Title
Effect of noisy channel estimates on the performance of convolutionally coded systems with transmit diversity

Permalink
https://escholarship.org/uc/item/184934jv

Author
Jootar, Jittra

Publication Date
2006

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Effect of noisy channel estimates on the performance of convolutionally coded systems with transmit diversity

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Electrical Engineering (Communications Theory and Systems) by Jittra Jootar

Committee in charge:

Professor James R. Zeidler, Chair
Professor William S. Hodgkiss
Professor William A. Kuperman
Professor Laurence B. Milstein
Professor John G. Proakis

2006
The dissertation of Jittra Jootar is approved, and it is acceptable in quality and form for publication on microfilm:

______________________________________________

______________________________________________

______________________________________________

______________________________________________

______________________________________________

Chair

University of California, San Diego

2006
To my family for their love and support.
# TABLE OF CONTENTS

Signature Page ................................................................. iii

Dedication Page ................................................................. iv

Table of Contents ............................................................... v

List of Figures ................................................................. viii

Acknowledgements ............................................................. x

Vita and Publications ........................................................ xii

Abstract of the Dissertation ............................................... xiv

1 Introduction ................................................................. 1

2 Performance of convolutional codes with finite-depth interleaving and noisy channel estimates ........................................... 8
   2.1 Introduction ............................................................ 8
   2.2 System model .......................................................... 10
      2.2.1 Transmitter ...................................................... 10
      2.2.2 Channel ......................................................... 12
      2.2.3 Receiver ........................................................ 12
   2.3 Analysis ............................................................... 13
      2.3.1 Characteristic function ....................................... 14
      2.3.2 The Chernoff bound ........................................... 15
      2.3.3 The pairwise error probability ............................. 16
      2.3.4 Verification for perfect CSI ................................ 16
   2.4 Numerical Results ................................................... 19
      2.4.1 The optimal normalized Doppler frequency ............... 19
      2.4.2 Comparing the effects of the pilot SNR on systems in fast and slowly fading channels ..................................... 21
      2.4.3 Improving the performance through the interleaving depth .... 23
      2.4.4 Filter choice ................................................... 24
3 Performance of concatenated convolutional code and Alamouti space-time code with noisy channel estimates and finite-depth interleaving 32

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>32</td>
</tr>
<tr>
<td>3.2</td>
<td>Characteristic function, residue theorem and Gauss-Chebyshev approximation</td>
<td>34</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Characteristic function of a Hermitian quadratic form of a complex Gaussian random vector</td>
<td>35</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Residue theorem</td>
<td>35</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Gauss-Chebyshev approximation</td>
<td>36</td>
</tr>
<tr>
<td>3.3</td>
<td>System model</td>
<td>36</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Transmitter</td>
<td>36</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Channels</td>
<td>38</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Receiver</td>
<td>38</td>
</tr>
<tr>
<td>3.4</td>
<td>Linear-combining scheme</td>
<td>39</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Alamouti space-time code with LC-STD</td>
<td>39</td>
</tr>
<tr>
<td>3.4.2</td>
<td>CCA with LC-STD and ML convolutional decoder</td>
<td>42</td>
</tr>
<tr>
<td>3.5</td>
<td>ML decoder</td>
<td>44</td>
</tr>
<tr>
<td>3.5.1</td>
<td>Alamouti space-time code with ML-STD</td>
<td>45</td>
</tr>
<tr>
<td>3.5.2</td>
<td>CCA with joint Alamouti and convolutional ML decoder</td>
<td>47</td>
</tr>
<tr>
<td>3.6</td>
<td>Multi-path</td>
<td>48</td>
</tr>
<tr>
<td>3.7</td>
<td>Numerical results</td>
<td>49</td>
</tr>
<tr>
<td>3.7.1</td>
<td>Alamouti space-time code</td>
<td>49</td>
</tr>
<tr>
<td>3.7.2</td>
<td>Concatenated convolutional code and Alamouti space-time code</td>
<td>54</td>
</tr>
<tr>
<td>3.8</td>
<td>Conclusions</td>
<td>58</td>
</tr>
<tr>
<td>3.9</td>
<td>Acknowledgement</td>
<td>59</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

1.1 Block diagram of an interleaved coded system ............................................ 1
1.2 Block interleaver ......................................................................................... 2
1.3 Block diagram of an interleaved coded system with the Alamouti space-time code ........................................................................................................... 5

2.1 System block diagram .................................................................................. 11
2.2 Pairwise error probability versus normalized Doppler frequency for different interleaving depths for Jakes PSD and Gaussian PSD. .................................................. 20
2.3 Pairwise error probability versus normalized Doppler frequency for realistic cases and ideal cases. ................................................................. 20
2.4 Pairwise error probability and Chernoff bounds versus pilot SNR for different normalized Doppler frequencies. ......................................................... 22
2.5 Pairwise error probability versus pilot SNR for different interleaving depths and normalized Doppler frequencies. ......................................................... 23
2.6 Pairwise error probability difference versus pilot SNR and normalized Doppler frequency for different channel estimators. .................................................. 24
2.7 Pairwise error probability versus normalized Doppler frequency for different interleaving depths and multi-path characteristics. ................................. 25
2.8 Pairwise error probability versus normalized Doppler frequency and data-to-pilot energy ratio for different multi-path characteristics. ................................. 26
2.9 Pairwise error probability versus pilot SNR for different Hamming weights, normalized Doppler frequency and Rician $K$ factor. ........................................... 27
2.10 Bit error probability and block error probability (Rayleigh) versus normalized Doppler frequency from Monte Carlo simulation and from the analysis. .................................................................................. 29
2.11 Bit error probability and block error probability (Rician) versus normalized Doppler frequency from Monte Carlo simulation and from the analysis. 30

3.1 System diagram of STTD over two-symbol period ...................................... 37
3.2 Illustration of an error codeword with Hamming weight 5 and 4-symbol interleaving depth ........................................................................................................... 43
3.3 Alamouti pairwise error probability versus pilot SNR for the LC-STD. ........ 50
3.4 Alamouti pairwise error probability versus pilot SNR for the ML-STD. ....... 50
3.5 3-D plot of the bit error probability for the SISO and Alamouti STC with the linear-combining STD. ................................................................. 51
3.6 The linear combining space-time decoder in the multi-path fading channels 51
3.7 Alamouti pairwise error probability versus pilot SNR for different normalized Doppler frequencies and channel estimators. 53
3.8 Performance of the linear combining space-time decoder using WCDMA parameters 53
3.9 Pairwise error probability versus normalized Doppler frequency for different antenna correlation factors assuming a large pilot SNR. 55
3.10 Pairwise error probability versus normalized Doppler frequency for different antenna correlation factors assuming a small pilot SNR. 55
3.11 Contour plot of the pairwise error probability versus normalized Doppler frequency and the data-to-pilot energy ratio. 56
3.12 Bit error probability and block error probability versus normalized Doppler frequency from Monte Carlo simulation and from the analysis. 58

4.1 Baseband representation of the CLTD system 63
4.2 Contour plot of the bit error probability for PA-CLTD and PO-CLTD. 72
4.3 Bit error probability versus data SNR for different pilot SNR’s. 73
4.4 Bit error probability versus lag time for Gaussian PSD and Jakes PSD for different filter lengths. 74
4.5 3-D plot of the bit error probability for SISO, Alamouti space-time code and PA-CLTD. 75
4.6 Contour plot of the pairwise error probability for different lag times . 77
4.7 Bit error probability and block error probability versus normalized Doppler frequency for PO-CLTD. 78
4.8 Bit error probability and block error probability versus normalized Doppler frequency for PA-CLTD. 78
4.9 Bit error probability versus normalized Doppler frequency for SISO, the Alamouti STC and PA-CLTD (5-symbol lag time) . 80
4.10 Block error probability versus normalized Doppler frequency for SISO, the Alamouti STC and PA-CLTD (5-symbol lag time) . 80
ACKNOWLEDGEMENTS

I am grateful to many people who have touched my life and made me the person I am today. First and foremost, I would like to thank my advisor, Professor James Zeidler, for his guidance and support throughout the duration of my graduate studies. I am also grateful to my other committee members, Professors John Proakis, Laurence Milstein, William Hodgkiss and William Kuperman, for their valuable time and interest in my thesis topic.

In addition, I would like to thank my fellow graduate students in Professors Zeidler and Proakis’s group for useful discussions and social interaction. Thanks to Steve Thompson and Arun Batra, who always provide me with useful technical discussion, Latex support and comical relieve. Thanks to Haichang Sui, whose technical expertise is always appreciated. Thanks to Yiannis Spyropoulos and Kostas Stamatiou for great conversations and lunches we have had over the years. Thanks to Adam Anderson for always volunteering to read my manuscripts.

A successful graduate life requires great love and moral support. I would like to express my sincere gratitude to my parents for bringing me up to be the person I am and for providing me with the best education I could have hoped for. Thanks to my older brother, Jay, for setting such a great example for me to follow. It is a privilege to have you as my brother. Thanks to my younger brother, Don, and his extended family for their support and encouragement. I would also like to thank my dear Bernd and Muzelpuz Meurers for their continued loving care and support. I cannot imagine how I could have accomplished this without you two.

I would also like to thank my friends at UCSD and elsewhere. All of you make this journey a pleasant one. Thanks to Cecile Levasseur, Elizabeth Gire, Azadeh Bozorgzadeh, Maziar Nezhad, Chandra Murthy and Periklis Liaskovitis, for keeping me sane during these stressful years. Thanks to Gundu and Bharat Raghunathan for their love and support. You have given me hope in the time I needed it most. Finally, I would like to thank my dear friend, Anand Subramaniam. We never got a chance to say goodbye. We only knew each other for a short time but I felt like I knew you forever. Wherever you are, Anand, I would like to let you know that you are always in my thoughts and prayers.
VITA

1976
Born, Nashville, Tennessee

1997
B. E. in Electrical Engineering
*Summa cum laude*, Chulalongkorn University

1999
M. S. in Electrical Engineering
Stanford University

1999–2003
Engineer/Senior Engineer
Qualcomm Inc.

2003–2006
Research assistant
Department of Electrical and Computer Engineering
University of California San Diego

Summer 2005
Summer intern
Texas Instrument Inc.

2006
Ph. D. in Electrical Engineering
University of California San Diego

PUBLICATIONS


ABSTRACT OF THE DISSERTATION

Effect of noisy channel estimates on the performance of convolutionally coded systems with transmit diversity

by

Jittra Jootar

Doctor of Philosophy in Electrical Engineering (Communications Theory and Systems)

University of California San Diego, 2006

Professor James R. Zeidler, Chair

This dissertation focuses on the effect of noisy channel estimates on the performance of finite-depth convolutionally coded system with different types of transmit diversity algorithms. Three transmit diversity algorithms analyzed are no transmit diversity, the Alamouti space-time code and the closed-loop transmit diversity.

The dissertation begins with the derivation of the theoretical pairwise error probability of a finite-depth convolutionally coded system without transmit diversity assuming that channel estimates are calculated by an FIR filter from noisy channel estimates. With the assumptions that an FIR filter is used as the channel estimator and the interleaving depth is finite, the estimation-diversity tradeoff resulting from the effects of the Doppler spread on the system performance via the channel estimation accuracy and the channel diversity can be investigated. In addition, this section also verifies that, in the limit when the channel estimates are perfect, the result presented in this dissertation is the same as the well-known result derived from a different mathematical approach. To verify the analysis, the analytical results are compared with results from Monte Carlo simulation and the comparison shows that the analytical results match well with the simulation results.

The dissertation continues with the analysis of the Alamouti space-time code with
and without convolutional codes in time-varying fading channels also assuming that channel estimates are noisy. Two types of receivers are investigated for the conventional Alamouti space-time code without convolutional codes, namely, the linear-combining space-time decoder (LC-STD) and the maximum-likelihood space-time decoder (ML-STD). Two types of receivers are investigated for the concatenated system, namely, the LC-STD with the ML convolutional decoder and the joint Alamouti and convolutional ML decoder. The results have shown that the LC-STD is more sensitive to the Doppler spread than the ML-STD. However, since the ML-STD is very sensitive to the channel estimation error, the gains provided by the decoder in fast fading channels will be offset unless an optimized channel estimator is employed. Performance comparisons between the Alamouti systems and the SISO systems indicate that, when the system environment is not ideal, the SISO systems may outperform the Alamouti systems. A comparison between the analytical results and the simulation results shows that the analysis can predict the simulation results accurately.

Finally, the closed-loop transmit diversity analyses are presented. The closed-form expressions for the uncoded bit error probability of closed-loop transmit diversity algorithms with two transmit antennas and noisy channel estimates in time-varying Rayleigh fading channels are derived. Two closed-loop transmit diversity algorithms considered are the phase-amplitude closed-loop transmit diversity (PA-CLTD), where the transmit antennas may transmit with different signal energy, and the phase-only closed-loop transmit diversity (PO-CLTD), where the transmit antennas must transmit with the same signal energy. The results have shown that PA-CLTD performs slightly better than PO-CLTD although PA-CLTD requires significantly more feedback information. Moreover, a comparison between PA-CLTD, the Alamouti space-time code and the SISO system indicates that PA-CLTD outperforms the other two systems when the Doppler spread is small and the pilot SNR is large. In addition to the uncoded bit error probability, this dissertation also derives the pairwise error probability when finite-depth interleaved convolutional codes are used with the closed-loop transmit diversity algorithms. The analytical results show that, when the Doppler spread is large, the performance of the closed-loop transmit diversity may degrade significantly. Finally, the analytical results are compared with results from Monte Carlo simulation and the comparison shows that the analytical results match well with the simulation results.

This work was done at UCSD’s California Institute for Telecommunications and
Information Technology (CalIT2), under the Core Grant No. 02-10109 sponsored by Ericsson and the U.S. Army Research Office under the Multi University Initiative (MURI) grant No. W911NF-04-1-0224.
Convolutional codes have long been used in communication systems as a way to improve the system performance. However, most well-known codes have been designed for an AWGN channel, where the signal-to-noise ratios (SNR) of the received signals within the same codeword are uncorrelated. When these codes are used in a wireless communication system where the channel undergoes time-varying fading, the received SNR is temporally correlated and the performance of convolutional codes degrades [12, 46]. As a result, interleaving has been used in most convolutionally coded wireless communication systems in order to mitigate the effect of temporal correlation as illustrated in fig. 1.1 [46].

The goal of the interleaver is to reduce the temporal correlation, hence to increase the time diversity, by reordering the coded bits such that the bits that are close to one another are transmitted far apart in time. At the receiver, the deinterleaver undoes the reordering done by the interleaver, therefore, the coded sequence is restored and is ready

![Figure 1.1 Block diagram of an interleaved coded system](image-url)
for the decoder. Since consecutive bits are transmitted far apart in time, the fading coefficients between consecutive bits are less correlated and the temporal correlation experienced at the decoder is reduced.

The interleaver under consideration in this dissertation is called a block interleaver, where the coded bits are written into a matrix form row by row and read out column by column [46] as illustrated in fig. 1.2. In this figure, it is assumed that the convolutional code is terminated and the block length is equal to 90 coded bits and the time indexes of the coded bits are indicated by the number in the matrix. After all the coded bits are written into the matrix (row by row), the output of the interleaver is generated by reading out from the matrix column by column. Therefore, the time indexes of the output sequence become [1, 11, 21, 31, 41, 51, 61, 71, 81, 2, 12, ..., 80, 90]. It is apparent that, because of the interleaving process, consecutive bits are at least 10 bits apart. This 10-bit spacing is equal to the number of columns and is often referred to as the interleaving depth.

As a result, the temporal correlation or the time diversity resulting from interleaving depends on two factors: the interleaving depth and the Doppler spread [48]. Ideally, an interleaver with an infinite interleaving depth can make the channel appear memoryless.

![Figure 1.2 Block interleaver](image-url)
at the receiver regardless of the Doppler spread. In reality, however, the interleaving depth is finite and the temporal correlation observed at the receiver is a function of the Doppler spread. As the interleaving depth decreases, a larger Doppler spread is required to achieve the maximum time diversity at the receiver [26, 48].

In addition to the temporal correlation, the Doppler spread also strongly affects the accuracy of the channel estimates. Since it is easier to estimate the fading coefficients when the channel changes slowly, a small Doppler spread is preferable over a large Doppler spread for accurate channel estimation [14].

Combining the effect of the Doppler spread on the channel memory and the channel estimates, it is known that increasing the Doppler spread improves the performance by lowering the channel memory but degrades the performance by increasing the channel estimation error [62, 66, 67]. Thus, the system performance in a practical system, where both the interleaving and the channel estimation are not perfect, is not trivial and system engineers often have to resort to lengthy simulation in order to find the error performance they should expect to see from the real system. Therefore, the goal of chapter 2 is to derive the pairwise error probability (PEP) of a convolutionally coded system with finite-depth interleaving and noisy channel state information (CSI) in a time-varying Rician fading channel assuming that the channel estimates are calculated by an FIR filter from noisy pilot signals [14]. Moreover, it is also proved in chapter 2 that, in the limit where CSI is perfect, our analysis provides the same result as the result from the perfect CSI analysis, which is derived by a different mathematical approach.

The results from chapter 2 describe many important system characteristics. For example, because of the tradeoff between diversity and estimation error, there exists an optimal Doppler spread that provides the best combination of the time diversity and the quality of the channel estimates. Increasing the Doppler spread beyond the optimal value causes the channel estimation error to increase to the point where the degradation from CSI error outweighs the performance gain from diversity. On the other hand, decreasing the Doppler spread lower than the optimal value reduces the time diversity such that the performance degradation caused by diversity reduction outweighs the gain from better channel estimation accuracy. Moreover, investigating the PEP performance as a function of the Rician K factor, it can be seen that increasing the Rician K factor does not always improve the performance since increasing the energy in the line-of-sight
implies decreasing the energy in the diffuse component, hence decreasing the gain from diversity.

Another method that has recently gained much interest as a way to improve the system performance above what is achievable by conventional error correcting codes alone is transmit diversity (TD). Numerous TD algorithms have emerged over the last decade and several have been included in wireless communication standards [2, 55]. Transmit diversity algorithms can be categorized into two types based on the requirement of the feedback information at the transmitter; namely, the open-loop TD (OLTD) and the closed-loop TD (CLTD). As the names suggest, the OLTD algorithms are the TD algorithms that the transmitters do not require any feedback information from the receivers and the closed-loop TD algorithms are the TD algorithms that the transmitters require feedback information from the receivers. Although there have been much interest on the performance analyses of complex TD systems with a large number of transmit antennas, in this dissertation, only TD systems with two transmit antennas are studied. In chapter 3, the OLTD system called the Alamouti space-time code (STC) is analyzed [7], and in chapter 4, two CLTD algorithms, namely, the PA-CLTD and the PO-CLTD, are analyzed. In both CLTD systems, the data signals are multiplied by complex weights (different weights for different antennas) prior to the transmission, and the received data SNR is maximized when the signals from different transmit antennas are combined in-phase at the receiver. The difference between PA-CLTD and PO-CLTD is that the transmit antennas in a PA-CLTD system may transmit with different signal energy (complex weights may have different magnitudes) while the transmit antennas in a PO-CLTD system must transmit with the same signal energy (complex weights must have the same magnitude).

Various works have been done to analyze the performance of the Alamouti STC. In his original paper [7], Alamouti derived the performance of the Alamouti STC with the linear-combining space-time decoder (LC-STD) assuming that the receiver has the perfect knowledge of the channel state information (CSI) and the channels are quasi-static over a two-symbol period. The low complexity LC-STD has been designed such that, under the perfect CSI and quasi-static fading assumptions, it can completely eliminate the interference from the other symbol in the codeword and achieve the same performance as the more complex ML space-time decoder (ML-STD) [7, 59]. However, when the channels are not quasi-static or when the CSI is not perfect, the LC-STD cannot
completely eliminate the interference from the other symbol in the codeword causing performance degradation [9, 13, 16, 27, 28, 49, 49, 53, 58, 65]. Moreover, the performance of the Alamouti STC can also be further degraded by spatial correlation between the transmit antennas [34, 58, 60]. Although numerous works have been done to analyze the effects of noisy CSI, non quasi-static fading channels and spatial correlation on the performance of the Alamouti STC. No one has defined the combined effect of all three system imperfections. The goal of chapter 3 is to derive the closed-form expression of the theoretical error performance of the Alamouti STC taking into account all of these imperfections. This dissertation focuses on two space-time decoder configurations; namely, the low-complexity LC-STD and the high-complexity ML-STD, assuming that the channel estimates used by these space-time decoders are calculated by FIR filters from noisy pilot symbols [14] and that the channels are time-varying Rayleigh fading channels.

In addition to the conventional Alamouti STC, this dissertation also analyzes a system employing both the Alamouti STC and the convolutional codes as illustrated in fig. 1.3. Two types of decoders analyzed in this work are the LC-STD with the ML convolutional decoder (LC-ML) and the joint Alamouti space-time and convolutional ML decoder (JML). The LC-ML is a simple decoder that is often used in practice, while
the JML is a decoder that is too complex to be implemented in practice. However, a comparison between the error performance of these two decoders helps quantify the performance loss from using the simple LC-ML instead of the complex JML. Similar to the analysis of the conventional Alamouti STC, it is assumed that the channel estimates are calculated by FIR filters from noisy pilot channels and that the channels are not quasi-static. In addition, similar to the assumption made in chapter 2, the interleaving depth is assumed to be finite.

For the conventional Alamouti STC without convolutional codes, the results indicate that the performance of the LC-STD degrades significantly with the Doppler spread, while the ML-STD appears to be more tolerant to the Doppler spread. However, comparing different types of channel estimators, the results indicate that the ML-STD is very sensitive to CSI error. Therefore, in a system where the channel estimators are not optimal, the degradation from using the simple LC-STD compared to the complex ML-STD is reduced. Comparing the PEP for the concatenated convolutional and Alamouti code as a function of the Doppler spread between the two decoders, it can be seen that the gain from using the high complexity JML is marginal at high Doppler spread and negligible at small Doppler spread. These results indicate that implementing the complex JML may not result in significant performance gains. Investigating the error performance when the two transmit antennas are correlated, the performance degrades with the spatial correlation as expected. It can also be seen from the numerical results that, when the CSI accuracy is poor, the concatenated convolutional and Alamouti code may be outperformed by the convolutional code without transmit diversity.

Various studies have also been done to analyze the theoretical performance of uncoded CLTD systems. The system performance of uncoded CLTD depends on many system parameters, such as, the feedback quantization error [18], the feedback error [44], the feedback rate [24], the lag (or delay) time [18, 19, 24, 36, 43, 47] and the channel estimation error [15, 36, 58]. In general, it is difficult to analyze all the above parameters jointly. Consequently, chapter 4 focuses on the effects of the channel estimation error and the lag time on the performance of PA-CLTD and PO-CLTD, while assuming the rest to be perfect. The importance of these two parameters will be shown by the analysis that, although the rest of parameters are assumed to be perfect, the imperfections of CSI and the lag time can cause the CLTD systems to degrade considerably such that it is outperformed by a system without transmit diversity. In this chapter, the channels
are assumed to be time-varying Rayleigh fading channels and the transmit antennas are assumed to be uncorrelated. Unlike the analyses presented in [15, 58], which derived approximations of the bit error probability of the PA-CLTD system, and [36], which derived the bit error probability of the PA-CLTD system, this chapter assumes that the channel estimates are calculated by FIR filters from noisy pilot symbols. Therefore, the effects of the filter coefficients and the Doppler spread (via the CSI accuracy) on the bit error performance of the PA-CLTD and the PO-CLTD systems are captured. Also note that, in addition to demodulation, the accuracy of CSI in the CLTD systems also affects the performance of CLTD via the complex weights. Finally, in addition to the uncoded CLTD systems, chapter 4 also analyzes the performance of the PA-CLTD or the PO-CLTD systems with finite-depth convolutional codes and noisy CSI in time-varying Rayleigh fading channels.

The results from chapter 4 indicate that, although PO-CLTD requires much less feedback information than PA-CLTD, the performance loss is not significant. Considering the complexity of the feedback quantization and the signalling in a practical system, this finding suggests that the performance gain versus complexity tradeoffs between these two approaches make PO-CLTD much more attractive than PA-CLTD. Using the results from chapters 2, 3 and 4, the performance of these three systems (no transmit diversity, the Alamouti TD and the CLTD) with noisy CSI and with or without finite-depth convolutional codes can be compared. The results show that CLTD systems are the most attractive when the CSI accuracy is good and the Doppler spread is small. When the Doppler spread increases, the Alamouti STC code becomes more attractive. However, when both the CSI accuracy is poor and the Doppler spread is large, the system without transmit diversity is the most robust.
Performance of convolutional codes with finite-depth interleaving and noisy channel estimates

2.1 Introduction

Previous analytical studies on the performance of convolutional codes in a time-varying fading channel have focused on either imperfect channel state information (CSI) or finite-depth interleaving (FID), while assuming the other to be perfect [20, 26, 35, 48]. Since fading coefficients can be estimated with better accuracy in a slowly fading channel, a perfectly interleaved system with noisy CSI in a slowly fading channel outperforms the system in a fast fading channel. However, when the CSI is perfect but the interleaving is imperfect due to finite interleaving depth, the performance is reversed, i.e., the system in a fast fading channel outperforms the system in a slowly fading channel [26, 35, 48]. This is because the number of independent fading realizations available for a codeword, the number referred to as the channel diversity [67], of a fast fading channel is greater than that of a slowly fading channel.

In a practical system where both CSI and interleaving are not perfect and the im-
perfections contribute to the performance degradation of the system, the performance analysis has to take into account both imperfections. Since increasing the Doppler spread improves the system performance by increasing the channel diversity but degrades the performance by worsening the channel estimation accuracy, the system performance is expected to display the estimation-diversity tradeoff as a function of the Doppler spread when both imperfect CSI and imperfect interleaving are considered [6,63,66,67].

In order to address system performance in realistic operating environment, there has recently been growing interest in the performance analysis of coded systems with imperfect CSI and imperfect interleaving. However, earlier analyses did not model the CSI accuracy as a function of the Doppler spread [54], or used simple assumptions such as non-interleaved codes [42], or discussed the tradeoff from simulation results without providing any analytical analysis [6]. The analysis on the estimation-diversity tradeoff was first presented in [63], where the authors derived the optimal memory lengths and the error exponent bounds for joint estimation and decoding assuming a block fading channel. The block fading assumption was also used in later works [66,67], where the pairwise error probability (PEP) for coded systems in a Rayleigh fading channel and a Rician fading channel was derived. Later, in [38], the block fading assumption was replaced with a more general channel model. However, the authors used the assumption that the noise components, after multiplying the received signals with the conjugate of noisy channel estimates, are Gaussian random variables. This assumption caused the analytical results in [38] to be just approximate, not the exact performance.

