Optimal Differential Detection and Performance Analysis of Orthogonal Space-Time Block Codes over Semi-Identical MIMO Fading Channels

Meixia Tao and Pooi Yuen Kam
Department of Electrical and Computer Engineering
National University of Singapore
Singapore 117576
Email: {mxtao, elekampy}@nus.edu.sg

Abstract—This paper considers the optimal detector and its error performance of differential orthogonal space-time block codes over independent and semi-identically distributed, block Rayleigh fading channels. This semi-identically distributed case refers to the situation where the channel gains associated with a common receive antenna are identically distributed, but the ones associated with a common transmit antenna are not. We first derive the optimal symbol-by-symbol differential detector, and show that the conventional differential detector is suboptimal. We then derive an exact bit error probability expression for both BPSK and QPSK constellations based on the optimal detector. The result is applicable for any number of transmit and receive antennas, and assumes arbitrary block fading rates. Simple and insightful upper bounds are also obtained.

I. INTRODUCTION

Antenna diversity is a recognized technique to mitigate multipath fading without sacrificing channel bandwidth. While performance analysis for diversity reception in single-input multi-output (SIMO) channels has been investigated for decades, the design and performance of joint transmit and receive antenna diversity in multi-input multi-output (MIMO) systems is a recent research topic. Of the existing transmit antenna diversity techniques, orthogonal space-time block codes (OSTBC) [1]–[3] are the most popular because of their elegant code structure and low decoding complexity. It is well known that, when instantaneous channel state information (CSI) can be estimated at the receiver, the performance of OSTBC is identical to that of maximum ratio combing in SIMO channels (apart from being scaled down by the number of transmit antennas and coding rate). When CSI is unavailable at the receiver, OSTBC can be used in a differential encoding manner and still possess low decoding complexity [4]–[6]. Existing work on performance analysis of differential OSTBC can be found in [7]–[9]. All these previous results are based on a common assumption that the channel path gains associated with different transmit-receive antenna pairs are statistically identical. In real propagation environments, however, this assumption may not hold. For instance, the environmental scattering on the propagation paths observed by different antennas may be different. This may incur shadowing effects and, consequently, unequal path loss on different antennas. In MIMO systems, this is especially common when the antenna spacing is relatively large (compared with the carrier wavelength) to ensure low correlation. The performance analysis of diversity reception in SIMO channels using various kinds of modulation schemes over non-identically distributed branches has been studied in, for example, [10] and [11].

In this paper, we consider the error probability analysis of differential OSTBC over independent, semi-identically distributed (i.s.i.d) and block-correlated Rayleigh fading channels. This i.s.i.d case refers to the situation where the channel path gains associated with a common receive antenna are identically distributed, but the ones associated with a common transmit antenna are not. Such situation would most likely occur in the uplink of a cellular system, where the antennas on the base station are mounted relatively far apart from one another, whereas the antennas on a mobile handset are inside a small antenna panel. We first propose an optimal symbol-by-symbol differential detector, and show that the conventional differential detector is suboptimal in the i.s.i.d case. We then derive an exact bit error probability (BEP) expression for the optimal detector. This result is applicable for any number of receive antennas and transmit antennas where OSTBCs exist. It also assumes arbitrary block fading rates. This BEP expression involves only a single integral over finite limits, and hence can be numerically computed easily. Moreover, simple and insightful upper bounds are also obtained.

II. CHANNEL MODEL

Let $H[k]$ denote the $M \times N$ channel gain matrix in an $M$-transmit and $N$-receive antenna system during the $k$-th block, with each block consisting of $L$ symbol intervals. The $(m, n)$-th entry $h_{m,n}[k]$ is the path gain from the $m$-th transmit antenna to the $n$-th receive antenna. The paths are all independent, and each sequence $\{h_{m,n}[k]\}_k$ is modeled as samples of a complex, zero-mean, Gaussian random process having autocorrelation function $2R_{m,n}[l] = E|h_{m,n}[k]|^2 h_{m,n}[k-l]$. In general, $R_{m,n}[l]$ is not necessarily the same for all $1 \leq m \leq M$ and $1 \leq n \leq N$. It takes into account the non-identical distribution of the channel gains between different transmit-receive antenna pairs. The fading correlation coefficient across
two adjacent blocks is given by $\rho_{m,n} = R_{m,n}[1]/R_{m,n}[0]$, and is a measure of the fluctuation rate of the channel fading process. In practice, since the antennas are usually fixed upon installation, the Doppler frequency shift experienced by the communication links due to relative movement of the sender and receiver would be the same for all transmit-receive antenna pairs. As a result, $\rho_{m,n}$ will be same for all $m$'s and $n$'s. Nevertheless, we keep the subscripts $m$ and $n$ for the sake of generality.

