Differential Coding for Non-Orthogonal Space-Time Block Codes with Non-Unitary Constellations over Arbitrarily Correlated Rayleigh Channels

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

Abstract—In this paper, we propose a maximum likelihood (ML) decoder for differentially encoded full-rank square non-orthogonal space-time block codes (STBCs) using unitary or non-unitary signal constellations, which is also applicable to full-ranked orthogonal STBC (OSTBC). As the receiver is jointly optimized with respect to the channel and the unknown data, it does not require any knowledge of channel power, signal power, or noise power to decode the signal, and the decision is purely based on two consecutively received data blocks. We analyze the effect of channel correlation on the performance of the proposed system in Rayleigh fading channels. Assuming a general correlation model, an upper bound of the pair-wise error probability (PEP) of the differential OSTBCs is derived. An approximate bound of the PEP for the differential non-orthogonal STBCs is also derived. We propose a precoder designing criterion for differential STBC over arbitrarily correlated Rayleigh channels. Precoding improves the system performance over the correlated Rayleigh MIMO channels. Our precoded differential codes differ from the previously proposed precoder designs for differential OSTBC in the following ways: 1) We propose a precoder design for arbitrarily correlated Rayleigh channels, whereas the previous works consider only for transmit correlation. 2) The previous work is only applicable to the OSTBCs with PSK constellations, whereas our precoder is applicable to any type of full-rank square STBCs with unitary and non-unitary signal constellations.

Index Terms: Differential space-time codes, Non-unitary constellations, Precoded differential codes, Non-orthogonal codes.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) communications systems have drawn a lot of attention [1]–[3]. They provide high capacity and diversity gains as compared to single-input single-output (SISO) systems and, therefore, MIMO systems are suitable for the high data-rates required by the next generation wireless communication systems. In MIMO systems, space-time block codes (STBCs) are used to provide the maximum diversity [1]–[3]. However, the coherent detection of STBC requires perfect channel state information (CSI), which is gained by transmission of training data at the cost of reduced effective data-rate. Furthermore, this method is suitable only for the channels that remain constant for several symbol durations. If the channel remains constant for small number of symbol durations, this method does not work well, since the channel estimates might be very wrong when it is going to be used by the receiver. Differential STBC may be used to overcome this difficulty and improve the data-rate. In [4]–[6], the differential coding for square orthogonal STBCs (OSTBCs) with unitary constellations like QPSK, which have all the points on a circle, is suggested. These differential OSTBCs have a simple encoding structure and their decoding does not need knowledge of the channel coefficients. However, this differential system does not work with square non-orthogonal high data-rate STBCs [7]–[11] and non-unitary constellations like QAM. Whereas, high data-rate STBC and QAM constellations are very useful in providing higher data-rates.

In [12], [13], differential codes for non-unity constellations are suggested but they are applicable to the square orthogonal space-time structures and rectangular QAM constellations only. The differential codes, suggested for non-unity constellations [14], [15] are also applicable to square orthogonal STBCs only and their decoding requires knowledge of channel power, signal power, and noise power. In [14], the channel power is estimated by buffering many samples (normally \( L \geq 100 \)) of the received data and then finding the average autocorrelation of the received data sequence. These methods introduce a delay in making decisions and the performance of the receiver may degrade in the fastly varying channels, which do not remain constant for many symbol durations. All these existing differential schemes cannot be applied to the full-rank square non-orthogonal STBCs like high-rate STBCs, which use optimized non-unity constellations for maximizing the coding gain [7]–[9] or achieving the diversity multiplexing trade-offs [10], [11]. High rate STBCs are very useful for improving the performance of the wireless communications systems in the terms of data-rates and capacity.

In this paper, our main contributions are as follows:

1) We derive an ML decoder for differential STBCs with non-unity constellations like QAM, which follow the full-rank square non-orthogonal structures like high-rate STBCs. The proposed decoder also provides better performance for OSTBCs as compared to the previously suggested differential decoder. Moreover, the proposed decoder does not require any knowledge of channel, signal, or noise power for decoding of the space-time data, and it performs better than the differential STBCs of [14], [15].
2) We derive an upper bound of the pairwise error probability (PEP) of differential coding for full-rank square OSTBC with unitary and non-unitary constellations over arbitrarily correlated Rayleigh channels. In addition, an approximate bound of PEP of differential full-rank square non-orthogonal STBCs is also derived.

3) We propose precoded differential coding to reduce the effect of channel correlations by introducing a full memoryless precoder matrix in the transmitter. This precoded differential coding is applicable to full-rank square orthogonal and non-orthogonal STBCs with unitary or non-unitary constellation and arbitrarily correlated Rayleigh channels. Whereas, the previously proposed precoder for differential codes [16] only considers the special case of the Kronecker model with transmit correlation only and full-rank square OSTBCs with M-PSK constellations.

The rest of this paper is organized as follows: In Section II, the system model is introduced. The non-unitary constellations used by orthogonal and non-orthogonal STBCs are discussed and the proposed encoding and decoding of the differential STBC for non-unitary constellations are explained in Section III. Section IV derives upper bound of PEP of the differential OSTBCs and approximate bound of PEP of the differential non-orthogonal STBCs with unitary and non-unitary constellations over arbitrarily correlated Rayleigh channels. The precoder is designed in Section V to improve the performance of the system over correlated Rayleigh channels. Simulation results are discussed in Section VI, and Section VII contains some conclusions.

