Double-Differential Coding for Orthogonal Space-Time Block Codes

Manav R. Bhatnagar, Are Hjørungnes, and Lingyang Song
UniK – University Graduate Center, University of Oslo
Instituttveien 25, P. O. Box 70, NO-0277 Kjeller, Norway
Email: {manav,arehj,lingyang}@unik.no

Abstract—Communications over multiple-input multiple-output (MIMO) channels with carrier offsets is an important practical and theoretical problem. Double-differential coding is a technique, which allows the receiver to decode the data without any channel or carrier offset knowledge. We propose a double-differential (DD) coding scheme which is applicable to any square orthogonal space-time block codes (OSTBC) using $M$-PSK constellation. The main advantages of the proposed DD coding scheme are: 1) The previously proposed DD codes are applicable only to the specific class of space-time block codes which follow the diagonal unitary group property, whereas our DD coding is applicable to any square OSTBC. 2) We propose a suboptimal decoder which preserves the linear decoding property of the OSTBC. We derive an upper bound of the pairwise error probability (PEP) of the proposed double-differential orthogonal space-time block codes (DDOSTBCs). The proposed DDOSTBC is able to achieve better performance than the similar rate existing DD coding scheme. In addition, the proposed DDOSTBC outperforms the conventional training based system.

Keywords: Double-Differential Modulation, Orthogonal Space-Time Block Codes, MIMO System, Carrier Offset, Differential Coding.

I. INTRODUCTION

Space-Time block codes (STBC) are one of the main coding techniques for multiple-input multiple-output (MIMO) systems, which enables them to exploit full diversity without any channel knowledge at the transmitter. The decoding complexity can be reduced by using OSTBC. Generally, it is assumed that the receiver can acquire the knowledge about the channel by using training data. However, this training based technique is only useful for channels which remain constant for several symbol durations. The receiver may use training data for the estimation of the channel coefficients at the cost of reduced data rate. However, training-based methods do not work well when the channel remains constant for relatively small number of channel usages. Differential coding is proposed in the literature [1], [2], [3], [4] to avoid the need of channel estimation at the receiver end.

Another challenge is the presence of the carrier offsets, which exists because of the movement of the receiver, transmitter, or scatterers, and the mismatch between the transmit and receive oscillators. The differential systems in [1], [2], [3], [4] fail to perform well in the presence of carrier offsets, because the carrier offsets make the flat fading channel behave as a time-varying channel. Hence, the channel does not remain constant over two consecutive STBC block transmission time-intervals, which is a basic assumption for differential systems.

Accurate estimation of the carrier offsets is difficult and needs intensive training data causing large delays in the decision process. Small residual carrier offset error can degrade the performance of the receiver substantially [5]. DD coding [6], [7] is a key technique which could be used to avoid the need of carrier offset and channel estimation. DD coding reflects its utility specially for channels which remain constant over a small number of block (symbol) durations and are perturbed with carrier offsets. A trained decoder [8] will reduce the data rate. DD coding enables the receiver to make decisions based on the three consecutively received data blocks/matrices without any carrier offset or channel knowledge. Hence, it is efficient in terms of data rate. DD coding for MIMO system was proposed in [9], however, it can only be applied to a specific class of OSTBC, which belong to the diagonal unitary group.

In this paper, our main contributions are as follows: 1) We propose a double-differential coding, which is applicable to any square OSTBC with $M$-PSK constellations. 2) A low complexity linear decoder of the DDOSTBC is obtained. 3) An upper bound for the PEP of the DDOSTBC is also derived.

The rest of the paper is organized as follows: In Section II DD encoding for OSTBC is explained while in Section III the channel model is discussed. Decoding of DDOSTBC is presented in Section IV, Section V performs the theoretical analysis of the performance of the DDOSTBC scheme. Simulation results and comparisons are discussed in Section VI. Some conclusions are drawn in Section VII. The article contains one appendix.