Here we consider only the semi-identical channels with $R_{m,n}[l] = R_{n}[l]$ and $\rho_{m,n} = \rho$, for all $m$. That is, the path gains associated with a common receive antenna and different transmit antennas are identically distributed. This channel model is general in the sense that non-identical channels can also be treated as the semi-identical using certain manipulations. Specifically, we assume $\rho_{m,n} = \rho$, for all $m$ and $n$, and define an $MN \times MN$ diagonal matrix $R$ whose $[(n-1)N+m]$-th diagonal entry is $2R_{m,n}[0]$, for $1 \leq m \leq M$ and $1 \leq n \leq N$. The matrix $R$ is essentially the covariance matrix of the $MN \times 1$ vector, vec($H[k]$), formed by stacking the columns of $H[k]$ under each other. If $R$ can be decoupled as $R = R_T \otimes R_R$, where $\otimes$ stands for the Kronecker product, and $R_T$ and $R_R$ are $M \times M$ and $N \times N$ diagonal matrices, respectively, then the statistical properties of $H[k]$ are identical to those of the matrix $R_T^{1/2}H[k]R_T^{1/2}$. We denote the statistical equivalence by

$$H[k] \sim R_T^{1/2}H[k]R_T^{1/2}. \quad (1)$$

Here, $R_T^{1/2}(R_T^{1/2})^H = R_T$, $R_T^{1/2}(R_T^{1/2})^H = R_R$, and the $M \times N$ matrix $H[k]$ contains independent and identically distributed (i.i.d) entries, each of which is modeled as samples of a complex, zero-mean, Gaussian random process having autocorrelation function $2R[l]$, with $2R[0] = 1$ and $2R[1] = \rho$. Therefore, a power allocation matrix $P = (R_T^{1/2})^{-1}$ can be applied at the transmitter such that the effective channel gain matrix $PH[k]$ becomes semi-identical. All the results obtained in this paper can then be used directly in the more general non-identical channels.

The channel-baseband channel input-output relationship can be modeled as

$$Y[k] = \sqrt{E_x}X[k]H[k] + W[k],$$

where $X[k]$, $Y[k]$, and $W[k]$ are the $L \times M$ transmitted signal matrix, $L \times N$ received signal matrix and $L \times N$ noise matrix, respectively, during the $k$-th block. $X[k]$ satisfies the average power constraint: $\mathbb{E} \{ ||X[k]||^2 \} = L$. The entries of $W[k]$ are i.i.d, each with mean zero and variance $N_0/2$ per dimension. $E_x$ is the average total energy transmitted on all antennas.

### III. DIFFERENTIAL OSTBC AND OPTIMAL DIFFERENTIAL DETECTOR

In the differential transmission of square OSTBC codeword, the block length equals the number of transmit antennas, i.e., $L = M$. Let $\{s_p\}_{p=1}^P$ denote a set of $P$ information symbols to be transmitted in the $k$-th block. They are complex scalars chosen from PSK (phase shift keying) constellations with $s_p = e^{j\theta_p}$ and encoded in an $M \times M$ OSTBC codeword as

$$D[k] = \frac{1}{\sqrt{P}} \sum_{p=1}^P \left( \Phi_p \cos \theta_p + j \Psi_p \sin \theta_p \right), \quad (2)$$

where the $M \times M$ encoding matrices $\Phi_p$ and $\Psi_p$ are linked to the theory of amicable orthogonal designs [12]–[14]. The data matrix $D[k]$ is differentially encoded in the transmitted signal matrix $X[k]$ as $X[k] = D[k]X[k-1]$. Since $D[k]$ is unitary, $X[k]$ is also unitary.