**Notation:** We have used the following notation throughout the paper: Uppercase bold letters $X$ are used for matrices, lowercase bold letters $x$ are used for vectors, $x, X$ denote scalar variables, $(\cdot)^H$ means the Hermitian of a matrix, $(\cdot)^T$ is used for the transpose of a matrix, $\mathbb{E} \{ \cdot \}$ represents expectation, $\Pr\{\cdot\}$ is used for probability, $\text{Tr}\{\cdot\}$ is the trace operator, Frobenius norm of a matrix is depicted by $\|\cdot\|_F$, $I$ stands for identity matrix, $|\cdot|$ denotes the determinant operator, vec $(\cdot)$ is vectorization operator, which stacks the columns of matrix in a long column vector, rank $(\cdot)$ denotes the rank of the matrix, $\max\{\ldots\}$ chooses larger out of the two quantities, and Re $(\cdot)$ is used as the real operator, which returns only the real part of the matrix it is applied to.

**II. System Model**

Consider a MIMO system with $n_t$ transmit and $n_r$ receive antennas as shown in Fig. 1. Let $H$ be a $n_r \times n_t$ channel matrix and $S_k$ be the $n_t \times n$ data matrix transmitted at block $k$. The received data $n_r \times n$ matrix will be

$$Y_k = HS_k + Q_k,$$

where $Q_k$ is an $n_r \times n$ matrix, containing 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 $H$ is assumed constant over the transmission period of at least two consecutive code matrices $(S_{k-1}$ and $S_k)$.

**A. Model of Correlated Rayleigh Channels**

We assume flat block-fading correlated Rayleigh fading channel model [17]. Let the channel $H$ have zero mean, complex Gaussian circularly distribution with positive semi-definite autocorrelation given by $R = \mathbb{E} \{ \text{vec} (H) \text{vec}^H (H) \}$ of size $n_r n_t \times n_r n_t$. A channel realization of the correlated Rayleigh channels can be found by vec $(H) = R^{1/2} \text{vec} (H_w)$, where $R^{1/2}$ is the unique positive semi-definite matrix square root [18] of $R$ and $H_w$ is $n_r \times n_t$ matrix consisting of complex circular Gaussian distributed elements with zero-mean and unit variance.

**Kronecker Model:** A special case of the model given above is the Kronecker model, which can be represented as [17]

$$R = R_{T}^{T} \otimes R_{r},$$

where $R_r$ is the $n_r \times n_r$ receive correlation matrix and $R_T$ is the $n_t \times n_t$ transmit correlation matrix. This model does not always render the multi-path structure correctly, and it might introduce artifact paths that are not present in the underlying measurement data [19]. The Kronecker model can underestimate the mutual information of the MIMO channel systematically. However, the generalized model that we used throughout this work considers that the receive (or transmit) correlation depends on at which transmit (or receive) antenna the measurements are performed. Hence, the generalized model depicts the behavior of MIMO channel more accurately than the Kronecker model. The previously proposed differential code based on eigen-beamforming precoder [16] only considers the Kronecker model with transmitter correlation and no receiver correlation and is not applicable for the general correlation model assumed here.
better than the OSTBCs. Fig. 2 shows the constellations of rate-2 code of [7] and 8-QAM STBC of [9], which utilize 4-QAM and 8-QAM, respectively, as the basic signal constellations to obtain the optimized constellations. These codes not only improve the spectral efficiency but also achieve the diversity-multiplexing trade-offs [21]. But these codes are non-orthogonal and the previously reported differential coding schemes [4]–[6], [12]–[15] cannot be applied for them.

C. Differential Encoding for Non-Unitary Constellations

Let us first, consider a case, when channel is uncorrelated, i.e., \( R = I_{n_t n_r} \) and no precoding \( F = I_{n_t} \) in Fig. 1. Let \( V_{k-1} \) be an \( n_t \times n_t \) non-orthogonal STBC matrix transmitted at block \( k - 1 \). The received data matrix at block \( k - 1 \) will be

\[
Y_{k-1} = HV_{k-1} + Q_{k-1}. \tag{4}
\]

In the next block \( k \), we transmit the \( V_k \) matrix as shown in Fig. 1

\[
V_k = \frac{1}{\beta_{k-1}} V_{k-1} C_k, \tag{5}
\]

where \( C_k \) is a full-rank square non-orthogonal STBC matrix of size \( n_t \times n_t \), which consists of data from the optimized non-unitary constellation (see Subsection III-B), \( \beta_{k-1} \) is a normalization factor, which limits the peak power of \( V_k \). The choice of \( \beta_{k-1} \) depends upon the specific code. Let \( s_1, s_2, \ldots, s_{n_s} \) be a block of \( n_s \) optimized non-unitary symbols transmitted at block \( k - 1 \) through differential non-orthogonal STBC \( V_{k-1} \), then one choice of \( \beta_{k-1} \) could be

\[
\beta_{k-1} = \left( a \sum_{i=1}^{n_s} |s_i|^2 \right)^{1/2}, \tag{6}
\]

where \( a \) is a constant, which depends upon the non-orthogonal STBC and number of transmit antennas. \( a = 1/2 \) for rate-2 STBCs of [7, Ex. (6)] and STBCs of [9, Eq. (5)] for two transmit antennas and QPSK constellation, and \( a = 1/5 \) for Golden codes [10, Eq. (10)] with any-QAM used for obtaining the optimized constellation and two transmit antennas. Another choice of \( \beta_{k-1} \) could be a constant, to put an average power constraint. The received data at block \( k \geq 1 \) will be

\[
Y_k = HV_k + Q_k = HV_{k-1} \tilde{C}_k + Q_k, \tag{7}
\]

where \( \tilde{C}_k = C_k/\beta_{k-1} \). Here, it is assumed that the channel \( H \) remains stationary for at least two consecutive code matrix transmission. In general, we transmit the following data matrix at any block \( p \geq 1 \)

\[
V_p = V_{p-1} \tilde{C}_p = V_{p-2} \tilde{C}_{p-1} \tilde{C}_p = V_0 \tilde{C}_1 \tilde{C}_2 \tilde{C}_3 \cdots \tilde{C}_p, \tag{8}
\]

therefore, this is a normal form of differential encoding (similar to that of unitary constellations) where the current transmitted code matrix has information about the previously transmitted code matrices.
D. Decoding of Non-Orthogonal Differential Space-Time Code with Non-Unitary Constellations

