5. Space-Time Trellis Codes

- 5.1 Introduction
- 5.2 Space-Time Coded Systems
- 5.3 Space-Time Code Word Design Criteria
- 5.4 Design of Space-Time Trellis Codes on Slow Fading Channels
- 5.5 Design of Space-Time Trellis Codes on Fast Fading Channels
- 5.6 Performance Analysis in a Slow Fading Channels
- 5.7 Performance Analysis in a Fast Fading Channel
- 5.8 The Effect of Imperfect Channel Estimation on Code Performance
- 5.9 Effect of Antenna Correlation on Performance
- 5.10 Delay Diversity as an STTC
- 5.11 Comparison of STBC and STTC
- 5.12 Simulation Exercises
- References
5.1 Introduction

- In Chapter 4, we discussed space-time block codes (STBC). These codes provided maximum diversity advantage using simple decoding techniques. However, STBC did not provide coding gain, and nonfullrate STBC introduces bandwidth expansion. In view of this, it becomes worthwhile to consider a joint design of error control coding, modulation, transmit, and receive diversity to develop an effective signaling scheme called space-time trellis codes (STTC), which is able to combat effects of fading. This concept was first introduced by Tarokh, Seshadri, and Calderbank. STTC can simultaneously offer coding gain with spectral efficiency and full diversity over fading channels. The coding gain is achieved through the inherent nature of the STTC itself and is distinct from the coding gain achieved by temporal block codes and convolution codes.

- In this chapter, we explore the basic theory leading to such code design using M-PSK schemes for various numbers of transmit antennas and spectral efficiencies, in both slow as well as fast fading channels. The code performance is examined with simulations and the effects of imperfect channel estimates and correlation are also considered.
5.2 Space-Time Coded Systems

- Consider a baseband space-time coded system with $M_T$ transmit antennas and $M_R$ receive antennas.
Figure 5.1 A block diagram of space-time trellis encoder
At each instant $t$, a block of $m$ binary information symbols denoted by

$$c_t = (c_t^1, c_t^2, \ldots, c_t^m)$$  \hspace{1cm} (5.1)$$
is fed into the space-time encoder.

The space-time encoder maps the block of $m$ binary input data into $M_T$ modulation symbols from a signal set of $M = 2^m$ points. The coded data are applied to a serial-to-parallel (S/P) converter to produce a sequence of $M_T$ parallel symbols, arranged as a $M_T \times 1$ column vector

$$s_t = (s_t^1, s_t^2, \ldots, s_t^{M_T})^T$$ \hspace{1cm} (5.2)$$

The $M_T$ parallel outputs are simultaneously transmitted from all the $M_T$ antennas, whereby symbol $s_t^i$, $1 \leq i \leq M_T$ is transmitted by antenna $i$ and all transmitted symbols have the same duration $T$ sec.

The vector of coded modulation symbols from different antennas, as shown in (5.2), is called as space-time symbol.
The spectral efficiency of the system is

\[ \eta = \frac{r_b}{B} = m \quad b/s/Hz \]  

(5.3)

where \( r_b \) is the data rate and \( B \) is the channel bandwidth.

- The spectral efficiency in (5.3) is equal to the spectral efficiency of a reference uncoded system with one transmit antenna.
The multiple antennas at both the transmitter and receiver create a MIMO channel. Assume that flat fading between each transmit and receive antenna and the channel is memoryless.

The channel matrix at any given time $t$ is given by

$$
H_t = \begin{bmatrix}
    h_{1,1}^t & h_{1,2}^t & \cdots & h_{1,M_r}^t \\
    h_{2,1}^t & h_{2,2}^t & \cdots & h_{2,M_r}^t \\
    \vdots & \vdots & \ddots & \vdots \\
    h_{M_t,1}^t & h_{M_t,2}^t & \cdots & h_{M_t,M_r}^t
\end{bmatrix}
$$

(5.4)

where the $j$th element, denoted by $h_{j,i}^t$, is the fading attenuation coefficient for the path from transmit antenna $i$ to receive antenna $j$.

The coefficients in (5.4) are assumed to be i.i.d Gaussian.

- Slow fading channel (i.e, the fading coefficients are constant during a frame and vary from one frame to another). This is also called quasi-static fading.

- Fast fading channel (i.e., the fading coefficients are constant within each symbol period and vary from one symbol to another).
At time $t$ the received signal at antenna $j$ ($j = 1,2,\ldots,M_R$) denoted by $r^j_t$ is given by

$$r^j_t = \sum_{i=1}^{M_T} h^j_i s^i_t + n^j_t$$

(5.5)

where $n^j_t$ is the noise component of receive antenna $j$ at time $t$, which is also i.i.d. Gaussian.

We represent

$$r_t = (r^1_t, r^2_t, \ldots, r^{M_R}_t)$$

(5.6)

and

$$n_t = (n^1_t, n^2_t, \ldots, n^{M_R}_t)$$

(5.7)

Thus the received signal vector can be represented as

$$r_t = H_t s_t + n_t$$

(5.8)
The decoder uses a maximum likelihood algorithm to estimate the transmitted information sequence and we assume that the receiver has complete knowledge of the channel. The transmitter, however, has no knowledge of the channel. The decision metric is computed based on the squared Euclidean distance between the received sequence as measured and the actual received sequence, as

$$\sum_{t}^{M_{k}} \left| r'_t - \sum_{i=1}^{M_{r}} h'_j s^i_t \right|^2$$

(5.9)

The decoder selects a code word with the minimum decision metric as the decoded sequence. This decoder is implemented as a Viterbi decoder.
5.3 Space-Time Code Word Design Criteria

- Assume that the transmitted data frame length is $L$ symbols long for each antenna. This leads to a space-time code word matrix $M_{T \times L}$,

$$S = [s_1, s_2, \ldots, s_L] = \begin{bmatrix}
    s_1^1 & s_2^1 & \cdots & s_L^1 \\
    s_1^2 & s_2^2 & \cdots & s_L^2 \\
    \vdots & \vdots & \cdots & \vdots \\
    s_1^{M_T} & s_2^{M_T} & \cdots & s_L^{M_T}
\end{bmatrix}$$

where each row corresponds to the data sequence transmitted from each antenna and each column is the space-time symbol at time $t$. 
The pairwise error probability (PEP) is the probability that the maximum likelihood decoder selects as its estimate a signal \( e = e_1^1 e_1^2 \ldots e_1^{MT} e_2^1 \ldots e_2^{MT} e_L^1 \ldots e_L^{MT} \) when in fact the signal \( s = s_1^1 s_1^2 \ldots s_1^{MT} s_2^1 \ldots s_2^{MT} s_L^1 \ldots s_L^{MT} \) was transmitted. This will occur if, summing over all symbols, antennas, and time periods

\[
\sum_{i=1}^{L} \sum_{j=1}^{M} \left| r_i^j - \sum_{i=1}^{M} h_{j,i}^i s_i^j \right|^2 \geq \sum_{i=1}^{L} \sum_{j=1}^{M} \left| r_i^j - \sum_{j=1}^{M} h_{j,i}^i e_i^j \right|^2
\]

which can be rewritten as

\[
\sum_{i=1}^{L} \sum_{j=1}^{M} 2 \Re \left( \eta_i^j \right) \sum_{i=1}^{M} h_{j,i}^i \left( e_i^j - s_i^j \right) \geq \sum_{i=1}^{L} \sum_{j=1}^{M} \left| \sum_{i=1}^{M} h_{j,i}^i \left( e_i^j - s_i^j \right) \right|^2
\]