We consider the maximum likelihood (ML) detection of $D[k]$ based on two consecutive received signal blocks $Y[k]$ and $Y[k-1]$. Let $Y = [Y^T[k-1], \ Y^T[k]]^T$ and $y_n = [y_n^T[k-1], \ y_n^T[k]]^T$, where $y_n[k]$ is the $n$-th column of $Y[k]$, for $n = 1, \ldots, N$. Conditioned on $D[k]$ and $X[k-1]$, the $2M \times 1$ vectors $y_n$’s can be shown to be independent for different $n$, and each has a complex Gaussian distribution with mean zero and covariance matrix

$$\Lambda_n = E_x \left[ 2R_n[0]I_M 2R_n[1]D^H \right] 2R_n[0]I_M + N_0 I_{2M}. \quad (3)$$

In (3), the block index in $D[k]$ is omitted for notational brevity. The conditional probability density function (PDF) of $Y$ can thus be written as

$$p(Y|D) = \frac{1}{2^{2MN}} \prod_{n=1}^N \det(\Lambda_n)^{-1} \exp \left[ -\sum_{n=1}^N y_n^H(\Lambda_n^{-1})y_n \right]. \quad (4)$$

Using certain matrix formulas, we can obtain

$$\Lambda_n^{-1} = \frac{1}{N_0} \left[ I - \frac{\gamma_n}{1 + \gamma_n} \right] \left[ 1 + \gamma_n(1 - \rho_n^2) \right] \left[ 1 + \gamma_n(1 - \rho_n^2) \right]^{-1} \left[ \frac{\rho_nD^H}{\rho_nD} \right] \left[ \frac{\rho_nD}{\rho_nD} \right]. \quad (5)$$

and

$$\det(\Lambda_n) = N_0^M \left[ 1 + \gamma_n^2 - (\rho_n\gamma_n)^2 \right]. \quad (6)$$

In both (5) and (6), $\gamma_n = 2E_xR_n[0]/N_0$ is the average received signal-to-noise ratio (SNR) per symbol interval on the $n$-th receive antenna. Taking the natural logarithm of (4) and ignoring the terms that are independent of $D$, we obtain the ML decision rule for $D$, after substantial simplification, as

$$D = \arg \max_D \left\{ \sum_{n=1}^N w_n y_n^H[k-1]D^H y_n[k] \right\},$$

where $\{\cdot\}$ represents the real part, and the weighting coefficient $w_n$ is given by

$$w_n = \frac{\rho_n\gamma_n}{1 + \gamma_n^2 - (\rho_n\gamma_n)^2}. \quad (7)$$

Here we assume that the channel second-order statistics and the total transmit SNR are known at the receiver. Now, using the inherent linear structure of the data matrix by design in (2), the above matrix-by-matrix detector reduces to a symbol-by-symbol detector

$$\hat{s}_p = \arg \max_{s_p} \left\{ \sum_{n=1}^N w_n z_{n,p}s_p \right\}, \quad p = 1, \ldots, P \quad (8)$$

where

$$z_{n,p} = \mathbb{R}\{y_n^H[k-1]\Phi_p^H y_n[k]\} - j\mathbb{R}\{y_n^H[k-1]\Psi_p^H y_n[k]\},$$

is the contribution to the decision variable from the $n$-th receive antenna. The contributions from different receive antennas are weighted by the $w_n$’s. Note that the optimal weights
in (7) for the semi-identical MIMO channels considered here coincide with the result in [10] for SIMO channels.

When the channel second-order statistics and the total transmit SNR are unknown at the receiver, one would let the contributions from different receive antennas be weighted equally in the detector. Thus, by letting $w_n = 1$ for all $n$, (8) reduces to

$$s_p = \arg \max_{p} \left\{ \Re \left( \sum_{n=1}^{N} z_{n,p} \phi_{p} \right) \right\}, \quad p = 1, \ldots, P,$$

which agrees with the decision rule in [6], [8] for i.i.d channels. Hence, the traditional differential detector is suboptimal in the i.s.i.d case.