It can be verified from (4) and (7), that if channel \( H \) and \( V_k \) are known,

\[
f(\vec{Y}_k | H V_k) = \frac{1}{\pi^{n_r n_t} \det(\Psi)} \exp\left(-\|\vec{Y}_k - \vec{H} \vec{V}_k\|^2\right),
\]

\[
\text{vec}(Y_k)^T \Psi^{-1} \text{vec}(Y_k) - \text{vec}(H V_k))
\]

where \( \Psi = \mathbb{E} \left[\text{vec}(Q_k) \text{vec}^T(Q_k)\right] \) is the covariance matrix of \( Q_k \). Since \( \text{vec}(Q_k) \) is AWGN, \( \Psi = \sigma^2 I_{n_r n_t} \).

Let \( y = [\text{vec}^T(Y_k), \text{vec}^T(Y_{k-1})]^T \) be a vector consisting of two consecutive blocks at the receiver. When \( H, V_{k-1}, \beta_{k-1}, \) and \( C_k \) are known,

\[
f(y | H, V_{k-1}, \beta_{k-1}, C_k) = \frac{1}{\pi^{2n_r n_t}} \left(\frac{\sigma^2}{2}\right)^{2n_r n_t} \times \exp\left(-\frac{1}{\sigma^2} \sum_{k=1}^{T} \|\vec{Y}_t - \vec{H} \vec{V}_t\|^2\right).
\]

In order to find a maximum-likelihood (ML) estimate of the unknown data \( C_k \), we need to maximize (10). If we assume that \( \sigma^2 \) is known, then the following metric is first minimized with respect to \( w.r.t. \) all unknown quantities \( H, V_{k-1}, \beta_{k-1}, \) and, subsequently, over \( C_k \):

\[
\Gamma = \sum_{l=k-1}^{k} \|\vec{Y}_t - \vec{H} \vec{V}_t\|^2_F, \quad k \geq 1,
\]

which is equivalent to maximization of (10). From the property of the Frobenius norm [22], (11) can be alternately expressed as

\[
\Gamma = \|Y_k - HV_{k-1} \hat{C}_k\|^2_F + \|Y_{k-1} - HV_{k-1}\|^2_F.
\]

It can be seen from (12) that the ML metric is obtained from maximization of joint p.d.f. of \( \text{vec}(Y_k) \) and \( \text{vec}(Y_{k-1}) \) and it consists of two different norms. Whereas, the differential codes of [12]–[15] use a least square decoder which assumes \( Y_{k-1} \) as an estimate of \( HV_{k-1} \) and finds the estimate of the OSTBC data by minimizing the following metric:

\[
\|Y_k - Y_{k-1} \hat{C}_k\|^2_F.
\]

However, this least square decoder does not work well in the case of non-orthogonal STBCs and its performance is very poor as compared to the optimal decoder. Therefore, we will proceed with the ML metric obtained in (12). Assuming that \( \hat{C}_k \) is known, the ML metric in (12) can be minimized \( w.r.t. \) \( HV_{k-1} \) by taking the derivative of (12) \( w.r.t. \) \( HV_{k-1} \) [23] and equating it to zero as

\[
HV_{k-1} = \left(Y_k \hat{C}_k^H + Y_{k-1}\right) \left(\hat{C}_k \hat{C}_k^H + I\right)^{-1}.
\]

After substituting the value of \( HV_{k-1} \) from (13) into (12) we get

\[
\Gamma(\hat{C}_k) = \|Y_k - \left(Y_k \hat{C}_k^H + Y_{k-1}\right) \left(\hat{C}_k \hat{C}_k^H + I\right)^{-1} \hat{C}_k\|^2_F

+ \|Y_{k-1} - \left(Y_k \hat{C}_k^H + Y_{k-1}\right) \left(\hat{C}_k \hat{C}_k^H + I\right)^{-1} \hat{C}_k\|^2_F.
\]

In order to find the ML estimate of \( C_k \), (14) can be jointly minimized \( w.r.t. \beta_{k-1} \) and \( C_k \). The physical interpretation of (14) is as follows: At a given time the receiver has information about two consecutively received data matrices, i.e., \( Y_k \) and \( Y_{k-1} \). Hence, in contrast to the conventional differential decoder it tries to minimize the euclidean distance between \( Y_k \) and \( HV_{k-1} \hat{C}_k \), and \( Y_{k-1} \) and \( HV_{k-1} \), see (12). As \( HV_{k-1} \) is not known, it first finds an estimate of \( HV_{k-1} \) through (13) in terms of \( \hat{C}_k \). It can be seen from (13) that the estimate of \( HV_{k-1} \) depends upon \( \hat{C}_k \), therefore, from (14) the receiver tries to utilize the available information about \( Y_k \) and \( Y_{k-1} \) in a good way as possible. It will be shown later in (17) and (18) that for OSTBC, the ML decoder reduces into a single norm which minimizes the euclidean distance between \( Y_k \) and \( Y_{k-1} \hat{C}_k \), and \( Y_k \) and \( Y_{k-1} \hat{C}_k \) for non-unitary and unitary constellations, respectively. If the receiver knows \( \beta_{k-1} \) perfectly, (14) can be minimized \( w.r.t. \) \( C_k \) to find the ML estimate of \( C_k \). Therefore, the ML detection of \( C_k \) leads to

\[
\hat{C}_k = \arg \min_{C_k \in \Xi} \Gamma(\hat{C}_k),
\]

where \( \Xi \) is the set of all full-rank square STBC matrices, consisting of symbols belonging to the finite non-unitary constellation.

