

Advanced Signal Processing Algorithms for Wireless Communications

Guest Editors: Erdal Panayırçı, Costas Georghiades,
Xiaodong Wang, and Hakan A. Çırpan



EURASIP Journal on
Wireless Communications and Networking

Advanced Signal Processing Algorithms for Wireless Communications

EURASIP Journal on
Wireless Communications and Networking

Advanced Signal Processing Algorithms for Wireless Communications

Guest Editors: Erdal Panayırıcı, Costas Georghiadés,
Xiaodong Wang, and Hakan A. Çırpan



Copyright © 2005 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in volume 2005 of “EURASIP Journal on Wireless Communications and Networking.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editor-in-Chief

Phillip Regalia, Institut National des Telecommunications, France

Associate Editors

Thushara Abhayapala, Australia
Farid Ahmed, USA
Alagan Anpalagan, Canada
Anthony C. Boucouvalas, UK
Jonathon Chambers, UK
Biao Chen, USA
Pascal Chevalier, France
Chia-Chin Chong, Korea
Soura Dasgupta, USA
PetarM. Djuric, USA
Abraham Fapojuwo, Canada
Michael Gastpar, USA
Alex B. Gershman, Canada
Wolfgang Gerstacker, Germany
David Gesbert, France

Fary Ghassemlooy, UK
Alfred Hanssen, Norway
Stefan Kaiser, Germany
G. K. Karagiannidis, Greece
Hyung-Myung Kim, Korea
Chi Chung Ko, Singapore
Richard J. Kozick, USA
Bhaskar Krishnamachari, USA
Vincent Lau, Hong Kong
Dave Laurenson, Scotland
Tho Le-Ngoc, Canada
Tongtong Li, USA
Wei (Wayne) Li, USA
Steve McLaughlin, UK
Marc Moonen, Belgium

Eric Moulines, France
Sayandev Mukherjee, USA
A. Nallanathan, Singapore
Kamesh Namuduri, USA
Athina Petropulu, USA
H. Vincent Poor, USA
Brian Sadler, USA
Ivan Stojmenovic, Canada
Lee Swindlehurst, USA
Sergios Theodoridis, Greece
Lang Tong, USA
Luc Vandendorpe, Belgium
Yang Xiao, USA
Lawrence Yeung, Hong Kong
Weihua Zhuang, Canada

Contents

Editorial, Erdal Panayırçı, Costas Georgiades, Xiaodong Wang, and Hakan A. Çırpan
Volume 2005 (2005), Issue 2, Pages 79-82

A Receiver for Differential Space-Time $\pi/2$ -Shifted BPSK Modulation Based on Scalar-MSDD and the EM Algorithm, Michael L. B. Riediger, Paul K. M. Ho, and Jae H. Kim
Volume 2005 (2005), Issue 2, Pages 83-91

The Extended-Window Channel Estimator for Iterative Channel-and-Symbol Estimation, Renato R. Lopes and John R. Barry
Volume 2005 (2005), Issue 2, Pages 92-99

Soft-In Soft-Output Detection in the Presence of Parametric Uncertainty via the Bayesian EM Algorithm, A. S. Gallo and G. M. Vitetta
Volume 2005 (2005), Issue 2, Pages 100-116

A Theoretical Framework for Soft-Information-Based Synchronization in Iterative (Turbo) Receivers, Nele Noels, Vincenzo Lottici, Antoine Dejonghe, Heidi Steendam, Marc Moeneclaey, Marco Luise, and Luc Vandendorpe
Volume 2005 (2005), Issue 2, Pages 117-129

Adaptive Blind Multiuser Detection over Flat Fast Fading Channels Using Particle Filtering, Yufei Huang, Jianqiu (Michelle) Zhang, Isabel Tienda Luna, Petar M. Djurić, and Diego Pablo Ruiz Padillo
Volume 2005 (2005), Issue 2, Pages 130-140

Blind Decoding of Multiple Description Codes over OFDM Systems via Sequential Monte Carlo, Zigang Yang, Dong Guo, and Xiaodong Wang
Volume 2005 (2005), Issue 2, Pages 141-154

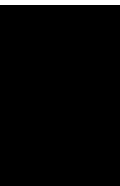
Adaptive Iterative Soft-Input Soft-Output Parallel Decision-Feedback Detectors for Asynchronous Coded DS-CDMA Systems, Wei Zhang, Claude D'Amours, and Abbas Yongaçoğlu
Volume 2005 (2005), Issue 2, Pages 155-162

A Low-Complexity KL Expansion-Based Channel Estimator for OFDM Systems, Habib Şenol, Hakan A. Çırpan, and Erdal Panayırçı
Volume 2005 (2005), Issue 2, Pages 163-174

Impact of Channel Estimation Errors on Multiuser Detection via the Replica Method, Husheng Li and H. Vincent Poor
Volume 2005 (2005), Issue 2, Pages 175-186

Factor-Graph-Based Soft Self-Iterative Equalizer for Multipath Channels, Ben Lu, Guosen Yue, Xiaodong Wang, and Mohammad Madihian
Volume 2005 (2005), Issue 2, Pages 187-196

Estimation of Directions of Arrival by Matching Pursuit (EDAMP), Güneş Z. Karabulut, Tolga Kurt, and Abbas Yongaçoğlu
Volume 2005 (2005), Issue 2, Pages 197-205



Blind Multiuser Detection for Long-Code CDMA Systems with Transmission-Induced Cyclostationarity, Tongtong Li, Weiguo Liang, Zhi Ding, and Jitendra K. Tugnait
Volume 2005 (2005), Issue 2, Pages 206-215

Adaptive Space-Time-Spreading-Assisted Wideband CDMA Systems Communicating over Dispersive Nakagami-m Fading Channels, Lie-Liang Yang and Lajos Hanzo
Volume 2005 (2005), Issue 2, Pages 216-230

Opportunistic Carrier Sensing for Energy-Efficient Information Retrieval in Sensor Networks, Qing Zhao and Lang Tong
Volume 2005 (2005), Issue 2, Pages 231-241

Editorial

Erdal Panayircı

Department of Electronics Engineering, IŞIK University, Maslak 80670, Istanbul, Turkey
Email: eepanay@isikun.edu.tr

Costas Georghiades

Electrical Engineering Department, Texas A&M University, College Station, TX 77843-3128, USA
Email: georghiades@tamu.edu

Xiaodong Wang

Department of Electrical Engineering, Columbia University, 500 West 120th Street, New York, NY 10027, USA
Email: wangx@ee.columbia.edu

Hakan A. Çirpan

Department of Electrical-Electronics Engineering, Istanbul University, Avcılar 34850, Istanbul, Turkey
Email: hcirpan@istanbul.edu.tr

Traditional wireless technologies are confronted with new challenges in meeting the ubiquity and mobility requirements of cellular systems. Hostile channel characteristics and limited bandwidths in wireless applications provide key barriers that future generation systems must cope with. Advanced signal processing methods, such as

- (i) the expectation-maximization (EM) algorithm,
- (ii) the SAGE algorithm,
- (iii) the Baum-Welch algorithm,
- (iv) per-survivor processing,
- (v) Kalman filters and their extensions,
- (vi) hidden Markov modeling,
- (vii) sequential Monte Carlo filters,
- (viii) stochastic approximation algorithms,

in collaboration with inexpensive and rapid computing power provide a promising avenue for overcoming the limitations of current technologies. Applications of the advanced signal processing algorithms mentioned above include, but are not limited to, joint/blind/sequence detection, decoding, synchronization, equalization, as well as channel estimation techniques employed in advanced wireless communication systems, such as OFDM/OFDMA, space-time-frequency coding, MIMO, CDMA, and multiuser detection *in time- and frequency-selective MIMO channels*. In particular, the development of suitable algorithms for wireless multiple-access systems in nonstationary and interference-rich environments presents major challenges to the system designer.

While considerable previous work has addressed many aspects of this problem separately, for example, single-user channel equalization, interference suppression for multiple-access channels, and tracking of time-varying channels, the problem of jointly combining these impairments in wireless channels has only recently become significant. On the other hand, the optimal solutions mostly cannot be implemented in practice because of their prohibitively high computational complexity. The statistical tools implemented by the advanced signal processing techniques above provide promising new routes for the design of low-complexity signal processing techniques with performance approaching the theoretical optimum for fast and reliable communication in the highly severe and dynamic wireless environment.

Although over the past decade such methods have been successfully applied in a variety of communication contexts, many technical challenges remain in emerging applications, whose solutions will provide the bridge between the theoretical potential of such techniques and their practical utility.

Key knowledge gaps here concern the following.

- (i) Theoretical performance and convergence analyses of these algorithms.
- (ii) New and efficient algorithms need to be developed for the problems mentioned above.
- (iii) Computational complexity problems of these algorithms when applied to on-line implementations of some algorithms running in digital receivers must be handled.

- (iv) Implementation of these algorithms based on batch processing and sequential (adaptive) processing depending on how the data are processed has not been completely solved for some of the techniques mentioned above.
- (v) Although research on sequential Monte Carlo signal processing has only recently begun, many optimal signal processing problems found in wireless communications, such as mitigation of various types of radio-frequency interference, tracking of fading channels, resolving multipath channel dispersion, space-time processing, and exploiting coded signal structures, represent a few problems waiting to be solved under the powerful Monte Carlo signal processing framework.

The call for papers for this issue solicited papers describing state-of-the-art research in advanced signal processing algorithms, that is, methods and techniques specifically designed for the next-generation wireless communication systems. Except for the two invited papers, the papers that follow this editorial were selected on the basis of blind peer review. The papers selected cover several key research topics, and specifically, the following:

- (i) EM algorithms and techniques,
- (ii) sequential Monte Carlo methods,
- (iii) iterative RLS techniques.

Four papers follow on the subject of EM algorithm applications. In the paper, "A receiver for differential space-time $\pi/2$ -shifted BPSK modulation based on scalar-MSDD and the EM algorithm," Riediger et al. address the problem of blind detection of Alamouti-type differential *space-time* (ST) modulation in static Rayleigh fading channels. They apply an iterative *expectation-maximization* (EM) algorithm which performs joint channel estimation and sequence detection. To further increase receiver performance, this algorithm uses minimum mean square estimation to obtain channel estimates and the maximum likelihood principle to detect the transmitted sequence, followed by differential decoding. The next paper, "The extended-window channel estimator for iterative channel-and-symbol estimation" by Lopes and Barry, considers the application of the EM algorithm to channel estimation which results in a well-known iterative channel-and-symbol estimator (ICSE). But, since the EM-ICSE has high complexity, and it is prone to mis-convergence, the authors propose a novel extended-window (EW) channel estimator for ICSE that can be used with any soft-output symbol estimator. Therefore, the symbol estimator may be chosen according to performance or complexity specifications. In the third paper, "Soft-in soft-output detection in the presence of parametric uncertainty via the Bayesian EM algorithm," Gallo and Vitetta investigate the application of the Bayesian expectation-maximization (BEM) technique to the design of soft-in soft-out (SISO) detection algorithms for wireless communication systems operating over channels affected by parametric uncertainty. In particular, the authors analyze the problems of SISO detection of spread-spectrum, single-carrier, and multicarrier space-time

block-coded signals and show that BEM-based detectors perform close to the maximum-likelihood receivers under perfect channel state information as long as channel variations are not too fast. The last paper on EM algorithms entitled "A theoretical framework for soft-information-based synchronization in iterative (turbo) receivers," by Noels et al., is concerned with turbo synchronization by an EM algorithm. The algorithm makes use of soft-data information to estimate parameters like carrier phase, frequency, or timing offsets within a turbo receiver. In the paper, a general theoretical framework for turbo synchronization is provided, which enables the derivation of parameter estimation procedures for carrier phase and frequency offsets, timing offset, and channel gain.

Sequential Monte Carlo technique with applications to wireless communications is examined in the following two papers. In the first paper, "Adaptive blind multiuser detection over flat fast fading channels using particle filtering," Huang et al. propose a method for blind multiuser detection (MUD) in synchronous systems over flat and fast Rayleigh fading channels employing a low-complexity particle filtering and a mixture Kalman filtering technique. To describe the dynamics of the addressed multiuser system, they suggest a novel time-observation state-space model (TOSSM) by adopting an autoregressive-moving-average (ARMA) process to model the temporal correlation of the channels. They further propose to use a more efficient PF algorithm known as the stochastic M -algorithm. In the second paper, "Blind decoding of multiple description codes over OFDM systems via sequential Monte Carlo," the authors Z. Yang et al. develop a blind soft-input soft-output OFDM detector, which is based on the sequential Monte Carlo method. Multiple description scalar quantization (MDSQ) is applied first to the continuous source signal, resulting in two correlated source descriptions. The two descriptions are then OFDM modulated and transmitted through two parallel frequency-selective fading channels. At the receiver, a blind turbo receiver is developed for joint OFDM demodulation and MDSQ decoding. Transformation of the extrinsic information of the two descriptions is exchanged between each other to improve system performance. Finally, they also treat channel-coded systems and develop a novel blind turbo receiver for joint demodulation, channel decoding, and MDSQ source decoding.

The following two papers deal with efficient design of adaptive detectors and channel estimators based on the least mean square, the recursive least squares, and the low-complexity minimum mean square batch estimation techniques. The first paper, "Adaptive iterative soft-input soft-output parallel decision-feedback detectors for asynchronous coded DS-CDMA systems" by Zhang et al., employs adaptive algorithms in the SISO multiuser detector in order to avoid the need for a priori information which is essential for the optimum and many suboptimum iterative soft-input soft-output (SISO) multiuser detectors. After deriving the optimum SISO parallel decision-feedback detector for asynchronous coded DS-CDMA systems, they propose two adaptive versions of this SISO detector, which are based

on the normalized least mean square (NLMS) and recursive least squares (RLS) algorithms which effectively exploit the a priori information of coded symbols, whose soft inputs are obtained from a bank of single-user decoders, to further improve their convergence performance. Furthermore, they consider how to select practical finite feedforward and feedback filter lengths to obtain a good tradeoff between the performance and computational complexity of the receiver. The second paper, entitled “A low-complexity KL expansion-based channel estimator for OFDM systems” by Şenol et al., proposes a computationally efficient, pilot-aided linear minimum mean square error (MMSE) batch channel estimation algorithm for OFDM systems in unknown wireless fading channels. The approach employs a convenient representation of the discrete multipath fading channel based on the Karhunen-Loeve (KL) orthogonal expansion and finds MMSE estimates of the uncorrelated KL series expansion coefficients. Based on such an expansion, no matrix inversion is required in the proposed MMSE estimator. Moreover, optimal rank reduction is achieved by exploiting the optimal truncation property of the KL expansion resulting in a smaller computational load on the estimation algorithm. The authors then consider the stochastic Cramér-Rao bound and derive a closed-form expression for the random KL coefficients and consequently exploit the performance of the MMSE channel estimator based on the evaluation of minimum Bayesian MSE. The effect of a modeling mismatch on the estimator performance is also analyzed.

The last six papers are concerned with the applications of general signal processing techniques on channel equalization, blind multiuser detection, direction-of-arrival estimation, and wideband CDMA systems. In the invited paper by H. Li and Poor, “Impact of channel estimation errors on multiuser detection via the replica method”, system performance is obtained in the large system limit for optimal MUD, linear MUD, and turbo MUD, and is validated by numerical results for finite systems. The paper by Lu et al., entitled “Factor-graph-based soft self-iterative equalizer for multipath channels,” considers factor-graph-based soft self-iterative equalization in wireless multipath channels. The performance of the considered self-iterative equalizer is analyzed in both single-antenna and multiple-antenna multipath channels. It is concluded that when factor graphs of multipath channels have no cycles or mild cycle conditions, the considered self-iterative equalizer can converge to optimum performance after a few iterations; but it may suffer local convergence in channels with severe cycle conditions. In the third paper, “Estimation of directions of arrival by matching pursuit (EDAMP)” by Karabulut et al., a novel system architecture is proposed that employs a matching pursuit-based basis selection algorithm for directions-of-arrival estimation. The proposed system does not require a priori knowledge of the number of angles to be resolved and uses a very small number of snapshots for convergence. The performance of the algorithm is not affected by correlation in the input signals. The algorithm is compared with well-known directions-of-arrival estimation methods with different branch-SNR levels, correlation levels, and

different angle-of-arrival separations. The fourth paper by T. Li et al., “Blind multiuser detection for long-code CDMA systems with transmission-induced cyclostationarity,” considers blind channel identification and signal separation in long-code CDMA systems. A long-code CDMA system is characterized using a time-invariant system model. Then a multistep linear prediction method is used to reduce the intersymbol interference introduced by multipath propagation, and channel estimation then follows by utilizing the nonconstant modulus precoding technique with or without the matrix-pencil approach. After channel estimation, equalization is carried out using a cyclic Wiener filter. Finally, since chip-level equalization is performed, the proposed approach can readily be extended to multirate cases, either with multicode or variable spreading factor. The fifth paper, “Adaptive space-time-spreading-assisted wideband CDMA systems communicating over dispersive Nakagami- m fading channels” by L.-L. Yang and Hanzo, investigates the performance of wideband code-division multiple-access (W-CDMA) systems using space-time-spreading (STS)-based transmit diversity, when frequency-selective Nakagami- m fading channels, multiuser interference, and background noise are considered. The analysis and numerical results suggest that the achievable diversity order is the product of the frequency-selective diversity order and the transmit diversity order. Furthermore, both the transmit diversity and the frequency-selective diversity have the same order of importance. Taking several facts into account, an adaptive STS-based transmission scheme is then proposed for improving the throughput of W-CDMA systems. The numerical results demonstrate that this adaptive STS-based transmission scheme is capable of significantly improving the effective throughput and the bit rate of W-CDMA systems. The last paper, “Opportunistic carrier sensing for energy-efficient information retrieval in sensor networks,” is an invited paper by Zhao and Tong which is concerned with sensor networks. The authors consider distributed information retrieval for sensor networks with cluster heads or mobile access points. A distributed opportunistic transmission protocol is proposed using a combination of carrier sensing and backoff strategy that incorporates channel state information of individual sensors.

Erdal Panayırçı
Costas Georghiadis
Xiaodong Wang
Hakan A. Çırpan

Erdal Panayırçı received the Diploma Engineering degree in electrical engineering from Istanbul Technical University, Istanbul, Turkey, in 1964, and the Ph.D. degree in electrical engineering and system science from Michigan State University, USA, in 1970. Between 1970–2000, he was with the Faculty of Electrical and Electronics Engineering, Istanbul Technical University, where he was a Professor and Head of the Telecommunications Chair. Currently, he is a Professor and

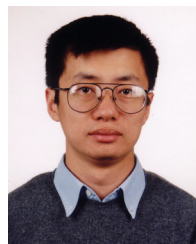


Head of the Electronics Engineering Department, IŞIK University, Istanbul, Turkey. He is engaged in research and teaching in digital communications and wireless systems, equalization and channel estimation in multicarrier (OFDM) communication systems, and efficient modulation and coding techniques (TCM and turbo coding). He spent two years (1979–1981) with the Department of Computer Science, Michigan State University, as a Fulbright-Hays Fellow and a NATO Senior Scientist. From August 1990 to December 1991, he was with the Center for Communications and Signal Processing, New Jersey Institute of Technology, as a Visiting Professor, and took part in the research project on interference cancellation by array processing. Between 1998–2000, he was a Visiting Professor at the Department of Electrical Engineering, Texas A&M University, and took part in research on developing efficient synchronization algorithms for OFDM systems. Between 1995–1999, Professor Panayircı was an Editor for the IEEE Transactions on Communications in the fields of synchronization and equalizations. He is currently a Director of the Network of Excellence on Wireless Communications (NEWCOM) established by the European Commission 6th Framework Programme in March 2002, representing IŞIK University. He is a Fellow of the IEEE.

Costas Georghiades received the B.E. degree with distinction from the American University of Beirut in June 1980, and the M.S. and D.S. degrees from Washington University in May 1983 and May 1985, respectively, all in electrical engineering. Since September 1985, he has been with the Electrical Engineering Department, Texas A&M University, where he is a Professor and holder of the Delbert A. Whitaker Endowed Chair. His general interests are in the application of information, communication, and estimation theories to the study of communication systems. Dr. Georghiades is a Fellow of the IEEE and a registered Professional Engineer in Texas. Over the years, he served in editorial positions with the IEEE Transactions on Communications, the IEEE Transactions on Information Theory, the IEEE Journal on Selected Areas in Communications, and the IEEE Communications Letters. He has been involved in organizing a number of conferences; recently he has been the General Cochair for the 2004 IEEE Information Theory Workshop. He currently serves as the Chair of the Communication Theory Technical Committee, as the Technical Program Cochair for the 2005 IEEE Communication Theory Workshop, and as the Chair of the 2005 SPIE Noise in Communication Systems Conference.

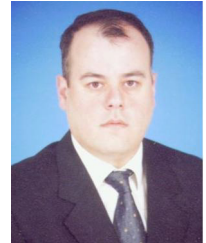


Xiaodong Wang received the B.S. degree in electrical engineering and applied mathematics (with the highest honors) from Shanghai Jiao Tong University, Shanghai, China, in 1992; the M.S. degree in electrical and computer engineering from Purdue University in 1995; and the Ph.D. degree in electrical engineering from Princeton University in 1998. From July 1998 to December 2001, he was on the faculty of the Department of Electrical Engineering, Texas A&M University. In January 2002, he joined the faculty of the Department of Electrical Engineering, Columbia University. Dr. Wang's research interests fall in the general areas of computing, signal processing, and communications. Among his publications is a recent book entitled *Wireless Communication Systems: Advanced Techniques for Signal Reception*. Dr. Wang received the 1999 NSF CAREER Award, and the 2001



IEEE Communications Society and Information Theory Society Joint Paper Award. He currently serves as an Associate Editor for the IEEE Transactions on Communications, the IEEE Transactions on Wireless Communications, the IEEE Transactions on Signal Processing, and the IEEE Transactions on Information Theory.

Hakan A. Çırpan received the B.S. degree in 1989 from Uludag University, Bursa, Turkey, the M.S. degree in 1992 from Istanbul University, Istanbul, Turkey, and the Ph.D. degree in 1997 from the Stevens Institute of Technology, Hoboken, NJ, USA, all in electrical engineering. From 1995 to 1997, he was a Research Assistant at the Stevens Institute of Technology, working on signal processing algorithms for wireless communication systems. In 1997, he joined the faculty of the Department of Electrical-Electronics Engineering, Istanbul University. His general research interests cover wireless communications, statistical signal and array processing, system identification, and estimation theory. His current research activities are focused on signal processing and communication concepts with specific attention to channel estimation and equalization algorithms for space-time coding and multicarrier (OFDM) systems. Dr. Çırpan received the Peskin Award from Stevens Institute of Technology as well as the Professor Nazim Terzioğlu Award from the Research Fund of Istanbul University. He is a Member of IEEE and Sigma Xi.



A Receiver for Differential Space-Time $\pi/2$ -Shifted BPSK Modulation Based on Scalar-MSDD and the EM Algorithm

Michael L. B. Riediger

School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6
Email: mlriedig@sfu.ca

Paul K. M. Ho

School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6
Email: paulho@cs.sfu.ca

Jae H. Kim

School of Mechatronics Engineering, Changwon National University, Changwon, Kyungnam 641-773, Korea
Email: hyung@changwon.ac.kr

Received 21 April 2004; Revised 10 September 2004

In this paper, we consider the issue of blind detection of Alamouti-type differential *space-time* (ST) modulation in static Rayleigh fading channels. We focus our attention on a $\pi/2$ -shifted BPSK constellation, introducing a novel transformation to the received signal such that this binary ST modulation, which has a second-order transmit diversity, is equivalent to QPSK modulation with second-order receive diversity. This equivalent representation allows us to apply a low-complexity detection technique specifically designed for receive diversity, namely, scalar *multiple-symbol differential detection* (MSDD). To further increase receiver performance, we apply an iterative *expectation-maximization* (EM) algorithm which performs joint channel estimation and sequence detection. This algorithm uses minimum mean square estimation to obtain channel estimates and the maximum-likelihood principle to detect the transmitted sequence, followed by differential decoding. With receiver complexity proportional to the observation window length, our receiver can achieve the performance of a coherent maximal ratio combining receiver (with differential decoding) in as few as a single EM receiver iteration, provided that the window size of the initial MSDD is sufficiently long. To further demonstrate that the MSDD is a vital part of this receiver setup, we show that an initial ST conventional differential detector would lead to a strange convergence behavior in the EM algorithm.

Keywords and phrases: multiple-symbol differential detection, Alamouti modulation, differential space-time codes, EM algorithm.

1. INTRODUCTION

Differential detection of a *differentially encoded phase-shift keying* (DPSK) signal is a technique commonly used to recover the transmitted data in a communication system, when channel information (on both the amplitude and phase) is absent at the receiver. The performance of DPSK in traditional wireless communication systems employing one transmit antenna and one or more receive antennas is well documented in the literature. In recent years, this encoding-

detection concept has been extended to cover the scenario where there is more than one transmit antenna. This leads to differential *space-time block codes* (STBCs), an extension of the STBCs originally proposed in [1]. Like conventional DPSK, differential STBCs enable us to decode the received signal without knowledge of channel information, provided that the channel remains relatively constant during the observation interval [2, 3, 4, 5, 6]. Another similarity between conventional DPSK and differential STBCs is that both suffer a loss in performance when compared to their respective ideal coherent receiver.

For conventional DPSK, one approach often used to improve receiver performance is to make decisions based on multiple symbols, that is, *multiple-symbol differential*

This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

detection (MSDD). Previous research has demonstrated that when there is only a single channel, that is, only one transmit antenna and one receive antenna, the performance of MSDD can approach that of the ideal coherent detector when N , the observation window length in a number of symbol intervals, is sufficiently large [7, 8]. This observation is true for both the *additive white Gaussian noise* (AWGN) channel and the Rayleigh fading channel. Moreover, the computational complexity of MSDD is only $N \log N$, provided that the channel is constant over the observation window of the detector and that the implementation procedure developed by Mackenthun is employed [9]. For receive-diversity only systems, Simon and Alouini demonstrated again that the performance of an MSDD combiner approaches that of a coherent *maximal ratio combining* (MRC) receiver with differential decoding, when N is sufficiently large [10]. The application of the MSDD concept to detect differentially encoded STBCs has been considered by a number of authors [11, 12, 13, 14, 15]. Their results indicate that *space-time MSDD* (ST-MSDD) can provide substantial performance improvement over the standard *space-time* (ST) differential detector in [2]. Unfortunately, for both the MSDD combiner and the ST-MSDD, there is no known efficient algorithm for the optimal implementation of these receivers. The complexity of both optimal receivers is exponential in N . In this paper, we will use the term *scalar-MSDD* to refer to the optimal MSDD for the single channel case [7, 9], and the term *vector-MSDD* to refer to either an MSDD combiner [10] or an ST-MSDD [11].

In light of the exponential complexity of the optimal vector-MSDD, several suboptimal, reduced-complexity variants have been proposed for detecting differential STBC. For example, Lampe et al. implemented a code-dependent technique with a complexity that is essentially independent of the observation window length of the detector [12, 13]. The concept of decision feedback was employed by Schober and Lampe in their MSDD for a system employing both transmit and receive diversity [6]. Similar ideas were also employed by Tarasak and Bhargava in a transmit-diversity only scenario [14], and by Lao and Haimovich in an interference suppression and receive-diversity setting [15]. In addition, Tarasak and Bhargava investigated reducing receiver complexity using a reduced search detection approach [14].

In this paper, we propose an iterative receiver for differential STBC employing a $\pi/2$ -shifted BPSK constellation, two transmit antennas, and an Alamouti-type code structure [16]. By employing a novel transformation to the received signal, it is shown that this STBC is equivalent to conventional differential QPSK modulation with second-order receive diversity. As a result, selection diversity and scalar-MSDD can be employed in the first pass of our iterative receiver. Due to the low complexity of the scalar-MSDD, a very large window size N (i.e., 64) can be employed to provide the receiver with very accurate initial estimates of the transmitted symbols. Successive iterations of the receiver operations are then based on the *expectation-maximization* (EM) algorithm [17] for joint channel estimation and sequence detection. Our results show that the iterative receiver we introduce can essentially achieve the performance of the ideal coherent

MRC receiver, with differential encoding, in as few as a single EM iteration (i.e., a total of two passes).

This paper is organized as follows. Section 2 presents the STBC adopted in this investigation, the channel model, and the transformation employed to convert this second-order transmit-diversity system into an equivalent second-order receive-diversity system. Details of the receiver operations, including that of the EM algorithm, which performs joint channel estimation and sequence detection, are described in Section 3. The bit error performance of the proposed receiver is given in Section 4, while conclusions of this investigation are made in Section 5.