II. DOUBLE-DIFFERENTIAL ENCODING

Let $S_k, k \geq 2$, be an $n_t \times n_t$ OSTBC matrix obtained from $s_k = [s_1, s_2, \ldots, s_{n_t}]^T$, $n_s \leq n_t$, $s_i$ is an $M$-PSK symbol. The first order differential matrix $F_k$ of size $n_t \times n_t$ can be obtained as

$$F_k = F_{k-1}S_k, \quad k \geq 2,$$

with $F_1 \in \mathbb{C}^{n_t \times n_t}$ as the initialization matrix. Next, the second order differential matrix $D_k$ can be obtained from $F_k$ as

$$D_k = D_{k-1}F_k, \quad k \geq 1,$$

with $D_0 \in \mathbb{C}^{n_t \times n_t}$ as initialization matrix. For $p \geq 2$, the second order differential matrix $D_p$ is obtained from the current OSTBC matrix $S_p$ from:

$$D_p = D_{p-2}F_{p-1}^2S_p = D_0F_1\cdots F_{p-2}^2F_{p-1}S_p. \quad (3)$$
**Lemma 1:** The double-differential encoding of (3) can be used if $S_p$ is a full-rank square matrix.

**Proof:** It can be seen from [10, Eqs. (1.7.4) and (1.7.5)] that the product of (3) follows a diminishing rank property and may results into an all zero matrix, if $S_p$, for all $p$, is not a full-rank and square matrix. It can be further observed from [10, Eq. (1.7.5)] that the rank of $D_p$ is equal to the rank of $D_0$ if $S_p$, for all $p$, is a full-rank square matrix.

Lemma 1 indicates an important result about the choice of initialization matrices $D_0$ and $F_1$. It is apparent that from the full diversity point of view [8] $D_0$ and $F_1$ must be full-rank matrices. Lemma 1 also indirectly puts an restriction on the signal alphabet which can be used for DD encoding since the signal alphabet cannot contain the origin.

It can be seen from (1), (2), (3), and Lemma 1 that the proposed DD encoding is applicable to any full-rank $n_t \times n_t$ OSTBC matrix, which can transmit $n_s$ ($n_s \leq n_t$) M-PSK symbols. Whereas, the DD encoding proposed in [9] can only be applied to the diagonal OSTBC, which can transmit only one M-PSK symbol by an $n_t \times n_t$ diagonal matrix. Hence, the proposed DD encoding is able to provide better data rate than [9].

**III. Channel Model**

Consider a MIMO channel with $n_t$ transmit and $n_r$ receive antennas. Let $h_{m,n}$ be the channel gain between $m$-th receive and $n$-th transmit antenna and $D_k$ be the $k$-th transmitted $n_t \times n_r$ DD encoded matrix. Let $\omega_m \in [-\pi, \pi]$ be the random carrier offset between all transmit antennas and the $m$-th receive antenna, which is assumed to be independent of time $k$ and constant over the transmission period of at least $D$ DD encoded matrices $D_k$. The received data $y_{m,k} \in \mathbb{C}^{1 \times n_t}$ at the $m$-th receiver antenna corresponding to $D_k$, $k \geq 0$, is [8, Eq. (9.7.1)]

$$y_{m,k} = \exp(\jmath \omega_m n_t k) h_{m} D_k \Omega_m + q_{m,k}, \quad (4)$$

where $\Omega_m \in \mathbb{C}^{n_t \times n_t}$ is the diagonal matrix $\Omega_m = \text{diag} \{ \exp(\jmath \omega_{m1}), \ldots, \exp(\jmath \omega_{m(n_t - 1)}) \}$, $h_{m} = [h_{m,1}, \ldots, h_{m,n_t}]$ is an $1 \times n_t$ row vector consisting of the channel coefficients between the $m$-th transmit antenna and the $m$-th receive antenna, $q_{m,k} \in \mathbb{C}^{1 \times n_t}$ contains zero mean additive white complex-valued Gaussian noise (AWGN), whose elements are i.i.d. Gaussian random variables with zero mean and variance $\sigma^2$. The channel $H = [h_1^T, h_2^T, \ldots, h_{n_r}^T]^T$ is assumed to be constant over the transmission period of at least three DD encoded matrices.

