Impact of Channel Estimation Errors on Space Time Trellis Codes

by

Rekha Menon

Thesis submitted to the faculty of
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree for

MASTER OF SCIENCE
in
Electrical Engineering

Dr. R. Michael Buehrer, Chair
Dr. Charles W. Bostian
Dr. Jeffrey H. Reed

December 15, 2003
Blacksburg, Virginia

Keywords: Transmit Diversity, Rayleigh Fading Channel, Channel Estimation, Flat Fading, Frequency Selective Fading

Impact of Channel Estimation Errors on Space Time Trellis Codes

Rekha Menon

Abstract

Space Time Trellis Coding (STTC) is a unique technique that combines the use of multiple transmit antennas with channel coding. This scheme provides capacity benefits in fading channels, and helps in improving the data rate and reliability of wireless communication. STTC schemes have been primarily designed assuming perfect channel estimates to be available at the receiver. However, in practical wireless systems, this is never the case. The noisy wireless channel precludes an exact characterization of channel coefficients. Even near-perfect channel estimates can necessitate huge overhead in terms of processing or spectral efficiency. This practical concern motivates the study of the impact of channel estimation errors on the design and performance of STTC.

The design criteria for STTC are validated in the absence of perfect channel estimates at the receiver. Analytical results are presented that model the performance of STTC systems in the presence of channel estimation errors. Training based channel estimation schemes are the most popular choice for STTC systems. The amount of training however, increases with the number of transmit antennas used, the number of multi-path components in the channel and a decrease in the channel coherence time. This dependence is shown to decrease the performance gain obtained when increasing the number of transmit antennas in STTC systems, especially in channels with a large Doppler spread (low channel coherence time). In frequency selective channels, the training overhead associated with increasing the number of antennas can be so large that no benefit is shown to be obtained by using STTC.

The amount of performance degradation due to channel estimation errors is shown to be influenced by system parameters such as the specific STTC code employed and the number of transmit and receive antennas in the system in addition to the magnitude of the estimation error. Hence inappropriate choice of system parameters is shown to significantly alter the performance pattern of STTC.

The viability of STTC in practical wireless systems is thus addressed and it is shown that that channel estimation could offset benefits derived from this scheme.

Contents

List of Figures

List of Tables

1. Introduction to STTC

Band-limited wireless channels provide an impediment to high data rates and consequently constrain attempts to improve the reliability of wireless communications by using error correction codes along with high data rate services. One method of increasing the capacity of wireless channels is the use of multiple transmit and receive antennas [Fosc1]. This additional capacity can be thought of as channel "broadening" which can be used to improve system performance in band-limited environments.

Space Time Trellis Coding (STTC) is a unique coding scheme that integrates channel coding and modulation with the use of multiple transmit antennas and optional receiver diversity. This scheme offers increased performance, and helps in improving the data rate and the reliability of communication over fading channels. However, most of the research on the design and performance analysis of STTC assumes the availability of perfect channel estimates at the receiver. In practical wireless systems this is almost impossible to achieve. This concern motivates the study of the impact of channel estimation errors on the design and performance of STTC. This thesis investigates the validity of design criteria for STTC, and attempts to characterize its performance in the presence of channel estimation errors. It also tries to resolve the factors that influence the way channel estimation errors affect STTC performance.

In this chapter, an introduction to Space Time Trellis Codes (STTC) and a description of the basic framework of a system that implements this coding technique is presented.

1.1. Diversity in Wireless Systems

Wireless channels are prone to large fluctuations. The channel may suffer attenuation due to destructive addition of multi-path components of the transmitted signal. This increases the burden on the receiver in determining the transmitted signal. Multiple, possibly less-attenuated copies of the transmitted signal, can improve the receiver decisions significantly. This is termed diversity and is the primary tool against fading. Diversity uses alternate independent or highly uncorrelated paths for communication. Hence if one of the paths undergoes a deep fade, a different path may have a stronger signal and can be exploited by the receiver. Diversity can be provided in different dimensions as described below.

- *Time Diversity*: This is achieved by transmitting copies of the signal in different time slots. This technique is not as effective in slow fading channels. Channel coding is also a form of temporal diversity.

- *Frequency Diversity*: Transmitting the same information over carriers of different frequencies achieves this. This exploits the fact that different frequencies experience different multi-path fading in the propagating media. This technique cannot be used when the delay spread is small. Equalizers exploit this type of diversity.

- *Spreading the Signal*: A signal that has a bandwidth much greater than the coherence bandwidth of the channel is used. Such a signal resolves the multi-path components and provides the receiver with several independent fading signal paths. This is a form of frequency diversity. Equalizers or Rake receivers are needed for this technique.

- *Antenna Diversity*: This can be achieved in different ways. Examples are,

    - *Polarization Diversity*: The independence of orthogonal polarizations is exploited.
    - *Space Diversity*: The antennas are spatially separated such that received signals are uncorrelated. This technique incurs no bandwidth penalty.

1.2. Motivation for Transmit Diversity

Receive diversity can be used at the base station or any at receiver that permits fixed complex structures for reception. But for the downlink, receive diversity is not as practical. Receive diversity is too expensive to be implemented in handsets because of the additional processing power and RF circuitry required. It might also not be very useful because of the electromagnetic interaction of antenna elements on small platforms and the resultant correlation between antenna components. These reasons form the motivation for the study of transmit diversity schemes.

Multiple transmit antennas are used to provide the required diversity advantage over a channel. This helps to shift the diversity burden from the mobile to the base station. It has been shown that the diversity gain provided by the use of transmit diversity scheme is comparable to that provided by Maximum Ratio Receive Combining (MRRC) techniques without the 3dB aperture gain (unless feedback is used). The same is verified in course of this thesis as well.

The transmit diversity schemes found in the literature can be broadly divided into two categories: those with and those without feedback. For schemes with feedback, transmitter sequences are weighted before transmission to ensure maximum benefit. Weights are chosen adaptively based on the information feedback from the receiver. An example of this scheme is switched diversity proposed in [Wint1]. The schemes without feedback use linear processing at the transmitter to spread the information across antennas. At the receiver, information is obtained by linear processing or Maximum Likelihood(ML) techniques. An example of this technique is the delay diversity scheme proposed in [Sesh1]. In this scheme, copies of the same symbol are transmitted through multiple antennas at different times. At the receiver these delays introduce a multi-path like distortion. This distortion can be resolved by using a Minimum Mean Square Error (MMSE) equalizer or a ML scheme to obtain a diversity gain. Space Time Block Codes (STBC) proposed in [Alam1] and generalized in [Taro3] is another example of this technique. These codes are defined by a mapping operation of a block of input symbols into space and time domains creating orthogonal sequences that are transmitted from different transmit antennas. The receiver uses a ML detection rule to decode the transmitted information. STTC is also a transmit diversity scheme that does not require feedback from the receiver. It provides coding gain in addition to the diversity advantage and is essentially an extension of the Trellis Coded Modulation scheme to multiple transmit antennas.

1.3. Trellis Coded Modulation

Channel coding and modulation are traditionally considered as two separate blocks in communication system design. Channel coding adds redundancy to the information bits and modulation maps these bits into appropriate symbols from the chosen signal constellation. The redundancy in the data bits can be accommodated either by increasing the bandwidth of the channel or by expanding the signal set over the un-coded system. Increasing the bandwidth is often not allowable. An expansion of the signal set requires additional power to maintain the same error rate as the Euclidean distance between code-words is otherwise reduced. Codes with large performance gains are required to overcome this power penalty.

Trellis coded modulation is a novel technique that combines coding and modulation into a composite operation. A conventional convolutional code is used and the redundancy for coding is provided by using an expanded signal set. Symbols are chosen from this expanded constellation such that the Euclidean distance (as opposed to Hamming distance in conventional channel code design) between code-words is maximized. This leads to significant performance gains and does not require an increase in the transmit power for the expanded signal set. Thus neither bandwidth nor power efficiency is compromised.

The general strategy for encoding for TCM schemes is as follows, [Wick1],

- Add one bit of redundancy for every m source bits
- Expand the signal constellation from 2^m to 2^{m+1} signals
- Encode the sequence using trellis code
- Use the $m+1$-bit encoded source blocks to select a signal from the expanded signal constellation.

The most important part of the encoding process is the mapping of m information bits to 2^{m+1} symbols from the expanded signal set. This mapping is achieved through set partitioning of the expanded signal constellation. The partitioning is done such that the resulting sub-constellations have larger minimum distances than the original constellation. Figure 1-1 shows the signal partitioning for the 8-PSK constellation. The binary labels at the leaf nodes provide a means for partition selection by coded information bits. The binary labels are arrived at by denoting each rightward branch by one and each leftward branch by a zero.

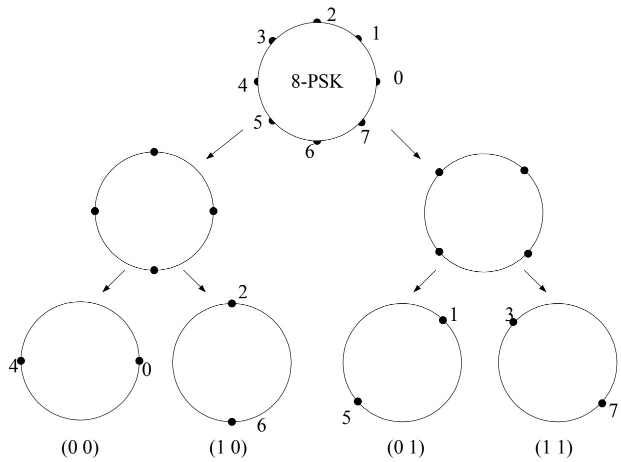

Figure 1-1: Partitioning of the 8-PSK constellation

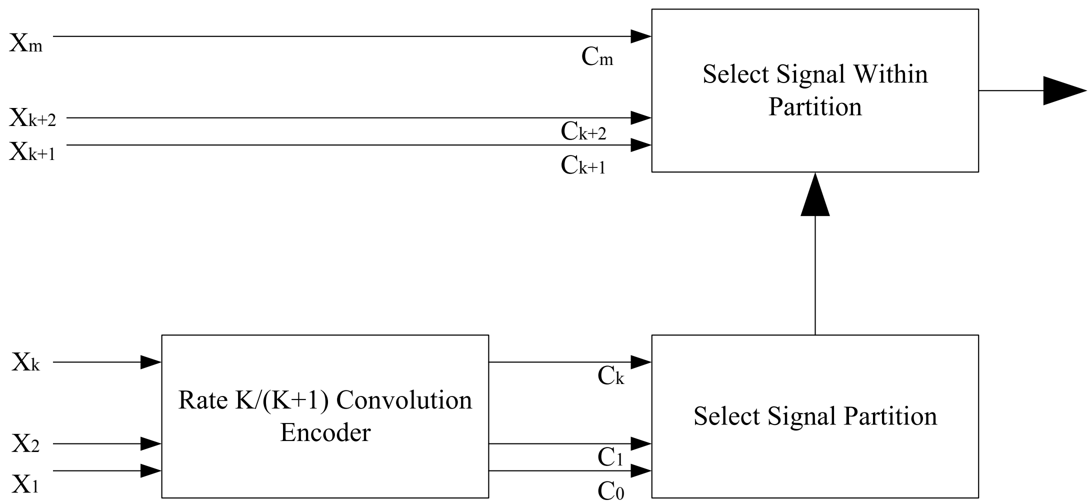

Figure 1-2: A TCM encoder

A TCM encoder is shown in Figure 1-2. k bits are selected from a block of m information bits, $x_1, x_2, \ldots, x_m$ and fed to a rate $\frac{k}{(k+1)}$ encoder. The resulting $k+1$ bits are then used to select a partition from the $(k+1)st$ level of the constellation's partition tree. The remaining $(m-k)$ bits are used to select a signal within the partition. For example, consider a system which encodes information in blocks of two bits using a TCM encoder (thus $m = 2$). Let the symbol constellation used be 8-PSK and let $k = 1$. The first information bit is fed to the convolutional encoder and the resulting $k+1=2$ coded bits are used to select a particular constellation from the last partition level of Figure 1-1. The second information bit is used to choose a particular symbol from the selected partition.

The resulting sequences have a trellis structure due to convolutional coding used in the encoder. Hence the encoder is called a Trellis Coded Modulation system. The trellis restricts the allowable sequences and thus increases the Euclidean distance between code words compared to an un-coded system. Additional design rules that maximize the distance between signals are given by:

- Signals in the same, lowest partition in the partition tree are assigned to parallel transitions
- Signals in the preceding partition are assigned to transitions that start or stop in the same state
- All signal are used equally often

Parallel transitions result in a reduction of the Euclidean distance between code-words and can be eliminated by passing all m information bits through the convolutional encoder. All $m+1$ bits are then used to select a symbol from a constellation. The design rules except for the rule dealing with parallel transitions remain relevant. The elimination of parallel transitions also increases the achievable coding gain.

As a consequence of the trellis structure imparted to TCM codes, the Viterbi decoder can be used for decoding TCM sequences. The Euclidean distance between symbols at a particular transition is used as the decision metric.

Reference [Bigl1] provides a more detailed study of TCM

1.4. Space-Time Trellis Codes

STTC combines the advantages of transmit diversity and TCM in an ingenious way to obtain reliable, high data rate transmission in wireless channels. Channel coding by using convolutional encoders adds redundancy to the information sequence. In TCM, the resulting coded bits are then used to choose symbols from an expanded signal constellation. But in STTC, the encoded data is used to select symbols from N_t separate but identical constellations, where N_t is the number of transmit antenna employed by the MIMO system. These N_t symbols are then simultaneously transmitted from the N_t transmit antennas. The choice of symbols, as in the TCM case, is made such that the Euclidean distance between code-words is maximized. Thus TCM principles are exploited to add channel coding without penalizing the information rate, decreasing the power efficiency or increasing the bandwidth of the system. The scheme concurrently introduces diversity by using multiple transmit antennas in the system. The coding across transmit antennas also aids in providing maximum benefit of the offered system diversity.

A system model for a Multi Input Multi Output (MIMO) system employing a Space Time Trellis Code (STTC) is shown in Figure 1-3. The system has N_t transmit and N_r receive antennas. Consider the signal constellation size of the system to be M . An encoder is used to map k (=$\log_2$M) data bits to N_t separate but identical constellations of size M after applying a convolutional code to it. The STTC encoder is dependant on the number of input data bits (k), the number of states of the convolutional encoder (v) and the number of transmit antennas (N_t). A generator matrix of size $(k+v)\times N_t$ can entirely define the STTC. It can also be represented by a trellis diagram with 2^v states. Each transition of the trellis defines the symbols transmitted from N_t transmit antennas. It also defines the beginning and end states of the encoder. The encoder is implemented by a feed forward shift register structure with memory order v. The Calderbank-Mazo algorithm can be used translate the trellis diagrams into closed form expressions [Taro2], which are very useful in designing the implementation.

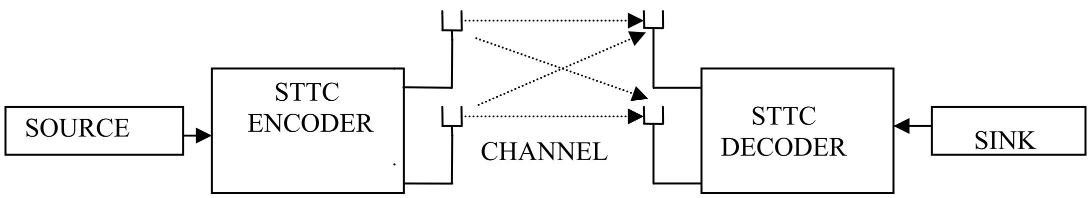

Figure 1-3: System model for a MIMO system employing an STTC

Some examples are used to illustrate the encoder. Consider the 4-state trellis code shown in Figure 1-4 (from [Taro1]). Each trellis transition defines the symbols to be transmitted from each transmit antenna (two in this case) for a particular combination of the state and input bits. The trellis also shows the start and end states after a transition. It can also be noticed that there are M ($M=4$ in this case) transitions from each state corresponding to all possible combinations of k ($k=2$ in this case) input bits. The symbols are chosen from the QPSK constellation shown in Figure 1-5. The encoder can be represented in a closed analytical form by the equations (from [Goza1]),

$$\begin{aligned} x_2(t) &= 2a_2(t)+a_1(t) \\ x_1(t) &= 2a_4(t)+a_3(t) \end{aligned} \qquad \text{1-1}$$

where, $x_1(t)$ and $x_2(t)$ are the symbols transmitted from the first and second transmit antennas. $a_1(t), a_2(t), a_3(t), a_4(t)$ are the input and state bits respectively. The generator matrix of the given code is therefore given by $\begin{bmatrix} 2\,1\,0\,0 \\ 0\,0\,2\,1 \end{bmatrix}^T$ and the given code can be implemented by the feed forward structure shown in Figure 1-6.

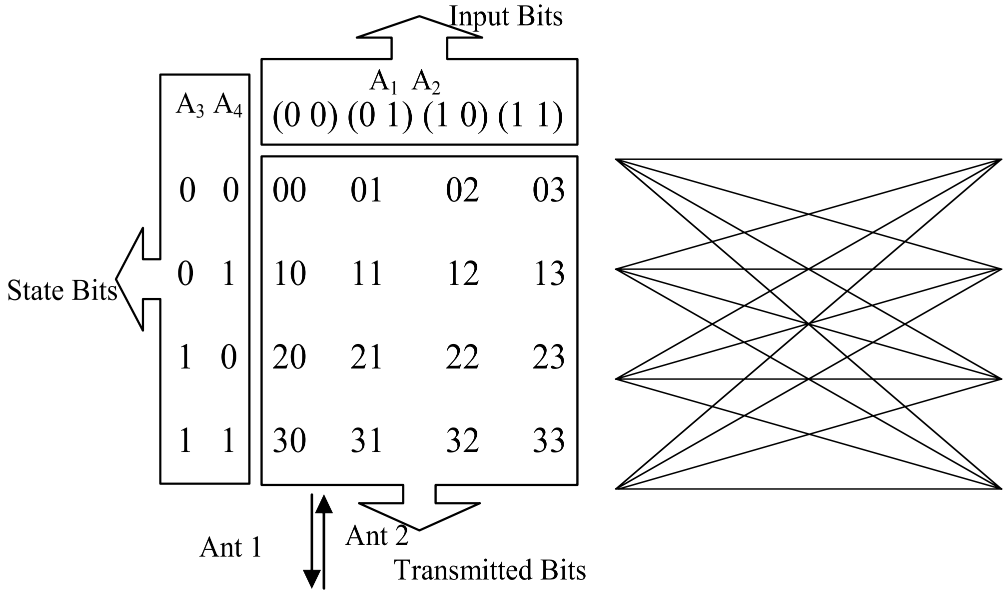

Figure 1-4: Trellis for a 4-state STTC

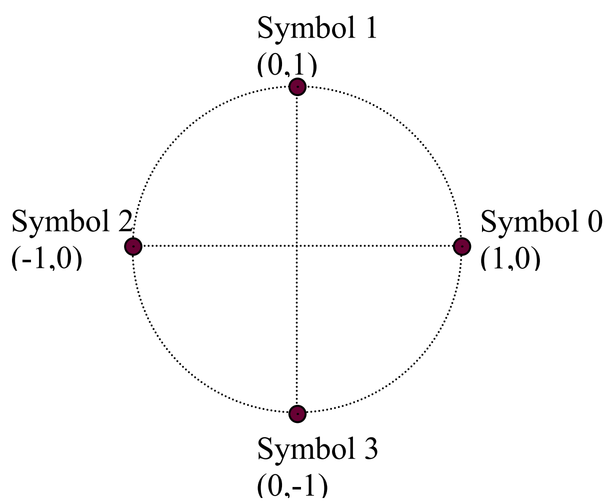

Figure 1-5: QPSK symbol constellation

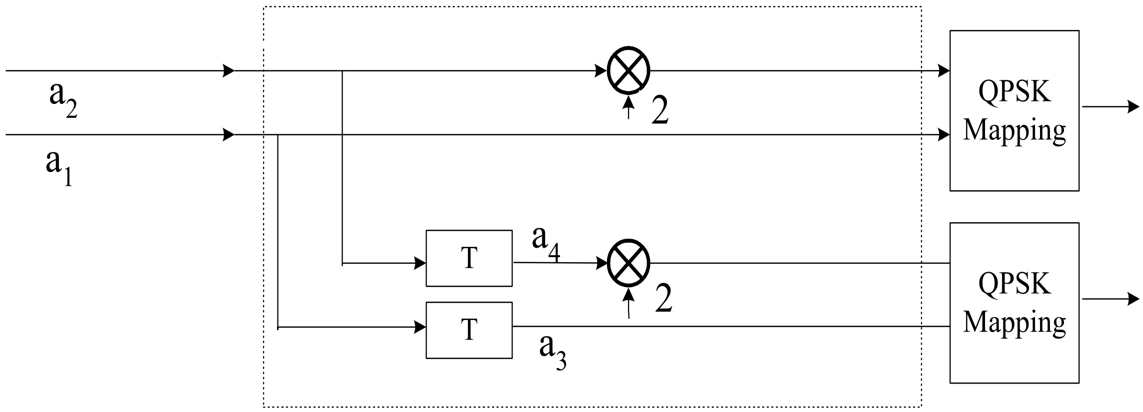

Figure 1-6: Feed forward encoder structure for 4-state STTC

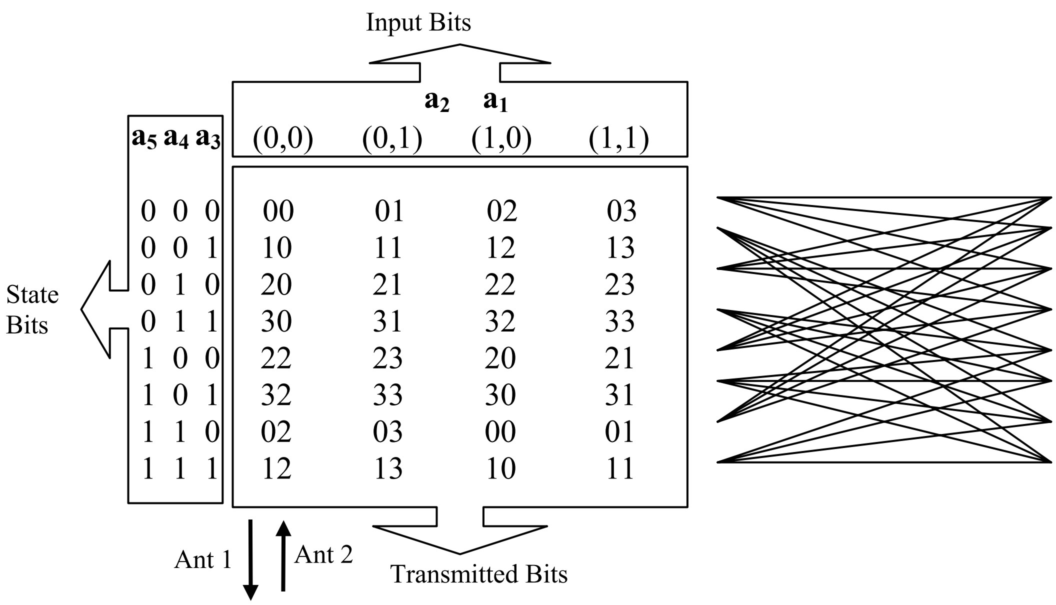

Figure 1-7: Trellis for 8-state STTC

The trellis diagram of an 8-state code (from [Taro1]) is shown in Figure 1-7. The encoder for the same can be represented by a closed analytical form given by

$$\begin{aligned} x_2(t) &= 2a_5(t) + 2a_2(t) + a_1(t) \\ x_1(t) &= 2a_5(t) + 2a_4(t) + a_3(t) \end{aligned} \qquad \text{1-2}$$

and implemented by feed forward structure as shown in Figure 1-8. The corresponding generator matrix of the code is $\begin{bmatrix} 1\,2\,0\,0\,2 \\ 0\,0\,1\,2\,2 \end{bmatrix}^T$.

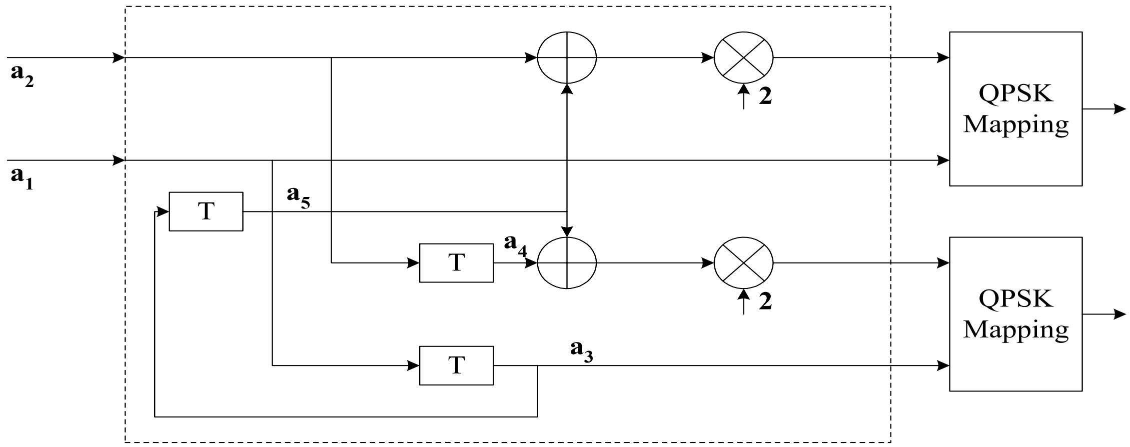

Figure 1-8: Feed forward encoder structure for 8-State STTC

The N_t symbols generated by the encoder are simultaneously transmitted from the N_t transmit antennas. The transmission occurs in frames of length L. The receiver front end consists of N_r receive antennas and the signal at each receive antenna is a noisy linear superposition of simultaneously transmitted N_t symbols weighted by the fade coefficients of the channel. The noise is assumed to be Additive White Gaussian Noise (AWGN). Consider $c_i(t)$ to be the symbol transmitted from the i^{th} transmit antenna at time instant t. Then, the received signal $r_j(t)$ at the j^{th} receive antenna at any time instant t is given by,

$$r_j(t) = \sum_{i=1}^{N_t} \alpha_{ij}(t) c_i(t) + \eta_j(t); 1 \le j \le N_r \qquad \text{1-3}$$

$\eta_j(t)$ is the Gaussian noise at time instant t and is modeled by zero mean complex Gaussian random process with variance $\frac{N_o}{2}$ per dimension. The channel is modeled by an $N_t \times N_r$ matrix $\Omega(t)$, whose entries $\alpha_{ij}(t)$ represent the complex Gaussian fading coefficient from the i^{th} transmit antenna to the j^{th} receive antenna at time instant t. The fade coefficients are assumed to have zero mean (for a Rayleigh fading channel) and variance of 0.5 per dimension. The receiver estimates the channel and uses a Maximum Likelihood Sequence Estimation (MLSE) decoder to decode the information bits. The MLSE decoder computes the lowest accumulated Euclidean distance metric over an entire frame to extract the most likely transmitted sequence. The branch metric used by the MLSE decoder in the presence of perfect channel estimates is given by

$$\sum_{j=1}^{N_r} \left| r_j(t) - \sum_{i=1}^{N_t} \alpha_{ij}(t) e_i(t) \right|^2 \qquad \text{1-4}$$

where $e_i(t), i = 1, 2, \ldots N_t$ is a candidate codeword.

The design of a STTC scheme seeks to maximize the Euclidean distance between codewords transmitted from different transmit antennas. A detailed discussion on the design criteria and optimal code-construction techniques for STTC over different fading channels is presented in the next chapter. The basic design principles involved in STTC are seen to be similar to TCM.

1.5. Organization of Thesis

This thesis analyzes the influence of channel estimation errors on the design and performance of STTC over different Rayleigh fading channels. It specifically investigates any alterations caused in the design criteria or the performance behavior due to the presence of estimation errors.

Design and code-construction criteria for STTC over different flat fading channels are presented in Chapter 2. The analyses in this chapter assume perfect channel estimates to be available at the receiver. The performance of various STTC schemes adhering to different design criteria are compared via simulations. An analytical method that determines the exact pair-wise error probability of STTC assuming perfect channel estimates and thus supplements the simulated performance analysis is also presented. Chapter 3 introduces several channel estimation schemes for multiple antenna systems and discusses their relative merits. It also presents a capacity analysis of the multi-input channel in the presence of channel estimation errors. This work is the summary of several previously published results.

Chapter 4 discusses the impact of channel estimation errors on the performance of STTC. It reevaluates the design criteria derived in Chapter 2 in the absence of perfect channel estimates at the receiver. The chapter derives an exact expression for the pair-wise error probability of STTC in the presence of channel estimation errors. It also attempts to identify and characterize factors (such as code-choice, the number of transmit/ receive antennas in the system, the amount of training used and the coherence time of the channel) that influence the extent of performance loss due to channel estimation errors in STTC systems.

Chapter 5 analyzes the performance of STTC in frequency selective channels. It presents a new design criterion for STTC that ensures a better performance than existing schemes over multi-path channels. It also evaluates the influence of channel estimation errors on the performance of STTC over multi-path channels.

Chapter 6 discusses the results and conclusions formed by the thesis and points to directions for future work.

2. STTC Performance Analysis and Design Criteria

Space Time Trellis Codes (STTC) are intended to exploit diversity in both space and time to improve the reliability of communications over fading channels using multiple antennas. A comprehensive study and analysis of the design criteria for STTC, which draws on the extensive literature available on the subject, is presented in this chapter. These code-construction criteria attempt to allow STTC to derive maximum benefit over different fading channels and are in general formulated by analyzing the expressions for pair-wise error probability of the codes. Some of the seminal work in this area was done by Tarokh et al in [Taro1]. Their paper dealt with design criteria for STTC over slow flat fading, fast flat fading and spatially correlated channels assuming high SNRs. Alternate criteria for STTC in channels with large possible diversities (larger number of transmit/receive antennas) were derived in [Chen1]. This chapter discusses these key results and their extensions. The performance of various STTC schemes adhering to different design criteria are compared via simulations. An analytical method of determining the exact pair-wise error probability of STTC derived in [Turi1], which provides a tool to verify and supplement the simulated performance analysis, is also presented. This chapter provides the necessary context for the main work in this thesis presented in Chapter 4.

2.1. System Model

A generalized model for Multi Input Multi Output (MIMO) systems employing STTC is presented in this section. All subsequent analyses in this thesis follow some form of this basic model.

A MIMO system with N_t transmit and N_r receive antennas is considered. The STTC encoder is defined by a generator matrix G and combines the encoding and symbol mapping procedure into a composite operation. The generator matrix is implemented by a feed-forward shift register with memory order v. According to the symbol constellation used, the input bit stream is subdivided into blocks of appropriate length and fed to the encoder. The input data is thus encoded and modulated into N_t parallel streams of symbols. The symbols are transmitted in frames of length L. The channel is modeled by an $N_t \times N_r$ matrix $\Omega(t)$, whose entry $\alpha_{ij}(t)$ represents the complex fading coefficient from the i^{th} transmit antenna to the j^{th} receive antenna at time instant t. The receiver consists of N_r receive antennas and a Maximum Likelihood Sequence Estimation (MLSE) decoder. The MLSE decoder computes the lowest accumulated Euclidean distance metric over an entire frame to extract the most likely transmitted sequence. The signal at the receiver is a noisy superposition of simultaneously transmitted symbols weighted by the fade coefficients. The noise is assumed to be Additive White Gaussian Noise (AWGN).

Consider $c_i(t)$ to be the symbol transmitted from the i^{th} transmit antenna at time instant t. To allow the validity of the following analysis to any modulation scheme, each symbol in the signal constellation is contracted by average symbol energy $\sqrt{E_s}$, such that the average energy of the signal constellation is 1.

The received signal $r_j(t)$ at the j^{th} receive antenna at any time instant t is given by,

$$r_j(t) = \sum_{i=1}^{N_t} \alpha_{ij}(t) c_i(t) \sqrt{E_s} + \eta_j(t) ; 1 \le j \le N_r \qquad \text{2-1}$$

where, $\eta_j(t)$ is the Gaussian noise at time instant t and is modeled by zero mean complex Gaussian random process with variance $\frac{N_o}{2}$ per dimension. E_s is the received energy per symbol per transmit and receive antenna. The average signal power at each receive antenna from each transmit antenna is assumed to be the same.

2.2. Design Criteria for STTC

2.2.1. Upper Bound on Pair-wise Error Probability and TSC Criteria

Design criteria for STTC over slow frequency non-selective, fast frequency non-selective and spatially correlated channels were presented in [Taro1]. These criteria were derived by examining expressions for the upper bound on pair-wise error probability. This analysis is presented below.

The transmitted codeword is assumed to be given by
$c = c_1(1)c_2(1)\ldots c_{N_t}(1)c_1(2)c_2(2)\ldots c_{N_t}(2)\ldots c_1(L)c_2(L)\ldots c_{N_t}(L)$
and the MLSE decoder is assumed to decide erroneously in favor of codeword
$e = e_1(1)e_2(1)\ldots e_{N_t}(1)e_1(2)e_2(2)\ldots e_{N_t}(2)\ldots e_1(L)e_2(L)\ldots e_{N_t}(L)$.

2.2.1. A. Quasi-static Channel with Independent Fade Coefficients

A quasi-static channel where the fade coefficients are constant over a frame and vary over consecutive frames is assumed. The sub-channels (from each transmit antenna to each receive antenna) are assumed to be mutually independent and are modeled as independent samples of a zero mean complex Gaussian random process with variance 0.5 per dimension.

Let α_{ij} be the fade coefficient over a frame between transmit antenna i and receive antenna j. Assuming perfect channel estimation, the Chernoff bound for the conditional pair-wise error probability is,

$$P\left(c \to e \mid \alpha_{ij}, i = 1,2,\ldots,N_t, j = 1,2,\ldots,N_r\right) \le \exp\left(\frac{-d^2(c,e)E_s}{4N_0}\right) \qquad \text{2-2}$$

where,

$$d^2(c,e)=\sum_{j=1}^{N_r}\sum_{t=1}^{L}\left|\sum_{i=1}^{N_t}\alpha_{ij}\left(c_i(t)-e_i(t)\right)\right|^2 \qquad \text{2-3}$$

Equation (2-3) can be rewritten as

$$d^2(c,e)=\sum_{j=1}^{N_r}\sum_{i=1}^{N_t}\sum_{i'=1}^{N_t}\alpha_{ij}\alpha_{i'j}^*\sum_{t=1}^{L}\left(c_i(t)-e_i(t)\right)\left(c_{i'}(t)-e_{i'}(t)\right)^* \qquad \text{2-4}$$

where, x^* denotes the complex conjugate of x. Substituting $\Omega_j=\left(\alpha_{1j},\alpha_{2j},\ldots,\alpha_{N_tj}\right)$ and

$$A_{pq}(c,e)=\sum_{t=1}^{L}\left(c_p(t)-e_p(t)\right)\left(c_q(t)-e_q(t)\right)^*, p=1,2,\ldots,N_t, q=1,2,\ldots,N_t,$$

$$d^2(c,e)=\sum_{j=1}^{N_r}\Omega_jA\Omega_j^* \qquad \text{2-5}$$

By definition, $A(c,e)$ is a Hermitian matrix. It can be written as $VA(c,e)V^*=D_e$, where V is unitary matrix, whose rows are made of the eigenvectors of $A(c,e)$ and D_e is a diagonal matrix whose elements are given by the eigenvalues $\lambda_i, i=1,2,\ldots,N_t$ of $A(c,e)$. By construction, the square root of $A(c,e)$ is given by the difference matrix,

$$B(c,e)=\begin{bmatrix} e_1(1)-c_1(1) & \cdots & \cdots & e_1(L)-c_1(L) \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ e_{N_t}(1)-c_{N_t}(1) & \cdots & \cdots & e_{N_t}(L)-c_{N_t}(L) \end{bmatrix} \qquad \text{2-6}$$

Hence the eigenvalues of $A(c,e)$ are nonnegative real numbers. Let $\left(\beta_{1j},\beta_{2j},\ldots,\beta_{N_tj}\right)=\Omega_jV^*$, then

$$d^2(c,e)=\sum_{j=1}^{N_r}\sum_{i=1}^{N_t}\lambda_i\left|\beta_{ij}\right|^2 \qquad \text{2-7}$$

As α_{ij} are samples of a complex Gaussian random process with zero mean and V is a unitary matrix, β_{ij} are also independent samples of a complex Gaussian random process with zero mean and variance 0.5 per dimension. $\left|\beta_{ij}\right|$ follows a Rayleigh distribution given by

$$p\left(\left|\beta_{ij}\right|\right)=2\left|\beta_{ij}\right|\exp\left(-\left|\beta_{ij}\right|^2\right)\text{ for }\left|\beta_{ij}\right|\geq 0 \tag{2-8}$$

The upper bound on the pair-wise error probability is the statistical average of $\exp\left(\dfrac{-E_s\sum_{j=1}^{N_r}\sum_{i=1}^{N_t}\lambda_i\left|\beta_{ij}\right|^2}{4N_0}\right)$ with respect to the distribution of $\left|\beta_{ij}\right|$ and is given by,

$$P\left(c\to e\right)\leq\left(\frac{1}{\prod_{i=1}^{N_t}\left(1+\lambda_i\dfrac{E_s}{4N_0}\right)}\right)^{N_r} \tag{2-9}$$

If r is the rank of A, then the matrix A has r nonzero eigenvalues, $\lambda_i, i=1,2,\ldots,r$. At high Signal to Noise Ratios (SNRs), $\lambda_i\dfrac{E_s}{4N_0}>>1\;\forall i$, Equation (2-9) can be written as

$$P\left(c\to e\right)\leq\left(\prod_{i=1}^{r}\lambda_i\right)^{-N_r}\left(\frac{E_s}{4N_0}\right)^{-rN_r} \tag{2-10}$$

The diversity advantage is equivalent to the power of the SNR in the denominator of the expression. The coding advantage is a measure of the gain of the system over an un-coded system offering the same level of diversity. It is seen from Equation (2-10) that the diversity advantage obtained is rN_r and can be increased by increasing the rank of the distance matrix or the number of receive and transmit antennas. The coding advantage obtained is given by $\left(\prod_{i=1}^{r}\lambda_i\right)^{-N_r}$. It can be improved by increasing the product of eigenvalues of the distance matrix. This leads to the following design criteria,

Rank Criterion: In order to maximize the diversity gain over the Rayleigh fading channel, the minimum rank of the distance matrix $A\left(c,e\right)$ over all pairs of distinct code-words c and e is to be maximized. Maximum diversity advantage of $N_t\times N_r$ is obtained if the matrix is full rank. Otherwise a diversity advantage of $r\times N_r$ is obtained where r is the rank of the matrix.

Determinant Criterion: The minimum of the product of eigenvalues $\lambda_i, i = 1,2,\ldots,r$ taken over all pairs of distinct code-words c and e must be maximized. For a full rank matrix this is equivalent to maximizing the minimum determinant of the distance matrix $A(c,e)$ over all possible code-word pairs.

These will be referred to in this thesis as the TSC–RD (Tarokh, Seshadri, Calderbank - Rank and Determinant) criteria.

It is seen from Equation (2-10) that the rank of the distance matrix is an exponent in the probability expression. Hence the minimum rank criterion becomes more important than the minimum determinant criteria in determining the code performance.

2.2.1. B. Fast Fading Channel with Independent Fade Coefficients

A fast-fading channel is assumed where the fade coefficients vary from one symbol to the next. The pair-wise error probability is approximated by

$$P\left(c \to e \mid \alpha_{ij}(t), i = 1,2,\ldots,N_t, j = 1,2,\ldots,N_r, t = 1,2,\ldots,L\right) \le \exp\left(\frac{-d^2(c,e)E_s}{4N_0}\right) \qquad \text{2-11}$$

where,

$$d^2(c,e) = \sum_{j=1}^{N_r}\sum_{t=1}^{L}\left|\sum_{i=1}^{N_t}\alpha_{ij}(t)\left(c_i(t) - e_i(t)\right)\right|^2 \qquad \text{2-12}$$

The channel coefficients $\alpha_{ij}(t)$ are assumed to vary independently from one symbol to the next in a fast fading channel and are hence included in the summation over the length of the frame in Equation (2-12) (unlike in Equation (2-4)).
Let $\Omega_j(t) = \left(\alpha_{1j}(t), \alpha_{2j}(t), \ldots, \alpha_{N_t j}(t)\right)$ and $C_{pq}(t) = \left(c_p(t) - e_p(t)\right)\left(c_q(t) - e_q(t)\right)^*$

From Equation (2-12),

$$d^2(c,e) = \sum_{j=1}^{N_r}\sum_{t=1}^{L}\Omega_j(t)C(t)\Omega_j^*(t) \qquad \text{2-13}$$

Since $C(t)$ is Hermitian, it can be written as $V(t)C(t)V^*(t) = D_e(t)$, where *V(t)* is a unitary matrix whose rows are made of the eigenvectors of $C(t)$ and $D_e(t)$ is a diagonal matrix whose elements are given by the eigenvalues $\lambda_i(t), i = 1,2,\ldots,N_t$ of $C(t)$. As $C(t)$ is a Hermitian matrix, its eigenvalues are real.

Let $\left(\beta_{1j}(t), \beta_{2j}(t), \ldots, \beta_{N_t j}(t)\right) = \Omega_j(t) V^*(t)$, then

$$\Omega_j(t) C(t) \Omega_j^*(t) = \sum_{i=1}^{N_t} \lambda_i(t) \left|\beta_{ij}(t)\right|^2 \tag{2-14}$$

As $\alpha_{ij}(t)$ are samples of a complex Gaussian random variable with mean zero and V is a unitary matrix, $\beta_{ij}(t)$ are independent samples of a complex Gaussian random process with mean zero and variance 0.5 per dimension. $\left|\beta_{ij}(t)\right|$ follows a Rayleigh distribution given by

$$p\left(\left|\beta_{ij}(t)\right|\right) = 2\left|\beta_{ij}(t)\right| \exp\left(-\left|\beta_{ij}(t)\right|^2\right) \text{for } \left|\beta_{ij}(t)\right| \geq 0 \tag{2-15}$$

The upper bound on the pair-wise error probability is found by averaging

$$P(c \to e) \leq \exp\left(\frac{-E_s \sum_{j=1}^{N_r} \sum_{t=1}^{L} \sum_{i=1}^{N_t} \lambda_i(t) \left|\beta_{ij}(t)\right|^2}{4N_0}\right) \tag{2-16}$$

with respect to the distribution of $\left|\beta_{ij}(t)\right|$ and is given by,

$$P(c \to e) \leq \left(\frac{1}{\prod_{t=1}^{L} \prod_{i=1}^{N_t} \left(1 + \lambda_i \frac{E_s}{4N_0}\right)}\right) \tag{2-17}$$

The columns of $C(t)$ are multiples of $c(t) - e(t) = \left(c_1(t) - e_1(t), c_2(t) - e_2(t), \ldots, c_{N_t}(t) - e_{N_t}(t)\right)$. Hence $C(t)$ has rank 1 only if $c_1(t) c_2(t) \ldots c_{N_t}(t) \neq e_1(t) e_2(t) \ldots e_{N_t}(t)$ and rank zero otherwise. Thus the only possible nonzero eigenvalue of $D(t)$ is $\left|c(t) - e(t)\right|^2$. Substituting in Equation (2-17),

$$P(c \to e) \leq \left(\frac{1}{\prod_{t=1}^{L} \left(1 + \left|c(t) - e(t)\right|^2 \frac{E_s}{4N_0}\right)}\right) \tag{2-18}$$

Let l be the number of time instances in a frame that $\left|c(t)-e(t)\right| \neq 0$, then Equation (2-18) can be expressed as

$$P(c \rightarrow e) \leq \left(\prod_{t=1}^{l} \left|c(t)-e(t)\right|^2 \right)^{-N_r} \left(\frac{E_s}{4N_0} \right)^{-lN_r} \qquad \text{2-19}$$

Diversity advantage is seen to be governed by lN_r. The design criteria follow as,

Distance Criterion: The diversity gain can be maximizing the number of time instances for which the strings $c_1(t)c_2(t)\ldots c_{N_t}(t)$ and $e_1(t)e_2(t)\ldots e_{N_t}(t)$ of any pair of distinct code-words differ during the duration of a frame. If the code-words differ for l time instances in a frame, a diversity advantage of $l \times N_r$ is obtained.

