CHAPTER 2

SPACE TIME/SPACE FREQUENCY/SPACE TIME FREQUENCY CODES

In order to take advantage of the spatial and temporal diversity, a number of space time (ST) codes [17-26] and modulation methods have been proposed. The ST codes can be implemented in two ways (a) ST Trellis (b) ST block coding (STBC). In ST coding, the maximum achievable diversity is equal to the product of number of transmit and receive antennas i.e. $M_{Tx}M_{Rx}$. ST coded OFDM was first introduced in [17]. It uses space time trellis codes over frequency tones. In ST trellis coding scheme, an information stream is encoded via $M_{Tx}$ convolutional encoders (or via one convolutional encoder with $M_{Tx}$ outputs) to obtain $M_{Tx}$ streams of symbols that are transmitted from $M_{Tx}$ antennas simultaneously. Delay diversity (DD) [26] is based upon ST trellis coding. In DD, first antenna will transmit data stream as $[x_n, x_{n+1},...]$, while second antenna will transmit the same stream but delayed by D symbol intervals and hence, the data stream is $[x_{n-D}, x_{n-D+1},...]$. The resulting codes achieve spatial diversity instead of full diversity with higher decoding complexity. The decoding complexity increases exponentially with diversity amount i.e. by increasing number of transmit, receive antennas and transmission rate [28]. Alamouti [20] proposes orthogonal ST block code (OSTBC) design for $2 \times 1$ and $2 \times 2$ systems to overcome this problem. The Alamouti code provides the full diversity of 2 with 2 transmit antennas and code rate of 1. Due to the orthogonal nature of codeword matrix, Alamouti code has fast ML decoding properties that simplify single symbol ML detection. The Alamouti codes can be generalized for more than two transmit antennas using the theory of orthogonal designs [18]. It is shown that OSTBC codes neither provide coding gain nor achieve a rate larger than $3/4$ [25] for more than two transmit antennas.

A general Alamouti coded systematic transceiver model of $M_{Tx}M_{Rx}$ MIMO-OFDM system is shown in Figure 2.1. Initially, the incoming bit stream is mapped into data symbols via modulation technique like BPSK. Then data symbol block is encoded into a codeword matrix $C$ of size $N_cT \times M_{Tx}$ which is sent through $M_{Tx}$ transmit antennas in T
OFDM blocks i.e. $c_1, c_2, \ldots, c_T^T$. Each OFDM block consists of $N_C$ subcarriers that are transmitted from $i^{th}$ transmitting antenna in OFDM blocks 1, 2, ..., $T$.

**Figure 2.1:** ST/ SF / STF coded $M_{Tx} \times M_{Rx}$ MIMO-OFDM transceiver structure

The codeword matrix $C_{ST}$ [26] can be generalized as

$$C_{ST} = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,M_{Tx}} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,M_{Tx}} \\ \vdots & \vdots & \ddots & \vdots \\ c_{T,1} & c_{T,2} & \cdots & c_{T,M_{Tx}} \end{bmatrix} \in \mathbb{C}^{N_C \times M_{Tx}} \quad (2.1)$$

where $c_{t,j}$ denotes a vector of length $N_C$, for all $j = 1, 2, \ldots, M_{Tx}$ and $t = 1, 2, \ldots, T$.

The codeword matrix in (2.1) encodes the data symbols in ST domain. This can be modified to form space frequency (SF) [29-36] codeword matrix $C_{SF}$ as

$$C_{SF} = \begin{bmatrix} c_1(0) & c_2(0) & \cdots & c_{M_{Tx}}(0) \\ c_1(1) & c_2(1) & \cdots & c_{M_{Tx}}(1) \\ \vdots & \vdots & \ddots & \vdots \\ c_1(N_C-1) & c_2(N_C-1) & \cdots & c_{M_{Tx}}(N_C-1) \end{bmatrix} \in \mathbb{C}^{N_C \times M_{Tx}} \quad (2.2)$$
A space time frequency (STF) codeword [37-41] has an additional dimension of time diversity added to the above SF codeword as shown below.

\[
C^K_{STF} = \begin{bmatrix}
  c^K_1(0) & c^K_2(0) & \cdots & c^K_{\text{MTx}}(0) \\
  c^K_1(1) & c^K_2(1) & \cdots & c^K_{\text{MTx}}(1) \\
  \vdots & \vdots & \ddots & \vdots \\
  c^K_1(N_c - 1) & c^K_2(N_c - 1) & \cdots & c^K_{\text{MTx}}(N_c - 1)
\end{bmatrix} \in \mathbb{C}^{N_cK \times \text{MTx}} \tag{2.3}
\]

The OFDM transmitter performs an \(N_c\) point IFFT to each column of matrix \(C\). ISI caused due to multiple delays of channel is removed by addition of CP to each OFDM symbol but its addition reduces spectral efficiency. The length of CP should be equal to or greater than delay spread of channel. The OFDM symbol corresponding to the \(i^{th}\) \((i=1, 2, \ldots, \text{MTx})\) column of \(C\) is transmitted by transmit antenna \(i\). The information is then passed through MIMO channel which is characterized by Jake’s model [42] for both Rayleigh frequency flat and selective channels. After removing the CP and applying FFT on frequency tones, the received signal at \(j^{th}\) receive antenna for \(n^{th}\) subcarrier during \(k^{th}\) block is given by

\[
y^k_j(n) = \sqrt{\frac{7}{\text{MTx}}} \sum_{i=1}^{\text{MTx}} H^k_{i,j}(n)C^k_{i,j}(n) + N^k_j(n) \tag{2.4}
\]

For \(1 \leq k \leq K\), \(0 \leq n \leq N_c-1\), \(1 \leq i \leq \text{MTx}\), and \(1 \leq j \leq \text{MRx}\), \(H^k_{i,j}(n)\) is given by

\[
H^k_{i,j}(n) = \sum_{l=0}^{L-1} a^k_{i,j}(l)e^{-j2\pi nMG_L} \tag{2.5}
\]

Equation (2.5) represents channel frequency response at the \(n^{th}\) subcarrier between transmit antenna \(i\) and receive antenna \(j\). In (2.5), \(\Delta f = 1/T_s\) is the subcarrier separation in frequency domain. It is assumed that CSI is perfectly known at receiver but unknown at the transmitter. Due to this assumption, there is no need to estimate channel coefficients. The channels between different transmit and receive antenna pairs are assumed to have same power delay profile. The channel is assumed to be quasi-static channel and hence the path gains are constant over a frame of time \(T_s\), and it changes from frame to frame. In (2.4) \(N^k_j(n)\) denotes the additive complex Gaussian noise with zero mean and unit
variance at the \( n^{th} \) subcarrier and at \( j^{th} \) receive antenna. The noise samples are assumed to be uncorrelated for different \( j \) and \( n \). The factor \( \sqrt{\gamma} M_{1x} \) ensures that, \( \gamma \) is the average SNR at each receive antenna, and is independent of the number of transmit antennas. The linear combiner combines the output from FFT’s to form composite output signal. The combined signal is equalized by applying equalizers like decision feedback equalization (DFE), maximum likelihood detector (MLD) and sphere decoder (SD).

### 2.1 Alamouti ST/STF/STF Codes

To transmit \( \beta \) bits/cycle, a modulation scheme is used that maps every \( \beta \) bits to one symbol from a constellation with \( 2\beta \) symbols. The constellation can be a real or complex constellation like PAM, PSK, and QAM. The transmitter picks \( x_1 \) and \( x_2 \) symbols from constellation using a block of \( 2\beta \) bits, which are then transmitted by antenna one and two. Then according to Alamouti [20] at time \( t+T \), it transmits \(-x_2^*\) and \( x_1^* \) from antennas one and two, respectively. The scheme is shown in Table 2.1.

<table>
<thead>
<tr>
<th>Time (( t ))</th>
<th>Transmitter Antenna 1</th>
<th>Transmitter Antenna 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 )</td>
<td>( x_2 )</td>
<td></td>
</tr>
<tr>
<td>( -x_2^* )</td>
<td>( x_1^* )</td>
<td></td>
</tr>
</tbody>
</table>

The transmitted codeword \( C \) or \( X \) is given by

\[
X = C = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix}
\]

(2.6)

To check whether the code provides full diversity or not, the rank of all possible difference matrices [13] \( \text{Diff}(C, C') \) is calculated. If the rank comes out to be 2 for every \( C \neq C' \), then it means code provides full diversity. Considering another pair of symbols \((x'_1, x'_2)\), codeword matrix becomes

\[
X' = C' = \begin{bmatrix} x'_1 & x'_2 \\ -x'_2^* & x'_1^* \end{bmatrix}
\]

(2.7)
The difference matrix $\text{Diff}(C, C')$ can be written as

$$\text{Diff}(C, C') = \begin{bmatrix} x_1' - x_1 & x_1' - x_2 \\ x_2' - x_2' & x_2' - x_1' \end{bmatrix}$$

(2.8)

The determinant of difference matrix, $\det[\text{Diff}(C, C')]$ is $|x_1' - x_1|^2 + |x_2' - x_2|^2$ and it is zero only, if $x_1' = x_1$ and $x_2' = x_2$. Thus, $\text{Diff}(C, C')$ is always a full rank matrix when $(C \neq C')$ and hence Alamouti code satisfies the determinant criterion. It provides a diversity of $2M_{Rx}$ for $M_{Rx}$ receive antennas. Therefore the Alamouti code is called full diversity code. According to Alamouti STBC [20], the received signals with two transmit and one receive antenna can be represented by

$$y_1 = y(t) = h_1 x_1 + h_2 x_2 + n_1$$

(2.9)

$$y_2 = y(t+T) = -h_1 x_2^* + h_2 x_1^* + n_2$$

(2.10)

where $y_1$ and $y_2$ are received signals at time $t$ and $t+T$, $h_1$ and $h_2$ are channel coefficients between particular receiver and transmitter antenna pair and $n_1$ and $n_2$ are Gaussian random variables with zero mean and variance $\sigma^2$. The receiver estimates of transmitted symbols $x_1$ and $x_2$ are given by

$$\hat{x}_1 = h_1^* y_1 + h_2^* y_2 = (\alpha_1^2 + \alpha_2^2) x_1 + h_1^* n_1 + h_2^* n_2$$

(2.11)

$$\hat{x}_2 = h_2^* y_1 - h_1^* y_2 = (\alpha_1^2 + \alpha_2^2) x_2 - h_1^* n_1 + h_2^* n_2$$

(2.12)

where, $\alpha_1$ and $\alpha_2$ are channel amplitudes or gains for $h_1$ and $h_2$. The transmitted signals are obtained from their receiver estimates by applying ML detection algorithm on it. In above case, ML detection has to minimize the following decision metric over all possible values of $x_1$ and $x_2$.

$$\left| y_1 - \alpha_1 x_1 - \alpha_2 x_2 \right|^2 + \left| y_2 + \alpha_1 x_2^* - \alpha_2 x_1^* \right|^2$$

(2.13)

Decoding scheme in (2.13) requires full search over all possible pairs of $x_1$ and $x_2$ and its complexity grows exponentially with increase in transmitting antennas. After deleting
the terms that are independent of codewords, the above minimization problem can be decomposed into two parts for separate decoding of $x_1$ and $x_2$ is shown below.

\[
|x_1|^2 \left[ (y_1^* \alpha_1^* + y_2^* \alpha_2 - x_1) + (\alpha_1^2 + \alpha_2^2 - 1) \right]
\] (2.14)

\[
|x_2|^2 \left[ (y_1^* \alpha_2^* - y_2^* \alpha_1 - x_2) + (\alpha_1^2 + \alpha_2^2 - 1) \right]
\] (2.15)

This decomposition reduces the decoder complexity; but it increases linearly instead of exponentially with increase in transmitting antennas. Alamouti received signals with 2 transmit and 2 receive antennas can be represented as follows.

\[
y_{11} = y_1(t) = h_{11}x_1 + h_{21}x_2 + n_{11}
\]

\[
y_{12} = y_1(t+T) = -h_{11}^*x_2^* + h_{21}^*x_1^* + n_{12}
\] (2.16)

\[
y_{21} = y_2(t) = h_{12}x_1 + h_{22}x_2 + n_{21}
\]

\[
y_{22} = y_2(t+T) = -h_{12}^*x_2^* + h_{22}^*x_1^* + n_{22}
\]

In (2.16), $y_{ij}(t)$ represents received signal at $j^{th}$ receiving antenna, $h_{ij}(t)$ is channel coefficient between $i^{th}$ transmit antenna and $j^{th}$ receive antenna and $n_{ij}$ is Gaussian noise of zero mean and variance $\sigma^2$. Receiver estimates of transmitted symbols $x_1$ and $x_2$ with two receivers are given as

\[
\hat{x}_1 = h_{11}^*y_{11} + h_{12}^*y_{21} + h_{21}^*y_{12} + h_{22}^*y_{22}
\]

\[
= (\alpha_{11}^2 + \alpha_{21}^2 + \alpha_{12}^2 + \alpha_{22}^2)x_1 + h_{11}^*n_{11} + h_{21}^*n_{12} + h_{12}^*n_{21} + h_{22}^*n_{22}
\] (2.17)

\[
\hat{x}_2 = h_{21}^*y_{11} - h_{11}y_{12}^* + h_{22}^*y_{21} - h_{12}y_{22}^*
\]

\[
= (\alpha_{11}^2 + \alpha_{21}^2 + \alpha_{12}^2 + \alpha_{22}^2)x_2 - h_{11}n_{12}^* + h_{21}n_{21}^* - h_{12}n_{22} + h_{22}n_{21}
\] (2.18)

As discussed earlier Alamouti code can provide full diversity of 2 with rate 1 and with simpler single symbol detection due to its orthogonal nature. It can be generalized for more transmit antenna case based upon theory of orthogonal design. In such cases OSTBC cannot provide rate more than $3/4$ [25]. In ST coding full multipath diversity can not be achieved. To exploit it coding is done across antennas and OFDM subcarriers
called SF coding [29]. Alamouti based SF coding can be achieved by spreading Alamouti code across two subcarriers [32] in one OFDM block as shown in Table 2.2.

