Optimum Space-Time Block Codes over Time-Selective Channels

Jun He and Pooi Yuen Kam
Department of Electrical and Computer Engineering
National University of Singapore, Republic of Singapore, S117576
Email: {hejun,elekampy}@nus.edu.sg

Abstract—We analyze the average bit error probability (BEP) of orthogonal space-time block codes (STBC) over time-selective channels. Exact BEP results are obtained in closed form. Through the analysis, we reveal the relationship between the inter-symbol interference (ISI) and the row positions in a STBC matrix. We then introduce one proposition and two design criteria for code design/search. Optimum/near-optimum codes are found designed with reduced ISI.

I. INTRODUCTION

Orthogonal space-time block codes (STBC) [1] are commonly used in MIMO systems, due to the simple maximum likelihood (ML) decoding structure. However, this decoding structure is based on the assumption that the channels are block-wise constant, that is not always true in practice. If the channels vary from symbol to symbol, the orthogonality will be corrupted by inter-symbol interference (ISI), so the conventional linear ML decoder [1] is no longer optimum.

Considering the time-selective channels, [2]–[10] have proposed different decoders for orthogonal STBC. Reference [2] first proposed a suboptimum detection scheme with the conventional linear decoder. It retains the linear decoding structure, but has an irreducible error floor in the high signal-to-noise ratio (SNR) region. Later, an elegant zero-forcing (ZF) decoder for two transmit antennas is presented in [3], where the ISI is completely removed. The ZF decoder is extended to three and four transmit antennas cases in [4]–[6]. Besides the linear decoders above, there are also nonlinear decoders, including parallel interference cancellation (PIC) decoder [6]–[8], ML decoder [8], successive interference cancellation (SIC) decoder [9] and decision-feedback decoder [10]. Instead of designing a new decoding scheme, a modified orthogonal STBC is developed in [11]. While keeping the full diversity order and the orthogonality, the modified orthogonal STBC reduces the ISI to a much lower level, compared with the original orthogonal STBC [1].

Among these works, however, surprisingly few have analyzed the performance of STBC over time-selective channels. Due to the lack of theoretical analysis, little insight can be gained, and it remains unclear how and to what extent the code structure affects the performance of STBC over a time-selective channel. Most of the works simply apply the existing STBC to the time-selective channels, while these codes only provide sub-optimum performances.

In this paper, through the performance analysis of orthogonal STBC over the time-selective channel with the conventional linear decoder, we reveal the relationship between the ISI and the structure of the code matrix. Based on this relationship, we introduce one proposition and two design criteria, following which it is easy to design or search for better STBC that generate less ISI, compared with the original code matrix. With reduced ISI, the new codes not only improve the performance of the conventional linear decoder, but also benefit the PIC, SIC, decision-feedback and other decoders in existing works.

II. SYSTEM MODEL

For the purpose of illustration, we consider a system with four transmit and one receive antenna, transmitting with a modified $G_4$ encoder [1]. The method we use here to analyze the $G_4$ system can be easily applied to other orthogonal STBC systems with arbitrary numbers of antennas.

Transmitting four information symbols $s = [s_1, s_2, s_3, s_4]^T$ in one STBC block, the original $G_4$ encoder generates an $8 \times 4$ code matrix, which is given as [1]

\[
G_4 = \begin{bmatrix}
    s_1 & s_2 & s_3 & s_4 \\
    -s_2 & s_1 & -s_4 & s_3 \\
    -s_3 & s_4 & s_1 & -s_2 \\
    -s_4 & -s_3 & s_2 & s_1 \\
    s_1^* & s_2^* & s_3^* & s_4^* \\
    -s_2^* & s_1^* & -s_4^* & s_3^* \\
    -s_3^* & s_4^* & s_1^* & -s_2^* \\
    -s_4^* & -s_3^* & s_2^* & s_1^*
\end{bmatrix}
\] (1)