Product criterion: To maximize coding gain, the minimum of the product of the distances between all pairs of code-words at, time instances when the distances are not zero, should be maximized.

These will be referred to as the TSC–DP (Tarokh, Seshadri, Calderbank – Distance and Product) criteria.

2.2.1. C. Dependent Quasi-static Fade Coefficients

The channels are assumed to be correlated and channel coefficients are modeled by samples of dependent zero mean complex Gaussian random variables with variance 0.5 per dimension. Let $Y(c,e)$ be a block diagonal matrix with dimension $N_t \times N_r$ whose diagonal elements are *A(c,e)* and let $\Omega = \left(\Omega_1, \Omega_2, \ldots, \Omega_{N_r}\right)$. The Pair-wise error probability is then given by

$$P\left(c \rightarrow e \mid \alpha_{ij}, i=1,2,\ldots,N_t, j=1,2,\ldots,N_r\right) \leq \exp\left(\frac{-\Omega Y(c,e) \Omega^* E_s}{4N_0} \right) \qquad \text{2-20}$$

Let Θ be the correlation matrix of Ω, $\Theta = E\left[\Omega^* \Omega\right]$. Θ is assumed to be full rank. As Θ is a nonnegative definite Hermitian matrix, it has a square root Ψ. The diagonal elements of Θ are equal to unity. Hence the rows of Ψ are also of length one. Let $\omega = \Omega\left(\Psi^*\right)^{-1}$, then the elements of ω are samples of uncorrelated complex Gaussian random process with variance of 0.5 per dimension. ω thus follows the Rayleigh distribution. From Equation (2-20),

$$P(c \to e) \le \exp\left(\frac{-E_s \omega \Psi^* Y(c,e) \Psi \omega^*}{4N_0}\right) \quad 2\text{-}21$$

This is similar to the case of independent fading coefficients with $A(c,e)$ replaced by $\Psi^* Y(c,e)\Psi$.

$$P(c \to e) \le \left(\prod_{i=1}^{r} \lambda_i\right)^{-N_r} \left(\frac{E_s}{4N_0}\right)^{-rN_r} \quad 2\text{-}22$$

Where r is the rank of $\Psi^* Y(c,e)\Psi$ and $\lambda_i, i = 1,2,\ldots,r$ are the eigenvalues of $\Psi^* Y(c,e)\Psi$. From Equation (2-22), as Ψ is full rank, maximizing the rank of $\Psi^* Y(c,e)\Psi$ is equivalent to maximizing the rank of $Y(c,e)$, which is in turn equivalent to maximizing the rank of $A(c,e)$. Maximizing the determinant of $\Psi^* Y(c,e)\Psi$ is also equivalent to maximizing the determinant of $A(c,e)$. Thus comparing with the first case (Section 2.2.1.A), it is seen that the design criteria for the case of independent quasi-static channel coefficients also holds for the case of dependent quasi-static channel coefficients.

2.2.2. Generalized Design Criteria

Generalized design criteria were formulated in [Gama1] that determined the diversity and coding advantage achieved by STTC in MIMO block-fading channels. By fixing the size of the blocks to appropriate values, the design criteria in [Taro1] (Section 2.2.1) for quasi-static and fast fading MIMO channels were derived as special cases of the new criteria.

Consider that each frame of size L has M blocks over which the channel coefficients are constant. Then the fade coefficients are constant over $\frac{L}{M}$ consecutive symbol durations. Let $\alpha_{ij}(m)$ be the fading coefficient for the m^{th} fading block. The other parameters can also be expressed accordingly over a fading block.

$$R_j[m] = \left[r_j\left(\frac{(m-1)L}{M}+1\right), \ldots, r_j\left(\frac{mL}{M}\right)\right]_{1\times\frac{L}{M}}$$

$$N_j[m] = \left[\eta_j\left(\frac{(m-1)L}{M}+1\right), \ldots, \eta_j\left(\frac{mL}{M}\right)\right]_{1\times\frac{L}{M}}$$

$$H_j[m] = \left[\alpha_{1j}[m], \ldots, \alpha_{N_t j}[m]\right]_{1\times N_t}$$

$$c[m]=\begin{bmatrix} c_1\left(\frac{(m-1)L}{M}+1\right) & \cdots & c_1\left(\frac{mL}{M}\right) \\ \vdots & \ddots & \vdots \\ c_{N_t}\left(\frac{(m-1)L}{M}+1\right) & \cdots & c_{N_t}\left(\frac{mL}{M}\right) \end{bmatrix}_{N_t\times\frac{L}{M}}$$

Then, for $1\le m\le M$, the received signal can be expressed as,

$$R_j[m]=H_j[m]c[m]+N_j[m] \tag{2-23}$$

The pair-wise error probability of transmitting a code word c and deciding erroneously in favor of a code-word e can be approximated by

$$P\left(c\to e\,|\,\alpha_{ij}, i=1,2,\ldots,N_t, j=1,2,\ldots,N_r\right)\le \exp\left(\frac{-d^2(c,e)E_s}{4N_0}\right) \tag{2-24}$$

where,

$$d^2(c,e)=\sum_{j=1}^{N_r}\sum_{m=1}^{M}\left\|H_j[m]\left(c[m]-e[m]\right)\right\|^2 \tag{2-25}$$

For a matrix X, $\|X\|^2=XX^*$. Following the approach in [Taro1], the upper bound on the pair-wise probability of error can be written as

$$P(c\to e)\le\prod_{m=1}^{M}\left(\frac{\mu_m E_s}{4N_0}\right)^{-d_m N_r} \tag{2-26}$$

where,

$d_m=rank\left(c[m]-e[m]\right)$, $\mu_m=\left(\lambda_1[m]\lambda_2[m]\ldots\lambda_{d_m}[m]\right)^{\frac{1}{d_m}}$ and $\lambda_1[m],\lambda_2[m],\ldots,\lambda_{d_m}[m]$ are eigenvalues of $A[m]=\left(c[m]-e[m]\right)\left(c[m]-e[m]\right)^H$.

The diversity order of the system is given by

$$d=\sum_{m=1}^{M}d_m=\sum_{m=1}^{M}rank\left(c[m]-e[m]\right) \tag{2-27}$$

and the coding gain of the system is given by,

$$\mu = \left(\prod_{m=1}^{M} \lambda_1[m]\lambda_2[m]\ldots\lambda_{d_m}[m] \right)^{\frac{1}{d}} \quad \text{2-28}$$

Hence the design criterion for block fading channels is as follows.

- *Block Fading Sum of Ranks Criterion*
  To maximize the diversity advantage, d is to be maximized over all pairs of distinct code-words c and e.

- *Block Fading Product Distance Criterion*
  To maximize the coding gain, μ is to be maximized over all pairs of distinct code-words c and e.

The design criteria for quasi-static and fast fading channels can be obtained from the block fading criteria by letting $M = 1$ *and* $M = L$ respectively.

2.2.3. Trace Criterion

It was shown in [Taro1] (Section 2.2.1) that the rank criterion is more important than the determinant criterion in determining code performance. However, the difference matrix, $B(c,e)$ (defined in Equation (2-6)) of dimension $N_t \times L$, has a maximum rank given by the $\min(N_t, v)$ where v is the constraint length of the code. Hence a full rank value of N_t is not always achievable. A new design criterion for STTC was introduced in [Chen1] which did not require the difference matrix to have full rank. It was shown in [Chen1] that when STTC is used in systems with a large product of the number of transmit and receive antennas (>3), the multiple fading sub-channels converge to an additive white Gaussian channel. The new design criterion (described below) takes advantage of this approximation.

Assume that a maximum likelihood receiver decides erroneously in favor of a signal $e = e_1(1)e_2(1)\ldots e_{N_t}(1)e_1(2)e_2(2)\ldots e_{N_t}(2)\ldots e_1(L)e_2(L)\ldots e_{N_t}(L)$ assuming that $c = c_1(1)c_2(1)\ldots c_{N_t}(1)c_1(2)c_2(2)\ldots c_{N_t}(2)\ldots c_1(L)c_2(L)\ldots c_{N_t}(L)$ was transmitted. Let r $(r \le N_t)$ be the rank of the difference matrix *B(c, e)* and let λ_i be the eigenvalues of the distance matrix $A(c,e)$. As mentioned earlier, $A(c,e)$ is the square of matrix *B(c, e)* by construction.

2.2.3. A. Quasi-static Channel

The conditional pair-wise probability is upper bounded by (from [Gama1] and Equation (2-2)),

$$p\left(c \rightarrow e \mid \alpha\right) \leq \exp\left(-d^2(c,e)\frac{E_s}{4N_0}\right)$$

where,

$$d^2(c,e) = \sum_{j=1}^{N_r}\sum_{t=1}^{L}\left|\sum_{i=1}^{N_t}\alpha_{i,j}\left(c_i\left(t\right) - e_i\left(t\right)\right)\right|^2$$

The conditional probability can also be expressed as (from [Taro1] or substituting Equation (2-7) in Equation (2-2)),

$$p\left(c \rightarrow e \mid \alpha\right) \leq \exp\left(-\sum_{j=1}^{N_r}\sum_{i=1}^{N_t}\lambda_i\left|\beta_{ij}\right|^2\frac{E_s}{4N_0}\right) \qquad \text{2-29}$$

$\left|\beta_{ij}\right|^2$ follows the central chi square distribution since $\left|\beta_{ij}\right|$ is Rayleigh. Its mean and variance is equal to 1 (From [Gama1]). For a large $N_t N_r$ (>3) value, according to Central Limit Theorem, the expression $\sum_{j=1}^{N_r}\sum_{i=1}^{N_t}\lambda_i\left|\beta_{ij}\right|^2$ approaches a Gaussian random variable *D* with mean

$$\mu_D = N_r\sum_{i=1}^{N_t}\lambda_i \qquad \text{2-30}$$

and variance

$$\sigma_D^2 = N_r\sum_{i=1}^{N_t}\lambda_i^2 \qquad \text{2-31}$$

Thus, the unconditional pair-wise error probability can be upper bounded by [Taro1],

$$p\left(c \rightarrow e\right) \leq \int_0^{\infty}\exp\left(-\frac{E_s}{4N_0}D\right)p(D)dD \qquad \text{2-32}$$

where, *p(D)* is a Gaussian distribution.

$$p\left(c \rightarrow e\right) \leq \exp\left(\frac{1}{2}\left(\frac{E_s}{4N_0}\right)^2\sigma_D^2 - \frac{E_s}{4N_0}\mu_D\right)Q\left(\frac{\frac{E_s}{4N_0}\sigma_D^2 - \mu_D}{\sigma_D}\right) \qquad \text{2-33}$$

By using $Q\left(x\right) \leq \frac{1}{2}e^{-\frac{x^2}{2}}\ x \geq 0$,

$$p\left(c \rightarrow e\right) \leq \frac{1}{2}\exp\left(-N_r\frac{E_s}{4N_0}\sum_{i=1}^{N_t}\lambda_i\right) \qquad \text{2-34}$$

The pair-wise error probability can be minimized by maximizing the sum of the eigenvalues of the matrix $A(c,e)$. For a square matrix the sum of the eigenvalues equals the trace of the matrix. The trace of matrix $A(c,e)$ can be written as

$$\text{tr(A)}=\sum_{i=1}^{N_t}\sum_{t=1}^{L}\left|e_i(t)-c_i(t)\right|^2 \qquad \text{2-35}$$

Thus the pair wise error probability can be minimized if the minimum Euclidean distance between any two code words is maximized. Hence it is shown that when the diversity order $rN_r \geq 3$ the maximum coding gain is governed by the Euclidean distance between any two code-words over all transmit antennas. This criterion was referred to by the authors of [Chen1] as the *trace* criterion. It is also noted that when the number of transmit antennas is equal to two, it is important for the distance matrix $A(c,e)$ to be full rank. If $N_t \geq 3$, then the full rank criterion is not necessary.

2.2.3. B. Fast-Fading Channels

An extension of the design criteria for systems with large diversities to fast fading channels was presented in [Vuce1] and is described in this section.
From the analysis for STTC over a fast fading channel discussed before, expression (2-14), $\Omega_j(t)C(t)\Omega_j^*(t)=\sum_{i=1}^{N_t}\lambda_i(t)\left|\beta_{ij}(t)\right|^2$, can be written as

$$\Omega_j(t)C(t)\Omega_j^*(t)=\sum_{i=1}^{N_t}\left|c(t)-e(t)\right|^2\left|\beta_{ij}(t)\right|^2 \qquad \text{2-36}$$

as the only possible nonzero eigen value of $D(t)$ is $\left|c(t)-e(t)\right|^2$. By averaging (2-36) over the Gaussian random variable, the PWEP can be upper bounded by,

$$p(c\to e)\leq\frac{1}{2}\exp\left(\frac{1}{2}\left(\frac{E_s}{4N_0}\right)^2 N_rD^4-\frac{E_s}{4N_0}N_rd_E^2\right)Q\left(\frac{E_s}{4N_0}\sqrt{N_rD^4}-\frac{\sqrt{N_r}d_E^2}{\sqrt{D^4}}\right) \qquad \text{2-37}$$

where, $d_E^2=\sum_{t=1}^{L}\left|c(t)-e(t)\right|^2$ and $D^4=\sum_{t=1}^{L}\left|c(t)-e(t)\right|^4$. Hence it is seen that in the presence of a large diversity order, the Trace criterion is valid for fast fading channels as well.

2.3. Code Construction

In this section an overview of the code construction methods employing the design criteria derived in the previous sections are given.

The STTC Encoder can be modeled by a feed-forward shift register with a memory order of v. This structural representation helps in code-construction and is described below. Consider a QPSK system with N_t transmit antennas. The encoder has two branches with memory orders v_1 and v_2. At any given time two binary inputs, $I_1(t)$ and $I_2(t)$ are fed to the encoder. These input streams are passed through their respective shift register branches and multiplied by coefficient vectors given by,

$a(p)=\left(a_1(p),a_2(p),\ldots,a_{N_t}(p)\right)$ and $b(q)=\left(b_1(q),b_2(q),\ldots,b_{N_t}(q)\right)$ respectively,

where, $a_i(p),b_i(q)\in\{0,1,2,3\}$, $i=1,2,\ldots,n_t$, $p=1,2,\ldots,v_1$, $q=1,2,\ldots,v_2$.

The symbol transmitted on the i^{th} antenna at time t is computed as

$$c_t^i=\left(\sum_{p=0}^{v_1} I_1(t-p)a_i(p)+\sum_{q=0}^{v_2} I_2(t-q)b_i(p)\right)\bmod 4 \qquad \text{2-38}$$

The associated feed forward structure is illustrated in Figure 2-1. The generator matrix G is formed by the branch weights and is given by,

$$G^T=\begin{bmatrix} a_1(0) & b_1(0) & \cdots & a_1(v) & b_1(v) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N_t}(0) & b_{N_t}(0) & \cdots & a_{N_t}(v) & b_{N_t}(v) \end{bmatrix} \qquad \text{2-39}$$

where, $v=\max(v_1,v_2)$.

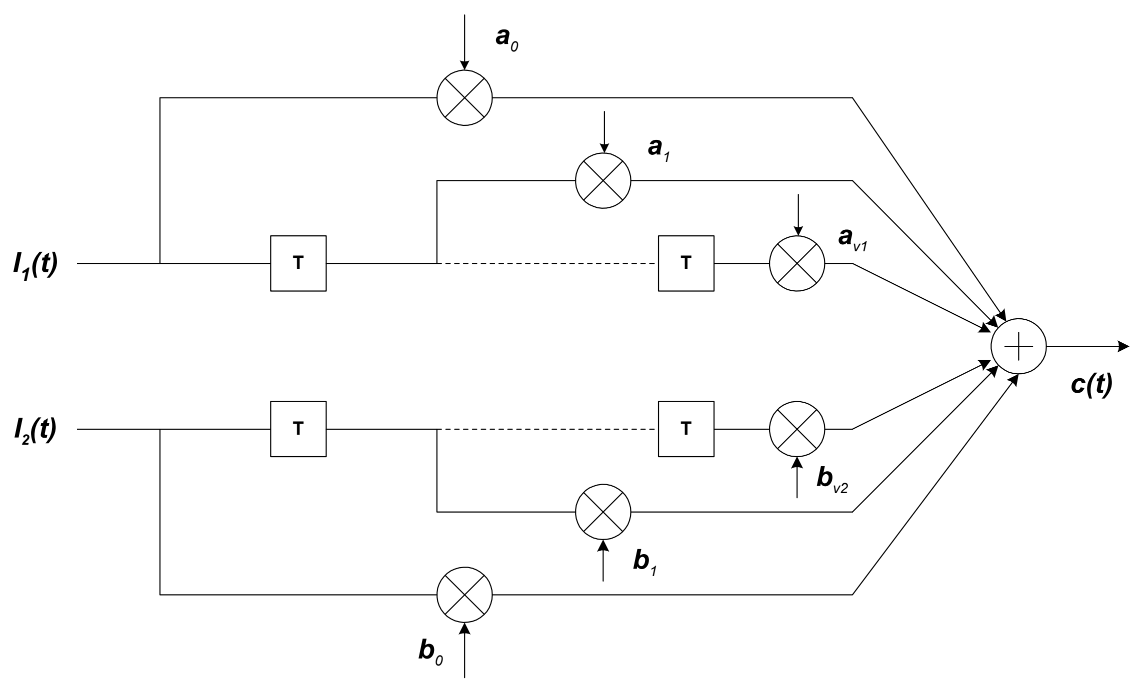

Figure 2-1: STTC feed forward encoder structure

If any branch has memory order less than v, then some of its columns are allowed to be absent. For instance, if the first branch has memory order v-1, then its column with elements $a_i(v)$, $i=1,2,\ldots,N_t$, is absent.

2.3.1. TSC Criteria

2.3.1. A. Quasi-static Channel

The TSC-RD criteria discussed earlier requires the code distance matrix to be full rank to guarantee full diversity advantage. The geometric uniformity of the code trellis can be used to ensure full rank for any given code. Two heuristic design rules were described in [Taro1] that guaranteed full rank and hence maximum diversity for a system with two transmit antennas.

- *Design Rule 1*: Transitions departing from the same state differ in the second symbol.
- *Design Rule 2*: Transitions arriving at the same state differ in the first symbol

When these rules are followed, the code-word difference matrix $B(c,e)$ has the form $B(c,e)=\begin{bmatrix} 0 & *\cdots* & e_1(L)-c_1(L) \\ e_2(1)-c_2(2) & *\cdots* & 0 \end{bmatrix}$, which guarantees full rank. Hence the space-time code achieves two-level spatial diversity. The example codes in [Taro1] are constructed using these design rules and will be referred to as the *TSC* codes in this thesis.

The design rules in [Taro1] were generalized in [Grim1] to any level of diversity. A code that has the property of zeros symmetry (i.e. every code-word difference matrix is upper and lower triangular) is full rank. This criterion ensures full rank and reduces the search for good codes. However, this rule might be overly restrictive at times. Example codes constructed according to this symmetry property are given in [Grim1].

Space Time Trellis Codes were represented by a much more amenable “Generator” form, than the trellis form in [Baro1]. A systematic code search was then carried out by varying the values of the generator matrix and generating code and error sequences specific to a generator matrix. As the rank criterion is predominant, a generator matrix was discarded as soon as it did not achieve full rank for any pair of sequences. It was found that geometric uniformity could be used to limit computer searches for good space-time codes, but could not be treated as a necessary condition for good codes. Codes were presented that have coding gain larger than the codes presented in [Taro1] for the same decoder complexity. These codes are referred to here as the *BBH* (Baro, Bauch and Hansmann) codes.

2.3.1. B. Fast Fading Channels

Space Time Trellis Codes which best satisfy the TSC-DP criterion were presented in [Firm1] (These codes will be referred to here as the *FVY -Firmanto, Vucetic and Yuan* codes). The STTC encoder was represented by a feed forward shift register with appropriate memory order. An exhaustive and systematic search was done by varying the multiplication coefficients of the shift registers. The new codes were shown to have a much larger minimum product distance than the *TSC* or *BBH* codes for the same memory order and give a better performance in fast fading channels while maintaining a comparable performance in slow fading channels.

The TSC-DP criterion for fast fading channels was improved upon in [Safa1]. The paper also proposed a systematic construction procedure for STTC in fast fading channels. Example codes, constructed according the new criterion, were shown to give improved performance in fast fading channels.

2.3.2. Trace Criterion

STTC codes constructed according to the Trace criterion were presented in [Chen1] (for two transmit antennas) and in [Spas1] for higher numbers of transmit antennas. The feed forward structure with the generator matrix described previously was utilized for a systematic and exhaustive computer search to identify the codes. These codes will be referred to as the CYV codes in this thesis.

2.4. Performance Results

The performance of space time trellis codes designed using different criteria are compared by simulating them over different fading environments. In the simulations, the modulation scheme used is QPSK and the length of the frame is 130 unless otherwise mentioned. The encoder is required to be in the zero state at the beginning and end of each frame. Perfect knowledge of the channel is assumed at the receiver. The branch metric used by the MLSE Viterbi decoder is

$$\sum_{j=1}^{N_r}\left|r_j\left(t\right)-\sum_{i=1}^{N_t}\alpha_{i,j}c_i\left(t\right)\right|^2 \qquad \text{2-40}$$

The performance curves are plotted against SNR, the symbol energy to noise ratio per receive antenna.

2.4.1. TSC Criteria

The performance of the codes designed using the *TSC* criteria is compared over different fading scenarios in this section.

2.4.1. A. Quasi-Static Channel with Independent Fade Coefficients

The performances of different STTC codes designed for optimum performance in Quasi-Static fading channels are compared. The generator matrices of the codes are given in Table 2-1. All the codes given in the table have full rank. The “Calderbank Mazo”

Algorithm described in [Goza1] is used to translate the code trellis structures in [Taro1] to closed analytical forms.

Code Type	n_t	No: of States	Generator Matrix	Min Det
TSC	2	4	$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix}$	2
BBH	2	4	$\begin{pmatrix} 2 & 0 & 1 & 3 \\ 2 & 2 & 0 & 1 \end{pmatrix}$	$\sqrt{8}$
TSC	2	8	$\begin{pmatrix} 2 & 1 & 0 & 0 & 2 \\ 0 & 0 & 2 & 1 & 2 \end{pmatrix}$	$\sqrt{12}$

Table 2-1: STTC codes for quasi-static channel

From Figure 2-2, it is clearly seen that diversity improves the performance of a system. The system with a diversity order of six (two transmit, three receive antennas) outperforms the system with a diversity of two (two transmit, one receive antenna) which is in turn better than a system with no diversity. Maximum Ratio Receive Combining (MRRC) scheme (without any coding) with two receive antennas is simulated to compare its performance with the corresponding STTC scheme with a diversity of two provided by two transmit antennas. The performance of the STTC scheme is about 3dB worse than the MRRC scheme. This can be attributed to the fact that in the case of STTC, the power transmitted from the two transmit antennas is halved so that the total transmitted power is the same for both STTC and MRRC schemes. In other words, MRRC benefits from a 3dB aperture gain. The BBH code has the same rank as the TSC code. Hence the diversity advantage of both codes is the same. This is confirmed by the identical slopes of their frame error rate curves. However, the BBH code has a larger minimum determinant value and hence is capable of providing more coding gain than the TSC code. This manifests itself as a horizontal shift in performance curve for the two receive antenna case. This performance improvement due to coding gain is not observed in the single receive antenna case, as the performance of low diversity systems is dominated by the rank criterion.

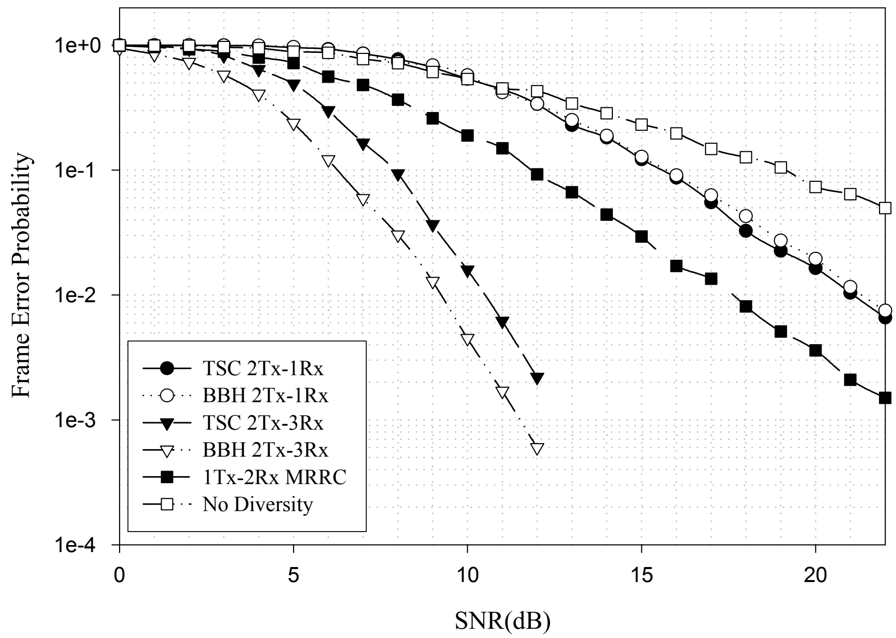

Figure 2-2: Performance comparison of 4 state-STTC codes in quasi-static channel

Figure 2-3 shows that the 8 State STTC performs better than the 4state STTC, due to the additional coding gain achieved by the 8 state STTC.

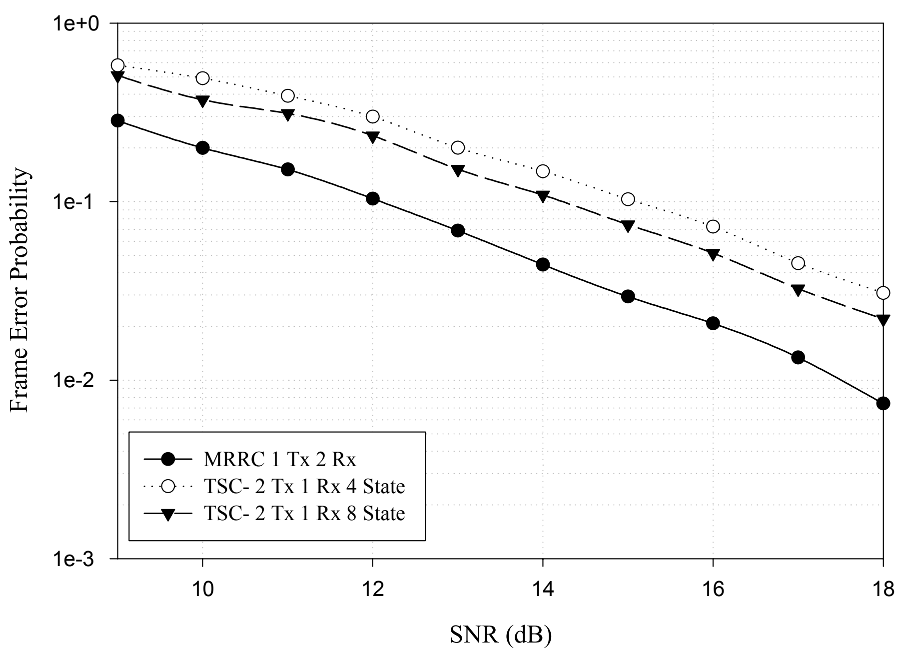

Figure 2-3: Performance comparison of 4 and 8 state STTC codes in quasi-static channel

2.4.1. B. Fast Fading Channel with Independent Fade Coefficients

Code Type	n_t	No: of States	Generator Matrix	Minimum Determinant	Minimum Product Distance
FVY	2	4	$\begin{pmatrix} 3 & 2 & 2 & 0 \\ 1 & 2 & 1 & 2 \end{pmatrix}$	$\sqrt{8}$	24
BBH	2	4	$\begin{pmatrix} 2 & 0 & 1 & 3 \\ 2 & 2 & 0 & 1 \end{pmatrix}$	$\sqrt{8}$	8

Table 2-2: STTC codes for fast fading channels

The FVY code from [Firm1], designed for optimum performance in fast fading channels by using the TSC-DP criteria is analyzed in this section. The generator matrices of the codes are given in Table 2-2. The minimum product distances of the codes are also compared.

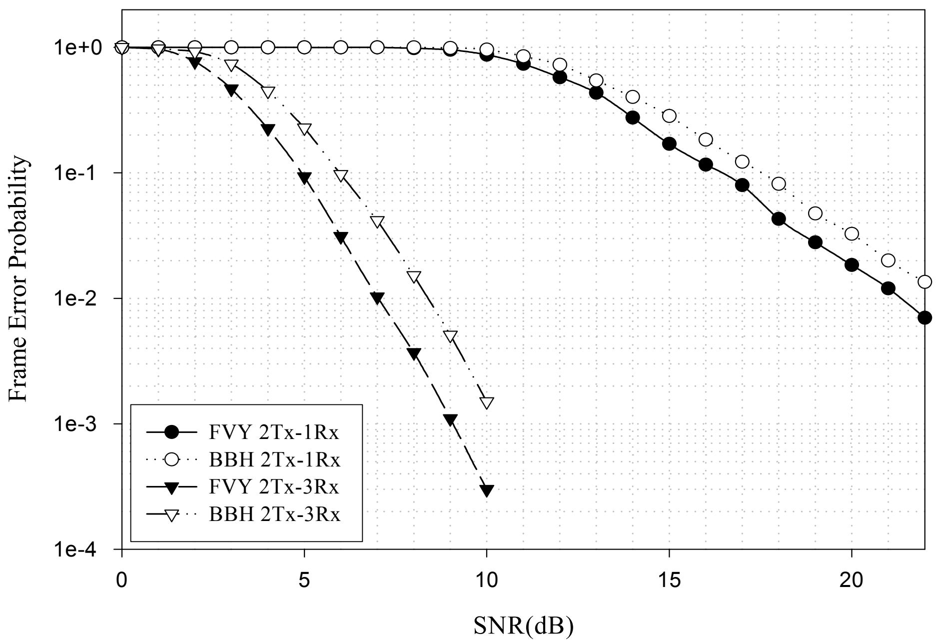

Figure 2-4: Performance comparison of STTC codes in fast fading channel

It is seen from Figure 2-4 that the FVY code performs better than the BBH code which has a larger value of the minimum determinant but a smaller value of the minimum product distance. Hence it is seen that minimum product distance is very important in fast fading channels. The effect is seen to be more pronounced as the number of receive antennas is increased. A clearer picture is presented by the bit error rate plots (Figure 2-5).

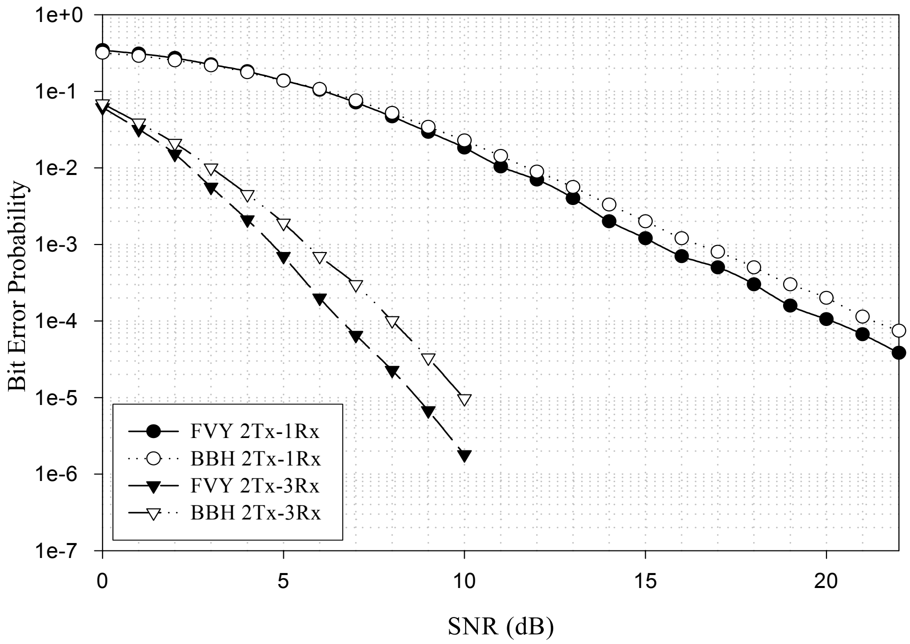

Figure 2-5: BER comparison of STTC codes in fast fading channel

2.4.1. C. *Spatially Correlated Channel*

The transmit antennas are assumed to be correlated by a specified factor and the receive antennas are assumed to undergo independent fading. The correlation across channels is modeled by the method given in [14, Appendix D]. The TSC code is simulated for a system with two transmit antennas and two receive antennas.

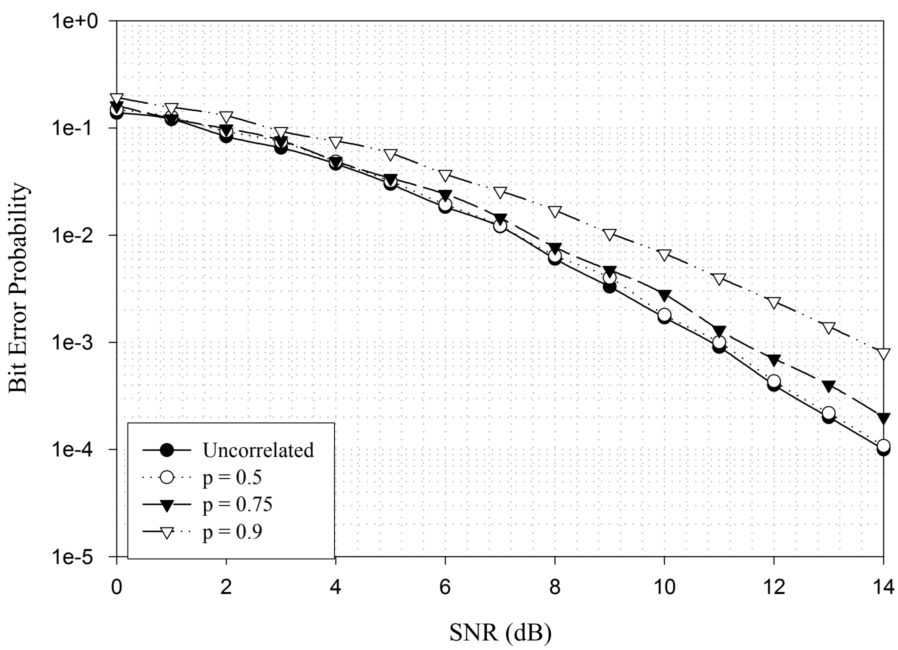

Figure 2-6: BER comparison of STTC codes in spatially correlated channel

From Figure 2-6, it is observed that for a correlation factor ≤ 0.75, the diversity advantage is maintained but a loss in coding gain is observed. For a higher correlation factor the system performance degrades considerably and a loss is seen in the diversity advantage as well.

2.4.2. Trace Criterion

2.4.2. A. Quasi-Static Channel

Code Type	n_t	No: of States	Generator Matrix	Minimum Determinant	Minimum Trace
TSC	2	4	$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix}$	2	4
CYV	2	4	$\begin{pmatrix} 0 & 2 & 1 & 2 \\ 2 & 3 & 2 & 0 \end{pmatrix}$	2	10

Table 2-3: STTC codes for quasi-static channel with large diversity order

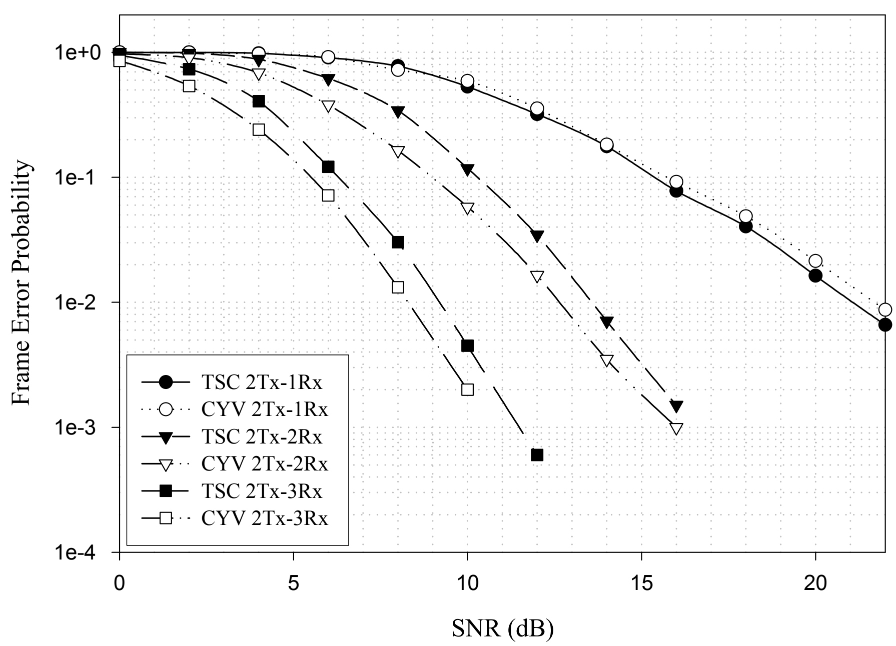

Figure 2-7: Performance comparison of STTC codes in quasi-static channel

The CYV code, designed according to the trace criteria, is analyzed in this section. Figure 2-7 compares the performance of the TSC code and the CYV code given in Table 2-3 over a quasi-static channel. Both codes have full rank.

For systems with low diversity orders, the rank criterion is dominant. TSC and CYV, have identical ranks of two. Hence their performance is comparable in the system with two-transmit and one receive antenna. However, in systems with diversity order >3, the CVY code is seen to outperform the TSC code. The advantage offered by the CYV code also increases with increasing number of receive antennas. Hence the trace criterion is appropriate for higher diversity order systems.

2.4.3. Comparison of Trace and TSC Criteria

When the diversity order of a system is very small, it becomes very important to maximize the diversity advantage offered by the code. Hence, in low diversity systems, the rank criterion is very important. It is seen through simulation analysis ([Gama1] and Figure 2-7) that after the rank of the code has been maximized, the determinant criterion offers very slight improvement in coding gain over the trace criterion in low diversity systems. A code that is not full rank but which has been optimized with respect to the trace criterion performs worse compared to codes that have full rank. This illustrates the significance of the rank criterion in low diversity systems.

In contrast, improving the coding advantage is more important in higher diversity order systems as the system already has large diversity gains. Also, in the presence of large diversities, the composite fading channel tends towards an Additive White Gaussian Noise channel. Hence the design criterion for maximizing coding gain in an AWGN environment (the Euclidean distance criterion or the trace criterion), becomes valid in a system with large diversities. It is shown in [Gama1] that in a system with a large diversity order, a code that is not full rank but has been optimized with respect to the trace criterion gives a much better performance than a code that is full rank and has been optimized with respect to the determinant criterion (and has a lower trace value than the former code). It is also shown in [Gama1] that for a system with two transmit antennas the maximum value of the minimum trace is the same with or without the codes having full rank. But for more than two transmit antennas, the maximum value of the minimum trace is much larger for codes without full rank as compared to codes with full rank. Hence the trace criterion is much more significant than the rank and determinant criterion in systems with a large diversity order.

2.4.3. B. Fast Fading Channel

A comparison of the performance of the CYV, FVY and BBH codes in fast fading channel is given in Figure 2-8. The CYV code outperforms the BBH code and gives a performance comparable to the FVY code which is designed specifically for fast fading channels using the TSC-DP criteria. This is expected since both TSC-DP and the trace criteria essentially try to maximize the minimum Euclidean distance between code-words. Also, the trace criterion is designed for high diversity order systems and the fast fading channel offers large temporal diversity. It is thus seen that the same design criteria can be used to give optimal performances in two different types of environments. This is unlike the design criteria suggested in [Taro1] that specifies different criteria for slow and fast fading channels.

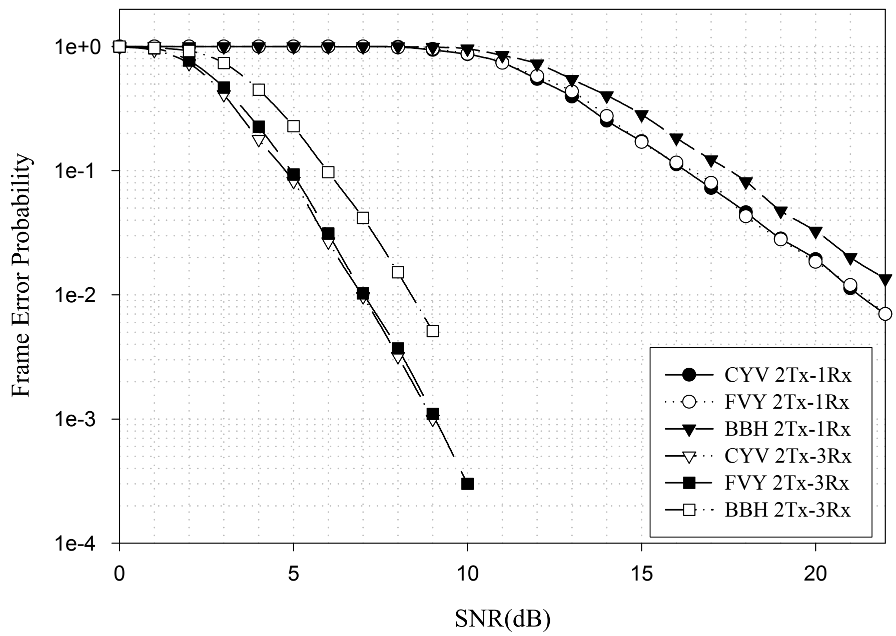

Figure 2-8: Performance comparison of STTC codes in fast fading channel

2.4.3. Performance Sensitivity of STTC to Coherence Time

Coherence time is a statistical measure of the time duration over which the channel impulse response is invariant. Shorter coherence times correspond to channel coefficients that change after shorter time intervals. Hence a decrease in the coherence time of the channel increases the temporal diversity provided by the channel. The performance of STTC over different channel coherence times and its capacity to derive benefit from the additional available temporal diversity is discussed in this section.

Equation (2-2) gives the Chernoff bound for the conditional pair-wise error probability. Consider that the fade coefficients are constant over l_1 symbols in a frame of length L. Then, $d^2(c,e)$ in Equation (2-2) can be expressed as,

$$d^2(c,e)=\sum_{j=1}^{N_r}\left(\sum_{t=1}^{l_1}\left|\sum_{i=1}^{N_t}\alpha_{1,i,j}\left(c_i(t)-e_i(t)\right)\right|^2+\sum_{t=l_1+1}^{L}\left|\sum_{i=1}^{N_t}\alpha_{2,i,j}\left(c_i(t)-e_i(t)\right)\right|^2\right) \qquad \text{2-41}$$

where, $\alpha_{1,i,j}$ is the channel coefficient from the i^{th} transmit antenna to the j^{th} receive antenna over the first l_1 symbols in the frame and $\alpha_{2,i,j}$ is the corresponding channel coefficient over the last $L-l_1$ symbols in the frame. Equation (2-41) can be rewritten as,

$$d^2(c,e) = \sum_{j=1}^{N_r} \left(\begin{array}{l} \sum_{i=1}^{N_t} \sum_{i'=1}^{N_t} \alpha_{1,i,j} \overline{\alpha_{2,i',j}} \sum_{t=1}^{l_1} \left(c_i(t) - e_i(t)\right) \overline{\left(c_{i'}(t) - e_{i'}(t)\right)} \\ + \sum_{i=1}^{N_t} \sum_{i'=1}^{N_t} \alpha_{2,i,j} \overline{\alpha_{2,i',j}} \sum_{t=l_1+1}^{L} \left(c_i(t) - e_i(t)\right) \overline{\left(c_{i'}(t) - e_{i'}(t)\right)} \end{array} \right) \qquad \text{2-42}$$

$$d^2(c,e) = \sum_{j=1}^{N_r} \left(\Omega_{1j} A_1 \Omega_{1j}^* + \Omega_{1j} A_2 \Omega_{1j}^* \right) \qquad \text{2-43}$$

where,

$$A_{1,p,q} = \sum_{t=1}^{l_1} \left(c_p(t) - e_p(t)\right) \overline{\left(c_q(t) - e_q(t)\right)}$$

$$A_{2,p,q} = \sum_{t=l_1}^{L} \left(c_p(t) - e_p(t)\right) \overline{\left(c_q(t) - e_q(t)\right)}$$

$\Omega_{1j} = \left(\alpha_{1,1j}, \alpha_{1,2j}, \ldots, \alpha_{1,N_t j}\right)$ and

$\Omega_{2j} = \left(\alpha_{2,1j}, \alpha_{2,2j}, \ldots, \alpha_{2,N_t j}\right)$.

