Chapter 1

Introduction

1.1 Motivation

The ever increasing demands for capacity in wireless communication systems has fuelled research efforts investigating methods to increase the system capacity. In the past two decades, wireless networks have evolved from an analog Advanced Mobile Phone Service (AMPS) system [1–3], digital second-generation (2G) communication networks such as Global System for Mobile communications (GSM) [4] and north American Interim Standard 95 (IS-95) [5] data services to the multimedia, advanced interactive and high data rate third-generation (3G) system [6] of today.

The main characteristics of 3G systems, known collectively as International Mobile Telecommunications 2000 (IMT-2000), are a single family of compatible standards that have the following features [7]:

- Wideband CDMA systems
- Spectrum bandwidth 5 MHz
- Full coverage and mobility for a data rate of 144 kbps to 384 kbps
- High flexibility to introduce new and multimedia services
1.1 Motivation

Future systems will be based on user’s demands as the beyond third-generation (B3G) [8] and fourth-generation (4G) [9] cellular systems. The B3G/4G mobile network are targeting high-speed mobile internet applications and demanding even higher data rates and quality of service (QoS). It will not be possible to accommodate future mobile communication traffic, unless a transmission capacity of at least ten times that of IMT-2000 is attained as shown in Fig. 1.1. The International Telecommunications Union (ITU) in July 2003, has made a requirement for 4G system as follows [9]:

- At a standstill condition, the transmission data rate should be 1 Gbps
- At a moving conditions, the transmission data rate should be 100 Mbps

The demand for these services is growing at an extremely rapid pace and these trends are likely to continue for several years. However, the radio spectrum available for wireless services is extremely scarce and expensive. As a consequence, a prime issue in current wireless
1.1 Motivation

systems is the conflict between the increasing demand for wireless services and the scarce electromagnetic spectrum. Spectral efficiency is therefore of primary concern in the design of future wireless data communication systems with the omnipresent bandwidth constraint. Together with this bandwidth constriction, reliable wireless communication is challenged by harsh propagation environment characterized by path loss, time-varying multipath fading and power limitations.

To overcome the multipath fading and the resulting problems, traditionally, equalization, coding and interleaving techniques have been used [10, 11]. However, these techniques have limits on their performance enhancement, because all of them are passive techniques in the sense that their main purpose is to mitigate the effect of multipath fading. Recently, multiple antennas and the resulting space diversity have become a popular solution for overcoming the multipath fading problem.

Unlike the traditional techniques, multiple antennas make it possible to actively take advantage of the multipath, and the resulting capacity was shown to be orders of magnitude larger than the capacity of single antenna based systems with the same total power and bandwidth [12–15]. Several techniques for achieving the multiple antenna advantages have been investigated [16]. The block diagram of a MIMO system model is shown in Fig. 1.2.

In [12] Alamouti proposed a simple scheme, known as a space-time block code (STBC), where he presented a simple scheme for transmission using two antennas transmitting orthogonal sequences. This scheme is included in both wideband code-division multiple-access (WCDMA) and CDMA-2000 standards [6, 17]. STBCs is still attractive in terms of its simplicity. However, the STBC does not provide coding gain and the bandwidth efficiency is the same as in a single antenna transmit system.

A novel space-time trellis coding (STTC) [18–21] which is a joint design of trellis coding, modulation, and transmit diversity is proposed by Tarokh, Seshadri and Calderbank (TSC). Although, the STTCs simultaneously offer a substantial coding gain, spectral efficiency and diversity gain to the system, it has a potential drawback that the maximum likelihood decoder complexity grows exponentially with the number of bits per symbol, thus limiting achievable data rates. A layered space-time architecture (LST), proposed by Foschini in [14], can attain a tight lower bound on the MIMO channel capacity. The distinguishing feature of this architecture is that it allows processing of multidimensional signals in the space domain by 1-D
1.1 Motivation

processing steps in a systematic fashion. LST architectures offer the advantage of utilizing the well established 1-D codec technology. Depending on the assignment of modulated symbols over various transmit antennas and the use of error control coding techniques, various LST architectures have been proposed.

The original uncoded LST architecture described in [22] is known as vertical Bell Laboratories layered space-time scheme (V-BLAST). In the V-BLAST, the input information sequence is first demultiplexed into $n_T$ sub-streams where $n_T$ denote the number of transmit antennas. Each sub-stream is subsequently modulated by an $M$-level modulation scheme and transmitted from a transmit antenna. The signal processing chain related to an individual sub-stream is referred to as a layer.

In the horizontal layered space-time (HLST) [14] architecture, the information sequence is first encoded by a channel code and subsequently demultiplexed into $n_T$ sub-streams. Each sub-stream is modulated, interleaved and assigned to a transmit antenna. By just employing 1-D encoders after S/P conversion and before transmission, the HLST scheme suffers from the presence of sub-channels in deep fade.

Another variant of BLAST, mentioned also in [14], is the diagonal layered space-time (DLST)
architecture where the code blocks from a constituent encoder are dispersed across the diagonal of the transmission array. The diagonal layering introduces space diversity and thus achieves a better performance than the horizontal one. However, the DLST structures incur a loss in spectral efficiency due to the zero padding which is done in order to implement the diagonal structure.

A similar effect can be achieved without incurring a spectral loss by interleaving the symbols over the antennas before transmission. The transmitter structure differs from DLST in that the transmitted symbols are not periodically cycled amongst the $n_T$ antennas and they are first fed into a spatial interleaver $S II$ and then transmitted by the $n_T$ antennas. The transmitted symbols appear to have been threaded over time and across the transmit antennas. This scheme is known as threaded layered space-time architecture (TLST) [23] and it has a similar performance as DLST. The TLST combines efficient algebraic code design with iterative signal processing techniques.

Furthermore, a different perspective of LST code is enumerated at an excellent groundbreaking paper in [24], summarized that there is a fundamental trade-off between diversity and multiplexing gain. The trade-off curve was used to evaluate the performance of space-time coding schemes computing the achievable diversity-multiplexing tradeoff curve and then comparing it against the optimal tradeoff curve. It was shown that, the tradeoff performance of VLST is limited due to the independence over space whereas diagonal BLAST (DLST), with coding over the signals transmitted on different antennas can approach the optimal diversity-multiplexing trade-off curve.

Despite a significant capacity increase in communication over wireless channels that can be achieved by multiple-input multiple-output (MIMO) systems, the presence of multiple antennas at the transmitter, along with multipath and multiple users inevitably introduces not only intersymbol interference (ISI) and multiple access interference (MAI), but also interference amongst the transmit antennas (TAI).

Many advance signal processing techniques are proposed to combat interference and multipath channel distortion. Integrated detection and decoding approaches have gained in significance since the introduction of the iterative (turbo) decoding principle in the joint detection and decoding algorithm [25–27]. The iterative multiuser receivers offer huge gains over the receivers with separate detection and decoding. However, the complexity increase from
performing soft-output decoding in each iteration is considerable. It is therefore of interest to consider low-complexity iterative receivers such as the iterative parallel interference canceller (PIC) proposed in [28–30]. A PIC has a lower linear complexity and near interference free performance [28].

Moreover, it is well known that the knowledge of the channel response can be used to improve the performance of digital communication systems, both in the case of single antenna [31] or multiple antenna transmission [32]. The problem of power allocation for single and multiple antenna transmission in fading channels has been studied under a variety of criteria [33–37]. However, all these algorithms require full channel state information (CSI) leading to the requirement of large feedback overhead. Another drawback of aforesaid schemes is that they assume perfect CSI at receiver and transmitter. However, this hypothesis is not a valid one as it is not always possible to have access to ideal channel state information.

**1.2 Methodology**

In this thesis, we consider optimization of transmit power of various layers through minimization of the average bit error rate of a threaded layered space-time (TLST) architecture with an iterative PIC detector at the receiver. The calculated power is transferred to the transmitter through an error free feedback channel. A convolutional code is used as the constituent code and BPSK modulation is used as the modulation scheme. At the receiver, the PIC detector performs both interference regeneration and suppression simultaneously for all layers. The iterative decoding principle is applied to pass the a posteriori probability estimates between the detector and decoders. The decoder is based on the maximum a posteriori (MAP) algorithms [38].

Considering the practical nature of communications systems, uncertainty in the channel estimation should be taken into account in order to quantify the degradation due to channel estimation errors of any system that assumes perfect CSI. We address the problem of the design of a power allocation strategy for LST architecture with iterative PIC considering imperfect channel estimation at the receiver since perfect channel state information (CSI) is never available in practice neither at the receiver nor at the transmitter.

We investigate the effect of channel estimation errors on LST architectures with iterative PIC.
receivers. It is shown that imperfect channel estimation at the LST receiver results in erroneous decision statistics at the very first iteration and this error propagates to the subsequent iterations which ultimately leads to severe degradation of the overall performance. At the receiver, we assume the channel estimation is conducted by the training symbols based on maximum-likelihood (ML) [10] method. We consider a model for the channel-estimation error using optimal training sequences derived in [39].

1.3 Thesis Outline

The focus of this thesis is on a joint transmit signal optimization and iterative detection and decoding of layered space-time architectures. The outline of each of the chapter is as follows.

Chapter 1, gives the motivation, methodology and thesis outline.

Chapter 2 introduces the forward error correcting (FEC) schemes which will be considered in later chapters. A maximum likelihood (ML), a maximum a posteriori probability (MAP) and a turbo decoding principle are described for the example of convolutional and turbo codes.

Chapter 3 reviews the layered space-time architectures, proposed by Foschini. The principles of LST codes and various transmitter structures are revisited, followed by the exposition of the signal processing techniques used to decouple and detect the LST signals.

Chapter 4 and Chapter 5 contain the main thesis contributions. Chapter 4 considers transmit signal weighting, iterative detection and decoding in an LST system. A convolutional code is used as the constituent code. Parallel interference cancellers (PIC) are applied for detection while maximum a posteriori probability (MAP) methods are utilized for decoding. In addressing the transmit signal optimization through power allocation, it is assumed that the receiver has ideal channel state information (CSI). The transmit power is determined at the receiver through the minimization of the overall bit error rate (BER) and transferred to the transmitter through an error free feedback channel. The superiority of the proposed structures over the conventional LST architectures without joint transmit signal weighting, iterative detection and decoding is substantiated by simulations.
1.3 Thesis Outline

Chapter 5 extends the results of Chapter 4, in which perfect CSI was assumed at the receiver, to a more realistic situation of imperfect CSI accounting for channel estimation errors. It is demonstrated that the proposed design can combat channel estimation errors substantially. This has been validated by extensive simulations.

Chapter 6 concludes the thesis summarizing the main obtained results and projecting possible future work.
Chapter 2

Iterative Decoding Principles

We devote this chapter to review the optimal decoding of convolutional codes and the iterative maximum a posteriori probability (MAP) decoding principles. Convolutional codes and turbo codes are chosen as they are used in the current and proposed next generation cellular mobile communication standards. In this revisiting of maximum likelihood (ML) decoding and iterative MAP decoding, we refer most of the materials to [38, 40–43].

2.1 System Model

We consider the simple model of a single user coded system as shown in Fig. 2.1. A digital information source generates messages bearing information to be transmitted. The output of the information source is converted to a sequence $u$ consisting of $N$ binary digits. The channel encoder converts the information sequence $u$ into an encoded sequence $v$. The encoding process introduces controlled redundancy into the information sequence to combat thermal noise and some other deleterious factors during the transformation through the channel. At the receiver, the decoder performs the inverse operation resulting in an estimate of the information sequence $\hat{u}$, which is denoted by $\hat{u}$. The modulator will transform the output of the encoder into appropriate electrical waveforms suitable for transmission over the channel based on the channel characteristics. The modulated sequence, $x$, will then be sent into the channel for transmission. The channel is the physical medium between the transmitter and
receiver, which might be wired lines, optical fiber cables and wireless links. In the process of transmission, noise is inevitably introduced by various mechanisms ranging from additive thermal noise generated by electronic devices, to atmospheric noise. The channel is modeled as an additive white Gaussian noise channel. At the receiver the demodulator processes the received waveforms and produces the discrete time output with an infinite number of quantization levels. The demodulator output sequence is called the received sequence and denoted by \( r \). The channel decoder attempts to recover the information sequence \( u \). It produces an estimated sequence \( \hat{u} \) which, ideally, is a replica of \( u \).

### 2.2 Convolutional Codes and Optimum Decoding

The \((n, k, m)\) convolutional code can be implemented as a \( k \) input \( n \), \((n > k)\), output linear circuit with input memory \( m \). The structure introduced to the signal by a convolutional code is defined by the generator polynomials which describe the connections between the encoder inputs and outputs. The performance of the code depends on a code rate, defined as \( R = \frac{k}{n} \), and a code memory \( m \). Fig. 2.2 shows the encoder for the example of a binary \((2, 1, 2)\) code with generator polynomials \((5_8, 7_8)\). The encoding equations can be written as

\[
\begin{align*}
v^1 &= u * g^{(1)} \\
v^2 &= u * g^{(2)}
\end{align*}
\]

where * denotes the convolution and all operations are modulo-2.
A convenient and common way of describing encoding and decoding operations is using trellis diagrams. A trellis stage for an input at time $t$ for a binary $(2, 1, 2)$ code is shown in Fig. 2.3. A trellis diagram consists of $N$ such stages, where $N$ is the number of input words, each consisting of $k$ input data bits. The state of the encoder is defined as the content of its shift register. For the encoder with a total memory $K$ the number of states is $2^K$. Each new block of $k$ inputs causes the transition to a new state. That is, there are $2^k$ branches leaving each state. Each branch is labeled by the $k$ inputs causing the transition at time unit $t$, denoted by $u_t = (u_{t,1}, \cdots, u_{t,k})$, and $n$ corresponding outputs, denoted by $v_t = (v_{t,0}, \cdots, v_{t,n-1})$.