IV. EXACT BEP AND UPPER BOUNDS

The analysis of the BEP for detecting symbol $s_p$ involves the computation of the following probability [15]

$$F(\alpha|\theta_p) = P\left( Z_p(\alpha) < 0 | s_p = e^{j\theta_p} \right),$$

where

$$Z_p(\alpha) = \Re \left( \sum_{n=1}^{N} w_n z_{n,p} e^{j\alpha} \right)$$

is the decision phasor for $s_p$, and $\alpha$ is some angle.

Given $Y[k-1]$, the decision phasor $Z_p(\alpha)$ in (9) is shown to be conditionally Gaussian distributed with mean

$$m_Z = \frac{\cos(\alpha + \theta_p)}{\sqrt{P}} \sum_{n=1}^{N} \rho_n \gamma_n w_n [y_n[k-1]]^2$$

and variance

$$\sigma_Z^2 = \frac{N_0}{2} \sum_{n=1}^{N} w_n^2 \left[ \frac{(1 + \gamma_n)^2 - (\rho_n \gamma_n)^2}{1 + \gamma_n} \right] [y_n[k-1]]^2,$$

where $\| \cdot \|^2$ stands for the squared Frobenius norm. The proof is outlined in the appendix. By substituting the optimal weights (7) into (10) and (11), the conditional probability of $Z_p(\alpha) < 0$ can be expressed using $Q$-function as

$$P\left( Z_p(\alpha) < 0 | D, Y[k-1] \right) = Q \left( \sqrt{c} \sum_{n=1}^{N} A_n \| y_n[k-1] \|^2 \right),$$

where

$$c = \frac{2 \cos^2(\alpha + \theta_p)}{N_0 P},$$

and

$$A_n = \frac{(\rho_n \gamma_n)^2}{(1 + \gamma_n)(1 + \gamma_n)^2 - (\rho_n \gamma_n)^2}.$$

We define $g_n = c A_n \| y_n[k-1] \|^2$. It is easy to see that each $g_n$ is distributed according to a chi-square distribution with $2M$ degrees of freedom. Averaging (12) over the distribution of $\{g_n\}$ yields $F(\alpha|\theta_p)$.

Using the alternative representation of $Q$-function [16]

$$Q(x) = \frac{1}{\pi} \int_{0}^{\pi/2} \exp\left( -\frac{x^2}{2 \sin^2 \theta} \right) d\theta, \quad x \geq 0,$$

we can evaluate $F(\alpha|\theta_p)$ as

$$F(\alpha|\theta_p) = E \left[ \frac{1}{\pi} \int_{0}^{\pi/2} \exp\left( -\frac{1}{2 \sin^2 \theta} \sum_{n=1}^{N} g_n \right) d\theta \right]$$

$$= \frac{1}{\pi} \int_{0}^{\pi/2} \prod_{n=1}^{N} \left[ \exp\left( -\frac{g_n}{2 \sin^2 \theta} \right) \right] d\theta$$

$$= \frac{1}{\pi} \int_{0}^{\pi/2} \prod_{n=1}^{N} \left( 1 + \frac{B_n}{2 \sin^2 \theta} \right)^{-M} d\theta,$$

where

$$B_n = \frac{2 \cos^2(\alpha + \theta_p)(\rho_n \gamma_n)^2}{P[(1 + \gamma_n)^2 - (\rho_n \gamma_n)^2]}.$$

Eqn. (13) involves only a single integral over finite limits and hence can be numerically computed easily.

Now we can compute the BEP. For BPSK and QPSK (with Gray mapping) constellations, the exact BEP is given using the half-plane decision region approach [15], respectively, by

$$P_b, \text{BPSK} = F(\alpha = 0 | \theta_p = 0)$$

and

$$P_b, \text{QPSK} = F(\alpha = -\pi/4 | \theta_p = 0).$$

An upper bound on the probability $F(\alpha|\theta_p)$ can be obtained by letting $\theta = \pi/2$ in (13). Doing this and using the expression for $B_n$ in (14) lead to

$$F(\alpha|\theta_p) \leq \frac{1}{2} \prod_{n=1}^{N} \left( 1 + \frac{\cos^2(\alpha + \theta_p)(\rho_n \gamma_n)^2}{P[(1 + \gamma_n)^2 - (\rho_n \gamma_n)^2]} \right)^{-M},$$