Because existing analyses are limited to specific assumptions such as block fading [63,66,67], which is not an accurate assumption for several wireless systems, or Gaussian noise component [38], the primary focus of this chapter is to derive the PEP with the assumptions that are more general and the model that incorporates implementation issues such as the choice of the pilot filter. Consequently, the system performance in a realistic scenario can be calculated from the analysis without having to resort to lengthy simulations, allowing optimization studies of various design parameters, such as the pilot filter coefficients, the interleaving depth and the pilot-to-signal power ratio. It should be noted that the analysis presented in this chapter builds mainly upon the work on imperfect CSI by Cavers [14] and the work on non-interleaved codes with imperfect CSI by Nobelen and Taylor [42]. In addition, because the mathematical model of this system is similar to the mathematical model of the correlated maximal ratio combining (MRC)
system, the method used here also resembles the ones used in [23,37,52].

The chapter is organized as follows. Section 2.2 introduces the system model, which includes the transmitter at the base station, the frequency-selective Rician fading channel and the receiver at the mobile unit. In section 2.3, the Chernoff bound of the pairwise error probability and the exact pairwise error probability of the system are derived. In addition, section 2.3 also verifies that, for a special case when CSI is perfect, the analysis agrees with the existing perfect CSI analysis. Discussions of the results and conclusions are presented in section 2.4 and section 2.5, respectively. Finally, section 2.6 addresses the acknowledgement.

2.2 System model

For the rest of this chapter, the following notation will be used. A lowercase bold letter denotes a vector and an uppercase bold letter denotes a matrix. The element in the $m^{th}$ row and the $n^{th}$ column of a matrix $X$ is denoted by $X(m,n)$ and the element in the $m^{th}$ row (column) of a column (row) vector $x$ is denoted by $x(m)$. The superscripts $\ast, T, H$ denote the complex conjugate, the matrix transpose and the matrix Hermitian operation, respectively. The determinant of a matrix $X$ is denoted by $|X|$. The length $m$ column vector of ones, the square identity matrix, the square zero matrix and the anti-diagonal matrix of order $m$ are denoted by $1_m, I_m, 0_m$ and $J_m$, respectively.

The system considered is a downlink BPSK DS-CDMA system. A complex baseband representation of the system is illustrated in fig. 2.1, where the base station’s transmitter, the frequency-selective fading channel and the mobile’s receiver are shown in the upper left corner, the upper right corner and the bottom section of the figure, respectively.

2.2.1 Transmitter

It is assumed that there are $K + 1$ signal streams transmitted from the base station. The $K + 1$ streams consist of one pilot stream ($0^{th}$ stream) and $K$ data streams assigned to $K$ users ($k^{th}$ stream for the $k^{th}$ user). The pilot stream $c_0[\lfloor n/N \rfloor] = 1$ is spread with the orthogonal code $w_0[n]$, where $n$ denotes the chip time index and $N$ denotes the period or the spreading gain of $w_0[n]$. Similarly, the $k^{th}$ BPSK data stream $c_k[\lfloor n/N \rfloor]$, which is
The interleaved convolutionally coded BPSK signal, is spread with the orthogonal code $w_k[n]$, which has the same period $N$. After spreading, the $k^{th}$ signal stream is scaled by $\sqrt{E_{c,k}}$, and the signals from all branches are combined and scrambled by the base station dependent complex long code $\tilde{w}[n]$. The signal after scrambling can be expressed as

$$u[n] = \tilde{w}[n] \sum_{k=0}^{K} \sqrt{E_{c,k}} w_k[n] c_k[[n/N]]. \quad (2.1)$$

The signal $u[n]$ is then passed through an impulse modulator and a pulse-shaping filter with frequency response $H(f)$, where $X(f) = |H(f)|^2$ satisfies the Nyquist condition for zero inter-symbol interference (ISI) [46], i.e.,

$$\mathcal{F}^{-1}\{X(f)\} = x(t) \begin{cases} 1, & \text{when } t = 0; \\ 0, & \text{when } t = m\tau_c \text{ and } m \text{ is any non-zero integer.} \end{cases} \quad (2.2)$$

where $\mathcal{F}^{-1}\{X(f)\}$ denotes the inverse Fourier transform of $X(f)$ and $\tau_c$ denotes the
chip period. Examples of functions with these properties are functions in the family of square-root raised cosine pulse-shaping filters [46].

2.2.2 Channel

The channel is assumed to be a time-varying frequency selective Rician fading channel with $L$ resolvable paths and impulse response

$$h(t) = \sum_{l=1}^{L} \alpha_l(t) \delta(t - \beta_l \tau_c),$$

(2.3)

where $\alpha_l(t)$ denotes the fading coefficient corresponding to the $l^{th}$ path, $\delta(t)$ denotes the dirac delta function and $\beta_l \tau_c$ denotes the delay associated with the $l^{th}$ path. It is also assumed that the $\beta_l$’s are integers, i.e., the delays are multiples of the chip period and that $\beta_1 < \beta_2 < \ldots < \beta_L$.

The fading coefficient $\alpha_l(t)$ is assumed to be a circularly symmetric complex Gaussian random variable with real-valued mean $E[\alpha_l(t)] = \mu_l$ and autocovariance function

$$1/2 E[(\alpha_l(t) - \mu_l)(\alpha_l(t - \tilde{\tau}) - \mu_l)^*] = \sigma_{\text{rl}}^2 R_l(\tilde{\tau}).$$

Note that $\mu_l^2$ represents the energy of the line-of-sight component of the fading channel and $\sigma_{\text{rl}}^2 R_l(\tilde{\tau})$ represents the autocorrelation function of the diffuse component of the fading channel. In addition, it is assumed that the fading coefficients from different paths are independent; thus, $E[(\alpha_{l_1}(t) - \mu_{l_1})(\alpha_{l_2}(t - \tilde{\tau}) - \mu_{l_2})^*] = 0$, when $l_1 \neq l_2$. An analysis for the Rayleigh fading channel is simply a special case when $\mu_l = 0$. Although the channel is time-varying, this analysis is restricted to the case when the channel changes slowly enough that the fading coefficients appear to be constant over one symbol period $\tau = N \tau_c$.

Finally, the thermal noise $n_w(t)$ is assumed to be additive white Gaussian noise (AWGN) with variance $\sigma_w^2$.

2.2.3 Receiver

At the mobile’s receiver of user 1, the received signal $r(t)$ can be expressed as

$$r(t) = \sum_{l=1}^{L} \alpha_l(t) \sum_{n=-\infty}^{\infty} u[n] h(t - n \tau_c - \beta_l \tau_c) + n_w(t).$$

(2.4)
After despreading, the pilot signal and the data signal at the output of the accumulator corresponding to the \(l\)th branch of the RAKE receiver can be approximated as \[21\]

\[
r_{p,l}[m] = \sqrt{E_p} \alpha_l (\beta_l \tau_c + m \tau) + n_{p,l}[m]
\]

\[
r_{s,l}[m] = \sqrt{E_s} c_1[m] \alpha_l (\beta_l \tau_c + m \tau) + n_{s,l}[m],
\]

where \(E_p = E_{c,0} N^2\), \(E_s = E_{c,1} N^2\), and \(n_{p,l}[m]\) and \(n_{s,l}[m]\) denote the summation of the self-noise and the thermal noise components of the pilot and the data signals, respectively. Conditioned on the transmitted data, the self-noise and the thermal noise components can be approximated as zero-mean Gaussian random variables with variances \(N \tilde{\sigma}_{c,l}^2 E_t\) and \(N \sigma_w^2\), respectively, where \(\tilde{\sigma}_{c,l}^2 = \sum_{i=1,i \neq l}^{L} (\sigma_{c,i}^2 + \mu_i^2 / 2)\) and \(E_t = \sum_{k=0}^{K} E_{c,k}\) \[21\].

Thus, the variances of both \(n_{p,l}[m]\) and \(n_{s,l}[m]\) are equal to \(\sigma_{n,l}^2 = N (\tilde{\sigma}_{c,l}^2 E_t + \sigma_w^2)\). For simplicity, this symbol rate model will be used for the rest of the chapter.

The channel estimator for the \(l\)th branch of the RAKE receiver \[46\] is assumed to be a \((2M + 1)\)-tap FIR filter with the filter coefficient vector \(h_l = [h_l[M] \ldots h_l[0] \ldots h_l[-M]]^T\).

As a result, the channel estimate can be written as

\[
\hat{\alpha}_l[m] = h_l^H r_{p,l}[m] = \sum_{i=-M}^{M} h_l^* [i] r_{p,l}[m - i],
\]

where \(r_{p,l}[m] = [r_{p,l}[m - M] \ldots r_{p,l}[m] \ldots r_{p,l}[m + M]]^T\).

### 2.3 Analysis

Without loss of generality, it is assumed that the transmitted codeword is an all-zero codeword, which is mapped to an all-one BPSK sequence \(c_0\). Due to interleaving, the pairwise error probability (PEP), which is the probability that the decoder chooses the coded sequence \(c_i (i \neq 0)\) instead of \(c_0\) given that \(c_0\) was transmitted, is a function of \(c_i\) and the structure of the interleaver. Finding the PEP for each error pattern with respect to a specific interleaver is tedious and adds little insight into the overall system performance \[48\]. Therefore, the approximation that an interleaving depth \(I\) of a block interleaver creates the same effect as separating consecutive symbol errors by \(I\) symbols will be used throughout this chapter \[26\]. As a result, the PEP can be simplified such that it depends only on the Hamming weight of the error codeword but not the structure of the interleaver nor the error codeword itself. Also note that this...
approximation is used here mainly to simplify the analysis and an extension to the case without this approximation can be done straightforwardly. The PEP, when a RAKE receiver is used with a mismatched ML (Viterbi) decoder (see [56] and references therein for the description of the mismatch decoder), can be expressed as

$$P_2 = \Pr \{ z < 0 \} = \int_{-\infty}^{0} p_z(x) \, dx,$$

(2.8)

where $P_2$ denotes the PEP, $z = \sum_{l=1}^{L} z_l$, $z_l = \Re \left[ \sum_{i=1}^{d} r_{s,i}[I] \hat{\alpha}_l[I] \right]$, $I$ is the interleaving depth, $d$ is the Hamming weight of the error codeword, $\Re[x]$ denotes the real part of the complex number $x$, and $p_z(x)$ is the probability density function of $z$.

2.3.1 Characteristic function

Following the approach used in [14], $z_l$ can be written in a quadratic form of complex Gaussian random variables $x_l$, i.e., $x_l^H Q x_l$, where $x_l = [r_{s,i}[I] \ldots r_{s,i}[dI] \hat{\alpha}_l[I] \ldots \hat{\alpha}_l[dI]]^T$ and

$$Q = \frac{1}{2} \begin{bmatrix} 0_d & I_d \\ I_d & 0_d \end{bmatrix}.$$  

(2.9)

The characteristic function of the Hermitian quadratic form of complex Gaussian random variables was derived by Turin to be [57]

$$\Phi_{z_l}(s) = \exp \left\{ -\frac{1}{2} \bar{x}_l^H \Sigma_l^{-1} [I_{2d} - (I_{2d} - 2s \Sigma_l Q)^{-1}] \bar{x}_l \right\},$$

(2.10)

where $\bar{x}_l$ and $\Sigma_l$ are the mean vector and the covariance matrix of $x_l$, respectively. Due to the assumption that the fading coefficients from different paths are independent, the characteristic function of $z$ is simply the product of the characteristic functions of $z_l$ for $l = 1, \ldots, L$. Thus, the characteristic function of $z$ becomes

$$\Phi_z(s) = \prod_{l=1}^{L} \Phi_{z_l}(s) = \exp \left\{ -\frac{1}{2} \sum_{l=1}^{L} \bar{x}_l^H \Sigma_l^{-1} [I_{2d} - (I_{2d} - 2s \Sigma_l Q)^{-1}] \bar{x}_l \right\} \prod_{l=1}^{L} |I_{2d} - 2s \Sigma_l Q|.$$  

(2.11)

Using the symbol rate model from section 2.2.3, the mean vector $\bar{x}_l$ and the covariance matrix $\Sigma_l$ can be found after some math to be

$$\bar{x}_l = \mu_l \begin{bmatrix} \sqrt{E_s} \, 1_d \\ \sqrt{E_p} \, \mathbf{I}_{12M+1}^H \, 1_d \end{bmatrix}, \quad \Sigma_l = \begin{bmatrix} \Sigma_l^{(11)} & \Sigma_l^{(12)} \\ \Sigma_l^{(21)} & \Sigma_l^{(22)} \end{bmatrix},$$

(2.12)
where
$$
\Sigma_{l}^{(11)}(m, n) = E_s\sigma_{c,l}^2 R((m - n)I\tau) + \sigma_{n,l}^2 \delta(m - n) \tag{2.13}
$$
$$
\Sigma_{l}^{(22)}(m, n) = h_l^H (E_p\sigma_{c,l}^2 D_{l,(m-n)} + \sigma_{n,l}^2 \delta(m-n)I) h_l \tag{2.14}
$$
$$
\Sigma_{l}^{(12)}(m, n) = \Sigma_{l}^{(21)}(n, m)^{*} = \sqrt{E_p} E_s \sigma_{c,l}^2 w_{H,l,(n-m)} h_l \tag{2.15}
$$
and $\delta_e$ is a square matrix of size $2M + 1$ with ones on the $e^{th}$ diagonal and zero elsewhere, $D_{l,e}$ is a square matrix of size $2M + 1$ with $D_{l,e}(m, n) = R_l((e + m - n)\tau)$ and $w_{l,e}$ is the $(M + 1)^{th}$ column of $D_{l,e}$.

### 2.3.2 The Chernoff bound

The Chernoff bound of the PEP is an upper bound that is often used in analytical studies due to its simplicity. For our system, the Chernoff bound can be expressed as

$$
\min_{\rho \geq 0} E_z[\exp(-\rho z)] \geq P_2 \tag{2.16}
$$

where $\rho \geq 0$ is a parameter to be optimized [46], and $E_z[f(z)]$ denotes the expectation of $f(z)$ over the random variable $z$. Notice that the expectation can be written as a function of $\Phi_z(s)$, i.e.,

$$
E_z[\exp(-\rho z)] = \int_{-\infty}^{\infty} p_z(x)e^{-\rho x}dx = \Phi_z(-\rho). \tag{2.17}
$$

Since $E_z[\exp(-\rho z)]$ is a convex function with respect to $\rho$ (the second order derivative of $E_z[\exp(-\rho z)]$ with respect to $\rho$ is equal to $\int_{-\infty}^{\infty} \rho^2 p_z(x)e^{-\rho x}dx$ and is always positive), it can be concluded that $\rho_{opt}$ is equal to $\arg\left\{\frac{d\Phi_z(-\rho)}{d\rho} = 0\right\}$. Using the characteristic function specified in (2.11), the set of $\rho_{opt}$ can be simplified as shown in Appendix A to be

$$
\rho_{opt} = \arg_{\rho} \left\{ \sum_{l=1}^{L} \hat{x}_l^H Q(I_{2d} + 2\rho \Sigma_l)Q^{-2}\hat{x}_l + \sum_{l=1}^{L} \sum_{i=1}^{2d} \frac{\hat{\lambda}_i^{(l)}}{1 + \rho \hat{\lambda}_i^{(l)}} = 0 \right\}, \tag{2.18}
$$

where $\hat{\lambda}_i^{(l)}$ for $i = 1, ..., 2d$ are the eigenvalues of $2\Sigma_lQ$. Also note that when the channel is a Rayleigh fading channel, which is the special case of the Rician fading channel when $\mu_l = 0$, (2.18) reduces to

$$
\rho_{opt} = \arg_{\rho} \left\{ \sum_{l=1}^{L} \sum_{i=1}^{2d} \frac{\hat{\lambda}_i^{(l)}}{1 + \rho \hat{\lambda}_i^{(l)}} = 0 \right\}. \tag{2.19}
$$
Substituting $s = -\rho_{opt}$ in (2.11), $\Phi_z(-\rho_{opt})$, which is equal to the Chernoff bound of the PEP, can be found.
2.3.3 The pairwise error probability

In addition to the Chernoff bound, $\Phi_z(s)$ can also be used to find the PEP directly. This is done by substituting the expression of $p_z(x)$ as a function of $\Phi_z(s)$,

$$p_z(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_z(s)e^{-sx}ds,$$  \hspace{1cm} (2.20)

into (2.8). Using Mellin’s inversion, the PEP becomes

$$P_2 = \int_{\epsilon-i\infty}^{\epsilon+i\infty} \frac{j \Phi_z(s)}{2\pi s} ds = -\sum_{i=1}^{n_q} \text{Res}\left[\frac{\Phi_z(s)}{s}\right]_{q_i},$$  \hspace{1cm} (2.21)

where $\epsilon$ lies between the left half-plane poles and the imaginary axis, $n_q$ denotes the number of negative poles of $\frac{\Phi_z(s)}{s}$, $q_i$ denotes the $i^{th}$ negative pole of $\frac{\Phi_z(s)}{s}$ and $\text{Res}[f(s) at q_i]$ denotes the residue of $f(s)$ at the pole $q_i$. The residue can be calculated by

$$\text{Res}[f(s) at a] = \lim_{s \to a} \frac{p^{(v-1)}(s)}{(v-1)!},$$  \hspace{1cm} (2.22)

where $p(s) = (s-a)^vf(s)$, $p^{(v)}(s)$ denotes the $v^{th}$ derivative of $p(s)$ and $v$ is the order of the pole at $a$.

Although the residue theorem leads to a desirable closed-form expression of the PEP, it turns out to be cumbersome for systems with large $d$, which are the systems normally seen in practice. For example, the 3$^{rd}$ generation UMTS-WCDMA standard uses convolutional codes rate 1/2 and rate 1/3 with $d_{\text{min}} = 12$ and 18, respectively. In order to find the PEP for systems with large $d$, a numerical approximation called the Gauss-Chebyshev approximation (suggested in [11] with a correction in [25]) can be used. The Gauss-Chebyshev can be expressed as

$$P_2 \approx \frac{1}{2m} \sum_{k=1}^{m} \left( \Re[\Phi_z(\vartheta_k)] + \tau_k \Im[\Phi_z(\vartheta_k)] \right),$$  \hspace{1cm} (2.23)

where $\Im[x]$ denotes the imaginary part of $x$, $\vartheta_k = \epsilon(1 + j\tau_k)$, $\tau_k = \tan((2k-1)\pi/4m)$ and, in general, $m$ between 16 and 32 is sufficient [11]. In addition, $\epsilon$ is the same $\epsilon$ as defined for (2.21).

2.3.4 Verification for perfect CSI

Although this chapter focuses on systems with noisy CSI, the analysis can also be used to find the performance of systems with perfect CSI, which is a special case of
our analysis when the pilot SNR is infinitely large and the pilot filter is well-designed. This subsection will verify that the PEP corresponding to this special case is equal to the PEP calculated from the perfect CSI analysis. For simplicity, only the flat fading channel is considered, thus, the subscript \( l \) can be dropped. Note that an extension to the frequency-selective fading channel is straightforward.

The perfect CSI assumption can be realized by using \( E_p = \infty \) and \( \mathbf{h} = [1/\sqrt{E_p}] \). Thus, \( \bar{\mathbf{x}} = [\sqrt{E_s} \mu \mathbf{1}_d \ \mu \mathbf{1}_d]^T \) and

\[
2\Sigma_Q = \begin{bmatrix}
\sqrt{E_s} \sigma_c^2 \mathbf{K} & E_s \sigma_c^2 \mathbf{K} + \sigma_n^2 \mathbf{I}_d \\
\sigma_c^2 \mathbf{K} & \sqrt{E_s} \sigma_c^2 \mathbf{K}
\end{bmatrix}
\]

(2.24)

where \( \mathbf{K} \) is a square matrix of size \( d \) with \( \mathbf{K}(m, n) = R((m-n)I_\tau) \). Finding a matrix inverse is usually a difficult task. Fortunately, the matrix inverse of \( \mathbf{I}_{2d} - 2s \Sigma_Q \) can be found easily to be

\[
(\mathbf{I}_{2d} - 2s \Sigma_Q)^{-1} = \begin{bmatrix}
\mathbf{I}_d - s \sigma_c^2 \sqrt{E_s} \mathbf{K} & s(E_s \sigma_c^2 \mathbf{K} + \sigma_n^2 \mathbf{I}_d) \\
\sigma_c^2 \mathbf{K} & \mathbf{I}_d - s \sigma_c^2 \sqrt{E_s} \mathbf{K}
\end{bmatrix} \mathbf{T}^{-1},
\]

(2.25)

where \( \mathbf{T} = \mathbf{I}_d - (2s \sigma_c^2 \sqrt{E_s} + s^2 \sigma_c^2 \sigma_n^2) \mathbf{K} \).

Substituting \( \bar{\mathbf{x}} = [\sqrt{E_s} \mu \mathbf{1}_d \ \mu \mathbf{1}_d]^T \) and (2.25) into (2.11), the characteristic function becomes

\[
\Phi_z^{(\text{perfCSI})}(s) = \frac{\exp \left\{ \frac{s \mu^2}{2} (2\sqrt{E_s} + s \sigma_n^2) \mathbf{1}_d^H \mathbf{T}^{-1} \mathbf{1}_d \right\}}{|\mathbf{T}|}.
\]

(2.26)

Using (2.21) and recalling from Appendix B that \( P_2 \) is invariant to the value of \( a \) in \( \Phi_z(a \mathbf{s})/s \), given that \( a \neq 0 \), \( \Phi_z^{(\text{perfCSI})}(s/\sigma_c^2 \sqrt{E_s}) \) will be used to find \( P_2 \) instead of \( \Phi_z^{(\text{perfCSI})}(s) \). The scaled characteristic function \( \Phi_z^{(\text{perfCSI})}(s/\sigma_c^2 \sqrt{E_s}) \) can be expressed as

\[
\Phi_z^{(\text{perfCSI})}\left(\frac{s}{\sqrt{E_s} \sigma_c^2}\right) = \frac{\exp \left\{ \frac{s \mu^2}{2 \sigma_c^2} (2 + \frac{s}{\tilde{\gamma}_s}) \mathbf{1}_d^H (\mathbf{I}_d - (2s + \frac{s^2}{\tilde{\gamma}_s}) \mathbf{K})^{-1} \mathbf{1}_d \right\}}{|\mathbf{I}_d - (2s + \frac{s^2}{\tilde{\gamma}_s}) \mathbf{K}|}
\]

(2.27)

\[
= \prod_{i=1}^{d} \left( 1 - s \left( \tilde{\lambda}_i - \sqrt{\tilde{\lambda}_i^2 + \tilde{\lambda}_i/\tilde{\gamma}_s} \right) \right) \left( 1 - s \left( \tilde{\lambda}_i + \sqrt{\tilde{\lambda}_i^2 + \tilde{\lambda}_i/\tilde{\gamma}_s} \right) \right)
\]

(2.28)

where \( \tilde{\gamma}_s = E_s \sigma_c^2 / \sigma_n^2 \) is the data SNR and \( \{ \tilde{\lambda}_1, \ldots, \tilde{\lambda}_d \} \) are the eigenvalues of \( \mathbf{K} \). From (2.28), it can be shown that the poles of \( \Phi_z^{(\text{perfCSI})}(s/\sigma_c^2 \sqrt{E_s}) \) are \( -\tilde{\gamma}_s \pm \sqrt{\tilde{\gamma}_s^2 + \tilde{\gamma}_s/\tilde{\lambda}_i} \).
Since $K$ is a positive definite matrix, $\lambda_i$ for $i = 1, \ldots, d$ are greater than zero. It can be shown that the negative poles are $-\bar{\gamma}_s - \sqrt{\bar{\gamma}_s^2 + \bar{\gamma}_s/\lambda_i}$ and the positive poles are $-\bar{\gamma}_s + \sqrt{\bar{\gamma}_s^2 + \bar{\gamma}_s/\lambda_i}$.

The requirement for (2.21) is that $\epsilon$ lies between the poles on the left half-plane and the imaginary axis. A value for $\epsilon$ that satisfies the constraint is $-\bar{\gamma}_s$. Substituting $\epsilon = -\bar{\gamma}_s$ in (2.21), when the scaled characteristic function $\Phi_{\epsilon}^{(\text{perfCSI})}(s/\sigma^2_{\epsilon} \sqrt{E_s})$ is used instead of $\Phi_{\epsilon}^{(\text{perfCSI})}(s)$, $P_2$ becomes

$$P_2 = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{(-\bar{\gamma}_s + jy)} \Phi_{\epsilon}^{(\text{perfCSI})}(\frac{-\bar{\gamma}_s + jy}{\sqrt{E_s\sigma^2_{\epsilon}}}) dy.$$  \hspace{1cm} (2.29)

After changing the dummy variable from $y$ to $\gamma_s t$, $P_2$ becomes

$$P_2 = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\bar{\gamma}_s}{(-\bar{\gamma}_s + j\gamma_s t)} \Phi_{\epsilon}^{(\text{perfCSI})}(\frac{-\bar{\gamma}_s + j\gamma_s t}{\sqrt{E_s\sigma^2_{\epsilon}}}) dt.$$  \hspace{1cm} (2.30)

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{(1 + jt)}{(1 + t^2)} \exp\left\{ -\frac{\bar{\gamma}_s t^2}{2\sigma^2_{\epsilon}} (1 + t^2) \begin{bmatrix} I_d & 1_H \end{bmatrix} (\begin{bmatrix} I_d & 1_H \end{bmatrix} + \bar{\gamma}_s (1 + t^2)K) \right\} dt.$$  \hspace{1cm} (2.31)

$$= \frac{1}{\pi} \int_{0}^{\infty} \frac{1}{1 + t^2} \exp\left\{ -\frac{\bar{\gamma}_s t^2}{2\sigma^2_{\epsilon}} (1 + t^2) \begin{bmatrix} I_d & 1_H \end{bmatrix} (\begin{bmatrix} I_d & 1_H \end{bmatrix} + (1 + t^2)\bar{\gamma}_s K) \right\} dt.$$  \hspace{1cm} (2.32)

In the following, it will be proven that the perfect CSI performance calculated by averaging the PEP over the distribution of the instantaneous data SNR is equal to (2.32).