Given that $r$ is received, the conditional error probability of the decoder is defined as

$$P(E|r) = P(\hat{u} \neq u) = P(\hat{v} \neq v)$$  \hspace{1cm} (2.2)

The probability of error in the decoder is then given by

$$P(E) = \sum_r P(E|r)P(r)$$  \hspace{1cm} (2.3)

The term $P(r)$ is independent of the decoding algorithm, so the minimum probability of error in Eq. (2.3) is achieved by minimizing $P(E|r) = P(\hat{v} \neq v|r)$ for all $r$. This is equivalent to maximizing $P(\hat{v} = v|r)$. That is, the decoder searches for $(\hat{v})$ which maximizes

$$P(\hat{v} \neq v|r) = \frac{P(r|v)P(v)}{P(r)}$$  \hspace{1cm} (2.4)

If the coded sequences are all equally likely, then for a discrete memory channel the optimal decoding maximizes

$$P(r|v) = \prod_i P(r_i|v_i)$$  \hspace{1cm} (2.5)

It is convenient to consider the logarithm of the expression in Eq. (2.5). As $\log(x)$ is a monotonically increasing function, maximizing the expression in Eq. (2.5) is equivalent to maximizing

$$\log P(r|v) = \sum_i \log P(r_i|v_i)$$  \hspace{1cm} (2.6)

The function $\log P(r|v)$ is known as log-likelihood function and the decoding algorithm that maximizes this function is known as maximum likelihood (ML) decoding algorithm.
Figure 2.2: A (2,1,2) convolutional encoder
2.2 Convolutional Codes and Optimum Decoding

2.2.1 Maximum Likelihood (ML) Algorithm

In 1967 Viterbi proposed an efficient algorithm for decoding of convolutional codes. This algorithm can achieve maximum likelihood decoding performance which minimizes the probability of error in decoding of the whole received sequence when the binary information bits are statistically independent and equally likely.

The following describes the principles of the Viterbi algorithm. For simplicity, we assume a binary code, \( k = 1 \) and consider BPSK modulation. The data sequence \( u = (u_1, \ldots, u_t, \ldots, u_N) \) is encoded by a convolutional \((n, 1, m)\) encoder. The trellis diagram has a total number of \( M = 2^m \) distinct states, indexed by integer \( l, l = 0, 1, \ldots, M - 1 \). The state of the trellis \( S_t \) represents the encoder register content at time \( t \). The state in a trellis diagram, from time \( t \) to \( t' \), is denoted by

\[
S_{t'} = (S_t, S_{t+1}, \ldots, S_{t'}) \quad (2.7)
\]

The corresponding encoder output sequence is

\[
v_{t'} = (v_t, v_{t+1}, \ldots, v_{t'}) \quad (2.8)
\]

where \( v_t = (v_{t,0}, v_{t,1}, \ldots, v_{t,n-1}) \) and \( n \) is the codeword length. The encoded sequence is BPSK modulated, generating the sequence

\[
x_{t'} = (x_t, x_{t+1}, \ldots, x_{t'}) \quad (2.9)
\]
where \( \mathbf{x}_t = (x_{t,0}, x_{t,1}, \cdots, x_{t,n-1}) \). The received bit corresponding to the transmitted bit \( x_{t,i} \) is a sum of the transmitted bit and additive white Gaussian noise (AWGN)

\[
r_{t,i} = x_{t,i} + n_{t,i}, \quad i = 0, 1, \cdots, n - 1
\]

(2.10)

where \( n_{t,i} \) is a zero mean Gaussian random variable with variance \( \sigma^2 \). The received sequence for a block of information bits from time \( t \) to time \( t' \) is denoted by

\[
\mathbf{r}_t' = (r_t, r_{t+1}, \cdots, r_{t'})
\]

(2.11)

where \( r_t = (r_{t,0}, r_{t,1}, \cdots, r_{t,n-1}) \). The Viterbi (ML) decoder selects the word \( \mathbf{x} = \mathbf{x}_1^N \) that maximizes the likelihood function

\[
\log P(\mathbf{r}|\mathbf{x}) = \sum_{t=1}^{N} \log P(r_t|x_t) = \sum_{t=1}^{N} \sum_{j=0}^{n-1} \log P(r_{t,j}|x_{t,j})
\]

(2.12)

For the Gaussian channel this becomes

\[
\log P(\mathbf{r}|\mathbf{x}) = -nN \log \sqrt{2\pi\sigma} - \sum_{t=1}^{N} \sum_{j=0}^{n-1} \frac{(r_{t,j} - x_{t,j})^2}{2\sigma^2}
\]

(2.13)

From Eq. (2.13) we notice that maximizing the log-likelihood function is equivalent to minimizing the Euclidean distance between the received sequence \( \mathbf{r} = \mathbf{r}_1^N \) and modulated sequence \( \mathbf{x} = \mathbf{x}_1^N \) in the trellis diagram. The Euclidean distance is given by

\[
\text{ED}(\mathbf{r}, \mathbf{x}) = \sum_{t=1}^{N} \sum_{j=0}^{n-1} \frac{(r_{t,j} - x_{t,j})^2}{2\sigma^2}
\]

(2.14)

For the procedures in the Viterbi algorithm it is convenient to define the branch and path metrics as

\[
\nu^{(x_t)}_t = \sum_{j=0}^{n-1} \frac{(r_{t,j} - x_{t,j})^2}{2\sigma^2} \quad \text{and} \quad \mu_M = \sum_{t=1}^{M} \nu^{(x_t)}_t
\]

(2.15)

The Viterbi algorithm is equivalent to using dynamic programming for finding the shortest path through a weighted graph. The objective of the algorithm is to find a maximum likelihood path in the trellis, that is, a path with a minimum path metric (also referred to as the minimum distance path or shortest path). The search for the maximum likelihood path is performed in a recursive manner. At each time \( t \), the path consisting of \( t \) branches with a
2.2 Convolutional Codes and Optimum Decoding

minimum metric is selected and stored for each state. This shortest partial path for a particular state is called the survivor. The decision on the data bit $u_t$ is made with a delay $\tau$. The minimum path metric at time $t + \tau$ is computed as the minimum metric path among the $M$ survivor path metrics, one for each of the $M$ trellis states. The survivor path metric for state $S_t$ is computed as

$$\mu_t = \min_{x_t} \left[ \mu_{t-1} + v_t^{(x_t)} \right]$$  \hspace{1cm} (2.16)

where $\mu_{t-1}$ is the survivor metric for the state $S_{t-1}$ which is connected with the state $S_t$ by a branch associated with a codeword $x_t$.

If the minimum path metric $\mu_{N,\text{min}}$ corresponds to path $(\hat{x})_t$, then the decoder selects the binary symbol on this path at time $t$, namely $(\hat{u})_t$, and delivers it as a hard estimate of the transmitted symbol $u_t$.

In the concatenated schemes where the output of the decoder is fed into another decoder or another soft-input module, it is of interest that the decoder produces the bits’ a posteriori probability estimates rather than the bits’ hard estimates.

The Viterbi algorithm can be modified to generate soft outputs by considering both the maximum likelihood path and the strongest competitor (SC) path. The strongest competitor path is the minimum path metric for the path obtained when the trellis symbol on the maximum likelihood path at time $t$ is replaced by its complementary symbol. If we denote the metric of the ML and SC paths containing the data bit $u_t$ by $M_{t}^{\text{sc}}$, respectively, then the a posteriori log-likelihood ratio for the data bit at time $t$ is estimated as

$$\Lambda(u_t) = \log \frac{P(u_t = 1|r)}{P(u_t = 0|r)} = \log \frac{e^{-M_{t}^{\text{sc}}}}{e^{-M_{t}^{\text{ml}}}}$$  \hspace{1cm} (2.17)

If the bit on a ML path is $u_t = 1$, or as

$$\Lambda(u_t) = \log \frac{P(u_t = 1|r)}{P(u_t = 0|r)} = \log \frac{e^{-M_{t}^{\text{sc}}}}{e^{-M_{t}^{\text{ml}}}}$$  \hspace{1cm} (2.18)

If the bit on a ML path is $u_t = 0$.

The a posteriori probabilities of the data bit can be calculated from Eq. (2.17) as

$$P(u_t = 1|r) = \frac{e^{\Lambda}}{1 + e^{\Lambda}}$$  \hspace{1cm} (2.19)
2.2 Convolutional Codes and Optimum Decoding

\[ P(u_t = 0|\mathbf{r}) = \frac{1}{1 + e^{\Lambda_t}} \]  

(2.20)

If the decision is made on a finite block length, then the proposed algorithm can be implemented as a bi-directional recursive method with forward and backward recursions.

### 2.2.2 MAP Decoding Algorithm

The maximum a posteriori probability (MAP) algorithm uses a decoding criteria that minimizes the bit error probability, whereas the Viterbi algorithm minimizes the sequence error probability. The MAP algorithm is computationally more complex than the Viterbi algorithm and requires knowledge of noise variance. However, MAP considers all possible paths in a trellis as opposed to the soft output Viterbi algorithm (SOVA) which considers only the ML and SC paths. This becomes a very important advantage of the MAP algorithm for the iterative decoding algorithms.

The soft-output MAP decoder calculates the a posteriori log-likelihood ratio for the data bit \( u_t \) as

\[ \Lambda(u_t) = \log \frac{P\{u_t = 1|\mathbf{r}\}}{P(u_t = 0|\mathbf{r})} \]  

(2.21)

where \( P\{u_t = i|\mathbf{r}\}, i = 0, 1 \) is the a posteriori probability (APP) of the data bit \( u_t \). The decoder makes the hard decision by comparing \( \Lambda(u_t) \) to zero

\[ u_t = \begin{cases} 
1 & \text{if } \Lambda(u_t) > 0 \\
0 & \text{otherwise.} 
\end{cases} \]  

(2.22)

The APPs in Eq. (2.21) can be computed from a trellis diagram as

\[ P(u_t = 0|\mathbf{r}) = \sum_{(m', m) \in B_t^0} P\{S_t-1 = m', S_t = m|\mathbf{r}\} \]  

(2.23)

\[ P(u_t = 1|\mathbf{r}) = \sum_{(m', m) \in B_t^1} P\{S_t-1 = m', S_t = m|\mathbf{r}\} \]  

(2.24)

where \( S_{t-1} \) and \( S_t \) are the encoder states at time \( t - 1 \) and \( t \), respectively, and \( B_t^0 \) and \( B_t^1 \) are set of transitions from state \( m' \) to state \( m \) caused by \( u_t = 0 \) and \( u_t = 1 \), respectively. Eqs.

16
2.2 Convolutional Codes and Optimum Decoding

(2.23) and (2.24) can be written as

\[
P(u_t = 0 | r) = \sum_{(m', m) \in B_0^t} \frac{P\{S_{t-1} = m', S_t = m, r\}}{P\{r\}} \tag{2.25}
\]

\[
P(u_t = 1 | r) = \sum_{(m', m) \in B_1^t} \frac{P\{S_{t-1} = m', S_t = m, r^N\}}{P\{r\}} \tag{2.26}
\]

where the \(P\{r\}\) is a constant. Since it does not affect the maximization, it will be dropped in further derivations.

In order to efficiently calculate the information bits APPs, the following probability functions are defined [38]

\[
\alpha_i(m) = P\{S_t = m, r_i^t\} \tag{2.27}
\]

\[
\beta_i(m) = P\{r_i^N | S_t = m\} \tag{2.28}
\]

\[
\gamma_i^I(m', m) = P\{u_t = i, S_t = m, r_t | S_{t-1} = m'\} \tag{2.29}
\]

where

\[
r_t = (r_{t, 0}, r_{t, i}, \cdots, r_{t, n-1}) \tag{2.30}
\]

\[
r_t^k = (r_t, r_{t+1}, \cdots, r_k) \tag{2.31}
\]

The joint transition probability, \(P\{S_{t-1} = m', S_t = m, r\}\), can be expressed as

\[
P\{S_{t-1} = m', S_t = m, r\} = \alpha_{t-1}(m') \sum_{i \in 0, 1} \gamma_i^I(m', m) \beta_i(m) \tag{2.32}
\]

where \(\alpha_i(m)\) and \(\beta_i(m)\) are obtained recursively as

\[
\alpha_i(m) = \sum_{m'} \alpha_{i-1}(m') \sum_{i \in 0, 1} \gamma_i^I(m', m) \tag{2.33}
\]

\[
\beta_i(m) = \sum_{m'} \beta_{i+1}(m') \sum_{i \in 0, 1} \gamma_i^I(m, m') \tag{2.34}
\]
and $\gamma_t^i(m', m)$ is a channel transition probability weighted by the information bit a priori probability $p_t(u_t = i), i = 0, 1$, where $u_t$ is the information symbol associated with the transition $S_{t-1} = m' \rightarrow S_t = m$. Coefficient $\gamma_t^i(m', m)$ can be written as

$$\gamma_t^i(m', m) = p_t(u_t = i) \prod_{j=0}^{j=n-1} P\{r_{t,j}|x_{t,j}\}$$

(2.35)

where $x_{t,j}, j = 0, \ldots, n-1$ is a BPSK modulated symbol in a codeword associated with transition $S_{t-1} = m' \rightarrow S_t = m$.

If we assume that the encoder starts and ends in a zero state the boundary conditions are

$$\alpha_0(0) = 1, \alpha_0(m) = 0 \quad \text{for} \quad m \neq 0$$

(2.37)

$$\beta_N(0) = 1, \beta_N(m) = 0 \quad \text{for} \quad m \neq 0$$

(2.38)

The log-likelihood ratio $\Lambda(u_t)$ can be written as

$$\Lambda(u_t) = \log \frac{\sum_{(m', m) \in B^1} \alpha_{t-1}(m') \gamma_t^i(m', m) \beta_t(m)}{\sum_{(m', m) \in B^0} \alpha_{t-1}(m') \gamma_t^i(m', m) \beta_t(m)}$$

(2.39)

The above algorithm is usually referred to as forward/backward algorithm, since the coefficients $\alpha_t(m)$ are calculated recursively starting from the beginning of the trellis (forward recursion), and the $\beta(m)$ coefficients are calculated recursively starting from the end of the trellis (backward recursion).

Fig. 2.4 shows the graphical representation of the forward and backward recursions. In this figure $\alpha_{t-1}(m_t')$ represents the $\alpha$ coefficient for state $m_t'$ in the $(t - 1)$-th stage which is connected with a state $m$ in the $t$-th trellis stage and where transition $S_{t-1} = m' \rightarrow S_t = m$ is caused by the information bit $u_t = i, i = 0, 1$. Similarly, the $\beta_{t+1}(m''_t)$ denotes the $\beta$ coefficient for state $m''_t$ in the $(t+1)$-th trellis stage which is connected with a state $m$ in $t$-th trellis stage and where the transition $S_t = m \rightarrow S_{t+1} = m''_t$ is caused by the information bit $u_t = i, i = 0, 1$.

