

Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be “used for any purpose other than private study, scholarship, or research.” If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of “fair use” that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select “Pages from: first page # to: last page #” on the print dialog screen

The Van Houten library has removed some of the personal information and all signatures from the approval page and biographical sketches of theses and dissertations in order to protect the identity of NJIT graduates and faculty.

ABSTRACT

EM-BASED ITERATIVE CHANNEL ESTIMATION AND SEQUENCE DETECTION FOR SPACE-TIME CODED MODULATION

by
Zbigniew Baranski

Reliable detection of signals transmitted over a wireless communication channel requires knowledge of the channel estimate. In this work, the application of *expectation-maximization* (EM) algorithm to estimation of unknown channel and detection of space-time coded modulation (STCM) signals is investigated. An STCM communication system is presented which includes symbol interleaving at the transmitter and iterative EM-based soft-output decoding at the receiver. The channel and signal model are introduced with a quasi-static and time-varying Rayleigh fading channels considered to evaluate the performance of the communication system. Performance of the system employing Kalman filter with per-survivor processing to do the channel estimation and Viterbi algorithm for sequence detection is used as a reference.

The first approach to apply the EM algorithm to channel estimation presents a design of an on-line receiver with sliding data window. Next, a block-processing EM-based iterative receiver is presented which utilizes soft values of *a posteriori probabilities* (APP) with *maximum a posteriori probability* (MAP) as the criterion of optimality in both: detection and channel estimation stages (APP-EM receiver). In addition, a symbol interleaver is introduced at the transmitter which has a great desirable impact on system performance. First, it eliminates error propagation between the detection and channel estimation stages in the receiver EM loop. Second, the interleaver increases the diversity advantage to combat deep fades of a fast fading channel.

In the first basic version of the APP-EM iterative receiver, it is assumed that noise variance at the receiver input is known. Then a modified version of the receiver

is presented where such assumption is not made. In addition to sequence detection and channel estimation, the EM iteration loop includes the estimation of unknown additive white Gaussian noise variance.

Finally, different properties of the APP-EM iterative receiver are investigated including the effects of training sequence length on system performance, interleaver and channel correlation length effects and the performance of the system at different Rayleigh channel fading rates.

**EM-BASED ITERATIVE CHANNEL ESTIMATION AND SEQUENCE
DETECTION FOR SPACE-TIME CODED MODULATION**

by
Zbigniew Baranski

**A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering**

Department of Electrical and Computer Engineering

August 2002

Copyright © 2002 by Zbigniew Baranski

ALL RIGHTS RESERVED

APPROVAL PAGE

**EM-BASED ITERATIVE CHANNEL ESTIMATION AND SEQUENCE
DETECTION FOR SPACE-TIME CODED MODULATION**

Zbigniew Baranski

Dr. Alexander M. Haimovich, Dissertation Advisor Date
Professor of Electrical and Computer Engineering, NJIT

Dr. ~~Yeheskel~~ Bar-Ness, Committee Member Date
Distinguished Professor of Electrical and Computer Engineering, NJIT

Dr. Zoi-Heleni Michalopoulou, Committee Member Date
Associate Professor of Mathematical Sciences, NJIT

Dr. Ali Abdi, Committee Member Date
Assistant Professor of Electrical and Computer Engineering, NJIT

Dr. Javier Garcia-Frias, Committee Member Date
Assistant Professor of Electrical and Computer Engineering, Univ. of Delaware

BIOGRAPHICAL SKETCH

Author: Zbigniew Baranski
Degree: Doctor of Philosophy
Date: August 2002

Undergraduate and Graduate Education:

- Doctor of Philosophy in Electrical Engineering,
New Jersey Institute of Technology, Newark, NJ, 2002
- Master of Science in Electronic Engineering,
Warsaw Technical University, Warsaw, Poland, 1981

Major: Electrical Engineering

Presentations and Publications:

- Z. Baranski, A.M. Haimovich, J. Garcia-Frias,
“Joint Data Detection, Channel and Noise Estimation for Space-Time Coding,”
submitted to IEEE Transactions on Communications.
- Z. Baranski, A.M. Haimovich, J. Garcia-Frias,
“EM-based Iterative Receiver for Space-Time Coded Modulation with Noise
Covariance Estimation,”
submitted to Global Telecommunications Conference (GLOBECOM), 2002.
- Z. Baranski, A.M. Haimovich, J. Garcia-Frias,
“Iterative Channel Estimation and Sequence Detection for Space-Time Coded
Modulation,”
*36th Conference on Information Sciences and Systems (CISS), Princeton, NJ,
March 2002.*
- Z. Baranski, A.M. Haimovich,
“Performance Analysis of VPSK Digital Modulation,”
Wireless Personal Communications, July 2001.

Blank Page

ACKNOWLEDGMENT

I would like to express my gratitude to my Advisor Dr. Alexander Haimovich for his guidance during the period I worked on this dissertation.

Special thanks also go to Dr. Yehekel Bar-Ness, Dr. Zoi-Heleni Michalopoulou, Dr. Ali Abdi and Dr. Javier Garcia-Frias for their time to participate in my dissertation Committee.

Finally, I would like to express my appreciation and gratitude to the Hashimoto Endowment for the financial support during the last two years of my studies.

TABLE OF CONTENTS

Chapter	Page
1 INTRODUCTION	1
1.1 Outline of the Dissertation	6
1.2 Contributions	7
2 SIGNAL AND CHANNEL MODELS	8
2.1 Model A - Random Constant Channel - Quasi-static Channel Model	9
2.2 Model B - Random Process Channel - Time-varying, Slow and Fast Fading Channel Model	10
2.3 Model C - Gauss-Markov Channel - Dynamic Channel Model	11
2.4 Jakes Channel - Rayleigh Fading Channel Simulation	12
3 KALMAN FILTER-PSP RECEIVER FOR STCM WITH ASD MODIFIED VA METRIC	15
3.1 Kalman Filter-PSP CSI Estimation	15
3.2 Numerical Results	17
4 EM-BASED ON-LINE RECEIVER WITH SLIDING DATA WINDOW	21
4.1 Adaptive Sequence Detector for STCM Signals	21
4.2 Introduction to EM Algorithm with ML Sequence Estimate as Criterion of Optimality	24
4.3 EM-based Sliding Data Window Receiver with ASD	27
4.4 Numerical Results	30
5 EM-BASED BLOCK-PROCESSING ITERATIVE RECEIVER WITH SYMBOL INTERLEAVING	35
5.1 EM Algorithm with MAP Channel Estimate as Criterion of Optimality	35
5.2 APP for STCM	38
5.3 Description of the Iterative Receiver with Symbol Interleaving	45
5.4 Numerical Results	48
6 ITERATIVE APP-EM STCM RECEIVER WITH WHITE NOISE VARIANCE ESTIMATION	52

TABLE OF CONTENTS
(Continued)

Chapter	Page
6.1 Initial Value of Noise Variance	54
6.2 Performance of the Iterative APP-EM Receiver with Noise Variance Estimation - Numerical Results	56
7 INVESTIGATION OF THE APP-EM RECEIVER PROPERTIES	61
7.1 Effects of Training Sequence Length on System Performance	61
7.2 Study of the Interleaver and Channel Correlation Length Effects	62
7.3 Evaluation of the Deterministic Interleaver Gain	63
7.4 Effect of Channel Fading Rate on System Performance	64
8 SUMMARY AND CONCLUSIONS	67
APPENDIX FLOWCHART OF THE SIMULATION PROGRAM	70
BIBLIOGRAPHY	74

LIST OF FIGURES

Figure	Page
2.1 STCM system with $N = 2$ transmit and $M = 2$ receive antennas used in computer simulations.	14
2.2 Trellis of 4-PSK 8-state space-time code used in computer simulations. .	14
3.1 Block diagram of the system using Kalman filter with per-survivor processing for channel estimation.	19
3.2 Kalman filter-PSP receiver - bit error probability vs. SNR per receive antenna for two transmit-two receive antennas, 8-state 4-PSK STCM over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$	19
3.3 Kalman filter-PSP receiver - frame error probability vs. SNR per receive antenna for two transmit-two receive antennas, 8-state 4-PSK STCM over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$	20
4.1 Block diagram of the EM on-line receiver with sliding data window. . . .	32
4.2 EM-based algorithm with sliding channel and data windows.	33
4.3 EM-based sliding windows receiver (EM-SW) - bit error probability vs. SNR per receive antenna for two transmit-two receive antennas, 8-state 4-PSK STCM over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$	33
4.4 EM-based sliding windows receiver (EM-SW) - frame error probability vs. SNR per receive antenna for two transmit-two receive antennas, 8-state 4-PSK STCM over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$	34
5.1 APP algorithm notation describing the STCM trellis states and transitions.	46
5.2 Block diagram of EM-APP block-processing iterative receiver with symbol interleaving.	47
5.3 EM-APP block-processing iterative receiver - bit error probability vs. SNR per receive antenna over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$ - no interleaving, 4 EM iterations.	49
5.4 EM-APP block-processing iterative receiver - bit error probability vs. SNR per receive antenna with known AWGN variance over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$ - with symbol interleaving - 1, 2 and 4 EM iterations.	50

LIST OF FIGURES
(Continued)

Figure	Page
5.5 EM-APP block-processing iterative receiver - frame error probability vs. SNR per receive antenna with known AWGN variance over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$ - with symbol interleaving - 1, 2 and 4 EM iterations.	50
5.6 EM-APP block-processing iterative receiver, AWGN variance known - bit error probability vs. SNR per receive antenna over Rayleigh fading channel, slow fading rate $f_d T_s = 0.001$ - with symbol interleaving - 1, 2 and 4 EM iterations.	51
5.7 EM-APP block-processing iterative receiver, AWGN variance known - frame error probability vs. SNR per receive antenna over Rayleigh fading channel, slow fading rate $f_d T_s = 0.001$ - with symbol interleaving - 1, 2 and 4 EM iterations.	51
6.1 Block diagram of the APP-EM iterative receiver with estimation of unknown white noise variance.	54
6.2 Bit error probability vs. initial value of white noise variance for SNR=8 dB and 12 dB.	56
6.3 APP-EM iterative receiver - bit error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over quasi-static channel.	58
6.4 APP-EM iterative receiver - frame error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over quasi-static channel.	58
6.5 APP-EM iterative receiver - bit error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over Rayleigh fading channel with slow fading rate $f_d T_s = 0.001$	59
6.6 APP-EM iterative receiver - frame error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over Rayleigh fading channel with slow fading rate $f_d T_s = 0.001$	59
6.7 APP-EM iterative receiver - bit error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over Rayleigh fading channel with slow fading rate $f_d T_s = 0.01$	60
6.8 APP-EM iterative receiver - frame error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over Rayleigh fading channel with slow fading rate $f_d T_s = 0.01$	60

LIST OF FIGURES
(Continued)

Figure		Page
7.1	Bit error probability vs. training sequence length - effect of training sequence length on system performance, Rayleigh fading channel $f_d T_s = 0.01$	62
7.2	Bit error probability vs. frame length with fixed percentage of training symbols - fast fading Rayleigh channel $f_d T_s = 0.01$	64
7.3	Bit error probability vs. SNR performance of the system with deterministic and random interleavers, Rayleigh fading channel $f_d T_s = 0.01$, channel and noise variance estimated.	65
7.4	Bit error probability vs. SNR for different Rayleigh channel fading rates $f_d T_s$, channel and noise variance estimated.	66
8.1	Comparison of bit error probability vs. SNR performance over Rayleigh channel with $f_d T_s = 0.01$ for different communication systems considered.	68
8.2	Comparison of frame error probability vs. SNR performance over Rayleigh channel with $f_d T_s = 0.01$ for different communication systems considered.	68
A.1	Simulation program flowchart - main program.	71
A.2	Simulation program flowchart - receiver (1).	72
A.3	Simulation program flowchart - receiver (2).	73

CHAPTER 1

INTRODUCTION

It is well known that multiple time-varying propagation paths of a wireless communication channel result in signal fading and its dispersion in time. Together with interference from other users they severely limit the performance of a wireless communication system. Various diversity techniques can be used to mitigate these impediments and make the reliability of such communication system acceptable in practice. These are: time diversity in the form of coding and interleaving, frequency diversity and spatial receive/transmit diversity where multiple receive and/or transmit antennas are used. Implementation of any of these techniques requires sophisticated signal processing algorithms to be employed in the receiver.

Forney had shown in [1] that *maximum likelihood sequence estimation* (MLSE) based on Viterbi algorithm (VA) [2] is the optimal method for decoding data sequence that has undergone coding and transmission over a dispersive and noisy channel, assuming that the receiver has perfect knowledge of channel state information (CSI). In practical mobile communication system however, the CSI is unknown and it has to be estimated jointly with data detection. Several joint data detection and channel estimation methods were proposed by combining Viterbi algorithm for data detection with various CSI estimation approaches [3, 4, 5, 6].

Usually in digital communication system the data transmission is block-oriented, that is data is transmitted in blocks or frames of known format, containing some known sequence of data symbols (training sequence). In the simplest approach, when the channel is assumed constant over the duration of a frame (quasi-static), the training sequence can be used by the receiver to estimate the CSI. A simple deterministic *least squares* (LS) scheme provides a minimum mean-square error, linear estimate of CSI. The estimated CSI is then used by VA to detect the unknown

information symbols comprising the rest of the frame. In most practical cases however, the length of a frame and the rate of channel time variations are such that the simplest LS estimation method with quasi-static channel model are not adequate and more complex time-varying channel model has to be considered. In such cases, adaptive MLSE receivers using VA for data detection combined with adaptive algorithms for CSI estimation and tracking are appropriate. The proposed adaptive solutions are usually developed from one of the known adaptive algorithms: *least mean squares* (LMS), *recursive least squares* (RLS) or Kalman filter. Among these, Kalman filter incorporating the prior statistical information about the channel is capable of delivering the best performance if the autoregressive process modeling the channel's time behavior is accurate. Kalman filter iterative receivers were proposed in [7, 4, 8].

The adaptive algorithms perform data-aided CSI estimation and consequently the quality of CSI estimation suffers as a result of decision delay inherent in the VA. In an attempt to eliminate the affect of VA decision delay on the quality of CSI estimation, the *per-survivor processing* (PSP) techniques were introduced [9]. In PSP, a separate CSI estimate is maintained and updated along each survivor path in the VA trellis. Adaptive CSI estimation algorithms combined with PSP exhibit improved performance since the CSI estimates are updated with no delay and the CSI along the best survivor path is always based on the right tentative data estimates. Implementation of PSP had shown improvement in channel tracking accuracy in comparison with a single channel estimator [10, 11]. The computational complexity of PSP techniques is considerably increased however. An alternative approach to PSP is presented in [12] where the authors propose an adaptive block processing method to joint data detection and channel estimation featuring less complexity than PSP and performance better than conventional MLSE adaptive algorithm.

In addition to development of more advanced signal processing algorithms to improve the reliability of a communication system, new modulation techniques have

been proposed that combine coding and spatial diversity to mitigate the fading effects of the channel: *space-time coded modulation* (STCM) [13, 14] and BLAST [15, 16]. This work concentrates on STCM, which combines the advantage of space-time diversity of multiple transmit/receive antennas with spectral efficiency of trellis coded modulation. Two different ways of performing CSI estimation and data detection of STCM signals are investigated. First method includes recursive on-line symbol by symbol decoding and CSI estimation incorporating *adaptive sequence detection* (ASD) algorithm [17]. It resulted in acceptable performance and high computational complexity. The search for an alternative method with lower complexity lead to the second, block processing method which maintains only one CSI estimate and reduces the computational complexity as compared with PSP techniques thus, allowing more sophisticated soft decisions detection/CSI estimation algorithm with symbol interleaving to be employed.

The *expectation-maximization* (EM) algorithm [18] is recently becoming an attractive estimation tool in data communications. It is a general method for iterative estimation of parameters of interest according to the selected criterion of optimality. In most published examples of EM, the *maximum likelihood* (ML) is the criterion of optimality. Its applications include joint CSI estimation and sequence detection for various modulation schemes and channel models [19, 20, 21, 22, 23]. It also has been applied to multiuser detection and parameter estimation problem in [24, 25]. Different types of RLS and Kalman filter algorithms are derived from EM in [19, 20] to compute CSI estimate recursively on-line. EM-derived iterative receiver proposed in [26] performs block oriented CSI estimation and VA detection of STCM encoded data symbols. For a flat time-varying channel, each EM iteration produces new CSI estimates based on the current estimate of transmitted STCM symbols. Subsequently, a new VA sequence detection is performed using current CSI estimates. This alternating process is continued until convergence. Over fast fading channels,

good frame error probability (FEP) performance is obtained only when the algorithm is initialized with pilot symbols interspersed throughout the data frame. Good bit error probability (BEP) performance of EM-based receiver is achieved in [23] but the known CSI is assumed as the initial condition to start the EM iterations, which makes such receiver impractical.

Unable to find in the literature the performance characteristics of a Kalman filter-PSP receiver for STCM signals, first such receiver has been implemented. Its performance has been investigated to have a reference for comparing the performance of other EM-based receivers. Next, a recursive on-line algorithm executing EM-derived data detection/CSI estimation on a limited fixed size window of data symbols has been developed. Assuming that data frame consisted of known training sequence followed by data symbols, the initial window was set such that it included the training symbols plus one data symbol to be detected. The EM iterations were run on the channel and data samples within the window using modified VA with ASD metric in detection stage. After convergence, the fixed length window was shifted by one symbol forward along the received data frame and the iterations were repeated. The performance achieved was about 5 dB worse than performance of the clairvoyant (known channel) receiver. It was approximately 2 dB worse than Kalman filter with PSP and the complexity of computations was quite high. In addition, it was noticed (what is also mentioned in [26]) that the BEP (and symbol error probability - SEP) performance of such implementation of the EM algorithm suffered from the effects of error propagation (positive feedback) between the detection and estimation stages. When a symbol error was made in the detection stage, it caused inaccurate CSI estimates in the vicinity of the erroneous symbol. This in turn led to even more symbol errors in this area at the next iteration. In effect, estimation-detection errors accumulated.

Continuing the work, a remedy to this problem was found by implementing a modified block processing EM-based iterative receiver that utilizes soft decisions with *maximum a posteriori probability* (MAP) as the criterion of optimality in the detection process and incorporates a symbol interleaver in the EM loop (APP-EM iterative receiver). It is well known that an interleaver introduces time diversity in transmitted sequence and improves performance of the trellis modulated codes over fading channels [27]. The role of the interleaver is to separate bursts of errors produced by the fading in the channel. It is shown in this work that this property of the symbol interleaver can be very useful when the interleaver is inserted in the iterations of the EM algorithm. The sequence of STCM encoded symbols in the transmitter is passed through an interleaver before being transmitted. At the receiver, the interleaved sequence is used to compute the channel estimate, while detection is done on the deinterleaved sequence. Such approach allows to decouple the time instances when symbol errors are made in the detection stage from the time instances where channel estimation errors occur in the estimation stage. Simulation results show that in the presence of additive white Gaussian noise the performance of the proposed receiver is within approximately 0.5 dB of the clairvoyant receiver for fast fading channel.

In most publications dealing with CSI estimation and sequence detection methods, the assumption is made that the variance of additive white Gaussian noise is known [19, 26, 28]. In the real communication systems, noise variance is unknown and has to be estimated. In this work, a modified version of the iterative APP-EM receiver is also presented in which, in addition to estimating CSI and detecting the data sequence, the unknown white noise variance is also estimated in the EM loop.

1.1 Outline of the Dissertation

In this work, the application of expectation-maximization algorithm to estimation of unknown, fast fading flat channel and detection of space-time coded modulation signals in an iterative manner is investigated.

In Chapter 2, the signal and channel models used throughout the remaining part of this document are introduced.

To have a reference for qualifying the performance of EM-based channel estimation methods, Chapter 3 presents the implementation of Kalman filter with per-survivor processing for estimation/detection of STCM signals corrupted with additive white Gaussian noise.

Next, in Chapter 4, an EM-derived iterative on-line receiver is presented with sliding data window using maximum likelihood as criterion of EM optimality. Combining it with adaptive sequence detection algorithm it adds the capability to do the detection in the presence of correlated additive Gaussian noise.

Then, in Chapter 5 a block-processing APP-EM iterative receiver is proposed which utilizes soft values of data symbols a posteriori probabilities with maximum a posteriori probability as the criterion of optimality in the detection stage. In addition, employing symbol interleaving has a great desirable impact on system performance. First, it eliminates error propagation between the detection and channel estimation stages in the EM loop. Second, the interleaver increases the diversity advantage to combat deep fades of a fast fading channel.

Chapter 6 present a modified version of the APP-EM iterative receiver introduced in Chapter 5. In Chapter 5, the assumption was made that the additive noise at the receiver antennas was white Gaussian with known variance. Here a more practical approach is taken and the noise variance estimation is included in the EM loop.

Chapter 7 presents the results of numerical simulations investigating the performance and properties of the APP-EM receiver.

Finally, Chapter 8 contains the summary of this Dissertation.

1.2 Contributions

- STCM communication system employing symbol interleaver to eliminate error propagation between sequence detection and channel estimation in the EM-based iterative receiver.
- EM algorithm for channel estimation and detection of STCM signals with MAP as a criterion of optimality for both.
- Noise variance estimation in the EM loop of the iterative STCM receiver in addition to channel estimation and sequence detection.
- Investigation of performance and other properties of the iterative EM-based STCM receiver.
- Performance evaluation of STCM receiver using Kalman filter with per-survivor processing for channel estimation and Viterbi algorithm for sequence detection.

CHAPTER 2

SIGNAL AND CHANNEL MODELS

In this chapter, the signal and channel models used throughout the rest of this work are introduced. Assuming that our STCM transmit/receive system has N transmit and M receive antennas, the basic STCM transmitter consists of : 1) Space-time encoder accepting blocks of m binary bits at time k , labeled u_k , where $u_k \in \mathcal{U}$ and $\mathcal{U} = 2^m$. Symbols u_k are encoded into N symbols $c_n(k)$, $n = 1, \dots, N$ following a multiple STCM trellis code. 2) Symbol interleaver which operates on N symbols $c_n(k)$ at a time, generated by the STCM encoder. 3) Memoryless modulator which generates 2^m -PSK complex symbols. 4) N antennas transmitting the modulated symbols $s_n(k)$, $n = 1, \dots, N$ simultaneously. The transmitted signals pass through the multiple input-multiple output (MIMO) channel with unknown CSI, then they are corrupted by additive Gaussian noise and enter the M receive antennas. Data is transmitted in frames of length L symbols and a frame consists of the sequence of known training symbols followed by data symbols. Block diagram of the generic STCM system with $N = 2$ transmit and $M = 2$ receive antennas used in the simulations is shown in Figure 2.1 and the 4-PSK 8-state space-time code presented in [13] in Figure 2.2. Depending on the available channel knowledge, different channel models can be considered when designing an STCM receiver for joint channel estimation and sequence detection.

Channel models used in the receivers investigated in our work are presented in Sections 2.1 - 2.3. In addition, to run the simulations we had to generate the MIMO channels. STCM modulated signals generated in the transmitter passed through the generated simulated channel and after being distorted by additive Gaussian noise, entered the receive antennas. The quasi-static channel was simulated as a complex random variable with mean value 0 and variance equal to $1/M$ per receive antenna. To simulate the time-varying Rayleigh fading channel with different fading rates, the

Jakes channel model was used [29, 30]. Generation of Jakes channel is described in Section 2.4.