where $\gamma_n$ is the geometrical mean of $\{\gamma_n\}$, given by $\gamma_{\text{gem}} = \left( \prod_{n=1}^{N} \gamma_n \right)^{1/N}$. Recall that, for a set of real positive values, the geometrical mean is always upper bounded by the arithmetic mean. Hence, when the total received SNR over all receive antennas, defined as $\sum_{n=1}^{N} \gamma_n$, is fixed, the semi-identical channel will degrade the BEP performance compared with the identical channel. Next we consider a time-varying channel with $\rho_n = \rho$ for all $n$. By letting $\gamma_n \to \infty$, the right-hand side of the inequality (15) approaches a constant:

$$F(\alpha|\theta_p) \to \frac{1}{2} \left[ \frac{\rho \cos^2(\alpha + \theta_p)}{(1 - \rho^2) P} \right]^{-MN}.$$

This indicates that the asymptotic irreducible error floor of differential OSTBC with optimal detection is regardless of the unequal average received power on different receive antennas. In other words, the asymptotic irreducible error floors of the semi-identical and identical channels are the same.
V. NUMERICAL RESULTS AND DISCUSSIONS

In this section, we use the differential OSTBC with $M = P = 2$ using QPSK constellation. The total average SNR per bit over $N$ receive antennas is defined as $\gamma_b = \sum_{n=1}^{N} \gamma_n / b$, where $b$ is the transmission rate in bits per channel use and equals 2 in this example. The block correlation coefficients are assumed to be identical, i.e., $\rho_n = \rho$, for all $n$.

In Figs. 1 and 2, we plot the analytical BEP results for $N = 2$ receive antennas. The quantity $\eta$ denotes the fraction of the total average received signal energy from receive antenna 1 (and hence $1 - \eta$ is from receive antenna 2). For instance, we have $\gamma_1 : \gamma_2 = 10\% : 90\%$ when $\eta = 0.1$ and $\gamma_1 : \gamma_2 = 49\% : 51\%$ when $\eta = 0.49$. The results for $N = 3$ receive antennas are plotted in Fig. 3, where the average received signal energy distribution among the three receive antennas is set to $\gamma_1 : \gamma_2 : \gamma_3 = 10\% : 30\% : 60\%$. The exact BEPs for the cases of i.i.d channels from [8] and i.s.i.d channels with suboptimal differential detector from [17] are plotted for comparison.

From these figures several observations can be made. First, the non-identical channel gain distribution among the receive antennas degrades the BEP performance. For instance, when the total average SNR per bit $\gamma_b$ is 20 dB and the block correlation coefficient $\rho$ is 0.99, the BEP is equal to $5.63 \times 10^{-5}$ for the i.i.d case ($\gamma_1 = \gamma_2 = 20$ dB), and $2.12 \times 10^{-4}$ for optimal detection with $\eta = 0.1$ ($\gamma_1 = 13$ dB and $\gamma_2 = 22.55$ dB). Second, the irreducible error floor of i.s.i.d channels as $\gamma_b \to \infty$ when the optimal detector is used approaches that of i.i.d channels. These two observations agree with the asymptotic behavior of the upper bound we obtained in Section IV. Third, when the block correlation coefficient $\rho$ is not too close to unity, the optimal detector can substantially improve the BEP performance, compared with the suboptimal detector, especially in the regime of high SNR. For example, for $\gamma_b = 30$ dB, $\rho = 0.99$ and $\gamma_1 : \gamma_2 : \gamma_3 = 10\% : 30\% : 60\%$ ($\gamma_1 = 23$ dB, $\gamma_2 = 27.78$ dB, and $\gamma_3 = 30.79$ dB), the BEP is $2.64 \times 10^{-7}$ for the suboptimal detector and is $6.852 \times 10^{-8}$ for the optimal detector, as shown in Fig. 3. Lastly, when $\rho$ approaches unity the optimal and suboptimal detectors are almost equivalent, regardless of the unequal received power among the receive antennas.