The a posteriori probabilities of the information bits can be calculated as

$$P\{u_t = 1|r\} = \frac{e^{\Lambda(u_t)}}{1 + e^{\Lambda(u_t)}}$$

(2.40)
The a posteriori probabilities of the transmitted bits can be calculated by adding the probabilities of the codewords that contain a particular transmitted bit. That is,
\begin{align}
P\{x_{t,j} = 1 | r\} &= \sum_{u_t = i, x_{t,j} = 1} P\{u_t = i | r\} \quad (2.42) \\
P\{x_{t,j} = -1 | r\} &= \sum_{u_t = i, x_{t,j} = -1} P\{u_t = i | r\} \quad (2.43)
\end{align}

\section{2.3 Turbo Codes and an Iterative Decoding Principle}

Turbo codes belong to the class of codes obtained by serial or parallel concatenation of recursive systematic convolutional codes [38]. These codes have gained a lot of attention since they can be decoded iteratively with a decoding performance which approaches the capacity limit. Furthermore, the iterative (turbo) decoding principle has been successfully applied in some other applications, such as joint decoding and detection in the Layered Space-Time structure, which we address in the following chapters. In order to describe
2.3 Turbo Codes and an Iterative Decoding Principle

the iterative decoding principle, we consider an example of a parallel turbo code consisting of two 1/2 rate recursive systematic convolutional codes (RSC). Turbo codes employ RSC codes since these have a weight distribution which results in a better bit error performance at low SNR than the equivalent nonsystematic feedforward codes [38].

The turbo-encoder for this example is shown in Fig. 2.5. It consists of two (2,1,4) encoders operating on the same block of information bits. The first encoder directly encodes the information sequence $u$, while the second encoder encodes the interleaved version of the information sequence. The interleaver is very important and both its size its design largely influence the performance. Its role is to generate a long block code from the small memory convolutional codes and to decorrelate the information which gets iteratively exchanged between the decoders in a turbo decoding process [38].
2.3 Turbo Codes and an Iterative Decoding Principle

The overall rate of turbo decoder for this example is 1/3 and the encoder output for each input data bit consists of an information bit followed by the two parity check bits generated by the RSC encoders.

For large interleaver sizes the complexity of the optimal MLSE or MAP decoding becomes very high. The practical importance of the turbo codes comes from the existence of an efficient iterative decoding algorithm.

The block diagram of the iterative (turbo) decoder is shown in Fig. 2.6.

In the turbo decoder each component code is decoded separately by the soft output decoding algorithm. The $i$-th, $i = 1, 2$ decoder calculates the a posteriori log-likelihood ratio (LLR) of the information bit as

$$\Lambda^i(u_t) = \log \frac{\sum_{(m',m) \in B^1} \alpha_{t-1}(m') p_i^1(1) \exp \left( -\sum_{j=0}^{n}(r_t,j - x_t,j)^2 \right) \beta_t(m)}{\sum_{(m',m) \in B^0} \alpha_{t-1}(m') p_i^1(0) \exp \left( -\sum_{j=0}^{n}(r_t,j - x_t,j)^2 \right) \beta_t(m)}$$

(2.44)

where $\alpha$ and $\beta$ are given by Eq.(2.34) and the a priori probability that bit $u_t = j$, $j = 0, 1$ is denoted by $p_i^j(j)$, for the decoder $i, i = 1, 2$.

From the log-likelihood ratio (LLR) for the data bit $u_t$, the decoder extracts a part of it which is not directly dependent on the a priori information or a systematic bit for the same data bit $u_t$. This information, referred to as extrinsic information, is then passed to the other decoder. Extrinsic information ratio (EIR) is given by

$$\Lambda^i_e(u_t) = \Lambda^i(u_t) - \log \frac{p_i^1(1)}{p_i^1(0)} + \frac{2}{\sigma^2} r_{t,0}$$

(2.45)

where $\log \frac{p_i^1(1)}{p_i^1(0)}$ is the a priori information ratio calculated in the $j$-th decoder $j \neq i$ in the previous iteration, and $\frac{2}{\sigma^2} r_{t,0}$ is the systematic bit contribution. The decoder for one component code takes as input extrinsic information supplied by the other decoder and uses it as the data bit a priori information in the MAP or SOVA decoding process.

The extrinsic information ratio (EIR) can be directly calculated by the forward/backward algorithm by excluding from Eq. (2.44) terms that contain bit a priori probabilities $p_i^j(j)$ and the systematic bit $r_{t,0}$. That is, in terms of the forward/backward coefficients, the EIR is
2.3 Turbo Codes and an Iterative Decoding Principle

Figure 2.6: Turbo decoder with MAP decoding of the composed codes.
2.3 Turbo Codes and an Iterative Decoding Principle

given by

\[
\Lambda_e(\epsilon_t) = \log \frac{\sum_{(m',m) \in B^1} \alpha_{t-1}(m') \exp \left( -\frac{\sum_{j=1}^n (r_{t,j} - x_{t,j})^2}{2\sigma^2} \right) \beta_t(m)}{\sum_{(m',m) \in B^0} \alpha_{t-1}(m') \exp \left( -\frac{\sum_{j=1}^n (r_{t,j} - x_{t,j})^2}{2\sigma^2} \right) \beta_t(m)}
\]  

(2.46)

The iterative exchange of information between decoders results in improvement of the individual decoder estimates. After a certain number of iterations the decoders stop producing further improvements. In the last stage of the decoding hard decisions are made in the second decoder according to the Eq. (2.22).
Chapter 3

Layered Space Time Architectures

Multiple antenna systems have a potential to achieve a high capacity compared to a single antenna system in a Rayleigh fading environment. Notwithstanding the ability of Space-Time Trellis Code (STTC) to achieve a full diversity order, its spectral efficiency is practically limited by the complex encoder structure as the maximum likelihood decoder complexity grows exponentially with the number of modulation levels. The layered space-time structure (LST), proposed by Foschini [14], can overcome this problem. With powerful signal processing techniques, the LST architecture allows processing of multidimensional signals in space by one-dimensional (1-D) processing steps, which facilitates the integration of the mature 1-D encoded technology into the MIMO framework. The complexity of the LST receivers grows linearly with the date rate.

Furthermore, the horizontally (HLST), diagonally (DLST) and threaded (TLST) layered space time structures were proposed to realize this concept [14, 22, 23, 44–46]. In the HLST architecture, an information sequence is split into $n_T$ substreams, where $n_T$ is the number of transmit antennas. Each substream is encoded by a channel encoder and then transmitted by a particular transmit antenna. However, better performance is achieved by the DLST architecture [14], in which the output codeword of each encoder is distributed among the $n_T$ antennas along the diagonal of the transmission array. The TLST [23] combines efficient algebraic code design with iterative signal processing techniques. It is assumed that the multipath channel parameters are not known at the transmitter but can be perfectly estimated at the receiver. The received signal is a superposition of the transmitted coded symbols
scaled by the multipath fading coefficients and corrupted by AWGN. At the receiver side, interference suppression and interference cancellation techniques [47] are employed in the detector and each 1-D constituent code can be decoded individually [14, 46, 48]. In addition, a great performance improvement can be realized by employing iterative detection/decoding techniques [23, 44, 49].

The remainder of the chapter is organized as follows. Section 3.1 presents the system model of the LST structure. Section 3.2 gives the transmit structures of four classical LST architectures (VLST, HLST, DLST and TLST). Section 3.3 revisits multiuser detection techniques for LST systems. VLST receiver, interference suppression combined with interference cancellation based on QR decomposition and iterative MMSE receivers are described. We refer most of the materials of this literature review to [50].

### 3.1 System Model

Fig. 3.1 shows the block diagram of the general LST architecture. We consider a wireless channel $H$, represented by an $n_r \times n_t$ complex matrix. The element of $H$, denoted by $h_{ij}$, $1 \leq i \leq n_T$, $1 \leq j \leq n_R$, is the fading attenuation for the path from transmit antenna $i$ to receive antenna $j$. For normalization purposes we assume that the received power for each of the $n_R$ receive branches is equal to the total transmitted power. Thus we obtain the normalization constraint for the elements of $H$, on a channel with fixed coefficients, as

$$\sum_{j=1}^{n_R} |h_{ij}|^2 = n_T, \quad i = 1, 2, \ldots, n_R \quad (3.1)$$

We assume that the channel matrix is known to the receiver but not at the transmitter. The channel matrix can be estimated at the receiver by transmitting a training sequence. The vector match-filtered signals at receive antennas is then given by

$$r = Hx + n \quad (3.2)$$

where $x$ is an $[n_T \times 1]$ vector of simultaneously transmitted symbols from the $n_T$ transmit antennas, and $n$ is an AWGN modeling of the additive effects of the channel and the receiver. The covariance matrix of the receiver noise is given by

$$R_{nn} = E\{nn^H\} \quad (3.3)$$
Figure 3.1: LST block diagram.
3.2 LST Transmitters

If there is no correlation between the components of \( \mathbf{n} \), the covariance matrix is obtained as

\[
\mathbf{R}_{nn} = \sigma_n^2 \mathbf{I}_{n_R}
\]  (3.4)

We assume synchronous transmission from various transmit antennas and a quasi-static flat Rayleigh fading channel, where the fading coefficients change independently and randomly every \( L \) symbols, where \( L \) is the number of symbols in a frame. The total transmitted power is constrained to \( P \), regardless of the number of transmit antennas \( n_T \) and can be represented as

\[
P = tr(\mathbf{R}_{xx})
\]  (3.5)

where \( tr(\mathbf{A}) \) denotes the trace of matrix, \( \mathbf{A} \), and \( \mathbf{R}_{xx} \) denotes the covariance matrix of the transmitted signal and is given by

\[
\mathbf{R}_{xx} = E\{\mathbf{x}\mathbf{x}^H\}
\]  (3.6)

\( E\{\cdot\} \) denotes the expectation, and \( (\mathbf{A})^H \) denotes the Hermitian of matrix \( \mathbf{A} \). The transmitted power of each layer is assumed to be equal as

\[
P_i = P/n_T \quad i = 1, 2, \ldots, n_T
\]  (3.7)

3.2 LST Transmitters

In this section, we first review the three classical LST systems [50]. These systems were proposed by Bell Laboratories and hence the name Bell Laboratories Layered Space Time (BLAST) architecture. They are vertical layered space time (VLST) [22], horizontal layered space time (HLST) [14], and diagonal layered space time (DLST) [14] systems. Then an improved LST, known as threaded layered space time (TLST) structure [23] which is designed for the joint iterative detection/decoding receiver is presented.

3.2.1 VLST

In VLST, a single binary information stream is demultiplexed into \( n_T \) substreams and each substream is then mapped into symbols and fed to its respective transmit antenna. The formation of the transmission matrix is described in (3.8). The signal processing chain related
3.2 LST Transmitters

Figure 3.2: A VLST architecture.

To an individual sub-stream is refereed to as a layer. We emphasize that there is no coding in VLST. The VLST structure is illustrated in Fig. 3.2.

$$
\begin{pmatrix}
I_1 & I_2 & \cdots & I_k & \cdots \\
\vdots & \vdots & \ddots & \vdots & \ddots \\
I_{n_T} & I_{2n_T} & \cdots \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
I_1 & I_{n_T+1} & \cdots \\
I_2 & I_{n_T+2} & \cdots \\
\vdots & \vdots & \ddots \\
I_{n_T} & I_{2n_T} & \cdots \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
x_1^1 & x_2^1 & \cdots \\
x_1^2 & x_2^2 & \cdots \\
\vdots & \vdots & \ddots \\
x_1^{n_T} & x_2^{n_T} & \cdots \\
\end{pmatrix} \quad (3.8)
$$

3.2.2 HLST

In the HLST architecture, shown in Figs. 3.3 and 3.4, the information sequence is first encoded by a channel code and subsequently demultiplexed into $n_T$ substreams. Each sub-stream is modulated, interleaved and transmitted from the $i$th transmit antenna. Therefore, coded symbols from the $i$th layer always occupy the $i$th row of transmission matrix.
3.2 LST Transmitters

Figure 3.3: HLST architecture with a single code.

Expression (3.9) shows an example of the HLST architecture. Information bits \( I_1, I_2, I_3, \ldots, I_k \) are demultiplexed into \( n_T \) layers. Each layer is encoded by a separate constituent encoder. The first layer, for example, containing bits \( I_1, I_{n_T+1}, I_{2n_T+1}, \ldots \) is encoded by a constituent encoder. The encoder output of length \( l \) on the first layer is denoted by sequence \( c^1 = (c^1_1, c^1_2, c^1_3, \ldots, c^1_l) \), which is the first row of the transmission matrix. With HLST, symbols from the sequence \( c^i \) are transmitted sequentially in time through transmit antenna \( i \). Therefore, at time \( t \), symbols \( c^1_t, c^2_t, \ldots, c^{n_T}_t, t \in 1, 2, \ldots, l \), which is the \( t \)th column of the transmission matrix, are transmitted simultaneously through \( n_T \) transmit antennas.
Figure 3.4: HLST architecture with a separate code in each layer.