<table>
<thead>
<tr>
<th>Transmitting Antenna 1</th>
<th>OFDM Subcarrier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_1$</td>
</tr>
<tr>
<td>$x_1$</td>
<td>-$x_2^*$</td>
</tr>
<tr>
<td>Transmitting Antenna 2</td>
<td>$x_2$</td>
</tr>
</tbody>
</table>

Table 2.2: Alamouti SF encoding scheme

Table 2.2 shows that two symbols $x_1$ and -$x_2^*$ are sent from subcarriers $K_1$ and $K_2$ of the same OFDM block through transmitting antenna 1. Similarly symbols $x_2$ and $x_1^*$ are sent from subcarriers $K_1$ and $K_2$ of the same OFDM block but through transmitting antenna 2. This code still can’t achieve full diversity especially in frequency selective channels. Performance can be enhanced by spreading Alamouti coding across space, time and frequency called STF codes. Table 2.3 shows that two symbols $x_1$ and -$x_2^*$ are sent from subcarriers $K_1$ and $K_2$ of OFDM block $n$ through transmitting antenna 1. Similarly symbols $x_2$ and $x_1^*$ are sent from subcarriers $K_1$ and $K_2$ of OFDM block $(n+1)$ through transmitting antenna 2.

<table>
<thead>
<tr>
<th>Transmitting Antenna 1</th>
<th>OFDM Subcarrier</th>
<th>OFDM Block (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_1$</td>
<td>$K_2$</td>
</tr>
<tr>
<td>$x_1$</td>
<td>-$x_2^*$</td>
<td></td>
</tr>
<tr>
<td>Transmitting Antenna 2</td>
<td>$x_2$</td>
<td>$x_1^*$</td>
</tr>
</tbody>
</table>

Table 2.3: Alamouti STF encoding scheme

In case of ST coding with Alamouti structure, the decoding metric can be decomposed into two separate parts for detecting each individual symbol i.e. ML decoding becomes single symbol decodable ML (SML) as shown in (2.14), (2.15), (2.17) and (2.18). In SF and STF coding, single symbol ML decoder doesn’t yield optimum results because the channel orthogonality is disturbed in case of frequency selective channels. In such cases, joint ML decoder (JML) is preferred which detects two symbols jointly. ML decision metric, which detects $x_1$ and $x_2$ jointly, is given by
\[ D_m(\hat{x}_1, \hat{x}_2) = ([y - XH])^2 \]
\[ = [y_0 - H_{11}x_1 - H_{21}x_2]^2 + [y_1 + H_{12}x_2^* - H_{22}x_1^*]^2 \]  

(2.19)

By neglecting the terms that are independent of transmitted symbols \( x_1 \) and \( x_2 \), the above decision metric reduces to

\[ D_m(\hat{x}_1, \hat{x}_2) = \left( [H_{11}]^2 + [H_{22}]^2 \right) \|x_1\|^2 + \left( [H_{12}]^2 + [H_{21}]^2 \right) \|x_2\|^2 - \]
\[ 2R \left( [H_{11}y_0 + H_{22}y_1]x_1 + (H_{12}y_0^* - H_{21}y_1^*)x_2 \right) + (H_{12}H_{22}^* - H_{11}H_{21}^*) x_1^* x_2^* \]

(2.20)

In [38], STF coded MIMO-OFDM can achieve full diversity and full rate. However, its decoding complexity is higher than ST and SF coding as shown in (2.20). To reduce the design complexity, a grouping method with precoding and bit-interleaving is proposed in [36-37]. In [33], the repetition mapping technique is proposed to transform existing ST codes to full diversity SF codes. In next section, the general performance analysis for MIMO-OFDM systems with Alamouti ST, SF and STF coding schemes in severely faded quasi-static Rayleigh frequency flat and selective channel is shown. In such channel environments, channel coefficients are close to signal amplitudes and the signal is severely faded. It is very difficult to decode signal in such a channel scenario. We achieved full rate and full diversity even in such channel environments using Alamouti STF codes with repetition techniques. Our decoding emphasize upon increasing the diversity order rather than employing precoding as done in existing literature. The simulation results verify this for higher SNR region. If channel changes independently from one block to another, STF coding provides a significant improvement compared to the SF coding approach. The increased decoder complexity is resolved by employing SD on receiver side.

### 2.2 PERFORMANCE CRITERIA

In this section ST, SF and STF coding approaches for MIMO-OFDM are analyzed and their performance criteria are derived [39]. These coding schemes are compared in terms of coding rate, diversity gain and decoding complexity. The code rate is generally defined as the ratio of total number of information symbols sent per channel and mathematically, it is approximately equal to \( N_c/N_cT_s \) symbols per channel use (PCU).
This means $N_g$ information symbols are sent over $N_C$ T channels where $N_C$ channels are used in $T_s$ times. Diversity gain is the number of faded replicas of same information that are sent to the receiver in various domains like space, time and frequency. Since the probability that all the signal replicas fade equally and simultaneously is extremely small, the receiver performance is enhanced significantly. In flat MIMO channels, full diversity gain is $M_{Tx}M_{Rx}$ whereas in frequency-selective MIMO channels it is $M_{Tx}M_{Rx}L$. Although ST coded MIMO-OFDM has a simple implementation with minimal decoding complexity, but it neither achieve multipath diversity nor high rate. SF coding can achieve maximum diversity and full rate over multipath fading channels with increased decoding complexity. A joint ML decoding method is needed for such cases. STF coded MIMO-OFDM achieve full diversity and full rate. However, its decoding complexity is higher than ST and SF coding.

### 2.2.1 Pairwise Error Probability (PEP) Criterion

The design criterion derived in [39] serve as a formulation to evaluate any coding scheme. The received signal in (2.4) can be rewritten in vector form as

$$Y = \sqrt{\frac{N}{M_{Tx}}} DH + N$$  

(2.21)

In (2.21), $D$ is a $K_N M_{Tx} \times KN_C M_{Tx} M_{Rx}$ matrix constructed from STF codeword in (2.3) and is given by

$$D = I_{M_{Rx}} \otimes \begin{bmatrix} D_1 & D_2 & D_3 & \ldots & D_{M_{Tx}} \end{bmatrix}$$  

(2.22)

and

$$D_i = \text{diag}\{C_i(0), C_i(1), \ldots, C_i(KN_C - 1)\}$$  

(2.23)

For any $i=1, 2 \ldots M_{Tx}$. Each $D_i$ in (2.23) is related to $i^{th}$ column of the STF codeword in (2.3). The channel vector $H$ of size $KN_C M_{Tx} M_{Rx} \times 1$ can be combined as

$$H = \begin{bmatrix} H_{1,1}^T \cdots H_{1,M_{Tx}}^T \cdots H_{M_{Tx},1}^T \cdots H_{M_{Tx},M_{Rx}}^T \end{bmatrix}^T$$  

(2.24)
where
\[
H_{i,j} = \begin{bmatrix} H_{i,j}(0) & H_{i,j}(1) & \ldots & H_{i,j}(KN_c - 1) \end{bmatrix}^T
\]  
(2.25)

The received signal vector Y of size \(KN_cM_{Rx}\times 1\) is given by
\[
Y = \begin{bmatrix} y_1(0) & \ldots & y_1(KN_c - 1) & y_2(0) & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ y_{M_{Rx}}(0) & \ldots & y_{M_{Rx}}(KN_c - 1) \end{bmatrix}^T
\]  
(2.26)

Also, the noise vector N is same as of Y, i.e.
\[
N = \begin{bmatrix} n_1(0) & \ldots & n_1(KN_c - 1) & n_2(0) & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ n_{M_{Rx}}(0) & \ldots & n_{M_{Rx}}(KN_c - 1) \end{bmatrix}^T
\]  
(2.27)

Suppose D and \(\tilde{D}\) are two different matrices related to different STF codewords C and \(\tilde{C}\) respectively. The PEP between D and \(\tilde{D}\) can be upper bounded as
\[
P_e(D \rightarrow \tilde{D}) \leq \left(\frac{2r - 1}{r}\right) \left(\prod_{i=1}^r \lambda_i\right)^{-\frac{1}{2}} \left(\frac{\mathcal{Y}}{M_{Ts}}\right)^{-r}
\]  
(2.28)

where \(r\) is the rank of \((D \rightarrow \tilde{D})\) \(\mathcal{R}\) \((D \rightarrow \tilde{D})^H\), \(\lambda_1, \lambda_2, \lambda_3\ldots \lambda_r\) are the non-zero eigen values of \((D \rightarrow \tilde{D})\) \(\mathcal{R}\) \((D \rightarrow \tilde{D})^H\), and \(\mathcal{R} = \mathbb{E}\{HH^H\}\) is the correlation matrix of H. Based on the upper bound on the PEP in (2.28), two general STF performance criteria are derived as follows:

**Diversity Criterion:** It is also called rank criterion, which says that minimum rank of \((D \rightarrow \tilde{D})\) \(\mathcal{R}\) \((D \rightarrow \tilde{D})^H\) over all pairs of different codewords C and \(\tilde{C}\) should be as large as possible.

**Product Criterion:** It says that minimum value of the product \(\prod_{i=1}^r \lambda_i\) over all pairs of different codewords C and \(\tilde{C}\) should be maximized.
2.2.2 Diversity Analysis Criterion

In spatially uncorrelated MIMO channels [39], the channel taps $\alpha_{i,j}^k(l)$ between each pair of transmit antenna $i$ and receive antenna $j$ are independent of each other. Thus, correlation matrix $R$ of size $KN_{C}M_{Tx}M_{Rx} \times KN_{C}M_{Tx}M_{Rx}$ can be combined as

$$
R = \text{diag} \left( R_{1,1}, \ldots, R_{M_{Tx},1}, \ldots, R_{1,2}, \ldots, \ldots, R_{M_{Tx},2}, \ldots, R_{1,M_{Rx}}, \ldots, R_{M_{Tx},M_{Rx}} \right)
$$

(2.29)

where

$$
R_{i,j} = E[H_{i,j}H_{i,j}^H]
$$

(2.30)

denotes the channel frequency response correlation matrix from transmit antenna $i$ to receive antenna $j$. Using notation $w = e^{-j2\pi f}$, $H_{i,j}$ can be decomposed as

$$
H_{i,j} = (I_K \otimes W)A_{i,j}
$$

(2.31)

with

$$
W = \begin{bmatrix}
1 & 1 & \ldots & 1 \\
\omega^{\tau_0} & \omega^{\tau_1} & \ldots & \omega^{\tau_{L-1}} \\
\omega^{(N_c-1)\tau_0} & \omega^{(N_c-1)\tau_1} & \ldots & \omega^{(N_c-1)\tau_{L-1}} \\
\ldots & \ldots & \ldots & \ldots
\end{bmatrix}_{N_c \times L}
$$

(2.32)

which is related to delay distribution, and

$$
A_{i,j} = \begin{bmatrix}
\alpha_{i,j}^0(0), \alpha_{i,j}^0(1), \ldots, \alpha_{i,j}^0(L-1), \ldots, \\
\ldots, \alpha_{i,j}^k(0), \alpha_{i,j}^k(1), \ldots, \alpha_{i,j}^k(L-1), \ldots
\end{bmatrix}^T
$$

(2.33)

This is related to the power distribution of the channel impulse response. In general, $W$ is not a unitary matrix. If all of the $L$ delay paths fall at the sampling instances of the receiver, then $W$ is a part of the DFT matrix which is unitary. The correlation matrix of the channel frequency response vector between transmit antenna $i$ and receive antenna $j$ can be calculated as
\[
R_{i,j} = E \left\{ (I_k \otimes W) A_{i,j} A_{i,j}^H (I_k \otimes W)^H \right\} 
\]  \hspace{1cm} (2.34)

\[
R_{i,j} = (I_k \otimes W) E \left\{ A_{i,j} A_{i,j}^H \right\} (I_k \otimes W)^H 
\]  \hspace{1cm} (2.35)

It is assumed that the path gains \( \alpha_{i,j}(1) \) are independent for different paths and different pairs of transmit and receive antennas. The second order statistics of the time correlation is same for all transmit and receive antenna pairs for all paths. Thus the correlation matrix \( E \{ A_{i,j} A_{i,j}^H \} \) can be expressed as

\[
E \{ A_{i,j} A_{i,j}^H \} = R \otimes \text{diag}(\delta_0^2, \delta_1^2, \ldots, \delta_{L-1}^2) 
\]  \hspace{1cm} (2.36)

where \( R \) is temporal correlation matrix, thus frequency correlation matrix \( R_F \) can also be expressed as

\[
R_F = E \{ H_{i,j}^K H_{i,j}^{K^T} \} 
\]  \hspace{1cm} (2.37)

where, \( H_{i,j}^K = [H_{i,j}^K(0) \ldots \ldots H_{i,j}^K(N_C - 1)^T \] then, \( R_F = W \text{diag}(\delta_0^2, \delta_1^2, \ldots, \delta_{L-1}^2) W^H \). As a result

\[
R_{i,j} = (I_k \otimes W) (R_T \otimes \text{diag}(\delta_0^2, \delta_1^2, \ldots, \delta_{L-1}^2)) (I_k \times W^H) 
\]  \hspace{1cm} (2.38)

\[
R_{i,j} = (R_T \otimes W \text{diag}(\delta_0^2, \delta_1^2, \ldots, \delta_{L-1}^2)) (W^H) = R_T \otimes R_F 
\]  \hspace{1cm} (2.39)

Finally, the expression for \((D \rightarrow \tilde{D}) R (D \rightarrow \tilde{D})^H\) can be rewritten as

\[
(D - \tilde{D}) R (D - \tilde{D})^H = I_{M_{\text{Rx}}} \otimes \sum_{i=1}^{M_{\text{Tx}}} (D_i - \tilde{D}_i)(R_T \otimes R_T) (D_i - \tilde{D}_i)^H 
\]  \hspace{1cm} (2.40)

\[
= I_{M_{\text{Rx}}} \otimes \left\{ \left( C_i - \tilde{C}_i \right) (C_i - \tilde{C}_i)^H \right\} (R_T \otimes R_F) \right\} 
\]  \hspace{1cm} (2.41)

Let \( \Delta \equiv (C_i - \tilde{C}_i)(C_i - \tilde{C}_i) \) and \( R \equiv R_T \otimes R_F \). Substitute (2.41) into (2.28) it becomes

\[
P_v(C \rightarrow \tilde{C}) \leq \left( \frac{2r_{\text{M}}}{\gamma M_{\text{Rx}}} - 1 \right) \left( \prod_{i=1}^{N_{\text{c}}} \lambda_i \right)^{-M_{\text{Rx}}} \left( \frac{7}{M_{\text{TX}}} \right)^{-r_{\text{M}} M_{\text{Rx}}} 
\]  \hspace{1cm} (2.42)
where \( r_v \) is rank of \( \Delta \circ \mathcal{R} \) and \( \lambda_1, \lambda_2, \lambda_3 \ldots \lambda_{nv} \) are the non-zero eigen values of \( \Delta \circ \mathcal{R} \). Consequently, diversity and product criteria are modified as follows.

