

MASTER COPY: PLEASE KEEP THIS "MEMORANDUM OF TRANSMITTAL" BLANK FOR REPRODUCTION PURPOSES. WHEN REPORTS ARE GENERATED UNDER THE ARO SPONSORSHIP, FORWARD A COMPLETED COPY OF THIS FORM WITH EACH REPORT SHIPMENT TO THE ARO. THIS WILL ASSURE PROPER IDENTIFICATION. NOT TO BE USED FOR INTERIM PROGRESS REPORTS; SEE PAGE 2 FOR INTERIM PROGRESS REPORT INSTRUCTIONS.

MEMORANDUM OF TRANSMITTAL

U.S. Army Research Office
ATTN: AMSRL-RO-BI (TR)
P.O. Box 12211
Research Triangle Park, NC 27709-2211

- | | |
|---|--|
| <input type="checkbox"/> Reprint (Orig + 2 copies) | <input type="checkbox"/> Technical Report (Orig + 2 copies) |
| <input checked="" type="checkbox"/> Manuscript (1 copy) | <input type="checkbox"/> Final Progress Report (Orig + 2 copies) |
| | <input type="checkbox"/> Related Materials, Abstracts, Theses (1 copy) |

CONTRACT/GRANT NUMBER: **W911NF0410224 (46637-CI-MUR)**

REPORT TITLE:

On the Performance of Concatenated Convolutional Code and Alamouti Space-Time Code with Noisy Channel Estimates and Finite-Depth Interleaving

Jittra Jootar, James R. Zeidler and John G. Proakis

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):
IEEE Transactions on Communications

Sincerely,

Dr. James Zeidler
Department of Electrical and Computer Engineering
University of California, San Diego

REPORT DOCUMENTATION PAGEForm Approved
OMB NO. 0704-0188

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188,) Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE 04/01/2006	3. REPORT TYPE AND DATES COVERED Manuscripts: 01 June 2004 – 01 April 2006	
4. TITLE AND SUBTITLE On the Performance of Concatenated Convolutional Code and Alamouti Space-Time Code with Noisy Channel Estimates and Finite-Depth Interleaving			5. FUNDING NUMBERS W911NF0410224	
6. AUTHOR(S) Jittra Jootar, James R. Zeidler, and John G. Proakis				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of California – San Diego Department of Electrical and Computer Engineering 9500 Gilman Dr., La Jolla, CA 92093			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.				
12 a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.			12 b. DISTRIBUTION CODE N/A	
13. ABSTRACT (Maximum 200 words) In this paper, performance of the Alamouti space-time code (STC) and performance of the concatenation between the convolutional code and the Alamouti STC are derived. In order to describe realistic performance issues, we assume that the channel estimates are calculated from linear filters using noisy pilot symbols. We also assume non-quasi-static channels, spatially correlated transmit antennas and finite-depth interleaving. Two types of receivers are investigated for the Alamouti STC, 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. Lastly, analytical results are compared with simulation results to illustrate the accuracy of the analysis.				
14. SUBJECT TERMS N/A			15. NUMBER OF PAGES 25	
			16. PRICE CODE N/A	
17. SECURITY CLASSIFICATION OR REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION ON THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	

Enclosure 1

On the Performance of Concatenated Convolutional Code and Alamouti Space-Time Code with Noisy Channel Estimates and Finite-Depth Interleaving

Jittra Jootar, *Student Member, IEEE*, James R. Zeidler, *Fellow, IEEE*,
and John G. Proakis, *Life Fellow, IEEE*

Abstract

In this paper, performance of the Alamouti space-time code (STC) [1] and performance of the concatenation between the convolutional code and the Alamouti STC are derived. In order to describe realistic performance issues, we assume that the channel estimates are calculated from linear filters using noisy pilot symbols. We also assume non-quasi-static channels, spatially correlated transmit antennas and finite-depth interleaving. Two types of receivers are investigated for the Alamouti STC, 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. Lastly, analytical results are compared with simulation results to illustrate the accuracy of the analysis.

I. 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 [2]. In this paper, we analyze the conventional Alamouti space-time coded system [1] 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 [3], [4].

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 [5], [6]. 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 [5], [6].

The Alamouti space-time code was proposed in 1998 [1] and has generated significant interest due to the low complexity linear-combining space-time decoder (LC-STD) suggested by Alamouti in [1]. 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) [1], [7]. 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 [8], [9]. 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 [7], [8].

Other factors that can significantly degrade the performance of the Alamouti space-time code are noisy CSI [9]–[17], which is caused by noise in the pilot symbols or time-varying fading channels, and spatial correlation of the transmit antennas [17]–[19]. 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]–[14], [17]. 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 [1] 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 [7], while analyses assuming noisy CSI in quasi-static fading channels were presented in [9]–[15], [17]. In addition, analyses on the performance of space-time codes over spatially correlated Rayleigh fading channels were presented in [17]–[19]. 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 [2], [20]. 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 [21]–[23]

The objective of this paper is to generalize the existing analyses and provide analytical performance of the conventional Alamouti space-time code [1] 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 [1] 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). We would like to note that a portion of the material presented in this paper was presented in [24].

The paper is organized as follows. Section II describes the mathematical tools used in this paper. In section III, the system model is introduced. Performance analyses of the systems with the LC-STD and the LC-ML are derived in section IV, while performance analyses of the systems with the ML-STD or the JML are derived in section V. In section VI, we describe how this analysis can be extended to evaluate performance in Rician and multi-path fading channels. Numerical results are presented and the comparison between the analytical and simulation results are discussed in section VII. Finally, the conclusion of our paper is presented in section VIII.

II. CHARACTERISTIC FUNCTION, RESIDUE THEOREM AND GAUSS-CHEBYSHEV APPROXIMATION

The following notation will be used for this analysis. 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 matrix \mathbf{X} is denoted by $\mathbf{X}(m, n)$ and the element in the m^{th} row (column) of a column (row)

vector \mathbf{x} is denoted by $\mathbf{x}(m)$. The superscripts $*$, T , H denote the complex conjugate, the matrix transpose and the matrix Hermitian operation, respectively. The determinant of a matrix \mathbf{X} is denoted by $|\mathbf{X}|$. The length m column vector of ones, the square identity matrix and the square zero matrix of order m are denoted by $\mathbf{1}_m$, \mathbf{I}_m and $\mathbf{0}_m$, respectively.

Analytical tools used in this paper 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 \mathbf{x} , i.e., $z = \mathbf{x}^H \mathbf{Q} \mathbf{x}$, then, with the knowledge of the covariance matrix of \mathbf{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.

A. Characteristic function of a Hermitian quadratic form of a complex Gaussian random vector

Let \mathbf{x} be a zero-mean complex Gaussian random vector of length m and \mathbf{Q} be a square matrix of order m , then the characteristic function of $z = \mathbf{x}^H \mathbf{Q} \mathbf{x}$ can be written as [25]

$$\Phi_z(s) = |\mathbf{I}_m - 2s\Sigma\mathbf{Q}|^{-1}, \quad (1)$$

where Σ denotes the covariance matrix of \mathbf{x} .

B. Residue theorem

The probability that z is less than zero can be found by Mellin's inversion formula and the residue theorem as follows

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

where ϵ lines 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 $\text{Res}[f(s) \text{ at } q_i]$ denotes the residue of $f(s)$ at q_i [26]. In addition, the residue can be calculated by

$$\text{Res}[f(s) \text{ at } q_i] = \lim_{s \rightarrow q_i} \frac{p^{(m-1)}(s)}{(m-1)!}, \quad (3)$$

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 $(m-1)^{\text{th}}$ derivative of $p(s)$. We would like to note 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 $\Pr\{z < 0\}$.

C. 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 [27] with a correction in [28]) as follows

$$\Pr\{z < 0\} \approx \frac{1}{2m} \sum_{k=1}^m (\epsilon \Re[\Phi_z(\epsilon + j\epsilon\tau_k)] + \tau_k \Im[\Phi_z(\epsilon + j\epsilon\tau_k)]), \quad (4)$$

where $\Re[x]$ and $\Im[x]$ denote the real and the imaginary parts of x , respectively, $\tau_k = \tan((2k - 1)\pi/4m)$, ϵ is as defined in section II-B and, in general, m between 16 and 32 is sufficient [27].

