Double-Differential Orthogonal Space-Time Block Codes for Arbitrarily Correlated Rayleigh Channels with Carrier Offsets

Manav R. Bhatnagar, Member, IEEE, Are Hjørungnes, Senior Member, IEEE, and Lingyang Song, Member, IEEE

Abstract—The presence of carrier offsets in the multiple-input multiple-output (MIMO) channels 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 square orthogonal space-time block codes (OSTBC) using $M$-PSK constellation. The main advantages of our 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. 3) A theoretical analysis is performed to find a pairwise error probability (PEP) upper bound of the proposed double-differential orthogonal space-time block codes (DDOSTBC). 4) In order to improve the performance of DDOSTBC over the arbitrarily correlated Rayleigh channels we propose a precoder which minimizes an upper bound of the PEP. The proposed DDOSTBC are able to achieve higher coding gain than the similar rate existing DD coding scheme. In addition, the proposed precoded DDOSTBC achieves performance gain for correlated channels as compared to the unprecoded DDOSTBC.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) communications have a substantial potential in terms of increasing the data rate, capacity, and diversity of wireless communication systems. Space-time block codes (STBC) are one of the main coding techniques for MIMO systems, which enables them to exploit full diversity without any channel knowledge at the transmitter. The decoding complexity can be further reduced by using orthogonal space-time block codes (OSTBC). It is often assumed that the receiver can acquire the knowledge about the channel by using training data. However, this technique is only useful for channels which remain constant for many symbol durations. The receiver may use training data for the estimation of the channel coefficients at the cost of reduced effective data rate. But training-based methods do not work well when the channel remains constant for a relatively small number of channel usages. Differential coding is proposed in the literature [1]–[4] to avoid the need of channel estimation at the receiver end.

Another challenge is the presence of the carrier offsets, which result because of the movement of the receiver or transmitter and/or scatterers, and mismatch between the transmit and receive oscillators. The differential systems in [1]–[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. Perfect estimation of the carrier offsets is difficult and needs intensive training data, fast Fourier transformation (FFT), and maxima searching procedures causing large delays in the decision process. Small residual carrier offset errors can degrade the performance of the receiver substantially [5]. Double-differential (DD) coding [6], [7] is a key technique which could be used to avoid the need of carrier offset and channel estimation from training data. It reflects its utility specially for those channels which remain constant for a small number of symbol durations and are perturbed with carrier offsets. We will name these channels rapidly varying channels with carrier offsets (RVCCO). We can apply a training-based decoder [8] which utilizes training data for estimation of carrier offsets and channel for RVCCO. Obviously, this type of decoder will reduce the effective data rate. Blind or semi-blind decoders can hardly provide any performance gain for RVCCO as they generally need the channel to remain constant for a significantly large number of symbol durations (much larger than three STBC blocks) before the decision is made. Iterative decoding also cannot improve the performance in RVCCO because of fast variations of the channel [9]. DD coding enables the receiver to make decisions based on three consecutively received data blocks/matrices without any carrier offset or channel knowledge. Hence, DD coding is efficient in terms of data rate and suitable for RVCCO. DD coding for MIMO system was proposed in [10], however, it can only be applied to a specific class of STBC, which belong to the diagonal unitary group.

Our main contributions are as follows: 1) We propose a double-differential coding, which is applicable to square OSTBC with $M$-PSK constellations. 2) A low complexity suboptimal linear decoder of double-differential orthogonal space-time block codes (DDOSTBC) is obtained based on maximum-likelihood (ML) metric. 3) We derive an upper bound of the PEP of the proposed DDOSTBC for arbitrarily correlated Rayleigh channel. 4) Based on this PEP bound, we
propose a precoder designed to improve the performance of DDOSTBC over correlated MIMO channels.

The rest of the paper is organized as follows: In Section II, DD encoding for OSTBC is explained while the channel model is discussed in Section III. Decoding of DDOSTBC is presented in Section IV. Section V performs the theoretical performance analysis of the DDOSTBC scheme. A precoder is designed in Section VI. Simulation results and comparisons are discussed in Section VII. Some conclusions are drawn in Section VIII. At the end of article, nine appendices give the details of the derivations.

Notation: \( X \) denotes matrix, \( x \) represents a vector, \( x, X \) are used for variables, \((\cdot)^H\) provides the Hermitian of a matrix, complex conjugation of a matrix is represented by \((\cdot)^*\), \((\cdot)^T\) gives the transpose of a matrix, \(\mathbb{E}\{\cdot\}\) represents expectation, \(\text{Pr}\{\cdot\}\) stands for probability, \(\|\cdot\|\) denotes the squared Frobenius norm of a matrix or the squared Euclidean norm of a vector, \(\{\cdot\}\) gives the transpose of a matrix, \(\arg\{\cdot\}\) returns the angle of a complex scalar within \([-\pi, \pi]\), and \(\text{Re}\{\cdot\}\) and \(\text{Im}\{\cdot\}\) represent the real and imaginary operators, respectively, which return the real and imaginary part, respectively, of the matrix/vector/scalar they are applied to.

II. DOUBLE-DIFFERENTIAL ENCODING

Let \( S_k, k \geq 2 \), be an \( n_t \times n_t \) OSTBC matrix obtained from\n\[
S_k = [s_1, s_2, \ldots, s_{n_t}]^T, \ n_s \leq n_t, s_i \in \mathbb{C}^{n_t}\text{-PSK constellation}.
\]
Moreover, \( S_kS_k^H = a\|s_k\|^2 I_{n_t} = an_\sigma s_k^2 I_{n_t} \), where \( a \) is a scaling factor and \( \sigma^2 \) is the average power of \( s_i \). In order to avoid the power fluctuations, we assume that \( a = 1 \) and the signal constellation is normalized such that \( \sigma^2_k = 1/n_s \) and \( S_kS_k^H = I_{n_t} \). The first order differential matrix \( C_k \) of size \( n_t \times n_t \) can be obtained as shown in Fig. 1
\[
C_k = C_{k-1}S_k, \ k \geq 2,
\]
with \( C_1 \in \mathbb{C}^{n_t \times n_t} \) as the initialization matrix. Next, the second order differential matrix \( D_k \) can be obtained from \( C_k \) as shown in Fig. 1
\[
D_k = D_{k-1}C_k, \ k \geq 1,
\]
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:
\[
D_p = D_{p-2}C_{p-1}S_p = D_0C_1 \cdot \ldots \cdot C_{p-2}C_{p-1}S_p,
\]
Lemma 1: The matrices \( D_p \) and \( C_p \) are full-rank when \( D_0 \) and \( C_1 \) are full-rank and the OSTBC matrix \( S_p \) is square.

See Appendix I for the proof of Lemma 1. Lemma 1 indicates an important result about the choice of the initialization matrices \( D_0 \) and \( C_1 \). It is apparent that from the full diversity point of view \([8] D_0 \) and \( C_1 \) must be full-rank matrices. Lemma 1 also indirectly puts a restriction on the signal alphabet which can be used for DD encoding since the alphabet cannot contain the origin. In addition, the channel codes used for single or double-differential modulation must be full-rank and square.