The above encoder, however, generates high ISI over a time-selective channel (which we will explain later), and therefore, we use a modified $G_4$ encoder in this paper. The modified encoder simply interchanges the rows of the original code matrix, and is given by

\[
G_4^{op} = \begin{bmatrix}
    s_1 & s_2 & s_3 & s_4 \\
    -s_4 & -s_3 & s_2 & s_1 \\
    -s_2 & s_1 & -s_4 & s_3 \\
    -s_3 & s_4 & s_1 & -s_2 \\
    s_1^* & s_2^* & s_3^* & s_4^* \\
    -s_2^* & s_1^* & -s_4^* & s_3^* \\
    -s_3^* & s_4^* & s_1^* & -s_2^* \\
    -s_4^* & -s_3^* & s_2^* & s_1^*
\end{bmatrix}
\] (2)
This modified encoder is optimum in the sense that it minimizes the ISI of this orthogonal STBC over a time-selective channel. Further explanation will be given in Section IV.

Defining \( \mathbf{r} = [r_1, \ldots, r_4, r_5^{\ast}, \ldots, r_8^{\ast}]^T \), the received signals can be written as

\[
\mathbf{r} = \mathbf{Hs} + \mathbf{n},
\]

where

\[
\mathbf{H} = \begin{bmatrix}
    h_1(1) & h_2(1) & h_3(1) & h_4(1) \\
    h_2(2) & h_3^\ast(2) & -h_2^\ast(2) & -h_1^\ast(2) \\
    h_2(3) & h_1^\ast(3) & h_3^\ast(3) & -h_3(3) \\
    h_2(4) & -h_4(4) & -h_1(4) & h_2(4) \\
    h_2(5) & -h_2^\ast(5) & -h_1^\ast(5) & h_2^\ast(5) \\
    h_2(6) & -h_1(6) & h_3(6) & -h_3(6) \\
    h_2(7) & h_3^\ast(7) & -h_2^\ast(7) & -h_1(7) \\
    h_2(8) & h_3^\ast(8) & h_5(8) & h_4^\ast(8)
\end{bmatrix}.
\]

Here, \( \mathbf{n}(k) = [n_1, \ldots, n_4, n_5^\ast, \ldots, n_8^\ast]^T \) is the noise vector whose elements are independent and identically distributed (i.i.d.), complex, Gaussian random variables, each with mean zero and variance \( N_0 \), and \( \mathbf{H} \) is the channel matrix. We assume all channels undergo frequency-flat, time-selective Rayleigh fading and the channel fading gains \( \{ h_i(t) \}_{1 \leq i \leq 8} \) are identically distributed Gaussian random variables with mean zero and autocorrelation function

\[
\frac{1}{2} E[ h_i(t) h_i^\ast(t + l)] = \sigma_n^2 R(l), \quad i = 1, \ldots, 4,
\]

where \( \sigma_n^2 \) is the average power of every fading process, which we normalize to 1/2. Applying Jakes' model \([12]\), we have \( R(l) = J_0(2\pi f_d T_s l) \). We assume that the channel fading processes in different transmit-receive links are i.i.d., i.e., the coefficient \( R(l) \) is common for all links, and for any \( i \neq j \), we have \( E[ h_i(t) h_j^\ast(t)] = 0 \). Finally, the knowledge of \( \mathbf{H} \) and the channel correlation coefficient \( R(l) \) are assumed to be perfectly known at the receiver end.

### III. PERFORMANCE ANALYSIS

The beauty of orthogonal STBC is that the optimum decoder can reduce to a linear symbol-by-symbol (SBS) decoder. If the same linear decoder is applied in the case of a time-selective channel, the decision vector is formed by multiplying the received vector \( \mathbf{r} \) by \( \mathbf{H}^H \) to give

\[
\mathbf{H}^H \mathbf{r} = \mathbf{H}^H \mathbf{Hs} + \mathbf{H}^H \mathbf{n}.
\]