**IV. Decoding of DDOSTBC**

It can be seen from (4) that if $\omega_m$, $h_{m}$, and $D_k$ are known for all $k$ and $m$, $y_{m,k}$ has the following probability density function (p.d.f.):

$$f(y_{m,k} | \omega_m, h_{m}, D_k) = \frac{1}{\pi \sigma^2 \text{det}(\Lambda)} \times \exp \left( - \frac{1}{\sigma^2} \| y_{m,k} - \exp(\jmath \omega_m n_t k) h_{m} D_k \Omega_m \|^2 \right) \Lambda^{-1} \times \left[ y_{m,k} - \exp(\jmath \omega_m n_t k) h_{m} D_k \Omega_m \right]^T \right), \quad (5)$$

where $\Lambda = E \left( q_{m,k}^T q_{m,k}^* \right)$ is the covariance matrix of zero mean noise $q_{m,k}$. It is assumed that $q_{m,k}$ is AWGN with $\Lambda = \sigma^2 I_{n_t}$. Let $y_{m,k} = [y_{m,k-2}, y_{m,k-1}, y_{m,k}] \in \mathbb{C}^{1 \times 3n_t}$ contains three consecutively received data vectors at the receive antenna $m$. When $h_{m}$, $\omega_m$, $D_{k-2}$, $F_{k-1}$, and $S_k$ are known,

$$f(y_{m,k} | \omega_m, h_{m}, D_{k-2}, F_{k-1}, S_k) = \frac{1}{\pi^3 \sigma^6} \times \exp \left( - \frac{1}{\sigma^2} \sum_{l=k-2}^k \| y_{m,l} - \exp(\jmath \omega_m n_t l) h_{m} D_l \Omega_m \|^2 \right). \quad (6)$$

In order to find a maximum likelihood (ML) estimate of the unknown data $S_k$, the joint probability distribution function (p.d.f.) in (6) is first maximized with respect to (w.r.t.) all unknown quantities $h_{m}$, $\omega_m$, $D_{k-2}$, and $F_{k-1}$, and subsequently over $S_k$, which results into minimization of the following metric:

$$\Gamma_{m,k} = \sum_{l=k-2}^k \| y_{m,l} - \exp(\jmath \omega_m n_t l) h_{m} D_l \Omega_m \|^2, k \geq 2. \quad (7)$$

Minimization of (7) w.r.t. $h_{m}$, $\omega_m$, $D_{k-2}$, and $F_{k-1}$, and $S_k$ is very complicated, therefore, we focus on a suboptimal decoder which makes the decision independent of channel and carrier offset knowledge.

**A. Suboptimal Decoder of Double-Differential OSTBC**

**Lemma 2:** Let $X_k$ is a square OSTBC matrix. Then $\hat{X}_k \triangleq \Omega_n^H X_k \Omega_m$ is also a square OSTBC matrix.

**Proof:** It can be verified using [8, Eqs. (7.1.1) and (7.4.4)] that $X_k$ is square OSTBC matrix.

To simplify the decision process, we consider a degenerated decision metric from (7) as follows:

$$D_{m,k} = \sum_{l=k-2}^{k-1} \| y_{m,l} - \exp(\jmath \omega_m n_t l) h_{m} D_l \Omega_m \|^2, k \geq 2. \quad (8)$$

By means of Lemma 2, the unitary property of the OSTBC matrix with M-PSK constellation [8], and the results given in [11] for matrix derivatives, (8) is minimized w.r.t. $g_{m,k} = \exp(\jmath \omega_m n_t (k - 2)) h_{m} D_{k-2} \Omega_m$ as

$$g_{m,k} = \frac{1}{2} \left( E_{\omega_m} y_{m,k-1}^\dagger \tilde{F}_{k-1} + y_{m,k-2} \right), \quad (9)$$

where $E_{\omega_m} = \exp(\jmath \omega_m n_t)$ and $\tilde{F}_{k-1} = \Omega_n^H F_{k-1} \Omega_m$. By substituting (9) into (8) and using the unitary property of OSTBC [8], (8) reduces into

$$D_{m,k} = \| y_{m,k-1} - E_{\omega_m} y_{m,k-2} \tilde{F}_{k-1} \|^2, k \geq 2. \quad (10)$$

It can be seen from (10) that if $\tilde{F}_{k-1}$ is unknown, it is not possible to find the estimate of $E_{\omega_m}$.