III. 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.

A. 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 uncoded symbols for the conventional Alamouti system, are then encoded by the Alamouti space-time encoder [1], [3]. Since we assume BPSK modulation, the complex conjugates used in the Alamouti code can be ignored.

In addition, we assume 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, we also assume 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 τ_c denotes the chip period. The symbol-rate baseband representation of this system is shown in fig. 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. The signals corresponding to the symbol time index 1 are shown in parentheses and the signals corresponding to the symbol time index 2 are shown in square brackets.

B. 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 α_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 β_k . Because of the Rayleigh fading assumption, α_k and β_k are circularly symmetric zero-mean complex Gaussian random variables. We also assume that α_k and β_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, we assume that the correlation between α_k and β_k can be written as $\frac{1}{2}E[\alpha_k\beta_{k-m}^*] = \rho\sigma_c^2 R(m\tau)$, where ρ is a real value indicating the spatial correlation between the two transmit antennas. We would also like to note that, although we assume time-varying channels, we restrict ourselves to the case when the channels change slowly enough that the fading coefficients appears to be constant over one symbol period τ .

After despreading, we assume 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 σ_p^2 , σ_p^2 and σ_s^2 , respectively.

C. Receiver

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

$$\begin{aligned} \mathbf{r}_{s,2k+1} &= \begin{bmatrix} r_{s,2k+1} \\ r_{s,2k+2}^* \end{bmatrix} = \sqrt{\frac{E_s}{2}} \begin{bmatrix} \alpha_{2k+1} & -\beta_{2k+1} \\ \beta_{2k+2}^* & \alpha_{2k+2}^* \end{bmatrix} \begin{bmatrix} s_{2k+1} \\ s_{2k+2}^* \end{bmatrix} + \begin{bmatrix} n_{s,2k+1} \\ n_{s,2k+2}^* \end{bmatrix} \\ r_{p,k}^{(1)} &= \sqrt{\frac{E_p}{2}} \alpha_k + n_{p,k}^{(1)} \rightarrow \mathbf{r}_{p,k}^{(1)} = [r_{p,k-M}^{(1)} \cdots r_{p,k}^{(1)} \cdots r_{p,k+M}^{(1)}]^T \\ r_{p,k}^{(2)} &= \sqrt{\frac{E_p}{2}} \beta_k + n_{p,k}^{(2)} \rightarrow \mathbf{r}_{p,k}^{(2)} = [r_{p,k-M}^{(2)} \cdots r_{p,k}^{(2)} \cdots r_{p,k+M}^{(2)}]^T, \end{aligned} \quad (5)$$

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

$$\hat{\alpha}_k = \mathbf{h}^H \mathbf{r}_{p,k}^{(1)}, \quad \hat{\beta}_k = \mathbf{h}^H \mathbf{r}_{p,k}^{(2)}, \quad (6)$$

where $\mathbf{h} = [h_M \dots h_0 \dots h_{-M}]^T$ 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.

IV. THE LINEAR-COMBINING SCHEME

The linear-combining space-time decoder (LC-STD) is the simple space-time decoder originally suggested by Alamouti in [1] and it is the most likely scheme to be implemented in practice due to its low complexity. In this section, we derive 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.

A. 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,

$$\begin{bmatrix} z_1 \\ z_2^* \end{bmatrix} = \begin{bmatrix} \hat{\alpha}_1^* & \hat{\beta}_2 \\ -\hat{\beta}_1^* & \hat{\alpha}_2 \end{bmatrix} \mathbf{r}_{s,1} = \sqrt{\frac{E_s}{2}} \mathbf{G}_1 \begin{bmatrix} s_1 \\ s_2^* \end{bmatrix} + \begin{bmatrix} \tilde{n}_{s,1} \\ \tilde{n}_{s,2} \end{bmatrix}, \quad (7)$$

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

$$\mathbf{G} = \begin{bmatrix} \alpha_1 \hat{\alpha}_1^* + \beta_2^* \hat{\beta}_2 & \alpha_2^* \hat{\beta}_2 - \beta_1 \hat{\alpha}_1^* \\ \beta_2^* \hat{\alpha}_2 - \alpha_1 \hat{\beta}_1^* & \beta_1 \hat{\beta}_1^* + \alpha_2^* \hat{\alpha}_2 \end{bmatrix}. \quad (8)$$

Notice that, in an ideal environment, where CSI is perfect and channels are quasi-static, we have $\hat{\alpha}_1 = \hat{\alpha}_2 = \alpha_1 = \alpha_2$, $\hat{\beta}_1 = \hat{\beta}_2 = \beta_1 = \beta_2$. Therefore, the linear-combining scheme can eliminate interference within the code block completely [1] and \mathbf{G} becomes

$$\mathbf{G}_{\text{perfect}} = \begin{bmatrix} |\alpha_1|^2 + |\beta_1|^2 & 0 \\ 0 & |\alpha_1|^2 + |\beta_1|^2 \end{bmatrix}. \quad (9)$$

Our goal 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_{Alamouti} , which is upper bounded by

$$P_{\text{Alamouti}} \leq 2P_b. \quad (10)$$

Since the bit error probabilities of the first symbol and the second symbol are equal, we can focus on the bit error probability of the first symbol only. Without loss of generality, we assume that $s_1 = 1$. Consequently, the bit error probability becomes

$$P_b = \frac{1}{2} (P_{b,s_2=1} + P_{b,s_2=-1}), \quad (11)$$

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

$$P_{b,s_2=a} = \Pr \{ \Re[z_1] < 0 \mid s_2 = a \}. \quad (12)$$

We can see 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

$$\mathbf{x}_1 = [r_{s,1} \ r_{s,2} \ \hat{\alpha}_1 \ \hat{\beta}_2]^T, \text{ and } \mathbf{Q}_1 = \frac{1}{2} \begin{bmatrix} \mathbf{0}_2 & \mathbf{I}_2 \\ \mathbf{I}_2 & \mathbf{0}_2 \end{bmatrix}. \quad (13)$$

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 (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

$$\Sigma_1 = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^H & \mathbf{A}_{22} \end{bmatrix}, \quad (14)$$

where

$$\mathbf{A}_{11} = E_s \sigma_c^2 \begin{bmatrix} (1 - \rho a) + \bar{\gamma}_s^{-1} & 0 \\ 0 & (1 + \rho a) + \bar{\gamma}_s^{-1} \end{bmatrix}, \mathbf{A}_{12} = \frac{\sqrt{E_s E_p} \sigma_c^2}{2} \begin{bmatrix} \mathbf{w}_0^H \mathbf{h} (1 - \rho a) & \mathbf{w}_1^H \mathbf{h} (\rho - a) \\ \mathbf{w}_{-1}^H \mathbf{h} (\rho + a) & \mathbf{w}_0^H \mathbf{h} (1 + \rho a) \end{bmatrix},$$

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

where $\bar{\gamma}_p$ and $\bar{\gamma}_s$ are the average pilot SNR and the average data SNR, respectively, \mathbf{D}_e is a square matrix of order $2M + 1$ with $D_e(m, n) = R((e + m - n)\tau)$ 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 (2) or the Gauss-Chebyshev approximation (4) as described in section II. In addition, P_b and the upper bound of P_{Alamouti} can be found from (11) and (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

$$\mathbf{h} = \sqrt{\frac{2}{E_p}} (\mathbf{D}_0 + 2\bar{\gamma}_p^{-1} \mathbf{I}_{2M+1})^{-1} \mathbf{w}_0. \quad (15)$$