### 3.2.3 DLST

Coded symbols from a layer will occupy a diagonal of the transmission matrix. Expression (3.10) shows the formation of the transmission matrix of the DLST architecture.

\[
\begin{pmatrix}
    c_1^1 & c_1^2 & \cdots & c_1^{n_T} \\
    c_2^1 & c_2^2 & \cdots & c_2^{n_T} \\
    \vdots & \vdots & \ddots & \vdots \\
    c_{n_T}^1 & c_{n_T}^2 & \cdots & c_{n_T}^{n_T}
\end{pmatrix}
\rightarrow
\begin{pmatrix}
    c_1^1 & c_1^2 & \cdots & c_1^{n_T} & 0 & \cdots & 0 \\
    0 & c_1^2 & c_1^3 & \cdots & c_1^{n_T} & \ddots & \vdots \\
    \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
    0 & \cdots & 0 & c_{n_T}^1 & c_{n_T}^2 & \cdots & c_{n_T}^{n_T}
\end{pmatrix}
\] (3.10)

As before, the information bits are demultiplexed into \( n_T \) layers and are individually encoded per layer. Let us denote the symbols of the \( i \)-th layer encoded by the channel encoder as \( c_i^1 = (c_i^1, c_i^2, c_i^3, \cdots, c_i^{n_T}) \). In the diagonalization process, the codeword from the first layer occupies the leftmost northwest-southeast diagonal of the transmission matrix, and the codeword symbol from the second layer is assigned to the next diagonal, and so on. Therefore, for the first layer, at time \( t = 1 \), the first transmit antenna transmits modulated symbol \( c_1^1 \), while all other antennas remain idle. At time \( t = 2 \), the second transmit antenna transmits modulated symbol \( c_2^1 \) and transmit antennas \((3, \cdots, n_T)\) are idle. Finally, symbols \( c_{n_T}^1 \) is
transmitted at time $n_T$ from antenna $n_T$. For the second layer, its first symbol $c_2^1$ is transmitted at $t = 2$ from the first antenna and the last symbol $c_2^{n_T}$ is transmitted at $t = n_T + 1$ from antenna $n_T$. In the above description, we assume the diagonal width $W = 1$, which means that $1 \times n_T$ encoded symbols from a layer are assigned to $1 \times n_T$ time slot and transmitted over $n_T$ different antennas. For $W = 1$, the length of the codeword can be $W \cdot n_T$, and each diagonal contains $W \cdot n_T$ symbols, in which $W$ consecutive symbols are transmitted over one of $n_T$ antennas within $W$ time slots. We notice that the efficiency of DLST is not maximized, since a portion of the transmission matrix is idle. The dimension of the transmission matrix is defined as $n_T \times (W_2n_T-1)$, in which $W_2n_T(n_T-1)$ entries are empty.

### 3.2.4 TLST

A TLST structure is obtained by introducing a spatial interleaver $S\Pi$ prior to time interleavers $\Pi_1, \cdots, \Pi_{n_T}$ as shown in Fig. 3.5. Hence the codeword symbols of each encoder are transmitted over different antennas. Let us consider a system with $n_T = 4$. The operation of
3.3 LST Receivers

STI can be expressed as

\[
\begin{pmatrix}
  c_1^1 & c_2^1 & c_3^1 & c_4^1 & \cdots \\
  c_1^2 & c_2^2 & c_3^2 & c_4^2 & \cdots \\
  c_1^3 & c_2^3 & c_3^3 & c_4^3 & \cdots \\
  c_1^4 & c_2^4 & c_3^4 & c_4^4 & \cdots \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
  c_1^1 & c_2^1 & c_3^1 & c_4^1 & \cdots \\
  c_1^2 & c_2^2 & c_3^2 & c_4^2 & \cdots \\
  c_1^3 & c_2^3 & c_3^3 & c_4^3 & \cdots \\
  c_1^4 & c_2^4 & c_3^4 & c_4^4 & \cdots \\
\end{pmatrix}
\] (3.11)

in which an element of the codeword matrix, denoted by \(c_i^t\), represents the encoded bit of layer \(i\) at time \(t\). The matrix on the right hand side of (3.11), denoted by \(C'\), is the TLST transmission matrix. That is, the codeword symbols \(c_1^1, c_2^2, c_3^3, c_4^4, \cdots\), generated by layers 1,4,3 and 2, respectively, are transmitted from antenna 1. The spatial interleaver of the TLST can be represented by a cyclic-shift interleaver as follows. If we denote the left-hand side matrix in (3.11) by \(C\), the first column of the transmission matrix \(C'\), is identical to first column of matrix \(C\). The second column of \(C'\) is obtained by a cyclic shift of the second column of \(C\) by one position from the top to bottom. The third and fourth column of \(C'\) is obtained by a cyclic shift of the respective column of \(C\) by two and three positions respectively. While the fifth column of \(C'\) is identical to the fifth column of \(C\) etc. In general, if we denote the entries of \(C'\) by \(c_i'^t\), the mapping of \(c_i^t\) to \(c_i'^t\) can be expressed as

\[
c_i'^t = c_i^t, \quad t' = [(i + t - 2) \mod n_T] + 1
\] (3.12)

The spectral efficiency of the HLST and TLST schemes is \(Rmn_T\), where \(R\) is the code rate and \(m\) is the number of bits in a modulated symbol, while the spectral efficiency of the DLST is slightly reduced due to zero padding in the transmission matrix.

3.3 LST Receivers

In this section, we consider the receiver structures for the layered space-time architecture. In order to simplify the analysis, horizontal layering with binary channel codes and BPSK modulation are assumed. Extension to nonbinary codes and to multilevel modulation schemes is straightforward. An LST structure can be viewed as a synchronous code division multiple access (CDMA) in which the number of transmit antennas is equal to the number of users. Similarly, the interference between transmit antennas is equivalent to multiple access interference (MAI) in CDMA systems, while the complex fading coefficients correspond
3.3 LST Receivers

Figure 3.6: Iterative HLST receiver with single decoder.

Figure 3.7: Iterative HLST receiver with separate decoders.
3.3 LST Receivers

to the spreading sequence. This analogy can be further extended to receiver strategies, so that multiuser receiver structures derived for CDMA can be directly applied to LST systems.

Under this scenario, the optimum receiver for an uncoded LST system is a maximum likelihood (ML) multiuser detector [47] operating on a trellis. It computes ML statistics as in the Viterbi algorithm. The complexity of this algorithm is exponential to the number of the transmit antennas.

For coded LST schemes, the optimum receiver performs joint detection and decoding on an overall trellis obtained by combining the trellis of the layered space-time code and the channel code. The complexity of the receiver is an exponential function of the product of the number of the transmit antennas and the code memory order. For many systems, the exponential increase in implementation complexity may make the optimal receiver impractical even for a small number of transmit antennas. Thus, we will examine a number of less complex receiver structures which have good performance/complexity trade-offs. Block diagrams of various LST receivers are shown in Figs. 3.6, 3.7 and 3.8.
3.3 LST Receivers

3.3.1 VLST Receiver

The original VLST receiver [45] is based on a combination of interference suppression and cancellation. Conceptually, each transmitted sub-stream is considered in turn to be the desired symbol and the remainder are treated as interferers. These interferers are suppressed by a zero forcing (ZF) approach [45]. This detection algorithm produces ZF based decision statistics for a desired sub-stream from the received signal vector \( r \), which contains residual interference from other transmitted sub-streams. Subsequently, a decision on the desired sub-stream is made from the decision statistics and its interference contribution is regenerated and subtracted out from the received vector \( r \). Thus \( r \) contains a lower level of interference and this will increase the probability of correct detection of other sub-streams. This operation is illustrated in Fig. 3.9. The ZF strategy is only possible if the number of receive antennas is as large as the number of transmit antennas. Another drawback of this approach is that achievable diversity depends on a particular layer. If the ZF strategy is used in removing interference and if \( n_R \) receive antennas are available, it is possible to remove

\[
n_i = n_R - d_0 \tag{3.13}
\]

interferers with diversity order of \( d_0 \) [51]. The diversity order can be expressed as

\[
d_0 = n_R - n_i \tag{3.14}
\]

If the interference suppression starts at layer \( n_T \), then at this layer \((n_T - 1)\) interferers need to be suppressed. Assuming that \( n_R = n_T \), the diversity order in this layer, according to (3.13) is 1. In the first layer, there are no interferers to be suppressed, so the diversity order is \( n_R = n_T \). As different layers have different diversity orders, diagonal layering is required to achieve equal performance of various encoded streams.

3.3.2 QR Decomposition Interference Suppression Combined with Interference Cancellation

Any \( n_R \times n_T \) matrix \( H \), where \( n_R \geq n_T \), can be decomposed as

\[
H = U_R R, \tag{3.15}
\]
Figure 3.9: VLST detection based on combined interference suppression and successive cancellation.
where $U_R$ is a unitary matrix and $R$ is an $n_T \times n_T$ upper triangular matrix. This is called QR factorization. Let us introduce an $n_T$-component column matrix $y$ obtained by multiplying the receive vector $r$, given by Eq. (3.2), by $U_R^T$

$$y = U_R^T r$$

or

$$y = U_R^T Hx + U_R^T n$$

Substituting the QR decomposition of $H$ from (3.15) into (3.17), we get for $y$

$$Rx + n'$$

where $n' = U_R^T n$ is an $n_T$-component column matrix of i.i.d AWGN noise signals. As $R$ is upper-triangular, the $i$-th component in $y$ depends only on the $i$-th and higher layer transmitted symbols at time $t$, as follows

$$y_t^i = (R_{i,i})_t x_t^i + n_t^i + \sum_{j=i+1}^{n_T} (R_{i,j})_t x_t^j$$

(3.19)

Consider $x_t^i$ as the current desired detected signal. Eq. (3.19) shows that $y_t^i$ contains a lower level of interference than in the received signal $r_t$, as the interference from $x_t^l$, for $l < i$, are suppressed. The third term in (3.19) represents contributions from other interferers, $x_t^{i+1}, x_t^{i+2}, \cdots, x_t^{n_T}$, which can be cancelled by using the available decisions $\hat{x}_t^{i+1}, \hat{x}_t^{i+2}, \cdots, \hat{x}_t^{n_T}$, assuming that they have been detected. The decision statistics on $x_t^i$, denoted by $y_t^i$, can be rewritten as

$$y_t^i = \sum_{j=1}^{n_T} (R_{i,j})_t x_t^j + n_t^i \quad i = 1, 2, \cdots, n_T$$

(3.20)

The estimate on the transmitted symbol $x_t^i$ is given by

$$\hat{x}_t^i = q \left( \frac{y_t^i - \sum_{j=i+1}^{n_T} (R_{i,j})_t \hat{x}_t^j}{(R_{i,i})_t} \right) \quad i = 1, 2, \cdots, n_T$$

(3.21)

where $q(x)$ denotes the hard decision on $x$. 

37
3.3.3 Interference Minimum Mean Square Error (MMSE) Suppression Combined with Interference Cancellation

In the MMSE detection algorithm, the expected value of the transmitted vector $x$ and a linear combination of the received vector $w^H r$, is minimized

$$\min E(x - w^H r)^2$$ (3.22)

where $w$ is an $n_R \times n_T$ matrix of linear coefficients given by $[xx]$

$$w^H = [H^H H + \sigma^2 I_{n_T}]^{-1} H^H$$ (3.23)

$\sigma^2$ is the noise variance and $I_{n_T}$ is an $n_T \times n_T$ identity matrix. The decision statistics for the symbol sent from antenna $i$ at time $t$ is obtained as

$$y^i_t = w^H_i r$$ (3.24)

where $w^H_i$ is the $i$-th row of $w^H$ consisting of $n_R$ components. The estimate of the symbol sent by antenna $i$, denoted by $\hat{x}^i_t$, is obtained by making a hard decision on $y^i_t$

$$\hat{x}^i_t = q(y^i_t)$$ (3.25)

In an algorithm with interference suppression only, the detector calculates the hard decision estimates by using (3.24) and (3.25) for all transmit antennas.

In combined interference suppression and interference cancellation, the receiver starts from antenna $n_T$ and computes its signal estimate by using (3.24) and (3.25). The received signal $r$ in this level is denoted by $r^{n_T}$ and this modified received signal denoted by $r^{n_T-1}$ is used in computing the decision statistics for antenna $(n_T-1)$ in Eq. (3.24) and its hard estimate from (3.25). In the next level, corresponding to antenna $(n_T-2)$, the interference from $(n_T-1)$ is subtracted from the received signal $r^{n_T-1}$ and this signal is used to calculate the decision statistics in (3.24) for antenna $(n_T-2)$. This process continues for all other levels up to the first antenna.

After detection of level $i$, the hard estimate $\hat{x}^i_t$ is subtracted from the received signal to remove its interference contribution, giving the received signal for level $i-1$

$$r^{i-1} = r^i - \hat{x}^i_t h_i$$ (3.26)
where \( h_i \) is the \( i \)-th column in the channel matrix \( H \), corresponding to the path attenuations from antenna \( i \). The operation \( \hat{x}_i^t h_i \) in (3.26) replicates the interference contribution caused by \( \hat{x}_i^t \) in the received vector. \( r^{i-1} \) is the received vector free from interference coming from \( \hat{x}_i^n, \hat{x}_i^{n-1}, \ldots, \hat{x}_i^1 \). For estimation of the next antenna signal \( x_{i-1}^t \), this signal \( r^{i-1} \) is used in (3.24) instead of \( r \). Finally, a deflated version of the channel matrix is calculated, denoted by \( H_{d}^{i-1} \), by deleting column \( i \) from \( H_d^i \). The deflated matrix \( H_{d}^{i-1} \) at the \( (n_T - i + 1) \)th cancellation step is given by

\[
H_{d}^{i-1} = \begin{bmatrix}
h_{1,1} & h_{1,2} & \cdots & h_{1,i-1} \\
h_{2,1} & h_{2,2} & \cdots & h_{2,i-1} \\
\vdots & \vdots & \ddots & \vdots \\
h_{n_R,1} & h_{n_R,2} & \cdots & h_{n_R,i-1}
\end{bmatrix} \tag{3.27}
\]

This deflation is needed as interference associated with the current symbol has been removed. This deflated matrix \( H_{d}^{i-1} \) is used in (3.23) for computing the MMSE coefficients and the signal estimate from antenna \( i-1 \). Once the symbols from each antenna have been estimated, the receiver repeats the process on the vector \( r_{t+1} \) received at time \( (t + 1) \).

The receiver can be implemented without the interference cancellation step (3.26). This will reduce system performance but some computational cost can be saved. With no cancellation, the MMSE coefficients are only computed once, as \( H \) remains unchanged. The most computationally intensive operation in the detection algorithm is the computation of the MMSE coefficients. A direct calculation of the MMSE coefficients based on the channel matrix inversion has a complexity polynomial in the number of transmit antennas. However, on slow fading channels, it is possible to implement adaptive MMSE receivers with the complexity being linear in the number of transmit antennas.

The described algorithm is for uncoded LST systems. The same detector can be applied to coded systems. The receiver consists of the described MMSE interference suppressor/canceller followed by the decoder. The decision statistics, \( y_i^t \), from (3.24), is passed to the decoder which makes a decision on the symbol estimate \( \hat{x}_i^t \). The performance of a QR decomposition receiver (QR), the linear MMSE detector and the performance of the last detected layer in an MMSE detector with successive interference cancellation (MMSE-IC) are shown for VLST structure with \( n_T=4, \ n_R=4 \) and BPSK modulation on a slow Rayleigh fading channel in Fig. 3.10.
Figure 3.10: VLST example, (4,4) Tx-Rx with QR decomposition, MMSE interference suppression and MMSE interference suppression/successive cancellation.
3.3 LST Receivers

3.3.4 Interference Cancellation

The basic idea of interference cancellation is to use estimated symbols to regenerate and subtract out interference from the received signal. Successive (SIC) [52, 53] and parallel (PIC) [54] interference cancellation are two major schemes. The idea of SIC is to order the signal coming from each transmit antenna from strongest to weakest, and to successively demodulate and remove interference from signals in that order. On the other hand, the PIC scheme estimates symbols from all transmit antennas simultaneously.