**Lemma 3:** If $\tilde{F}_{k-1} = I_{n_t}$, the estimator of $E_{\omega_m}$ is given by

$$\hat{E}_{\omega_m,k-1} = \exp(\jmath \arg\{ y_{m,k-1}^\dagger y_{m,k-2}^\dagger \}). \quad (11)$$
Proof: If \( F_{k-1} = I_n \), then \( \hat{F}_{k-1} = I_n \) and (10) reduces into the following form:

\[
D_{m,k} = \| y_{m,k-1} - \mathcal{E}_{\omega_m} y_{m,k-2} \|^2, \quad k \geq 2. \tag{12}
\]

Equation (12) can be minimized to find an estimate of the carrier offset and results into (11).

However, we cannot take the liberty of assuming \( F_{k-1} = I_n \) as the identity matrix as we are using OSTBC matrices. Nevertheless, we may assume that the initialization matrix \( F_1 \) is equal to the identity matrix. Then, the estimate of \( \mathcal{E}_{\omega_m} \) can be found from (11) at \( k = 2 \), such that \( \mathcal{E}_{\omega_m,1} \) can be used in the place of \( \mathcal{E}_{\omega_m} \) in the further analysis.

Remark: For double-differential modulation, it can be assumed that the receiver has perfect knowledge about the initialization matrix \( F_1 \). The receiver can reconstruct the subsequent \( \hat{F}_{k-1} \), for \( k > 2 \), from the estimated data and (1). However, \( \hat{F}_{k-1} \), for \( k > 2 \), is a noisy version of the original \( F_{k-1} \), for \( k > 2 \). Hence, the estimates of \( \mathcal{E}_{\omega_m} \) at \( k > 2 \), based on \( \hat{F}_{k-1} \), for \( k > 2 \), will be worse than \( \mathcal{E}_{\omega_m,1} \). Hence, \( \mathcal{E}_{\omega_m,1} \) should be used for all \( k \) over which the carrier offset remains constant as the estimate of carrier offset.

From (1), (2), and (4), the received vector at the \( k \)-th time instant and of the \( m \)-th receive antenna can be written in the terms of \( \mathcal{E}_{\omega_m}, g_{m,k} \), and \( F_{k-1} \) as

\[
y_{m,k} = \mathcal{E}_{\omega_m} g_m \mathcal{H} \hat{F}_{k-1}^2 S_k \Omega_m + q_{m,k}. \tag{13}
\]

Hence, we can minimize the following decision metric to find the estimate of \( S_k \), \( k \geq 2 \):

\[
\hat{S}_k = \arg \min_{\hat{S}_k \in \Xi} \| y_{m,k} - \mathcal{E}_{\omega_m,1} g_m \mathcal{H} \hat{F}_{k-1}^2 \hat{S}_k \Omega_m \|^2, \tag{14}
\]

where \( \Xi \) is the set of all OSTBC matrices consisting of symbols from the \( M \)-PSK constellation, \( \Omega_m \) is obtained from \( \mathcal{E}_{\omega_m,1} \), and \( \hat{S}_k = \hat{\Omega}_m^H S_k \Omega_m \). Let \( \hat{h}_m = \mathcal{E}_{\omega_m,1} g_m \mathcal{H} \hat{F}_{k-1} \hat{\Omega}_m \), then (14) can be written as

\[
\hat{S}_k = \arg \min_{\hat{S}_k \in \Xi} \| y_{m,k} - \hat{h}_m \hat{S}_k \|^2. \tag{15}
\]

