bits/(channel use), (c) $R = 6$ bits/(channel use) are valid.](image-url)
TD is exploited instead of SD, especially for fast fadings.

Figure 5.9: BER vs. $10\log_{10}(E_b/N_0)$ for time diversity with $N_B = 3$ and space diversity with $N_T = 3$. $R = 4$ bits/(channel use) is valid.

Similar observations as can be made from Fig. 5.8 can be made from Fig. 5.9, which is valid for $N_T = 3$ and $N_B = 3$. The performances of SD and TD with $R = 4$ bits/(channel use) are illustrated. SD exhibits error floors of BER = 0.18 for $B_f T = 0.02$ and at BER = $3.3 \times 10^{-3}$ for $B_f T = 0.005$. TD performs better than SD, resulting in lower error floors of BER = $8 \times 10^{-3}$ and BER = $10^{-5}$ for $B_f T = 0.02$ and $B_f T = 0.005$, respectively.
Chapter 6

Conclusions and Recommendations

6.1 Conclusions

We have presented a differential amplitude/phase space-time modulation scheme based on diagonal signals over flat Rayleigh fading channels when neither the transmitter nor the receiver has access to CSI. The proposed DAPSTM extends unitary space-time modulation and combines phase and amplitude modulation while keeping the group property. DAPSTM keeps the diagonal structure in both spatial and temporal dimensions and achieves full diversity and full rate.

We analyze the performance of DAPSTM with SH-DD. The performance of calculation results are consistent with the simulation results.

Compared with the existing DUSTM proposed by Hochwald and Sweldens [5], we relax the restriction of differential phase modulation for data symbols to achieve significant improvement in error performance for high data rates. We use the maximized PEP to optimize the parameter in our signal design.

Compared with the STC-DAPSK [8] proposed by Xia, we extend the STC-DAPSK from
two transmit antennas to any number of transmit antennas, while keeping full diversity and full rate to achieve high bandwidth efficiency. Because of the diagonal structure of both phase and amplitude, DAPSTM can be used to exploit the pure time diversity while only one transmit antenna is employed. STC-DAPSK, on the other hand, can achieve full diversity and full rate only when two transmit antennas are employed, and can only be used in space diversity. However, DAPSTM is not as power efficient as STC-DAPSK when two transmit antennas are employed.

Two corresponding DD schemes, SH-DD and ML-DD are derived for a conventional non-coherent receiver. To further reduce the loss of performance for fast fading, two improved noncoherent DD schemes, i.e., MSD and DF-DD, with lower complexity, are investigated in this thesis.

6.2 Recommendations for Future Work

The DAPSTM constellation design and validation have been presented in this thesis. However, there is still much work to be done before perfection is achieved. Recommendations for possible future work are as follows:

- Further reducing complexity of ML-DD.

- Optimizing design parameters $\rho$, $u_m$ and $\theta_m$ with lower complexity to achieve better performance.

- Combining of space and time diversity to yield high performance for a wide range of fading velocities.
Glossary

Operators

\( \text{argmax}\{ \cdot \} \)  \hspace{1em} \text{argument maximizing the expression in brackets}
\( \text{argmin}\{ \cdot \} \)  \hspace{1em} \text{argument minimizing the expression in brackets}
\( \text{diag}\{ \cdot \} \)  \hspace{1em} \text{diagonal matrix with diagonal entries of vector argument}
\( \mathcal{E}\{ \cdot \} \)  \hspace{1em} \text{expectation}
\( \Re\{x\}, \Im\{x\} \)  \hspace{1em} \text{real and imaginary part of } x
\( P_e\{ \cdot \} \)  \hspace{1em} \text{error probability}
\( [\cdot]^H \)  \hspace{1em} \text{Hermitian transpose}
\( [\cdot]^T \)  \hspace{1em} \text{transpose}
\( (\cdot)^* \)  \hspace{1em} \text{complex conjugate}
\( \| \cdot \|^2 \)  \hspace{1em} \text{the Frobenius norm}
\( | \cdot | \)  \hspace{1em} \text{magnitude of a complex number}

Sets

\( \mathcal{A} \)  \hspace{1em} \text{signal alphabet}
\( \mathbb{Z} \)  \hspace{1em} \text{integer numbers}
\( \mathbb{N} \)  \hspace{1em} \text{natural numbers}
\( \mathbb{R} \)  \hspace{1em} \text{real numbers}
\( \mathbb{R}^+ \) positive real numbers

**Constants**

\( j \) imaginary unit: \( j^2 = -1 \)

\( \pi \) the number pi \( \pi = 3.14159265358979 \)

\( e \) Euler number \( e = 2.718281828 \)

\( I_n \) \( n \times n \) identity matrix

\( O_n \) \( n \times n \) zero matrix

\( 0_{n\times1} \) \( n \)-dimensional all zero row vector

**Other Functions**

\( \det \{ \cdot \} \) determinant of a matrix

\( \text{tr} \{ \cdot \} \) trace of a matrix

\( \exp(\cdot) \) exponential function

\( J_0(\cdot) \) zeroth order Bessel function of the first kind

\( \ln(\cdot) \) logarithm to base \( e \)

\( \log_{10}(\cdot) \) logarithm to base 10

\( \tan(\cdot) \) tangent function

**Acronyms**

ACF Autocorrelation Function

AWGN Additive White Gaussian Noise

BER Bit Error Rate

CSI Channel State Information

DASK Differential Amplitude Shift Keying
**GLOSSARY**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAPSK</td>
<td>Differential Amplitude/Phase Shift Keying</td>
</tr>
<tr>
<td>DAPSTM</td>
<td>Differential Amplitude/Phase Space-Time Modulation</td>
</tr>
<tr>
<td>DD</td>
<td>Differential Detection</td>
</tr>
<tr>
<td>DF-DD</td>
<td>Decision-Feedback Differential Detection</td>
</tr>
<tr>
<td>DPSK</td>
<td>Differential Phase Shift Keying</td>
</tr>
<tr>
<td>DUSTM</td>
<td>Differential Unitary Space-Time Modulation</td>
</tr>
<tr>
<td>DSTBC</td>
<td>Differential Space-Time Block Coding</td>
</tr>
<tr>
<td>PEP</td>
<td>Pairwise Error Probability</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple-Input Multiple-Output</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum Likelihood</td>
</tr>
<tr>
<td>MSD</td>
<td>Multiple-Symbol Detection</td>
</tr>
<tr>
<td>i.i.d.</td>
<td>Independent and Identically Distributed</td>
</tr>
<tr>
<td>OSTC</td>
<td>Orthogonal Space-Time Code</td>
</tr>
<tr>
<td>SH</td>
<td>Simple Heuristic</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>STC-DAPSK</td>
<td>Differential Amplitude/Phase Shift Keying</td>
</tr>
<tr>
<td>QAM</td>
<td>Quadrature Amplitude Modulation</td>
</tr>
</tbody>
</table>
Bibliography