From the analysis in Section 2.2.2, the diversity order, is seen to be given by the sum of the ranks of matrices A_1 and A_2. Code construction criteria in Section 2.4.1, set up design rules for STTC based on the geometry of the code trellis, that ensure full rank for any given code. In the two transmit antenna case for instance, transitions from the same state are required to differ in the second symbol and transitions that arrive at the same state are required to differ in the first symbol. As the code sequence over a frame is required to start and end at the same state, the code difference matrix has the form $B(c,e) = \begin{bmatrix} 0 & * \cdots * & e_1(L) - c_1(L) \\ e_2(1) - c_2(1) & * \cdots * & 0 \end{bmatrix}$, which guarantees full rank.

However, consider matrix $A_1(c,e)$, which has square root $B_1(c,e)$ given by,

$$B_1(c,e) = \begin{bmatrix} 0 & * \cdots * & e_1(l_1) - c_1(l_1) \\ e_1(2) - c_1(2) & * \cdots * & e_2(l_1) - c_2(l_1) \end{bmatrix}$$

As the l_1^{th} node and not the L_1^{th} node is considered, an STTC that is designed the quasi-static scenario (where the channel coefficients are assumed to be constant over the length of the frame) might not have differ in the first symbol. Hence $B_1(c,e)$ and consequently $A_1(c,e)$ might not be full rank. Similarly as the code words do not start from the same node in the $(l_1+1)^{th}$ stage $A_2(c,e)$ might also not be full rank. Hence STTC codes designed for quasi-static channels might not be able to exploit the temporal diversity offered by slightly lower coherence times. A loss in spatial diversity might be seen due to the possible reduction in rank of the difference matrix between code-word block. For instance, if the sum of the ranks of $A_1(c,e)$ and $A_2(c,e)$ do not add up to the rank of $A(c,e)$, the spatial diversity offered by the code would be reduced. A loss in coding gain might also be observed if the product of the eigenvalues of $A_1(c,e)$ and $A_2(c,e)$ are less than the product of the eigenvalues of $A(c,e)$.

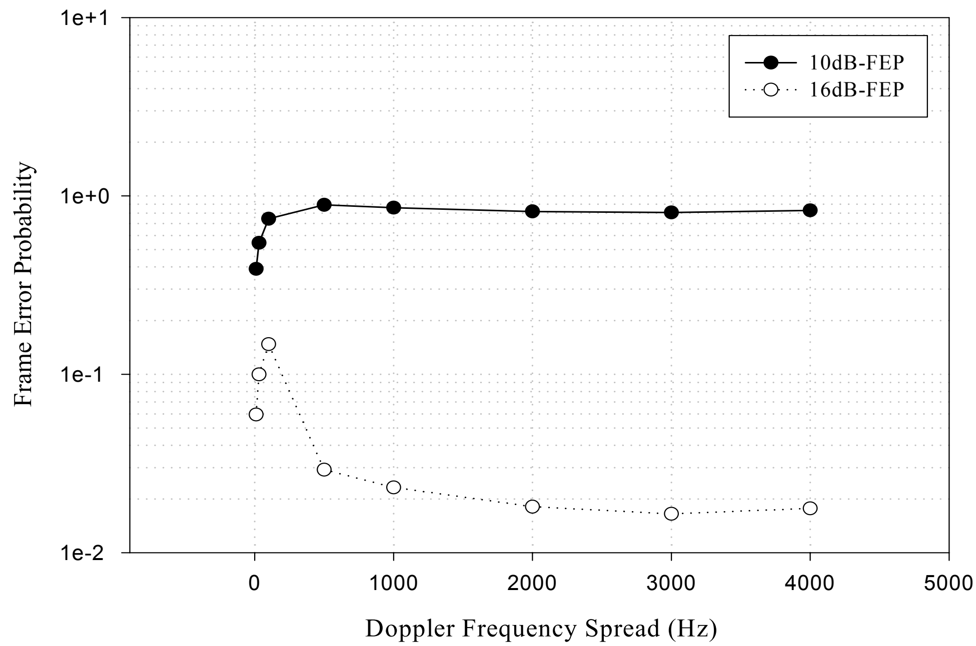

Figure 2-9: Performance of 2Tx-1Rx 8State-TSC over Doppler spreads

However, when the coherence times are very small, the performance of STTC is dominated by the number of time instances for which distinct code-words differ during the duration of a frame (Section 2.2.1.B). Hence higher gains are possible from the additional temporal diversity offered.

These results are illustrated in Figure 2-9, which shows the performance of 8-state STTC designed for quasi-static channel over different Doppler spreads of the channel (corresponding to different coherence times, Transmission frequency is assumed to be around 10,000Hz. 4000Hz Doppler spread roughly corresponds to the channel being constant over one symbol and 30Hz to the channel being constant over a frame.). The performance of the code at high SNRs is seen to initially degrade with decreasing coherence time, but to eventually improve. For small reductions in coherence time, the STTC is not able to exploit the additional temporal diversity, as shown in the analysis. The performance degrades due to a loss in coding gain. However, the performance begins to improve for very small coherence intervals (about 20 symbol intervals in this case) and represents the scenario when the loss in coding gain is compensated by the increased available temporal diversity.

2.5. Analytical Performance Analysis of STTC

Design criteria for STTC have been derived based on an upper bound on the Pair-wise Error Probability (PWEP) as shown in previous sections. The upper bound is derived from the Chernoff bound for conditional PWEP. However, this upper bound tends to be loose over the range of SNRs of interest and also over quasi-static channels in general. An exact measure of the exact PWEP is hence useful. An estimate of the bit error performance of a system can be calculated from the exact PWEP by taking in to account few dominant error lengths.

Cavers and Ho derived an exact expression for the PWEP of TCM transmitted over interleaved Rayleigh (i.e. fast fading) fading channels in [Cave1]. This derivation made use of the residue technique combined with characteristic function approach illustrated in [Proa1]. The estimate of the bit error probabilities calculated from the expression for the exact PWEP in [Cave1] is seen to be very accurate for high SNRs but not as accurate for low SNRs. This analysis was extended in [Uysa1] to include multiple transmit and receive antennas. A more generalized expression for the exact PWEP of STTC over Rayleigh fading channels was derived in [Turi1] and is discussed in this section. The performance of STTC over quasi-static, fast fading and spatially correlated channels are obtained as special cases of this expression.

2.5.1. Exact PWEP for STTC with Perfect Channel Estimates

An expression for the exact PWEP was derived in [Cave1] for STTC over Rayleigh fading channels and assuming perfect channel estimation. The PWEP derivation is through a residual method based on the characteristic function of the quadratic form of complex Gaussian random variables [Turi1]. The result in [Turi1] allows for the derivation of the characteristic function for the more general case unlike [Proa1], which is restricted to the scalar form.

From initial definitions, the received signal $r_j(t)$ at the j^{th} receive antenna at a time instant t, can be expressed as,

$$r_t^j = \sum_{i=1}^{N_t} \alpha_{ij}(t) c_i(t) + \eta_j(t); 1 \le j \le N_r \qquad \text{2-44}$$

Let,

$$c(l) = \left(c_1(l), c_2(l), ..., c_{N_t}(l)\right), X = diag\left(c(1), c(2), ..., c(L)\right)$$
$$\alpha_j(l) = \left(\alpha_{1j}(l), \alpha_{2j}(l), ..., \alpha_{N_t j}(l)\right)^T, \alpha_j = \left(\alpha_j(1), \alpha_j(2), ..., \alpha_j(L)\right)^T$$
$$r_j = \left(r_j(1), r_j(2), ..., r_j(L)\right)^T, r = \left(r_1, r_2, ..., r_{N_r}\right) \text{ and}$$
$$\eta_j = \left(\eta_j(1), \eta_j(2), ..., \eta_j(L)\right)^T, \eta = \left(\eta_1, \eta_2, ..., \eta_{N_r}\right)$$

The received signal can be written as,

$$r_j = X\alpha_j + \eta_j \qquad \text{2-45}$$

Assuming coherent detection, the decision metric to be minimized is,

$$\mu(r,x)=\sum_{j=1}^{N_r}\left\|r_j - X\alpha_j\right\|^2 \qquad \text{2-46}$$

The Pair-wise Error Probability (PWEP) $P(c\to e)$ represents the probability of incorrectly choosing sequence c, when in fact sequence e was transmitted.

Let, $e(l)=\left(e_1(l),e_2(l),...,e_{N_t}(l)\right), \widehat{X}=diag\left(e(1),e(2),...,e(L)\right)$ and let a random variable D be defined as

$$D=\sum_{j=1}^{N_r}\left(\left\|r_j-\widehat{X}\alpha_j\right\|^2-\left\|r_j-X\alpha_j\right\|^2\right) \qquad \text{2-47}$$

The PWEP is given by,

$$P(c\to e)=P\left(\mu\left(r,\widehat{X}\right)\le\mu(r,X)\right)=P(D\le 0) \qquad \text{2-48}$$

Let $\phi_D(s)$, be the Laplace transform of D. Then the probability density function of D is given by an inverse transform,

$$P(c\to e)=P(D\le 0)=-residue\left[e^{s\delta}\phi_D(s)/s\right]_{RP,\delta=0} \qquad \text{2-49}$$

Expression (2-47) can be expanded to,

$$D=\sum_{j=1}^{N_r}\alpha_j^H\left(X-\widehat{X}\right)^H\left(X-\widehat{X}\right)\alpha_j+\alpha_j^H\left(X-\widehat{X}\right)^H\eta_j+\left(\eta_j\right)^H\left(X-\widehat{X}\right)\alpha_j \qquad \text{2-50}$$

Defining, $y_j=\left[\left(X-\widehat{X}\right)\alpha_j \quad \eta_j\right]$ and $A=\begin{bmatrix} I_L & -I_L \\ -I_L & 0_L \end{bmatrix}$

$$D=\sum_{j=1}^{N_r}\left(y_j\right)^H Ay_j \qquad \text{2-51}$$

D is a summation over the quadratic form of complex variables and A is a Hermitian matrix by definition. By virtue of the results in [Turi1], the characteristic function of D can be written as,

$$\phi_D(s)=\prod_{j=1}^{N_r}\frac{1}{\det\left(I_{2L}+sC_{y_j}A\right)} \qquad \text{2-52}$$

C_{y_j} is the covariance matrix of y_j and is given by, $C_{y_j} = \begin{bmatrix} C_{\alpha_j} & 0_L \\ 0_L & N_0 I_L \end{bmatrix}$ where, C_{α_j} is the covariance matrix of $\left(X - \widehat{X}\right)\alpha_j$.

$$C_{\alpha_j}\left(m,q,l,k\right) = E\left(\sum_{i=1}^{N_t}\sum_{q=1}^{N_t}\alpha_{i,j}\left(l\right)\alpha_{q,j}\left(k\right)\left(c_i\left(l\right)-e_i\left(l\right)\right)\left(c_q\left(k\right)-e_q\left(k\right)\right)^*\right) \qquad \text{2-53}$$

The expression for C_{α_j} can be simplified according to the channel type. Four different channels are discussed below.

2.5.1. A. Fast Fading Channel

The fade coefficients are assumed to be constant over a symbol interval and vary from one symbol to the next. The fade coefficients across symbols and across transmit/receive antenna pairs are assumed to be independent.

$$C_{\alpha_j}\left(i,q,l,k\right) = \begin{Bmatrix} E_s\left(\sum_{i=1}^{N_t}\left|\left(c_i\left(l\right)-e_i\left(l\right)\right)\right|^2\right), l = k \; and \; i = q \\ 0, otherwise \end{Bmatrix} \qquad \text{2-54}$$

as $E\left(\alpha_{ij}\left(l\right)\alpha_{ij}^*\left(k\right)\right) = \delta\left(l-k\right)$ for fast fading channels. Consequently, C_{α^j} is a diagonal matrix, i.e. $C_{\alpha_j} = diag\left(\beta(1),\beta(2),...,\beta(L)\right) \; with \; \beta\left(l\right) = E_s\sum_{i=1}^{N_t}\left|e_i\left(l\right)-c_i\left(l\right)\right|^2$. In this case,

$$I_{2L} + sC_{y_j}A = \begin{bmatrix} \left(I_L + diag\left(\beta(1),\beta(2),...,\beta(L)\right)\right) & -s * diag\left(\beta(1),\beta(2),...,\beta(L)\right) \\ -sN_0 I_L & I_L \end{bmatrix} \qquad \text{2-55}$$

The characteristic function can be reformatted in a product form,

$$\phi_D\left(s\right) = \prod_{j=1}^{N_r}\prod_{l=1}^{L}\frac{1}{1+s\beta\left(l\right)-s^2\beta\left(l\right)} = \prod_{j=1}^{N_r}\prod_{l=1}^{L}\frac{1}{\beta\left(l\right)}\frac{1}{s^2 - s - \dfrac{1}{\beta\left(l\right)}} \qquad \text{2-56}$$

Expanding the denominator and exploiting independency across antennas,

$$\phi_D\left(s\right) = \left(\prod_{l=1}^{|\Omega|}\frac{E_s}{4N_o}\beta\left(l\right)\right)^{-N_r}\left(\prod_{l=1}^{|\Omega|}\frac{-1}{4\left(s-p_{1k}\right)\left(s-p_{2k}\right)}\right) \qquad \text{2-57}$$

$$\begin{bmatrix} p_{1k} \\ p_{2k} \end{bmatrix} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{1}{\frac{E_s}{N_o}\beta(l)}}$$

where $|\Omega|$ is the number of times that the transmitted and decoded sequences differ. Let the Effective Code Length (ECL) of a code be defined as the length of the shortest error event. At high SNRs the code performance is dominated by the probability of the shortest error event. Hence from expression (2-57), it can be concluded that the diversity order will be dictated by the product of the ECL of the code and the number of receive antennas being used in the system. It is also seen from Equation (2-57) that in fast fading channels, the number of transmit antennas contributes to the coding gain achieved and does not affect the diversity order. Hence in fast fading channels, the temporal diversity offered by the channel dominates performance rather than the spatial diversity offered by transmit antennas.

2.5.1. B. Quasi-static Case

The channel coefficients are assumed to be constant over a frame and to vary independently across from one frame to the next. They are also assumed to vary independently across transmit/receive antenna pairs. In this case, C_{α_j} is given by,

$$C_{\alpha_j}(i,q,l,k) = \begin{Bmatrix} E_s\left(\sum_{i=1}^{N_t}\left|\left(c_i(l) - e_i(l)\right)\right|^2\right), l = k \text{ and } i = q \\ E_s\sum_{i=1}^{N_t}\left(c_i(l) - e_i(l)\right)\left(c_i(k) - e_i(k)\right)^*, l \neq k \text{ and } i = q \end{Bmatrix} \tag{2-58}$$

C_{α_j} is not a diagonal matrix due to introduced temporal correlation between fading coefficients. Hence a derivation similar to the above cannot be done. The general formula is used instead.

$$\phi_D(s) = \left(\frac{1}{\det\left(I_{2L} + sC_{y_j}A\right)}\right)^{N_r} = \prod_{i=1}^{r}\left(\frac{1}{1 + s\lambda_i\left(\frac{E_s}{N_o}\right)}\right)^{N_r} \tag{2-59}$$

where r is the number of non-zero eigenvalues λ_i ,of $C_{y_j}A$. It is seen that the diversity order is determined by the product of r and the number of receive antennas in the system.

2.5.1. C. Spatially Correlated Fast-fading Scenario

The channel coefficients are assumed to vary between consecutive symbols. The transmit antennas are assumed to be correlated and the correlation between transmit antenna i and q is given by $\rho_{iq} = E(\alpha_{i,j}\alpha_{q,j}^{*})$. The receive antennas are still assumed to be independent.

$$C_{a_j}(l,k) = \begin{cases} E_s\left(\sum_{i=1}^{N_t}\left|\left(c_i(l)-e_i(l)\right)\right|^2\right) + E_s\sum_{\substack{i=1\\i\neq q}}^{N_t}\sum_{q=1}^{N_t}\rho_{iq}\left(c_i(l)-e_i(l)\right)\left(c_q(l)-c_q(l)\right)^{*} & l=k \\ 0 \quad \textit{otherwise} & \end{cases} \qquad \text{2-60}$$

It is seen that the cross-correlation terms tend to decrease the value of the diagonal elements of C_{α^j} and consequently the coding gain. But the presence of cross-correlation terms does not affect the diversity order of the system.

2.5.1. D. Spatially Correlated Quasi-static Scenario

The channel coefficients are assumed to be constant over a frame and to vary from one frame to the next. The transmit antennas are assumed to be correlated.

$$C_{\alpha_j}(i,q,l,k) = \begin{cases} E_s\left(\sum_{i=1}^{N_t}\left|\left(c_i(l)-e_i(l)\right)\right|^2\right) + E_s\sum_{\substack{i=1\\i\neq q}}^{N_t}\sum_{q=1}^{N_t}\rho_{iq}\left(c_i(l)-e_i(l)\right)\left(c_q(l)-e_q(l)\right)^{*} \quad l=k \\ E_s\sum_{i=1}^{N_t}\left(c_i(l)-e_i(l)\right)\left(c_i(k)-e_i(k)\right)^{*} + E_s\sum_{\substack{i=1\\i\neq q}}^{N_t}\sum_{q=1}^{N_t}\rho_{iq}\left(c_i(l)-e_i(l)\right)\left(c_q(k)-e_q(k)\right)^{*}, \\ \qquad l\neq k \textit{ and } i=q \end{cases} \qquad \text{2-61}$$

It is again seen that the presence of cross correlation terms does not affect the diversity order of the system.

2.5.1. E. Pair-wise Error Probability and Average Bit Error Probability

The probability density function of D is the inverse Laplace transform of $\phi_D(s)$.

$$P(c \to e) = P(D \le 0) = -residue\left[e^{s\delta}\phi_D(s)/s\right]_{RP,\delta=0} \qquad \text{2-62}$$

Due to the nature of the expression of PWEP, the calculation cannot be extended to give a closed form expression for the bit error probability. An estimate of the actual bit error probability, which is very useful in evaluating the performance of systems, can be obtained by averaging the PWEP over a limited number of dominant error events. The estimate of the bit error probability is given by the expression,

$$P_b = \frac{1}{b}\sum_{c \neq e} q\left(c \rightarrow e\right) P\left(c \rightarrow e\right) \tag{2-63}$$

where, b is the number of input bits per trellis transition and $q(c \rightarrow e)$ is the Hamming distance between the code word and the error event. The estimate is obtained by only considering error events up to a particular error event length v. This represents a truncation of the infinite series used in the evaluation of the Union Bound. The choice of v is very critical. It should be done such that all the dominant error events in a particular SNR range are included, while keeping a check on the computational complexity which exponentially increases with v. It is also noted, that expression (2-63), does not provide an upper bound on the bit error probability. The actual results could be lower or higher than the approximation.

2.5.2. Analytical Performance Results

The analytical performance of the 4-state TSC code described in the previous section is evaluated from expression (2-63) over error events of different lengths and different channel conditions. These are compared with corresponding performances obtained from simulations.

The length of the shortest possible error event length of a code is called the *Effective Code Length* (ECL) of the code. The TSC code has an ECL of two. In fast fading channels, performance of the STTC code is dictated by the *ECL* of the code. This is reflected in Figure 2-10 which shows that the simulated performance is well-approximated by the analytical evaluation for error events of length two. It is also seen that the analytical estimates for different error event lengths converge at asymptotically high SNRs. It is thus seen that an upper bound on the bit error probability can be obtained by taking into account error events of all lengths.

Figure 2-11 shows the performance of the code for quasi-static fading channels. The diversity of the code over quasi-static channel is decided by the minimum rank of the code (the TSC-RD criteria). It is seen from the plots that in the system being considered (2Tx-1Rx with TSC code), the fast fading channel does not give the code any added diversity advantage over the quasi-static fading channel. As noted earlier, in fast-fading channels, diversity performance of the code is overseen by the ECL of the code. But in quasi-static channels it is overseen by the minimum rank of the code. In the case of the TSC code, the minimum rank of the code and the ECL are both equal to 2. Hence the diversity advantage offered by the code is the same in both cases.

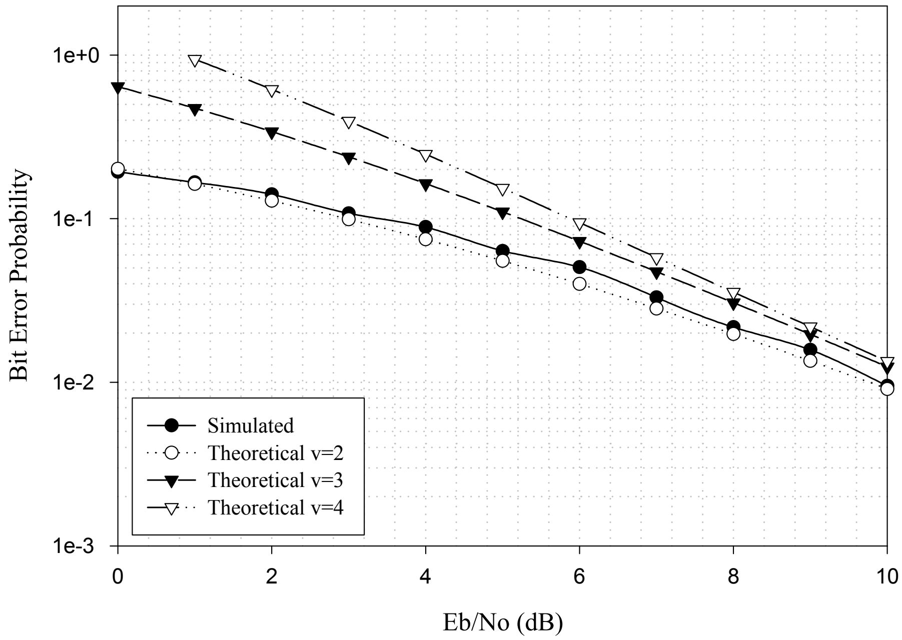

Figure 2-10: Analytical and simulated performance comparison of 2Tx-1Rx TSC code over fast fading channels

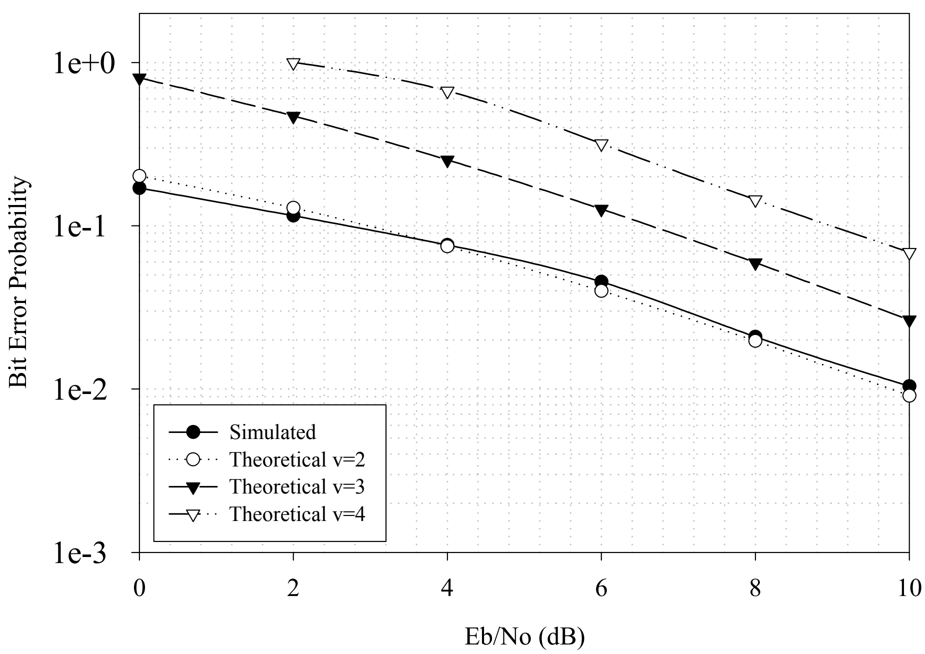

Figure 2-11: Analytical and simulated performance comparison of 2Tx-1Rx TSC code over quasi-static fading channels

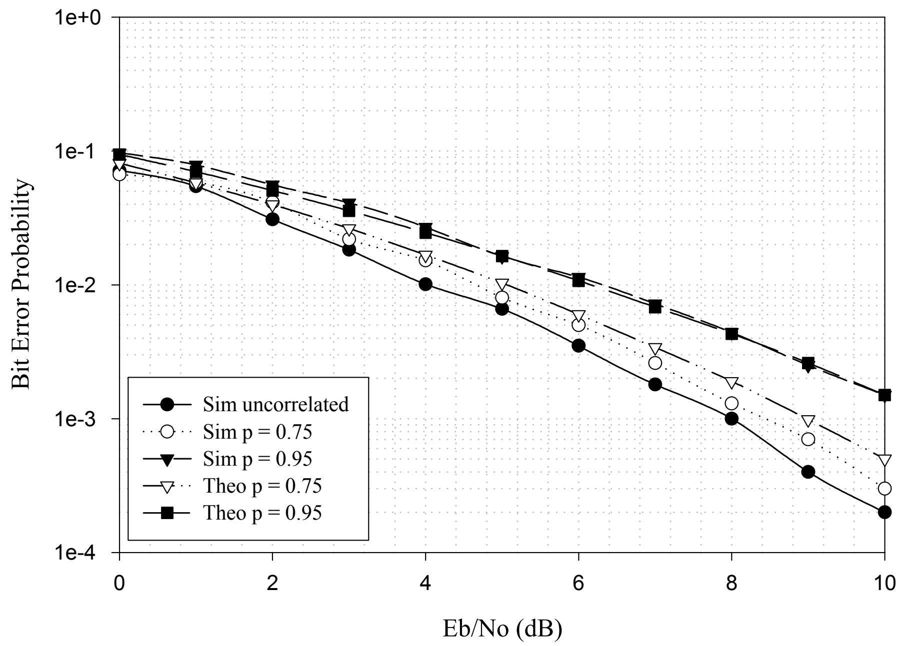

Figure 2-12: Analytical and simulated performance comparison of 2Tx-2Rx TSC code over spatially correlated fading channels

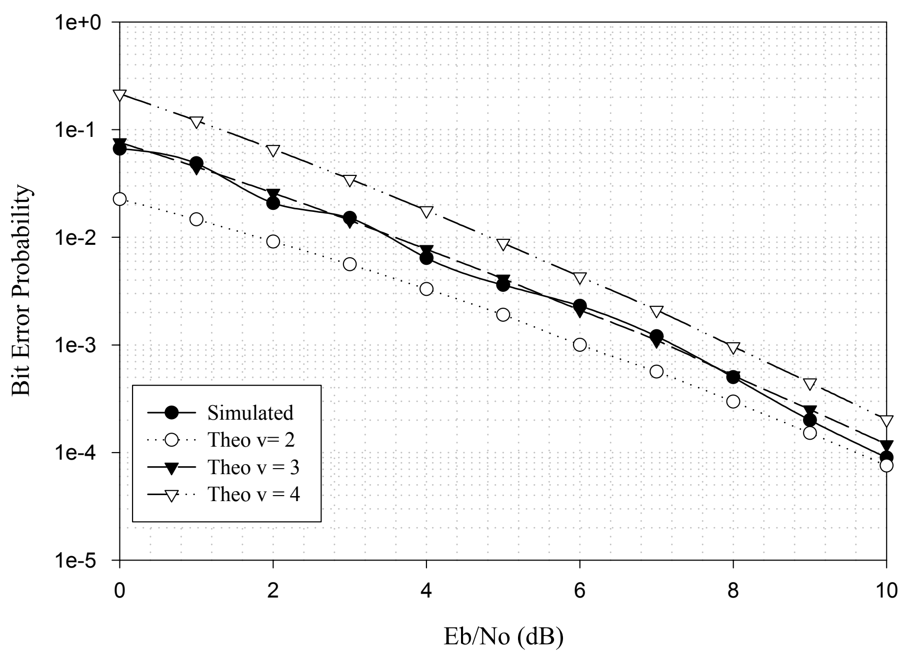

Figure 2-13: Analytical and simulated performance comparison of 2Tx-2Rx CYV code over quasi-static fading channels

The performance of the code is plotted over spatially correlated channel with different degrees of correlation in Figure 2-12. The analytical results are seen to provide good approximations of the simulated performance. As can be expected, a loss in performance is observed with an increase in the correlation between antennas.

The performance of the CVY code is analyzed in Figure 2-13. The predicted performance matches the simulated performance. It is also seen that the dominant error event length is 3 compared to 2 in the case of the TSC code. This is because the Effective Code Length of the CVY code is 3 while the ECL of the TSC code is 2.

2.6. Chapter Summary

This chapter discussed and compared the code design criteria of space-time codes over different channel conditions and assuming perfect channel estimates to be available at the receiver. Codes are analyzed in quasi-static and fast fading channels. Quasi-static channels assume that fade coefficients remain constant over a frame and vary from one frame to the next. Fast fading channels assume that fade coefficients vary over consecutive symbols in a frame (thus offering additional time diversity compared to quasi-static channels). Both represent extreme scenarios and the actual channel response is expected to be in between. Code design criteria should thus ideally intend to construct codes that can achieve optimal performance (derive maximum benefit) in these two extreme channel conditions.

Design criteria are derived by analyzing expressions for the upper bound on pair-wise error probability of STTC based on the Chernoff bound. When the diversity order of a system is small, the TSC-RD criteria (rank criterion) proposed in [Taro1], which recommends maximizing the minimum rank of the distance matrix between any two code-words to maximize the diversity gain of a scheme, is the most appropriate. In contrast, improving the coding gain is more important for systems with high diversity orders. The trace criterion which maximizes the Euclidean distance between code-words was proposed in [Chen1] for large diversity systems (diversity >3). The trace criterion betters the performance of the TSC-RD criteria in systems with a large diversity order. It also does better in fast fading channels irrespective of the diversity order of the system. This can be expected as fast fading essentially translates to large diversities. The trace criterion is thus found to be more useful than the TSC-RD criteria except for low diversity order systems operating over quasi-static channels.

This chapter also presented an analysis of the exact pair-wise error probability of STTC. Estimates of the bit error rate of STTC systems operating over different channel conditions are obtained from the exact PWEP expression. This analytical performance analysis is useful in corroborating the simulated performance of the system and also provides additional insights into the performance of the code.

3. Channel Estimation Techniques for Multiple Transmit Antenna Systems

Multiple Input Multiple Output (MIMO) systems help in increasing the capacity offered by a channel. But a MIMO channel also results in increased number of channel parameters to be estimated. This makes the scheme especially sensitive to channel estimation errors and the choice of an appropriate channel estimation technique is of particular importance. Also, as the signal at the receiver in a multiple transmit antenna system is a superposition of signals transmitted from the multiple antennas, single transmit antenna channel estimation schemes cannot be directly used and specialized schemes need to be designed. This chapter introduces several channel estimation techniques for multiple antenna systems found in the literature and discusses their associated tradeoffs and suitability. It also presents capacity analysis of the multi-input channel when training is used to perform channel estimation.

3.1. Channel Estimation using Training Sequences

For quasi-static or slowly varying channels, the most obvious and popular choice of channel estimation is by using training sequences. Also, current wireless packet communication systems provide for a training sequence to be inserted in each packet to aid in channel estimation at the receiver end. But as mentioned, multiple antenna systems impose additional constraints and properties on the training sequences as opposed to single antenna systems. These challenges in the design of training sequences are analyzed and discussed in this section.

3.1.1. Training Model and Training Sequences

3.1.1. A Flat fading Channel

A system with N_t transmit and N_r receive antennas is considered. The channel is assumed to undergo flat fading. It is constant over a frame of length L and varies from one frame to the next. A known sequence $s_i(t)$ of length L_t is transmitted from the i^{th} transmit antenna at the beginning of each frame to estimate the channel. The multiple antenna channel model is given by,

$$r_j = S\alpha_j + \eta_j;\ j = 1,2,...,N_t \qquad \text{3-1}$$

where,

$r_j = \left[r_j(1) \quad \dots \quad r_j(L_t) \right]^T$ is the received vector,

$S = \begin{bmatrix} s_1(1) & \cdots & s_{N_t}(1) \\ \vdots & \ddots & \vdots \\ s_1(L_t) & \cdots & s_{N_t}(L_t) \end{bmatrix}$ is the transmitted training sequence matrix,

$\eta_j = \begin{bmatrix} \eta_j(1) & \ldots & \eta_j(L_t) \end{bmatrix}^T$ is the noise vector and

$\alpha_j = \begin{bmatrix} \alpha_{1j} & \ldots & \alpha_{N_t j} \end{bmatrix}^T$ are the channel coefficients.

Sequences used for estimating single-input flat fading channels require impulse like auto-correlation properties and have been widely investigated in the literature. But in the multi-input channel scenario, the receiver observes the super-position of training sequences transmitted from different transmit antennas. This necessitates the additional requirement of very low (ideally zero) cross-correlation between the sequences transmitted from different antennas. This result is reiterated by the following analysis.
Let $\widehat{\alpha}_j$ denote linear least square channel estimates. These are calculated by,

$$\widehat{\alpha}_j = \left(S^H S\right)^{-1} S^H r_j \tag{3-2}$$

Mean square channel estimation error is given by,

$$MSE = E\left[\left(\alpha_j - \widehat{\alpha}_j\right)^H \left(\alpha_j - \widehat{\alpha}_j\right)\right] = N_o trace\left(S^H S\right)^{-1} \tag{3-3}$$

Minimum mean square error is given by,

$$MMSE = \frac{N_o}{L_t} \tag{3-4}$$

This is achieved if and only if,

$$S^H S = L_t I \tag{3-5}$$

It follows that optimal sequences which guarantee the minimum possible mean square error have impulse like auto-correlation properties and zero cross correlation.

A simple solution for achieving zero cross correlation is to transmit training symbols only from one antenna at a time. But this leads to a large loss in bandwidth efficiency and high peak-to-average power ratio and is not preferred. Training sequences that satisfy the above properties can be classified into two categories. The first approach presented in [Spas1] constructed optimal sequences from an N^{th} root-of-unity alphabet, $A_N = \left\{ \exp\left(\frac{j2\pi k}{N}\right); k = 1, 2, ..., N \right\}$, without constraining the alphabet size N. For any training sequence length, there exist optimal training sequences that belong to an N^{th}

root-of-unity alphabet. But these Perfect Root of Unity Sequences (PRUS) might not belong to a standard constellation. The second approach constrains the training sequences to belong to a specific constellation (e.g. BPSK, QPSK) which aids in simple transmitter/receiver implementation. In this case, optimal sequences do not exist for all training lengths. Instead, sub-optimal sequences are identified by an exhaustive search in a space of size $C^{N_t L_t}$, where L_t is the length of the training sequence and C is the alphabet size. The search space could be reduced by restricting the training sequence scheme to a particular constellation (when the data constellation is different). But this leads to an increase in the minimum achievable mean square error. The authors in [Frag1] introduced "L-Perfect" sequences. A single "L-Perfect" sequence could be used instead of multiple training sequences in a multiple transmit antenna system. An alternative method to exhaustive search is also identified that could be used when optimal or "L-Perfect" training sequences do not exist for a particular constellation. Performance bounds show that these sequences achieve performances close to those produced by optimal sequences. In [Frag2], the problem of identifying multiple training sequences is reduced to that of identifying a single sequence with impulse like autocorrelation. This consequently reduces the search space to C^{L_t}. This scheme is discussed in the next section.

The variance of the channel estimation error assuming that orthogonal sequences are transmitted can be determined in the following way. Expression (3-2) can be written as,

$$\hat{\alpha}_j = \alpha_j + \left(S^H S\right)^{-1} S^H \eta_j \tag{3-6}$$

The channel estimation error is seen to be given by $\left(S^H S\right)^{-1} S^H \eta_j$ which has zero mean. The variance of the estimation error for each channel coefficient, for a given training length, L_t, is given by $\frac{N_o}{2L_t E_s}$, where E_s is the symbol energy.

Figure 3-1, shows the performance of a training based channel estimation system for a two-transmit, one receive antenna MIMO system. The variance of the channel estimation error is plotted for varying training sequence lengths over a range of channel SNRs. It is seen that the variance of the channel estimation errors in the simulated system closely match the theoretical measure (calculated as $\frac{N_o}{2L_t E_s}$). It can also be observed from the figure that the variance of channel estimation error decreases with increasing training length and channel SNR. But at high SNRs, the performances are comparable and using longer training sequences does not necessarily provide an advantage.

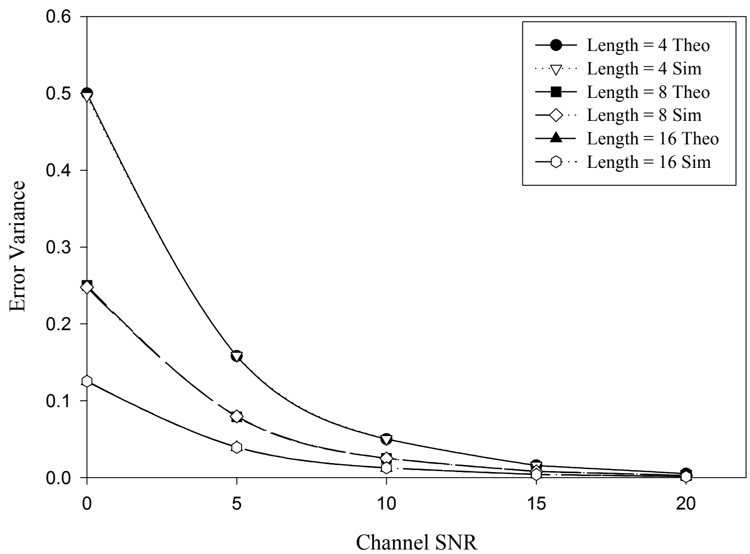

Figure 3-1: Channel estimation error variance for different training lengths and channel SNRs

3.1.1. B Training in Frequency Selective Channels

Training sequences to be used in frequency selective channel have the additional requirement that the cross correlation between the N_t training sequences be very low (ideally zero) over time lags equal to the channel memory in addition to good auto-correlation properties. This requirement is illustrated in the following analysis [Frag3].

A quasi-static frequency selective channel with L_c multi-paths is considered. The multi-paths are assumed to arrive at multiples of the symbol time. L_t is assumed to be the length of the training sequence and $L_t + L_c - 1$ is the length of the training interval.
The received signal can be represented as,

$$r_j(t) = \sum_{l=0}^{Lc} \sum_{i=1}^{N_t} s_i(t-l) a_{i,j}(l) + \eta_j(t),\ j = 1,2,...,N_r,\ t = 1,2,...,L_t + Lc - 1 \qquad \text{3-7}$$

Further defining,

$$R_j = \begin{bmatrix} r_j(1) & r_j(2) & \dots & r_j(L_t) \end{bmatrix}^T,$$
$$S = \begin{bmatrix} S_1(0) & S_2(0) & \dots & S_{N_t}(0) & \dots & S_1(Lc-1) & S_2(Lc-1) & \dots & S_{N_t}(Lc-1) \end{bmatrix},$$
$$S_i(l) = \begin{bmatrix} 0_{l*1} & s_i(1) & s_i(2) & \dots & s_i(L_t - l) \end{bmatrix}^T,$$
$$a_j = \begin{bmatrix} a_{1,j}(0) & a_{2,j}(0) & \dots & a_{N_t,j}(0) & \dots & a_{1,j}(L_c-1) & a_{2,j}(L_c-1) & \dots & a_{N_t,j}(L_c-1) \end{bmatrix}^T,$$

and $\eta_j = \left[\eta_j(1) \quad \eta_j(2) \quad \dots \quad \eta_j(L_t)\right]$, the received signal can be expressed in matrix notation as,

$$R_j = Sa_j + \eta_j \tag{3-8}$$

Linear least square estimates, assuming S has full column rank, are given by,

$$\widehat{\alpha}_j = \left(S^H S\right)^{-1} S^H R_j \tag{3-9}$$

The channel estimation Mean Square Error is given by,

$$MSE = E\left[\left(\alpha_j - \widehat{\alpha}_j\right)^H \left(\alpha_j - \widehat{\alpha}_j\right)\right] = 2N_o trace\left(\left(S^H S\right)^{-1}\right) \tag{3-10}$$

The Minimum MSE is equal to

$$MMSE = 2N_o \frac{L_c}{L_t} \tag{3-11}$$

which is achieved only if

$$S^H S = L_t I \tag{3-12}$$

The training sequences for frequency selective channels are thus seen to require very low cross-correlation over time lags equal to the delay spread of the channel. PRUS are optimal over frequency selective channels as well, as they have zero cross correlation over time lags equal to the delay spread of the channel. Suitable sub-optimal sequences are identified by exhaustive searches. The variance of channel estimation error is similar to that calculated for the flat fading case as training sequences are designed to have low cross-correlation over time lags corresponding to multi-path delays.

3.1.2. Complexity Reduction Techniques for Training Sequences

The structure of space-time codes can be exploited to reduce the problem of identifying multiple training sequences to that of finding a single training sequence. A space-time encoder used in conjunction with a single training sequence can be used to generate N_t sequences that though not always independent (depending on the space-time encoder used) are constrained in their mutual properties by the encoder.

3.1.2. A Block Code for Training Symbols

A sequence s with impulse like auto correlation properties is encoded by a space time block code to produce two sequences s_1 and s_2 which have zero cross-correlation properties. The training sequence for N_t transmit antennas can be encoded by using any orthogonal matrix U of dimension $N_t \times N_t$ such that $U^H U = N_t I_n$, where I_n is an identity matrix of dimension N_t. The received signal at the j^{th} receive antenna can be written as,

$$r_j = sU\alpha_j + \eta_j \tag{3-13}$$

The N_t training sequences are given by the N_t rows of the signal matrix sU. The linear least square channel estimates are given by,

$$\widehat{\alpha}_j = \frac{(sU)^H (sU)}{\left((sU)^H (sU)\right)}\widehat{\alpha}_j + \frac{(sU)^H}{\left((sU)^H (sU)\right)}\eta_j \tag{3-14}$$

$$(sU)^H (sU) = s^H s = L_t \tag{3-15}$$

Let

$$\widehat{\eta}_j = \frac{(sU)^H}{\left((sU)^H (sU)\right)}\eta_j \tag{3-16}$$

The mean square channel estimation error is given by,

$$E\left(\widehat{\eta}_j^H \widehat{\eta}_j\right) = \frac{N_o}{L_t} \tag{3-17}$$

This agrees with the lower bound for the MSE found earlier. Hence the N_t training sequences formed by passing a single optimal sequence through a space-time encoder are also optimal (have ideal correlation properties).

Figure 3-2 compares the BER performance of an 8-state 8-PSK trellis code in an EDGE TU environment ($L_t = 26; L_c = 4$) with an optimal PRUS and the proposed scheme with a sub-optimal sequence [Frag2]. The BER results are for active states, M=16 and M=32. It is seen that the optimal and sub-optimal schemes achieve similar performance.

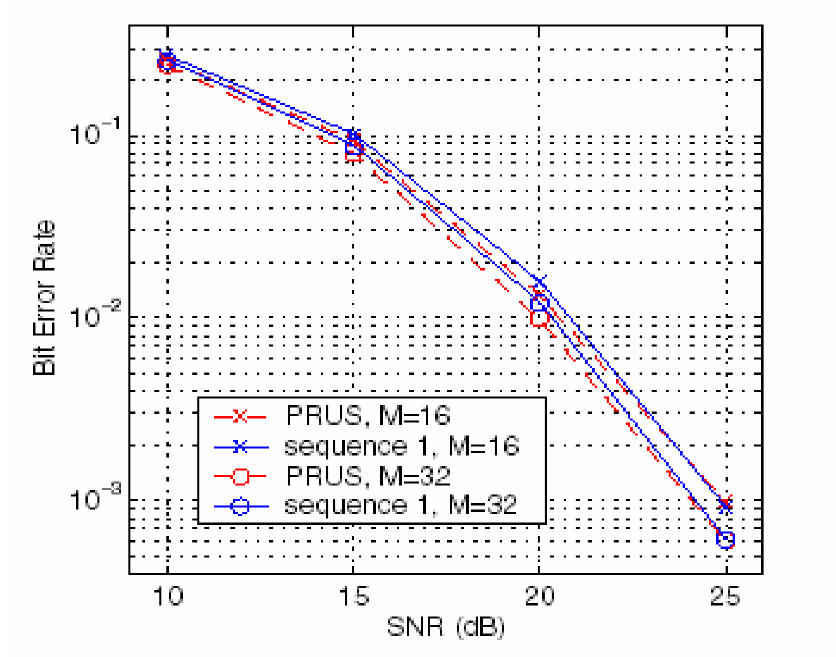

Figure 3-2: [Frag2] Performance of 8-state 8-PSK STC in a 2Tx-1Rx system with optimal and sub-optimal training sequences of length=13.