It is obvious that the off-diagonal elements of

\[
\mathbf{H}^H \mathbf{H} = \begin{bmatrix}
    \| \mathbf{h}_1 \|^2 & \beta_{12} & \beta_{13} & \beta_{14} \\
    \beta_{21} & \| \mathbf{h}_2 \|^2 & \beta_{23} & \beta_{24} \\
    \beta_{31} & \beta_{32} & \| \mathbf{h}_3 \|^2 & \beta_{34} \\
    \beta_{41} & \beta_{42} & \beta_{43} & \| \mathbf{h}_4 \|^2
\end{bmatrix}
\]

are non-zero, i.e., \( \beta_{ij} \neq 0, \; i, j = 1, \ldots, 4 \). Therefore, they cause ISI. Due to symmetry, each \( s_i, \; i = 1 \cdots 4 \), has the same bit error probability (BEP), so we can focus on \( s_1 \). The decision statistic for \( s_1 \) is given by

\[
z_1 = \| \mathbf{h}_1 \|^2 s_1 + \beta_{12} s_2 + \beta_{13} s_3 + \beta_{14} s_4 + \hat{n},
\]

where \( \mathbf{h}_1 \) is the first column of the code matrix \( \mathbf{H} \) and

\[
\beta_{12} = (h_1^1(1) h_2^2(1) - h_3^3(3) h_4^4(3)) + (h_4^4(2) h_3^3(2) - h_5^5(4) h_4^4(4)) + (h_4^4(7) h_5^5(7) - h_5^5(3) h_4^4(5)) + (h_1^1(8) h_2^2(8) - h_5^5(6) h_1^1(6)),
\]

\[
\beta_{13} = (h_1^1(1) h_3^3(1) - h_3^3(5) h_4^4(5)) + (h_2^2(3) h_3^3(3) - h_4^4(7) h_2^2(7)) + (h_3^3(6) h_4^4(6) - h_4^4(2) h_2^2(2)) + (h_1^1(8) h_2^2(8) - h_4^4(4) h_1^1(4)),
\]

\[
\beta_{14} = (h_1^1(1) h_4^4(1) - h_4^4(2) h_1^1(2)) + (h_3^3(2) h_4^4(2)) + (h_1^1(5) h_2^2(5) - h_2^2(6) h_3^3(6)) + (h_1^1(8) h_2^2(8) - h_4^4(7) h_1^1(7)).
\]

Here, \( \hat{n} \) is a complex, Gaussian random variable with mean zero and variance \( \| \mathbf{h}_1 \|^2 N_0 \).

For M-PSK modulation with equiprobable symbols, we can let \( s_1 = \sqrt{E_s} \), without loss of generality. The BEP of \( s_1 \) can be computed from the probability \( \Pr\{z_1 e^{-j\alpha} < 0 \} \) \([13]\), where \( \alpha \) is an angle that depends on the modulation scheme. For BPSK modulation, the BEP result is given by setting \( \alpha = 0 \), and for Q-PSK modulation with Gray coding, by setting \( \alpha = \frac{\pi}{4} \) \([13]\). Conditioning on \( \mathbf{h}_1 \), i.e.,

\[
\begin{bmatrix}
    h_1(1), h_2^\ast(2), h_3^\ast(3), h_3(4), h_5^\ast(5), h_2(6), h_4(7), h_4^\ast(8)
\end{bmatrix},
\]

we first evaluate the conditional BEP of \( s_1 \). Since the channel gains \( h_i(t) \) and \( h_i(t + l) \) are jointly Gaussian for any \( i = 1, \ldots, 4 \), conditioning on the channel gain \( h_i(t) \), the channel gain \( h_i(t + l) \) is a conditional, complex, Gaussian random variable with mean \( m_i(t + l) \) and variance \( 1 - E[m_i(t + l)^2] \) \([14]\). Here, the mean \( m_i(t + l) \) is proportional to the channel gain \( h_i(t) \), and is given by

\[
m_i(t + l) = R(l) h_i(t), \quad i = 1, \ldots, 4.
\]

Therefore, conditioning on the channel gain \( h_i(t) \), we can write \( h_i(t + l) \) as

\[
h_i(t + l)|_{h_i(t)} \sim \mathcal{CN}(R(l) h_i(t), 1 - |R(l)|^2).
\]

Similarly, \( h_i(t) \) is also a complex Gaussian random variable conditioning on the channel gain \( h_i(t + l) \), and is given by

\[
h_i(t)|_{h_i(t+1)} \sim \mathcal{CN}(R(l) h_i(t) + 1 - |R(l)|^2).
\]