Since the covariance matrix is independent of the value of s_2 when $\rho = 0$, we assume, without loss of generality, that $s_2 = 1$. Therefore, the covariance matrix is simplified to

$$\hat{\Sigma}_1 = \begin{bmatrix} E_s \sigma_c^2 (1 + \bar{\gamma}_s^{-1}) \mathbf{I}_2 & \frac{\sigma_c^2}{2} \sqrt{\frac{E_s}{2}} \begin{bmatrix} \varepsilon_0 & -\varepsilon_1 \\ \varepsilon_1 & \varepsilon_0 \end{bmatrix} \\ \frac{\sigma_c^2}{2} \sqrt{\frac{E_s}{2}} \begin{bmatrix} \varepsilon_0 & \varepsilon_1 \\ -\varepsilon_1 & \varepsilon_0 \end{bmatrix} & \sigma_c^2 \frac{\varepsilon_0}{2} \mathbf{I}_2 \end{bmatrix}, \quad (16)$$

where $\varepsilon_0 = 2\mathbf{w}_0^H (\mathbf{D}_0 + 2\bar{\gamma}_p^{-1} \mathbf{I}_{2M+1})^{-1} \mathbf{w}_0$ and $\varepsilon_1 = 2\mathbf{w}_1^H (\mathbf{D}_0 + 2\bar{\gamma}_p^{-1} \mathbf{I}_{2M+1})^{-1} \mathbf{w}_0$. 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

$$\Upsilon = \left(\frac{4(1 + \bar{\gamma}_s^{-1})}{\varepsilon_0} - \left(\frac{\varepsilon_1}{\varepsilon_0} \right)^2 \right)^{1/2}. \quad (17)$$

To use the residue theorem, we need to identify the poles of $\frac{\Phi_{z_1}(s)}{s}$, i.e., 0 , λ_1^{-1} and λ_2^{-1} , that are in the left half-plane. Since we know that the minimum mean square error (MMSE) of the channel estimate is equal to $2\sigma_c^2(1 - \varepsilon_0/2)$ and $\mathbf{D}_0 + 2\bar{\gamma}_p^{-1} \mathbf{I}_{2M+1}$ is positive definite, we can conclude that $2 \geq \varepsilon_0 \geq 0$. In addition, since the MMSE is smaller than $E[\|\alpha_{k+1} - \hat{\alpha}_k\|^2]$, we can conclude 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 $\Upsilon > 1$, λ_1^{-1} is positive and λ_2^{-1} is negative.

From the residue theorem, the bit error probability is equal to the negative of the summation of the residues of $\frac{\Phi_{\mathfrak{R}[z_1]}(s)}{s}$ at the negative poles. Using the fact that the residue is invariant to the scaling of the poles [6], we scale the original poles with $\sigma_c^2 \varepsilon_0$ and the compact form of the bit error probability can be calculated as follows

$$P_b = \lim_{s \rightarrow \frac{\sigma_c^2 \varepsilon_0}{\lambda_2}} \frac{d}{ds} \left(\frac{(s - \frac{\sigma_c^2 \varepsilon_0}{\lambda_2})^2}{s(1 - \frac{s\lambda_1}{\sigma_c^2 \varepsilon_0})^2 (1 - \frac{s\lambda_2}{\sigma_c^2 \varepsilon_0})^2} \right) = \frac{1}{4} (2 + \Upsilon^{-1}) (1 - \Upsilon^{-1})^2. \quad (18)$$

As a sanity check, we compare (18) with the perfect CSI result in the existing papers by setting $\bar{\gamma}_p = \infty$. Since $\mathbf{D}_0^{-1}\mathbf{D}_0 = \mathbf{I}_{2M+1}$ and \mathbf{w}_0 is the $M + 1^{th}$ column of \mathbf{D}_0 , we get $\mathbf{D}_0^{-1}\mathbf{w}_0 = [0 \dots 0 \ 1 \ 0 \dots 0]^T$ (with one in the $(M + 1)^{th}$ row). Therefore, the values of $\varepsilon_0, \varepsilon_1$ and Υ 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 Υ into (18), we can obtain the closed-form expression of the bit error probability when $\rho = 0$ with the Wiener filter as the channel estimator.

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

$$R(\tau) = 0 \rightarrow P_b = \frac{1}{4} \left(2 + \sqrt{\frac{\bar{\gamma}_s}{2\bar{\gamma}_s + 2}} \right) \left(1 - \sqrt{\frac{\bar{\gamma}_s}{2\bar{\gamma}_s + 2}} \right)^2, \quad (19)$$

$$R(\tau) = 1 \rightarrow P_b = \frac{1}{4} \left(2 + \sqrt{\frac{\bar{\gamma}_s}{\bar{\gamma}_s + 2}} \right) \left(1 - \sqrt{\frac{\bar{\gamma}_s}{\bar{\gamma}_s + 2}} \right)^2. \quad (20)$$

It is apparent that (20) agrees with the result derived in [7], [13].

B. 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 \mathbf{c}_j instead of the transmitted sequence \mathbf{c}_i , where $i \neq j$. For the system model described in section III, 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 [29]. Therefore, we adopt similar simplifications to those used in [6], 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 [30]. 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 6 and the interleaving depth of 4 symbols is shown in fig. 2. Using the LC-STD with the ML convolutional decoder, the PEP, denoted by P_2 , can be written as

$$P_2 = \Pr \left\{ \sum_{k=0}^{d-1} \Re\{z_{kI+1}\} < 0 \right\}. \quad (21)$$

Notice that in (21) we assume that the error symbol starts at the first symbol. By inspection, we can write $\sum_{k=0}^{d-1} \Re\{z_{kI+1}\}$ in a Hermitian quadratic form $\sum_{k=0}^{d-1} \Re\{z_{kI+1}\} = \mathbf{x}_2^H \mathbf{Q}_2 \mathbf{x}_2$, where

$$\mathbf{x}_2 = [\mathbf{x}_2^{(1)} \ \mathbf{x}_2^{(2)} \ \mathbf{x}_2^{(3)} \ \mathbf{x}_2^{(4)}]^T, \mathbf{Q}_2 = \frac{1}{2} \begin{bmatrix} \mathbf{0}_{2d} & \mathbf{I}_{2d} \\ \mathbf{I}_{2d} & \mathbf{0}_{2d} \end{bmatrix}, \quad (22)$$

and $\mathbf{x}_2^{(1)}, \mathbf{x}_2^{(2)}, \mathbf{x}_2^{(3)}, \mathbf{x}_2^{(4)}$ are row vectors of length d with $\mathbf{x}_2^{(1)}(m) = r_{s,g^{(1)}(m)}$, $\mathbf{x}_2^{(2)}(m) = r_{s,g^{(2)}(m)}$, $\mathbf{x}_2^{(3)}(m) = \hat{\alpha}_{g^{(1)}(m)}$, $\mathbf{x}_2^{(4)}(m) = \hat{\beta}_{g^{(2)}(m)}$,

$$g^{(1)}(k) = \begin{cases} (k-1)I + 1 & ; \text{ when } I \text{ is even} \\ 4\lfloor k/2 \rfloor + 6\lfloor (k-1)/2 \rfloor + 1 & ; \text{ otherwise,} \end{cases} \quad (23)$$

$g^{(2)}(k) = g^{(1)}(k) + 1$ and $\lfloor x \rfloor$ denotes the closest integer that is less or equal to x . Since \mathbf{x}_2 is a zero-mean complex Gaussian random vector, the characteristic function of $\mathbf{x}_2^H \mathbf{Q}_2 \mathbf{x}$ can be used with the residue theorem or the Gauss-Chebyshev approximation to find the PEP. The covariance matrix of \mathbf{x}_2 , denoted by Σ_2 , is needed to find the characteristic function. From the above expression, it can be shown that

$$\Sigma_2 = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{0}_d & \mathbf{B}_{13} & \mathbf{B}_{14} \\ \mathbf{0}_d & \mathbf{B}_{22} & \mathbf{B}_{23} & \mathbf{B}_{24} \\ \mathbf{B}_{13}^H & \mathbf{B}_{23}^H & \mathbf{B}_{33} & \mathbf{B}_{34} \\ \mathbf{B}_{14}^H & \mathbf{B}_{24}^H & \mathbf{B}_{34}^H & \mathbf{B}_{33} \end{bmatrix}, \quad (24)$$