3.1.2. B. STTC Encoded Training Sequence

Consider a multi-input system with two transmit antennas. Some specific space-time trellis codes used in conjunction with this system can transform it into an equivalent single input system with a data–dependent channel response. For example, the code shown in Figure 3-3 results in a response with a D-transform given by,

$$\alpha_{eqvl}^{STTC}(k,D)=\alpha_1(D)+p_k\alpha_2(D) \quad \text{3-18}$$

where $p_k=\pm 1$ is data-dependent. Hence the input sequence determines the equivalent channel. By transmitting only even symbols from the constellation, the equivalent channel response is given by

$$\alpha_e(D)=\alpha_1(D)+D\alpha_2(D) \quad \text{3-19}$$

By transmitting only odd symbols the equivalent channel is given by,

$$\alpha_o(D)=\alpha_1(D)-D\alpha_2(D) \quad \text{3-20}$$

The channel can be calculated as

$$\alpha_1(D)=\frac{\alpha_e(D)+\alpha_o(D)}{2} \quad \text{3-21}$$

and

$$\alpha_2(D) = \frac{\alpha_e(D) - \alpha_o(D)}{2} \qquad 3\text{-}22$$

Hence only a single training sequence is required to be found instead of N_t training sequences. The search space is restricted to C^{L_t} from $C^{N_t L_t}$.

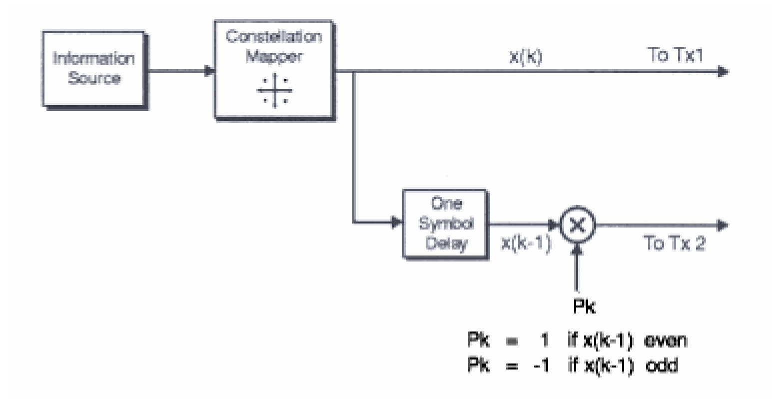

Figure 3-3: [Frag3] Encoder for the 8-state 8-PSK STTC with 2 transmit antennas

3.2. Information Theoretic Results for the Amount of Training

Assuming perfect channel estimates, an increase in the number of antennas causes an increase in system capacity. But when training is used to estimate the channel, an increase in the number of antennas translates to an increase in the number of channel parameters to be estimated and also a reduction in the transmitted power (as transmission power is divided between N_t transmit antennas). This results in an increase in the number of training symbols which could obviate any advantage in spectral efficiency offered by using STTC.

An information theoretic approach is used in [Hass1] to find the amount of training that could provide maximum benefit to multiple antenna systems. The analysis identifies optimal system configurations (for e.g. the number of transmit antennas, the training sequence transmitted power).

3.2.1. Flat Fading Scenario

Foshini and Gans in [Fosc1] have shown that extraordinary capacity is available in MIMO systems with high probability. Compared to the $N_t = 1$ case, which scales by one bit/hertz for every 3dB of SNR increase, the multi-antenna case scales by N_t bits for every 3dB of SNR increase. But these capacity predictions are based on the assumption of perfect channel estimation which is almost impossible to achieve in practical wireless systems. Hence a study of the effect of channel estimation errors on capacity is of import.

Hassibi and Hochwald in [Hass1] have shown how channel estimation using training sequences affects the capacity of fading channels. A lower bound on the information theoretic capacity achievable with training based schemes is computed and is maximized as a function of signal-to-noise ratio (ρ), fading coherence time (T) and the number of transmitter antennas (N_t). Consider that the training symbols of length T_t are transmitted with an SNR of ρ_t . Let ρ_d be the SNR of the data symbols and let T_d be the length of the data symbols. The total transmitted energy of the system can be expressed by the relation,

$$\rho T = \rho_d T_d + \rho_t T_t \tag{3-23}$$

$$T = T_d + T_t \tag{3-24}$$

The capacity lower bound can be written as

$$C_T \geq E\left(\frac{T - T_t}{T} \log\det\left(I_{N_r} + \frac{\rho_d \sigma_{\hat{H}}^2}{1 + \rho_d \sigma_{\tilde{H},R_s}^2} \frac{R_V^{-1} \bar{H}^* R_S \bar{H}}{N_t} \right) \right) \tag{3-25}$$

where,

$\bar{H}$ is the normalized channel estimate matrix,
$\sigma_{\hat{H}}^2$ is the variance of the channel estimates.
$R_S = E\left(S^* S\right)$, S is the transmitted training sequence as mentioned earlier.
$\sigma_{\tilde{H},R_s}^2 = \frac{1}{N_t N_r} E\left(tr\left(\tilde{H}^* R_S \tilde{H} \right) \right)$, $\tilde{H}$ is the channel estimate error matrix.

This bound is maximized with respect to the training data (S), training SNR (ρ_t) and training length (T_t). It is shown that the optimal solution for the choice of training sequence is given by, $S^H S = T_t I$. Hence the sequences must have zero cross-correlation. This agrees with results in the previous sections.

The optimal power allocation $\alpha = \frac{\rho_d T_d}{\rho T}$ is given by,

$$\alpha = \begin{cases} \gamma - \sqrt{\gamma(\gamma - 1)} \;\; for\, T_d > N_t \\ \frac{1}{2} \;\; for\, T_d = N_t \\ \gamma + \sqrt{\gamma(\gamma - 1)} \;\; for\, T_d < N_t \end{cases} \tag{3-26}$$

where, $\gamma = \dfrac{N_t + \rho T}{\rho T\left(1 + \dfrac{N_t}{T_d}\right)}$

The optimum length of the training sequence is the minimum possible, $T_t = N_t$ provided that the training and data powers are allowed to vary. If the training and data powers are required to be the same, as is usually the case in practical wireless systems, the optimum training length might be larger than N_t. The optimum value can be calculated by evaluating the capacity lower bound given below either analytically or via Monte Carlo simulations for the channel estimate matrix.

$$C_T \geq E\left(\frac{T - T_t}{T}\log\det\left(I_{N_t} + \frac{\rho^2 T_t / N_t}{1 + \left(1 + T_t / N_t\right)\rho}\frac{\bar{H}\bar{H}^*}{N_t}\right)\right) \qquad \text{3-27}$$

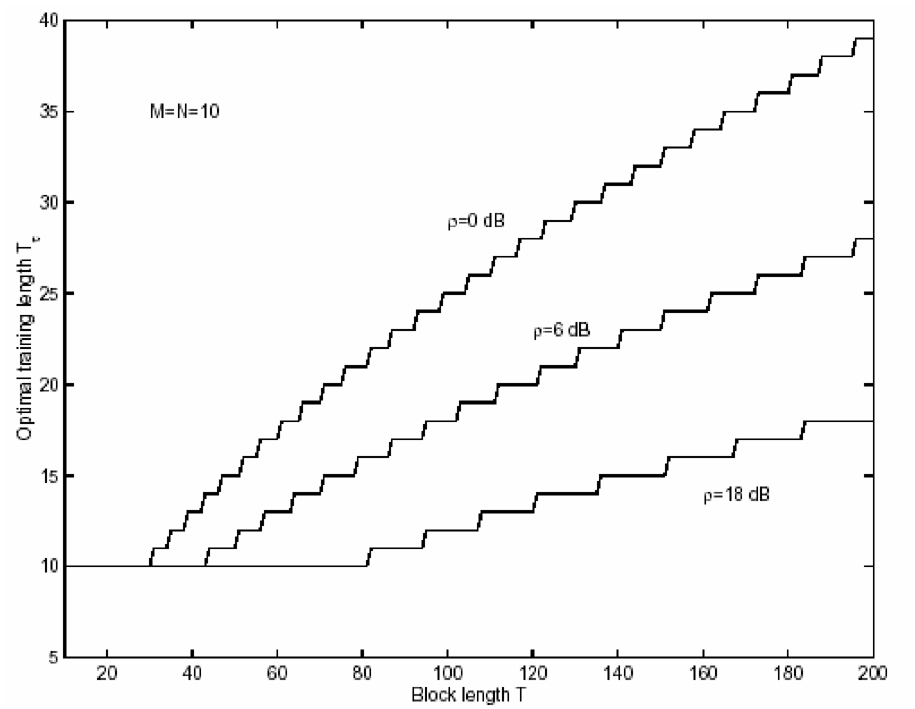

Figure 3-4: [Hass1] Optimal amount of training as a function of channel coherence time (N_t = N_r = *10*)

Figure 3-4 shows the optimal amount of training T_t as a function of block length T (equivalent to the coherence time of the channel) for different SNRs, for $N_t = N_r = 10$ and constraining the data and training powers to be equal. It can be seen that when SNR decreases the amount of training increases. At low SNRs the length might be as much as half the coherence time (half the frame length).

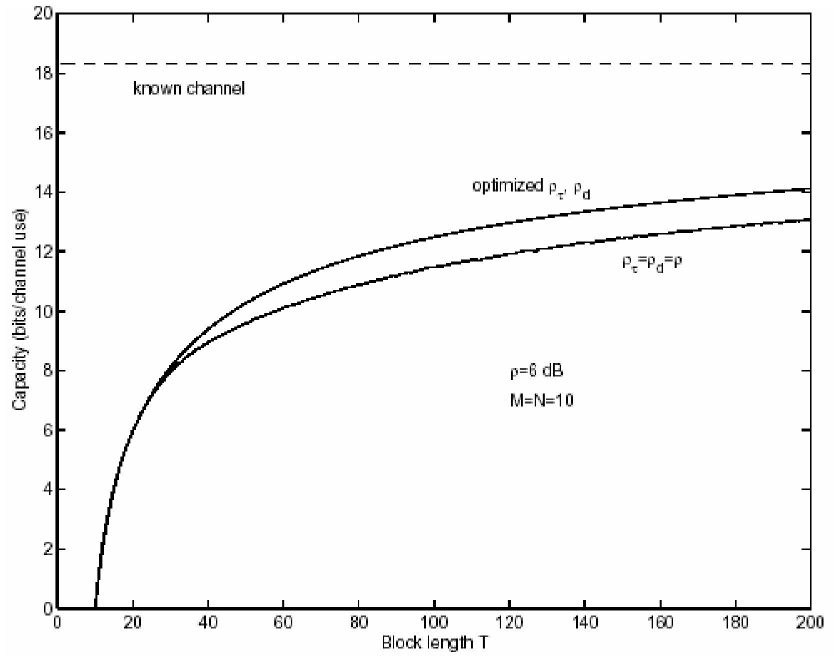

Figure 3-5:[Hass1]Capacity of training based system as a function of coherence time ($N_t = N_r = 10$)

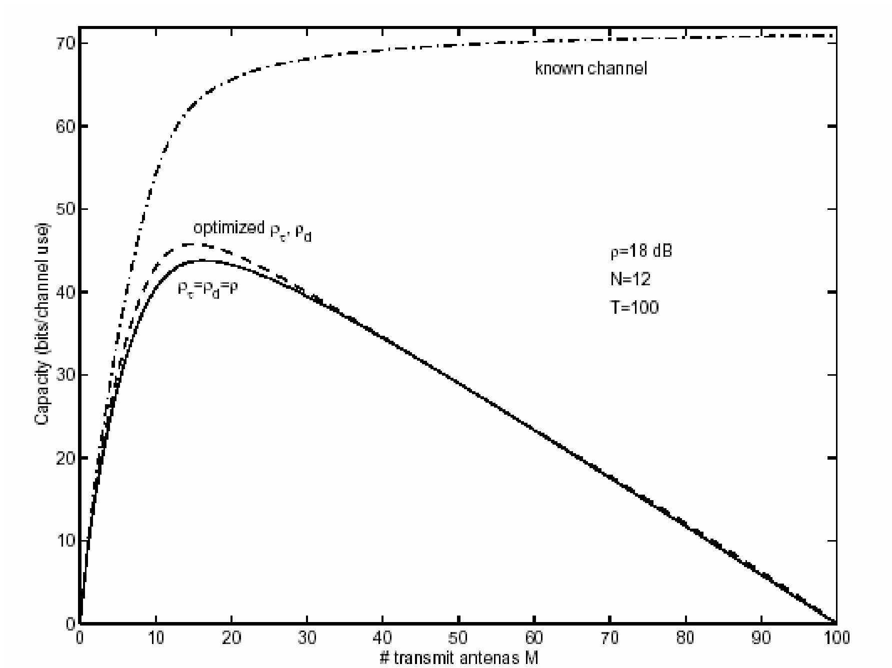

Figure 3-6: [Hass1]Capacity as a function of number of transmit antennas for $N_r = 12$ coherence time of 100 symbols

Figure 3-5 shows the lower bound on capacity as a function of T and when $N_t = N_r = 10$. It is seen that higher capacity is achievable if the transmit and data powers are allowed to vary independently. Figure 3-6 and Figure 3-7 show capacity as a function of transmit antennas. The number of receive antennas is 12. The solid line is optimized over T_t for equal training and data powers. The dashed line is optimized over the power allocation and $T_t = N_t$. It is seen that when $T = 100$, the capacity curve peaks for $N_t \approx 15$ and when $T = 20$ for $N_t \approx 7$. It is thus shown that the number of transmit antennas that maximizes capacity is often relatively small and choosing the wrong number of antennas can reduce the maximum achievable data rates.

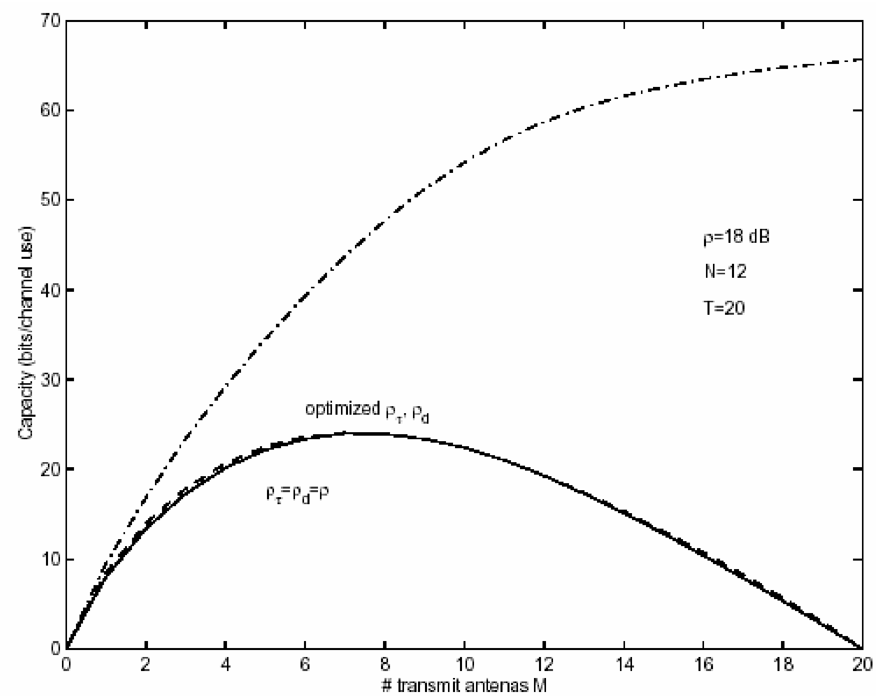

Figure 3-7: [Hass1] Capacity as a function of number of transmit antennas for N_r = 12 and coherence time of 20 symbols

3.2.2. Frequency Selective Scenario

The following analysis attempts to find the optimal number of transmit antennas to be used in a frequency selective channel. Consider a frequency selective channel with L_c multi-paths. To obtain a meaningful estimate of the channel,

$$\begin{aligned} T_t N_r &\geq N_t L_c N_r \\ T_t &\geq N_t L_c \end{aligned} \tag{3-28}$$

By using the concept of virtual antennas, the capacity of a MIMO system with N_t-transmit antennas, N_r receive antennas and training is given by

$$C_T \geq E\left(\frac{T-T_t}{T}\log\det\left(I_{N_v} + \frac{\rho^2 T_t / N_t}{1+\left(1+T_t/N_t\right)\rho}\frac{\overline{H}\overline{H}^*}{N_t}\right)\right) \quad 3\text{-}29$$

where, $N_v = N_t\ L_c$ is the number of virtual antennas. The transmit power for the data symbols and the training symbols is assumed to be the same ($\rho_d = \rho_t$). The optimal value for T_t is obtained by evaluating the capacity bound via Monte Carlo simulations. The capacity for varying number of transmit-antennas is plotted in Figure 3-8. The plot is optimized for the length of the training sequence.

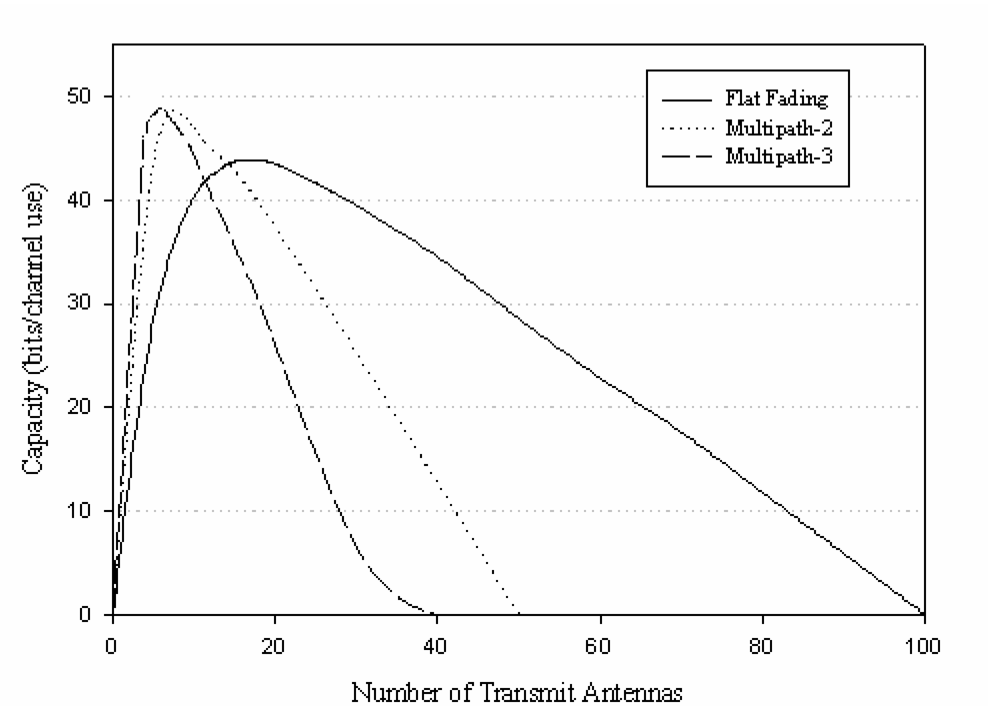

Figure 3-8: Capacity as a function of number of transmit antennas for $N_r = 12$ and channel coherence time of 100 symbols in the presence of multi-path

As can be expected, the optimal number of antennas is observed to decrease with an increase in the number of taps of the frequency selective channel. For the case of a two-tap channel, the optimal number of transmit antennas is nine and for the three-tap case, it is five, both less than number of antennas for the flat fading case. Reduction in the optimal number of antennas translates into a reduction in the maximum achievable capacity of the channel. Hence it is seen that in the presence of multi-path and training, the maximum achievable capacity of a channel is further reduced.

3.3. Blind and Semi-Blind Techniques

3.3.1. Iterative Channel Estimation

Usually channel estimates are obtained by using training sequences. The estimates are assumed to perfect and are used to decode the transmitted information. This can result in a significant loss in performance. In the techniques presented here, the redundancy offered by a space time code is exploited to enhance the estimates obtained by employing training sequences. A small amount of training is initially used to obtain preliminary channel estimates. These estimates are then utilized to decode the received signal and yield data estimates. Coding at the transmitter end makes these estimates sufficiently robust and these are used as uncertain training sequences for improved channel estimates. This process can be continued iteratively until some criterion is met. The performance of the system is found to be very close to that obtained by using perfect channel estimates.

The iterative receiver can be implemented in various ways. Figure 3-9 shows the structure of the receiver. $r_j(t)$ is the received signal at receive antenna j at time t. $\tilde{\alpha}_{ij}(t)$ and $\tilde{c}_i(t)$ are the channel and transmitted data estimates respectively.

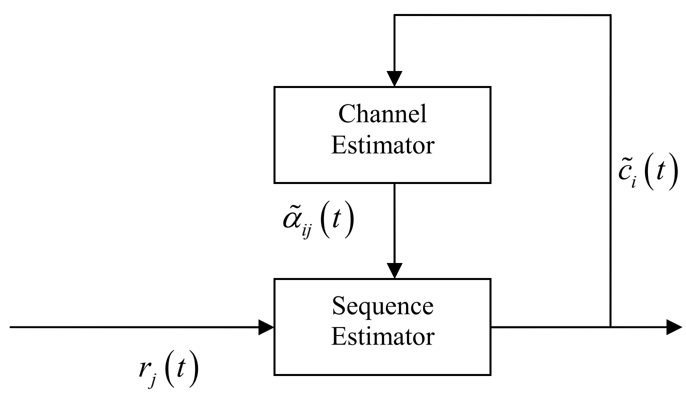

Figure 3-9 : Model for an Iterative channel estimator

The receiver uses a small amount of training to form the initial channel estimates. These initial estimates are then used to decode the received sequence using the Viterbi decoder. The resultant data estimates are then used to compute new channel estimates according to the expression[Gran1],

$$\tilde{\alpha}_{ij}(t) = \tilde{c}_i^*(t)\left(r_j(t) - \sum_{m \neq i} \tilde{\alpha}_{mj}(t)\tilde{c}_m(t)\right) \qquad 3\text{-}30$$

This may be done in parallel or serially. If the channel estimates are constant over a time period L, then the channel estimate $\tilde{\alpha}_{ij}$ over time period L is given by

$$\tilde{\alpha}_{ij} = \frac{\sum_{t=1}^{L} \tilde{\alpha}_{ij}(t)}{\|\tilde{c}_i\|} \qquad \text{3-31}$$

A linear Minimum Mean Square estimate [Ranh1], given below, may also be used to compute new channel estimates.

$$\tilde{\alpha}_{ij}(t) = \left(\hat{c}_i^H(t)\hat{c}_i(t)\right)^{-1} \hat{c}_i^H(t) r_j(t) \qquad \text{3-32}$$

Figure 3-10 and Figure 3-11 compare the performance of the TSC scheme with and without perfect channel estimates. The system has two transmit and two receive antennas. For systems with two transmit antennas, a training sequence of length of at least eight is required to obtain the transmit diversity advantage of two over medium SNR range. This is illustrated in the figures, where, the performance slopes are similar for the two cases reflecting an equivalent diversity advantage. It is also seen that the case with the training sequence length of eight performs within 1dB of the case with perfect channel estimates.

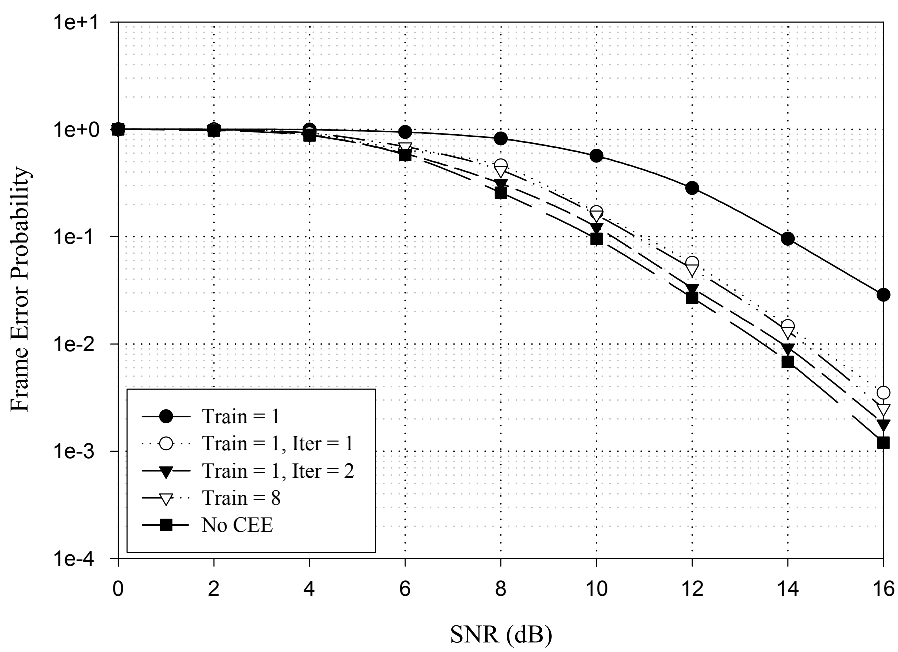

Figure 3-10: Performance of 2Tx-2Rx TSC STTC with iterative training

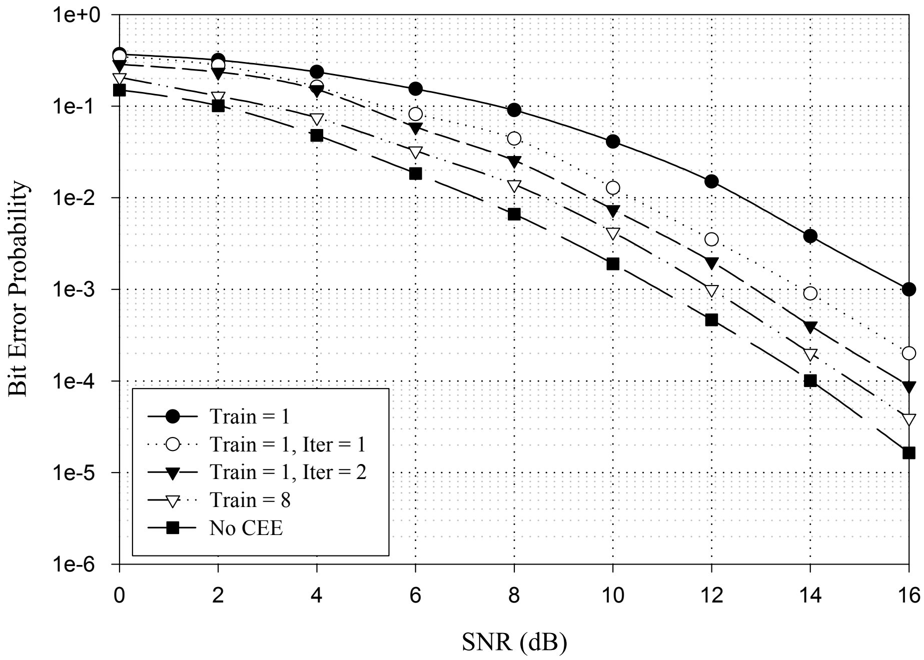

Figure 3-11: BER performance of 2Tx-2Rx TSC STTC with iterative training

A training sequence of length one gives an unacceptable performance degradation of more than 4dB (compared to the 3dB performance degradation incurred when a differential scheme is used and channel estimates are not required.).

The performance using the iterative scheme with the initial channel estimates given by a training sequence of length one is also plotted in the figures. It is seen from the Frame Error Rate curves in Figure 3-10, that a performance comparable to that of the case using eight training symbols is generated in the first iteration. The second iteration leads to improved results and the performance is within 0.5dB of the case with perfect channel estimates. Convergence is weaker in the Bit Error Rate curves as frames in error will iteratively lead to poorer channel estimates and consequently a large number of bit errors.

3.3.2. HMM Based Blind Channel Estimation

Channel estimation using training sequences leads to considerable loss in bandwidth and power efficiency of space-time codes especially under low SNR conditions. A method that jointly estimates the channel and decodes the transmitted information without using training sequences or any statistical assumptions on the channel characteristics is presented in [Perr1]. The structure of the space-time codes is exploited to formulate a Hidden Markov Model. The MAP algorithm is then used to detect the transmitted bits. The fading coefficients are important parameters of this model and are computed using the Expectation-Maximization (EM) algorithm. The algorithm allows iterative estimation of the parameters of the model and its good parameter estimation properties make it particularly suited to the HMM/MAP scheme. The algorithm can also exploit the information in training sequences if present. A description of the scheme is given below.

The received signal $r_j(t)$ at the j^{th} receive antenna at any time instant t is given by,

$$r_j(t) = \sum_{i=1}^{N_t} \alpha_{ij}(t) c_i(t) + \eta_j(t); 1 \le j \le N_r \tag{3-33}$$

η_t^j is the Gaussian noise at time instant t and is modeled by zero mean complex Gaussian variables with variance of $\frac{N_o}{2}$ per dimension. The transmitted symbols $c_i(t)$ can be expressed as functions of the non-coded sequence of information bits $b(t)$ by

$$c_i(t) = f_i\big(b(t) \quad b(t-1) \quad \dots \quad b(t-M+1)\big) \tag{3-34}$$

Functions f_i model the encoder with memory M and modulation. The received signal can now be written as

$$r_j(t) = \sum_{i=1}^{N_t} \alpha_{ij}(t) f_i\big(b(t) \quad b(t-1) \quad \dots \quad b(t-M+1)\big) + \eta_j(t); 1 \le j \le N_r \tag{3-35}$$

Let, $B(t) = \big[b(t) \quad b(t-1) \quad \dots \quad b(t-M+1)\big]^T$. It is seen that $B(t)$ can be a state of Markov chain. Assuming the knowledge of channel coefficients, the bits $b(t)$ can be detected by calculating the aposteriori probabilities for $B(t)$ to be equal to one of the permissible state levels $\varepsilon_i \{0,1\}^M$ conditioned on the set of past observations, $Y(t) = \big(y(1) \quad \dots \quad y(t)\big)$ using the forward recursions of the HMMs.

$$\Pr\big(B(t) = \varepsilon_i, Y(t)\big) = \Pr\left(\frac{y(t)}{B(t)} = \varepsilon_i\right) \sum_j A_{ji} \Pr\big(B(t) = \varepsilon_i, Y(t-1)\big) \tag{3-36}$$

where, A_{ji} is the transition probability matrix associated with Markov chain $B(t)$. Bit $b(t-M+1)$ can be detected by evaluating the marginal probability

$$\Pr\big(b(t), Y(t)\big) = \sum_{b(t-1),\dots,b(t-M+2)} \Pr\big(B(t), Y(t)\big) \tag{3-37}$$

These calculations require the knowledge of fading coefficients $\alpha_{ij}(t)$ which are iteratively estimated by using the Estimation-Maximization (EM) algorithm.

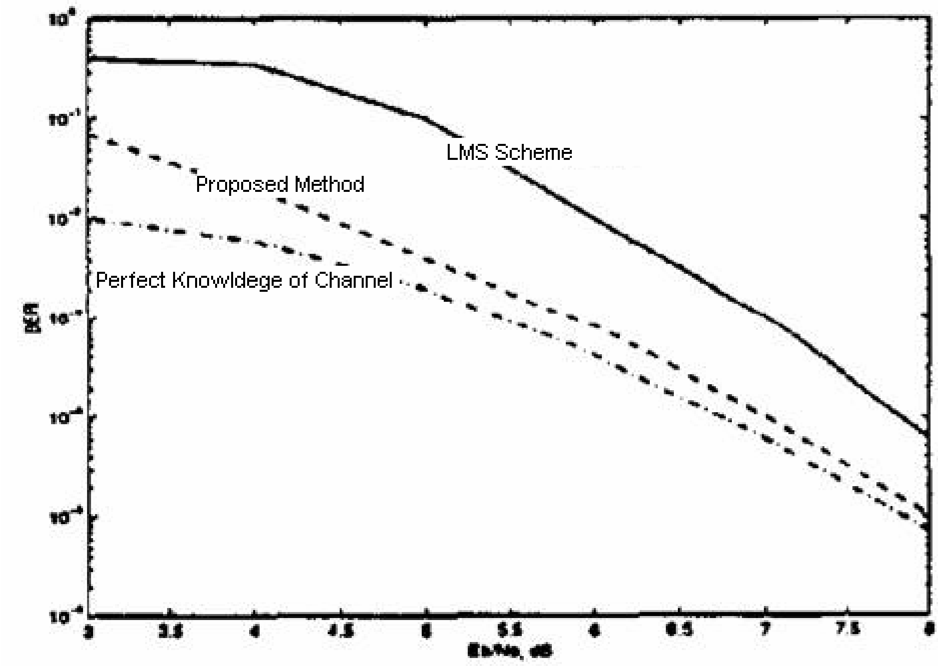

Figure 3-12: Comparison of the HMM/MAP based blind decoder performance with that of a decoder with perfect channel information

The EM algorithm performs the estimation by recursively maximizing the cost function

$$J_t(\alpha)=\sum_{k=0}^{t}\lambda^{t-k}E\left\{\log N\left(y_k-\alpha f(B);\sigma^{\varepsilon}\right)|Y\right\} \tag{3-38}$$

where $N\left(;\sigma^{\varepsilon}\right)$denotes the Gaussian function of mean zero and variance σ^2, and $0<\lambda<1$is a forgetting constant which is chosen to permit tracking of the channel variations. The resulting algorithm has the form

$$\widehat{\alpha}(t)=\widehat{\alpha}(t-1)+\left(y(t)-\widehat{\alpha}(t-1)f\left(\widehat{B}(t)\right)\right)f\left(\widehat{B}(t)\right)^{T}R(t)^{-1} \tag{3-39}$$

$$R(t)=\lambda R(t-1)+E\left\{f\left(B(t)\right)f\left(B(t)\right)^{T}|Y\right\} \tag{3-40}$$

where $\widehat{B}(t)$is the conditional mean estimate. The conditional moments are computed from the aposteriori probabilities in (3-37).

It is seen from Figure 3-12, that this scheme provides performance comparable to the case with perfect channel estimates for SNRs greater than 4dB. The figure also shows the performance when a gradient scheme like LMS is used for estimation. This leads to considerable reduction in computational complexity but, as can be observed, significant degradation in performance as well. The HMM/MAP scheme can be easily extended to multi-path channels as well.

3.4. Chapter Summary

This chapter provides an overview of channel estimation schemes for STTC. Training sequence based estimation schemes are widely favored as their implementation is relatively simple and most current wireless standards already provide for their use. The properties that are required by training sequences to be used in multiple transmit antenna systems in flat and frequency selective channels are derived and discussed. It is seen that the sequences transmitted from different transmit antennas must have zero (or very low) cross-correlation. Frequency-selective channels in addition necessitate zero (or very low) correlation between versions of the training sequence delayed by time intervals equal to the multi-path delays in the channel. The chapter presented capacitive analysis of the optimal amount of training required in MIMO channels and showed that the number of transmit antennas that can provide increased benefit in systems using training sequences to estimate the channel, is quite small. The chapter also presented a semi-blind iterative channel estimation scheme and a blind channel estimation scheme based on the Expectation-Maximization algorithm. Both schemes exploit the benefit provided by the redundancy and structure of a Space-Time Trellis code. The iterative scheme significantly reduces the amount of training required by a system. The EM based scheme requires no training, but can derive benefit from training, if present. The drawback of both schemes is the increased amount of processing at the receiver.

This thesis analyzes the influence of channel estimation errors on the performance of STTC. The channel estimation errors used in the analyses are mostly modeled assuming training based channel estimation. This estimation technique is the most widely used and hence is a valid choice. Also, a majority of the performance results and observations in this thesis are general enough to make sense for any alternate channel estimation technique.

4. Design and Performance of STTC in the Absence of Perfect Channel Estimates

The design criteria and performance analyses for STTC presented in Chapter 2 assume that perfect channel estimates are available at the receiver. However, in practical wireless systems, this is never the case. The noisy wireless channel precludes an exact characterization of channel coefficients. Even near-perfect channel estimates can necessitate huge overhead in terms of processing or spectral efficiency (due to long training sequences) which could offset any benefit obtained by using STTC. Hence it becomes important to study and compare the robustness of space-time trellis codes to channel estimation errors.

This chapter first re-evaluates the design criteria for STTC (Section 2.1) in the absence of perfect channel estimates at the receiver. Secondly, an expression for the exact pair-wise error probability of STTC in the presence of channel estimation errors is derived, which provides an analytical tool to evaluate the performance of STTC in the presence of imperfect channel estimates. Finally, the factors that influence the way channel estimation errors (CEEs) affect the performance of STTC are studied.

4.1. Design Criteria in the Presence of CEE

In chapter two, optimum design criteria were identified for STTC under different channel and system configurations. The TSC-RD and the TSC-DP design criteria were found to be appropriate for low diversity systems over quasi-static and fast, flat fading channels respectively. The trace criterion is found to be appropriate for systems with large diversity orders. The suitability of these criteria are re-examined here in the absence of perfect channel estimates.

4.1.1. TSC-RD Criteria

The TSC-RD criteria were proposed by Tarokh et al in [Taro1] (Section 2.2.1 A) for quasi-static flat fading channels. The authors also re-validated their design criteria in the presence of channel estimation errors in [Taro2]. This analysis is presented in this section.

Let the estimates of the channel coefficients be represented by $\hat{\alpha}_{ij} = \alpha_{ij} + e_{ij}$, where, e_{ij} is the estimation error. α_{ij} represents the channel coefficient from the i^{th} transmit antenna to the j^{th} receive antenna and is modeled in baseband by a complex Gaussian random variable with mean zero and variance 0.5 per dimension. Estimation error e_{ij} is assumed to be a complex Gaussian random variable with mean zero and variance σ_e^2. Consequently, $\hat{\alpha}_{ij}$ is also a zero mean complex Gaussian random variable. It is assumed to have a variance of σ^2 per dimension. $\hat{\alpha}_{ij}$ depends on α_{ij} with correlation μ given by

$$\mu = \frac{1}{\sqrt{1+\sigma_e^2}} \qquad 4\text{-}1$$

Conditioned on $\hat{\alpha}_{ij}$, the random variable α_{ij} has mean $\frac{\mu\hat{\alpha}_{ij}}{\sqrt{2}\sigma}$ and variance $\frac{\left(1-|\mu|^2\right)}{2}$ per dimension. The pair-wise probability of error for quasi-static fading channels, considering the mean and variance of the received vector to be conditioned on the channel estimates, $\hat{\alpha}_{ij}$ is given by (from [Taro2]),

$$p\left(c \to e \mid \hat{\alpha}\right) \le \exp\left(-\mu^2 d^2(c,e)\frac{E_s}{4N_0 + 4N_t\left(1-|\mu|^2\right)E_s}\right) \qquad 4\text{-}2$$

where,

$$d^2(c,e) = \sum_{j=1}^{N_r}\sum_{t=1}^{L}\left|\sum_{i=1}^{N_t}\frac{\hat{\alpha}_{i,j}}{\sqrt{2}\sigma}\left(c_i(t)-e_i(t)\right)\right|^2 \qquad 4\text{-}3$$

N_t, N_r and L denote the number of transmit antennas, the number of receive antennas and the length of the frame respectively. Let $\gamma_{ij} = \frac{\hat{\alpha}_{i,j}}{\sqrt{2}\sigma}$. γ_{ij} is a Gaussian with zero mean and variance 0.5 per dimension.

$$d^2(c,e) = \sum_{j=1}^{N_r}\sum_{t=1}^{L}\left|\sum_{i=1}^{N_t}\gamma_{ij}\left(c_i(t)-e_i(t)\right)\right|^2 \qquad 4\text{-}4$$

If $\Omega_j = \left(\gamma_{1j},...,\gamma_{N_tj}\right)$,

$$d^2(c,e) = \sum_{j=1}^{N_r}\Omega_j A\left(c,e\right)\Omega_j^* \qquad 4\text{-}5$$

$A\left(c,e\right)$ can be written as $VA\left(c,e\right)V^* = D$, where V is unitary matrix, the rows of which are made of the eigen vectors of $A\left(c,e\right)$ and D is a diagonal matrix whose elements are given by the eigen values $\lambda_i, i = 1,2,\ldots,N_t$ of $A\left(c,e\right)$. Let $\left(\nu_{1j},...,\nu_{N_tj}\right) = \Omega_j V^*$, then

$$d^2\left(c,e\right) = \sum_{j=1}^{N_r}\sum_{i=1}^{N_t}\lambda_i\left|v_{ij}\right|^2 \qquad 4\text{-}6$$

The conditional probability can now be represented as,

$$p(c \to e \mid \alpha) \le \exp\left(-\mu^2 \sum_{j=1}^{N_r} \sum_{i=1}^{N_t} \lambda_i \left| v_{ij} \right|^2 \frac{E_s}{4N_0 + 4N_t \left(1 - |\mu|^2\right) E_s} \right) \quad \text{4-7}$$

By virtue of V being a unitary matrix, v_{ij} are independent complex Gaussian random variables with zero mean and variance 0.5 per dimension. Thus $\left|v_{ij}\right|$ are also independent samples of a complex Gaussian random variable with mean zero and variance 0.5 per dimension and follows a Rayleigh distribution given by,

$$p\left(\left|v_{ij}\right|\right) = 2\left|v_{ij}\right| \exp\left(-\left|v_{ij}\right|^2\right) \text{ for } \left|v_{ij}\right| \ge 0 \quad \text{4-8}$$

The upper bound on the pair-wise error probability is the statistical average of Equation (4-7) with respect to the distribution of $\left|v_{ij}\right|$ and is given by,

$$p(c \to e \mid \alpha) \le \left(\frac{1}{\prod_{i=1}^{N_t} 1 + \mu^2 \lambda_i \frac{(E_s / 4N_0)}{1 + N_t \left(1 - |\mu|^2\right)(E_s / N_0)}} \right) \quad \text{4-9}$$

At high SNRs, $\frac{E_s}{4N_0} >> 1$ and hence Equation (4-9) can be expressed as

$$p(c \to e \mid \alpha) \le \left(\prod_{i=1}^{r} \lambda_i \right)^{-N_r} \left(\frac{\mu^2 (E_s / 4N_0)}{1 + N_t \left(1 - |\mu|^2\right)(E_s / N_0)} \right)^{-rN_r} \quad \text{4-10}$$

where, r is the rank of the matrix $A(c,e)$. From Equation (4-10) and Equation (4-1), it can be seen that for a given variance of the channel estimation error and a given channel SNR, the diversity gain and the coding gain can be maximized by maximizing the minimum rank and the minimum determinant of the distance matrix $A(c,e)$ between any two code-words. Hence the design criteria for STTC derived assuming perfect channel estimation [Taro1] hold in the presence of channel estimation errors as well. It can also be seen from Equation (4-10) that the diversity advantage offered by the code does not change in the presence of channel estimation errors.

The above conclusions are verified in Figure 4-1, which compares the simulated performance of a two-transmit and two-receive antenna system employing 4-state TSC code with and without prefect channel estimates. The channel is assumed to be constant over the length of the frame (130 symbols). Channel estimation is assumed to be done by appending a training sequence before the start of each transmitted frame. The training symbols are assumed to be transmitted at the same power as the data symbols. The length of the training sequence is a function of the channel SNR, the diversity order of the system and the coherence time of the channel, as determined from the capacity expression (3-27). The variance of the channel estimation error is equal to $\frac{N_o}{2L_tE_t}$ where, L_t is the length of the training sequence and E_t is the energy of each transmitted training symbol ([Taro2] and 3.1.1. A). The effect of channel estimation is modeled in the simulations by introducing errors of appropriate variance in the channel fade coefficients input to the receiver module.