III. CHANNEL MODEL

Consider a MIMO system 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 \( n_t \times n_t \) DD encoded matrix. Let \( \omega_m \in [-\pi, \pi] \) be the random carrier offset between the transmit antennas and the \( m \)-th receive antenna, which is independent of time \( k \), and assumed to be constant over the transmission period of at least three DD encoded matrices \( D_k \). The received data \( y_{m,k} \in \mathbb{C}^{1 \times n_t} \) at the \( m \)-th receive antenna corresponding to the \( D_k, k \geq 0 \), is \([8, \text{Eq. (9.7.14)}]\)
\[
y_{m,k} = \exp (j\omega_m n_t k) h_m F D_k \Omega_m + q_{m,k},
\]
where \( \Omega_m \in \mathbb{C}^{n_t \times n_t} \) is a diagonal matrix \( \Omega_m = \text{diag}[1, \exp(j\omega_m), \ldots, \exp(j\omega_m(n_t - 1))] \), \( h_m = [h_{m,1}, \ldots, h_{m,n_t}] \) is an \( n_t \times 1 \) row vector consisting of channel coefficients between the transmit antennas and \( m \)-th receive antenna. \( d_{m,k} \in \mathbb{C}^{1 \times n_t} \) contains additive white complex-valued Gaussian noise (AWGN), whose elements are i.i.d. Gaussian random variables with zero mean and variance \( \sigma^2 \), and \( F \in \mathbb{C}^{n_t \times n_t} \) is a precoder matrix which is used to improve the performance of the DDOSTBC over correlated channels. The channel \( H = [h_1^T, h_2^T, \ldots, h_n^T]^T \) is assumed to be constant over the transmission period of at least three consecutively DD encoded matrices.

A. Model of Correlated Rayleigh Channels

We assume flat block-fading correlated Rayleigh channel model \([11]\) for the matrix \( H \) which is the fading part of our channel model. Let the channel \( H \) have zero mean, complex Gaussian circularly distribution with positive semidefinite autocorrelation given by \( R = E[\text{vec}(H)\text{vec}^H(H)] \) of size \( n_t n_r \times n_t n_r \). A channel realization of the correlated Rayleigh
channels can be found by $\text{vec}(H) = R^{1/2} \text{vec}(H_w)$, where $R^{1/2}$ is the unique positive semidefinite matrix square root [12] of $R$ and $H_w$ is an $n_r \times n_t$ matrix consisting of complex circular Gaussian distributed elements with zero mean and $\gamma^2$ variance.

**Kronecker Model:** A special case of the model given above is the Kronecker model, which can be represented as [11, Eq. (3.26)]

$$R = R_t^T \otimes R_r,$$  

where, $R_t$ is the $n_r \times n_r$ receive correlation matrix and $R_r$ is the $n_t \times n_t$ transmit correlation matrix.

### IV. Decoding of DDOSTBC

The double-differential encoding at the transmitter may be utilized to obtain a decoder based on heuristics (ad-hoc ways) which is different from the ML approach as shown in [6], [7], [10]. In this paper, we try to follow a systematic way of obtaining a decoder of DD encoded OSTBC by minimizing the ML metric. However, this is very difficult, and we make some simplifying approximations.

We first consider the case of uncorrelated channel, i.e., $R = I_{n_r n_t}$ and without precoding $F = I_{n_t}$. From (4), it can be seen that if $\omega_m$, $D_k$, and $C_k$ are known for all $k$ and $m$, $y_{m,k}$ is distributed as

$$f(y_{m,k}|\omega_m,h_m,D_k) = \frac{1}{\pi^{n_r} \det(\Psi)} \times \exp \left( - \left[ y_{m,k} - \exp(j\omega_m n_k) h_m D_k \Omega_m \right]^* \Psi^{-1} \left[ y_{m,k} - \exp(j\omega_m n_k) h_m D_k \Omega_m \right] \right),$$

(6)

where $\Psi$ is the covariance matrix of $q_{m,k}$. Since $q_{m,k}$ is AWGN, $\Psi = \sigma^2 I_{n_r}$. Let $y_{m,k} = [y_{m,k-2} \ldots y_{m,k-1} y_{m,k}]$ be the vector consisting of the received data samples in three consecutive time intervals at receive antenna $m$. When $h_m$, $\omega_m$, $D_{k-2}$, $C_{k-1}$, and $S_k$ are known,

$$f(y_m|\omega_m,h_m,D_{k-2},C_{k-1},S_k) = \frac{1}{\pi^{n_r(m)}(\sigma^2)^{3m_t}} \times \exp \left( - \frac{1}{\sigma^2} \sum_{i=k-2}^k \|y_{m,i} - \exp(j\omega_m n_i) h_m D_i \Omega_m\|^2 \right).$$

(7)

In order to find an estimate of the unknown data $S_k$, the joint probability distribution function (p.d.f.) of (7) is first maximized with respect to (w.r.t) all unknown quantities $\omega_m$, $h_m$, $D_{k-2}$, and $C_{k-1}$ and, subsequently, over $S_k$, which results into minimization of the following ML metric:

$$\Gamma_k = \sum_{i=k-2}^k \|y_{m,i} - \exp(j\omega_m n_i) h_m D_i \Omega_m\|^2, \quad k \geq 2.$$  

(8)

Minimization of (8) w.r.t. $h_m$, $\omega_m$, $D_{k-2}$, and $C_{k-1}$, and $S_k$ is very complicated. Therefore, we focus on a suboptimal decoder which makes the decision independent of the channel and carrier offset knowledge.

### A. Suboptimal Decoder of DDOSTBC

**Lemma 2:** Let $X_k$ is a square OSTBC matrix, then $\tilde{X}_k = \Omega_m^T X_k \Omega_m$ is also a square OSTBC matrix.

Refer Appendix II for the proof of Lemma 2.

To simplify the decision process, we consider a degenerated decision metric from (8) as follows:

$$D_{m,k} = \sum_{l=k-2}^{k-1} \|y_{m,l} - \exp(j\omega_m n_l) h_m D_l \Omega_m\|^2, \quad k \geq 2,$$  

(9)

i.e., out of the three data vectors received at time $k-2$, $k-1$, and $k$ in (8), we consider only first two data vectors received at time $k-2$ and $k-1$. By means of Lemma 2, the unitary property of the OSTBC matrix using the $M$-PSK constellation [8], and the results given in [13] for matrix derivatives, (9) is minimized w.r.t. $g_{m,k} = \exp(j\omega_m n_l (k-2)) h_m D_{k-2} \Omega_m$

$$\hat{g}_{m,k} = \frac{1}{2} \left( \mathcal{E}_{\omega_m} y_{m,k-1} \hat{C}_{k-1} + y_{m,k-2} \right),$$

(10)

where $\mathcal{E}_{\omega_m} = \exp(j\omega_m n_l)$ and $\hat{C}_{k-1} = \Omega_m^T \hat{C}_{k-1} \Omega_m$.