In (9)-(11), we have expressed each of the interference parameters \( \beta_{ij}, \; i = 2, 3, 4 \), as the sum of four terms. Applying (14) and (15) to each of these terms, we find that they are conditional Gaussian random variables. Therefore, conditioning on \( \mathbf{h}_1 \), \( \beta_{ij}, \; i = 2, 3, 4 \), are given by

\[
\beta_{12}|_{h_1} \sim \mathcal{CN}(0, \| \mathbf{h}_1 \|^2 (1 - |R(2)|^2)),
\]

\[
\beta_{13}|_{h_1} \sim \mathcal{CN}(0, \| \mathbf{h}_1 \|^2 (1 - |R(4)|^2)),
\]

\[
\beta_{14}|_{h_1} \sim \mathcal{CN}(0, \| \mathbf{h}_1 \|^2 (1 - |R(1)|^2)).
\]

Conditioning on \( s_i, \; i = 2, 3, 4 \), the total interference term, \( \beta_{12}s_2 + \beta_{13}s_3 + \beta_{14}s_4 \), is a zero-mean Gaussian random variable, whose variance is independent of the values of the \( s_i \)'s, as the \( s_i \)'s are M-PSK modulated and have the same amplitude. Since \( \beta_{12}s_2 + \beta_{13}s_3 + \beta_{14}s_4 \) is independent of the
effective signal $\|h_1\|^2s_1$ and the noise $\tilde{n}$, we can treat the interference as Gaussian noise. Using these results in (8), it is now straightforward to show that the conditional BEP is given by

$$P_b|_{h_i} = \frac{1}{2} \text{erfc}\left( \frac{E_r\|h_i\|^2\cos^2\alpha}{\sqrt{\eta E_s + N_o}} \right)^2,$$

where $\eta = 3 - |R(1)|^2 - |R(2)|^2 - |R(4)|^2$. Having this conditional BEP, we can average over channel vector $h_1$ and obtain the average BEP

$$P_b = \frac{1}{2} \sum_{l=1}^{8} \left( 1 - \sqrt{c_l} \right) \prod_{n=1, n \neq l}^{8} \left( c_l - c_n \right),$$

where

$$c_l = \begin{cases} \frac{(1-|R(1)|E_r\cos^2\alpha)}{\sqrt{\eta E_s + N_o}}, & l = 1, 3, 5, 7, \\ \frac{(1+|R(l-1)|E_r\cos^2\alpha)}{\sqrt{\eta E_s + N_o}}, & l = 2, 4, 6, 8. \end{cases}$$

Due to the space limit, the final BEP result is given here directly. Interested readers can refer to [15] for the detailed derivation. A special case of two transmit antennas is in [16].

IV. OPTIMUM CODE DESIGN

As mentioned in Section II, the original code matrix (1) introduces high ISI. It is important to reduce the ISI, while keeping the orthogonality of the code matrix. One simple but effective method is by changing the positions of the rows in the code matrix [11]. However, the question of how to obtain an optimum modified code matrix, in the sense that the ISI is minimized is not addressed in [11], where only an intuitive 'every-other-line' scheme is introduced.

One way to find the optimum code is to simulate the performances of all the code matrices, which is difficult, if not impossible. Alternatively, we can evaluate the performance of each code matrix, using the BEP results obtained in the previous section. This method is simpler compared with the first one, but still too complex. For example, there are $8!$ possible code matrices for a $G_4$ system, and $16!$ for a $G_8$ system. Even though we can cancel some symmetric equivalent matrices, the computational burden is prohibitive. Therefore, it is important to know how the code structure affects the ISI.

To this end, we first rewrite the original code matrix in the form of row vectors, which is given as

$$G_4 = [s_1^T, \cdots, s_8^T, s_4^H, \cdots, s_8^H]^T,$$

and the corresponding channel matrix is given by

$$H_4 = [h_1^T(1), \cdots, h_4^T(4), h_5^H(5), \cdots, h_8^H(8)]^T.$$

where $h_i^{(4)}$ and $h_i^{(8)}(t)$ are $1 \times 4$ row vectors, and $t$ is the symbol time slot.