3.3.4.1 Successive Interference Canceller (SIC)

The key idea in a SIC based receiver is serial cancellation of the transmit antennas interference (TAI), where the individual data streams are successively decoded and stripped away layer-by-layer. The algorithm starts by detecting the symbol of an arbitrarily chosen layer, assuming that the other symbols from the remaining layers are interference. Upon detection of the chosen symbol, its contribution from the received signal vector is subtracted and the procedure is repeated until all symbols are detected. In the absence of error propagation the SIC converts the MIMO channel into a set of parallel SISO channels with increasing diversity order at each successive stage [15, 26]. In practice, error propagation will be encountered, especially in the absence of an adequate temporal coding for each layer. The error rate performance will therefore be dominated by the first stream decoded by the receiver (which is also the stream experiencing the smallest diversity order).

An improved SIC processor is obtained by selecting the stream with the highest signal to interference plus noise ratio (SINR) at each decoding stage. Such receivers are known as ordered SIC (OSIC) receivers or in the MIMO literature as V-BLAST detectors [17, 56], because they have been used successfully for BLAST architectures. OSIC receivers reduce the probability of error propagation by realizing a selection diversity gain at each decoding step.

The OSIC algorithm requires slightly higher complexity than the SIC algorithm resulting from the need to compute and compare the SINRs of the remaining streams at each stage. A major problem with the SIC or the OSIC methods for TAI cancellation is the delay inherent in the implementation of the canceller, since it requires one symbol delay per layer [14]. This
problem may be alleviated to some extent by devising methods that perform interference cancellation in parallel [15].

3.3.4.2 Parallel Interference Canceller (PIC)

In SIC receivers, the interference estimates are created and removed from the received signal before making decisions on the transmitted symbol estimates. In detectors based on PICs, the interference estimation and the cancellation process are executed simultaneously at each stage for all the layers. In [15], it was shown for an asynchronous Direct Sequence CDMA (DS-CDMA) that SIC receivers are superior to PIC receivers in a Rayleigh fading channel without power control. PIC based detectors, on the other hand, exhibit better performance under ideal power control. This is not surprising, since the parallel scheme treats all the users fairly and simultaneously. Therefore, if all the users’ powers at the receiver are the same, they all experience the same amount of interference. When dealing with point to point MIMO architectures, it is plausible to assume that the signal transmitted from different antennas have a similar power at the receiver, or alternatively power control techniques can be easily employed, since all the symbols are transmitted by one user only. Therefore, PIC receivers appear to be a better option than SIC detectors when dealing with TAI.

The original work on PICs in [55] employs standard detection techniques at each stage, such as matched filter (MF) or zero forcing (ZF) detectors, to estimate the MAI. The interference was then simultaneously removed from all the users for the next stage. Between two stages, demodulation of the users’ data and hard decision were performed to regenerate the MAI. Decoding was only performed at the last stage of the cancellation process, if a channel encoder was used. As previously mentioned, this might lead to decision errors at each stage, thus decreasing the reliability of the estimated interference and hence the overall performance of the receiver. If a channel encoder is employed, at each stage, it is possible to perform hard output decoding or soft output decoding of the codewords to obtain a more reliable estimate of the MAI and hence reduce the error propagation. This results in a greater interference cancellation and better performance [25, 28].

However, if decoding is required at each stage, not only is a longer processing delay introduced in the system, but also hardware complexity is increased, as it is necessary to replicate the detector and the decoder as many times as the number of the receiver’s stages. There-
3.3 LST Receivers

Therefore, in an iterative implementation of PIC receivers where the output of the decision device at each iteration is fed back to the PIC for the following iteration, we would require only one realization of the detection chain at the expense of a longer processing delay. A detailed analysis and block diagrams of multistage and iterative detectors based on PICs will be given in the next chapter.
Chapter 4

Weighted Layered Space-Time Code with Iterative Parallel Interference Canceller

There has been a great deal of challenges in the detection of space-time signal; especially to design a low-complexity detector, which can efficiently remove multilayer interference and approach the interference free bound. The application of iterative processing principle to joint detection and decoding has been a promising approach [25, 27, 28]. In [28] it has been shown that, the iterative receiver with parallel interference canceller (PIC) has a linear complexity and near interference free performance.

Besides, as we reviewed the layered space-time architectures in Chapter 3, the different substreams are transmitted with equal power assuming that the transmitter has no knowledge of the channel. But, once the optimized allocated transmit powers for all transmit antennas or the channel state information (CSI) that can be exploited for transmit signal weighting through allocating power across various transmit antennas are available at the transmitter, the performance of layered space time codes can be considerably improved.

In this chapter, we address the problem of the design of a power allocation strategy in LST architecture to simultaneously optimize coding, diversity and weighting gains. A convolutional code is used as the constituent code and an iterative parallel interference canceller (PIC) is employed at the receiver. The receiver is assumed to have the perfect CSI. We will discuss a more practical scenario by assuming imperfect CSI at the receiver in Chapter 5.
The transmit power of various layers is optimized through minimization of the average bit error rate of the LST architecture with a low complexity iterative PIC detector [28].

At the receiver, the PIC detector performs both interference regeneration and cancellation simultaneously for all layers. The iterative decoding principle is applied to pass the a posteriori probability estimates between the detector and decoders. The decoder is based on the maximum a posteriori (MAP) algorithms [38]. Extensive simulation results are provided to validate the performance superiority of the proposed scheme over the conventional LST system without joint optimization of coding, diversity and weighting gain.

The remainder of this chapter is organized as follows. In Section 4.1, we present the system model of proposed LST architecture. The detection and transmit signal weighting technique for an iterative PIC receiver is described in Section 4.2 followed by the simulation results and discussion in Section 4.3. Finally, we present the conclusion in Section 4.4.

4.1 System Model

We consider a wireless system equipped with $n_T$ transmit and $n_R$ receive antennas in a quasi-static flat Rayleigh fading channel. We assume that the fading coefficients remain the same over a frame and change independently from frame to frame. It is assumed that the channel state information (CSI) is perfectly available at the receiver. The information bits are encoded by a convolutional encoder to generate a matrix $C$ of $n_T$ rows. Then each row of $C$ is interleaved independently, modulated and transmitted by a separate antenna. The transmitted symbol of antenna $i$, $1 \leq i \leq n_T$, at time $t$, is denoted by $x^t_i$. $H$ represents the complex channel matrix with dimension $n_R \times n_T$ and entries, denoted by $h_{ji}$, represent the fading coefficient from transmit antenna $i$ to receive antenna $j$ and are assumed to be independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance.

Based on the feedback information, the transmit power $P_i$ ($i = 1, 2, \cdots, n_T$) is assigned to the data symbol $x^t_i$ and the symbol is transmitted through the $i$th antenna. The signal at
4.2 Iterative Detection, Decoding and Transmit Signal Weighting

receive antenna $j$ at time $t$ is given by

$$r^j_t = \sum_{i=1}^{n_T} h_{ji} \sqrt{P_i} x^i_t + n^j_t$$

(4.1)

where, $1 \leq j \leq n_R$ and $n^j_t$ is a statistically independent sample of AWGN with zero mean and variance of $\sigma^2_n$. It is assumed that the receiver determines the transmit power $P_i$ for $n_T$ transmit antennas with the total power constraint $\sum_{i=1}^{n_T} P_i = n_T$, and sends the calculated power coefficients $P_i$ ($i = 1, 2, \cdots, n_T$) to the transmitter through an error-free feedback channel. For the conventional LST system without feedback, $P_i = 1$ for all $i$.

The proposed structure of the LST system is depicted in Fig. 4.1. The information stream is first demultiplexed into $n_T$ parallel sequences, each of which is independently encoded. The encoded sequences are then modulated and fed into a spatial interleaver, $\text{SII}$ and a bank of time interleavers, $\Pi_1, \cdots, \Pi_{n_T}$. The interleaved codewords are transmitted over $n_T$ antennas.

In receiver part, the received sequences are denoted by $r^1, \cdots, r^n_R$ where $r^i = (r^i_1, r^i_2, \cdots, r^i_l)$ represents the received sequence of length $l$ for $i$th receive antenna. The detector estimates the transmitted sequences, $\hat{x}^1, \cdots, \hat{x}^i, \cdots, \hat{x}^{n_T}$, where $\hat{x}^i$ is the estimated sequence transmitted by antenna $i$. After demodulation and de-interleaving, the transmitted sequence is estimated from the detector outputs. The soft output of the decoder is interleaved and feedback to the detector. In the next iteration, the interference caused by all other antennas is cancelled by subtracting the soft decoder output on the symbols from the received signal. The receiver is partitioned into a detector and a decoder for each layer. The iterative principle is applied to pass the probability estimates between the detector and decoders.

4.2 Iterative Detection, Decoding and Transmit Signal Weighting

In this section, we present the iterative detection, decoding and transmit signal weighting scheme for the LST architecture. The iterative parallel interference canceller (PIC) receiver is used for detection and decoding with a convolutional code as the constituent code as in [50].
In this receiver, the detector provides joint soft-decision estimates of the $n_T$ transmitted symbol sequences. Each of the detected sequences is decoded by a separate channel decoder with soft inputs/outputs. At each iteration, the decoder soft outputs are used as estimates of the transmitted signals.

Another important algorithm that determines the receiver complexity is the detector. Hence, architectures that provide low complexity are considered. We consider a standard parallel interference canceller (PIC-STD) [30] and its improved version known as PIC with decision statistics combiner (PIC-DSC) [28]. These detectors are chosen because they offer a good performance-complexity trade-off, particularly when the number of transmit antennas is high and the optimal joint detection and decoding becomes impractical. The contents and equations therein of subsections 4.2.1 and 4.2.2 have been taken from [50].

Finally, the transmit signal weighting scheme for LST architecture with PICs is designed where we consider minimization of bit error rate as an optimization criterion. We assume the perfect channel estimation at receiver.
4.2 Iterative Detection, Decoding and Transmit Signal Weighting

4.2.1 Iterative Detection with PIC-STD

A block diagram of a PIC-STD system is shown in Fig. 4.2. In the first iteration, the PIC detectors are equivalent to a bank of matched filters. The detectors provide decision statistics of the \( n_T \) transmitted symbol sequences. The decision statistics in the first iteration, for antenna \( i \) and time \( t \), denoted by

\[
y_{i,1}^t = \sqrt{P_i} h_i^H r_t
\]  

(4.2)

where \( h_i^H \) is the \( i \)th row of matrix \( H^H \) and \( r_t \) is an \( n_R \) component column matrix of the received signals across the \( n_R \) antennas.

These decision statistics are passed to the respective decoders, which generate soft estimates on the transmitted symbols. In the second and later iterations, the estimated transmitted symbols, from the decoder output is used for interference cancellation at the detector. The PIC-STD detector output in the \( k \)th iteration for the symbol transmitted at time \( t \), for transmit antenna \( i \), denoted by \( y_{i,k}^t \) is given by

\[
y_{i,k}^t = \sqrt{P_i} h_i^H (r_t - H P \hat{X}_{k-1}^{i,t})
\]  

(4.3)

where \( r_t = (r_1^t, \ldots, r_n^t)^T \) is a \( n_R \times 1 \) vector of the received signals and \( \hat{X}_{k-1}^{i,t} \) is a \( n_T \times 1 \) column vector with the symbol estimates from the \((k-1)\)th iteration as elements, except for the \( i \)th element is set to zero. It can be written as

\[
\hat{X}_{k-1}^{i,t} = \left( \hat{x}_{1,i,k-1}^t, \ldots, \hat{x}_{i-1,k-1}^t, 0, \hat{x}_{i+1,k-1}^t, \ldots, \hat{x}_{n_T,k-1}^t \right)^T
\]  

(4.4)

\( P \) represents diagonal transmit power matrix constructed from the feedback information and is given by \( P = diag(\sqrt{P_1}, \sqrt{P_2}, \ldots, \sqrt{P_{n_T}}) \).
4.2 Iterative Detection, Decoding and Transmit Signal Weighting

Figure 4.2: Block diagram of an iterative receiver with PIC-STD.
4.2 Iterative Detection, Decoding and Transmit Signal Weighting

The detection outputs for layer $i$ is then interleaved and passed to the $i$th decoder. The soft estimates of the transmitted BPSK symbols are calculated at the detector output as follows

$$\hat{x}_{i}^{i,k} = 1.P(x_{i}^{i,k} = 1|y_i^{i,k}) + (-1)P(x_{i}^{i,k} = -1|y_i^{i,k})$$ (4.5)

where $y_i^{i,k}$ is a vector of detector outputs for layer $i$ and $P(x_{i}^{i,k} = j|y_i^{i,k})$, $j = 1, -1$, are the symbol a posteriori probabilities (APP) calculated by the decoder in the $k$-th iteration.