A detailed derivation of (10) is given in Appendix III. By substituting (10) into (9) and using the unitary property of OSTBC [8], (9) reduces into

$$D_{m,k} = \|y_{m,k-1} - \mathcal{E}_{\omega_m} y_{m,k-2} \hat{C}_{k-1}\|^2, \quad k \geq 2.$$  

(11)

From (11), it can be seen that if $\hat{C}_{k-1}$ is unknown, it is impossible to find the estimate of $\mathcal{E}_{\omega_m}$.

**Lemma 3:** If $\hat{C}_{k-1} = I_{n_t}$, the estimator of $\mathcal{E}_{\omega_m}$ is given by

$$\hat{\mathcal{E}}_{\omega_m,k-1} = \exp\left\{j \arg\left\{ y_{m,k-1} y_{m,k-2}^H \right\} \right\}. $$

(12)

Refer to Appendix IV for the proof.

However, we cannot take the liberty of assuming $\hat{C}_{k-1}$ as the identity matrix as we are using OSTBC matrices. Nevertheless, we may assume that the matrix $C_1$ is equal to the identity matrix. Then, the estimate of $\mathcal{E}_{\omega_m}$ can be found from (12) at $k = 2$, such that $\hat{\mathcal{E}}_{\omega_m,1}$ can be used in the place of $\mathcal{E}_{\omega_m}$ in the further analysis.

**Remark 1:** For double-differential modulation, it can be assumed that the receiver has perfect knowledge about the matrix $C_1$, whereas, $D_0$ can be arbitrarily chosen full-rank matrix unknown to the receiver and it can be seen from (2) that $D_1$ is also unknown to the receiver. Hence, no training data or pilot symbols are used in the proposed DD scheme. The receiver can reconstruct the subsequent $C_{k-1}$, $k > 2$, from the estimated data and (1). However, $C_{k-1}$, $k > 2$, are noisy versions of the original $C_{k-1}$, $k > 2$. Hence, the estimates of $\mathcal{E}_{\omega_m}$ at $k > 2$, based on $\hat{C}_{k-1}$, $k > 2$, will be worse than $\hat{\mathcal{E}}_{\omega_m,1}$. Hence, $\hat{\mathcal{E}}_{\omega_m,1}$ should be used for all $k$ (within the duration of a frame) as the estimate of carrier offset.

\(^1\) A frame consists of all OSTBC matrices which are consecutively transmitted while the channel remains constant. The minimum value of three block usages can be used in this work.
From (1), (2), and (4), the data received at the $k$-th time instant and at the $m$-th receive antenna can be written in the terms of $\mathcal{E}_{\omega_m}$, $g_{m,k}$, and $C_{k-1}$ as
\[ y_{m,k} = \mathcal{E}_{\omega_m} g_{m,k} \Omega_m^H C_{k-1} S_k \Omega_m + q_{m,k}. \] (13)
Hence, we can minimize the following decision metric to find the estimate of $S_k, k \geq 2$:
\[ \hat{S}_k = \arg \min_{S_k \in \Xi} \left\| y_{m,k} - \mathcal{E}_{\omega_m} \hat{g}_{m,k} \Omega_m^H C_{k-1} \hat{S}_k \Omega_m \right\|^2, \] (14)
where $\Xi$ is the set of all OSTBC matrices consisting of symbols from the $M$-PSK constellation. $\Omega_m$ is obtained from $\mathcal{E}_{\omega_m}$, and $\hat{S}_k = \Omega_m^H \hat{S}_k \Omega_m$. Let $h_m = \mathcal{E}_{\omega_m} \hat{g}_{m,k} \hat{S}_k$, then (14) can be written as
\[ \hat{S}_k = \arg \min_{S_k \in \Xi} \left\| y_{m,k} - h_m \hat{S}_k \right\|^2. \] (15)
Equation (15) can be seen as the decoding of $S_k$ given that the receiver knows the channel $h_m$, the received vector $y_{m,k}$, and the carrier offset matrix $\Omega_m$. It is shown in Appendix V that the decoding of $S_k$ can be done linearly. Hence, the proposed suboptimal decoder (15) is computationally much less complex than the suboptimal decoder of [10, Eq. (29)].
Further, the decision is based on three consecutively received data blocks, hence, the decoding complexity is also lower than the coherent decoder, which utilizes specific training matrices, FFT, and maxima searching procedure for estimation of the carrier offsets [8, Eq. (9.7.13)], and training based channel coefficients estimator [8, Eq. (9.7.5)] before decoding the OSTBC data. Apparently, the decoding of $S_k$ from (14) requires the knowledge of $C_{k-1}$. It can be assumed that the receiver has full knowledge about $C_1$ and in the subsequent time intervals, it obtains $C_{k-1}$ from (1).
The decoding of (14) can be generalized for $k \geq 2$ to the $n_r$ receive antennas case as
\[ \hat{S}_k = \arg \min_{S_k \in \Xi} \sum_{m=1}^{n_r} \left\| y_{m,k} - \mathcal{E}_{\omega_m} \hat{g}_{m,k} \hat{S}_k \Omega_m \right\|^2. \] (16)
From the discussion above, it is clear that we have not been able to find a true ML decoder since it is very complicated, however, we have obtained an approximation.

V. PERFORMANCE ANALYSIS OF DDOSTBC

In this section, we analyze 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 $S_k^0$ where $S_k \neq S_k^0$ can be written as
\[ \Pr \{ S_k^0 \rightarrow S_k \} = \Pr \left\{ \sum_{m=1}^{n_r} \left\| y_{m,k}^0 - \mathcal{E}_{\omega_m} g_{m,k} \Omega_m^H C_{k-1} S_k^0 \Omega_m \right\|^2 \right\} \]
\[ \leq \sum_{m=1}^{n_r} \left\| y_{m,k}^0 - \mathcal{E}_{\omega_m} g_{m,k} \Omega_m^H C_{k-1} S_k^0 \Omega_m \right\|^2, \] (17)
where $y_{m,k}^0 = \exp (j \omega_m n_k) h_{m,k} + D_{k-2} C_{k-1}^0 \Omega_m + q_{m,k}$ and $C_{k-1}$ is the estimate of $C_{k-1}$. It is shown in Appendix VI that
\[ \Pr \{ S_k^0 \rightarrow S_k \} \leq \exp \left\{ - \frac{\sum_{m=1}^{n_r} \left\| \lambda_{\min}(C_{k-1}) \right\| \left\| y_{m,k} S_k^0 \right\|^2}{6 \sigma^2} \right\}, \] (18)
where $S_k = S_k^0 - S_k$, $\lambda_{\min}(C_{k-1})$ is the minimum eigenvalue of $C_{k-1}$, $W_{m,k} = \mathcal{E}_{\omega_m} \Omega_m^H C_{k-1}^2$.

Theorem 1: The average pairwise error probability of DDOSTBC is bounded by
\[ \mathbb{E}_H [ \Pr \{ S_k^0 \rightarrow S_k \} ] \leq \left( S_k^0 - S_k \right) A_k \left( S_k^0 - S_k \right)^H \right)^{-n_r} \times \left( \gamma^2 \right)^{-n_r}, \] (20)
where $A_k = \lambda_{\min}(C_{k-1}) I_{n_r}$.
See Appendix VII for the proof of Theorem 1.
It is apparent from (20) that the proposed DDOSTBC achieves full diversity ($n_t n_r$).