2.1 Model A - Random Constant Channel - Quasi-static Channel Model

The measurement model representing signals received at time k at M receive antennas can be expressed:

$$\mathbf{x}_k = \mathbf{S}_k \mathbf{h} + \mathbf{z}_k \quad (2.1)$$

where $\mathbf{x}_k = \begin{bmatrix} x_1(k) & x_2(k) & \dots & x_M(k) \end{bmatrix}^T$ is the $M \times 1$ vector of signals received at M receive antennas at time k and the superscript T denotes transposition. \mathbf{S}_k is $M \times NM$ matrix $\mathbf{S}_k = \mathbf{I}_M \otimes \mathbf{s}_k^T$, $\mathbf{s}_k = \begin{bmatrix} s_1(k) & s_2(k) & \dots & s_N(k) \end{bmatrix}^T$ is $N \times 1$ vector of symbols transmitted by N transmit antennas at time k , \mathbf{I}_M is $M \times M$ identity matrix and \otimes denotes the Kronecker product. The $NM \times 1$ channel vector $\mathbf{h} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{NM} \end{bmatrix}^T$ remains constant for the duration of 1 frame and h_{nm} is the complex scalar channel from transmit antenna n to receive antenna m . Components of \mathbf{h} are random variables. $\mathbf{z}_k = \begin{bmatrix} z_1(k) & z_2(k) & \dots & z_M(k) \end{bmatrix}^T$ is $M \times 1$ vector of noise samples received at M receive antennas at time k .

Description of the full frame of length L symbols is:

$$\mathbf{x} = \mathbf{S} \mathbf{h} + \mathbf{z} \quad (2.2)$$

where $\mathbf{x} = [\mathbf{x}_1^T \quad \mathbf{x}_2^T \quad \dots \quad \mathbf{x}_L^T]^T$ is $LM \times 1$ concatenated vector of received symbols, \mathbf{S} is a $LM \times MN$ matrix $\mathbf{S} = [\mathbf{S}_1^T \quad \mathbf{S}_2^T \quad \dots \quad \mathbf{S}_L^T]^T$ and $\mathbf{z} = [\mathbf{z}_1^T \quad \mathbf{z}_2^T \quad \dots \quad \mathbf{z}_L^T]^T$ is $LM \times 1$ noise vector.

2.2 Model B - Random Process Channel - Time-varying, Slow and Fast Fading Channel Model

The signals received at M receive antennas at time k is expressed:

$$\mathbf{x}_k = \mathbf{S}_k \mathbf{h}_k + \mathbf{z}_k \quad (2.3)$$

where \mathbf{x}_k , \mathbf{S}_k and \mathbf{z}_k are defined the same as in (2.1). The $NM \times 1$ channel vector at time k is denoted $\mathbf{h}_k = \left[h_{11}(k) \ h_{12}(k) \ \dots \ h_{NM}(k) \right]^T$, where $h_{nm}(k)$ is the complex scalar channel from transmit antenna n to receive antenna m at time k . Signal model in (2.3) can be extended to include a full frame of length L :

$$\mathbf{x} = \mathbf{S} \mathbf{h} + \mathbf{z} \quad (2.4)$$

where \mathbf{x} and \mathbf{z} are defined the same as in (2.2), \mathbf{S} is $LM \times LMN$ matrix with diagonal elements $\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_L$. Channel vector of dimension $LMN \times 1$ is $\mathbf{h} = [\mathbf{h}_1^T \ \mathbf{h}_2^T \ \dots \ \mathbf{h}_L^T]^T$. Channel gains $h_{nm}(k)$, $k = 1 \dots L$ form a stationary zero-mean complex Gaussian random process (Rayleigh fading) with autocorrelation $E[h_{nm}(k)h_{nm}^*(k+l)] = r_l \delta_{im} \delta_{jn}$ where $\delta_{im} = 1$ if $i = m$ and $\delta_{im} = 0$ otherwise. In matrix notation, a single channel covariance matrix is:

$$\begin{aligned} \mathbf{R} &= E[\mathbf{h}_{nm} \mathbf{h}_{nm}^*] \\ &= \begin{bmatrix} r_0 & r_1 & \dots & r_{L-1} \\ r_{-1} & r_0 & \dots & r_{L-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{-L+1} & r_{-L+2} & \dots & r_0 \end{bmatrix} \end{aligned} \quad (2.5)$$

where $\mathbf{h}_{nm} = [h_{nm}(1) \ h_{nm}(2) \ \dots \ h_{nm}(L)]^T$. Noting that:

$$\begin{aligned} E \left\{ \left[\begin{array}{c} \mathbf{h}_1 \\ \mathbf{h}_2 \end{array} \right] \left[\begin{array}{cc} \mathbf{h}_1^* & \mathbf{h}_2^* \end{array} \right] \right\} &= \\ &= \begin{bmatrix} r_0 \mathbf{I}_2 & r_1 \mathbf{I}_2 \\ r_{-1} \mathbf{I}_2 & r_0 \mathbf{I}_2 \end{bmatrix} = \\ &= \mathbf{R} \otimes \mathbf{I}_2 \end{aligned} \quad (2.6)$$

In general, the MIMO $MNL \times MNL$ channel covariance matrix is:

$$\mathbf{K} = E [\mathbf{h}\mathbf{h}^H] = \mathbf{R} \otimes \mathbf{I}_{MN} \quad (2.7)$$

Channel density function is:

$$p(\mathbf{h}) = \pi^{-MNL} |\mathbf{K}|^{-1} \exp[-\mathbf{h}^H \mathbf{K}^{-1} \mathbf{h}] \quad (2.8)$$

2.3 Model C - Gauss-Markov Channel - Dynamic Channel Model

Dynamic channel model is given by:

$$h_{nm}(k) = -ah_{nm}(k-1) + w(k) \quad (2.9)$$

With channel vector at time k defined the same as in (2.3), the process model is given by:

$$\mathbf{h}_k = \mathbf{F}\mathbf{h}_{k-1} + \mathbf{w}_k \quad (2.10)$$

where \mathbf{w}_k is $MN \times 1$ process noise and \mathbf{F} is the $MN \times MN$ matrix:

$$\mathbf{F} = -a\mathbf{I}_{MN} \quad (2.11)$$

The process noise vector is $\mathbf{w}_k = \begin{bmatrix} w_{11}(k) & w_{12}(k) & \cdots & w_{NM}(k) \end{bmatrix}^T$. By assumption, \mathbf{w}_k is complex-valued, Gaussian, zero-mean and

$$E[\mathbf{w}_k \mathbf{w}_k^H] = \beta \mathbf{I}_{MN} \quad (2.12)$$

where β is the variance of the white noise process.

The STCM signal observation model at time k is:

$$\mathbf{x}_k = \mathbf{S}_k \mathbf{h}_k + \mathbf{z}_k \quad (2.13)$$

where \mathbf{x}_k , \mathbf{S}_k and \mathbf{z}_k are defined the same as in (2.1).

2.4 Jakes Channel - Rayleigh Fading Channel Simulation

The Jakes fading channel model is a deterministic method introduced in [29] and modified in [30] for simulating time correlated Rayleigh fading waveforms. The model assumes that D equal-strength rays arrive at a moving receiver with uniformly distributed arrival angles α_n such that ray n experiences a Doppler shift $\omega_n = \omega_d \cos \alpha_n$, where $\omega_d = \frac{2\pi f_c \nu}{c}$ is the maximum Doppler shift, ν -vehicle speed in [$\frac{m}{s}$], f_c -carrier frequency in [Hz] and c -speed of light in [$\frac{m}{s}$].

Using $\alpha_n = \frac{2\pi(n-0.5)}{D}$ the fading waveform can be modeled with D_0 complex equal power oscillators, where $D_0 = \frac{D}{4}$. This leads to the following waveform model:

$$T(t) = \sqrt{\frac{2}{D_0}} \sum_{n=1}^{D_0} [\cos \beta_n + j \sin \beta_n] \cos(\omega_n t + \theta_n) \quad (2.14)$$

where $j = \sqrt{-1}$.

If D_0 (and D) is large enough, we may invoke the Central Limit Theorem to conclude that $T(t)$ is approximately a complex Gaussian process, so that $|T|$ is Rayleigh distributed. The normalization factor $\sqrt{\frac{2}{D_0}}$ gives $E\{T(t)T^*(t)\} = 1$. By using $\beta_n = \pi n D_0$, the real and imaginary parts of $T(t)$ have equal power and are

uncorrelated. Making θ_n a uniformly distributed random number in the range $0 \dots 2\pi$ it provides for different waveform realizations.

The time autocorrelation of $T(kT_s)$ can be closely approximated by the zeroth-order first-kind Bessel function $R_k = J_0(k2\pi f_d T_s)$ where the product $f_d T_s$ is the waveform fading rate and T_s is one symbol time.

To generate L consecutive samples of the Rayleigh fading channel with the fading rate $f_d T_s$, the following algorithm steps have to be executed:

1. Given the carrier frequency f_c , vehicle speed ν and baud rate $R_s = \frac{1}{T_s}$ compute the maximum Doppler shift ω_d . The fading rate is $f_d T_s = \frac{\omega_d}{2\pi} T_s$.
2. Choose D_0 , according to [29] D_0 should be large, not less than 8,
3. Compute vectors α , β , θ , ω and \mathbf{S} each of dimension $D_0 \times 1$:

for $n = 1 : D_0$

$$\alpha(n) = \frac{2\pi}{D_0}(n - 0.5)$$

$$\beta(n) = \frac{\pi}{D_0}n$$

$\theta(n)$ - generate uniformly distributed number on the range $0 \dots 2\pi$

$$\omega(n) = \omega_d \cos \alpha(n)$$

$$S(n) = \cos \beta(n) + j \sin \beta(n)$$

end

4. Compute vector \mathbf{h} of dimension $L \times 1$ with L samples of the fading waveform:

for $k = 1 : L$

$$h(k) = \sqrt{\frac{2}{D_0}} \sum_{n=1}^{D_0} S(n) \cos [\omega(n)(k - 1)T_s + \theta(n)]$$

end

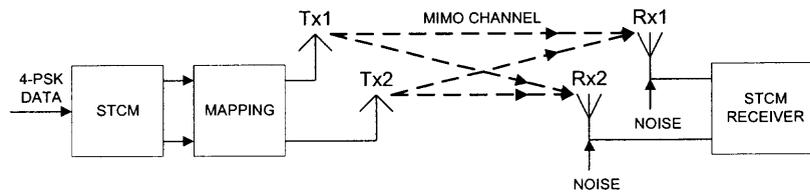


Figure 2.1 STCM system with $N = 2$ transmit and $M = 2$ receive antennas used in computer simulations.

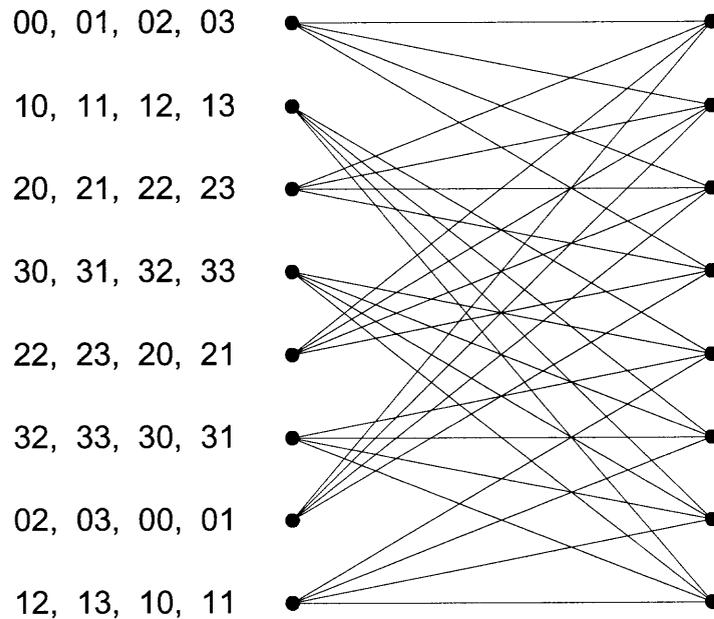


Figure 2.2 Trellis of 4-PSK 8-state space-time code used in computer simulations.

CHAPTER 3

KALMAN FILTER-PSP RECEIVER FOR STCM WITH ASD MODIFIED VA METRIC

Several different implementations of CSI estimation based on Kalman filter approach have been published in literature over the years. In general Kalman filter theory, the estimate of the system state (in our case CSI) is computed recursively using the previous state estimate and the new input data. The Kalman filter theory was applied to do the estimation of time-varying MIMO channel in the receiver of STCM encoded signals. The detection was done using the standard Viterbi algorithm.

3.1 Kalman Filter-PSP CSI Estimation

Gauss-Markov channel description presented in Section 2.3 is used to model the channel estimated by Kalman filter. For the purpose of CSI estimation, the behavior of the system is described by two equations given in Section 2.3 and repeated here for convenience:

1. Process equation:

$$\mathbf{h}_k = \mathbf{F}\mathbf{h}_{k-1} + \mathbf{w}_k \quad (3.1)$$

2. Measurement equation:

$$\mathbf{x}_k = \mathbf{S}_k\mathbf{h}_k + \mathbf{z}_k \quad (3.2)$$

Following the Kalman filter theory [58] the set of equations presented below computes the *mean square* estimate of the CSI $\hat{\mathbf{h}}_{k+1}$ to be used by VA detection stage

at time $k + 1$:

$$\mathbf{R}_k = \mathbf{S}_k \mathbf{K}_{k|k-1} \mathbf{S}_k^H + \mathbf{C}_k \quad (3.3)$$

$$\mathbf{G}_k = \mathbf{F} \mathbf{K}_{k|k-1} \mathbf{S}_k^H \mathbf{R}_k^{-1} \quad (3.4)$$

$$\alpha_k = \mathbf{x}_k - \mathbf{S}_k \hat{\mathbf{h}}_k \quad (3.5)$$

$$\hat{\mathbf{h}}_{k+1} = \mathbf{F} \hat{\mathbf{h}}_k + \mathbf{G}_k \alpha_k \quad (3.6)$$

$$\mathbf{K}_k = \mathbf{K}_{k|k-1} - \mathbf{F} \mathbf{G}_k \mathbf{S}_k \mathbf{K}_{k|k-1} \quad (3.7)$$

$$\mathbf{K}_{k+1|k} = \mathbf{F} \mathbf{K}_k \mathbf{F}^H + \mathbf{Q}_k \quad (3.8)$$

where

\mathbf{R}_k is $M \times M$ correlation matrix of the innovations process at time k ,

\mathbf{S}_k and $\hat{\mathbf{h}}_k$ are the same as \mathbf{S}_k and \mathbf{h}_k in Equation (2.13),

$\mathbf{K}_{k|k-1}$ is $MN \times MN$ correlation matrix of the predicted state error (CSI error),

\mathbf{C}_k is $M \times M$ correlation matrix of the Gaussian additive noise \mathbf{z}_k ,

\mathbf{G}_k is $MN \times M$ Kalman gain matrix,

\mathbf{F} is $MN \times MN$ channel state transition matrix,

\mathbf{Q}_k is $MN \times MN$ correlation matrix of the process noise,

$\hat{\mathbf{h}}_k$ is $MN \times 1$ predicted system state (CSI) vector,

\mathbf{x}_k is $M \times 1$ observation vector,

α_k is $M \times 1$ innovations process vector.

In a standard receiver implementation using VA for signal detection and Kalman filter based CSI estimator, the order of operations is as follows: assuming that at any time k the CSI estimate $\hat{\mathbf{h}}_k$ is already known and VA survivor paths entering each trellis node and their metrics are already computed, the equations (3.3) through (3.8) are executed using the data \mathbf{S}_k associated with the survivor path having the lowest metric. The result is $\hat{\mathbf{h}}_{k+1}$ - the CSI estimate at time $k + 1$ which VA will use to find the next symbol extending the survivor path from time k to $k + 1$.

As mentioned earlier, the quality of detection and CSI estimation can be significantly improved at the expense of increased computational complexity by introducing PSP technique. With PSP, a separate CSI estimate is maintained for each survivor path entering each trellis node at all times. At any given time k , the set of equations (3.3) through (3.8) are executed for each survivor path using the previously computed data $(\mathbf{S}_k, \hat{\mathbf{h}}_k)$ along that path. Consequently, the result is as many new $\hat{\mathbf{h}}_{k+1}$ -s as there are survivor paths (trellis nodes). The final decision on the most likely data symbol and CSI vector is made with the delay D symbols (in practice $D \geq 5$), which means that at time k the decision is made on symbol $k - D$. The symbol at time $k - D$ is chosen which belongs to the survivor path having the lowest VA metric at time k . Alternatively, such decision on the most likely data sequence of length L can be made at the end of the frame.

3.2 Numerical Results

Numerical simulations of the Kalman filter-PSP algorithm were run using the 4-PSK 8-state space-time code presented in [13] and shown in Figure 2.2 with $N = 2$ transmit and $M = 2$ receive antennas. The block diagram of the Kalman filter-PSP system is shown in Figure 3.1. The transmitted frame consisted of $L_{Tr} = 14$ training symbols followed by $L_D = 116$ data symbols to the total frame length of $L = 130$ symbols. The MIMO channel consisted of independent flat paths with Rayleigh distribution represented by Jakes model with autocorrelation function being the zeroth-order first-kind Bessel function $r_k = J_0(k2\pi f_d T_s)$, [29, 30] and the fading rate $f_d T_s = 0.01$. To obtain the initial conditions for starting the Kalman filter CSI estimation algorithm, first the Least Squares channel estimate was computed for the duration of the training sequence. The assumption was made that during the L_{Tr} training symbols interval the channel was constant. The LS CSI estimate over the training sequence was computed

from [59]:

$$\mathbf{H}_{Tr} = (\mathbf{S}_{Tr}^H \mathbf{S}_{Tr})^{-1} \mathbf{S}_{Tr}^H \mathbf{X}_{Tr} \quad (3.9)$$

where

$$\mathbf{H}_{Tr} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \text{ - CSI matrix over training sequence, dimension } N \times M,$$

$$\mathbf{S}_{Tr} = \begin{bmatrix} \mathbf{s}_1^T & \mathbf{s}_2^T & \cdots & \mathbf{s}_{Tr}^T \end{bmatrix}^T \text{ - matrix of known training symbols, dimension } L_{Tr} \times N,$$

$$\mathbf{X}_{Tr} = \begin{bmatrix} \mathbf{x}_1^T & \mathbf{x}_2^T & \cdots & \mathbf{x}_{Tr}^T \end{bmatrix}^T \text{ - matrix of received data samples corresponding to training sequence, dimension } L_{Tr} \times M.$$

The additive noise was Gaussian and white (AWGN). Subsequent CSI estimates at times $k = L_{Tr} + 1, \dots, L$ were computed from Kalman equations (3.3) through (3.8) for each node in the trellis separately. BEP and FEP performance plots are shown in Figure 3.2 and Figure 3.3 respectively. It can be seen from BEP vs. SNR plot that at $\text{BEP}=10^{-3}$ performance of the Kalman filter-PSP algorithm is about 3 dB worse than the case with known channel. Similarly, the FEP vs. SNR plot shows that at $\text{FEP}=10^{-2}$ the FEP performance of Kalman filter-PSP is about 2 dB worse than the known channel case.

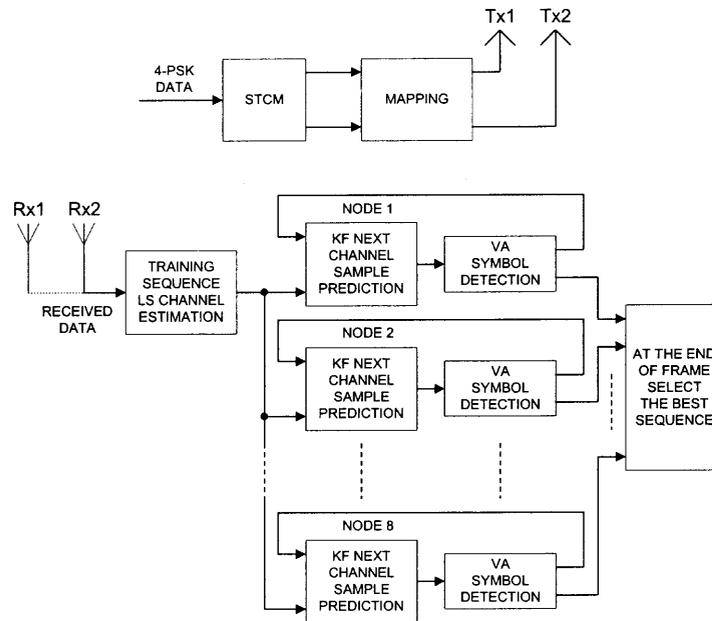


Figure 3.1 Block diagram of the system using Kalman filter with per-survivor processing for channel estimation.

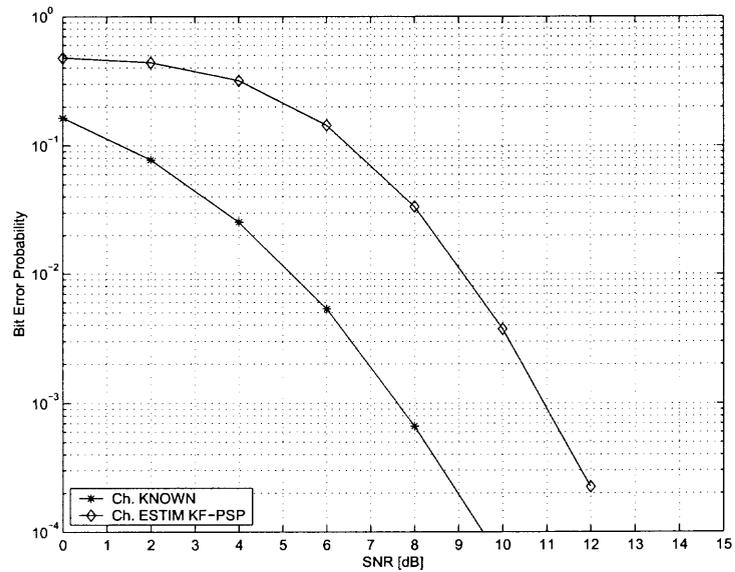


Figure 3.2 Kalman filter-PSP receiver - bit error probability vs. SNR per receive antenna for two transmit-two receive antennas, 8-state 4-PSK STCM over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$.