Let \( s_i \in M\)-PSK constellation, then by setting \( |s_i|^2 = 1/(an_s) \) it can be seen from (6) that \( \beta_{k-1} = 1 \), and ML decoder of (15) reduces into

\[
\hat{C}_k = \arg \min_{C_k \in \mathcal{A}} \left\{\|Y_k - (Y_k \hat{C}_k^H + Y_{k-1}) (C_k \hat{C}_k^H + I)^{-1} C_k\|^2_F + \|Y_{k-1} - (Y_k \hat{C}_k^H + Y_{k-1}) (C_k \hat{C}_k^H + I)^{-1} C_k\|^2_F\right\},
\]

where \( \mathcal{A} \) is the set of all full-rank square STBC matrices, consisting of symbols belonging to the \( M\)-PSK constellation.

If \( C_k \) is a full-rank square OSTBC with \( M\)-QAM or \( M\)-PAM constellation, it can be shown after some manipulations that (14) reduces into the following form:

\[
\Gamma(\hat{C}_k) = \left(\frac{\beta_{k-1}}{\beta_{k-1} + \beta_{k-1}}\right)^2 \|Y_k - Y_{k-1} \hat{C}_k\|^2_F.
\]

If \( C_k \) is a full-rank square OSTBC with \( M\)-PSK constellation, then \( \beta_k = \beta_{k-1} = 1 \) and (17) reduces into

\[
\Gamma(\hat{C}_k) = \frac{1}{4} \|Y_k - Y_{k-1} \hat{C}_k\|^2_F,
\]

which is an optimal decoder of \( C_k \) with \text{unitary} constellations, already obtained in the literature [20, Eq. (9.6.18)]. It is clear from (14), (15), (16), (17), and (18), that the proposed decoder is applicable to unitary, and standard and optimized non-unitary constellations, and orthogonal and non-orthogonal full-rank space STBCs. Hence, our differential decoder is more general than the previously suggested differential decoder for full-rank square OSTBCs with unitary or non-unitary constellations. In addition, the proposed decoder does not need any information about the channel power, signal power, or noise power for detection of \( C_k \).
E. Application of the Proposed ML Decoder to Differential Quasi-OSTBC

Quasi-OSTBC (QOSTBC) has been proposed in [24]–[28]. The QOSTBC is a full-rank and square non-orthogonal STBC with pairwise decoding complexity, i.e., symbols can be decoded in pair. Differential schemes for QOSTBC have been reported in [29], [30]. Recently, an optimized differential modulation scheme based on the QOSTBC of [27], [28] was proposed in [31]. In [31], an optimized constellation is designed such that the differential encoding and suboptimal decoding1 proposed in [12] for unitary orthogonal matrices can be applied to QOSTBC and full diversity, and rate-1 can be achieved with maximized coding gain. However, as we have derived ML decoder for differentially encoded non-orthogonal STBCs, therefore, it is now possible to apply the simple encoding of (5) to the QOSTBC and the receiver can use the ML decoder of (14). We will show the comparison of the performance of the proposed differential encoding and ML decoding for QOSTBC of [25] with the same rate optimized differential QOSTBC scheme of [31] in Subsection VI-B. In addition, the proposed differential scheme is more general than the differential schemes in [29]–[31], since we also consider the effect of the channel correlation on the performance of the differential system and propose a precoder design for improving the system performance.

IV. PEP PERFORMANCE ANALYSIS OF THE PRECODED DIFFERENTIAL STBC

In this section, we consider the effect of channel correlation on the system performance. So far, the effect of correlated channels for differential STBCs has not been studied in detail [16]. We present theoretical analysis, which considers the effect of channel correlation on the PEP performance of the differential communication system. The idea is as follows: To improve the performance, we can apply precoding before the differential signals are transmitted as done in the coherent non-differential case [32], [33]. The block diagram of the precoded differential coding scheme is shown in Fig. 1. Here, $F$ is the $n_t \times n$ precoder matrix multiplied with the differential code matrix before transmission. The precoder matrix does not change the ML decoding as the precoding matrix can be absorbed into the channel as $H' = HF$. Hence, the ML decoding of $C_k$ in a precoded system can still be done by (15) and (16) without the knowledge of the precoder in the receiver.

A. PEP Bound for Precoded Differential STBC

We are using a precoder before transmission of the differential STBC as shown in Fig. 1, to reduce the effect of channel correlation and improve the performance of the system. Hence, the $n_r \times n$ received data matrices at the $(k-1)$-th and the $k$-th block, corresponding to the precoded differential STBC are

$$Y_{k-1} = HFV_{k-1} + Q_{k-1}, \quad (19)$$

where $F$ is a memoryless precoding matrix of size $n_r \times n$, $C_k$ of size $n \times n$ is the full-rank square non-orthogonal STBC matrix, and $V_{k-1}$ and $Q_k$ are differential code matrices of size $n \times n$. We may write (19) and (20) as follows:

$$Y_k = [Y_k, Y_{k-1}]_{n_r \times 2n} = HFV_{k-1} \left[ \frac{C_k}{\beta_{k-1}}, I \right]_{n \times 2n} + [Q_k, Q_{k-1}]_{n_r \times 2n}, \quad (21)$$

where $[ \cdot ]_{l \times m}$ indicates the size of the matrix inside the parenthesis as $l \times m$. We can write (21) alternatively as

$$\tilde{Y}_k = [Y_k, Y_{k-1}]_{n_r \times 2n} = HFV_{k-1} \left[ C_k, \beta_{k-1} I \right]_{n \times 2n} + [Q_k, Q_{k-1}]_{n_r \times 2n}, \quad (22)$$

where $V_{k-1} = \tilde{V}_{k-2} C_{k-1}$ has size $n \times n$ and is defined as:

$$V_{k-1} \triangleq \frac{V_{k-1} - V_0}{\beta_{k-1}} = [C_k, \beta_{k-1} I]_{n \times 2n}.$$  \quad (23)