where

$$\begin{aligned} \mathbf{B}_{11}(m, n) &= E_s \sigma_c^2 (1 - \rho) R(q^{(1)}(m, n) \tau) + \sigma_s^2 \delta(m - n), \quad \mathbf{B}_{13}(m, n) = \frac{\sigma_c^2 \sqrt{E_s E_p}}{2} (1 - \rho) \mathbf{w}_{q^{(1)}(n, m)}^H \mathbf{h} \\ \mathbf{B}_{14}(m, n) &= \frac{\sigma_c^2 \sqrt{E_s E_p}}{2} (\rho - 1) \mathbf{w}_{q^{(1)}(n, m) + 1}^H \mathbf{h}, \quad \mathbf{B}_{22}(m, n) = E_s \sigma_c^2 (1 + \rho) R(q^{(1)}(m, n) \tau) + \sigma_s^2 \delta(m - n) \\ \mathbf{B}_{23}(m, n) &= \frac{\sigma_c^2 \sqrt{E_s E_p}}{2} (1 + \rho) \mathbf{w}_{q^{(1)}(n, m) - 1}^H \mathbf{h}, \quad \mathbf{B}_{24}(m, n) = \frac{\sigma_c^2 \sqrt{E_s E_p}}{2} (1 + \rho) \mathbf{w}_{q^{(1)}(n, m)}^H \mathbf{h} \\ \mathbf{B}_{33}(m, n) &= \mathbf{h}^H \left(\frac{E_p \sigma_c^2}{2} \mathbf{D}_{q^{(1)}(m, n)} + \sigma_p^2 \boldsymbol{\delta}_{q^{(1)}(m, n)} \right) \mathbf{h}, \quad \mathbf{B}_{34}(m, n) = \frac{\rho E_p \sigma_c^2}{2} \mathbf{h}^H \mathbf{D}_{q^{(1)}(m, n) - 1} \mathbf{h}, \\ q^{(1)}(m, n) &= g^{(1)}(m) - g^{(1)}(n) \text{ and } \boldsymbol{\delta}_e \text{ is a square matrix of size } 2M + 1 \text{ with ones on the } e^{th} \text{ diagonal.} \end{aligned}$$

V. 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.

A. 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, we assume 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

$$P_{\text{Alamouti}} \leq P_{[1 \ -1]} + P_{[-1 \ 1]} + P_{[-1 \ -1]}. \quad (25)$$

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 $i = 1, \dots, 3$ can be simplified to

$$\tilde{P}_i = \Pr \{ \Re [z^{(i)}] < 0 \}, \quad (26)$$

where $z^{(1)} = -\sqrt{\frac{E_s}{2}} (r_{s,1}\hat{\beta}_1^* - r_{s,2}\hat{\alpha}_2^*) + \frac{E_s}{2} (\hat{\alpha}_1\hat{\beta}_1^* - \hat{\alpha}_2\hat{\beta}_2^*)$, $z^{(2)} = \sqrt{\frac{E_s}{2}} (r_{s,1}\hat{\alpha}_1^* + r_{s,2}\hat{\beta}_2^*) + \frac{E_s}{2} (\hat{\alpha}_1\hat{\beta}_1^* - \hat{\alpha}_2\hat{\beta}_2^*)$ and $z^{(3)} = \sqrt{\frac{E_s}{2}} r_{s,1}(\hat{\alpha}_1 - \hat{\beta}_1)^* + \sqrt{\frac{E_s}{2}} r_{s,2}(\hat{\alpha}_2 + \hat{\beta}_2)^*$. By inspection, we can write $\Re[z^{(i)}]$ in the Hermitian quadratic form $\Re[z^{(i)}] = \mathbf{x}_3^H \mathbf{Q}_3^{(i)} \mathbf{x}_3$, where

$$\mathbf{x}_3 = \left[r_{s,1} \ r_{s,2} \ \sqrt{\frac{E_s}{2}} \hat{\alpha}_1 \ \sqrt{\frac{E_s}{2}} \hat{\alpha}_2 \ \sqrt{\frac{E_s}{2}} \hat{\beta}_1 \ \sqrt{\frac{E_s}{2}} \hat{\beta}_2 \right]^T, \quad (27)$$

and

$$\mathbf{Q}_3^{(1)} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}, \mathbf{Q}_3^{(2)} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \end{bmatrix}, \mathbf{Q}_3^{(3)} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Since \mathbf{x}_3 is a zero-mean complex Gaussian random vector, the characteristic function of $\mathbf{x}_3^H \mathbf{Q}_3^{(i)} \mathbf{x}_3$ can be used with the residue theorem or the Gauss-Chebyshev approximation to find \tilde{P}_i . In addition, the covariance matrix of \mathbf{x}_3 can be found to be

$$\Sigma_3 = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{12}^H & \mathbf{C}_{22} & \mathbf{C}_{23} \\ \mathbf{C}_{13}^H & \mathbf{C}_{23}^H & \mathbf{C}_{22} \end{bmatrix}, \quad (28)$$

where

$$\mathbf{C}_{11} = E_s \sigma_c^2 \begin{bmatrix} 1 - \rho + \bar{\gamma}_s^{-1} & 0 \\ 0 & 1 + \rho + \bar{\gamma}_s^{-1} \end{bmatrix}, \quad \mathbf{C}_{12} = \frac{E_s \sigma_c^2}{2} \sqrt{\frac{E_p}{2}} \begin{bmatrix} (1 - \rho) \mathbf{w}_0^H \mathbf{h} & (1 - \rho) \mathbf{w}_1^H \mathbf{h} \\ (1 + \rho) \mathbf{w}_{-1}^H \mathbf{h} & (1 + \rho) \mathbf{w}_0^H \mathbf{h} \end{bmatrix}$$

$$\mathbf{C}_{13} = \frac{E_s \sigma_c^2}{2} \sqrt{\frac{E_p}{2}} \begin{bmatrix} (\rho - 1) \mathbf{w}_0^H \mathbf{h} & (\rho - 1) \mathbf{w}_1^H \mathbf{h} \\ (\rho + 1) \mathbf{w}_{-1}^H \mathbf{h} & (\rho + 1) \mathbf{w}_0^H \mathbf{h} \end{bmatrix}, \quad \mathbf{C}_{23} = \frac{\rho E_s E_p \sigma_c^2}{4} \begin{bmatrix} \mathbf{h}^H \mathbf{D}_0 \mathbf{h} & \mathbf{h}^H \mathbf{D}_{-1} \mathbf{h} \\ \mathbf{h}^H \mathbf{D}_1 \mathbf{h} & \mathbf{h}^H \mathbf{D}_0 \mathbf{h} \end{bmatrix},$$

$$\mathbf{C}_{22} = \frac{E_s E_p \sigma_c^2}{4} \begin{bmatrix} \mathbf{h}^H (\mathbf{D}_0 + 2\bar{\gamma}_p^{-1} \mathbf{I}_{2M+1}) \mathbf{h} & \mathbf{h}^H (\mathbf{D}_{-1} + 2\bar{\gamma}_p^{-1} \boldsymbol{\delta}_{-1}) \mathbf{h} \\ \mathbf{h}^H (\mathbf{D}_1 + 2\boldsymbol{\delta}_1 \bar{\gamma}_p) \mathbf{h} & \mathbf{h}^H (\mathbf{D}_0 + 2\bar{\gamma}_p^{-1} \mathbf{I}_{2M+1}) \mathbf{h} \end{bmatrix}.$$

B. 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 IV-B, 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

$$P_2 = \Pr\{\mathbf{x}_4^H \mathbf{Q}_4 \mathbf{x}_4 < 0\} \quad (29)$$

where $\mathbf{x}_4 = [\mathbf{x}_4^{(1)} \ \mathbf{x}_4^{(2)} \ \mathbf{x}_4^{(3)} \ \mathbf{x}_4^{(4)} \ \mathbf{x}_4^{(5)} \ \mathbf{x}_4^{(6)}]^T$ and $\mathbf{x}_4^{(i)}$ for $i = 1, \dots, 6$ are row vectors of length d with $\mathbf{x}_4^{(1)}(m) = r_{s,g^{(1)}(m)}$, $\mathbf{x}_4^{(2)}(m) = r_{s,g^{(2)}(m)}$, $\mathbf{x}_4^{(3)}(m) = \sqrt{\frac{E_s}{2}} \hat{\alpha}_{g^{(1)}(m)}$, $\mathbf{x}_4^{(4)}(m) = \sqrt{\frac{E_s}{2}} \hat{\alpha}_{g^{(2)}(m)}$, $\mathbf{x}_4^{(5)}(m) = \sqrt{\frac{E_s}{2}} \hat{\beta}_{g^{(1)}(m)}$ and $\mathbf{x}_4^{(6)}(m) = \sqrt{\frac{E_s}{2}} \hat{\beta}_{g^{(2)}(m)}$. In addition, \mathbf{Q}_4 is a square matrix with

$$\mathbf{Q}_4 = \frac{1}{2} \begin{bmatrix} \mathbf{0}_d & \mathbf{0}_d & \mathbf{0}_d & \mathbf{0}_d & -\mathbf{I}_d & \mathbf{0}_d \\ \mathbf{0}_d & \mathbf{0}_d & \mathbf{0}_d & \mathbf{I}_d & \mathbf{0}_d & \mathbf{0}_d \\ \mathbf{0}_d & \mathbf{0}_d & \mathbf{0}_d & \mathbf{0}_d & \mathbf{I}_d & \mathbf{0}_d \\ \mathbf{0}_d & \mathbf{I}_d & \mathbf{0}_d & \mathbf{0}_d & \mathbf{0}_d & -\mathbf{I}_d \\ -\mathbf{I}_d & \mathbf{0}_d & \mathbf{I}_d & \mathbf{0}_d & \mathbf{0}_d & \mathbf{0}_d \\ \mathbf{0}_d & \mathbf{0}_d & \mathbf{0}_d & -\mathbf{I}_d & \mathbf{0}_d & \mathbf{0}_d \end{bmatrix} \quad (30)$$