A. Coding Gain Comparisons

The double-differential coding scheme proposed in [10] utilizes a diagonal OSTBC matrix. The diagonal OSTBC matrix belongs to the unitary diagonal group specified by $G = \{ I_n, \mathcal{C}, \ldots, \mathcal{C}^{M-1} \}$ [10, Eq. (46)], where $C$ is the $n_t \times n_r$ diagonal OSTBC matrix [10, Eq. (47)]
\[ C = \begin{bmatrix} z_{k_1} & 0 & \ldots & 0 \\ 0 & z_{k_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & z_{k_{n_r}} \end{bmatrix}, \] (21)
where $z_{M} = \exp (2 \pi j / M)$ and $k_1, k_2, \ldots, k_{n_r}$ are arbitrary powers which can be decided by optimizing the coding advantage [10, Eq. (44)]. It can be seen that the diagonal OSTBC of (21) utilizes $n_t$ time intervals to transmit one $M$-PSK symbol. On the other hand, the proposed DDOSTBC transmits $n_s$ $M$-PSK symbols in $n_r$ time intervals through OSTBC matrix. Hence, we can compare the two codes on the basis of the same spectral efficiency. Comparing (20) and [10, Eq. (42)], and using the definition given in [10, Eq. (44)], the
The coding advantage and rate of the proposed DDOSTBC is numerically calculated and compared with the previously proposed DD coding for MIMO channels [10] in Table I for different number of transmit antennas and different constellation sizes. The proposed DDOSTBC uses Alamouti code [14] for \( n_t = 2 \) and the OSTBC given by [8, Eq. (7.4.8)] for \( n_t = 4 \). We obtain PEP optimal diagonal unitary codes [10] by maximizing the coding advantage [10, Eq. (49)] for different transmit antennas and constellations. These are listed in Table I for comparison with the proposed DDOSTBC. It can be seen from Table I that for the same spectral efficiency, the proposed DDOSTBC provides higher coding gain as compared to the previously proposed DDOSTBC [10] for all the cases considered in the Table I.

VI. PRECODED DDOSTBC

In order to improve the performance of DDOSTBC over arbitrarily correlated Rayleigh channel, we introduce a full memoryless precoder \( F \in \mathbb{C}^{n_t \times n_t} \) as shown in Fig. 1 before the DDOSTBC matrix \( D_k \in \mathbb{C}^{n_t \times n_t} \) is transmitted over the channel. Hence, in place of \( D_k \), we now transmit precoded DDOSTBC matrix \( FD_k \in \mathbb{C}^{n_t \times n_t} \), see Fig. 1. The precoder matrix does not change the decoder of the DDOSTBC as the precoding matrix can be absorbed into the channel as \( h_n = h_n F \). Therefore, the decoding can still be done by using (16).

**Theorem 2:** The pairwise error probability of the precoded DDOSTBC over arbitrarily correlated Rayleigh channel is bounded by

\[
\mathbb{E}_{\mathbf{H}} \left[ \Pr \{ S_0^k \to S_k \} \right] \leq \left[ I_{n_t n_t} + \frac{\lambda_{\min}(\mathbf{X}_k)}{12 \sigma^2} \left( (S_0^k - S_k)^T \mathbf{F}^T \otimes I_{n_t} \right) \mathbf{R} \right. \\
\times \left. \left( (S_0^k - S_k)^T \mathbf{F}^T \otimes I_{n_t} \right)^{-1} \right].
\]

See Appendix VIII for the proof of Theorem 2.

### A. Design of Precoder for DDOSTBC

In this subsection, we will discuss the design of the precoder for the DSDOTBC. Since the channel statistics varies far more slowly than the channel coefficients, we assume that the receiver can estimate the channel correlation matrix and noise variance, and feed these back to the transmitter. Under the assumption that the transmitter knows the channel correlations and the noise variance perfectly, we will discuss the precoder design for the DDOSTBC. By using the properties of OSTBC [8], it can be shown that the average power per DDOSTBC block is

\[
\mathbb{E} \{ D_k D_k^H \} = I_{n_t}.
\]

Hence, the average power constraint on the transmitted block \( FD_k \) can be expressed as

\[
\text{Tr} \left\{ \mathbf{F} \mathbf{F}^H \right\} = P,
\]

where \( P \) is the average power given by the transmitted block \( FD_k \). The goal is to design a precoder \( F \) such that an upper bound of the pairwise error probability (UBPEP) is minimized with a constraint over the average power on the transmitted block \( FD_k \). We can express this optimization problem as [15]–[18]

\[
\min_{\mathbf{F} \in \mathbb{C}^{n_t \times n_t}\{ \text{Tr}[\mathbf{F} \mathbf{F}^H] = P \}} \text{UBPEP}.
\]

It can be seen from (26), that a closed-form precoder is difficult for the general case of arbitrary correlation. Hence, we focus on a numerical method to design the precoder in the next subsection.

### B. Precoder Design

In order to find an optimized precoder for minimizing the upper bound of PEP, we can follow a similar method as given in [15]. The constrained minimization problem of Subsection VI-A can be converted into an unconstrained minimization problem by introducing a Lagrange multiplier \( \mu' \):

\[
\mathcal{L}(\mathbf{F}) = \ln(\text{UBPEP}) + \mu' \text{Tr} \left\{ \mathbf{F} \mathbf{F}^H \right\}.
\]

Since the objective function should be minimized, \( \mu' > 0 \). It can be observed that the use of natural logarithm of UBPEP, i.e., \( \ln(\text{UBPEP}) \), does not change the nature of the objective function because \( \ln() \) is a monotonically increasing function. Let \( \mathcal{L}_{\text{UBPEP}}(\mathbf{F}) = \ln(\text{UBPEP}) \), then it can be shown after some manipulations that [13]

\[
\text{vec}^T(\mathbf{F}) = \mu \mathcal{D}_F \cdot \mathcal{L}_{\text{UBPEP}}(\mathbf{F}),
\]
where \( D_F \cdot L_{\text{UBPEP}}(F) \) is the first order derivative of \( L_{\text{UBPEP}}(F) \) with respect to \( F^* \) [13], and \( \mu \) is a scalar chosen such that the power constraint in (25) is satisfied. The derivative can be found by using the results in [13], and the details of the derivation are given in Appendix IX. We can summarize the results of the derivative as follows:

\[
D_F \cdot L_{\text{UBPEP}}(F) = \text{vec}^H \left( R^H S^H \right) \Pi Q^H, \tag{29}
\]

where

\[
\begin{align*}
R &= \left( \left( S_k^0 - S_k \right) F^T \otimes I_{n_r} \right) R, Q = \left( S_k^0 - S_k \right) \otimes I_{n_r} \otimes \text{vec} \left( I_{n_r} \right), \\
I_{n_t}, \Pi &= \left( I_{n_t} \otimes K_{n_r,n_t} \otimes I_{n_r} \right) \left( I_{n_t} \otimes \text{vec} \left( I_{n_r} \right) \right), \\
S &= -\frac{\lambda^2_{\text{min}} \left( \mathbf{X}_k \right)}{12\sigma^2} \left[ I_{n_t n_r} + \frac{\lambda^2_{\text{min}} \left( \mathbf{X}_k \right)}{12\sigma^2} \left( \left( S_k^0 - S_k \right)^T \right) \times F^T \otimes I_{n_r} \right] R \left( \left( S_k^0 - S_k \right)^T F^T \otimes I_{n_r} \right)^{-1},
\end{align*}
\]

where \( K_{n_r,n_t} \) is the commutation matrix of size \( n_r n_t \times n_r n_t \) [13]. A description of the procedure of minimizing the upper bound of the PEP with respect to the precoder \( F \) is summarized with pseudo-code in Table II.