The differential detection may be treated as a problem of detecting $\Phi_k = [C_k, \beta_{k-1} I]_{n \times 2n}$ with an unknown channel $HFV_{k-1}$. It can be proved with the help of the results given in [34, Section III] that

$$E_{H} \left[ \Pr \{ C_k \rightarrow C_k \} \right] \leq \left| a_{11} \right| + \frac{1}{\lambda_{2}^{\frac{1}{2}}} \left( F^\top \tilde{V}_{k-1} (\Phi_k^0)^\top \right) \Pi_{\Phi_k^0}^{\frac{1}{2}} \left( F^\top \tilde{V}_{k-1} (\Phi_k^0)^\top \right) \Pi_{\Phi_k^0} \left( F^\top \tilde{V}_{k-1} (\Phi_k^0)^\top \right) \Pi_{\Phi_k^0}^{-1} \left( F^\top \tilde{V}_{k-1} (\Phi_k^0)^\top \right) \Pi_{\Phi_k^0}^{\frac{1}{2}} \right|^{-1}, \quad (24)$$

where $\Phi_k^0 = [C_k^0, \beta_{k-1} I]_{n \times 2n}$ and $\Pi_{\Phi_k^0} = I_{2n} - \Phi_k^0 (\Phi_k^0)^\top \Phi_k^0 (\Phi_k^0)^\top \Pi_{\Phi_k^0}^{-1} \Phi_k^0 (\Phi_k^0)^\top$ is the orthogonal projector onto the complement of the range space of $\Phi_k^0$. It can be seen from (24) that the PEP of the differential non-orthogonal STBC depends upon the product matrix $V_{k-1}$, which makes the PEP bound of (24) impractical for precoder design. As $V_{k-1}$ is a product of $k$ normalized matrices which depends upon $k$, therefore, $V_{k-1}$ will be different at different block $k$ and one need to determine the value of $V_{k-1}$ which maximizes the PEP for every block $k$ which is computationally complex. If we design a precoder based on this bound, then the precoder matrix will depend upon $V_{k-1}$, i.e., in turn it depends upon the block $k$. This means that the precoder matrix $F$ will be different for different block transmissions, hence, the effective channel will be block varying and the proposed differential scheme will fail. Therefore, for designing the precoder we need to form the optimization problem such that it does not depend upon $V_{k-1}$. Further, it can be seen from (24) that the term $(\Phi_k^0)^\top \Pi_{\Phi_k^0}^{-\frac{1}{2}} \Phi_k^0 (\Phi_k^0)^\top$ depends upon STBC matrices $C_k$ and $C_k^0$, and $\beta_{k-1}$, which can be a vector norm $\beta_{k-1} = \frac{1}{2} \sqrt{\sum_{i=1}^{n} |s_i|^2}$ containing elements from a fixed finite QAM constellations or a fixed constant to satisfy the average power constraint. Therefore, fixed values of $C_k, C_k^0$, and $\beta_{k-1}$ can be obtained which maximize PEP irrespective of $k$.

In order to obtain a more practical bound of PEP we should average (24) over $V_{k-1}$. Applying Minkowski’s determinant

$$Y_k = HFV_{k} + Q_k = HFV_{k-1} \frac{C_k}{\beta_{k-1}} + Q_k, \quad (20)$$
inequality [35, Part II, Eq. (4.1.8.6)] to (24), we can further upper bound the PEP of differential STBC as

$$E_{H,V_k-1} \left[ \Pr \left\{ C_k^0 \to C_k \right\} \right] \leq E_{V_k-1} \left[ \left( 1 + \frac{1}{4\delta^2} \left| \left( F^* \right)^* \right| F^T \right. \right.$$  
$$\times \left( \Theta_k^0 \right)^* \Pi_{\phi_k^*}^\perp \left( \Theta_k^0 \right)^T \left. F^T \otimes I_{n_r} \right) R \left| \frac{1}{\sigma_r^2} \right|^{-n_{inr}} \right],$$  
(25)

By using the property of the determinant $|AB| = |BA| = |A| |B| |A|$, (25) can be written as

$$E_{H,V_k-1} \left[ \Pr \left\{ C_k^0 \to C_k \right\} \right] \leq E_{V_k-1} \left[ \left( 1 + \frac{1}{4\delta^2} \left| \left( F^* \right)^* \right| F^T \right. \right.$$  
$$\times \left( \Theta_k^0 \right)^* \Pi_{\phi_k^*}^\perp \left( \Theta_k^0 \right)^T \left. F^T \otimes I_{n_r} \right) R \left| \frac{1}{\sigma_r^2} \right|^{-n_{inr}} \right],$$  
(26)

In the case of OSTBC utilizing non-unitary constellation, it can be shown by using the unitary property of normalized OSTBC that $\tilde{V}_{k-1} \tilde{V}_{k-1} = I_n$. Therefore, from (26) the upper bound of PEP (UBPEP) for differential OSTBC with non-unitary constellation can be given as

$$UBPEP_0 = \left( 1 + \frac{1}{4\delta^2} \left| \left( F^* \right)^* \right| F^T \right. \right.$$  
$$\otimes I_{n_r} \left. \right) R \left| \frac{1}{\sigma_r^2} \right|^{-n_{inr}},$$  
(27)

which is free from $\tilde{V}_{k-1}$. By using the inequality $1 + a \geq a$, we can obtain UBPEP of non-orthogonal STBC from (26) as

$$UBPEP_{NO} = E_{V_k-1} \left[ \left| \tilde{V}_{k-1} \tilde{V}_{k-1} \right|^{-1} \right] \left( \left( F^* \right)^* \Pi_{\phi_k^*}^\perp \left( F^* \right)^T \left. \right. \right.$$  
$$\otimes I_{n_r} \left. \right) R \left| \frac{1}{\sigma_r^2} \right|^{-n_{inr}}.$$  
(28)