Since \mathbf{x}_4 is a zero-mean complex Gaussian random vector, the PEP can be computed by means of the residue theorem (2) or the Gauss-Chebyshev approximation (4). From the above expressions, the covariance matrix can be found to be

$$\Sigma_4 = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{0}_d & \sqrt{\frac{E_s}{2}} \mathbf{B}_{13} & -\sqrt{\frac{E_s}{2}} \mathbf{B}_{14} & -\sqrt{\frac{E_s}{2}} \mathbf{B}_{15} & \sqrt{\frac{E_s}{2}} \mathbf{B}_{14} \\ \mathbf{0}_d & \mathbf{B}_{22} & \sqrt{\frac{E_s}{2}} \mathbf{B}_{23} & \sqrt{\frac{E_s}{2}} \mathbf{B}_{24} & \sqrt{\frac{E_s}{2}} \mathbf{B}_{23} & \sqrt{\frac{E_s}{2}} \mathbf{B}_{24} \\ \sqrt{\frac{E_s}{2}} \mathbf{B}_{13}^H & \sqrt{\frac{E_s}{2}} \mathbf{B}_{23}^H & \frac{E_s}{2} \mathbf{B}_{33} & \mathbf{E}_{34} & \frac{E_s}{2} \mathbf{B}_{34} & \frac{E_s}{2} \mathbf{B}_{34} \\ -\sqrt{\frac{E_s}{2}} \mathbf{B}_{14}^H & \sqrt{\frac{E_s}{2}} \mathbf{B}_{24}^H & \mathbf{E}_{34}^H & \frac{E_s}{2} \mathbf{B}_{33} & \mathbf{E}_{45} & \mathbf{E}_{46} \\ -\sqrt{\frac{E_s}{2}} \mathbf{B}_{15}^H & \sqrt{\frac{E_s}{2}} \mathbf{B}_{23}^H & \frac{E_s}{2} \mathbf{B}_{34}^H & \mathbf{E}_{45}^H & \frac{E_s}{2} \mathbf{B}_{33} & \mathbf{E}_{34} \\ \sqrt{\frac{E_s}{2}} \mathbf{B}_{14}^H & \sqrt{\frac{E_s}{2}} \mathbf{B}_{24}^H & \frac{E_s}{2} \mathbf{B}_{34}^H & \mathbf{E}_{46}^H & \mathbf{E}_{34}^H & \frac{E_s}{2} \mathbf{B}_{33} \end{bmatrix}, \quad (31)$$

where \mathbf{B}_{11} , \mathbf{B}_{22} , \mathbf{B}_{13} , \mathbf{B}_{23} , \mathbf{B}_{14} , \mathbf{B}_{24} , \mathbf{B}_{33} and \mathbf{B}_{34} are as defined in section IV-B. In addition, $\mathbf{E}_{34}(m, n) = \frac{E_s}{2} \mathbf{h}^H \left(\frac{E_p \sigma_c^2}{2} \mathbf{D}_{q^{(1)}(m,n)-1} + \sigma_p^2 \delta_{q^{(1)}(m,n)-1} \right) \mathbf{h}$, $\mathbf{E}_{45}(m, n) = \frac{\rho E_s E_p \sigma_c^2}{4} \mathbf{h}^H \mathbf{D}_{q^{(1)}(m,n)+1} \mathbf{h}$ and $\mathbf{E}_{46}(m, n) = \frac{\rho E_s E_p \sigma_c^2}{4} \mathbf{h}^H \mathbf{D}_{q^{(1)}(m,n)} \mathbf{h}$

VI. 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, we assume 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 [31], 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.

A. Alamouti space-time code

Fig. 3 and 4 compare the simulation results to the performance bounds of P_{Alamouti} employing the LC-STD and the ML-STD (from (10) and (25), respectively) for a 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, we can see that the bounds are reasonably tight especially at low probability of error. Comparing fig. 3 and 4, we can see 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 [32]. Using the Wiener filter as the channel estimator, the bit error probability of the SISO system can be written as [32]

$$P_b = \frac{1}{2} \left(1 - \sqrt{\frac{\mathbf{w}_0^H (\mathbf{D}_0 + \bar{\gamma}_p^{-1} \mathbf{I}_d)^{-1} \mathbf{w}_0}{1 + \bar{\gamma}_s^{-1}}} \right), \quad (32)$$

where \mathbf{D}_0 , \mathbf{w}_0 , $\bar{\gamma}_p$, $\bar{\gamma}_s$ are as previously defined. In fig. 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 (small $f_d\tau$) 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. We can also see that at small pilot SNR or large $f_d\tau$, the SISO system provides better performance than the Alamouti coded system with the LC-STD.

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 (10) and (25)) of P_{Alamouti} for a Wiener filter and for a moving average filter are illustrated in fig. 6. Both filters are assumed to have 11 taps and the upper bounds of P_{Alamouti} corresponding to $f_d\tau = 0.0005$ and 0.05 are illustrated. From the figure, we notice 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, we can see a bigger performance loss 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.

B. Concatenated convolutional code and Alamouti space-time code

As defined above, 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. 7, 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 [6] (convolutional code without Alamouti space-time code) is also illustrated in this figure by the dashed line. Comparing the performance of the LC-ML and the JML, we can see 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, we can conclude that, considering the complexity and the performance, the LC-ML is a very good choice for a practical CCA system.

We can see 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' power spectral density function [6]. In addition, the estimation-diversity tradeoff can be seen in the CCA as well as the SISO system (see [5], [6] 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 ρ is small. However, the performance degrades with ρ . When the transmit antennas are fully correlated ($\rho = 1$), we can see that the performance of the CCA is worse than that of the SISO.

It is also important to point out that the pilot SNR used in fig. 7 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 ρ 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. 8. In this figure, we observe

that, even when there is no spatial correlation ($\rho = 0$) the SISO outperforms the CCA at large $f_d\tau$. 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. 9, 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, we can separate the figure 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, we can see that 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 our analysis, analytical results are compared with results from Monte Carlo simulations in fig. 10, 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 [4] 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 [33] with Jakes' power spectral density. The channel estimator is an 11-tap Wiener filter. For the analytical results, we calculated the bit error probability and the frame error probability by taking into account the PEP corresponding to the five smallest Hamming weights ($d = 18, 20, 22, 24, 26$) [34]. From the figure, we can see that the analytical results and the simulation results match well.

VII. CONCLUSION

In this paper, 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. Realistic assumptions such as noisy channel estimates and non-quasi-static fading channels allow us to evaluate the effect of these undesirable factors on the system performance. 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 scheme. However, the difference diminishes when the channel estimator is not optimized. In addition, comparison between the linear-combining scheme and the SISO system reveals that the Alamouti space-time code does not always outperform the SISO system.

We have 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 scheme with the ML convolutional decoder (LC-ML) and the joint Alamouti space-time and convolutional ML decoder (JML). Our 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 simulations and results from the analysis has verified that the analysis can predict the system performance with high accuracy.