2. DIFFERENTIAL ST $\pi/2$ -SHIFTED BPSK AND EQUIVALENT RECEIVE DIVERSITY

2.1. System model

We consider a wireless communications system operating over a slow, flat Rayleigh fading channel, in which space-time block-coded symbols are sent from two transmit antennas and received by a single receive antenna. The space-time block code employed falls into the class of the popular two-branch transmission-diversity scheme introduced by Alamouti [16]. Specifically, if $c_1[k]$ and $c_2[k]$ are, respectively, the complex symbols transmitted by the first and second antennas, in the first subinterval of the k th coded interval, then the transmitted symbols in the second subinterval by the same two antennas are, respectively, $-c_2^*[k]$ and $c_1^*[k]$. Note that throughout this paper, the notations $(\cdot)^*$ and $(\cdot)^\dagger$ are used to represent the complex conjugate of a complex number and the conjugate (Hermitian) transpose of a complex vector/matrix. The various coded symbols are taken from the $\pi/2$ -shifted BPSK constellation $S = \{+1, -1, +j, -j\}$, where the subsets $S_1 = \{+1, -1\}$ and $S_2 = \{+j, -j\}$ are used alternately in successive subintervals at each transmit antenna. This alternation between S_1 and S_2 not only reduces envelope fluctuation, but it also enables us to transform the proposed second-order transmit-diversity BPSK system into an equivalent second-order receive-diversity QPSK system. Assuming that $c_1[k]$ is chosen from S_1 , it follows that $c_2[k]$ must be chosen from S_2 . Then, the transmitted code matrix in the k th coded interval becomes

$$\mathbf{C}[k] = \begin{bmatrix} c_1[k] & c_2[k] \\ -c_2^*[k] & c_1^*[k] \end{bmatrix} = \begin{bmatrix} c_1[k] & c_2[k] \\ c_2[k] & c_1[k] \end{bmatrix}, \quad (1)$$

where $\mathbf{C}[k]$ is a member of the set $V = \{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4\}$, with

$$\begin{aligned} \mathbf{V}_1 &= \begin{bmatrix} 1 & j \\ j & 1 \end{bmatrix}, & \mathbf{V}_2 &= \begin{bmatrix} 1 & -j \\ -j & 1 \end{bmatrix}, \\ \mathbf{V}_3 &= \begin{bmatrix} -1 & -j \\ -j & -1 \end{bmatrix}, & \mathbf{V}_4 &= \begin{bmatrix} -1 & j \\ j & -1 \end{bmatrix}. \end{aligned} \quad (2)$$

Note that the columns of $\mathbf{C}[k]$ correspond to the two transmit antennas, while the rows of $\mathbf{C}[k]$ correspond to the coded subintervals.

TABLE 1: Logic table showing the ST differential encoding rule for $\mathbf{C}[k]$, given $\mathbf{C}[k-1]$ and $\mathbf{D}[k]$.

$\mathbf{C}[k-1]$	$\mathbf{D}[k]$			
	\mathbf{U}_1	\mathbf{U}_2	\mathbf{U}_3	\mathbf{U}_4
\mathbf{V}_1	\mathbf{V}_1	\mathbf{V}_2	\mathbf{V}_3	\mathbf{V}_4
\mathbf{V}_2	\mathbf{V}_2	\mathbf{V}_3	\mathbf{V}_4	\mathbf{V}_1
\mathbf{V}_3	\mathbf{V}_3	\mathbf{V}_4	\mathbf{V}_1	\mathbf{V}_2
\mathbf{V}_4	\mathbf{V}_4	\mathbf{V}_1	\mathbf{V}_2	\mathbf{V}_3

Since we will be using MSDD in the first pass of our iterative receiver, it is necessary for the $\mathbf{C}[k]$'s to be differentially encoded ST symbols. The $\mathbf{C}[k]$'s are related to the actual data symbols, the $\mathbf{D}[k]$'s, according to

$$\mathbf{C}[k] = \mathbf{D}[k]\mathbf{C}[k-1], \quad (3)$$

where $\mathbf{D}[k]$ is from the set $U = \{\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \mathbf{U}_4\}$, with

$$\begin{aligned} \mathbf{U}_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & \mathbf{U}_2 &= \begin{bmatrix} 0 & -j \\ -j & 0 \end{bmatrix}, \\ \mathbf{U}_3 &= \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, & \mathbf{U}_4 &= \begin{bmatrix} 0 & j \\ j & 0 \end{bmatrix}. \end{aligned} \quad (4)$$

Without loss of generality, the initial transmitted symbol $\mathbf{C}[0]$, which carries no information and serves only as an initialized reference, is chosen to be \mathbf{V}_1 . It can be easily verified that the \mathbf{U}_n 's are unitary matrices, and that for any \mathbf{V}_m in set V and any \mathbf{U}_n in the set U , the product $\mathbf{U}_n\mathbf{V}_m$ is a member of the set V . The relations between $\mathbf{C}[k-1]$, $\mathbf{D}[k]$, and $\mathbf{C}[k]$, which arise from the differential encoding rule, are explicitly depicted in Table 1.

The transmitted symbols at each transmit antenna will be pulse-shaped by a *square-root raised cosine* (SQRC) pulse, and then transmitted over a wireless link to the receiver. Each link introduces fading to the associated transmitted signal, and the receiver's front end introduces AWGN. The composite received signal from the two links is matched-filtered and sampled, twice per encoded interval, to provide the receiver with sufficient statistics to detect the transmitted data. Assuming the channel gains in the two links, f_1 and f_2 , are constant within the observation window of the data detector, the two received samples in the k th interval can be modeled as

$$\mathbf{R}[k] = [r_1[k], r_2[k]]^T = \mathbf{C}[k]\mathbf{F} + \mathbf{N}[k], \quad (5)$$

where

$$\mathbf{F} = [f_1, f_2]^T \quad (6)$$

is the vector of complex channel gains,

$$\mathbf{N}[k] = [n_1[k], n_2[k]]^T \quad (7)$$

is a noise vector containing the two complex Gaussian noise terms $n_1[k]$ and $n_2[k]$, and $(\cdot)^T$ denotes the transpose of a matrix. The channel fading gains are assumed to be *independent and identically distributed* (i.i.d.) zero-mean complex Gaussian random variables, with unit variance. In addition, these channel gains are assumed to be constant over the observation window of N symbol intervals. The static fading channel has been frequently considered when investigating systems with transmit and receive diversity [10, 18, 19, 20, 21, 22, 23]. On the other hand, the sequence of noise samples, $\{\dots, n_1[k], n_2[k], n_1[k+1], n_2[k+1], \dots\}$, is a complex, zero-mean white Gaussian process, with a variance of N_0 . It should be pointed out that the fading gains and the noise samples are statistically independent.

To recover the data contained in the $\mathbf{R}[k]$'s, the receiver can employ the ST differential detector in [2]. The metric adopted by this simple detector can be expressed in the form $I = |\mathbf{R}^\dagger[k]\tilde{\mathbf{D}}[k]\tilde{\mathbf{C}}[k-1] + \mathbf{R}^\dagger[k-1]\tilde{\mathbf{C}}[k-1]|^2$, where $\tilde{\mathbf{D}}[k] \in U$ represents a hypothesis for the data symbol $\mathbf{D}[k]$, $\tilde{\mathbf{C}}[k-1] \in V$ represents a hypothesis for transmitted symbol $\mathbf{C}[k-1]$, and $|\cdot|$ denotes the magnitude of a complex vector. Since I is actually independent of $\tilde{\mathbf{C}}[k-1]$, the hypothesis on $\mathbf{D}[k]$ that maximizes the metric I is chosen as the most likely transmitted data symbol. Though simple, this detector was shown to exhibit a 3 dB loss in power efficiency when compared to the ideal coherent receiver. To narrow this performance gap, a vector-MSDD can be used instead [11]. This detector organizes the $\mathbf{R}[k]$ s into overlapping blocks of size N , with the last vector in the previous block being the first vector in the current block. For the block starting at time zero, the decoding metric can be expressed in the form $J = |\sum_{k=0}^{N-1} \mathbf{R}^\dagger[k](\prod_{i=1}^k \tilde{\mathbf{D}}[i])\tilde{\mathbf{C}}[0]|^2$. Like the metric I , this vector-MSDD metric is independent of $\tilde{\mathbf{C}}[0]$. Consequently, the detector selects the hypothesis $(\tilde{\mathbf{D}}[1], \tilde{\mathbf{D}}[2], \dots, \tilde{\mathbf{D}}[N-1])$ that maximizes J , as the most likely transmitted pattern in this interval. It is clear from the expression of J that there are altogether 4^{N-1} hypotheses to consider. So far, there does not exist any algorithm that performs this search in an efficient and yet optimal fashion.

The approach we adopt to mitigate the complexity issue in the vector-MSDD is to first transform the received signal vector in (5) into one that we would encounter in a receive-diversity only system. Although the optimal vector-MSDD in this latter case still has an exponential complexity [10], we now have the option of using selection combining in conjunction with a scalar-MSDD [18]. Although there is still a substantial gap between selection combining MSDD and the MRC, this gap can be closed by employing additional processing based on the iterative EM algorithm described in the next section. In this case, the decisions made by the selection combining MSDD are used to initialize the EM processing unit. The following subsection provides details about the transformation required to turn our second-order transmit-diversity system into an equivalent second-order receive-diversity system.

TABLE 2: Logic table showing the equivalent QPSK differential encoding rule for $b[k]$, given $b[k-1]$ and $\mathbf{D}[k]$.

$\mathbf{C}[k-1]$ $b[k-1]$	$\mathbf{D}[k], a[k]$			
	$\mathbf{U}_1, y_1 = 1$	$\mathbf{U}_2, y_2 = -j$	$\mathbf{U}_3, y_3 = -1$	$\mathbf{U}_4, y_4 = j$
$\mathbf{V}_1, x_1 = 1 + j$	x_1	x_2	x_3	x_4
$\mathbf{V}_2, x_2 = 1 - j$	x_2	x_3	x_4	x_1
$\mathbf{V}_3, x_3 = -1 - j$	x_3	x_4	x_1	x_2
$\mathbf{V}_4, x_4 = -1 + j$	x_4	x_1	x_2	x_3

2.2. From transmit diversity to receive diversity

To assist in the development of transformation, we first expand (5) to obtain

$$\begin{aligned} r_1[k] &= f_1 c_1[k] + f_2 c_2[k] + n_1[k], \\ r_2[k] &= f_1 c_2[k] + f_2 c_1[k] + n_2[k]. \end{aligned} \quad (8)$$

This equation clearly illustrates the structure of the received signal samples. Moreover, we can deduce from the equation that the average SNR in the received sample $r_1[k]$ is

$$\gamma = \frac{(1/2)E\{|f_1 c_1[k] + f_2 c_2[k]|^2\}}{(1/2)E\{|n_1[k]|^2\}} = \frac{2}{N_0}, \quad (9)$$

where $E\{\cdot\}$ is the expectation operator. The same SNR also appears in the received sample $r_2[k]$.

Next, we introduce the new variables

$$\begin{aligned} p_1[k] &= r_1[k] + r_2[k] = g_1 b[k] + w_1[k], \\ p_2[k] &= r_1^*[k] - r_2^*[k] = g_2 b[k] + w_2[k], \end{aligned} \quad (10)$$

where

$$g_1 \equiv f_1 + f_2, \quad g_2 \equiv f_1^* - f_2^* \quad (11)$$

are two new fading gains,

$$b[k] \equiv c_1[k] + c_2[k] \quad (12)$$

is an equivalent transmitted symbol, and

$$\begin{aligned} w_1[k] &\equiv n_1[k] + n_2[k], \\ w_2[k] &\equiv n_1^*[k] - n_2^*[k] \end{aligned} \quad (13)$$

are two new noise terms. It can be shown that the new fading gains g_1 and g_2 are independent Gaussian random variables, with a variance of 2. Similarly, it can also be shown that the new noise samples $w_1[k]$ and $w_2[k]$ are independent and have variance $2N_0$. These results mean that the SNR in the samples $p_1[k]$ and $p_2[k]$ is also γ , in other words, the original SNR is preserved. Of foremost interest, note the new symbol $b[k]$ is shared by $p_1[k]$ and $p_2[k]$. Consequently, (10) corresponds to the received signal encountered in a second-order receive-diversity system. Furthermore, $b[k]$ belongs to the QPSK signal set $X = \{x_1, x_2, x_3, x_4\}$, where

$$\begin{aligned} x_1 &= 1 + j, & x_2 &= 1 - j, \\ x_3 &= -1 - j, & x_4 &= -1 + j. \end{aligned} \quad (14)$$

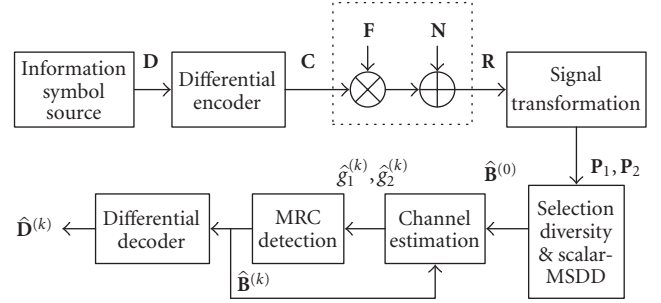


FIGURE 1: Block diagram of transmitter, channel model, and EM-based receiver performing joint channel estimation and sequence detection. Note that the matrix multiplication and addition operations are indexed by time.

In comparing (2) with (14), we can quickly see that x_i is simply the row (or column) sum of \mathbf{V}_i . Furthermore, for all $\mathbf{V}_n = \mathbf{U}_m \mathbf{V}_k$, $x_n = y_m x_k$, where y_m is the row (or column) sum of the unitary matrix \mathbf{U}_m in (4). This latter property implies that differential encoding of ST $\pi/2$ -shifted BPSK symbols is equivalent to differential encoding of scalar QPSK symbols. The respective QPSK encoding rule is $b[k] = a[k]b[k-1]$, where $a[k] \in \{1, j, -1, -j\}$ is the equivalent data symbol and $b[k] \in \{\pm 1 \pm j\}$ is the equivalent transmitted symbol. Note that x_n , the row/column sum of \mathbf{V}_n , can be expressed as $x_n = \mathbf{1}_2 \mathbf{V}_n \mathbf{1}_2^T / 2$ or as $x_n = \mathbf{1}_2 \mathbf{U}_m \mathbf{V}_k \mathbf{1}_2^T / 2$, where $\mathbf{1}_2 = [1, 1]$ is an all-one row vector of length two. However, we can also deduce that $\mathbf{1}_2 \mathbf{U}_m = y_m \mathbf{1}_2$ and $\mathbf{V}_k \mathbf{1}_2^T = x_k \mathbf{1}_2^T$, implying that $\mathbf{1}_2 \mathbf{U}_m \mathbf{V}_k \mathbf{1}_2^T / 2 = y_m x_k$. Table 2 shows this equivalent differential encoding rule. By comparing Table 1 and Table 2, it is evident that the indexings of the respective symbols are identical. The advantage of transforming the original STBC into an equivalent second-order receive-diversity QPSK system will be clearly demonstrated in the next section.

3. THE MSDD-AIDED EM-BASED ITERATIVE RECEIVER

The previous section demonstrated how an STBC $\pi/2$ -shifted BPSK system can be transformed into an equivalent receive-diversity system. This section describes how an iterative receiver based on selection diversity, scalar-MSDD, and the EM algorithm [17] processes the equivalent received signal and attains the equivalent performance to that of an ideal coherent receiver (with differential decoding). Figure 1 provides a quick overview of this proposed receiver.

3.1. First pass—selection diversity and scalar-MSDD

Given the new received variables in (10), we can use, in principle, an MSDD combiner [10] to detect the transmitted data. The decoding metric of this receiver is of the form

$$K = |\mathbf{P}_1^\dagger \tilde{\mathbf{B}}|^2 + |\mathbf{P}_2^\dagger \tilde{\mathbf{B}}|^2, \quad (15)$$

where

$$\mathbf{P}_i = [p_i[0], p_i[1], \dots, p_i[N-1]]^T = g_i \mathbf{B} + \mathbf{W}_i, \quad i = 1, 2, \quad (16)$$

are the equivalent received vectors, N is the window width of the MSDD combiner,

$$\mathbf{B} = [b[0], b[1], \dots, b[N-1]]^T \quad (17)$$

is the equivalent transmitted pattern,

$$\mathbf{W}_i = [w_i[0], w_i[1], \dots, w_i[N-1]]^T, \quad i = 1, 2, \quad (18)$$

are the equivalent noise patterns, and $\tilde{\mathbf{B}}$ represents a hypothesis of \mathbf{B} . The MSDD combiner searches through all possible hypotheses; the hypothesis which maximizes K is declared the most likely transmitted pattern. This most likely hypothesis is then differentially decoded to obtain the data symbols. This operation therefore makes the decision independent of the first symbol in $\tilde{\mathbf{B}}$. Consequently, we can simply assume all hypotheses start with the symbol x_1 in (14). Thus, as with the case of the vector-MSDD, there are 4^{N-1} candidates to consider. This exponential complexity prevents the use of a large N in (15). However, for suboptimal implementation, we can use selection diversity followed by scalar-MSDD [18], an option which is unavailable in vector-MSDD. It will be shown in the next section that an EM-based iterative receiver initiated by selection diversity scalar-MSDD has better performance and convergence properties than those initiated by conventional *space-time differential detection* (ST-DD).

A selection-diversity scalar-MSDD receiver obtains an estimate of the equivalent transmitted pattern \mathbf{B} according to

$$\hat{\mathbf{B}}^{(0)} = \arg \max_{\mathbf{B} \in \mathbb{B}} |\mathbf{Z}^\dagger \tilde{\mathbf{B}}|^2, \quad (19)$$

where \mathbb{B} is the collection of all possible length- N equivalent QPSK sequences, and

$$\mathbf{Z} = \begin{cases} \mathbf{P}_1, & |\mathbf{P}_1|^2 > |\mathbf{P}_2|^2, \\ \mathbf{P}_2, & \text{otherwise.} \end{cases} \quad (20)$$

The solution to (19) is easily found using the algorithm developed by Mackenthun [9], as the channel is constant over the observation interval. It is important to stress that this algorithm has a complexity of only $N \log N$.

The decision $\hat{\mathbf{B}}^{(0)}$ in (19) is used to initialize the EM algorithm described in the next section. This algorithm performs iterative channel estimation and data detection, by passing information back and forth between the channel estimator and the data detector. At this point, we want to point out that

other options for initializing the EM algorithm include using pilot symbols to acquire a channel fading estimate [19, 20], or using differential detection to acquire a transmitted signal estimate [21]. Although using pilot symbols provides a reliable reference to estimate the channel gains, it results in a power loss, and even after several iterations, the performance of coherent detection may not be reached [19, 20]. In the case of initializing the EM algorithm with differentially detected sequence [21], it was determined that the transmitted sequence estimate reconstructed from a vector-MSDD information sequence estimate does not yield good channel estimates due to differential reencoding. Hence, there was a consistent performance loss when compared to a coherent receiver.

3.2. Successive passes—joint estimation and detection using the EM algorithm

It was shown in [18] that with a large N (i.e., 64), the selection-diversity scalar-MSDD receiver, described in Section 3.1, experiences a 1.5 dB degradation in power efficiency when compared to MRC. To narrow this performance gap, we propose to adopt the EM algorithm to further process the initial estimate $\hat{\mathbf{B}}^{(0)}$ provided by the selection-diversity scalar-MSDD receiver.

The EM algorithm was first introduced by Dempster et al. [17]. It is suited for problems where there are random variables other than a desired component contributing to the observable data. The *complete set* of data consists of the *desired* data and the *nuisance* data. In the context of the problem at hand, the complete set of data is the (equivalent) transmitted pattern \mathbf{B} and the channel gains g_1 and g_2 ; the sequence \mathbf{B} is the desired data, and the channel gains are the nuisance parameters. To initialize the EM algorithm, it is necessary to provide an estimate of either component of the complete set. In our case, this will be the decision $\hat{\mathbf{B}}^{(0)}$ in (19). The accuracy of this initial estimate often determines the effectiveness of the EM algorithm and the average number of iterations necessary for convergence. An excellent description of the algorithm and the breadth of its applications can be found in [24]. A detailed application of the EM algorithm to joint channel estimation and sequence detection situations can be found in [25]. The scope of the description given below is restricted to our joint channel estimation and sequence detection problem.

The EM algorithm consists of two steps per iteration; an *expectation step* (E-step) and a *maximization step* (M-step). At the k th E-step, the algorithm estimates the fading gains by computing their means when conditioned on the received data \mathbf{P}_1 and \mathbf{P}_2 , and the most recent estimate $\hat{\mathbf{B}}^{(k-1)}$ of the equivalent QPSK symbols. Using the *minimum mean square estimation* (MMSE) principle, these conditional means can be expressed as [19, 20]

$$\begin{aligned} \hat{g}_i^{(k)} &= E\{g_i | \mathbf{P}_1, \mathbf{P}_2, \hat{\mathbf{B}}^{(k-1)}\} \\ &= E\{g_i | \mathbf{P}_i, \hat{\mathbf{B}}^{(k-1)}\} \\ &= \frac{1}{N + 1/\gamma} (\hat{\mathbf{B}}^{(k-1)})^\dagger \mathbf{P}_i, \quad i = 1, 2. \end{aligned} \quad (21)$$

Immediately following the k th E-step is the k th M-step. Here the algorithm assumes the fading gain estimates in (21) are perfect and performs MRC and data detection according to

$$\hat{\mathbf{B}}^{(k)} = \arg \max_{\tilde{\mathbf{B}} \in \mathbb{B}} \operatorname{Re} \{ (\hat{g}_1^{(k)} \mathbf{P}_1^\dagger + \hat{g}_2^{(k)} \mathbf{P}_2^\dagger) \tilde{\mathbf{B}} \}, \quad (22)$$

where $\operatorname{Re}\{\cdot\}$ is the real operator. In other words, the M-step updates the decision on \mathbf{B} according to the most recent estimates of the fading gains. It should be pointed out that (22) can easily be solved on a symbol-by-symbol basis. Furthermore, the estimated symbols in $\hat{\mathbf{B}}^{(k)}$ are then differentially decoded to obtain estimates of the information symbols. If it is desired to perform another EM iteration, the channel will be reestimated using (21), and hence another sequence estimate will be obtained using (22). The iterations cease when the sequence estimate does not change during two subsequent iterations, or after a prespecified number of iterations have occurred. A maximum of 10 iterations are considered in this research.

As the E-step is essentially an average of N variables, and the M-step maps each derotated statistic to the nearest QPSK signal, the complexity of each iteration is linearly proportional to N . We note that while it is possible to implement conventional ST-DD to initialize the EM algorithm, our results in the next section show that it is not an effective option.

4. RESULTS

This section details the results obtained via simulation of our system. MSDD of length $N = 16, 32, 64$, and 128 are considered. The results are shown in Figures 2, 3, 4, and 5, along with the performance of conventional ST-DD, equivalent to conventional *equal gain combining* (EGC), and the coherent detection lower bound (i.e., MRC with differential encoding). In these figures, the integer n in the notation EM- n refers to the number of EM iterations. When $n = 0$, we simply have a selection-diversity scalar-MSDD receiver. Note that SNR denotes the average signal-to-noise ratio per bit. Lastly, we remind the reader that simulations were performed using a complex Gaussian, static fading channel, as outlined in Section 2.1.

The results in Figures 2, 3, 4, and 5 indicate that there is a significant improvement in performance from the initial selection-diversity sequence estimate, to the first estimate provided by the EM algorithm. Although they are not included, it should be known that the performance curves of the EM-2 to EM-9 receivers lie consecutively within the curves for the EM-1 and EM-10 receivers. For N equal to 128, the first iteration of the EM receiver essentially meets the lower bound given by coherent reception. Further simulation results not included here indicate that the EM receiver is able to meet the lower bound within a single EM iteration, for all N greater than 128.

The authors stress that the success of this receiver depends strongly on the initial sequence estimate provided by (19), which in turn provides an excellent channel estimate using (21). To elaborate, note in Figures 2, 3, 4, and 5 that the performance of the conventional differential detector is comparable to that of the standard selection-diversity receiver. One might suppose an EM-based receiver using an initial conventional ST-DD sequence estimate (obtained without using selection diversity or MSDD) could yield the same performance results as those shown here; however, this is not the case. The performance curves for an EM-based receiver initialized using a conventional ST-DD sequence estimate are shown in Figure 6. Clearly, the performance of the first iteration is substantially inferior to that of the conventional ST-DD initialization. In this case, the observation window for the conventional detector is only 2 symbol intervals, and the frame length from which the channel estimates are constructed is much larger (i.e., 64 symbol intervals). The inferior performance can be explained by noting that the transmitted sequence must be regenerated before the channel estimates are made. Due to the differential encoding, a single information symbol error may result in a significant number of incorrect transmitted symbol errors and hence a poor transmitted sequence estimate [21]. As the number of iterations increases, the performance improves, however it takes many iterations to approach that of a coherent receiver, and there is still a 0.25 dB performance gap after 10 iterations. This explains why using a conventionally detected sequence as an initialization to the EM-based receiver does not yield such good results. When the selection-diversity MSDD sequence estimate is used as an initialization to the EM-based receiver, the sequence decision rule is based on the entire received sequence, and received statistics are derotated together in an optimal fashion (19). Hence, propagated errors in the regenerated transmitted sequence do not occur.

An assumption we have made is that the channel is constant (static) over N symbol intervals. In the more general situation of a time-varying channel, the methodology proposed here can still be considered, with minor modification to the receiver structure. Firstly, the appropriate, straightforward adjustments must be made to the channel estimation (21) and MRC detection (22) units in the iterative section of the receiver. Secondly, as the Mackenthun algorithm can only be applied to static channels, the scalar-MSDD component would need to be replaced. An appropriate replacement would be a low-complexity, suboptimal MSDD, suited for a time-varying channel [26, 27]. Compared to the optimal MSDD for time-varying channels in [8], these suboptimal detectors have much lower computational complexity. Although there is a small SNR penalty (in the neighborhood of 1 to 2 dB), these detectors exhibit no irreducible error floor, even when the fading rate is as high as a few percent of the symbol rate. Consequently, the initial sequence decision provided by these detectors will be of reasonable quality, and we expect good convergence properties in subsequent EM iterations, similar to that seen in the static fading case.

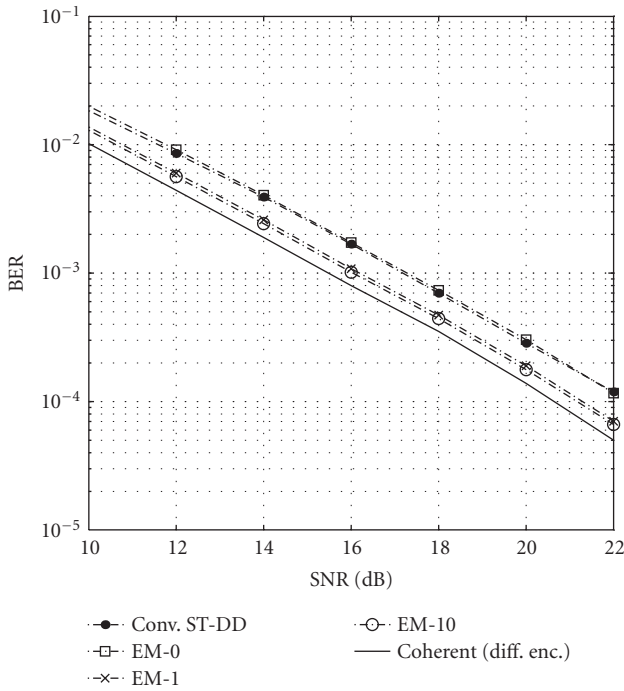


FIGURE 2: BER comparison (conventional ST-DD, selection-diversity EM-based receiver, MRC); $N = 16$.

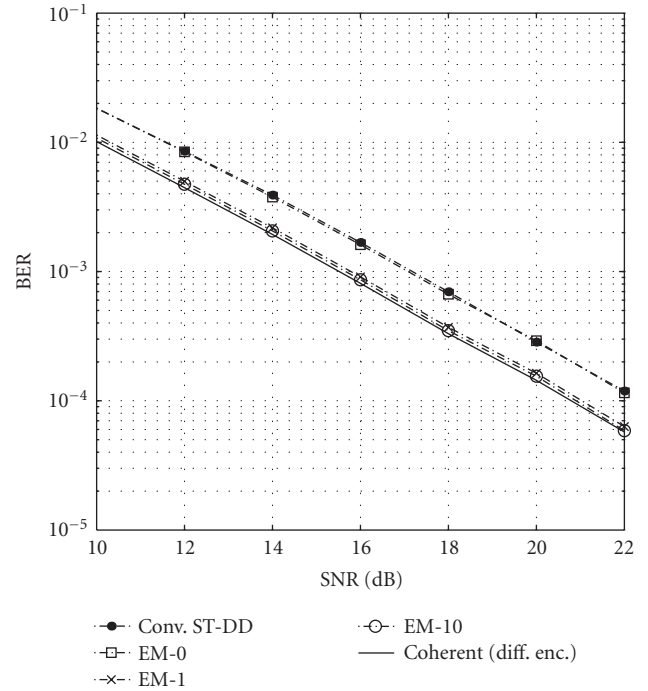


FIGURE 4: BER comparison (conventional ST-DD, selection-diversity EM-based receiver, MRC); $N = 64$.