It can be seen from (28) that $E_{V_k-1} \left[ \left| \tilde{V}_{k-1} \tilde{V}_{k-1} \right|^{-1} \right]$ is a constant which does not depend upon $F$, therefore, the PEP can be minimized with respect to (w.r.t.) $F$ by only maximizing the following term w.r.t. $F$:  
$$\left( \left( F^* \right)^* \Pi_{\phi_k^*}^\perp \left( F^* \right)^T \otimes I_{n_r} \right) R,$$  
which is free from $\tilde{V}_{k-1}$. Similarly, it can be seen from (27) that in the case of OSTBC also PBEP can be minimized by maximizing  
$$\left( \left( F^* \right)^* \Pi_{\phi_k^*}^\perp \left( F^* \right)^T \otimes I_{n_r} \right) R.$$  

The goal is to design a precoder $F$ that maximizes the PEP with the trivial precoder $F = \sqrt{\frac{P}{2\delta^2}} I_{max(n_t,n)} |n_t \times n_t$ and invertible channel correlation matrix. Here, $I_{max(n_t,n)}$ is a matrix of the size $n_t \times n_t$, taking the first $n_t$ rows and first $n$ columns from an identity matrix $I_{max(n_t,n)}$, where max ($n_t, n$) returns the maximum value of $n_t$ and $n$. Therefore, from (27), (28), and (29) $\Theta$ can be obtained as

$$\Theta = \min \left\{ \phi_k^0 \right\} \Pi_{\phi_k^*}^\perp \left( \phi_k^0 \right)^T \left( \Theta \otimes I_{n_r} \right) R,$$  
(30)

where $\phi_k^0 = [C_k \beta_k - I_n]_{n_t \times n_t}$, $\phi_k = [C_k \beta_k - I_n]_{n_t \times n_t}$, $\phi_k^0$ and $\phi_k$ are two different STBC matrices, and $\beta_k - 1$ is either a vector norm of $n_s$ symbols belonging to the unitary or non-unitary constellation, or a fixed constant.

V. DESIGN OF PRECODER FOR DIFFERENTIAL STBCS

In this section, we will discuss the design of the precoder for differential STBCs. 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 differential STBCs using unitary or non-unitary constellations and full-rank square orthogonal or non-orthogonal STBC.

A. Problem Formulation

By using the properties of full-rank square orthogonal and non-orthogonal STBC [7]–[11], it can be proved that the average power transmitted per orthogonal or non-orthogonal STBC block is

$$E \left\{ C_k C_k^H \right\} = a_{in} \sigma_n^2 \mathbf{I}_n,$$  
(31)

where $\sigma_n^2$ is the average power of each symbol in the $n_s \times 1$ data vector $s$ to be encoded into non-orthogonal STBC. The average power constraint on the transmitted block $S_k = FV_k$ can be expressed as

$$a_{in} \sigma_n^2 \text{Tr} \left\{ FF^H \right\} = P,$$  
(32)

where $P$ is the average power used by the transmitted block $S_k$. The goal is to design a precoder $F$ such that an upper bound of the pairwise error probability (UBPEP) is minimized under the constraint over the average power on the transmitted block $S_k$ given in (32). It can be seen from (27) and (28) that UBPEP can be minimized by
maximizing \( \left| \left( F^* \left( \Phi_k^0 \right)^\dagger \Pi_{\Phi_k^0} \left( \Phi_k^0 \right)^T F T \otimes I_{n_r} \right) \right| \). However, from the property of the determinant \( |AB| = |B| |A| \) it can be shown that the trivial precoder \( F = \sqrt{\frac{P}{m_n \sigma^2}} I_{\max \{n_r, n_t\}} \) is a solution of the optimization problem of maximizing \( \left| \left( F^* \left( \Phi_k^0 \right)^\dagger \Pi_{\Phi_k^0} \left( \Phi_k^0 \right)^T F T \otimes I_{n_r} \right) \right| \) under the power constraint of (32). Therefore, for obtaining a non-trivial precoding solution for differential STBCs, we need to modify the optimization problem. By using the determinant inequality \( |I + A| \geq |A| \) [20, Eq. (A.4.26)], it can be shown that

\[
\left| \left( F^* \left( \Phi_k^0 \right)^\dagger \Pi_{\Phi_k^0} \left( \Phi_k^0 \right)^T F T \otimes I_{n_r} \right) \right| 
\leq \left| I_{n_r n_t} + \left( F^* \left( \Phi_k^0 \right)^\dagger \Pi_{\Phi_k^0} \left( \Phi_k^0 \right)^T F T \otimes I_{n_r} \right) \right|. \tag{33}
\]

We propose to maximize the upperbound given in the right side of (33) for designing the precoder for differential STBC. Apparently, utilizing the upper bound of (33) for minimizing the PEP may result into a suboptimal precoding solution and obtaining an optimal precoder is still an open research problem. However, it is shown by the simulation results in Subsections VI-D and VI-E that the proposed precoder is able to provide significant performance improvement to the differential STBCs over correlated MIMO channels and better results are obtained with the precoder found from the right hand side of (33) as compared to the trivial precoder. We can express the optimization problem as [32], [33], [37], [38]

\[
\max_{\{F \in \mathbb{C}^{n_r \times n_t}\}} \ln \left| I_{n_r n_t} + \left( F^* \left( \Phi_k^0 \right)^\dagger \Pi_{\Phi_k^0} \left( \Phi_k^0 \right)^T F T \otimes I_{n_r} \right) \right| \] 
subject to \( \text{Tr} \left\{ FF^H \right\} = \frac{P}{m_n \sigma^2} \). \tag{34}

It can be observed that the use of natural logarithm does not change the nature of the objective function because \( \ln(\cdot) \) is a monotonically increasing function. It can be seen from (34), that a closed-form precoder is difficult for the general case of arbitrary correlation. Hence, we focus on fast-converging numerical method to design the precoder in the next subsection.