**Diversity (rank) criterion:** The minimum rank of \( \Delta \circ \mathcal{R} \) over all pairs of distinct codewords \( \mathcal{C} \) and \( \tilde{\mathcal{C}} \) should be as large as possible.

**Product criterion:** The minimum value of the product \( \prod_{i=1}^{r_v} \lambda_i \) over all pairs of distinct codewords \( \mathcal{C} \) and \( \tilde{\mathcal{C}} \) should also be maximized.

If the minimum rank of \( \Delta \circ \mathcal{R} \) is \( v \) for any pair of distinct STF codewords \( \mathcal{C} \) and \( \tilde{\mathcal{C}} \), then STF code achieves a diversity order of \( r_v M_{Rx} \) for a fixed number of OFDM blocks, transmitting antennas and correlation matrices \( \mathcal{R}_T \) and \( \mathcal{R}_F \). According to the rank inequalities on Hadamard product and Kronecker product we have

\[
\text{rank} (\Delta \circ \mathcal{R}) \leq \text{rank} (\Delta) \cdot \text{rank} (\mathcal{R}_T) \cdot \text{rank} (\mathcal{R}_F) \quad (2.43)
\]

Since the rank \( \Delta \) have maximum value \( M_{Tx} \), the rank of \( \mathcal{R}_F \) is less than or equal to \( L \), and the rank of \( \Delta \circ \mathcal{R} \) is less than or equal to \( KN_C \), we get

\[
\text{rank} (\Delta \circ \mathcal{R}) \leq \min \{ \text{rank}(\mathcal{R}_T), KN_C \} \quad (2.44)
\]

Thus, the full achievable diversity is upper bounded as \( \min [LM_{TX} M_{Rx} \text{rank}(\mathcal{R}_T), KN_C M_{Rx}] \). In section 2.2.3, it is shown that this diversity can be achieved. If the channel remains constant over multiple OFDM blocks, rank of time correlation matrix would be close to 1. Thus, the full diversity in such cases becomes \( \min [LM_{TX} M_{Rx} \text{rank}(\mathcal{R}_T), KN_C M_{Rx}] \) that is same as in case of SF coding.

**2.2.3 STF Code Design Criterion**

In this section, a full diversity STF code criterion is derived from SF coding. For this purpose, it is assumed that the number of subcarriers \( N_C \) is not less than \( LM_{TX} \). Our objective is to show that the maximum achievable diversity order is \( \min [LM_{TX} M_{Rx} \text{rank}(\mathcal{R}_T)] \). Suppose \( C_{SF} \) is a full diversity code of size \( N_C \times M_{TX} \). We can construct a STF code by repeating \( C_{SF} \) codeword \( K \) times (over \( K \) OFDM blocks) as shown below.
\[ C_{STF} = 1_{K \times 1} \otimes C_{SF} \quad (2.45) \]

Let \( \Delta_{STF} = \left( C_{STF} - \tilde{C}_{STF} \right)^{\dagger} \left( C_{STF} - \tilde{C}_{STF} \right) \) and \( \Delta_{SF} = \left( C_{SF} - \tilde{C}_{SF} \right)^{\dagger} \left( C_{SF} - \tilde{C}_{SF} \right) \). Also we have

\[ \Delta_{STF} = \left[ 1_{K \times 1} \otimes \left( C_{SF} - \tilde{C}_{SF} \right) \right] \left[ 1_{K \times 1} \otimes \left( C_{SF} - \tilde{C}_{SF} \right)^{\dagger} \right] \quad (2.46) \]
\[ = 1_{K \times 1} \otimes \Delta_{SF} \quad (2.47) \]

Thus

\[ \Delta_{STF} \circ \mathcal{R} = \left( 1_{K \times K} \otimes \Delta_{SF} \right) \circ \left( \mathcal{R}_T \otimes \mathcal{R}_F \right) \quad (2.48) \]
\[ = \mathcal{R}_T \otimes \left( \Delta_{STF} \circ \mathcal{R} \right) \quad (2.49) \]

Since the SF code \( C_{SF} \) achieves full diversity in each OFDM block, the rank of \( \Delta_{STF} \circ \mathcal{R} \) is \( \text{LM}_{TX} \). Therefore, the rank of \( \Delta_{STF} \circ \mathcal{R} \) is \( \text{LM}_{TX} \text{rank} \left( \mathcal{R}_T \right) \). It means \( C_{STF} \) achieves full diversity of \( \text{LM}_{TX} \text{M}_{Rx} \text{rank} \left( \mathcal{R}_T \right) \).

It is observed that the maximum achievable diversity depends on the rank of the temporal correlation matrix \( \mathcal{R}_T \). If the fading channels are constant during \( K \) OFDM blocks, i.e. \( \text{rank} \left( \mathcal{R}_T \right) = 1 \), the maximum achievable diversity order for STF codes (coding among several OFDM blocks) is the same as that for SF codes (coding within one OFDM block). Moreover, if the channel changes independently in time, i.e. \( \mathcal{R}_T = I_K \), the repetition structure of STF code \( C_{STF} \) is sufficient, but not necessary to achieve the full diversity. We have

\[ (\Delta \circ \mathcal{R}) = \text{diag} \left( \Delta_1 \circ \mathcal{R}_F, \Delta_2 \circ \mathcal{R}_F, \ldots, \Delta_K \circ \mathcal{R}_F \right) \quad (2.50) \]

where \( \Delta_k = \left( C_K - \tilde{C}_K \right)^{\dagger} \left( C_K - \tilde{C}_K \right) \) for \( 1 \leq k \leq K \). Thus, the necessary and sufficient condition to achieve full diversity of \( K \text{LM}_{TX} \text{M}_{Rx} \) is to make \( \Delta_k \circ \mathcal{R}_F \) of rank \( \text{LM}_{TX} \) over all pairs of distinct codewords for \( 1 \leq k \leq K \). This STF code design achieves maximum diversity with rate 1.
2.3 QUASI OSTBC AND ROTATED QUASI OSTBC

Full rate orthogonal designs with complex elements in its transmission matrix are impossible for more than two transmit antennas [25]. The only example of a full-rate full-diversity complex STBC using orthogonal designs is Alamouti schemes [20]. The generator matrix [20] of Alamouti code as mentioned in section 2.1 is given as

$$G(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{pmatrix}$$  \hspace{1cm} (2.51)

To design full rate codes with complex constellation, we consider codes with decoding pair of symbols [34]. Such codes are called QOSTBC as shown below.

$$G_{QOSTBC} = \begin{pmatrix} G(x_1, x_2) & G(x_3, x_4) \\ -G^*(x_3, x_4) & G^*(x_1, x_2) \end{pmatrix}$$

$$= \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_2^* & x_1^* & -x_4^* & x_3^* \\ -x_3^* & x_4^* & x_1 & x_2 \\ x_4 & -x_3 & -x_2 & x_1 \end{pmatrix}$$  \hspace{1cm} (2.52)

If the $i^{th}$ column of above matrix is denoted by $v_i$, for any intermediate variable $x_1, x_2, x_3, x_4$ we have,

$$\langle v_1, v_2 \rangle = \langle v_1, v_3 \rangle = \langle v_2, v_4 \rangle = \langle v_3, v_4 \rangle = 0$$  \hspace{1cm} (2.53)

where $\langle V_i, V_j \rangle$ is the inner product of vectors $V_i$ and $V_j$. Therefore, the subspace created by $V_1$ and $V_4$ is orthogonal to the subspace created by $V_2$ and $V_3$. This is the rationale behind the name “quasi-orthogonal” for the code. The minimum rank of the difference matrix $\text{Diff}(C, C')$ could be two. Therefore, the diversity of the code is two, which is less than the maximum possible diversity of four. Thus, to design full-diversity QOSTBCs, the encoding should be done in such a way that at time $t$, the four elements of $t^{th}$ row of $C$ are transmitted from the four transmit antennas as shown in (2.52). Thus, for such QOSTBC codes, the codeword matrix can be represented as
C = G (x_1, x_2, x_3, x_4). Since four symbols are transmitted in four time slots, the code rate is 1. The ML decoding matrix for QOSTBC is shown below.

\[
\min_{x_1, x_2, x_3, x_4} \{H^H C^H CH - H^H C^H Y - Y^H CH \}
\]

After simple calculations, ML decoding minimizes the following sum.

\[
f_{14}(x_1, x_4) + f_{23}(x_2, x_3)
\]

where \(f_{14}(x_1, x_4)\) and \(f_{23}(x_2, x_3)\) are given as

\[
f_{14}(x_1, x_4) = \sum_{j=1}^{M_{Rx}} \left( |x_1|^2 + |x_4|^2 \right) + \left( \sum_{i=1}^{4} |h_{i,j}|^2 \right)
+ 2R \left\{ -h_{i,j} y_{i,j}^* - h_{2,j} y_{2,j}^* - h_{3,j} y_{3,j}^* - h_{4,j} y_{4,j}^* \right\} x_1
+ \left\{ -h_{4,j} y_{4,j}^* + h_{3,j} y_{3,j}^* + h_{2,j} y_{2,j}^* - h_{1,j} y_{1,j}^* \right\} x_4
+ 4R \left\{ h_{1,j} h_{4,j}^* - h_{2,j} h_{3,j}^* \right\} R \{ s_1, s_4 \} \tag{2.56}
\]

and

\[
f_{23}(x_2, x_3) = \sum_{j=1}^{M_{Rx}} \left( |x_2|^2 + |x_3|^2 \right) + \left( \sum_{i=1}^{4} |h_{i,j}|^2 \right)
+ 2R \left\{ -h_{2,j} y_{2,j}^* + h_{i,j} y_{2,j}^* - h_{3,j} y_{3,j}^* + h_{4,j} y_{4,j}^* \right\} x_2
+ \left\{ -h_{3,j} y_{3,j}^* - h_{4,j} y_{4,j}^* + h_{2,j} y_{2,j}^* + h_{1,j} y_{1,j}^* \right\} x_3
+ 4R \left\{ h_{2,j} h_{3,j}^* - h_{1,j} h_{4,j}^* \right\} R \{ s_2, s_3 \} \tag{2.57}
\]

\(f_{14}(x_1, x_4)\) and \(f_{23}(x_2, x_3)\) are independent from \((x_2, x_3)\) and \((x_1, x_4)\). The pairs \((x_2, x_3)\) and \((x_1, x_4)\) are decoded separately. Further \(f_{14}(x_1, x_4)\) and \(f_{23}(x_2, x_3)\) can be rewritten as

\[
f_{14}(x_1, x_4) = f_1(x_1) + f_4(x_4)
\]

\[
f_{23}(x_2, x_3) = f_2(x_2) + f_3(x_3)
\]

\[42\]
Thus, we can decode \( x_1, x_2, x_3 \) and \( x_4 \) separately. Another QOSTBC with similar properties of (2.52) is given as

\[
G_{\text{QOSTBC}} = \begin{pmatrix} G(x_1, x_2) & G(x_3, x_4) \\ G(x_3, x_4) & G(x_1, x_2) \end{pmatrix}
\]  

(2.60)

For regular symmetric constellations like PSK and QAM, it is easy to show that the minimum rank of the difference matrix \( \text{Diff}(C, C') \) is two for QOSTBCs in (2.52). Therefore, for \( M_{\text{Rx}} \) receive antennas, a diversity of \( 2M_{\text{Rx}} \) is achieved while the code rate is one. The maximum diversity of \( 4M_{\text{Rx}} \) for a rate one complex orthogonal code is impossible in this case if all symbols are chosen from the same constellation. To provide full diversity, we use different constellations for different transmitted symbols i.e. we may rotate symbols \( x_3 \) and \( x_4 \) before transmission. Let us denote \( \tilde{x}_3 \) and \( \tilde{x}_4 \) as the rotated versions of \( x_3 \) and \( x_4 \). The generator matrix for such codes are given as

\[
G_{\text{R-QOSTBC}} = \begin{pmatrix} x_1 & x_2 & \tilde{x}_3 & \tilde{x}_4 \\ -x_2^* & x_1^* & -\tilde{x}_4^* & \tilde{x}_3^* \\ -\tilde{x}_3^* & -\tilde{x}_4^* & x_1 & x_2^* \\ \tilde{x}_4 & -\tilde{x}_3 & -x_2 & x_1 \end{pmatrix}
\]  