Under the perfect CSI assumption, the PEP is a function of $\gamma_{\epsilon} = \sum_{i=1}^{d} \gamma_{s,i}$, where $\gamma_{s,i}$ is the instantaneous data SNR corresponding to the $i^{th}$ error symbol. The characteristic function of $\gamma_{\epsilon}$ was given in [51] to be

$$\Phi_{\gamma_{\epsilon}}(s) = \exp\left\{ \frac{s \bar{\gamma}_s t^2}{2\sigma^2_{\epsilon}H} \begin{bmatrix} I_d & 1_H \end{bmatrix} \left( \begin{bmatrix} I_d & 1_H \end{bmatrix} - s\bar{\gamma}_s K \right)^{-1} 1_d \right\},$$  \hspace{1cm} (2.33)

where $\sigma^2_{\epsilon}, \mu, \bar{\gamma}_s$ and $K$ are as previously defined. Since $P_2$ given $\gamma_{\epsilon}$ is equal to $Q(\sqrt{2\gamma_{\epsilon}})$ [51], where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-t^2} dt$ for $x \geq 0$, the average PEP can be found by averaging $Q(\sqrt{2\gamma_{\epsilon}})$ over the distribution of $\gamma_{\epsilon}$, i.e.,

$$P_2 = \int_{0}^{\infty} Q(\sqrt{2\gamma_{\epsilon}}) p(\gamma_{\epsilon}) d\gamma_{\epsilon}.$$  \hspace{1cm} (2.34)

Using an alternative form of the complementary error function [5]

$$\text{erfc}(x) = \frac{2}{\pi} \int_{0}^{\infty} e^{-x^2(1+t^2)} \frac{dt}{1 + t^2} \text{ for } x > 0,$$  \hspace{1cm} (2.35)
where \(2Q(x) = \text{erfc}(x/\sqrt{2})\), (2.34) can be simplified to

\[
P_2 = \frac{1}{\pi} \int_{0}^{\infty} \int_{0}^{\infty} \frac{e^{-\gamma_s(1+t^2)}}{1+t^2} p(\gamma_s) d\gamma_s \, dt
\]

\[
= \frac{1}{\pi} \int_{0}^{\infty} \Phi_{\gamma_s}(s)|s = -(1+t^2)| \, dt
\]

\[
= \frac{1}{\pi} \int_{0}^{\infty} \frac{1}{1+t^2} \exp \left\{ - \frac{\gamma_s \mu^2}{2\sigma_c^2} (1+t^2) \mathbf{1}_d^H (\mathbf{I}_d + (1+t^2)\gamma_s \mathbf{K})^{-1} \mathbf{1}_d \right\} dt,
\]

Comparing (2.32) and (2.38), it is obvious that they are identical. Therefore, this subsection has successfully shown that, in the limit when the channel estimates are perfect, the results from this chapter agree with the results from the perfect CSI analysis.

### 2.4 Numerical Results

In this section, analytical results calculated by the Gauss-Chebyshev approximation are discussed and a comparison between analytical results and results from Monte Carlo simulation is used to illustrate the accuracy of the analysis.

#### 2.4.1 The optimal normalized Doppler frequency

One of the effects from the estimation-diversity tradeoff is the optimal channel memory or the optimal normalized Doppler frequency \(f_d\tau\), which is the channel memory that utilizes the tradeoff in the most effective way [63]. At the optimal \(f_d\tau\), the performance is such that, if \(f_d\tau\) increases, the performance degradation due to worse channel estimates outweighs the benefit from the increase of the channel memory. On the other hand, if \(f_d\tau\) decreases from the optimal value, the degradation due to smaller channel diversity outweighs the benefit from better channel estimates. These optimal \(f_d\tau\)’s can easily be seen in fig. 2.2, which illustrates the PEP of systems in Rayleigh fading channels as a function of \(f_d\tau\) for \(I = 1, 5, 10, 15, 20, 25\) for two types of power spectral density (PSD), namely, the Jakes PSD and the Gaussian PSD. In addition, it is assumed that \(d = 18\), data SNR = 7 dB, pilot SNR = 0 dB and an 11-tap Wiener filter is used as the channel estimator.

One striking difference between the Jakes and the Gaussian PSD’s that should be mentioned is the oscillation seen in the plots corresponding to the Jakes PSD but not
Figure 2.2 Pairwise error probability versus normalized Doppler frequency for different interleaving depths for Jakes PSD and Gaussian PSD.

Figure 2.3 Pairwise error probability versus normalized Doppler frequency for realistic cases and ideal cases.
the Gaussian PSD. This is because the autocorrelation function of the Jakes PSD (the zeroth-order Bessel function of the first kind) is not monotonically decreasing. For a non-monotonically decreasing function, increasing the symbol spacing does not always decrease the correlation or increase the channel diversity. As a result, the plot corresponding to a non-monotonically decreasing autocorrelation function oscillates while the plot corresponding to a monotonically decreasing function does not. In addition, the oscillation can also be seen when the PEP is plotted as a function of the interleaving depth for the same reason.

To better understand the system performance as a function of $f_d \tau$, consider fig. 2.3 where the PEP assuming noisy CSI and finite-depth interleaving, the PEP assuming perfect CSI and finite-depth interleaving and the PEP assuming perfect interleaving and noisy CSI are compared. It is also assumed that $d = 18$, data SNR = 7 dB, an 11-tap Wiener filter is used as the channel estimator and the channel has the Gaussian PSD. From this figure, it can be seen that the solid lines (noisy CSI, finite-depth interleaving) corresponding to 10 dB pilot SNR are very close to the dotted lines (perfect CSI) at small $f_d \tau$. This is because, with this pilot SNR and the Wiener pilot filter, the receiver can accurately estimate the channel. Thus, the performance is close to the perfect CSI case. In addition, it can also be seen that at 0 dB pilot SNR, the performance is much further away from the perfect CSI because of bad channel estimates. The figure also shows that as $f_d \tau$ increases and the channel estimates are less accurate, the solid lines diverge more and more from perfect CSI, and when $f_d \tau$ is large enough such that the channel diversity is equal to the code diversity, the solid lines merge with the perfect interleaving performance (dashed lines). Since the system with a larger $I$ can reach the channel diversity at a smaller $f_d \tau$ [48], the solid line corresponding to a larger $I$ merges with the perfect interleaving line at a smaller $f_d \tau$.

### 2.4.2 Comparing the effects of the pilot SNR on systems in fast and slowly fading channels

In fig. 2.4, the effect of the pilot SNR, which is equal to $E_p \sigma_c^2/\sigma_n^2$, on the PEP and the Chernoff bound of systems in Rayleigh fading channels is illustrated. It is assumed that data SNR = 7 dB, $I = 30$, $d = 18$, the channel has the Gaussian PSD and an 11-tap Wiener filter is used as the channel estimator. Two fading channels shown are
Figure 2.4 Pairwise error probability and Chernoff bounds versus pilot SNR for different normalized Doppler frequencies.

the fast fading channel with $f_d\tau = 0.1$, referred to as system A, and the slowly fading channel with $f_d\tau = 0.01$, referred to as system B. It can be seen from the figure that at small pilot SNR, system B outperforms system A, and vice versa at large pilot SNR. The behavior agrees with the finding in [63] which can be explained as follows. According to [63], when the system operates at a rate close to the capacity, the CSI accuracy is crucial. But when the system operates at a rate much lower than the capacity, the channel diversity is crucial. Since the rate is constant and the capacity at small pilot SNR is smaller than the capacity at large pilot SNR, the CSI accuracy and the channel diversity dominate the performance of the system at small pilot SNR and large pilot SNR, respectively. In addition to the PEP, the figure also illustrates the Chernoff bounds of the PEP. It is clearly seen that the bounds follow the exact PEP nicely.

Although the performance at small and large pilot SNR is predictable, it should be noted that the performance when the pilot SNR is moderate is not easily predicted and must be found through calculation because it depends on other parameters such as the interleaving depth and the data SNR. An analysis such as the one presented in this chapter is needed to quantify the performance in this moderate pilot SNR region.
Figure 2.5 Pairwise error probability versus pilot SNR for different interleaving depths and normalized Doppler frequencies.

2.4.3 Improving the performance through the interleaving depth

In addition to the Doppler spread, the channel diversity can be increased by increasing the interleaving depth $I$, which, unlike the Doppler spread, is a controllable parameter limited only by the delay constraint of the system. Fig. 2.5 illustrates the effect of $I$ on the PEP in a Rayleigh fading channel with $f_d \tau = 0.005$ and $0.0005$ given that data SNR = 7 dB, $d = 18$, the Jakes PSD and an 11-tap Wiener filter is used as the channel estimator. It can be seen that increasing $I$ can significantly improve the performance at large pilot SNR but not as much at small pilot SNR. The reason is the same as the one stated in subsection 2.4.2 that, the accuracy of CSI, not the channel diversity, dominates the performance at small pilot SNR [63]. Therefore, increasing the channel diversity via the interleaving depth does not improve the performance much at small pilot SNR.
2.4.4 Filter choice

Up until now, the dynamic Wiener filter, which calculates $h$ according to the system’s pilot SNR and the channel statistic as the channel estimator, has been used. In order to do this, the receiver must have knowledge of the pilot SNR and the channel statistics of the system. In a real system, this knowledge may not be available or it may not be accurate. To get around this problem, instead of using a dynamic filter, a simple receiver may use a static filter which never changes its filter tap coefficients. Fig. 2.6 compares the performance of the two filters where the fixed filter is randomly chosen to be the Wiener filter corresponding to pilot SNR = -10 dB and $f_d\tau = 0.0482$ in a Rayleigh fading channel. It is assumed that data SNR = 7 dB, $d = 18$, $I = 10$, the Jakes PSD, and that an 11-tap Wiener filter is used as the channel estimator. In this figure, the z-axis represents $(P_2(\text{fixed}) - P_2(\text{dynamic}))/P_2(\text{dynamic})$, the value which is positive when the dynamic filter outperforms the fixed filter. It can be seen from this plot that the dynamic filter is sometimes outperformed by the fixed filter. This behavior is expected because the optimal receiver, which results in the smallest PEP, performs
Figure 2.7 Pairwise error probability versus normalized Doppler frequency for different interleaving depths and multi-path characteristics.

joint estimation-decoding; thus, using the combination of the optimal estimator and the optimal decoder does not guarantee the optimal result. It is also apparent that the filter choice is critical when the pilot SNR and \( f_d \tau \) are large as \( P_2(\text{fixed}) \) is almost 3 times larger than \( P_2(\text{dynamic}) \).

### 2.4.5 Frequency-selective fading channel

Fig. 2.7 shows the PEP when the system is in a frequency-selective Rayleigh fading channel assuming that there are two resolvable paths and each path has half of the average power of the path in the flat fading case. The parameters are set such that \( E_s = 1, E_p = 0.01, N = 128, \sigma_w^2 = -31 \text{ dB}, d = 18 \), the channel has the Gaussian PSD and an 11-tap Wiener filter as the channel estimator.

It can be seen from the figure that path diversity added by the multi-path causes the performance of the frequency-selective fading to be better than the flat fading when \( f_d \tau \) is small. But, due to smaller pilot SNR per path, the channel estimation accuracy of the frequency-selective fading deteriorates much faster as \( f_d \tau \) increases, leading to worse
2.4.6 Effect from the data-to-pilot ratio

Fig. 2.8 illustrates the effects of the data-to-pilot ratio and $f_d\tau$ on the PEP of systems in a flat Rayleigh fading channel (solid plane) and a frequency-selective Rayleigh fading channel with two resolvable paths (dotted plane). The total energy $E_s + E_p$ is assumed to be 1.01. In addition, it is also assumed that $N = 128$, $\sigma_w^2 = -31$ dB, $I = 23$, $d = 18$, the channel has the Gaussian PSD and an 11-tap Wiener filter is used as the channel estimator.

Consider the figure from small $E_s/E_p$ toward $E_s/E_p = 100$. When $E_s/E_p$ is very small (small data energy, large pilot energy), the pilot SNR is large enough that even the multi-path system, which is the system with worse channel estimates, has accurate CSI. Since both systems have good channel estimates at small $E_s/E_p$, the diversity is the dominant factor. Therefore, the multi-path system which has more diversity outperforms the flat-fading system at small $E_s/E_p$. When $E_s/E_p$ is large, however, the CSI accuracy
Figure 2.9 Pairwise error probability versus pilot SNR for different Hamming weights, normalized Doppler frequency and Rician $K$ factor.

of the multi-path system becomes worse especially at high $f_d \tau$. Therefore, it can be seen from the figure that, at large $E_s/E_p$ and large $f_d \tau$, the performance of the multi-path system is worse than the flat-fading system. From the figure, it can be noticed that the gain from allocating appropriate energy to the data and the pilot channels can be significant. For example, changing $E_s/E_p$ from 40 to 1 can improve the performance up to 4 orders of magnitude. Lastly, it should be pointed out that the line corresponding to $I = 23$ in fig. 3.6 illustrates the cross-section of fig. 2.8 when $E_s/E_p = 100$.

2.4.7 Rician fading channel

Fig. 2.9 illustrates the PEP of systems in time-varying Rician fading channels with $f_d \tau = 0.01$ and 0.1 for the Rician $K$ factor (also denoted by $\mu^2/2\sigma_c^2$) 0.1, 1 and 4, and for $d = 3$ and 18. It is also assumed that data SNR = 7 dB, $I = 30$, the channel has the Jakes PSD, and 11-tap Wiener filter is used as the channel estimator.

It can be seen that the performance improves with the Rician $K$ factor as expected for all cases except the case when $d = 18$ and large pilot SNR. This is because, when
the pilot SNR is large, the dominant factor is the channel diversity [63], and with \( d = 18 \), the performance is tremendously improved by the channel diversity and, as a result, it is also very sensitive to the decrease in the channel diversity. When the the Rician \( K \) factor increases by a little bit (from 0.1 to 1), the channel diversity which comes from the diffuse component of the fading channel is reduced, while the line-of-sight component is not large enough to compensate for the performance loss due to the decrease in the channel diversity. But when the Rician \( K \) factor is large enough (\( \mu^2/2\sigma_c^2 = 4 \) in this example) that the line-of-sight component can compensate for the performance loss due to the decrease in the channel diversity, the performance of the system with \( d = 18 \) also improves with the Rician \( K \) factor. Also note that the same behavior is observed for the system with \( d = 3 \) but at smaller Rician \( K \) factor, i.e., at \( \mu^2/2\sigma_c^2 = 0.1 \rightarrow 0.2 \) instead of \( \mu^2/2\sigma_c^2 = 0.1 \rightarrow 1 \).

In addition, it can be observed that the performance when \( f_d \tau = 0.1 \) is sometimes outperformed by \( f_d \tau = 0.01 \) at large pilot SNR when the Rician \( K \) factor is nonzero. This observation was also seen in [67]. As explained in [67], the reason is that, as the line-of-sight component becomes stronger, the need for channel diversity diminishes.

Finally, when the Rician \( K \) factor approaches \( \infty \), the performance of \( f_d \tau = 0.01 \) and 0.1 converge to the same performance, which is \( P_2 \) corresponding to the AWGN channel.

2.4.8 Comparisons with Monte Carlo simulation

To show the accuracy of the analysis, the truncated union bound calculated from the analytical PEP are compared with results from Monte Carlo simulation for a Rayleigh fading channel in fig. 2.10. The code used in the simulation is the rate 1/3 convolutional code specified in the UMTS-WCDMA standard [3] with \( d_{\text{min}} = 18 \) and 256 states. The interleaver (with \( I = 23 \) in this simulation) is also specified in the UMTS-WCDMA standard [3]. The number of information bits per block is assumed to be 220 and each block is terminated with 8 zeros such that the encoder is set back to the all-zero state at the end of each block. The fading coefficients are generated by method of exact Doppler spread (MEDS) suggested in [45] with the autocorrelation of the Jakes model. The channel estimator is an 11-tap Wiener filter. In addition, data SNR = 2.22 dB and pilot SNR = 0.97 dB.
For the truncated union bound, only five smallest Hamming weights \((d = 18, 20, 22, 24, 26)\) are used to calculate the bit error probability and the block error probability, denoted by \(P_e\) and \(P_f\), respectively. In addition, two values of interleaving depths used in the analysis are 23 and 36. Comparing the simulation results and the analytical results with \(I = 23\), it can be seen that the analytical results match well with the simulation results especially when the probability of error is small. Comparing between the analytical results with \(I = 23\) and 36, it can be seen that the two interleaving depths provide similar performance especially when the Doppler spread is large.

A similar comparison between the simulation results and the analytical results for a Rician fading channel with \(\mu^2/2\sigma_c^2 = 0.5\) and \(\mu^2 + 2\sigma_c^2 = 1\) is illustrated in fig. 2.11. However, to illustrate the effect of a mismatch channel estimator, it is assumed in this simulation that an 11-tap moving average filter is used as the channel estimator. It can clearly be seen from the simulation results that performance rapidly degrades with the normalized Doppler spread as a result from the mismatch channel estimator. Comparing the truncated union bound of the bit error probability calculated from the minimum Hamming distance \((d_{\text{min}} = 18)\) for \(I = 23\) and 36 (the dotted and the dashed-dotted
Comparison between analytical and simulation results (Rician)

Figure 2.11 Bit error probability and block error probability (Rician) versus normalized Doppler frequency from Monte Carlo simulation and from the analysis.

lines, respectively), it is clear that the system performance in a Rician fading channel strongly depends on the interleaving depth and the spacing of error symbols. Since the spacing of error symbols in a real system depends on the interleaving pattern and the error sequence, the simple assumption that all consecutive error symbols are $I$ symbols apart does not lead to an accurate performance prediction in the Rician fading channel. Therefore, instead of calculating the PEP assuming that consecutive error symbols are $I$ symbols apart, the PEP is calculated by taking into account the error sequence of the convolutional code and the interleaving structure. (This modification can be done with minor change on the covariance matrix.) The PEP’s corresponding to all of the error patterns are then used to calculate the approximation of $P_e$ and $P_f$, which are shown in the figure as the solid and the dashed lines, respectively. Also note that three smallest Hamming weights, i.e., 18, 20 and 22, are used to calculate the $P_e$ and $P_f$ without the equally spaced error symbols assumption. It can be seen from the figure that the analytical results match well with the simulation results especially when $f_d\tau$ is less than 0.04.
2.5 Conclusions

This chapter has derived the Chernoff bound of the pairwise error probability and the exact pairwise error probability of coded systems with finite-depth interleaving and noisy channel estimates. The analysis provides an insight into the system performance in a realistic environment. For example, it has been shown that there exists an optimal channel memory length, which is a result of the estimation-diversity tradeoff as a function of the Doppler spread. Also, it has been observed that, in a fast fading channel, increasing the pilot SNR can improve the performance more effectively than increasing the interleaving depth. This chapter has also investigated the system performance as a function of the Rician $K$ factor and found that the system with a large $d$ is more sensitive to the decrease of the channel diversity resulting from increasing $K$.

In addition to gaining more understanding of the system behavior, it has also been shown that the analysis presented in this chapter is a great tool for system designs. For example, the analysis can be used to compare the performance between different pilot filters, data-to-pilot ratios, or coding schemes. Finally, to illustrate the accuracy of the analysis, the truncated union bounds calculated from the analytical pairwise error probabilities are compared with the results from Monte Carlo simulation and the bounds have been shown to match well with the simulation results when the probability of error is small (union bound limitation) and, for the Rician case, when the Doppler spread is not too large.

2.6 Acknowledgement

This chapter, in part, is a reprint of the material as it appears in “Performance of finite-depth interleaved convolutional codes in a Rayleigh fading channel with noisy channel estimates,” J. Jootar, J.R. Zeidler, J.G. Proakis, in Proc. IEEE Vehicular Technology Conference, vol. 1, Stockholm, May 2005, pp. 600-605 and “Performance of convolutional codes with finite-depth interleaving and noisy channel estimates,” J. Jootar, J.R. Zeidler, J.G. Proakis, to be appeared in IEEE Transactions on Communications. The dissertation author was the primary author of these papers.
Performance of concatenated convolutional code and Alamouti space-time code with noisy channel estimates and finite-depth interleaving

3.1 Introduction

Over the last decade, space-time coding has gained attention both from the research community, as seen in a large number of space-time code related papers published, and the industrial community, where several space-time coding schemes are included as parts of future wireless communication standards. However, space-time coding schemes with large diversity gains require a greater number of transmit antennas than are generally available in practice. Therefore, in order to achieve the bit error rate performance levels required in most wireless systems, space-time codes are often used in concatenation with channel codes or trellis-coded modulation [39]. This chapter analyzes the conventional Alamouti space-time coded system [7] as well as a modified system which is a concatenation of the convolutional and Alamouti space-time coded system (CCA). This system
is a part of the space-time transmit diversity (STTD) specifications proposed in the 3rd generation UMTS-WCDMA [2, 3].

The performance of convolutional codes in fading channels is well known. However, only recently have the concurrent effects of noisy channel state information (CSI) and finite-depth interleaving been considered (see chapter 2 or [33, 63]). Such studies have shown that there exists an estimation error and diversity gain performance tradeoff as a function of Doppler spread. The tradeoff is a result of degraded CSI accuracy with concurrent improvements on diversity as the Doppler spread increases (see chapter 2 or [33, 63]).

The Alamouti space-time code was proposed in 1998 [7] and has generated significant interest due to the low complexity linear-combining space-time decoder (LC-STD) suggested by Alamouti in [7]. The LC-STD was designed such that, under perfect CSI and quasi-static fading assumptions, it can completely eliminate interference from the other symbol in the codeword and achieve the same performance as the more complex ML space-time decoder (ML-STD) [7, 59]. When the channels are not quasi-static, however, the interference is not completely eliminated by the LC-STD (even if the CSI is perfect) causing performance degradation [49, 65]. Under this condition, either an interference suppression technique or the ML-STD should be used to mitigate the effect of the time-varying fading channels [59, 65].

Other factors that can significantly degrade the performance of the Alamouti space-time code are noisy CSI [9, 13, 16, 27, 28, 49, 53, 58, 61], which is caused by noise in the pilot symbols or time-varying fading channels, and spatial correlation of the transmit antennas [34, 58, 60]. It has been shown that, when CSI is noisy, the LC-STD cannot correctly eliminate the signal contribution from the other symbol resulting in interference even in the quasi-static channels [9, 13, 16, 28, 49, 53, 58]. These degradations are known to limit performance regardless of the space-time decoding scheme employed.

Given the close relationship between the Doppler spread and CSI accuracy and their effects on the system performance, it is desirable to consider both parameters when the system performance is evaluated. However, previous studies on the Alamouti space-time code [7] have focused on each of these factors in isolation, e.g., the analytical performance of a zero-forcing linear detector and a decision-feedback detector assuming perfect CSI and non-quasi-static fading channels was derived in [59], while analyses
assuming noisy CSI in quasi-static fading channels were presented in [9, 13, 16, 27, 28, 49, 53, 58]. In addition, analyses on the performance of space-time codes over spatially correlated Rayleigh fading channels were presented in [34, 58, 60]. Similarly, only limited studies are available for concatenated space-time and channel coded systems and most of these studies presented simulation results instead of analytical results [30, 39]. In addition, the papers that derived analytical performance used at least one idealistic assumption such as perfect CSI, quasi-static channels, no spatial correlation or perfect interleaving [10, 17, 50].

The objective of this chapter is to generalize the existing analyses and provide analytical performance of the conventional Alamouti space-time code [7] and the CCA by taking into account imperfect CSI, spatial correlation and finite-depth interleaving in time-varying fading channels. Two decoding schemes for the conventional Alamouti space-time coded system are considered, namely, the low complexity LC-STD [7] and the high complexity ML-STD. Two decoding schemes for the CCA are evaluated, namely, the LC-STD with an ML convolutional decoder (LC-ML) and the joint Alamouti and convolutional ML decoder (JML).

The chapter is organized as follows. Section 3.2 describes the mathematical tools used in this paper. In section 3.3, the system model is introduced. Performance analyses of the systems with the LC-STD and the LC-ML are derived in section 3.4, while performance analyses of the systems with the ML-STD or the JML are derived in section 3.5. In section 3.6, an extension to multi-path fading channels is described. Numerical results are presented and the comparison between the analytical and simulation results are discussed in section 3.7. Finally, the conclusion and the acknowledgement are presented in section 3.8 and 3.9, respectively.

3.2 Characteristic function, residue theorem and Gauss-Chebyshev approximation

For the rest of this chapter, the following notation will be used. A lowercase bold letter denotes a vector and an uppercase bold letter denotes a matrix. The element in the $m^{th}$ row and the $n^{th}$ column of a matrix $X$ is denoted by $X(m, n)$ and the element in the $m^{th}$ row (column) of a column (row) vector $x$ is denoted by $x(m)$. The superscripts
\(*, T, H\) denote the complex conjugate, the matrix transpose and the matrix Hermitian operation, respectively. The determinant of a matrix \(X\) is denoted by \(|X|\). The length \(m\) column column vector of ones, the square identity matrix, the square zero matrix and the anti-diagonal matrix of order \(m\) are denoted by \(1_m, I_m, 0_m\) and \(J_m\), respectively.

Analytical tools used in this chapter are the characteristic function of a Hermitian quadratic form of complex Gaussian random variables, the residue theorem, and the Gauss-Chebyshev approximation. The main idea is that, if the metric of interest \(z\) can be written as a Hermitian quadratic form of a zero-mean complex Gaussian random vector \(x\), i.e., \(z = x^H Q x\), then, with the knowledge of the covariance matrix of \(x\), the characteristic function of \(z\) is known. In addition, the probability that \(z\) is less than zero can be found by using the characteristic function of \(z\) with the residue theorem or the Gauss-Chebyshev approximation.

### 3.2.1 Characteristic function of a Hermitian quadratic form of a complex Gaussian random vector

Let \(x\) be a zero-mean complex Gaussian random vector of length \(m\) and \(Q\) be a square matrix of order \(m\), then the characteristic function of \(z = x^H Q x\) can be written as [57]

\[
\Phi_z(s) = |I_m - 2s\Sigma Q|^{-1}, \tag{3.1}
\]

where \(\Sigma\) denotes the covariance matrix of \(x\).