B. Precoder Design

The constrained minimization problem of Subsection V-A can be converted into an unconstrained minimization problem by introducing a Lagrange multiplier \( \mu' \):

\[
\mathcal{L}(F) = \ln \left| I_{n_r n_t} + \left( F^* \left( \Phi_k^0 \right)^\dagger \Pi_{\Phi_k^0} \left( \Phi_k^0 \right)^T F T \otimes I_{n_r} \right) \right| 
+ \mu' \text{Tr} \left\{ FF^H \right\}. \tag{35}
\]

Since the objective function should be maximized, \( \mu' < 0 \). It can be shown after some manipulations that [23] the necessary conditions for optimality can be written as:

\[
\text{vec}^T (F) = \mu D_F \ln \left| I_{n_r n_t} + \left( F^* \left( \Phi_k^0 \right)^\dagger \Pi_{\Phi_k^0} \left( \Phi_k^0 \right)^T F T \otimes I_{n_r} \right) \right|, \tag{36}
\]

where \( D_F \) is the first order derivative with respect to \( F^* \) [23], and \( \mu \) is a scalar chosen such that the power constraint in (32) is satisfied. The derivative can be found with the help of results in [23]. We can summarize the results of the derivative as follows [23]:

\[
D_F \ln \left| I_{n_r n_t} + \left( F^* \left( \Phi_k^0 \right)^\dagger \Pi_{\Phi_k^0} \left( \Phi_k^0 \right)^T F T \otimes I_{n_r} \right) \right| = \text{vec}^H \left( R^H \Sigma^H \right) \Pi, \tag{37}
\]

where \( R = \left( \left( \Phi_k^0 \right)^\dagger \Pi_{\Phi_k^0} \left( \Phi_k^0 \right)^T F T \otimes I_{n_r} \right) R, \Pi = \left( I_{n_r} \otimes K_{n_r n_t} \right) I_{n_r n_t} \otimes \text{vec}(I_{n_r}), \Sigma = \left( I_{n_r n_t} + \left( \left( \Phi_k^0 \right)^\dagger \Pi_{\Phi_k^0} \left( \Phi_k^0 \right)^T F T \otimes I_{n_r} \right) \right)^{-1}, \)

and \( K_{n_r n_t} \) is the commutation matrix of size \( n_t n_r \times n_t n_r \) [22].

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 I.

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initialization</td>
</tr>
<tr>
<td>2</td>
<td>Precoder Optimization</td>
</tr>
</tbody>
</table>

VI. SIMULATION RESULTS

The simulations are performed with 4, 8, 16, and 64 standard QAM and optimized non-unitary constellations, FSK constellations, \( n_t \in \{2, 3, 4\} \), and \( n_r \in \{1, 2\} \). The channel is assumed to be circular symmetric complex Gaussian with zero mean, and with uncorrelated coefficients except in Subsections VI-D and VI-E, where we consider the channel to be correlated. The channel is also assumed stationary over the transmission period of two consecutive code matrices. The initialization matrix is kept as identity, i.e., \( V_0 = I_2 \) and assumed to be perfectly known to the receiver. The SERs at each SNR level are obtained through 10⁸ Monte-Carlo simulation runs.
A. Performance of the Proposed Differential Scheme with OSTBCs

Fig. 3 shows the performance of the proposed differential STBC with OSTBC (Alamouti code), $n_t = 2$, and $n_r = 1$. The symbols are chosen from 8, 16, and 64 standard QAM and PSK constellations. Fig. 3 shows that our differential code works well with OSTBCs. The performance of the proposed differential code is also compared with unitary differential code [4], [5] with similar rate PSK constellations. The proposed code works better than unitary differential codes in the case of higher order QAM like 16 and 64-QAM. However, there is no performance gain in the case of 8-QAM compared to the 8 PSK as the two lower curves (8-QAM and 8-PSK) in Fig. 3 are almost on the top of each other.

B. Performance of the Proposed Differential Scheme with Non-Orthogonal STBCs

Fig. 4 shows the performance of the proposed differential scheme with $n_t = 2$, $n_r = 1$, different high-rate STBCs [7], [9], [10], and the ML decoder of (15). The simulations are performed with rate-2 STBCs from field extension [7], 4-QAM and 8-QAM STBCs based on number theory [9], and 4-QAM Golden codes of [10]. The codes of [7] use the symbols from higher order non-unitary optimized 16-point constellation, extended over QPSK. Similarly, the 4-QAM and 8-QAM STBCs of [9] use 16-point optimized constellation obtained from 4-QAM and 64-point optimized constellation obtained from 8-QAM, respectively, and 4-QAM Golden codes of [10] use optimized 16-point constellation obtained from 4-QAM. The performance of the proposed differential scheme for these codes is compared with the same rate PSK constellations, where in place of the optimized constellation, the high-rate STBCs [9] use the symbols drawn from PSK constellation. From Fig. 4, it is seen that differential coding for non-unitary constellations performs better than the differential coding for unitary constellations using the same rates. Moreover, the proposed differential scheme works better in the case of Golden code than the other high data-rate codes. In addition, our differential code is working well for unitary and non-unitary constellations. We have also plotted the SER performance of a suboptimal decoder for 4-QAM Golden codes in Fig. 4. The suboptimal decoder assumes that the previously transmitted data is a noisy estimate of the channel and tries to minimize the following metric to find the estimate of current data: $\|Y_k - Y_{k-1} \hat{C}_k\|_F^2$. From Fig. 4, it can be seen that the suboptimal decoder does not perform well as compared to the ML decoder for Golden code and at high SNR it looses diversity.