### VII. Simulation Results

The simulations are performed with \( n_t \in \{2, 4, 8\} \), \( n_r \in \{1, 2\} \), and complex unit variance AWGN noise. The complex valued Gaussian 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. The initialization matrices \( D_0 \) and \( C_1 \) are chosen as the identity matrices and it is assumed that receiver perfectly knows \( C_1 \). The simulation results are obtained from \( 10^6 \) channel realizations for each value of SNR.

![Fig. 2. SER versus SNR performance of the proposed DDOSTBC as compared to the existing DDOSTBC [10], trained decoder [8], and single-differential decoder.](image)

**A. Comparison of the Proposed DDOSTBC with the Existing DDOSTBC [10], Trained Decoder [8], and Single-Differential Decoder**

The channel is assumed to be uncorrelated, i.e., \( R = I_{n_t n_r} \) and \( F = I_{n_t} \). The proposed DDOSTBC are applied over Alamouti STBC [14] with QPSK constellation, \( n_t = 2 \), and \( n_r = 1 \). The carrier offset is assumed to be randomly distributed over \([-\pi, \pi] \). The DD code of [10] utilizes diagonal OSTBC matrices, which can transmit only one M-PSK symbol per OSTBC block, hence the DDOSTBC of [10] provides low data rates. Therefore, for a fair comparison at the same spectral efficiency, the simulations are performed for (16:1:7) DDOSTBC of [10] which transmits the following diagonal matrix in place of 16-PSK symbols

\[
\mathbf{C} = \begin{bmatrix} z_{16} & 0 \\ 0 & z_{16}^{-1} \end{bmatrix}, \tag{30}
\]

where \( z_{16} = \exp(2\pi j/16) \). It can be seen from Fig. 2 that our proposed DDOSTBC outperforms the existing DDOSTBC [10] by a large margin, for example at SER=10^{-3}, the gain is approximately 4 dB. We have also plotted the performance of a same rate trained decoder [8], 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. 2 that the proposed DDOSTBC also outperforms the trained receiver [8] with QPSK constellation for moderate to high SNRs. In addition, the performance of a single-differential modulation scheme with carrier offset estimator is also shown in Fig. 2, where in place of the proposed DD encoding, the single-differential encoding [1]–[4] is used. The single-differential system uses one initialization matrix which is kept identity (similar to \( C_1 \) in the proposed DD scheme). It is assumed that this

**TABLE II**

<table>
<thead>
<tr>
<th>Step 1: Initialization</th>
<th>Step 2: Precoder Optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose OSTBC, values for ( P, n_t, n_r, \epsilon, ) and ( M )</td>
<td>( i := 0 ) repeat ( i := i + 1 ) until ( | F_i - F_{i-1} | &lt; \epsilon )</td>
</tr>
<tr>
<td>Estimate the channel correlation matrix ( R ) and noise variance ( \sigma^2 )</td>
<td>Calculate the right hand side of (28).</td>
</tr>
<tr>
<td>Initialize the precoder ( F ) to an already optimized precoder or to ( F_0 = \sqrt{P/I_{n_t}} )</td>
<td>Normalize the precoder matrix ( F_i ) such that: ( | F_i | = \sqrt{P/I_{n_t}} ), to satisfy the power constraint.</td>
</tr>
<tr>
<td>summarized with pseudo-code in Table II.</td>
<td>The optimized precoder is given by ( F = F_i )</td>
</tr>
</tbody>
</table>

**Pseudo-code of the numerical optimization algorithm for the precoder.**

*Initialization*

Choose OSTBC, values for \( P, n_t, n_r, \epsilon, \) and \( M \). Estimate the channel correlation matrix \( R \) and noise variance \( \sigma^2 \). Initialize the precoder \( F \) to an already optimized precoder or to \( F_0 = \sqrt{P/I_{n_t}} \).

*Precoder Optimization*

\( i := 0 \) repeat

\( i := i + 1 \)

Calculate the right hand side of (28).

Normalize the precoder matrix \( F_i \) such that: \( \| F_i \| = \sqrt{P/I_{n_t}} \), to satisfy the power constraint.

until \( \| F_i - F_{i-1} \| < \epsilon \)

The optimized precoder is given by \( F = F_i \).
initialization matrix is known to the receiver and it can utilize it to find the estimate of the carrier offset. Unfortunately, no efficient estimator is available in the literature to estimate the carrier offset based on one initialization block only. Therefore, for simulation we have considered a hypothetical estimator with mean square error: \( \text{MSE} = \mathbb{E} \{|\hat{\omega}_m - \omega_m|^2\} \). 

\( \hat{\omega}_m \) is the estimate of the carrier offset \( \omega_m \) and \( \Delta \omega_m = \omega_m - \hat{\omega}_m \) is uniformly distributed error in the estimation. The MSE of the estimator varies inversely with SNR, i.e., when \( \text{SNR} \to \infty \), \( \text{MSE} \to 0 \), \( \Delta \omega_m \to 0 \) and when \( \text{SNR} \to 0 \), \( \text{MSE} \to (1/3)\pi^2 \), \( \Delta \omega_m \to \pm \pi \) which maximizes the error in \( \hat{\omega}_m \). It can be seen from Fig. 2 that the single-differential decoder performs poorer than the proposed DD scheme at all SNRs.

The SER versus SNR performance of the proposed DDOSTBC with different number of transmit antennas (Tx) and single receive antenna is shown in Fig. 3. The simulations are performed with \( n_t = 2 \) and Alamouti STBC [14], \( n_t = 4 \) and OSTBC of [8, Eq. (7.4.10)], and \( n_t = 8 \) and OSTBC of [8, Eq. (7.4.11)]. It can be seen from Fig. 3 that the diversity of the proposed DDOSTBC improves with the increase in number of the transmit antennas.

**B. Performance of the Precoded DDOSTBC over Correlated Rayleigh Channels**

We have performed simulations for the Alamouit code with \( n_t = 2 \) and \( n_r = 2 \), and QPSK constellation. The channel is assumed to be a correlated Rayleigh channel with \( \mathbf{R}_{i,j} = (\rho)^{i-j}, \quad 1 \leq \{i,j\} \leq n_t n_r, \quad \text{with} \quad \rho \in \{0.9, 0.99, 0.99999\} \).