**Lemma I:** The sequence of rows in the channel matrix is the same as the one in the code matrix. When we interchange the positions of two rows in the code matrix, the channel matrix changes accordingly.

**Proof:** It can be easily seen through the construction of the channel matrix. As an example, the modified code matrix (2) we used in the last section can be written as

$$G_4^{op} = [s_1^T, s_4^H, s_2^H, s_3^H, s_5^T, s_4^H, s_1^T, s_4^H]^T,$$

and the corresponding channel matrix (4) is given by

$$H = [h_1^T(1), h_2^T(2), h_3^H(3), h_4^T(4), h_5^H(5), h_6^T(6), h_7^T(7), h_8^H(8)]^T.$$

Notice that the row vector $h_i^{(4)}(t)$ here is different from the column vector defined in (12).

For the SBS decoder, each $s_i$ is decoded independently with similar interferences from $s_j$, $j \neq i, j = 1, \cdots, 4$. Due to symmetry, we can take $s_1$ as an example, without loss of generality. From (6) and (8), it can be seen that the noise term for $s_1$ is always a complex, Gaussian random variable with mean zero and variance $\|h_1\|^2N_o$. No matter how we change the position of each row. Similarly, the effective signal is always given by $\|h_1\|^2s_1$. Therefore, the modification of the code matrix only affects the ISI terms.

For a $G_4$ system, there are twenty four interference terms (as can be seen from (9) to (11)), since each row of the channel matrix generates three interference terms from $s_2, s_3$ and $s_4$, respectively. For example, the first row of channel matrix $H$ in (25) is given by

$$h_1(1) = [h_1(1), h_2(1), h_3(1), h_4(1)],$$

so the interferences generated are $h_1^H(1)h_2(1)s_2$, $h_1^H(1)h_3(1)s_3$ and $h_1^H(1)h_4(1)s_4$, respectively. As we did in (9) to (11), we can group these twenty four terms into twelve pairs, each of which is a conditional Gaussian random variable.

For the same example above, we find three more terms from the rows $h_1^H(3)$, $h_3^H(5)$ and $h_4^H(2)$, and we group them with three terms from the first row, which are generated by $h_1^H(1)h_2(1)-h_2(3)h_3(3))s_2$, $h_1^H(1)h_3(1)-h_3(5)h_5(5))s_3$ and $h_1^H(1)h_4(1)-h_4(2)h_2(2))s_4$, respectively. Conditioning on $h_1$ in (12), it is easy to see that these interference terms are Gaussian random variables with means zero and variances $(h_1(1)^2 + h_2(3))((1-|R(2)|^2)E_s)$, $(h_1(1)^2 + h_5(5))((1-|R(4)|^2)E_s)$, and $(h_1(1)^2 + h_4(2)^2)((1-|R(2)|^2)E_s)$, respectively. Since the variances of the channel gains are normalized to 1, the mean values of these ISI’s are obtained as $2(1-|R(1)|^2)E_s$, $2(1-|R(4)|^2)E_s$ and $2(1-|R(2)|^2)E_s$, respectively. Notice that these values depend on and only on the difference of row numbers between $h_1(1)$ and each of $h_1^H(3)$, $h_4^H(5)$ and $h_4^H(2)$. The same observation can be made for the rest of the interference terms. Following the discussion above and applying lemma I, we have Proposition I below:

**Proposition I:** The mean ISI is minimized by minimizing

$$I = \sum_{i=1}^{P} \sum_{j=1, j \neq i}^{P} (1 - |R(D[s_i, s_j^*])|^2)$$

where $2P$ is the number of rows and $D[s_i, s_j^*]$ is the distance between rows $s_i$ and $s_j^*$, given by the difference of their row numbers.
Proposition I can be applied to any \( G_i \) system. Now, the

design of the modified code matrix is simplified to a \( 2P \)-rows

and \( 2P \)-positions problem, i.e. how to arrange \( 2P \) rows in \( 2P \) positions in order to minimize the value of \( I \) in (27). A

computer search can be easily applied to find the optimum code. For the \( G_4 \) system, the optimum code matrix we found has been given in (2) and (24). For a \( G_8 \) system, similarly, if we rewrite the original code given by [1] as

\[
G_8 = \begin{bmatrix}
    s_1^T & s_2^T & s_3^T & s_4^T & s_5^T & s_6^T & s_7^T & s_8^T \\
    s_1^H & s_2^H & s_3^H & s_4^H & s_5^H & s_6^H & s_7^H & s_8^H
\end{bmatrix},
\]