### 3.2.2 Residue theorem

The probability that \(z\) is less than zero can be found by Mellin’s inversion formular and the residue theorem as follows

\[
\Pr\{z < 0\} = \int_{c - \infty j}^{c + \infty j} \frac{j \Phi_z(s)}{2\pi s} ds = -\sum_{i=1}^{n_q} \text{Res} \left[ \frac{\Phi_z(s)}{s} \right] \text{ at } q_i, \tag{3.2}
\]

where \(c\) lines between the left half-plane poles and the imaginary axis, \(n_q\) denotes the number of negative poles of \(\Phi_z(s)\), \(q_i\) denotes the \(i^{th}\) negative pole of \(\Phi_z(s)\) and \(\text{Res}[f(s) \text{ at } q_i]\) denotes the residue of \(f(s)\) at \(q_i\) [42]. In addition, the residue can be calculated by

\[
\text{Res}[f(s) \text{ at } q_i] = \lim_{s \to q_i} \frac{p^{(m-1)}(s)}{(m-1)!}, \tag{3.3}
\]
where \( m \) is the order of the pole at \( q_i \), \( p(s) = (s - q_i)^m f(s) \) and \( p^{(m-1)}(s) \) denotes the \((m - 1)^{th}\) derivative of \( p(s) \). It should be noted that, if the system is simple and does not have many poles and the poles are single, the residue theorem can be used with the characteristic function to find the closed-form expression of \( \Pr\{z < 0\} \).

### 3.2.3 Gauss-Chebyshev approximation

In the case that poles are close together or there are poles with large orders, it may not be feasible to use the residue theorem. However, if the characteristic function of \( z \) is known, numerical results can be calculated from the Gauss-Chebyshev approximation (suggested in [11] with a correction in [25]) as follows

\[
\Pr\{z < 0\} \approx \frac{1}{2m} \sum_{k=1}^{m} \left( \Re[\Phi_z(\epsilon + j\tau_k)] + \tau_m \Im[\Phi_z(\epsilon + j\tau_k)] \right),
\]

where \( \Re[x] \) and \( \Im[x] \) denote the real and the imaginary parts of \( x \), respectively, \( \tau_k = \tan((2k-1)\pi/4m) \), \( \epsilon \) is as defined in section 3.2.2 and, in general, \( m \) between 16 and 32 is sufficient [11].

### 3.3 System model

The system under consideration is a BPSK DS-CDMA system with two transmit antennas and one receive antenna. In this section, the transmitter, the channels and the receiver are discussed in detail.

#### 3.3.1 Transmitter

For the concatenated system, information bits are convolutionally encoded then interleaved with a finite-depth interleaver. Two consecutive interleaved symbols for the concatenated system, or two consecutive uncoded symbols for the conventional Alamouti system, are then encoded by the Alamouti space-time encoder [2, 7]. Since this analysis considers BPSK modulation, the complex conjugates used in the Alamouti code can be ignored.

In addition, it is assumed that data signals transmitted from the two transmit antennas use a common orthogonal code and are detected as one combined signal at the
receiver. This is different, however, for the pilot signals. Since the receiver has to estimate the fading coefficients for each link independently, the pilot signal transmitted from each antenna has to use a distinct orthogonal code. For simplicity, it is also assumed that all of the orthogonal codes used in the system have the same spreading gain $N$, which results in the symbol period $\tau = N\tau_c$, where $\tau_c$ denotes the chip period. The symbol-rate baseband representation of this system is shown in fig. 3.1, where $s_k$ denotes the data symbol during the symbol time index $k$, $E_s$ and $E_p$ denote transmit energy per data symbol and pilot symbol (equivalent to $E_s/2$ and $E_p/2$ per symbol per antenna), respectively.

Figure 3.1 System diagram of STTD over two-symbol period
3.3.2 Channels

The channels are assumed to be time-varying Rayleigh flat-fading channels. The fading coefficient corresponding to the channel between transmit antenna A and the receive antenna during the symbol time index \( k \) is denoted by \( \alpha_k \) and the fading coefficient corresponding to the channel between transmit antenna B and the receive antenna during the symbol time index \( k \) is denoted by \( \beta_k \). Because of the Rayleigh fading assumption, \( \alpha_k \) and \( \beta_k \) are circularly symmetric zero-mean complex Gaussian random variables. It is also assumed that \( \alpha_k \) and \( \beta_k \) have identical autocorrelation function

\[
\frac{1}{2}E[\alpha_k\alpha^*_k-m] = \frac{1}{2}E[\beta_k\beta^*_k-m] = \sigma_c^2 R(m\tau),
\]

where \( R(0) \) is normalized to unity.

In addition, it is assumed that the correlation between \( \alpha_k \) and \( \beta_k \) can be written as

\[
\frac{1}{2}E[\alpha_k\beta^*_k-m] = \rho \sigma_c^2 R(m\tau),
\]

where \( \rho \) is a real value indicating the spatial correlation between the two transmit antennas. It should also be noted that, although time-varying channels is assumed, this analysis is restricted to the case when the channels change slowly enough that the fading coefficients appears to be constant over one symbol period \( \tau \).

After despreading, it is assumed that the noise in the pilot channels from transmit antenna A, transmit antenna B, and the data channel at the symbol time index \( k \), denoted by \( n_{p,k}^{(1)} \), \( n_{p,k}^{(2)} \) and \( n_{s,k} \), are zero-mean circularly symmetric complex white Gaussian variables with variances \( \sigma_p^2 \), \( \sigma_p^2 \) and \( \sigma_s^2 \), respectively.

3.3.3 Receiver

The discrete baseband representation of the received signals over one Alamouti code block can be expressed as

\[
\begin{align*}
\mathbf{r}_{s,2k+1} &= \begin{bmatrix} r_{s,2k+1} & r_{s,2k+2}^* \end{bmatrix} = \sqrt{E_s} \begin{bmatrix} \alpha_{2k+1} & -\beta_{2k+1}^* \\ \beta_{2k+2}^* & \alpha_{2k+2}^* \end{bmatrix} \begin{bmatrix} s_{2k+1} \\ s_{2k+2}^* \end{bmatrix} + \begin{bmatrix} n_{s,2k+1} \\ n_{s,2k+2}^* \end{bmatrix} \\
\mathbf{r}_{p,k}^{(1)} &= \sqrt{\frac{E_p}{2}} \alpha_k + n_{p,k}^{(1)} \rightarrow \mathbf{r}_{p,k}^{(1)} = [r_{p,k-M}^{(1)} \ldots r_{p,k}^{(1)} \ldots r_{p,k+M}^{(1)}]^T \\
\mathbf{r}_{p,k}^{(2)} &= \sqrt{\frac{E_p}{2}} \beta_k + n_{p,k}^{(2)} \rightarrow \mathbf{r}_{p,k}^{(2)} = [r_{p,k-M}^{(2)} \ldots r_{p,k}^{(2)} \ldots r_{p,k+M}^{(2)}]^T,
\end{align*}
\]
where \( r_{s,k} \) denotes the received data signal during the symbol time index \( k \), \( r_{p,k}^{(1)} \) and \( r_{p,k}^{(2)} \) denote the received pilot signals from transmit antennas A and B, respectively, during the symbol time index \( k \). In addition, the channel estimation is assumed to be performed by a \((2M+1)\)-tap FIR filter. The channel estimates, which are the output of the FIR filters, can be expressed as

\[
\hat{\alpha}_k = h^H r_{p,k}^{(1)}, \quad \hat{\beta}_k = h^H r_{p,k}^{(2)},
\]

(3.6)

where \( h = [h_{M} \ldots h_{0} \ldots h_{-M}]^T \) denotes the filter coefficient vector, \( \hat{\alpha}_k \) and \( \hat{\beta}_k \) denote the channel estimates corresponding to the channel from the transmit antennas A and B, respectively.

### 3.4 Linear-combining scheme

The linear-combining space-time decoder (LC-STD) is the simple space-time decoder originally suggested by Alamouti in [7] and it is the most likely scheme to be implemented in practice due to its low complexity. This section derives the bit error probability of the conventional Alamouti space-time code when the LC-STD is used and the pairwise error probability (PEP) of the CCA when the LC-ML is used.

#### 3.4.1 Alamouti space-time code with LC-STD

The linear-combining scheme performed over an Alamouti space-time code block corresponding to the symbol time index 1 and 2 can be written in a matrix form as,

\[
\begin{bmatrix}
z_1 \\
z_2^* \\
\end{bmatrix} =
\begin{bmatrix}
\hat{\alpha}_1^* & \hat{\beta}_2 \\
\hat{\beta}_1^* & \hat{\alpha}_2 \\
\end{bmatrix}
\begin{bmatrix}
r_{s,1} \\
\end{bmatrix} = \sqrt{\frac{E_s}{2}} \begin{bmatrix}
s_1 \\
s_2^* \\
\end{bmatrix} + \begin{bmatrix}
\tilde{n}_{s,1} \\
\tilde{n}_{s,2} \\
\end{bmatrix},
\]

(3.7)

where \( z_1 \) and \( z_2 \) are the output of the linear combiner corresponding to the first and the second data symbols, respectively. In addition, \( \tilde{n}_{s,1} = \hat{\alpha}_1^* n_{s,1} + \hat{\beta}_2 n_{s,2}^* \), \( \tilde{n}_{s,2} = -\hat{\beta}_1 n_{s,1} + \hat{\alpha}_2 n_{s,2}^* \) and

\[
G = \begin{bmatrix}
\alpha_1 \hat{\alpha}_1^* + \beta_2^* \hat{\beta}_2 & \alpha_2^* \hat{\beta}_2 - \beta_1 \hat{\alpha}_1^* \\
\beta_2^* \hat{\alpha}_2 - \alpha_1 \hat{\beta}_1^* & \beta_1 \hat{\beta}_1^* + \alpha_2^* \hat{\alpha}_2 \\
\end{bmatrix}.
\]

(3.8)

Notice that, in an ideal environment, where CSI is perfect and channels are quasi-static, the condition becomes \( \hat{\alpha}_1 = \alpha_1 = \hat{\alpha}_2 = \alpha_2 , \hat{\beta}_1 = \beta_1 = \hat{\beta}_2 = \beta_2 \). Therefore, the
linear-combining scheme can eliminate interference within the code block completely [7] and \( \mathbf{G} \) becomes
\[
\mathbf{G}_{\text{perfect}} = \begin{bmatrix}
|\alpha_1|^2 + |\beta_1|^2 & 0 \\
0 & |\alpha_1|^2 + |\beta_1|^2
\end{bmatrix}.
\] (3.9)

The goal of this work is to derive the bit error probability, denoted by \( P_b \), and the block error probability of the Alamouti space-time code, denoted by \( P_{\text{Alamouti}} \), which is upper bounded by
\[
P_{\text{Alamouti}} \leq 2P_b.
\] (3.10)

Since the bit error probabilities of the first symbol and the second symbol are equal, only the bit error probability of the first symbol is derived. Without loss of generality, it is assumed that \( s_1 = 1 \). Consequently, the bit error probability becomes
\[
P_b = \frac{1}{2} (P_{b,s_2=1} + P_{b,s_2=-1}),
\] (3.11)
where \( P_{b,s_2=a} \) denotes the bit error probability given that \( s_2 = a \) and \( a \) is equal to 1 or -1. Using ML symbol-by-symbol detection, \( P_{b,s_2=a} \) can be written as
\[
P_{b,s_2=a} = \Pr \{ \Re[z_1] < 0 \mid s_2 = a \}.
\] (3.12)

It can be seen that \( \Re[z_1] \) can be written in a Hermitian quadratic form \( \Re[z_1] = \mathbf{x}_1^H \mathbf{Q}_1 \mathbf{x}_1 \), where
\[
\mathbf{x}_1 = [r_{s,1} \ r_{s,2} \ \hat{\alpha}_1 \ \hat{\beta}_2]^T, \quad \text{and} \quad \mathbf{Q}_1 = \frac{1}{2} \begin{bmatrix}
\mathbf{0}_2 & \mathbf{I}_2 \\
\mathbf{I}_2 & \mathbf{0}_2
\end{bmatrix}.
\] (3.13)

In addition, it can be shown that \( \mathbf{x}_1 \) is a zero-mean complex Gaussian random vector. Therefore, the characteristic function of \( \Re[z_1] \) can be found from (3.1), which requires only the knowledge of the covariance matrix of \( \mathbf{x}_1 \). The covariance matrix of \( \mathbf{x}_1 \) can be expressed as
\[
\Sigma_1 = \begin{bmatrix}
\mathbf{A}_{11} & \mathbf{A}_{12} \\
\mathbf{A}_{12}^H & \mathbf{A}_{22}
\end{bmatrix},
\] (3.14)
where
\[
\begin{align*}
\mathbf{A}_{11} &= E_s \sigma_c^2 \begin{bmatrix}
(1 - \rho a) + \bar{\gamma}_s^{-1} & 0 \\
0 & (1 + \rho a) + \bar{\gamma}_s^{-1}
\end{bmatrix}, \\
\mathbf{A}_{12} &= \frac{\sqrt{E_s} E_p \sigma_c^2}{2} \begin{bmatrix}
w_0^H h(1 - \rho a) & w_0^H h(\rho - a) \\
w_{-1}^H h(\rho + a) & w_{-1}^H h(1 + \rho a)
\end{bmatrix},
\end{align*}
\]
With the Wiener filter used as the channel estimator, the filter coefficients become

\[
\mathbf{A}_{22} = \frac{E_p \sigma_c^2}{2} \begin{bmatrix}
\mathbf{h}^H \left( \mathbf{D}_0 + 2 \gamma_p^{-1} \mathbf{I}_{2M+1} \right) \mathbf{h} & \rho \mathbf{h}^H \mathbf{D}_{-1} \mathbf{h} \\
\rho \mathbf{h}^H \mathbf{D}_1 \mathbf{h} & \mathbf{h}^H \left( \mathbf{D}_0 + 2 \gamma_p^{-1} \mathbf{I}_{2M+1} \right) \mathbf{h}
\end{bmatrix}
\]

where \( \gamma_p \) and \( \gamma_s \) are the average pilot SNR and the average data SNR, respectively, \( \mathbf{D}_e \) is a square matrix of order \( 2M+1 \) with \( \mathbf{D}_e(m,n) = R((\epsilon+m-n)\tau) \) and \( \mathbf{w}_e \) is the \((M+1)th \) column of \( \mathbf{D}_e \). With the knowledge of the characteristic function of \( \mathbf{x}_1 \), \( P_{b,s_2=a} \) can be found by means of the residue theorem (3.2) or the Gauss-Chebyshev approximation (3.4) as described in section 3.2. In addition, \( P_{b} \) and the upper bound of \( P_{\text{Alamouti}} \) can be found from (3.11) and (3.10), respectively.

Consider now the special case when there is no spatial correlation between the transmit antennas, i.e., \( \rho = 0 \), and when the Wiener filters are used as the channel estimators. With the Wiener filter used as the channel estimator, the filter coefficients become

\[
\mathbf{h} = \sqrt{\frac{2}{E_p}} \left( \mathbf{D}_0 + 2 \gamma_p^{-1} \mathbf{I}_{2M+1} \right)^{-1} \mathbf{w}_0.
\]

Since the covariance matrix is independent of the value of \( s_2 \) when \( \rho = 0 \), it is assumed, without loss of generality, that \( s_2 = 1 \). Therefore, the covariance matrix is simplified to

\[
\hat{\Sigma}_1 = \begin{bmatrix}
E_s \sigma_c^2 (1 + \gamma_p^{-1}) \mathbf{I}_2 & \sigma_c^2 \sqrt{\frac{E_s}{2}} \begin{bmatrix}
\varepsilon_0 & -\varepsilon_1 \\
\varepsilon_1 & \varepsilon_0
\end{bmatrix} \\
\sigma_c^2 \sqrt{\frac{E_s}{2}} \begin{bmatrix}
\varepsilon_0 & \varepsilon_1 \\
-\varepsilon_1 & \varepsilon_0
\end{bmatrix} & \sigma_c^2 \varepsilon_0 \mathbf{I}_2
\end{bmatrix}
\]

where \( \varepsilon_0 = 2 \mathbf{w}_0^H \left( \mathbf{D}_0 + 2 \gamma_p^{-1} \mathbf{I}_{2M+1} \right)^{-1} \mathbf{w}_0 \) and \( \varepsilon_1 = 2 \mathbf{w}_1^H \left( \mathbf{D}_0 + 2 \gamma_p^{-1} \mathbf{I}_{2M+1} \right)^{-1} \mathbf{w}_0 \). The eigenvalues of \( 2 \hat{\Sigma}_1 \mathbf{Q} \) can be found to be \( \lambda_1 = \sigma_c^2 \varepsilon_0 (1 + \Upsilon) \) and \( \lambda_2 = \sigma_c^2 \varepsilon_0 (1 - \Upsilon) \) (both with order two) where

\[
\Upsilon = \left( \frac{4(1 + \gamma_p^{-1})}{\varepsilon_0} - \left( \frac{\varepsilon_1}{\varepsilon_0} \right)^2 \right)^{1/2}.
\]

It can be shown that the poles of \( \frac{\Phi_{z1}(s)}{s} \) are 0, \( \lambda_1^{-1} \) and \( \lambda_2^{-1} \). To use the residue theorem, poles of \( \frac{\Phi_{z1}(s)}{s} \) in the left half-plane have to be identified. Since it is known that the minimum mean square error (MMSE) of the channel estimate is equal to \( 2 \sigma_c^2 (1 - \varepsilon_0/2) \) and \( \mathbf{D}_0 + 2 \gamma_p^{-1} \mathbf{I}_{2M+1} \) is positive definite, it can be concluded that \( 2 \geq \varepsilon_0 \geq 0 \). In addition, since the MMSE is smaller than \( E[\|\mathbf{\alpha}_{k+1} - \hat{\mathbf{\alpha}}_k\|^2] \), it can be concluded that \( \varepsilon_0 \geq \varepsilon_1 \) and,
consequently, $\Upsilon \geq 1$ with equality when $\bar{\gamma}_s = \infty$ and $\varepsilon_0 = \varepsilon_1 = 2$. In the case that $\Upsilon > 1$, $\lambda_1^{-1}$ is positive and $\lambda_2^{-1}$ is negative.

From the residue theorem, the bit error probability is equal to the negative of the summation of the residues of $\Phi_R(z_1)/(s)$ at the negative poles. Using the fact that the residue is invariant to the scaling of the poles (Appendix B), the original poles are scaled with $\sigma^2 \varepsilon_0$ and the compact form of the bit error probability can be calculated as follows

$$P_b = \lim_{s \to -\lambda_2^{-1}} \frac{d}{ds} \left( \frac{(s-\sigma^2 \varepsilon_0)^2}{s(1-s \lambda_1 / \sigma^2 \varepsilon_0)^2(1-s \lambda_2 / \sigma^2 \varepsilon_0)^2} \right) = \frac{1}{4} \left( 2 + \Upsilon^{-1} \right) \left( 1 - \Upsilon^{-1} \right)^2. \quad (3.18)$$

As a sanity check, (3.18) is compared with the perfect CSI result in the existing papers by setting $\bar{\gamma}_p = \infty$. Since $D_0^{-1} \mathbf{D}_0 = \mathbf{I}_{2M+1}$ and $\mathbf{w}_0$ is the $M+1$th column of $\mathbf{D}_0$, $\mathbf{D}_0^{-1} \mathbf{w}_0$ is equal to $[0 \ldots 0 1 0 \ldots 0]^T$ (with one in the $(M+1)$th row). Therefore, the values of $\varepsilon_0, \varepsilon_1$ and $\Upsilon$ when $\bar{\gamma}_p = \infty$ are equal to $2, 2R(\tau)$ and $(2(1 + \bar{\gamma}_s^{-1}) - R(\tau)^2)^{1/2}$, respectively. Substituting the value of $\Upsilon$ into (3.18), the closed-form expression of the bit error probability when $\rho = 0$ with the Wiener filter as the channel estimator can be obtained.

In the limit when $R(\tau) = 0$ (very fast fading) or $R(\tau) = 1$ (static channel) with perfect CSI, the bit error probability can be written in compact forms as

$$R(\tau) = 0 \rightarrow P_b = \frac{1}{4} \left( 2 + \sqrt{\frac{\bar{\gamma}_s}{2\bar{\gamma}_s + 2}} \right) \left( 1 - \sqrt{\frac{\gamma_s}{2\gamma_s + 2}} \right)^2, \quad (3.19)$$

$$R(\tau) = 1 \rightarrow P_b = \frac{1}{4} \left( 2 + \sqrt{\frac{\gamma_s}{\gamma_s + 2}} \right) \left( 1 - \sqrt{\frac{\gamma_s}{\gamma_s + 2}} \right)^2. \quad (3.20)$$

It is apparent that (3.20) agrees with the result derived in [28, 59].

### 3.4.2 CCA with LC-STD and ML convolutional decoder

When the convolutional code is used as the outer code, the output from the linear-combiner can be de-interleaved and convolutionally decoded by a ML convolutional decoder.

The pairwise error probability (PEP) is the probability that the decoder chooses the coded sequence $c_j$ instead of the transmitted sequence $c_i$, where $i \neq j$. For the system
model described in section 3.3, the PEP is a function of the transmitted codeword, the error pattern and the structure of the interleaver. Finding the PEP for each transmitted sequence and each error pattern for a specific interleaving structure is tedious and adds little insight into the overall system performance [48]. Therefore, similar simplifications to those used in chapter 2 are adopted. These simplifications are as follows. The transmitted codeword is assumed to be an all-zero codeword, which is mapped to an all-one BPSK sequence $c_0$, the error codeword is the codeword with $d$ consecutive error symbols and it corresponds to the BPSK sequence $c_i(i \neq 0)$, and an interleaving depth $I$ of a block interleaver creates the same effect as separating consecutive symbol errors by $I$ symbols [26]. As a result, the PEP depends only on the Hamming weight of the error codeword and the interleaving depth. An illustration of an error codeword with Hamming weight 5 and the interleaving depth of 4 symbols is shown in fig. 3.2. Using the LC-STD with the ML convolutional decoder, the PEP, denoted by $P_2$, can be written as

$$P_2 = \Pr \left\{ \sum_{k=0}^{d-1} \Re[z_{kI+1}] < 0 \right\}. \quad (3.21)$$

Notice that, (3.21) assumes that the error symbol starts at the first symbol. By inspection, $\sum_{k=0}^{d-1} \Re[z_{kI+1}]$ can be written in a Hermitian quadratic form $\sum_{k=0}^{d-1} \Re[z_{kI+1}] = x_2^H Q_2 x_2$, where

$$x_2 = [x_2^{(1)}, x_2^{(2)}, x_2^{(3)}, x_2^{(4)}]^T, Q_2 = \frac{1}{2} \begin{bmatrix} 0_{2d} & I_{2d} \\ I_{2d} & 0_{2d} \end{bmatrix}, \quad (3.22)$$

and $x_2^{(1)}, x_2^{(2)}, x_2^{(3)}, x_2^{(4)}$ are row vectors of length $d$ with $x_2^{(1)}(m) = r_{s,g(1)(m)}$, $x_2^{(2)}(m) = \ldots$

\begin{figure}
\centering
\includegraphics[width=\textwidth]{alamouti_space_time_code_blocks.png}
\caption{Alamouti space-time code blocks}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{error_symbols_interleaving_depth.png}
\caption{Error symbols when the interleaving depth is equal to 4 symbols}
\end{figure}

Figure 3.2 Illustration of an error codeword with Hamming weight 5 and 4-symbol interleaving depth
$r_{s,g^{(2)}(m)}, x_2^{(3)}(m) = \hat{\alpha}_{g^{(1)}(m)}, x_2^{(4)}(m) = \hat{\beta}_{g^{(2)}(m)}$, \\
g^{(1)}(k) = \begin{cases} (k - 1)I + 1 & \text{when } I \text{ is even} \\ 4\lceil k/2 \rceil + 6\lfloor (k - 1)/2 \rfloor + 1 & \text{otherwise,} \end{cases} \tag{3.23}
\\
g^{(2)}(k) = g^{(1)}(k) + 1 \text{ and } \lfloor x \rfloor \text{ denotes the closest integer that is less or equal to } x. \text{ Since } x_2 \text{ is a zero-mean complex Gaussian random vector, the characteristic function of } x_2^H Q_2 x \text{ can be used with the residue theorem or the Gauss-Chebyshev approximation to find the PEP. The covariance matrix of } x_2, \text{ denoted by } \Sigma_2, \text{ is needed to find the characteristic function. From the above expression, it can be shown that}
\\
\Sigma_2 = \begin{bmatrix}
B_{11} & 0_d & B_{13} & B_{14} \\
0_d & B_{22} & B_{23} & B_{24} \\
B_{13}^H & B_{23}^H & B_{33} & B_{34} \\
B_{14}^H & B_{24}^H & B_{34}^H & B_{33}
\end{bmatrix}, \tag{3.24}
\\
\text{where}
\\
B_{11}(m,n) = E_s\sigma_c^2(1 - \rho)R(q^{(1)}(m,n)\tau ) + \sigma_s^2\delta(m - n), \\
B_{13}(m,n) = \frac{\sigma_c^2}{2}\sqrt{E_sE_p}(1 - \rho)w_{q^{(1)}(n,m)}^H h \\
B_{14}(m,n) = \frac{\sigma_c^2}{2}\sqrt{E_sE_p}(\rho - 1)w_{q^{(1)}(n,m)+1}^H h, \\
B_{22}(m,n) = E_s\sigma_c^2(1 + \rho)R(q^{(1)}(m,n)\tau ) + \sigma_s^2\delta(m - n) \\
B_{23}(m,n) = \frac{\sigma_c^2}{2}\sqrt{E_sE_p}(1 + \rho)w_{q^{(1)}(n,m)-1}^H h, \\
B_{24}(m,n) = \frac{\sigma_c^2}{2}\sqrt{E_sE_p}(1 + \rho)w_{q^{(1)}(n,m)}^H h \\
B_{33}(m,n) = h^H \left( \frac{E_p\sigma_c^2}{2}D_{q^{(1)}(m,n)} + \sigma_p^2\delta_{q^{(1)}(m,n)} \right) h, \\
B_{34}(m,n) = \frac{\rho E_p\sigma_c^2}{2}h^H D_{q^{(1)}(m,n)-1} h, \\
q^{(1)}(m,n) = g^{(1)}(m) - g^{(1)}(n) \text{ and } \delta_e \text{ is a square matrix of size } 2M + 1 \text{ with ones on the } e^{th} \text{ diagonal.}
\\
3.5 \text{ ML decoder}
\\
In this section, the performance of the conventional Alamouti space-time code, when the ML space-time decoder (ML-STD) is used, as well as the performance of the CCA,
when the joint Alamouti and convolutional ML decoder (JML) is used, are derived.