(5.11)

where \( \Re \{ . \} \) refers to the real part of the argument.
If we assume that the receiver has perfect knowledge of the channel, then for a given instance of channel path gains \( \{h_{i,j}\} \), the term on the right of (5.11) is a constant equal to \( d^2(e,s) \) and the term on the left is a zero-mean Gaussian random variable with variance \( 4\sigma^2 d^2(e,s) \). Hence, the PEP conditioned on knowing \( \{h_{i,j}\} \) is given by

\[
P(s \rightarrow e \mid h_{i,j}, i = 1,2,\ldots,M_T, j = 1,2,\ldots,M_R) = Q\left( \frac{d(s,e)}{2\sigma} \right) \leq \exp\left( -\frac{d^2(s,e)E_s}{4N_0} \right)
\]
Now, $d^2(e,s)$ can be rewritten

$$d^2(s,e) = \sum_{i=1}^{L} \sum_{j=1}^{M_T} \sum_{i' = 1}^{M_T} h_{i,j} h_{i',j}^* (s_i' - e_i') (s_i' - e_i')^*$$

(5.12)

If we denote $\Omega_j = (h_{1,j}, h_{2,j}, \ldots, h_{M_T,j})$, we can rewrite

$$d^2(s,e) = \sum_{j=1}^{M_T} \Omega_j A \Omega_j^+$$

(5.13)

where $\Omega_j^+$ denotes the Hermitian transpose of $\Omega_j$ and $A = A(e,s)$ is an $M_T \times M_T$ matrix independent of time and contains entries.

The PEP becomes

$$P(s \rightarrow e \mid h_{i,j}, i = 1,2,\ldots,M_T, j = 1,2,\ldots,M_R) \leq \prod_{j=1}^{M_R} \exp \left( - \frac{\Omega_j A \Omega_j^+ E_j}{4N_0} \right)$$

(5.14)
Since $A$ is Hermitian, there exists a unitary matrix $V$ that satisfies $VV^H = I$ and a real diagonal matrix $D$ such that

$$VAV^* = D$$

The rows $\{v_1, v_2, \ldots, v_{M_T}\}$ of $V$ are the eigenvectors of $A$ and form a complete orthonormal basis of an $M_T$-dimensional vector space. Furthermore, the diagonal elements of $D$ are the eigenvalues $\lambda_i$, $i=1,2,\ldots,M_T$ of $A$, including multiplicities, and are nonnegative real numbers since $A$ is Hermitian. By construction of $A$, the code word difference matrix, $B$ where

$$B(s,e) = \begin{pmatrix} e_1^1 - s_1^1 & e_1^1 - s_1^1 & \cdots & e_L^1 - s_L^1 \\ e_1^2 - s_1^2 & e_2^2 - s_2^2 & \cdots & e_L^2 - s_L^2 \\ \vdots & \vdots & \ddots & \vdots \\ e_1^{M_T} - s_1^{M_T} & e_2^{M_T} - s_2^{M_T} & \cdots & e_L^{M_T} - s_L^{M_T} \end{pmatrix}$$

is a square root of $A$ (i.e., $A = BB^H$).
Next, we express $d^2(s,e)$ in terms of the $\{\lambda_j\}$'s. Let $(\beta_{1,j}, \beta_{2,j}, \ldots, \beta_{M_T,j}) = \Omega_j V^+$, then

$$\Omega_j A \Omega_j^+ = \Omega_j V^+ D V \Omega_j^+ = \sum_{i=1}^{M_T} \lambda_i |\beta_{i,j}|^2 \Rightarrow d^2(s,e) = \sum_{i=1}^{M_T} \sum_{j=1}^{M_T} \lambda_i |\beta_{i,j}|^2$$

(5.15)

Since $h_{i,j}$ are samples of a complex Gaussian random variable with mean $E h_j$, let

$$K^j = (E h_{1,j}, E h_{2,j}, \ldots, E h_{M_T,j})$$

Since $V$ is unitary, this implies $\beta_{i,j}$ are independent complex Gaussian random variables with variance 0.5 per dimension and with mean $K^j v_i$. 
5.4 Design of Space-Time Trellis Codes on Slow Fading Channels

- 5.4.1 Error Probability on Slow Fading Channels
- 5.4.2 Design Criteria for Slow Rayleigh Fading STTCs
- 5.4.3 Encoding/Decoding of STTCs for Quasi-Static Flat Fading Channels
- 5.4.4 Code Construction for Quasi-Static Flat Fading Channels
- 5.4.5 Example Using 4PSK
5.4.1 Error Probability on Slow Fading Channels

- In the case of flat Rayleigh fading, \( E h_{i,j} = 0 \) for all \( i \) and \( j \). To obtain an upper bound on the average probability of error, we average

\[
\prod_{j=1}^{M_j} \exp \left( - \left( \frac{E_s}{4N_0} \sum_{i=1}^{M_i} \lambda_i |\beta_{i,j}|^2 \right) \right)
\]

(5.16)

With respect to the independent Rayleigh distributions of \( |\beta_{i,j}| \) with probability density

\[
p(|\beta_{i,j}|) = 2 |\beta_{i,j}| \exp(-|\beta_{i,j}|^2).
\]
Letting $c = \frac{E_s}{4N_0}$, we get

$$E\prod_{j=1}^{M_R} e^{-c\sum_{i=1}^{M_T} |\varphi_{j,i}|^2} = \prod_{j=1}^{M_R} \prod_{i=1}^{M_T} E\left(e^{-c\lambda_{j,i}}\right)$$

By independence and

$$E\left(e^{-c\lambda_{j,i}}\right) = 2\int_0^\infty e^{-cw} we^{-w^2} dw = \frac{1}{1 + c\lambda_i}$$

so,

$$P(s \to e) \leq \frac{1}{\left(\prod_{i=1}^{M_T} \left(1 + \frac{E_s}{4N_0} \lambda_i\right)\right)^{M_R}}$$

(5.17)
Let $r$ be the rank of $A$. Then the eigenvalue $0$ has a multiplicity of $M_T - r$. Let the nonzero eigenvalues of $A$ be $\lambda_1, \lambda_2, ..., \lambda_r$. Then for high enough SNR, the PEP in (5.17) becomes

$$P(s \rightarrow e) \leq \left( \prod_{i=1}^{r} \lambda_i \right)^{-M_R} \left( \frac{E_s}{4N_0} \right)^{-sM_R} = \left( \prod_{i=1}^{r} \lambda_i^{1/r} \right)^{-M_R} \left( \frac{E_s}{4N_0} \right)^{-sM_R}$$

(5.18)

Thus a diversity advantage of $M_R r$ and a coding advantage of $(\prod_{i=1}^{r} \lambda_i)^{1/r}$ are obtained.
5.4.2 Design Criteria for Slow Rayleigh Fading STTCs

- Rank and determinant criteria
  - The rank criterion: To maximize the minimum rank $r$ of the matrix $B$ over all pairs of distinct code words.
  - The determinant criterion: To maximize the minimum determinant $\prod_{i=1}^{r} \lambda_i$ of the matrix $B$ along the pairs of distinct code words with that minimum rank.