REFERENCES

- [1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.
- [2] T. H. Liew and L. Hanzo, "Space-time codes and concatenated channel codes for wireless communications," *Proc. IEEE*, vol. 90, pp. 187–219, Feb. 2002.
- [3] 3GPP, "Technical specification group radio access network, physical channels and mapping of transport channels onto physical channels (FDD), 25.211."
- [4] 3GPP, "Technical specification group radio access network, physical layer procedure (FDD), 25.212."
- [5] A. P. Worthen and W. E. Stark, *On the channel memory-diversity tradeoff in communication systems*. in Information, coding and mathematics Kluwer Academic Publishers, 2002.
- [6] J. Jootar, J. R. Zeidler, and J. G. Proakis, "Performance of convolutional codes with finite-depth interleaving and noisy channel estimates," to be published in *IEEE Trans. Commun.*
- [7] A. Vielmon, Y. Li, and J. R. Barry, "Performance of Alamouti transmit diversity over time-varying Rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1369 – 1373, Sept. 2004.
- [8] Y. Zhang and D. Li, "Performance of space-time transmit diversity over time-selective fading channels," in *Proc. 9th IEEE Asia Pacific Conference on Communications (APCC)*, pp. 402–405, 2003.
- [9] P. Schulz-Rittich, J. Baltersee, and G. Fock, "Channel estimation for DS-CDMA with transmit diversity over frequency selective fading channels," in *Proc. IEEE VTC Spring 2001*, pp. 1973 – 1977, May 2001.
- [10] M. Stege, M. Bronzel, and G. Fettweis, "On the performance of space-time blockcodes," in *Proc. IEEE VTC Spring 2001*, pp. 2282–2286, May 2001.
- [11] R. M. Buehrer and N. A. Kumar, "The impact of channel estimation error on space-time block codes," in *Proc. IEEE VTC Fall 2002*, pp. 1921–1925, Sept. 2002.
- [12] H. Cheon and D. Hong, "Performance analysis of space-time block codes in time-varying Rayleigh fading channels," in *Proc. IEEE ICASSP'02*, pp. 2357 – 2360, May 2002.
- [13] D. Gu and C. Leung, "Performance analysis of transmit diversity scheme with imperfect channel estimation," *Electronics Letters*, vol. 39, pp. 402–403, Feb. 2003.
- [14] W. C. Ang and C. S. Ng, "Analysis of a transmit diversity scheme in fading channels and imperfect channel estimation," in *Proc. International Conference on Communication Systems 2002. (ICCS 2002)*, pp. 188–192, Nov. 2002.
- [15] P. Garg, R. K. Mallik, and H. M. Gupta, "Performance analysis of space-time coding with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 257–265, Jan. 2005.
- [16] X. Wang and J. Wang, "Effect of imperfect channel estimation on transmit diversity in CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 53, pp. 1400–1412, Sept. 2004.
- [17] K. Vanganuru and A. Annamalai, "Analysis of transmit diversity schemes: Impact of fade distribution, spatial correlation and channel estimation errors," in *Proc. Wireless Communications and Networking Conf., 2003*, pp. 247–251, Mar. 2003.

- [18] J. Wang, M. K. Simon, M. P. Fitz, and K. Yao, "On the performance of space-time codes over spatially correlated rayleigh fading channels," *IEEE Trans. Commun.*, vol. 52, pp. 877–881, June 2004.
- [19] E. A. Jorswieck and A. Sezgin, "Impact of spatial correlation on the performance of orthogonal space-time block codes," *IEEE Commun. Lett.*, vol. 8, pp. 21–23, Jan. 2004.
- [20] S. K. Jayaweera and H. V. Poor, "Turbo (iterative) decoding of a unitary space-time code with a convolutional code," in *Proc. IEEE VTC Spring 2002*, pp. 1020 – 1024, May 2000.
- [21] H. Schulze, "Performance analysis of concatenated space-time coding with two transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 669–679, July 2003.
- [22] G. Bauch and J. Hagenauer, "Analytical evaluation of space-time transmit diversity with FEC coding," in *Proc. IEEE Globecom 2001*, pp. 435–439, Nov. 2001.
- [23] M. Chiani, A. Conti, and V. Tralli, "Design and performance of bit-interleaved pragmatic space-time codes in block fading channels," in *Personal, Indoor and Mobile Radio Communications, 2002 13th IEEE International Symposium on*, pp. 683–687, Sept. 2002.
- [24] J. Jootar, J. R. Zeidler, and J. G. Proakis, "Performance of Alamouti space-time code in time-varying channels with noisy channel estimates," in *Proc. Wireless Communications and Networking Conf.*, pp. 498–503, Mar. 2005.
- [25] G. L. Turin, "The characteristic function of hermitian quadratic forms in complex normal variables," *Biometrika*, vol. 47, pp. 199–201, June 1960.
- [26] R. V. Nobelen and D. P. Taylor, "Analysis of the pairwise error probability of noninterleaved codes on the Rayleigh-fading channel," *IEEE Trans. Commun.*, vol. 44, pp. 456–463, Apr. 1996.
- [27] E. Biglieri, G. Caire, G. Taricco, and J. Ventura-Traveset, "Simple method for evaluating error probabilities," *Electronics Letters*, vol. 32, pp. 191–192, Feb. 1996.
- [28] J. R. Foerster and L. B. Milstein, "Coded modulation for a coherent DS-CDMA system employing an MMSE receiver in a fading channel," *IEEE Trans. Commun.*, vol. 48, pp. 1909–1918, Nov. 2000.
- [29] M. Rice and E. Perrins, "A simple figure of merit for evaluating interleaver depth for the land-mobile satellite channel," *IEEE Trans. Commun.*, vol. 49, pp. 1343–1353, Aug. 2001.
- [30] F. Gagnon and D. Haccoun, "Bound on the error performance of coding for nonindependent Rician-fading channels," *IEEE Trans. Commun.*, vol. 40, pp. 351–360, Feb. 1992.
- [31] W. C. Jakes, *Microwave Mobile Communications*. New Jersey: IEEE Press, 1994.
- [32] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.
- [33] M. Patzold, *Mobile Fading Channels*. New York: John Wiley and Sons, 1 ed., 2002.
- [34] E. Biglieri, D. Divsalar, P. J. McLane, and M. K. Simon, *Introduction to Trellis-Coded Modulation with Applications*. New York: MacMillan, 1991.

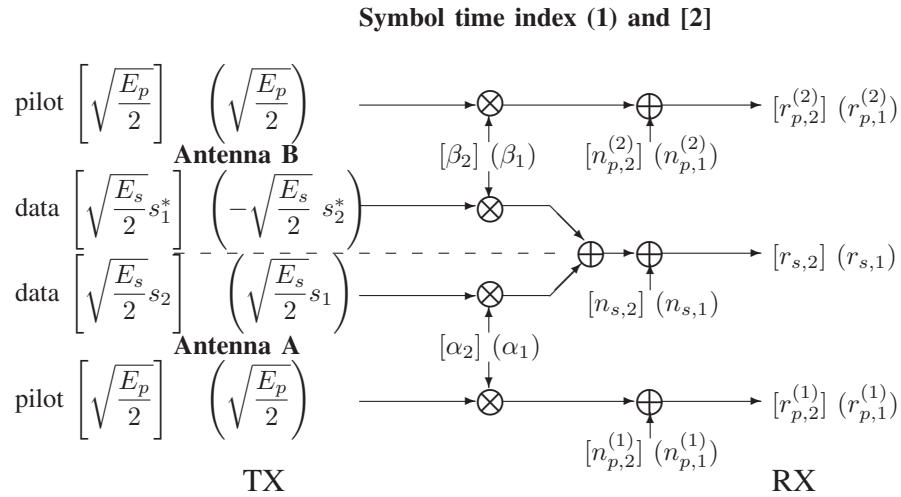


Fig. 1. System diagram of STTD over two-symbol period

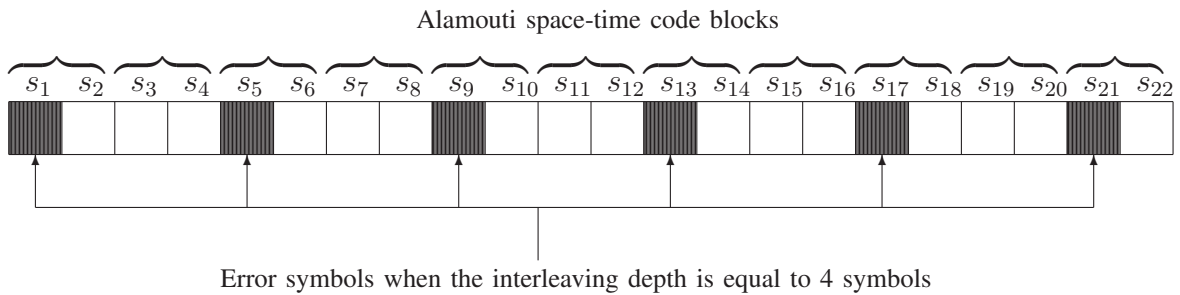


Fig. 2. Illustration of an error codeword with Hamming weight 6 and 4-symbol interleaving depth