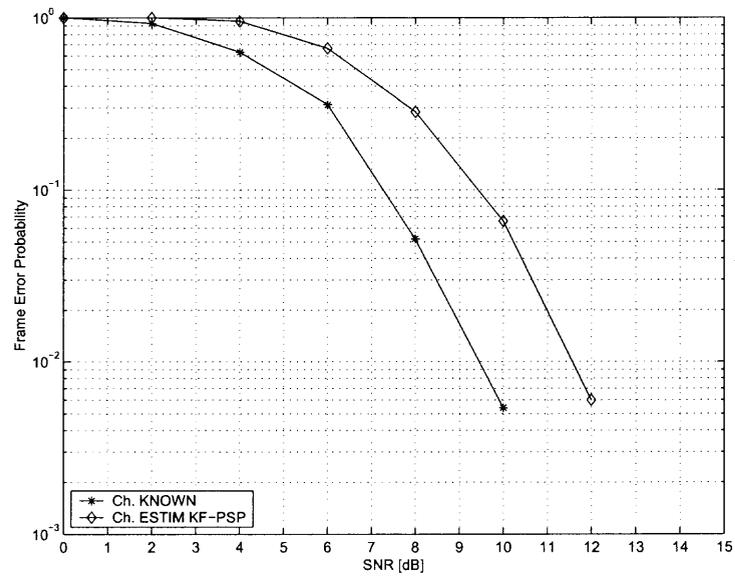


Figure 3.3 Kalman filter-PSP receiver - frame error probability vs. SNR per receive antenna for two transmit-two receive antennas, 8-state 4-PSK STCM over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$.

CHAPTER 4

EM-BASED ON-LINE RECEIVER WITH SLIDING DATA WINDOW

The EM-based receiver with sliding data window incorporating adaptive sequence detector is introduced in this chapter. It was the first attempt to do the sequence detection of STCM signals distorted by unknown fast fading channel and additive correlated Gaussian noise. The noise was modeled as autoregressive process of order $(Q - 1)$. In this approach, the CSI is estimated jointly with the data and initially the covariance matrix of the noise is assumed known. The EM-based on-line receiver with sliding data window can perform the same functions as Kalman filter-PSP described in the previous chapter and in addition it has the capability to handle time correlated noise, that is AWGN with interference.

The next two sections introduce the adaptive sequence detector for STCM, EM algorithm with ML sequence estimation as the criteria of optimality followed by the derivation, functional description and performance evaluation of the sliding data window receiver with ASD.

4.1 Adaptive Sequence Detector for STCM Signals

Adaptive sequence detector is a modification of the VA for sequence detection incorporating space-time processing of the interference. The ASD was initially introduced in [17] for known channels and Gaussian interference with known space-time covariance matrix. We use fast fading channel model B and introduce a modified signal notation for the purpose of presenting the ASD for STCM. The presentation is based on [17].

Rewrite (2.2) - the measurement model for a full frame of length L :

$$\mathbf{x}_{1 \rightarrow L} = \mathbf{S}_{1 \rightarrow L} \mathbf{h}_{1 \rightarrow L} + \mathbf{z}_{1 \rightarrow L} \quad (4.1)$$

where notation $\mathbf{x}_{1 \rightarrow L}$ denotes a sequence of received samples at times $1, \dots, L$.

Define the following space-time quantities for later use: the $MQ \times 1$ vector $\underline{\mathbf{x}}_k = [\mathbf{x}_k^T \quad \mathbf{x}_{k-1}^T \quad \cdots \quad \mathbf{x}_{k-Q+1}^T]^T$, the $MQ \times MNQ$ matrix $\underline{\mathbf{S}}_k$ with diagonal elements $\mathbf{S}_k, \mathbf{S}_{k-1}, \dots, \mathbf{S}_{k-Q+1}$, the $MNQ \times 1$ channel vector $\underline{\mathbf{h}}_k = [\mathbf{h}_k^T \quad \mathbf{h}_{k-1}^T \quad \cdots \quad \mathbf{h}_{k-Q+1}^T]^T$ and the $MQ \times 1$ noise vector $\underline{\mathbf{z}}_k = [\mathbf{z}_k^T \quad \mathbf{z}_{k-1}^T \quad \cdots \quad \mathbf{z}_{k-Q+1}^T]^T$. Then the $MQ \times MQ$ space-time noise covariance matrix can be expressed $\mathbf{C} = cov[\underline{\mathbf{z}}_k]$ and it is positive definite by assumption.

The received signals \mathbf{x}_k conditioned on the transmitted signals \mathbf{S}_k are realizations of a Gaussian random process. The ML detector for the full frame symbol sequence $\mathbf{S}_{1 \rightarrow L}$ can be found from:

$$\widehat{\mathbf{S}}_{1 \rightarrow L} = \arg \max_{\mathbf{S}_{1 \rightarrow L}} p(\mathbf{x}_{1 \rightarrow L} | \mathbf{S}_{1 \rightarrow L}) \quad (4.2)$$

where the maximum is taken over all possible sequences within $\mathbf{S}_{1 \rightarrow L}$.

By Bayes rule, the joint conditional probability density function (pdf) representing the likelihood function can be factored into a product of conditional pdf's:

$$\mathcal{L} = p(\mathbf{x}_{1 \rightarrow L} | \mathbf{S}_{1 \rightarrow L}) = \prod_{k=1}^L p(\mathbf{x}_k | \mathbf{x}_{1 \rightarrow k-1}, \mathbf{S}_{1 \rightarrow L}) \quad (4.3)$$

where the notation $\mathbf{x}_{1 \rightarrow p}$ is for a concatenated vector of dimension $pM \times 1$. Note that in this expression, the current observation \mathbf{x}_k is conditioned on past observations $\mathbf{x}_{1 \rightarrow k-1}$. To continue, recall the assumptions made in the signal model:

- The interference is modeled by an autoregressive process of order $(Q - 1)$. This means that the random vector \mathbf{x}_k is independent on \mathbf{x}_{k-Q} . The densities on the right hand of (4.3) can then be written:

$$p(\mathbf{x}_k | \mathbf{x}_{1 \rightarrow k-1}, \mathbf{S}_{1 \rightarrow L}) = p(\mathbf{x}_k | \mathbf{x}_{k-Q+1 \rightarrow k-1}, \mathbf{S}_{1 \rightarrow L}) \quad (4.4)$$

- The channel is flat fading, implying that since \mathbf{x}_k depends on $\mathbf{x}_{k-Q+1 \rightarrow k-1}$, it also depends on the symbols $\underline{\mathbf{S}}_k = \mathbf{S}_{k-Q+1 \rightarrow k-1}$, where the notation refers to a

$MQ \times MNQ$ matrix with diagonal blocks $\mathbf{S}_k, \mathbf{S}_{k-1}, \dots, \mathbf{S}_{k-Q+1}$. This leads to the conditional density:

$$p(\mathbf{x}_k | \mathbf{x}_{k-Q+1 \rightarrow k-1}, \mathbf{S}_{1 \rightarrow L}) = p(\mathbf{x}_k | \mathbf{x}_{k-Q+1 \rightarrow k-1}, \mathbf{S}_{k-Q+1 \rightarrow k-1}) \quad (4.5)$$

Substituting (4.5) into (4.3) and applying the Bayes rule once more, the factored form of the likelihood function is obtained:

$$\mathcal{L} = \prod_{k=Q}^L \frac{p(\mathbf{x}_{k-Q+1 \rightarrow k} | \mathbf{S}_{k-Q+1 \rightarrow k})}{p(\mathbf{x}_{k-Q+1 \rightarrow k-1} | \mathbf{S}_{k-Q+1 \rightarrow k-1})} \quad (4.6)$$

Maximizing the likelihood function \mathcal{L} in (4.6) is equivalent to minimizing its negative logarithm. Thus, the maximum likelihood detector is given by:

$$\hat{\mathbf{S}}_{1 \rightarrow L} = \arg \min_{\mathbf{S}_{1 \rightarrow L}} \left(\sum_{k=Q}^L T_k(\underline{\mathbf{x}}_k, \underline{\mathbf{S}}_k) \right) \quad (4.7)$$

where the notation was simplified using earlier definitions of $\underline{\mathbf{x}}_k$ and $\underline{\mathbf{S}}_k$, and by observing that $\underline{\mathbf{x}}_k = \mathbf{x}_{k-Q+1 \rightarrow k}$ and $\underline{\mathbf{S}}_k = \mathbf{S}_{k-Q+1 \rightarrow k}$. The quantity $T_k(\underline{\mathbf{x}}_k, \underline{\mathbf{S}}_k)$ is the branch metric used in the Viterbi algorithm. Starting with (4.6) and using the expression for the multivariate Gaussian density:

$$p(\underline{\mathbf{x}}_k | \underline{\mathbf{S}}_k) = \pi^{-MQ} \exp \left(-(\underline{\mathbf{x}}_k - \underline{\mathbf{S}}_k \underline{\mathbf{h}}_k)^H \mathbf{C}^{-1} (\underline{\mathbf{x}}_k - \underline{\mathbf{S}}_k \underline{\mathbf{h}}_k) \right) \quad (4.8)$$

where H denotes Hermitian operation and after some manipulations, it can be shown that the branch metric for the Viterbi algorithm is given by:

$$\begin{aligned} T_k(\underline{\mathbf{x}}_k, \underline{\mathbf{S}}_k) &= (\underline{\mathbf{x}}_k - \underline{\mathbf{S}}_k \underline{\mathbf{h}}_k)^H \mathbf{C}^{-1} (\underline{\mathbf{x}}_k - \underline{\mathbf{S}}_k \underline{\mathbf{h}}_k) \\ &\quad - (\underline{\mathbf{x}}_k - \underline{\mathbf{S}}_k \underline{\mathbf{h}}_k)^H \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{11}^{-1} \end{bmatrix} (\underline{\mathbf{x}}_k - \underline{\mathbf{S}}_k \underline{\mathbf{h}}_k) \\ &= \underline{\mathbf{x}}_k^H \mathbf{A} \underline{\mathbf{x}}_k - \underline{\mathbf{h}}_k^H \underline{\mathbf{S}}_k^H \mathbf{A} \underline{\mathbf{x}}_k - \underline{\mathbf{x}}_k^H \mathbf{A} \underline{\mathbf{S}}_k \underline{\mathbf{h}}_k + \underline{\mathbf{h}}_k^H \underline{\mathbf{S}}_k^H \mathbf{A} \underline{\mathbf{S}}_k \underline{\mathbf{h}}_k \end{aligned} \quad (4.9)$$

where $\mathbf{A} = \mathbf{C}^{-1} - \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{11}^{-1} \end{bmatrix}$ and \mathbf{C}_{11}^{-1} is the inverse of \mathbf{C}_{11} - a cofactor matrix of noise covariance matrix \mathbf{C} .

The expression in (4.9) represents the branch metric of the adaptive sequence detector. The metric amount to a space time processor embedded in the Viterbi algorithm.

Next, two special cases are considered that simplify (4.9) and reduce it to familiar results.

- The interference has no time correlation but is provided by distinct sources (thus it has spatial correlation). In this case, $Q = 1$ and (4.8) simplifies to $p(\underline{\mathbf{x}}_k | \underline{\mathbf{S}}_k) = p(\mathbf{x}_k | \mathbf{S}_k)$, where $p(\mathbf{x}_k | \mathbf{S}_k) = \pi^{-M} |\mathbf{C}|^{-1} \exp(-\mathbf{z}_k^H \mathbf{C}^{-1} \mathbf{z}_k)$, $\mathbf{C} = E[\mathbf{z}_k \mathbf{z}_k^H]$ and \mathbf{z}_k is defined in (2.1). Then (4.9) becomes:

$$T_k(\mathbf{x}_k, \mathbf{S}_k) = \mathbf{z}_k^H \mathbf{C}^{-1} \mathbf{z}_k \quad (4.10)$$

This branch metric accounts only for spatial correlation of the interference.

- The interference is both spatially and temporally white. This is the case treated in most of the literature on space-time coding. In this case, the branch metric reduces to that of a multi-dimensional Viterbi algorithm in white Gaussian noise:

$$T_k(\mathbf{x}_k, \mathbf{S}_k) = \mathbf{z}_k^H \mathbf{z}_k \quad (4.11)$$

4.2 Introduction to EM Algorithm with ML Sequence

Estimate as Criterion of Optimality

Expectation-Maximization algorithm first introduced in [18] found applications in the communication systems as an iterative method for finding ML estimates of parameters

of interest where the complexity of analytical solution makes it impractical because the available data does not provide complete information.

In general, if the observed data vector is denoted as \mathbf{b} and \mathbf{a} the vector of parameters to be estimated from \mathbf{b} then given the conditional density $p(\mathbf{b}|\mathbf{a})$, the ML criterion for the estimation of \mathbf{a} from the observed data \mathbf{b} is given by:

$$\hat{\mathbf{a}} = \arg \max_{\mathbf{a}} p(\mathbf{b}|\mathbf{a}) \quad (4.12)$$

Defining the likelihood function of \mathbf{a} as $\mathcal{L}'(\mathbf{a}) = \log p(\mathbf{b}|\mathbf{a})$ the equivalent expression for ML estimate of \mathbf{a} is:

$$\hat{\mathbf{a}} = \arg \max_{\mathbf{a}} \mathcal{L}'(\mathbf{a}) = \arg \max_{\mathbf{a}} \log p(\mathbf{b}|\mathbf{a}) \quad (4.13)$$

In most of the practical cases, it happens that the expression for conditional density $p(\mathbf{b}|\mathbf{a})$ is difficult to obtain analytically or even if such expression is found, the complexity of maximization in (4.13) is intractable. To apply the EM algorithm, let's assume that there is access to another data vector \mathbf{c} in addition to observed data vector \mathbf{b} . The data \mathbf{c} is chosen such that \mathbf{b} can be obtained from it through many-to-one mapping $\mathbf{b} = \mathbf{f}(\mathbf{c})$ and knowledge of \mathbf{c} makes the maximization:

$$\hat{\mathbf{a}} = \arg \max_{\mathbf{a}} \log p(\mathbf{c}|\mathbf{a}) \quad (4.14)$$

easier than maximization in (4.13). The data \mathbf{c} is referred to as *complete data* and observed data \mathbf{b} as *incomplete data*. The *complete data* set can be expressed as $\mathbf{c} = \{\mathbf{b}, \mathbf{d}\}$ where \mathbf{d} is the *completing data* needed for estimating \mathbf{a} . We note that:

$$p(\mathbf{c}|\mathbf{a}) = p(\mathbf{b}, \mathbf{d}|\mathbf{a}) = p(\mathbf{b}|\mathbf{a}) p(\mathbf{d}|\mathbf{a}, \mathbf{b}) \quad (4.15)$$

Using (4.15) the likelihood of \mathbf{a} is:

$$\mathcal{L}'(\mathbf{a}) = \log p(\mathbf{b}|\mathbf{a}) = \log p(\mathbf{c}|\mathbf{a}) - \log p(\mathbf{d}|\mathbf{a}, \mathbf{b}) = \log p(\mathbf{b}, \mathbf{d}|\mathbf{a}) - \log p(\mathbf{d}|\mathbf{a}, \mathbf{b}) \quad (4.16)$$

$\mathcal{L}'(\mathbf{a})$ has to be maximized with respect to \mathbf{a} but because the *completing data* \mathbf{d} is not known, then instead of computing $\mathcal{L}'(\mathbf{a})$ its conditional expectation is computed with respect to \mathbf{d} given the observed data \mathbf{b} and the last estimate $\hat{\mathbf{a}}^{(l)}$ (or initial estimate $\hat{\mathbf{a}}^{(0)}$) of parameter \mathbf{a} :

$$\mathcal{L}(\mathbf{a}) = E_{\mathbf{d}} \{ \log p(\mathbf{b}, \mathbf{d}|\mathbf{a}) | \mathbf{b}, \hat{\mathbf{a}}^{(l)} \} - E_{\mathbf{d}} \{ \log p(\mathbf{d}|\mathbf{a}, \mathbf{b}) | \mathbf{b}, \hat{\mathbf{a}}^{(l)} \} \quad (4.17)$$

If the completing data \mathbf{d} is chosen such that it is independent on the estimated parameter \mathbf{a} , then the second term in (4.17) is not a function of \mathbf{a} then from the maximization point of view it can be dropped. The subsequent EM iterations cycle consists of two steps:

Expectation:

$$Q(\mathbf{a}|\hat{\mathbf{a}}^{(l)}) = E_{\mathbf{d}} \{ \log p(\mathbf{c}|\mathbf{a}) | \mathbf{b}, \hat{\mathbf{a}}^{(l)} \} = E_{\mathbf{d}} \{ \log p(\mathbf{b}, \mathbf{d}|\mathbf{a}) | \mathbf{b}, \hat{\mathbf{a}}^{(l)} \} \quad (4.18)$$

The next estimate $\hat{\mathbf{a}}^{(l+1)}$ of parameter \mathbf{a} is obtained from the Maximization step:

Maximization:

$$\hat{\mathbf{a}}^{(l+1)} = \arg \max_{\mathbf{a}} Q(\mathbf{a}|\hat{\mathbf{a}}^{(l)}) = \arg \max_{\mathbf{a}} E_{\mathbf{d}} \{ \log p(\mathbf{b}, \mathbf{d}|\mathbf{a}) | \mathbf{b}, \hat{\mathbf{a}}^{(l)} \} \quad (4.19)$$

Taking expectation $E_{\mathbf{d}}[\dots]$ over the completing unknown parameter \mathbf{d} removes it from (4.17). Then $Q(\mathbf{a}|\hat{\mathbf{a}}^{(l)})$ is only a function of $(\mathbf{b}, \mathbf{a}$ and $\hat{\mathbf{a}}^{(l)})$ and not \mathbf{d} .

Consequently, an iterative procedure is obtained for computing a sequence of estimates $\hat{\mathbf{a}}^{(l)}$, $l = 1, 2, \dots$ starting from a known initial estimate $\hat{\mathbf{a}}^{(0)}$. Following [18] it can be shown that the sequence of likelihoods of these estimates $\mathcal{L}(\hat{\mathbf{a}}^{(l+1)})$, $l = 0, 1, \dots$

is monotonically increasing with $\mathcal{L}(\hat{\mathbf{a}}^{(l+1)}) > \mathcal{L}(\hat{\mathbf{a}}^{(l)})$ when $Q(\hat{\mathbf{a}}^{(l+1)}|\hat{\mathbf{a}}^{(l)}) > Q(\hat{\mathbf{a}}^{(l)}|\hat{\mathbf{a}}^{(l)})$. The iterations (4.18) and (4.19) are repeated until convergence, that is until $\hat{\mathbf{a}}^{(l+1)} = \hat{\mathbf{a}}^{(l)} = \hat{\mathbf{a}}$. Then, if $\mathcal{L}(\mathbf{a})$ has only one maximum the ML criterion is satisfied:

$$\hat{\mathbf{a}} = \arg \max_{\mathbf{a}} \mathcal{L}(\mathbf{a}) \quad (4.20)$$

However, in general, there is no guarantee that $\hat{\mathbf{a}}^{(l)}$ converges to the global maximum. The convergence properties of the EM algorithm depend on the density $p(\mathbf{b}|\mathbf{a})$ and also on the choice of the complete data \mathbf{c} and the initial estimate $\hat{\mathbf{a}}^{(0)}$.

4.3 EM-based Sliding Data Window Receiver with ASD

The window of length $(T + 1)$ symbols is defined. At time k it extends over the time samples $k - T, \dots, k$. Applying the generic EM algorithm introduced in Section 4.2 to this case of STCM data detection/CSI estimation over time-varying fast fading channel the following EM related definitions are introduced:

- received data $\mathbf{x}_{k-T \rightarrow k}$ - *incomplete data*,
- CSI vector $\mathbf{h}_{k-T \rightarrow k}$ - *completing data*,
- $\mathbf{y}_{k-T \rightarrow k} = \{\mathbf{x}_{k-T \rightarrow k}, \mathbf{h}_{k-T \rightarrow k}\}$ - *complete data*,
- transmitted sequence $\mathbf{S}_{k-T \rightarrow k}$ - parameter to be estimated.

Similarly to Section 4.1, for clarity we simplify the time indices notation:

- $\underline{\mathbf{x}}_k$ is equivalent to $\mathbf{x}_{k-T \rightarrow k}$,
- $\underline{\mathbf{x}}_{k-1}$ is equivalent to $\mathbf{x}_{k-T \rightarrow k-1}$.

The same time indices notation applies to other variables: $\underline{\mathbf{S}}_k$, $\underline{\mathbf{z}}_k$, $\underline{\mathbf{y}}_k$ and $\underline{\mathbf{h}}_k$.

Signal model can then be written:

$$\underline{\mathbf{x}}_k = \underline{\mathbf{S}}_k \underline{\mathbf{h}}_k + \underline{\mathbf{z}}_k \quad (4.21)$$

The following derivations assume that:

- noise $\underline{\mathbf{z}}_k$ is a random AR process with covariance matrix \mathbf{C} and correlation length Q so that the dimension of matrix \mathbf{C} and \mathbf{C}^{-1} is $M(Q+1) \times M(Q+1)$,
- channel vector $\underline{\mathbf{h}}_k$ is a random process with covariance matrix \mathbf{K} and correlation length P so that the dimension of matrix \mathbf{K} and \mathbf{K}^{-1} is $MN(P+1) \times MN(P+1)$.

Matrix operations require that \mathbf{C}^{-1} and \mathbf{K}^{-1} represent the process of the same correlation length. Let's make this correlation length equal to data window length $T = \max(Q, P)$. Then the required dimensions of \mathbf{C} and \mathbf{K} are:

- dimension of \mathbf{C} is $M(T+1) \times M(T+1)$,
- dimension of \mathbf{K} is $MN(T+1) \times MN(T+1)$.

To have the required dimensions of \mathbf{C} and \mathbf{K} , the smaller is expanded to become a band matrix.

Incorporating ASD into the pdf-s of EM algorithm and skipping the lengthy derivations, the Expectation and Maximization steps are given below. Maximization is done using VA with metric $Q(\mathbf{S}_{1 \rightarrow k} | \widehat{\mathbf{S}}_{1 \rightarrow k}^{(l)})$.

Expectation:

$$Q(\mathbf{S}_{1 \rightarrow k} | \widehat{\mathbf{S}}_{1 \rightarrow k}^{(l)}) = Q(\mathbf{S}_{1 \rightarrow k-T-1} | \widehat{\mathbf{S}}_{1 \rightarrow k-T-1}^{(\max)}) + Q(\underline{\mathbf{S}}_k | \widehat{\underline{\mathbf{S}}}_k^{(l)}) \quad (4.22)$$

According to (4.22), function $Q(\mathbf{S}_{1 \rightarrow k} | \widehat{\mathbf{S}}_{1 \rightarrow k}^{(l)})$ that will be maximized in the Maximization step can be split into two terms. The first term $Q(\mathbf{S}_{1 \rightarrow k-T-1} | \widehat{\mathbf{S}}_{1 \rightarrow k-T-1}^{(\max)})$ is a VA metric over time instances $1, \dots, k-T-1$ and $\widehat{\mathbf{S}}_{1 \rightarrow k-T-1}^{(\max)}$ is the detected sequence that maximized this term at time $k-1$. It will remain constant during EM iterations at time k . The second term $Q(\underline{\mathbf{S}}_k | \widehat{\underline{\mathbf{S}}}_k^{(l)})$ is the expression that has to be maximized with respect to $\underline{\mathbf{S}}_k$ at time k .

Maximization:

$$\widehat{\mathbf{S}}_{1 \rightarrow k}^{(l+1)} = \arg \max_{\mathbf{S}} Q(\mathbf{S}_{1 \rightarrow k} | \widehat{\mathbf{S}}_{1 \rightarrow k}^{(l)}) = \widehat{\mathbf{S}}_{1 \rightarrow k-T-1}^{(\max)} + \widehat{\underline{\mathbf{S}}}_k^{(l+1)} \quad (4.23)$$

The first term in (4.23) represents the part of the sequence detected from the beginning of the frame until the beginning of the current data window at time k , that is over time slots $1, \dots, k - T - 1$. The second term is the sequence inside the current data window over time slots $k - T, \dots, k$, that is the sequence that maximizes (4.22) at time k .