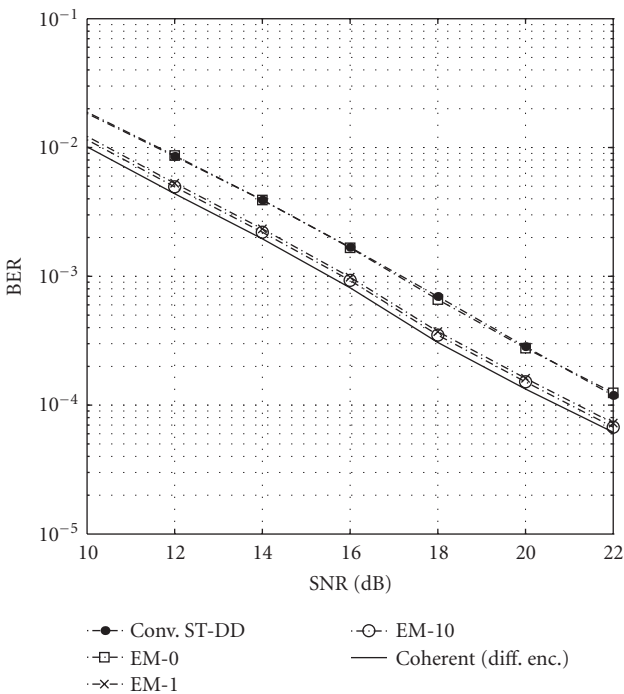


FIGURE 3: BER comparison (conventional ST-DD, selection-diversity EM-based receiver, MRC); $N = 32$.

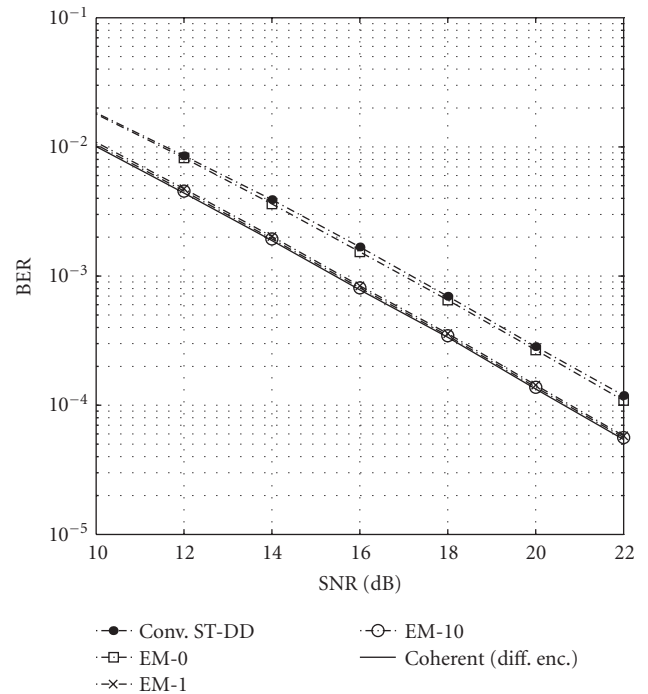


FIGURE 5: BER comparison (conventional ST-DD, selection-diversity EM-based receiver, MRC); $N = 128$.

Finally, we would like to draw some qualitative comparisons between the proposed iterative receiver and those based on pilot symbols [19, 20]. From a bandwidth efficiency

point of view, our pilotless (noncoherent) receiver is more attractive as there is no need to transmit any pilot symbols for channel sounding purposes. Although the gain in

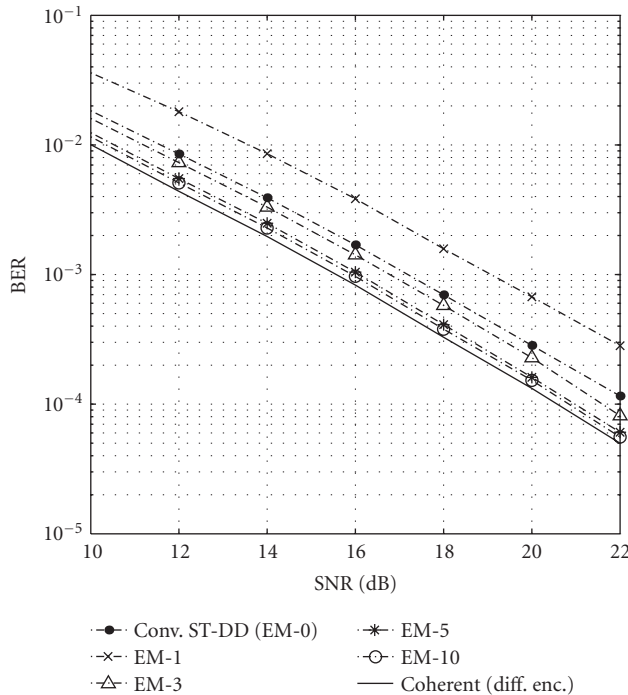


FIGURE 6: BER comparison (EM-based receiver initialized with conventional ST-DD, MRC); frame length of 64 ST symbols.

bandwidth efficiency is minimal for the static fading environment, it can be significant for a time-varying channel. As mentioned in the previous paragraph, the proposed receiver methodology can also be used in a fast fading environment, provided that a suitable MSDD replaces the Mackenthun MSDD. From a power efficiency point of view, we believe our noncoherent receiver and a pilot-aided receiver [19] will have similar performance in the steady state (i.e., after a sufficient number of iterations). We notice a performance gap, in the neighborhood of 1.5 dB, between the receiver for a coded system in [19] and the respective ideal coherent bound without differential encoding. Conversely, our noncoherent receiver can attain the performance indicated by the coherent bound with differential encoding. Recall that there is a 1.5 dB difference between the two coherent bounds for a second-order diversity system. The last performance measure is the computational complexity. We note that the initial pass of our noncoherent EM receiver requires approximately the same amount of signal processing as a pilot-symbol-based system, and the successive iterations require an identical amount of computational resources. However, it may take many iterations to reach the steady-state performance for a pilot-aided system [19, 20], while the noncoherent EM receiver can meet the coherent detection (with differential encoding) lower bound in a single iteration. Thus it appears that the proposed receiver requires less computation, due to its better convergence behavior arising from block detection.

5. CONCLUSION

In summary, we present a novel transformation on a specific Alamouti-type space-time modulation, and obtain a scalar, receive-diversity equivalent. With this transformation, it is simple to apply low-complexity, high-performance, receive-diversity techniques. The results show that when using the sequence estimate from selection-diversity scalar-MSDD as an initialization to an iterative channel and sequence estimator, it is possible to achieve the performance of coherent detection.

Using STBC-MSDD to obtain the lower-performance bound of coherent detection would require implementing an algorithm with complexity 4^{N-1} , where 4 is the cardinality of the transmission symbol set and N is a large number of transmitted space-time symbols. For the system discussed in this paper, the coherent detection lower bound is achieved using a receiver with complexity of essentially $N \log N$, given by the complexity of the scalar-MSDD [9] used to initialize the EM algorithm. Clearly, the scalar equivalent system using the EM algorithm employed in this paper offers a low-complexity method to achieve the performance of coherent detection.

ACKNOWLEDGMENTS

This research was supported by the Natural Sciences and Engineering Research Council (NSERC) and the Canadian Wireless Telecommunications Association (CWTA). This paper was presented in part at VTC'04 Fall, Los Angeles, California, USA, September 26–29, 2004.

REFERENCES

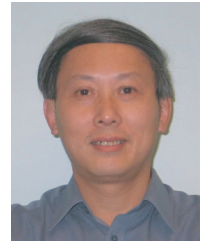
- [1] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.
- [2] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Select. Areas Commun.*, vol. 18, no. 7, pp. 1169–1174, 2000.
- [3] H. Jafarkhani and V. Tarokh, "Multiple transmit antenna differential detection from generalized orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2626–2631, 2001.
- [4] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2567–2578, 2000.
- [5] R. Schober and L. H.-J. Lampe, "Differential modulation diversity," *IEEE Trans. Veh. Technol.*, vol. 51, no. 6, pp. 1431–1444, 2002.
- [6] R. Schober and L. H.-J. Lampe, "Noncoherent receivers for differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 768–777, 2002.
- [7] D. Divsalar and M. K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300–308, 1990.
- [8] P. Ho and D. Fung, "Error performance of multiple-symbol differential detection of PSK signals transmitted over correlated Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 40, no. 10, pp. 1566–1569, 1992.
- [9] K. M. Mackenthun Jr., "A fast algorithm for multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 42, no. 234, pp. 1471–1474, 1994.

- [10] M. K. Simon and M.-S. Alouini, "Multiple symbol differential detection with diversity reception," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1312–1319, 2001.
- [11] C. Gao, A. M. Haimovich, and D. Lao, "Multiple-symbol differential detection for space-time block codes," in *Proc. 36th Annual Conference on Information Sciences and Systems (CISS '02)*, Princeton University, Princeton, NJ, USA, March 2002.
- [12] L. H.-J. Lampe, R. Schober, and R. F. H. Fischer, "Coded differential space-time modulation for flat fading channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 3, pp. 582–590, 2003.
- [13] L. H.-J. Lampe, R. Schober, and R. F. H. Fischer, "Differential space-time modulation—coding and capacity results," in *Proc. IEEE International Conference on Communications (ICC '03)*, vol. 4, pp. 2593–2597, Anchorage, Alaska, USA, May 2003.
- [14] P. Tarasak and V. K. Bhargava, "Reduced complexity multiple symbol differential detection of space-time block code," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC '02)*, vol. 1, pp. 505–509, Orlando, Fla, USA, March 2002.
- [15] D. Lao and A. M. Haimovich, "Multiple-symbol differential detection with interference suppression," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 208–217, 2003.
- [16] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [17] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM-algorithm," *J. Royal Statistical Society: Series B*, vol. 39, no. 1, pp. 1–38, 1977.
- [18] J. H. Kim, P. K. M. Ho, and M. L. B. Riediger, "Suboptimal multiple-symbol differential detection of MPSK with diversity reception," to appear in *IEE Proceedings, Communications*.
- [19] Y. Li, C. N. Georghiades, and G. Huang, "Iterative maximum-likelihood sequence estimation for space-time coded systems," *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 948–951, 2001.
- [20] C. Cozzo and B. L. Hughes, "Joint channel estimation and data detection in space-time communications," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1266–1270, 2003.
- [21] M. L. B. Riediger and P. K. M. Ho, "A differential space-time code receiver using the EM-algorithm," in *Proc. IEEE Canadian Conference on Electrical and Computer Engineering (CCECE '04)*, pp. 185–188, Niagara Falls, Ontario, Canada, May 2004.
- [22] G. Caire and G. Colavolpe, "On low-complexity space-time coding for quasi-static channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 6, pp. 1400–1416, 2003.
- [23] D. M. Ionescu, "On space-time code design," *IEEE Transactions on Wireless Communications*, vol. 2, no. 1, pp. 20–28, 2003.
- [24] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Mag.*, vol. 13, pp. 47–60, November 1996.
- [25] C. N. Georghiades and J. Han, "Sequence estimation in the presence of random parameters via the EM algorithm," *IEEE Trans. Commun.*, vol. 45, no. 3, pp. 300–308, 1997.
- [26] P. Pun and P. Ho, "The performance of Fano-multiple symbol differential detection," to appear in *Proc. IEEE International Conference on Communications (ICC'05)*, Seoul, Korea, May 2005.
- [27] P. Kam and C. Teh, "Reception of PSK signals over fading channels via quadrature amplitude estimation," *IEEE Trans. Commun.*, vol. 31, no. 8, pp. 1024–1027, 1983.

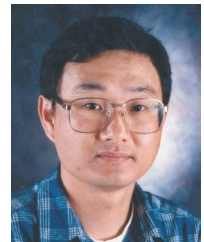
Michael L. B. Riediger was born in Transcona, Manitoba, Canada, in 1978. He received his B.S. degree in computer engineering in 2000 and his M.S. in electrical and computer engineering in 2002, from the University of Manitoba (Winnipeg, Manitoba, Canada). At present, he is a Ph.D. student in the School of Engineering Science, Simon Fraser University, Burnaby, British Columbia, Canada. His current research interests include low-complexity, noncoherent detection in MIMO systems. Recently, he was the coauthor of a best paper award at IEEE CCECE'04 in Niagara Falls. Michael has been awarded Natural Sciences and Engineering Research Council (NSERC) Scholarships at the bachelors, masters, and doctoral levels.



Paul K. M. Ho received his B.A.Sc degree from University of Saskatchewan in 1981, and his Ph.D. degree in electrical engineering from Queen's University, Kingston, Ontario, in 1985, both in electrical engineering. He joined the School of Engineering Science at Simon Fraser University, in 1985, where he is currently a Professor. Between 1991 and 1992, he was a Senior Communications Engineer at Glenayre Electronics, Vancouver, and was on leave at the Electrical and Computer Engineering Department, the National University of Singapore between 2000 and 2002. His research interests include noncoherent detection in fading channels, coding and modulation, space-time processing, channel estimation, and performance analysis. He was the coauthor of a best paper award at IEEE VTC'04 Fall in Los Angeles, and at IEEE CCECE'04 in Niagara Falls. Paul is a registered Professional Engineer in the province of British Columbia, and has been a consultant to a number of companies in Canada and abroad.



Jae H. Kim was born in Seoul, Korea. He received his B.S. and M.S. degrees in electronics engineering from Korea University, Seoul, Korea, in 1983 and 1985, respectively. He received the Ph.D. degree in communication engineering from Korea University in August, 1989. Since 1991, he has been with Changwon National University, where he is currently a Professor of the School of Mechatronics Engineering. His current research interests include wireless modem design and implementation.



The Extended-Window Channel Estimator for Iterative Channel-and-Symbol Estimation

Renato R. Lopes

DSPCom, DECOM, FEEC, University of Campinas (UNICAMP), 400 Albert Einstein Avenue, 13083-970 Campinas, Sao Paulo, Brazil
Email: rlopes@decom.fee.unicamp.br

John R. Barry

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, USA
Email: barry@ece.gatech.edu

Received 29 April 2004; Revised 21 September 2004

The application of the expectation-maximization (EM) algorithm to channel estimation results in a well-known iterative channel-and-symbol estimator (ICSE). The EM-ICSE iterates between a symbol estimator based on the forward-backward recursion (BCJR equalizer) and a channel estimator, and may provide approximate maximum-likelihood blind or semiblind channel estimates. Nevertheless, the EM-ICSE has high complexity, and it is prone to misconvergence. In this paper, we propose the extended-window (EW) estimator, a novel channel estimator for ICSE that can be used with any soft-output symbol estimator. Therefore, the symbol estimator may be chosen according to performance or complexity specifications. We show that the EW-ICSE, an ICSE that uses the EW estimator and the BCJR equalizer, is less complex and less susceptible to misconvergence than the EM-ICSE. Simulation results reveal that the EW-ICSE may converge faster than the EM-ICSE.

Keywords and phrases: blind channel estimation, EM algorithm, maximum-likelihood estimation, iterative systems.

1. INTRODUCTION

Channel estimation is an important part of communications systems. Channel estimates are required by equalizers that minimize the bit error rate (BER), and can be used to compute the coefficients of suboptimal but lower-complexity equalizers such as the minimum mean-squared error (MMSE) linear equalizer (LE) [1], or the decision-feedback equalizer (DFE) [1]. Traditionally, a sequence of known bits, called a training sequence, is transmitted for the purpose of channel estimation [1]. These known symbols and their corresponding received samples are used to estimate the channel. However, this approach, known as trained estimation, ignores received samples corresponding to the information bits, and thus does not use all the information available at the receiver. Alternatively, semiblind estimators [2] use every available channel output for channel estimation. Thus, they outperform estimators based solely on the channel outputs corresponding to training symbols, and require a shorter training sequence. Channel estimation is still

possible even if no training sequence is available, using a technique known as blind channel estimation.

An important class of algorithms for blind and semiblind channel estimation is based on the iterative strategy depicted in Figure 1 [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14], which we call *iterative channel-and-symbol estimation* (ICSE). In these algorithms, an initial channel estimate is used by a *symbol estimator* to provide initial estimates of the first-order (and possibly also the second-order) statistics of the transmitted symbol sequence. These estimates are used by a *channel estimator* to improve the initial channel estimates. The process is then repeated. The hope is that several iterations between these two low-complexity estimators will lead to estimates that nearly maximize the joint likelihood function.

The application of the expectation-maximization (EM) algorithm, also known as the Baum-Welch algorithm [15, 16], to the blind channel estimation problem results in the canonical ICSE that fits the framework of Figure 1. An EM iterative channel-and-symbol estimator (EM-ICSE) was first reported in [4], and it has some useful properties. First, it generates a sequence of estimates with nondecreasing likelihood, so that the channel estimates are capable of approaching the maximum-likelihood (ML) estimates.

This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

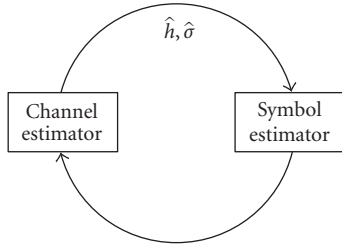


FIGURE 1: Iterative channel-and-symbol estimator.

Second, its symbol estimator is based on the forward-backward recursion of Bahl et al. (BCJR) [17], which minimizes the probability of decision error. Third, the EM-ICSE may be easily modified to exploit, in a natural and nearly optimal way, any a priori information the receiver may have about the transmitted symbols. This a priori information may arise because of pilot symbols (e.g., in semiblind estimation) or error-control coding (e.g., in the context of turbo equalization [6, 7, 8, 9]).

The application of iterative channel estimation to turbo equalization is particularly important, since it leads to channel estimates that benefit from the presence of channel coding, thus performing well at low signal-to-noise ratios [6, 7, 8, 9]. This is particularly important because powerful codes such as turbo codes [18, 19] allow reliable communication at extremely low signal-to-noise ratios, which only exacerbates the estimation problem for traditional channel estimators that ignore the existence of coding, as is the case with most blind channel estimation techniques.

The EM-ICSE has two main drawbacks that we address in this paper: its tendency to converge to inaccurate channel estimates, and its high computational complexity. The problem of convergence to inaccurate estimates arises because the EM-ICSE necessarily generates a sequence of estimates with nondecreasing likelihood. This property makes the EM-ICSE susceptible to getting trapped in a local maximum of the likelihood function. Also, the EM-ICSE has two sources of complexity. First, the EM channel estimator involves the computation and inversion of a square matrix whose order is equal to the channel length. Second, and more important, the complexity of the EM symbol estimator is exponential in the channel length. In [11, 12], ICSEs are proposed that reduce the complexity of the EM-ICSE by introducing a low-complexity symbol estimator. However, these works focus only on the symbol estimator, and use the same channel estimator as the EM-ICSE, resulting in a computational complexity that grows with the square of the channel memory.

In this work, we focus on the channel estimator of Figure 1. We will propose the simplified EM channel estimator (SEM), a channel estimator for ICSE that avoids the matrix inversion of the EM channel estimator. More importantly, an ICSE based on the SEM channel estimator does not require the BCJR equalizer, and thus may be implemented with any number of low-complexity alternatives to the BCJR algorithm, such as those proposed in [20, 21]. Since the complexity of the SEM channel estimator is linear in the channel memory, the overall complexity of an ICSE based on the SEM

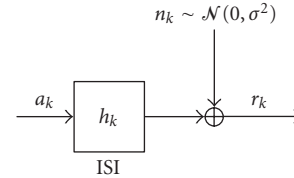


FIGURE 2: Channel model.

channel estimator is also linear if a linear-complexity equalizer is used.

We will also investigate the convergence of an ICSE based on the SEM estimator. We will see that, after misconvergence, the SEM channel estimates may have a structure that can be exploited to escape the local maximum of the likelihood function. We then propose the extended-window (EW) channel estimator, a simple modification to the SEM channel estimator that exploits this structure and greatly decreases the probability of misconvergence, without significantly affecting the computational complexity.

This paper is organized as follows. In Section 2 we present the channel model and describe the problem we will investigate. In Section 3, we briefly review the EM-ICSE. In Section 4, we propose the SEM estimator, a linear-complexity channel estimator for ICSE that is not intrinsically linked to a symbol estimator. In Section 5, we propose the EW estimator, an extension to the SEM estimator of Section 4 that is less likely than EM to get trapped in a local maximum of the joint likelihood function. In Section 6, we present some simulation results, and we draw some conclusions in Section 7.

2. CHANNEL MODEL AND PROBLEM STATEMENT

Consider the transmission of K zero-mean, uncorrelated symbols a_k belonging to some alphabet \mathcal{A} , with unit energy $E[|a_k|^2] = 1$, across a dispersive channel with memory μ and additive-white Gaussian noise. The received signal at time k can be written as

$$r_k = \mathbf{h}^T \mathbf{a}_k + n_k, \quad (1)$$

where $\mathbf{h} = (h_0, h_1, \dots, h_\mu)^T$ represents the channel impulse response, $\mathbf{a}_k = (a_k, a_{k-1}, \dots, a_{k-\mu})^T$, and n_k represents white Gaussian noise with variance σ^2 . Let $\mathbf{a} = (a_0, a_1, \dots, a_{K-1})$ and $\mathbf{r} = (r_0, r_1, \dots, r_{N-1})$ denote the input and output sequences, respectively, where $N = K + \mu$. The resulting channel model is depicted in Figure 2.

Notice that, as far as channel estimation is concerned, the assumption that the transmitted symbols are uncorrelated is not too restrictive. Indeed, most training sequences are chosen so as to satisfy this assumption (thus minimizing the Cramér-Rao bound [22]) and the presence of an interleaver in most coded systems also ensures that the transmitted sequence is approximately uncorrelated. In other words, for channel estimation purposes, assuming that the transmitted

symbols are uncorrelated does not exclude the presence of a training sequence or of a channel code. As we will see, it is the symbol estimator in Figure 1 that exploits the presence of a training sequence or of a channel code.

This paper concerns the joint estimation of \mathbf{a} , \mathbf{h} , and σ relying solely on the received signal \mathbf{r} . Ideally, we would like to solve the joint-ML channel estimation and symbol detection problem, that is, find

$$(\hat{\mathbf{a}}_{\text{ML}}, \hat{\mathbf{h}}_{\text{ML}}, \hat{\sigma}_{\text{ML}}) = \arg \max \log p_{\mathbf{h}, \sigma}(\mathbf{r}|\mathbf{a}), \quad (2)$$

where $\log p_{\mathbf{h}, \sigma}(\mathbf{r}|\mathbf{a})$ is the log-likelihood function, defined as the logarithm of the pdf of the received signal \mathbf{r} conditioned on the channel input \mathbf{r} and parameterized by \mathbf{r} and σ . Intuitively, the ML estimates are those that best explain the received sequence, in the sense that we are less likely to observe the channel output if we assume any other set of parameters to be correct, that is, $p_{\mathbf{h}, \sigma}(\mathbf{r}|\mathbf{a}) \geq p_{\mathbf{h}_{\text{ML}}, \sigma_{\text{ML}}}(\mathbf{r}|\mathbf{a}_{\text{ML}})$ for all \mathbf{h} , σ , \mathbf{a} . Besides this intuitive interpretation, ML estimates have many interesting theoretical properties [22].

It is noteworthy that the maximization in (2) should be performed over the set of valid transmitted sequences. Thus, the joint-ML channel-and-symbol estimation problem in (2) incorporates all possible scenarios: fully trained estimation (all of \mathbf{a} is known); semiblind estimation without coding (parts of \mathbf{a} are known, unknown parts of \mathbf{a} can be any sequence of symbols); semiblind estimation with coding (parts of \mathbf{a} are known, \mathbf{a} must be a valid codeword); blind estimation without coding (none of \mathbf{a} is known, \mathbf{a} can be any sequence of symbols); and blind estimation with coding (none of \mathbf{a} is known, \mathbf{a} must be a valid codeword).

Unfortunately, a direct solution to the problem in (2) is too complex. Therefore, this paper focuses on iterative techniques that provide an approximate solution to (2) with reasonable computational complexity. In the sequel, we review the EM-ICSE, an ICSE that computes a sequence of estimates with nondecreasing likelihood and that, with proper initialization or if the likelihood function is well-behaved, will converge to the ML estimates.

3. THE EM-ICSE

The EM algorithm [15, 16] provides an iterative solution to the blind identification problem in (2) that fits the paradigm of Figure 1, as first reported in [4]. The EM channel estimator (see Figure 1) for the $(i+1)$ th iteration of the EM-ICSE is defined by

$$\hat{\mathbf{h}}_{(i+1)} = \mathbf{R}_i^{-1} \mathbf{p}_i, \quad (3)$$

$$\begin{aligned} \hat{\sigma}_{(i+1)}^2 &= \frac{1}{N} \sum_{k=0}^{N-1} E \left[|r_k - \hat{\mathbf{h}}_{(i+1)}^T \mathbf{a}_k|^2 | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2 \right] \\ &= \frac{1}{N} \sum_{k=0}^{N-1} |r_k|^2 - 2\hat{\mathbf{h}}_{(i+1)}^T \mathbf{p}_i + \hat{\mathbf{h}}_{(i+1)}^T \mathbf{R}_i \hat{\mathbf{h}}_{(i+1)}, \end{aligned} \quad (4)$$

where

$$\mathbf{R}_i = \frac{1}{N} \sum_{k=0}^{N-1} E [\mathbf{a}_k \mathbf{a}_k^T | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2], \quad (5)$$

$$\mathbf{p}_i = \frac{1}{N} \sum_{k=0}^{N-1} r_k E [\mathbf{a}_k | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2]. \quad (6)$$

The EM symbol estimator (see Figure 1) provides the values of $\hat{\mathbf{a}}_k^{(i)} = E[\mathbf{a}_k | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2]$ and $E[\mathbf{a}_k \mathbf{a}_k^T | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2]$ that are required by (5) and (6). The a posteriori expected values in (5) and (6) are computed assuming that $\hat{\mathbf{h}}_{(i)}$ and $\hat{\sigma}_{(i)}^2$ are the actual channel parameters. Notice that $\hat{a}_k = E[a_k | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2]$ is the a posteriori MMSE estimate of a_k , and we refer to \hat{a}_k as a *soft symbol estimate*.

Also, note that \mathbf{R}_i and \mathbf{p}_i of (5) and (6) can be viewed as estimates of the a posteriori autocorrelation matrix of the transmitted sequence and the cross-correlation vector between the transmitted and received sequences, respectively. Thus, (3) and (4) are similar to the MMSE-trained channel estimator [22], in which \mathbf{R}_i and \mathbf{p}_i are computed with the actual transmitted sequence.

The computation of the expected values in (5) and (6) require the knowledge of the a posteriori probabilities $E[\mathbf{a}_k | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2]$ and $E[\mathbf{a}_k \mathbf{a}_k^T | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2]$. For an uncoded system, these can be exactly computed with the forward-backward recursion or BCJR algorithm [17]. Because the computational complexity of this algorithm grows exponentially with the channel length, some authors [11, 12] have proposed lower-complexity alternatives that compute approximations to these a posteriori probabilities. In other words, the algorithms of [11, 12] are approximations to the EM-ICSE that also fit the framework of Figure 1, and that are also based on the channel estimator of (3), (4), (5), and (6).

Unfortunately, in the presence of a channel code, an exact computation of \mathbf{R}_i and \mathbf{p}_i is prohibitively complex. The most common solution in this case is to modify the EM-ICSE, using a turbo equalizer as the symbol estimator [6]. In other words, for coded systems, $E[\mathbf{a}_k | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2]$ and $E[\mathbf{a}_k \mathbf{a}_k^T | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2]$ are based on the decoder output. Similarly, the presence of training symbols is easily handled by the symbol estimator, which only has to set the training symbols as deterministic constants when computing \mathbf{R}_i and \mathbf{p}_i . Based on these two observations, we see that the channel estimator of the EM-ICSE always ignores the presence of a training sequence or of a channel code. It is the symbol estimator that exploits the structure of the transmitted symbols to improve their estimates.

4. A SIMPLIFIED EM CHANNEL ESTIMATOR

In this section, we propose the simplified EM estimator (SEM), an alternative to the EM channel estimator in (3), (4), (5), and (6) that avoids the computation of \mathbf{R}_i and the matrix inversion of (3). To derive the SEM estimator, we note that, from channel model (1) and the uncorrelatedness assumption, we get $h_n = E[r_k a_{k-n}]$. This expected value may

be computed by conditioning on \mathbf{r} , yielding

$$E[r_k a_{k-n}] = E[E[r_k a_{k-n} | \mathbf{r}]] = E[r_k E[a_{k-n} | \mathbf{r}]], \quad (7)$$

where the last equality follows from the fact that r_k is a constant given \mathbf{r} . Note that the channel estimator has no access to $E[a_k | \mathbf{r}]$, which requires exact channel knowledge. However, based on the iterative paradigm of Figure 1, at the i th iteration the channel estimator does have access to $\tilde{a}_k^{(i)} = E[a_k | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2]$. Replacing this value in (7), and also replacing a time average for the ensemble average, leads to the following channel estimator:

$$\hat{h}_n^{(i+1)} = \frac{1}{N} r_k \tilde{a}_{k-n}^{(i)} \quad \text{for } n = 0, 1, \dots, \mu. \quad (8)$$

Notice that in (8) the channel is estimated by correlating the received signal with the soft symbol estimates \tilde{a}_k . This is similar to the fully trained channel estimator of [23, 24], known as channel probing, except that the training symbols have been replaced by their soft estimates.

As for estimating the noise variance, let $\hat{a}_k^{(i)}$ be a hard decision of the k th transmitted symbol, chosen as the element of \mathcal{A} closest to $\tilde{a}_k^{(i)}$. Also, define the vector $\hat{\mathbf{a}}_k^{(i)} = (\hat{a}_k^{(i)}, \hat{a}_{k-1}^{(i)}, \dots, \hat{a}_{k-\mu}^{(i)})^T$. We propose to compute $\hat{\sigma}_{(i+1)}^2$ using

$$\hat{\sigma}_{(i+1)}^2 = \frac{1}{N} \sum_{k=0}^{N-1} |r_k - \hat{\mathbf{h}}_{(i+1)}^T \hat{\mathbf{a}}_k^{(i)}|^2. \quad (9)$$

Notice that in (9) we use hard instead of soft symbol estimates. In our simulations, we found that doing so improved convergence speed.