In Fig. 5, we have compared the performance of the differential code of [39] with the proposed differential scheme with STBC of [7] for three transmit and one receive antennas. The differential code of [39] uses the unitary matrices given in [39, Subsection V-C] for three transmit antennas with data-rate of 1.9924 bits/sec/Hz. Whereas the STBC of [7] utilizes the full rank code for three transmit antennas given by [7,
Eq. (6)] with the indeterminate \( z = e^{j.05} \) and data-rate of 2 bits/sec/Hz. From Fig. 5, it can be seen that the STBC of [7] with the proposed differential encoding and decoding not only provides better data-rate but also performs better than the differential code of [39] for all values of SNR.

The performance plot of the proposed differential scheme for the QOSTBC of [25] with four transmit and one receive antennas is shown in Fig. 6. The symbols are chosen from two different QPSK constellations which have a relative rotational angle of \( \phi \) between them. The value of \( \phi \) is kept as 0.5344 as suggested in [25, Fig. 1]. We have also plotted the performance of the same rate optimized differential QOSTBC of [31] which utilizes the symbols chosen from the four point optimized constellation shown in [31, Fig. 4 (a)]. From Fig. 6 it can be seen that the proposed differential coding works approximately 2 dB better than the differential QOSTBC of [31] at SER=10^{-4}.

### C. Comparison with Conventional Differential Codes for QAM Constellations

Fig. 7 shows the comparison of the proposed differential code with the differential codes of [14] and [15]. The simulations are done for 16-QAM constellation, the Alamouti code, \( n_t = 2 \), and \( n_r = 1 \). The decoding of differential code of [14] depends upon the knowledge of the channel power. The method of channel power estimation, suggested in [14], requires the channel to be constant for more than 100 symbol durations (\( L \geq 100 \)), for obtaining good estimate of channel power. If the channel remains constant for less than 100 symbol durations, then the performance of [14], degrades substantially as shown in Fig. 7. We have shown the performance of [14] with \( L=2 \) and \( L=6 \). Since our code is independent of channel power knowledge, it works equally well with any length of channel if it is constant over at least two consecutive block channel uses. The code in [15] uses a soft decoder with a complex two level decoding structure.

Nevertheless, the performance of this code is 3.7 dB poorer than the coherent detection, even when the receiver perfectly knows the signal power and the noise power. The proposed code performs better than the differential STBC of [15] as well.

### D. Performance of Precoded Differential OSTBC with Unitary Constellations

For making a comparison with the previously proposed precoded differential system [16] we have plotted the SER plots considering Kronecker correlation model with transmit correlation only. i.e., \( R_{i,j} = (0.99999)^{i-j}, \ 1 \leq \{i,j\} \leq n_t \) and \( R_t = I_{n_t}, \ n_t = 2, \ n_r = 1, \) and QPSK constellation in Fig. 8. It can be seen from Fig. 8 that the differential system with our precoder works similar to the differential system with the eigen-beamforming precoder [16]. Fig. 9 shows the performance of precoded differential code for OSTBC (Alamouti Code) with unitary constellation in a correlated non-Kronecker
channel with $[R]_{i,j} = (0.99999)^{i-j}$, $1 \leq \{i,j\} \leq n_t n_r$. We have plotted the results for the QPSK constellation and a MIMO system with $n_t = 2$ and $n_r = 2$. It can be seen from Fig. 9 that the application of the proposed precoder improves the SER versus SNR performance of the differential system for correlated non-Kronecker channels by approximately 2.5 dB. However, the precoder based on the eigen-beamforming [16] is not applicable for arbitrarily correlated channels as it is designed by making the assumption of a Kronecker model with transmit channel correlations only. Our precoded differential system works for any correlation model for Rayleigh channel that has an invertible correlation matrix.

E. SER Performance of Precoded Differential STBC with Non-Unitary Constellations

Fig. 10 shows the symbol error rate (SER) performance of the precoded differential space-time scheme for Golden code and Alamouti code with $n_t = 2$ and $n_r = 1$. We have performed simulations for a correlated channel using $[R]_{i,j} = (0.99999)^{i-j}$, $1 \leq \{i,j\} \leq n_t$ and $R_r = I_{n_r}$. The comparisons are made against a system not employing any precoding, i.e., $F = \frac{P}{\sigma_n^2} I_{n_r}$. The precoding matrix is found by the procedure explained in Subsection V-B. It can be seen from Fig. 10 that the proposed precoded differential coding outperforms the differential coding without a precoder for both orthogonal and non-orthogonal STBCs with non-unitary constellations. The previously proposed precoder for differential STBCs [16] is not applicable in this case, as it only works with OSTBCs using unitary (PSK) constellations.

VII. Conclusions

We have proposed differential space-time block codes that are applicable for full-rank square orthogonal and non-orthogonal STBCs, and standard or optimized signal constellations. The decoder of the proposed code does not need to acquire any knowledge about the channel power, signal power, or noise power before the actual decision. Therefore, the decision is based on two consecutive received data samples only, which is an inherent property of a differential decoder. We have also proposed a PEP based precoder design criterion for arbitrarily invertible correlation matrix and orthogonal and non-orthogonal full-rank square STBC with unitary and non-unitary constellations, where existing precoder design cannot be used.

REFERENCES