Each receive antenna is assumed to have different carrier offset randomly distributed over \([-\pi, \pi]\). It can be seen from Fig. 4 that the proposed precoded DDOSTBC achieves performance gain over the correlated Rayleigh channel as compared to the unprecoded DDOSTBC of Section II. From Fig. 4, it can be seen that when the channel is highly correlated with \( \rho = 0.99999 \), the precoded DDOSTBC performs approximately 3 dB better than the DDOSTBC with trivial precoder at SER of \( 10^{-2} \).

**VIII. Conclusions**

We have proposed double-differential coding for square orthogonal space-time block codes over arbitrarily correlated Rayleigh channels. The proposed double-differential codes are able to decode the space-time data without knowing the carrier offsets or channel coefficients and achieve higher coding gain as compared to the previously proposed double-differential space-time block codes. The proposed double-differential scheme also works better than a similar rate conventional trained decoding scheme, which relies on the estimates of the channel coefficients and carrier offsets obtained through training data. We have also proposed a PEP based precoder design for the proposed double-differential codes over arbitrarily correlated Rayleigh channels.

**Appendix I**

**Proof of Lemma 1**

It can be observed from [19, Eq. (1.7.4)] that if \( \mathbf{Y} \) and \( \mathbf{Z} \) are two arbitrary matrices with sizes such that the product \( \mathbf{Y} \mathbf{Z} \) is defined, then

\[
\text{rank} \left( \mathbf{Y} \mathbf{Z} \right) \leq \min \left( \text{rank} \left( \mathbf{Y} \right), \text{rank} \left( \mathbf{Z} \right) \right). \tag{31}
\]

It is apparent from (1), (3), and (31) that

\[
\text{rank} \left( \mathbf{C}_p \right) \leq \min \left\{ \text{rank} \left( \mathbf{C}_1 \right), \text{rank} \left( \mathbf{S}_1 \right), \ldots, \text{rank} \left( \mathbf{S}_{p-1} \right), \text{rank} \left( \mathbf{S}_p \right) \right\}. \tag{32}
\]
rank \((D_p)\) \leq \min \{\text{rank} \((D_0)\), \text{rank} \((C_1)\), \ldots , \text{rank} \((C^2_{p-1})\), \text{rank} \((S_p)\)\}. \quad (33)

It follows from (1) and (32) that if \(S_p\) for all \(p \geq 2\) is not a full-rank matrix, the product of matrices in (1) will have diminishing rank property and could lead to a zero rank matrix, which is an all zero matrix. Similar observations can be pointed out from (3) and (33). It can also be seen from [19, Eq. (1.7.5)] that

\[
\text{rank} \((YZ)\) = \text{rank} \((Y)\),
\]

if \(Z\) is a square matrix of full-rank. Hence, it is clear from (32) and (34) that if \(S_p\) for all \(p \geq 2\) is full-rank square matrix,

\[
\text{rank} \((C_p)\) = \text{rank} \((C_1)\).
\]

It can be seen from (3), that \(D_p\) is obtained from \(C^2_{p-1}\), which is only possible if \(C_{p-1}\) is a square matrix. From (31)-(35), it can be observed that if \(S_p\) for all \(p \geq 2\) is full-rank square matrix and \(C_1\) is also a full-rank square matrix then

\[
\text{rank} \((D_p)\) = \text{rank} \((D_0)\). \quad (36)
\]

Hence, the differential encoding of (1), (2), and (3) will provide full-rank first and second order differential matrices when \(D_0\) and \(C_1\) are full-rank and the OSTBC matrix \(S_p\) is square.

\section*{APPENDIX II}

\textbf{PROOF OF LEMMA 2}

From Section III, it follows that \(\Omega_m\) is a diagonal matrix with \((\Omega_m)_{p,q} = \exp \((p \omega_{m} q \pi (p - 1))\) and \((\Omega_m)_{p,p} = 0\) for \(p \neq q\). Therefore, \(\Omega_m \Omega^T_m = \text{I}_n\) and \(X_k \Omega_m = \Omega_m \Omega^T_m X_k \Omega_m = \Omega_m X_k \Omega_m\), where \(X_k = \Omega^T_m X_k \Omega_m\). As \(X_k\) is an square OSTBC matrix it can be expressed as [8, Eq. (7.1.1)]

\[
X_k = \sum_{n=1}^{n} \left( \text{Re} \{s_n\} A_n + j \text{Im} \{s_n\} B_n \right),
\]

where \(A_n\) and \(B_n\) are independent fixed (in general complex-valued) code matrices of size \(n_t \times n_t\), which satisfy the following properties:

\[
A_n A^T_n = I_{n_t}, B_n B^T_n = I_{n_t}, A_n B^T_n = -B_n A^T_n, \quad \forall n, n',
\]

\[
A_n A^T_n = -A_n A^T_n, B_n B^T_n = -B_n B^T_n, \quad n \neq n', \quad (38)
\]

for \(n = 1, \ldots, n_s, r = 1, \ldots, n_s\). We can express \(X_k\) in the terms of \(A_n\) and \(B_n\) as

\[
\tilde{X}_k = \Omega^T_m \left( \sum_{n=1}^{n} \left( \text{Re} \{s_n\} A_n + j \text{Im} \{s_n\} B_n \right) \right) \Omega_m
\]

\[
= \sum_{n=1}^{n_s} \left( \text{Re} \{s_n\} \Omega^T_m A_n \Omega_m + j \text{Im} \{s_n\} \Omega^T_m B_n \Omega_m \right)
\]

\[
= \sum_{n=1}^{n_s} \left( \text{Re} \{s_n\} \tilde{A}_n + j \text{Im} \{s_n\} \tilde{B}_n \right).
\]

It can be verified that \(\tilde{A}_n\) and \(\tilde{B}_n\) satisfy (38). Hence, \(\tilde{X}_k\) is also an square OSTBC.

\section*{APPENDIX III}

\textbf{DERIVATION OF (10)}

The decision metric of (9) can be expressed in the terms of \(g_{m,k}\) as follows:

\[
D_{m,k} = \left\| y_{m,k-2} - g_{m,k} \right\|^2 + \left\| y_{m,k-1} - \mathcal{E}_{\omega_m} g_{m,k} \tilde{C}_k^{-1} \right\|^2.
\]

(40)

As \(\mathcal{E}_{\omega_m} = 1\) and \(\tilde{C}_k^{-1} = I_{n_t}\), it can be shown that

\[
D_{m,k} = \left\| y_{m,k-2} \right\|^2 + \left\| y_{m,k-1} \right\|^2 + 2 \left\| g_{m,k} \right\|^2 - 2 \text{Re}\left\{ y_{m,k-2} g_{m,k}^* + \mathcal{E}_{\omega_m} y_{m,k-1} \tilde{C}_k^{-1} g_{m,k}^* \right\}.
\]

(41)

Differentiating (41) w.r.t. \(g_{m,k}^*\) and equating the result to zero leads to (10).