Again, skipping the detailed derivations, the sequence sought for inside the data window of length T at time k is:

$$\hat{\underline{\mathbf{S}}}_k^{(l+1)} = \arg \min_{\underline{\mathbf{S}}} \sum_{u=k-T}^k T_u \quad (4.24)$$

T_u is a modified VA metric including ASD:

$$\begin{aligned} T_u = & \left[\left(\underline{\mathbf{x}}_u - \underline{\mathbf{S}}_u \hat{\underline{\mathbf{h}}}_u^{(l)} \right)^H \mathbf{C}^{-1} \left(\underline{\mathbf{x}}_u - \underline{\mathbf{S}}_u \hat{\underline{\mathbf{h}}}_u^{(l)} \right) \right] + \text{tr} \left\{ \underline{\mathbf{S}}_u^H \mathbf{C}^{-1} \underline{\mathbf{S}}_u \underline{\Sigma}_u^{(l)} \right\} - \\ & - \left[\left(\underline{\mathbf{x}}_{u-1} - \underline{\mathbf{S}}_{u-1} \hat{\underline{\mathbf{h}}}_{u-1}^{(l)} \right)^H \mathbf{C}_{11}^{-1} \left(\underline{\mathbf{x}}_{u-1} - \underline{\mathbf{S}}_{u-1} \hat{\underline{\mathbf{h}}}_{u-1}^{(l)} \right) \right] - \\ & - \text{tr} \left\{ \underline{\mathbf{S}}_{u-1}^H \mathbf{C}_{11}^{-1} \underline{\mathbf{S}}_{u-1} \underline{\Sigma}_{u-1}^{(l)} \right\} \end{aligned} \quad (4.25)$$

where

$$\hat{\underline{\mathbf{h}}}_u^{(l)} = \underline{\Sigma}_u^{(l)} \left(\underline{\mathbf{S}}_u^{(l)} \right)^H \mathbf{C}^{-1} \underline{\mathbf{x}}_u \quad (4.26)$$

$$\underline{\Sigma}_u^{(l)} = \left[\left(\underline{\mathbf{S}}_u^{(l)} \right)^H \mathbf{C}^{-1} \underline{\mathbf{S}}_u^{(l)} + \mathbf{K}^{-1} \right]^{-1} \quad (4.27)$$

$$\hat{\underline{\mathbf{h}}}_{u-1}^{(l)} = \underline{\Sigma}_{u-1}^{(l)} \left(\underline{\mathbf{S}}_{u-1}^{(l)} \right)^H \mathbf{C}_{11}^{-1} \underline{\mathbf{x}}_{u-1} \quad (4.28)$$

$$\underline{\Sigma}_{u-1}^{(l)} = \left[\left(\underline{\mathbf{S}}_{u-1}^{(l)} \right)^H \mathbf{C}_{11}^{-1} \underline{\mathbf{S}}_{u-1}^{(l)} + \mathbf{K} \mathbf{1}^{-1} \right]^{-1} \quad (4.29)$$

\mathbf{C}_{11} is a $MT \times MT$ cofactor matrix of noise covariance matrix \mathbf{C} ,

$\mathbf{K} \mathbf{1}$ is a $MNT \times MNT$ low principle submatrix of channel covariance matrix \mathbf{K} .

To summarize, with the definitions of EM algorithm *complete/completing data* as presented in this section, the EM iterations compute the following:

- ML estimate of the data sequence using VA - Equation (4.24),

- as a by-product in this form of EM we compute the MAP estimate of the CSI expected value - Equations (4.26) and (4.27).

The sliding data window algorithm steps to be executed at time k are as follows:

1. Initial conditions after recursions performed up to (and including) time $k - 1$:
 - $\widehat{\mathbf{S}}_{1 \rightarrow k-1}^{(l)}$ - sequence of detected symbols transmitted at times $1, \dots, k - 1$,
 - $\widehat{\mathbf{h}}_{1 \rightarrow k-1}$ - estimated CSI samples over time slots $1, \dots, k - 1$,
 - $\widehat{\mathbf{x}}_{1 \rightarrow k}$ - sequence of received data samples over time slots $1, \dots, k$.
2. Create the initial sequence estimate $\widehat{\mathbf{S}}_k^{(0)} = \widehat{\mathbf{S}}_{k-T \rightarrow k}^{(0)}$ by appending $\widehat{\mathbf{S}}_k^{(0)} = 0$ to $\widehat{\mathbf{S}}_{k-T \rightarrow k-1}^{(\max)}$.
3. Compute the first CSI estimate $\widehat{\mathbf{h}}_k^{(0)} = \widehat{\mathbf{h}}_{k-T \rightarrow k}^{(0)}$ over time slots $k - T, \dots, k$ using (4.26) with $u = k$ and $l = 0$.
4. Find the best sequence $\widehat{\mathbf{S}}_{k-T \rightarrow k}^{(l+1)}$ using VA with metric from (4.25) - (4.29) and the current CSI estimate $\widehat{\mathbf{h}}_{k-2T+1 \rightarrow k}^{(l)} = \widehat{\mathbf{h}}_{k-2T+1 \rightarrow k-T-1}^{(\max)} + \widehat{\mathbf{h}}_{k-T \rightarrow k}^{(l)}$.
5. Compute the next CSI estimate $\widehat{\mathbf{h}}_{k-T \rightarrow k}^{(l+1)}$ using (4.26) and (4.27) with $u = k$ and $\widehat{\mathbf{S}}_{k-T \rightarrow k}^{(l+1)}$. Make $l = l + 1$.
6. Keep iterating between p.4 and p.5 until convergence, that is until $\widehat{\mathbf{S}}_{k-T \rightarrow k}^{(l+1)} = \widehat{\mathbf{S}}_{k-T \rightarrow k}^{(l)}$.

4.4 Numerical Results

Computer simulations have been done using frames of length $L = 130$ symbols with the first $L_{Tr} = 14$ symbols considered known training sequence followed by $L_D = 116$ data symbols. The block diagram of the EM on-line receiver with sliding data window is shown in Figure 4.1. The same as in the Kalman filter-PSP simulations, the 4-PSK 8-state space-time code shown in Figure 2.2 was used with $N=2$ transmit and $M=2$

receive antennas. Also, the MIMO channel was assumed to consist of independent flat paths with Rayleigh distribution and time variation governed by the Jakes model with autocorrelation function represented by the zeroth-order first-kind Bessel function $r_k = J_0(k2\pi f_d T_s)$, [29, 30]. The simulations were done for the fading rate $f_d T_s = 0.01$. Additive noise was Gaussian and white (interference correlation length $Q = 0$).

After running experimental simulations with different sliding window sizes, a satisfactory compromise between computational complexity and probability of error performance was achieved when two sliding windows of different dimensions were used. First, the window of dimension $U = 40$ symbols over which the CSI was computed. Second, the window of length $T = 5$ symbols over which the Viterbi Algorithm was run. Both windows are shown in Figure 4.2. Symbol detection starts at time $k = L_{Tr} + 1$ - the first data symbol following the training sequence. The operations performed at time k are as follows. The initial estimate of symbols \mathbf{s}_k is assumed to be $\mathbf{0}$. Next, using the received sequence $\mathbf{x}_{k-U \rightarrow k}$ and detected data sequence $\mathbf{s}_{k-U \rightarrow k}$ the CSI estimate $\mathbf{h}_{k-U \rightarrow k}$ is computed over time slots $k - U \dots k$. With the new CSI estimates, the VA is run over time slots $k - T \dots k$ and a new data sequence $\mathbf{s}_{k-T \rightarrow k}$ is found. Then the iteration is repeated. It was observed that in most cases after 2 iterations the algorithm converged. At time instances where k was less or equal to 40, to avoid negative indices, the dimension of channel updates window was selected as $U = \min(k, 40)$ for each time instance k .

Bit error probability and frame error probability plots of the sliding window algorithm are shown in Figure 4.3 and Figure 4.4 respectively. The BEP performance at $\text{BEP} = 10^{-3}$ was found to be about 2 dB worse than Kalman filter-PSP algorithm. Also, the FEP performance at $\text{FEP} = 10^{-2}$ was about 1 dB worse than Kalman filter-PSP.

The simulation results and in-depth investigation of the EM-based sliding window algorithm operation led to the conclusion that the system performance was

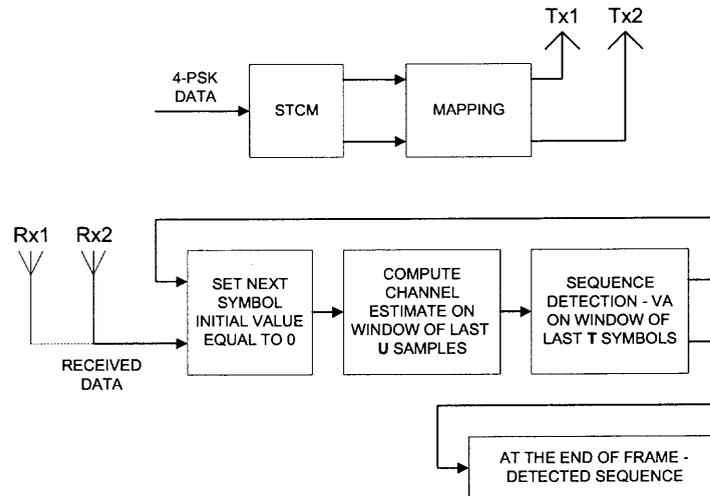


Figure 4.1 Block diagram of the EM on-line receiver with sliding data window.

diminished by the fact that the EM iterations were run on a window of limited length. Consequently, part of the available data (some elements of the received vector and some elements of the channel covariance matrix) were not used in each iteration. The result was that often the EM algorithm converged to one of its local rather than the global maximum and together with error propagation between channel estimation and sequence detection stages several data frames were observed with long burst errors.

These results motivated further research and lead to EM-based block-processing iterative algorithm with symbol interleaving presented in Chapter 5.

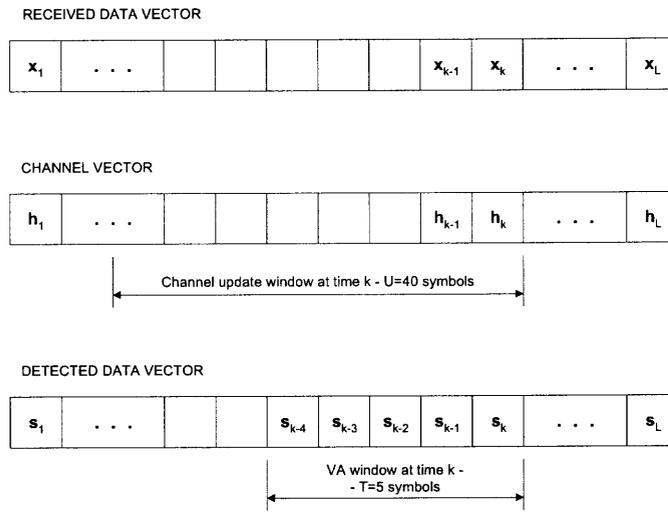


Figure 4.2 EM-based algorithm with sliding channel and data windows.

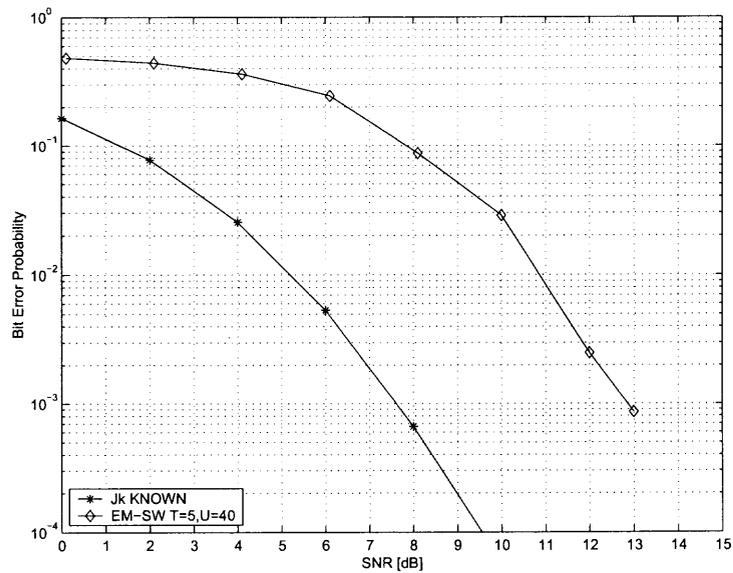


Figure 4.3 EM-based sliding windows receiver (EM-SW) - bit error probability vs. SNR per receive antenna for two transmit-two receive antennas, 8-state 4-PSK STCM over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$.

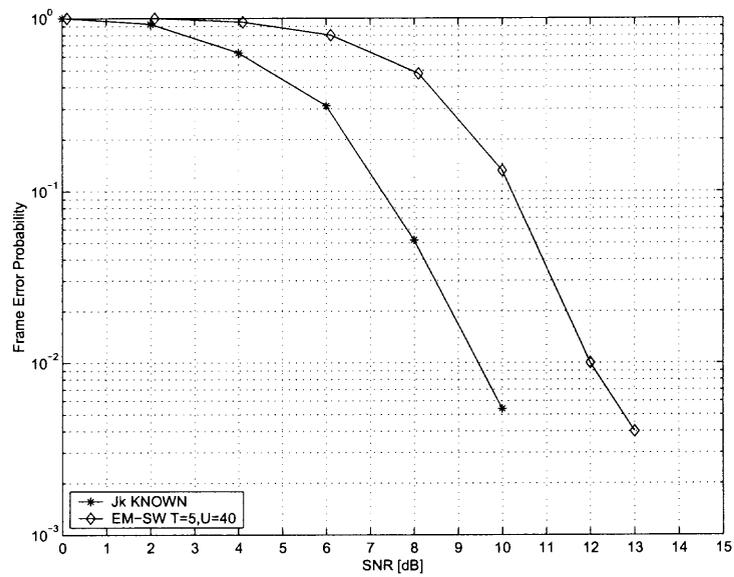


Figure 4.4 EM-based sliding windows receiver (EM-SW) - frame error probability vs. SNR per receive antenna for two transmit-two receive antennas, 8-state 4-PSK STCM over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$.

CHAPTER 5

EM-BASED BLOCK-PROCESSING ITERATIVE RECEIVER WITH SYMBOL INTERLEAVING

As mentioned earlier, the drawbacks of the sliding data window receiver are high computational complexity and error propagation. These problems were coped with by implementing a modified block processing EM-based iterative receiver described in this chapter. All data processing is done on full data frame of length L symbols and signal models represented by (2.2) and (2.4) apply. Also, a different EM *complete/completing data* definition is chosen as compared with the case of sliding data window receiver. The parameter to estimate is not the data sequence but the CSI with MAP as criterion of optimality. The by-product that has to be computed when executing this form of EM algorithm is the MAP estimate of the data sequence. Introduction of symbol interleaving serves to remove the error propagation between data detection and CSI estimation stages in the EM iteration loop.

The following sections presents the details of EM algorithm with MAP estimate of the CSI as criterion of optimality and the *a posteriori probabilities* (APP) algorithm for STCM data detection.

5.1 EM Algorithm with MAP Channel Estimate as Criterion of Optimality

In this version of the EM algorithm, the *complete/completing data* definitions are:

- received data \mathbf{x} - *incomplete data*,
- transmitted sequence \mathbf{S} - *completing data*,
- $\mathbf{y} = \{\mathbf{x}, \mathbf{S}\}$ - *complete data*,
- CSI vector \mathbf{h} - parameter to be estimated.

Using the MAP criterion for CSI estimation, the optimal solution given the observed data can be expressed:

$$\hat{\mathbf{h}} = \arg \max_{\mathbf{h}} p(\mathbf{h}|\mathbf{x}) \quad (5.1)$$

However, the complexity of maximization in (5.1) is intractable. Applying the EM algorithm, it is expected that maximizing the conditional density $p(\mathbf{h}|\mathbf{y})$ - CSI given the *complete data* should be easier. That is, the optimum CSI estimate sought for is given by:

$$\hat{\mathbf{h}} = \arg \max_{\mathbf{h}} p(\mathbf{h}|\mathbf{y}) \quad (5.2)$$

Following the generic EM derivations presented in section 4.2, the Expectation step can be written:

$$Q(\mathbf{h}|\hat{\mathbf{h}}^{(l)}) = E_{\mathbf{S}}[\log p(\mathbf{h}|\mathbf{y})|\mathbf{x}, \hat{\mathbf{h}}^{(l)}] \quad (5.3)$$

Noting that:

$$p(\mathbf{h}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{h})p(\mathbf{h})}{p(\mathbf{y})} \quad (5.4)$$

the Expectation step (5.3) becomes:

$$\begin{aligned} Q(\mathbf{h}|\hat{\mathbf{h}}^{(l)}) &= E_{\mathbf{S}}[\log p(\mathbf{y}|\mathbf{h}) + \log p(\mathbf{h}) - \log p(\mathbf{y})|\mathbf{x}, \hat{\mathbf{h}}^{(l)}] \\ &= E_{\mathbf{S}}[\log p(\mathbf{y}|\mathbf{h})|\mathbf{x}, \hat{\mathbf{h}}^{(l)}] + \log p(\mathbf{h}) - \log p(\mathbf{y}) \end{aligned} \quad (5.5)$$

where the term $p(\mathbf{y})$ is not a function of the parameter being estimated \mathbf{h} and can be dropped. Then the EM Expectation and Maximization steps are:

Expectation:

$$Q(\mathbf{h}|\hat{\mathbf{h}}^{(l)}) = E_{\mathbf{S}}[\log p(\mathbf{y}|\mathbf{h})|\mathbf{x}, \hat{\mathbf{h}}^{(l)}] + \log p(\mathbf{h}) \quad (5.6)$$

Maximization:

$$\hat{\mathbf{h}}^{(l+1)} = \arg \max_{\mathbf{h}} \{E_{\mathbf{S}}[\log p(\mathbf{y}|\mathbf{h})|\mathbf{x}, \hat{\mathbf{h}}^{(l)}] + \log p(\mathbf{h})\} \quad (5.7)$$

Substituting the complete data definition into (5.6):

$$\begin{aligned} Q(\mathbf{h}|\hat{\mathbf{h}}^{(l)}) &= E_{\mathbf{S}}[\log p(\mathbf{x}, \mathbf{S}|\mathbf{h})|\mathbf{x}, \hat{\mathbf{h}}^{(l)}] + \log p(\mathbf{h}) \\ &= E_{\mathbf{S}}[\log p(\mathbf{S}|\mathbf{h})|\mathbf{x}, \hat{\mathbf{h}}^{(l)}] + E_{\mathbf{S}}[\log p(\mathbf{x}|\mathbf{S}, \mathbf{h})|\mathbf{x}, \hat{\mathbf{h}}^{(l)}] + \log p(\mathbf{h}) \end{aligned} \quad (5.8)$$

The transmitted data sequence \mathbf{S} is independent on the channel vector \mathbf{h} so the first term in (5.8) is not a function of \mathbf{h} and after dropping it, (5.8) becomes:

$$Q(\mathbf{h}|\hat{\mathbf{h}}^{(l)}) = E_{\mathbf{S}}[\log p(\mathbf{x}|\mathbf{S}, \mathbf{h})|\mathbf{x}, \hat{\mathbf{h}}^{(l)}] + \log p(\mathbf{h}) \quad (5.9)$$

Assuming that the additive noise entering the M receive antennas is white Gaussian with variance N_0 , the conditional density of the received signal in (5.9) can be expressed:

$$p(\mathbf{x}|\mathbf{S}, \mathbf{h}) = \pi^{-ML} N_0^{-1} \exp[-N_0^{-1}(\mathbf{x} - \mathbf{S}\mathbf{h})^H(\mathbf{x} - \mathbf{S}\mathbf{h})] \quad (5.10)$$

Following the Rayleigh fading assumption about the channel, its density function is expressed by (2.8) and repeated here for convenience:

$$p(\mathbf{h}) = \pi^{-MNL} |\mathbf{K}|^{-1} \exp[-\mathbf{h}^H \mathbf{K}^{-1} \mathbf{h}] \quad (5.11)$$

Substituting (5.11) and (5.10) into (5.9) and dropping the constants the final version of the EM algorithm Expectation step is:

$$Q(\mathbf{h}|\hat{\mathbf{h}}^{(l)}) = -N_0^{-1} E_{\mathbf{S}}[(\mathbf{x} - \mathbf{S}\mathbf{h})^H(\mathbf{x} - \mathbf{S}\mathbf{h})|\mathbf{x}, \hat{\mathbf{h}}^{(l)}] - \mathbf{h}^H \mathbf{K}^{-1} \mathbf{h} \quad (5.12)$$

In the EM Maximization step, (5.12) is maximized with respect to \mathbf{h} by setting the derivative of $Q(\mathbf{h}|\hat{\mathbf{h}}^{(l)})$ equal to 0 with $\mathbf{h} = \hat{\mathbf{h}}^{(l+1)}$:

$$\frac{\partial}{\partial \mathbf{h}^*} [Q(\mathbf{h}|\hat{\mathbf{h}}^{(l)})] = -N_0^{-1} E_{\mathbf{S}} [(-\mathbf{S}^H \mathbf{x} + \mathbf{S}^H \mathbf{S} \mathbf{h}) | \mathbf{x}, \hat{\mathbf{h}}^{(l)}] - \mathbf{K}^{-1} \mathbf{h} = 0 \quad (5.13)$$

or equivalently:

$$N_0^{-1} E_{\mathbf{S}} [\mathbf{S}^H \mathbf{x} | \mathbf{x}, \hat{\mathbf{h}}^{(l)}] = N_0^{-1} E_{\mathbf{S}} [\mathbf{S}^H \hat{\mathbf{S}} \hat{\mathbf{h}}^{(l+1)} | \mathbf{x}, \hat{\mathbf{h}}^{(l)}] + \mathbf{K}^{-1} \hat{\mathbf{h}}^{(l+1)} \quad (5.14)$$

Then the next CSI estimate is:

$$\hat{\mathbf{h}}^{(l+1)} = (N_0^{-1} E_{\mathbf{S}} [\mathbf{S}^H \mathbf{S} | \mathbf{x}, \hat{\mathbf{h}}^{(l)}] + \mathbf{K}^{-1})^{-1} E_{\mathbf{S}} [\mathbf{S}^H | \mathbf{x}, \hat{\mathbf{h}}^{(l)}] N_0^{-1} \mathbf{x} \quad (5.15)$$

The quantity $E_{\mathbf{S}} [\mathbf{S} | \mathbf{x}, \hat{\mathbf{h}}^{(l)}]$ in (5.15) represents the conditional expected value of the detected data sequence given the observed data \mathbf{x} and the last CSI estimate $\hat{\mathbf{h}}^{(l)}$. The APP algorithm is employed in data detection stage so that its soft outputs allow to compute the expected values $E_{\mathbf{S}} [\mathbf{S} | \mathbf{x}, \hat{\mathbf{h}}^{(l)}]$ of the detected data symbols. The STCM implementation of the APP algorithm is presented in the next section. The EM algorithm described here iterates between the estimation stage, where CSI estimate is computed using (5.15) and the detection stage executing the APP algorithm until convergence. At the end of the EM iterations the final result is a MAP estimate of the CSI and the last sequence detected with APP algorithm which used MAP as its criterion of optimality.