Remark 1. Combining the estimates (8) into a single vector, we find that $\hat{\mathbf{h}}_{(i+1)} = (\hat{h}_0^{(i+1)}, \dots, \hat{h}_\mu^{(i+1)})^T = \mathbf{p}_i$. Thus, we may view (8) as a simplification of the EM estimate $\mathbf{R}_i^{-1} \mathbf{p}_i$ that avoids matrix inversion by approximating \mathbf{R}_i by \mathbf{I} . This approximation is reasonable, since \mathbf{R}_i is a posteriori estimate of the autocorrelation matrix of the transmitted vector, which, due to the uncorrelatedness assumption, is close to the identity for large N . Furthermore, since this approximation results in a channel estimator that is less complex than the EM channel estimator defined in (3) and (4), we refer to the channel estimator defined by (8) and (9) as the simplified EM estimator (SEM).

Remark 2. The SEM channel estimator requires only the soft symbol estimates $\tilde{a}_k^{(i)}$, so that an ICSE based on the SEM estimator may be represented as in Figure 3. Note that any equalizer that produces soft symbol estimates can be used, which allows for an even lower-complexity implementation of an SEM-based ICSE, using equalizers such as those proposed in [20, 21].

Remark 3. It is interesting to notice that, while substituting the actual values of \mathbf{h} or \mathbf{a} for their estimates will always improve the performance of the iterative algorithm, the same is not true for σ . Indeed, substituting σ for $\hat{\sigma}$ will often result

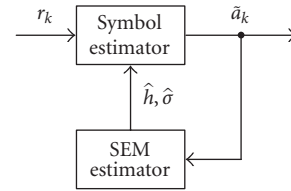


FIGURE 3: Iterative channel-and-symbol estimation with the SEM channel estimator.

in performance degradation. Intuitively, one can think of $\hat{\sigma}$ as playing two roles: in addition to measuring σ , it also acts as a measure of reliability of the channel estimate $\hat{\mathbf{h}}$. In fact, consider a decomposition of the channel output:

$$r_k = \hat{\mathbf{h}}^T \mathbf{a}_k + (\mathbf{h} - \hat{\mathbf{h}})^T \mathbf{a}_k + n_k. \quad (10)$$

The term $(\mathbf{h} - \hat{\mathbf{h}})^T \mathbf{a}_k$ represents the contribution to r_k from the estimation error. By using $\hat{\mathbf{h}}$ to model the channel in the BCJR algorithm, we are in effect lumping the estimation error with the noise, resulting in an effective noise sequence with variance larger than σ^2 . It is thus appropriate that $\hat{\sigma}$ should exceed σ whenever $\hat{\mathbf{h}}$ differs from \mathbf{h} . Alternatively, it stands to reason that an unreliable channel estimate should translate to an unreliable symbol estimate, regardless of how well $\hat{\mathbf{h}}^T \mathbf{a}_k$ matches r_k . Using a large value of $\hat{\sigma}$ in the BCJR equalizer ensures that its output will have a small reliability. Fortunately, the noise variance estimate produced by (9) measures the energy of both the second and the third term in (10). If $\hat{\mathbf{h}}$ is a poor channel estimate, $\tilde{\mathbf{a}}$ will also be a poor estimate for \mathbf{a} , and convolving $\tilde{\mathbf{a}}$ and $\hat{\mathbf{h}}$ will produce a poor match for \mathbf{r} , so that (9) will produce a large estimated noise variance.

5. THE EXTENDED-WINDOW CHANNEL ESTIMATOR

Misconvergence is a common characteristic of ICSEs, especially in blind systems. To illustrate this problem, consider estimating the channel $\mathbf{h} = (1, 2, 3, 4, 5)^T$ with a BPSK constellation and $\text{SNR} = \|\mathbf{h}\|^2 / \sigma^2 = 20$ dB. An ICSE based on the BCJR symbol estimator and the SEM channel estimator converges to $\hat{\mathbf{h}}^{(20)} = (2.1785, 3.0727, 4.1076, 5.0919, 0.1197)^T$ after 20 iterations, with $K = 1000$ bits, with initialization $\hat{\mathbf{h}}^{(0)} = (1, 0, 0, 0, 0)^T$ and $\hat{\sigma}_{(0)}^2 = 1$. Although the algorithm fails, $\hat{\mathbf{h}}^{(20)}$ is seen to roughly approximate a shifted (or delayed) and truncated version of the actual channel. A possible explanation for this behavior is that the channel is maximum phase, while we used a minimum phase initialization. This phase mismatch between $\hat{\mathbf{h}}$ and the initialization $\hat{\mathbf{h}}^{(0)}$ introduces a delay that cannot be compensated for by the iterative scheme. In fact, after convergence, a_k is approximately $\text{sign}(\tilde{a}_{k+1})$, and h_0 can be accurately estimated by correlating r_k with \tilde{a}_{k+1} . However, because the delay n in (8) is limited to the narrow window $0, \dots, \mu$, this correlation is never computed. This observation leads us to propose

the *extended-window* (EW) channel estimator, in which (8) is computed for a broader range of n .

To determine how much the correlation window must be extended, consider two extreme cases. First, suppose $\mathbf{h} \approx (0, \dots, 0, 0, 1)^T$, so that $r_k \approx a_{k-\mu} + n_k$. Also, assume that $\hat{\mathbf{h}} \approx (1, 0, 0, \dots, 0)^T$. In this case, assuming a BPSK constellation, the symbol estimator output is $\hat{a}_k = \tanh(r_k/\sigma^2)$. Hence, assuming a large SNR, $\hat{a}_k \approx a_{k-\mu}$, so to estimate h_0 and h_μ we must compute (8) for $n = -\mu$ and $n = 0$, respectively. Likewise, if $\mathbf{h} \approx (1, 0, 0, \dots, 0)^T$ and $\hat{\mathbf{h}} \approx (0, \dots, 0, 0, 1)^T$, the symbol estimator output \hat{a}_k is such that $\hat{a}_k \approx a_{k+\mu}$, so to estimate h_0 and h_μ we must compute (8) for $n = \mu$ and $n = 2\mu$, respectively. These observations, based on two extreme cases, suggest the *extended-window* (EW) channel estimator, which computes

$$g_n = \frac{1}{N} \sum_{k=0}^{N-1} r_k \hat{a}_{k-n}^{(i)} \quad \text{for } n = -\mu, \dots, 2\mu. \quad (11)$$

By doing this, we ensure that $\mathbf{g} = (g_{-\mu}, \dots, g_{2\mu})^T$ has $\mu + 1$ entries that estimate the desired correlations $E[r_k a_{k-n}]$, for $n \in \{0, \dots, \mu\}$. Its remaining terms are an estimate of $E[r_k a_{k-n}]$ for $n \notin \{0, \dots, \mu\}$, and hence should be close to zero. Therefore, we define the EW channel estimates by

$$\hat{\mathbf{h}}_{(i+1)} = (g_\delta, \dots, g_{\delta+\mu})^T, \quad (12)$$

where the delay parameter $\delta \in \{-\mu, \dots, \mu\}$ is chosen so that $\hat{\mathbf{h}}_{(i+1)}$ represents the $\mu + 1$ consecutive coefficients of \mathbf{g} with highest energy. In other words, δ is chosen to maximize $\|\hat{\mathbf{h}}_{(i+1)}\|^2$.

Notice that after convergence we expect that $g_\delta \approx h_0$. Comparing (7) and (11), we note that this is equivalent to saying that $a_k \approx \hat{a}_{k-\delta}^{(i)}$. This delay must be taken into account in the estimation of the noise variance. With that in mind, we propose to estimate σ^2 using a modified version of (9), namely

$$\hat{\sigma}_{(i+1)}^2 = \frac{1}{N} \sum_{k=0}^{N-1} |r_k - \hat{\mathbf{h}}_{(i+1)}^T \hat{\mathbf{a}}_{k-\delta}^{(i)}|^2. \quad (13)$$

5.1. Computational complexity

We now compare the computational complexity of the EW channel estimator of (11), (12), and (13) to that of the EM channel estimator of (3) and (4). We ignore the cost of computing \hat{a}_k , and we consider the complexity in terms of sums and multiplications per received symbol.

For each received symbol, the EW algorithm performs $3\mu + 1$ multiplications and $3\mu + 1$ additions to compute the vector \mathbf{g} in (11). The division by N , as well as the computation of δ , is done only once per block of N received symbols, and thus can be ignored. The computation of each term in the summation in (13) involves $\mu + 2$ multiplications and the same number of sums. Hence, the total computational cost of the EW channel estimator is $4\mu + 4$ multiplications and $4\mu + 4$ sums.

For the EM channel estimator, we consider that $E[\mathbf{a}_k \mathbf{a}_k^T | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2] \approx E[\mathbf{a}_k | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2] E[\mathbf{a}_k | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2]^T$. This approximation is used in [11, 12], and allows for a simpler complexity comparison. With this simplification, and noting that $E[\mathbf{a}_k \mathbf{a}_k^T | \mathbf{r}; \hat{\mathbf{h}}_{(i)}, \hat{\sigma}_{(i)}^2]$ is a symmetric matrix, we see that the computation of \mathbf{R}_i in (5) requires $(\mu + 1)\mu/2$ multiplications and an equal number of sums per received symbol. On the other hand, the computation of \mathbf{p}_i in (6) requires $\mu + 1$ multiplications and sums per received symbol. The linear system in (3) is solved only once, so that its cost can be ignored. The same can be said about most of the operations in (4), except for its first term, which requires 1 multiplication and sum per received symbol. Thus, the total cost of this approximate EM channel estimator is $\mu^2/2 + 3\mu/2 + 2$ multiplications and sums per received symbol.

6. SIMULATION RESULTS

In this section, we use simulations to compare the performance of the fully blind EM-ICSE and the fully blind EW-ICSE, assuming both ICSEs use the BCJR symbol estimator. The results presented in this section all correct for the aforementioned shifts in the estimates. In other words, when computing channel estimation error or BER, the channel and symbol estimates were shifted to best match the actual channel or the transmitted sequence. Note that this shift was done only for the purpose of computing the errors, and hence did not affect the estimates in the iterative procedure.

For comparison purposes, we also consider fully trained channel estimators, in which all the transmitted bits are assumed known by the channel estimator. We consider the fully trained MMSE estimator which, as discussed in Section 3, can be seen as a trained version of the EM channel estimator. We also consider channel probing which, as discussed in Section 4, can be seen as the trained counterpart of the EW channel estimator. In the simulations of the trained estimators, we use the same block of received samples to estimate the channel (assuming that all transmitted symbols are known) and to estimate the transmitted symbols (with the BCJR equalizer, using the trained channel estimates).

As a first test of the EW-ICSE, we simulated the transmission of $K = 600$ BPSK symbols over the channel $\mathbf{h} = (-0.2287, 0.3964, 0.7623, 0.3964, -0.2287)^T$ from [12]. To stress the fact that the EW-ICSE is not sensitive to initial conditions, we initialized $\hat{\mathbf{h}}$ randomly using $\hat{\mathbf{h}}_{(0)} = \mathbf{u}\hat{\sigma}_{(0)}/\|\mathbf{u}\|$, where $\mathbf{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\hat{\sigma}_{(0)}^2 = \sum_{k=0}^{N-1} |r_k|^2/2N$. By assigning half of the received energy to the signal and half to the noise, we are essentially initializing the SNR estimate to 0 dB. In Figure 4, we show the convergence behavior of the EW-ICSE estimates, averaged over 100 independent runs of this experiment using $\text{SNR} = \|\mathbf{h}\|^2/\sigma^2 = 9$ dB. Only the convergence of \hat{h}_0 , \hat{h}_1 , and \hat{h}_2 is shown; the behavior of \hat{h}_3 and \hat{h}_4 is similar to that of \hat{h}_2 and \hat{h}_0 , respectively, but we show only the coefficients with worse convergence. The shaded regions around the channel estimates correspond to plus and minus one standard deviation. For comparison, we show the average behavior of the EM channel estimates in Figure 5.

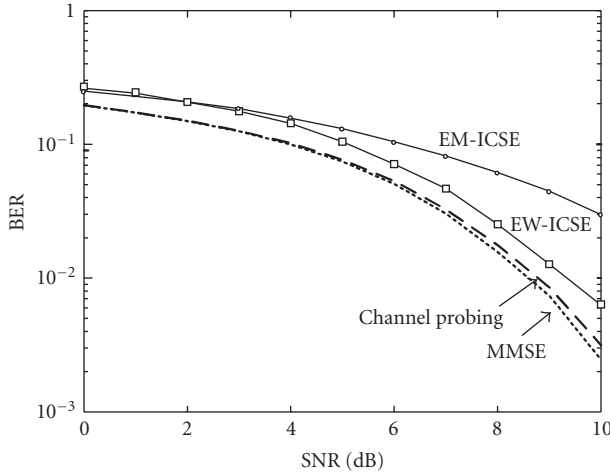


FIGURE 8: Bit error rate versus SNR using EM and EW estimates after 20 iterations. Also shown is the performance resulting from the use of the trained channel probing and trained MMSE estimates.

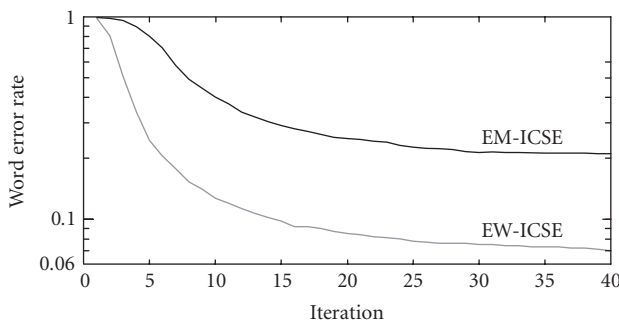


FIGURE 9: WER for the EW-ICSE and the EM-ICSE, averaged over 1000 random channels.

similar. In other words, the channel probing estimates are “good enough,” and the added complexity of the MMSE estimator does not have much impact on the BER performance in the SNR range considered here. Finally, we observed that the BER performance of a BCJR equalizer with channel knowledge cannot be distinguished from that of a BCJR equalizer using the MMSE estimates.

To further support the claim that the EW-ICSE avoids most of the local maxima of the likelihood function that trap the EM-ICSE, we ran both the EM-ICSE and the EW-ICSE on 1000 random channels of memory $\mu = 4$, generated as $\mathbf{h} = \mathbf{u}/\|\mathbf{u}\|$, where $\mathbf{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. The estimates were initialized to $\hat{\sigma}_{(0)}^2 = \sum_{k=0}^{N-1} |r_k|^2 / 2N$ and $\hat{\mathbf{h}}^{(0)} = (0, \dots, 0, \hat{\sigma}_{(0)}, 0, \dots, 0)^T$, that is, the center tap of $\hat{\mathbf{h}}^{(0)}$ is initialized to $\hat{\sigma}_{(0)}$. We used SNR = 18 dB, and blocks of $K = 1000$ BPSK symbols. In Figure 9 we show the word error rate (WER) (fraction of blocks detected with errors) of the EW-ICSE and the EM-ICSE versus iteration. It is again clear that the EW-ICSE outperforms the EM-ICSE. It should be noted that in this example the equalizer based on the channel probing estimates was able to detect all transmitted sequences correctly.

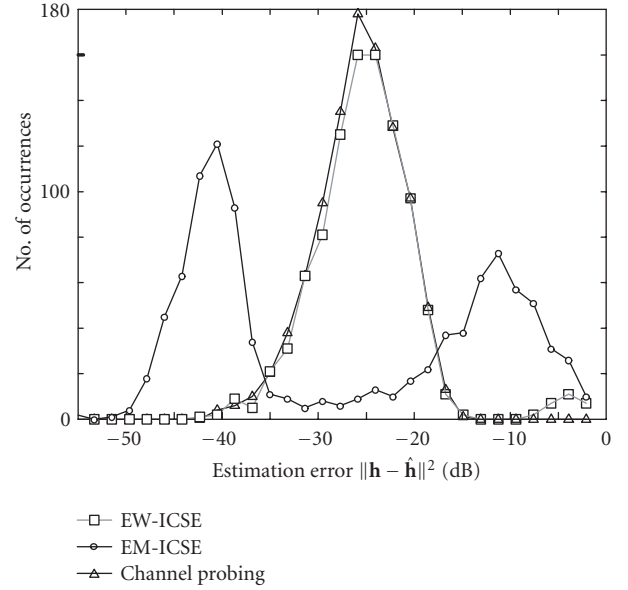


FIGURE 10: Histograms of estimation errors for the EW-ICSE and the EM-ICSE over an ensemble of 1000 random channels.

The better performance of the EW estimates can also be seen in Figure 10, where we show histograms of the estimation errors (in dB) for the channel probing, the EW, and the EM estimates, computed after 50 iterations. We see that while only 3% of the EW estimates have an error larger than -16 dB, 35% of the EM estimates have an error larger than -16 dB. In fact, the histogram for the EW estimates is very similar to that of the channel probing estimates, which again shows the good convergence properties of the EW-ICSE. It is also interesting to note in Figure 10 that the EM estimates have a bimodal behavior: the estimation errors produced by the EM-ICSE are grouped around -11 dB and -43 dB. These groups are respectively better than and worse than the channel probing estimates. This bimodal behavior can be explained by the fact that the EM algorithm often converges to inaccurate estimates, leading to large estimation errors. On the other hand, when the EM algorithm does work, it works very well.

7. CONCLUSIONS

We presented the EW channel estimator, a linear-complexity channel estimator for ICSE. We have shown that this technique can be seen as a modification of the EM channel estimator. A key feature of the EW estimator is its extended window, which greatly improves the convergence behavior of ICSEs based on the EW estimator, avoiding most of the local maxima of the likelihood function that trap the EM-ICSE. Furthermore, the computational complexity of the EW estimator grows linearly with the channel memory, as opposed to the quadratic complexity of the EM channel estimator. Additionally, the EW estimator may be used with any soft-output equalizer. This allows for even further complexity reduction when compared to the EM-ICSE, which requires

a BCJR equalizer. However, simulations show that, despite its good convergence properties, the EW-ICSE is not globally convergent. The problem of devising an iterative strategy that is guaranteed to always avoid misconvergence, regardless of initialization, remains open.

REFERENCES

- [1] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communications*, Kluwer Academic Publishers, Norwell, Mass, USA, 3rd edition, 2003.
- [2] J. Ayadi, E. de Carvalho, and D. T. M. Slock, "Blind and semi-blind maximum likelihood methods for FIR multichannel identification," in *Proc. IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP '98)*, vol. 6, pp. 3185–3188, Seattle, Wash, USA, May 1998.
- [3] M. Feder and J. A. Catipovic, "Algorithms for joint channel estimation and data recovery-application to equalization in underwater communications," *IEEE J. Oceanic Eng.*, vol. 16, no. 1, pp. 42–55, 1991.
- [4] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels," *IEEE Trans. Commun.*, vol. 42, no. 7, pp. 2406–2413, 1994.
- [5] C. Anton-Haro, J. A. R. Fonollosa, and J. R. Fonollosa, "Blind channel estimation and data detection using hidden Markov models," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 241–247, 1997.
- [6] J. Garcia-Frias and J. D. Villasenor, "Combined turbo detection and decoding for unknown ISI channels," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 79–85, 2003.
- [7] K.-D. Kammeyer, V. Kühn, and T. Petermann, "Blind and nonblind turbo estimation for fast fading GSM channels," *IEEE J. Select. Areas Commun.*, vol. 19, no. 9, pp. 1718–1728, 2001.
- [8] A. O. Berthet, B. S. Ünäl, and R. Visoz, "Iterative decoding of convolutionally encoded signals over multipath Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 19, no. 9, pp. 1729–1743, 2001.
- [9] R. R. Lopes and J. R. Barry, "Exploiting error-control coding in blind channel estimation," in *IEEE Global Communications Conference (GLOBECOM '01)*, vol. 2, pp. 1317–1321, San Antonio, Tex, USA, November 2001.
- [10] V. Krishnamurthy and J. B. Moore, "On-line estimation of hidden Markov model parameters based on the Kullback-Leibler information measure," *IEEE Trans. Signal Processing*, vol. 41, no. 8, pp. 2557–2573, 1993.
- [11] L. B. White, S. Perreau, and P. Duhamel, "Reduced computation blind equalization for FIR channel input Markov models," in *IEEE International Conference on Communications (ICC '95)*, vol. 2, pp. 993–997, Seattle, Wash, USA, June 1995.
- [12] M. Shao and C. L. Nikias, "An ML/MMSE estimation approach to blind equalization," in *Proc. IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP '94)*, vol. 4, pp. 569–572, Adelaide, SA, Australia, April 1994.
- [13] H. A. Cirpan and M. K. Tsatsanis, "Stochastic maximum likelihood methods for semi-blind channel estimation," *IEEE Signal Processing Lett.*, vol. 5, no. 1, pp. 21–24, 1998.
- [14] B.-P. Paris, "Self-adaptive maximum-likelihood sequence estimation," in *IEEE Global Communications Conference (GLOBECOM '93)*, vol. 4, pp. 92–96, Houston, Tex, USA, November–December 1993.
- [15] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *Annals of Mathematics Statistics*, vol. 41, no. 1, pp. 164–171, 1970.
- [16] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistics Society*, vol. 39, no. 1, pp. 1–38, 1977.
- [17] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, 1974.
- [18] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: turbo-codes," in *IEEE International Conference on Communications (ICC '93)*, vol. 2, pp. 1064–1070, Geneva, Switzerland, May 1993.
- [19] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: performance analysis, design, and iterative decoding," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 909–926, 1998.
- [20] M. Tüchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, 2002.
- [21] R. R. Lopes and J. R. Barry, "Soft-output decision-feedback equalization with a priori information," in *IEEE Global Communications Conference (GLOBECOM '03)*, vol. 3, pp. 1705–1709, San Francisco, Calif, USA, December 2003.
- [22] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, New York, NY, USA, 2nd edition, 1994.
- [23] C. A. Montemayor and P. G. Flikkema, "Near-optimum iterative estimation of dispersive multipath channels," in *IEEE 48th Vehicular Technology Conference (VTC '98)*, vol. 3, pp. 2246–2250, Ottawa, ON, Canada, May 1998.
- [24] M. Sandell, C. Luschi, P. Strauch, and R. Yan, "Iterative channel estimation using soft decision feedback," in *IEEE Global Communications Conference (GLOBECOM '98)*, vol. 6, pp. 3728–3733, Sydney, NSW, Australia, November 1998.

Renato R. Lopes received the B.S. and M.S. degrees from the University of Campinas (UNICAMP), Brazil, in 1995 and 1997, and the Ph.D. degree from the Georgia Institute of Technology, USA, in 2003, all in electrical engineering. He also received an M.A. degree in applied mathematics from the Georgia Institute of Technology, USA, in 2001. During his studies, he was supported by the Brazilian agencies CNPq and CAPES, and held teaching and research assistant positions from 1999 to 2003. He is currently a postdoctoral researcher at UNICAMP, under a grant from FAPESP. His research interests are in the general area of communications theory, including equalization, identification, iterative receivers, and coding theory.



John R. Barry received the B.S. degree in electrical engineering from the State University of New York, Buffalo, in 1986, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1987 and 1992, respectively. Since 1992, he has been with the Georgia Institute of Technology, Atlanta, where he is an Associate Professor in the School of Electrical and Computer Engineering. Currently he is visiting Georgia Tech Lorraine, Metz, France. His research interests include wireless communications, equalization, and multiuser communications. He is a coauthor with E. A. Lee and D. G. Messerschmitt of *Digital Communications*, third edition, Kluwer, Norwell, Mass, 2004, and the author of *Wireless Infrared Communications*, Kluwer, Norwell, Mass, 1994.



Soft-In Soft-Output Detection in the Presence of Parametric Uncertainty via the Bayesian EM Algorithm

A. S. Gallo

Department of Information Engineering, University of Modena and Reggio Emilia, via Vignolese 905, 41100 Modena, Italy
Email: asgallo@unimo.it

G. M. Vitetta

Department of Information Engineering, University of Modena and Reggio Emilia, via Vignolese 905, 41100 Modena, Italy
Email: giorgio.vitetta@unimo.it

Received 30 April 2004; Revised 6 October 2004

We investigate the application of the *Bayesian expectation-maximization* (BEM) technique to the design of *soft-in soft-out* (SISO) detection algorithms for wireless communication systems operating over channels affected by parametric uncertainty. First, the BEM algorithm is described in detail and its relationship with the well-known *expectation-maximization* (EM) technique is explained. Then, some of its applications are illustrated. In particular, the problems of SISO detection of spread spectrum, single-carrier and multicarrier space-time block coded signals are analyzed. Numerical results show that BEM-based detectors perform closely to the *maximum likelihood* (ML) receivers endowed with perfect channel state information as long as channel variations are not too fast.

Keywords and phrases: expectation-maximization algorithm, soft-in soft-out detection, fading channels, space-time coding, OFDM.

1. INTRODUCTION

In recent years, many research efforts have been devoted to the study of detection algorithms for digital signals transmitted over channels affected by random parametric uncertainty, like multipath fading channels and AWGN channels with phase jitter (see, e.g., [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13] and references therein). In this field the attention has been progressively shifting from *maximum likelihood* (ML) sequence detection [2, 3, 4] to *maximum a posteriori* (MAP) symbol detection techniques [5, 6, 7, 8, 9, 10, 11, 12, 13] producing *a posteriori probabilities* (APPs) on the possible data decisions. This has been mainly due to the need of robust receiver structures for coded modulations and, more specifically, to the advent of the *turbo processing principle* applied to efficient iterative decoding of concatenated coding structures [14, 15, 16, 17, 18, 19, 20, 21, 22]. Such a principle has been also exploited to design iterative detection/equalization/decoding algorithms for interleaved coded signals transmitted over channels with memory

[10, 11, 12, 13, 23]. In all these cases good error performance is achieved by means of concatenated detection/decoding structures exchanging among each other soft information about the detected data. The basic building blocks of these structures are the so-called *soft-in soft-out* (SISO) modules [18, 22].

A wealth of technical papers on the design techniques for ML sequence detectors operating on channels with parametric uncertainty is available (see [1, 2, 3, 4] and references therein). Since in many problems the implementation of the ML strategy is prohibitively complicated, general tools, like the principle of *per-survivor processing* (PSP) [2] and the *expectation-maximization* (EM) algorithm [3, 4, 24, 25], have been proposed to devise quasioptimal receivers. The EM technique is an iterative algorithm generating the ML estimate of a set of deterministic unknown parameters, if properly initialized. It has been successfully applied to a number of problems and, in particular, to the ML detection of digital signals transmitted over fading channels [3, 4, 6, 26] and to carrier phase recovery [3, 7, 27, 28]. The EM algorithm, however, being a technique for ML estimation, is unable to incorporate any statistical information about the unknown parameters to be estimated, even if such information are available.

Recently, an extension of the EM, dubbed *Bayesian EM* (BEM), has been applied to solve MAP estimation problems and to derive SISO receivers [29, 30, 31, 32] for single-user detection over frequency-flat Rayleigh fading channels. The BEM algorithm allows to design SISO modules estimating the channel state, incorporating the symbol *a priori probabilities* (APRPs) and the statistics of the channel uncertainty, and generating the symbol APPs. Therefore, it can be easily employed in iterative equalization/decoding structures for coded transmissions [17, 23]. The favorable features of the BEM technique have suggested to further investigate its application to other communication scenarios.

This paper offers both a tutorial introduction to BEM-based estimation techniques and some recent research results about its applications. In fact, in its first part it describes the BEM technique, its relationship with the EM algorithm, and how it can be used to derive SISO algorithms for the detection of digital data transmitted over channels having memory and affected by parametric uncertainty. Then, in the second part of the paper, the application of the BEM approach to some detection problems of current interest is illustrated. In particular, we consider

- (1) the multiuser detection of *direct sequence spread spectrum* (DSSS) signals in a synchronous CDMA system [33];
- (2) the detection of single-carrier space-time block coded signals transmitted over frequency-flat fading channels [34];
- (3) the detection of multicarrier space-time block coded signals transmitted over frequency-selective fading channels [35].

For each specific problem, in the third scenario, a BEM-based SISO algorithm is described and some numerical results are illustrated. Moreover, the use of a BEM-based SISO module in an iterative receiver is described in detail.

The paper is organized as follows. The EM and BEM techniques are described in Section 2. The use of the BEM technique to devise SISO modules for channels with parametric uncertainty and memory is illustrated in Section 3. Specific applications of the BEM tool are analyzed in Section 4. Finally, Section 5 offers some conclusions.