- Maximizing the minimum rank $r$ of the matrix $B$ implies that we need to find a space-time code that achieves the full rank of the matrix $B$ (i.e., $M_T$). This is not always achievable due to the restriction of the trellis code structure.
• For a space-time trellis code with memory order of \(v\), the length of an error event, denoted by \(L\), can be lower bounded as

\[
L \geq \lceil \frac{v}{2} \rceil + 1
\]

• For an error event path of length \(L\) in the trellis, \(B(s,e)\) is a matrix of size \(M_T \times L\), which results in the maximum achievable rank of \(\min(M_T, L)\). This yields the upper bound as given by \(\min(M_T, \lceil \frac{v}{2} \rceil + 1)\).

• There is an interaction between the maximum achievable rank, the number of transmit antennas, and the memory order of an STTC.
Table 5.1 Upper bound of the rank values for STTC

<table>
<thead>
<tr>
<th></th>
<th>$M_T = 2$</th>
<th>$M_T = 3$</th>
<th>$M_T = 4$</th>
<th>$M_T = 5$</th>
<th>$M_T \geq 6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v = 2$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$v = 3$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$v = 4$</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$v = 5$</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$v = 6$</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
Substituting (5.15) in (5.11),

\[
P(s \rightarrow e \mid H) \leq \frac{1}{2} \exp \left( - \sum_{j=1}^{M_R} \sum_{i=1}^{M_T} \lambda_i |\beta_{i,j}|^2 \frac{E_s}{4N_0} \right)
\]  

(5.21)

|\beta_{j,i}|^2 follows the central Chi-square distribution with the mean value and the variance given by

\[
\begin{align*}
\mu_{|\beta_{j,i}|^2} &= 1 \\
\sigma_{|\beta_{j,i}|^2}^2 &= 1
\end{align*}
\]

For a large \(rM_R (>3)\) value according to the central limit theorem, the expression

\[
\sum_{j=1}^{M_R} \sum_{i=1}^{M_T} \lambda_i |\beta_{i,j}|^2
\]  

(5.22)

approaches a Gaussian random variable \(D\) with the mean value and the variance

\[
\begin{align*}
\mu_D &= M_R \sum_{i=1}^{M_T} \lambda_i \\
\sigma_D^2 &= M_R \sum_{i=1}^{M_T} \lambda_i
\end{align*}
\]
Thus the unconditional pairwise error probability can be upper-bounded by

\[
P(s \rightarrow e \mid H) \leq \frac{1}{2} \int_{d=0}^{\infty} \exp \left(-\frac{E_s}{4N_0} \right) p(D) dD
\]

(5.23)

\[
\Rightarrow P(s \rightarrow e \mid H) \leq \frac{1}{4} \exp \left(-M_r \frac{E_s}{4N_0} \sum_{i=1}^{M_T} \lambda_i \right)
\]

(5.24)
To minimize the PEP, it follows from (5.24) that the sum of the eigenvalues of matrix $A(s,e)$ where $A(s,e) = B(s,e)B^*(s,e)$ should be maximized. For a square matrix the sum of all the eigenvalues is equal to the trace of the matrix, denoted by $\text{tr}(v)$. It can be written as

$$\text{tr}(v) = \sum_{i=1}^{M_r} \lambda_i = \sum_{i=1}^{M_r} A_{ii}$$

(5.25)

where $A_{i,i}$ are the elements on the main diagonal of matrix $A(s,e)$. The trace of matrix $A(s,e)$ can be expressed as

$$\text{tr}(v) = \sum_{i=1}^{M_r} \sum_{j=1}^{L} |e_j^i - s_j^i|^2$$

(5.26)

As (5.26) shows, the trace of $A(s,e)$ is equivalent to the Euclidian distance between code words $s$ and $e$ over all the transmit antennas. In other words, the pairwise error probability is minimized if the Euclidian distance is maximized. This is consistent with the conclusions on the convergence of a fading channel to an AWGN channel for a large number of diversity branches and with the design criteria for trellis coded modulation on fading channels. Note that a larger $rM_R$ value provides faster convergence to an ideal AWGN channel.
Summarizing, if $rM_R$ is sufficiently large ($>3$), the performance of STTC codes is dominated by the minimum trace of $A(s,e)$ taken over all pairs of distinct code words $s$ and $e$. 
Trace criterion

- The coding gain is maximized if the minimum trace of $A(s,e)$ over all code word pairs is maximized.
- Note that the condition of the trace criterion, $rM_R > 3$, is met for most of the combinations of the transmit and receive antenna numbers.

In most cases, the condition $rM_R > 3$ can translate into $M_TM_R > 3$. When $M_T = 2$, it is important for $A(s,e)$ to have full rank of $r=2$. When $M_T \geq 4$, the full rank requirement is not necessary to minimize the PEP. This is because, in such cases, the largest minimum trace dominates the performance of the code.

When $rM_R$ is small ($<4$), (5.22) no longer approaches a Gaussian distribution. In this case one can follow the rank and determinant criteria. Simulations have justified this, as it was found that as long as $rM_R \geq 4$, the best codes based on the trace criterion outperform the best codes based on the rank and determinant criterion.
### Table 5.2 Parameters of the codes with two transmit antennas

<table>
<thead>
<tr>
<th></th>
<th>( (a_0^1, a_0^2) )</th>
<th>( (a_1^1, a_1^2) )</th>
<th>( (b_0^1, b_0^2) )</th>
<th>( (b_1^1, b_1^2) )</th>
<th>det</th>
<th>tr</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code A</td>
<td>(0, 2)</td>
<td>(2, 0)</td>
<td>(0, 1)</td>
<td>(1, 0)</td>
<td>4.0</td>
<td>4.0</td>
<td>2</td>
</tr>
<tr>
<td>Code B</td>
<td>(0, 2)</td>
<td>(1, 2)</td>
<td>(2, 3)</td>
<td>(2, 0)</td>
<td>4.0</td>
<td>10.0</td>
<td>2</td>
</tr>
<tr>
<td>Code C</td>
<td>(0, 2)</td>
<td>(2, 2)</td>
<td>(2, 3)</td>
<td>(0, 2)</td>
<td>0</td>
<td>10.0</td>
<td>1</td>
</tr>
</tbody>
</table>
Consider three QPSK codes. These are 4-state codes with two transmit antennas. The three codes are denoted by A, B, and C, respectively. These codes have the same bandwidth efficiency of 2 bit/s/Hz. It is shown that codes A and B have full rank and the same determinant of 4, whereas code C is not of full rank and, therefore, its determinant is 0. On the other hand, the minimum trace for codes B and C is 10, whereas code A has a smaller minimum trace of 4.
Figure 5.2 FER performance of the 4-state space-time coded QPSK with two transmit antennas. The solid line indicates one receive antenna. The dash indicates four receive antennas.
Simulation results

- Codes A and B outperform code C if one receive antenna is employed, which indicates that the minimum rank is much more important in determining the code performance for systems with a small number of independent subchannels.

- When the number of receive antennas is four, code C performs better than code A, which indicates that the minimum trace is much more important in determining the code performance for systems with a large number of independent subchannels.

- Code B is better than code C, although they have the same minimum trace. This is due to the fact that code B has a larger rank than code C.
Figure 5.3 Flow chart: is code A better than code B?
The value of $rM_R$ determines whether the rank and determinant criteria or the trace criteria should be used.