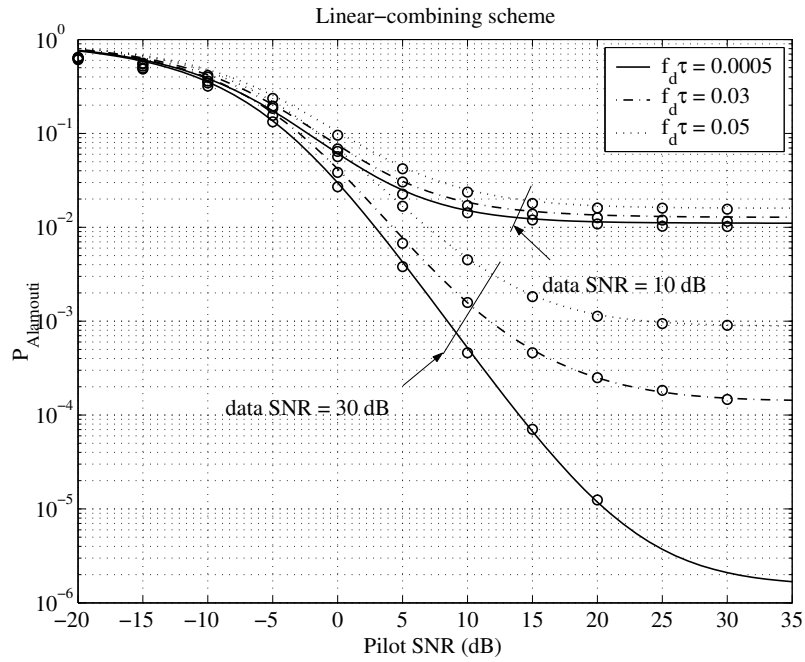


Fig. 3. Performance of the Alamouti space-time code with the LC-STD with 11-tap Wiener filter and Jakes PSD.

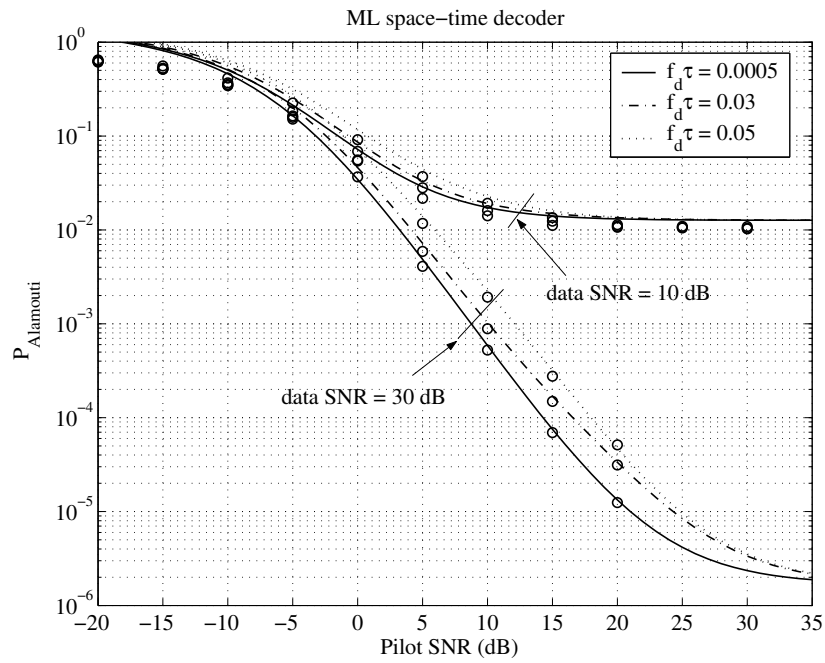


Fig. 4. Performance of the Alamouti space-time code with the ML-STD with 11-tap Wiener filter and Jakes PSD.

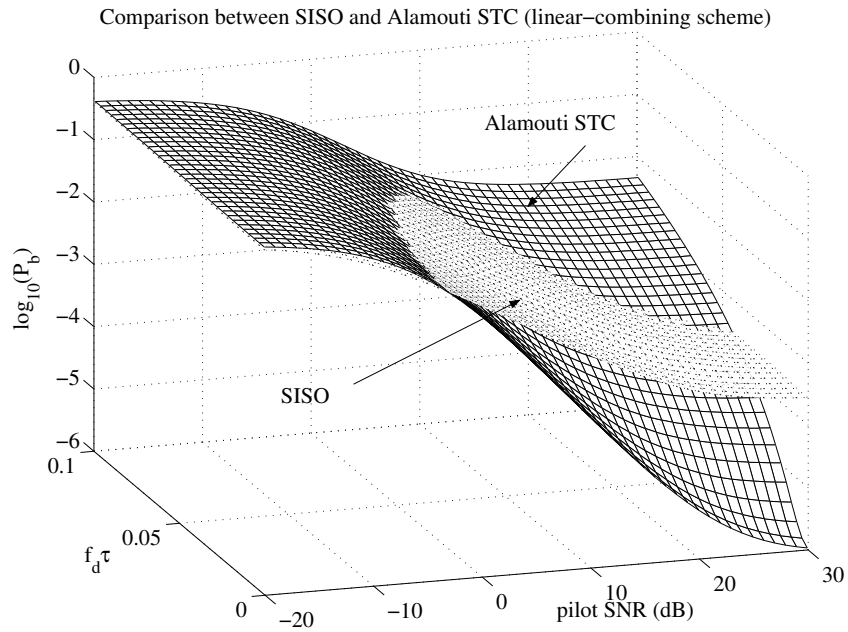


Fig. 5. Comparison between a SISO system and the Alamouti space-time code with the LC-STD: data SNR = 30 dB, 11-tap Wiener filter and Jakes PSD.

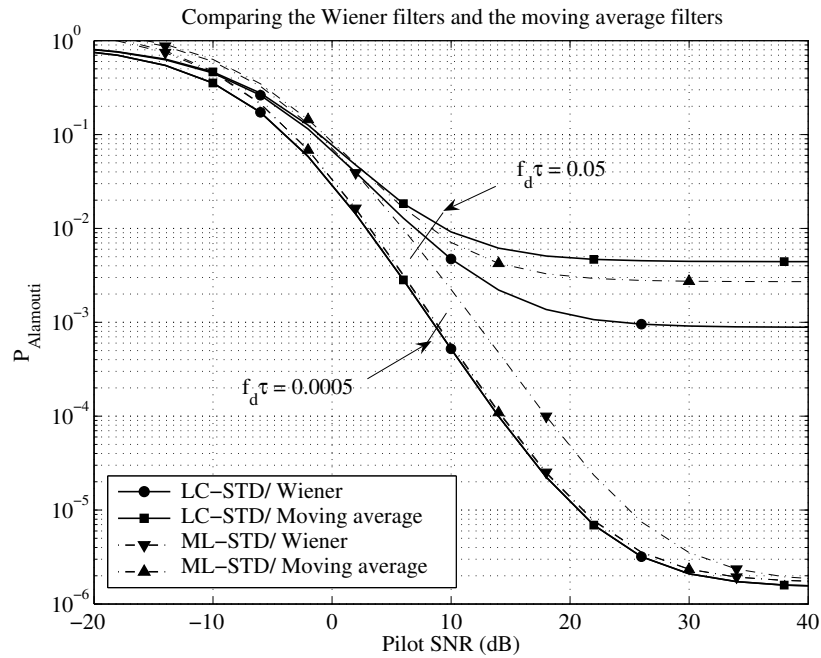


Fig. 6. Comparison between 11-tap Wiener filters and 11-tap moving average filters for the Alamouti space-time code: data SNR = 30 dB.