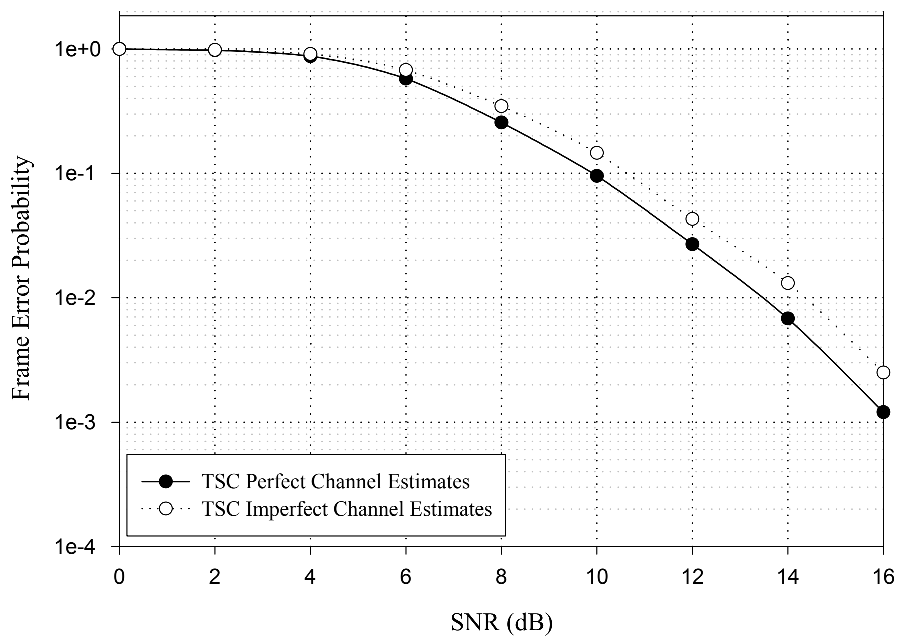

Figure 4-1: Performance of 2Tx-2Rx TSC code over quasi-static channel with optimal training

From Figure 4-1 it is observed that the diversity advantage, reflected in the slope of the performance curve, is preserved in the presence of channel estimation errors. However, a loss in coding gain is observed. Hence the TSC-RD criteria are appropriate for quasi-static flat-fading channels with imperfect channel information as well as perfect channel knowledge.

4.1.2. TSC-DP Criteria

This section extends the previous analysis (Section 4.1.1 and reference [Taro2]) to the TSC-DP criteria for fast fading channels and examines its validity in the absence of perfect channel estimates at the receiver.

Fast fading channels are modeled by making the channel coefficients α_{ij} vary from one symbol to the next. The pair-wise error probability in the absence of perfect channel estimates can then be expressed as,

$$p\left(c \to e \mid \hat{\alpha}\right) \leq \exp\left(-\mu^2 d^2(c,e) \frac{E_s}{4N_0 + 4N_t\left(1-|\mu|^2\right)E_s}\right) \qquad \text{4-11}$$

where,

$$d^2(c,e) = \sum_{j=1}^{N_r} \sum_{t=1}^{L} \left| \sum_{i=1}^{N_t} \frac{\hat{\alpha}_{i,j}(t)}{\sqrt{2}\sigma} \left(c_i(t) - e_i(t)\right) \right|^2 \qquad \text{4-12}$$

Substituting, $\gamma_{ij}(t) = \frac{\hat{\alpha}_{i,j}(t)}{\sqrt{2}\sigma}$ ($\gamma_{ij}(t)$ is a Gaussian with zero mean and variance 0.5 per dimension),

$$d^2(c,e) = \sum_{j=1}^{N_r} \sum_{t=1}^{L} \left| \sum_{i=1}^{N_t} \gamma_{ij}(t)\left(c_i(t) - e_i(t)\right) \right|^2 \qquad \text{4-13}$$

Following the analysis in Section *2.2.1.B*, the upper bound on the pair wise error probability can be written as,

$$P\left(c \to e\right) \leq \left(\prod_{t=1}^{l} \left|c(t) - e(t)\right|^2\right)^{-N_r} \left(\frac{\mu^2\left(E_s/4N_0\right)}{1 + N_t\left(1-|\mu|^2\right)\left(E_s/N_0\right)}\right)^{-lN_r} \qquad \text{4-14}$$

where, l is the number of time instances in a frame that $\left|c(t) - e(t)\right| \neq 0$

From expression (4-14) it can be seen the diversity advantage offered by the code is maximized, when the minimum number of time instances when any two code-words differ, is maximized. This translates into the TSC-DP criteria derived earlier for fast fading channels with perfect channel knowledge at the receiver. Thus the TSC-DP criteria hold in the presence of channel estimation errors as well. This is further illustrated by Figure 4-2 that compares the performance of the FVY code (designed according to TSC-DP criteria and introduced in the previous chapter) in the presence and absence of perfect channel estimates at the receiver. The variance of the channel estimation errors is assumed to be inversely proportional to the channel SNR and a proportionality factor of eight is arbitrarily chosen for the simulations. It is observed from the figure that CEEs do not cause a loss in the diversity advantage offered by the code.

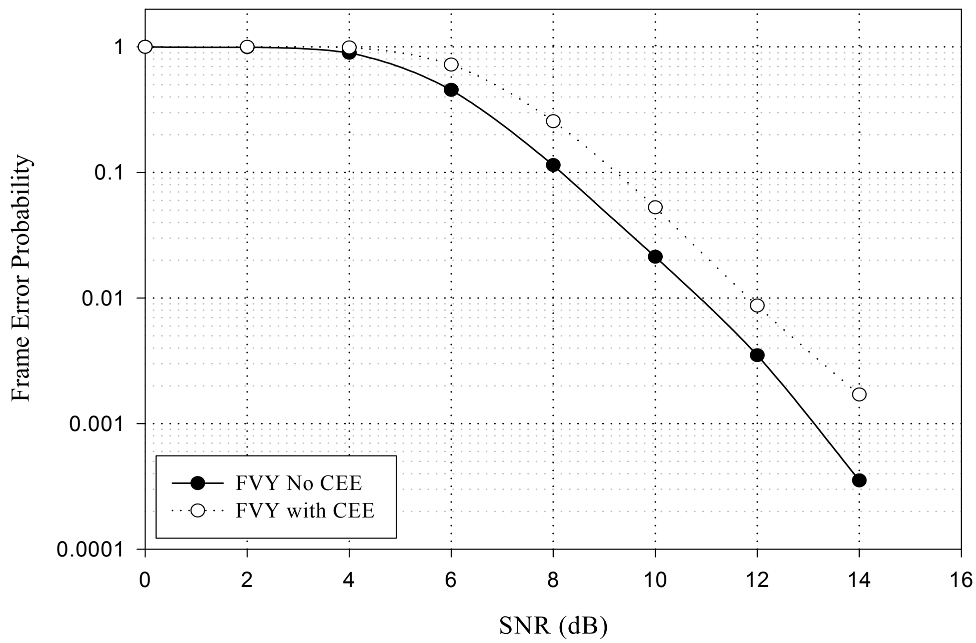

Figure 4-2: Performance of 2Tx-2Rx FVY code over fast fading channel with optimal training

4.1.3. Trace Criterion

The trace criterion was proposed for designing codes for systems with high diversity orders operating over both quasi-static and fast, non-frequency selective fading channels ([Chen1] and [Vuce1], Section 2.2.3) and assuming perfect channel estimates to be available at the receiver. This section evaluates the design criterion in the absence of perfect channel estimates.

The conditional pair-wise error probability with imperfect channel estimates can be expressed as (from Section 4.1.1),

$$p\left(c \to e \,|\, \alpha\right) \le \exp\left(-\mu^2 \sum_{j=1}^{N_r} \sum_{i=1}^{N_t} \lambda_i \left|v_{ij}\right|^2 \frac{E_s}{4N_0 + 4N_t\left(1-\left|\mu\right|^2\right)E_s} \right) \qquad 4\text{-}15$$

For a large $N_t N_r$ value (>3), according to Central Limit Theorem, the expression $\sum_{j=1}^{N_r} \sum_{i=1}^{N_t} \lambda_i \left|v_{ij}\right|^2$ approaches a Gaussian random variable D with mean

$$\mu_D = N_r \sum_{i=1}^{N_t} \lambda_i \qquad 4\text{-}16$$

and variance

$$\sigma_D^2 = N_r \sum_{i=1}^{N_t} \lambda_i^2 \tag{4-17}$$

Thus the unconditional pair-wise error probability can be upper-bounded by,

$$p\left(c \to e\right) \le \int_0^\infty \exp\left(-\mu^2 \frac{E_s}{4N_0 + 4N_t\left(1-\left|\mu\right|^2\right)E_s} D\right) p(D)dD \tag{4-18}$$

$$p\left(c \to e\right) \le \exp\left(\frac{1}{2}\left(\mu^2 \frac{E_s}{4N_0 + 4N_t\left(1-\left|\mu\right|^2\right)E_s}\right)^2 \sigma_D^2 - \mu^2 \frac{E_s}{4N_0 + 4N_t\left(1-\left|\mu\right|^2\right)E_s}\mu_D\right)$$
$$Q\left(\frac{\mu^2 \frac{E_s}{4N_0 + 4N_t\left(1-\left|\mu\right|^2\right)E_s}\sigma_D^2 - \mu_D}{\sigma_D}\right) \tag{4-19}$$

By using $Q\left(x\right) \le \frac{1}{2} e^{-\frac{x^2}{2}} \; x \ge 0$,

$$p\left(c \to e\right) \le \frac{1}{2}\exp\left(-N_r\mu^2 \frac{E_s \sum_{i=1}^{N_t} \lambda_i}{4N_0 + 4N_t\left(1-\left|\mu\right|^2\right)E_s}\right) \tag{4-20}$$

Expression (4-20) shows that pair-wise error probability can be minimized by maximizing the sum of eigen values of the matrix $A(c,e)$ or equivalently the trace of $A(c,e)$, which is given by,

$$\mathrm{tr(A)}=\sum_{\mathrm{i=1}}^{N_\mathrm{t}} \sum_{\mathrm{t=1}}^{\mathrm{L}} \left|\mathrm{e}_i\left(t\right) - c_i\left(t\right)\right|^2 \tag{4-21}$$

The trace of matrix $A(c,e)$ is equivalent to the Euclidean distance between the code words. It can be concluded that the pair-wise error probability between two code-words can be minimized if the Euclidean distance between the code words is maximized. Hence the design criterion for diversity orders >3 shown in [Chen1], remains valid in the presence of channel estimation errors.

This is further illustrated by the simulation results in Figure 4-3. The CYV code is implemented in a system with two-transmit and two-receive antennas. Training sequences of optimal length are used to estimate the channel. In the presence of channel estimation errors, the slope of the curve and hence the diversity advantage of the code is preserved but a loss in coding gain is observed. Hence the trace criterion is appropriate for systems with high diversity order in the absence of perfect channel information at the receiver as well as with perfect channel knowledge.

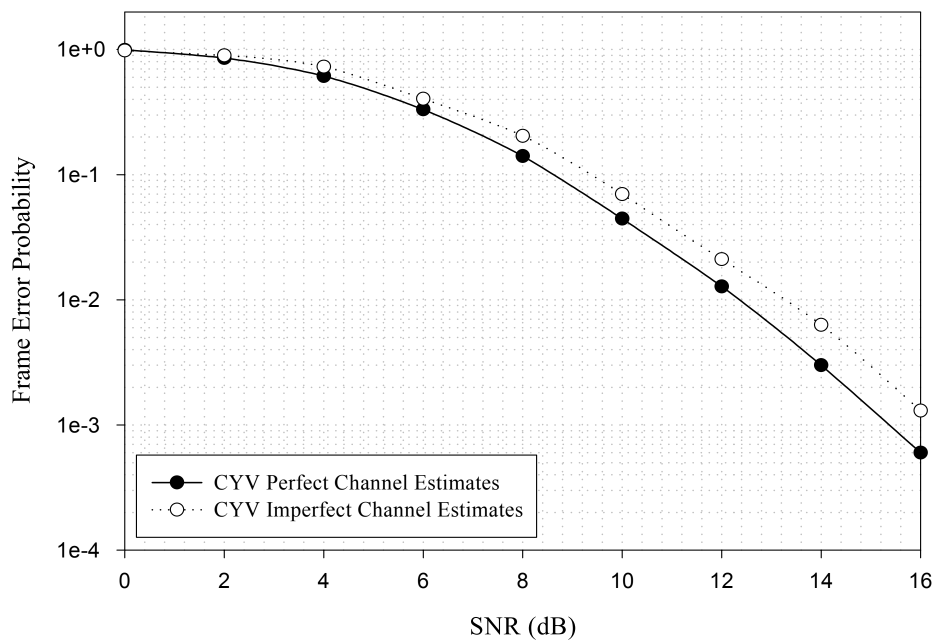

Figure 4-3 : Performance of 2Tx-2Rx CYV code over quasi-static channel with optimal training

4.2. Analytical Performance Analysis in the Presence of Channel Estimation Errors

Performance criteria for STTC were re-evaluated in the presence of channel estimation errors in the previous sections. These criteria were derived based on the upper bound on pair-wise error probability (PWEP) for quasi-static and fast fading channels with imperfect channel estimation. But as noted previously, this upper bound might be loose over SNR ranges of interest and quasi-static channels. Hence an exact evaluation of the PWEP is useful as it can be utilized to calculate an estimate of the bit error probability. This could help in the efficient design of systems employing these codes. The inclusion of channel estimation errors in an exact evaluation of PWEP is especially valuable, since this can be used to model real-world systems more closely. This section attempts to formulate an expression for the exact pair-wise error probability of STTC in the presence of channel estimation errors. From the exact probabilities, a good estimate of the overall bit error rate can be obtained by considering only a small number of dominant bit errors.

4.2.1. Exact PWEP for STTC with Imperfect Channel Estimates

An expression for the exact PWEP of STTC assuming perfect channel estimates is derived in [Uysa1] (Section 2.5). The derivation makes use of the residue technique combined with the characteristic function approach. This technique is generalized here to include the effects of imperfect channel estimation.

Following the definitions in Section 2.5.1, the received signal can be written as,

$$r_j = X\alpha_j + \eta_j \tag{4-22}$$

The channel estimates are given by,

$\hat{\alpha}_j(t) = \left(\hat{\alpha}_{1,j}(t), \hat{\alpha}_{2,j}(t), \ldots, \hat{\alpha}_{N_t,j}(t)\right)^T, \hat{\alpha}_j = \left(\hat{\alpha}_j(1), \hat{\alpha}_j(2), \ldots, \hat{\alpha}_j(L)\right)^T$

Assuming coherent detection, the Maximum Likelihood receiver minimizes the decision metric given by,

$$\mu(r,x) = \sum_{j=1}^{N_r} \left\| r_j - X\hat{\alpha}_j \right\|^2 \tag{4-23}$$

Let a random variable D be defined as

$$D = \sum_{j=1}^{N_r} \left(\left\| r_j - \hat{X}\hat{\alpha}_j \right\|^2 - \left\| r_j - X\hat{\alpha}_j \right\|^2 \right) \tag{4-24}$$

The Pair-wise Error Probability (PWEP) represents the probability of incorrectly choosing sequence *c*, when sequence *e* was transmitted. It is given by,

$$P(c \to e) = P\left(\mu\left(r, \hat{X}\right) \le \mu(r, X)\right) = P(D \le 0) \tag{4-25}$$

Let $\phi_D(s)$, be the Laplace transform of D. Then the probability density function of D is given by an inverse transform,

$$P(c \to e) = P(D \le 0) = -residue\left[e^{s\delta}\phi_D(s)/s\right]_{RP,\delta=0} \tag{4-26}$$

The channel estimate $\hat{\alpha}_j$ can be written as,

$$\hat{\alpha}_j = \alpha_j + \varepsilon_j \tag{4-27}$$

where, ε_j is the estimation error in the calculation of the channel estimates. It is modeled as independent samples of zero mean Gaussian random variables with variance $\sigma_e^2/2$ per dimension. Substituting and expanding for D,

$$D=\sum_{j=1}^{N_r}\alpha_j^H\left(X-\widehat{X}\right)^H\left(X-\widehat{X}\right)\alpha_j+\alpha_j^H\left(X-\widehat{X}\right)^H\eta+\eta^H\left(X-\widehat{X}\right)\alpha_j$$
$$+\varepsilon_j^H\left(X-\widehat{X}\right)^H\eta_j+\eta_j^H\left(X-\widehat{X}\right)\varepsilon-\varepsilon_j^H\widehat{X}^H\left(X-\widehat{X}\right)\alpha_j-\alpha_j^H\left(X-\widehat{X}\right)^H\widehat{X}\varepsilon_j \qquad 4\text{-}28$$

Consider the following matrices defined as,

$$y_j=\left[\left(\left(X-\widehat{X}\right)\alpha_j\right)^T\quad\left(X\varepsilon_j\right)^T\quad\left(\widehat{X}\varepsilon_j\right)^T\quad\eta_j^T\right]^T \text{ and } A=\begin{bmatrix} I_L & 0_L & I_L & -I_L \\ 0_L & 0_L & 0_L & I_L \\ I_L & 0_L & 0_L & -I_L \\ -I_L & I_L & -I_L & 0_L \end{bmatrix}.$$

where, 0_L and I_L are zero and identity matrices respectively with dimension L. From the definitions, it can be seen that y_j is zero mean and A is a Hermitian matrix. D can be written in a quadratic form of variable y_j as

$$D=\sum_{j=1}^{N_r}y_j^H A y_j \qquad 4\text{-}29$$

From results in [Turi1], the characteristic function of D (given that y_j is zero mean and A is a Hermitian) is,

$$\phi_D(s)=\prod_{j=1}^{N_r}\frac{1}{\det\left(I_{4L}+sC_{y_j}A\right)}=\prod_{i=1}^{r}\left(\frac{1}{1+s\lambda_i}\right)^{N_r} \qquad 4\text{-}30$$

where, C_{y_j} is the covariance matrix of y_j and r is the number of non-zero eigen values, λ_i, of $C_{y_j}A$. C_{y_j} can be expressed as,

$$C_{y_j}=\begin{bmatrix} C_{xaxa} & 0_L & 0_L & 0_L \\ 0_L & C_{xexe} & C_{xe\widehat{x}e} & 0_L \\ 0_L & C_{\widehat{x}exe} & C_{\widehat{x}e\widehat{x}e} & 0_L \\ 0_L & 0_L & 0_L & N_0 I_L \end{bmatrix} \qquad 4\text{-}31$$

where,

$$C_{xaxa}(l,k)=E\left(\sum_{m=1}^{N_t}\sum_{q=1}^{N_t}\alpha_{m,j}(l)\alpha_{q,j}(k)\left(c_m(l)-e_m(l)\right)\left(c_q(k)-e_q(k)\right)^*\right)$$
$$C_{xexe}(l,k)=E\left[\sum_{m=1}^{N_t}\sum_{q=1}^{N_t}\varepsilon_{m,j}(l)\varepsilon_{q,j}^*(k)c_m(l)c_q^*(k)\right]$$

$$C_{\hat{x}e\hat{x}e}(l,k) = E\left[\sum_{m=1}^{N_t}\sum_{q=1}^{N_t}\varepsilon_{m,j}(l)\varepsilon_{q,j}^{*}(k)e_m(l)e_q^{*}(k)\right] \qquad \text{4-32}$$

$$C_{xe\hat{x}e}(l,k) = E\left[\sum_{m=1}^{N_t}\sum_{q=1}^{N_t}\varepsilon_{m,j}(l)\varepsilon_{q,j}^{*}(k)c_m(l)e_{q,j}^{*}(k)\right]$$

$$C_{\hat{x}exe}(l,k) = E\left[\sum_{m=1}^{N_t}\sum_{q=1}^{N_t}\varepsilon_{m,j}(l)\varepsilon_{q,j}^{*}(k)e_m(l)c_{q,j}^{*}(k)\right]$$

Different channel scenarios lead to different forms for the covariance matrix C_{y_j} and are discussed below.

4.2.1. A. Quasi-static Flat Fading Scenario

The fade coefficients are assumed to be constant over the length of the frame and to change from one frame to the next. The values of the covariance matrices in this scenario can be expressed as,

$$C_{xaxa}(l,k) = \left(\sum_{m=1}^{N_t}\left(c_m(l)-e_m(l)\right)\left(c_m(k)-e_m(k)\right)^{*}\right)$$

$$C_{xexe}(l,k) = \sigma_e^2\sum_{m=1}^{N_t}c_m(l)c_m^{*}(k)$$

$$C_{\hat{x}e\hat{x}e}(l,k) = \sigma_e^2\sum_{m=1}^{N_t}e_m(l)e_m^{*}(k) \qquad \text{4-33}$$

$$C_{xe\hat{x}e}(l,k) = \sigma_e^2\left[\sum_{m=1}^{N_t}c_m(l)e_m^{*}(l)\right]$$

$$C_{\hat{x}exe}(l,k) = \sigma_e^2\left[\sum_{m=1}^{N_t}e_m(l)c_m^{*}(l)\right]$$

The characteristic function is expressed as,

$$\phi_D(s) = \left(\frac{1}{\det\left(I_{2L}+sC_{y_j}A\right)}\right)^{N_r} = \prod_{i=1}^{r}\left(\frac{1}{1+s\lambda_i}\right)^{N_r} \qquad \text{4-34}$$

r is the number of non-zero eigen values, λ_i, of $C_{y_j}A$ and gives the diversity order of the system. Comparing with results in Section 2.5.1.A , it can be seen that the diversity order is not reduced in the presence of channel estimation errors.

4.2.1. B. Fast Fading Scenario

The fade coefficients are constant over a single symbol period and differ over consecutive symbol periods independently. Thus only the diagonal elements of, each covariance matrix, are non zero.

$$
\begin{aligned}
C_{xaxa}(l,k) &= \begin{cases} \left(\sum_{m=1}^{N_t} \left| \left(c_m(l) - e_m(l) \right) \right|^2 \right) & l = k \\ 0 & otherwise \end{cases} \\
C_{xexe}(l,k) &= \begin{cases} \sigma_e^2 \sum_{m=1}^{N_t} \left| c_m(l) \right|^2 & l = k \\ 0 & otherwise \end{cases} \\
C_{\hat{x}e\hat{x}e}(l,k) &= \begin{cases} \sigma_e^2 \sum_{m=1}^{N_t} \left| e_m(l) \right|^2 & l = k \\ 0 & otherwise \end{cases} \\
C_{xe\hat{x}e}(l,k) &= \begin{cases} \sigma_e^2 \left[\sum_{m=1}^{N_t} c_m(l) e_m^*(l) \right] & l = k \\ 0 & otherwise \end{cases} \\
C_{\hat{x}exe}(l,k) &= \begin{cases} \sigma_e^2 \left[\sum_{m=1}^{N_t} e_m(l) c_m^*(l) \right] & l = k \\ 0 & otherwise \end{cases}
\end{aligned}
\qquad 4\text{-}35
$$

4.2.1.C. Spatially Correlated Quasi-static Scenario

The transmit antennas are assumed to be correlated. The correlation between transmit antenna m and q is given by $\rho_{mq} = E(\alpha_{m,n}\alpha_{q,n}^*)$. The receive antennas are assumed to be independent and the channel quasi-static.

$$
\begin{aligned}
C_{xaxa}(l,k) &= \sum_{m=1}^{N_t} \left(c_m(l) - e_m(l) \right) \left(c_m(k) - e_m(k) \right)^* + \sum_{\substack{m=1 \\ m \neq q}}^{N_t} \sum_{q=1}^{N_t} \rho_{mq} \left(c_m(l) - e_m(l) \right) \left(c_q(k) - e_q(k) \right)^* \\
C_{xexe}(l,k) &= \sigma_e^2 \left(\sum_{m=1}^{N_t} c_m(l) c_m^*(k) + \sum_{\substack{m=1 \\ m \neq q}}^{N_t} \sum_{q=1}^{N_t} \rho_{mq} c_m(l) c_q^*(k) \right) \\
C_{\hat{x}e\hat{x}e}(l,k) &= \sigma_e^2 \left(\sum_{m=1}^{N_t} e_m(l) e_m^*(k) + \sum_{\substack{m=1 \\ m \neq q}}^{N_t} \sum_{q=1}^{N_t} \rho_{mq} e_m(l) e_q^*(k) \right)
\end{aligned}
$$

4-36

$$C_{xe\hat{x}e}(l,k)=\sigma_e^2\left[\sum_{m=1}^{N_t}c_m(l)e_m^*(k)+\sum_{\substack{m=1\\m\neq q}}^{N_t}\sum_{q=1}^{N_t}\rho_{mq}c_m(l)e_q^*(k)\right]$$

$$C_{\hat{x}exe}(l,k)=\sigma_e^2\left[\sum_{m=1}^{N_t}e_m(l)c_m^*(k)+\sum_{\substack{m=1\\m\neq q}}^{N_t}\sum_{q=1}^{N_t}\rho_{mq}e_m(l)c_q^*(k)\right]$$

4.2.1. D. Spatially Correlated Fast Fading Scenario

The channel is assumed to undergo fast fading and the transmit antennas are assumed to be correlated by a factor ρ_{mq}.

$$C_{xaxa}(l,k)=\begin{cases}\left(\sum_{m=1}^{N_t}\left|\left(c_m(l)-e_m(l)\right)\right|^2\right)+\sum_{\substack{m=1\\m\neq q}}^{N_t}\sum_{q=1}^{N_t}\rho_{mq}\left(c_m(l)-e_m(l)\right)\left(c_q(l)-e_q(l)\right)^* & l=k\\ 0 \quad otherwise\end{cases}$$

$$C_{xexe}(l,k)=\begin{cases}\sigma_e^2\left(\sum_{m=1}^{N_t}\left|c_m(l)\right|^2+\sum_{\substack{m=1\\m\neq q}}^{N_t}\sum_{q=1}^{N_t}\rho_{mq}c_m(l)c_q^*(l)\right) & l=k\\ 0 \quad otherwise\end{cases}$$

$$C_{\hat{x}e\hat{x}e}(l,k)=\begin{cases}\sigma_e^2\left(\sum_{m=1}^{N_t}\left|e_m(l)\right|^2+\sum_{\substack{m=1\\m\neq q}}^{N_t}\sum_{q=1}^{N_t}\rho_{mq}e_m(l)e_q^*(l)\right) & l=k\\ 0 \quad otherwise\end{cases} \tag{4-37}$$

$$C_{xe\hat{x}e}(l,k)=\begin{cases}\sigma_e^2\left[\sum_{m=1}^{N_t}c_m(l)e_m^*(l)+\sum_{\substack{m=1\\m\neq q}}^{N_t}\sum_{q=1}^{N_t}\rho_{mq}c_m(l)e_q^*(l)\right] & l=k\\ 0 \quad otherwise\end{cases}$$

$$C_{\hat{x}exe}(l,k)=\begin{cases}\sigma_e^2\left[\sum_{m=1}^{N_t}e_m(l)c_m^*(l)+\sum_{\substack{m=1\\m\neq q}}^{N_t}\sum_{q=1}^{N_t}\rho_{mq}e_m(l)c_q^*(l)\right] & l=k\\ 0 \quad otherwise\end{cases}$$

4.2.1. E. Pair-wise Error Probability and Average Bit Error Probability

As mentioned earlier PWEP is the probability density function of D which is given by a simple inverse Laplace transform of $\phi_D(s)$.

$$P\left(c \rightarrow e\right) = P\left(D \le 0\right) = -residue\left[e^{s\delta}\phi_D\left(s\right)/s\right]_{RP,\delta=0} \qquad \text{4-38}$$

As the PWEP is given in terms of a residue computation, it does not lend itself for an evaluation of the upper bound on the bit error probability using classical transfer function bound approach which implicitly takes into account error event of all lengths. An approximation of the actual bit error probability can be obtained by taking into consideration error events up to a particular pre-determined length.

Probability of bit error is approximated by,

$$P_b = \frac{1}{b}\sum_{c \ne e} q\left(c \rightarrow e\right) P\left(c \rightarrow e\right) \qquad \text{4-39}$$

Where, b is the number of input bits per trellis transition and $q\left(c \rightarrow e\right)$ is the Hamming distance between the code word and the error event. An estimate for the bit error probability is obtained by only considering error events up to a particular error event length v. The choice of v is very critical. It should be done such that all the dominant error events in a particular SNR range are included, while keeping a check on the computational complexity which exponentially increases with v.

4.2.2. Analytical Performance Results

Analytical performance-curves are generated for STTC systems with two transmit antennas and employing the *TSC* code in the presence channel estimation errors of a defined variance. Figure 4-4 shows the performance of the STTC scheme over a quasi-static channel with two receive antennas and a CEE of 10dB. Figure 4-5 shows the code performance in a system with one receive antenna operating over a fast fading channel and with CEE variance of 20dB. The performance over spatially correlated channels is shown in Figure 4-6. These are compared to actual simulation results of the systems. Error lengths (v) up to two, three and four are considered. Performance is plotted over average energy per bit to noise ratio at each receive antenna.

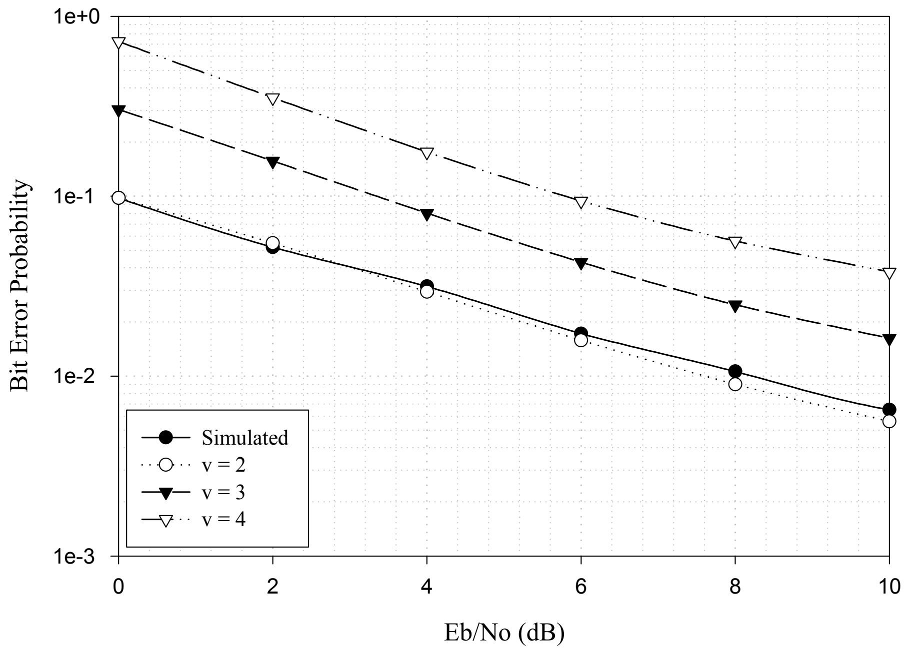

Figure 4-4 : Performance of 2Tx-2Rx TSC code over quasi-static channel with 10dB CEE

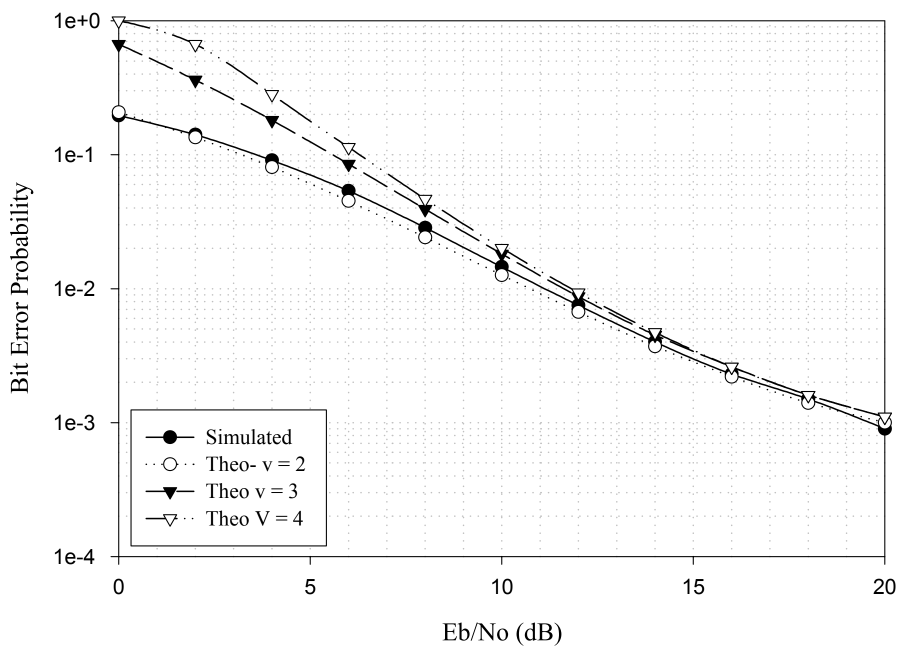

Figure 4-5 : Performance of 2Tx-1Rx TSC code over fast fading channel with 20dB CEE

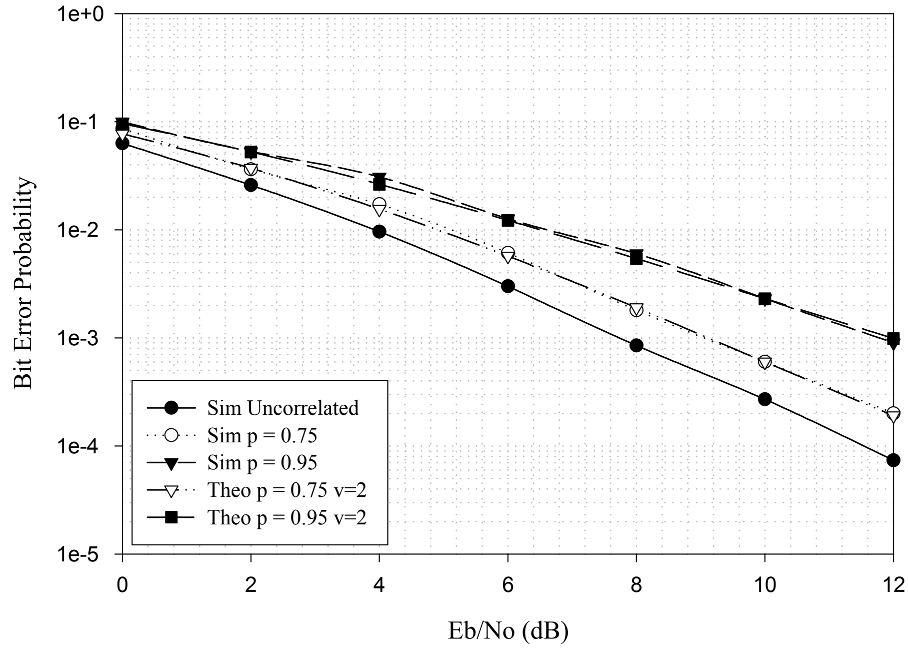

Figure 4-6 : Performance of 2Tx-2Rx TSC code over quasi-static spatially correlated channel with 20dB CEE

The analytical results are seen to approximate the simulated performance of the code In the presence of channel estimation errors of defined variance, for error events of length two. This agrees with the previously observed fact that the dominant error length of the TSC code is two (Section 2.6). It is observed from the slopes of the curves in Figure 4-4 and Figure 4-5, that the PWEP calculation converges to a union bound faster in the fast fading channel than in the quasi-static fading channel. Figure 4-6 shows the performance the TSC code in a spatially correlated channel with correlation factors of 0.75 and 0.95. It is seen that the performance advantage of using two transmit antennas over a single transmit antenna is lost when the correlation between the transmit antennas is 0.95.

The discussion shows that the analytical results model the performance of STTC in the presence of CEE well and re-iterate earlier observations of the same.

4.3. Effect of Channel Estimation Errors on STTC Performance

In the previous sections it is seen that though the diversity advantage offered by space time trellis codes is maintained in the presence of channel estimation errors, they suffer significant loss in performance. This section attempts to characterize factors that influence the extent of this performance loss. The influence of code-choice, the number of transmit/ receive antennas in the system, the amount of training used and the coherence time of the channel on the performance degradation due to channel estimation errors is studied. This analysis highlights issues that must be taken into consideration for optimal design and choice of components for practical wireless communication systems.

4.3.1. Optimal Amount of Training

An information theoretic approach to the amount of training required for multiple antenna systems is discussed in [Hass1] (Section 3.2). A lower bound for the performance of space-time codes is found based on the capacity that can be achieved by using training sequences to estimate the channel (Equation 3-27). This capacity expression is a function of the SNR of the system, the number of transmit and receive antennas, the coherence time of the channel and the length of the training sequence. The analysis also shows that if the SNR of the training sequence is allowed to vary independent of the SNR of the symbols, the optimal number of training symbols required is equal to the number of transmit antennas. But when the training sequence and the data sequence are required to be transmitted at the same power, the optimum training length might be larger than the number of transmit antennas. The former is seen to be more efficient than the latter case, but is not very feasible. Hence the analyses in this chapter assume data and training symbols to be transmitted at the same power.

The optimal amount of training, when data and training symbols are transmitted at the same power, is calculated by finding the training sequence length that maximizes Equation (3-27) via Monte-Carlo simulations. Figure 4-7 plots the optimal amount of training for two-transmit, two-receive antenna MIMO systems as a function of the coherence time of the channel and for different channel SNRs.

Figure 4-8 shows the simulated performance of a two-transmit, two-receive antenna system employing the CYV code, operating over a channel with coherence interval of 120 symbols and an SNR of 12dB, for various training lengths. The energy spent on training is compensated by a reduction in the energy available for data transmission. The performance is seen to be the best for a training length of 16 and to degrade for shorter or longer training lengths. It can be observed from Figure 4-7, that the optimal amount of training for a two-transmit, two receive antenna system for a coherence interval of 120 and a channel SNR of 12dB is 16. The optimal length results thus correspond to the maximum achievable performance of a system with training.

It is observed from Figure 4-7 that the proportion of training compared to the length of the data sequence in a frame, increases with decreasing coherence time. It can also be seen that the optimal amount of training increases with channel SNRs. At low SNRs the required training interval might increase to as much as half the coherence time. This can be noticed in Figure 4-10 which shows the optimal training length for a MIMO system with eight transmit and two receive antennas.

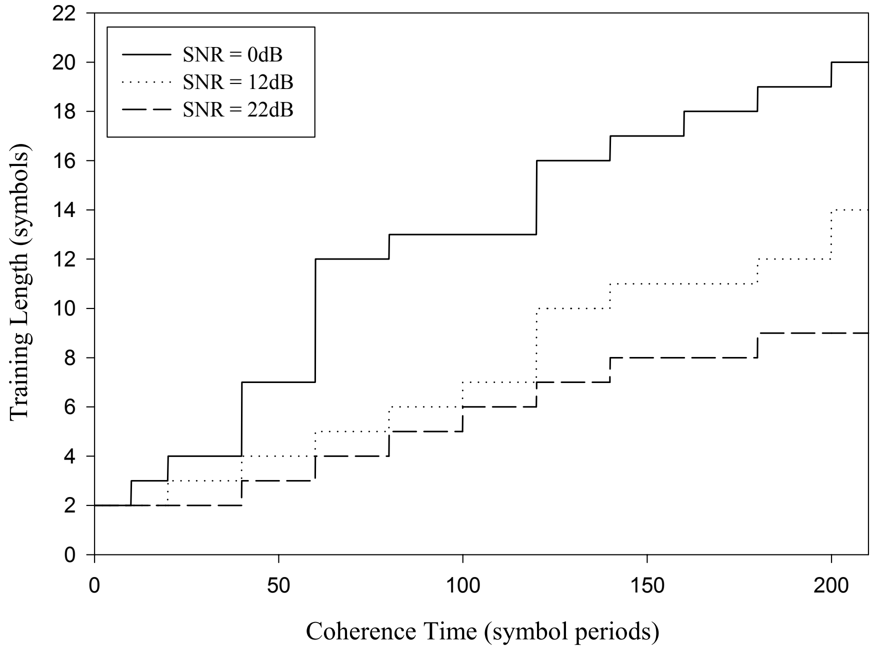

Figure 4-7: Optimal training length for 2Tx-2Rx MIMO system over different channel coherence times and difference channel SNRs

The dependence of the required training sequence length on the SNR of the system is further illustrated in Figure 4-9. The BER performance of a two-transmit, two-receive antenna system (of diversity order 4) employing a *CYV* code is plotted for varying training sequence lengths, over a quasi-static channel. These are compared with the performance of a diversity-two Maximum Ratio Receive Combining scheme (with one transmit and two receive antennas and a diversity order of 2) to highlight any diversity advantage obtained. The training symbols and the data symbols are transmitted at the same transmit power. The energy overhead due to training in a frame is not compensated for in these simulations. It is seen that at high SNRs the diversity-four scheme employing a training sequence of length two gives a better performance compared to that of the diversity-two scheme. But at low SNRs a training sequence of length of at least eight is required to achieve a better performance.

Increasing the diversity of a system by adding transmit antennas increases the burden on the channel estimation scheme and longer training sequences are required. This can be observed in Figure 4-10 which shows the optimal amount of training required by a eight-transmit, two-receive antenna system. The optimal amount of training compared to Figure 4-7, is considerably larger.