(2.61)

The general conditions for QOSTBC and R-QOSTBC [21] codes to achieve full diversity can be postulated based upon coding gain distance (CGD) analysis. The CGD between a pair of codewords \( C = G(x_1, x_2, \tilde{x}_3, \tilde{x}_4) \) and \( C' = G(x_1', x_2', \tilde{x}_3', \tilde{x}_4') \) from the QOSTBC in (2.52) is given by

\[
\text{CGD}(C, C') = \det \left| \text{Diff}(C, C')^\dagger \text{Diff}(C, C') \right|
\]  

(2.62)

For the code in (2.52) we have

\[
G^H G = \begin{bmatrix} a & 0 & 0 & b \\ 0 & a & -b & 0 \\ 0 & -b & a & 0 \\ b & 0 & 0 & a \end{bmatrix}
\]  

(2.63)

where, \( a = \sum_{k=1}^{4} |x_k|^2 \) and \( b = 2R(x_1x_4^* - x_2x_3^*) \). Using determinant equality i.e.
\[
\det\begin{pmatrix}
A & B \\
C & D
\end{pmatrix} = \det(A)\det(D - CA^{-1}B)
\]

(2.64)

Thus, \(\det(G^H G) = (a^2 - b^2)^2\). By applying simple manipulations we have

\[
\det(G^H G) = \left| |x_1 - x_4'| + |x_2 + x_3'| \right|^2 \left| |x_1 + x_4'| + |x_2 - x_3'| \right|^2
\]

(2.65)

By replacing \(x_1, x_2, x_3, x_4\) with \((x_1 - x_1'), (x_2 - x_2'), (\bar{x}_3 - \bar{x}_3')\) and \((\bar{x}_4 - \bar{x}_4')\) respectively, we get

\[
\text{CGD} \left( C, C' \right) = \left[ \left| (x_1 - x_1') - (\bar{x}_4 - \bar{x}_4') \right|^2 + \left| (x_2 - x_2') + (\bar{x}_3 - \bar{x}_3') \right|^2 \right]^{1/2}
\]

(2.66)

\[
\left[ \left| (x_1 - x_1') + (\bar{x}_4 - \bar{x}_4') \right|^2 + \left| (x_2 - x_2') - (\bar{x}_3 - \bar{x}_3') \right|^2 \right]^{1/2}
\]

The unified formula for codewords in (2.52) and (2.60) is given as

\[
\text{CGD} \left( C, C' \right) = \left[ \sum_{k=1}^{2} \left| (x_k - x_k') + (\bar{x}_{k+2} - \bar{x}_{k+2}') \right|^2 \right]^{1/2} \left[ \sum_{k=1}^{2} \left| (x_k - x_k') - (\bar{x}_{k+2} - \bar{x}_{k+2}') \right|^2 \right]^{1/2}
\]

(2.67)

Full diversity is achieved if and only if the CGD in (2.67) is not zero. If the CGD becomes zero, then at least one of the two factors in (2.67) should be zero. To find the minimum CGD, we consider all possible values of \(x_k, x_k', \bar{x}_{k+2}\) and \(\bar{x}_{k+2}'\) for \(k=1,2\). If there is a set of values such that \(\sum_{k=1}^{2} \left| (x_k - x_k') + (\bar{x}_{k+2} - \bar{x}_{k+2}') \right|^2 = 0\), without any loss of generality, we need to consider the case that the second sum becomes zero. This sum contains two non-negative terms, and the sum is zero if and only if both terms are zero. If none of the two pairs of symbols results in \((x_k - x_k') = (\bar{x}_{k+2} - \bar{x}_{k+2}')\), we conclude that the second, and consequently the first, sum are not equal to zero. Also if the first and the second sums are not zero i.e. \((x_k - x_k') \neq (\bar{x}_{k+2} - \bar{x}_{k+2}')\) for any pairs of symbols, Thus collectively we can state that quasi-orthogonal codes will provide full diversity if and only if the rotated constellation is such that \((x_k - x_k') = (\bar{x}_{k+2} - \bar{x}_{k+2}')\) for \(k=1, 2\) is impossible.
Till now we described how to design a rotated quasi-orthogonal STBC that provides full diversity and rate one while decoding pairs of symbols separately. This code structure has one degree of freedom in choosing the rotation parameter. Now we have to look for optimal choice of rotation angles for different constellations. We will use the determinant criterion mentioned above to optimize the rotation parameter. In other words, we will find the rotation angle that maximizes the minimum possible value of the CGD in (2.67) among all possible constellation points. The encoder uses one block of 4b input bits to pick symbols \((x_1, x_2, x_3, \text{ and } x_4)\) from a constellation. Then, \(\tilde{x}_3 = e^{j\theta}x_3\) and \(\tilde{x}_4 = e^{j\theta}x_4\) are the rotated symbols that are used in generator matrix \(G\) of the QOSTBC. For a given constellation, we will find the optimal value of \(\phi\) to maximize the minimum CGD. We denote the minimum CGD of a rotated QOSTBC by \(\text{CGD}_{\min}(\phi)\). Our goal is to maximize \(\text{CGD}_{\min}(\phi)\) among all possible rotations. \(\text{CGD}_{\min}(\phi)\) can be calculated as follows

\[
\text{CGD}_{\min}(\phi) = \min_{(x_i, \tilde{x}_i) \in \{x_i, \tilde{x}_i\}} \left| (x_i - x_i')^2 + (\tilde{x}_i - \tilde{x}_i')^2 \right|^{\frac{1}{2}} \tag{2.68}
\]

where \(x_i\) and \(x_i'\) are symbols from the original constellation and \(\tilde{x}_i\) and \(\tilde{x}_i'\) are symbols from the rotated constellation. The term \(\text{CGD}_{\min}(\phi)\) in (2.68) can be used as a distance between a constellation and its rotated version. Let us denote \(d_{\min} = \min |x_i - x_i'|\) as the minimum Euclidean distance among all constellation points. This distance is same for original and rotated constellations. Choosing \(\tilde{x}_3 = \tilde{x}_3'\) makes right hand side of (2.68) equal to the power of eight of the minimum Euclidean distance of the constellation i.e.

\[
\text{CGD}_{\min}(\phi) \leq d_{\min}^8 \tag{2.69}
\]

Thus, \(d_{\min}^8\) is an upper bound on \(\text{CGD}_{\min}(\phi)\). If there exists a rotation \(\phi^*\) for which the minimum possible CGD is equal to \(d_{\min}^8\). Then \(\max_{\phi} [\text{CGD}_{\min}(\phi)] = \text{CGD}_{\min}(\phi^*) = d_{\min}^8\) and \(\phi^*\) is the optimal rotation. In general the optimal rotation may not be unique. Also, the upper bound is not always achievable. So we will derive another upper bound on \(\text{CGD}_{\min}(\phi)\). Let us assume that the closest points in the constellation are \(\hat{x}\) and \(\hat{x}'\).
Thus, \(d_{\text{min}} = |\hat{x} - \hat{x}'|\), and using the definition of \(\text{CGD}_{\text{min}}(\phi)\), the minimum CGD cannot exceed the CGD between the pairs \((\hat{x}, \hat{x}e^{j\theta})\) and \((\hat{x}', \hat{x}'e^{j\theta})\). Thus

\[
\text{CGD}_{\text{min}}(\phi) \leq \left| (\hat{x} - \hat{x}')^2 + (\hat{x}e^{j\theta} - \hat{x}'e^{j\theta})^2 \right|^{1/2}
\]

\[
= \left| 1 - e^{j2\theta} \right| \left| (\hat{x} - \hat{x}')^2 \right|
\]

\[
= |2\sin \phi|^2 \ d_{\text{min}}^2
\]

(2.70)

The upper bound of \(\text{CGD}_{\text{min}}(\phi)\) in (2.70) is smaller than the upper bound in (2.69) or in other words, (2.70) is a tighter bound than that of (2.69). In case of L-PSK constellation with \(L > 6\) and \(\phi < \frac{\pi}{6}\), the upper bound in (2.69) is not achievable. For BPSK constellation, \(\phi = \frac{\pi}{2}\) achieves the upper bound (2.69). In this case, the original constellation consists of \((-1, 1)\) and the rotated constellation consists of \((-j, j)\). For the optimal \(\phi = \frac{\pi}{2}\), we have \(\text{CGD}_{\text{min}}\left(\frac{\pi}{2}\right) = 256\), which is the maximum possible \(\text{CGD}_{\text{min}}(\phi)\).

\(\text{CGD}_{\text{min}}(\phi)\) is a non-decreasing function of the rotation \(0 \leq \phi \leq \frac{\pi}{2}\) i.e. as rotation \(\phi\) increases in the range of \(0 \leq \phi \leq \frac{\pi}{2}\) the performance improves. For QPSK constellation, we achieved the upper bound of (2.69) with optimal rotation \(\phi = \frac{\pi}{4}\) and maximum possible \(\text{CGD}_{\text{min}}\left(\frac{\pi}{4}\right) = 16\).

In sections 2.1-2.3, we discussed Alamouti ST/SF/STF codes along with quasi and rotated quasi OSTBC's with their ML decoding metrics. In MIMO channels, multiple receivers will receive the signals from all transmitters. These signals are combined to form composite signal and then equalized by applying different equalizer algorithms which is given in the following section.
2.4 EQUALIZATION AND DETECTION

Delay spread causes ISI, which in turn produces an irreducible error floor in most digital modulation techniques. There are several techniques [43] proposed to limit problems produced by delay spread. These techniques fall in two broad categories namely, signal processing and antenna solutions. In a broad sense, equalization is a signal processing technique used at the receiver to alleviate the ISI problem caused by delay spread. Signal processing can also be used at the transmitter to make the signal less susceptible to delay spread, spread spectrum and multicarrier modulation belongs to this category of transmitter signal processing techniques. Our main focus is on equalization.

As mentioned earlier the goal of equalization is to mitigate the effects of ISI. However, this goal must be balanced in the sense that noise power in the received signal should not be enhanced. Figure 2.2 shows generalized block diagram of any communication receiver employing equalizer.

![Figure 2.2: Receiver block diagram employing equalizer](image)

Consider a signal $x(t)$ passed through a channel with frequency response $H(f)$. At the receiver front end, if white Gaussian noise $n(t)$ is added to the signal, the signal input to receiver can be written as

$$Y = HX + N$$  \hspace{1cm} (2.71)

To mitigate channel affects we multiply (2.71) with $H^{-1}$. This multiplication enhances the noise by factor $n/H$. In such scenario, though ISI effects are removed, the equalization performs poorly due to reduced SNR. Thus the true goal of equalization is to balance ISI mitigation with maximizing the SNR of the post equalization signal.
For an equalizer to mitigate ISI introduced by the channel, the channel impulse or frequency response estimation must be known. Since the wireless channel is time variant, the equalizer must estimate the frequency response of the channel (training) and then update its coefficients accordingly (tracking). The process of equalizer training and tracking is often referred to as adaptive equalization, since the equalizer adapts to the changing channel. In general, the training is done by sending a fixed length known bit sequence over the channel. The equalizer at the receiver uses the known training sequence to adapt its filter coefficients to match the channel frequency response. Specifically, the equalizer filter coefficients are updated to minimize the error between the actual channel output and the output resulting from the known training sequence transmitted through channel frequency response estimation. The training process assumes that the channel is relatively constant over the length of the training sequence. If the channel changes relatively slowly, adaptive algorithms are usually sufficient to track the channel changes. However, if the channel is changing quickly, the training sequence may be transmitted periodically to insure that the equalizer coefficients do not drift significantly from their optimal values. It is clearly desirable to avoid periodic retraining because this increases overhead. The length of the training sequence depends on the equalizer structure and its tap update algorithm as well as the channel delay spread and coherence time.

Equalization techniques [43] fall into two broad categories: linear and non linear. The linear techniques are generally the simplest to implement and to understand conceptually. However, linear equalization techniques typically suffer from noise enhancement in frequency-selective fading channels, and therefore they are not preferred in most wireless applications. Among non linear equalization techniques, DFE is the most common, since it is fairly simple to implement and does not suffer from noise enhancement. However, for channels with low SNR, the DFE suffers from error propagation when bits are decoded in error, leading to poor performance. The optimal equalization technique to use is ML Equalizer, but the ML complexity grows exponentially with memory length, which makes it impractical to use in such cases. However, the BER performance of the MLE is
often used as an upper bound for other equalizers. Figure 2.3 summarizes the different equalizer types, along with their corresponding structures and tap updating algorithm.

![Equalizer Tree Diagram]

**Figure 2.3**: Various equalizer types

Equalizer can be symbol by symbol (SBS) or sequence estimators (SE). SBS equalizers remove ISI from each symbol and then detect each symbol individually. All linear equalizers remove ISI from each symbol and then detect each symbol individually. All linear equalizers shown in Figure 2.3 as well as DFE are SBS equalizers. SE equalizers detect sequence of symbols, so the effect of ISI is part of the estimation process. Maximum likelihood sequence estimation (MLSE) is the optimal form of sequence detection, but it is highly complex.

Linear equalizer or transversal filter can be implemented as an FIR filters. In such an equalizer, current and past values of the received signal are linearly weighted by the filter coefficient and summed to produce the output. Non linear equalizers are based upon channel estimation. It updates the equalizer parameters (such as the filter coefficients) as
it processes the data. Typically, it uses the MSE cost function; it assumes that it makes the correct symbol decisions, and uses its estimate of the symbols to compute error. In the subsequent sections, these algorithms are described in context to MIMO channels.