\section*{APPENDIX IV}

\textbf{PROOF OF LEMMA 3}

If \(C_{k-1} = I_{n_t}\), then \(\tilde{C}_k = I_{n_t}\). We can write the decision metric of (11) as follows:

\[
D_{m,k} = \left\| y_{m,k-1} \right\|^2 + \left\| y_{m,k-2} \right\|^2 - \mathcal{E}_{\omega_m} y_{m,k-1} \tilde{Y}_{m,k-2} - \mathcal{E}_{\omega_m} y_{m,k-2} \tilde{Y}_{m,k-1}.
\]

(42)

Let \(\mathcal{E}_{\omega_m} = \exp (j \theta)\). Differentiating \(D_{m,k}\) w.r.t. the real scalar parameter \(\theta\) and equating the result to zero we get following:

\[
\exp (j \theta) = \frac{y_{m,k-1} y_{m,k-2} - y_{m,k-2} y_{m,k-1}}{\exp \left( j \arg \left\{ y_{m,k-1} y_{m,k-2} \right\} \right) - \exp \left( -j \arg \left\{ y_{m,k-1} y_{m,k-2} \right\} \right)}
\]

\[
= \exp \left( j 2 \arg \left\{ y_{m,k-1} y_{m,k-2} \right\} \right). \quad (43)
\]

From (43), it can be seen that \(\theta = \arg \left\{ y_{m,k-1} y_{m,k-2} \right\}\), from which (12) can be obtained.

\section*{APPENDIX V}

\textbf{PROOF OF DECOUPLED DECODING}

Equation (15) can be simplified as

\[
\hat{S}_k = \arg \min_{\hat{S}_k \in \mathbb{C}^{n_s}} \left\| y_{m,k} \right\|^2 + \Re \left\{ h_{m} \hat{S}_k \hat{S}_k^{H} \right\} - 2 \Re \left\{ h_{m} \hat{S}_k y_{m,k}^{H} \right\}.
\]

(44)

The first term in (44) does not contribute into the decision. It can be seen from Lemma 2 that \(\hat{S}_k\) is a square OSTBC matrix, therefore, \(\hat{S}_k\) can be expressed as (39). Hence, (44) reduces into

\[
\hat{s}_k = \arg \min_{\hat{s}_k \in \mathbb{C}^{n_s}} \left\| s_k \right\|^2 \left\| h_{m} \right\|^2 - 2 \sum_{n=1}^{n_s} \left( \Re \left\{ h_{m} \hat{A}_n y_{m,k}^{H} \right\} \times \Re \left\{ s_n \right\} + 2 \sum_{n=1}^{n_s} \left( \Re \left\{ h_{m} \hat{B}_n y_{m,k}^{H} \right\} \right) \text{Im} \left\{ s_n \right\} \right) \right).
\]

(45)
After some manipulations it can be shown that decoding of (45) reduces into
\[
\hat{s}_k = \arg \min_{s_k \in \mathcal{X}} \sum_{n=1}^{n_s} \left| s_n - \frac{\text{Re}\left\{ h_m \bar{A}_n y_{m,k}^\dagger \right\} - j \text{Im}\left\{ h_m \bar{B}_n y_{m,k}^\dagger \right\}}{\| h_m \|^2} \right|^2,
\]
(46)

where \( \mathcal{X} \) is the \( M \)-PSK alphabet. Equation (46) demonstrates that the ML metric in (44) decouples into a sum of \( n_s \) terms, where each term depends on exactly one complex symbol \( s_n \). Consequently, detection of \( s_n \) is decoupled from the detection of \( s_p \) for \( n \neq p \), hence, the decoding complexity is linear in \( n_s \).

**APPENDIX VI**

**DERIVATION OF (18)**

Let us define the following row vectors:
\[
w_{m,k} = \mathcal{E}_{\omega}^2 g_{m,k} \Omega_{m}^\dagger C_{k-1}^2,
\]
\[
\eta_{m,k} = -\left( \frac{\mathcal{E}_{\omega,1}^2}{2} q_{m,k-1} \Omega_{m}^\dagger + \frac{\mathcal{E}_{\omega,1}^2}{2} q_{m,k-2} \Omega_{m}^\dagger \right),
\]
(47)

where \( \eta_{m,k} \) is a vector of circular complex Gaussian variables satisfying
\[
\begin{align*}
\mathbb{E}\left\{ \eta_{m,k} | h_m \right\} &= 0_{1 \times n_s}, \\
\mathbb{E}\left\{ \eta_{m,k}^\dagger \eta_{m,k} | h_m \right\} &= \sigma^2 I_{n_s}.
\end{align*}
\]
(48)

Hence, the distribution of \( \eta_{m,k} \) is not changed by multiplying \( \eta_{m,k} \) with a unitary matrix like OSTBC using \( M \)-PSK. Substituting the value of \( y_{m,k} \) in (17) and after some manipulations it can be shown that
\[
\begin{align*}
\Pr \left\{ S_{k}^0 \rightarrow S_k \right\} &= \Pr \left\{ \sum_{n=1}^{n_s} \left[ w_{m,k} S_{k}^0 \Omega_{m}^\dagger \hat{v}_{m,k}^\dagger \eta_{m,k} \right] + v_{m,k} S_{k} \hat{v}_{m,k}^\dagger \Omega_{m}^\dagger \left( S_{k}^0 \right)^\dagger \eta_{m,k} \right\} < 0 \right. \\
&+ 2\text{Re}\left\{ v_{m,k} S_{k} \hat{v}_{m,k}^\dagger \left( q_{m,k} + \eta_{m,k} \right)^\dagger \right\} < 0 \right),
\end{align*}
\]
(49)

where \( S_k = S_{k}^0 - S_k \) and \( v_{m,k} \) is defined in (19). Let us define an intermediate variable \( \tau_k \) as
\[
\tau_k = \sum_{n=1}^{n_s} 2\text{Re}\left\{ v_{m,k} S_{k} \hat{v}_{m,k}^\dagger \left( q_{m,k} + \eta_{m,k} \right)^\dagger \right\}.
\]
(50)

By using the orthogonal properties of \( C_{k-1} \), \( S_{k}^0 \), and \( S_k \), it can be shown that \( \tau_k \), conditioned on the channels \( H = \left[ h_1^T, h_2^T, \ldots, h_{n_s}^T \right]^T \) is a zero mean complex Gaussian random variable satisfying
\[
\begin{align*}
\mathbb{E}\left\{ \tau_k | H \right\} &= 0, \\
\mathbb{E}\left\{ \tau_k \tau_k^\dagger | H \right\} &= 3 \sum_{n=1}^{n_s} \| v_{m,k} S_k \|^2 \sigma^2.
\end{align*}
\]
(51)