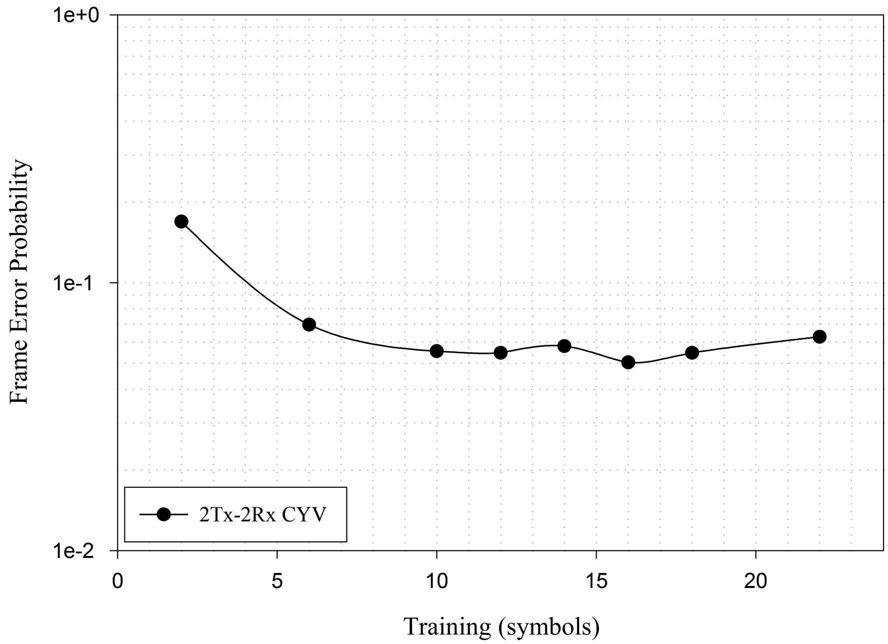

Figure 4-8: Performance comparison with varying length of training, for a 2Tx-2Rx CYV Code at 12dB channel SNR

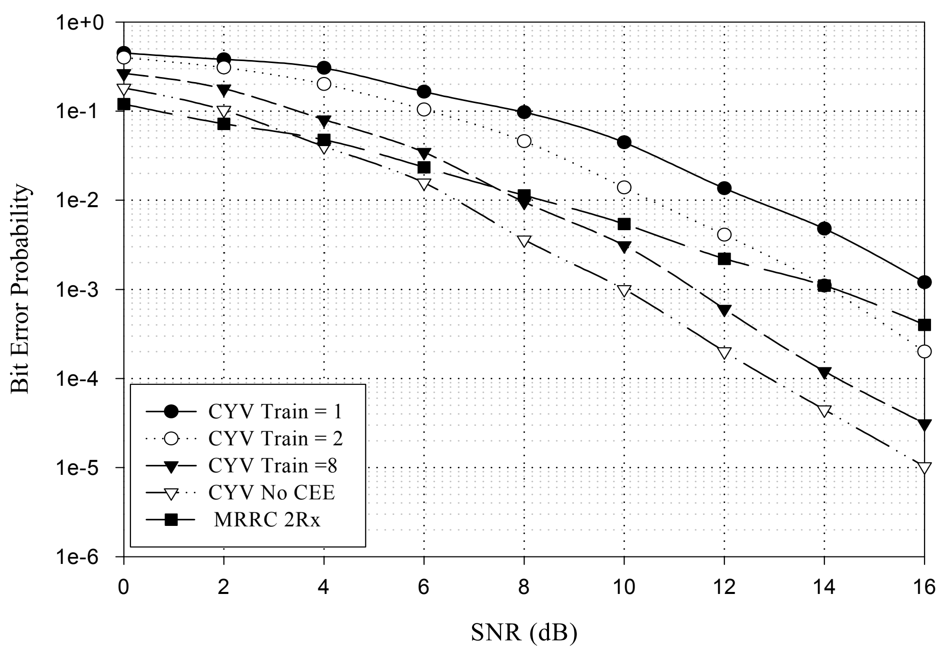

Figure 4-9 : Performance of 2Tx-2Rx CYV Code over quasi-static channel with varying training lengths

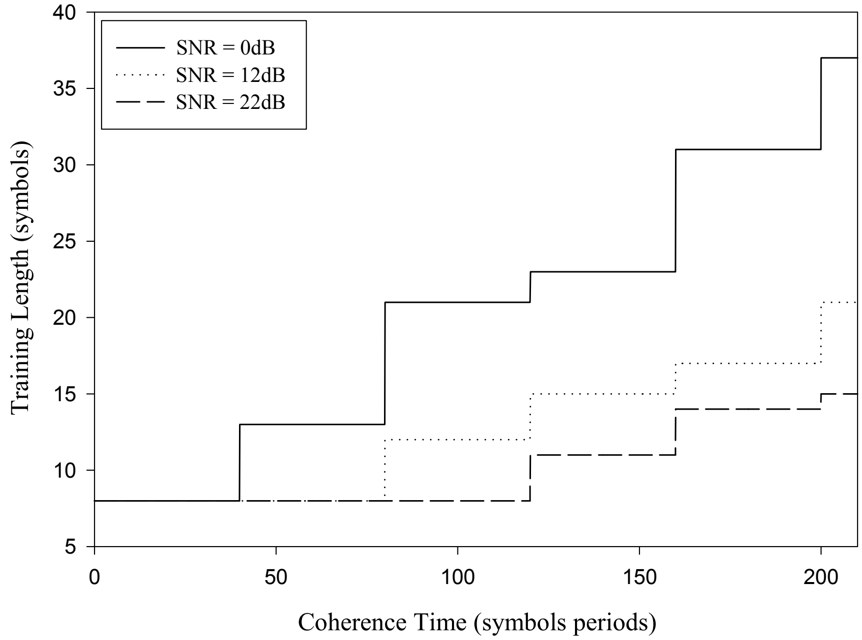

Figure 4-10: Optimal training length for 8Tx-2Rx MIMO system over different channel coherence times and difference channel SNRs

This required increase in training is further illustrated in Figure 4-11 which compares the performance of *CYV* codes for three and four transmit antennas [Chen2] in the presence and absence of perfect channel estimates, over a quasi-static channel. Two receive antennas are assumed. The training sequences for both systems are of length eight. The training symbols are transmitted at the same power as the data symbols from each of the transmit antennas. The variance of the channel estimation error is given by $\frac{N_0}{2L_tE_t}$, where N_0 is the channel noise variance, E_t is the energy of the transmitted training symbol and L_t is the length of the training sequence. The total transmitted energy for both schemes are the same in the simulations. Consequently, the CEE variance is larger for a four antenna system than a three antenna system and the performance of the four antenna system degrades much more than the three antenna system and this is reflected in Figure 4-11. Hence an increase in the number of transmit antennas being used results in an increase in the required length of the training sequence.

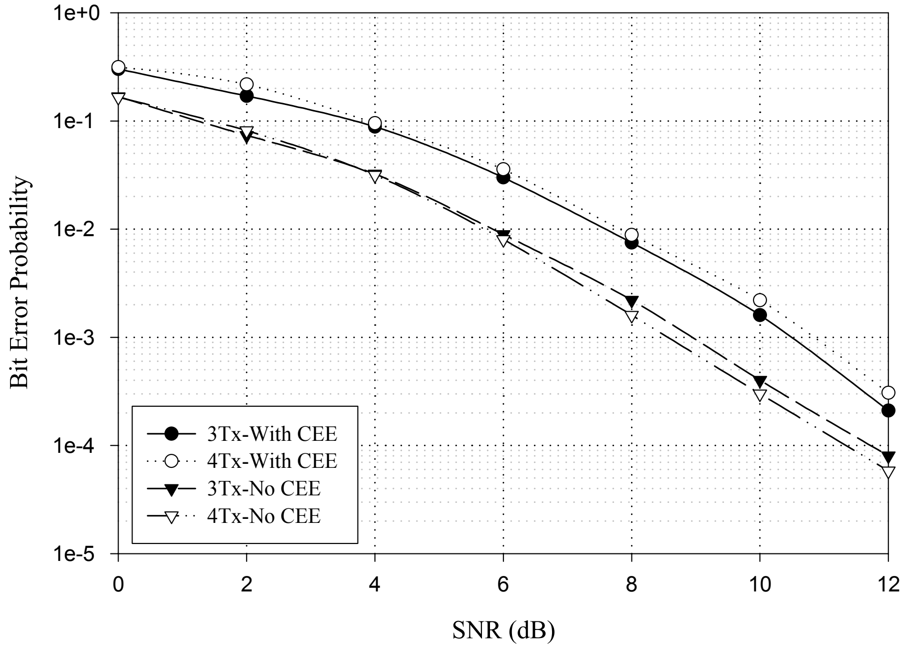

Figure 4-11: **Performance comparison of CYV Code for different number of transmit antennas; in a system with two receive antennas and training length of eight**

It is also interesting to note in Figure 4-11 that when a training sequence of length eight is used, a system with transmit diversity of three starts to perform better than a system with a transmit diversity of four. This reveals the fact that it is possible for the diversity advantage offered by a system to be lost by an inappropriate choice of training sequence length. In some cases, (for e.g. when there is a constraint on the length of the training sequence because of the required throughput), it might in fact be more beneficial to use lower diversity systems.

4.3.2. Performance Degradation due to CEE

In the presence of channel estimation errors, the pair-wise error probability between two code-words is upper-bounded by, (Equation 4-20),

$$p(c \to e) \leq \frac{1}{2}\exp\left(-N_r \mu^2 \frac{E_s \sum_{\mathrm{i=1}}^{N_t} \sum_{\mathrm{t=1}}^{\mathrm{L}} \left| \mathrm{e}_i(t) - c_i(t) \right|_i^2}{4N_0 + 4N_t\left(1 - |\mu|^2\right) E_s} \right) \qquad \text{4-40}$$

The above expression can also be written as

$$p(c \to e) \leq \frac{1}{2}\exp\left(-N_r \alpha \sum_{\mathrm{i=1}}^{N_t} \sum_{\mathrm{t=1}}^{\mathrm{L}} \left| \mathrm{e}_i(t) - c_i(t) \right|_i^2 \right) \qquad \text{4-41}$$

where,

$$\alpha = \mu^2 \frac{\left(\frac{E_s}{4N_0}\right)}{1 + N_t\left(1 - |\mu|^2\right)\left(\frac{E_s}{N_0}\right)} \qquad \text{4-42}$$

It is seen that for a constant channel SNR, a change in the variance of the channel estimation error will lead to a variation in the natural log of the probability, proportional to the minimum trace of the distance matrix. Thus performance loss is proportional to the Euclidean distance between code-words. Consequently, a scheme with a larger value of "minimum trace" will suffer a larger loss in performance when channel estimation errors are introduced or increased.

The *TSC* code (Section 2.5.1) for two transmit antennas has a minimum trace of 4 and the corresponding *CYV* code has a minimum trace of 10. The BER performance of both codes in a system with diversity four is simulated and plotted in Figure 4-12. The channel is assumed to be quasi-static and the channel coefficients are estimated using a training sequence of length eight transmitted at the same power as the data sequence from the transmit antennas (A two-transmit antenna, two-receive antenna system has an optimal training length of eight at around 13dB). As predicted, the *CYV* code suffers a greater loss in performance in the presence of CEE than *TSC*. The performance of the *CYV* code degrades by around 1dB while the performance of the *TSC* code degrades by only about 0.5dB. The FEP curves in Figure 4-1and Figure 4-3 exhibit similar behavior. The BER curves are used for comparison as the difference in performance is more perceivable.

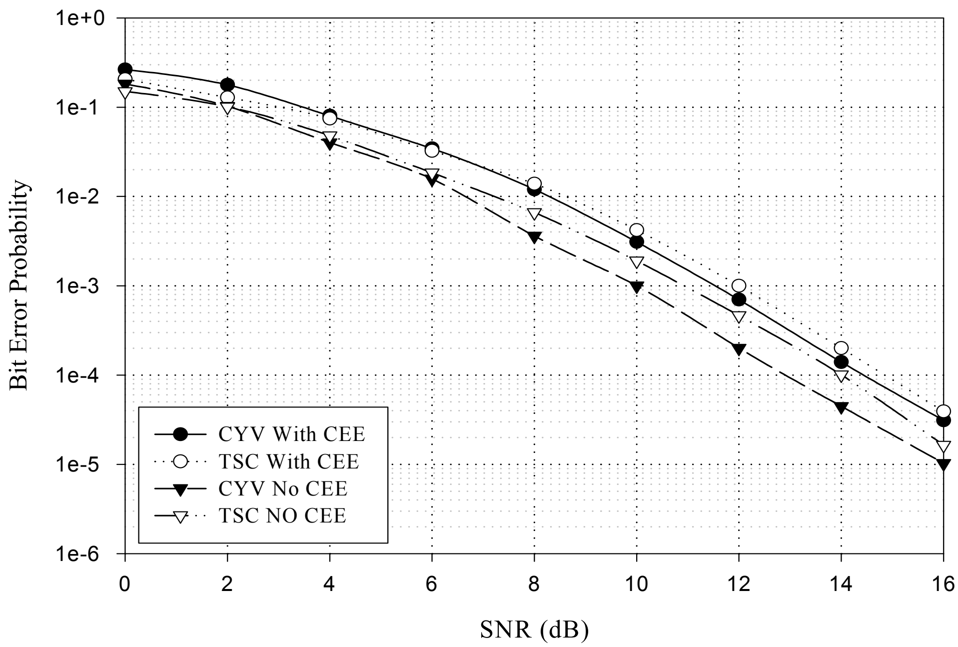

Figure 4-12 : Performance comparison of 2Tx-2Rx CYV and TSC codes over quasi-static channel with training length of eight

The performance of the *BBH* code, which has a minimum trace of 6, is compared in the presence and absence of perfect channel estimates in Figure 4-13. As expected, the degradation in performance is larger than the *TSC* code but less than the *CYV* code.

For optimal training lengths, it can be observed from Figure 4-1 and Figure 4-3 that the CYV code always performs better than the TSC code. Figure 4-14 shows the performance of the TSC and CYV code using optimal amount of training over a range of coherence intervals. The energy spent on training is compensated by a reduction in the energy available for data transmission. It is seen that the performance in the presence of training degrades with decreasing coherence interval, though this not the case in the presence of perfect channel knowledge at the receiver. This is because low coherence intervals warrant additional training resulting in significant reduction of the energy available for data transmission. It should be noted that with perfect channel estimation coherence time is irrelevant. Imperfect channel estimates changes this dramatically. It is also observed that in the presence of optimal training, the CYV code performs better than the TSC code over the range of coherence intervals considered. However, constraints on the amount of training or the generation of good channel estimates might be present in practical systems. This prompts the study of these codes in the presence of large CEEs. The performance of the *TSC* and the *CYV* code in the presence of training sequences of different lengths is shown in Figure 4-15. A longer training sequence usually translates to better channel estimates and lower variance of channel estimation errors. As noted earlier, the performance degradation corresponding to an increase in channel estimation error is more pronounced in the case of the *CYV* code. It is also seen that at low CEE SNRs the *CYV* code performs worse than the *TSC* code though this is not the case at high

CEE SNRs or in the presence of perfect channel estimates. This is reiterated in Figure 4-16, which plots the performance of the *TSC* and *CYV* code for varying CEE SNRs at a channel SNR of 8dB

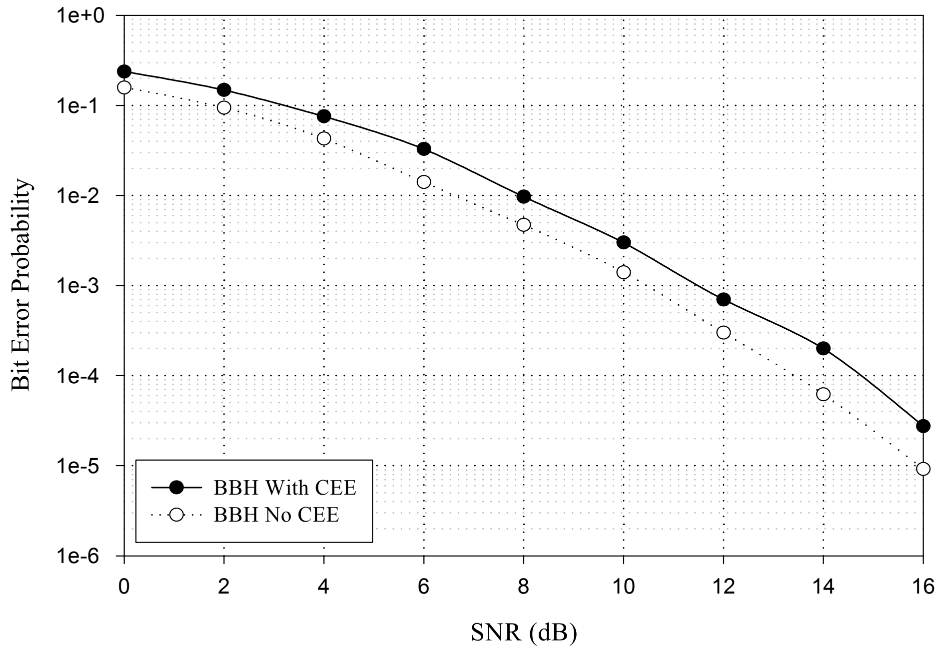

Figure 4-13 : Performance comparison of 2Tx-2Rx BBH code over quasi-static channel with training length of eight

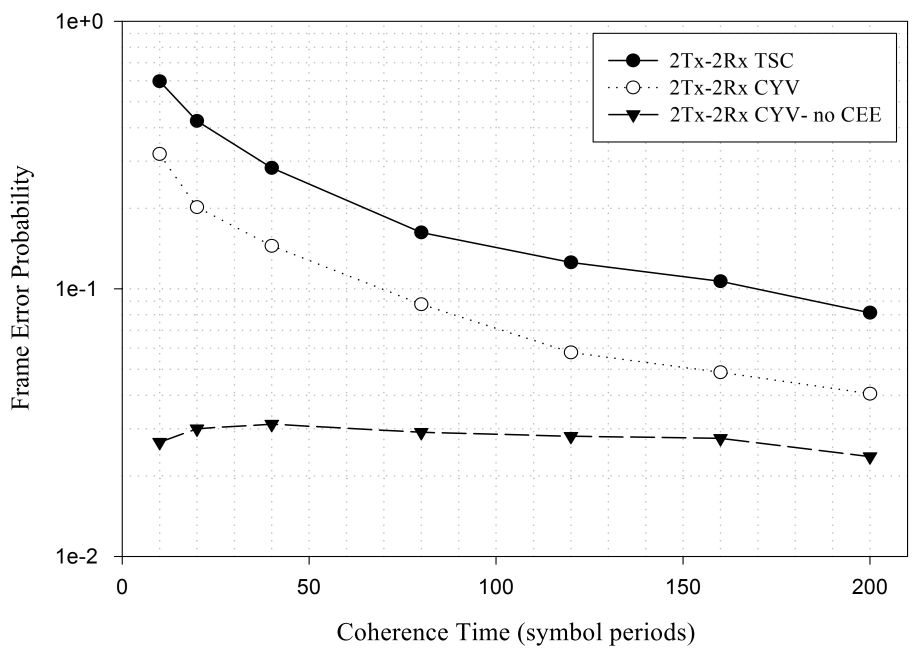

Figure 4-14: Performance comparison of 2Tx-2Rx, CYV and TSC codes over varying channel coherence times, 10dB channel SNR and optimal training

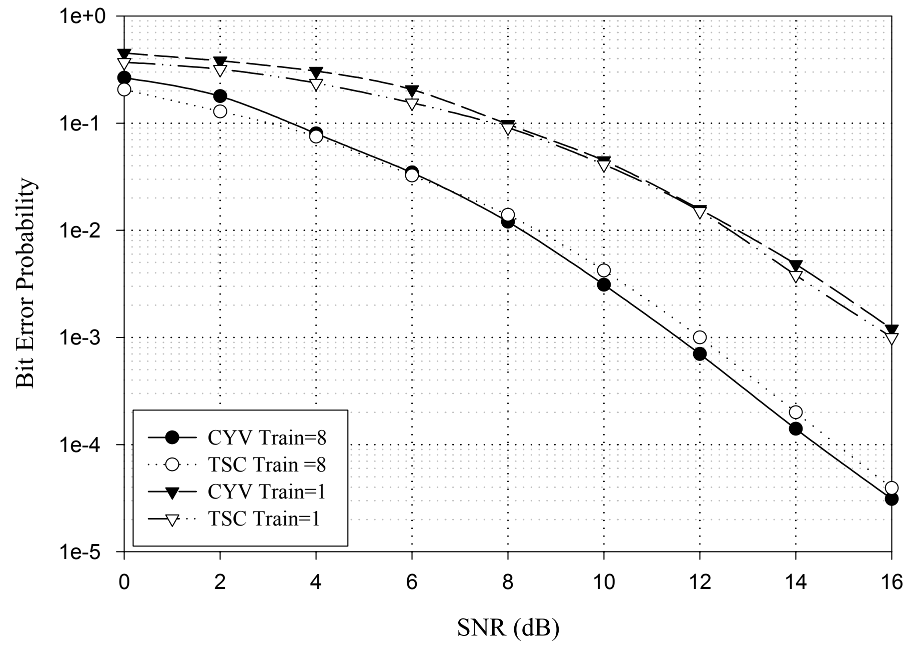

Figure 4-15 : Performance comparison of 2Tx-2Rx, CYV and TSC codes over quasi-static channel and for different training lengths

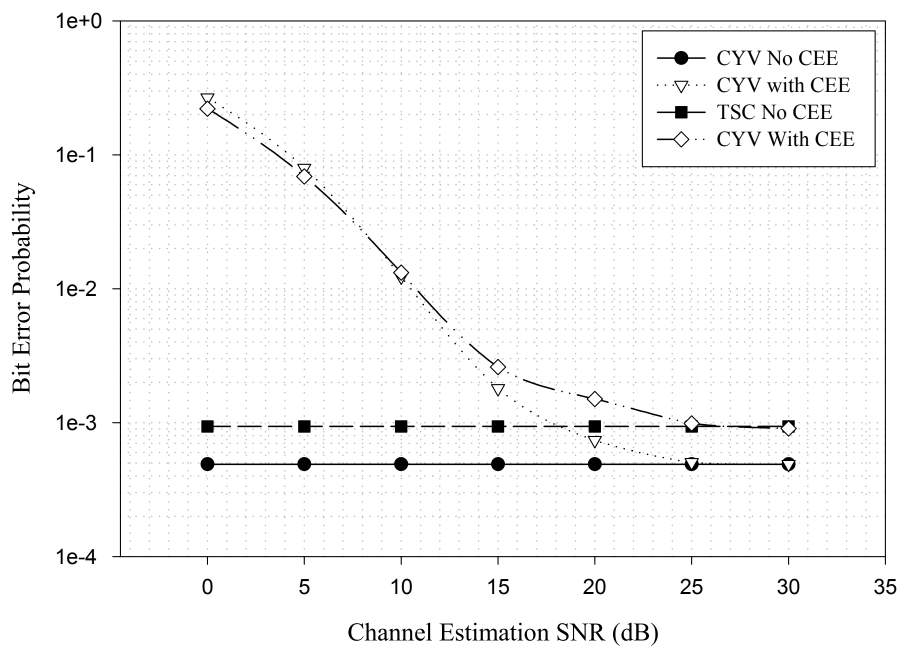

Figure 4-16 : Performance comparison of 2Tx-2Rx, CYV and TSC codes over quasi-static channel and for different CEE SNRs

The performance results from Chapter 2 (Section 2.4), show that in the presence of large diversities and perfect channel estimates, the *CYV* codes performs better that the *TSC* codes. But, as can be seen from the results presented in this section, large channel estimation errors alter the behavior of these codes and the TSC code is seen to perform better than the *CYV* code.

The analysis in this section thus shows that results in the presence of perfect channel estimates do not necessarily translate into equivalent results in the absence of these perfect estimates. The accuracy of channel estimates should be an important consideration during the design of optimal codes for a practical communication system.

4.3.3. Performance Sensitivity to Coherence Time

Coherence time quantifies the similarity of channel responses at different times and is a statistical measure of the time duration over which the channel impulse response is invariant [Rapp1]. Hence shorter coherence times correspond to shorter intervals over which the channel coefficients have strong potential for amplitude correlation. This translates into an increase in the proportion of training in a frame with a decrease in the coherence time of the channel, as the receiver has to re-estimate the channel at smaller intervals. This relation between the required amount of training and the coherence time is reiterated in Figure 4-7 and Figure 4-10. As a result of this dependence the capacity of a channel can be predicted to decrease with decreasing coherence time of the channel, when the channel is estimated by sending training sequences. The capacity analysis of the channel in the presence of training carried out by Hassibi and Hochwald in [Hass1] reflects this result (Chapter 3). This section analyses the influence of this dependence on different diversity order systems and coding schemes and examines its role in altering the design choice.

Different STTC schemes are simulated for a specific channel SNR and over coherence intervals of 10-200 symbols periods. Different coherence intervals are simulated by changing the Doppler spreads of the Rayleigh fading channel. The channel is assumed to be estimated by training sequences of lengths, optimized for the specific MIMO system, channel SNR and coherence time of the channel. The energy spent on training in a frame is compensated for, by reducing the energy available for sending the data bits in the frame. The new SNR of the system is given by, $SNR_{new} = SNR_{actual} \times \frac{L - L_t}{L}$ where L is the frame length and L_t is the required amount of training. The length of the frame is fixed at 200 symbols. Normalizing the SNR thus, provides a tool to compare the performance of systems with different configurations and different training lengths.

As noted before, the capacity of a MIMO system with training decreases with an increase in the coherence time of the channel. This capacity result is mirrored in the comparison between Figure 4-17 and Figure 4-19. Figure 4-17 shows the performance of one and two transmit antenna schemes over the coherence time of the channel assuming perfect channel estimates to be available at the receiver. Figure 4-19 shows the performance of these schemes with the channel assumed to be estimated by training sequences of optimal length. The loss in performance due to training increases with decreasing coherence time.

This roughly translates into a loss in capacity of a channel with training with decreasing coherence time of the channel.

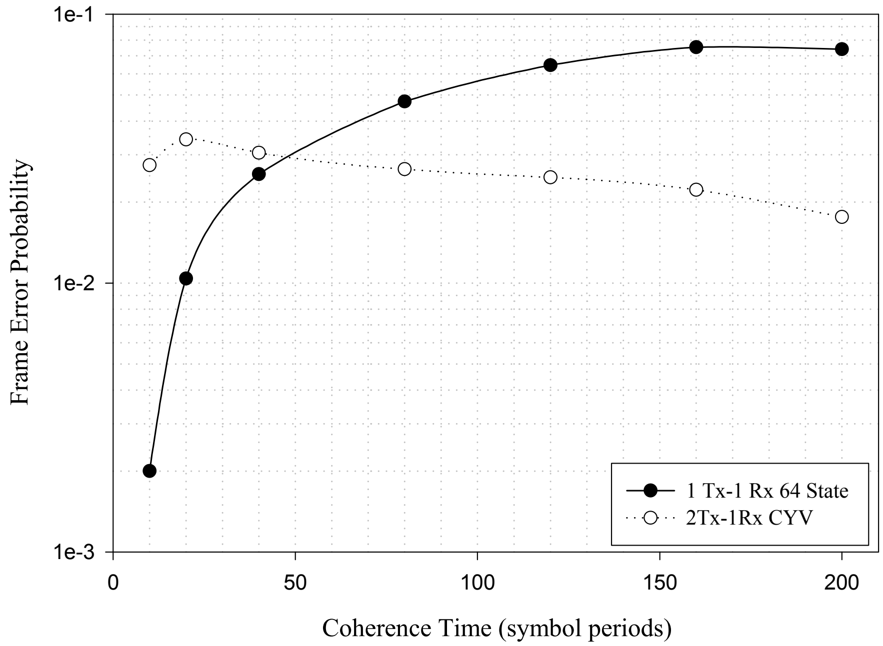

Figure 4-17: Performance comparison of 2Tx and 1Tx Schemes over channel coherence time, with 22dB channel SNR and perfect channel estimates at the receiver

Figure 4-17, as mentioned before, compares the performance of one and two transmit antenna schemes over the coherence time of the channel assuming perfect channel estimates to be available at the receiver. The single antenna system employs a 64-state convolution channel code (with generator matrix $\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$) and the QPSK modulation scheme. It is observed that the performance of the single antenna system improves with a decrease in the coherence time of the channel and hence takes advantage of the additional time diversity offered by a faster fading channel. But the performance of the 4-state STTC (CYV STTC code is used as previous results indicate that this code performs well over both fast and quasi-static flat fading channels) initially degrades with decreasing coherence time, but eventually starts to improve. This can attributed to the specific design methodology employed for STTC (Section 2.5.3). The STTC is unable to exploit the temporal diversity provided by coherence times slightly lower than the frame size and also suffers a loss in coding gain. However, the performance begins to improve for very small coherence intervals (about 20 symbol intervals in this case) and represents the scenario when the loss in coding gain is compensated by the increased available temporal diversity. Figure 4-18 shows the performance of the CYV code with and without the inclusion of an interleaver in the system. The interleaver increases the temporal diversity of the system, hence enabling the system to compensate for the loss in coding gain at larger coherence intervals as compared to the system that does not use an interleaver.

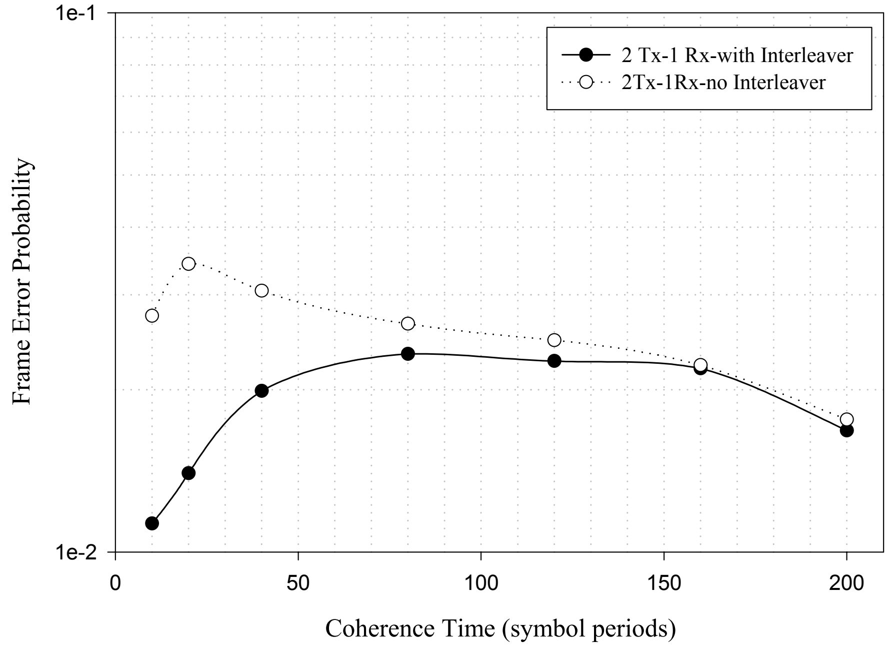

Figure 4-18: Performance comparison of 2Tx-2Rx CYV code over coherence time in the presence of perfect channel estimates, with and without an interleaver

Figure 4-19 compares the normalized performance of an STTC scheme with transmit diversity of two with different single antenna systems with the channel assumed to be estimated by training sequences of optimal length. By comparing Figure 4-17 and Figure 4-19, it is observed that the degradation of performance due to training and channel estimation is greater for the two-transmit antenna scheme than the single antenna scheme. This is attributed to fact that an increase in the number of antennas, increases the required amount of training and consequently increases the energy overhead of the system.

Figure 4-19 also shows that though the two-transmit antenna scheme is penalized more than the single antenna system, it always performs better than a single antenna scheme with no channel coding. But the performance of the single antenna system employing some form channel coding approaches the performance of the two-transmit antenna system for small coherence times of the channel. This is observed from the plots of the normalized performance of the single transmit antenna system employing a 4-state code (with generator matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$) and the 64-state code discussed earlier. The performance of the 64-state single antenna code is seen to be better than the two-transmit STTC over small coherence times and with perfect channel estimates from Figure 4-17, though over large coherence times the STTC scheme offers a benefit over the 64-state code. As noted earlier, the 64-state code profits from the temporal diversity available at small coherence times of the channel but the 4-state STTC does not seem to. However, it provides spatial diversity which is very helpful when no temporal diversity is available (i.e. at long coherence times). Figure 4-19 shows that increased cost of training for the two-transmit antenna system, decreases the margin of this benefit rendering the performance of the two systems comparable. The increased temporal diversity offered by using an interleaver is seen to reduce the performance degradation suffered by the two transmit antenna system at low channel coherence times. However the performance is still worse than that of the 64-state code at low coherence intervals. In practical wireless systems, adding channel codes (and increasing the complexity of the channel code) for single transmit antenna systems is regarded to more cost-efficient than adding an extra antenna (and hence an extra RF chain) at the transmitter end. Hence the comparison of the two-transmit system employing a four state STTC to a single antenna system employing a 64-state convolution code is justified. It is thus seen that at low coherence times and with the option of increasing the complexity of the channel codes employed available, channel training offsets the benefit of transmit diversity provided by STTC.

Figure 4-20 compares the performance of multiple antenna systems with varying number of transmit antennas and employing STTC. It is seen that for very large Doppler spreads (fast fading channels), a system with four transmit antennas and a system with two transmit antennas give comparable performance. Hence using four transmit antennas in these channel conditions is of no benefit

The capacity results from Chapter 3 [Hass1], show that the optimal number of antennas for a MIMO system decreases with decreasing coherence time of the channel. This trend can also be observed from Figure 4-20. Systems with larger number of transmit antennas suffer more degradation in channels with small coherence intervals as the comparative increase in training is larger. It is thus seen that the benefit due to increased diversity is offset by the penalty incurred due to the corresponding increase in training in channel with small coherence intervals. Hence the coherence interval over which the system is required to operate should be an important consideration in system design.

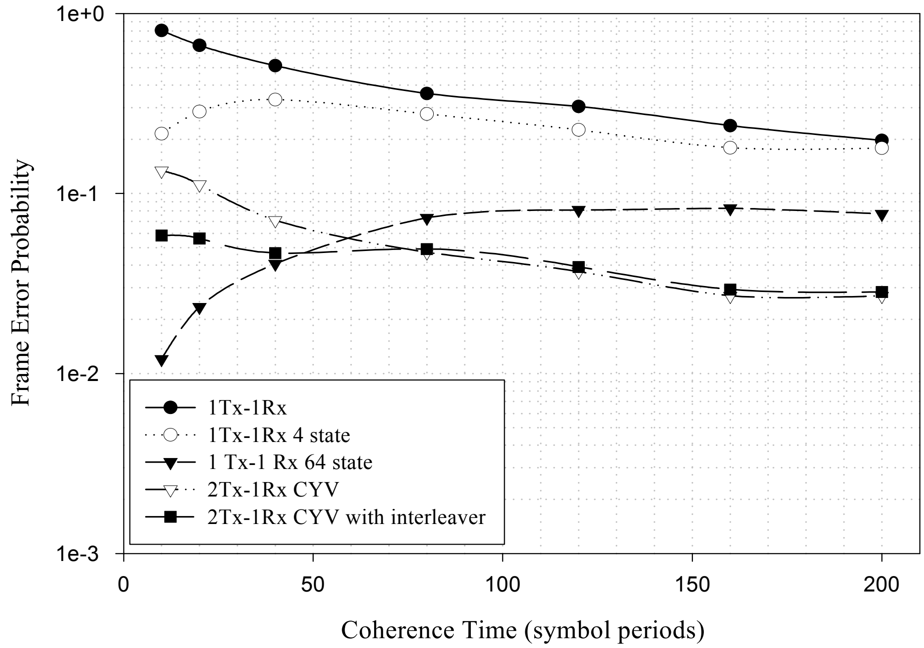

Figure 4-19: Performance comparison of different transmit schemes with 22dB channel SNR, 1 receive antennas and optimal training

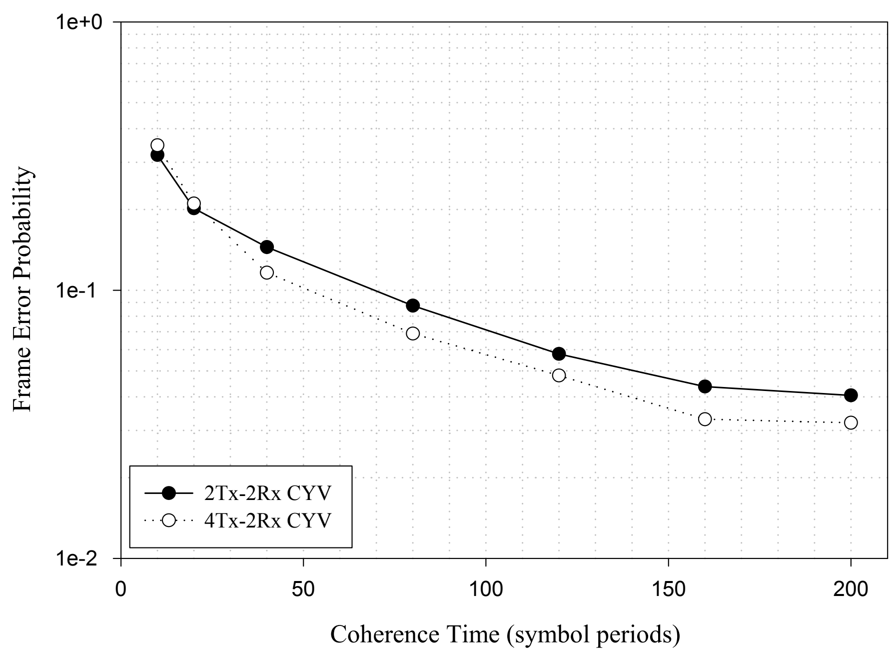

Figure 4-20: Performance comparison of different transmit schemes with 12dB channel SNR, 2 receive antennas and optimal training

4.3.4. Performance Sensitivity to Varying Number of Antennas

From the expression for the upper bound of PWEP of STTC in the presence of STTC 4-20, it is observed that variation in probability of error due to changes in the variance of the channel estimation errors depends on the number of transmit and receive antennas in addition to the Euclidean distance between code-words (as seen in Section 4.3.1). This is investigated in this section.

Figure 4-21 compares the performance of a two-transmit antenna and a four-transmit antenna system employing the *CYV* STTC code in the presence and absence of channel estimation error. The system has two receive antennas. The channel estimation error is assumed to be 15dB in both systems. It is seen that the performance degradation in the four-transmit antenna system is slightly more than the performance degradation in the two-transmit antenna system. Hence an increase in the number of transmit antennas is observed to cause an increase in the degradation due to channel estimation errors. The same can be inferred from Equation (4-20) as an increase in the number of transmit-antennas generally leads to an increase in the factor $\sum_{i=1}^{N_t} \lambda_i$.

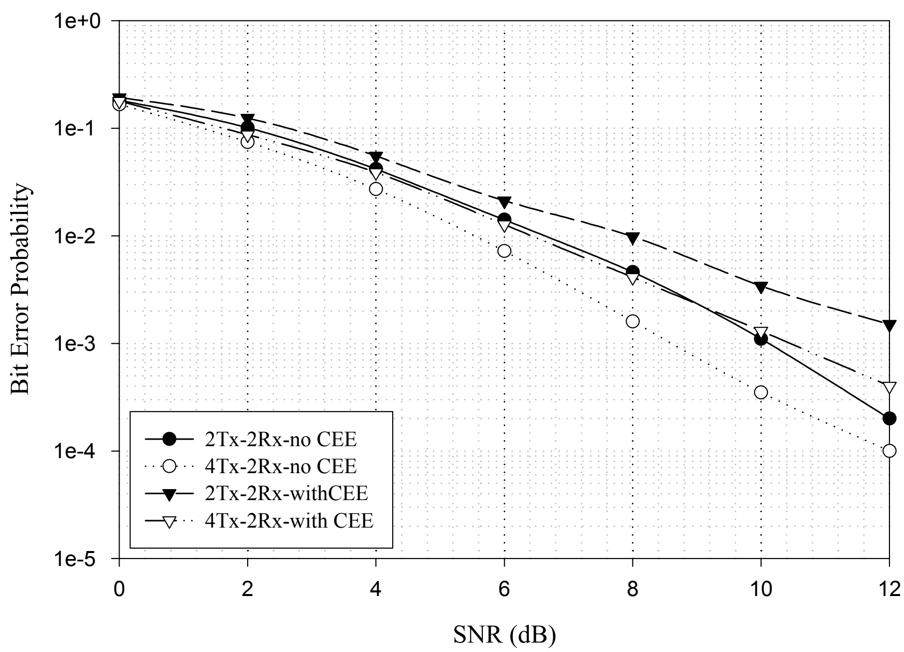

Figure 4-21 : Performance of CYV code with different transmit antennas, over a quasi-static channel and CEE of 15dB

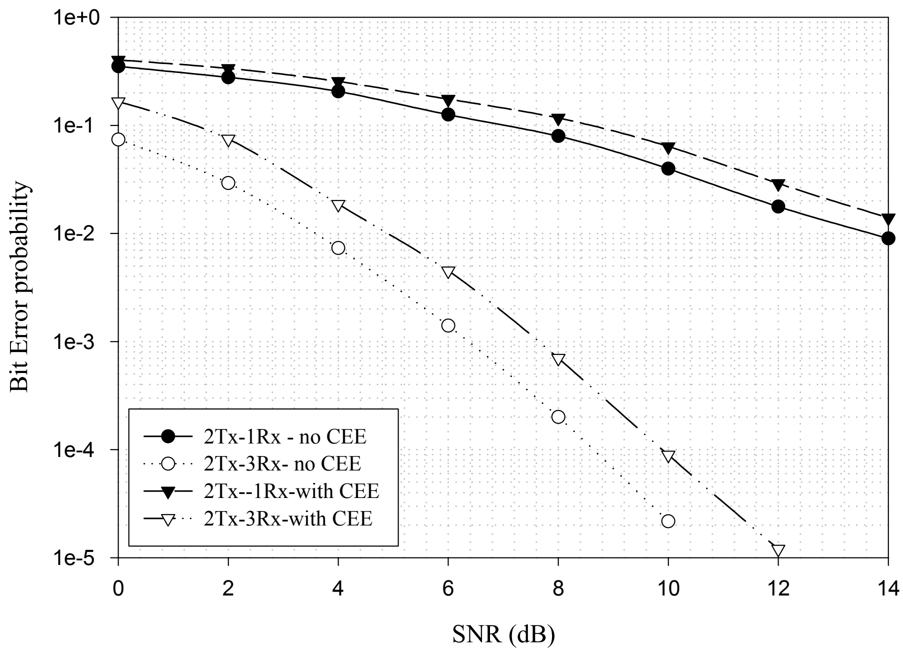

Figure 4-22: Performance of CYV code with different receive antennas, over a quasi-static channel and training of length eight

It can also be noted from Equation (4-20) that an increase in the number of receive-antennas will lead to similar results. This is verified in Figure 4-22, which compares the performance a two receive antenna and a three receive antenna system. Both have two transmit antennas and the *CYV* STTC code. The channel is estimated by a training sequence of length eight. The degradation in performance is much larger for the three-receive antenna system than the one-receive antenna system. However, the large diversity benefits preclude any possibility of the single receive antenna system performing better than the system with three receive antennas.

4.3.5. Performance Comparison with Differential-STBC

The decoding complexity of STTC (measured by the number of trellis states at the decoder), increases exponentially as a function of both the diversity level and the transmission rate. A novel space time coding scheme for two transmit antennas is presented in [Alam1] by Alamouti that uses only linear processing the receiver. Space Time Block Coding (STBC) scheme introduced in [Taro3] generalizes Alamouti's scheme to an arbitrary number of transmit antennas and also has low receiver complexity. Absence of perfect channel estimates at the receiver motivates the use of differential schemes. A differential scheme is presented in [Taro4] that utilizes the basic structure of a STB code. This scheme retains the STBC property of low receiver complexity.

Figure 4-23 (from [Taro4]) compares the performance of this scheme with coherent STBC. The simulated system employs two transmit antennas. It is observed that the D-STBC scheme exploits the diversity offered by the system but incurs a 3dB loss in performance due to non-coherent detection. Both Differential STBC and STTC are thus seen to be capable of exploiting the diversity advantage offered by multiple transmit antenna systems. D-STBC schemes have the advantage of requiring a simple receiver structure that does not need to implement channel estimation. But STTC schemes are still valuable as they (unlike STBC schemes) offer coding gain in addition to the diversity gain. It is observed in the earlier sections that channel estimation errors cause a loss in coding gain. Hence it becomes important to analyze the performance of STTC in the presence of CEE as compared to the D-STBC scheme.

Figure 4-24 plots the performance the *CYV* STTC code and the D-STBC scheme in a two transmit and two receive antenna system. Channel estimation is done by using training sequences of length eight, two and one which result in appropriately increasing variance of the channel estimation error. It is seen that the performance of STTC degrades substantially when a training sequence of length one is used as a training sequence of length one is not sufficient to train a two transmit antenna system (Section 3.2.1). It is observed that in all the other cases, STTC outperforms the D-STBC scheme. Therefore in the presence of adequate training, STTC is better than D-STBC.