It can be seen from Lemma 2 that \( \hat{S}_k \) is a square OSTBC matrix. Hence, \( \hat{S}_k \) can be written as \( \hat{S}_k = \sum_{n=1}^{n_t} (s_n A_n + j s_n B_n) \) [8, Eq. (7.1.1)], where \( s_n \) and \( s_n \) are the real and the imaginary part, respectively, of the complex symbol \( s_n \), and \( A_n \) and \( B_n \) are fixed (in general complex-valued) code matrices of size \( n_t \times n_t \), which satisfy the properties given in [8, Eq. (7.4.4)]. Expanding the right hand side of (15) and using the orthogonal property \( \hat{S}_k \hat{S}_k^H = \sum_{n=1}^{n_t} |s_n|^2 I_{n_r} \), it can be shown that the decoding of \( \hat{S}_k \) can be done linearly. Decoding of (14) can be generalized to \( n_r \) receive antennas as

\[
\hat{S}_k = \arg \min_{\hat{S}_k \in \Xi} \sum_{m=1}^{n_r} \| y_{m,k} - \mathcal{E}_{\omega_m,1} \hat{g}_m \mathcal{H} \hat{F}_{k-1}^2 S_k \Omega_m \|^2. \tag{16}
\]

V. PERFORMANCE ANALYSIS OF DDOSTBC

In this section, we are analyzing the pairwise error probability (PEP) of DDOSTBC over flat fading MIMO channels with carrier offset. From (16), the probability of detecting \( S_k \) in place of \( S_0^k \) where \( S_k \neq S_0^k \) can be written as

\[
\Pr \{ S_0^k \to S_k \} = \Pr \left\{ \sum_{m=1}^{n_r} \| y_{m,k} - \mathcal{E}_{\omega_m,1} \hat{g}_m \mathcal{H} \hat{F}_{k-1}^2 S_k \Omega_m \|^2 < \sum_{m=1}^{n_r} \| y_{m,k} - \mathcal{E}_{\omega_m,1} \hat{g}_m \mathcal{H} \hat{F}_{k-1}^2 S_0^k \Omega_m \|^2 \right\}, \tag{17}
\]

where \( y_{m,k} = \exp (j \omega_m n_t k) h_m \mathcal{D}_{k-2} \mathcal{F}_{k-1}^2 \Omega_m + q_{m,k} \) and \( \hat{F}_{k-1} \) is the estimate of \( F_{k-1} \). After many manipulations and using [8, Theorem 4.2], the Chernoff bound [12, Eq. (2.172)], and Fischer’s matrix inequality [10, Eq. (11.8.1)], it can be shown to be

\[
\Pr \{ S_0^k \to S_k \} \leq \exp \left( - \frac{\sum_{m=1}^{n_r} \left( \lambda_{\min}( \lambda_{\mu}(s_{m,k}) \mathcal{S}_m ) \right)^2}{6 \sigma^2 \sum_{m=1}^{n_r} \| v_{m,k} S_m \|^2} \right), \tag{18}
\]

where \( \bar{S} = S - S_k \), \( W_{m,k} = \mathcal{E}_{\omega_m} \mathcal{H} \hat{F}_{k-1}^2 \),

\[
v_{m,k} = \frac{\hat{E}_{\omega_m,1}}{2} |\mathcal{E}_{\omega_m} \hat{g}_m \mathcal{H} \hat{F}_{k-1}^2 S_k \Omega_m |
\]

\[
\mathcal{X}_{m,k} = \mathcal{H} \left( S_0^k \right)^\mathcal{H} W_{m,k} \left( V_{m,k} \right)^{-1} \hat{\Omega}_m + \hat{\Omega}_m V_{m,k} \mathcal{H} S_0^k \Omega_m,
\]

\[
V_{m,k} = \frac{\hat{E}_{\omega_m,1}}{2} |\mathcal{E}_{\omega_m} \hat{g}_m \mathcal{H} \hat{F}_{k-1}^2 S_0^k \Omega_m |
\]

\[
\lambda_{\min}( \mathcal{X}_{m,k} ) \text{ is the minimum eigenvalue of } \mathcal{X}_{m,k}.
\]

Theorem 1: The average pairwise error probability of DDOSTBC can be bounded as

\[
\mathcal{E}_{\mathcal{H}} \left[ \Pr \{ S_0^k \to S_k \} \right] \leq \left( S_0^k - S_k \right) \Pi_k ( S_0^k - S_k )^{\mathcal{H}} - n_r \times \left( \frac{\gamma^2}{12 \sigma^2} \right)^{-n_r}, \tag{20}
\]

where \( \Pi_k = \lambda_{\min}( \mathcal{X}_k ) I_{n_t} \).