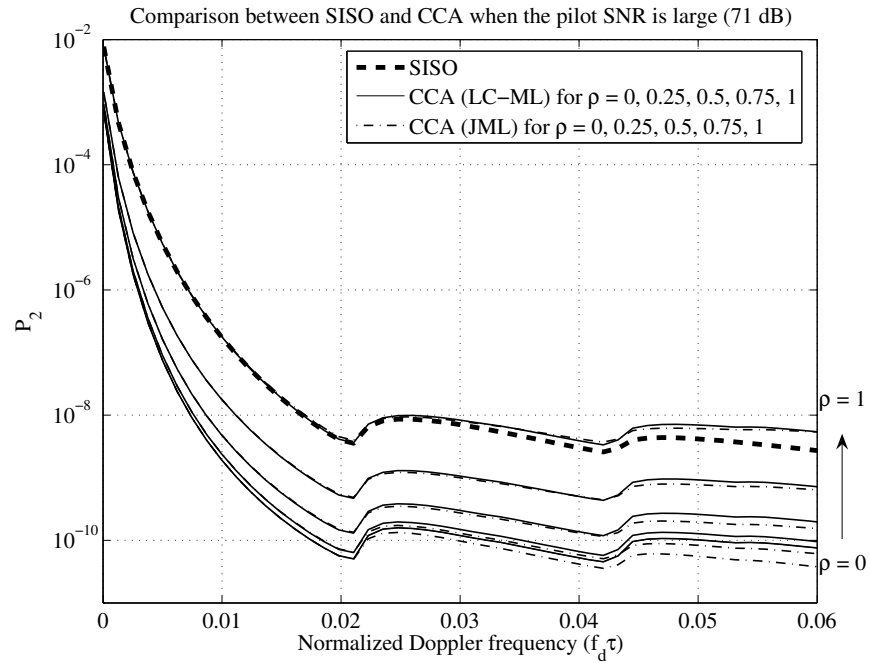


Fig. 7. Comparison between the convolutionally coded SISO system and the CCA scheme with the LC-ML and the JML: data SNR = 2.22 dB, pilot SNR = 71 dB, $d = 18$, Jakes PSD, 11-tap Wiener filter, $I = 23$.

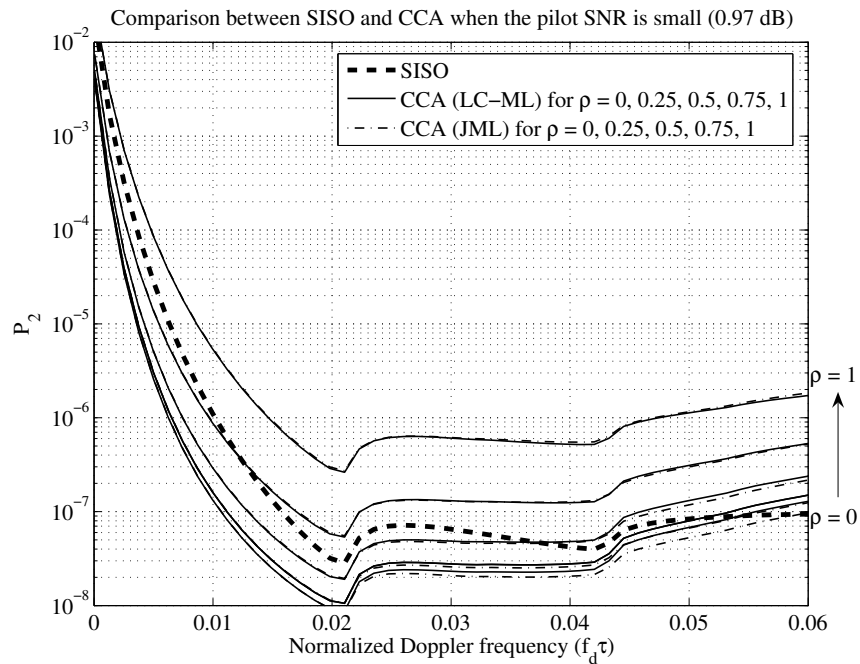


Fig. 8. Comparison between the convolutionally coded SISO system and the CCA scheme with the LC-ML and the JML: data SNR = 2.22 dB, pilot SNR = 0.97 dB, $d = 18$, Jakes PSD, 11-tap Wiener filter, $I = 23$

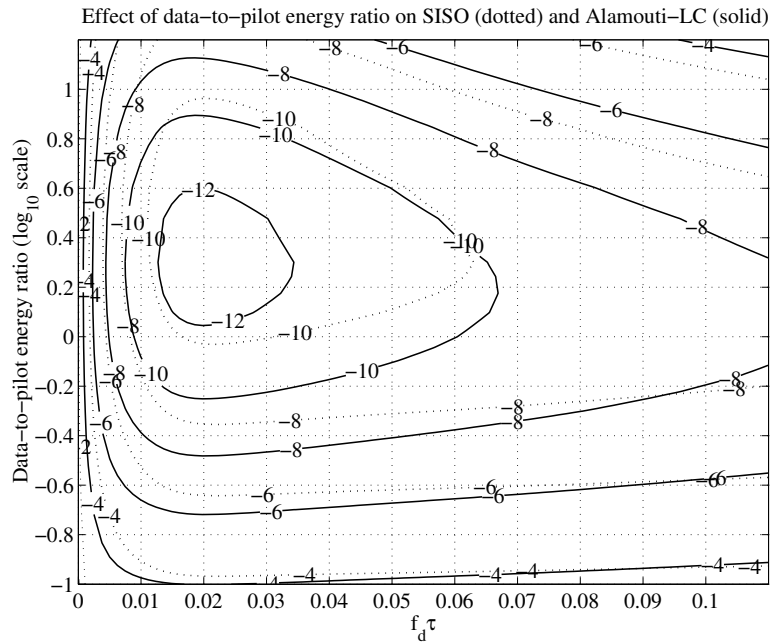


Fig. 9. Performance of the convolutionally coded SISO system and the CCA scheme with the LC-ML as a function of the data-to-pilot energy ratio and $f_d\tau$: $E_s + E_p = 1.01$, $\sigma_s^2 = \sigma_p^2 = 0.1$, $I = 23$, $d = 18$, Gaussian PSD, 11-tap Wiener filter

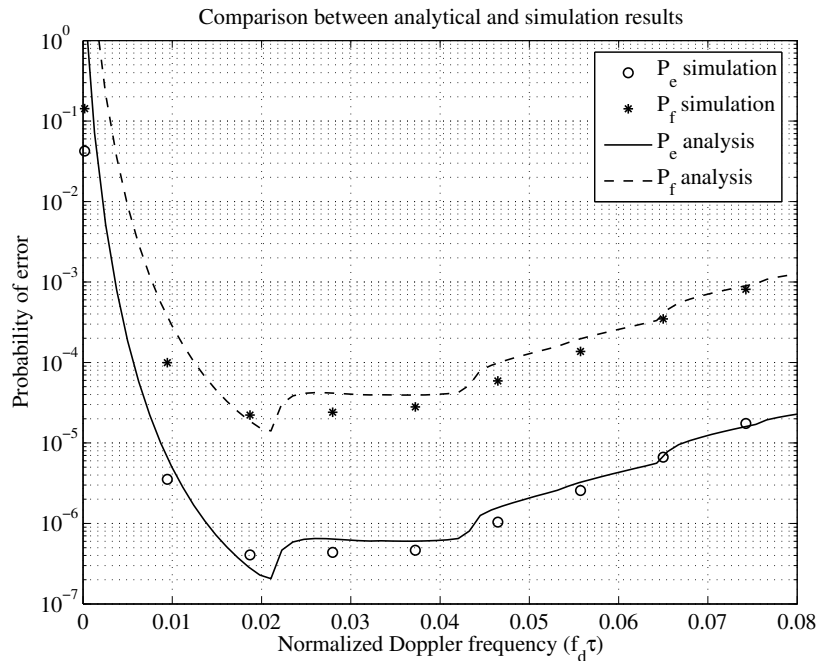


Fig. 10. Comparison between analytical and simulation results of the CCA system with LC-ML: data SNR = 2.22 dB, pilot SNR = 0.97 dB, Jakes PSD, 11-tap Wiener filter, rate 1/3 convolutional code, $d_{\min} = 18$, 220 information bits per block, 8 bits zero-padding