- When $rM_R < 4$, the rank and determinant criteria are applicable.
- When $rM_R > 3$, the trace criterion is applicable.
- However, in the code design, the number of receive antennas is not considered a design parameter.
Figure 5.4 The boundary for applicability of the TSC and the trace criteria

(a) $2 \leq v \leq 5$

(b) $v \leq 6$
- The points in the rectangular blocks are the cases where rank and determinant criteria are to be employed.
- The trace criterion can be used for all other cases.
- The rank and determinant criteria only apply to the systems with one receive antenna.
5.4.3 Encoding/Decoding of STTCs for Quasi-Static Flat Fading Channels

- The encoding for STTCs is similar to trellis coded modulation except that at the beginning and end of each frame, the encoder is required to be in the zero state.
- At each time t, depending on the state of the encoder and the input bits, a transition branch is selected. If the label of the transition branch is \( s_t^1, s_t^2, \ldots, s_t^{MT} \), then transmit antenna I is used to send the constellation symbol \( s_t^i, i = 1, 2, \ldots, M_T \) and all these transmissions are in parallel. The encoder coefficient set, denoted by

\[
g' = [(g'_{0,1}, g'_{0,2}, \ldots, g'_{0,M_r}),(g'_{1,1}, g'_{1,2}, \ldots, g'_{1,M_r}), \ldots,(g'_{v_1,1}, g'_{v_1,2}, \ldots, g'_{v_1,M_r})]\]  

is usually found next to the trellis diagram of the trellis code. Each \( g'_{i,L,k} \) is an element of the 4-PSK constellation set \( \{0,1,2,3\} \) and \( v_i \) is the memory order of the \( i \)th shift register. Multiplier outputs are added modulo 4.
Figure 5.5 Space-time trellis code encoder for 4PSK
The encoder can be described in generator polynomial format. The input binary sequence to the upper shift register can be represented as

$$u^1(D) = u_0^1 + u_1^1 D + u_2^1 D^2 + u_3^1 D^3 + ...$$  \hspace{1cm} (5.28)

Similarly, the binary input sequence to the lower shift register can be written as

$$u^2(D) = u_0^2 + u_1^2 D + u_2^2 D^2 + u_3^2 D^3 + ...$$  \hspace{1cm} (5.29)

where $u_{j}^{k}$, $j = 0,1,2,3,...$, $k = 1,2$ are binary symbols 0,1.
The feed-forward generator polynomial for the upper encoder and transmit antenna i, where i=1,2 can be written as

\[ G_i^1(D) = g^1_{0,i} + g^1_{1,i} D + \ldots + g^1_{v_1,i} D^{v_1} \]  \hspace{1cm} (5.30)

where \( g^1_{j,i}, j=0,1,\ldots,v_1 \) are nonbinary coefficients that can take values from a constellation such as 4PSK as 1, -j, -1, j and \( v_1 \) is the memory order of the upper encoder. Similarly, the feed-forward generator polynomial for the lower encoder and transmit antenna i, where i=1,2 can be written as

\[ G_i^2(D) = g^2_{0,i} + g^2_{1,i} D + \ldots + g^2_{v_2,i} D^{v_2} \]  \hspace{1cm} (5.31)

where \( g^2_{j,i}, j=0,1,\ldots,v_2 \) are nonbinary coefficients that can take values from a constellation such as 4PSK as 1, -j, -1, j and \( v_2 \) is the memory order of the upper encoder.
The encoded symbol sequence transmitted from antenna \( i \) is given by

\[
s'(D) = u'(D)G'_i(D) + u''(D)G''_i(D) \mod 4 \tag{5.32}
\]

We can also express this as

\[
s'(D) = \begin{bmatrix} u'(D) & u''(D) \end{bmatrix} \begin{bmatrix} G'_i(D) \\ G''_i(D) \end{bmatrix} \mod 4 \tag{5.33}
\]

Assume that \( r_{tj} \) is the received signal at antenna \( j \) at time \( t \), the branch metric is given by

\[
\sum_{j=1}^{M_j} \left| r_j - \sum_{i=1}^{M_i} h_{i,j} q_i \right|
\]

The Viterbi algorithm is used to compute the path with the lowest branch metric. In the absence of ideal CSI, we estimate the CSI based on training symbols.
5.4.4 Code Construction for Quasi-Static Flat Fading Channels

- If a space-time trellis code guarantees a diversity advantage of $r$ for the quasi-static flat fading channel model (given one receive antenna), then it is called an $r$-STTC.
- The constraint length of an $r$-STTC is at least $r-1$.
- If the diversity advantage is $M_T M_R$, then the transmission rate is at most $b$ bit/s/Hz with a $2^b$ signal constellation. Thus 4PSK, 8PSK, and 16QAM will be upper-bounded by 2, 3, 4 bit/s/Hz, respectively.
- If $b$ is the transmission rate, the trellis complexity is at least $2^{b(r-1)}$.
- An STTC is shown to be geometrically uniform and its performance is independent of the transmitted code word.
5.4.5 Example Using 4PSK

- STTCs are an extension of conventional trellis codes to multiantenna systems. These codes are handcrafted to extract diversity gain and coding gain using the criteria described in Section 5.4.4. Each STTC can be described using a trellis. The number of nodes in a trellis diagram corresponding to the number of states in the trellis.
Figure 5.6 Example of 2 transmit space-time trellis code with r states (4PSK constellation with spectral efficiency of 2 bit/s/Hz)

<table>
<thead>
<tr>
<th>Input bits</th>
<th>00</th>
<th>01</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 0</td>
<td>00</td>
<td>01</td>
<td>02</td>
<td>03</td>
</tr>
<tr>
<td>Output for</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>antenna 1,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>antenna 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State 1</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>Output for</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>antenna 1,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>antenna 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State 2</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
</tr>
<tr>
<td>Output for</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>antenna 1,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>antenna 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

State #
0
1
2
3
The trellis has four nodes corresponding to four states. There are four groups of symbols at the left of every node since there are four possible inputs (4PSK constellation). Each group has two entries corresponding to symbols to be output through the two transmit antennas. At the top of the diagram we have the binary input bits that drive these symbols, which are output from the transmit antennas. These symbols come in pairs (for a two-antenna transmitter), wherein the first digit corresponds to the symbol transmitted from antenna 1 and the second from antenna 2. The encoder is required to be in the zero state at the beginning and at the end of each frame.
Mathematically, if \((b_t, a_t)\) are the input binary sequence, the output signal pair \(s_1 t^{t} s_2 t^{t}\) at time \(t\) is given by
\[
(s'_1, s'_2) = b_{t-1} (2,0) + a_{t-1} (1,0) + b_t (0,2) + a_t (0,1) \mod 4
\]
\[
= ((2b_{t-1} + a_{t-1}) \mod 4, (2b_t + a_{t-1}) \mod 4)
\] (5.34)

For the diversity advantage to be 2 (to qualify as a 2-space-time code), the rank of \(B(s,e)\) has to be 2. This can be seen from (5.34), since if the paths corresponding to code words \(s\) and \(e\) diverge at time \(t_1\) and remerge at time \(t_2\), then the vectors \((e_1^{t_1} s_1^{t_1}, e_2^{t_1} s_2^{t_1})\) and \((e_1^{t_2} s_1^{t_2}, e_2^{t_2} s_2^{t_2})\) are linearly dependent on each other and with \(e_1^{t_1} s_1^{t_1} = e_2^{t_1} s_2^{t_1} = 0\) and \(e_1^{t_2} s_1^{t_2} \neq 0\) and \(e_2^{t_2} s_2^{t_2} \neq 0\).