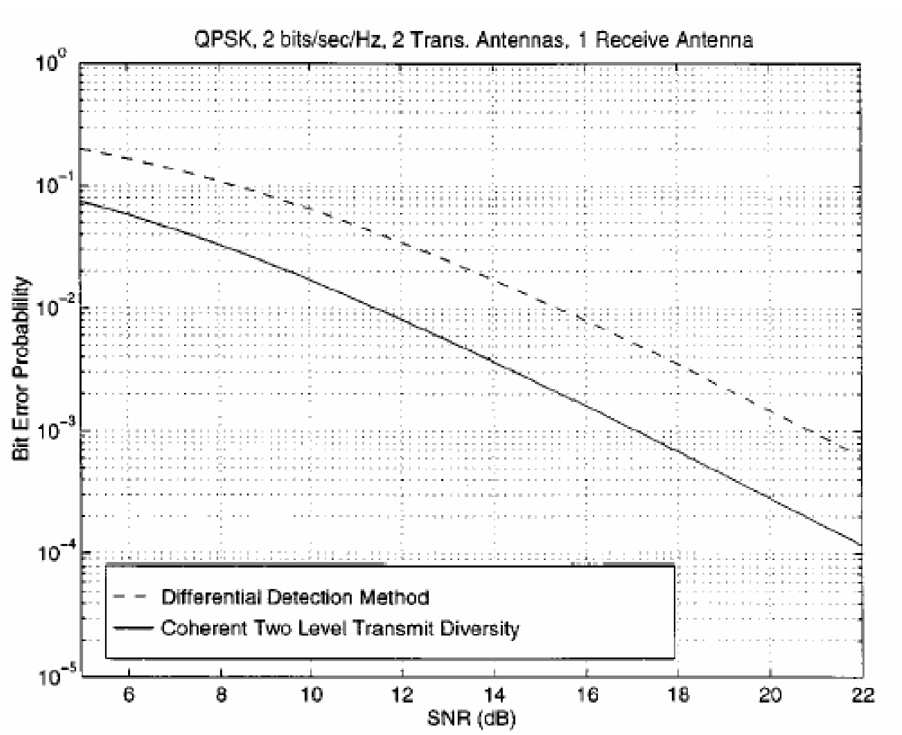

Figure 4-23: [Taro4] Performance of 2Tx-2Rx D-STBC scheme with QPSK over quasi-static channel

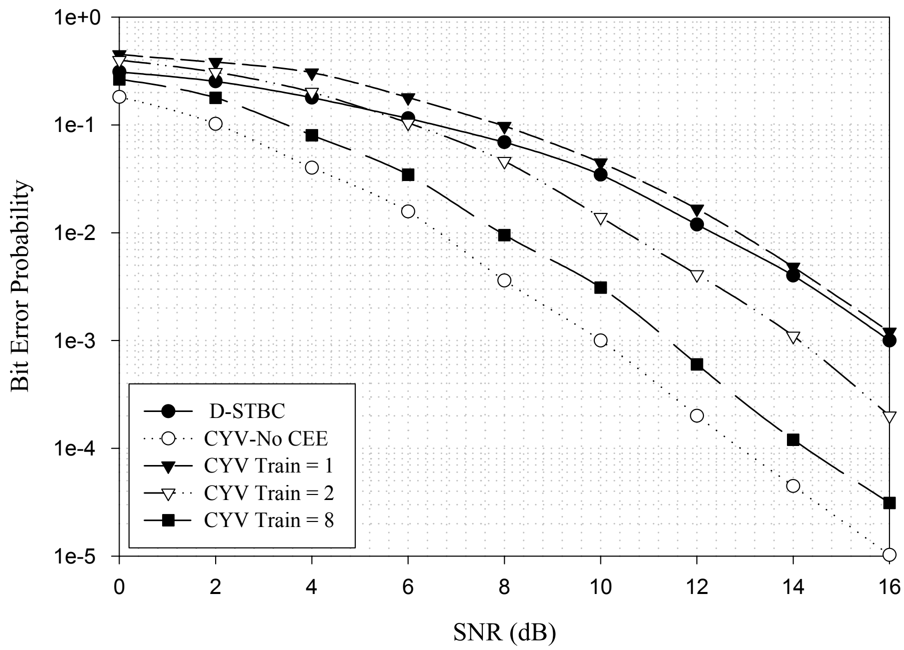

Figure 4-24: Performance comparison of D-STBC with 2Tx-2Rx CYV Code over a quasi-static channel and different training lengths

4.4. Chapter Summary

The effect of channel estimation errors on the performance of STTC in flat fading channels is studied in this chapter. The design criteria for STTC are re-validated in the presence of channel estimation errors. The analysis in [Taro2] showed the validity of the TSC-RD criteria for quasi-static flat fading channels, in the presence of estimation errors. This analysis is extended in this chapter to show that the TSC-DP criteria for fast fading channels and the trace criterion for systems with a large diversity order also hold in the presence of channel estimation errors (CEEs). It is shown that the diversity advantage offered by these criteria is maintained even in the presence of CEEs.

An exact expression for the pair-wise error probability of STTC in the presence of channel estimation errors is derived in this chapter. This expression can be used to find an approximate measure of the bit error probability of STTC codes in the presence of estimation errors of a specified variance.

The performance of different coding schemes and system configurations is studied and compared in the presence of CEEs. The CYV code (designed according to the trace criterion) performs better than the TSC code (designed according to the TSC-RD criterion) in the presence of perfect channel estimates as seen in the previous chapter. This performance advantage is maintained when optimal training is used to estimate the channel. But in general, the degradation due to channel estimation errors is larger for the CYV code. Hence in the presence of large channel estimation errors CYV performs worse than TSC code. Similarly, an increase in the number of transmit antennas is observed to increase the degradation due to estimation errors (even without compensating for the energy spent on training). In the presence large channel estimation errors, a system with a smaller number of transmit antennas performs better than a system with a larger number of transmit antennas. It is thus seen that CEEs could alter the performance behavior of codes and systems.

The optimal amount of training required for channel estimation for different MIMO system configurations is also calculated. The optimal amount of training increases with an increase in number of transmit antennas, a decrease in channel SNR and a decrease in coherence interval of the channel. These dependencies are used to show that at low channel coherence intervals, increasing the number of transmit antennas might not benefit the system due to increased training overhead. At low coherence intervals (when sufficient temporal fading is available) and with the option of using convolutional channel coding available, a system with no spatial diversity is able to perform better than a system with transmit diversity of two, employing a four state STTC. Similarly, performance of a system with two transmit antennas and a system with four transmit antennas employing corresponding STTCs are comparable at low coherence intervals negating the need for increased transmit diversity.

Hence the coherence interval of the channel over which a system is intended to operate and the channel estimation error variance (amount of training) possible within given system constraints should be important considerations in system design.

5. STTC in Frequency Selective Channels

Space Time Trellis Codes have been designed primarily for flat fading channels. Therefore, it is necessary to analyze these codes in frequency selective channels particularly in the presence of channel estimation errors, as CEE has been seen to alter the performance pattern of STTC codes. This chapter presents a discussion on the design criteria for STTC in frequency-selective channels. The trace criterion is extended to derive a new design criterion for frequency selective channels that improves the performance offered by the TSC-RD criteria in the presence of multi-path. This criterion is derived by modeling multi-paths as virtual antennas. This chapter also analyzes the impact of channel estimation errors on the performance of STTC in frequency selective channels.

5.1. Design Criteria in Frequency Selective Channels

5.1.1. TSC-RD Design Criteria

The TSC-RD criteria for STTC shows that diversity gain in a quasi-static flat fading Rayleigh channel can be maximized by maximizing the rank of the distance matrix between any two code-words and the coding gain, by maximizing the minimum determinant. This section follows the analysis in [Taro2] and examines the suitability of these criteria to frequency-selective channels. The analysis assumes a two-ray multi-path channel, but can easily be generalized to higher number of paths.

The MIMO system considered has N_t transmit and N_r receive antennas. The channel is represented by $N_t \times N_r$ sub-channels which are in turn modeled by the impulse function, $\alpha\delta(t) + \kappa\delta(t-\tau)$, where α and κ are independent complex Gaussian random variables with variances of 1 and σ_κ^2 respectively. $\delta(\ )$ is the Dirac-Delta function. The delay parameter τ is a random variable with a probability distribution function of $f(\tau)$. Let the signal transmitted from antenna i at time t be represented by $\sum_{k=0}^{L-1} \sqrt{E_s} c_i (k+1) u(t-kT_s)$ where, L is the length of a frame and $u(t)$ is a function time-limited to $[0, T_s]$, with total energy $\int_0^T |u(t)|^2 \, dt = 1$. The channel coefficients $\alpha_{i,j}$ are assumed to be constant for the duration of a frame and to vary between consecutive frames. The received signal at the j^{th} antenna at time t is,

$$r_j(t) = \sqrt{E_s} \sum_{k=0}^{L-1} \sum_{i=1}^{N_t} \alpha_{i,j} c_i (k+1) u(t-kT_s) + \kappa_{i,j} c_i (k+1) u\left(t-kT_s-\tau_{i,j}\right) + \eta(t) \qquad \text{5-1}$$

Consider that the transmitted code-word is

$c = c_1(1)c_2(1)\ldots c_{N_t}(1)c_1(2)c_2(2)\ldots c_{N_t}(2)\ldots c_1(L)c_2(L)\ldots c_{N_t}(L)$. The decoder makes an error if it decides erroneously in favor of a code-word
$e = e_1(1)e_2(1)\ldots e_{N_t}(1)e_1(2)e_2(2)\ldots e_{N_t}(2)\ldots e_1(L)e_2(L)\ldots e_{N_t}(L)$.

From the standard approximation for the tail of the Gaussian distribution, the upper bound on the probability of the decoder making an error is given by,

$$P\left(c \to e \mid \alpha_{i,j}, \kappa_{i,j}, \tau_{i,j}, i = 1,2,\ldots,N_t, j = 1,2,\ldots,N_r\right) \le \prod_{j=1}^{N_r} \exp\left(\frac{-\Lambda_j \mathrm{H}_j \Lambda_j^* E_s}{4N_0}\right) \qquad \text{5-2}$$

where,

$$\Lambda_j = \left(\alpha_{1,j}, \ldots, \alpha_{N_t,j}, \frac{\kappa_{1,j}}{\sigma_k}, \ldots, \frac{\kappa_{N_t,j}}{\sigma_k}\right)$$

$$\mathrm{H}_j = \begin{pmatrix} A(c,e) & \sigma_k D(c,e,\tau) \\ \sigma_k D(c,e,\tau)^* & \sigma_k^2 A(c,e) \end{pmatrix}$$

$A(c,e)$ is as defined in Section 2.2.1 and $D(c,e,\tau)$ is any matrix which depends only on c, e *and* τ. An analysis similar to Section 2.2.1 can be followed to arrive at,

$$P(c \to e) \le \prod_{j=1}^{N_r} \left(\frac{1}{\prod_{k=1}^{2N_t} \left(1 + \lambda_k\left(\mathrm{H}_j\right) \frac{E_s}{4N_0}\right)} \right) \qquad \text{5-3}$$

where, $\lambda_k\left(\mathrm{H}_j\right), k = 1,2,\ldots,2N_t$, are the eigen values of H_j counting multiplicities.
It can be proved that the rank of H_j is at least equal to the rank of $A(c,e)$ (Section 2.2.1). Hence a minimum diversity advantage of $N_r rank\left(A(c,e)\right)$ will be achieved in a multi-path channel. Consequently, STC codes designed according to the TSC criteria for Quasi-static channels will continue to perform well in frequency selective channels. The achievable diversity is not reduced in a frequency selective environment.

The performance of a 4-state TSC-RD code in a two-tap multi-path channel is shown in Figure 5-1. The system has two transmit antennas and one or two receive antennas. Multi-path arrivals are assumed to occur at integral multiples of the symbol interval and the channels are assumed to be quasi-static. The transmitted symbols are decoded by using the MLSE Combined Trellis Equalizer and Decoder (CT-ED) [Heik1], which combines the equalizer and decoder trellis into a super trellis structure. Consider N_{st} to be the number of states of the channel code. Then, the number of states of the super-trellis is given by $N_{st}M^{L_c-1}$, where M is the size of the symbol constellation used. Hence each state of the super trellis corresponds to a possible state of the code and L_c-1 previously transmitted. The branch metric is given by $\sum_{j=1}^{N_r}\left|r_j(t)-\sum_{i=1}^{N_t}\sum_{l=1}^{L_c}\alpha_{ij}(l)e_i(t-l)\right|^2$, where L_c is the number of multi-paths in the channel. The decoder chooses the sequence that minimizes the metric over the length of the frame. However, CT-ED might not be an appropriate choice when the number of multi-paths in the channel is higher as its complexity grows exponentially with channel tap length. A sub-optimal detector is proposed in [Heik1], which uses a reduced state-space method with decision feedback. Complexity can also be reduced by exploiting the structure of STTC as described in [Nagu1]. Performance can be improved in multi-path channels by using a channel interleaver. In this case, a Maximum Aposteriori Probability Equalizer/Detector (MAP-ED) [Bauc1] is required at the receiver. A reduced complexity version of the MAP-ED is presented in [Frag4].

It is seen from Figure 5-1, that diversity advantage does not decrease in frequency selective channels. The performance of the STTC scheme in fact improves because of the additional diversity offered by the multi-path components. But the TSC criteria might not be able to fully exploit the additional diversity offered by multiple paths in each sub-channel. Fitz et al in [Youj1] present a design methodology aimed at utilizing this additional available diversity. The design criteria are referred to as the YFT criteria and are presented in the next section.

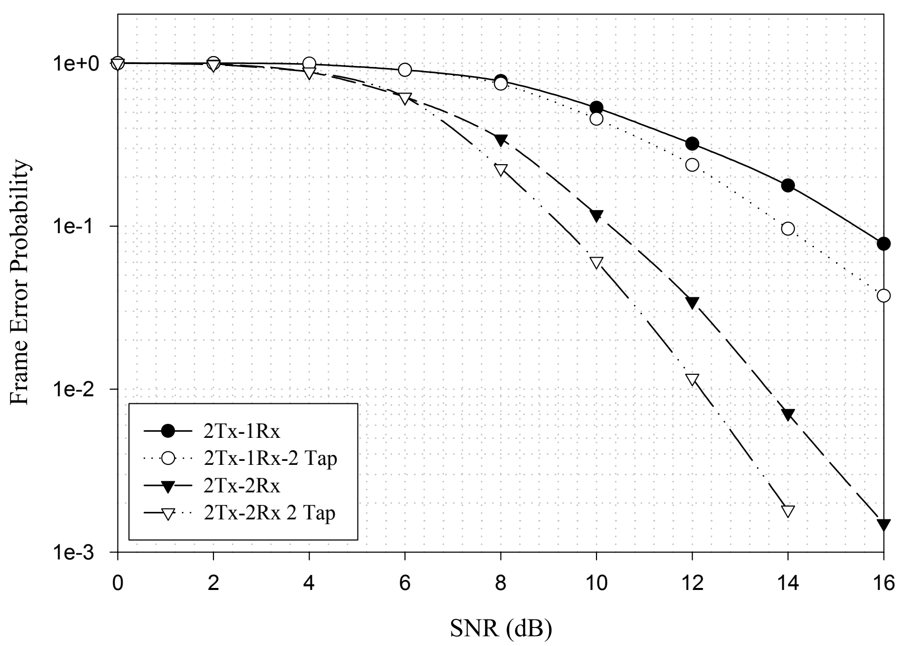

Figure 5-1: Performance of TSC code in a two-tap frequency selective channel

5.1.2. Modified RD Design Criteria

Code design criteria are formulated for channels with Inter Symbol Interference (ISI) in [Youj1] by introducing the concept of virtual antennas. The criteria guarantee full diversity in frequency selective channels. The code design takes into consideration two scenarios: with and without a channel interleaver.

The signal model used is described below. A $N_t \times N_r$ multi antenna system with L_c multi-paths per sub channel is considered. The multi-paths are assumed to arrive at multiples of the symbol time (The analysis in the previous section takes into consideration fractional delays.). From previous definitions, the received signal can be represented as,

$$r_j(t) = \sum_{l=0}^{Lc} \sum_{i=1}^{N_t} c_i(t-l) a_{i,j}(t,l) + \eta_j(t),\ j = 1,2,...,N_r,\ t = 1,2,...,L+Lc-1 \qquad 5\text{-}4$$

Further defining, $R_j = \begin{bmatrix} r_j(1) & r_j(2) & \dots & r_j(L+Lc-1) \end{bmatrix}^T$,

$X = \begin{bmatrix} X_1(0) & X_2(0) & \dots & X_{N_t}(0) & \dots & X_1(Lc-1) & X_2(Lc-1) & \dots & X_{N_t}(Lc-1) \end{bmatrix}$,

$X_i(l) = diag\left(\begin{bmatrix} 0_{l*1} & c_i(1) & c_i(2) & \dots & c_i(L) & 0_{(Lc-1-l)*1} \end{bmatrix}\right)$,

$a_j = \begin{bmatrix} a_{1,j}(0) & a_{2,j}(0) & \dots & a_{N_t,j}(0) & \dots & a_{1,j}(L_c-1) & a_{2,j}(L_c-1) & \dots & a_{N_t,j}(L_c-1) \end{bmatrix}^T$,

$a_{i,j}(l) = \begin{bmatrix} a_{i,j}(1,l) & a_{i,j}(2,l) & \dots & a_{i,j}(L+Lc-1,l) \end{bmatrix}^T$

and $\eta_j = \begin{bmatrix} \eta_j(1) & \eta_j(2) & \dots & \eta_j(L+Lc-1) \end{bmatrix}$, the received signal can be expressed in Matrix Notation as,

$$R_j = Xa_j + \eta_j \qquad 5\text{-}5$$

The code word vector X contains the code-word transmitted through the transmit antennas and zero-padded versions of the code-word transmitted through virtual antennas. The vector X will thus play the same role as the code-word matrix in flat fading channels. Let $Z = X^\alpha - X^\beta$ be the code difference matrix and $C_{cj} = E\left[\alpha_j \alpha_j^H\right]$ be the covariance matrix of channel coefficients.

$$C_{sj} = ZC_{cj}Z^H \qquad 5\text{-}6$$

The pair wise error probability can be upper bounded by

$$P\left(D^\alpha \to D^\beta\right) \le \frac{K}{\left(\frac{E_s}{N_o}\right)^{rN_r} \left(\prod_{i=1}^{r} \lambda_i\right)^{N_r}} \qquad 5\text{-}7$$

where, r is the number of non-zero Eigen values λ_i of C_{sj} and K is a constant. The product of non-zero Eigen values gives the coding gain.

If the multi-paths are assumed to be independent, the diversity advantage can be maximized by maximizing the minimum rank of the code difference matrix Z for any two code-words. The maximum achievable diversity gain is $N_t \times L_c \times N_r$ (compared to the maximum achievable diversity $N_t \times N_r$ by the TSC code). Hence this design criterion can lead to codes that exploit all the diversity possible in the multi-path channel. The coding gain can be maximized by maximizing the product of eigen values of the distance matrix taken over all pairs of distinct code-words. These criteria are thus seen to be an extension of the TSC-RD criteria for frequency selective channels. The algebraic Σ_0-rank theory [Youj2] is used to design full diversity STTC codes. By utilizing the concept of Virtual antennas, similar representations can be used for signals transmitted in frequency selective fading and flat fading channels. This lends itself to similar code design procedures (illustrated in [Safa1]). Some additional constraints are also placed on code design to ensure r-level receive diversity.

Theorem1: (without Channel Interleaver)
The code must have at least $2^{R(r-L_c)}$ states, where R is the transmission rate in bits/symbol and the constraint length is at least $r - L_c - 1$, where the constraint length is defined as minimum length of the error path of the original trellis code.

Theorem2: (with Channel Interleaver)
The code must have at least $2^{R\left(\frac{r}{L_c}-1\right)}$ states and the constraint length is at least $\frac{r}{L_c}$, where the constraint length is defined as minimum length of error path of the original trellis code.

These constraints follow from similar constraints in [Taro1] for code-design in flat fading channels. Tarokh et al showed in [Taro1] that the constraint length of an r-space-time trellis code is at least $r-1$ and that if b is the transmission rate, the trellis complexity of the code is at least $2^{b(r-1)}$.

For a given number of multi-paths, a code with a larger number of states is thus capable of exploiting the diversity advantage offered by a multi-path system. This is illustrated in Figure 5-2. The frequency selective channel is assumed to have two independent taps with equal powers. The 16-state code (presented in [Youj1] and denoted here by “YFT”) has a larger slope than the 4-state code illustrating the fact that it is able to exploit additional diversity in a frequency selective channel.

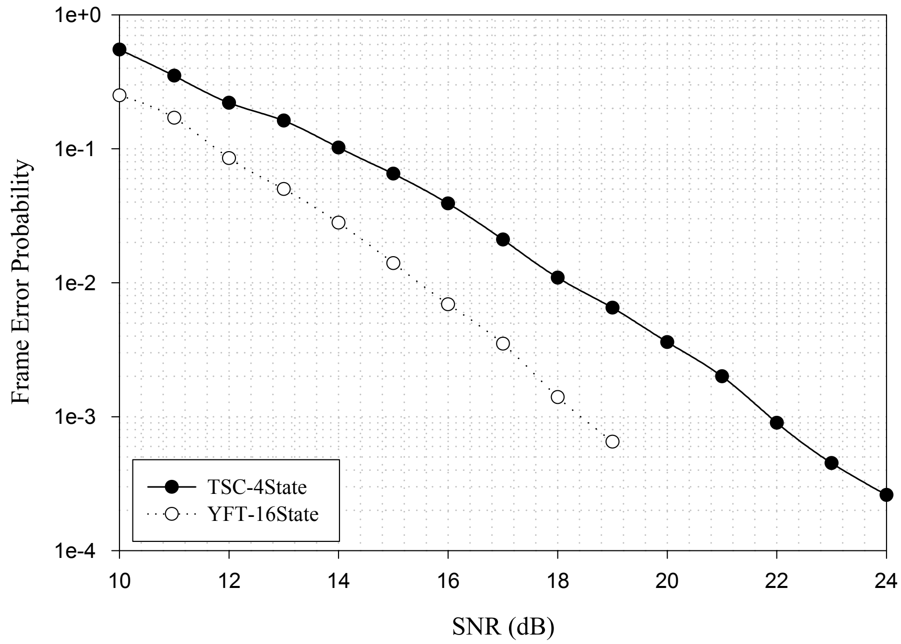

Figure 5-2: (Adapted from [Youj1]) Performance comparison of 16-state YFT code and 4-state TSC-code in a 2-tap multi-path channel

These code design criteria thus lead to codes that have a larger number of states, depending on the multi-path components in the channel. An increase in the number of states leads to more complexity especially in a frequency selective channel with equalization requirements. In the next section a new design criterion is developed for STTC in frequency selective channels that gives improved performance as compared to the TSC-RD codes but does not necessitate an increase in the number of states. The analysis follows results in [Chen1] and is similar to the CYV criterion. The CYV criterion is subject to the constraint that the product of the rank of the code and the number of receive antennas, be greater than three. But by utilizing the concept of virtual antennas introduced in this section, the existence of even a single multi-path eliminates this requirement.

5.1.3. New Design Criteria

The design of STTC presented in Section 2.2.1 is based on the maximization of the minimum rank and minimum determinant of the distance matrices. It is shown here, that in the presence of multi-path, the design of STTC codes with maximum coding gain is governed by the minimum trace of the distance matrices, or the Euclidean distance between any two code words over all transmit antennas. The trace design criterion for flat fading channels is derived by minimizing pair-wise error probability. This criteria is extended to for multi-path channels by introducing “Virtual Antennas” [Chen1].

The probability of transmitting a signal *c* and deciding in favor of a signal *e* in a quasi-static multi-path channel, assuming ideal channel state information, is given by,

$$p\left(c \to e \mid \alpha_{ij}(l), i=1,2,\dots N_t, j=1,2,\dots N_r, l=0,1,\dots l=L_c-1\right) \le \exp\left(-d^2(c,e)\frac{E_s}{4N_0}\right)$$

5-8

where $N_0/2$ is the noise variance per dimension and

$$d^2(c,e)=\sum_{j=1}^{N_r}\sum_{t=1}^{L}\left|\sum_{l=0}^{L_c-1}\sum_{i=1}^{N_t}\alpha_{ij}(l)\left(c_i(t-l)-e_i(t-l)\right)\right|^2 \quad 5\text{-}9$$

Setting,

$$\Omega_j=\left[a_{1j}(0) \quad a_{2j}(0) \quad \dots \quad a_{N_t j}(0) \quad \dots \quad a_{1j}(L_c-1) \quad a_{2j}(L_c-1) \quad \dots \quad a_{N_t j}(L_c-1)\right]$$
$$B(c,e)=\left[X_1(0) \quad X_2(0) \quad \dots \quad X_{N_t}(0) \quad \dots \quad X_1(Lc-1) \quad X_2(Lc-1) \quad \dots \quad X_{N_t}(Lc-1)\right]^T$$

$$X_i(l)=\left[0_{l*1} \quad e_i(1)-c_i(1) \quad e_i(2)-c_i(2) \quad \dots \quad e_i(L)-c_i(L) \quad 0_{(Lc-1-l)*1}\right]$$
$$A(c,e)=B(c,e)B^*(c,e)$$
$$d^2(c,e)=\sum_{j=1}^{Nr}\Omega_j A(c,e)\Omega_j^*$$

The code-word in the above case can be viewed as being sent through virtual antennas. There exists a real diagonal matrix D, whose diagonal elements are the Eigen values of A, such that,

$$VA(c,e)V^*=D \quad 5\text{-}10$$

where the rows of V are the eigen vectors of A. Let,

$$\left[\beta_{1,j} \quad \beta_{2,j} \quad \dots \quad \beta_{(N_t*L_c)-1,j} \quad \beta_{N_t*L_c,j}\right]=\Omega_j V^* \quad 5\text{-}11$$

$$\Omega_j A(c,e)\Omega_j^*=\sum_{i=1}^{N_t*L_c}\lambda_i\left|\beta_{ij}\right|^2 \quad 5\text{-}12$$

Thus the conditional pair-wise error probability can be expressed as,

$$p(c \to e \mid \alpha) \le \exp\left(-\sum_{j=1}^{N_r}\sum_{i=1}^{N_t*L_c}\lambda_i\left|\beta_{ij}\right|^2\frac{E_s}{4N_0}\right) \quad 5\text{-}13$$

$|\beta ij|^2$ follows the central chi square distribution. Its mean and variance is equal to 1.

Let r be the rank of the matrix A. For a large rN_r (>3) value, according to Central Limit Theorem, the expression $\sum_{j=1}^{N_r}\sum_{i=1}^{N_t}\lambda_i\left|\beta_{ij}\right|^2$ approaches a Gaussian random variable D with mean

$$\mu_D = N_r\sum_{i=1}^{N_t}\lambda_i \qquad 5\text{-}14$$

and variance

$$\sigma_D^2 = N_r\sum_{i=1}^{N_t}\lambda_i^2 \qquad 5\text{-}15$$

Let the code matrix be defined as,

$$Bt(c,e) = \begin{bmatrix} e_1(1)-c_1(1) & \cdots & \cdots & e_1(L)-c_1(L) \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ e_{N_t}(1)-c_{N_t}(1) & \cdots & \cdots & e_{N_t}(L)-c_{N_t}(L) \end{bmatrix} \qquad 5\text{-}16$$

Let r_t be the rank of the matrix Bt. Then, the rank of A will be r_tL_c. Thus in the presence of a full rank code and multi-path, rN_r is seen to be greater than 3 in all cases. The upper bound on the average probability of error can be obtained by averaging w. r. t the probability distribution function of D.

$$p(c\rightarrow e)\le\int_0^\infty \exp\left(-\frac{E_s}{4N_0}D\right)p(D)dD$$

$$p(c\rightarrow e)\le\exp\left(\frac{1}{2}\left(\frac{E_s}{4N_0}\right)^2\sigma_D^2-\frac{E_s}{4N_0}\mu_D\right)Q\left(\frac{\frac{E_s}{4N_0}\sigma_D^2-\mu_D}{\sigma_D}\right) \qquad 5\text{-}17$$

By using, $Q(x)\le\frac{1}{2}e^{-\frac{x^2}{2}}\ x\ge 0$,

$$p(c\rightarrow e)\le\frac{1}{2}\exp\left(-N_r\frac{E_s}{4N_0}\sum_{i=1}^{N_t}\lambda_i\right) \qquad 5\text{-}18$$

The pair-wise error probability can be minimized by maximizing the sum of the Eigen values of the matrix $A(c,e)$. For a square matrix the sum of the Eigen values equals the trace of the matrix. The trace of matrix $A(c,e)$ can be written as

$$\mathrm{tr(A)=L_c}\sum_{i=1}^{N_t}\sum_{t=1}^{L}\left|\mathrm{e}_i\left(t\right)-c_i\left(t\right)\right|^2 \tag{5-19}$$

Thus the pair wise error probability can be minimized if the Euclidean distance between code-words is maximized.

The CYV code introduced previously maximizes the Euclidean distance and is used to demonstrate the performance of the new scheme in Figure 5-3 . It is seen that in the frequency selective channel additional improvement is obtained as compared to the flat fading channel by exploiting the multi-path components.

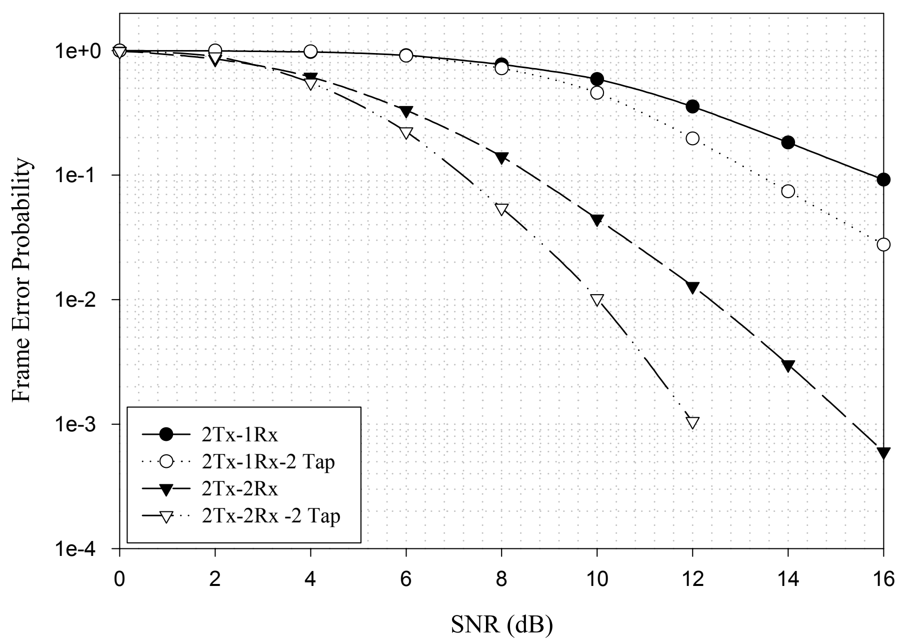

Figure 5-3: Performance of the new scheme in a two-tap frequency selective channel

Figure 5-4 and Figure 5-5 compare the performance of the TSC scheme and the new scheme. The system has two transmit antennas and one receive antenna. The TSC code and the CYV code have full ranks and equal minimum determinant values of 4. However, the minimum trace of the CYV code is 10 and that of the TSC code is 4. From the analysis it is observed that when the diversity order is greater than 3 or in the presence of multi-path, the dominant performance criterion is the maximization of the minimum trace between two code-words. When the diversity order is less than 3 and in the absence of multi-path, the dominant performance criterion is the minimum rank criterion. The simulations reflect this.

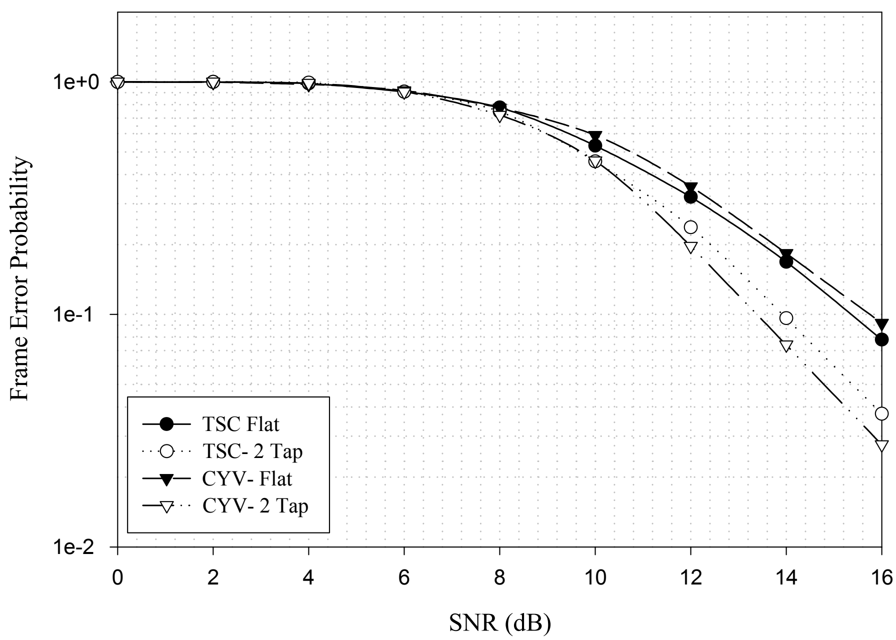

Figure 5-4: FER performance comparison of 2Tx-1Rx TSC code and the new scheme in multi-path channel

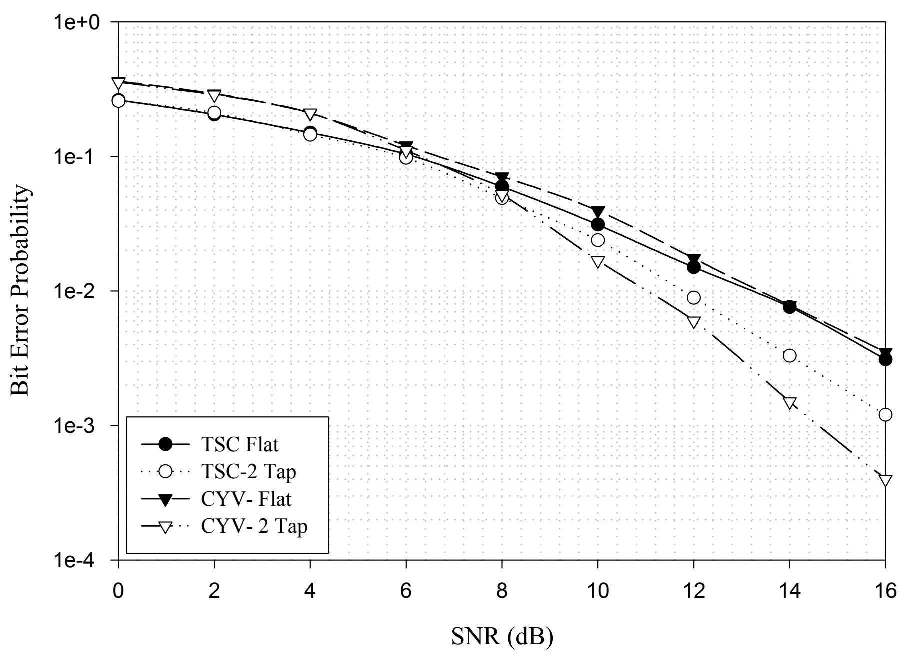

Figure 5-5: BER performance comparison of 2Tx-1Rx TSC code and the new scheme in multi-path channel

It can be seen from the figures that in the absence of multi-path the TSC and CYV codes perform almost similarly in the diversity-2 scheme considered above, as both have equal rank and determinant values. But in the presence of multi-path, the CYV code performs much better than the TSC code, as the CYV Code has a larger minimum trace than the TSC code. The above analysis and the theoretical results from before show that in the presence of multi-path the dominant performance criterion is the minimum trace criterion.

5.2. Effect of Channel Estimation Errors on STTC Performance

5.2.1. Optimal Amount of Training

An expression for the available capacity of a MIMO channel utilizing training and assuming the training symbols and data symbols to be transmitted at the same power is derived in Chapter 3 (Section 3.3.2 and Equation 3-29). The optimal training required for a specific MIMO system over a frequency selective channel and for a given SNR can be evaluated from this expression by calculating the length that maximizes the capacity. Figure 5-6 plots the optimal amount of training required by two transmit and two receive antenna system operating over a two-tap channel for different channel SNRs and different coherence times. Figure 5-7 shows the simulated performance of a two-transmit, two-receive antenna system employing the TSC code, operating over a two-tap channel with coherence interval of 120 symbols and an SNR of 12dB, for various training lengths. The energy spent on training is compensated by a reduction in the energy available for data transmission. The performance is seen to be the good for a range of training lengths from nine to twenty and to degrade for shorter or longer training lengths. It can be observed from Figure 5-6, that the optimal amount of training for a two-transmit, two receive antenna system over a two tap channel with a coherence interval of 120 and a channel SNR of 12dB is nine. The optimal length results thus give a conservative estimate of the training length that maximizes the performance of a system.

Frequency selective channels have the additional constraint that the length of the training sequence cannot be less than the product of the number of multi-paths (L_c) and the number of transmit antennas (Equation 5-20).

$$T_t \geq N_t L_c \tag{5-20}$$

This minimum length requirement considerably increases the amount of training required at small channel SNRs and coherence times compared to the training required in systems operating over flat fading channels. This can be observed by comparing Figure 5-6 with Figure 4-7. The proportion of the required training to the number of data symbols in a frame is seen to be significantly increased at small coherence intervals in frequency selective channels. But at larger coherence times the performance improvement caused by multi-path channels (because of the additional available diversity in time) is able to offset the demand for increased training. These observations are reiterated in Figure 5-8 which shows the optimal training length for the two-transmit, two-receive antenna MIMO system over a frequency selective channel with three taps. The performance improvement due to the additional available multi-path causes a decrease in the required training at large coherence intervals. But the optimal length of training increases at small coherence intervals, due to an increase in the minimum required training.

Expression 5-20 also shows that increasing the number of transmit antennas would cause a considerable increase in the required amount of training in frequency selective channels (larger than the corresponding increase in flat-fading channels). This is illustrated in Figure 5-9 which plots the optimal required training for a system with four transmit antenna over a two-tap frequency selective channel. This large increase in training overhead could offset any diversity advantage offered by increasing the number of transmit antennas in frequency selective channels.

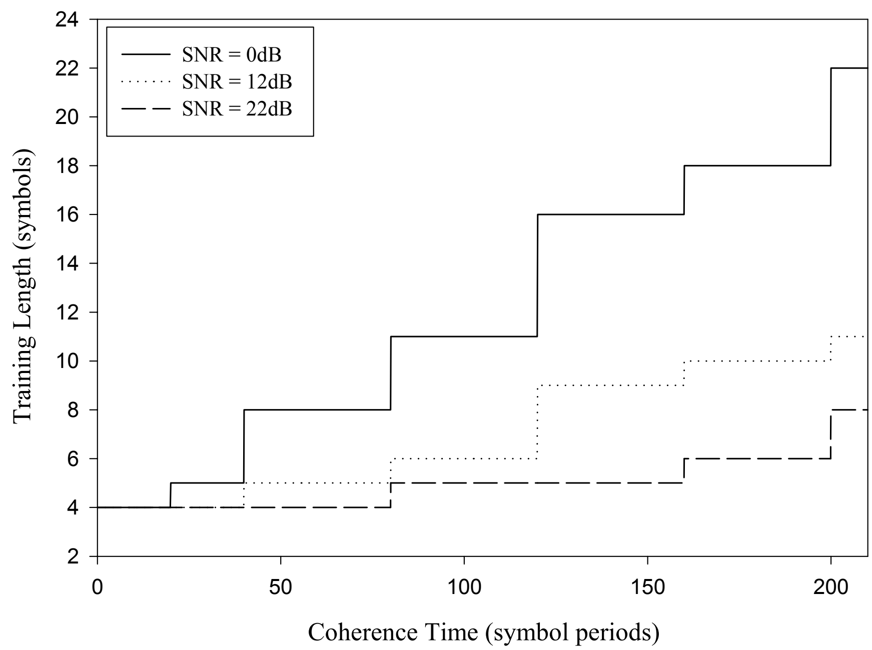

Figure 5-6: Optimal training length for 2Tx-2Rx MIMO system over multi-path channel with 2-taps

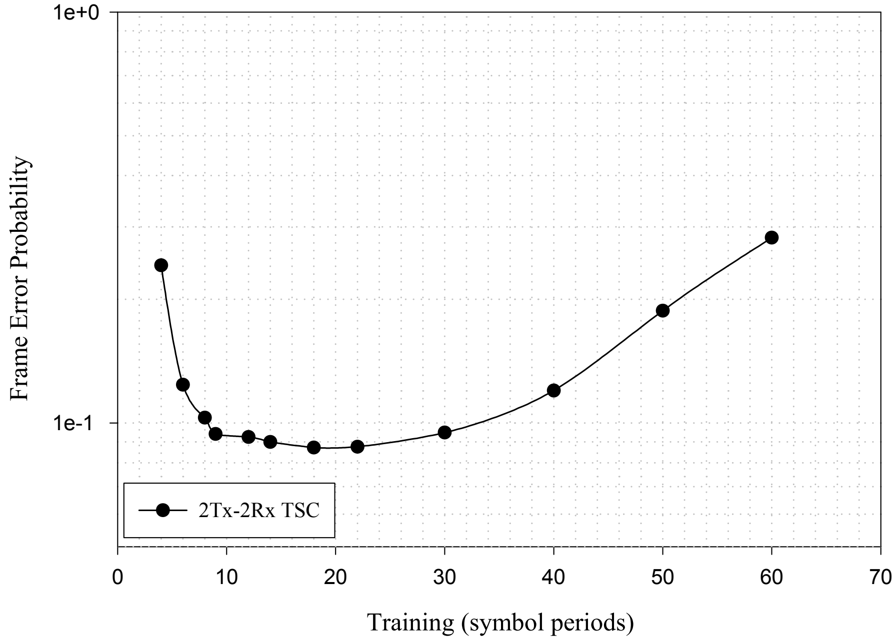

Figure 5-7: Performance comparison with varying length of training, for a 2Tx-2Rx TSC Code at 12dB channel SNR and over a two-tap channel

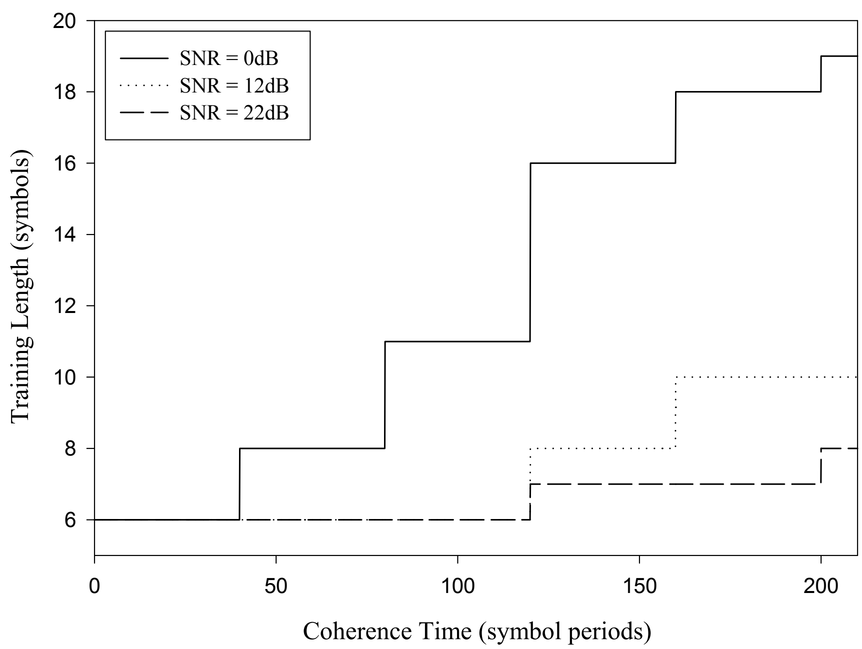

Figure 5-8: Optimal training length for 2Tx-2Rx MIMO system over multi-path channel with 3-taps

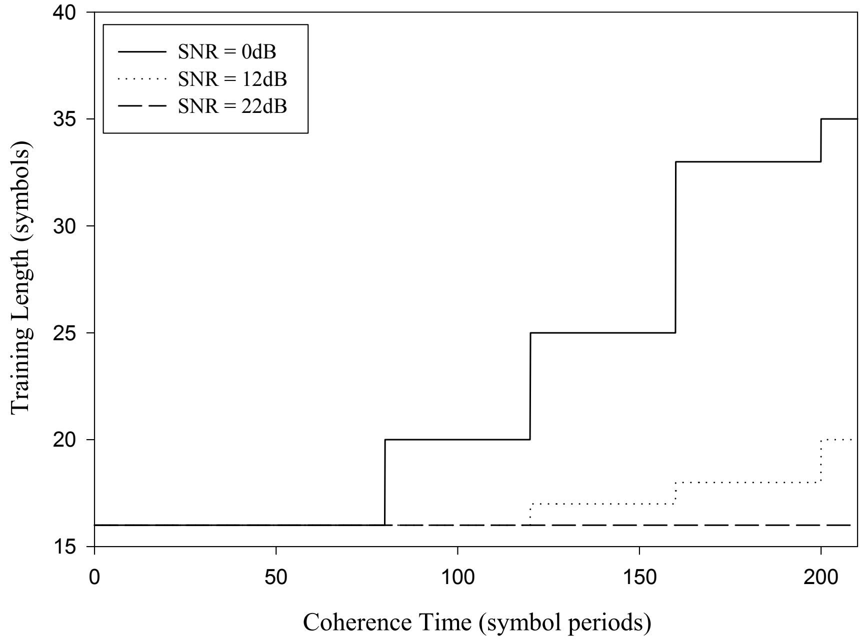

Figure 5-9: Optimal training length for 8Tx-2Rx MIMO system over multi-path channel with 3-taps

5.2.2. Performance Degradation due to CEE

The performance loss caused by imperfect channel estimation on an STTC system operating over a frequency selective channel is studied. Figure 5-10 shows the performance of a two-transmit, two-receive antenna scheme over a two tap channel selective in the presence of CEE. A training sequence of length eight is used to estimate the channel. The training sequence length is twice the product of the number of transmit antennas and the number of multi-path components and hence sufficient to estimate the channel. It is observed that CEE causes a loss in coding gain of around 2dB but does not affect the diversity advantage offered by the code. The loss in performance due to CEE in a frequency selective channel is seen to be much larger than the corresponding performance loss in a flat fading channel. The CYV code, in the presence of multi-paths and channel estimation errors and at low channel SNRs, is seen to perform worse than the CEE affected code in a flat fading channel. This performance degradation is despite the additional time diversity offered by the multi-path channel. Estimation errors are thus seen to have a larger impact on performance in multi-path channels and sufficient care must be taken during system design such that they do not alter the required performance behavior.