When transmit antennas of MIMO systems transmit different symbols simultaneously, the signals are summed on each receive antenna and the receiver has to separate them in order to detect individual signals. If we consider each transmit antenna as a user, the system can be thought of as a multiple access system and multi-user detection techniques can be employed by the MIMO system. Equalization can be done in both time and frequency domain. Equalization in frequency domain is simpler compared to its time domain counterpart. The important parameter in equalization design is to choose number of taps. The number of taps is limited by maximum time delay spread offered by the channel. An equalizer can equalize for delay intervals less than or equal to the maximum delay within the filter structure. Therefore it is important to know about the number of taps before selecting an equalizer structure and its algorithm. Main algorithms used for MIMO equalization are reported in the following text.

2.4.1 Zero Forcing (ZF) Equalizer

An ISI channel may be modeled by an equivalent finite impulse response (FIR) filter plus noise. A zero-forcing equalizer [44-46] uses an inverse filter to compensate for the channel response function. At the output of the equalizer, an overall response function is equal to one for the symbol that is being detected and an overall zero response for other symbols. This results in removal of interference from all other symbols. It is a linear equalization method that does not consider the effects of noise. In fact, the noise may be enhanced in the process of eliminating the interference.

The inverse of the channel matrix $H$ exists if we assume that $M_{TX} = M_{RX}$ and $H$ is a full rank square matrix. After multiplying (2.71) with $H^{-1}$ we get

$$YH^{-1} = X + NH^{-1}$$

(2.72)

In case if $M_{TX}$ and $M_{RX}$ are not same, we multiply (2.71) by Moore Penrose inverse or pseudo inverse of $H$ to achieve similar results. Here we need to calculate mean square error (MSE) to solve for $X$ i.e. we have to minimize $\|Y - HX\|^2$ as
\[
\text{MSE} = \left\| Y - HX \right\|^2 = (Y - HX)^H (Y - HX)
\]

(2.73)

where \( Y \) is \( M_{R_3 \times 1} \) received vector, \( H \) is \( M_{R_3 \times M_{T_X}} \) channel matrix and \( X \) is \( M_{T_X \times 1} \) transmitted vector. On solving (2.73) we get

\[
\text{MSE} = \left\| Y - HX \right\|^2 = (Y - HX)^H (Y - HX)
\]

\[
= Y^H Y - X^H H^H Y - Y^H H X + X^H H^H H X
\]

(2.74)

To minimize (2.74), take derivative of (2.74) w.r.t \( X \) and equate it to zero, we get

\[
\frac{d \left( \left\| Y - HX \right\|^2 \right)}{dX} = -2H^H Y + 2H^H H X = 0
\]

(2.75)

\[-2H^H Y + 2H^H H X = 0\]

So the estimated value of \( X \) can be written as

\[
\hat{X} = \left( H^H H \right)^{-1} H^H Y
\]

(2.76)

The matrix \( \left( H^H H \right)^{-1} H^H \) is called pseudo inverse of \( H \). The covariance matrix of the effect of noise is given as

\[
E \left[ \left( N H^H \right)^H N H^{-1} \right] = \left( H^{-1} \right)^H E(N^H N) H^{-1}
\]

(2.77)

\[= n \left( H H^H \right)^{-1}\]

It is clear from (2.77) that noise power gets increased by factor \( \left( H H^H \right)^{-1} \). Thus, ZF equalizer filter compensates for the channel-induced ISI as well as the ISI generated by the transmitter and receiver filters. The ZF filter designed using the equation above does not eliminate all ISI because the filter is of finite length.

### 2.4.2 Minimum Mean Square Error (MMSE) Equalizer

MMSE minimizes the MSE [47] calculated for values of a dependent variable, which is a common measure of estimator quality. The term MMSE more specifically refers to estimation in a Bayesian setting with quadratic cost function. The basic idea behind the Bayesian approach is to estimate from practical situations where we often have some
prior information about the parameter to be estimated. For instance, we may have prior information about the range that the parameter can assume; or we may have an old estimate of the parameter that we want to modify when a new observation is made available. In this Bayesian approach, such prior information is captured by the prior pdf of the parameters and is based directly on Bayes theorem \[48\]. It allows us to make better posterior estimates as more observations become available. The Bayesian estimator seeks to estimate a parameter that is itself a random variable. Furthermore, Bayesian estimation can also deal with situations where the sequences of observations are not necessarily independent.

If we choose \(X\) to be \(M_{Tx \times 1}\) unknown (hidden) random vector variable, and \(Y\) be \(M_{Rx \times 1}\) known random vector variable (the measurement or observation) and both of them not necessarily have same dimension. An estimator \(\hat{X}(Y)\) of \(X\) is function of the measurement \(Y\). The estimation error vector is given by \(e = \hat{X} - X\) and its MSE is given by the trace of error covariance matrix.

\[
\text{trace}\left[\hat{X} - X\right]\left[\hat{X} - X\right]^H = E\left[\hat{X} - X\right]\left[\hat{X} - X\right]^H \right] = E\left\{W_{\text{MMSE}}^H Y - X\right\}\left[W_{\text{MMSE}}^H Y - X\right]^H \right]
\]

(2.78)

or

\[
E\left\{W_{\text{MMSE}}^H Y - X\right\}\left[W_{\text{MMSE}}^H Y - X\right]^H \right]
\]

Thus in MMSE \[49\], we need to find coefficient \(W_{\text{MMSE}}\) that minimizes the above criterion. On solving (2.78) we get

\[
E\left[\hat{X} - X\right]\left[\hat{X} - X\right]^H = E\left\{W_{\text{MMSE}}^H YY^H W_{\text{MMSE}} - XY^H W_{\text{MMSE}} - W_{\text{MMSE}}^H YX + XX^H \right\}
\]

(2.79)

Taking expectation operator inside, (2.79) can be written as

\[
\left\{W_{\text{MMSE}}^H E\left\{YY^H\right\} W_{\text{MMSE}} - E\left(XY^H\right) W_{\text{MMSE}} - W_{\text{MMSE}}^H E\left(YX\right) + E\left(XX^H\right) \right\}
\]

(2.80)

We define \(E\left\{YY^H\right\} = R_{YY}\) as covariance matrix of \(Y\). \(E\left\{XY^H\right\} = R_{XY}\) as cross covariance of \(X\) and \(Y\), \(E\left\{YX^H\right\} = R_{YX}\) as cross covariance of \(Y\) and \(X\) and
\[ E(XX^H) = R_{XX} \] as covariance of X. Also, the cross covariance of Y and X i.e. \( R_{YX} \) can be represented in terms of cross-covariance of X and Y i.e. \( R_{XY} \) as shown below.

\[ R_{YX} = R_{XY}^H \]  \hspace{1cm} (2.81)

Thus, (2.80) can be written as

\[ = \{ W_{MMSE}^H R_{YY} W_{MMSE} - R_{XY} W_{MMSE} - W_{MMSE}^H R_{YY} + R_{XX} \} \]  \hspace{1cm} (2.82)

Equation (2.82) can be further modified as

\[ = \{ W_{MMSE}^H R_{YY} W_{MMSE} - 2W_{MMSE}^H R_{YX} + R_{XX} \} \]  \hspace{1cm} (2.83)

To calculate minimum values of (2.83), differentiate it w.r.t to \( W_{MMSE} \) and equate it to zero i.e.

\[ 2R_{YY} W_{MMSE} - 2R_{YX} = 0 \]  \hspace{1cm} (2.84)

Hence optimal \( W_{MMSE} \) is given by

\[ W_{MMSE_{opt}} = \frac{R_{YX}}{R_{YY}} \]  \hspace{1cm} (2.85)

or

\[ W_{MMSE_{opt}} = R_{XY} (R_{YY})^{-1} \]

The estimated X i.e. \( \hat{X} \) can be represented as

\[ \hat{X} = W_{MMSE_{opt}}^H Y \]  \hspace{1cm} (2.86)

or

\[ \hat{X} = R_{XY} (R_{YY})^{-1} Y \]

In case of MIMO channels, \( R_{XX} \), \( R_{YX} \) and \( R_{YY} \) can be computed as follows

\[ R_{XX} = E(XX^H) \]
\[
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_{M_{\text{Tx}}}
\end{bmatrix}
\begin{bmatrix}
  x_1^* \\
  x_2^* \\
  \vdots \\
  x_{M_{\text{Tx}}}^*
\end{bmatrix}
\]

\[
= E\begin{bmatrix}
  x_1 x_1^* & x_1 x_2^* & \cdots & x_1 x_{M_{\text{Tx}}}^* \\
  x_2 x_1^* & x_2 x_2^* & \cdots & x_2 x_{M_{\text{Tx}}}^* \\
  \vdots & \vdots & \ddots & \vdots \\
  x_{M_{\text{Tx}}} x_1^* & x_{M_{\text{Tx}}} x_2^* & \cdots & x_{M_{\text{Tx}}} x_{M_{\text{Tx}}}^*
\end{bmatrix}
\]

\[
(2.87)
\]

Taking expectation operator inside we get

\[
\begin{bmatrix}
  E(x_1 x_1^*) & E(x_1 x_2^*) & \cdots & E(x_1 x_{M_{\text{Tx}}}^*) \\
  E(x_2 x_1^*) & E(x_2 x_2^*) & \cdots & E(x_2 x_{M_{\text{Tx}}}^*) \\
  \vdots & \vdots & \ddots & \vdots \\
  E(x_{M_{\text{Tx}}} x_1^*) & E(x_{M_{\text{Tx}}} x_2^*) & \cdots & E(x_{M_{\text{Tx}}} x_{M_{\text{Tx}}}^*)
\end{bmatrix}
\]

\[
(2.88)
\]

If we assume \( x_1, x_2 \ldots x_{M_{\text{Tx}}} \) are independent and uncorrelated, then (2.88) becomes

\[
R_{XX} = \begin{bmatrix}
  (x_1^2) & 0 & \cdots & 0 \\
  0 & (x_2^2) & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & (x_{M_{\text{Tx}}}^2)
\end{bmatrix} = P_d I_{M_{\text{Tx}}}
\]

\[
(2.89)
\]

Covariance of \( Y \) can be computed as

\[
R_{YY} = E(Y Y^H)
\]

\[
= E\left[(H X + N)(H X + N)^H\right]
\]

\[
= E[H X X^H H^H + N X^H H^H + H X N^H + N N^H]
\]

\[
(2.90)
\]

Taking expectation operator inside, (2.90) can be written as
\[ R_{YY} = \left[ HE(XX^H)H^H + E(NX^H)H^H + HE(XN^H) + E(NN^H) \right] \] (2.91)

The transmitted symbols X and channel noise n are uncorrelated. Thus, (2.91) can be rewritten as

\[ R_{YY} = \left[ HR_{xx} H^H + E(NN^H) \right] \] (2.92)

In (2.92), N is a noise vector of size \( M_{\text{Rx}} \times 1 \). Its expectation i.e. \( E(NN^H) \) can be expanded as

\[
E(NN^H) = E\left[ \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_{M_{\text{Rx}}} \end{bmatrix} \begin{bmatrix} n_1^* \\ n_2^* \\ \vdots \\ n_{M_{\text{Rx}}}^* \end{bmatrix} \right]
\]

\[
= E\left[ \begin{bmatrix} n_1n_1^* & n_1n_2^* & \cdots & n_1n_{M_{\text{Rx}}}^* \\ n_2n_1^* & n_2n_2^* & \cdots & n_2n_{M_{\text{Rx}}}^* \\ \vdots & \vdots & \ddots & \vdots \\ n_{M_{\text{Rx}}}n_1^* & n_{M_{\text{Rx}}}n_2^* & \cdots & n_{M_{\text{Rx}}}n_{M_{\text{Rx}}}^* \end{bmatrix} \right]
\] (2.93)

Taking expectation operator inside, we get

\[
= \begin{bmatrix} E(n_1n_1^*) & E(n_1n_2^*) & \cdots & E(n_1n_{M_{\text{Rx}}}^*) \\ E(n_2n_1^*) & E(n_2n_2^*) & \cdots & E(n_2n_{M_{\text{Rx}}}^*) \\ \vdots & \vdots & \ddots & \vdots \\ E(n_{M_{\text{Rx}}}n_1^*) & E(n_{M_{\text{Rx}}}n_2^*) & \cdots & E(n_{M_{\text{Rx}}}n_{M_{\text{Rx}}}^*) \end{bmatrix}
\] (2.94)

If we assume \( n_1, n_2, \ldots, n_{M_{\text{Rx}}} \) are independent and uncorrelated, then (2.94) becomes

\[
E(NN^H) = \begin{bmatrix} n_1^2 & 0 & \cdots & 0 \\ 0 & n_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n_{M_{\text{Rx}}}^2 \end{bmatrix} = \sigma_n^2 I_{M_{\text{Rx}}}
\] (2.95)
Thus, (2.92) becomes

\[ R_{YY} = \left[ P_d HH^H + \sigma_n^2 I_{M_R} \right] \]  

(2.96)

Cross covariance of Y with X i.e. \( R_{YX} \) for MIMO channels can be computed as

\[ R_{YX} = E(YX^H) \]
\[ = E[(HX + N)X^H] \]
\[ = E\left[ (HXX^H + NX^H) \right] \]  

(2.97)

Take expectation operator inside, we get

\[ R_{YX} = \left[ HE\left(XX^H\right) + E\left(NX^H\right) \right] \]  

(2.98)

The transmitted symbols X and channel noise N are uncorrelated. Thus, (2.98) can be rewritten as

\[ R_{YX} = P_d H \]  

(2.99)

Put values of \( R_{YX} \) and \( R_{YY} \) in (2.85), we get

\[ W_{\text{MMSE,ap}} = P_d H \left( P_d HH^H + \sigma_n^2 I_{M_R} \right)^{-1} \]  

(2.100)

The estimated X i.e. \( \hat{X} \) can be represented as

\[ \hat{X} = P_d H^H \left( P_d HH^H + \sigma_n^2 I_{M_R} \right)^{-1} Y \]  

(2.101)

Equation (2.101) can be equivalently written as

\[ \hat{X} = P_d \left( P_d H^H H + \sigma_n^2 I_{M_R} \right)^{-1} H^H Y \]  

(2.102)

In (2.102), we observed that there is no noise enhancement as unlike ZFE. Thus, MMSE equalizer is also called optimum equalizer. At high SNR, (2.102) can be approximated as

\[ \hat{X} \approx P_d \left( P_d H^H H \right)^{-1} H^H Y \]  

(2.103)
\[ \hat{X} \equiv (H^H H)^{-1} H^H Y \] (2.104)

Equation (2.104) is same as of (2.76); Thus MMSE equalizer can be reduced to ZFE at high SNR’s. At low SNR’s (2.102) can be approximated as

\[ \hat{X} = P_d \left( P_d H^H H + \sigma_n^2 I_{M_k} \right)^{-1} H^H Y \]

\[ \hat{X} = P_d \left( \sigma_n^2 I_{M_k} \right)^{-1} H^H Y \]

or

\[ \hat{X} = \frac{P_d}{\sigma_n^2} H^H Y \] (2.105)

Above expression is similar to matched filter (MF) expression. In other words, MMSE equalizer turns out to be MF at low SNR values.