To compute the coding advantage, we need to find code words \(s\) and \(e\) such that the determinant
\[
\det \left( \sum_{t=1}^{T} \left( e_i^{t} - s_i^{t} , e_i^{2} - s_i^{2} \right)^{\top} \left( e_i^{t} - s_i^{t} , e_i^{2} - s_i^{2} \right) \right)
\] (5.35)
is minimized.

- Using the theorem that the code is geometrically uniform, we can assume that we start with the all-zeros code word.
We can express the edge labels \((s_1 s_2)\) by the complex equivalent \((j^{s_1}, j^{s_2})\) where \(j^2 = -1\). By substituting \((j^{s_1}, j^{s_2})\) into (5.35), we obtain

\[
\det \left( \sum_{n=1}^{L} (j^{s_1} - 1, j^{s_2} - 1)^* (j^{s_1} - 1, j^{s_2} - 1) \right)
\]

(5.36)

Since a zero code word maps to \(j^0 \leftrightarrow 1\). Then

\[
\det \left( \sum_{n=1}^{L} \left( j^{-s_1} - 1 \right) \left( j^{s_1} - 1, j^{s_2} - 1 \right) \right)
\]

(5.37)

And, therefore, the inner product takes the form

\[
\begin{pmatrix}
(j^{-s_1} - 1)(j^{s_1} - 1) & (j^{-s_1} - 1)(j^{s_2} - 1) \\
(j^{-s_2} - 1)(j^{s_1} - 1) & (j^{-s_2} - 1)(j^{s_2} - 1)
\end{pmatrix}
\]

(5.38)

If we transpose this matrix (it does not affect the determinant value), we obtain

\[
\begin{pmatrix}
(j^{-s_1} - 1)(j^{s_1} - 1) & (j^{-s_2} - 1)(j^{s_1} - 1) \\
(j^{-s_1} - 1)(j^{s_2} - 1) & (j^{-s_2} - 1)(j^{s_2} - 1)
\end{pmatrix}
\]

(5.39)
Figure 5.7 State diagram for 4PSK example
The reader is advised to refer to Figure 5.6. In that example, we dealt with the case of a transition from state 0 to state 2, given that the input bits were 10. In such a case, $s_1=0$ and $s_2=2$. Substituting these values into (5.39), we obtain $[0 \ 0; \ 0 \ 4]$. This is clearly seen in the state diagram in Figure 5.7 when we follow the arrow from state 0 (00) to state 2 (10) at the diagonally opposite end of the figure.
Diverging from the zero state contributes a matrix of the form
\[
\begin{bmatrix}
0 & 0 \\
0 & t
\end{bmatrix}
\]
and remerging to the zero state contributes a matrix of the form
\[
\begin{bmatrix}
s & 0 \\
0 & 0
\end{bmatrix}
\]
with \( s, t \geq 2 \).

(5.35) can be written as
\[
\text{det} \left( \begin{pmatrix} s & 0 \\ 0 & t \end{pmatrix} + \begin{pmatrix} a & b \\ \bar{b} & d \end{pmatrix} \right)
\]
with \( a, d \geq 0, |b|^2 \leq ad \). So the minimum determinant is 4.
Given that the diversity is \( rM_R \), we wish to maximize the minimum determinant. We have achieved this goal by ensuring that this minimum determinant is 4. The rank of the matrix \( B \) is full rank (i.e., \( r=2 \)). If we require full diversity of \( M_T M_R \), then it is important that this rank criterion is satisfied. In such a case, for this example, the minimum determinant will be 4. If it were less, then the code is useless. Recall that the value of the minimum determinant defines the coding gain. The higher this minimum determinant, the more the coding gain. We should, therefore, strive to make this value as high as possible. It does not, however, provide an accurate estimate of the true coding gain (i.e., there is no direct relationship for us to actually be able to predict the realizable coding gain).

The design rules that guarantee diversity for the 4PSK and 8PSK code are:

- Transitions departing from the same state differ in the second symbol.
- Transitions arriving at the same state differ in the first symbol.
Using (5.33), (5.34) can also be expressed as

\[ s' s^2 = \begin{bmatrix} b & a \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} = \begin{bmatrix} b & a \end{bmatrix} \begin{bmatrix} (02), (20) \\ (01), (10) \end{bmatrix} \]

where \( g_1 \) and \( g_2 \) are generator polynomials.

The second column of this matrix is the \( t-1 \) state. If we expand the matrix, we obtain (5.34) as

\[ (s_1', s_2') = b_{i-1}(2,0) + a_{i-1}(1,0) + b_i(0,2) + a_i(0,1) \mod 4 \]

This method of representing codes as generator polynomials is useful and compact and is used extensively.

We show some useful generator polynomials using the rank and determinant criteria, as well as trace criterion as appropriate.
Table 5.3 Generator sequences for varying number of transmit antennas based on rank and determination criteria

<table>
<thead>
<tr>
<th>Modulation</th>
<th>v</th>
<th>Number of Transmit Antennas</th>
<th>Generator Sequences</th>
<th>Rank(r)</th>
<th>Det</th>
<th>tr</th>
</tr>
</thead>
</table>
| QPSK       | 2 | 2                           | g\(_1\)[(0, 2), (2, 0)]  
               |    |                             | g\(_2\)[(0, 1), (1, 0)] | 2  | 4.0 |    |
| QPSK       | 4 | 2                           | g\(_1\)[(0, 2), (2, 0), (0, 2)]  
               |    |                             | g\(_2\)[(0, 1), (1, 2), (2, 0)] | 2  | 12.0|    |
| QPSK       | 4 | 3                           | g\(_1\)[(0, 0, 2), (0, 1, 2), (2, 3, 1)]  
               |    |                             | g\(_2\)[(2, 0, 0), (1, 2, 0), (2, 3, 3)] | 3  | 32  | 16 |
| 8PSK       | 3 | 2                           | g\(_1\)[(0, 4), (4, 0)]  
               |    |                             | g\(_2\)[(0, 2), (2, 0)]  
               |    |                             | g\(_3\)[(0, 1), (5, 0)] | 2  | 2   | 4  |
| 8PSK       | 4 | 2                           | g\(_1\)[(0, 4), (4, 4)]  
               |    |                             | g\(_2\)[(0, 2), (2, 2)]  
               |    |                             | g\(_3\)[(0, 1), (5, 1), (1, 5)] | 2  | 3.515 | 6  |
Table 5.4 Generator sequences for varying number of transmit antennas based on trace criterion