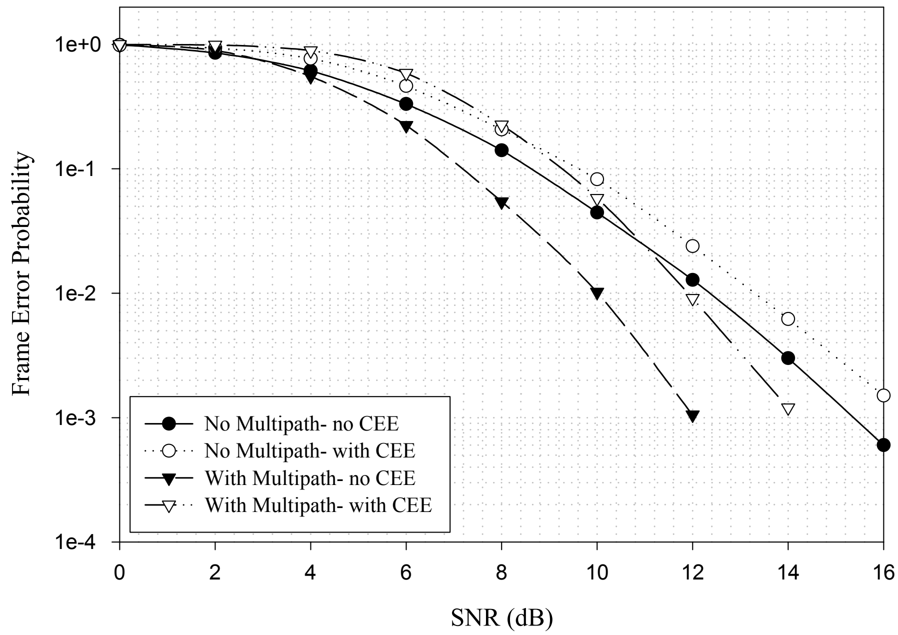

Figure 5-10: Performance of 2TX-2Rx CYV code in a flat fading and 2-Tap multi-path channel with training length of 8

Figure 5-11 and Figure 5-12 compare the performance of the TSC and CYV STTC in the presence of channel estimation errors and multi-paths. The system has two transmit antennas and one receive antenna. A training sequence of length eight is used to estimate the two tap multi-path channel. It is seen that the degradation due to CEE is larger for the CYV code than the TSC code. This is consistent with results in (Section 4.3.2), that shows that the CYV code degrades more, due to a larger value of the minimum trace of the code-word matrix. However, it still performs better than the TSC code when training of length eight is used. Figure 5-13 compares the performance of two codes for optimal training lengths over a range of coherence times. It is seen that the performance advantage of the CYV code over the TSC code is consistent.

Figure 5-14 shows the performance of the two codes for varying values of channel estimation error SNR and reiterates the fact that the performance degradation caused by CEE is more pronounced for the CYV code. This leads to the CYV code performing worse than the TSC code for very large channel estimation errors, as can be seen from the figure. It is also noted by examining Figure 5-14 and Figure 4-15 that the crossover point occurs for a larger value of the channel estimation SNR.

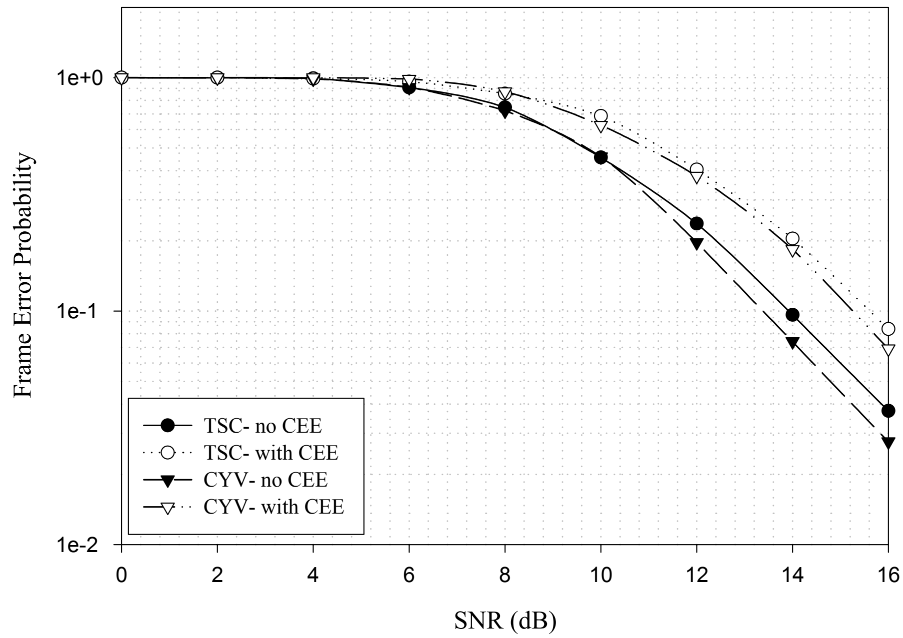

Figure 5-11: FEP Performance Comparison of 2Tx-1Rx TSC code and the new scheme with training length eight

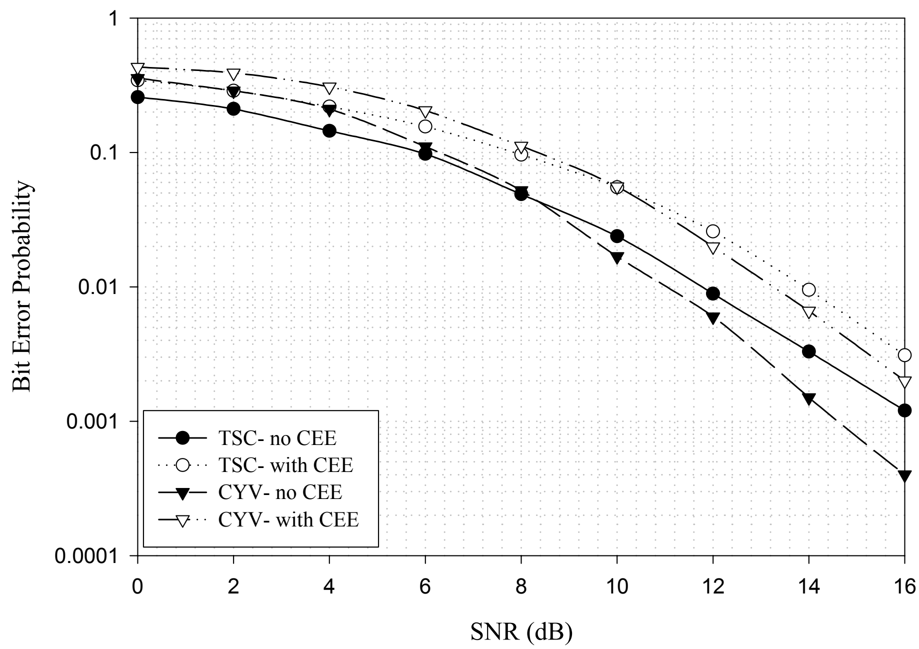

Figure 5-12: BER Performance Comparison of 2Tx-1Rx TSC code and the new scheme with training length eight

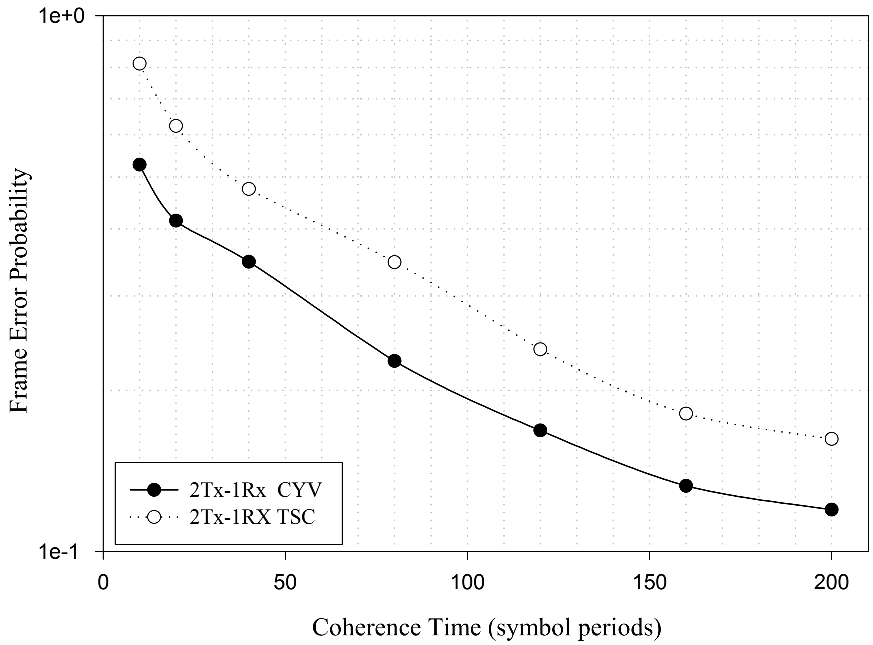

Figure 5-13: Comparison of 2TX-1Rx TSC and CYV code with optimal training and channel SNR of 16dB

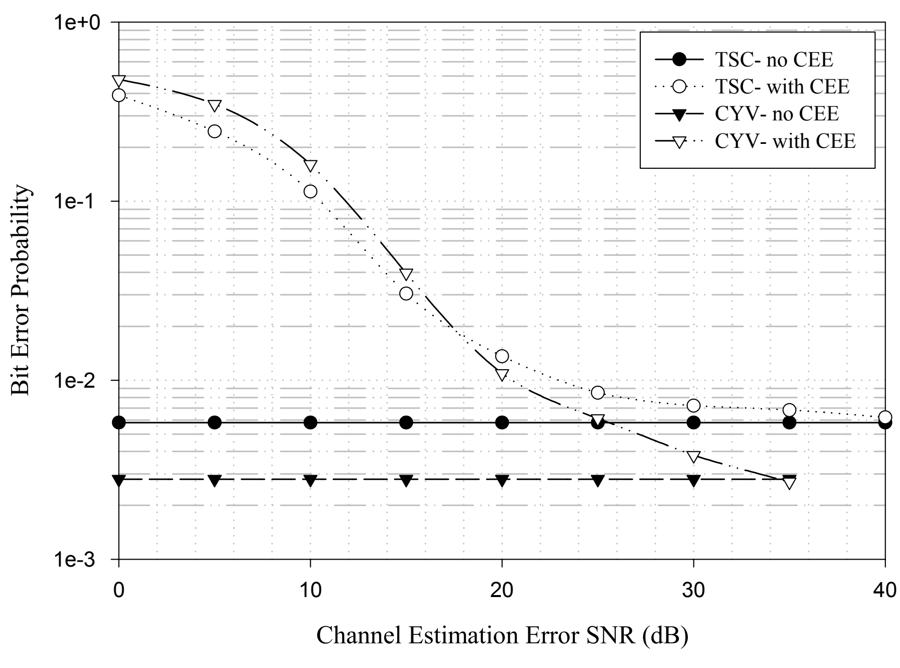

Figure 5-14: Comparison of 2Tx-2Rx TSC and CYV in 2-Tap channel with SNR of 10dB and varying channel estimation SNR

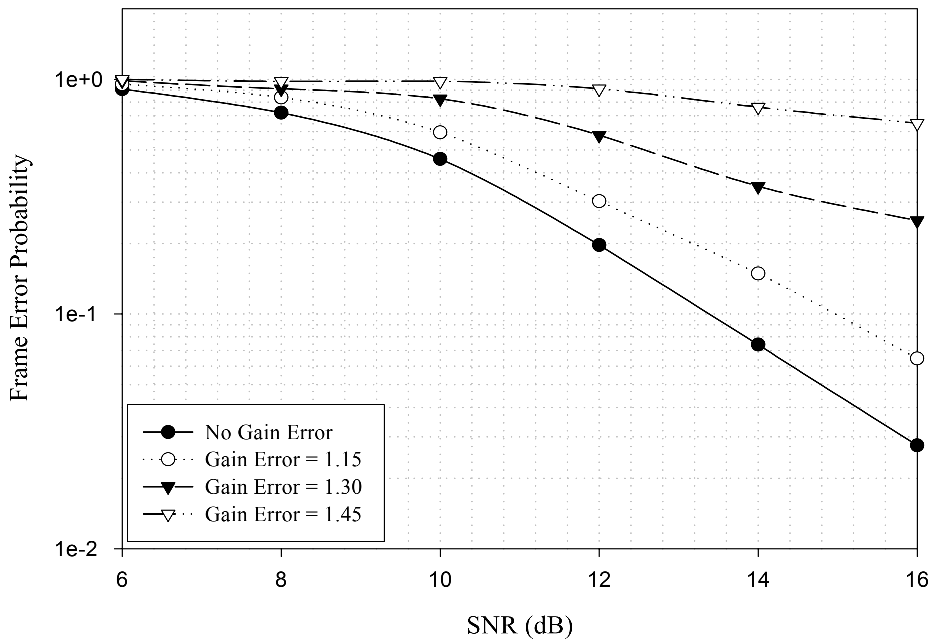

Figure 5-15: Performance of 2Tx-1Rx CYV over 2 Tap channel for varying gain errors of the channel estimates

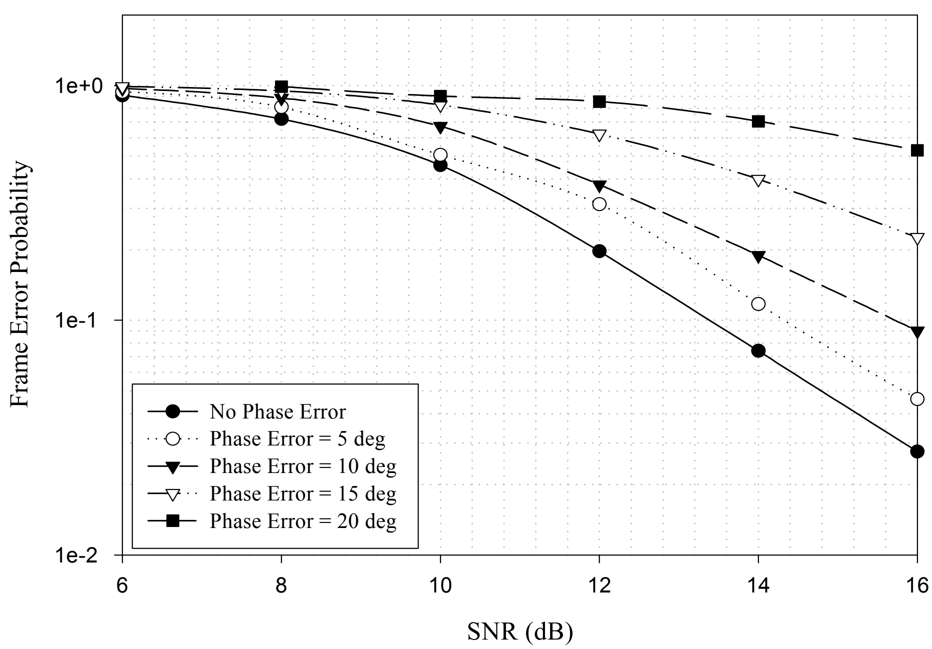

Figure 5-16: Performance of 2Tx-1Rx CYV over 2 Tap channel for varying phase errors of the channel estimates

The performance of the CYV STTC scheme in frequency selective channels is analyzed in the presence of gain errors (Figure 5-15) and phase errors (Figure 5-16) in channel estimation. It is seen that the STTC scheme is robust to gain errors of up to 30% (the gain of a perfect channel estimate is assumed to be one in the graphs.). The diversity offered the code does not decrease for gain errors less than 30% in magnitude. The code is also seen to be robust to phase errors of up to 10 degrees and deterioration in performance is seen for larger phase errors.

5.2.3. Performance Sensitivity to Coherence Time

The capacitive analysis in chapter three shows the proportion of training in a frame increases with a decrease in the coherence time of the channel. This leads to a decrease in the capacity of the channel for small coherence times. In a frequency selective channel, this dependence is pronounced. Frequency selective channels require larger training sequences than flat fading channel especially at low coherence times. Hence the loss in capacity is increased.

Figure 5-18 shows the performance of different transmit antenna systems with the channel assumed to be estimated by training sequences of optimal length. The energy spent on training for a particular scheme is compensated for by reducing the energy available for transmitting data symbols. Figure 5-17 shows the performance of these schemes assuming perfect channel estimation. Comparing Figure 5-17 and Figure 5-18, it can be seen that schemes with a larger number of transmit antennas experience larger degradation in performance. This can be attributed to the comparatively larger training required for systems with larger number of transmit antennas. The increase in training caused by the increase in transmit antennas is also compounded by the number of multi-paths in frequency selective channels (expression 5-20). This is reiterated by comparing Figure 5-19 and Figure 5-20.

It is observed from Figure 5-18 that in the presence of optimal training, the performance of the single transmit antenna system using a four state code, is comparable with the performance of the two-transmit antenna system employing a four state CYV code over the range of lengths of the coherence interval. It is also seen that the a single transmit antenna system employing an 8-state code betters the performance of the CYV code over the entire range of the coherence intervals considered when training is employed to estimate the channel. This is despite the fact that the single antenna eight state code performs worse than the two transmit antenna system when perfect channel estimates are assumed (Figure 5-17). It can also be observed from Figure 5-20 that the three and four-transmit antenna systems start to perform worse than the two-antenna systems over the entire range of the coherence interval when training is used to estimate the channel. In flat fading systems such behavior is noticed only for small coherence intervals. These results are a consequence of the comparatively larger training overhead associated with larger transmit diversity systems in frequency selective channels.

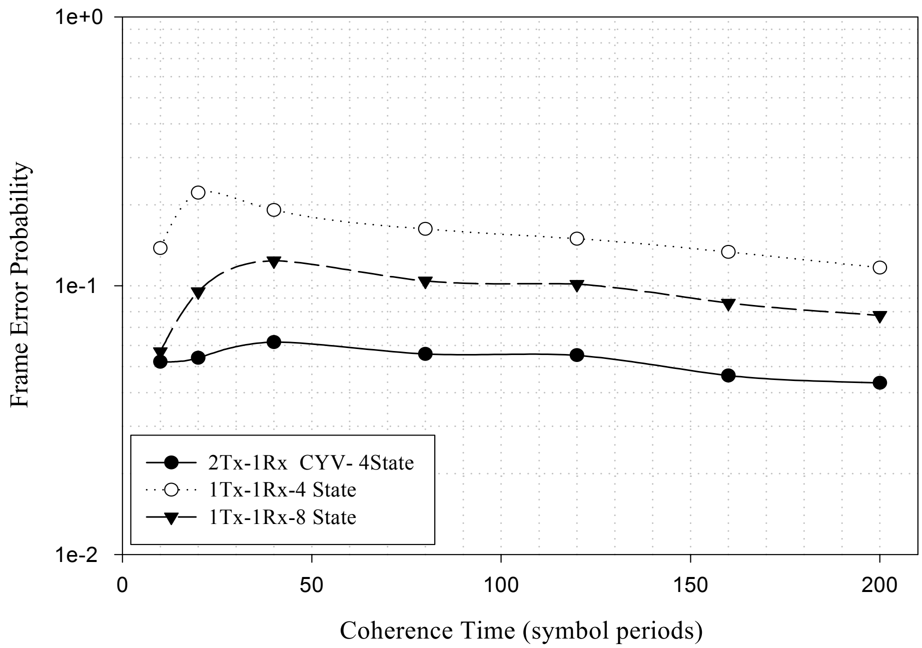

Figure 5-17: Comparison of different transmit schemes in a 2-Tap channel with SNR of 16dB and assuming perfect channel estimates

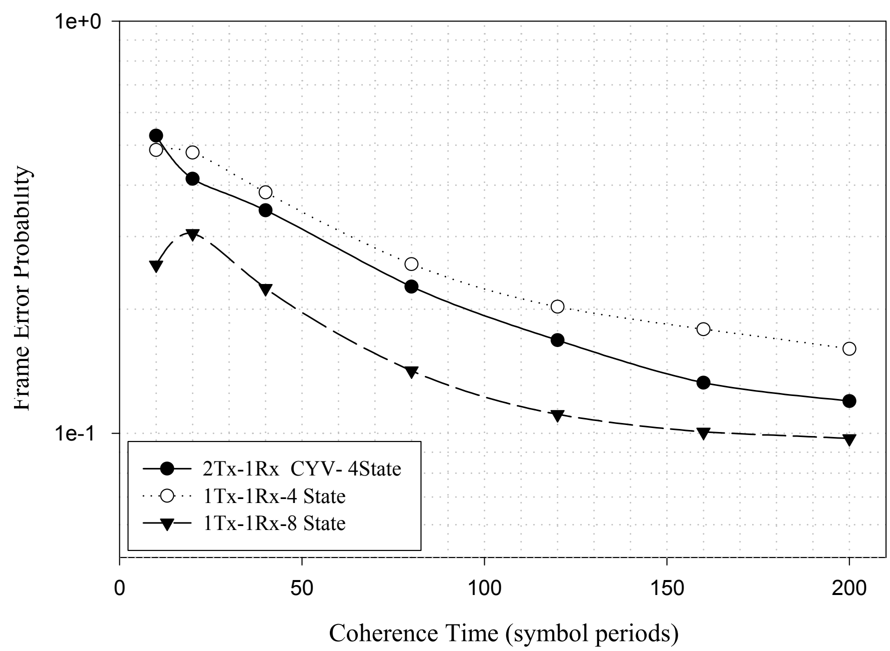

Figure 5-18: Comparison of different transmit schemes in a 2-Tap channel with SNR of 16dB and optimal training

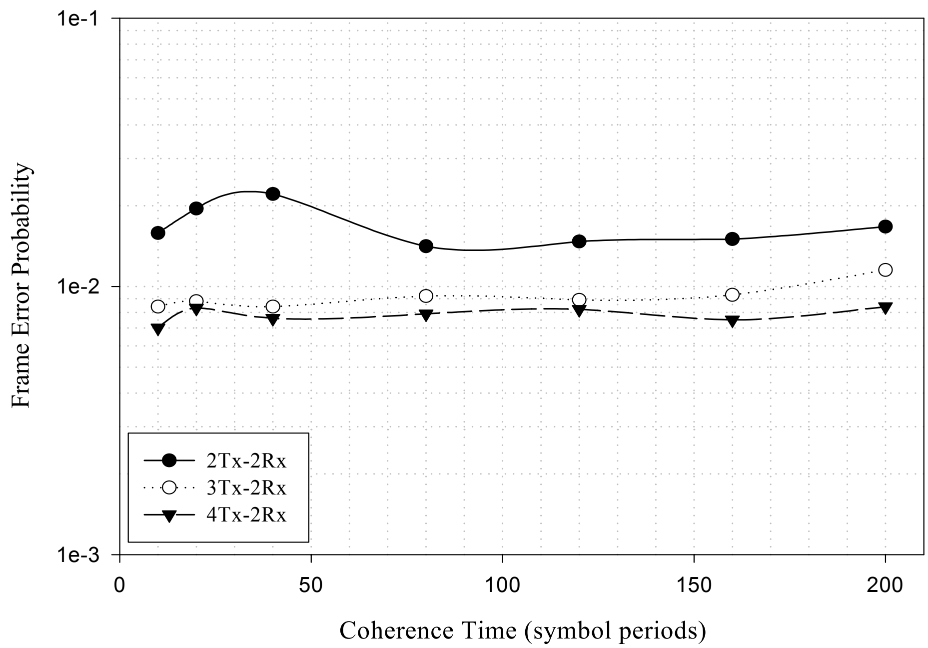

Figure 5-19: Comparison of different transmit schemes in a 2-Tap Channel with SNR of 10dB and perfect channel estimates

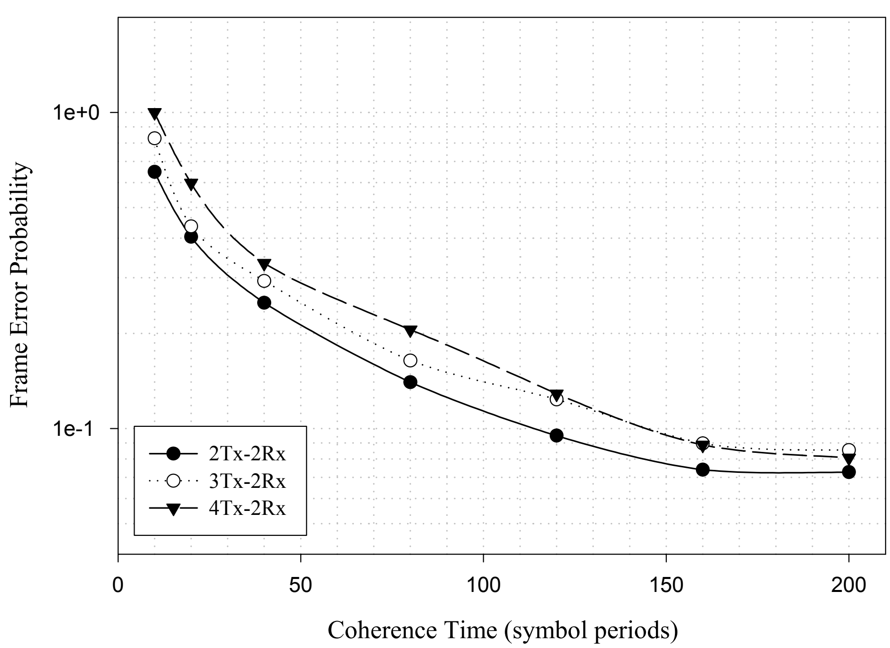

Figure 5-20: Comparison of different transmit schemes in a 2-Tap channel with SNR of 10dB and optimal training

It can thus be concluded from the analysis in this section that STTC schemes does not provide any benefit in frequency selective channels when training sequences are used to estimate the channel and when the option of utilizing a low complexity channel code is available for the single antenna system.

5.3. Chapter Summary

The performance of STTC schemes operating over frequency selective channels in the presence and absence of channel estimation errors is studied in this chapter. A discussion of the design criteria of STTC over frequency selective channels is presented. The TSC-RD design criteria are shown to be appropriate for frequency selective channels as well (from [Taro2]). A modified design criteria (based on the TSC-RD and introduced in [Youj1]) is presented that attempts to exploit the diversity offered by multi-path. The trace criterion is extended in this chapter to derive a new design criterion for frequency selective channels that improves the performance offered by the TSC-RD criteria in the presence of multi-path. This criterion is derived by modeling multi-paths as virtual antennas. CYV codes satisfy the new design criterion and are used for performance analyses.

Channel estimation errors are shown to cause a larger degradation in multi-path channels than in flat fading channels. Hence sufficient care must be taken during system design such that the additional time diversity offered in multi-path channels is not lost. The CYV codes perform better than TSC codes when optimal amount of training is used. But for high error variances, it is shown that the behavior is inverted and TSC codes start performing better. The performance of STTC is also shown to be resistant to estimation errors up to 30% of the magnitude and 10 degrees of the phase of the channel coefficients.

The optimal amount of training required for estimation of frequency selective channels for different MIMO configurations, specified SNRs and different number of multi-paths is calculated. The optimal length of training increases with an increase in the number of multi-paths and antennas. Hence the training overhead for frequency selective channels is seen to be larger than that for flat fading channels. Analyses show that transmit-diversity employing low complexity STTCs does not provide any benefit in frequency selective channels due to large capacity losses incurred due to training. Systems with no diversity and with the option of implementing low complexity channel codes are shown to perform better than a system with two-transmit antennas and employing a low complexity STTC. It is also shown that three-transmit and four-transmit antenna systems with STTC perform worse than a two-transmit antenna system with STTC. Hence transmit diversity employing STTC is seen to be unsuitable for frequency selective channels as large training overheads offset the diversity and coding benefits.

6. Thesis Summary and Future Work

Space Time Trellis Coding is a technique than intends to provide high data rates and more reliable communication over wireless channels. However, STTC have been primarily designed assuming perfect channel estimates to be available at the receiver. This thesis studied the impact of channel estimation errors (CEEs) on the design and performance of STTC. It investigated the validity of design criteria for STTC in the presence of CEEs. A detailed analysis of the influence of channel estimation on the performance benefits provided by STTC was also presented. An overview of the results and conclusions formed by the thesis is presented in this chapter

6.1. STTC Performance Analysis and Design Criteria

A comprehensive study and analysis of design and code-construction criteria for STTC over different fading channels in the presence of perfect channel estimates was presented in Chapter 2. The TSC-RD criteria are shown to be appropriate for systems with low diversity orders and operating over quasi-static and spatially correlated quasi-static flat fading channels. The criteria recommend maximizing the minimum rank and minimum determinant of the distance matrix between any two code-words, to maximize the diversity and coding gain respectively of a scheme. The TSC-DP criteria maximize the number of times the corresponding symbols of two code-words differ and are appropriate for fast-fading channels. The Trace criterion maximizes the Euclidean distance between code-words and are optimal for systems with a large diversity order (>3). An analytical evaluation of the exact pair-wise error probability of STTC was also presented. The exact PWEP expression is useful in obtaining estimates of bit error rates of STTC systems over different channel conditions.

6.2. Channel Estimation Techniques for Multiple Transmit Antenna Systems

An overview of channel estimation schemes for multiple transmit antenna systems in general and STTC in particular was presented in Chapter 3. Training sequence based estimation schemes are widely favored as their implementation is relatively simple and most current wireless standards already provide for their use. Training sequences for multiple transmit antennas require properties of zero (or very low) cross correlation between sequences transmitted from different transmit antennas. Frequency-selective channels in addition necessitate zero correlation between versions of the training sequence delayed by time-lags corresponding to multi-path delays in the channel. Capacitive analysis of the optimal amount of training required in MIMO channels was also presented.

6.3. Design and Performance of STTC in the Absence of Perfect Channel Estimates

Chapter 4 investigated the effect of channel estimation errors on the performance of STTC in flat fading channels. The design criteria for STTC were reevaluated in the presence of channel estimation errors. The analysis in [Taro2] that showed the validity of the TSC-RD criteria in quasi-static flat fading channels was discussed. This analysis was extended in the chapter, to show that both the TSC-DP criteria for fast fading channels and the trace criterion for large diversity order systems are valid in the presence of CEEs as well. The diversity advantage offered by these schemes is shown to be maintained even in the presence of estimation errors.

An exact expression for pair-wise error probability of STTC in the presence of channel estimation errors, which provides an analytical tool to evaluate the performance of STTC in the presence of imperfect channel estimates, was derived. The performance of different coding schemes and system configurations was evaluated in the presence of CEEs. It was shown that the degradation due to channel estimation errors for a specific system, in addition to the magnitude of the error also depended upon the particular choice of STTC code, the number of transmit/ receive antennas in the system and the coherence time of the channel. For instance, the Trace criterion causes larger degradation than the TSC-RD criteria. The degradation also increases with the number of antennas used in the system. An inappropriate choice of system parameters and the resultant variation in the performance degradation due to channel estimation errors are shown to alter the expected performance pattern.

This chapter also analyzed the capacity of STTC schemes using training to estimate the channel. The optimal amount of training required for channel estimation for different MIMO system configurations was calculated and was observed to increase with an increase in number of transmit antennas, a decrease in channel SNR and a decrease in coherence interval of the channel. These dependencies were used to show that for low channel coherence times and in the presence of training, increasing the number of transmit antennas leads to a decrease in the capacity offered by the system when employing STTC.

6.4. STTC in Frequency Selective Channels

Chapter 5 examined the performance of STTC schemes operating over frequency selective channels in the presence and absence of perfect channel estimates at the receiver. A new design criterion for STTC operating over frequency selective channels was derived by modeling multi-paths as virtual antennas. The criterion is based on the trace criterion. The new criterion is observed to perform better than existing schemes over multi-path channels.

Channel estimation errors are shown to cause a larger degradation in multi-path channels than in flat fading channels. The new criterion was shown to offer better performance than the TSC-RD criteria when the channel was estimated by training sequences of optimal length. But the degradation due to channel estimation errors was found to be more pronounced in the new criterion and for large estimation errors the new criteria performs worse than the TSC-RD criteria.

The optimal amount of training required for estimation of frequency selective channels for different MIMO configurations and different number of multi-paths was calculated. The optimal length of training was observed to increase with an increase in the number of taps of the frequency selective channel. This leads to larger training overhead in frequency selective channels than flat fading channels especially at low channel coherence intervals. As a consequence, in the presence of training and when the option of using channel codes is available to single antenna systems, no improvement in capacity was shown to be obtained by using STTC in frequency selective channels.

6.5. Conclusions and Directions for Future work

This thesis validated the different design criteria for STTC in the presence of channel estimation errors. It presented analytical results that modeled the performance of STTC systems in the presence of CEEs. Training based channel estimation schemes are the most popular choice for STTC systems. The amount of training however, increases with number of transmit antennas used, the number of multi-paths in the channel and a decrease in the channel coherence time. This dependence was shown to decrease the performance gain obtained by increasing the number of transmit antennas in STTC systems, especially in channels with a large Doppler spread (low channel coherence time). In multi-path channels, the training overhead associated with increasing the number of transmit antennas was shown to be so large that no benefit is obtained by using STTC.

The amount of performance degradation due to channel estimation errors was shown to be influenced by system parameters such as the specific STTC code employed and the number of transmit and receive antennas in the system in addition to the magnitude of the estimation error. Hence inappropriate choice of system parameters was shown to significantly alter the performance pattern of STTC.

This thesis thus addressed the viability of STTC in practical wireless systems and showed that channel estimation could offset benefits derived from this scheme. Hence an investigation of differential schemes for STTC might be of import as these schemes can help avoid the capacity and performance losses due to channel estimation. Also, huge losses in capacity are incurred due to the use of training sequences to estimate the channel. Blind and semi-blind channel estimation schemes which exploit the redundancy offered by STTC provide interesting alternatives and are a promising area for future research.

References

[Alam1] Alamouti, S.M, “A simple transmit diversity technique for wireless communications”, IEEE Journal on Selected Areas in Communications, Volume: 16 Issue: 8, Oct. 1998 Page(s): 1451 -1458

[Baro1] Baro, S.; Bauch, G.; Hansmann, A, “Improved codes for space-time trellis-coded modulation”, IEEE Communications Letters, Volume: 4 Issue: 1, Jan. 2000 Page(s): 20 -22

[Bauc1] Bauch, G.; Naguib, A.F, “MAP equalization of space-time coded signals over frequency selective channels”, IEEE Wireless Communications and Networking Conference, 21-24 Sept. 1999, Volume:1, Page(s): 261 -265

[Bigl1] Biglieri, E; Divsalar, D; Mclane, P.J and Simon, M.K, “Introduction to Trellis-Coded Modulation with Applications”, Macmillan Publishing Company, 1991.

[Bueh1] Buehrer, R.M.; Kumar, N.A., “The impact of channel estimation error on space-time block codes”, IEEE 56th Proceedings of Vehicular Technology Conference, Fall. 2002 Volume: 3, 24-28 Sept. 2002 Page(s): 1921 -1925 vol.3

[Cave1] Cavers, J.K.; Ho, P, “Analysis of the error performance of trellis-coded modulations in Rayleigh-fading channels”, IEEE Transactions on Communications, Volume: 40 Issue: 1, Jan 1992 Page(s): 74 -83

[Chen1] Chen, Z.; Yuan, J.; Vucetic, B, “An improved space-time trellis coded modulation scheme on slow Rayleigh fading channels”, IEEE International Conference on Communications, 2001, Volume: 4, 2001 Page(s): 1110 -1116

[Chen2] Chen, Z; Vucetic, B.; Jinhong Yuan; Ka Leong Lo, “Space-time trellis codes with two, three and four transmit antennas in quasi-static flat fading channels”, ICC 2002. Volume: 3, 28 April-2 May 2002, Page(s): 1589 -1595

[Firm1] Firmanto, W.; Vucetic, B.S.; Yuan, J, “Space-time TCM with improved performance on fast fading channels”, IEEE Communications Letters, Volume: 5 Issue: 4, Apr 2001 Page(s): 154 -156

[Fosc1] Foschini, G.J and Gans, M.J, “On limits of wireless communications in a fading environment when using multiple antennas,” Wireless Personal Communications, Volume: 6, Page(s): 311 —335, March 1998.

[Frag1] Fragouli, C.; Al-Dhahir, N.; Turin, W., “Finite-alphabet constant-amplitude training sequence for multiple-antenna broadband transmissions”, IEEE International Conference on Communications, Volume: 1, 28 April-2 May 2002 Page(s): 6 -10

[Frag2] Fragouli, C.; Al-Dhahir, N.; Turin, W., "Reduced-complexity training schemes for multiple-antenna broadband transmissions", IEEE Wireless Communications and Networking Conference, Volume: 1, 17-21 March 2002 Page(s): 78 -83 vol.1

[Frag3] Fragouli, C.; Al-Dhahir, N.; Turin, W, "Training-based channel estimation for multiple-antenna broadband transmissions", IEEE Transactions on Wireless Communications, Volume: 2 Issue: 2 , March 2003 Page(s): 384 -391

[Frag4] Fragouli, C.; Al-Dhahir, N.; Diggavi, S.N.; Turin, W, "Prefiltered space-time M-BCJR equalizer for frequency-selective channels", IEEE Transactions on Communications, Volume: 50 Issue: 5, May 2002 Page(s): 742 -753

[Gama1] Gamal, H.E.; Hammons, A.R., Jr., "On the design of algebraic space-time codes for MIMO block-fading channels", IEEE Transactions on Information Theory, Volume: 49 Issue: 1, Jan. 2003 Page(s): 151 -163

[Goza1] Gozali, R.; Woerner, B.D, "Applying the Calderbank-Mazo algorithm to space-time trellis coding", Proceedings of the IEEE Southeastcon 2000, 7-9 April 2000 Page(s): 309 -314

[Goza2] Gozali, R, "Space time codes for high data rate wireless communications", PhD Dissertation, Virginia Tech, 2002

[Gran1] Grant, A.; "Joint decoding and channel estimation for space-time codes," Vehicular Technology Conference,, Volume: 1, 24-28 Sept. 2000 Page(s): 416 -420 vol.1

[Grim1] Grimm, J; Fitz, M.P; and Krogmeier, J.V, "Further Results in space-time coding for Rayleigh fading", in Proc. 36th Annul. Allerton Conference on Communication, Control and Computing, Monticello, IL, Sept 1998, Pages(s) 391-400.

[Hass1] Hassibi, B.; Hochwald, B.M.; "How much training is needed in multiple-antenna wireless links?" IEEE Transactions on Information Theory, Volume: 49 Issue: 4, April 2003 Page(s): 951 -963

[Heik1] Heikkila, M.J.; Majonen, E.; Lilleberg, J, " Decoding and performance of space-time trellis codes in fading channels with intersymbol interference", The 11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Volume: 1 , 18-21 Sept. 2000 Page(s): 490 -494 vol.1

[Luo1] Luo, P; Guan, Y, "Optimum receiver for space-time trellis code in multi-path fading channel", International Zurich Seminar on Transmission Access, Networking Broadband Communications, 19-21 Feb. 2002, Page(s): 43-1 -43-5

[Meix1] Meixia, T; Cheng, R.S, "Improved design criteria and new trellis codes for space-time coded modulation in slow flat fading channels", IEEE Communications Letters, Volume: 5 Issue: 7, July 2001 Page(s): 313 -315

[Nagu1] Naguib, A.F.; Seshadri, N., "MLSE and equalization of space-time coded signals" IEEE 51st Vehicular Technology Conference Proceedings, Volume: 3, 15-18 May 2000, Page(s): 1688 -1693

[Perr1] Perreau, S.; White, L.B.; "Channel estimation and symbol detection for space-time codes in mobility conditions", Proceedings of International Zurich Seminar on Broadband Communications, 15-17 Feb. 2000 Page(s): 105 -108

[Proa1] Proakis, J. G, "Digital Communications", Appendix B-7, McGraw-Hill,1995.

[Proa2] Proakis, J. G, "Digital Communications", McGraw-Hill,1995.

[Ranh1] Ranheim, A.; des Rosiers, A.P.; Siegel, P.H, "An iterative receiver algorithm for space-time encoded signals", Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers, Volume: 1, 29 Oct.-1 Nov. 2000 Page(s): 516 -520 vol.1

[Rapp1] Rappaport, T.S, "Wireless Communications - Principles and Practice", Prentice Hall1996

[Safa1] Safar, Z.; Liu, K.J.R, "Space-time trellis code construction for fast fading channels", IEEE International Conference on Communications, 2002, Volume: 1, 28 April-2 May 2002 Page(s): 563 -567

[Sesh1] Seshadri, N and Winters, J.H, "Two signaling schemes for improving the error per-formance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," International. Journal on Wireless Information Networks, vol.1 1994

[Spas1] Spasojevic, P.; Georghiades, C.N, "Complementary sequences for ISI channel estimation", IEEE Transactions on Information Theory, Volume: 47 Issue: 3 , March 2001 Page(s): 1145 -1152

[Taro1] Tarokh, V, Seshadri, N and Calderbank, A.R, "Space-Time Trellis Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction", IEEE Transactions on Information Theory, Volume 44, No 2, March 1998.

[Taro2] Tarokh, V.; Naguib, A.; Seshadri, N.; Calderbank, A.R., "Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility, and multiple paths", IEEE Transactions on Communications, Volume: 47 Issue: 2, Feb 1999 Page(s): 199 -207

[Taro3] Tarokh, V.; Jafarkhani, H.; Calderbank, A.R., "Space-time block codes from orthogonal designs", IEEE Transactions on Information Theory, Volume: 45 Issue: 5, July 1999 Page(s): 1456 -1467

[Taro4] Tarokh, V.; Jafarkhani, H, "A differential detection scheme for transmit diversity", IEEE Journal on Selected Areas in Communications, Volume: 18 Issue: 7, July 2000, Page(s): 1169 -1174

[Turi1] Turin, G.L, "The Characteristic function of Hermitian Quadratic forms in complex normal random variables", Biometrika, Pages(s) 199-201, June 1960.

[Uysa1] Uysal, M.; Georghiades, C.N, "Error performance analysis of space-time codes over Rayleigh fading channels", Vehicular Technology Conference, 2000.

[Vuce1] Vucetic, B; Yuan, Y, "Space-Time Coding", John Wiley and Sons, 2003

[Wick1] Stephen. B. Wicker, "Error Control Systems for Digital Communication and Storage", Prentice Hall 1995

[Wint1] Winters, J.H, "Switched diversity with feedback for DPSK mobile radio systems", IEEE Transactions on Vehicular Technology, Volume: 32, pp134-150, 1983.

[Youj1] Youjian, L; Fitz, M.P.; Takeshita, O.Y., "Space-time codes performance criteria and design for frequency selective fading channels", IEEE International Conference on Communications, Volume: 9, 2001 Page(s): 2800 -2804 vol.9

[Youj2] Youjian, L; Fitz, M.P.; Takeshita, O.Y, "A rank criterion for QAM space-time codes", IEEE Transactions on Information Theory, Volume: 48 Issue: 12 , Dec. 2002 Page(s): 3062 -3079

Vita

Rekha Menon was born on March 13th 1979 in Chennai, India. She received her Bachelors degree in Electronics and Communication Engineering from Regional Engineering College, Trichy, India in 2000. She received her Masters degree in Electrical Engineering specializing in wireless communications from Virginia Tech in 2003. While at Virginia Tech, she was a Graduate Research Assistant (GRA) with the Center for Wireless Telecommunication from May 2001 to May 2002 during which she was part of Team "Hokiesat", a Nanosatellite to be built by Virginia Tech as a part of the Ionspheric Observation Nanosatellite Project. During the summer and fall of 2002, she interned with the RF and Network Operations Division of Cingular Wireless in Hanover, MD. She also served as a GRA for the Mobile Portable and Radio Research Group (MPRG), Virginia Tech, in fall 2003 during the course of which she worked on developing OSSIE (Open-Source SCA Implementation::Embedded), a Software Communication Architecture (SCA) tool. She is continuing at MPRG as a Ph.D. student. Her research interests include wireless signal processing and information theory.