(28)

the optimum \( G_8 \) code we found is given by

\[
G_8^* = \begin{bmatrix}
    s_1^T & s_4^T & s_5^T & s_7^T & s_2^T & s_3^T & s_6^T & s_8^T \\
    s_1^H & s_4^H & s_5^H & s_7^H & s_2^H & s_3^H & s_6^H & s_8^H
\end{bmatrix}^T.
\]

(29)

It is easy to see that with reduced ISI, the modified code

matrix not only improves the performance of the conventional linear decoder, but also benefits the ZF, PIC, SIC and decision-feedback decoders in existing works.

As mentioned above, the number of possible code matrices is given by \((2N_2)!\). Although the calculation of \( I \) in (27) is easy, it is still time consuming to implement an exhaustive search, when the number of transmit antennas is large. From proposition I, we can observe that only the distance between conjugate rows and non-conjugate rows with different subscripts will affect the ISI, therefore, we propose below two design criteria for the modified code matrix.

**Criterion I:** Conjugate rows and non-conjugate rows should be adjacent to each other.

**Criterion II:** Based on criterion I, the rows with the same subscript should be put as far apart as possible.

Using these two criteria, we can easily design code matrices by hand. In Figure 1, we show how to systematically design a \( G_4 \) code, starting from row \( s_1 \) to \( s_4 \). After completing all the rows, the hand-designed \( G_4 \) code matrix is given by

\[
G_4^h = \begin{bmatrix}
    s_1^T & s_2^T & s_3^T & s_4^T \\
    s_1^H & s_2^H & s_3^H & s_4^H
\end{bmatrix}.
\]

(30)

Following the same steps, a hand-designed \( G_8 \) code is given by

\[
G_8^h = \begin{bmatrix}
    s_1^T & s_2^T & s_3^T & s_4^T & s_5^T & s_6^T & s_7^T & s_8^T \\
    s_1^H & s_2^H & s_3^H & s_4^H & s_5^H & s_6^H & s_7^H & s_8^H
\end{bmatrix}^T.
\]

(31)

As we can see, the design of the code matrix is straightforward.

In the next section, we will show that the performance of these hand-designed code matrices are close to the optimum ones.

Finally, we would like to mention that the decoders for the modified code matrices remain the same as the one for the original code matrix, therefore, no extra costs are incurred.

**V. NUMERICAL EXAMPLES**

In Figure 2, we compare the normalized ISI of the original \( G_4 \) code matrix (22), hand-designed code matrix (30) and the optimum code matrix (24), compared with that of code matrix from [11].

As shown in Figure 2, the original code matrix has a much larger ISI compared with the other three. The optimum code matrix has the smallest ISI, and the hand-designed code matrix has a near-optimum ISI level. Both our hand-designed and our optimum code matrices perform better than the one from [11].

In order to further show the effects of reduced ISI of the modified code matrices, we plot the BEP of these codes in Figure 3. In this and the next examples, \( B \)-PSK modulation is used with conventional linear decoders. Here, we use a similar method (which is also detailed in [15]) to approximate the performances of the latter three code matrices, as introduced in Section III.

In Figure 4, we compare the BEP of the optimum \( G_4 \) code matrix and the original \( G_4 \) code matrix, against different values of the channel fade rate. The BEP of the SISO system is also plotted as a reference. As shown in Figure 4, the performances of the two code matrices converge to each other when the channel is static, i.e., \( f_d T_s = 0 \), as there is no ISI introduced in this case. However, if the channel fade rate becomes larger, the performance of the original code matrix degrades much faster than that of the optimum code matrix.
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2008 proceedings.