<table>
<thead>
<tr>
<th>Modulation</th>
<th>v</th>
<th>Number of Transmit Antennas</th>
<th>Generator Sequences</th>
<th>Rank(r)</th>
<th>Det</th>
<th>tr</th>
</tr>
</thead>
<tbody>
<tr>
<td>QPSK</td>
<td>2</td>
<td>2</td>
<td>$g_1[(0, 2), (1, 2)]$</td>
<td>2</td>
<td>4.0</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$g_2[(2, 3), (2, 0)]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QPSK</td>
<td>4</td>
<td>2</td>
<td>$g_1[(1, 2), (1, 3), (3, 2)]$</td>
<td>2</td>
<td>8.0</td>
<td>16.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$g_2[(2, 0), (2, 2), (2, 0)]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QPSK</td>
<td>2</td>
<td>4</td>
<td>$g_1[(0, 2, 2, 0), (1, 2, 3, 2)]$</td>
<td>2</td>
<td></td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$g_2[(2, 3, 3, 2), (2, 0, 2, 1)]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8PSK</td>
<td>4</td>
<td>2</td>
<td>$g_1[(2, 4), (3, 7)]$</td>
<td>2</td>
<td>0.686</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$g_2[(4, 0), (6, 6)]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$g_3[(7, 2), (0, 7), (4, 4)]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8PSK</td>
<td>4</td>
<td>4</td>
<td>$g_1[(2, 4, 2, 2), (3, 7, 2, 4)]$</td>
<td>2</td>
<td></td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$g_2[(4, 0, 4, 4), (6, 6, 4, 0)]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$g_3[(7, 2, 2, 0), (0, 7, 6, 3), (4, 4, 0, 2)]$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.8 Example of 2 transmit space-time trellis code with 8 states (4PSK constellation with spectral efficiency of 2 bit/s/Hz)
5.5 Design of Space-Time Trellis Codes on Fast Fading Channels

- At each time $t$, we define a space-time symbol difference vector, $F(s,e)$ as
  \[
  F(s,e) = [s_t^1 - e_t^1, s_t^2 - e_t^2, \ldots, s_t^{M_t} - e_t^{M_t}]^T
  \]

- Consider a $M_T \times M_T$ matrix $S=S(s,e)$, defined as $S=F(s,e)F^+(s,e)$. It is clear that $S$ is Hermitian, so there exists a unitary matrix $V_t$ and a real diagonal matrix $D_t$ such that
  \[
  V_t S V_t^+ = D_t
  \]

- The diagonal entries of $D_t$, $\{D_t^i, i=1,2,\ldots,M_T\}$ are the eigenvalues of $S$; the rows of $V_t$, $\{v_t^i, i=1,2,\ldots,M_T\}$ are the eigenvalues of $S$, which form a complete orthonormal basis of an $M_T$-dimensional vector space.
Note that $S$ is a rank 1 matrix (since we are dealing on a symbol basis) if $s \neq e$ and is rank 0 otherwise. It follows that $M_T - 1$ elements in the list $\{D_t^i, i=1,2,\ldots,M_T\}$ are zeros and, consequently, we can let the single nonzero element in this list be $D_t^1$, which is equal to the squared Euclidian distance between the two space-time symbols $s_t$ and $e_t$.

$$D_t^i = |s_t - e_t|^2 = \sum_{i=1}^{M_T} |s_t^i - e_t^i|^2$$

The eigenvector of $S(s_t,e_t)$ corresponding to the nonzero eigenvalue $D_t^1$ is denoted by $v_t^i$.
We define $h_t^j$ as
\[ h_t^j = (h_{t,1}^j, h_{t,2}^j, \ldots, h_{t,M_t}^j) \quad (5.42) \]

Now,
\[ d^2(s,e) = \|H \cdot (E - S)\|^2 = \sum_{i=1}^{L} \sum_{j=1}^{M_t} h_{t,j}^i \left( e_t^i - s_t^i \right) \]

This can be rewritten as
\[ d^2(s,e) = \sum_{i=1}^{L} \sum_{j=1}^{M_t} |\beta_{t,i,j}|^2 \cdot D_t \quad (5.43) \]

where $\beta_{t,i,j} = h_{t,j}^i \cdot v_t^i$.

Since at each time $t$ there is, at most, only one nonzero eigenvalue, $D_t$, the expression (5.43) can be represented by
\[ d^2(s,e) = \sum_{t \in \rho(s,e)} \sum_{j=1}^{M_t} |\beta_{t,j}^i|^2 \cdot D_t = \sum_{t \in \rho(s,e)} \sum_{j=1}^{M_t} |\beta_{t,j}^i|^2 \cdot |s_t^i - e_t^i|^2 \quad (5.44) \]

Where $\rho(s,e)$ denotes the set of time instances $t=1,2,\ldots,L$ such that $|s_t^i - e_t^i| \neq 0$. 
Substituting (5.44) into (5.19), we obtain
\[ P(s \rightarrow E \mid H) \leq \frac{1}{2} \exp \left( - \sum_{i \in \mathcal{P}(s,e)} \sum_{j=1}^{M_t} |\beta_{i,j}^t|^2 \cdot |s_i - e_i|^2 \frac{E_s}{4N_0} \right) \]  
(5.45)

Since \( h_{i,j} \) are samples of a complex Gaussian random variable with mean \( \mathbb{E} h_{i,j} \), let
\[ K_i^j = (E_{h_{1,j}}, E_{h_{2,j}}, \ldots, E_{h_{M_t,j}}) \]

Since \( V \) is unitary, this implies \( \beta_{i,j} \) are independent complex Gaussian random variables with variance 0.5 per dimension and with mean \( K_i \cdot v_i \).

If we define \( \delta_H \) as the number of space-time symbols in which the two code words \( s \) and \( e \) differ, then at the right-hand side of inequality (5.45), there are \( \delta_H M_R \) independent random variables.

Once again, like in the slow fading cases, we discuss two situations, such as when \( \delta_H M_R < 4 \) and \( \delta_H M_R \geq 4 \).
Case when $\delta_H M_R \geq 4$

According to the central limit theorem, the expression $d^2(s,e)$ in (5.43) can be approximated by a Gaussian random variable with the mean

$$\mu_d = \sum_{t \in \rho(s,e)} \sum_{j=1}^{M_R} |s_t - e_t|^2$$

(5.46)

and the variance

$$\sigma_d^2 = \sum_{t \in \rho(s,e)} \sum_{j=1}^{M_R} |s_t - e_t|^4$$

(5.47)
By averaging (5.45) over the Gaussian random variable and using

\[ \int_{D=0}^{\infty} \exp(-\gamma D)p(D)dD = \exp\left( \frac{1}{2} \gamma^2 \sigma_D^2 - \gamma \mu_D \right) Q\left( \frac{\gamma \sigma_D^2 - \mu_D}{\sigma_D} \right) \quad \gamma > 0 \]

we obtain the PEP as

\[ P(s \rightarrow e) \leq \frac{1}{2} \exp\left( \frac{1}{2} \left( \frac{E_s}{4N_0} \right)^2 \sigma_D^2 - \frac{E_s}{4N_0} \mu_d \right) Q\left( \frac{E_s}{4N_0} \sqrt{M_r D^4} - \frac{\sqrt{M_r d_E^2}}{\sqrt{D^4}} \right) \] (5.48)

where \( d_E^2 \) is the accumulated squared Euclidian distance between the two space-time symbol sequences, given by

\[ d_E^2 = \sum_{t \in \rho(s,e)} |s_t - e_t|^2 \] (5.49)

and \( D^4 \) defined as

\[ D^4 = \sum_{t \in \rho(s,e)} |s_t - e_t|^4 \] (5.50)
Case when $\delta_H M_R < 4$