5.2 APP for STCM

Computation of the CSI estimate from (5.15) involves computation of the conditional mean values of symbols transmitted by each transmit antenna. That means, the computation requires knowledge of the a posteriori probabilities $P(\mathbf{s}_k = \mathbf{s}^{(i)} | \mathbf{x}, \hat{\mathbf{h}}^{(l)})$ where $\{\mathbf{s}^{(i)}\} = \mathcal{A}$. It is assumed that CSI estimate from the previous EM iteration (or its initial value if $l = 0$) $\hat{\mathbf{h}}^{(l)}$ is available and fixed for the purpose of APP decoding

and the notation can be simplified by dropping $\widehat{\mathbf{h}}^{(l)}$. Below a generic APP algorithm is presented for computing the a posteriori probabilities of data symbols u_k at the input to STCM encoder $P(u_k = u^{(i)}|\mathbf{x})$. The detailed derivation of the algorithm can be found in [54]. It is based on BCJR algorithm [55] but it is different from it in two respects: non-binary trellis and multiple input-multiple output (MIMO) channel. The extension of the algorithm is also presented, which computes the APPs of symbols at the N outputs of STCM encoder - that is the probabilities of STCM modulated symbols transmitted by each transmit antenna $P(s_n(k) = s^{(i)}|\mathbf{x})$.

Consider the sequence of symbols u_k entering the STCM encoder. The APPs of symbols at the input to the encoder at time k can be expressed:

$$P(u_k = u^{(i)}|\mathbf{x}) = \sum_{\sigma_k} \sum_{\sigma_{k+1}} P(u_k = u^{(i)}, \sigma_k, \sigma_{k+1}|\mathbf{x}) \quad (5.16)$$

where $P(u_k = u^{(i)}, \sigma_k, \sigma_{k+1}|\mathbf{x})$ is the joint probability that given the received data \mathbf{x} , the transition from state σ_k to σ_{k+1} is associated with encoder input symbol $u_k = u^{(i)}$. The notations are illustrated in Figure 5.1. With no parallel transitions between states, any two elements from $(\sigma_k, \sigma_{k+1}, u_k)$ uniquely define a transition. Let's define \mathbf{x}_k^- as the sequence received before time k , and \mathbf{x}_k^+ the sequence received after time k . Then there is $\mathbf{x} = (\mathbf{x}_k^-, \mathbf{x}_k, \mathbf{x}_k^+)$. Using the procedure similar to [55] the joint probability $P(u_k = u^{(i)}, \sigma_k, \sigma_{k+1}|\mathbf{x})$ can be expressed in a recursive form. Applying Bayes rule to (5.16):

$$P(u_k = u^{(i)}, \sigma_k, \sigma_{k+1}|\mathbf{x}) = hP(u_k = u^{(i)}, \sigma_k, \sigma_{k+1}, \mathbf{x}) \quad (5.17)$$

where the normalizing coefficient h is chosen such that:

$$\begin{aligned} h &= \frac{1}{P(\mathbf{x})} \\ &= \frac{1}{\sum_i \sum_{\sigma_k} \sum_{\sigma_{k+1}} P(u_k = u^{(i)}, \sigma_k, \sigma_{k+1}, \mathbf{x})} \end{aligned} \quad (5.18)$$

To compute the joint probability $P(u_k = u^{(i)}, \sigma_k, \sigma_{k+1}, \mathbf{x})$, we first rewrite it:

$$P(u_k = u^{(i)}, \sigma_k, \sigma_{k+1}, \mathbf{x}) = P(u_k = u^{(i)}, \sigma_k, \sigma_{k+1}, \mathbf{x}_k^-, \mathbf{x}_k, \mathbf{x}_k^+) \quad (5.19)$$

Applying the chain rule to (5.19):

$$\begin{aligned} P(u_k = u^{(i)}, \sigma_k, \sigma_{k+1}, \mathbf{x}_k^-, \mathbf{x}_k, \mathbf{x}_k^+) &= p(\mathbf{x}_k^+ | u_k = u^{(i)}, \sigma_k, \sigma_{k+1}, \mathbf{x}_k^-, \mathbf{x}_k) P(u_k = u^{(i)}, \sigma_k, \sigma_{k+1}, \mathbf{x}_k^-, \mathbf{x}_k) \\ &= p(\mathbf{x}_k^+ | u_k = u^{(i)}, \sigma_k, \sigma_{k+1}, \mathbf{x}_k^-, \mathbf{x}_k) P(u_k = u^{(i)}, \sigma_{k+1}, \mathbf{x}_k | \sigma_k, \mathbf{x}_k^-) P(\sigma_k, \mathbf{x}_k^-) \end{aligned} \quad (5.20)$$

From the Markov chains properties if state σ_{k+1} is known, \mathbf{x}_k^+ does not depend on any of the other parameters:

$$p(\mathbf{x}_k^+ | u_k = u^{(i)}, \sigma_k, \sigma_{k+1}, \mathbf{x}_k^-, \mathbf{x}_k) = p(\mathbf{x}_k^+ | \sigma_{k+1}) \quad (5.21)$$

Also, if state σ_k is known, quantities at times k and on do not depend on \mathbf{x}_k^- :

$$P(u_k = u^{(i)}, \sigma_{k+1}, \mathbf{x}_k | \sigma_k, \mathbf{x}_k^-) = P(u_k = u^{(i)}, \sigma_{k+1}, \mathbf{x}_k | \sigma_k) \quad (5.22)$$

Defining the following quantities:

$$\begin{aligned}
\gamma_k^{(i)}(\sigma_k, \sigma_{k+1}, \mathbf{x}_k) &= P(u_k = u^{(i)}, \sigma_{k+1}, \mathbf{x}_k | \sigma_k) \\
\alpha_k(\sigma_k) &= P(\sigma_k, \mathbf{x}_k^-) \\
\beta_{k+1}(\sigma_{k+1}) &= p(\mathbf{x}_k^+ | \sigma_{k+1})
\end{aligned} \tag{5.23}$$

the APP in (5.17) can be expressed:

$$P(u_k = u^{(i)}, \sigma_k, \sigma_{k+1} | \mathbf{x}) = h\alpha_k(\sigma_k)\gamma_k^{(i)}(\sigma_k, \sigma_{k+1}, \mathbf{x}_k)\beta_{k+1}(\sigma_{k+1}) \tag{5.24}$$

Next, it is shown that $\alpha_k(\sigma_k)$ and $\beta_{k+1}(\sigma_{k+1})$ can be computed recursively. It is also shown how to compute $\gamma_k^{(i)}(\sigma_k, \sigma_{k+1}, \mathbf{x}_k)$ having the received sequence \mathbf{x} and CSI estimate $\hat{\mathbf{h}}^{(l)}$.

- $\gamma_k^{(i)}(\sigma_k, \sigma_{k+1}, \mathbf{x}_k)$ is the joint probability of the transition from state σ_k to σ_{k+1} with the observation \mathbf{x}_k . Using the definition in (5.23) it can be expressed:

$$\begin{aligned}
\gamma_k^{(i)}(\sigma_k, \sigma_{k+1}, \mathbf{x}_k) &= P(u_k = u^{(i)}, \sigma_{k+1}, \mathbf{x}_k | \sigma_k) \\
&= p(\mathbf{x}_k | \sigma_k, \sigma_{k+1}, u_k = u^{(i)})P(u_k = u^{(i)}, \sigma_{k+1} | \sigma_k) \\
&= p(\mathbf{x}_k | \sigma_k, \sigma_{k+1}, u_k = u^{(i)})P(u_k = u^{(i)} | \sigma_k, \sigma_{k+1})P(\sigma_{k+1} | \sigma_k)
\end{aligned} \tag{5.25}$$

where the conditional pdf $p(\mathbf{x}_k | \sigma_k, \sigma_{k+1}, u_k = u^{(i)}) = p(\mathbf{x}_k | \sigma_k, u^{(i)})$ can be computed from:

$$p(\mathbf{x}_k | \sigma_k, u^{(i)}) = (\pi N_0)^{-M} \exp[-(\mathbf{x}_k - \mathbf{S}_k \hat{\mathbf{h}}_k^{(l)})^H N_0^{-1} (\mathbf{x}_k - \mathbf{S}_k \hat{\mathbf{h}}_k^{(l)})] \tag{5.26}$$

\mathbf{S}_k is a matrix of STCM modulated symbols defined in (2.1). The term $P(u_k = u^{(i)} | \sigma_k, \sigma_{k+1}) = 1$ if there is a transition $\sigma_k \rightarrow \sigma_{k+1}$ due to input $u^{(i)}$ and it is zero otherwise. If there is such transition, we can also write $P(\sigma_{k+1} | \sigma_k) = P(u_k = u^{(i)})$ where $P(u_k = u^{(i)})$ is the a priori probability of input symbol $u^{(i)}$. Now (5.25) can be written:

$$\gamma_k^{(i)}(\sigma_k, \sigma_{k+1}, \mathbf{x}_k) = p(\mathbf{x}_k | \sigma_k, u^{(i)}) P(u_k = u^{(i)}) \quad (5.27)$$

for existing transitions $\sigma_k \rightarrow \sigma_{k+1}$ associated with input $u^{(i)}$.

- $\alpha_k(\sigma_k) = P(\sigma_k, \mathbf{x}_k^-)$ are the joint probabilities of the trellis states at time k and the data received until time $k - 1$, \mathbf{x}_k^- . These probabilities are obtained from the forward recursion:

$$\begin{aligned} \alpha_{k+1}(\sigma_{k+1}) &= P(\sigma_{k+1}, \mathbf{x}_{k+1}^-) \\ &= \sum_{\sigma_k} P(\sigma_k, \sigma_{k+1}, \mathbf{x}_k^-, \mathbf{x}_k) \\ &= \sum_{\sigma_k} P(\sigma_{k+1}, \mathbf{x}_k | \sigma_k, \mathbf{x}_k^-) P(\sigma_k, \mathbf{x}_k^-) \\ &= \sum_{\sigma_k} P(\sigma_{k+1}, \mathbf{x}_k | \sigma_k) P(\sigma_k, \mathbf{x}_k^-) \\ &= \sum_{\sigma_k} \alpha_k(\sigma_k) \gamma(\sigma_k, \sigma_{k+1}, \mathbf{x}_k) \end{aligned} \quad (5.28)$$

where $P(\sigma_{k+1}, \mathbf{x}_k | \sigma_k, \mathbf{x}_k^-) = P(\sigma_{k+1}, \mathbf{x}_k | \sigma_k)$ was used, which says that events after $k - 1$ are not influenced by \mathbf{x}_k^- if σ_k is known. Also, the following definition was used:

$$\gamma(\sigma_k, \sigma_{k+1}, \mathbf{x}_k) = P(\sigma_{k+1}, \mathbf{x}_k | \sigma_k) \quad (5.29)$$

From (5.23) and (5.29) it can be seen that:

$$\gamma(\sigma_k, \sigma_{k+1}, \mathbf{x}_k) = \sum_i \gamma_k^{(i)}(\sigma_k, \sigma_{k+1}, \mathbf{x}_k) \quad (5.30)$$

- $\beta_{k+1}(\sigma_{k+1}) = p(\mathbf{x}_k^+ | \sigma_{k+1})$ are the densities of the observations after time k conditioned on the states at time k . The recursive expression for $\beta_{k+1}(\sigma_{k+1})$ can be obtained as follows:

$$\begin{aligned} \beta_k(\sigma_k) &= p(\mathbf{x}_{k-1}^+ | \sigma_k) \\ &= \sum_{\sigma_{k+1}} P(\sigma_{k+1}, \mathbf{x}_{k-1}^+ | \sigma_k) \\ &= \sum_{\sigma_{k+1}} P(\sigma_{k+1}, \mathbf{x}_k, \mathbf{x}_k^+ | \sigma_k) \\ &= \sum_{\sigma_{k+1}} p(\mathbf{x}_k^+ | \sigma_k, \sigma_{k+1}, \mathbf{x}_k) P(\sigma_{k+1}, \mathbf{x}_k | \sigma_k) \\ &= \sum_{\sigma_{k+1}} p(\mathbf{x}_k^+ | \sigma_{k+1}) P(\sigma_{k+1}, \mathbf{x}_k | \sigma_k) \\ &= \sum_{\sigma_{k+1}} \beta_{k+1}(\sigma_{k+1}) \gamma(\sigma_k, \sigma_{k+1}, \mathbf{x}_k) \end{aligned} \quad (5.31)$$

where the Markov property was used $p(\mathbf{x}_k^+ | \sigma_k, \sigma_{k+1}, \mathbf{x}_k) = p(\mathbf{x}_k^+ | \sigma_{k+1})$ and definition (5.29).

By combining (5.28), (5.31) and (5.27) and substituting into (5.24) and then into (5.16), the APPs of the symbols u_k at the input to the STCM encoder can be expressed:

$$P(u_k = u^{(i)}|\mathbf{x}) = h \sum_{\sigma_k} \sum_{\sigma_{k+1}} \alpha_k(\sigma_k) \gamma_k^{(i)}(\sigma_k, \sigma_{k+1}, \mathbf{x}_k) \beta_{k+1}(\sigma_{k+1}) \quad (5.32)$$

The APPs of STCM modulated symbols transmitted by antenna n at time k are:

$$P(s_n(k) = s^{(i)}|\mathbf{x}) = \sum_{\sigma_k} \sum_{\sigma_{k+1}} P(s_n(k) = s^{(i)}, \sigma_k, \sigma_{k+1}|\mathbf{x}) \quad (5.33)$$

where $P(s_n(k) = s^{(i)}, \sigma_k, \sigma_{k+1}|\mathbf{x})$ is the joint probability that given the received data \mathbf{x} , the transition from state σ_k to σ_{k+1} is associated with symbol $s^{(i)}$ emitted from transmit antenna n , $n = 1, \dots, N$. Recalling that one encoder input symbol $u^{(i)}$ generates a vector of STCM modulated transmitted symbols $\mathbf{s}_k^{(i)} = \left[s_1^{(i)}(k) \ s_2^{(i)}(k) \ \dots \ s_N^{(i)}(k) \right]^T$ and following the procedure to derive the APPs of the encoder input symbols u_k , the probability in (5.33) can be expressed similarly to that in (5.24):

$$P(s_n(k) = s^{(i)}, \sigma_k, \sigma_{k+1}|\mathbf{x}) = q \alpha_k(\sigma_k) \gamma_k^{(i)}(n, \sigma_k, \sigma_{k+1}, \mathbf{x}_k) \beta_{k+1}(\sigma_{k+1}) \quad (5.34)$$

where $\gamma_k^{(i)}(n, \sigma_k, \sigma_{k+1}, \mathbf{x}_k)$ is the joint probability of the transition from state σ_k to σ_{k+1} with observation \mathbf{x}_k associated with symbol $s_n(k) = s^{(i)}$ transmitted from antenna n and q is a normalizing coefficient. By analogy to (5.25):

$$\gamma_k^{(i)}(n, \sigma_k, \sigma_{k+1}, \mathbf{x}_k) = p(\mathbf{x}_k|\sigma_k, \sigma_{k+1}, u_k = u^{(i)}) P(s_n(k) = s^{(i)}|\sigma_k, \sigma_{k+1}) P(\sigma_{k+1}|\sigma_k) \quad (5.35)$$

The first term $p(\mathbf{x}_k|\sigma_k, \sigma_{k+1}, u_k = u^{(i)})$ is given by (5.26). The term $P(s_n(k) = s^{(i)}|\sigma_k, \sigma_{k+1}) = 1$ if transition $\sigma_k \rightarrow \sigma_{k+1}$ is associated with symbol $s^{(i)}$ being tran-

mitted from antenna n and it is zero otherwise. If there is such transition, there is also $P(\sigma_{k+1}|\sigma_k) = P(s_n(k) = s^{(i)})$ which is the a priori probability of symbol $s^{(i)}$ being transmitted from antenna n . Now (5.35) can be written:

$$\gamma 1_k^{(i)}(n, \sigma_k, \sigma_{k+1}, \mathbf{x}_k) = p(\mathbf{x}_k|\sigma_k, u^{(i)})P(s_n(k) = s^{(i)}) \quad (5.36)$$

Combining (5.28), (5.31), (5.36) and substituting into (5.34) and then into (5.33) the APPs of STCM modulated symbols transmitted by antenna n at time k can be expressed:

$$P(s_n(k) = s^{(i)}|\mathbf{x}) = q \sum_{\sigma_k} \sum_{\sigma_{k+1}} \alpha_k(\sigma_k) \gamma 1_k^{(i)}(n, \sigma_k, \sigma_{k+1}, \mathbf{x}_k) \beta_{k+1}(\sigma_{k+1}) \quad (5.37)$$

Computation of the APPs from (5.37) for all transmit antennas $n = 1, \dots, N$ and full length data frame $k = 1, \dots, L$ results in soft-output values of sequences transmitted by all N STCM antennas. Those are the quantities needed to determine the conditional expectations of the N transmitted sequences which are required for computing the CSI estimate from (5.15) in the next EM iteration.

5.3 Description of the Iterative Receiver with Symbol Interleaving

The block diagram of the receiver is shown in the bottom part of Figure 5.2. Receiver operates on data frames of length L symbols at a time. In the transmitter, a frame of length L symbols consisting of a preamble of L_t known training symbols had been STCM encoded and interleaved before being transmitted. The length of the interleaver is chosen such that the training symbols are spread evenly across the

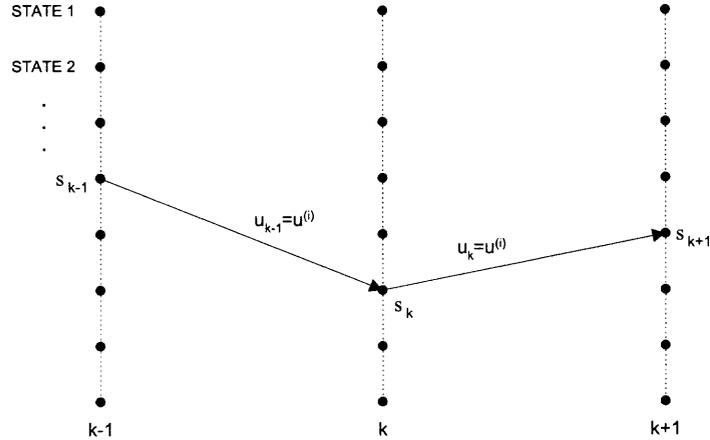


Figure 5.1 APP algorithm notation describing the STCM trellis states and transitions.

full frame. The observation sequence received at the receiver is made up of the interleaved transmitted data distorted by the fading MIMO channel and additive noise. After reception, the observation data of length L is fed into EM derived CSI estimator represented by (5.15). The estimator also utilizes the APPs of the transmitted data. The first EM iteration is carried out assuming probabilities of 1 for the known training symbols and 2^{-m} for the unknown data symbols. Equivalently, for the 2^m -PSK modulated symbols considered in this paper the unknown symbols are assumed to be equal to zero. In case of time-varying channel model, the CSI estimate obtained from (5.15) is associated with the transmitted interleaved data. Before the observation sequence and CSI estimate are fed into the APP data decoder, both have to be deinterleaved. Only then the data sequence being decoded will match the trellis code used to encode it in the transmitter. APP decoder soft outputs computed according to (5.37) are interleaved again and together with the original sequence of observations they are fed to CSI estimator to carry out the next EM iteration.

It can be seen that the EM iteration steps, namely CSI estimation and data detection are executed in a closed loop. Much of the published EM joint estimation-

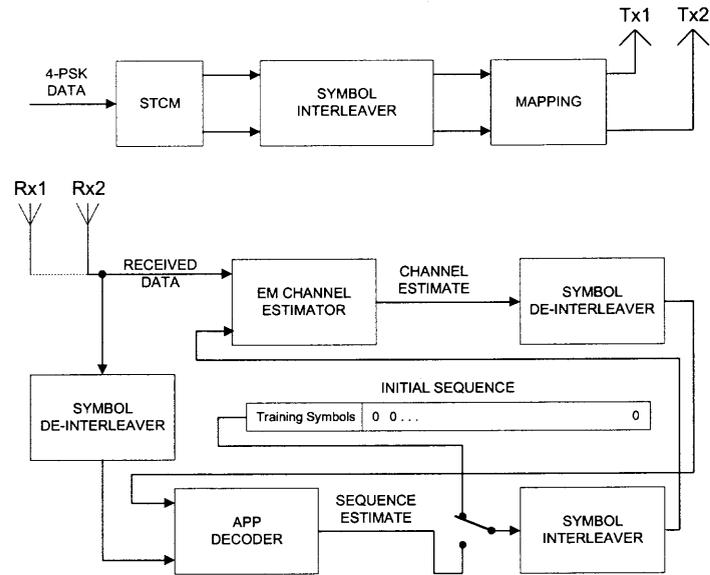


Figure 5.2 Block diagram of EM-APP block-processing iterative receiver with symbol interleaving.

detection methods (such as [26]) suffer from the effects of error propagation between the detection and estimation stages. When estimating the time-varying channel Model B and a symbol error is made in the detection stage, it causes inaccurate CSI estimates in the vicinity of the erroneous symbol. This in turn leads to even more symbol errors in this area in the next iteration. In effect, estimation/detection errors accumulate. The result is that several frames contain long bursts of errors. Implementing a symbol interleaver after the information symbols are STCM encoded in the transmitter solves the described above error propagation problem. With symbol interleaving, the interleaved sequence is used at the receiver to compute the CSI estimate while detection is done on the deinterleaved sequence. Such approach allows to decouple the time instances when symbol errors are made in the detection process from the time instances where CSI estimation errors occur. The interleaver also has the added benefit of increasing the diversity advantage to combat deep fades over a fast fading channel [27].

5.4 Numerical Results

Numerical results were obtained for the 4-PSK 8-state space-time code presented in [13] and shown in Figure 2.2, with $N = 2$ transmit and $M = 2$ receive antennas. The transmitted symbols were assembled in frames of length $L = 130$. Each frame consisted of 14 training symbols followed by 116 data symbols. The additive noise used in the simulations was white, Gaussian with known variance and the effective SNR has been obtained by incorporating into its computation the transmitted power lost for transmission of training symbols:

$$SNR_{EFF} = SNR \frac{L}{L - L_t} \quad (5.38)$$

The interleaver used was a simple deterministic 14×10 matrix interleaver. The MIMO channel was assumed to consist of independent flat paths with Rayleigh distribution and time variation governed by the Jakes model with autocorrelation function modeled by the zeroth-order first-kind Bessel function $r_k = J_0(k2\pi f_d T_s)$, [29, 30]. The fading rate was $f_d T_s = 0.01$, which corresponds to a vehicle moving at approximately 90 mph for a carrier at 2 GHz and a symbol rate of 24.3 ksymbols/sec (current TDMA IS-136 standard).