2. EXPECTATION-MAXIMIZATION ALGORITHMS FOR PARAMETER ESTIMATION

2.1. The EM algorithm

Let $\Theta \doteq [\Theta_0, \Theta_1, \dots, \Theta_{L-1}]^T$ denote an L -dimensional *deterministic* vector to be estimated from an N -dimensional received vector $\mathbf{R} \doteq [R_0, R_1, \dots, R_{N-1}]^T$ of noisy data (with $N \geq L$).¹ The ML estimate of Θ is the solution of the problem [36]

$$\theta_{\text{ML}} = \arg \max_{\theta} L_{\mathbf{r}}(\tilde{\theta}), \quad (1)$$

¹In the following, a random vector and its realizations are always denoted by an uppercase letter and the corresponding lowercase letter, respectively.

where $L_{\mathbf{r}}(\tilde{\theta}) \doteq \log f(\mathbf{r}|\tilde{\theta})$ is a log-likelihood function and $f(\mathbf{x}|\mathbf{y})$ denotes the *probability density function* (pdf) of the random vector \mathbf{X} conditioned on the event $\{\mathbf{Y} = \mathbf{y}\}$. Solving the problem (1) in a direct fashion requires a closed form expression for $L_{\mathbf{r}}(\tilde{\theta})$ but, even if this expression is available, the search for its maximum may entail an unacceptable computational burden. When this occurs, a feasible alternative can be offered by the EM algorithm [3, 25]. The EM approach develops from the assumption that a *complete* data vector $\mathbf{C} = [C_0, C_1, \dots, C_{P-1}]^T$ (with $P \geq N$) is observed in place of the *incomplete* data set \mathbf{R} . The vector \mathbf{C} is characterized by a couple of relevant properties: (1) it is not observed directly but, if available, would ease the estimation of Θ ; (2) \mathbf{R} can be obtained from \mathbf{C} through a many-to-one mapping $\mathbf{C} \rightarrow \mathbf{R}(\mathbf{C})$. In practice, in communication problems, \mathbf{C} is always chosen as a superset of the incomplete data [3], that is,

$$\mathbf{C} = [\mathbf{R}^T, \mathbf{I}^T]^T, \quad (2)$$

where the so-called *imputed* data \mathbf{I} are properly selected to simplify the ML estimation problem [25]. In particular, when Θ consists of all the transmitted channel symbols, \mathbf{I} collects all the unwanted random parameters (fading, phase jitter, etc.) affecting the communication channel [3, 25]. These choices lead to *hard* detection algorithms often having an acceptable complexity and capable of incorporating the statistical properties of the channel parameters. In the following the complete data vector \mathbf{C} will be always structured as in (2).

Given \mathbf{C} , the *auxiliary function*

$$\begin{aligned} Q_{\text{EM}}(\theta, \tilde{\theta}) &\doteq E_{\mathbf{I}}\{L_c(\theta) | \mathbf{R} = \mathbf{r}, \Theta = \tilde{\theta}\} \\ &= E_{\mathbf{I}}\{\log f(\mathbf{C}|\theta) | \mathbf{R} = \mathbf{r}, \Theta = \tilde{\theta}\} \\ &= \int_{\mathcal{S}_{\mathbf{I}}} \log f(\mathbf{r}, \mathbf{i}|\theta) f(\mathbf{i} | \mathbf{r}, \tilde{\theta}) d\mathbf{i} \end{aligned} \quad (3)$$

is evaluated, where $E_{\mathbf{X}}\{\cdot\}$ denotes the statistical average with respect to a random vector \mathbf{X} and $\mathcal{S}_{\mathbf{I}}$ is the space of \mathbf{I} . Then, this function is employed in the following two-step procedure generating successive approximations $\{\theta^{(k)}, k = 1, 2, \dots\}$ of θ_{ML} (1):

- (1) *expectation step*— $Q_{\text{EM}}(\theta, \tilde{\theta})$ in (3) is evaluated for $\tilde{\theta} = \theta_{\text{EM}}^{(k)}$;
- (2) *maximization step*—given $\theta_{\text{EM}}^{(k)}$, the next estimate $\theta_{\text{EM}}^{(k+1)}$ is computed as

$$\theta_{\text{EM}}^{(k+1)} = \arg \max_{\theta} Q_{\text{EM}}(\theta, \theta_{\text{EM}}^{(k)}), \quad k = 0, 1, \dots \quad (4)$$

An initial estimate $\theta_{\text{EM}}^{(0)}$ of θ must be provided for the algorithm start-up. In digital communication problems, proper initialization of the EM algorithm is usually accomplished exploiting the information provided by known (pilot) symbols [3]. It can be proved that, under mild conditions, the sequence $\{\theta_{\text{EM}}^{(k)}\}$ converges to the true ML estimate θ_{ML} of (1), provided that the existence of local maxima does not prevent it. Avoiding this requires an accurate initial estimate $\theta_{\text{EM}}^{(0)}$ whose choice, for this reason, is critical [25].

2.2. The BEM algorithm

The unknown vector $\Theta = [\Theta_0, \Theta_1, \dots, \Theta_{L-1}]^T$ mentioned in the previous paragraph can be also modeled as a *random* quantity, when its joint pdf $f(\theta)$ is available. In this case the MAP estimate θ_{MAP} of Θ , given the observed data vector \mathbf{r} , can be evaluated as [36]

$$\theta_{\text{MAP}} = \arg \max_{\tilde{\theta}} M_{\mathbf{r}}(\tilde{\theta}), \quad (5)$$

where $M_{\mathbf{r}}(\tilde{\theta}) \doteq \log f(\mathbf{r}, \tilde{\theta})$. Solving (5) may be a formidable task for the same reasons previously illustrated for the ML problem (1). In principle, however, an improved estimate of Θ can be evaluated via the MAP approach since statistical information about channel uncertainty are exploited.

Since there is a strong analogy between the ML problem (1) and the MAP one (5), it is not surprising that an expectation-maximization procedure, dubbed *Bayesian EM* (BEM) [29, 37], for solving the latter, is available. The BEM algorithm evolves through the same iterative procedure as the EM, but with a different auxiliary function [29], namely,

$$\begin{aligned} Q_{\text{BEM}}(\theta, \tilde{\theta}) &= E_{\mathbf{C}} \{ M_{\mathbf{c}}(\theta) \mid \mathbf{R} = \mathbf{r}, \Theta = \tilde{\theta} \} \\ &= E \{ \log f(\mathbf{C}, \theta) \mid \mathbf{R} = \mathbf{r}, \Theta = \tilde{\theta} \} \\ &= \int_{\mathbf{s}_i} \log f(\mathbf{r}, \mathbf{i}, \theta) f(\mathbf{i} \mid \mathbf{r}, \tilde{\theta}) d\mathbf{i}. \end{aligned} \quad (6)$$

A clear relationship can be established between the BEM and the EM algorithms. In fact, factoring the pdf $f(\mathbf{r}, \mathbf{i}, \theta)$ as

$$f(\mathbf{r}, \mathbf{i}, \theta) = f(\mathbf{r}, \mathbf{i} \mid \theta) f(\theta) \quad (7)$$

and substituting (7) into (6) produces

$$Q_{\text{BEM}}(\theta, \tilde{\theta}) = Q_{\text{EM}}(\theta, \tilde{\theta}) + I(\theta), \quad (8)$$

where

$$I(\theta) \doteq \log f(\theta). \quad (9)$$

Equation (8) shows that the difference between $Q_{\text{BEM}}(\theta, \tilde{\theta})$ (6) and $Q_{\text{EM}}(\theta, \tilde{\theta})$ (3) is simply a *bias* term $I(\theta)$ (9) favoring the most likely values of Θ . It is worth noting that, if a priori information about Θ were unavailable and, consequently, a uniform pdf was selected for $f(\theta)$, the contribution from $I(\theta)$ would turn into a constant in (8), that is, it could be neglected. Therefore, the BEM encompasses the EM as a special case and, since the former benefits by the statistical information about Θ , it is expected to provide improved accuracy with respect to the latter. For the same reason, an increase in the speed of convergence and an improved robustness against the choice of the initial conditions could be offered by the BEM.

3. SISO DATA DETECTION IN THE PRESENCE OF PARAMETRIC UNCERTAINTY VIA THE BEM TECHNIQUE

In this section we show how the BEM technique can be employed to derive SISO algorithms for detecting digital signals transmitted over channels with parametric uncertainty and memory. A single-user transmission over a *single-input single-output* channel is considered for simplicity, but, as shown in the following section, the proposed approach can be extended to an arbitrary number of users and to a *multiple-input multiple-output* (MIMO) system without any substantial conceptual problem.

Here we assume that the k th component of the received data vector \mathbf{R} can be expressed as²

$$R_k = g_k(\mathbf{D}, \mathbf{A}) + N_k, \quad (10)$$

where $\mathbf{D} \doteq [D_0, D_1, \dots, D_{N-1}]^T$ is a vector of independent channel symbols belonging to a constellation $\Sigma = \{s_0, s_1, \dots, s_{M-1}\}$ of cardinality M and average energy E_s , $\mathbf{A} \doteq [A_0, A_1, \dots, A_{L-1}]^T$ is a vector of random channel parameters *independent* of \mathbf{D} and *with known statistical properties*, $\{N_k\}$ is an AWGN sequence with variance σ_N^2 , and $g_k(\cdot, \cdot)$ expresses the known functional dependence of the channel on both the transmitted symbols and its parametric uncertainty. In particular, we concentrate on *conditional finite memory* channels, that is, on random channels such that

$$g_k(\mathbf{D}, \mathbf{A}) = g_k(D_k, D_{k-1}, D_{k-2}, \dots, D_{k-L_c}, \mathbf{A}), \quad (11)$$

where L_c denotes the *channel memory*.

Our target is devising MAP SISO detection algorithms [18, 22], given the observed data $\mathbf{R} = \mathbf{r}$ and a statistically known parameter vector \mathbf{A} . In data detection problems involving the EM technique, two different choices have been usually suggested for the imputed data \mathbf{I} (see (2)) and the parameter vector Θ :

- (1) $\mathbf{I} = \mathbf{A}$ and $\Theta = \mathbf{D}$ [3];
- (2) $\mathbf{I} = \mathbf{D}$ and $\Theta = \mathbf{A}$ [6, 8, 29].

It is extremely important to comment now on the meaning and the consequences of these choices.

In the first case, both EM and BEM-based algorithms aim at producing *hard* estimates of the transmitted data. The only substantial difference between these two classes of strategies is that BEM allows to exploit the data statistics, that is, their APRPs, in the detection algorithm, since $I(\theta)$ in (8) turns into (see (9))

$$I(\theta) = I(\mathbf{D}) = \sum_{n=0}^{N-1} \log \Pr(d_n), \quad (12)$$

²Here we concentrate on detection algorithms processing one sample per channel symbol. The extension of the following ideas to multisampling detection is straightforward.

where $\Pr(d_n)$ denotes the probability of the event $\{D_n = d_n\}$. In other words, employing the EM (BEM) technique leads to *hard-in (soft-in) hard-output* detection algorithms.

In the second case, both EM- and BEM-based algorithms estimate the random parameters of the communication channel in a direct fashion. Nonetheless, they can be considered as SISO detectors, since they generate *soft* estimates (i.e., the APPs) of the transmitted data as a by-product of the estimation procedure and can also incorporate the data APRPs. BEM-based estimators, however, also make use of channel statistics, whereas EM-based estimators do not, that is, they operate in a *blind* fashion. Since blind detection techniques can be substantially outperformed by their counterparts exploiting channel statistics (see, e.g., [4, 38, 39]), this offers a strong motivation for preferring BEM-based strategies to EM-based ones when such statistical information are available. To further clarify these ideas, we derive now the BEM estimator of $\Theta = \mathbf{A}$, given $\mathbf{I} = \mathbf{D}$. In (6) the joint pdf $f(\mathbf{r}, \mathbf{i}, \theta)$ can be factored as

$$f(\mathbf{r}, \mathbf{i}, \theta) = f(\mathbf{r}, \mathbf{d}, \mathbf{a}) = f(\mathbf{r}|\mathbf{d}, \mathbf{a})f(\mathbf{d})f(\mathbf{a}) \quad (13)$$

as the data \mathbf{D} are independent of the channel parameters \mathbf{A} . Here

$$f(\mathbf{d}) = \sum_{\mathbf{d}_l \in \Lambda} \Pr(\mathbf{d}_l) \delta_N(\mathbf{d} - \mathbf{d}_l). \quad (14)$$

Λ is the set of all the M^N possible data sequences of length N , $\delta_N(\cdot)$ is the N -dimensional Dirac delta function, and $\Pr(\mathbf{d}) = \prod_{n=0}^{N-1} \Pr(d_n)$ denotes the APRP of the channel symbol vector \mathbf{d} . If we define the *channel state* vector $\Delta_k \doteq (d_{k-1}, d_{k-2}, \dots, d_{k-L_c})$, the conditional pdf $f(\mathbf{r}|\mathbf{d}, \mathbf{a})$ in (13) can be expressed as

$$f(\mathbf{r}|\mathbf{d}, \mathbf{a}) = \prod_{k=0}^{N-1} \frac{1}{\pi \sigma_N^2} \exp \left[- \frac{|r_k - g_k(d_k, \Delta_k, \mathbf{a})|^2}{\sigma_N^2} \right] \quad (15)$$

since the k th sample r_k depends on \mathbf{d} through the couple (d_k, Δ_k) only, and the random variables $\{R_k\}$, conditioned on \mathbf{D} and \mathbf{A} , are independent. Moreover, the conditional pdf $f(\mathbf{i}|\mathbf{r}, \theta)$ in (6) is given by

$$f(\mathbf{i}|\mathbf{r}, \theta) = f(\mathbf{d}|\mathbf{r}, \bar{\mathbf{a}}) = \sum_{\mathbf{d}_l \in \Lambda} \Pr(\mathbf{d}_l|\mathbf{r}, \bar{\mathbf{a}}) \delta_N(\mathbf{d} - \mathbf{d}_l), \quad (16)$$

where $\Pr(\mathbf{d}_l|\mathbf{r}, \bar{\mathbf{a}})$ is the probability of the event $\{\mathbf{d} = \mathbf{d}_l\}$, given $\mathbf{R} = \mathbf{r}$ and $\mathbf{A} = \bar{\mathbf{a}}$. Substituting (14) and (15) into (13) and substituting (13) and (16) into (6) and dropping the un-relevant terms produces, after some manipulations,

$$\begin{aligned} Q_{\text{BEM}}(\mathbf{a}, \bar{\mathbf{a}}) &= -\frac{1}{\sigma_N^2} \sum_{k=0}^{N-1} \sum_{\Delta_k \in \Pi} \sum_{d_k \in \Sigma} \Pr(d_k, \Delta_k|\mathbf{r}, \bar{\mathbf{a}}) |r_k - g_k(d_k, \Delta_k, \mathbf{a})|^2 \\ &\quad + \log f(\mathbf{a}), \end{aligned} \quad (17)$$

where Π denotes the set of M^{L_c} possible channel state vectors. We define now the estimate vector $\mathbf{a}[k] \doteq [a_0[k], a_1[k], \dots, a_{L-1}[k]]^T$ generated, at the k th iteration, by the BEM estimation algorithm based on $Q_{\text{BEM}}(\mathbf{a}, \bar{\mathbf{a}})$ (17). Such an algorithm operates as follows. First, $Q(\mathbf{a}, \mathbf{a}[k])$ is evaluated (E step). The next estimate $\mathbf{a}[k+1]$ corresponds to the maximum of $Q(\mathbf{a}, \mathbf{a}[k])$ with respect to \mathbf{a} . Then, taking the gradient of (17) with respect to \mathbf{a} and setting it to zero produces the recursive relation

$$\begin{aligned} \frac{1}{\sigma_N^2} \sum_{k=0}^{N-1} \sum_{\Delta_k \in \Pi} \sum_{d_k \in \Sigma} \Pr(d_k, \Delta_k|\mathbf{r}, \mathbf{a}[k]) &\times 2 \operatorname{Re} \{ (g_k^*(d_k, \Delta_k, \mathbf{a}) - r_k^*) \\ &\quad \times \nabla_{\mathbf{a}} g_k(d_k, \Delta_k, \mathbf{a}) \}_{\mathbf{a}=\mathbf{a}[k+1]} \\ &- \frac{1}{f(\mathbf{a})} \nabla_{\mathbf{a}} f(\mathbf{a}) \Big|_{\mathbf{a}=\mathbf{a}[k+1]} = 0 \end{aligned} \quad (18)$$

expressing a set of nonlinear equations for evaluating $\mathbf{a}[k+1]$, given $\mathbf{a}[k]$ (M-step). It is worth noting that complexity of solving (18) depends on the type of functional dependence of $g_k(\cdot)$ on \mathbf{a} and on the inner structure of $\log f(\mathbf{a})$.

We us now explain why the estimator based on (18) can be also interpreted as a SISO algorithm. First of all, we note that the contribution from $\Pr(\mathbf{d}_l)$ (coming from (14)), being independent of \mathbf{a} , has been dropped in $Q_{\text{BEM}}(\mathbf{a}, \bar{\mathbf{a}})$ (17). The contribution from the APRPs of the channel symbols, however, has not really disappeared since such probabilities are used in the evaluation of the APPs $\{P(d_k, \Delta_k|\mathbf{r}, \bar{\mathbf{a}})\}$. This means that, in its $(k+1)$ th iteration, the BEM-based estimation algorithm requires the evaluation of the new APPs starting from the available APRPs and the last estimate $\mathbf{a}[k]$ of channel parameters. Generally speaking, on channels with memory, these APPs can be evaluated by means of a *forward-backward* recursive procedure operating on the trellis diagram of the channel states [6, 20, 40] and which can be derived as follows. To begin, we note that the couple (Δ_k, d_k) uniquely identifies a transition (Δ_k, Δ_{k+1}) in the channel state, so that $P(d_k, \Delta_k|\mathbf{r}, \bar{\mathbf{a}}) = P(\Delta_k, \Delta_{k+1}|\mathbf{r}, \bar{\mathbf{a}})$. Applying the Bayes' rule to the evaluation of $P(\Delta_k, \Delta_{k+1}|\mathbf{r}, \bar{\mathbf{a}})$ gives

$$\begin{aligned} P(\Delta_k, \Delta_{k+1}|\mathbf{r}, \bar{\mathbf{a}}) &= \frac{f(\mathbf{r}, \Delta_k, \Delta_{k+1}|\bar{\mathbf{a}})}{f(\mathbf{r}|\bar{\mathbf{a}})} \\ &= \frac{f(\mathbf{r}, \Delta_k, \Delta_{k+1}|\bar{\mathbf{a}})}{\sum_{\bar{\Delta}_k, \bar{\Delta}_{k+1} \in \Pi} f(\mathbf{r}, \bar{\Delta}_k, \bar{\Delta}_{k+1}|\bar{\mathbf{a}})}. \end{aligned} \quad (19)$$

Following [6, 20, 40] it can be proved that

$$\begin{aligned} f(\mathbf{r}, \Delta_k, \Delta_{k+1}|\bar{\mathbf{a}}) &= \alpha_k(\Delta_k) f(r_k|\Delta_k, \Delta_{k+1}, \bar{\mathbf{a}}) \beta_{k+1}(\Delta_{k+1}) \Pr(\Delta_{k+1}|\Delta_k) \end{aligned} \quad (20)$$

where $\mathbf{r}_j^T \doteq [r_j, r_{j+1}, \dots, r_l]^T$, $\alpha_k(\Delta_k) \doteq f(\mathbf{r}_0^{k-1}, \Delta_k|\bar{\mathbf{a}})$, $\beta_{k+1}(\Delta_{k+1}) \doteq f(\mathbf{r}_{k+1}^{N-1}|\Delta_{k+1}, \bar{\mathbf{a}})$, $\Pr(\Delta_{k+1}|\Delta_k)$ is the probability of the state transition $\Delta_k \rightarrow \Delta_{k+1}$, and $f(r_k|\Delta_k, \Delta_{k+1}, \bar{\mathbf{a}}) = [\pi \sigma_N^2]^{-1} \exp[-|r_k - g_k(d_k, \Delta_k, \bar{\mathbf{a}})|^2/\sigma_N^2]$. The quantities $\{\alpha_k(\Delta_k)\}$, and $\{\beta_{k+1}(\Delta_{k+1})\}$ are evaluated by means of the

following recursive equations:

$$\alpha_k(\Delta_k) = \sum_{\tilde{\Delta}_{k-1} \in S(\tilde{\Delta}_{k-1}, \Delta_k)} \alpha_{k-1}(\tilde{\Delta}_{k-1}) f(r_{k-1} | \Delta_k, \tilde{\Delta}_{k-1}, \tilde{\mathbf{a}}) \times \Pr(\Delta_k | \tilde{\Delta}_{k-1}), \quad (21)$$

$$\beta_{k+1}(\Delta_{k+1}) = \sum_{\tilde{\Delta}_{k+2} \in S(\Delta_{k+1}, \tilde{\Delta}_{k+2})} \beta_{k+2}(\tilde{\Delta}_{k+2}) f(r_{k+1} | \Delta_{k+1}, \tilde{\Delta}_{k+2}, \tilde{\mathbf{a}}) \times \Pr(\tilde{\Delta}_{k+2} | \Delta_{k+1}), \quad (22)$$

where $S(\Delta_i, \Delta_j)$ is the subset of states Δ_i such that the transition $\Delta_i \rightarrow \Delta_j$ is admissible. The initial conditions $\{\alpha_0(\Delta_0) = \Pr(\Delta_0); \Delta_0 \in \Pi\}$ and $\{\beta_N(\Delta_N) = 1; \Delta_N \in \Pi\}$ need to be fixed before starting the forward (21) and the backward iterations (22), respectively.

After K iterations the BEM algorithm stops, producing a final estimate $\mathbf{a}_{\text{BEM}} = \mathbf{a}[K]$ and the APPs $\{\Pr(d_k, \Delta_k | \mathbf{r}, \mathbf{a}_{\text{BEM}})\}$ of the channel symbols. The symbol APPs $\{\Pr(d_k | \mathbf{r}, \mathbf{a}_{\text{BEM}})\}$ can be easily derived from these quantities as

$$\Pr(d_k | \mathbf{r}, \mathbf{a}_{\text{BEM}}) = \sum_{\Delta_k \in \Omega(d_k)} \Pr(d_k, \Delta_k | \mathbf{r}, \mathbf{a}_{\text{BEM}}), \quad (23)$$

where $\Omega(d_k)$ denotes the subset of all the state transitions $\Delta_k \rightarrow \Delta_{k+1}$ labeled by the channel symbol d_k . Then, decisions on the channel symbols can be taken according to the MAP decision strategy [6]

$$\hat{d}_k = \arg \max_{d_k} \Pr(d_k | \mathbf{r}, \mathbf{a}_{\text{BEM}}) \quad (24)$$

with $k = 0, 1, \dots, N-1$. Alternatively, if channel coding is employed, the APPs $\{\Pr(d_k | \mathbf{r}, \mathbf{a}_{\text{BEM}})\}$ can be delivered to soft decoding stages (see, e.g., [30, 31]) to improve the error performance of a digital receiver (see Section 4.4.3).

Finally, we note that substantial simplifications of the BEM-based procedure based on (18) can be found when the communication channel does not have memory, that is, $L_c = 1$, since in this case the forward-backward procedure is no more required. Specific examples of BEM-based algorithms for memoryless channels can be found in [30, 31, 32], where frequency-flat fading and phase jitter are considered as channel impairments.

4. SPECIFIC APPLICATIONS

In this section, three specific applications of the BEM strategy are briefly illustrated. In particular, SISO detectors are developed for the following three different scenarios: (1) a synchronous multiuser CDMA system; (2) a single-carrier system employing an orthogonal *space-time block code* (STBC); (3) an *orthogonal frequency division multiplexing* (OFDM) system using an orthogonal STBC on a subcarrier-by-subcarrier basis. For each scenario we provide a brief introduction citing a set of key references about the specific problem, a description of the signal and channel models, an analysis of the corresponding BEM-based SISO algorithm, and some numerical results.

4.1. Multiuser detection of synchronous DSSS signals over frequency-flat fading channels

4.1.1. Introduction

One of the most challenging problems in receiver design for DSSS-CDMA systems is the derivation of reduced-complexity multiuser detectors. This is due to the fact that the complexity of optimal multiuser detection grows exponentially with the number of users [41]. One of the interesting applications of the EM technique has been the derivation of multiuser detectors for synchronous DS-CDMA systems operating over frequency-flat fading channels [42, 43, 44]. However, all the solutions proposed in the cited papers produce hard estimates of the data. A BEM-based soft detector is illustrated in the following.

4.1.2. Channel and signal models

Multiuser detection on synchronous uplink of a J -user DS-CDMA system is considered here. In the presence of slow frequency-flat fading the output of the receiver matched filter bank in the l th symbol interval can be expressed as [42, 43]

$$\mathbf{Z}(l) = \mathbf{R}\mathbf{B}[l]\mathbf{A}[l] + \mathbf{N}[l], \quad (25)$$

where $\mathbf{Z}[l] \doteq [Z_1[l], \dots, Z_J[l]]^T$, $\mathbf{B}[l] \doteq \text{diag}(B_1[l], \dots, B_J[l])$ is the channel symbol matrix, $\mathbf{A}[l] \doteq [A_1[l], \dots, A_J[l]]^T$ is the channel fading vector, $\mathbf{R} = [r_{mn}]$ ($m, n = 1, 2, \dots, J$) is the $J \times J$ matrix of signature cross-correlations, and $\mathbf{N}[l]$ is a complex Gaussian noise vector having zero mean and covariance matrix $\sigma_w^2 \mathbf{R}$, with $\sigma_w^2 = 2N_0$. Here $B_j[l] \in \{\pm \sqrt{2E_{b,j}}\}$ ($E_{b,j}$ is the average transmitted energy per bit) is the BPSK channel symbol transmitted by the j th user in the l th signaling interval, $A_j[l]$ is the fading distortion affecting $B_j[l]$, and $r_{mn} = \int_0^{T_s} p_m(t)p_n(t)dt$ ($m, n = 1, 2, \dots, J$), where T_s is the symbol interval and $p_n(t)$ is the signature waveform³ of the n th user. In the following it is assumed that the J fading processes $\{A_j[l]\}$ are independent, identically distributed and zero mean Gaussian (Rayleigh fading) with autocorrelation function $R_a[m]$ ($R_a[0] = 1$).

If \mathbf{R} is positive definite, it can be Cholesky factored as $\mathbf{R} = \mathbf{\Gamma}^H \mathbf{\Gamma}$, where $\mathbf{\Gamma}$ is a lower triangular matrix. Then, pre-multiplying $\mathbf{Z}(l)$ (25) by $(\mathbf{\Gamma}^H)^{-1}$ produces [43]

$$\mathbf{Y}[l] = [Y_1[l], \dots, Y_J[l]]^T \doteq (\mathbf{\Gamma}^H)^{-1} \mathbf{Z}[l] = \mathbf{C}\mathbf{B}[l]\mathbf{A}[l] + \mathbf{W}[l]. \quad (26)$$

Here the noise vector $\mathbf{W}[l] = [W_1[l], \dots, W_J[l]]^T$ is white Gaussian since its covariance matrix is $\sigma_w^2 \mathbf{I}_J$ (\mathbf{I}_J is the $J \times J$ identity matrix).

Extending the one-shot model (26) to an observation interval of N consecutive symbols (with $l = 1, \dots, N$) yields

$$\mathbf{Y} = \text{diag}(\mathbf{\Gamma})\mathbf{B}\mathbf{A} + \mathbf{W}, \quad (27)$$

³We assume that its support is the interval $[0, T_s]$.

where $\mathbf{Y} \doteq [\mathbf{Y}^T[1], \dots, \mathbf{Y}^T[L]]^T$, $\mathbf{A} \doteq [\mathbf{A}^T[1], \dots, \mathbf{A}^T[L]]^T$, $\mathbf{W} \doteq [\mathbf{W}^T[1], \dots, \mathbf{W}^T[L]]^T$, and $\mathbf{B} \doteq \text{diag}(\mathbf{B}[l], l = 1, 2, \dots, L)$ is an $NJ \times NJ$ block matrix having $\{\mathbf{B}[l]\}$ on its main diagonal. Following [45], we decompose the noise vector $\mathbf{W}[l]$ as $\sum_{j=1}^J \mathbf{W}_j[l]$, where $\{\mathbf{W}_j[l], l = 1, 2, \dots, N\}$ are independent Gaussian vectors having zero mean and covariance matrix $E\{\mathbf{W}_j[l]\mathbf{W}_j^H[l]\} = \sigma_{w,j}^2 \mathbf{I}_J$, with $\sigma_{w,j}^2 = \beta_j \sigma_w^2$. Here $\{\beta_j\}$ are real positive coefficients satisfying the constraint $\sum_{j=1}^J \beta_j = 1$ in order to ensure statistical equivalence. Then, $\mathbf{Y}[l]$ (26) can be decomposed as $\sum_{j=1}^J \mathbf{U}_j[l]$, where

$$\mathbf{U}_j[l] = [U_1[l], \dots, U_J[l]]^T \doteq \mathbf{\Gamma}_j b_j[l] a_j[l] + \mathbf{W}_j[l] \quad (28)$$

and $\mathbf{\Gamma}_j$ is the j th column ($j = 1, 2, \dots, J$) of $\mathbf{\Gamma}$.