### 2.4.3 Decision Feedback Equalizer (DFE)

As discussed earlier, the channel equalizers are either linear or non linear. Non linear equalization is needed when the channel distortion is too severe for the linear equalizer to mitigate the channel impairments. Linear equalizer forces ISI to become zero for every symbol decision. A ZFE enhances noise and results in performance degradation. On the other hand, MMSE minimizes the error between the received symbol and the transmitted symbol without enhancing the noise. Although MMSE performs better than ZFE, its performance is not enough for channels with severe ISI. An obvious choice for channels with severe ISI is a non linear equalizer. It is well known [43] that a MLSE gives optimum performance. It tests all possible data sequences and decodes for the data with maximum probability as the output. Generally, the Viterbi algorithm provides a solution to the problem with MLSE equalizer. However, the computational complexity of an MLSE increases with channel spread and signal constellation size. The numbers of states of the Viterbi decoder are \(M^L\). where \(M\) is constellation size and \(L\) is channel spread length \(-1\). Therefore for an 8-PSK constellation with a channel span of 5, that is translated to an \(8^4 = 4096\) states Viterbi decoder. Thus, the cost effective implementation is impossible. In this situation, the obvious choice is to use sub-optimal solutions such as a DFE or a DFE followed by a Viterbi equalizer.
DFE [50] is a nonlinear equalizer that uses previous detector decision to eliminate the ISI on pulses that are currently being demodulated. In other words, the distortion on a current pulse that was caused by previous pulses is subtracted. Figure 2.4 shows simplified block diagram of a DFE that employ forward and feedback filters. Each of these filters can be linear such as transversal filter.

Figure 2.4: Schematic of decision feedback equalizer

The nonlinearity of the DFE appears from the nonlinear characteristic of the detector that provides an input to the feedback filter. The DFE works on the principle that that if the values of the symbols previously detected are known, then ISI contributed by these symbols can be canceled out exactly at the output of the forward filter by subtracting past symbol values with appropriate weighting. The forward and feedback tap weights can be adjusted simultaneously to fulfill a criterion like MSE minimization. The advantage of a DFE [51] implementation is the feedback filter that remove ISI. This operates on noiseless quantized levels, and thus its output is free of channel noise. In Figure 2.4, forward filter is driven by decision on the output of the detector, and its coefficients can be adjusted to cancel the ISI on the current symbol based on past detected symbols. RLS
(recursive least squares) [51] algorithm is used to determine the coefficient of an adaptive filter. It uses information from all past input symbols to estimate the autocorrelation matrix of the input vector. To decrease the influence of input symbols from the past symbols, a weighting factor is used for the influence of each symbol. DFE equalizer combines filtering and adaptive process. In filtering process, algorithm computes the output of a linear filter in response to an input signal is given by

\[ y = \tilde{w}HC \]  \hspace{1cm} (2.106)

where \( y \) is output of a linear filter. All subscripts are omitted for simplification. Output \( y \) is compared with desired response \( d \) to generate estimation error \( e \) which is given by

\[ e = y_{\text{desired}} - y \]  \hspace{1cm} (2.107)

In next phase tap-weight vector is updated by incrementing its old value by an amount equal to the complex conjugate of the estimation error as described in (2.107).

\[ \tilde{w}(n_t + 1) = \tilde{w}(n_t) + \mu Ce^* \]  \hspace{1cm} (2.108)

where \( n_t \) and \( \mu \) controls the convergence rate and stability of algorithm.

2.4.4 Successive Interference Cancellation (SIC)

When the signals are detected successively, the outputs of previous detectors can be used to detect upcoming signals. Such algorithms are termed as decision directed detection algorithms which include successive interference cancellation (SIC), parallel interference cancelation (PIC) and multistage detection. SIC [52] approach can be combined with ZF after optimal ordering and with MMSE after equal power allocation. In SIC, the first symbol is detected by decorrelator and its output is used to cancel the interference from the received signal vector assuming that first symbol is detected correctly.

In ZF-SIC, after applying ZF equalizer receiver estimates of transmitted symbols \( x_1 \) and \( x_2 \) are given by
\[
\begin{bmatrix}
\hat{x}_1 \\
\hat{x}_2
\end{bmatrix} = \left( H^H H \right)^{-1} H^H \begin{bmatrix}
y_1 \\
y_2
\end{bmatrix}
\]
(2.109)

Thereafter, receiver chooses one of the received estimates and subtracts its effect from symbols \( y_1 \) and \( y_2 \). Receiver picks \( \hat{x}_1 \) or \( \hat{x}_2 \) based upon the value of received power due to \( x_1 \) and \( x_2 \). The received power at both antennas corresponding to the transmitted symbol \( x_1 \) and \( x_2 \) is given as

\[
P_{x_1} = \left| h_{1,1} \right|^2 + \left| h_{2,1} \right|^2
\]
(2.110)

\[
P_{x_2} = \left| h_{1,2} \right|^2 + \left| h_{2,2} \right|^2
\]
(2.111)

If we assume that \( P_{x_1} > P_{x_2} \), the receiver will decide to remove the effect of \( \hat{x}_1 \) from the received vector \( y_1 \) and \( y_2 \) and then re-estimate \( \hat{x}_2 \) as shown below:

\[
\begin{bmatrix}
\hat{y}_1 \\
\hat{y}_2
\end{bmatrix} = \begin{bmatrix}
y_1 - h_{1,1}\hat{x}_1 \\
y_2 - h_{1,2}\hat{x}_1
\end{bmatrix} = \begin{bmatrix}
h_{1,2}x_2 + n_1 \\
h_{2,2}x_2 + n_2
\end{bmatrix}
\]
(2.112)

where \( \hat{y}_1 \) and \( \hat{y}_2 \) are received signals after subtracting effects of \( x_1 \). Equation (2.112) can be equivalently written as

\[
\hat{Y} = h\hat{x}_2 + \hat{N}
\]
(2.113)

From (2.113), \( \hat{x}_2 \) can be estimated as

\[
\hat{x}_2 = \frac{h^H \hat{Y}}{h^H h}
\]
(2.114)

For the ZF-SIC, as the interference is already nullified, the significance of SIC is to reduce the noise amplification. For MMSE-SIC, the significance of SIC is not only to minimize the amplification of noise but also to cancel the interference from other antennas. System performance can be further improved by optimally re-ordering the ZF-SIC and MMSE-SIC process. It ensures that the reliability of the symbol which is decoded first is guaranteed to have a lower error probability than the other symbol. This results in
lowering the chances of incorrect decisions resulting in erroneous interference cancellation. Hence it gives lower error rate than simple SIC process.

2.4.5 Maximum Likelihood (ML) Detection

The MSE based linear equalizers are optimum [53] when channel does not introduce large amplitude distortion. For larger amplitude distortions ML based equalizer are used which tests all possible data sequences and chooses the data sequence with maximum probability at the output. These equalizers require knowledge of channel characteristics in order to compute the metrics for decision making. They also require knowledge of statistical distribution of the noise, which determines the metric for optimum demodulation of the received signal. Maximum likelihood (ML) decoding [54] calculates the codeword \( \hat{X} \) that solves the following minimization problem.

\[
\hat{X} = \arg \min_X \| Y - XH \|_F^2
\]  
(2.115)

Equation (2.115) can be expanded using Frobenius norm as follows:

\[
\hat{X} = \arg \min_X \left[ \text{Trace} \left( (Y - XH)^H Y - XH \right) \right]
\]

\[
\hat{X} = \arg \min_X \left[ \text{Trace} \left( Y^H Y + H^H X^H XH - H^H X^H Y - Y^H XH \right) \right] 
\]  
(2.116)

If \( Y^H Y \) is independent of the transmitted codeword, (2.116) can be written as

\[
\hat{X} = \arg \min_X \left[ \text{Trace} \left( H^H X^H XH \right) - 2 \text{Real} \left( \text{Tr} \left( H^H X^H Y \right) \right) \right]
\]  
(2.117)

Also (2.117) can be generalized for multiple receivers as follows:

\[
\text{Trace} \left( H^H X^H XH \right) = \sum_{m=1}^{M_R} H_m^H X^H XH_m
\]  
(2.118)

\[
\text{Trace} \left( H^H X^H Y \right) = \sum_{m=1}^{M_R} H_m^H X^H Y_m
\]  
(2.119)

Applying (2.118) and (2.119) in (2.117) we get

\[
\hat{X} = \arg \min_X \left[ \sum_{m=1}^{M_R} H_m^H X^H XH_m - 2 \text{Real} \left( \sum_{m=1}^{M_R} H_m^H X^H Y_m \right) \right]
\]  
(2.120)
In case of one receive antenna, the minimization function reduces to

\[ H_1^H X^H X H_1 - 2 \text{Real}(H_1^H X y_1) \]

(2.121)

For multiple receivers, we can write the function for one receiver and the result is extended to achieve the ML decoding formulae for general case of \( M_{RX} \) receiving antennas. It is equivalent to maximum ratio combining (MRC). In case of ST coding with Alamouti structure, the above metric can be decomposed into two separate parts for detecting individual symbols i.e. ML decoding becomes single symbol decodable ML (SML) as shown in (2.14) and (2.15). In SF coding, single symbol ML decoder doesn’t yield optimum results because channel orthogonality is disturbed in case of frequency selective channels. In such cases, joint ML decoder (JML) is preferred which detects two symbols jointly as shown in (2.20). For MIMO channels, (2.115) can be rewritten as

\[ \hat{X} = \arg \min_X \left\{ \sum_{i=1}^{M_{RX}} y_i - \sum_{j=1}^{M_{TX}} h_{i,j} x_j \right\} \]

(2.122)

where \( h_{i,j} \) is the \( i^{th}, j^{th} \) component of the matrix \( H \). It is observed that the exact calculation of the right hand side of the above equations requires the summation of \( M_{RX} M_{TX} \) terms for each value of \( c \). The \( M_{TX} \times 1 \) vector \( c \) can take on \( 2^{BM_{TX}} \) values, where \( 2^B \) is the size of the modulation format used. This is the number of computations of the norm in (2.122) necessary for minimization if exhaustive search is used. The complexity of an ML receiver grows exponentially with \( 2^B \) and \( M_{TX} \) which restricts its implementation to small signal constellations and small numbers of transmit antennas.