To evaluate the effect of the interleaver on the performance of the receiver, first the simulation of the system described above was ran with interleaving/deinterleaving operations removed. The bit-error probability (BEP) vs. SNR for such configuration is shown in Figure 5.3 for CSI assumed known and for CSI estimated using four EM iterations. The case of unknown estimated CSI can be considered an alternative implementation of the system in [26]. The main difference between the system considered and [26] is the format of the data block used in EM iterations: a frame of 130 symbols with 14 training symbols instead of 3 consecutive 25-symbol frames with 2-symbol pilot blocks interspersed between them. It can be seen from Figure 5.3 that with the frame 130 symbols and no interleaving, the algorithm fails to

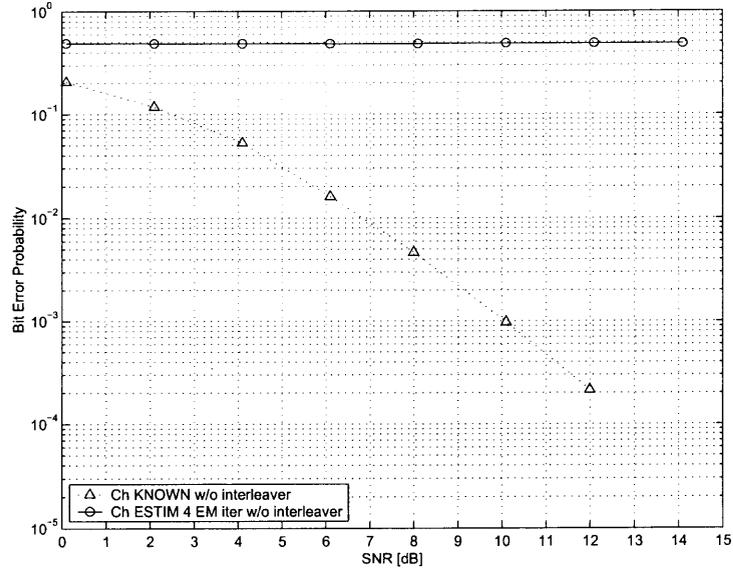


Figure 5.3 EM-APP block-processing iterative receiver - bit error probability vs. SNR per receive antenna over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$ - no interleaving, 4 EM iterations.

produce a satisfactory CSI estimate and consequently fails to deliver a reasonable BEP performance even at high SNR values. Introducing symbol interleaving/deinterleaving has a huge impact on system performance. With interleaver/deinterleaver in place, Figures 5.4 through 5.7 show the BEP and frame error probability (FEP), respectively, versus signal-to-noise ratio per receive antenna for 1-4 iterations for slow and fast fading Rayleigh channels. The performance of the receiver with known CSI and with symbol interleaving was used for reference. From Figure 5.4 and 5.6, it is observed that at $\text{BEP} = 10^{-3}$ and after four iterations, the performance of the proposed EM-APP receiver is within 0.5 dB of the performance of the system with known CSI. Similarly, from Figure 5.5 and 5.7, it is seen that at $\text{FEP} = 10^{-2}$ and after four iterations, the performance has only a 0.5 dB gap to the case of known CSI. Note by comparing Figure 5.3 with 5.4 and 5.6 that for $\text{BEP} = 10^{-3}$ about 0.5 dB advantage is evident even after two iterations over the performance with known channel but no interleaver.

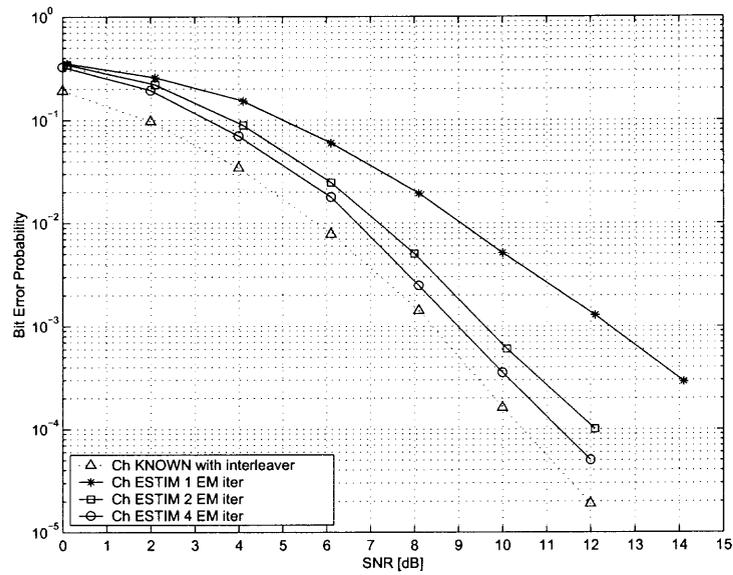


Figure 5.4 EM-APP block-processing iterative receiver - bit error probability vs. SNR per receive antenna with known AWGN variance over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$ - with symbol interleaving - 1, 2 and 4 EM iterations.

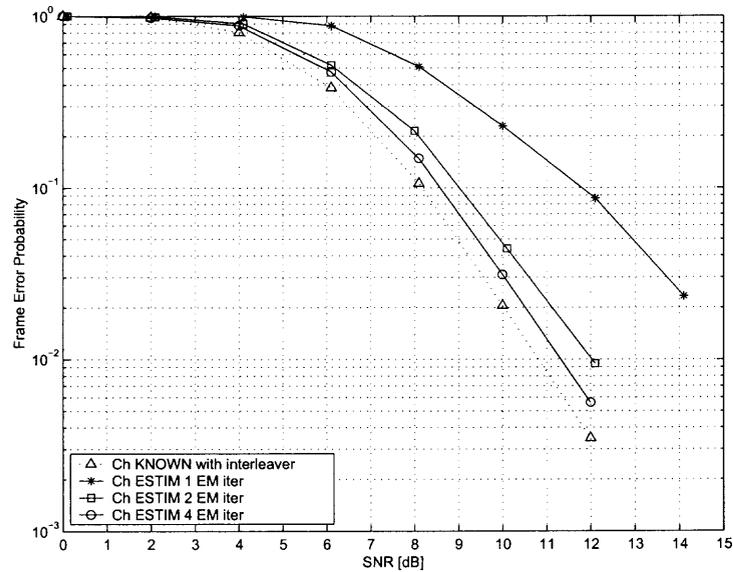


Figure 5.5 EM-APP block-processing iterative receiver - frame error probability vs. SNR per receive antenna with known AWGN variance over Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$ - with symbol interleaving - 1, 2 and 4 EM iterations.

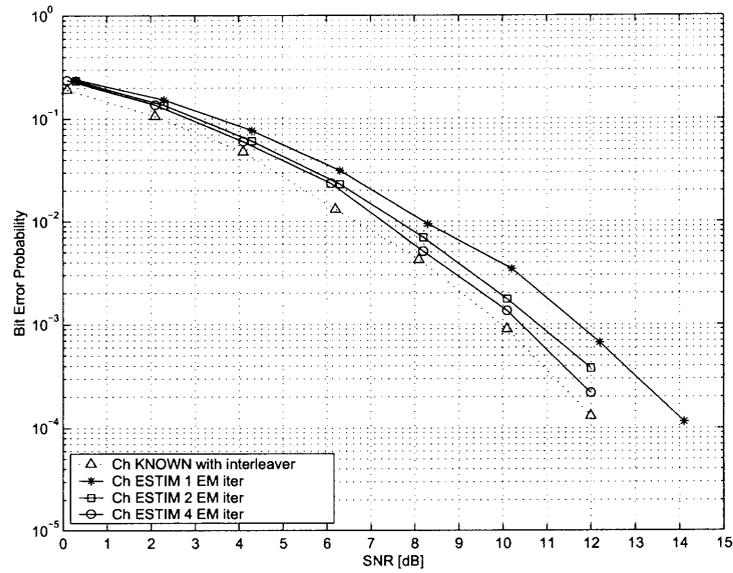


Figure 5.6 EM-APP block-processing iterative receiver, AWGN variance known - bit error probability vs. SNR per receive antenna over Rayleigh fading channel, slow fading rate $f_d T_s = 0.001$ - with symbol interleaving - 1, 2 and 4 EM iterations.

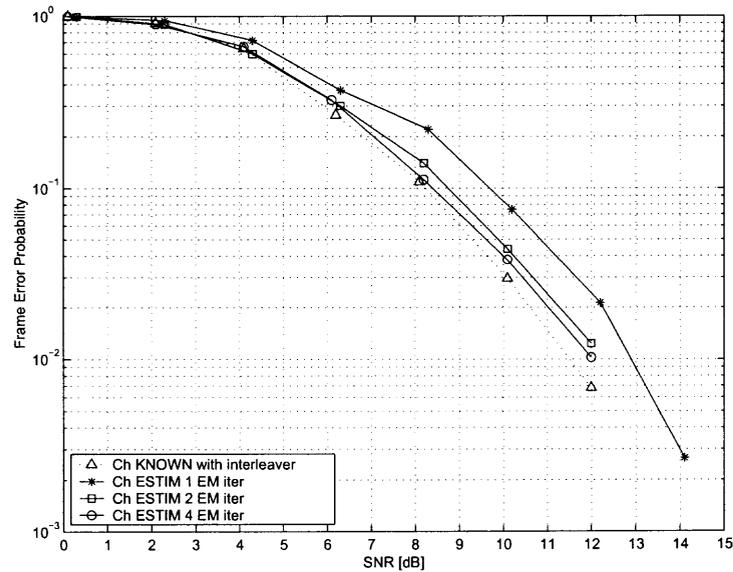


Figure 5.7 EM-APP block-processing iterative receiver, AWGN variance known - frame error probability vs. SNR per receive antenna over Rayleigh fading channel, slow fading rate $f_d T_s = 0.001$ - with symbol interleaving - 1, 2 and 4 EM iterations.

CHAPTER 6

ITERATIVE APP-EM STCM RECEIVER WITH WHITE NOISE VARIANCE ESTIMATION

The iterative APP-EM receiver presented in Chapter 5 computes the CSI estimate using (5.15). In derivations leading to (5.15), the assumption was made that the additive noise was Gaussian, white in time and space and its variance N_0 was known. In this chapter, a more practical approach is taken to do sequence detection and compute CSI estimate assuming that the noise variance is unknown and has to be estimated. We would like to incorporate the noise variance estimation into the EM iterations loop, which suggests that signal processing in our modified receiver should be performed in three rather than two stages:

1. CSI estimation,
2. APP signal detection,
3. Noise variance estimation.

The generic EM algorithm introduced in Section 3.1 should be modified. If instead of estimating a single parameter \mathbf{a} , we need to estimate a set of different parameters $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$, following [20] we can partition the set of parameters into smaller subsets and find the estimates of the subsets separately. Then a single EM iteration can be split into several steps in which only one parameter/subset is estimated at a time while the remaining parameters are assumed available and kept fixed:

$$\hat{\mathbf{a}}_1^{(l+1)} = \arg \max_{\mathbf{a}_1} Q(\mathbf{a}_1, \hat{\mathbf{a}}_2^{(l)}, \dots, \hat{\mathbf{a}}_k^{(l)} | \hat{\mathbf{A}}^{(l)}) \quad (6.1)$$

$$\hat{\mathbf{a}}_2^{(l+1)} = \arg \max_{\mathbf{a}_2} Q(\hat{\mathbf{a}}_1^{(l+1)}, \mathbf{a}_2, \dots, \hat{\mathbf{a}}_k^{(l)} | \hat{\mathbf{A}}^{(l)}) \quad (6.2)$$

$$\hat{\mathbf{a}}_{\mathbf{k}}^{(l+1)} = \arg \max_{\mathbf{a}_{\mathbf{k}}} Q(\hat{\mathbf{a}}_1^{(l+1)}, \hat{\mathbf{a}}_2^{(l+1)}, \dots, \mathbf{a}_{\mathbf{k}} | \hat{\mathbf{A}}^{(l)}) \quad (6.3)$$

where $(\hat{\mathbf{a}}_1^{(l+1)}, \hat{\mathbf{a}}_2^{(l+1)}, \dots, \hat{\mathbf{a}}_{\mathbf{k}}^{(l+1)}) = \hat{\mathbf{A}}^{(l+1)}$ is a new estimate of the full parameters set \mathbf{A} .

Applying the approach of parameters set estimation (6.1)-(6.3) to modify our APP-EM iterative receiver, instead of estimating a single parameter \mathbf{h} , we need to estimate a set of parameters $\mathbf{P} = (\mathbf{h}, \mathbf{z})$ where \mathbf{z} is noise vector defined in (2.2). Then the three stages of signal processing in our modified receiver are:

1. CSI estimation from:

$$\hat{\mathbf{h}}^{(l+1)} = \arg \max_{\mathbf{h}} Q(\mathbf{h} | \hat{\mathbf{h}}^{(l)}, \hat{N}_0^{(l)}) \quad (6.4)$$

using Equation (5.15).

2. APP signal computation $E_{\mathbf{S}} [\mathbf{S} | \mathbf{x}, \hat{\mathbf{h}}^{(l+1)}, \hat{N}_0^{(l)}]$ using the density of received data given by (5.26).
3. Noise estimation:

- (a) Noise samples vector computation:

$$\hat{\mathbf{z}}^{(l+1)} = \mathbf{x} - E_{\mathbf{S}} [\mathbf{S} | \mathbf{x}, \hat{\mathbf{h}}^{(l+1)}, \hat{N}_0^{(l)}] \hat{\mathbf{h}}^{(l+1)} \quad (6.5)$$

- (b) Noise variance $\hat{N}_0^{(l+1)}$ computation:

$$\hat{N}_0^{(l+1)} = \frac{1}{LM} \sum_{j=1}^{LM} \hat{z}^{(l+1)}(j)^* \cdot z^{(l+1)}(j) \quad (6.6)$$

The modified APP-EM receiver including the noise variance estimation is shown in Figure 6.1. The channel estimator takes at its input the received observations sequence, the last estimate of noise variance and the last estimate of data symbols APPs computed from (5.37) and produces the sequence of new CSI estimates. The

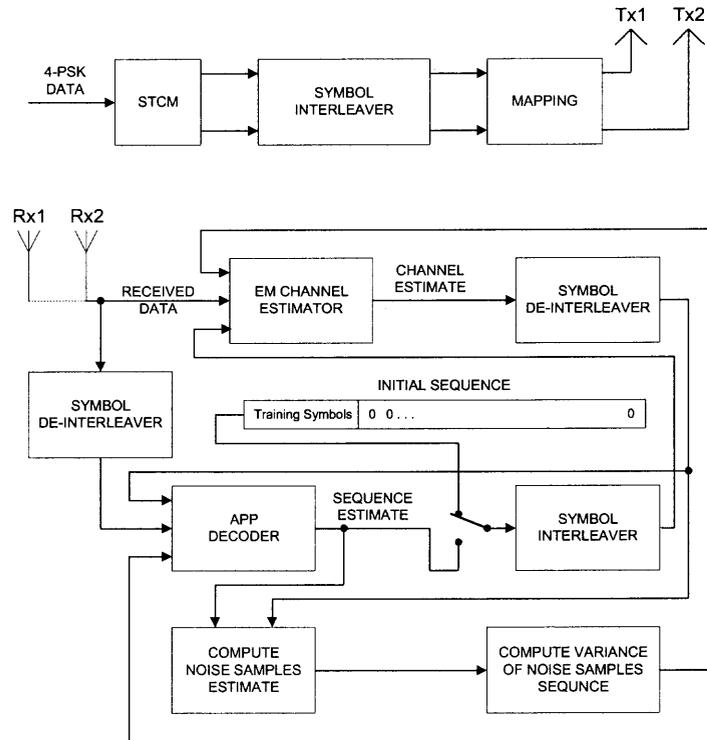


Figure 6.1 Block diagram of the APP-EM iterative receiver with estimation of unknown white noise variance.

new CSI estimates sequence after deinterleaving is fed to the APP decoder together with the deinterleaved sequence of observations and last estimate of noise variance. The output of the APP decoder is a sequence of new APPs of data symbols transmitted by all N transmit antennas. Having the new estimates of CSI and data symbols APPs, a new estimate of noise variance is computed which will be used in the next EM iteration. At the end of EM iterations, (5.32) is used for making the final hard decision on the most likely data u_k , $k = 1, \dots, L$ at the input to STCM encoder.

6.1 Initial Value of Noise Variance

To start the EM iterations, initial value of noise variance is needed. Simulations were run with different initial noise variance values in the range 0.1 through 30 for

$SNR = 8$ dB and 12 dB to determine the range of these values resulting in the best performance. With SNR 8 dB and 12 dB and signal power $P_S = 1$, the noise variances P_N are:

$$\begin{aligned} SNR = 8 \text{ dB: } & 10 \log \frac{P_S}{P_N} = 8 & P_N = 0.16 \\ SNR = 12 \text{ dB: } & 10 \log \frac{P_S}{P_N} = 12 & P_N = 0.06 \end{aligned}$$

Figure 6.2 shows the BEP vs. initial noise variance plots. It can be observed from the figure that the best BEP performance is obtained when initial noise variance is greater than the true noise variance. That is, it is better to overestimate the initial noise variance rather than underestimate it. To justify this conclusion, let's express the equation for computing CSI estimate (5.15) in a slightly different form:

$$\begin{aligned} \hat{\mathbf{h}}^{(l+1)} &= (N_0^{-1} E_{\mathbf{S}}[\mathbf{S}^H \mathbf{S} | \mathbf{x}, \hat{\mathbf{h}}^{(l)}] + \mathbf{K}^{-1})^{-1} E_{\mathbf{S}}[\mathbf{S}^H | \mathbf{x}, \hat{\mathbf{h}}^{(l)}] N_0^{-1} \mathbf{x} \\ &= (E_{\mathbf{S}}[\mathbf{S}^H \mathbf{S} | \mathbf{x}, \hat{\mathbf{h}}^{(l)}] + N_0 \mathbf{K}^{-1})^{-1} E_{\mathbf{S}}[\mathbf{S}^H | \mathbf{x}, \hat{\mathbf{h}}^{(l)}] \mathbf{x} \end{aligned} \quad (6.7)$$

In the first EM iteration, \mathbf{S} represents the initial data sequence estimate made up of only the training symbols. All the remaining yet unknown data symbols are set equal to 0. \mathbf{K} is the known correlation matrix of Rayleigh fading channel. The noise variance N_0 can be looked at as a weight factor which balances the contributions of two terms: $E_{\mathbf{S}}[\mathbf{S}^H \mathbf{S} | \mathbf{x}, \hat{\mathbf{h}}^{(l)}]$ and \mathbf{K}^{-1} to the final result in (6.7). We will get more accurate initial channel estimate result $\hat{\mathbf{h}}^{(l+1)}$ when we give more weight to the known term \mathbf{K}^{-1} rather than relying on inaccurate initial data sequence estimate \mathbf{S} . Consequently, as the initial noise estimate we decided to use the power of the received sequence divided by 2. Below we justify this choice.

Designate $P_X = P_S + P_N$ to be the power of the received signal, where P_S is power of data symbols and P_N is power of noise. Then SNR is $SNR = \frac{P_S}{P_N}$. We consider the range of $SNR > 1$ (0 dB) then with the assumption of $P_S = 1$ there is $P_N < 1$. It can be observed that when $P_N < 1$ and $P_{N_{init}} = \frac{P_X}{2}$ there is always

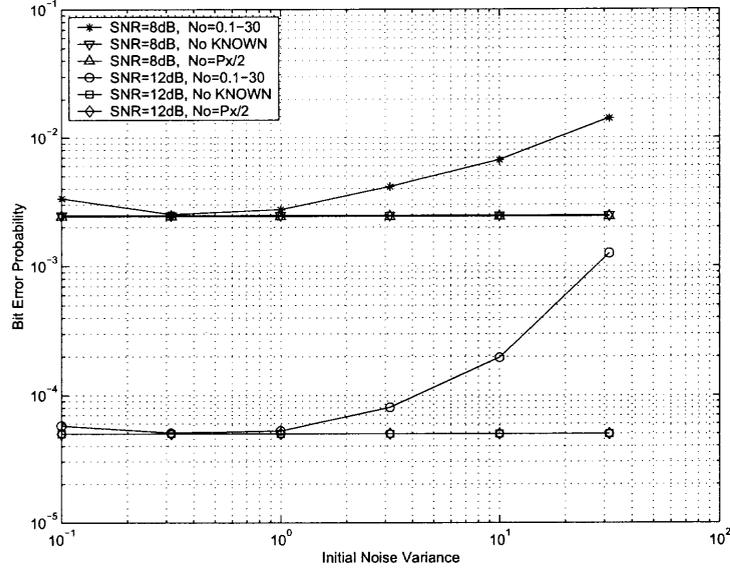


Figure 6.2 Bit error probability vs. initial value of white noise variance for SNR=8 dB and 12 dB.

P_{Ninit} - initial noise variance greater than true noise variance P_N :

$$P_{Ninit} = \frac{P_S + P_N}{2} = \frac{1 + P_N}{2} > P_N \quad (6.8)$$

when $P_N < 1$.

6.2 Performance of the Iterative APP-EM Receiver with Noise Variance Estimation - Numerical Results

The same as in Chapter 5, numerical results were obtained for the 4-PSK 8-state space time code presented in [13] and shown in Figure 2.2 with $N = 2$ transmit and $M = 2$ receive antennas. Each frame consisted of 14 training symbols followed by 116 data symbols to the total of $L = 130$ symbols per frame. Deterministic 14×10 symbol interleaver was used. We considered performance of the system over the quasi-static channel (Model A) introduced in Section 2.1 and over the time-varying Rayleigh fading channel (Model B) introduced in Section 2.2. The simulations were

run at two fading rates of the Rayleigh channel: slow fading $f_d T_s = 0.001$ and fast fading $f_d T_s = 0.01$.

The BEP and FEP vs. SNR performance plots were obtained for the iterative APP-EM receiver with CSI and noise variance being estimated. For reference, we also included in the figures two other performance plots: for the receiver with known CSI and known noise variance and for the receiver with CSI estimated and noise variance known.

Figures 6.3 and 6.4 show BEP and FEP vs. SNR performance of our receiver over a quasi-static channel. It can be observed that in both cases BEP and FEP, the performance of our receiver with unknown noise variance and CSI estimation is only a fraction of a dB worse than the performance of the system with known CSI and known noise variance.

Figures 6.5 and 6.6 show the BEP and FEP vs. SNR performance over the Rayleigh fading channel with slow fading rate $f_d T_s = 0.001$. In this case, the BEP performance of our receiver is within 0.2 dB and FEP performance is within 0.1 dB of the system with known CSI and known noise variance.

Figures 6.7 and 6.8 present the BEP and FEP vs. SNR performance over the Rayleigh fading channel with fast fading rate $f_d T_s = 0.01$. It can be observed that at $BEP = 10^{-3}$ the performance of our receiver with CSI and noise estimation is only 0.5 dB worse than performance of the reference system with known CSI and known noise variance. At $FEP = 10^{-2}$ the gap between our receiver and the reference system is 0.2 dB.

A flowchart of the simulation program used to obtain the performance plots is presented in the Appendix A.