### 3.5.1 Alamouti space-time code with ML-STD

Unlike the LC-STD, the ML-STD does not assume that the channels are quasi-static. Without loss of generality, it is assumed that the transmitted sequence is $[s_1 \ s_2] = [1 \ 1]$ and $P_{[\hat{s}_1 \ \hat{s}_2]}$ denotes the probability that the decoder chooses $[\hat{s}_1 \ \hat{s}_2]$ as the output instead of $[1 \ 1]$. The upper bound of the block error probability of the Alamouti space-time code can be written as

$$P_{\text{Alamouti}} \leq P_{[1 - 1]} + P_{[-1 1]} + P_{[-1 -1]}. \quad (3.25)$$

Let $\tilde{P}_1 = P_{[1 - 1]}, \tilde{P}_2 = P_{[-1 1]}$ and $\tilde{P}_3 = P_{[-1 -1]}$, the probabilities $\tilde{P}_i$ for $i = 1, \ldots, 3$ can be simplified to

$$\tilde{P}_i = \Pr \left\{ \Re[z^{(i)}] < 0 \right\}, \quad (3.26)$$

where

$$z^{(1)} = -\sqrt{\frac{E_s}{2}} (r_{s,1}\hat{\beta}_1^* - r_{s,2}\hat{\alpha}_2^*) + \frac{E_s}{2} (\hat{\alpha}_1\hat{\beta}_1^* - \hat{\alpha}_2\hat{\beta}_2^*)$$

$$z^{(2)} = \sqrt{\frac{E_s}{2}} (r_{s,1}\hat{\alpha}_1^* + r_{s,2}\hat{\beta}_2^*) + \frac{E_s}{2} (\hat{\alpha}_1\hat{\beta}_1^* - \hat{\alpha}_2\hat{\beta}_2^*)$$

$$z^{(3)} = \sqrt{\frac{E_s}{2}} r_{s,1}(\hat{\alpha}_1 - \hat{\beta}_1)^* + \sqrt{\frac{E_s}{2}} r_{s,2}(\hat{\alpha}_2 + \hat{\beta}_2)^*.$$

By inspection, $\Re[z^{(i)}]$ can be written in the Hermitian quadratic form $\Re[z^{(i)}] = x_3^H Q_3^{(i)} x_3$, where

$$x_3 = \begin{bmatrix} r_{s,1} & r_{s,2} & \sqrt{\frac{E_s}{2}} \hat{\alpha}_1 & \sqrt{\frac{E_s}{2}} \hat{\alpha}_2 & \sqrt{\frac{E_s}{2}} \hat{\beta}_1 & \sqrt{\frac{E_s}{2}} \hat{\beta}_2 \end{bmatrix}^T, \quad (3.27)$$

and

$$Q_3^{(1)} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}.$$
\[ Q_3^{(2)} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \end{bmatrix} \]

\[ Q_3^{(3)} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \]

Since \( x_3 \) is a zero-mean complex Gaussian random vector, the characteristic function of \( x_3^H Q_3^{(i)} x_3 \) can be used with the residue theorem or the Gauss-Chebyshev approximation to find \( \tilde{P}_i \). In addition, the covariance matrix of \( x_3 \) can be found to be

\[ \Sigma_3 = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{12}^H & C_{22} & C_{23} \\ C_{13}^H & C_{23}^H & C_{22} \end{bmatrix}, \tag{3.28} \]

where

\[ C_{11} = E_s \sigma_c^2 \begin{bmatrix} 1 - \rho + \gamma_s^{-1} & 0 \\ 0 & 1 + \rho + \gamma_s^{-1} \end{bmatrix}, \]

\[ C_{12} = \frac{E_s \sigma_c^2}{2} \sqrt{\frac{E_p}{2}} \begin{bmatrix} (1 - \rho)w_0^H h & (1 - \rho)w_1^H h \\ (1 + \rho)w_{-1}^H h & (1 + \rho)w_0^H h \end{bmatrix}, \]

\[ C_{13} = \frac{E_s \sigma_c^2}{2} \sqrt{\frac{E_p}{2}} \begin{bmatrix} (\rho - 1)w_0^H h & (\rho - 1)w_1^H h \\ (\rho + 1)w_{-1}^H h & (\rho + 1)w_0^H h \end{bmatrix}, \]

\[ C_{23} = \frac{E_s E_p \sigma_c^2}{4} \begin{bmatrix} h^H D_0 h & h^H D_{-1} h \\ h^H D_1 h & h^H D_0 h \end{bmatrix}, \]

\[ C_{22} = \frac{E_s E_p \sigma_c^2}{4} \begin{bmatrix} h^H (D_0 + 2\gamma_p^{-1} I_{2M+1}) h & h^H (D_{-1} + 2\gamma_p^{-1} \delta_{-1}) h \\ h^H (D_{1} + 2\delta_1 \gamma_p) h & h^H (D_0 + 2\gamma_p^{-1} I_{2M+1}) h \end{bmatrix}. \]
3.5.2 CCA with joint Alamouti and convolutional ML decoder

Ideally, the convolutional code and the Alamouti space-time code used in the CCA system should be decoded jointly by the ML decoder to minimize the PEP. However, in reality, large block sizes of convolutional codes cause the iterative space-time and convolutional decoder, which performs very close to the joint ML decoder, to be too complex to implement. The analysis for this decoder is included here mainly to define the theoretical limit of the system performance and to assess the gains relative to the use of the LC-ML decoder.

Similar to the LC-ML described in section 3.4.2, the PEP of the JML is also a function of the transmitted sequence, the error sequence and the interleaving structure. Therefore, the same simplifications are made here to simplify the analysis. Assuming that the transmitted sequence is an all-one BPSK sequence and the Hamming weight of the error codeword is \( d \), the PEP can be written as

\[
P_2 = \Pr \{ x_4^H Q_4 x_4 < 0 \}
\]

(3.29)

where \( x_4 = [x_4^{(1)} \ x_4^{(2)} \ x_4^{(3)} \ x_4^{(4)} \ x_4^{(5)} \ x_4^{(6)}]^T \) and \( x_4^{(i)} \) for \( i = 1, \ldots, 6 \) are row vectors of length \( d \) with

\[
x_4^{(1)}(m) = r_{s,g^{(1)}(m)}, \quad x_4^{(2)}(m) = r_{s,g^{(2)}(m)};
\]

\[
x_4^{(3)}(m) = \sqrt{\frac{E_s}{2}} \hat{\alpha}_{g^{(1)}(m)};
\]

\[
x_4^{(4)}(m) = \sqrt{\frac{E_s}{2}} \hat{\beta}_{g^{(2)}(m)};
\]

\[
x_4^{(5)}(m) = \sqrt{\frac{E_s}{2}} \hat{\beta}_{g^{(1)}(m)};
\]

\[
x_4^{(6)}(m) = \sqrt{\frac{E_s}{2}} \hat{\beta}_{g^{(2)}(m)}.
\]

In addition, \( Q_4 \) is a square matrix which can be written as

\[
Q_4 = \frac{1}{2} \begin{bmatrix}
  0_d & 0_d & 0_d & 0_d & -I_d & 0_d \\
  0_d & 0_d & 0_d & I_d & 0_d & 0_d \\
  0_d & 0_d & 0_d & 0_d & I_d & 0_d \\
  0_d & I_d & 0_d & 0_d & 0_d & -I_d \\
  -I_d & 0_d & I_d & 0_d & 0_d & 0_d \\
  0_d & 0_d & 0_d & -I_d & 0_d & 0_d
\end{bmatrix}
\]

(3.30)

Since \( x_4 \) is a zero-mean complex Gaussian random vector, the PEP can be computed by means of the residue theorem (3.2) or the Gauss-Chebyshev approximation (3.4). From
the above expressions, the covariance matrix can be found to be

\[
\Sigma_4 = \begin{bmatrix}
B_{11} & 0_d & \sqrt{\frac{E_s}{2}} B_{13} & -\sqrt{\frac{E_s}{2}} B_{14} & -\sqrt{\frac{E_s}{2}} B_{15} & \sqrt{\frac{E_s}{2}} B_{14} \\
0_d & B_{22} & \sqrt{\frac{E_s}{2}} B_{23} & \sqrt{\frac{E_s}{2}} B_{24} & \sqrt{\frac{E_s}{2}} B_{23} & \sqrt{\frac{E_s}{2}} B_{24} \\
-\sqrt{\frac{E_s}{2}} B_{13}^H & \sqrt{\frac{E_s}{2}} B_{23}^H & \frac{E_s}{2} B_{33} & E_{34} & \frac{E_s}{2} B_{34} & \frac{E_s}{2} B_{34} \\
-\sqrt{\frac{E_s}{2}} B_{14}^H & \sqrt{\frac{E_s}{2}} B_{24}^H & E_{45} & \frac{E_s}{2} B_{33} & \frac{E_s}{2} B_{34} & \frac{E_s}{2} B_{34} \\
\sqrt{\frac{E_s}{2}} B_{14}^H & \sqrt{\frac{E_s}{2}} B_{24}^H & E_{46} & \frac{E_s}{2} B_{34} & \frac{E_s}{2} B_{34} & \frac{E_s}{2} B_{34} \\
\end{bmatrix}
\]

(3.31)

where \(B_{11}, B_{22}, B_{13}, B_{23}, B_{14}, B_{24}, B_{33}\) and \(B_{34}\) are as defined in section 3.4.2. In addition,

\[
E_{34}(m,n) = \frac{E_s}{2} h^H \left( \frac{E_p \sigma_c^2}{2} D_{q^{(1)}}(m,n) - 1 + \sigma_c^2 \delta_{q^{(1)}}(m,n) - 1 \right) h,
\]

\[
E_{45}(m,n) = \frac{\rho E_s E_p \sigma_c^2}{4} h^H D_{q^{(1)}}(m,n) + 1 h,
\]

\[
E_{46}(m,n) = \frac{\rho E_s E_p \sigma_c^2}{4} h^H D_{q^{(1)}}(m,n) h.
\]

### 3.6 Multi-path

Assuming that all paths fade independently and there is no multi-path interference, the extension from the single-path to the multi-path can be done similarly to the case without transmit diversity presented in chapter 2. The characteristic function of \(\Re[z]\) (or \(\Re[z^{(i)}]\)) of the multi-path channels is basically the product of the characteristic functions of all paths. The error probability can be found using the residue theorem or the Gauss-Chebyshev approximation suggested before. Also, in order to make a fair comparison, the fading variance should be normalized such that the sum of the fading variance in the multi-path case is equal to the fading variance in the flat-fading case, i.e., \(\sum_{l=1}^{L} \sigma_{c,l}^2 = \sigma_c^2\), where \(\sigma_{c,l}^2\) is the fading variance of the \(l^{th}\) path, and \(\sigma_c^2\) is the fading variance in the flat-fading case.
3.7 Numerical results

In this section, numerical examples are presented to illustrate the effect of the Doppler spread, the finite-depth interleaving and the channel estimation error on the performance of the conventional Alamouti space-time code and the CCA. In addition, the system performance is also compared with the performance of a single-input single-output (SISO) system to illustrate the scenarios when the SISO system is favorable. Lastly, analytical results of the concatenated system are compared with results from Monte Carlo simulations to verify the accuracy of the analysis.

Throughout this section, unless stated otherwise, it is assumed that the fading autocorrelation function is the zeroth-order Bessel function of the first kind $J_0(2\pi f_d \tau)$, which is derived from Jakes PSD [29], where $f_d$ denotes the normalized Doppler frequency. To provide a reasonable balance between performance and implementation complexity, an 11-tap Wiener filter is used as the channel estimator unless stated otherwise.

3.7.1 Alamouti space-time code

Fig. 3.3 and 3.4 compare the simulation results of $P_{\text{Alamouti}}$ to the performance bounds of $P_{\text{Alamouti}}$ employing the LC-STD and the ML-STD (from (3.10) and (3.25), respectively) for data SNR of 10 dB and 30 dB. The simulation results are represented by circles while the bounds for $f_d \tau = 0.0005$, 0.03 and 0.05 are represented by solid, dashed-dotted and dotted lines, respectively. Comparing the simulation results and the analytical bounds, it can be seen that the bounds are reasonably tight especially at low probability of error. Comparing fig. 3.3 and 3.4, it can be seen that the ML-STD performs better than the LC-STD, especially when the normalized Doppler frequency, the pilot SNR and the data SNR are large.

The effect of channel estimation error on the SISO system is well known [14]. Using the Wiener filter as the channel estimator, the bit error probability of the SISO system can be written as [14]

$$P_b = \frac{1}{2} \left( 1 - \sqrt{\frac{\mathbf{w}_0^H \left( \mathbf{D}_0 + \bar{\gamma}_p^{-1} \mathbf{I}_d \right)^{-1} \mathbf{w}_0}{1 + \bar{\gamma}_s^{-1}}} \right),$$

(3.32)

where $\mathbf{D}_0, \mathbf{w}_0, \bar{\gamma}_p, \bar{\gamma}_s$ are as previously defined.
Figure 3.3 Alamouti pairwise error probability versus pilot SNR for the LC-STD.

Figure 3.4 Alamouti pairwise error probability versus pilot SNR for the ML-STD.
Comparison between SISO and Alamouti STC (linear-combining scheme)

Figure 3.5 3-D plot of the bit error probability for the SISO and Alamouti STC with the linear-combining STD.

The linear combining scheme with multi-path [0.5 0.5]

Figure 3.6 The linear combining space-time decoder in the multi-path fading channels
In fig. 3.5, the bit error probability of the SISO system and the bit error probability of the Alamouti space-time coded system employing the LC-STD are compared in log$_{10}$ scale. The data SNR for both systems is set to $30$ dB. As expected, the LC-STD performs well when the channels are close to quasi-static ($f_d\tau$) and the CSI accuracy is reasonably good (large pilot SNR and small $f_d\tau$). However, the performance of the LC-STD degrades significantly when the pilot SNR decreases or $f_d\tau$ increases. It can also be seen that at small pilot SNR or large $f_d\tau$, the SISO system provides better performance than the Alamouti space-time coded system with the LC-STD.

Fig. 3.6 shows the performance of the Alamouti space-time code employing the LC-STD in the frequency-selective fading channels, assuming that the two resolvable paths have the same average energy, which is equal to half of the average energy of the single-path. Although the system still suffers from the Doppler spread, the performance gain from the path diversity is evidenced as seen in better performance floors at high pilot SNR, which are significantly lower than the performance floors corresponding to the flat-fading channels in fig. 3.3.

Although it is desirable to implement the Wiener filter as the channel estimator, it may be unrealistic due to lack of information and the complexity limitations in a practical system. To address the effect of a suboptimal channel estimator, the upper bounds (from (3.10) and (3.25)) of $P_{\text{Alamouti}}$ for a Wiener filter and for a moving average filter are illustrated in fig. 3.7 assuming data SNR $= 30$ dB. Both filters are assumed to have 11 taps and the upper bounds of $P_{\text{Alamouti}}$ corresponding to $f_d\tau = 0.0005$ and $0.05$ are illustrated. From the figure, it can be noticed that both filters provide similar performance for the two decoders when the Doppler spread is small ($f_d\tau = 0.0005$). This behavior is expected because, when the Doppler spread is small, the filter tap coefficients of the Wiener filter are similar to those of the moving average filter. However, as the Doppler spread increases, the errors in the moving average filter increases relative to the Wiener filter. Therefore, performance loss from using the moving average filter grows larger as the Doppler spread increases. Also note that the degradation resulting from the use of the moving average filter at high Doppler is significantly larger for the ML-STD than for the LC-STD. Therefore, in a practical system where a sub-optimal channel estimator is employed, the gain from using the ML-STD is mitigated by the effects of estimation errors.
Comparing the Wiener filters and the moving average filters

Figure 3.7 Alamouti pairwise error probability versus pilot SNR for different normalized Doppler frequencies and channel estimators.

Performance of the LC–STD using WCDMA parameters

Figure 3.8 Performance of the linear combining space-time decoder using WCDMA parameters
Although it has been shown that the LC-STD is greatly affected by the channel estimation error and the Doppler frequency, it should be pointed out that the choice of the space-time decoder should depend on the operating condition of the respective system. For example, the bit error performance of the Alamouti space-time coded system with the LC-STD corresponding to the UMTS-WCDMA 12.2 kbps downlink reference measurement channel assuming that the spreading factor is 128 (or $\tau^{-1} = 30$ KHz) is shown in fig. 3.8. In this figure, the bit error probability corresponding to the mobile speed of 3 kmph and 120 kmph for perfect CSI and noisy CSI (pilot SNR = 10 dB, Wiener filter with $M = 5$) are compared. Fig. 3.8 clearly shows that, if the desired bit error probability of the system is greater than $10^{-4}$, the degradation from using the LC-STD is very small.

3.7.2 Concatenated convolutional code and Alamouti space-time code

As described earlier, when a convolutional code is used as the outer code and the Alamouti space-time code is used as the inner code with an interleaver in between, the CCA system performance is improved due to coding and time diversity. In fig. 3.9, the PEP of the CCA employing the LC-ML and the JML are illustrated by the solid lines and the dashed-dotted lines, respectively. In this figure, the PEP is plotted as a function of $f_d\tau$ for the spatial correlation $\rho = 0, 0.25, 0.5, 0.75$ and 1. In addition, the PEP of the SISO system [33] (convolutional code without the Alamouti space-time code) is also illustrated in this figure by the dashed line. Comparing the performance of the LC-ML and the JML, it can be seen that the performance loss from using the LC-ML is very small. Since the complexity of LC-ML is significantly less than that of JML, it can be concluded that, considering the complexity and the performance, the LC-ML is a very good choice for a practical CCA system.

It can be seen from the figure that the PEP oscillates for both the SISO or the CCA systems. This is a result of the non-monotonically decreasing nature of the zeroth-order Bessel function of the first kind, which is the autocorrelation function corresponding to the Jakes PSD [33]. In addition, the estimation-diversity tradeoff can be seen in the CCA as well as the SISO system (see chapter 2 for more details on the tradeoff). It can also be seen that the CCA system performs much better than the SISO system when the spatial correlation $\rho$ is small. However, the performance degrades with $\rho$. When the
Comparison between SISO and CCA when the pilot SNR is large (71 dB)

Figure 3.9 Pairwise error probability versus normalized Doppler frequency for different antenna correlation factors assuming a large pilot SNR.

Comparison between SISO and CCA when the pilot SNR is small (0.97 dB)

Figure 3.10 Pairwise error probability versus normalized Doppler frequency for different antenna correlation factors assuming a small pilot SNR.
transmit antennas are fully correlated ($\rho = 1$), it can be seen that the performance of the CCA is worse than that of the SISO when the Doppler is large.

It is also important to point out that the pilot SNR used in fig. 3.9 is 71 dB, which is a very large SNR. Therefore, CSI accuracy is reliable until the channels change very rapidly and, among the values of $\rho$ illustrated, only the CCA with $\rho = 1$ performs worse than the SISO system. To illustrate the effect of the pilot SNR on the system performance when the LC-ML is used, the PEP when the pilot SNR is equal to 0.97 dB is illustrated in fig. 3.10. In this figure, it can be observed that, even when there is no spatial correlation ($\rho = 0$) the SISO outperforms the CCA at large Doppler spread. This example emphasizes the importance of the system environment on the performance of the CCA.

In most wireless communication systems, the total transmit energy is limited and allocating more energy to the pilot symbols means that less energy is allocated to the data symbols. It is important that the energy is allocated properly in order to achieve the best performance. In fig. 3.11, a contour plot illustrating the PEP in $\log_{10}$ scale is shown. In this figure, the dotted lines represent the PEP of the SISO system and
the solid lines represent the PEP of the CCA system employing the LC-ML. Instead of the Jakes PSD, the Gaussian PSD is used in this figure to avoid the confusion from the oscillation associated with the Jakes' PSD. From the figure, it can be seen that the best performance of the concatenated scheme is less than $10^{-12}$, which is better than the best performance of the SISO system. Comparing the performance of the two systems, the figure can be separated into two sections, the upper right triangle and the lower left triangle. The upper right triangle represents the situation where the channels change too rapidly and the pilot SNR is too small for the CCA to outperform the SISO. On the other hand, the lower left triangle represents the case where the channels change slowly enough and the pilot SNR is large enough for the CCA to outperform the SISO. Also note that the optimal data-to-pilot energy ratio in this figure is $E_s/E_p \approx 2$ or 0.3 in the $\log_{10}$ scale. However, keep in mind that this value applies for this particular system setting and is a function of other system parameters such as the interleaving depth, and the Hamming weight.

In addition, it can be seen that, when $E_s/E_p$ is small (less than $10^{-0.6}$ for this particular scenario), the PEP is less sensitive to $f_d\tau$ than when $E_s/E_p$ is large. The reason is that, when $E_s/E_p$ is small, the pilot SNR is large enough that CSI is still accurate at large $f_d\tau$, therefore, the PEP, which is a function of the CSI accuracy, is also not very sensitive to $f_d\tau$. However, the behavior is different when $E_s/E_p$ is large. In this case, the pilot SNR is small and increasing $f_d\tau$ can cause a significant degradation in CSI, therefore, the PEP is more sensitive to $f_d\tau$ in this region. Further investigation of this figure also reveals that the optimal $E_s/E_p$ decreases with $f_d\tau$. This is expected because higher pilot energy can help improve the CSI accuracy at high Doppler spread.

Lastly, to verify the accuracy of this analysis, analytical results are compared with results from Monte Carlo simulation in fig. 3.12, where $P_e$ denotes the bit error probability and $P_f$ denotes the block error probability of the convolutional code. The rate 1/3 convolutional code and the block interleaver specified in the UMTS-WCDMA standard [3] are used in the simulation. In addition, the minimum Hamming weight of the convolutional code is equal to eighteen and the constraint length is equal to nine. The number of information bits per block is assumed to be 220 and each code block is terminated with eight zeros to set the encoder back to the all-zero state at the end of each block. Fading coefficients are generated by method of exact Doppler spread (MEDS) suggested in [45] with Jakes PSD. The channel estimator is an 11-tap Wiener filter, the data SNR is set
Comparison between analytical and simulation results

- $P_e_{\text{simulation}}$
- $P_f_{\text{simulation}}$
- $P_e_{\text{analysis}}$
- $P_f_{\text{analysis}}$

Figure 3.12 Bit error probability and block error probability versus normalized Doppler frequency from Monte Carlo simulation and from the analysis.

To 2.22 dB and the pilot SNR is set to 0.97 dB. For the analytical results, the bit error probability and the frame error probability are calculated by taking into account the PEP corresponding to the five smallest Hamming weights ($d = 18, 20, 22, 24, 26$) [12]. From the figure, it can be seen that the analytical results and the simulation results match well.

### 3.8 Conclusions

In this chapter, the bit error probability of the Alamouti space-time code with the linear-combining scheme and the pairwise error probability of Alamouti space-time code with the ML space-time decoder are derived. By assuming that the channel estimates are calculated from noisy pilot signals and the channels are not quasi-static, the effects of the channel estimation error and the time-varying fading channels on the system performance are evaluated. The results have shown that, when the channel estimator is optimized and the pilot SNR is large, the ML space-time decoder is significantly more tolerant to fast fading channels than the linear-combining space-time decoder.
However, the difference diminishes when the channel estimator is not optimized. In addition, comparisons between the linear-combining space-time decoder and the SISO system reveal that the Alamouti space-time code does not always outperform the SISO system.

This chapter has also investigated the performance of the concatenated convolutional code and Alamouti space-time code. The decoding schemes considered are extensions of those of the conventional Alamouti space-time code, namely, the linear-combining space-time decoder with the ML convolutional decoder (LC-ML) and the joint Alamouti space-time and convolutional ML decoder (JML). The results have indicated that, although the LC-ML is significantly less complex than JML, the performance difference between these two decoders is not significant. In addition, the analytical results have emphasized the importance of various system parameters on the performance of the concatenated system and that, without desirable system environments, it may be more beneficial to use only the convolutional code rather than the more complex concatenated convolutional code and Alamouti space-time code. In addition, comparison between results from Monte Carlo simulation and results from the analysis has verified that the analysis can predict the system performance with high accuracy.

3.9 Acknowledgement

4

Performance of closed-loop transmit diversity with noisy channel estimates and finite-depth interleaved convolutional codes

4.1 Introduction

In recent years, transmit diversity (TD) has been proposed as a powerful technique to improve the performance of wireless communication systems. The impact of TD is evidenced in both the research community, where hundreds of TD related publications emerged, and the industrial community, where various TD schemes have recently been included as parts of various wireless communication standards [4, 55].

This chapter focuses on a particular TD algorithm called the closed-loop transmit diversity (CLTD) or the transmit antenna array (TxAA) [22]. As the name suggests, the transmitters using CLTD algorithms need feedback information sent from the receivers in order to operate. Two CLTD systems under consideration are the phase-amplitude CLTD (PA-CLTD) and the phase-only CLTD (PO-CLTD), both with two transmit antennas and one receive antenna. In both PA-CLTD and PO-CLTD, the data signal is multiplied by complex weights (different values for different antennas) prior to the
transmission and these complex weights are calculated by the receiver and sent to the transmitter via the feedback control channel. The goal of this process is to maximize the received data SNR by changing the phases or the magnitudes of the transmitted data signals (through the complex weights) such that, after going through the fading channels, data signals from different transmit antennas are combined in-phase at the receiver. For PA-CLTD, the complex weights may not have the same magnitude (transmit antennas may transmit with different data energy), however, for PO-CLTD, the complex weights must have the same magnitude (transmit antennas must transmit with the same data energy). Also note that PA-CLTD and PO-CLTD can be considered as the generalized versions of CLTD mode 2 and mode 1 described in 3GPP WCDMA-UMTS standard, respectively [4], and thus the results obtained can be used to assess the issues that must be addressed in order to implement these algorithms successfully.