Let $\lambda_i^{i,k}$ be the log-likelihood ratios (LLR) in the $k$-th iteration for antenna $i$, at time $t$, defined as

$$\lambda_i^{i,k} = \log \frac{P(x_{i}^{i,k} = 1|y_i^{i,k})}{P(x_{i}^{i,k} = -1|y_i^{i,k})}$$ (4.6)

The symbol a posteriori probabilities $P(x_{i}^{i,k} = j|y_i^{i,k})$, $j = 1, -1$, can then be calculated by $\lambda_i^{i,k}$ as follows

$$P(x_{i}^{i,k} = 1|y_i^{i,k}) = \frac{e^{\lambda_i^{i,k}}}{1 + e^{\lambda_i^{i,k}}}$$ (4.7)

$$P(x_{i}^{i,k} = -1|y_i^{i,k}) = \frac{1}{1 + e^{\lambda_i^{i,k}}}$$ (4.8)

By combining Eqs. (4.5), (4.7) and (4.8), we get for the symbol estimates

$$\hat{x}_{i}^{i,k} = \frac{e^{\lambda_i^{i,k}} - 1}{e^{\lambda_i^{i,k}} + 1}$$ (4.9)

When the LLR is calculated on the basis of the a posteriori probabilities, it is obtained as

$$\lambda_i^{i,k} = \log \frac{\sum_{m,m'=0}^{M_s-1} \alpha_{j-1}(m') p_t(x_i^i = 1) \exp \left(-\frac{\sum_{l=(i-1)n}^{jn} (y_{i}^{i,k} - x_l)^2}{2(\sigma_{i,k}^2)}\right) \beta_j(m)}{\sum_{m,m'=0}^{M_s-1} \alpha_{j-1}(m') p_t(x_i^i = -1) \exp \left(-\frac{\sum_{l=(i-1)n}^{jn} (y_{i}^{i,k} - x_l)^2}{2(\sigma_{i,k}^2)}\right) \beta_j(m)}$$ (4.10)

where $\lambda_i^{i,k}$ denotes the LLR ratio for $p$-th symbol within the $j$-th codeword transmitted at time $t = (j - 1)n + p$ and $n$ is the code symbol length. $m'$ and $m$ are the pair of states connected in the trellis, $x_i^i$ is the $t$-th BPSK modulated symbol in a code symbol in a code symbol connecting the states $m'$ and $m$, $y_i^{i,k}$ is the detector output in iteration $k$, for antenna $i$, at time $t$, $(\sigma_{i,k}^2)$ is the noise variance for layer $i$ and iteration $k$, $M_s$ is the number of states in the trellis and $\alpha(m')$ and $\beta(m)$ are the feed-forward and feedback recursive variables, defined as for LLR.
4.2 Iterative Detection, Decoding and Transmit Signal Weighting

4.2.2 Iterative Detection with PIC-DSC

In computing the LLR value in \( (4.10) \) the decoder uses two inputs. The first input is the decision statistics, \( y_{i,k}^{i,k} \), which depends on the transmitted signal \( x_i^t \). The second input is the a priori probability on the transmitted signal \( x_i^t \), computed as

\[
p_t(x_i^t = l) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y_{i,k}^{i,k} - P_l\mu_{i}^t)^2}{2\sigma^2}}, \quad l = 1, -1
\]

where \( \mu_{i}^t \) is the mean of received amplitude after matched filtering, given by \( \mu_{i}^t = h_i^H h_i \).

When the LLR is calculated on the basis of the APPs, the decision statistics mean value, conditional on \( x_i^t \) is biased due to the correlation of the two inputs of decoder and the bias has always a sign opposite of \( x_i^t \). That is, the bias reduces the useful signal term and degrades the system performance. This bias is particularly significant for a large number of interferers.

The bias effect can be eliminated by estimating the mean of the transmitted symbols based on the a posteriori extrinsic information ratio (EIR) instead of LLR. The EIR does not include the metric for the symbol \( x_i^t \) that is being estimated. That is

\[
\lambda_{i,k}^{t,e} = \log \frac{\sum_{m,m'=0,x_i^t=1}^{m_M-1} \alpha_{j-1}(m')p_t(x_i^t = 1) \exp \left( -\frac{\sum_{l=(j-1)n, l \neq t}^{jn} (y_{i,k}^{i,k} - x_i^l)^2}{2(\sigma_{i,k})^2} \right) \beta_j(m)}{\sum_{m,m'=0,x_i^t=-1}^{m_M-1} \alpha_{j-1}(m')p_t(x_i^t = -1) \exp \left( -\frac{\sum_{l=(j-1)n, l \neq t}^{jn} (y_{i,k}^{i,k} - x_i^l)^2}{2(\sigma_{i,k})^2} \right) \beta_j(m)}
\]

where \( \lambda_{i,k}^{t,e} \) denotes the EIR ratio for the \( p \)-th symbol transmitted at time \( t = (j - 1)n + p \) and within the \( j \)-th codeword, \( n \) is the code symbol length, \( \alpha(m') \) and \( \beta(m') \) are defined as for the LLR [38]. However, excluding the contribution of the bit of interest when estimating the mean of the transmitted symbols reduces the output SNR, leading again to a degraded system performance. A decision statistics combining (DSC) method is effective in minimizing these effects. In the iterative parallel interference canceller with decision statistics combining (PIC-DSC) [28], shown in Fig. 4.3, a DSC module is added to the PIC-STD structure. The decision statistics of the PIC-DSC is obtained at the DSC output, generated as a weighted sum of the current PIC output and the DSC output from the previous operation. The weighting coefficients are estimated by minimizing the output signal-to-noise plus interference ratio (SINR). In each stage, except in the first one, the PIC output for transmit antenna \( i \) and stage \( k \), denoted by \( y_{i,k}^{i,k} \), is passed to the DSC module. The DSC module performs recursive linear combining of the detector output in iteration \( k \) for layer \( i \), denoted by \( y_{i,k}^{i,k} \), with the DSC
output from the previous iteration for the same layer, denoted by \( y_{t,c}^{i,k-1} \). The output of the decision statistics combiner, in iteration \( k \) and for layer \( i \), denoted by \( y_{t,c}^{i,k} \), is given by

\[
y_{t,c}^{i,k} = p_1^{i,k} y_{t,c}^{i,k} + p_2^{i,k} y_{t,c}^{i,k-1}
\]

(4.13)

where \( p_1^{i,k} \) and \( p_2^{i,k} \) are the DSC weighting coefficients in stage \( k \), respectively. They are estimated by maximizing the signal-to-noise plus interference ratio (SINR) at the output of DSC in iteration \( i \) under the assumption that \( y_{t,c}^{i,k} \) and \( y_{t,c}^{i,k-1} \) are Gaussian random variables with the conditional means \( \mu_{i,c}^{i,k} \) and \( \mu_{i,c}^{i,k-1} \), given that \( x_i \) is the transmitted symbol for antenna \( i \), and variances \( (\sigma_{i,c}^{i,k})^2 \) and \( (\sigma_{i,c}^{i,k-1})^2 \), respectively. The maximization of SINR with respect to \( p_1^{i,k} \) and \( p_2^{i,k} \) yields linearly dependant solutions for these coefficients. Thus, these coefficients can be normalized in the following way

\[
E\{y_{t,c}^{i,k}\} = p_1^{i,k} \mu_{i,c}^{i,k} + p_2^{i,k} \mu_{i,c}^{i,k-1} = 1
\]

(4.14)

At the first iteration, the output of the PIC-STD and PIC-DSC layer are same. That is,

\[
y_{t,c}^{i,1} = y_{t,c}^{i,1}
\]

(4.15)

The SINR at the output of the DSC for layer \( i \) and in iteration \( k \) is then given by

\[
\text{SINR}_{i,k} = \frac{P_i}{(p_1^{i,k})^2(\sigma_{i,c}^{i,k})^2 + 2p_1^{i,k} \left( \frac{1-p_1^{i,k}}{\mu_{i,c}^{i,k-1}} \right) \rho_{k,k-1}^{i,k} (\sigma_{i,c}^{i,k-1})^2 + \left( \frac{1-p_1^{i,k}}{\mu_{i,c}^{i,k-1}} \right)^2 (\sigma_{i,c}^{i,k-1})^2}
\]

(4.16)

where \( \rho_{k,k-1}^{i} \) is the correlation coefficient between the detector output in \( k \)-th and \( (k-1) \)-th iteration defined as

\[
\rho_{k,k-1}^{i} = \frac{E\{(y_{t,c}^{i,k} - \mu_{i,c}^{i,k}) (y_{t,c}^{i,k-1} - \mu_{i,c}^{i,k-1}) | x_i \}}{(\sigma_{i,c}^{i,k})^2 (\sigma_{i,c}^{i,k-1})^2}
\]

(4.17)

The optimal combining coefficient is given by

\[
p_1^{i,k}_{\text{opt}} = \frac{(\mu_{i,c}^{i,k})^2 (\sigma_{i,c}^{i,k-1})^2 - \frac{1}{\mu_{i,c}^{i,k-1}} \rho_{k,k-1}^{i,k} (\sigma_{i,c}^{i,k-1})^2}{(\sigma_{i,c}^{i,k})^2 - 2 \frac{\mu_{i,c}^{i,k}}{\mu_{i,c}^{i,k-1}} \rho_{k,k-1}^{i,k} (\sigma_{i,c}^{i,k-1})^2 + \left( \frac{\mu_{i,c}^{i,k}}{\mu_{i,c}^{i,k-1}} \right)^2 (\sigma_{i,c}^{i,k-1})^2}
\]

(4.18)

The parameters required for the calculation of the optimal combining coefficients in Eq. (4.18) are difficult to estimate, apart from the signal variances. However, in a system with a large number of interferers, which happens when the number of the transmit antennas is large
4.2 Iterative Detection, Decoding and Transmit Signal Weighting

Figure 4.3: Block diagram of an iterative receiver with PIC-DSC.
relative to the number of the receive antennas, and for the APP based symbol estimates, the DSC inputs in the first few iterations are low correlated. Thus, it is possible to combine them, in a way similar to receive diversity maximum ratio combining. Under these conditions, the weighting coefficient in this receiver can be obtained from Eq. (4.18) by assuming that the correlation coefficient is zero and neglecting the reduction of the received signal conditional mean caused by interference. The DSC coefficient are then given by

\[ p_{i,k}^1 = \frac{(\sigma_{c}^{i,k-1})^2}{(\sigma_{c}^{i,k-1})^2 + (\sigma_{i,k})^2} \quad (4.19) \]

The DSC output, in the second and higher iterations, with coefficients from (4.19) can be expressed as

\[ y_{i,c}^{i,k} = \frac{(\sigma_{c}^{i,k-1})^2}{(\sigma_{c}^{i,k-1})^2 + (\sigma_{i,k})^2} y_{i,k}^{i,k} + \frac{(\sigma_{i,k})^2}{(\sigma_{c}^{i,k-1})^2 + (\sigma_{i,k})^2} y_{i,c}^{i,k-1} \quad i > 1 \quad (4.20) \]

The complexity of both PIC-STD and PIC-DSC is linear in the number of transmit antennas.

### 4.2.3 Transmit Signal Weighting Algorithm

The post detection signal-to-interference plus noise ratio (SINR) at the output of both PIC-STD and PIC-DSC for layer \( i \) and at first iteration can be calculated from (4.2) as

\[ \text{SINR}^i = \frac{P_i}{\sigma_{T,i}^2} \quad (4.21) \]

\( \sigma_{T,i}^2 \) denotes the total interference plus noise variance for the layer \( i \).

Approximating residual interference as Gaussian as in [56, 57], \( \sigma_{T,i}^2 \) can be estimated as

\[ \sigma_{T,i}^2 = \frac{1}{L} \sum_{i=1}^{L} (y_i^i - \mu_i^i S_i^i)^2 \quad (4.22) \]

where \( S_i^i \) is the transmitted training sequence, \( L \) is the frame size, \( \mu_i^i = h_i h_i^H \) is the nominal mean of the received amplitudes after the maximum-ratio combining (MRC) and \( y_i^i \) is the PIC output. Exploiting the independency of transmitted symbol from one another, it can be calculated the overall bit error rate (BER) \( P_b(e) \) as an arithmetic mean of the BER for every symbol. For BPSK modulation, the overall BER is given by [10]

\[ P_b(e) = \frac{1}{n_T} \sum_{i=1}^{n_T} Q \left( \frac{P_i}{\sigma_{T,i}^2} \right) \quad (4.23) \]
4.2 Iterative Detection, Decoding and Transmit Signal Weighting

where \( Q(x) \) is the error-function. We consider that the receiver has a perfect channel estimate and therefore, the decision boundaries are correctly defined.

Now, our objective is to provide a transmit signal weighting scheme through an effective power allocation across all the transmit antennas minimizing the mean BER subject to the total transmit power constraint given by

\[
\sum_{i=1}^{n_T} P_i = n_T \tag{4.24}
\]

It is a very difficult case to have a closed form solution of the mean BER in terms of computational complexity. The direct minimization exists only for some certain scenario such as in a block transmission system considering additional zero forcing constraints. In this work, we will consider the minimization of the Chernoff upper bound [10] of the mean BER instead of exact mean BER, which is known to be a tight approximation for high SINR. The Chernoff upper bound is defined as:

\[
Q(x) \leq \frac{1}{2} e^{-x^2/2} \tag{4.25}
\]

The solution of the stated minimization problem to find the transmit power \( \{P_i\} \) that minimizes the overall BER in (4.23) under the total transmit power constraints in (4.24) is carried out by means of the Lagrange multiplier method. The cost function \( J \) to be minimized is expressed as

\[
J = P_b(e) + \lambda \left( \sum_{i=1}^{n_T} P_i - n_T \right) \tag{4.26}
\]

where \( J = f(P_1, P_2, \cdots, P_{n_T}) \) and \( \lambda \) is the Lagrange multiplier.

By taking partial derivatives of the cost function \( J \) with respect to transmit weighting coefficients \( P_i \) and equating to zero as \( \frac{\delta J}{\delta P_i} = 0 \), we have a set of \( n_T \) equations

\[
\frac{dQ}{dP_i} \left( \frac{P_i}{\sigma^2_{T,i}} \right) = -n_T \lambda \tag{4.27}
\]

Approximating \( Q \) function to an upper bound as in (4.25) and by substituting this approximated value in (4.27), we have the solutions of transmit signal coefficients as

\[
P_i = -2\sigma^2_{T,i} \ln(4n_T\lambda\sigma^2_{T,i}) \quad i = 1, 2, 3, \cdots, n_T \tag{4.28}
\]
4.3 Numerical Results

The Lagrange multiplier constant, $\lambda$ can be calculated by using the total power constraints $\sum_{i=1}^{n_T} P_i = n_T$ as

$$\lambda = \exp \left( \frac{n_T + 2 \sum_{i=1}^{n_T} \sigma_T^2 \ln(4n_T \sigma_T^2)}{2 \sum_{i=1}^{n_T} \sigma_T^2} \right)$$

(4.29)

4.3 Numerical Results

4.3.1 Transmit Signal Weighting, Iterative Detection with PIC-STD and Decoding

In this section, we illustrate the performance of proposed transmitter and receiver optimization techniques for PIC-STD detection scheme.