4.1.3. The CDMA-BEM algorithm

We define now the vector $\mathbf{U} \doteq [\mathbf{U}^T[1], \dots, \mathbf{U}^T[N]]^T$, with $\mathbf{U}[l] \doteq [U_1[l], \dots, U_J[l]]^T$. Then, in applying the BEM technique, we select $\mathbf{C} = \{\mathbf{B}, \mathbf{U}\}$ and $\mathbf{\Theta} = \mathbf{A}$ (see Section 2.2) as the complete and parameter vectors, respectively. This leads to the auxiliary function (further analytical details are available in [33])

$$\begin{aligned} Q(\mathbf{a}, \tilde{\mathbf{a}}) &= \sum_{j=1}^J \sum_{l=1}^N \frac{1}{\sigma_{w,j}^2} \sum_{\tilde{\mathbf{b}}[l] \in \Omega} 2 \text{Re} \left\{ \mathbf{\Gamma}_j^H \hat{\mathbf{u}}_j[l] a_j^* [l] \tilde{b}_j^* [l] \right\} \\ &\quad \times \Pr(\tilde{\mathbf{b}}[l] | \mathbf{y}, \tilde{\mathbf{a}}) \\ &\quad - \sum_{j=1}^J \sum_{l=1}^N \frac{2E_{b,j}}{\sigma_{w,j}^2} \|a_j[l]\|^2 - \sum_{j=1}^J \mathbf{a}_j^H \mathbf{C}_A^{-1} \mathbf{a}_j, \end{aligned} \quad (29)$$

where $\tilde{b}_j[l]$ is the j th component of $\tilde{\mathbf{b}}[l] = [\tilde{b}_1[l], \tilde{b}_2[l], \dots, \tilde{b}_J[l]]^T$, $\Pr(\tilde{\mathbf{b}}[l] | \mathbf{y}, \tilde{\mathbf{a}})$ is the probability of the event $\{\mathbf{b}[l] = \tilde{\mathbf{b}}[l]\}$ conditioned on $\mathbf{Y} = \mathbf{y}$ and $\mathbf{A} = \tilde{\mathbf{a}}$, and

$$\begin{aligned} \hat{\mathbf{u}}_j[l] &\doteq E\{\mathbf{u}_j[l] | \mathbf{b}[l] = \tilde{\mathbf{b}}[l], \mathbf{y}, \tilde{\mathbf{a}}\} \\ &= \mathbf{\Gamma}_j \tilde{a}_j[l] \tilde{b}_j[l] + \beta_j \left(\mathbf{y}[l] - \sum_{i=1}^J \mathbf{\Gamma}_i \tilde{a}_i[l] \tilde{b}_i[l] \right). \end{aligned} \quad (30)$$

Given $Q(\mathbf{a}, \tilde{\mathbf{a}})$ (29), the expectation-maximization can be expressed as follows [33]. Given the fading estimates $\mathbf{a}_j^k = [a_j^k[1], \dots, a_j^k[N]]^T$, with $j = 1, 2, \dots, J$, at the k th iteration, the new estimate \mathbf{a}_j^{k+1} is evaluated as

$$\mathbf{a}_j^{k+1} = (\mathbf{P}_j)^{-1} \mathbf{v}_j^k, \quad (31)$$

where

$$\mathbf{P}_j \doteq 2E_{b,j} \mathbf{I}_L + \sigma_{w,j}^2 \mathbf{C}_A^{-1} \quad (32)$$

and $\mathbf{v}_j^k = [v_j^k[1], v_j^k[2], \dots, v_j^k[L]]^T$, with

$$v_j^k[l] \doteq \sum_{\tilde{\mathbf{b}}[l] \in \Omega} \mathbf{\Gamma}_j^H \hat{\mathbf{u}}_j[l] \tilde{b}_j^* [l] \Pr(\tilde{\mathbf{b}}[l] | \mathbf{y}, \tilde{\mathbf{a}}^k). \quad (33)$$

It is worth noting that the inverse of \mathbf{P}_j (32) does not need to be recomputed as long as the channel statistics do not

change, and that such matrix depends on j , that is, on the considered user, through $E_{b,j}$ and $\sigma_{w,j}^2$ only. The APPs $\Pr(\tilde{\mathbf{b}}[l] | \mathbf{y}, \mathbf{a}^k)$ in (33) can be evaluated as

$$\begin{aligned} \Pr(\mathbf{b}[l] = \tilde{\mathbf{b}}[l] | \mathbf{y}, \mathbf{a}^k) &= \frac{f(\mathbf{y}[l] | \tilde{\mathbf{b}}[l], \mathbf{a}^k[l]) \Pr(\tilde{\mathbf{b}}[l])}{\sum_{\tilde{\mathbf{b}}[l] \in \Omega} f(\mathbf{y}[l] | \tilde{\mathbf{b}}[l], \mathbf{a}^k[l]) \Pr(\tilde{\mathbf{b}}[l])}, \end{aligned} \quad (34)$$

where

$$\begin{aligned} f(\mathbf{y}[l] | \mathbf{b}[l], \mathbf{a}[l]) &= \frac{1}{(\pi \sigma_w^2)^J} \exp \left(\frac{-\|\mathbf{y}[l] - \mathbf{\Gamma} \mathbf{B}[l] \mathbf{A}[l]\|^2}{\sigma_w^2} \right). \end{aligned} \quad (35)$$

Moreover, the data APRP $\Pr(\mathbf{b}[l])$ of (34) can be expressed as

$$\Pr(\mathbf{b}[l]) = \prod_{j=1}^J \Pr(b_j[l]) \quad (36)$$

for the independence of the J users.

After K iterations the BEM-based algorithm based on (31)–(36) (dubbed CDMA-BEM in the following) stops producing a channel estimate $\mathbf{a}_{\text{BEM}} = \mathbf{a}^{(K+1)}$ and the data APPs $\{P(b_j[l] | \mathbf{y}, \mathbf{a}_{\text{BEM}})\}$. Then, data decisions can be taken according to a MAP decision strategy (see (24)) or, if channel coding is used, can be delivered to soft decoding stages.

4.1.4. Numerical results

Computer simulations have been carried out in order to assess the *bit error rate* (BER) performance of the CDMA-BEM multiuser detector. In the following it is always assumed that (1) the autocovariance function of the fading process $\{A_j[l]\}$ (with $j = 1, \dots, J$) is $R_a[m] = J_0(2\pi m B_D T_s)$ (Clarke's fading [46]), where $J_0(x)$ is the zeroth-order Bessel function of the first kind and B_D is the fading Doppler bandwidth; (2) each user continuously transmits packets containing $N = 14$ consecutive symbols; (3) each packet consists of 12 information symbols and is preceded by a couple of pilot symbols (used for channel estimation), so that the pilot symbol rate is $R_p = 1/7$; (4) Wiener filtering techniques are exploited at the receiver side in order to evaluate the channel estimates needed for the initialization of the CDMA-BEM [29]; (5) the CDMA-BEM processes a block of $(2 \cdot N + 2) = 30$ received signal samples corresponding to 2 consecutive packets (plus the first two samples of the next packet) and carries out $K = 3$ iterations; (6) the *signal-to-noise ratio* for the j th user (SNR_j) is defined as $E_{b,j}/N_0$, where $E_{b,j}$ is the average received energy per bit for the j th user and $N_0/2$ is the noise two-sided power spectral density; (7) the receiver is provided with an ideal estimate of the SNR for all the active users so that the parameters $\{\beta_j, j = 1, \dots, J\}$ can be selected as [42]

$$\beta_j = \frac{E_{b,j}}{\sum_{i=1}^J E_{b,i}}. \quad (37)$$

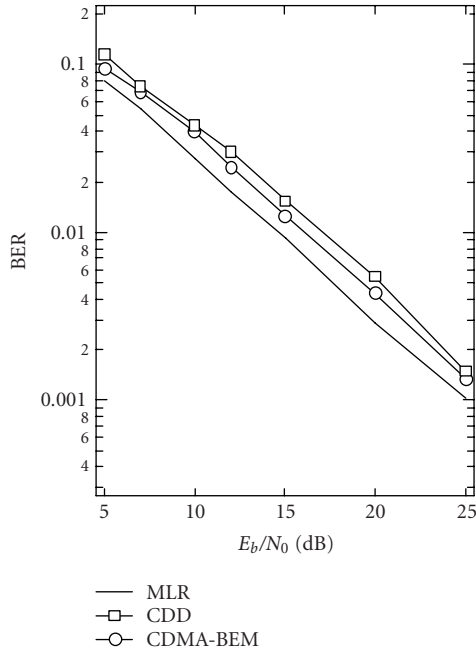


FIGURE 1: BER performance of the CDMA-BEM algorithm with $B_D T_s = 5 \cdot 10^{-3}$, $J = 4$, $N = 14$, and $K = 3$. The BER performance of the MLR and CDD is also shown for comparison.

In the following, we consider a four-user scenario ($J = 4$) characterized by the matrix of signature cross-correlations [43]:

$$\mathbf{R}_4 = \frac{1}{7} \begin{bmatrix} 7 & -1 & 3 & 3 \\ -1 & 7 & 3 & -1 \\ 3 & 3 & 7 & -1 \\ 3 & -1 & -1 & 7 \end{bmatrix}. \quad (38)$$

The BER performance of the CDMA-BEM receiver is illustrated in Figure 1. Here it is assumed that the normalized Doppler bandwidth is $B_D T_s = 5 \cdot 10^{-3}$ and that all the users have the same SNR. In this figure the performance of the *maximum likelihood receiver* (MLR) endowed with ideal *channel state information* (CSI) and that of the *coherent decorrelator detector* (CDD) [47] are also shown for comparison. It is interesting to note that, in these scenarios, the CDMA-BEM almost achieves the same performance of the MLR and outperforms the CDD by about 1.5 dB in SNR.

Figure 2 shows the performance of CDMA-BEM versus the normalized Doppler bandwidth for $B_D T_s \in (5 \cdot 10^{-3}, 5 \cdot 10^{-2})$, under the assumption that $\text{SNR}_j = 15, 20, 25$ dB for $j = 1, \dots, 4$. The error performance of the proposed algorithm slightly worsens as the Doppler bandwidth increases because of the poorer quality of the initial channel estimates.

Finally, the near-far resistance of the CDMA-BEM receiver is illustrated in Figure 3. The SNR of the first user (SNR_1) is set to 20 dB, whereas the other three SNRs (SNR_j , $j = 2, 3, 4$) are equal and vary in the range (5, 25) dB.

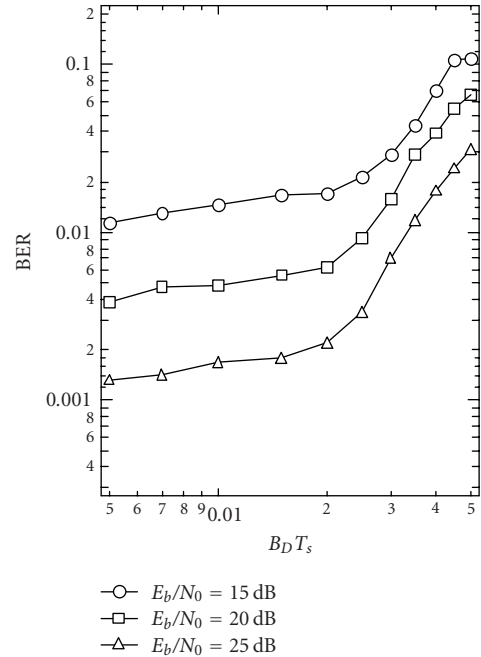


FIGURE 2: BER performance of the CDMA-BEM algorithm versus $B_D T_s$, $J = 4$, $E_{b,k}/N_0 = 20$ dB, $N = 14$, and $K = 3$.

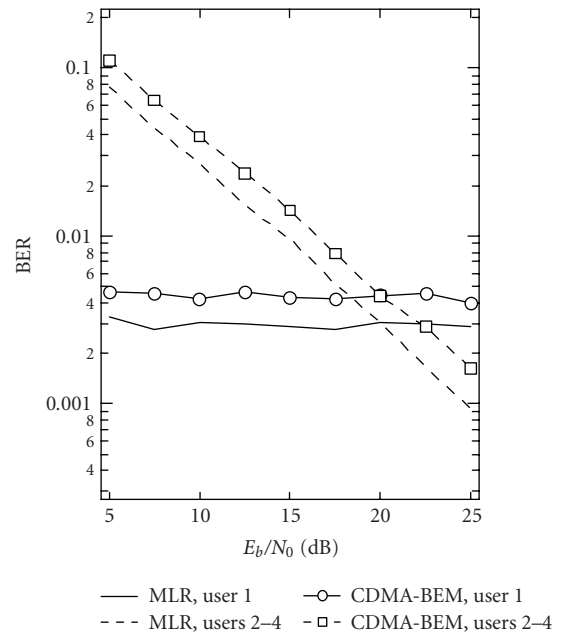


FIGURE 3: Near-far resistance of the CDMA-BEM algorithm. $J = 4$, $\text{SNR}_1 = 20$ dB, $\text{SNR}_k \in (5, 25)$ dB ($k = 2, 3, 4$), and $B_D T_s = 5 \cdot 10^{-3}$.

The performance of the MLR is also shown for comparison. These results show that, in this case, the CDMA-BEM exhibits a performance which is substantially independent of the energies of the interfering users.

4.2. SISO detection of space-time block coded signals

4.2.1. Introduction

In the last years it has been shown that the information capacity of wireless communication systems can be substantially increased by employing antenna arrays [48], jointly with proper coding [49] and signal processing techniques [50]. One of the most promising results in this research area has been the development of new block and trellis codes for multiple antennas, known as *space-time codes* (STCs) [49, 51]. Such codes provide significant diversity gains without bandwidth expansion. Exact knowledge of the CSI is often assumed in devising space-time decoding algorithms even if channel estimation may represent a serious problem, especially in time-varying environments [52]. EM-based hard detectors for STCs have been derived in [52, 53, 54]. In this section a BEM-based soft detector for orthogonal STBCs is illustrated.

4.2.2. Signal and channel models

Here we focus on a space-time block coded system employing N_T transmit and N_R receive antennas [49]. The set of channel symbols transmitted during the n th block⁴ is denoted by the $L \times N_T$ matrix $\mathbf{S}[n] = [s_{l,i}[n]]$ (with $l = 1, 2, \dots, L$, $i = 1, 2, \dots, N_T$), where L is the overall duration of the block in channel symbols and $s_{l,i}[n]$ is the channel symbol feeding the i th antenna in the symbol interval $(l + nL)$.

In the following we assume that the multiple channels involved in the communication system are (a) affected by frequency-flat Rayleigh fading and (b) *quasi-static*, that is, channel variations within each block are negligible, whereas changes from block to block are taken into account. Then the path gain $a_{i,j}[n]$ (with $i = 1, 2, \dots, N_T$ and $j = 1, 2, \dots, N_R$) from the i th transmit antenna to the j th receive antenna during the n th block is a complex Gaussian random process having zero mean and correlation function $R_a[m] \doteq E\{a_{i,j}[n+m]a_{i,j}^*[n]\}$ (with $R_a[0] = 1$). Moreover, the gain processes $\{a_{i,j}[n]\}$ are *independent* (rich scatterer environment).

Let $r_{l,j}[n]$ denote the received signal sample taken at the output of the j th receive antenna in the $(l + nL)$ th symbol interval, with $j = 1, \dots, N_R$ and $l = 1, \dots, L$. Then the $L \times N_R$ received signal matrix $\mathbf{R}[n] = [r_{l,j}[n]]$ is given by [52]

$$\mathbf{R}[n] = \mathbf{S}[n]\mathbf{A}[n] + \mathbf{W}[n]. \quad (39)$$

Here $\mathbf{S}[n] \in \Omega$, where $\Omega = \{\mathbf{S}_m, m = 1, \dots, M\}$ is an M -ary alphabet of unitary matrices (i.e., $(\mathbf{S}_m)^H \mathbf{S}_m = \mathbf{I}_{N_T}$, where \mathbf{I}_n is the $n \times n$ identity matrix) [49, 51]. Moreover $\mathbf{A}[n] = [a_{i,j}[n]]$ and $\mathbf{W}[n] = [w_{l,j}[n]]$ are the $N_T \times N_R$ fading matrix and the $L \times N_R$ noise matrix, respectively. The elements $\{w_{l,j}[n]\}$ of $\mathbf{W}[n]$ are independent Gaussian random variables, all having zero mean and variance $\sigma_w^2 = 2N_0$.

⁴Throughout the section, the parameter n denotes the block index, whereas k specifies the location of a channel symbol within each block.

A set of N consecutive vectors (39) (with $n = 0, \dots, N - 1$) can be grouped as $\mathbf{R} \doteq [\mathbf{R}^H[0], \mathbf{R}^H[1], \dots, \mathbf{R}^H[N - 1]]^H$ ($(\mathbf{A})^T$ and $(\mathbf{A})^H$ denote transpose and conjugated transpose of \mathbf{A} , resp.), with

$$\mathbf{R} = \mathbf{D}(\mathbf{S})\mathbf{A} + \mathbf{W}, \quad (40)$$

where $\mathbf{A} \doteq [\mathbf{A}^H[0], \mathbf{A}^H[1], \dots, \mathbf{A}^H[N - 1]]^H$ and $\mathbf{W} \doteq [\mathbf{W}^H[0], \mathbf{W}^H[1], \dots, \mathbf{W}^H[N - 1]]^H$, respectively, and $\mathbf{D}(\mathbf{S}) \doteq \text{diag}\{\mathbf{S}[0], \mathbf{S}[1], \dots, \mathbf{S}[N - 1]\}$.

4.2.3. A BEM-based SISO algorithm for space-time block coded systems

Following the same indications illustrated in the previous application, we set $\Theta = \mathbf{A}$ and $\mathbf{C} = \{\mathbf{R}, \mathbf{S}\}$ in applying the BEM technique. Then the auxiliary function is (analytical details can be found in [55])

$$Q(\mathbf{A}, \tilde{\mathbf{A}}) = - \sum_{j=1}^{N_R} \mathbf{A}_j^H [\mathbf{C}_A^{-1} + \left(\frac{1}{\sigma_w^2}\right) \mathbf{I}_{N_{NT}}] \mathbf{A}_j - \left(\frac{2}{\sigma_w^2}\right) \text{Re}\{\tilde{\mathbf{V}}_j^H \mathbf{A}_j\}, \quad (41)$$

where \mathbf{A}_j is the j th column of \mathbf{A} , $\mathbf{C}_A \doteq E\{\mathbf{A}_j \mathbf{A}_j^H\}$ is a fading covariance matrix, and $\tilde{\mathbf{V}}_j$ is the j th column of the matrix

$$\tilde{\mathbf{V}} \doteq \mathbf{D}^H(\tilde{\mathbf{S}})\mathbf{R} \quad (42)$$

with $\tilde{\mathbf{S}} = \{\tilde{\mathbf{S}}[n], n = 0, 1, \dots, N - 1\}$. Here

$$\tilde{\mathbf{S}}[n] = \sum_{\mathbf{S}_m \in \Omega} \mathbf{S}_m \Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \tilde{\mathbf{A}}), \quad (43)$$

where $\Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \tilde{\mathbf{A}})$ is the APP of the event $\{\mathbf{S}[n] = \mathbf{S}_m\}$, given \mathbf{R} and $\mathbf{A} = \tilde{\mathbf{A}}$. Starting from (41), the following BEM-based recursive channel estimator can be derived. Given the channel estimate $\mathbf{A}^{(k)}$ at the k th iteration, the next estimate $\mathbf{A}^{(k+1)}$ is evaluated as

$$\mathbf{A}_j^{(k+1)} = [\mathbf{P}]^{-1} \mathbf{V}_j^{(k)}, \quad (44)$$

where $\mathbf{P} \doteq \mathbf{I}_{N_{NT}} + \sigma_w^2 \mathbf{C}_A^{-1}$. The APPs $\{\Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \tilde{\mathbf{A}})\}$ needed for the evaluation of (42) can be computed using the Bayes formula

$$\Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \tilde{\mathbf{A}}) = \frac{f(\mathbf{R}[n] | \mathbf{S}_m, \tilde{\mathbf{A}}[n]) \Pr(\mathbf{S}_m)}{\sum_{\tilde{\mathbf{S}}_m \in \Omega} f(\mathbf{R}[n] | \tilde{\mathbf{S}}_m, \tilde{\mathbf{A}}[n]) \Pr(\tilde{\mathbf{S}}_m)}, \quad (45)$$

where $\Pr(\mathbf{S}_m)$ is the probability of the event $\{\mathbf{S}[n] = \mathbf{S}_m\}$, and

$$f(\mathbf{R}[n] | \mathbf{S}_m, \tilde{\mathbf{A}}[n]) = \frac{1}{\det(\pi \sigma_w^2 \mathbf{I}_L)^{N_R}} \exp\left[-\frac{h(\mathbf{R}[n], \mathbf{S}_m, \tilde{\mathbf{A}}[n])}{\sigma_w^2}\right] \quad (46)$$

with $h(\mathbf{R}[n], \mathbf{S}_m, \tilde{\mathbf{A}}[n]) \doteq \text{tr}\{(\mathbf{R}[n] - \mathbf{S}_m \tilde{\mathbf{A}}[n])^H (\mathbf{R}[n] - \mathbf{S}_m \tilde{\mathbf{A}}[n])\}$.

It is important to note that (a) \mathbf{P} does not depend on the index of the receive antenna; (b) the inverse of \mathbf{P} does not need to be recomputed as long as the channel statistics do not change; (c) (44) can be simplified factoring \mathbf{C}_A as

$$\mathbf{C}_A = \tilde{\mathbf{C}}_a \otimes \mathbf{I}_{N_T}, \quad (47)$$

where $\tilde{\mathbf{C}}_a$ is the covariance matrix of the vector $\mathbf{a}_{i,j} = [a_{i,j}[0], a_{i,j}[1], \dots, a_{i,j}[N-1]]^T$ and \otimes is the Kronecker product, so that $\mathbf{P} = (\mathbf{I}_N + \sigma_w^2 \tilde{\mathbf{C}}_a^{-1}) \otimes \mathbf{I}_{N_T}$.

After K iterations the BEM algorithm stops producing a channel estimate $\mathbf{A}_{\text{BEM}} = \mathbf{A}^{(K)}$ and the APPs $\{\Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \mathbf{A}_{\text{BEM}})\}$ which can be processed exactly like in the previous application. In the following the BEM-based estimation algorithm (43)–(46) is dubbed STBC-BEM.

4.3. Numerical results

The error performance of the STBC-BEM algorithm has been assessed by computer simulation for the Alamouti's space-time block code [51]. Then we have

$$\mathbf{S}[n] \doteq \begin{bmatrix} s_n^1 & s_n^2 \\ -(s_n^2)^* & (s_n^1)^* \end{bmatrix}, \quad (48)$$

where the symbols $\{s_n^1, s_n^2\}$ belong to a BPSK constellation.⁵ In the following we assume that (1) $R_a[m] = J_0(2\pi m L B_D T)$, where $J_0(x)$ is the zeroth-order Bessel function of the first kind, B_D is the fading Doppler bandwidth, and T is the signaling interval; (2) the SNR is defined as E_b/N_0 , where E_b is the average received energy per receive antenna and information bit; (3) each *packet* of $(N_B - 1)$ consecutive information blocks is followed by one pilot block, so that the pilot symbol rate is $R_p = 1/N_B$.

The STBC-BEM algorithm processes a sample set \mathbf{R} consisting of $N \cdot L$ consecutive received signal samples, corresponding to N transmitted symbol blocks. It is assumed that the first and last L samples of \mathbf{R} always correspond to a pilot block. This entails that (a) $N = N_p N_B + 1$, if N_p packets are processed, and (b) the last block of each set is in common with the first of the next one. The information provided by the pilot symbols is exploited to initialize the BEM algorithm. In particular the initial channel estimate for the j th receive antenna is evaluated as $\mathbf{A}_j = \mathbf{F} \mathbf{R}_j$, where \mathbf{R}_j is the j th column of \mathbf{R} , with $j = 1, 2, \dots, N_R$. Here \mathbf{F} is an optimal $NN_T \times NL$ matrix that can be easily derived by standard methods (Wiener filtering) [29, 36], under the assumptions that (a) the information channel symbols are independent and identically distributed and (b) the pilot symbols are exactly known.

In all the following results it is assumed that the BEM algorithm processes $N_p = 4$ consecutive packets, each consisting of $N_B = 10$ consecutive blocks.

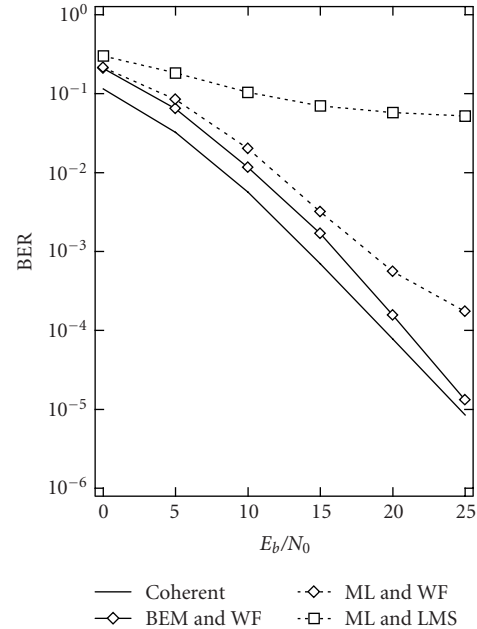


FIGURE 4: BER performance of various detection algorithms with Alamouti's STBC. $N_R = 1$ and $B_D T = 2 \cdot 10^{-2}$.

In Figure 4 the error performance of the STBC-BEM (with $K = 3$) is compared with that provided by an ML receiver using WF channel estimation⁶ and an ML receiver using decision-directed *least mean square* (LMS) channel tracking with step size $\mu = 0.5$ (the tracker is initialized for each packet using the pilot block at its beginning in order to avoid runaway problems) for single receive diversity ($N_R = 1$) and $B_D T = 2 \cdot 10^{-2}$. The BER performance of a coherent receiver endowed with ideal CSI is also shown. These results evidence that (1) since the energy loss due to pilot symbols is 0.45 dB, the BEM performs very well if the fading rate is not too large; (2) the BEM substantially outperforms the other detectors. Further simulations have also shown that a blind SISO detector based on the EM-based approach illustrated in [6] and initialized by a WF does not outperform the ML detector endowed with the same channel estimator.

Figure 5 shows the error performance of the STBC-BEM with a different number of iterations, that is, with $K = 1, 2$, and 3, in the same scenario as the previous figure. These results evidence the usefulness of running three full iterations in the BEM procedure, in order to approach the performance of a coherent receiver endowed with ideal CSI. We also found, however, that negligible gains are offered by $K > 3$.

The comments already expressed about the results of Figure 4 also apply to Figure 6, referring to double receive diversity ($N_R = 2$), channel estimation based on WF and $B_D T = 5 \cdot 10^{-3}, 10^{-2}$, and $2 \cdot 10^{-2}$ for the BEM ($B_D T = 2 \cdot 10^{-2}$ only is considered for the ML detector). This figure

⁵Further results (not shown for space limitations) evidence that the comments expressed for a BPSK system also apply to larger constellations.

⁶Its error performance coincides with that offered by the BEM without iterations.

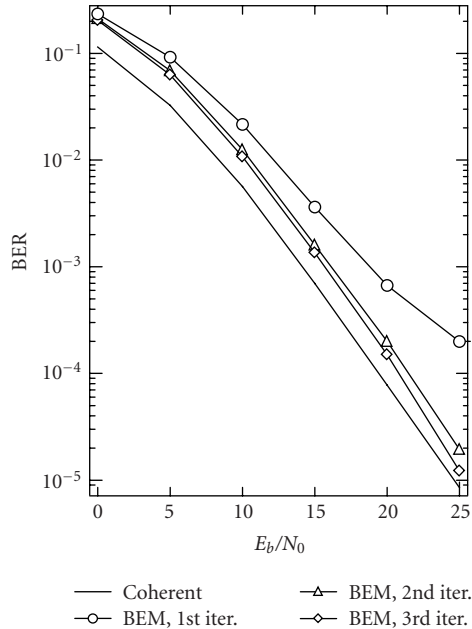


FIGURE 5: BER performance of the BEM detection algorithm with Alamouti's STBC. The error performance of the coherent detector is also shown for comparison. $N_R = 1$, $B_D T = 2 \cdot 10^{-2}$, and $K = 1, 2$, and 3 .

also evidences that the BEM performance is not substantially affected by a change in the Doppler rate, provided that $B_D T \leq 2 \cdot 10^{-2}$.

In Figure 7 the BEM and the ML detector BER versus the normalized Doppler bandwidth $B_D T$ is shown for $B_D T \in (10^{-2}, 5 \cdot 10^{-2})$ and $E_b/N_0 = 10$ dB (WF is used in both cases). It is worth noting that the performance degradation increases for larger Doppler bandwidths as the quality of the initial estimate of the BEM becomes poorer and this prevents BEM convergence to the global maximum, at least over some data blocks. Simulation results have also evidenced that, in this case, increasing the number of BEM iterations provides a negligible improvement.

4.4. SISO detection of space-time block coded OFDM signals

4.4.1. Introduction

The use of OFDM is often suggested to simplify channel equalization in the presence of appreciable frequency selectivity. When employed in MIMO wireless systems, the OFDM technique can be also easily combined with channel codes devised for multiple transmit antennas, that is, with *space-time* (ST) codes. A further improvement in the system performance can be achieved when conventional outer channel codes, like *convolutional codes* [56, 57] or *low-density parity-check* (LDPC) codes [58], are used in conjunction with proper ST symbol mappers.