The performance of CLTD systems is affected by many system imperfections, such as the feedback information error [44], the feedback quantization error [18], the feedback rate [24], the feedback delay [18,19,24,36,43,47] and the channel state information (CSI) accuracy [15,36,58]. However, this chapter focuses on only two system imperfections, namely, the CSI accuracy and the feedback delay while assuming the rest to be perfect. The importance of these two parameters will be shown by the analysis presented herein that, although the rest of parameters are assumed to be perfect, just the imperfections of CSI and the feedback delay can cause CLTD systems to degrade considerably such that it is outperformed by the system without transmit diversity.

Various works have been done to analyze the performance of CLTD systems with noisy CSI and the feedback delay. For example, SNR degradation as a function of the feedback delay for PA-CLTD and PO-CLTD was derived in [47], while the performance of PA-CLTD with the feedback delay assuming perfect coherent detection was derived in [18,19,24]. Approximations of the BEP for PA-CLTD assuming that channel estimation noise is Gaussian were found in [15,58] and the exact BEP of PA-CLTD as a function of the feedback delay with and without noisy CSI was derived in [36,43]. Unlike most of the existing papers, this chapter assumes that channel estimates are calculated by FIR filters from noisy pilot symbols [14], therefore, the effect of the Doppler spread (via the CSI accuracy and the feedback delay) on the BEP is captured in the analysis. In addition, unlike [18,19,24], this analysis assumes that the coherent detection is not perfect because of the channel estimation error. It should also be noted that, unlike [36]
where the BEP of PA-CLTD with any number of transmit antennas is in an integral form, this chapter focuses on systems with only two transmit antennas, therefore, the closed-form expressions of the BEP for uncoded PA-CLTD and PO-CLTD with noisy CSI and feedback delay can be found in this analysis. In addition to the uncoded CLTD systems, since error correcting codes are used in most wireless communication systems and CLTD is often used as a supplement to the existing codes, the uncoded analysis is extended to the analysis for the CLTD systems with finite-depth convolutional codes.

The chapter is organized as follows. The system model and assumptions are described in section 4.2. The bit error probability (BEP) of PA-CLTD and the pairwise error probability (PEP) of PA-CLTD with convolutional codes are derived in section 4.3. In section 4.4, the BEP of PO-CLTD and the PEP of PO-CLTD with convolutional codes are derived. Discussions of results and conclusions are presented in section 4.5 and section 4.6, respectively. Finally, section 4.7 addresses the acknowledgement.

### 4.2 System model

For the rest of this chapter, the following notation will be used. A lowercase bold letter denotes a vector and an uppercase bold letter denotes a matrix. The element in the $m^{th}$ row and the $n^{th}$ column of a matrix $X$ is denoted by $X(m, n)$ and the element in the $m^{th}$ row (column) of a column (row) vector $x$ is denoted by $x(m)$. The superscripts $\ast$, $T$, $H$ denote the complex conjugate, the matrix transpose and the matrix Hermitian operation, respectively. The determinant of a matrix $X$ is denoted by $|X|$. The length $m$ column column vector of ones, the square identity matrix, the square zero matrix and the anti-diagonal matrix of order $m$ are denoted by $\mathbf{1}_m$, $\mathbf{I}_m$, $\mathbf{0}_m$ and $\mathbf{J}_m$, respectively.

For simplicity, this chapter considers a BPSK DS-CDMA system with two transmit antennas (antenna A and antenna B) at the transmitter and one receive antenna at the receiver. However, extensions to other systems such as TDMA with a pilot training sequence can be done with minor modifications. The baseband representation of the system under consideration is shown in fig. 4.1. In this system, it is assumed that the pilot sequences transmitted from antenna A and antenna B use distinct orthogonal codes, while the data sequences use the same orthogonal code. Therefore, three channels detected at the receiver are two pilot channels and one combined data channel. For
simplicity, all of the orthogonal codes are assumed to have the same spreading gain $N$, which results in the symbol period $\tau$. The total energy per data symbol is assumed to be $E_s$ and the total energy per pilot symbol is assumed to be $E_p$ ($E_p/2$ for each antenna). Prior to the transmission, the data signal corresponding to the symbol time index $k$, denoted by $s_k$, is multiplied by the weights $w_{1,k}$ and $w_{2,k}$, both of which are determined by the receiver. It is also assumed that the weights are normalized such that the total energy per data symbol is preserved at $E_s$, i.e., $|w_{1,k}|^2 + |w_{2,k}|^2 = 1$.

The channels are assumed to be time-varying Rayleigh flat-fading channels, where the fading coefficients are zero-mean circularly symmetric complex Gaussian random variables with autocorrelation $\sigma_c^2 R(t)$ and $R(0)$ is normalized to unity. With sufficient transmit antenna spacing, the fading coefficients from different transmit antennas are assumed to be uncorrelated. The fading coefficients corresponding to antenna A and antenna B at the symbol time index $k$ are denoted by $\alpha_k$ and $\beta_k$, respectively. Therefore, the fading correlations can be expressed as $\frac{1}{2}E[\alpha_k \alpha_{k-m}^*] = \frac{1}{2}E[\beta_k \beta_{k-m}^*] = \sigma_c^2 R(\tau m)$, and $E[\alpha_k \beta_{k-m}^*] = 0$. The noise in the pilot channels and the data channel at the symbol time index $k$, denoted by $n_{p,k}^{(1)}$, $n_{p,k}^{(2)}$ and $n_{s,k}$, respectively, are assumed to be zero-mean circularly symmetric complex white Gaussian noise with variance $\sigma_p^2, \sigma_p^2$ and $\sigma_s^2$, respectively.

At the receiver, the channel estimation is performed by two $M$-tap causal FIR filters, whose outputs are the channel estimates $\hat{\alpha}_k$ and $\hat{\beta}_k$. These channel estimates are used
to find the feedback weights and to perform coherent detection. Note that since this chapter focuses on the effect of the channel estimation error and the feedback delay on the performance, other system imperfections such as the feedback quantization and the feedback error are neglected.

Using the system model explained above, the received data signal, denoted by \( r_{s,k} \), the received pilot signal from antenna A and antenna B, denoted by \( r_{p,k}^{(1)} \) and \( r_{p,k}^{(2)} \), respectively, can be written as

\[
\begin{align*}
\beta_{k} \sqrt{2/2} + n_{p,k}^{(2)} &\rightarrow r_{p,k}^{(2)} = [r_{p,k}^{(2)}_{p,k-(M-1)} \cdots r_{p,k}^{(2)}]^{T}. \\
\end{align*}
\]

In addition, the channel estimates from the \( M \)-tap filters can be written as \( \hat{\alpha}_{k} = h^{H} r_{p,k-d}^{(1)} \) and \( \hat{\beta}_{k} = h^{H} r_{p,k-d}^{(2)} \), where \( h \) denotes the real-valued filter coefficient vector \([h_{M} \ldots h_{1}]^{T}\) and \( \delta \) denotes the lag time due to processing delay and feedback delay. After the data signal bearing \( s_{k} \) is received at the receiver, the same \( \hat{\alpha}_{k} \) and \( \hat{\beta}_{k} \) that have been used to calculate the weight \( w_{1,k} \) and \( w_{2,k} \) are then used to demodulate the signal.

### 4.3 Phase-Amplitude CLTD

#### 4.3.1 Without Convolutional Codes

The optimal weights, which maximize the received SNR, when the two transmit antennas are allowed to transmit with different signal energy can be expressed as

\[
\begin{align*}
\hat{w}_{1,k}^{(PA)} &= \frac{\hat{\alpha}_{k}^{*}}{\sqrt{|\hat{\alpha}_{k}|^{2} + |\hat{\beta}_{k}|^{2}}} \quad \text{and} \quad \hat{w}_{2,k}^{(PA)} &= \frac{\hat{\beta}_{k}^{*}}{\sqrt{|\hat{\alpha}_{k}|^{2} + |\hat{\beta}_{k}|^{2}}}. \\
\end{align*}
\]

The bit error probability (BEP), which is the probability that the receiver chooses \( \hat{s}_{k} = -s_{k} \), can then be written as

\[
P_{e}^{(PA)} = Pr\{z_{1}^{(PA)} < 0\},
\]

where \( z_{1}^{(PA)} = x_{1}^{H} Q_{1} x_{1} + n_{1}, x_{1} = [\alpha_{k} \hat{\alpha}_{k} \beta_{k} \hat{\beta}_{k}]^{T} \) and \( n_{1} \) is a real-valued Gaussian noise with variance \( x_{1}^{H} Q_{2} x_{1} \). In addition,
\[ Q_1 = \frac{1}{2} (I_2 \otimes J_2) \text{ and } Q_2 = \frac{\sigma_s^2}{E_s} \left( I_2 \otimes \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right), \quad (4.6) \]

where \( \otimes \) denotes the Kronecker product operation \([1]\).

The joint probability density function of \( x_1 \) and \( n_1 \) can be found from Bayes’ rule to be

\[
p(x_1, n_1) = p(n_1|x_1) p(x_1) = \frac{\exp \left\{ -\frac{n_1^2}{2x_1^H Q_2 x_1} \right\}}{\sqrt{2\pi x_1^H Q_2 x_1}} \times \frac{\exp \left\{ -\frac{x_1^H \Sigma_1^{-1} x_1}{2} \right\}}{(2\pi)^{\frac{1}{2}}|\Sigma_1|}, \quad (4.7)\]

where \( \Sigma_1 \) is the covariance matrix of \( x_1 \) and it is equal to

\[
\Sigma_1 = \sigma_c^2 \left( I_2 \otimes \begin{bmatrix} 1 \\ \sqrt{\frac{E_s}{2}} h^H u_{0,\delta} \\ h^H \left( \frac{E_s}{2} D_0 + \frac{\sigma^2}{\sigma_e^2} I_M \right) h \end{bmatrix} \right), \quad (4.8)\]

where \( D_e(m, n) = R((e + m - n)\tau) \) and \( u_{e,\delta} = [R((e - \delta - M + 1)\tau) \ldots R((e - \delta)\tau)]^T \).

The characteristic function of \( z_1^{(PA)} \) can then be calculated using \( p(x, n_1) \) to be

\[
\tilde{\Phi}_{z_1^{(PA)}}(s) = \int_{x_1} p(x_1) e^{s x_1^H Q_2 x_1} \int_{n_1} p(n_1|x_1) e^{s n_1} dn_1 \, dx_1 = |I_4 - (2s Q_1 + s^2 Q_2) \Sigma_1|^{-1}. \quad (4.9)\]

Notice from (4.9) that there are two poles of order 2 at

\[
\sqrt{\frac{2}{E_s}} \left( \frac{-\zeta_2 + \sqrt{\zeta_2^2 + \zeta_1}}{\sigma_c^2 \z_1} \right) \text{ and } -\sqrt{\frac{2}{E_s}} \left( \frac{\zeta_2 + \sqrt{\zeta_2^2 + \zeta_1}}{\sigma_c^2 \z_1} \right),
\]

where \( \zeta_1 = h^H \Upsilon h, \zeta_2 = u_{0,\delta}^H h, \Upsilon = 2(\bar{\gamma}_s^{-1} + 1) (D_0/2 + \bar{\gamma}_p^{-1} I_M) - u_{0,\delta} u_{0,\delta}^H, \) and \( \bar{\gamma}_s = E_s \sigma_c^2 / \sigma_e^2 \) is the data SNR, \( \bar{\gamma}_p = \sigma_c^2 / \sigma_p^2 \) is the pilot SNR. Since the residue is invariant to the scaling of poles \([31]\), a modified characteristic function that gives the same probability of error as \( \Phi_{z_1^{(PA)}}(s) \) is considered instead. The modified characteristic function with poles scaled by \( \sigma_c^2 \sqrt{E_s}/2 \) can be expressed as

\[
\tilde{\Phi}_{z_1^{(PA)}}(s) = \left( \frac{-\zeta_1}{-\zeta_2 + \sqrt{\zeta_2^2 + \zeta_1}} \right)^{-2} \left( \frac{-\zeta_1}{-\zeta_2 + \sqrt{\zeta_2^2 + \zeta_1}} \right)^{-2} \quad (4.10)
\]

Knowing the locations of the poles, \( P_c^{(PA)} \) can be found from the residue theorem

\[
Pr\{z_1^{(PA)} < 0\} = -\sum_{m=1}^{N_k} \text{Res} \left[ \frac{-\sqrt{\frac{2}{E_s}} \left( \frac{-\zeta_2 + \sqrt{\zeta_2^2 + \zeta_1}}{\sigma_c^2 \z_1} \right)}{s} \right. \text{ at } q_m, \quad (4.11)
\]
where \( N^k \) denotes the number of the negative poles of \( \Phi_1(PA) / s \) and \( q_m \) denotes the \( m^{th} \) negative pole of \( \Phi_1(PA) / s \). Since \( \Sigma_1 \) is a covariance matrix, it is positive semi-definite and, consequently, \( 2h^H(D_0/2 + \bar{\gamma}^{-1}_p I_M)h \geq |h^H u_{0,\delta}|^2 \). As a result, it can be concluded that \( \zeta_1 \) is greater than zero and that there is a negative pole of order two at \( -\zeta_1^{-1}(\zeta_2 + \sqrt{\zeta_2^2 + \zeta_1}) \). The bit error probability calculated from the residue theorem becomes

\[
P_e^{(PA)} = \frac{3\zeta_3 + 1}{(\zeta_3 + 1)^3},
\]

where \( \zeta_3 = \zeta_1^{-1}(\zeta_2 + \sqrt{\zeta_2^2 + \zeta_1})^2 \).

### Wiener Filter

Consider a special case when the Wiener filters are used as the channel estimators. Substituting the expression of the filter coefficient vector of the Wiener filter,

\[
h = \frac{1}{\sqrt{2}} \left( \frac{D_0}{2} + \bar{\gamma}^{-1}_p I_M \right)^{-1} u_{0,\delta},
\]

into (4.12), the closed-form expression of the bit error probability becomes

\[
P_e^{(PA,\text{Wiener})} = \frac{1}{2} \left( 1 + \frac{1}{2} \sqrt{\frac{\bar{\gamma}_s E}{2(1 + \bar{\gamma}_s)}} \right) \left( 1 - \sqrt{\frac{\bar{\gamma}_s E}{2(1 + \bar{\gamma}_s)}} \right)^2,
\]

where \( E = u_{0,\delta}^H(D_0/2 + \bar{\gamma}^{-1}_p I_M)^{-1} u_{0,\delta} \).

### Perfect CSI

When the channel estimates are perfect (\( \bar{\gamma}_p = \infty \) with the Wiener filters as the channel estimators), it can be shown that \( E \) is equal to 2. Thus, the closed-form expression of the BEP for the perfect CSI case can be written as

\[
P_e^{(PA,\text{perfCSI})} = \frac{1}{2} \left( 1 + \frac{1}{2} \sqrt{\frac{\bar{\gamma}_s}{1 + \bar{\gamma}_s}} \right) \left( 1 - \sqrt{\frac{\bar{\gamma}_s}{1 + \bar{\gamma}_s}} \right)^2.
\]

### 4.3.2 With Convolutional Codes

In addition to the spatial diversity, the system performance can also be improved by error correcting codes. One way to do this is to convolutionally code information bits and interleave these coded bits prior to applying a transmit diversity algorithm.
The received signal can be decoded by the ML decoder and the error performance of interest is the pairwise error probability (PEP) which is defined as the probability that the decoder chooses the coded sequence \( \mathbf{c}_j \) instead of the transmitted sequence \( \mathbf{c}_i \), where \( i \neq j \). Although the PEP is a function of the error pattern and the structure of the interleaver, finding the PEP for each error pattern for a specific interleaving structure is tedious and adds little insight into the overall system performance [48]. Therefore, similar simplifications to those used in chapter 2 and chapter 3 are used, i.e., that the transmitted codeword is an all-zero codeword, which is mapped to an all-one BPSK sequence \( \mathbf{c}_0 \), that the error codeword with \( d \) consecutive error symbols corresponds to the BPSK sequence \( \mathbf{c}_i (i \neq 0) \) and that an interleaving depth \( I \) of a block interleaver creates the same effect as separating consecutive symbol errors by \( I \) symbols [26]. As a result, the PEP depends only on the Hamming weight of the error codeword and the interleaving depth.

Given that the all-zero codeword is transmitted, the PEP corresponding to PA-CLTD, denoted by \( P_2^{(PA)} \), assuming that the ML decoder is used can be written as

\[
P_2^{(PA)} = \Pr \left\{ \Re \left[ \sum_{k=1}^{d} r_{s,kI} \sqrt{|\hat{\alpha}_{kI}|^2 + |\hat{\beta}_{kI}|^2} \right] < 0 \right\},
\]

(4.16)

where \( \Re[x] \) denotes the real part of \( x \). In addition, the real part in (4.16), denoted by \( z_2^{(PA)} \), can be simplified further to \( z_2^{(PA)} = \mathbf{x}_2^H \mathbf{Q}_3 \mathbf{x}_2 + n_2 \), where

\[
\mathbf{x}_2 = [\alpha_I \ldots \alpha_{dI} \hat{\alpha}_{dI} \beta_I \ldots \beta_{dI} \hat{\beta}_{dI} \beta_I \ldots \beta_{dI}]^T,
\]

(4.17)

\( n_2 \) is a zero-mean real-valued Gaussian random variable with variance \( \mathbf{x}_2^H \mathbf{Q}_4 \mathbf{x}_2 \).

Using the same approach as in section 4.3.1, the characteristic function of \( z_2^{(PA)} \) can be found to be

\[
\Phi_{z_2^{(PA)}}(s) = |\mathbf{I}_{4d} - (2s \mathbf{Q}_3 + s^2 \mathbf{Q}_4) \mathbf{\Sigma}_2|^{-1},
\]

(4.18)

where \( \mathbf{\Sigma}_2 \) denotes the covariance matrix of \( \mathbf{x}_2 \) and it is equal to \( \mathbf{\Sigma}_2 = \mathbf{I}_2 \otimes \hat{\mathbf{\Sigma}}_2 \), where

\[
\hat{\mathbf{\Sigma}}_2 = \sigma_c^2 \begin{bmatrix}
\mathbf{B}_{11} & \mathbf{B}_{12} \\
\mathbf{B}_{12}^H & \mathbf{B}_{22}
\end{bmatrix},
\]

(4.19)
with \( B_{11}(m, n) = R((m - n)I \tau) \), \( B_{22}(m, n) = h^H \left( \frac{E_p}{2} D_{(m-n)I} + \frac{\sigma_c^2}{\bar{\gamma}_s^2} \delta_{(m-n)I} \right) h \) and \( B_{12}(m, n) = \sqrt{\frac{E_p}{2}} u^H_{(n-m)I, \delta} h \). Using the definition of \( \Sigma_2 \) and the knowledge that the PEP is invariant to the scaling of the poles, the poles can be scaled by \( \sigma_c^2 \) and still provide the same PEP (see Appendix B for proof). After the scaling, the modified characteristic function can be simplified further to

\[
\tilde{\Phi}_{\tilde{z}_1^{(PA)}}(s) = \left| I_{2d} - s \begin{bmatrix} B_{12}^H & B_{22} \\ B_{11} + \bar{\gamma}_s^{-1} I_d & B_{12} \end{bmatrix} \right|^{-2}. \tag{4.20}
\]

With the knowledge of the characteristic function, the residue theorem explained in section 4.3.1 can be used to find the PEP. Alternatively, numerical results can also be found from the Gauss-Chebyshev approximation (suggested in [11] with a correction in [25]), i.e.,

\[
P_2^{(PA)} \approx \frac{1}{2m} \sum_{k=1}^{m} \left( \Re[\tilde{\Phi}_{\tilde{z}_1^{(PA)}}(\varrho_k)] + \tau_k \Im[\tilde{\Phi}_{\tilde{z}_1^{(PA)}}(\varrho_k)] \right), \tag{4.21}
\]

where \( \Im[x] \) denotes the imaginary part of \( x \), \( \varrho_k = \epsilon (1 + j \tau_k) \), \( \tau_k = \tan((2k - 1)\pi/4m) \) and, in general, \( m \) between 16 and 32 is sufficient [11]. In addition, \( \epsilon \) should lie between the left half-plane poles and the imaginary axis.

### 4.4 Phase-Only CLTD

Unlike PA-CLTD, PO-CLTD has the constraint that equal signal energy is transmitted from the two transmit antennas. With this energy constraint, it is expected that the performance of PO-CLTD is inferior to that of PA-CLTD.

#### 4.4.1 Without Convolutional Codes

For the PO-CLTD algorithm, the optimal weights that maximize the receive SNR are

\[
w_{1,k}^{(PO)} = \frac{\hat{\alpha}_k^*}{\sqrt{2} |\hat{\alpha}_k|} \quad \text{and} \quad w_{2,k}^{(PO)} = \frac{\hat{\beta}_k^*}{\sqrt{2} |\hat{\beta}_k|}. \tag{4.22}
\]

Therefore, the bit error probability can be written as

\[
P_e^{(PO)} = \Pr\{z_1^{(PO)} < 0\}, \tag{4.23}
\]
where \( z_{1}^{(PO)} = x_{1} + x_{2} + n_{3}, \) \( x_{1} = \mathbb{R}[\alpha_{k} \hat{\alpha}_{k}^{*}]/|\hat{\alpha}_{k}|, \) \( x_{2} = \mathbb{R}[\beta_{k} \hat{\beta}_{k}^{*}]/|\hat{\beta}_{k}| \) and \( n_{3} \) is a zero-mean real-valued Gaussian random variable with variance \( 2\sigma_{s}^{2}/E_{s}. \)

Since \( x_{1} \) and \( x_{2} \) are independent identically distributed due to the assumption that the transmit antennas are far enough from each other such that fading coefficients from different transmit antennas are uncorrelated, the characteristic function of \( z_{1}^{(PO)} \) is the product of the characteristic functions of \( x_{1}, x_{2} \) and \( n_{3}. \) From Appendix C, the characteristic functions of \( x_{1} \) and \( x_{2} \) are equal to

\[
\Phi_{x_{1}}(iv) = \Phi_{x_{2}}(iv) = (\Psi_{i}(v) + i\Psi_{q}(v)) \exp \left\{ -\frac{v^{2}|\Sigma_{3}|}{2\Sigma_{3}(2, 2)} \right\}, \tag{4.24}
\]

where

\[
\Psi_{i}(v) = 1 + v\rho \sqrt{\pi} \ e^{-v^{2}\rho^{2}} \text{erfi}(v\rho), \quad \Psi_{q}(v) = v\rho \sqrt{\pi} \ e^{-v^{2}\rho^{2}},
\]

and \( \rho = \Sigma_{3}(1, 2)/\sqrt{2\Sigma_{3}(2, 2)}. \) In addition, \( \text{erfi}(x) \) denotes the imaginary error function of \( x \) \([1], i.e.,\)

\[
\text{erfi}(y) = \frac{2}{\sqrt{\pi}} e^{y^{2}} \int_{0}^{\infty} e^{-t^{2}} \sin(2yt) \, dt \quad \text{where} \quad y \in \mathbb{R}, \tag{4.25}
\]

\( \Sigma_{3} \) denotes the covariance matrix of \([\alpha_{k} \ \hat{\alpha}_{k}]^{T} \) or \([\beta_{k} \ \hat{\beta}_{k}]^{T} \) and is equal to

\[
\Sigma_{3} = \sigma_{c}^{2} \begin{bmatrix}
1 & \frac{1}{\sqrt{2}} \ h^{H} u_{0, \delta} \\
\frac{1}{\sqrt{2}} \ h^{H} u_{0, \delta} & \ h^{H} \left( \frac{1}{2} D_{0} + \bar{\gamma}_{p}^{-1} I_{M} \right) \ h
\end{bmatrix}. \tag{4.26}
\]

Therefore, the characteristic function of \( z_{1}^{(PO)}, \) denoted by \( \Phi_{z_{1}^{(PO)}}(iv), \) can be expressed as

\[
\Phi_{z_{1}^{(PO)}}(iv) = \Phi_{x_{1}}(iv)\Phi_{x_{2}}(iv)\Phi_{n_{3}}(iv) = g_{i}(v) + ig_{q}(v), \tag{4.27}
\]

where

\[
g_{i}(v) = (\Psi_{i}(v)^{2} - \Psi_{q}(v)^{2}) \exp \left\{ -v^{2} \left( \frac{\sigma_{s}^{2}}{E_{s}} + \frac{|\Sigma_{3}|}{\Sigma_{3}(2, 2)} \right) \right\}, \tag{4.28}
\]

\[
g_{q}(v) = 2\Psi_{i}(v)\Psi_{q}(v) \exp \left\{ -v^{2} \left( \frac{\sigma_{s}^{2}}{E_{s}} + \frac{|\Sigma_{3}|}{\Sigma_{3}(2, 2)} \right) \right\}. \tag{4.29}
\]

As a result, the probability density function \( z_{1}^{(PO)} \) can be found from the characteristic function to be

\[
p(z_{1}^{(PO)}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (g_{i}(v) + ig_{q}(v)) e^{-ivz_{1}^{(PO)}} \, dv = p_{e}(z_{1}^{(PO)}) + p_{o}(z_{1}^{(PO)}), \tag{4.30}
\]
where \( p_e(z_1^{(PO)}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g_i(v)e^{-ivz_1^{(PO)}} dv \) and \( p_0(z_1^{(PO)}) = \frac{i}{2\pi} \int_{-\infty}^{\infty} g_q(v)e^{-ivz_1^{(PO)}} dv \) denote the even part and the odd part of \( p(z_1^{(PO)}) \), respectively. The probability that \( z_1^{(PO)} \) is less than zero can be found by integrating the probability density function of \( z_1^{(PO)} \) from \(-\infty\) to zero. The result of this integration turns out to be (see Appendix D for details),

\[
P_e^{(PO)} = \Pr\{z_1^{(PO)} < 0\} = \frac{1}{2} - \rho \left( \frac{\sigma_s^2}{E_s} + \Sigma_3(1,1) \right)^{-1} \sqrt{\rho^2 + \frac{\sigma_s^2}{E_s} + \frac{\Sigma_3}{\Sigma_3(2,2)}},
\]

(4.31)

Substituting (4.26) into (4.31), the closed-form expression of the bit error probability can be obtained

\[
P_e^{(PO)} = \frac{1}{2} - \frac{u_{0,\delta}^H}{\sqrt{\mathcal{M}}} \sqrt{1 - \frac{|u_{0,\delta}^H|^2}{\mathcal{M}}},
\]

(4.32)

where \( \mathcal{M} = 4(1 + \bar{\gamma}_s^{-1})h^T \left( \frac{1}{2} D_0 + \bar{\gamma}_p^{-1} I_M \right) h \).