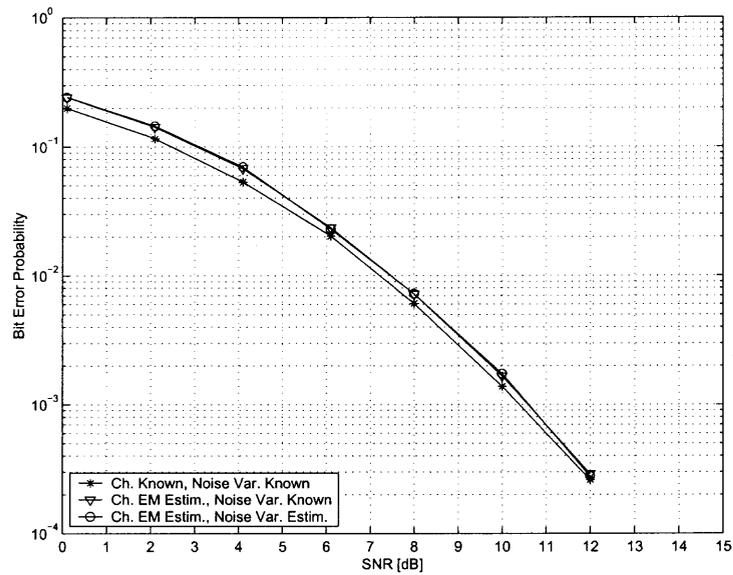


Figure 6.3 APP-EM iterative receiver - bit error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over quasi-static channel.

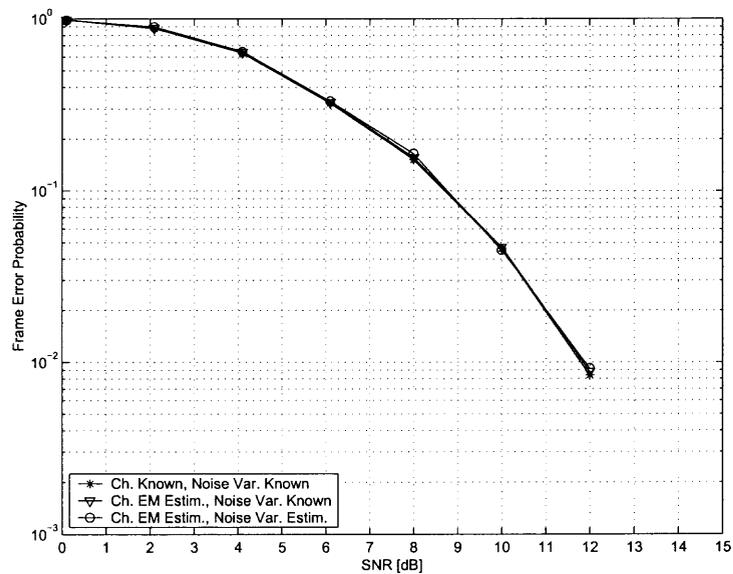


Figure 6.4 APP-EM iterative receiver - frame error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over quasi-static channel.

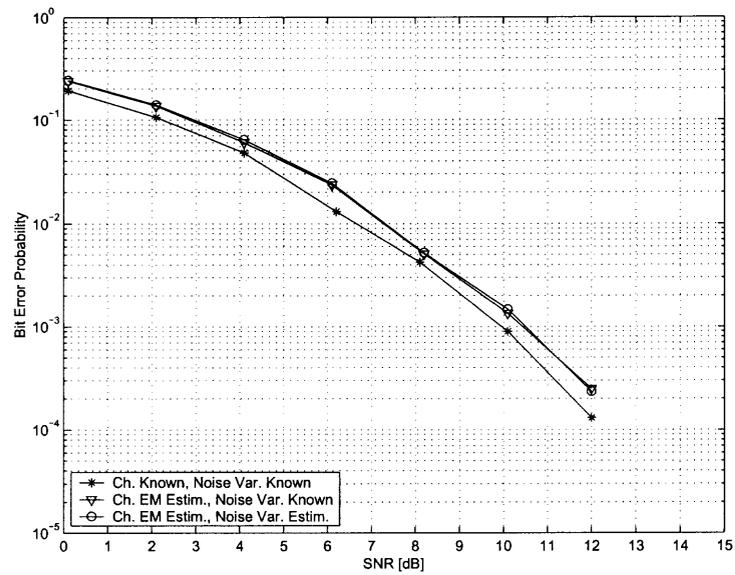


Figure 6.5 APP-EM iterative receiver - bit error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over Rayleigh fading channel with slow fading rate $f_d T_s = 0.001$.

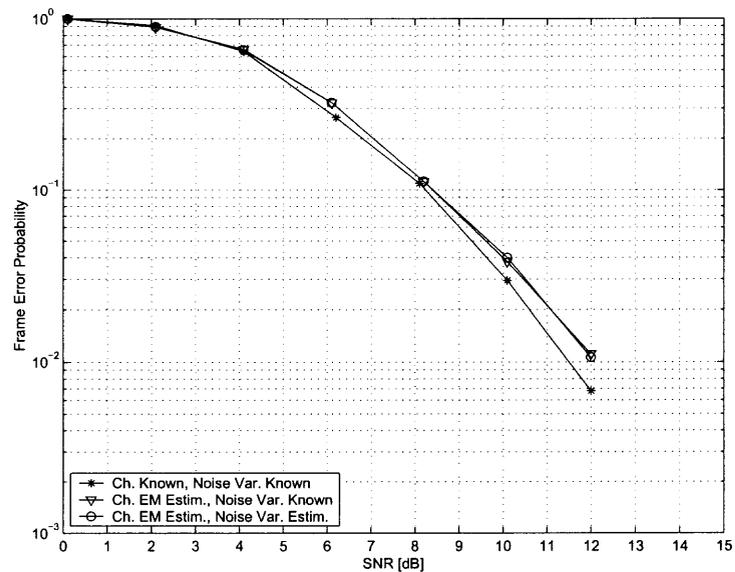


Figure 6.6 APP-EM iterative receiver - frame error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over Rayleigh fading channel with slow fading rate $f_d T_s = 0.001$.

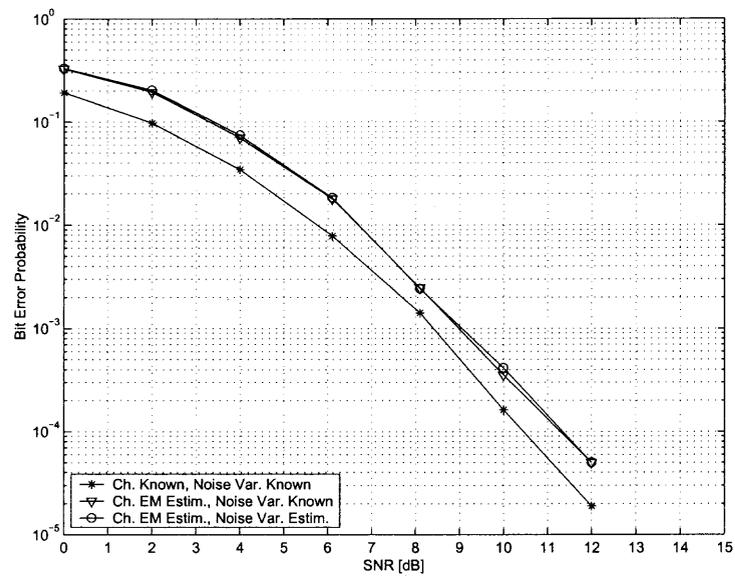


Figure 6.7 APP-EM iterative receiver - bit error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over Rayleigh fading channel with slow fading rate $f_d T_s = 0.01$.

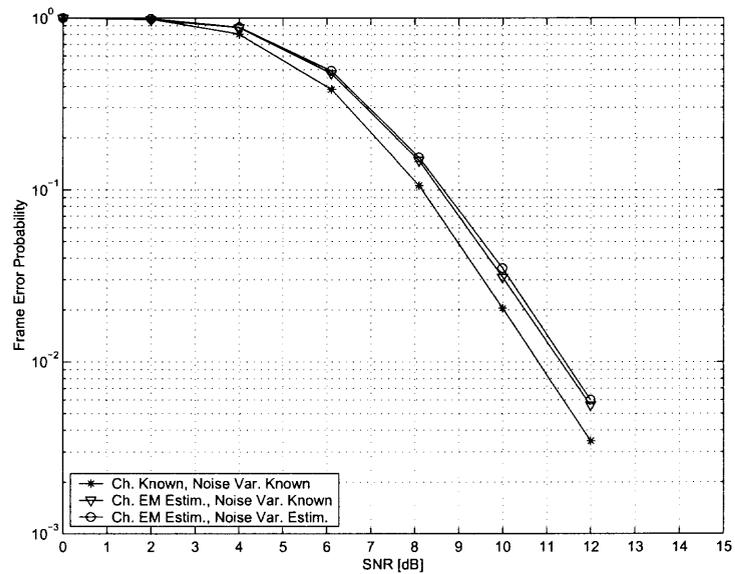


Figure 6.8 APP-EM iterative receiver - frame error probability vs. SNR per receive antenna with estimation of unknown AWGN variance over Rayleigh fading channel with slow fading rate $f_d T_s = 0.01$.

CHAPTER 7

INVESTIGATION OF THE APP-EM RECEIVER PROPERTIES

In this chapter, different properties of the APP-EM receiver are investigated. These include how the system performance depends on the training sequence length and the length of interleaver. In addition, the system performance with deterministic and random interleavers are compared and the effects of channel fading rate on the performance are investigated.

7.1 Effects of Training Sequence Length on System Performance

The trade-off between the quality of CSI/noise estimation and the SNR loss resulting from different number of training symbols used was evaluated. Figure 7.1 presents the plots BEP vs. number of training symbols expressed as a percentage of total frame length L with L kept constant at 130 symbols and number of training symbols set to 6, 10, 14, 18, 22, 30 and 38. For each training sequence length, the interleaver matrix dimension used is given in the table below.

Training Sequence Length	Interleaver Matrix Dimension
6	6×22
10	10×13
14	14×10
18	18×8
22	22×6
30	30×5
38	38×4

Three plots are shown for SNRs: 8, 10 and 12 dB over fast fading ($f_d T_s = 0.01$) channel. It can be seen from Figure 7.1 that the optimum number of training symbols resulting in the most efficient channel usage is when the training symbols make up

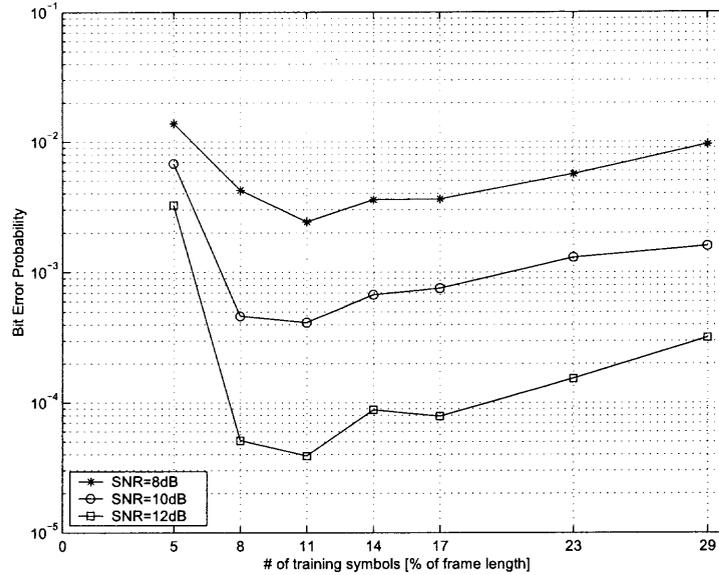


Figure 7.1 Bit error probability vs. training sequence length - effect of training sequence length on system performance, Rayleigh fading channel $f_d T_s = 0.01$.

about 10 – 11% of frame length. At $L = 130$ symbols it is equivalent to the number of training symbols $L_t = 14$ which agrees with TDMA IS-136 standard.

7.2 Study of the Interleaver and Channel Correlation Length Effects

The effect of interleaver length and channel correlation length considered in the computations was investigated by running the simulations with different lengths of training sequence L_t and the frame length L keeping the ratio $\frac{L}{L_t}$ constant. The dimension of the interleaver matrix was $L_t \times [\text{int}(\frac{L}{L_t}) + 1]$. Simulations were run for training sequence length L_t and frame length L given in the table below with the interleaver matrix dimensions adjusted appropriately.

Training Sequence Length	Frame Length	Interleaver Matrix Dimension
6	56	6×10
10	93	10×10
14	130	14×10
18	168	18×10
22	204	22×10
26	241	26×10
30	279	30×10

BEP vs. frame length L is shown in Figure 7.2 for fast fading Rayleigh channel ($f_d T_s = 0.01$). It can be seen from the figure that using frame lengths smaller than $L = 130$ symbols becomes inefficient because the short interleaver length and truncation of channel correlation function causes the increase of BEP. On the other hand, increasing the frame length beyond $L = 170$ symbols does not result in significant gain in performance but increases the decoding delay and complexity of computations (dimensions of matrices become greater). The optimum frame lengths are in the range $L = 130 - 170$ symbols which also agrees with TDMA IS-136 standard where $L = 130$.

7.3 Evaluation of the Deterministic Interleaver Gain

To evaluate the quality of the deterministic matrix interleaver used in most of the simulations, the case of random interleaver was considered. With the frame length $L = 130$ and training sequence $L_t = 14$, a random interleaver was generated different for each frame. Then, after averaging the BEP over all transmitted frames, the resulting performance was not a function of the interleaver. The BEP vs. SNR performance of the system with deterministic and random interleavers is shown in Figure 7.3. It can be observed that at $\text{BEP} = 10^{-3}$ the gain of about 0.5 dB is

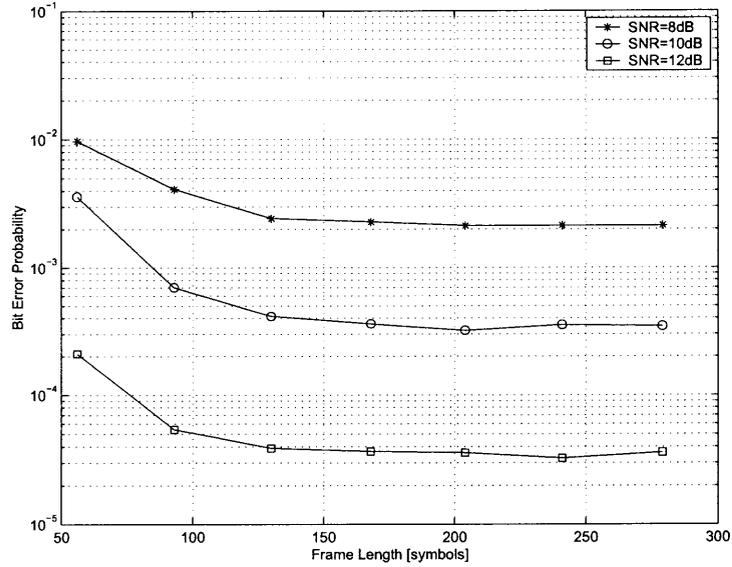


Figure 7.2 Bit error probability vs. frame length with fixed percentage of training symbols - fast fading Rayleigh channel $f_d T_s = 0.01$.

achieved for both cases of known and estimated CSI that can be attributed to the deterministic matrix interleaver over the random interleaver.

7.4 Effect of Channel Fading Rate on System Performance

The BEP performance of the receiver as a function of the channel fading rate has been investigated. Using frame length $L = 130$ symbols and training sequence length $L_t = 14$ symbols, the simulations were run with both the CSI and white noise variance being estimated at different channel fading rates. Figure 7.4 shows the BEP vs. SNR plots for channel fading rates $f_d T_s = 0.001, 0.01, 0.03, 0.05$ and 0.1 . It can be observed from the figure that at channel fading rates $f_d T_s$ greater than 0.03 the performance of the estimation algorithm becomes unacceptable. At lower fading rates in the range 0.001 to 0.01 , the performance is established by two main factors: diversity advantage and the quality of initial CSI estimate. Diversity advantage results from the use of the interleaver. It increases with the increase of the channel fading rate, which drives

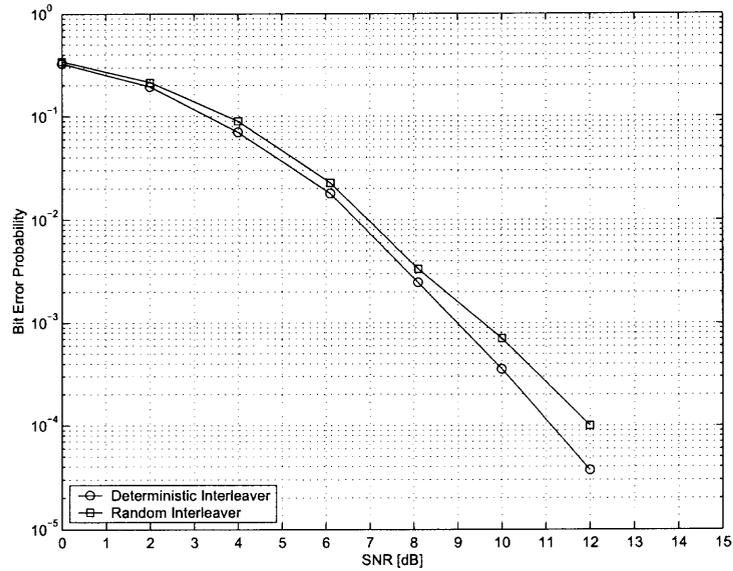


Figure 7.3 Bit error probability vs. SNR performance of the system with deterministic and random interleavers, Rayleigh fading channel $f_d T_s = 0.01$, channel and noise variance estimated.

the BEP down. At the same time, as the number of training symbols is kept fixed, the quality of initial CSI estimate gets worse with the increasing channel fading rate. It can be observed from Figure 7.4 that the diversity gain at $f_d T_s = 0.01$ is greater and the performance is slightly better than at $f_d T_s = 0.001$ for SNR greater than 5 dB. The iterative APP-EM receiver using frame length $L = 130$ symbols and $L_t = 14$ training symbols achieves the best performance when the channel fading rate $f_d T_s$ is in the range 0.001 to 0.03 with 0.01 being the optimum. In particular, the proposed receiver is not suitable to maintain reliable communications in the environment where the channel fading rates are greater than 0.03-0.05.

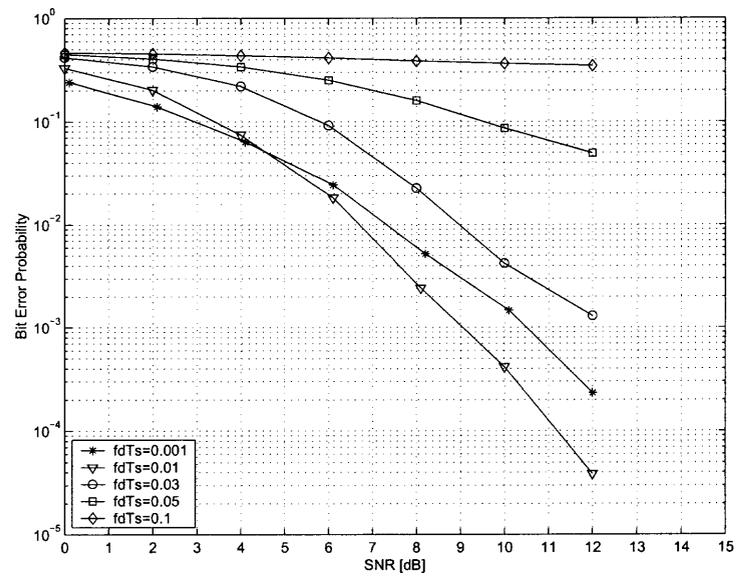


Figure 7.4 Bit error probability vs. SNR for different Rayleigh channel fading rates $f_d T_s$, channel and noise variance estimated.

CHAPTER 8

SUMMARY AND CONCLUSIONS

In this dissertation, the communication systems for space-time coded signals have been presented. The BEP versus SNR and FEP versus SNR performance plots of the different systems are shown in Figure 8.1 and 8.2 respectively. The best performance is achieved from the communication system employing symbol interleaving at the transmitter and the APP-EM receiver utilizing EM algorithm to do soft outputs sequence detection and channel estimation iteratively. The results of the simulations are presented showing the performance of the system over unknown quasi-static and Rayleigh fading channels. The operation of the proposed APP-EM receiver is based on EM iterations that produce channel and noise variance estimates utilizing symbol APPs. Noise covariance and channel estimates are subsequently used to get more accurate APPs. It is shown that for the channel models considered, the algorithm converges to within 0.2-0.5 dB of the BEP performance with known channel and known noise variance. Some of the system properties were also investigated, including how its performance depends on the number of training symbols and on the interleaver length. The performance as a function of Rayleigh channel fading rate suggests that the proposed system can serve as a powerful method for communications over unknown quasi-static and fading channels in the environments where the channel fading rates do not exceed $f_d T_s = 0.03$.

The following is the summary of contributions presented in this dissertation:

1. Employing symbol interleaver to eliminate error propagation between sequence detection and channel estimation in the EM-based iterative STCM receiver.
2. Implementation of the EM algorithm for channel estimation and data detection for STCM encoded signals with MAP as a criterion of optimality in both: estimation and detection.

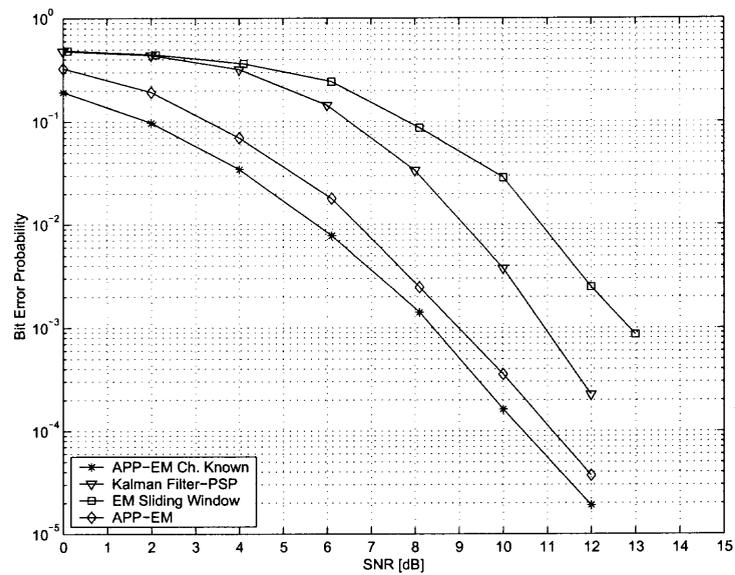


Figure 8.1 Comparison of bit error probability vs. SNR performance over Rayleigh channel with $f_d T_s = 0.01$ for different communication systems considered.

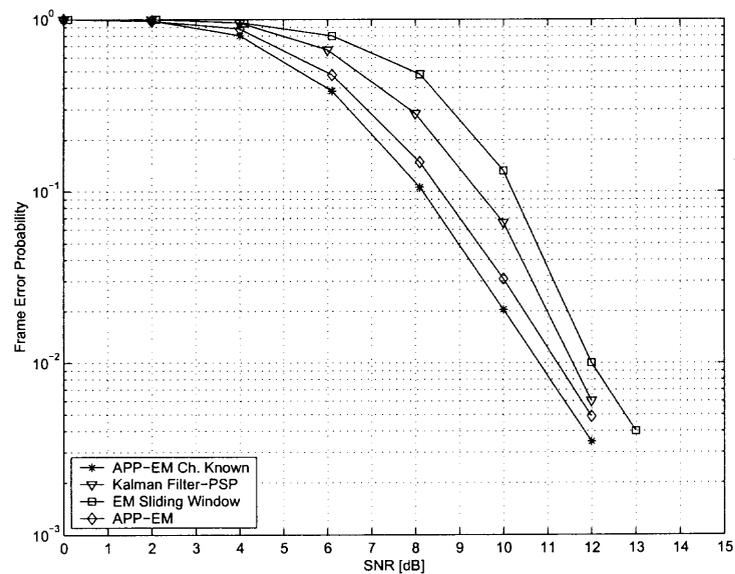


Figure 8.2 Comparison of frame error probability vs. SNR performance over Rayleigh channel with $f_d T_s = 0.01$ for different communication systems considered.