2.4.6 Sphere Decoder (SD)

The main idea behind sphere decoding [55-58] approach is to limit the possible codewords by considering only those codewords that are within a sphere centered at the received signal vector. Sphere decoding method finds the transmitted signal vector with minimum ML metric. However it considers only a small set of vectors within a given sphere rather than all possible transmitted signal vectors. Sphere decoder adjusts the sphere radius to the extent that there exists a single vector (ML solution vector) within a sphere. It increases the radius when no vector exists within the sphere, and decreases the
radius when there exist multiple vectors within the sphere. In the sequel, we sketch the idea of the SD through an example. Consider a square QAM in a $2 \times 2$ MIMO channel. The complex system may be converted into an equivalent real system. Let $y_{jR}$ and $y_{jI}$ denote the real and imaginary parts of the received signal at the $j^{th}$ receiving antenna, that is, $y_{jR} = \text{Re}\{y_j\}$ and $y_{jI} = \text{Im}\{y_j\}$. Similarly, the input signal $x_i$ from the $i^{th}$ antenna can be represented by $x_{iR} = \text{Re}\{x_i\}$ and $x_{iI} = \text{Im}\{x_i\}$. For $2 \times 2$ MIMO channel, the received signal can be expressed in terms of its real and imaginary parts as follows:

\[
\begin{pmatrix}
y_{1R} + jy_{1I} \\
y_{2R} + jy_{2I}
\end{pmatrix} =
\begin{pmatrix}
h_{11R} + jh_{11I} & h_{12R} + jh_{12I} \\
h_{21R} + jh_{21I} & h_{22R} + jh_{22I}
\end{pmatrix}
\begin{pmatrix}
x_{1R} + jx_{1I} \\
x_{2R} + jx_{2I}
\end{pmatrix} +
\begin{pmatrix}
n_{1R} + jn_{1I} \\
n_{2R} + jn_{2I}
\end{pmatrix}
\quad (2.123)
\]

The real and imaginary parts of (2.123) can be separated as

\[
\begin{pmatrix}
y_{1R} \\
y_{2R}
\end{pmatrix} =
\begin{pmatrix}
h_{11R} & h_{12R} \\
h_{21R} & h_{22R}
\end{pmatrix}
\begin{pmatrix}
x_{1R} \\
x_{2R}
\end{pmatrix} -
\begin{pmatrix}
h_{11I} & h_{12I} \\
h_{21I} & h_{22I}
\end{pmatrix}
\begin{pmatrix}
x_{1I} \\
x_{2I}
\end{pmatrix} +
\begin{pmatrix}
n_{1R} \\
n_{2R}
\end{pmatrix}
\quad (2.124)
\]

and

\[
\begin{pmatrix}
y_{1I} \\
y_{2I}
\end{pmatrix} =
\begin{pmatrix}
h_{11R} & h_{12I} \\
h_{21R} & h_{22I}
\end{pmatrix}
\begin{pmatrix}
x_{1R} \\
x_{2R}
\end{pmatrix} -
\begin{pmatrix}
h_{11I} & h_{12R} \\
h_{21I} & h_{22R}
\end{pmatrix}
\begin{pmatrix}
x_{1I} \\
x_{2I}
\end{pmatrix} +
\begin{pmatrix}
n_{1I} \\
n_{2I}
\end{pmatrix}
\quad (2.125)
\]

Equations (2.125) and (2.124) can be combined to yield the following expression.

\[
\begin{pmatrix}
y_{1R} \\
y_{2R} \\
y_{1I} \\
y_{2I}
\end{pmatrix} =
\begin{pmatrix}
h_{11R} & h_{12R} & -h_{11I} & -h_{12I} \\
h_{21R} & h_{22R} & -h_{21I} & -h_{22I}
\end{pmatrix}
\begin{pmatrix}
x_{1R} \\
x_{2R} \\
x_{1I} \\
x_{2I}
\end{pmatrix} +
\begin{pmatrix}
n_{1R} \\
n_{2R} \\
n_{1I} \\
n_{2I}
\end{pmatrix}
\quad (2.126)
\]

\[
Y = \begin{bmatrix} 
X & N 
\end{bmatrix}
\]
For \(Y, H, X\) and \(N\) defined above, (2.115) can be rewritten as

\[
\hat{X} = \arg \min_X \| Y - XH \|_F^2
\]

\[
\| Y - XH \|_F^2 = \| Y - HX - H\hat{X} + H\hat{X} \|_F^2
\]

\[
= \left( Y - HX - H\hat{X} + H\hat{X} \right)^T \left( Y - HX - H\hat{X} + H\hat{X} \right)
\]

\[
= \{ (Y - H\hat{X})^T + (H\hat{X} - HX)^T \} \{ (Y - H\hat{X}) + (H\hat{X} - HX) \}
\]

\[
= (Y - H\hat{X})^T (Y - H\hat{X}) + (H\hat{X} - HX)^T (H\hat{X} - HX)
\]

\[
+ (H\hat{X} - HX)^T (Y - H\hat{X}) + (Y - H\hat{X})^T (H\hat{X} - HX)
\]

(2.128)

Since \(\hat{X}\) is the least square solution, \((H\hat{X} - HX)^T (Y - H\hat{X}) = (Y - H\hat{X})^T (H\hat{X} - HX) = 0\). Thus, (2.128) can be modified as

\[
\| Y - XH \|_F^2 = (Y - H\hat{X})^T (Y - H\hat{X}) + (H\hat{X} - HX)^T (H\hat{X} - HX)
\]

(2.129)

If we consider solution of (2.129) as \(\hat{X} = (H^T H)^{-1} H^T Y\), (2.129) becomes

\[
\left\{ Y - H(H^T H)^{-1} H^T Y \right\}^T \left\{ Y - H(H^T H)^{-1} H^T Y \right\} + (\hat{X} - X)^T H^T H (\hat{X} - X)
\]

(2.130)

First term of (2.130) can be written as

\[
\left\{ Y - H(H^T H)^{-1} H^T Y \right\}^T \left\{ Y - H(H^T H)^{-1} H^T Y \right\}
= Y^T \left\{ I - H(H^T H)^{-1} H^T \right\} \left\{ I - H(H^T H)^{-1} H^T \right\} Y
\]

\[
= Y^T \left\{ I - H(H^T H)^{-1} H^T \right\} \left\{ I - H(H^T H)^{-1} H^T \right\} Y
\]

\[
= Y^T \{ I - H(H^T H)^{-1} H^T \} \{ I - H(H^T H)^{-1} H^T \} Y
\]

(2.131)

Equation (2.131) is constant with respect to \(X\). Thus, from (2.131) and (2.130) our relation in (2.127) can be modified as

64
\[
\arg\min_x \|Y - XH\|_2^2 = \arg\min_x (X - \hat{X})^T H^T H (X - \hat{X})
\]  
(2.132)

where \(\hat{X} = \left(H^T H\right)^{-1} H^T Y\) is the unconstrained solution of real system shown in (2.126). Consider the following sphere with radius \(R_{SD}\).

\[
(X - \hat{X})^T H^T H (X - \hat{X}) \leq R_{SD}^2
\]  
(2.133)

The SD method takes into account the vectors inside a sphere defined by (2.133). Figure 2.5 illustrates a sphere with the center \(\hat{X} = \left(H^T H\right)^{-1} H^T Y\) and radius \(R_{SD}\).

![Figure 2.5: Original sphere in SD](image)

In such cases, this sphere includes four candidate vectors, one of which is the ML solution vector. We note that no vector outside the sphere can be ML solution vector because their ML metric values are bigger than the ones occurring inside the sphere. If we are able to choose the closest one among the four candidate vectors, the radius in (2.133) can be reduced in such a way that a sphere exist within which a single vector remains. In other words, the ML solution vector is now contained in this sphere with a reduced radius as shown in Figure 2.6.
Figure 2.6: New sphere with reduced radius

The new metric in (2.132) can be expressed as

\[
(X - \hat{X})^T H^T H (X - \hat{X}) = (X - \hat{X})^T R^T R (X - \hat{X}) = \|R(X - \hat{X})\|^2
\]  

(2.134)

where \( R \) is obtained from QR decomposition of the real channel matrix \( H=QR \). When \( M_{\text{Tx}} \) and \( M_{\text{Rx}} \) is 2, the metric in (2.134) can be expressed as

\[
\|R(X - \hat{X})\|^2 = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \end{bmatrix} \begin{bmatrix} X_1 - \hat{X}_1 \\ X_2 - \hat{X}_2 \\ X_3 - \hat{X}_3 \\ X_4 - \hat{\hat{X}_4} \end{bmatrix} \begin{bmatrix} X_1 - \hat{X}_1 \\ X_2 - \hat{X}_2 \\ X_3 - \hat{X}_3 \\ X_4 - \hat{\hat{X}_4} \end{bmatrix} 
\]

\[
= r_{44}(X_4 - \hat{X}_4)^2 + r_{33}(X_3 - \hat{X}_3)^2 + r_{24}(X_4 - \hat{X}_4)^2 \\
+ r_{22}(X_2 - \hat{X}_2)^2 + r_{23}(X_3 - \hat{X}_3)^2 + r_{24}(X_4 - \hat{X}_4)^2 \\
+ r_{11}(X_1 - \hat{X}_1)^2 + r_{12}(X_2 - \hat{X}_2)^2 + r_{13}(X_3 - \hat{X}_3)^2 + r_{14}(X_4 - \hat{X}_4)^2
\]

(2.135)

From (2.135) and (2.134), the sphere in (2.133) can be expressed as
\[\begin{align*}
&\left| r_{44}(X_4 - \hat{X}_4)^2 + r_{33}(X_3 - \hat{X}_3)^2 + r_{54}(X_4 - \hat{X}_4)^2 + r_{22}(X_2 - \hat{X}_2)^2 + r_{43}(X_4 - \hat{X}_4)^2 + r_{43}(X_4 - \hat{X}_4)^2 \right|^2 + r_{11}(X_1 - \hat{X}_1)^2 + r_{12}(X_2 - \hat{X}_2)^2 + r_{13}(X_3 - \hat{X}_3)^2 + r_{14}(X_4 - \hat{X}_4)^2 \leq R_{SD}^2
\end{align*}\]  
(2.136)

Using the sphere in (2.136), the SD method can be described with the following four steps.

**Step 1:** Referring to (2.136), a value for \(x_4\) is considered which is arbitrarily chosen from the points in the sphere \(\left| r_{44}(X_4 - \hat{X}_4) \right|^2 \leq R_{SD}^2\). In other words, this point must be chosen in the range given by

\[\hat{X}_4 - \frac{R_{SD}}{r_{44}} \leq X_4 \leq \hat{X}_4 + \frac{R_{SD}}{r_{44}}\]  
(2.137)

Let \(\hat{X}_4\) denote the point chosen in Step 1. If there exists no candidate point satisfying the inequalities, the radius needs to be increased. It is assumed that a candidate value was successfully chosen. Then we proceed to Step 2 as follows.

**Step 2:** Referring to (2.136), a candidate value for \(x_3\) is chosen from the points in the following sphere.

\[\left| r_{44}(\tilde{X}_4 - \hat{X}_4)^2 + r_{33}(X_3 - \hat{X}_3)^2 + r_{54}(\tilde{X}_4 - \hat{X}_4)^2 \right| \leq R_{SD}^2\]  
(2.138)

Thus, the point must be chosen in the following range:

\[\hat{X}_3 - \frac{\sqrt{R_{SD}^2 - |r_{44}(\tilde{X}_4 - \hat{X}_4)|^2} - r_{34}(\tilde{X}_4 - \hat{X}_4)}{r_{33}} \leq X_3 \leq \hat{X}_3 + \frac{\sqrt{R_{SD}^2 - |r_{44}(\tilde{X}_4 - \hat{X}_4)|^2} - r_{34}(\tilde{X}_4 - \hat{X}_4)}{r_{33}}\]  
(2.139)

Resultant \(\tilde{X}_4\) in (2.139) is already chosen in Step 1. If a candidate value for \(X_3\) does not exist, we go back to Step 1 and choose other candidate value of \(\tilde{X}_4\). Then again search for \(X_3\) that meets the inequalities in (2.139) for given \(\tilde{X}_4\). If no candidate value of
X₃ exists with all possible values of \( \tilde{X}_4 \), increase the radius of sphere and repeat the Step 1. Let \( \tilde{X}_4 \) and \( \tilde{X}_3 \) denote the final points chosen from step 1 and step 2 respectively.

**Step 3:** Given \( \tilde{X}_4 \) and \( \tilde{X}_3 \), a candidate value for \( X_2 \) is chosen from the points in the following sphere.

\[
\left| r_{44}(\tilde{X}_4 - \hat{X}_4) \right|^2 + \left| r_{33}(\tilde{X}_3 - \hat{X}_3) + r_{44}(\tilde{X}_4 - \hat{X}_4) \right|^2 \\
+ \left| r_{22}(X_2 - \hat{X}_2) + r_{23}(\tilde{X}_3 - \hat{X}_3) + r_{44}(\tilde{X}_4 - \hat{X}_4) \right|^2 \leq R_{SD}^2 \quad (2.140)
\]

Candidate value is chosen for \( X_2 \) inside the sphere of (2.140). If no candidate value for \( X_2 \) exists, then go to step 2 and choose another candidate value of \( \tilde{X}_3 \). In case if no candidate value for \( X_2 \) exists with all possible candidate values for \( \tilde{X}_3 \), go back to step 1 and choose another candidate value for \( \tilde{X}_4 \). The final points chosen from step 1 through step 3 are represented as \( \tilde{X}_4, \tilde{X}_3 \) and \( \tilde{X}_2 \).

**Step 4:** Now, a candidate value for \( X_1 \) is chosen from the points in the following sphere.

\[
\left| r_{44}(\tilde{X}_4 - \hat{X}_4) \right|^2 + \left| r_{33}(\tilde{X}_3 - \hat{X}_3) + r_{44}(\tilde{X}_4 - \hat{X}_4) \right|^2 + \left| r_{22}(X_2 - \hat{X}_2) + r_{23}(\tilde{X}_3 - \hat{X}_3) + r_{44}(\tilde{X}_4 - \hat{X}_4) \right|^2 \\
+ \left| r_{11}(X_1 - \hat{X}_1) + r_{12}(\tilde{X}_2 - \hat{X}_2) + r_{13}(\tilde{X}_3 - \hat{X}_3) + r_{14}(\tilde{X}_4 - \hat{X}_4) \right|^2 \leq R_{SD}^2 \quad (2.141)
\]

Candidate value satisfying (2.141) is chosen for \( X_1 \). If no candidate value for \( X_1 \) exists, go back to step 3 to choose other candidate value for \( \tilde{X}_2 \). In case if no candidate value for \( X_1 \) exists with all possible candidate values for \( \tilde{X}_2 \), go back to step 2 to choose another value for \( \tilde{X}_3 \). Let \( \tilde{X}_1 \) denote the candidate value for \( X_1 \). Hence the candidate values are \( \tilde{X}_4, \tilde{X}_3, \tilde{X}_2 \) and \( \tilde{X}_1 \). The corresponding radius value is calculated using (2.141). Using the new reduced radius, step 1 is repeated. If \( \tilde{X}_1, \tilde{X}_2, \tilde{X}_3 \) and \( \tilde{X}_4 \) turns out to
be a single point inside a sphere with that radius, this is declared as ML solution vector and the procedure is terminated.