**Wiener Filter**

Using the Wiener filters described in (4.13) as the channel estimators, the closed-form expression for the bit error probability of PO-CLTD can be expressed as

\[
P_e^{(PO,\text{Wiener})} = \frac{1}{2} \left( 1 - \frac{\bar{\gamma}_s \mathcal{E}}{1 + \bar{\gamma}_s} \left( 1 - \frac{\bar{\gamma}_s \mathcal{E}}{4(1 + \bar{\gamma}_s)} \right) \right),
\]

(4.33)

where \( \mathcal{E} \) is the same as the one defined in section 4.3.1.

**Perfect CSI**

Assuming perfect CSI, where \( \bar{\gamma}_p = \infty \) and the Wiener filters are used as the channel estimators, it can be shown that \( \mathcal{E} = 2 \). Therefore, the bit error probability becomes

\[
P_e^{(PO,\text{perfCSI})} = \frac{1}{2} \left( 1 - \frac{2\bar{\gamma}_s}{1 + \bar{\gamma}_s} \left( 1 - \frac{\bar{\gamma}_s}{2(1 + \bar{\gamma}_s)} \right) \right),
\]

(4.34)

which is the same result as the result from the equal-gain combining with two diversity branches described in [64].

### 4.4.2 With Convolutional Codes

Similar to PA-CLTD with convolutional codes described in section 4.3.2, the PEP of PO-CLTD with convolutional codes is a function the error sequence and the interleaving
structure. Therefore, the same assumptions are made here to simplify the analysis. Assuming that the transmitted codeword is an all-zero codeword, the interleaving depth is \( I \) and the Hamming weight of the error codeword is \( d \), the PEP can be written as

\[
P_2^{(PO)} = \Pr \left\{ \mathbb{R} \left[ \sum_{k=1}^{d} r_{s,kI} (|\hat{\alpha}_{kI}| + |\hat{\beta}_{kI}|) \right] < 0 \right\}
\]

\[
= \Pr \left\{ n_4 + \sum_{k=1}^{d} (x_{3,k} + x_{4,k}) < 0 \right\},
\]

where \( x_{3,k} = (1 + f_k) \Re \{ \alpha_{kI} \hat{\alpha}_{kI}^* \} \), \( x_{4,k} = (1 + 1/f_k) \Re \{ \beta_{kI} \hat{\beta}_{kI}^* \} \) and \( f_k = |\hat{\beta}_{kI}|/|\hat{\alpha}_{kI}| \). In addition, \( n_4 \) is a zero-mean real-valued Gaussian random variable with variance \( (2\sigma^2_s/E_s) \sum_{k=1}^{d} (|\hat{\alpha}_{kI}| + |\hat{\beta}_{kI}|)^2 \). Let

\[
\hat{\alpha} = [\hat{\alpha}_I \ldots \hat{\alpha}_{dI}]^T, \quad \hat{\beta} = [\hat{\beta}_I \ldots \hat{\beta}_{dI}]^T
\]

\[
\tilde{\alpha} = [\alpha_I \ldots \alpha_{dI} \hat{\alpha}_I \ldots \hat{\alpha}_{dI}]^T, \quad \tilde{\beta} = [\beta_I \ldots \beta_{dI} \hat{\beta}_I \ldots \hat{\beta}_{dI}]^T
\]

and \( \Sigma_4 \) be the covariance matrix of \( \tilde{\alpha} \) or \( \tilde{\beta} \). (4.35) can be simplified to (see Appendix E for details)

\[
P_2^{(PO)} = E_{\hat{\alpha},\hat{\beta}} \left[ Q \left( \frac{\varphi_1 + \varphi_2}{\sqrt{\chi_1 + \chi_2 + \sigma_2^2}} \right) \right],
\]

where

\[
\chi_1 = a^H \tilde{\Sigma}_4 a, \quad \chi_2 = b^H \tilde{\Sigma}_4 b, \quad \sigma_2^2 = (2\sigma^2_s/E_s)a^H a
\]

\[
\varphi_1 = \Re [a^H C_{12} C_{22}^{-1} \tilde{\alpha}] \quad \text{and} \quad \varphi_2 = \Re [b^H C_{12} C_{22}^{-1} \tilde{\beta}].
\]

Moreover, \( a \) is a column vector where \( a(k) = (|\hat{\alpha}_{kI}| + |\hat{\beta}_{kI}|) \exp \{ i \angle \hat{\alpha}_{kI} \} \), \( b \) is a column vector where \( b(k) = (|\hat{\alpha}_{kI}| + |\hat{\beta}_{kI}|) \exp \{ i \angle \hat{\beta}_{kI} \} \), the operation \( \angle x \) denotes the phase of the complex number \( x \) and \( \tilde{\Sigma}_4 = C_{11} - C_{12} C_{22}^{-1} C_{21} \). The covariance matrix of \( \tilde{\alpha} \) or \( \tilde{\beta} \), denoted by \( \Sigma_4 \), is equal to

\[
\Sigma_4 = \begin{bmatrix} C_{11} & C_{12} \\ C_{12}^H & C_{22} \end{bmatrix},
\]

where \( C_{11}(m,n) = \sigma^2_c R ((m-n)I \tau) \), \( C_{22}(m,n) = \sigma^2_c h^H \left( \frac{1}{2} D_{(m-n)I} + \gamma_p^{-1} d_{(m-n)I} \right) h \) and 
\( C_{12}(m,n) = \sigma^2_c \frac{u^H}{\sqrt{2}} u_{(n-m)I}^{\delta} h \).

Due to the complexity of the PEP expression (4.36), there is no closed form solution for this PEP. However, numerical results can be found via the the Monte Carlo approximation [41].
4.5 Numerical Results

4.5.1 Comparison between PA-CLTD and PO-CLTD

A contour plot comparing the BEP of PA-CLTD and PO-CLTD as a function of the normalized CSI error variance and the data SNR $\tilde{\gamma}_s$ for one symbol lag time is shown in fig. 4.2. The value of the BEP for each contour line is listed on the right side of the figure. Assuming that Wiener filters are used as the channel estimators, the normalized CSI error variance is equal to

$$E[|\alpha_k - \hat{\alpha}_k|^2] = 1 - \frac{1}{2}w_H^H \left( \frac{D_0}{2} + \tilde{\gamma}_p^{-1}I_M \right)^{-1} w_d. \quad (4.38)$$

It is clearly seen from the figure that PA-CLTD performs better than PO-CLTD. This is expected because PO-CLTD has the signal power constraint that PA-CLTD does not have. The results also indicate that, although PA-CLTD requires much more feedback information than PO-CLTD, the gain from PA-CLTD is not significant. Therefore, it is not surprising that the closed-loop mode transmit diversity mode 2 (simplified PA-CLTD) was removed from the 3GPP UMTS-WCDMA standard while mode 1 (simplified

![Figure 4.2 Contour plot of the bit error probability for PA-CLTD and PO-CLTD.](image-url)
Figure 4.3 Bit error probability versus data SNR for different pilot SNR’s.

PO-CLTD) is still being used.

Although the performance difference between PA-CLTD and PO-CLTD in fig. 4.2 appears to be small, this is true only when the operating point is not close to the performance floor. For example, consider the CSI error variance -17.25 dB. It can be seen from fig. 4.2 that PO-CLTD needs as much as 5 dB of data SNR more than PA-CLTD to achieve $10^{-4}$ BEP. For a better understanding, fig. 4.3 illustrates this performance floor behavior for different pilot SNR values. This figure assumes that the channel has the Gaussian power spectral density (PSD), $f_d \tau = 0.0037$ (or 60kmph for the symbol rate 30kHz and the carrier frequency 2GHz), one symbol lag time and the Wiener filters are used as the channel estimator. It can be seen that the performance gain from using PA-CLTD instead of PO-CLTD is small when the BEP of interest is not close to the performance floor and the gain becomes significant near the performance floor. In addition to the numerical results, fig. 4.3 also illustrates the results from Monte Carlo simulation by the solid circles. It is clearly seen that the analytical results match the simulation results perfectly.
Figure 4.4 Bit error probability versus lag time for Gaussian PSD and Jakes PSD for different filter lengths.

4.5.2 Lag time

In addition to the filter length and the normalized Doppler frequency, the quality of CSI also depends on the lag time \( \delta \). Fig. 4.4 illustrates the performance degradation as a function of the lag time for channels with the Gaussian PSD and the Jakes PSD and for the number of Wiener filter taps \( M = 11 \) or 101. In addition, other parameters are set as follows: pilot SNR = 10 dB, data SNR = 10 dB and \( f_d \tau = 0.0037 \). It is clearly seen that the performance degrades with the lag time. In addition, the performance degrades faster when the channel estimation filters have smaller number of taps. It should also be noted that the degradation in the Jakes PSD case is not monotonically decreasing due to the non-monotonically decreasing characteristic of the autocorrelation function of the Jakes model (the zeroth-order Bessel function of the first kind). This oscillation is not seen in the Gaussian PSD case where the autocorrelation function is the monotonically decreasing Gaussian function.
4.5.3 Comparison between SISO, Alamouti space-time code and PA-CLTD

One important goal of this work is to compare the performance of different transmit diversity schemes with noisy CSI. This goal is achieved as shown in fig. 4.5 where the performance of a system without transmit diversity (SISO) [14], the performance of a system with the Alamouti space-time code (STC) employing the linear-combining scheme space-time decoder (LC-STD) (chapter 3) and the performance of a system with PA-CLTD with 5-symbol lag time are compared. It is assumed that data SNR is equal to 30 dB, the channels have the Gaussian PSD and that all of these systems use 11-tap Wiener filters to estimate the channel coefficients. However, due to the feedback constraint of PA-CLTD, the Wiener filters used in PA-CLTD are causal while the Wiener filters used in SISO and Alamouti STC are non-causal.

We can see from the figure that the performance of SISO degrades with $f_d\tau$, but the degradation is clearly not as severe as the Alamouti STC or PA-CLTD. The LC-STD used with the Alamouti STC operates based on the assumption that the channels are
quasi-static. When the channels change rapidly, the assumption is no longer true and the performance degrades significantly from this inaccurate assumption. In addition to the demodulation, the performance of the Alamouti STC also degrades with the pilot SNR because the LC-STD cannot completely eliminate the interference from the other symbol in the Alamouti codeword when the channel estimates are noisy. For the PA-CLTD system, large Doppler spread degrades the accuracy of CSI, which is used for both the weight calculation and the demodulation, and increases the sensitivity of the system performance on the lag time. Each system depends on the Doppler spread and the CSI accuracy in different ways and relative performance is not trivial. Our analysis allows a quantitative comparison between different TD systems and the results have shown that PA-CLTD is significantly more sensitive to the Doppler spread than the SISO system. Compared to the Alamouti STC, PA-CLTD outperforms the Alamouti STC when the pilot SNR is large or when the Doppler spread is small, however, PA-CLTD is not as robust as the Alamouti STC when both the pilot SNR and the Doppler spread are not ideal.

For completion, it should be noted that, in addition to the LC-STC, the ML space-time decoder (ML-STD) can also be used to decode the Alamouti STC. In fig. 4.5, the LC-STD is chosen because it is the system most likely to be used in practice due to its low complexity. However, since the ML-STD does not use the quasi-static assumption, the ML-STD is expected to perform better the LC-STD at large $f_d \tau$ (chapter 3).

### 4.5.4 Performance as a function of data-to-pilot energy ratio

In most wireless communication systems, the total transmit energy is limited and allocating more energy to the pilot symbols means that less energy is available for the data symbols. It is important that the energy is allocated properly in order to achieve the best performance. In fig. 4.6, a contour plot illustrating the effect of the data-to-pilot ratio and the normalized Doppler frequency on the PEP (in $\log_{10}$ scale) of PA-CLTD with 1-symbol (solid) and 5-symbol (dotted) lag times. In addition, it is assumed that $E_p + E_s = 11$, $\sigma_c^2 = 0.5$, $\sigma_s^2 = \sigma_p^2 = 0.01$ and 11-tap Wiener filters are used as the channel estimators. The Gaussian PSD is used in this figure to avoid the confusion from the oscillation associated with the Jakes PSD.

From the figure, it is clearly seen that 1-symbol lag time performs much better than
Effect of data-to-pilot energy ratio on PA-CLTD with 1 (dotted) or 5 (solid) symbols lag time

Figure 4.6 Contour plot of the pairwise error probability for different lag times

the 5-symbol lag time especially when the Doppler spread is large. For example, consider the crossing point between the solid line (-2) and the dotted line (-8) at $f_d\tau = 0.0575$ and $E_s/E_p = -0.6$ dB. It can be seen that at this operating point, the PEP corresponding to the system with 5-symbol lag time is equal to $10^{-2}$ while the PEP corresponding to the system with 1-symbol lag time is as small as $10^{-8}$. In addition, it can be observed that the optimal data-to-pilot ratio decreases with the Doppler spread. This behavior is expected because when the Doppler spread is large, the CSI accuracy dominates the system performance, therefore, improving the CSI accuracy by allocating more energy into the pilot channel is beneficial.

4.5.5 Comparisons with Monte Carlo simulation

To evaluate the accuracy of our analysis, fig. 4.7 and 4.8 compare the truncated union bounds of the BEP, denoted by $P_e$, and the block error probability, denoted by $P_f$, calculated from the analytical PEP with $P_e$ and $P_f$ from Monte Carlo simulation. The code used in the simulation is the rate 1/3 convolutional code specified in the UMTS-WCDMA standard [3] with $d_{\text{min}} = 18$ and 256 states. The interleaver (with $I = 23$ in
Figure 4.7 Bit error probability and block error probability versus normalized Doppler frequency for PO-CLTD.

Figure 4.8 Bit error probability and block error probability versus normalized Doppler frequency for PA-CLTD.
this simulation) is also specified in the UMTS-WCDMA standard [3]. The number of information bits per block is 220 and each block is terminated with 8 zeros such that the encoder is set back to the all-zero state at the end of each block. The fading coefficients are generated by method of exact Doppler spread (MEDS) suggested in [45] with the autocorrelation of the Jakes model. The channel estimator is an 11-tap Wiener filter. In addition, it is assumed that data SNR = 2.22 dB and pilot SNR = 0.97 dB.

For the truncated union bounds, bounds calculated from the minimum Hamming weight ($d_{\text{min}} = 18$) and bounds calculated from the four smallest Hamming weights ($d = 18, 20, 22, 24$) are illustrated. Comparing the simulation results and the analytical results, it can be seen that the analytical results match well with the simulation results.

In addition, the analytical performance of convolutionally codes without transmit diversity [33], with the Alamouti STC (with the LC-ML decoder) [32] and with PA-CLTD are compared. The bit error probability and the block error probability of these three systems are illustrated in fig. 4.9 and 4.10, respectively. In these figures, it is assumed that the lag time of PA-CLTD is equal to 5 symbols, which is more practical than the 1 symbol lag time used in fig. 4.7 and 4.8. It can be seen from the figures that PA-CLTD with 5-symbol lag time performs much better than SISO or the Alamouti STC when the Doppler spread is small. However, the performance of PA-CLTD degrades significantly with the Doppler spread and is eventually outperformed by other two systems. Although the Alamouti STC does not perform as well as PA-CLTD when the Doppler spread is small, it may be more desirable than PA-CLTD for a system that has an operating point in moderate to large Doppler spread.

### 4.6 Conclusions

In this chapter, the bit error probability of PA-CLTD and PO-CLTD without convolutional codes and the pairwise error probability of PA-CLTD and PO-CLTD with convolutional codes are derived. The analysis takes into account system imperfections, such as noisy channel estimates, lag time and finite-depth interleaving, allowing us to evaluate the effect of these undesirable factors on the system performance.

The results have shown that PA-CLTD performs slightly better than PO-CLTD, although PA-CLTD requires more feedback information than PO-CLTD. Therefore, PO-
Figure 4.9 Bit error probability versus normalized Doppler frequency for SISO, the Alamouti STC and PA-CLTD (5-symbol lag time)

Figure 4.10 Block error probability versus normalized Doppler frequency for SISO, the Alamouti STC and PA-CLTD (5-symbol lag time)
CLTD is more appealing in a practical system where the feedback information may require significant resource. It has also been observed that the performance of CLTD algorithms strongly depends on the CSI accuracy and the lag-time. A comparison between the SISO system, the Alamouti STC system and the PA-CLTD system has shown that PA-CLTD performs well when the Doppler spread is small and the pilot SNR is large. However, when the Doppler spread increases and the pilot SNR decreases, the performance of PA-CLTD may degrade rapidly such that it is outperformed by the SISO or the Alamouti STC systems. Finally, analytical results are compared with the results from Monte Carlo simulation and the analytical results have been shown to match well with the simulation results.

4.7 Acknowledgement

5

Conclusions

In this dissertation, the theoretical error performance of finite-depth convolutional coded systems in time-varying fading channels for different types of transmit diversity algorithms has been derived. Three transmit diversity algorithms studied are

- no transmit diversity,
- the Alamouti space-time code (STC) and
- closed-loop transmit diversity (CLTD).

In addition, it is assumed throughout this dissertation that the channel estimation is performed by FIR filters from noisy pilot signals, therefore, the effects of the Doppler spread and the pilot SNR on the channel estimation accuracy is captured.

Chapter 2 focuses on the performance analysis of finite-depth convolutional codes without transmit diversity. It has been shown that there exists the estimation-diversity tradeoff as a function of the Doppler spread when both imperfect channel state information (CSI) and finite-depth interleaving are considered. The tradeoff is a result of the fact that increasing the Doppler spread improves the system performance by increasing the time diversity but degrades the performance by worsening the channel estimation accuracy. In addition, the results have shown that, in a fast fading channel, increasing the pilot SNR can improve the performance more effectively than increasing the interleaving depth. Comparing the truncated union bounds of the bit error probability and the block error probability calculated from the analytical pairwise error probability with
the simulation results, the simulation results have been shown to match well with the analytical results. Therefore, the analysis can be used to accurately predict the simulation performance.

Chapter 3 focuses on the performance analysis of the Alamouti STC with and without finite-depth convolutional codes. It has been shown that, although the linear-combining space-time decoder (LC-STD) performs worse than the maximal-likelihood space-time decoder (ML-STD), the performance difference when the CSI accuracy is poor is not significant. Therefore, in a system where CSI cannot be measured with high accuracy, it is more efficient to use the LC-STD than the ML-STD. When finite-depth convolutional codes are used in conjunction with the Alamouti STC, the system performance demonstrates the similar estimation-diversity tradeoff seen in the case without the Alamouti STC. In addition, the performance of the convolutionally coded system with the Alamouti STC has been compared to the performance of the convolutionally coded system with no transmit diversity. The results have shown that, when there is no correlation between the two transmit antennas, the coded system with no transmit diversity may outperform the coded system with the Alamouti STC when the Doppler spread is large and the pilot SNR is small. However, the correlation between transmit antennas can severely degrade the performance of the Alamouti STC. When the pilot SNR is small (poor CSI accuracy) and the antenna correlation is large, the Alamouti STC can be outperformed by the no transmit diversity even at small Doppler spread. In addition, a comparison between the analytical results and the simulation results has shown that the analysis can predict the simulation results accurately.

Chapter 4 focuses on the performance analysis of the closed-loop transmit diversity with noisy CSI and the feedback delay. Two types of CLTD algorithms analyzed in this chapter are the phase-amplitude CLTD (PA-CLTD) and the phase-only CLTD (PO-CLTD). It has been shown that the performance difference between PA-CLTD and PO-CLTD is generally small. Therefore, since PO-CLTD requires much less feedback information than PA-CLTD and it is much easier to implement, PO-CLTD will be the preferred CLTD algorithm in practical systems. The results have also shown that the performance of both PA-CLTD and PO-CLTD strongly degrade with the feedback delay and the degradation happens faster when the channel estimation filters have smaller number of taps. When finite-depth convolutional codes are used with CLTD algorithms, the system performance also displays the similar estimation-diversity tradeoff seen in
the convolutionally coded systems with and without the Alamouti STC. In addition, the truncated union bounds of the bit error probability and the block error probability calculated from the pairwise error probability are compared with the bit error probability and the block error probability from Monte Carlo simulation. The comparison has shown that the analytical and the simulation results match well. Finally, a comparison between convolutionally coded performance of three transmit diversity systems is presented. The results have shown that CLTD provides the best performance when the Doppler spread is small. As the Doppler spread increases, the performance of CLTD degrades significantly to the point that the Alamouti STC outperforms the CLTD system. As the Doppler spread increases, the Alamouti STC is eventually outperformed by the system without transmit diversity.
Finding the optimal parameter for Chernoff bound

The optimal $\rho$ can be found by taking a derivative of $\Phi_z(-\rho)$ with respect to $\rho$, the derivative can be expressed as

$$\frac{d}{d\rho} \Phi_z(-\rho) = -\Phi_z(-\rho) \left( \frac{dW(\rho)}{d\rho} + \sum_{l=1}^{L} \sum_{i=1}^{2d} \frac{\lambda_i^l}{1 + \rho \lambda_i^l} \right) \quad (A.1)$$

where $W(\rho) = \frac{1}{2} \sum_{l=1}^{L} \bar{x}_l^H \Sigma_l^{-1} [I_{2d} - (I_{2d} + 2\rho \Sigma_l Q)^{-1}] \bar{x}_l = \rho \sum_{i=1}^{2d} \bar{x}_l^H Q (I_{2d} + 2\rho \Sigma_l Q)^{-1} \bar{x}_l$ and $\lambda_i^l$ is the $i^{th}$ eigenvalue of $2\Sigma_l Q$. It can be shown, after some math, that

$$\frac{dW(\rho)}{d\rho} = \sum_{l=1}^{L} \bar{x}_l^H Q (I_{2d} + 2\rho \Sigma_l Q)^{-2} \bar{x}_l.$$

Substituting (A.3) into (A.1) and conditioning that $\Phi_z(-\rho_{opt}) \neq 0$, the optimal $\rho$ can then be expressed as

$$\rho_{opt} = \arg_{\rho} \left\{ \sum_{l=1}^{L} \bar{x}_l^H Q (I_{2d} + 2\rho \Sigma_l Q)^{-2} \bar{x}_l + \sum_{l=1}^{L} \sum_{i=1}^{2d} \frac{\lambda_i^l}{1 + \rho \lambda_i^l} = 0 \right\}. \quad (A.4)$$
This section proves that the residue of $\Phi_z(s)/s$ at a pole $p$ is equal to the residues of $\Phi_z(as)/s$ at a pole $p/a$, where $a$ is any non-zero scaling factor.

From (2.10), the characteristic function $\Phi_z(s)/s$ can be simplified as follows,

$$
\Phi_z(s) = \exp \left\{ \bar{x}_l^H s Q (I_{2d} - 2s \Sigma_l Q)^{-1} \bar{x}_l \right\} = \frac{1}{|I_{2d} - 2s \Sigma_l Q|} \left( \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \right), \quad (B.1)
$$

where $\beta = \bar{x}_l^H s Q (I_{2d} - 2s \Sigma_l Q)^{-1} \bar{x}_l = |I_{2d} - 2s \Sigma_l Q|^{-1} \bar{x}_l^H s Q M \bar{x}_l$ and $M$ is a matrix such that $M(I_{2d} - 2s \Sigma_l Q) = (I_{2d} - 2s \Sigma_l Q)M = |I_{2d} - 2s \Sigma_l Q| I_{2d}$. Thus, the poles of $\Phi_z(s)/s$ are zero and $1/\lambda_i$ for $i = 1, \ldots, 2d$, where $\lambda_i$ are the eigenvalues of $2 \Sigma_l Q$.