3. Estimation of noise variance in the EM loop of the STCM iterative receiver in addition to channel estimation and data detection.
4. Investigation of performance and other properties of the STCM communication system with symbol interleaver and EM-based iterative APP-EM receiver.
5. Evaluation of performance of the STCM receiver using Kalman filter with per-survivor processing for channel estimation and Viterbi algorithm for sequence detection.

APPENDIX

FLOWCHART OF THE SIMULATION PROGRAM

A flowchart of the Matlab simulation program is shown in Figures A.1, A.2 and A.3. The main program in Figure A.1 makes calls to several functions to generate the data to be transmitted within each frame, MIMO channels and noise. Then a function `fStcmAPPEMInt()` is called which contains all the signal processing at the receiver. After the received data has been decoded, the number of errors is counted and the control is transferred back to the beginning of the program to generate and transmit the next data frame. Figures A.2 and A.3 show the details of data processing in the receiver function `fStcmAPPEMInt()`.

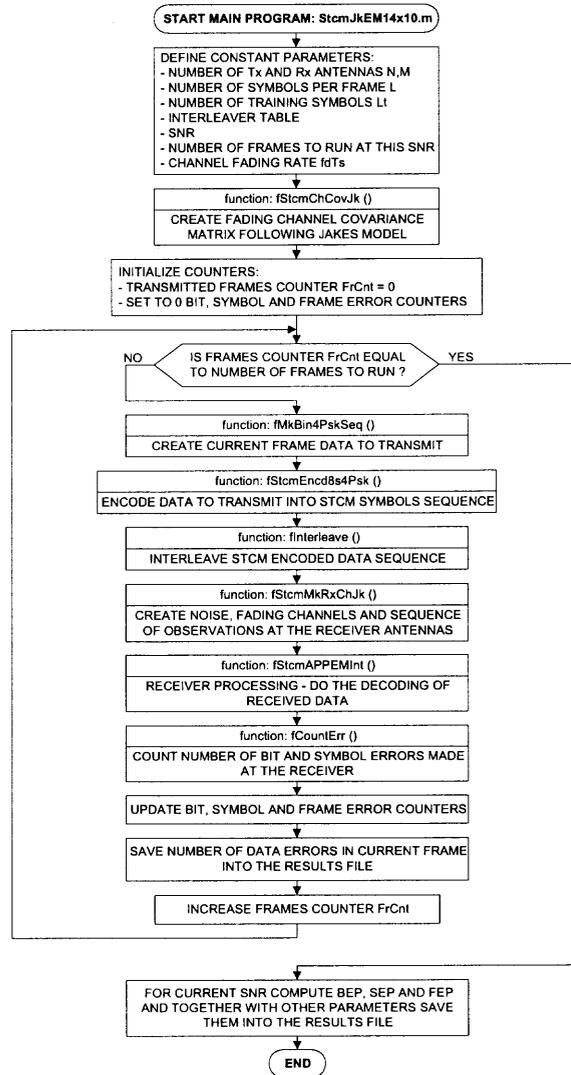


Figure A.1 Simulation program flowchart - main program.

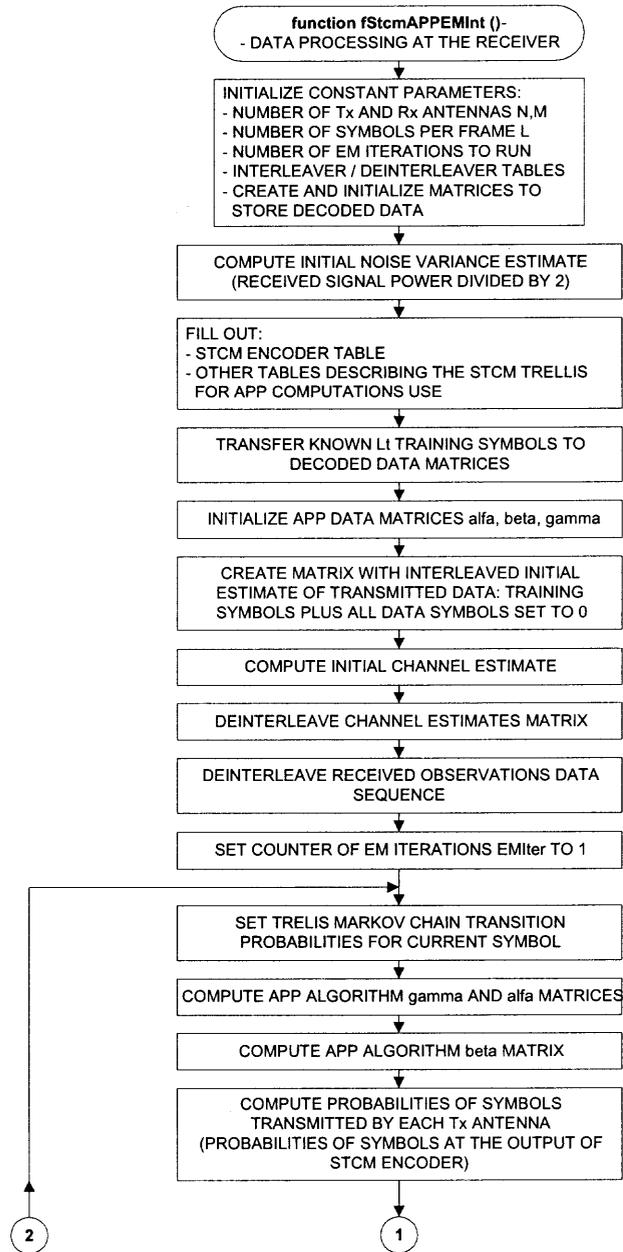


Figure A.2 Simulation program flowchart - receiver (1).

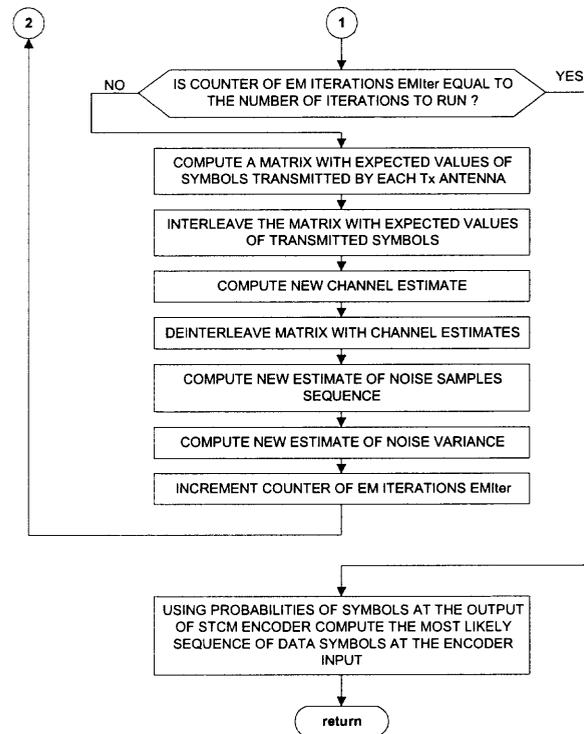


Figure A.3 Simulation program flowchart - receiver (2).

BIBLIOGRAPHY

- [1] G.D. Forney, "Maximum-likelihood sequence estimation of digital sequences in presence of intersymbol interference", *IEEE Trans. on Information Theory*, vol. 18, pp. 363-378, May 1972.
- [2] A.J. Viterbi, "Error Bound for Convolutional Codes an Asymptotically Optimum Decoding Algorithm", *IEEE Trans. on Information Theory*, vol. 13, pp. 260-269, Apr. 1967.
- [3] N. Seshadri, "Joint Data and Channel Estimation Using Blind Trellis Search Techniques", *IEEE Trans. on Commun.*, vol. 42, pp. 1000-1011, Feb./Mar./Apr. 1994.
- [4] Q. Dai, E. Shwedyk, "Detection of Bandlimited Signals over Frequency Selective Rayleigh Fading Channels", *IEEE Trans. on Commun.*, vol. 42, pp. 941-950, Feb. 1994.
- [5] F.R. Magee, J.G. Proakis, "Adaptive Maximum-Likelihood Sequence Estimation for Digital Signaling in Presence of Intersymbol Interference", *IEEE Trans. on Information Theory*, vol. 19, pp. 448-457, Jul. 1973.
- [6] S.U.H. Qureshi, E.E. Newhall, "An Adaptive Receiver for Data Transmission over Time-Dispersive Channels", *IEEE Trans. on Information Theory*, vol. 19, pp. 120-124, Jan. 1973.
- [7] G.E. Bottomley, K.J. Molnar, "Adaptive Channel Estimation for Multichannel MLSE Receivers", *IEEE Communications Letters*, vol. 3, No. 2, pp. 40-42, Feb. 1999.
- [8] L. Lindbom, "Simplified Kalman Estimation of Fading Mobile Radio Channels: High Performance at LMS Computational Load", ICASSP-93, vol. 3, pp. 352-355, 1993.
- [9] R. Raheli, A. Polydoros, C.K. Tzou, "Per-Survivor Processing: A General Approach to MLSE in Uncertain Environments", *IEEE Trans. on Commun.*, vol. 43, pp. 354-364, Feb./Mar./Apr. 1995.
- [10] K.M. Chugg, A. Polydoros, "MLSE for an Unknown Channel - Part I: Optimality Considerations", *IEEE Trans. on Commun.*, vol. 44, pp. 836-846, Jul. 1996.
- [11] K.M. Chugg, A. Polydoros, "MLSE for an Unknown Channel - Part II: Tracking Performance", *IEEE Trans. on Commun.*, vol. 44, pp. 949-958, Aug. 1996.
- [12] T. Vaidis, C.L. Weber, "Block Adaptive Techniques for Channel Identification and Data Demodulation Over Band-Limited Channels", *IEEE Trans. on Commun.*, vol. 46, pp. 232-243, Feb. 1998.

- [13] V. Tarokh, N. Seshadri, A.R. Calderbank, "Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction", *IEEE Trans. on Information Theory*, vol. 44, pp. 744-765, Mar. 1998.
- [14] A.F. Naguib, V. Tarokh, N. Seshadri, A.R. Calderbank, "Space-Time Coding Modem for High-Data-Rate Wireless Communications", *IEEE Journal on Selected Areas in Commun.*, vol. 16, pp. 1459-1478, Oct. 1998.
- [15] G.J. Foschini, "Layered Space-Time Architecture for Wireless Communications in a Fading Channel Environment when Using Multi-Element Antennas", *AT&T, Bell-Labs Technical Journal*, pp. 41-59, Autumn 1996.
- [16] P.W. Wolniansky, G.J. Foschini, G.D. Golden, R.A. Valenzuela, "V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel", *IEEE International Symposium on Signals, Systems and Electronics 1998*, pp. 295-300.
- [17] D. Cui, A. Haimovich, "Adaptive Sequence Detection for Space-Time Coded Modulation with Cochannel Interference", Proceedings of the 33rd CISS99, The John Hopkins University, March 1999..
- [18] A.P. Dempster, N.M. Laird, D.B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm", *Journal of the Royal Statistical Society*, vol. B39, pp. 1-38, 1977.
- [19] H. Zamiri-Jafarian, S. Pasupathy, "EM-Based Recursive Estimation of Channel Parameters", *IEEE Trans. on Commun.*, vol. 47, pp. 1297-1302, Sep. 1999.
- [20] H. Zamiri-Jafarian, S. Pasupathy, "Adaptive MLSDE Using the EM Algorithm", *IEEE Trans. on Communications*, vol. 47, pp. 1181-1193, Aug. 1999.
- [21] C.N. Georghiades, J.C. Han, "Sequence Estimation in the Presence of Random Parameters via the EM Algorithm", *IEEE Trans. on Commun.*, vol. 45, pp. 300-307, Mar. 1997.
- [22] Y. Li, C.N. Georghiades, G. Huang, "Iterative Maximum-Likelihood Sequence Estimation for Space-Time Coded Systems", *IEEE Trans. on Commun.*, vol. 49, pp. 948-951, Jun. 2001.
- [23] R. Perry, W.A. Berger, K. Buckley, "EM Algorithm for Sequence Estimation over Gauss-Markov ISI Channels", 2000 IEEE International Conference on Communications, vol. 1, pp. 16-20.
- [24] L.B. Nelson, H.V. Poor, "Iterative Multiuser Receivers for CDMA Channels: an EM-based Approach", *IEEE Trans. on Commun.*, vol. 44, pp. 1700-1710, Dec. 1996.

- [25] M. Kobayashi, J. Boutros, G. Caire, "Successive Interference Cancellation with SISO Decoding and EM Channel Estimation", *IEEE Journal on Selected Areas in Commun.*, vol. 19, pp. 1450-1460, Aug. 2001.
- [26] C. Cozzo, B.L. Hughes, "An Iterative Receiver for Space-Time Communications", 2000 Conference on Information Sciences and Systems, Princeton, Mar. 2000.
- [27] D. Divsalar, M. Simon, "The Design of Trellis Coded MPSK for Fading Channel: Performance Criteria", *IEEE Trans. Commun.*, vol. 36, pp. 1004-1012, Sep. 1988.
- [28] E. Chiavaccini, G.M. Vitetta, "MAP Symbol Estimation on Frequency-Flat Rayleigh Fading Channels Via a Bayesian EM Algorithm", *IEEE Trans. on Commun.*, vol. 49, pp. 1869-1872, Nov. 2001.
- [29] W.C. Jakes, "Microwave Mobile Communications", Wiley, New York, 1974.
- [30] P. Dent, G.E. Bottomley, T. Croft, "Jakes Fading Model Revisited", *Electronic Letters*, vol. 29, pp. 1162-1163, 24th Jun. 1993.
- [31] S.M. Kay, "Fundamentals of Statistical Signal Processing - vol.1 - Estimation Theory", Prentice-Hall 1993.
- [32] S.M. Kay, "Fundamentals of Statistical Signal Processing - vol.2 - Detection Theory", Prentice-Hall 1994.
- [33] G.M. Vitetta, D.P. Taylor, "Maximum Likelihood Decoding of Uncoded and Coded PSK Signal Sequences Transmitted Over Rayleigh Flat-Fading Channels", *IEEE Trans. on Commun.*, vol. 43, pp. 2750-2758, Nov. 1995.
- [34] H. Kong, E. Shwedyk, "On Channel Estimation and Sequence Detection of Interleaved Coded Signals over Frequency Nonselective Rayleigh Fading Channels", *IEEE Trans. on Vehicular Technol.*, vol. 47, pp. 558-565, May. 1998.
- [35] S. Gazor, A.M. Rabiei, S. Pasupathy, "Synchronized per Survivor MLSD Receiver Using a Differential Kalman Filter", *IEEE Trans. on Commun.*, vol. 50, pp. 364-368, Mar. 2002.
- [36] Y.C. Wang, J.T. Chen, "Space-Time Channel Tracking for Multipath Fast-Fading Channels", IEEE International Conference on Communications ICC 2001, vol. 2, pp. 382-386.
- [37] K. Park, J.W. Modestino, "An EM-based Procedure for Iterative Maximum-Likelihood Decoding and Simultaneous Channel State Estimation on Slow Fading Channels", Proceedings of IEEE International Symposium on Information Theory 1994, p. 27.
- [38] G.K. Kaleh, R. Vallet, "Joint Parameter Estimation and Symbol Detection for Linear or Nonlinear Unknown Channels", *IEEE Trans. on Commun.*, vol. 42, pp. 2406-2413, Jul. 1994.

- [39] T.K. Moon, "The Expectation-Maximization Algorithm", *IEEE Signal Processing Magazine*, pp. 47-60, Nov. 1996.
- [40] A. Ansari, R. Viswanathan, "Application of Expectation-Maximization Algorithm to the Detection of a Direct-Sequence Signal in Pulsed Noise Jamming", *IEEE Trans. on Commun.*, vol. 41, pp. 1151-1154, Aug. 1993.
- [41] X. Ma, M. Kobayashi, S.C. Schwartz, "EM-based Channel Estimation for OFDM", *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing PACRIM 2001*, vol. 2 pp.449-452. Direct-Sequence Signal in Pulsed Noise Jamming", *IEEE Trans. on Commun.*, vol. 41, pp. 1151-1154, Aug. 1993.
- [42] H. Zamiri-Jafarian, S. Pasupathy, "Recursive Channel Estimation for Wireless Communication via the EM Algorithm", *IEEE International Conference on Personal Wireless Communications 1997*, pp. 33-37.
- [43] J.K. Cavers, "An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels", *IEEE Trans. on Vehicular Technol.*, vol. 40, pp. 686-693, Nov. 1991
- [44] J.A. Fessler, A.O. Hero, "Space-Alternating Generalized Expectation-Maximization Algorithm", *IEEE Trans. on Signal Proc.*, vol. 42, pp. 2664-2677, Oct. 1994.
- [45] C.N. Georghiades, D.L. Snyder, "The Expectation-Maximization Algorithm for Symbol Unsynchronized Sequence Detection", *IEEE Trans. on Commun.*, vol. 39, pp. 54-61, Jan. 1991.
- [46] J.C. Han, C.N. Georghiades, "Maximum-Likelihood Sequence Estimation for Fading Channels via the EM Algorithm", *IEEE Global Telecommunications Conference GLOBECOM 1993*, vol. 4, pp. 133-137.
- [47] Y. Li, C.N. Georghiades, G. Huang, "EM-based Sequence Estimation for Space-Time Codes Systems", *IEEE International Symposium on Information Theory 2000*, p. 315.
- [48] Q. Li, C.N. Georghiades, X. Wang, "An Iterative Receiver for Turbo-Coded Pilot-Assisted Modulation in Fading Channels", *Commun. Letters*, vol. 5, pp. 145-147, Apr. 2001.
- [49] Q. Li, C.N. Georghiades, X. Wang, "An Iterative Decoding Scheme for Pilot-Assisted Modulation in Fading Channels", *IEEE Global Telecommunications Conference GLOBECOM 2000*, vol. 2, pp. 807-811.
- [50] C. Cozzo, B.L. Hughes, "Joint Detection and Estimation in Space-Time Coding and Modulation", *33rd Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA 1999*, vol. 1, pp. 613-617.
- [51] M. Feder, E. Weinstein, "Parameter Estimation of Superimposed Signals Using the EM Algorithm", *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 36, pp. 477-489, Apr. 1988.

- [52] A. Anastasopoulos, K.M. Chugg, "Adaptive Soft-input Soft-output Algorithms for Iterative Detection with Parametric Uncertainty", *IEEE Trans. on Commun.*, vol. 48, pp. 1638-1649, Oct. 2000.
- [53] L.M. Davis, I.B. Collings, P. Hoeher, "Joint MAP Equalization and Channel Estimation for Frequency-Selective and Frequency-Flat Fast Fading Channels", *IEEE Trans. on Commun.*, vol. 49, pp. 2106-2114, Dec. 2001.
- [54] D. Cui, A. Haimovich, "Parallel Concatenated Turbo Multiple Antenna Coded Modulation: Principles and Performance Analysis", *IEEE Journal on Selected Areas of Commun.* (submitted).
- [55] L.R. Bahl, J. Cocke, F. Jelinek, J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate", *IEEE Trans. on Inf. Theory*, pp. 284-287, Mar. 1974.
- [56] C. Komninakis, R.D. Wesel, "Pilot-Aided Joint Data and Channel Estimation in Flat Correlated Fading", *IEEE Global Telecommunications Conference GLOBECOM 1999*, pp. 2534-2539.
- [57] S. Haykin, "Communication Systems", John Wiley & Sons, Inc., 1994.
- [58] S. Haykin, "Adaptive Filter Theory", Prentice-Hall, 1996.
- [59] J. M. Mendel, "Lessons in Estimation Theory for Signal Processing, Communications and Control", Prentice-Hall, 1995.
- [60] W. Turin, "Digital Transmission Systems: Performance Analysis and Modeling", McGraw-Hill, 1998.
- [61] C. Komninakis, R.D. Wesel, "Joint Iterative Channel Estimation and Decoding in Flat Correlated Rayleigh Fading", *IEEE Journal on Selected Areas in Commun.*, vol. 19, pp. 1706-1717, Sep. 2001.
- [62] W. Turin, "MAP Decoding Using the EM Algorithm", pp. 1866-1870, *Proceedings of Vehicular Technology Conference 1999*.
- [63] A.L. Garcia, "Probability and Random Processes for Electrical Engineering", Addison-Wesley Publishing Co., 1993.
- [64] R. Matsuo, T. Ohtsuki, I. Sasase, "Simple EM Based Channel Estimation and Data Detection Using Estimation in the Preceding Frame in Space-Time Coded Signals", *ISIT2001*, Washington DC, Jun. 2001, pp. 222.
- [65] C. Brutel, J. Boutros, P. Mege, "Iterative Joint Channel Estimation and Detection of Coded CPM", *2000 International Zurich Seminar on Broadband Communications 2000*, pp. 287-292.

- [66] P.J. Chung, J.F. Bohme, "Recursive EM and SAGE Algorithms", Proceedings of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing 2001, pp. 540-543.
- [67] H.J. Su, E. Geraniotis, "Low-Complexity Joint Channel Estimation and Decoding for Pilot Symbol-Assisted Modulation and Multiple Differential Detection Systems with Correlated rayleigh Fading", *IEEE Trans. on Commun.*, vol. 50, pp. 249-261, Feb. 2002
- [68] J. Miguez, L. Castedo, "Iterative Space-Time Soft Detection in Time-Varying Multiaccess Wireless Channels", Proceedings of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing 2001, pp. 114-117.
- [69] A. Dogandzic, A. Nehorai, "Space-Time Fading Channel Estimation and Symbol Detection in Unknown Spatially Correlated Noise", *IEEE Trans. on Signal Proc.*, vol. 50, pp. 457-474, Mar. 2002.
- [70] C.H. Aldana, J. Cioffi, "Channel Tracking for Multiple Input Single Output Systems Using EM Algorithm", IEEE International Conference on Communications, vol. 2 pp. 586-590, ICC 2001.
- [71] S.J. Orfanidis, "Optimum Signal Processing", McGraw-Hill, 1988.
- [72] J.G. Proakis, "Digital Communications", McGraw-Hill, 1995.
- [73] J.G. Proakis, D.G. Manolakis, "Digital Signal Processing: Principles, Algorithms and Applications", Macmillan Publishing Co., 1992.
- [74] T.S. Rappaport, "Wireless Communications", Prentice-Hall, 1996.
- [75] G.J. McLachlan, T. Krishnan, "The EM Algorithm and Extensions", Wiley John & Sons, Inc., Oct. 1996.