The complexity of SD can be measured in terms of floating point operations (FLOPS) which include all arithmetic operations. For 2×2 MIMO system SD complexity depends on how well the initial radius is chosen. The initial radius can be determined in various ways like

\[
R_{SD}^2 = \sum_{i=1}^{4} \sum_{k=1}^{4} r_{ik} \left( \overline{X}_k - \hat{X}_k \right)^2
\]  

(2.142)

where \( \hat{X} = [\hat{X}_1, \hat{X}_2, \hat{X}_3, \hat{X}_4]^T \) is unconstrained least square solution, and \( \overline{X}_i = Q(\overline{X}_i) \) for \( i = 1, 2, 3, 4 \). The computation of the initial radius in (2.142) requires 14 real multiplications. Using this initial radius, the inequality conditions for the candidate selection for \( X_i, i = 4-s+1 \) in step \( s, s \in [1,4] \) can be generalized as

\[
\hat{X}_i + \frac{-\alpha_i - \beta_i}{r_{ii}} \leq X_i \leq \hat{X}_i + \frac{\alpha_i - \beta_i}{r_{ii}}
\]  

(2.143)

where \( \alpha_i = \sqrt{R_{SD}^2 - \sum_{k=1}^{4} \sum_{p=1}^{4} r_{kp} \left( \overline{X}_p - \hat{X}_p \right)^2} \) and \( \beta_i = \sum_{k=1}^{4} r_{ik} \left( \overline{X}_k - \hat{X}_k \right) \). In (2.143), \( \overline{X}_k \) are the selected symbols in the previous steps. Using the fact that each symbol \( \overline{X}_k \) is an integer, and substituting the results in the previous steps, one multiplication, two divisions, and one square root operation are required for calculating (2.143). In the first step \( (s=1) \), \( \beta_4 = 0 \) and thus, one division and one square root operation are required. To order to calculate the new radius using a new vector of length \( 2M_{Rx} = 4 \), only one multiplication is required, because the results in the previous steps can be reused. Table 2.4 summarizes the complexity of SD. In Table 2.4, the calculation \( \hat{X} = (H)^{-1}Y \) refers to the multiplication of \( (H)^{-1} \) and \( Y \), excluding the calculation of \( (H)^{-1} \). The ML complexity corresponds to the ML metric calculation of \( 16^2 = 256 \) times. Assuming that four real multiplications are required for each ML metric calculation, a total of \( 256 \times 4 = 1024 \) real multiplications are required. The main drawback of SD is that its complexity depends on
SNR. Further, the worst case complexity is same as that of ML detection, although the average complexity is significantly reduced.

**Table 2.4: Sphere decoder complexity**

<table>
<thead>
<tr>
<th></th>
<th>Multiplications</th>
<th>Divisions</th>
<th>Square Roots</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{X} = (H)^{-1} Y$</td>
<td>16</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$R_{SD}^2$ in (2.142)</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Step 1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Step 2-4 each</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>$R_{SD}^2$ update</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**2.5 SIMULATION RESULTS**

In this section, simulation results are obtained for Alamouti ST/SF/STF codes and QOSTBC, R-QOSTBC codes mentioned in section 2.1 and 2.3.

**2.5.1 Alamouti ST/SF/STF Results**

Various parameters used for simulation of MIMO-OFDM transceiver model described in Figure 2.1 are listed in Table 2.5. To illustrate analytical results derived in section 2.2, we plotted simulation results in terms of BER vs SNR. Simulations are done in two phases. While in phase 1, results are plotted for two transmit antennas and one receiving antenna, in phase 2, two transmit antennas and two receiving antennas are considered. In both cases the channel between transmitter and receiver antenna is assumed to be quasi-static Rayleigh flat and selective channel with higher tap order. The channel is modeled by Jake’s model which takes into account the effects of Doppler shift and Doppler spread existing due to relative motion between transmitter and receiver. Further, the channel experienced by each transmitting antenna is assumed to be independent. In addition, the transmitting power of each antenna is considered equal. The receiver is assumed to have perfect knowledge of the channel and transmitter does not have CSI information.
Table 2.5: Simulation parameters for Alamouti ST/SF/STF codes

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Bandwidth</td>
<td>1MHz</td>
</tr>
<tr>
<td>Number of transmitting antenna</td>
<td>2</td>
</tr>
<tr>
<td>Number of receiving antenna</td>
<td>1 and 2</td>
</tr>
<tr>
<td>Maximum Doppler freq.(f_{\text{d}})</td>
<td>200 Hz</td>
</tr>
<tr>
<td>Channel model used [39]</td>
<td>Six-ray urban TU channel model</td>
</tr>
<tr>
<td>Channel coefficient's amplitudes</td>
<td>0.2, 0.5, 0.4, 0.1, 0.06, 0.4</td>
</tr>
<tr>
<td>Channel delay spreads (in (\mu) secs)</td>
<td>0, 0.2, 0.5, 1.6, 2.3, 5</td>
</tr>
<tr>
<td>Carrier modulation used</td>
<td>BPSK</td>
</tr>
<tr>
<td>Spectral efficiency</td>
<td>1bit/sec/Hz</td>
</tr>
<tr>
<td>Number of data subcarriers</td>
<td>128</td>
</tr>
<tr>
<td>IFFT size</td>
<td>128</td>
</tr>
<tr>
<td>Cyclic prefix length</td>
<td>32</td>
</tr>
<tr>
<td>Number of channel Taps</td>
<td>12 and 34 for flat and selective channels</td>
</tr>
<tr>
<td>Channel Fading</td>
<td>Rayleigh independent frequency flat and selective fading i.e. rank (\mathcal{R}_T = I_k) (identity matrix). Therefore, the maximum achievable diversity in flat and selective channels with 2×1 configuration is 24 and 68 respectively. SF and STF schemes have same spectral efficiency because full diversity STF coding is obtained from full diversity SF coding via repetition mapping.</td>
</tr>
</tbody>
</table>

Figure 2.7 shows BER performance of Alamouti ST codes with ZF and MMSE equalizers. The BER curve is decreasing monotonically with increase in SNR for both equalizers. BER for MMSE equalizer is better, but performance gain is not that significant. In subsequent results BER improvement is shown for equalizers like DFE, ML and SD. Figures 2.8 and 2.9 shows BER performance of 2×1 MIMO-OFDM system with different equalizers in frequency flat and frequency selective channels. The STF coding scheme clearly dominates ST and SF coding. In other words, STF coding provides higher diversity order than other schemes. In section 2.2, it is explained that maximum achievable diversity in case of STF coding is \(LM_{R_x}M_{T_x}\), assuming independent channel fading i.e. temporal coefficient \(\mathcal{R}_T = I_k\) (identity matrix). Therefore, the maximum achievable diversity in flat and selective channels with 2×1 configuration is 24 and 68 respectively. SF and STF schemes have same spectral efficiency because full diversity STF coding is obtained from full diversity SF coding via repetition mapping.
Figure 2.7: BER comparison for 2×1 Alamouti ST coded MIMO-OFDM system with MMSE and ZF equalizers

Figure 2.8: BER comparison for 2×1 MIMO-OFDM system in quasi-static Rayleigh frequency flat channel for ST/SF/STF codes with different equalizers
Figure 2.9: BER comparison for 2×1 MIMO-OFDM system in quasi-static Rayleigh frequency selective channel for ST/SF/STF codes with different equalizers

Among equalizers, STF with SD outperforms all other combinations i.e. STF-ML and STF-DFE by almost 0.5dB to 1dB. Among all coding schemes, STF-SD outperforms SF-SD and ST-SD by approximately 2dB to 3dB. Thus, we conclude that STF coding with SD is the best combinations among all coding schemes. Figure 2.10 and 2.11 shows BER performance of 2×2 MIMO-OFDM system in both flat and selective channels. The achievable diversity is of 48 and 136 in case of flat and selective channels for 2×2 configuration, higher than that of 24 and 68 with 2×1 configuration. The results also show that employing 2 antennas on receiver side greatly enhance the system performance with increased diversity order. Further, the BER performance is better in flat channels than selective channels. Hence STF-SD is the best coding scheme in both cases.

Figure 2.12 shows the BER comparison of ST, SF and STF codes with existing codes. Results show that existing codes dominate the performance in lower SNR region. In higher SNR region, our ST, SF and STF codes outperforms existing counterparts as the impact of higher diversity order is more compared to precoding matrix.
Figure 2.10: BER comparison for 2×2 MIMO-OFDM system in quasi-static Rayleigh frequency flat channel for ST/SF/STF codes with different equalizers

Figure 2.11: BER comparison for 2×2 MIMO-OFDM system in quasi-static Rayleigh frequency selective channel for ST/SF/STF codes with different equalizers
Figure 2.12: BER comparison for various rate-1 STF and SF codes for 2×1 MIMO-OFDM system in Rayleigh frequency flat channel

The superiority of our STF code over that of [39] and [41] is evident from Figure 2.12. In Figure 2.12, for almost every value of SNR, our STF code outperforms codes given in [39] and [41] by almost 2-3 dB. Our SF codes also dominates SF codes in [36] and [41] by 1-1.5 dB. The comparison results are presented in Table 2.6.

Table 2.6: Result comparison for various rate-1 STF and SF codes for 2×1 MIMO-OFDM system with our codes

<table>
<thead>
<tr>
<th>Code</th>
<th>BER $= 10^{-4}$ at SNR</th>
<th>Diversity Order</th>
<th>Decoder Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate-1 ST code [ours]</td>
<td>18 dB</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Alamouti ST code [20]</td>
<td>&gt; 25 dB</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Rate-1 SF code [ours]</td>
<td>16.5 dB</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>SF code [36]</td>
<td>19 dB</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>SF code [41]</td>
<td>18.5 dB</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Rate-1 STF code [ours]</td>
<td>15 dB</td>
<td>Maximum</td>
<td>Moderate</td>
</tr>
<tr>
<td>STF code [39]</td>
<td>16 dB</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>STF code [41]</td>
<td>18 dB</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>
2.5.2 Quasi and Rotated Quasi OSTBC Results

The list of simulation parameters used for simulating QOSTBC and R-QOSTBC are shown in Table 2.6. Initially, the BER performance of Alamouti coding with BPSK, QPSK and 16-QAM modulations is shown in Figure 2.13. Each Figure has three subplots correspond to BPSK, QPSK and 16-QAM modulations. Figure 2.13 shows BER performance of Alamouti coding with different modulations. The BER decreases monotonically to $10^{-4}$ at SNR of about 19dB, 22dB and 23dB for BPSK, QPSK and 16-QAM modulation techniques respectively. Therefore the BER performance is better in Alamouti scheme with BPSK modulations.

Table 2.7: Simulation parameters for QOSTBC and R-QOSTBC

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Bandwidth</td>
<td>20MHz</td>
</tr>
<tr>
<td>Number of transmitting antenna</td>
<td>2 and 4</td>
</tr>
<tr>
<td>Number of receiving antenna</td>
<td>1 and 2</td>
</tr>
<tr>
<td>No. of Doppler shift ($N_0$)</td>
<td>8</td>
</tr>
<tr>
<td>Maximum Doppler freq. ($f_m$)</td>
<td>200 Hz</td>
</tr>
<tr>
<td>Maximum Doppler shift</td>
<td>$2\pi f_m = 1.256 \times 10^3$</td>
</tr>
<tr>
<td>Channel model used [39]</td>
<td>Six-ray urban TU channel model</td>
</tr>
<tr>
<td>Carrier modulation used</td>
<td>BPSK, QPSK and 16-QAM</td>
</tr>
<tr>
<td>Sampling frequency ($f_s$)</td>
<td>8000Hz</td>
</tr>
<tr>
<td>Sampling time ($t_s$)</td>
<td>$1/f_s$</td>
</tr>
<tr>
<td>Optimum rotation angle</td>
<td>$\pi/2$ (for BPSK), $\pi/4$ (QPSK), $\pi/4$ (16-QAM)</td>
</tr>
</tbody>
</table>
Figure 2.13: BER comparison for Alamouti STBC encoded MIMO system for various modulation schemes.

Figure 2.14 and 2.15 shows BER performance with QOSTBC and R-QOSTBC techniques. In Figure 2.14, BER is reduced to $10^{-4}$ at SNR of about 17dB, 19dB and 20dB for BPSK, QPSK and 16-QAM respectively. In Figure 2.15, BER is reduced to $10^{-4}$ at SNR of about 16dB, 17dB and 18dB for BPSK, QPSK and 16-QAM techniques respectively. Therefore by comparing results of Figures 2.14 and 2.15 it can be concluded that results are better with BPSK modulation among all cases. Also, the results using R-QOSTBC shows better performance then QOSTBC due to full rate and full diversity exhibited by R-QOSTBC. Figure 2.16 compares different STBC techniques for BPSK modulation and results that R-QOSTBC with optimal rotation gives better performance among other STBC techniques due to its full diversity and full transmission rate.
Figure 2.14: BER comparison for QOSTBC encoded MIMO system for various modulation schemes

Figure 2.15: BER comparison for R-QOSTBC encoded MIMO system for various modulation schemes
2.6 CONCLUSION

In this chapter, a general performance analysis for MIMO-OFDM systems with Alamouti ST, SF and STF coding is provided for both quasi static Rayleigh frequency flat and selective channel with higher tap order. Employing higher tap order helps in to achieve higher diversity order, which further helps to combat deep fades which occur due to harsh wireless environments. We determined the maximum achievable diversity with Alamouti ST, SF and STF codes. It is concluded that we achieved full rate and full diversity even in such channel environments using Alamouti STF codes. Employing higher tap order increases decoder complexity which is resolved by employing SD at receiver. We also studied the performance of STBC encoded MIMO systems under mobile radio channels using jakes model. Further, system performance is compared with different STBC techniques under different modulation schemes. It is concluded that R-QOSTBC with BPSK modulation shows full rate and full diversity. R-QOSTBC also achieves lower decoder complexity using ML decoder.