Therefore, the partial fraction of $\Phi_z(s)/s$ can be written as

$$
\frac{\Phi_z(s)}{s} = \frac{k_0}{s} + \frac{k_1}{s - 1/\lambda_1} + \frac{k_2}{s - 1/\lambda_2} + \frac{k_2^2}{(s - 1/\lambda_2)^2} + \frac{k_3^2}{(s - 1/\lambda_2)^3} + \ldots \quad (B.2)
$$

For generality, no assumption is made on the orders of the poles. Substituting $s = a\bar{s}$ into (B.2), the equation becomes

$$
\frac{\Phi_z(a\bar{s})}{a\bar{s}} = \frac{k_0}{a\bar{s}} + \frac{k_1}{a\bar{s} - 1/\lambda_1} + \frac{k_2}{a\bar{s} - 1/\lambda_2} + \frac{k_2^2}{(a\bar{s} - 1/\lambda_2)^2} + \frac{k_3^2}{(a\bar{s} - 1/\lambda_2)^3} + \ldots \quad (B.3)
$$

Multiplying both sides with $a$, resulting in

$$
\frac{\Phi_z(a\bar{s})}{\bar{s}} = \frac{k_0}{\bar{s}} + \frac{k_1}{\bar{s} - 1/a\lambda_1} + \frac{k_2}{\bar{s} - 1/a\lambda_2} + \frac{k_2^2/a}{(\bar{s} - 1/a\lambda_2)^2} + \frac{k_3^2/a^2}{(\bar{s} - 1/a\lambda_2)^3} + \ldots \quad (B.4)
$$
It is obvious that, changing from $\Phi_{zi}(s)/s$ to $\Phi_{zi}(as)/s$, we change the poles from $1/\lambda_i$ to $1/a\lambda_i$. In addition, the residue of $\Phi_{zi}(s)/s$ at the pole $1/\lambda_i$ in (B.2) is equal to the coefficient $k_i^1$, this residue is also equal to the residue of $\Phi_{zi}(as)/s$ at the pole $1/a\lambda_i$ in (B.4). Therefore, it can be concluded that the residue of $\Phi_{zi}(as)/s$ at $1/a\lambda_i$ is equal to the residue of $\Phi_{zi}(s)/s$ at $1/\lambda_i$. And since the PEP is equal to the summation of the residues of the poles in the left half-plane, the PEP is invariant to the value of $a$ in $\Phi_{zi}(as)/s$. 
Derivation of (4.24)

The goal of this section is to derive the characteristic function of \( z = \Re[y_1y_2^*]/|y_2| \) where \( y = [y_1 \ y_2]^T \) is a complex Gaussian random vector with zero-mean and covariance matrix \( \Sigma \). Since the joint probability density function (PDF) in the spherical coordinates, where \( r_1 = |y_1|, r_2 = |y_2| \) and \( \Re[y_1y_2^*] = r_1r_2 \cos \theta \) for \( 0 \leq \theta < \pi \), can be written as [40]

\[
p(r_1, r_2, \theta) = \frac{r_1r_2}{\pi|\Sigma|} e^{-\frac{1}{2}(W(1,1)r_1^2 + W(2,2)r_2^2) - W(1,2)r_1r_2 \cos \theta}, \tag{C.1}
\]

where \( W \) is the matrix inverse of \( \Sigma \). To find the PDF of \( z = r_1 \cos \theta \), we change the variables from \( \{r_1, r_2, \theta\} \) to \( \{z = r_1 \cos \theta, h_1 = r_1 \sin \theta, h_2 = r_2\} \). The Jacobian \( J \) is equal to \((z^2 + h_2^2)^{-1/2}\) and the joint PDF of \( \{z, h_1, h_2\} \) is equal to

\[
p(z, h_1, h_2) = Jp(r_1, r_2, \theta) = \frac{h_2}{\pi|\Sigma|} e^{-\frac{1}{2}(W(1,1)(z^2 + h_1^2) + W(2,2)h_2^2) - W(1,2)zh_2}. \tag{C.2}
\]

The marginal PDF of \( z \) can be found from integrating (C.2) with respect to \( h_1 \) and \( h_2 \), the integral results in

\[
p(z) = \int_{0}^{\infty} \int_{0}^{\infty} p(z, h_1, h_2) \, dh_2 \, dh_1 \tag{C.3}
\]

\[
= e^{-\frac{1}{2}W(1,1)z^2} \int_{0}^{\infty} e^{-\frac{1}{2}W(1,1)h_1^2} \, dh_1 \int_{0}^{\infty} \frac{h_2}{|\Sigma|} e^{-\frac{1}{2}W(2,2)h_2^2 - W(1,2)zh_2} \, dh_2 \tag{C.4}
\]

\[
= e^{-\frac{1}{2}W(1,1)z^2} \frac{\sqrt{2\pi W(1,1)}}{|\Sigma|} \int_{0}^{\infty} \frac{h_2}{|\Sigma|} e^{-\frac{1}{2}W(2,2)h_2^2 - W(1,2)zh_2} \, dh_2. \tag{C.5}
\]

The characteristic function of \( z \) can then be found to be

\[
\Phi_z(iv) = \int_{0}^{\infty} \left( \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2}W(1,1)z^2 + ivz - W(1,2)zh_2 \right\} \, dz \right) \frac{h_2}{|\Sigma|} e^{-\frac{1}{2}W(2,2)h_2^2} \, dh_2. \tag{C.6}
\]
Since the characteristic function of a Gaussian random variable \( x \) with mean \( \bar{x} \) and variance \( \sigma^2_x \) is equal to
\[
\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2_x}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2_x}} \, dx = \exp \left\{ i\bar{x}v - \frac{v^2\sigma^2_x}{2} \right\}.
\]
Eq. (C.6) can be simplified further to
\[
\Phi_z(iv) = \exp \left\{ -\frac{v^2}{2W(1,1)} \right\} \int_{0}^{\infty} \frac{h_2}{W(1,1)|\Sigma|} \exp \left\{ -\frac{1}{2} \frac{h_2^2}{W(1,1)|\Sigma|} - i\frac{vW(1,2)h_2}{W(1,1)} \right\} \, dh_2.
\]

Eq. (C.7) can be simplified further by using our finding that
\[
\int_{0}^{\infty} xe^{-\frac{1}{2}ax^2} e^{-ibx} \, dx = \frac{1}{a} \left( 1 - b\sqrt{\frac{\pi}{2a}} e^{-b^2/2a} \text{erfi} \left( \frac{b}{\sqrt{2a}} \right) \right) - i\frac{b}{a} \sqrt{\frac{\pi}{2a}} e^{-b^2/2a},
\]
where \( \text{erfi}(x) \) is an imaginary error function defined as [1]
\[
\text{erfi}(y) = \frac{2}{\sqrt{\pi}} e^{y^2} \int_{0}^{\infty} e^{-t^2} \sin(2yt) \, dt \text{ where } y \in \mathbb{R}.
\]

Applying (C.8) to the integral in (C.7) with \( x = \frac{h_2}{a}, a = W(1,1)|\Sigma|, b = vW(1,2)|\Sigma| \) and \( dh_2 = a \, dx \), the characteristic function can be obtained
\[
\Phi_z(iv) = (\Psi_i(v) + i\Psi_q(v)) \exp \left\{ -\frac{v^2|\Sigma|}{2\Sigma(2,2)} \right\},
\]
where \( \Psi_i(v) = 1 + v\rho\sqrt{\pi} e^{-v^2\rho^2} \text{erfi} (-v\rho), \Psi_q(v) = v\rho\sqrt{\pi} e^{-v^2\rho^2} \) and \( \rho = \frac{\Sigma(1,2)}{\sqrt{2\Sigma(2,2)}} \).
D

Derivation of (4.31)

The goal of this section is to derive the probability that \( z \) is less than zero assuming that the PDF of \( z \) is defined in (4.30). Recall from (4.30) that

\[
p(z) = p_e(z) + p_o(z),
\]

where \( p_e(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g_i(v)e^{-ivz}dv \) and \( p_o(z) = \frac{i}{2\pi} \int_{-\infty}^{\infty} g_q(v)e^{-ivz}dv \) denote the even and the odd parts of \( p(z) \), respectively. In addition, \( g_i(v) \) and \( g_q(v) \) are as defined in (4.28) and (4.29). The contributions to \( \Pr\{z < 0\} \) from \( \int_{-\infty}^{0} p_e(z)dz \) and \( \int_{-\infty}^{0} p_o(z)dz \) will be considered separately as follows.

Since \( p_e(z) \) is an even function, \( \int_{-\infty}^{0} p_e(z)dz = \frac{1}{2} \int_{-\infty}^{\infty} p_e(z)dz \). And since \( g_i(v) \) is the inverse Fourier transform of \( p_e(z) \), the contribution from the even function can be found by using the inverse Fourier transform identity \( g_i(v) = \int_{-\infty}^{\infty} p_e(z)e^{-ivz}dz \), i.e.,

\[
\int_{-\infty}^{0} p_e(z)dv = \frac{g_i(0)}{2} = \frac{1}{2}. \quad (D.2)
\]

Consider now the contribution from the odd function \( p_o(z) \), the contribution can be written as

\[
\int_{-\infty}^{0} p_o(z)dz = \frac{i}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{0} \Psi_i(v)\Psi_q(v) \exp \left\{ -v^2 \left( \frac{\sigma_s^2}{E_s} + \frac{|\Sigma|}{\Sigma(2, 2)} \right) - ivz \right\} dz \ dv \quad (D.3)
\]

\[
= \frac{i}{\pi} \int_{-\infty}^{\infty} \Psi_i(v)\Psi_q(v) \exp \left\{ -v^2 \left( \frac{\sigma_s^2}{E_s} + \frac{|\Sigma|}{\Sigma(2, 2)} \right) \right\} \left( \pi\delta(v) - \frac{1}{iv} \right) dv, \quad (D.4)
\]
where \( \Sigma \) is equal to \( \Sigma_3 \) specified in section IV-A. Note that (D.4) is obtained from utilizing the Fourier transform of the Heaviside step function \( H(x) \), i.e., \( \int_{-\infty}^{\infty} e^{-i2\pi f} H(x) dx = \frac{1}{2} (\delta(f) - \frac{1}{\pi f}) \). Using the sifting property of the Dirac delta function and the knowledge that \( \Psi_q(0) = 0 \), it can be shown that

\[
i \int_{-\infty}^{\infty} \frac{\pi \delta(v)}{\nu} \Psi_i(v) \Psi_q(v) \exp \left\{ -v^2 \left( \frac{\sigma_s^2}{E_s} + \frac{|\Sigma|}{\Sigma(2,2)} \right) \right\} dv = i \Psi_i(0) \Psi_q(0) = 0. \tag{D.5}
\]

Therefore, (D.4) can be simplified to

\[
\int_{-\infty}^{0} p_o(z) dz = -\frac{1}{\pi} \int_{-\infty}^{\infty} \Psi_i(v) \Psi_q(v) \exp \left\{ -v^2 \left( \frac{\sigma_s^2}{E_s} + \frac{|\Sigma|}{\Sigma(2,2)} \right) \right\} dv. \tag{D.6}
\]

Substitute the expressions for \( \Psi_i(v) \) and \( \Psi_q(v) \) into (D.6), (D.6) can be simplified further to

\[
\int_{-\infty}^{0} p_o(z) dz = -\frac{\rho}{\sqrt{\rho^2 + \frac{\sigma_s^2}{E_s} + \frac{|\Sigma|}{\Sigma(2,2)}}} - \int_{-\infty}^{\infty} v \rho^2 e^{-v^2 \left( \frac{\sigma_s^2}{E_s} + \frac{|\Sigma|}{\Sigma(2,2)} \right)} \text{erfi}(-v \rho) \ dv, \tag{D.7}
\]

where \( \rho = \frac{\Sigma(1,2)}{\sqrt{2 \Sigma(2,2)}} \) as defined in section 4.4.1. Using integral by paths, the integral in (D.7) can be written as

\[
- \int_{-\infty}^{\infty} v \rho^2 e^{-v^2 \left( \frac{\sigma_s^2}{E_s} + \frac{|\Sigma|}{\Sigma(2,2)} \right)} \text{erfi}(-v \rho) \ dv = \int_{-\infty}^{\infty} \frac{\rho^2 \text{erfi}(-v \rho)}{2 \left( \frac{\sigma_s^2}{E_s} + \Sigma(1,1) \right)} \left| e^{-v^2 \left( \frac{\sigma_s^2}{E_s} + \Sigma(1,1) \right)} \right| \left| e^{-\rho \rho^2 e^{-v^2 \left( \frac{\sigma_s^2}{E_s} + \Sigma(1,1) \right)}} \right| \left| \text{erfi}(-v \rho) \right|
\]

\[
= \rho^2 \text{erfi}(-v \rho) \frac{e^{-v^2 \left( \frac{\sigma_s^2}{E_s} + \Sigma(1,1) \right)}}{2 \left( \frac{\sigma_s^2}{E_s} + \Sigma(1,1) \right)} \int_{-\infty}^{\infty} \frac{\rho^2 e^{-v^2 \left( \frac{\sigma_s^2}{E_s} + \Sigma(1,1) \right)}}{2 \left( \frac{\sigma_s^2}{E_s} + \Sigma(1,1) \right)} \text{erfi}(-v \rho) \tag{D.8}
\]

From [1], it is known that \( \text{erfi}(x) = \left(2e^{x^2}/\sqrt{\pi} \right) \int_{0}^{\infty} e^{-t^2} \sin(2xt)dt \), therefore, it can be concluded that the first term in (D.9) is equal to zero. In addition, since \( \frac{\text{erfi}(x)}{dx} = \frac{2e^{x^2}}{\sqrt{\pi}} \) [1], the integral term in (D.9) can be simplified to

\[
- \int_{-\infty}^{\infty} \frac{\rho^2 e^{-v^2 \left( \frac{\sigma_s^2}{E_s} + \Sigma(1,1) \right)}}{2 \left( \frac{\sigma_s^2}{E_s} + \Sigma(1,1) \right)} \text{erfi}(-v \rho) = \int_{-\infty}^{\infty} \frac{\rho^3 e^{-v^2 \left( \frac{\sigma_s^2}{E_s} + \Sigma(1,1) - \rho^2 \right)}}{\sqrt{\pi} \left( \frac{\sigma_s^2}{E_s} + \Sigma(1,1) \right)} dv \tag{D.10}
\]

\[
= \frac{\rho^3}{\sqrt{\rho^2 + \frac{\sigma_s^2}{E_s} + \frac{|\Sigma|}{\Sigma(2,2)}}}. \tag{D.11}
\]
Finally, the probability of error then becomes

\[
\Pr\{z < 0\} = \int_{-\infty}^{0} (p_e(z) + p_c(z))dz = \frac{1}{2} - \rho \left( \frac{\sigma_s^2}{E_s} + \Sigma(1, 1) \right)^{-1} \sqrt{\rho^2 + \frac{\sigma_s^2}{E_s} + \frac{\|\Sigma\|}{\Sigma(2, 2)}}.
\]

(D.12)
Derivation of (4.36)

The goal of this section is to derive the PEP of PO-CLTD with convolutional codes as shown in (4.36). Let \( \tilde{\alpha} = [\alpha_I \ldots \alpha_dI \ \hat{\alpha}_I \ldots \hat{\alpha}_dI]^T \) and \( \tilde{\beta} = [\beta_I \ldots \beta_dI \ \hat{\beta}_I \ldots \hat{\beta}_dI]^T \) be independent identically distributed zero-mean complex Gaussian random vectors with variance \( \Sigma \). Also, let \( \tilde{n}_s \) be an independent zero-mean real-valued Gaussian random variable with variance \( \tilde{\sigma}^2_s = 2\sigma^2_s \sum_{k=1}^{d} (|\hat{\alpha}_{kI}| + |\hat{\beta}_{kI}|)^2 \). The joint PDF of \( \tilde{\alpha}, \tilde{\beta} \) and \( \tilde{n}_s \) can be found by Bayes’ rule to be

\[
p(\tilde{\alpha}, \tilde{\beta}, \tilde{n}_s) = p(\tilde{\alpha})p(\tilde{\beta})p(\tilde{n}_s|\tilde{\alpha}, \tilde{\beta}),
\]

where

\[
p(\tilde{\alpha}) = \frac{\exp \left\{ -\frac{1}{2} \tilde{\alpha}^H \Sigma^{-1} \tilde{\alpha} \right\} }{(2\pi)^{2d} |\Sigma|},
\]

\[
p(\tilde{\beta}) = \frac{\exp \left\{ -\frac{1}{2} \tilde{\beta}^H \Sigma^{-1} \tilde{\beta} \right\} }{(2\pi)^{2d} |\Sigma|},
\]

\[
p(\tilde{n}_s|\tilde{\alpha}, \tilde{\beta}) = \frac{\exp \left\{ -\frac{\tilde{n}_s^2}{2\tilde{\sigma}^2_s} \right\} }{\sqrt{2\pi\tilde{\sigma}^2_s}}.
\]

From (4.35), the pairwise error probability we are interested in equal to \( \Pr\{ \tilde{z}_1 + \tilde{z}_2 + \tilde{n}_s < 0 \} \), where \( \tilde{z}_1 = \sum_{k=1}^{d} (1 + f_k) \Re\{\alpha_{kI}\hat{\alpha}_{kI}^*\} \), \( \tilde{z}_2 = \sum_{k=1}^{d} (1 + f_k^{-1}) \Re\{\beta_{kI}\hat{\beta}_{kI}^*\} \) and \( f_k = |\hat{\beta}_{kI}|/|\hat{\alpha}_{kI}|. \)

E

Derivation of (4.36)

The goal of this section is to derive the PEP of PO-CLTD with convolutional codes as shown in (4.36). Let \( \tilde{\alpha} = [\alpha_I \ldots \alpha_dI \ \hat{\alpha}_I \ldots \hat{\alpha}_dI]^T \) and \( \tilde{\beta} = [\beta_I \ldots \beta_dI \ \hat{\beta}_I \ldots \hat{\beta}_dI]^T \) be independent identically distributed zero-mean complex Gaussian random vectors with variance \( \Sigma \). Also, let \( \tilde{n}_s \) be an independent zero-mean real-valued Gaussian random variable with variance \( \tilde{\sigma}^2_s = 2\sigma^2_s \sum_{k=1}^{d} (|\hat{\alpha}_{kI}| + |\hat{\beta}_{kI}|)^2 \). The joint PDF of \( \tilde{\alpha}, \tilde{\beta} \) and \( \tilde{n}_s \) can be found by Bayes’ rule to be

\[
p(\tilde{\alpha}, \tilde{\beta}, \tilde{n}_s) = p(\tilde{\alpha})p(\tilde{\beta})p(\tilde{n}_s|\tilde{\alpha}, \tilde{\beta}),
\]

where

\[
p(\tilde{\alpha}) = \frac{\exp \left\{ -\frac{1}{2} \tilde{\alpha}^H \Sigma^{-1} \tilde{\alpha} \right\} }{(2\pi)^{2d} |\Sigma|},
\]

\[
p(\tilde{\beta}) = \frac{\exp \left\{ -\frac{1}{2} \tilde{\beta}^H \Sigma^{-1} \tilde{\beta} \right\} }{(2\pi)^{2d} |\Sigma|},
\]

\[
p(\tilde{n}_s|\tilde{\alpha}, \tilde{\beta}) = \frac{\exp \left\{ -\frac{\tilde{n}_s^2}{2\tilde{\sigma}^2_s} \right\} }{\sqrt{2\pi\tilde{\sigma}^2_s}}.
\]

From (4.35), the pairwise error probability we are interested in equal to \( \Pr\{ \tilde{z}_1 + \tilde{z}_2 + \tilde{n}_s < 0 \} \), where \( \tilde{z}_1 = \sum_{k=1}^{d} (1 + f_k) \Re\{\alpha_{kI}\hat{\alpha}_{kI}^*\} \), \( \tilde{z}_2 = \sum_{k=1}^{d} (1 + f_k^{-1}) \Re\{\beta_{kI}\hat{\beta}_{kI}^*\} \) and \( f_k = |\hat{\beta}_{kI}|/|\hat{\alpha}_{kI}|. \)
Using Bayes’ rule, the characteristic function of \( z = \bar{z}_1 + \bar{z}_2 + \bar{n}_s \) can be written as

\[
\Phi_z(v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(\hat{\alpha}) p(\hat{\beta}) e^{iv(\bar{z}_1 + \bar{z}_2)} \int_{-\infty}^{\infty} p(\bar{n}_s | \hat{\alpha}, \hat{\beta}) e^{iv\bar{n}_s} \, d\bar{n}_s \, d\hat{\alpha} \, d\hat{\beta}
\]

(E.2)

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}v^2t^2} p(\hat{\alpha}) p(\hat{\beta}) \int_{-\infty}^{\infty} p(\alpha | \hat{\alpha}) e^{ivz_1} \, d\alpha \int_{-\infty}^{\infty} p(\beta | \hat{\beta}) e^{ivz_2} \, d\beta \, d\hat{\alpha} \, d\hat{\beta},
\]

(E.3)

where the conditional PDF are \( p(\alpha | \hat{\alpha}) = (2\pi)^{-d} |\hat{\Sigma}|^{-1} \exp \left\{ -\frac{1}{2}(\alpha - \hat{\alpha})^H \hat{\Sigma}^{-1}(\alpha - \hat{\alpha}) \right\} \)

and \( p(\beta | \hat{\beta}) = (2\pi)^{-d} |\hat{\Sigma}|^{-1} \exp \left\{ -\frac{1}{2}(\beta - \hat{\beta})^H \hat{\Sigma}^{-1}(\beta - \hat{\beta}) \right\} \) [8]. In addition, \( \hat{\alpha} = \Sigma_{12} \Sigma_{22}^{-1} \bar{\alpha}, \hat{\beta} = \Sigma_{12} \Sigma_{22}^{-1} \bar{\beta}, \hat{\Sigma} = \Sigma_1 - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_2 \) and [8]

\[
\Sigma = \begin{bmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{12}^H & \Sigma_{22}
\end{bmatrix}.
\]

(E.4)

The expression \( \int_{-\infty}^{\infty} p(\alpha, \hat{\alpha}) e^{ivz_1} \, d\alpha \) can be simplified further by using the properties of Gaussian random variables that

\[
p(\alpha | \hat{\alpha}) = \frac{\exp \left\{ -\frac{1}{2}(\alpha - \hat{\alpha})^H \hat{\Sigma}^{-1}(\alpha - \hat{\alpha}) \right\}}{(2\pi)^d |\hat{\Sigma}|}.
\]

(E.5)

where \( \hat{\alpha} = \Sigma_{12} \Sigma_{22}^{-1} \bar{\alpha}, \hat{\Sigma} = \Sigma_1 - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_2 \) [8] and

\[
\Sigma = \begin{bmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{12}^H & \Sigma_{22}
\end{bmatrix}.
\]

(E.6)

Now, consider the integral containing \( p(\alpha, \hat{\alpha}) \). Rewriting the equation into the real and the imaginary parts, \( p(\alpha, \hat{\alpha}) e^{ivz_1} \) becomes

\[
p(\alpha | \hat{\alpha}) e^{ivz_1} = \frac{\exp \left\{ \frac{1}{2} \begin{bmatrix} \Re[\alpha - \hat{\alpha}] \\ \Im[\alpha - \hat{\alpha}] \end{bmatrix}^H \hat{\Sigma}^{-1} \begin{bmatrix} \Re[\alpha - \hat{\alpha}] \\ \Im[\alpha - \hat{\alpha}] \end{bmatrix} + jv \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}^H \begin{bmatrix} \Re[\alpha] \\ \Im[\alpha] \end{bmatrix} \right\}}{(2\pi)^d |\hat{\Sigma}|},
\]

(E.7)

where \( m_1(k) = (1 + f_k) \Re[\hat{\alpha}_{kl}] \) and \( m_2(k) = (1 + f_k) \Im[\hat{\alpha}_{kl}] \). Since it is known that \( E[e^{jHx}] = e^{jH\mu - \frac{1}{2}H\Sigma x} \), where \( x \) is a Gaussian random vector with mean \( \mu \) and variance \( \Sigma \), the integral \( \int_{-\infty}^{\infty} p(\alpha, \hat{\alpha}) e^{ivz_1} \, d\alpha \) becomes

\[
\int_{-\infty}^{\infty} p(\alpha, \hat{\alpha}) e^{ivz_1} \, d\alpha = p(\hat{\alpha}) \exp \left\{ jv(m_1^T \Re[\hat{\alpha}] + m_2^T \Im[\hat{\alpha}]) - \frac{v^2}{2} \left( m_1^T \hat{\Sigma} m_1 + m_2^T \hat{\Sigma} m_2 \right) \right\}.
\]

(E.8)
Using the same method, the compact form of $\int_{-\infty}^{\infty} p(\beta, \hat{\beta}) e^{jvz} \, d\beta$ can also be found to be

$$\int_{-\infty}^{\infty} p(\beta, \hat{\beta}) e^{jvz} \, d\beta = p(\hat{\beta}) \exp \left\{ jv(m_3^T \Re[\hat{\beta}] + m_4^T \Im[\hat{\beta}]) - \frac{v^2}{2} \left( m_3^T \hat{\Sigma} m_3 + m_4^T \hat{\Sigma} m_4 \right) \right\},$$

(E.9)

where $m_3(k) = (1 + 1/f_k) \Re[\hat{\beta}_k]$ and $m_4(k) = (1 + 1/f_k) \Im[\hat{\beta}_k]$. Substituting (E.8) and (E.9) into (E.3), the characteristic function becomes

$$\Phi_z(v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(\hat{\alpha}) p(\hat{\beta}) \exp \left\{ -\frac{v^2}{2} (\chi + 2\sigma_s^2) + jv(\varphi_1 + \varphi_2) \right\} \, d\hat{\alpha} \, d\hat{\beta},$$

(E.10)

where $\chi = (m_1 + jm_2) H \hat{\Sigma} (m_1 + jm_2), \varphi_1 = m_1^T \Sigma_{12} \Sigma_{22}^{-1} \Re[\hat{\alpha}] + m_2^T \Sigma_{12} \Sigma_{22}^{-1} \Re[\hat{\beta}] + m_1^T \Sigma_{12} \Sigma_{22}^{-1} \Im[\hat{\beta}]$. Using the characteristic function in (E.10), the PDF of $z$ can be found to be

$$p(z) = \frac{1}{\sqrt{2\pi(\chi_1 + \chi_2 + \sigma_s^2)}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(\hat{\alpha}) p(\hat{\beta}) e^{-(z - \varphi_1 - \varphi_2)^2/2(\chi_1 + \chi_2 + \sigma_s^2)} \, d\hat{\alpha} \, d\hat{\beta}.$$

(E.11)

As a result, the PEP, which is the probability that $z < 0$, is equal to

$$\Pr\{z < 0\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(\hat{\alpha}) p(\hat{\beta}) Q \left( \frac{\varphi_1 + \varphi_2}{\sqrt{\chi_1 + \chi_2 + \sigma_s^2}} \right) \, d\hat{\alpha} \, d\hat{\beta},$$

(E.12)

$$= E_{\hat{\alpha}, \hat{\beta}} \left[ Q \left( \frac{\varphi_1 + \varphi_2}{\sqrt{\chi_1 + \chi_2 + \sigma_s^2}} \right) \right].$$

(E.13)

Since the PEP is in the form of an expectation, it can easily be calculated by the Monte Carlo method. In other words, the PEP can be found by averaging over a large number of $Q \left( \frac{\varphi_1 + \varphi_2}{\sqrt{\chi_1 + \chi_2 + \sigma_s^2}} \right)$, where $\hat{\alpha}$ and $\hat{\beta}$ are generated according to the Gaussian PDF with zero mean and covariance matrix $\Sigma_{22}$. It should also be noted that $\varphi_1, \varphi_2, \chi_1, \chi_2$ and $\sigma_s^2$ can be rewritten in compact forms: $\chi_1 = y_1^H \tilde{\Sigma} y_1, \chi_2 = y_2^H \tilde{\Sigma} y_2, \sigma_s^2 = 2\sigma_s^2 [y_1]^2 = 2\sigma_s^2 [y_2]^2, \varphi_1 = \Re[y_1^H \Sigma_{12} \Sigma_{22}^{-1} \hat{\alpha}], \varphi_2 = \Re[y_2^H \Sigma_{12} \Sigma_{22}^{-1} \hat{\beta}],$ where $y_1(k) = (|\hat{\alpha}_k| + |\hat{\beta}_k|) \exp\{i\angle \hat{\alpha}_k\}$ and $y_2(k) = (|\hat{\alpha}_k| + |\hat{\beta}_k|) \exp\{i\angle \hat{\beta}_k\}$.
Bibliography


[2] 3GPP. Technical specification group radio access network, physical channels and mapping of transport channels onto physical channels (FDD), 25.211.