- The central limit theorem argument is no longer valid and the average error probability can be expressed as

$$P(s \to e) \leq \int_{|\beta_j^1| = 0} \cdots \int_{|\beta_j^L| = 0} P(s \to E | H) p\left( |\beta_{1,1}^1|^2 \right) p\left( |\beta_{1,2}^1|^2 \right) \cdots p\left( |\beta_{M_R,1}^L|^2 \right) d|\beta_{1,1}^1| d|\beta_{2,1}^1| \cdots d|\beta_{M_R,1}^L|$$  \hspace{1cm} (5.51)

where $|\beta_{j,t}^1|$, $t = 1, 2, \ldots, L$ and $j = 1, 2, \ldots, M_R$ are independent Rayleigh distributed random variables. By integrating (5.51) term by term, the PEP becomes

$$P(s \to e) \leq \frac{1}{\prod_{t \in S, e} \left( 1 + |s_t - e_t|^2 \frac{E_s}{4N_0} \right)^{M_R}} \leq \left( d_p^2 \right)^{-M_R} \left( \frac{E_s}{4N_0} \right)^{-\delta_H M_R}$$  \hspace{1cm} (5.52)

where $d_p^2$ is the product of the squared Euclidian distances between the two space-time symbol sequences, given by

$$d_p^2 = \prod_{t \in S, e} |s_t - e_t|^2$$  \hspace{1cm} (5.53)
At high SNRs, the frame error probability is dominated by the PEP with the minimum product $\delta H M_R$. The exponent of the SNR term, $\delta H M_R$, is called the diversity gain for fast Rayleigh fading channels and

$$G_s = \frac{(d_p^2)^{1/\delta_H}}{d_u^2}$$

(5.54)

Is called the coding gain for fast Rayleigh fading channels, where $d_u^2$ is the squared Euclidean distance of the reference uncoded system.

Note that both diversity and coding gains are obtained as the minimum $\delta H M_R$ and $(d_p^2)^{1/\delta_H}$ over all pairs of distinct code words because this becomes the worst case.
Design criteria for fast Rayleigh fading channels

- Case when $\delta_H M_R < 4$
  
  - Maximize the minimum space-time symbol-wise Hamming distance $\delta_H$ between all pairs of distinct code words.
  
  - Maximize the minimum product distance $d_p^2$, along the path with the minimum symbol-wise Hamming distance $\delta_H$. 
Case when $\delta_{HM_R} \geq 4$

- The pairwise error probability is upper-bounded by (5.48). In the case of high SNRs,

$$\frac{E_s}{4N_0} \geq \frac{d_E^2}{D^4}$$

Where $d_E^2$ and $D^4$ are given by (5.49) and (5.50), respectively. (5.48) can be approximated by

$$P(s \rightarrow e) \leq \exp\left(-M_R \frac{E_s}{4N_0} \sum_{i=1}^{L} \sum_{i=1}^{M} |s_i^j - e_i^j|^2 \right) = \exp\left(-M_R \frac{E_s}{4N_0} d_E^2 \right)$$

(5.55)

- The frame error rate probability at high SNRs is dominated by the PEP with the minimum squared Euclidian distance $d_E^2$. To minimize the PEP on fading channels, the codes should satisfy the following criteria:
  - Make sure that the product of the minimum space-time symbol-wise Hamming distance and the number of receive antennas, $\delta_{HM_R}$, is large enough (larger than or equal to 4).
  - Maximize the minimum Euclidian distance among all pairs of distinct code words.
5.6 Performance Analysis in a Slow Fading Channel

- The performance of the STTC on slow fading channels is evaluated through simulations.
Figure 5.9 Performance comparison of 4PSK codes based on the rank and determinant criteria on slow fading channels with two transmit and one receive antennas.
Simulation results:
- All the codes achieve the same diversity order of 2, demonstrated by the same slope of the FER performance.
Figure 5.10 Performance comparison of 4PSK codes based on trace criterion on slow fading channels with two transmit and two receive antennas and four transmit and two receive antennas.
Simulation results:

- The code performance is improved by increasing the number of states. Figure 5.10 shows that increasing the number of transmit antennas also increases the margin of the coding gain compared with the coding gains in the top-half graph.

- This is evident from (5.18), wherein the value of $rM_R$ (ideally $M_T M_R$), defines the amount of coding gain. Similarly, if we keep $M_T$ constant and increase $M_R$, we achieve the same result with something more, in that part of this margin is also due to the array gain through multiple receive antennas. Furthermore, the diversity order realized with this scheme in the lower half is twice that in the upper half. Proceeding logically, as the number of receiver antennas increases, the diversity order increases proportionately and the channel tends to AWGN due to the increased diversity. A similar effect can be observed by keeping the number of receive antennas constant and increasing the number of transmit antennas. Finally, in the presence of a large number of receive antennas, increasing the number of transmit antennas does not produce that much of an increase in performance, as seen in the case when the number of receive antennas is limited and we increase the number of transmit antennas.
5.7 Performance Analysis in a Fast Fading Channel

- The performance of the STTC on fast fading channels is evaluated through simulations.
  - Figure 5.11 show that the FER performance of QPSK STTC with a bandwidth efficiency of 2 bit/s/Hz in Rayleigh channel.
Figure 5.11 Performance of the QPSK STTC on fast fading channels with two transmit and one receive antennas
Simulation results:

- We can see that 16-state QPSK codes are better than 4-state codes by 5.9 dB at a FER of $10^{-2}$ for two transmit antennas. Once again, as the number of states increases, the coding gain increases and so does the performance. The error rate curves of the codes are parallel, as predicted by the same value of $\delta_H$. Different values of $d_p^2$ yield different coding gains, which are represented by the horizontal shifts of the FER curves.
Figure 5.12 Performance of QPSK STTC on fast fading channels with three transmit and one receive antennas
● Simulation results:
  – 16-state QPSK codes are superior to 4-state codes by 6.8 dB at a FER of $10^{-2}$. This means that the performance relative to two transmit antennas has improved.
  – The conclusion here is that as the number of the transmit antennas gets larger, the performance gain achieved from increasing the number of states becomes larger.
5.8 The Effect of Imperfect Channel Estimation on Code

- We carry out imperfect estimation using the MMSE technique discussed in Chapter 4.
Figure 5.13 Performance of the 4-state 4PSK code on slow Rayleigh fading channels with two transmit and two receive antennas and imperfect channel estimation
● Simulation models:
  - 10 orthogonal signals in each data frame are used as pilot sequence to estimate the channel state information at the receiver.

● Simulation results:
  - From the figure, we can see that the deterioration due to imperfect channel estimation is about 5 dB throughput.
5.9 Effect of Antenna Correlation on Performance

- The effect of antenna correlation on performance is evaluated through simulations.
Figure 5.14 Performance of the 4PSK 4-state code on correlated slow Rayleigh fading channels with two transmit and two receive antennas
● Simulation models:
  – This example has been implemented using the code based on trace criterion.

● Simulation results:
  – The performance gap is 0.5 dB throughput for both a correlation factor of 0.75 as well as unity.
5.10 Delay Diversity as an STTC

- The delay diversity scheme discussed in Chapter 4 can be recast as an STTC. Assume a system with two transmit antennas and one receive antenna. In the delay diversity scheme, it will be recalled, we transmit one symbol from one antenna and then transmit the same symbol from the second antenna, but after a short delay of one symbol.
Figure 5.15 Trellis diagram for delay diversity code with 8PSK transmission and $M_T=2$
If we assume the input sequence as
\[ s = (010, 101, 111, 000, 001, \ldots) \]
The output sequence generated by the space-time trellis encoder is given by
\[ s = (02, 25, 57, 70, 01, \ldots) \]
Using the technique discussed earlier.