VI. CONCLUSIONS

This paper presents the analysis of bit error probabilities for differential orthogonal space-time block codes over semi-identical MIMO Rayleigh fading channels. We take into account both unequal received SNR distribution among the receive antennas and arbitrary fluctuation rates of the block channel fading processes. We show that the optimal symbol-by-symbol differential detector involves weighting the output from each receive antenna according to its channel statistical information. The exact BEP expressions for any number of transmit antennas for which OSTBCs exist are derived, which are well suited for numerical evaluation. Our results reveal that the non-identical channel statistics degrade the error performance compared with the identical case. We also find that...
the proposed optimal detector outperforms the conventional suboptimal detector in time-varying fading channels but they perform similarly in near-static fading channels. Though they are derived for semi-identical MIMO channels, our results can be directly applied to non-identical MIMO channels with proper channel processing.

APPENDIX

We rewrite $Z_p(\alpha)$ in (9) as

$$Z_p(\alpha) = \frac{1}{2} \sum_{n=1}^{N} w_n y_n^H[k - 1] Q y_n[k]$$

$$+ \frac{1}{2} \sum_{n=1}^{N} w_n y_n^H[k] Q^H y_n[k - 1],$$

where $Q = \cos \alpha \Phi^H_H + j \sin \alpha \Psi^H_H$. It follows from (3) that the joint PDF of $y_n[k]$ and $y_n[k - 1]$ is complex Gaussian with mean zero and covariance matrix:

$$A_n = \mathbb{E} \left\{ \begin{bmatrix} y_n[k] \\ y_n[k - 1] \end{bmatrix} \begin{bmatrix} y_n^H[k] \\ y_n^H[k - 1] \end{bmatrix} \right\}$$

$$= A_{n,11} A_{n,12} A_{n,21} A_{n,22},$$

where $A_{n,11} = A_{n,22} = N_0(1 + \gamma_n) I_M$, and $A_{n,12} = A^H_{n,21} = N_0 \rho_n \gamma_n D$. Therefore, the conditional distribution $p(y_n[k]|y_n[k - 1])$ is also complex Gaussian with mean

$$u_n = A_{n,12} A_{n,22}^{-1} y_n[k - 1]$$

and covariance matrix

$$\Sigma_n = A_{n,11} - A_{n,12} A_{n,22}^{-1} A_{n,21}. $$

After manipulation, the mean vector and covariance matrix can be simplified as

$$u_n = \frac{\rho_n \gamma_n}{1 + \gamma_n} D y_n[k - 1]$$

and

$$\Sigma_n = N_0 \left( \frac{(1 + \gamma_n)^2}{1 + \gamma_n} \right) I_M,$$

respectively. Thus, $Z_p(\alpha)$ is also conditionally Gaussian distributed. Using (16), we obtain the conditional mean of $Z_p(\alpha)$ as

$$m_Z = \frac{1}{2} \sum_{n=1}^{N} \rho_n \gamma_n w_n y_n^H[k - 1] \left( QD + D^H Q^H \right) y_n[k - 1].$$

From the orthogonality property of the encoding matrices in (2), we can have

$$QD + D^H Q^H = \frac{2}{\sqrt{P}} \Phi(\alpha + \theta_p) I_M.$$

Thus, $m_Z$ reduces to (10). The conditional variance of $Z_p(\alpha)$ is computed as

$$\mathcal{E} \left\{ [Z_p(\alpha) - m_Z]^2 \right\} = \frac{1}{4} \mathcal{E} \left\{ |a_1|^2 + |a_2|^2 + a_1^* a_2 + a_1 a_2^* \right\}$$

where

$$a_1 = a_2^* = \sum_{n=1}^{N} w_n y_n^H[k - 1] Q \left( y_n[k] - u_n \right).$$

The correlation between $a_1$ and $a_2$ is

$$\mathcal{E} [a_1 a_2^*] = \sum_{n=1}^{N} \sum_{\nu=1}^{N} w_n w_\nu y_n^H[k - 1] Q \cdot Q^T y_\nu^* [k - 1].$$

Since the real and imaginary parts of $y_n[k]$ are uncorrelated and both have the same covariance matrix, it follows that $\mathcal{E} [(y_n[k] - u_n)(y_\nu^*[k] - u_\nu)^T] = 0$. Hence $\mathcal{E} [a_1 a_2^*] = 0$, $\mathcal{E} [|a_1|^2]$ and $\mathcal{E} [|a_2|^2]$ can be obtained using (17). To summarize, the conditional variance of $Z_p(\alpha)$ is computed as (11).

REFERENCES