Decoding of ST codes usually requires an accurate knowledge of CSI at the receiver. In MIMO OFDM systems, however, channel estimation may represent a serious problem,

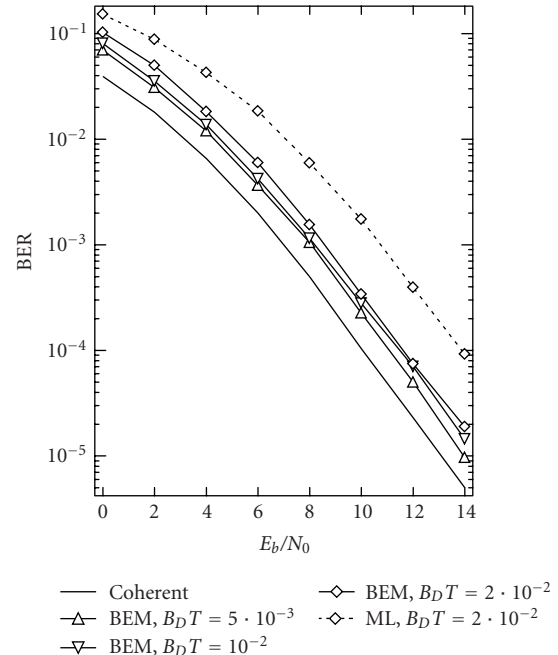


FIGURE 6: BER performance of various detection algorithms with Alamouti's STBC. $N_R = 2$.

especially in time-varying environments, because of the high complexity needed to achieve a satisfying accuracy [59], even if simplified pilot-based channel estimators can be devised [60]. Recently, it has been shown that, when OFDM is combined with ST block coding [51] and a pilot-based channel estimate is available at the receiver, the EM technique can be applied to devise accurate channel estimators [61] and that such estimators can be used for *soft-in hard-output* detection [54]. In the last case, hard decisions are then converted to soft data information which can be exploited in iterative receiver architectures when outer coding is employed at the transmitter. In this part we tackle the same problem, but from a different perspective. In fact, we derive a SISO module based on the BEM technique. Preliminary simulation results suggest that this algorithm offers better performance than that derived in [54] with a lower overall computational burden.

4.4.2. Signal and channel models

In this paper we consider an ST block coded OFDM system employing N subcarriers jointly with N_T transmit and N_R receive antennas. The block diagram of the communication system is illustrated in Figure 8a. The coding scheme results from the concatenation of a convolutional or an LDPC code with an orthogonal STBC. It is worth noting that that LDPC codes have some relevant properties [62], like low decoding complexity and excellent performance, which make them a promising coding technique for ST coded OFDM systems [58].

The input bit stream is partitioned into blocks, each independently encoded by means of a channel encoder. After (optional) bit interleaving (II) the coded bits are mapped

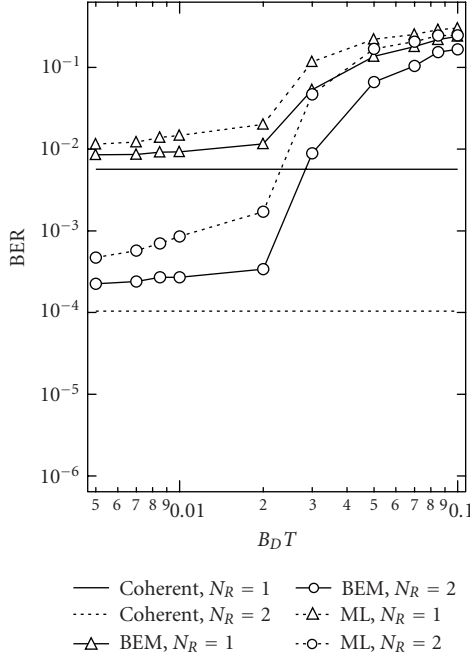


FIGURE 7: BER versus the normalized Doppler bandwidth $B_D T$ for various detection algorithms with STBC. $E_b/N_0 = 10$ dB. $N_R = 1$ and 2.

into channel symbols belonging to an M -ary PSK constellation. The resulting symbol sequence feeds an ST orthogonal block encoder. In the following, we consider, for simplicity, the Alamouti's STBC [51], even if the proposed detection algorithm can be easily extended to any orthogonal ST block code. The output sequence of the ST encoder is passed through a bank of N_T *inverse discrete Fourier transform* (IDFT) processors, which generate an ST-OFDM codeword spanning L OFDM symbol intervals. For instance, with Alamouti's STBC, we have $L = 2$ and, if $c_0[l, n]$ and $c_1[l, n]$ denote the channel symbols transmitted on the n th OFDM subcarrier (with $n = 0, \dots, N - 1$) in the l th OFDM symbol interval (with l even) by the first and the second transmit antenna, respectively, then $c_0[l + 1, n] = -c_1^*[l, n]$ and $c_1[l + 1, n] = c_0^*[l, n]$ are sent in the next symbol interval. In other words, the resulting codeword associated with the n th subcarrier is represented by the matrix

$$\mathbf{S}[n] = \begin{bmatrix} c_0[l, n] & c_1[l, n] \\ c_0[l + 1, n] & c_1[l + 1, n] \end{bmatrix} \quad (49)$$

belonging to an alphabet $\Omega = \{\mathbf{S}_p, p = 1, \dots, P\}$ (with $P = M^2$) of unitary matrices [51].

The OFDM signal is transmitted over a *wide sense stationary uncorrelated scattering* (WSS-US) MIMO channel [63]. In the following it is assumed that (a) all the single-input single-output channels associated with different transmit/receive antenna pairs are mutually independent, identically distributed and are affected by Rayleigh fading; (b) in the propagation scenario, frequency dispersion is independent of time dispersion. Under these hypotheses a full

statistical description of the MIMO channel is provided by its *power delay profile* (PDP) and its *Doppler power density spectrum* (PDS) or, equivalently, by its frequency correlation function $R_H(f)$ and its time correlation function $R_D(t)$, respectively [63]. At the receiver (see Figure 8b) a bank of N_R DFT processors (one per receive antenna) is fed by N_R distinct discrete-time signal sequences produced by matched-filtering and symbol-rate sampling. The outputs of the DFTs are processed by a BEM-based SISO detection algorithm (see the following paragraph) operating on a codeword-by-codeword basis. For this reason, in the following, we concentrate on the detection of a single ST-OFDM codeword. In particular, if $r_j[l, n]$ denotes the received signal sample taken at the output of the j th DFT for the n th subcarrier frequency in the l th OFDM symbol interval, with $j = 0, 1, \dots, N_R - 1$ and $n = 0, \dots, N - 1$, we always take a couple of consecutive received signal samples for $l = 0, 2, 4, \dots$. If we assume that the fading process remains constant over an ST codeword (i.e., over two adjacent OFDM symbol intervals with Alamouti's STBC), the $L \times N_R$ matrix $\mathbf{R}[l, n] = [r_j[l, n]]$ collecting the received signal samples over the observation interval for the n th subcarrier can be expressed as [54]

$$\mathbf{R}[l, n] = \mathbf{S}[l, n]\mathbf{H}[l, n] + \mathbf{W}[l, n]. \quad (50)$$

Here, $\mathbf{S}[n, l]$ is the $L \times N_T$ transmitted codeword matrix (see (49)), $\mathbf{H}[l, n] = [H_{i,j}[l, n]]$ is an $N_T \times N_R$ channel response matrix ($H_{i,j}[l, n]$ represents the complex channel gain between the i th transmit and the j th receive antenna at the n th subcarrier frequency), and $\mathbf{W}[l, n] = [w_{i,j}[l, n]]$ is an $L \times N_R$ noise matrix. The elements $\{w_{i,j}[l, n]\}$ of $\mathbf{W}[l, n]$ are independent complex zero mean Gaussian random variables with variance $\sigma_w^2 = 2N_0$. We also note that $\{H_{i,j}[l, n]\}$ are complex Gaussian random variables with zero mean and that the correlation between $H_{i,j}[l, n + m]$ and $H_{i,j}[l, n]$ is given by $R_H[m] = E\{H_{i,j}[l, n + m]H_{i,j}^*[l, n]\} = R_H(mf_\Delta)$, where f_Δ is the subcarrier spacing.

For a given l , the matrices (50) associated with all the different subcarriers ($n = 0, \dots, N - 1$) can be grouped in an $LN \times N_R$ matrix $\mathbf{R}[l] \doteq [\mathbf{R}^H[l, 0], \mathbf{R}^H[l, 1], \dots, \mathbf{R}^H[l, N - 1]]^H$. If the dependence on l is dropped, for simplicity, this vector can be expressed as

$$\mathbf{R} = \mathbf{D}(\mathbf{S})\mathbf{H} + \mathbf{W}, \quad (51)$$

where $\mathbf{H} \doteq [\mathbf{H}^H[0], \mathbf{H}^H[1], \dots, \mathbf{H}^H[N - 1]]^H$, $\mathbf{W} \doteq [\mathbf{W}^H[0], \mathbf{W}^H[1], \dots, \mathbf{W}^H[N - 1]]^H$, and $\mathbf{D}(\mathbf{S}) \doteq \text{diag}\{\mathbf{S}[0], \mathbf{S}[1], \dots, \mathbf{S}[N - 1]\}$.

4.4.3. A BEM-based SISO algorithm for OFDM systems

Following the same approach as the previous two scenarios, we choose $\Theta = \mathbf{H}$ and $\mathbf{I} = \mathbf{S}$. Then, as shown in [35], the BEM auxiliary function (6) can be expressed as

$$Q(\mathbf{H}, \hat{\mathbf{H}}) = - \sum_{j=1}^{N_R} \mathbf{H}_j^H \mathbf{M} \mathbf{H}_j - \frac{2}{\sigma_w^2} \text{Re} \{ \hat{\mathbf{V}}_j^H \mathbf{H}_j \}, \quad (52)$$

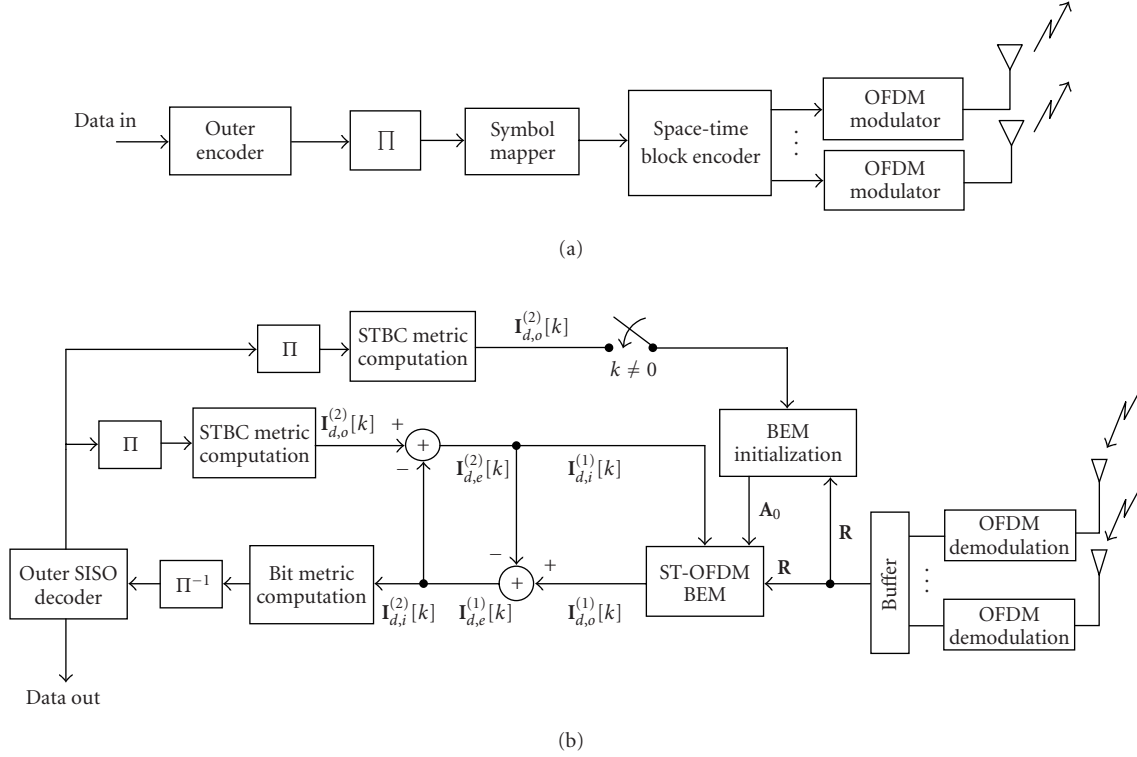


FIGURE 8: Block diagrams of the space-time block coded OFDM: (a) transmitter and (b) receiver.

where \mathbf{H}_j is the j th column of the matrix \mathbf{H} , $\mathbf{M} \doteq \mathbf{C}_H^{-1} + (1/\sigma_w^2)\mathbf{I}_{N_I}$ with $\mathbf{C}_H = E\{\mathbf{H}_j\mathbf{H}_j^H\}$, $\tilde{\mathbf{V}}_j$ is the j th column of the matrix $\tilde{\mathbf{V}} \doteq \mathbf{D}^H(\tilde{\mathbf{S}})\mathbf{R}$, and the matrix $\tilde{\mathbf{S}}$ results from the ordered concatenation of the matrices $\{\tilde{\mathbf{S}}[n], n = 0, 1, \dots, N-1\}$, with

$$\tilde{\mathbf{S}}[n] \doteq \sum_{\mathbf{S}_m \in \Omega} \mathbf{S}_m \Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \tilde{\mathbf{H}}). \quad (53)$$

The APPs $\{\Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \tilde{\mathbf{H}})\}$ can be evaluated using the Bayes formula

$$\Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \tilde{\mathbf{H}}) = \frac{f(\mathbf{R}[n] | \mathbf{S}_m, \tilde{\mathbf{H}}[n])P(\mathbf{S}_m)}{\sum_{\tilde{\mathbf{S}}_m \in \Omega} f(\mathbf{R}[n] | \tilde{\mathbf{S}}_m, \tilde{\mathbf{H}}[n])P(\tilde{\mathbf{S}}_m)}, \quad (54)$$

where

$$f(\mathbf{R}[n] | \mathbf{S}_m, \tilde{\mathbf{H}}[n]) = C_R \exp\left[-\frac{h(\mathbf{R}[n], \mathbf{S}_m, \tilde{\mathbf{H}}[n])}{\sigma_w^2}\right] \quad (55)$$

with $C_R \doteq \det(\pi\sigma_w^2\mathbf{I}_L)^{-N_R}$ and $h(\mathbf{R}[n], \mathbf{S}_m, \tilde{\mathbf{H}}[n]) \doteq \text{tr}\{(\mathbf{R}[n] - \mathbf{S}_m\tilde{\mathbf{A}}[n])^H \cdot (\mathbf{R}[n] - \mathbf{S}_m\tilde{\mathbf{A}}[n])\}$. The BEM algorithm operates as follows. Given the channel estimate $\mathbf{H}^{(k)}$ at the k th iteration, the next estimate $\mathbf{H}^{(k+1)}$ is evaluated as

$$\mathbf{H}_j^{(k+1)} = \mathbf{P}^{-1}\mathbf{V}_j^{(k)} \quad (56)$$

with $\mathbf{P} \doteq \mathbf{I}_{NN_T} + \sigma_w^2\mathbf{C}_H^{-1}$ and $j = 1, 2, \dots, N_R$. It is important to note that (a) \mathbf{P} does not depend on the index of the receive antenna; (b) the inverse of \mathbf{P} does not need to be recomputed as long as the channel statistics do not change; (c) (56) can be simplified factoring \mathbf{C}_H as

$$\mathbf{C}_H = \tilde{\mathbf{C}}_H \otimes \mathbf{I}_{N_T}, \quad (57)$$

where $\tilde{\mathbf{C}}_H$ is the covariance matrix of the vector $\mathbf{H}_{i,j} = [H_{i,j}[0], H_{i,j}[1], \dots, H_{i,j}[N-1]]^T$ and \otimes is the Kronecker product, so that $\mathbf{P} = (\mathbf{I}_N + \sigma_w^2\tilde{\mathbf{C}}_H^{-1}) \otimes \mathbf{I}_{N_T}$. In the following the BEM-based estimation algorithm (53)–(56) is dubbed ST-OFDM BEM.

After K iterations the BEM algorithm stops producing a channel estimate $\mathbf{H}_{\text{BEM}} = \mathbf{H}^{(K)}$ and the APPs $\{\Pr(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \mathbf{H}_{\text{BEM}})\}$. These can be exploited to take MAP decisions or for soft decoding of an outer code in a concatenated scheme. In our work, we have considered the iterative receiver structure as shown in Figure 8b. This structure operates as follows. After OFDM demodulation, the ST-OFDM BEM module takes as input the received signal vector $\mathbf{R} \doteq [\mathbf{R}^H[0], \mathbf{R}^H[1], \dots, \mathbf{R}^H[N-1]]^H$, an initial channel estimate matrix $\mathbf{H}^{(0)}$ (consisting of $N \cdot N_T \times N_R$ matrices $\mathbf{H}^{(0)}[n]$) and the $N \times P$ a priori information matrices $\{\mathbf{I}_{d,i}^{(1)}[k] = [(I_{d,i}^{(1)}[k])_{n,m}]\}$. Here $(I_{d,i}^{(1)}[k])_{n,m} = \log \Pr^{(l)}(\mathbf{S}[n] = \mathbf{S}_m)$, where $\Pr^{(l)}(\mathbf{S}[n] = \mathbf{S}_m)$ denotes the APRP that $\mathbf{S}[n]$ is equal to the m th codeword of the alphabet Ω at the k th step. After K iterations the BEM algorithm produces

the $N \times P$ output matrices $\{\mathbf{I}_{d,o}^{(1)}[k] = [(\mathbf{I}_{d,o}^{(1)}[k])_{n,m}]\}$ with $(\mathbf{I}_{d,o}^{(1)}[k])_{n,m} = \log \Pr^{(l)}(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \mathbf{H}_{\text{BEM}})$, where $\Pr^{(l)}(\mathbf{S}[n] = \mathbf{S}_m | \mathbf{R}, \mathbf{H}_{\text{BEM}})$ represents the APP of the event $\{\mathbf{S}[n] = \mathbf{S}_m\}$ at the k th step. Then the *extrinsic information* matrices $\{\mathbf{I}_{d,e}^{(1)}[k]\}$ are evaluated as $\mathbf{I}_{d,e}^{(1)}[k] = \mathbf{I}_{d,o}^{(1)}[k] - \mathbf{I}_{d,i}^{(1)}[k]$. Since interleaving is performed at the bit level, before sending the extrinsic information to the deinterleaver (Π^{-1}) and to the SISO decoder, the evaluation of the *soft* bit metrics is needed (see [64, Section II-C]). The channel decoder produces the a posteriori *bit* information matrices and, after bit interleaving and probability recombination, the a posteriori *symbol* information matrices $\{\mathbf{I}_{d,o}^{(2)}[k]\}$ (in log form). Finally, at the last iteration, the SISO decoder computes the APP matrix $\{\mathbf{P}_b\}$ together with a bit estimate vector. Subtracting $\{\mathbf{I}_{d,i}^{(2)}[k]\}$ from $\{\mathbf{I}_{d,o}^{(2)}[k]\}$ produces the extrinsic information matrices $\{\mathbf{I}_{d,e}^{(2)}[k]\}$ of the channel symbols which are fed back as input to the ST-OFDM BEM decoder.

In our simulations both convolutional and LDPC codes have been employed. With convolutional codes the bit APRPs produced by the ST-OFDM BEM feed a *Bahl Cocke Jelinek Raviv* (BCJR) algorithm [20] implemented in its log MAP form [65]. With LDPC codes bit *log-likelihood ratios* (LLRs) are evaluated on the basis of the bit APRPs and sent to an LDPC decoder based on the *belief propagation* (BP) algorithm [62, 66]. It is important to point out that (a) the parity check matrices of the LDPC codes employed in our work have been generated in a random fashion [67], avoiding cycles of length 4 in the code graph in order to improve the code distance properties; (b) due to the random generation of the encoding matrix, no external interleaver (deinterleaver) is needed at the output (input) of the LDPC encoder (decoder) [58].

Finally, we note that, in the proposed receiver structure, the APPs $\{\mathbf{I}_{d,o}^{(2)}[k]\}$, after interleaving, are also used to evaluate the estimate $\mathbf{H}^{(k+1)}$ needed for the initialization of the ST-OFDM BEM in the $(k+1)$ th iteration of the receiver. At the beginning of the first iteration, however, no *a priori* information on the channel symbols is available. For this reason the initial fading estimate $\mathbf{H}^{(0)}$ of the ST-OFDM BEM is evaluated by means of the pilot-based channel estimation algorithm derived in [60].

4.5. Numerical results

In this paragraph some BER results are illustrated. In our computer simulations the reduced complexity model for WSS-US channels proposed in [68] has been used for the generation of a MIMO multipath fading channel. In particular, for a given Doppler bandwidth B_D , the Doppler PDS has been defined as $S_D(f) = 1 - 1.72f_0^2 + 0.785f_0^4$ for $f_0 \leq 1$ and $S_D(f) = 0$ for $f_0 > 1$, where $f_0 = f/B_D$ [69]. Moreover, the multipath MIMO channel has been modeled as a 3-tap delay line approximating an exponential PDP $P_h(\tau) = \tau_0^{-1} \exp(-\tau/\tau_0)u(\tau)$, with $\tau_0 = 1.56$ microseconds (the corresponding frequency correlation function is $R_H(f) = 1/(1 + j2\pi f\tau_0)$). Then, in accordance with the OFDM physical layer specifications for the broadband radio access networks

(BRAN) in [70], the following parameters have been selected for the ST block coded OFDM system: (a) the DFT order is $N = 256$; (b) the number of useful OFDM subcarriers is equal to 192, since the total number of subcarriers N includes 27 suppressed carriers on the upper frequencies, 28 suppressed carriers on the lower frequencies, 8 BPSK pilot symbols, and 1 DC carrier set to 0; (c) the OFDM symbol interval is $T_S = 0.125$ microseconds; (d) the length of the cyclic prefix in the OFDM modulator has been set to 64; (d) with convolutional codes, a 4-state rate 1/2 convolutional code with generators $g_1 = (5)_8$ and $g_2 = (7)_8$ has been adopted, when used; (e) with LDPC codes, a regular (3,6) code with rate $R = 1/2$ and a BP algorithm with a maximum number of iterations equal to 10 have been adopted, when used; (f) QPSK modulation has been employed for both uncoded and coded transmission; (g) a single frame consists of 9 ST block coded OFDM information codewords plus one pilot ST block coded OFDM codeword appended at its beginning. Moreover a single receive antenna, that is, $N_R = 1$ and a Doppler bandwidth $B_D = 200$ Hz have been chosen for our simulations.

In addition, the following assumptions have been made at the receive side: (a) the SNR is defined as E_b/N_0 , where E_b is the average captured energy per receive antenna and information bit; (b) the BEM algorithm processes a block consisting of 192 Alamouti's space-time block codewords, and accomplishes $K = 3$ complete iterations; (c) the last channel estimate generated by the the BEM algorithm for each ST-OFDM codeword is used as an initial estimate of the same algorithm for the next codeword.

Figure 9 shows the ST-OFDM BEM algorithm performance without outer channel coding. Comparison is made with an ML detector endowed with ideal CSI (*genie bound*) and with an ML detector endowed with the same pilot-based *channel estimator* (CE) as the BEM [60]. These results evidence that the ST-OFDM BEM algorithm substantially outperforms a realistic ML detector. We also note that the energy loss due to pilot symbol insertion is 0.45 dB, so that the energy gap between the genie bound and the ST-OFDM BEM is about 1.5 dB [71].

Some simulation results referring to a convolutionally encoded system are shown in Figure 10, comparing the BER performance provided by the iterative receiver described in the previous paragraph (with 0, 1, and 2 iterations) with that offered by a BCJR decoder endowed with ideal CSI. We have also considered a receiver structure in which the likelihoods produced by the above-mentioned ML detector with pilot-based CE are exploited to generate soft data information feeding, after deinterleaving, the SISO outer decoder. The proposed iterative architecture substantially outperforms the latter and, if the energy loss due to pilot symbols is neglected, it approaches closely the genie bound. It is also worth noting that, in this scenario, carrying out global iterations provides a very small gain. This result can be explained as follows. The ST-OFDM BEM, starting from a pilot-based channel estimate, produces a good channel estimate and a good estimate of the data APPs since the beginning, that is, even in the absence of the APRPs produced by the BCJR, despite

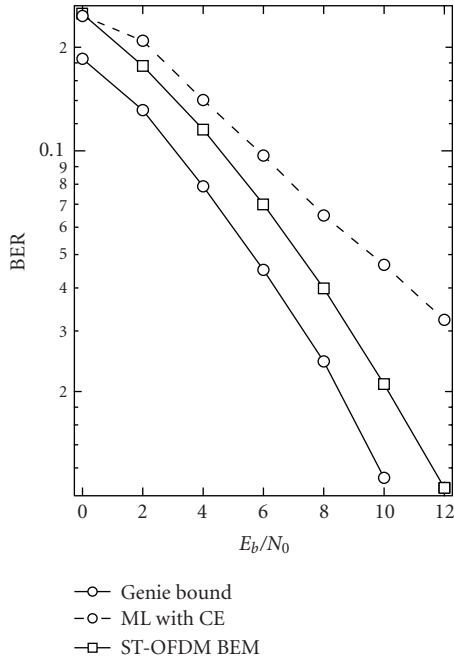


FIGURE 9: BER performance of the ST-OFDM BEM algorithm without outer coding. $N_R = 1$ and $K = 3$.

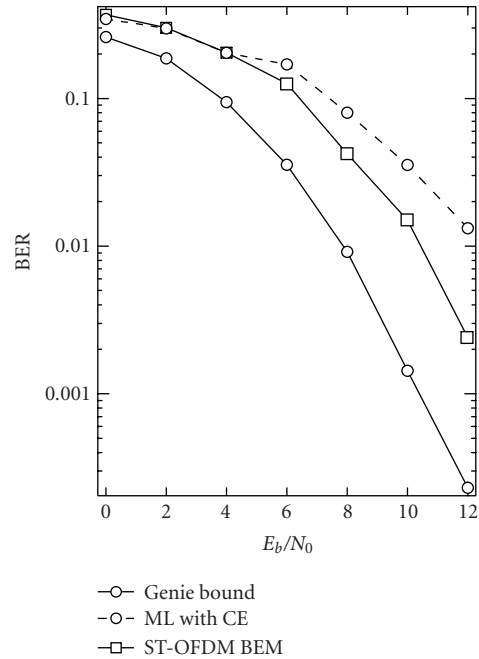


FIGURE 11: BER performance of the ST-OFDM BEM receiver. LDPC coding, $N_R = 1$, and $K = 3$.

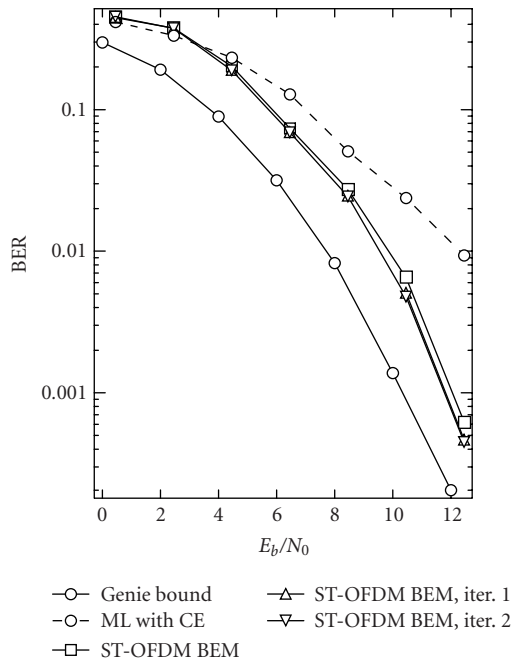


FIGURE 10: BER performance of the ST-DFDM BEM iterative receiver. Convolutional coding, $N_R = 1$, and $K = 3$.

the appreciable Doppler rate. These results are substantially different than those illustrated in [54, page 223], evidencing, for instance, a strong gap between the performance in the absence of iterations and that achieved after one iteration and suggesting the use of 3–5 global iterations. On the basis of these preliminary results, since the complexity (per iteration)

of the ST-OFDM BEM and that of the EM algorithm derived in [54] are comparable, the use of the former should be preferred to the latter, since it ensures faster convergence, that is, a smaller overall complexity.

Finally, in Figure 11 the performance of the ST-OFDM BEM receiver for LDPC-coded signals is illustrated. The BER performance of the proposed algorithm is compared with that obtained by a BP algorithm endowed with perfect CSI. The curve labeled as “ML with CE” represents the BER performance of an ML detector endowed with pilot-based CE and followed by the LDPC decoder. Even without turbo decoding, the ST-OFDM BEM algorithm brings a substantial gain against the ML-based symbol detection approach. It is worth noting that, in this scenario, the BER performances given by the LDPC and convolutional coding schemes are widely comparable. This poor behavior obtained by LDPC coding is mainly due to the small dimension of the parity-check matrix employed in our simulations.

5. CONCLUSIONS

In this paper the BEM technique has been proposed to solve MAP estimation problems. In particular, we have shown that it represents a useful tool to derive novel SISO detectors for communication channels with random parametric uncertainty and memory. As an application of these concepts, SISO modules for the iterative detection of coded digital signals transmitted over fading channels have been derived in three specific scenarios and their error performance has been assessed. Applications of the BEM technique to other communication scenarios are the subject of ongoing research activities.