The transmitted signal sequences from the two transmit antennas are
\[ \mathbf{s}^1 = (0, 2, 5, 7, 0, \ldots) \]
\[ \mathbf{s}^2 = (2, 5, 7, 0, 1, \ldots) \]
Very clearly, this is delay diversity, since the signal sequence transmitted from the first antenna is a delayed version of the signal sequence from the second antenna.

If we express \( \mathbf{s}^1 \) and \( \mathbf{s}^2 \) as a matrix,
\[
\mathbf{S} = \begin{bmatrix}
0 & 2 & 5 & 7 & 0 & \ldots \\
2 & 5 & 7 & 0 & 1 & \ldots
\end{bmatrix}
\]
It is easily verified that the rank of this matrix is 2. Hence, applying the rank criterion for the space-time code word design discussed earlier, the delay diversity transmission extracts the full diversity order of \( 2M_R \).
5.11 Comparison of STBC and STTC

- Space-time block codes and space-time trellis codes are two very different transmit diversity schemes.
  - Space-time block codes are constructed from known orthogonal designs, achieve full diversity, and are easily decodable by maximum likelihood decoding via linear processing at the receiver, but they suffer from a lack of coding gain.
  - Space-time trellis coders possess both diversity and coding gain, yet are complex to decode (since we use joint maximum likelihood sequence estimation) and arduous to design.
  - In both cases, the code design for a large number of transmit antennas remains an open question.
It is only fair to use concatenated STBC since STBC inherently lacks coding gain. Concatenated codes that have been used so far include AWGN Trellis codes or turbo codes.

- Sumeet et al. attempted a fair comparison of the performance of STBC with STTC over the flat fading quasi-static channel, presenting results in terms of the FER while keeping the transmit power and spectral efficiency constant. We shall now discuss her results.
In general, any STC can be analyzed in the same way as STTC using diversity advantage and coding advantage. Both of these advantages affect the performance curve differently.

- Diversity advantage causes the slope of the FER versus SNR graph to change in such a way that the larger the diversity, the more negative the slope.
- Coding advantage shifts the graph horizontally: the greater the coding advantage, the larger is the left shift.
- Full diversity codes were used and, hence, the slopes of their FER graphs were identical.
To further examine the coding advantage aspect, consider a high SNR regime (typically 4 dB to 18 dB). We first take the logarithm of the PEP expression in (5.18) for the kth code. This yields

$$P_k = \log(\text{PEP}) \approx -M_T M_R s_k - M_T M_R c_k$$

where $M_T M_R$ is the full diversity advantage, $s_k = \log(E_s/4N_0)$ is the SNR term and

$$c_k = \log\left(\prod_{i=1}^{M_r} \lambda_i^{U_{M_r}}\right)$$

is the coding advantage term. If we let $\delta_p = P_k - P_L$, $\delta_c = c_k - c_L$ and $\delta_s = s_k - s_L$ for the kth and Lth code, then

$$\delta_p = -M_T M_R \delta_s - M_T M_R \delta_c$$

If k is a better code, then $\delta_c > 0$. At a given SNR, $\delta_s = 0$ and the PEP for k is less than that for L by $\delta_p \approx M_T M_R \delta_c$. Clearly, this difference increases with $M_R$, the number of receive antennas. Thus, the effect of coding advantage improves when more receive antennas were used.
Simulation models:

- In the simulations conducted by Sumeet, the performance comparison between STBC and STTC was carried out for a system with two transmit and one, two, and three receive antennas with 4PSK modulation using STTC-Grimm and STTC-Yan. The STTC codes were used instead of STTC-Tarokh because they have the best possible coding advantage in the class of feed-forward convolutional (FEC) codes.
- The block code used was Alamouti, which is a full-rate code. This was concatenated with outer AWGN trellis codes.
- The spectral efficiency was maintained at 2 bit/s/Hz throughout the simulation. The FER is given for 130 symbols/frame. The channel is a quasi-static Rayleigh fading channel.
Figure 5.16 Performance of STBC, STBC+TCM, and STTC using 4 and 8 state codes with two transmit and one receive antenna
Figure 5.17 Performance of STBC, STBC+TCM, and STTC using 4 and 8 state codes with two transmit and two receive antenna.
Figure 5.18 Performance of STBC, STBC+TCM, and STTC using 4 and 8 state codes with two transmit and three receive antenna.
Simulation results:
- In Fig. 5.16 (4-state code), STBC by itself performs as well or better than all the STTCs, even though it provides no coding gain. This can be explained by the multidimensional structure of STBCs, since each code word spans two time symbols and averages out the noise over time. Other interesting observations are:
  - With the same number of trellis states, concatenated STBCs outperform STTCs at SNRs of interest (i.e., 4 to 12 dB for the case of two receivers and 10 to 18 dB for the case of one receiver).
  - With increasing number of antennas and trellis states, STTC begins to outperform concatenated STBC.
If the number of receive antennas are one or, at most, two, a simple concatenation of STBC with traditional AWGN trellis codes can significantly outperform STTC with the same number of state. This has very important implications for design and implementation of MIMO systems.

STBC + TCM curves lose in performance with increasing receive antennas, because STBCs incur a loss in capacity over channels with rank greater than one. Since TCM codes are outer codes, they are unable to recover performance after the signal has been encoded and decoded using space-time block codes.
**Table 5.5 Comparison of STBC versus STTC**

<table>
<thead>
<tr>
<th>STBC</th>
<th>STTC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No coding gain.</td>
<td>Has coding gain.</td>
</tr>
<tr>
<td>Easily decodable by maximum likelihood decoding via linear processing</td>
<td>Complex to decode (since we use joint maximum likelihood sequence estimation)</td>
</tr>
<tr>
<td>Simple to design based on orthogonal sequences.</td>
<td>Difficult to design.</td>
</tr>
<tr>
<td>For one receive antenna and 4-state code, performance is similar to STTC.</td>
<td>STTC outperforms with increasing antennas and trellis states.</td>
</tr>
<tr>
<td>Easily lends itself to industrial applications because of its simplicity.</td>
<td>Complex to deploy.</td>
</tr>
<tr>
<td>Lose capacity with two or more receive antennas.</td>
<td>Preserves capacity irrespective of the number of antennas.</td>
</tr>
</tbody>
</table>
5.12 Simulation Exercises

- If the number of transmit antennas is fixed and we increase the number of receive antennas, the margin of coding gain increases. Prove this statement using the generator codes given in Section 5.4.5. Remember the rule, that the higher the trace value with multiple receive antennas, the better the performance, regardless of the rank and determinant of other codes.

- Figure 5.13 shows the performance of a 4-state 4PSK code on slow Rayleigh fading channels with two transmit and two receive antennas and imperfect channel estimation using MMSE. Repeat this exercise using a 16-state 4PSK code.

- Figure 5.14 shows the performance of 4PSK codes in the presence of correlation between receive antennas. Check out the performance using higher trace codes from the tables given in Section 5.4.5. How do STTC codes compare with STBC codes from the point of view of robustness in the presence of correlation?
References