A rate 1/2 convolutional component encoders with information length $K=130$ and block length $N=266$ is chosen as the constituent code. The generator polynomial in octal form of this code is $(15, 17)$. The bits are BPSK modulated prior to transmission. In simulations, decoding is performed by a MAP algorithm. The LST scheme with transmit and receive antenna is denoted as an $(n_T, n_R)$ LST. The channel is modeled as frequency flat slow Rayleigh fading channel. The results are shown in the form of frame error rate (FER) versus $E_b/N_0$.

The numbers of iterations used between the detector and the decoder are mentioned in figures. These systems parameters are assumed in all simulations unless otherwise stated.

Figs. 4.4 and 4.5 show the performance for (4,4) and (2,2) LST system with PIC-STD respectively. With (4,4) configuration, proposed scheme attain a gain of slightly less than 2dB where as the SNR gain is about 6dB in case of (2,2) configuration at $10^{-3}$ at final iteration over the conventional LST system without transmit power allocation.

Fig. 4.6 illustrates the performance of a (4,4) LST system on a two-path slow Rayleigh fading channel with and without transmit power allocation where it is also shown that the proposed scheme attain improved FER performance. The overall performance is better than on single path Rayleigh fading channel due to a diversity gain.

With (2,2) LST in Fig. 4.5, the relative SNR gain of proposed scheme over conventional LST
4.3 Numerical Results

Figure 4.4: FER performance of a (4,4) LST PIC-STD system with and without adaptive power allocation.
Figure 4.5: FER performance of a (2,2) LST PIC-STD system with and without adaptive power allocation.
4.3 Numerical Results

Figure 4.6: FER performance of a (4,4) LST PIC-STD system on a two path slow Rayleigh fading channel with and without adaptive power allocation.

is higher than that with (4,4) in Fig. 4.4 over conventional LST. This can be attributed to a higher diversity gain of original (4,4) LST than that of original (2,2) LST scheme. Notably, in both cases, the proposed scheme is observed with faster convergence in fewer number of iterations which is an evidence of the ability of proposed scheme to control the error floor due to the residual interference.
4.3 Numerical Results

4.3.2 Transmit Signal Weighting, Iterative Detection with PIC-DSC and Decoding

In this section, we present the performance of proposed transmitter and receiver optimization techniques for PIC-DSC detection scheme. A rate 1/2 convolutional component encoders with information length $K=130$ and block length $N=266$ is chosen as the constituent code and the bits are BPSK modulated prior to transmission as in the previous section.

![Graph showing FER performance of a (4,4) LST PIC-DSC system with and without adaptive power allocation.](image)

Figure 4.7: FER performance of a (4,4) LST PIC-DSC system with and without adaptive power allocation.

Figs. 4.7 and 4.8 show the performance for (4,4) and (2,2) LST system with PIC-DSC respectively. With (4,4) configuration, proposed scheme attain a gain of slightly less than 2dB where as the SNR gain is about 6dB in case of (2,2) configuration at FER of $10^{-3}$ at final iteration over the conventional LST system without transmit power allocation.
4.3 Numerical Results

It is observed that likewise PIC-STD, the PIC-DSC also achieve the performance superiority with proposed scheme. We can draw the pretty similar interpretation as in PIC-STD technique. With (2,2) LST in Fig. 4.8, the relative SNR gain of proposed scheme over conventional LST is higher than that with (4,4) in Fig. 4.7 over conventional LST. This can be attributed to a higher diversity gain of original (4,4) LST than that of original (2,2) LST scheme. No error floor is seen in the simulation results.

Figure 4.8: FER performance of a (2,2) LST PIC-DSC system with and without adaptive power allocation.
4.3 Numerical Results

4.3.3 Comparison of Various Detection Techniques

In this section, we will compare the performance of the proposed scheme for different detection schemes.

Figs. 4.9 and 4.10 show the performance comparison for (4,4) and (2,2) LST system respectively. From both of these figures, it is clear that the proposed scheme is able to achieve performance superiority over the conventional system without adaptive power allocation.

![Figure 4.9](image-url)

Figure 4.9: Performance comparison of a (4,4) PIC-STD and PIC-DSC LST system with and without adaptive power allocation.
4.3 Numerical Results

Figure 4.10: Performance comparison of a (2,2) PIC-STD and PIC-DSC LST system with and without adaptive power allocation.
4.3 Numerical Results

4.3.4 FER Performance Comparison for Different numbers of Transmit and Receive Antennas

In this section, we will compare the performance of the proposed scheme with various numbers of $n_T$ and $n_R$. We use the same modulation and coding scheme as in the previous section.

![Figure 4.11: Performance comparison of a (4,6) PIC-STD system with and without adaptive power allocation.](image)

Figs. 4.11 and 4.12 show the performance comparison of proposed scheme for (4,6) LST with PIC-STD and PIC-DSC detection scheme respectively where as Fig. 4.13 shows the FER performance of PIC-STD scheme for (2,4) configuration.

Fig. 4.13 shows about 4dB SNR gain of proposed system over the conventional one for
Figure 4.12: Performance comparison of a (4,6) PIC-DSC system with and without adaptive power allocation.

(2,4) LST system in contrast to around 1.25 dB improvement with (4,6) LST configuration illustrated in Fig. 4.11. The large difference of SNR gain of proposed scheme with these two antenna configurations can be understood by the fact that the conventional system achieve enhanced performance with increasing number of receive antennas because with \( n_R > n_T \), the detector can provide a better estimation of the transmitted symbols by each antenna to the channel decoders.
Figure 4.13: Performance comparison of a (2,4) PIC-STD system with and without adaptive power allocation.
4.4 Conclusion

In this chapter, we investigated transmit signal weighting, iterative detection and decoding of LST architecture using a convolutional code as the constituent code. A standard parallel interference canceller (PIC-STD) and an improved version of PIC with decision statistics combiner known as PIC-DSC were used for iterative detection while maximum a posteriori probability (MAP) methods were applied for decoding. Perfect CSI was assumed at receiver. The calculated powers for various transmit antennas were transferred to the transmitter through an error free feedback channel.

Compared to the LST system without adaptive power allocation, detection and decoding technique, the proposed scheme can dramatically improve the overall system performance. It was also demonstrated that the proposed scheme not only can combat the error propagation effect due to residual interference but it can also substantially limit the inter transmit antenna interference.

Small feedback overhead was the another attractive feature of the proposed scheme as the feedback information contains only the transmit power of each antennas rather the full channel state information.
Chapter 5

Weighted Layered Space-Time Code with Imperfect Channel Estimation

In chapter 4, we assumed that the channel is perfectly known at the receiver. This hypothesis does not hold in real systems since perfect channel state information (CSI) is never available in practice neither at receiver nor at transmitter. In a realistic wireless environment, however, CSI has to be periodically estimated due to the random nature of the channel and, because of channel estimation errors, only imperfect CSI can be obtained [58, 59]. As this effect is of paramount importance in practical implementations, it is always necessary to consider the uncertainty in the channel estimation and to quantify the degradation due to channel estimation errors of any system that assumes perfect CSI. However, the meticulous approach is to obtain robust solutions by directly taking into account the existence of channel estimation errors when designing the system.

The performance evaluation and optimization of transmit signal weighting scheme considering channel estimation errors have been received researchers attention [35, 37, 60–65]. The contributions of these previous efforts can be shown into two perspectives viz. consideration of imperfect estimates of the channel impulse response to design the transmitter and design of optimal transmitter schemes based on the knowledge of channel statistics. Design that is based on SVD [66] i.e., beamforming approach, require the full channel state information leading large feedback requirement.
5.1 System Model

In this chapter, we investigate the effect of channel estimation errors on LST architecture with an iterative PIC [28, 30] receiver. It is shown that imperfect channel estimation at an LST receiver results in erroneous decision statistics at the very first iteration and this error propagates to the subsequent iterations, which ultimately leads to severe degradation of the overall performance.

We redefine the transmit power allocation policy to take into account the imperfection in the channel estimate. A closed-form optimal solution in terms of minimum BER is obtained. Small feedback overhead is an attracting feature of the proposed scheme as the feedback information contains only transmit power of each antennas rather full channel state information. Simulation results show the robustness of the proposed scheme against the channel estimation errors.

5.1 System Model

We consider a model of a wireless communication system impaired by slow-variant Rayleigh fading and equipped with \( n_T \) transmit and \( n_R \) receive antennas. \( H \) represents the complex channel matrix with dimension \( n_R \times n_T \) and entries, denoted by \( h_{ji} \), represent the path gain from transmit antenna \( i \) to receive antenna \( j \) and are assumed to be independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance. We assume that the path gains remain constant over a transmitted frame and change independently from one frame to another.

It is assumed that an imperfect estimate of the channel, \( \hat{h}_{ji} \), is available at the receiver. The information bits are encoded by the convolutional code to generate a matrix \( C \) of \( n_T \) rows. Then each row of \( C \) is interleaved independently, modulated and transmitted by a separate antenna. The transmitted symbol of antenna \( i, 1 \leq i \leq n_T \), at time \( t \), is denoted by \( x_i^t \).

Based on the feedback information, the transmit power \( P_i \) \((i = 1, 2, \ldots, n_T)\) is assigned to the data symbol \( x_i^t \) and the symbol is transmitted through the \( i \)th antenna. The received signal at time instant \( t \), is given by

\[
r_t = HPx_t + n_t \tag{5.1}
\]
5.2 Channel Estimation Error Model

where, \( r_t \) is an \( n_R \) component column matrix of the received signals across the \( n_R \) antennas, \( x_t \) is the \( t \)th column in the transmission matrix \( X \) and \( n_t \) is an \( n_R \) component column matrix of AWGN noise signals from the receive antennas with zero mean and variance of \( \sigma_n^2 \). \( P \) represents diagonal transmit power matrix constructed from the feedback information and is given by \( P = diag(\sqrt{P_1}, \sqrt{P_2}, \cdots, \sqrt{P_{n_T}}) \). It is assumed that the receiver determines the transmit power \( P_i \) for \( n_T \) transmit antennas with the total power constraint \( \sum_{i=1}^{n_T} P_i = n_T \), and sends the calculated power coefficients \( P_i \) (\( i = 1, 2, \cdots, n_T \)) to the transmitter through an error-free feedback channel. For the conventional LST system without feedback, \( P_i = 1 \) for all \( i \).

5.2 Channel Estimation Error Model

We now consider a model for channel estimation error as shown in Fig. 5.1. At the receiver, we assume that the channel estimation is conducted by the training symbols based on maximum-likelihood (ML) [10] method. We consider a model for the channel-estimation error using optimal training sequences derived in [39]. We assume that, \( L \geq n_T \) symbols are used for sending known training signals and the dimension of the training matrix \( S = [s_1 \cdots s_L] \) is \( n_T \times L \).

The training symbols subject to the total energy constraint as

\[
\text{tr}(SS^H) = \rho L
\]  

(5.2)

The superscripts \( H \) denotes the conjugate transpose and \( \rho \) is the total transmit energy. The corresponding \( n_T \times L \) matrices of received vectors and noise vectors are, respectively, \( R = [r_1 \cdots r_L] \) and \( N = [n_1 \cdots n_L] \). Taking the channel estimation error into consideration, the MIMO system equation stands during the estimation period then,

\[
R = HS + N
\]  

(5.3)

The maximum-likelihood (ML) [10] estimate of \( H \) is given by

\[
\hat{H} = RS^H(SS^H)^{-1}
\]

\[
= H + NS^H(SS^H)^{-1}
\]

\[
= H + \Delta H \quad \text{where, } \Delta H = NS^H(SS^H)^{-1}
\]  

(5.4)
5.2 Channel Estimation Error Model

Let us define the following notations and assumptions:

- $\xi_i$ is the $i$th column of $\Delta H$.
- The elements of $N$ are the i.i.d. Gaussian variables with variance $N_0$.
- The estimate is unbiased and the estimation error is uncorrelated among the $L$ columns of error matrix $\Delta H$.
- $\Sigma$ is the covariance of error matrix of $\Delta H$.

Hence, $ij$th element, $\Sigma_{i,j}$ of covariance matrix, $\Sigma$ is

$$
\Sigma_{i,j} = E[\xi_i \xi_j^H] = N_0 (SS^H)^{-1} \delta_{i,j}
$$

(5.5)

where $\delta_{i,j}$ is the Kronecker delta function.

It can be shown that the matrix $S$ is of the form $S = \sqrt{\rho L/n_T} U$, where the $n_T \times L$ matrix $U$ satisfies $UU^H = I_{n_T}$.

Figure 5.1: MIMO system with imperfect CSI
5.3 Effect of Channel Estimation Errors in Iterative Detection of LST

The resulting error covariance is

\[ E[\xi_i\xi_j^H] = (N_0n_T/\rho L)I_n \delta_{i,j} \]  \hspace{1cm} (5.6)

Thus, the channel estimation error matrix \( \Delta \mathbf{H} \) can be modeled as a matrix of i.i.d. complex Gaussian variables. To evaluate LST performance with various amount of training symbols, we compute the frame error rate (FER) for different values of the variance \( \sigma_\xi^2 \) of the elements \( \Delta \mathbf{H} \). We define channel estimation uncertainty by

\[ \sigma_\xi^2 = \frac{E[|\Delta \mathbf{H}_{i,j}|^2]}{E[|\mathbf{H}_{i,j}|^2]} \]  \hspace{1cm} (5.7)

From equation (5.5) and (5.7), we can calculate the number of training symbols for a given \( \sigma_\xi^2 \) as

\[ L = \frac{n_TN_0}{E[|\mathbf{H}_{i,j}|^2]\sigma_\xi^2\rho} \]  \hspace{1cm} (5.8)

5.3 Effect of Channel Estimation Errors in Iterative Detection of LST

5.3.1 Analytical Approach

We now examine the effect of channel estimation errors on the iterative detection with a PIC receiver in LST architecture. The decision statistics in the first iteration, for antenna \( i \) and time \( t \) with channel estimation errors is given by

\[ \hat{y}^{i,1}_t = \hat{h}^H_i \mathbf{r} \]
\[ = \hat{h}^H_i(\hat{\mathbf{H}} - \Delta \mathbf{H})\mathbf{x} + \hat{h}^H_i \mathbf{n} \]
\[ = \mu_i^t \mathbf{x} - \hat{h}^H_i \Delta \mathbf{H} \mathbf{x} + \hat{h}^H_i \mathbf{n} \]  \hspace{1cm} (5.9)

where \( \hat{h}_i^H \) is the \( i \)th row of matrix \( \mathbf{H}^H \) and \( \mu_i^t = \hat{h}_i \hat{h}_i^H \) is the nominal mean of the received amplitudes after the maximum-ratio combining (MRC). The second and third part of right hand side in (5.9) are channel estimation errors statistics and filtered noise. These decision
5.3 Effect of Channel Estimation Errors in Iterative Detection of LST

statistics are passed to the respective decoders, which generate soft estimates on the transmitted symbols. In the second and later iterations, the estimated mean of the transmitted symbols, from the decoder output is used for interference cancellation.