Proof: See Appendix I for the proof of Theorem 1.

From (20), it follows that the proposed DDOSTBC achieves full diversity \( (n_t n_r) \).

VI. SIMULATION RESULTS

The simulations are performed with \( n_t = 2, n_r = 1 \), uncorrelated complex Gaussian channel and unit variance complex AWGN noise. The channel coefficients are assumed to be constant over five transmitted blocks of DDOSTBC. The total power transmitted by all antennas in one time-interval is kept
unity, $D_0 = I_2$ and $F_1 = I_2$, and it is assumed that receiver perfectly knows $F_1$. The simulation results are obtained from $10^5$ channel realizations. The proposed DDOSTBC are applied over the Alamouti STBC [13] with QPSK constellation. The carrier offset is assumed to be uniformly distributed over $[-\pi, \pi]$. The DD code of [9] utilizes diagonal OSTBC matrices, which can transmit only one $M$-level symbol per OSTBC block, hence, the DDOSTBC of [9] provides low data rates. Therefore, for a fair comparison at the same spectral efficiency, the simulations are performed for (16:1.7) DDSTBC of [9] which transmits one 16-PSK symbol by a $2 \times 2$ diagonal matrix [9, Eq. (45)]

$$\text{diag}\left\{e^{i2\pi/16}, (e^{i2\pi/16})^7\right\},$$

where $i = 0, 1, \ldots, 15$. It can be seen from Fig. 1 that our proposed DDOSTBC outperforms the existing DDOSTBC [9] by a substantial margin, for example at $\text{SER}=10^{-3}$, the gain is approximately 4 dB. We have also plotted the performance of a trained decoder [8] using QPSK constellation, which utilizes the training sequence transmitted in the starting two blocks for the estimation of carrier offset and channel gains. These estimates are used to recover the unknown data in the subsequent time intervals. It can be seen from Fig. 1 that the proposed DDOSTBC also outperforms the trained receiver [8].

**VII. CONCLUSIONS**

We have proposed double-differential coding for square orthogonal space-time block codes. The proposed double-differential scheme is able to decode the space-time data without knowing the carrier offsets or channel coefficients. The proposed double-differential code exhibits performance gain as compared to the previously proposed double-differential space-time block codes. In addition, it outperforms the previously proposed trained decoder.

**APPENDIX I**

**Proof of Theorem 1**

If all links are perturbed by the same carrier offset, i.e., $\omega_m = \omega$, then (18) reduces to:

$$\text{Pr}\{S_k^0 \rightarrow S_k\} \leq \exp\left(-\frac{\lambda^2_{\min}(M)}{6\sigma^2} \sum_{m=1}^{n_d} ||v_{m,k}^* \omega||^2\right), \quad (21)$$

where $X^k = X_{m,k} | \omega_m = \omega$ and $v_{m,k} = \pi_{m,k, \omega_m = \omega}$. We can average (21) over all Rayleigh uncorrelated channels that channel average PEP upper bound as

$$E_{\text{H}}\left[\text{Pr}\{S_k^0 \rightarrow S_k\}\right] \leq \prod_{m=1}^{n_r} \left\{ I_{n_1} + (\Theta_{S_k^0} - \Theta_{S_k})_\Sigma K \left(\Theta_{S_k^0} - \Theta_{S_k}\right)^{\Sigma K}^{-1} \right\}, \quad (22)$$

where $\Sigma K = \frac{\sum_{m=1}^{n_r} (X^k)_{m,k}^* I_{n_1}, E\left[h^m_s h^m_s\right]}{\gamma^2 I_{n_1}, \Theta S_k = \Phi K S_k^0, \Theta S_k = \Phi K S_k}$, and

$$\Phi K = \frac{\hat{\omega}_{\text{min}}^2}{2} \left[\hat{E}_{\omega}^2 \hat{E}_{\omega^2} \exp\left(\omega n_k (k - 2)\right) D_{k-2} F_{k-1}^0 \Omega^2 \hat{F}_{k-1}^0 \right] \hat{F}_{k-1}^0.$$

Let us assume that all the links between the transmit and receive antennas are perturbed with the worst carrier offset,