We can rewrite (49) as
\[
\begin{align*}
\Pr \left\{ S_{k}^0 \rightarrow S_k \right\} &= \Pr \left\{ \sum_{n=1}^{n_s} \left[ w_{m,k} S_{k}^0 \Omega_{m}^\dagger \hat{v}_{m,k}^\dagger \eta_{m,k} \right] + v_{m,k} S_{k} \hat{v}_{m,k}^\dagger \Omega_{m}^\dagger \left( S_{k}^0 \right)^\dagger \eta_{m,k} \right\} < 0 \right) \\
&+ 2\text{Re}\left\{ v_{m,k} S_{k} \hat{v}_{m,k}^\dagger \left( q_{m,k} + \eta_{m,k} \right)^\dagger \right\} < 0 \right) \\
&= \Pr \left\{ \sum_{n=1}^{n_s} \left[ w_{m,k} S_{k}^0 \Omega_{m}^\dagger \hat{v}_{m,k}^\dagger \eta_{m,k} \right] + v_{m,k} S_{k} \hat{v}_{m,k}^\dagger \Omega_{m}^\dagger \left( S_{k}^0 \right)^\dagger \eta_{m,k} \right\} < 0 \right) \\
&+ 2\text{Re}\left\{ v_{m,k} S_{k} \hat{v}_{m,k}^\dagger \left( q_{m,k} + \eta_{m,k} \right)^\dagger \right\} < 0 \right).
\end{align*}
\]
(52)

Hence, using [8, Theorem 4.2] it can be shown that the PEP can be expressed as (53). Applying Chernoff bound [20, Eq. (2.1.172)] over (53), we can obtain (54). By using Fischer’s matrix inequality [19, Eq. (11.8.1)], it can be shown that
\[
\begin{align*}
&\mathcal{W}_{m,k} S_{k}^0 \Omega_{m}^\dagger \hat{v}_{m,k}^\dagger \Omega_{m}^\dagger \left( S_{k}^0 \right)^\dagger \eta_{m,k} + v_{m,k} \hat{v}_{m,k}^\dagger \Omega_{m}^\dagger \Omega_{m}^\dagger \left( S_{k}^0 \right)^\dagger \eta_{m,k} \\
&\geq \lambda_{\min} (\mathcal{X}_{m,k}) \| v_{m,k} S_k \|^2.
\end{align*}
\]
(55)

From (54) and (55), we can obtain an upper bound for the error probability as shown in (18).

**APPENDIX VII**

**PROOF OF Theorem 1**

If all links are perturbed by the same carrier offset, i.e., \( \omega_m = \omega \), then (18) reduces into:
\[
\begin{align*}
\Pr \left\{ S_{k}^0 \rightarrow S_k \right\} &\leq \exp \left( -\frac{\lambda_{\min} (\mathcal{X}_{k})}{6\sigma^2} \sum_{m=1}^{n_s} \| v_{m,k} S_k \|^2 \right),
\end{align*}
\]
(56)

where \( \mathcal{X}_{k} = \mathcal{X}_{m,k} | \omega_m = \omega \) and \( v_{m,k} \) is the transmit power. We can average (56) over all the channels to find an average PEP upper bound as
\[
\begin{align*}
\mathbb{E}_H \left[ \Pr \left\{ S_{k}^0 \rightarrow S_k \right\} \right] &\leq \prod_{m=1}^{n_s} \left[ I + \left( \mathcal{W}_{m,k} - \mathcal{W}_{m,0} \right) \mathcal{K}_{m,k} \right] \left( \mathcal{W}_{m,k} - \mathcal{W}_{m,0} \right)^{-1},
\end{align*}
\]
(57)

where \( \mathcal{W}_{k} = \mathcal{W}_{m,k} | \omega_m = \omega \), \( \mathcal{K}_{m,k} = \mathcal{K}_{m,k} | \omega_m = \omega \), and \( \mathcal{K}_{m,k} \) is the estimate of \( \omega \). Let us assume that all the links between the transmit and receive antennas are perturbed with the worst carrier offset, i.e., \( \omega = \pm \pi \). From (12), it can be seen that the error in the estimate of \( \omega \) is \( \omega_{\text{est}} = \omega_m | \omega_m = \omega \). Therefore, the worst estimate will be \( \omega_{\text{est}} = \omega_m | \omega_m = \omega \). Substituting \( \omega_{\text{est}} = \omega_m | \omega_m = \omega \), and using the bound given in [8, Eq. (4.2.18)], we can obtain the relation given in (20).
Pr \{ S_k^0 \to S_k \} = Q \left( \frac{\sum_{m=1}^{n_r} \left( w_{m,k}^H S_m^0 \Omega_m \hat{S}_k^0 \hat{S}_k^0 \hat{S}_m^H v_{m,k}^H + v_{m,k}^H S_k^0 \hat{S}_k^0 \hat{S}_m^H v_{m,k}^H \right)}{\sqrt{3\sigma^2 \sum_{m=1}^{n_r} \| v_{m,k}^H S_k^0 \|^2}} \right). 

(53)

Pr \{ S_k^0 \to S_k \} \leq \exp \left( -\frac{3\lambda_{\min}(X_k)}{6\sigma^2} \left\| \left( S_k^0 - S_k \right)^T \Phi_k^T F^T \otimes I_{n_t} \right\|^2 \right). 

(54)

APPENDIX VIII
PROOF OF THEOREM 2

For the precoded DDOSTBC \( g_{m,k} = \exp (j\omega_m n_t (k-2))h_m F D_{k-2} \Omega_m \). Therefore, from (19), (56), and after some manipulations, it can be shown that

\[
\Pr \{ S_k^0 \to S_k \} \leq \exp \left( -\frac{\lambda_{\min}(X_k)}{6\sigma^2} \left\| \left( S_k^0 - S_k \right)^T \Phi_k^T F^T \otimes I_{n_t} \right\|^2 \right). \tag{58}
\]

Substituting \( \hat{E}_\omega = \exp (j\omega n_t) \), \( \hat{E}_{\omega,1} = \exp (j\omega / 2) \), and \( C_{k-1} = C_{k-1} = I_{n_t} \) into (58) and averaging (58) over \( \hat{H} \), we can obtain (23).

APPENDIX IX
FIRST ORDER DERIVATIVE OF OBJECTIVE FUNCTION IN (28)

From [13, Table 2], it can be shown that

\[
d_{\text{UBPEP}}(F) = \Tr \left\{ -I_{n_t} + \frac{\lambda_{\min}(X_k)}{12\sigma^2} \left( S_k^0 - S_k \right)^T F^T \otimes I_{n_t} \right\}^{-1} \left( S_k^0 - S_k \right)^T F^T \otimes I_{n_t} \times \left[ \left( S_k^0 - S_k \right)^T F^T \otimes I_{n_t} \right] \left( S_k^0 - S_k \right)^T F^T \otimes I_{n_t} \right\}. \tag{59}
\]

Next, using the relationship [13, Table 2]: \( dZ_0 Z = (dZ_0 Z) Z + Z_0 dZ \). [13, Eq. (22)], and the following properties of trace operator, vector operator, and Kroncker product [12, [12], [19]:

\[
\Tr \{ YZ \} = \vec{\vec{Y}} \vec{Z},
\]

\[
\Tr \{ YZ \} = \Tr \{ ZY \},
\]

\[
\vec{\vec{Y}U} = \left( U^T \otimes Y \right) \vec{Z},
\]

\[
(YZ \otimes UV) = (Y \otimes U)(Z \otimes V), \tag{60}
\]

we can find the first order derivative of the objective function as given in (29).

REFERENCES