The PIC detector output in the $k$th iteration for the symbol transmitted at time $t$, for transmit antenna $i$, denoted by $\hat{y}_t^{i,k}$ is given by

$$\hat{y}_t^{i,k} = \hat{h}_t^H (r - \hat{H} \bar{X}_t^{k-1})$$

(5.10)

where $\bar{X}_t^{k-1}$ is an $n_T \times 1$ column matrix with the symbol estimates from the $(k-1)$th iteration as elements, except for the $i$th element being set to zero. It can be written as

$$\bar{X}_t^{k-1} = (\hat{x}_t^{1,k-1}, \ldots, \hat{x}_t^{i-1,k-1}, 0, \hat{x}_t^{i+1,k-1}, \ldots, \hat{x}_t^{n_T,k-1})^T$$

(5.11)

The detection outputs for layer $i$ is interleaved and then passed to the $i$th decoder. The estimates of the transmitted BPSK symbols are calculated by finding the mean

$$\hat{x}_t^{i,k} = \frac{1}{P(x_t^{i,k} = 1|\hat{y}_t^{i,k})} P(x_t^{i,k} = 1) + \frac{1}{1 - P(x_t^{i,k} = 1|\hat{y}_t^{i,k})} P(x_t^{i,k} = -1|\hat{y}_t^{i,k})$$

(5.12)

where $\hat{y}_t^{i,k}$ is a vector of detector outputs for layer $i$ for a whole block of transmitted symbols and $P(x_t^{i,k} = j|\hat{y}_t^{i,k})$, $j = 1, -1$, are the symbol a posteriori probabilities (APP) calculated by the decoder in the $k$th iteration.

When the LLR is calculated on the basis of the a posteriori probabilities, it is obtained as

$$\lambda_t^{i,k} = \log \frac{\sum_{m,m' = 0, x_t^{i} = 1}^{M_s - 1} \alpha_j - 1(m') p_i(x_t^{i}) \exp \left( -\frac{\sum_{l=1}^{2} (\hat{y}_t^{i,k} - x_t^{i})^2}{2\sigma_t^{i,k}^2} \right) \beta_j(m)}{\sum_{m,m' = 0, x_t^{i} = -1}^{M_s - 1} \alpha_j - 1(m') p_i(x_t^{i}) \exp \left( -\frac{\sum_{l=1}^{2} (\hat{y}_t^{i,k} - x_t^{i})^2}{2\sigma_t^{i,k}^2} \right) \beta_j(m)}$$

(5.13)

where $\lambda_t^{i,k}$ denotes the LLR ratio for $p$-th symbol within the $j$-th codeword transmitted at time $t = (j - 1)n + p$ and $n$ is the code symbol length. $m'$ and $m$ are the pair of states connected in the trellis, $x_t^{i}$ is the $t$-th BPSK modulated symbol in a code symbol in a code symbol connecting the states $m'$ and $m$, $\hat{y}_t^{i,k}$ is the detector output in iteration $k$, for antenna $i$, at time $t$, $(\sigma_t^{i,k})^2$ is the noise variance for layer $i$ and iteration $k$, $M_s$ is the number of states in the trellis and $\alpha(m')$ and $\beta(m)$ are the feed-forward and feedback recursive variables, defined as for LLR.

It is evident from eqs. (5.9), (5.10) and (5.13) that the erroneous decision due to the imperfection in the channel estimation in the first iteration will propagate to the later iterations and hence performance degradation will be occurred.
5.3 Effect of Channel Estimation Errors in Iterative Detection of LST

5.3.2 Illustration By Simulation

The effect of channel estimation error is illustrated by means of extensive simulations in this subsection for the various values of channel uncertainty as defined in equation (5.7). Figs. 5.2, 5.3 and 5.4 show the effect of channel estimation error on iterative 4X4 LST with -10dB, -20dB and -30dB channel uncertainty at receiver respectively.

For 2X2 configuration, this effect is shown in Figs. 5.5 and 5.6 with -20dB and -30dB channel estimation error respectively. Finally, effect of channel estimation errors is examined in Fig. 5.7 for 4X6 transmit-receive antenna configuration with -20dB imperfection in channel estimation. All the simulations in this subsection is carried out with iterative PIC-STD [28, 30] at receiver. The channel uncertainty at receiver is created by varying the number of pilot symbols for channel estimation according to the equation 5.8.

The impact of channel estimation errors for various antenna configuration and different values of channel uncertainty is clearly seen in all these simulation results. The SNR gap observed due to channel imperfection varied from 0.5 dB to 2dB depending upon the value of channel uncertainty and transmit-receive antenna configurations.
Figure 5.2: Effect of channel estimation error for 4X4 configuration with -10dB receiver channel uncertainty.
5.3 Effect of Channel Estimation Errors in Iterative Detection of LST

Figure 5.3: Effect of channel estimation error for 4X4 configuration with -20dB receiver channel uncertainty.
Figure 5.4: Effect of channel estimation error for 4X4 configuration with -30dB receiver channel uncertainty.
5.3 Effect of Channel Estimation Errors in Iterative Detection of LST

Figure 5.5: Effect of channel estimation error for 2X2 configuration.
5.3 Effect of Channel Estimation Errors in Iterative Detection of LST

Figure 5.6: Effect of channel estimation error for 2X2 configuration.
5.3 Effect of Channel Estimation Errors in Iterative Detection of LST

Figure 5.7: Effect of channel estimation error for 4X6 configuration.
In this section, we present a transmit adaptive power allocation technique taking into account that the CSI is noisy and imperfect. Basically, this is a special case of Eq. (4.28) with modified SINR.

With imperfect CSI at LST receiver, the post detection signal-to-interference plus noise ratio (SINR) at the output of PIC for layer \( i \) and at first iteration can be calculated as

\[
\text{SINR}^i = \frac{P_i}{\sigma_{T,i}^2} = \frac{P_i}{\frac{1}{L} \sum_{l=1}^{L} (\hat{y}_i^l - \hat{h}_i^H S_i^l)^2} = \frac{P_i}{\frac{1}{L} \sum_{l=1}^{L} (\hat{y}_i^l - \hat{h}_i^H S_i^l)^2}
\]  

(5.14)

\( \sigma_{T,i}^2 \) denotes the total interference plus noise variance for the layer \( i \).

With this modified SINR, we have the solutions of transmit signal coefficients as in Eq. (4.28)

\[
P_i = -2 \left\{ \frac{1}{L} \sum_{l=1}^{L} (\hat{y}_i^l - \hat{h}_i^H S_i^l)^2 \right\} \ln \left[ 4n_T \lambda \left\{ \frac{1}{L} \sum_{l=1}^{L} (\hat{y}_i^l - \hat{h}_i^H S_i^l)^2 \right\} \right]
\]  

(5.15)

where \( i = 1, 2, 3, \ldots, n_T \).

The Lagrange multiplier constant, \( \lambda \) can be calculated by using the total power constraints \( \sum_{i=1}^{n_T} P_i = n_T \) as

\[
\lambda = \exp \left[ \frac{n_T + 2 \sum_{i=1}^{n_T} \left\{ \frac{1}{L} \sum_{l=1}^{L} (\hat{y}_i^l - \hat{h}_i^H S_i^l)^2 \right\} \ln \left[ 4n_T \left\{ \frac{1}{L} \sum_{l=1}^{L} (\hat{y}_i^l - \hat{h}_i^H S_i^l)^2 \right\} \right]}{2 \sum_{i=1}^{n_T} \left\{ \frac{1}{L} \sum_{l=1}^{L} (\hat{y}_i^l - \hat{h}_i^H S_i^l)^2 \right\}} \right]
\]  

(5.16)
5.5 Numerical Results

In this section, we will enumerate the effectiveness of our proposed transmit signal weighting scheme taking into account the error due to the imperfect channel estimation at the receiver.

A rate 1/2 convolutional component encoders with information length K=130 and block length N=266 is chosen as the constituent code. The generator polynomial in octal form of this code is (15, 17). The bits are BPSK modulated prior to transmission. In simulations, decoding is performed by a MAP algorithm. The LST scheme with transmit and receive antenna is denoted as an \((n_T, n_R)\) LST. The channel is modeled as frequency flat slow Rayleigh fading channel. The results are shown in the form of frame error rate (FER) versus \(E_b/N_0\). The numbers of iterations used between the detector and the decoder are mentioned in figures. These systems parameters are assumed in all simulations unless otherwise stated.

To evaluate the sensitivity of LST system to channel estimation errors, we evaluate the FER performance for a range of channel estimation uncertainty. For a given channel estimation uncertainty, the required number of pilot symbols can be computed according to the equation (5.8).

5.5.1 Performance Results of Robust Transmit Signal Weighting Scheme

Figs. 5.8, 5.9 and 5.10 present the simulation results of proposed scheme versus PIC-STD LST system with -10dB, -20dB and -30dB channel estimation error respectively for 4X4 configuration.
Figure 5.8: FER performance of robust power allocation for 4X4 configuration with -10dB receiver channel uncertainty.
5.5 Numerical Results

From Figs. 5.8, 5.9 and 5.10, it is shown that the proposed scheme achieve around 2dB SNR gain over the 4X4 LST system in all the cases of channel uncertainty from relatively benign case of -30dB to the worst case of -10dB estimation error which demonstrate the robustness of the proposed scheme against the error caused by the channel estimation uncertainty. For 2X2 system, the effectiveness of proposed scheme is verified with -20dB channel estimation error in Fig. 5.11.

![Graph of FER performance of robust power allocation for 4X4 configuration with -20dB receiver channel uncertainty.](image)

Figure 5.9: FER performance of robust power allocation for 4X4 configuration with -20dB receiver channel uncertainty.
Error floor is observed at around $E_b/N_0 = 3 \text{dB}$ for 4X4 Tx-Rx configuration but after achieving a pretty acceptable frame error rate of $10^{-3}$. No error floor is seen for 2X2 system.

Figure 5.10: FER performance of robust power allocation for 4X4 configuration with -30dB receiver channel uncertainty.
5.5 Numerical Results

Figure 5.11: FER performance of robust power allocation for 2X2 configuration with -20dB receiver channel uncertainty.
5.5 Numerical Results

5.5.2 Performance Comparison of Robust Scheme with Original PIC-STD

Figs. 5.12, 5.13 and 5.14 present comparative results of proposed scheme versus PIC-STD LST system with and without channel estimation error. The first two figures demonstrate the comparison for 4X4 system with -20dB and -30dB channel estimation uncertainty while the last one shows the comparison for 4X6 system with -20dB uncertainty.

Figure 5.12: Comparison for 4X4 configuration -10dB receiver channel uncertainty.

87
5.5 Numerical Results

These comparison undoubtedly establish the ability of proposed scheme in combating the channel estimation errors. For 4X4 system, the robust design attain a SNR gain of slightly greater that 0.5dB with the worst case uncertainty of -10dB while it attain a SNR gain of around 1.5dB with the benign case uncertainty of -30dB over original LST with perfect channel estimation at receiver. The robustness of proposed scheme is also tested for 4X6 system where the robust design attain around 1dB SNR gain over original LST system with perfect channel estimation.
Figure 5.13: Comparison for 4X4 configuration -30dB receiver channel uncertainty.
5.5 Numerical Results

Figure 5.14: Comparison for 4X6 configuration with -20dB receiver channel uncertainty.
5.6 Conclusion

In this chapter we presented an effective design of an LST architecture with iterative detection and decoding that are robust to channel estimation errors. The effect of channel estimation errors was investigated. It was shown that imperfect channel estimation at LST receiver results erroneous decision statistics at the very first iteration and this error propagates to the subsequent iterations which ultimately leads to severe degradation of the overall performance.

We derived a closed-form optimal solution for power allocation in terms of the minimum BER under total transmit power constraint. The minimum bit error rate design has been formulated based on a generalized exponential bound of the function $Q(\sqrt{x})$.

Through extensive simulation the robustness of proposed scheme was verified. It was demonstrated that the proposed design can combat the channel estimation errors substantially. The effectiveness was tested with various values of channel uncertainty. It was observed that the new technique is less sensitive to the channel estimation errors.
Chapter 6

Conclusions

This thesis has considered communications through wireless multi-antenna systems which includes specific scenario of layered space-time architecture with an iterative parallel interference canceller at the receiver. An optimal transmit signal optimization, iterative detection and decoding has been reported assuming perfect channel state information. These results were later extended to account for the imperfect CSI due, for example, to channel estimation errors.

6.1 Summary

After giving the motivation of the thesis in Chapter 1, Chapter 2 briefly reviewed the various decoding principles of forward error correcting codes we used in this thesis. In particular, we discussed the optimal decoding of convolutional codes and iterative maximum a posteriori probability (MAP) decoding algorithm.

A review of the LST coding is given in Chapter 3. Various LST transmitters such as VLST, HLST, DLST and TLST and receiver architectures such as QR decomposition interference suppression combined with interference cancellation and MMSE suppression combined with interference cancellation techniques were reviewed.

Chapter 4 investigated a new weighted LST architecture with iterative detection and decod-
6.2 Future Work

In order to further enrich the future work, we suggest several future directions.

- In this thesis, we have presented transmit signal weighting, detection and decoding for narrow band LST system characterized by frequency non-selective flat fading channels. Recently, there has been an increasing interest in providing high data rate services such as video conference, multimedia and mobile computing over wideband wireless channels. In wideband wireless communications, the symbol period becomes smaller relative to the channel delay spread and consequently, the transmitted signals experience frequency-selective fading. Space-time coding technique could be used to...
achieve very high data rates in wideband systems. Therefore, an interesting topic is to investigate the incorporation of our proposed technique to wideband techniques such as OFDM and CDMA system.

- In order to have a complete robust design of transmit power allocation, spatial correlation of MIMO channels can be considered in future line of works.
Bibliography