ACKNOWLEDGMENTS

This work has been performed in the framework of the project STINGRAY IST-2000-30173, which is funded by the European Community. The authors would like to acknowledge the contributions of their colleagues from the Intracom Hellenic Telecommunications and Electronics Industry S.A., the University of Modena and Reggio Emilia, the Institute of Accelerating Systems and Applications, the Technical Research Centre of Finland, and the National Technical University of Athens.

REFERENCES

- [1] D. P. Taylor, G. M. Vitetta, B. D. Hart, and A. Mämmelä, "Wireless channel equalization," *European Transactions on Telecommunications*, vol. 9, no. 2, pp. 117–143, 1998.
- [2] R. Raheli, A. Polydoros, and C.-K. Tzou, "Per-survivor processing: a general approach to MLSE in uncertain environment," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 354–364, 1995.
- [3] C. N. Georghiades and J. C. Han, "Sequence estimation in the presence of random parameters via the EM algorithm," *IEEE Trans. Commun.*, vol. 45, no. 3, pp. 300–308, 1997.
- [4] H. Zamiri-Jafarian and S. Pasupathy, "Adaptive MLSDE using the EM algorithm," *IEEE Trans. Commun.*, vol. 47, no. 8, pp. 1181–1193, 1999.
- [5] A. Anastasopoulos and K. M. Chugg, "Adaptive soft-input soft-output algorithms for iterative detection with parametric uncertainty," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1638–1649, 2000.
- [6] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels," *IEEE Trans. Commun.*, vol. 42, no. 7, pp. 2406–2413, 1994.
- [7] G. K. Kaleh, "Joint carrier phase estimation and symbol decoding of trellis codes," *European Transactions on Telecommunications*, vol. 4, no. 2, pp. 157–163, 1993.
- [8] C. A. Haro, J. A. R. Fonollosa, and J. Fonollosa, "Blind channel estimation and data detection using hidden Markov models," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 241–247, 1997.
- [9] W. Turin, "Map decoding in channels with memory," *IEEE Trans. Commun.*, vol. 48, no. 5, pp. 757–763, 2000.
- [10] L. Davis, I. Collings, and P. Hoeher, "Joint MAP equalization and channel estimation for frequency-selective and frequency-flat fast-fading channels," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2106–2114, 2001.
- [11] P. Hoeher and J. Lodge, "Turbo DPSK: iterative differential DPSK demodulation and channel decoding," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 837–843, 1999.
- [12] I. D. Marsland and P. T. Mathiopoulos, "Multiple differential detection of parallel concatenated convolutional (turbo) codes in correlated fast Rayleigh fading," *IEEE J. Select. Areas Commun.*, vol. 16, no. 2, pp. 265–274, 1998.
- [13] M. J. Gertsman and J. H. Lodge, "Symbol-by-symbol MAP demodulation of CPM and PSK signals on Rayleigh flat-fading channels," *IEEE Trans. Commun.*, vol. 45, no. 7, pp. 788–799, 1997.
- [14] J. Hagenauer, "The turbo principle: tutorial introduction and state of the art," in *Proc. International Symposium on Turbo Codes & Related Topics*, pp. 1–11, Brest, France, September 1997.
- [15] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 429–445, 1996.
- [16] B. Sklar, "A primer on turbo code concepts," *IEEE Commun. Magazine*, vol. 35, no. 12, pp. 94–102, 1997.
- [17] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, 1996.
- [18] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "A soft-input soft-output APP module for iterative decoding of concatenated codes," *IEEE Communications Letters*, vol. 1, no. 1, pp. 22–24, 1997.
- [19] Y. Lin, B. Vucetic, and Y. Sato, "Optimum soft-output detection for channels with intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 41, no. 3, pp. 704–713, 1995.
- [20] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (Corresp.)," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, 1974.
- [21] V. Franz and J. B. Anderson, "Concatenated decoding with a reduced-search BCJR algorithm," *IEEE J. Select. Areas Commun.*, vol. 16, no. 2, pp. 186–195, 1998.
- [22] S. Benedetto, G. Montorsi, D. Divsalar, and F. Pollara, "Soft-input, soft-output modules for the construction and distributed iterative decoding of code networks," *European Transactions on Telecommunications*, vol. 9, no. 2, pp. 155–172, 1998.
- [23] C. Douillard, A. Picart, M. Jézéquel, P. Didier, C. Berrou, and A. Glavieux, "Iterative correction of intersymbol interference: turbo-equalization," *European Transactions on Telecommunications*, vol. 6, no. 5, pp. 507–511, 1995.
- [24] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B*, vol. 39, no. 1, pp. 1–38, 1977.
- [25] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Mag.*, vol. 13, no. 6, pp. 47–60, 1996.
- [26] H. Z. Jafarian and S. Pasupathy, "Em-based recursive estimation of channel parameters," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1297–1302, 1999.
- [27] V. Lottici and M. Luise, "Embedding carrier phase recovery into iterative decoding of turbo-coded linear modulations," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 661–669, 2004.
- [28] N. Noels, C. Herzet, A. Dejonghe, et al., "Turbo synchronization: an EM algorithm interpretation," in *Proc. IEEE International Conference on Communications (ICC '03)*, pp. 2933–2937, Anchorage, Alaska, USA, May 2003.
- [29] E. Chiavaccini and G. M. Vitetta, "MAP symbol estimation in the presence of random parameters via a generalized EM algorithm," in *Proc. IEEE International Conference on Communications (ICC '01)*, vol. 4, pp. 1057–1061, Helsinki, Finland, June 2001.
- [30] E. Chiavaccini and G. M. Vitetta, "A BEM-based detector for CPM signals transmitted over frequency-flat fading channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 3, pp. 409–412, 2003.
- [31] E. Chiavaccini and G. M. Vitetta, "MAP symbol estimation on frequency-flat Rayleigh fading channels via a Bayesian EM algorithm," *IEEE Trans. Commun.*, vol. 49, no. 11, pp. 1869–1872, 2001.
- [32] E. Chiavaccini and G. M. Vitetta, "A per-survivor phase-estimation algorithm for detection of PSK signals," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2059–2061, 2001.
- [33] A. S. Gallo, E. Chiavaccini, and G. M. Vitetta, "BEM-based multiuser detection for synchronous DS-CDMA systems over frequency flat fading channels," in *Proc. 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication Conference (PIMRC '02)*, vol. 3, pp. 1285–1289, Lisbon, Portugal, September 2002.
- [34] E. Chiavaccini, A. S. Gallo, and G. M. Vitetta, "BEM-Based SISO detection of space-time block coded signals transmitted over frequency-flat fading channels," in *Proc. IEEE Global*

- Telecommunications Conference (GLOBECOM '02)*, vol. 1, pp. 374–378, Taipei, Taiwan, November 2002.
- [35] A. S. Gallo and G. M. Vitetta, “A novel soft-in soft-out detection algorithm for space-time concatenated coding schemes operating over multipath fading channels,” submitted to the *IEEE Global Telecommunications Conference (GLOBECOM '04)*, 2004.
- [36] C. W. Helstrom, *Elements of Signal Detection & Estimation*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1995.
- [37] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, London, UK, 1995.
- [38] G. M. Vitetta, D. P. Taylor, and U. Mengali, “Double-filtering receivers for PSK signals transmitted over Rayleigh frequency-flat fading channels,” *IEEE Trans. Commun.*, vol. 44, no. 6, pp. 686–695, 1996.
- [39] G. M. Vitetta, U. Mengali, and D. P. Taylor, “Blind detection of CPM signals transmitted over frequency-flat fading channels,” *IEEE Trans. Veh. Technol.*, vol. 47, no. 3, pp. 961–968, 1998.
- [40] R. W. Chang and J. C. Hancock, “On receiver structures for channels having memory,” *IEEE Trans. Inform. Theory*, vol. 12, no. 4, pp. 463–468, 1966.
- [41] S. Verdù, *Multisuser Detection*, Cambridge University Press, Cambridge, UK, 1998.
- [42] A. Kocian and B. H. Fleury, “Iterative joint symbol detection and channel estimation for DS/CDMA via the sage algorithm,” in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '00)*, vol. 2, pp. 1410–1414, London, UK, September 2000.
- [43] L. B. Nelson and H. V. Poor, “Iterative multisuser receivers for CDMA channels: an EM-based approach,” *IEEE Trans. Commun.*, vol. 44, no. 12, pp. 1700–1710, 1996.
- [44] D. Raphaeli, “Suboptimal maximum-likelihood multisuser detection of synchronous CDMA on frequency-selective multipath channels,” *IEEE Trans. Commun.*, vol. 48, no. 5, pp. 875–885, 2000.
- [45] M. Feder and E. Weinstein, “Parameter estimation of superimposed signals using the EM algorithm,” *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 477–489, 1988.
- [46] R. H. Clarke, “A statistical theory of mobile radio reception,” *Bell System Technical Journal*, vol. 47, no. 6, pp. 957–1000, 1968.
- [47] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 2nd edition, 1989.
- [48] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.
- [49] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, “Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility and multiple paths,” *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 199–207, 1999.
- [50] A. Paulraj and C. B. Papadias, “Space-time processing for wireless communications,” *IEEE Signal Processing Mag.*, vol. 14, no. 6, pp. 49–83, 1997.
- [51] S. M. Alamouti, “A simple transmit diversity technique for wireless communications,” *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [52] C. Cozzo and B. L. Hughes, “Joint channel estimation and data symbol detection in space-time communications,” in *Proc. IEEE International Conference on Communications (ICC '00)*, vol. 1, pp. 287–291, New Orleans, La, USA, June 2000.
- [53] Y. Li, C. N. Georghiadis, and G. Huang, “Iterative maximum-likelihood sequence estimation for space-time coded systems,” *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 948–951, 2001.
- [54] B. Lu, X. Wang, and Y. Li, “Iterative receivers for space-time block-coded OFDM systems in dispersive fading channels,” *IEEE Transactions on Wireless Communications*, vol. 1, no. 2, pp. 213–225, 2002.
- [55] A. S. Gallo, E. Chiavaccini, and G. M. Vitetta, “A novel BEM-based multisuser detector for co-channel signals transmitted over frequency-flat fading channels,” in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '03)*, pp. 241–246, San Francisco, Calif, USA, December 2003.
- [56] Y. Gong and K. B. Letaief, “Concatenated space-time block coding with trellis coded modulation in fading channels,” *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 580–590, 2002.
- [57] T. H. Liew and L. Hanzo, “Space-time codes and concatenated channel codes for wireless communications,” *Proc. IEEE*, vol. 90, no. 2, pp. 187–219, 2002.
- [58] B. Lu, X. Wang, and K. R. Narayan, “LDPC-based space-time coded OFDM systems over correlated fading channels,” *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 74–88, 2002.
- [59] Y. Li, N. Seshadri, and S. Ariyavisitakul, “Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels,” *IEEE J. Select. Areas Commun.*, vol. 17, no. 3, pp. 461–471, 1999.
- [60] W. G. Jeon, K. H. Paik, and Y. S. Cho, “An efficient channel estimation technique for OFDM systems with transmitter diversity,” *IEICE Transactions Communications*, vol. E84-B, no. 4, pp. 967–974, 2001.
- [61] E. Panayirci and H. A. Cirpan, “Channel estimation for space-time block coded OFDM systems in the presence of multipath fading,” in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '02)*, vol. 2, pp. 1157–1161, Taipei, Taiwan, November 2002.
- [62] T. J. Richardson and R. L. Urbanke, “The capacity of low-density parity-check codes under message-passing decoding,” *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 599–618, 2001.
- [63] P. A. Bello, “Characterization of randomly time-variant linear channels,” *IEEE Transactions Communication Systems*, vol. 11, no. 4, pp. 360–393, 1963.
- [64] G. Taricco, G. Caire, and E. Biglieri, “Bit-interleaved coded modulation,” *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 927–946, 1998.
- [65] P. Robertson, E. Villebrun, and P. A. Hoeher, “A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain,” in *Proc. IEEE International Conference on Communications (ICC '95)*, vol. 2, pp. 1009–1013, Seattle, Wash, USA, June 1995.
- [66] R. G. Gallager, “Low-density parity-check codes,” *IEEE Trans. Inform. Theory*, vol. 8, no. 1, pp. 21–28, 1962.
- [67] D. J. C. MacKay, “Good error-correcting codes based on very sparse matrices,” *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 399–431, 1999.
- [68] E. Chiavaccini and G. M. Vitetta, “GQR Models for multipath Rayleigh fading channels,” *IEEE J. Select. Areas Commun.*, vol. 19, no. 6, pp. 1009–1018, 2001.
- [69] V. Erceg, K. V. S. Hari, M. S. Smith, et al., “Channel models for fixed wireless applications,” *IEEE 802.16.3 Task Group Contributions, Doc. IEEE 802.16.3.c-01/29r4*, 2001, available at <http://www.ieee802.org/16/tg3>.
- [70] ETSI, “Broadband Radio Access Networks (BRAN)—HIPERMAN, OFDM Physical (PHY) Layer,” draft, 2002.
- [71] The STINGRAY Consortium, “Benchmarking and Performance Criteria for STC-OFDM,” *Deliverable 3D3 (IST-2000-30173 STINGRAY)*, 2003.

A. S. Gallo was born in Nettuno, Rome, Italy, in 1974. He received the Dott. Ing. degree (cum laude) in electrical engineering from the University of Rome "La Sapienza," Italy, in 2000, defending a thesis on optimization techniques for neuro-fuzzy networks. In 2004 he received the Ph.D. degree in information engineering from the University of Modena and Reggio Emilia, Italy. His research interests lie in the areas of wireless communication and signal processing, with main emphasis on low-cost detection/equalization techniques, space-time coding for wideband communication systems, and neuro-fuzzy networks.



G. M. Vitetta was born in Reggio Calabria, Italy, in April 1966. He received the Dr.-Ing. degree in electronic engineering (cum laude) in 1990 and the Ph.D. degree in 1994, both from the University of Pisa, Italy. In 1992/1993, he spent a period at the University of Canterbury, Christchurch, New Zealand, doing research for digital communications on fading channels. From 1995 to 1998, he was a Research Fellow at the Department of Information Engineering of the University of Pisa. From 1998 to 2001, he held the position of Associate Professor of telecommunications at the University of Modena and Reggio Emilia. He is now a Full Professor of telecommunications in the same university. His main research interests lie in the broad area of communication theory, with particular emphasis on coded modulation, synchronization, statistical modeling of wireless channels and channel equalization. He is serving as an Editor of both the IEEE Transactions on Communications and the IEEE Transactions on Wireless Communications.



A Theoretical Framework for Soft-Information-Based Synchronization in Iterative (Turbo) Receivers

Nele Noels,¹ Vincenzo Lottici,² Antoine Dejonghe,³ Heidi Steendam,¹ Marc Moeneclaey,¹
Marco Luise,² Luc Vandendorpe³

¹Department of Telecommunications and Information Processing, Ghent University, 9000 Gent, Belgium
Emails: nnoels@telin.ugent.be, hs@telin.ugent.be, mm@telin.ugent.be

²Department of Information Engineering, University of Pisa, 56122 Pisa, Italy
Emails: v.lottici@iet.unipi.it, m.luise@iet.unipi.it

³Communications and Remote Sensing Laboratory, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium
Email: dejonghe@imec.be, vandendorpe@tele.ucl.ac

Received 13 May 2004; Revised 29 September 2004

This contribution considers turbo synchronization, that is to say, the use of soft data information to estimate parameters like carrier phase, frequency, or timing offsets of a modulated signal within an iterative data demodulator. In turbo synchronization, the receiver exploits the soft decisions computed at each turbo decoding iteration to provide a reliable estimate of some signal parameters. The aim of our paper is to show that such “turbo-estimation” approach can be regarded as a special case of the expectation-maximization (EM) algorithm. This leads to a general theoretical framework for turbo synchronization that allows to derive parameter estimation procedures for carrier phase and frequency offset, as well as for timing offset and signal amplitude. The proposed mathematical framework is illustrated by simulation results reported for the particular case of carrier phase and frequency offsets estimation of a turbo-coded 16-QAM signal.

Keywords and phrases: turbo synchronization, iterative detection, turbo codes, parameter estimation.

1. INTRODUCTION

The impressive performance of turbo codes [1] has triggered in the last decade a lot of research addressing the application of this powerful coding technique to digital communications [2]. More recently, the associated idea of iterative decoding has been extended to other receiver functions. This led to the so-called *turbo principle* which enables to perform (sub)optimal joint detection and decoding through the iterative exchange of soft information between soft-input/soft-output (SISO) stages. See [3, 4] for a review of some existing turbo receivers.

In addition to detection/decoding a receiver has also to perform signal synchronization, that is, to estimate a number of parameters like carrier phase offset, frequency offset, timing offset, and so forth. Synchronization for turbo-encoded systems is a challenging task since the receiver usually

operates at low SNR values (which can be defined as the ratio between the mean bit energy and the noise spectral density). In the technical literature a great effort is thus being devoted to the development of efficient estimation techniques to perform the above-mentioned synchronization functions within turbo receivers. We outline here at least two categories of algorithms.

(i) The first category consists of algorithms that try to modify classical SISO iterative detection/decoding in order to embed parameter estimation. In [5, 6], for instance, combined iterative decoding and estimation is performed with modified forward and backward recursions in the SISO decoders using a sort of per-survivor parameter estimation technique. In [7], the conventional turbo decoder structure is modified through the use of a simple phase estimation error model. A different approach is pursued in [8] wherein a method (having only polynomial complexity in the sequence length) of generating soft-decision metrics is illustrated and specifically applied to the problem of adaptive iterative detection of LDPC codes in the presence of time-varying unknown carrier phase offset. Further, simpler approximate

This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

receivers are proposed in [9] based on the insertion into each transmitted coded block of a number of pilot symbols with the aim of helping the joint phase estimation and decoding process.

(ii) The second category consists of algorithms that try to use the soft information provided at each iteration by a conventional turbo decoder. This approach will be referred to as *turbo synchronization* in the sequel. In [10], a carrier phase recovery algorithm operating in conjunction with the SISO decoders and exploiting the extrinsic information generated at each iteration is proposed. Furthermore, in [11, 12], for instance, it is proposed to combine soft-decision-directed carrier phase estimation with turbo decoding. Tentative decision-aided synchronization within a turbo decoder is reported in [13, 14].

Algorithms in the latter category seem to be promising but they often do not rely on any theoretical basis. The purpose of this paper is therefore to give a mathematical interpretation of such turbo synchronization algorithms and to generalize them. This can be done by means of the expectation-maximization (EM) algorithm. Such an algorithm has been applied to various problems, as in [15], for instance, wherein it is used for channel and noise variance estimation in combination with optimal BCJR-based detection. The same is done in [16] in combination with a suboptimal filter-based equalizer and in [17] for a coded CPM system. In [18], channel gain, and delay estimation is performed in an uncoded CDMA system with a hard-output iterative serial interference canceller. These ideas have been extended to turbo receivers in [19] (see also references therein) and [20] for channel and noise variance estimation in turbo-CDMA and turbo-MIMO contexts, respectively.

In the present paper, we will focus on the specific problem of synchronization. Section 2 will give a general formulation of iterative ML estimation of unknown parameters in the presence of nuisance parameters by means of the EM algorithm. The particular issue of synchronization (i.e., carrier phase, frequency offset, channel gain, and timing estimation) for a digital data-modulated passband signal will then be addressed in Section 3. This implementation will then be extended to the turbo context by showing that the EM algorithm iterations (for parameter estimation) can be combined with those of a turbo receiver (for symbol detection/decoding). This will lead to a general theoretical framework for turbo synchronization. In particular, it will turn out that algorithms introduced in an ad hoc fashion, such as the blind soft-decision-directed carrier phase turbo synchronizer recently proposed in [11], actually correspond to a particular instance of the general scheme proposed here. In order to illustrate the mathematical considerations, in Section 4 we consider as a case study the practical problem of carrier phase and frequency offsets estimation for a turbo-coded 16-QAM system. The relevant simulation results show that the proposed scheme enables to perform blind reliable synchronization and almost ideal coherent detection at very low SNR as required in a turbo receiver. Section 5 considers the computational complexity of the proposed algorithm, whereas a concluding section eventually ends up the paper.

2. ML ESTIMATION IN THE PRESENCE OF A NUISANCE VECTOR

We denote with \mathbf{r} a random vector obtained by expanding the received modulated signal $r(t)$ onto a suitable basis, and we indicate with \mathbf{b} a deterministic vector of parameters to be estimated from the observation of the received vector \mathbf{r} . Assume that \mathbf{r} also depends on a random nuisance parameter vector \mathbf{a} independent of \mathbf{b} and with a priori probability density function (pdf) $p(\mathbf{a})$. The problem addressed in this section is to find the ML estimate $\hat{\mathbf{b}}$ of \mathbf{b} , that is to say, the solution of

$$\hat{\mathbf{b}} = \underset{\tilde{\mathbf{b}}}{\operatorname{argmax}} \{ \ln p(\mathbf{r}|\tilde{\mathbf{b}}) \}. \quad (1)$$

The likelihood function to be maximized with respect to the trial value $\tilde{\mathbf{b}}$ of \mathbf{b} is obtained after elimination of the nuisance parameter vector \mathbf{a} as follows:

$$p(\mathbf{r}|\tilde{\mathbf{b}}) = \int_{\mathbf{a}} p(\mathbf{a})p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}})d\mathbf{a}. \quad (2)$$

In order to solve (1), we take the derivative of $\ln p(\mathbf{r}|\tilde{\mathbf{b}})$ with respect to $\tilde{\mathbf{b}}$ and we equate it to zero, that is,

$$\begin{aligned} \frac{\partial}{\partial \tilde{\mathbf{b}}} \ln p(\mathbf{r}|\tilde{\mathbf{b}}) &= \frac{\int_{\mathbf{a}} p(\mathbf{a})p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) (\partial/\partial \tilde{\mathbf{b}}) \ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) d\mathbf{a}}{\int_{\mathbf{a}} p(\mathbf{a})p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) d\mathbf{a}} \\ &= \int_{\mathbf{a}} \frac{p(\mathbf{a})p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}})}{p(\mathbf{r}|\tilde{\mathbf{b}})} \frac{\partial}{\partial \tilde{\mathbf{b}}} \ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) d\mathbf{a} = \mathbf{0}. \end{aligned} \quad (3)$$

Now, it is easily seen using Bayes' rule that the first factor in the integrand into (3) is nothing but the a posteriori conditional pdf $p(\mathbf{a}|\mathbf{r}, \tilde{\mathbf{b}})$ of the nuisance vector

$$\frac{p(\mathbf{a})p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}})}{p(\mathbf{r}|\tilde{\mathbf{b}})} = p(\mathbf{a}|\mathbf{r}, \tilde{\mathbf{b}}). \quad (4)$$

Therefore, the ML estimation problem given by (1), (2), and (3) is turned into

$$\begin{aligned} \frac{\partial}{\partial \tilde{\mathbf{b}}} \ln p(\mathbf{r}|\tilde{\mathbf{b}}) &= \int_{\mathbf{a}} p(\mathbf{a}|\mathbf{r}, \tilde{\mathbf{b}}) \frac{\partial}{\partial \tilde{\mathbf{b}}} \ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) d\mathbf{a} \\ &= E_{\mathbf{a}} \left\{ \frac{\partial}{\partial \tilde{\mathbf{b}}} \ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) \mid \mathbf{r}, \tilde{\mathbf{b}} \right\} = \mathbf{0}. \end{aligned} \quad (5)$$

In other words, the ML estimate $\hat{\mathbf{b}}$ of \mathbf{b} is that value that nulls the conditional a posteriori expectation of the derivative with respect to $\tilde{\mathbf{b}}$ of the conditional log-likelihood function (LLF) $\ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}})$.

Finding the solution of (5) is not trivial, since $\tilde{\mathbf{b}}$ appears in *both* factors of the integrand. Thus, we try an iterative method that produces a sequence of values $\hat{\mathbf{b}}^{(n)}$ hopefully converging to the desired solution. In particular, we use the previous sequence value $\hat{\mathbf{b}}^{(n-1)}$ to resolve the conditioning on the first factor of the integrand, and we find the current

solution $\hat{\mathbf{b}}^{(n)}$ by solving the resulting simplified equation that follows:

$$\int_{\mathbf{a}} p(\mathbf{a}|\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) \left[\frac{\partial}{\partial \tilde{\mathbf{b}}} \ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) \Big|_{\tilde{\mathbf{b}}=\hat{\mathbf{b}}^{(n)}} \right] d\mathbf{a} = \mathbf{0}. \quad (6)$$

If the sequence of estimates $\hat{\mathbf{b}}^{(n)}$ yielded by (6) converges to a finite value, that value is a solution of ML equation (5) [21].

Observe now that the first factor of the integrand in (6) does not depend on $\hat{\mathbf{b}}^{(n)}$. Therefore, we can bring the derivative back out of the integral and obtain the equivalent equation

$$\hat{\mathbf{b}}^{(n)} : \frac{\partial}{\partial \tilde{\mathbf{b}}} \left\{ \int_{\mathbf{a}} p(\mathbf{a}|\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) \ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) d\mathbf{a} \right\} \Big|_{\tilde{\mathbf{b}}=\hat{\mathbf{b}}^{(n)}} = \mathbf{0}, \quad (7)$$

that is, the estimate $\hat{\mathbf{b}}^{(n)}$ maximizes the conditional a posteriori expectation of the conditional LLF $\ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}})$:

$$\hat{\mathbf{b}}^{(n)} = \underset{\tilde{\mathbf{b}}}{\operatorname{argmax}} \{ \Lambda(\tilde{\mathbf{b}}, \hat{\mathbf{b}}^{(n-1)}) \}, \quad (8a)$$

$$\begin{aligned} \Lambda(\tilde{\mathbf{b}}, \hat{\mathbf{b}}^{(n-1)}) &= E_{\mathbf{a}} \{ \ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)} \} \\ &= \int_{\mathbf{a}} p(\mathbf{a}|\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) \ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) d\mathbf{a}. \end{aligned} \quad (8b)$$

Formulation (8a)-(8b) of our iterative solution can also be derived by means of the EM algorithm [21, 22, 23]. Consider \mathbf{r} as the ‘‘incomplete’’ observation and $\mathbf{z} \triangleq (\mathbf{r}^T, \mathbf{a}^T)^T$ as the ‘‘complete’’ observation. The EM algorithm states that the sequence $\hat{\mathbf{b}}^{(n)}$ defined by

(i) expectation step (E-step):

$$Q(\tilde{\mathbf{b}}, \hat{\mathbf{b}}^{(n-1)}) = E_{\mathbf{a}} \{ \ln p(\mathbf{z}|\tilde{\mathbf{b}}) | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)} \}, \quad (9a)$$

(ii) maximization step (M-step):

$$\hat{\mathbf{b}}^{(n)} = \underset{\tilde{\mathbf{b}}}{\operatorname{argmax}} \{ Q(\tilde{\mathbf{b}}, \hat{\mathbf{b}}^{(n-1)}) \} \quad (9b)$$

converges to the ML estimate under mild conditions [21, 22]. To make (9a)-(9b) equivalent to (8a)-(8b), we observe that, by using the Bayes rule and considering that the distribution of \mathbf{a} does not depend on the parameter vector to be estimated,

$$\begin{aligned} p(\mathbf{z}|\tilde{\mathbf{b}}) &= p(\mathbf{r}, \mathbf{a}|\tilde{\mathbf{b}}) = p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) p(\mathbf{a}|\tilde{\mathbf{b}}) \\ &= p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) p(\mathbf{a}). \end{aligned} \quad (10)$$

Therefore, substituting (10) in (9a), we get

$$\begin{aligned} Q(\tilde{\mathbf{b}}, \hat{\mathbf{b}}^{(n-1)}) &= \int_{\mathbf{a}} p(\mathbf{a}|\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) \ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) d\mathbf{a} \\ &\quad + \underbrace{\int_{\mathbf{a}} p(\mathbf{a}|\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) \ln p(\mathbf{a}) d\mathbf{a}}_{\zeta}. \end{aligned} \quad (11)$$

The second term ζ in (11) does not depend on $\tilde{\mathbf{b}}$, and as far as the M-step is concerned, it can be dropped. Consequently, the estimation procedure given by (8a)-(8b) and the EM algorithm, defined by (9b) and (11), yield the same sequence of estimates. We explicitly observe that the solution of (1) can be found iteratively by only using a posteriori probabilities $p(\mathbf{a}|\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ and the LLF $\ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}})$.

3. APPLICATION TO SYNCHRONIZATION FOR SOFT-INFORMATION-BASED RECEIVERS

3.1. EM-based synchronization

In this section, we will show how to apply the general framework of the previous section to the estimation of the synchronization parameters for a digital data-modulated band-pass signal. In this context, the nuisance parameter vector \mathbf{a} contains the values of the N unknown (hence random) transmitted symbols, that is, $\mathbf{a}^T = (a_0, \dots, a_{N-1})$. Those symbols take values in an M -point constellation \mathcal{A} (such as M -PSK, M -QAM, etc.) according to some rule. Thus, the vector \mathbf{a} has a probability mass function (pmf) $P(\mathbf{a} = \boldsymbol{\mu})$, with $\boldsymbol{\mu}^T = (\mu_0, \dots, \mu_{N-1})$ and $\boldsymbol{\mu} \in \mathcal{A}^N$. The vector \mathbf{b} contains the synchronization parameters to be estimated, that is, $\mathbf{b}^T = (A, \tau, \nu, \vartheta)$ where A , τ , ν , ϑ are the channel gain, symbol timing, carrier frequency, and phase offsets, respectively. Here, the synchronization parameters are assumed as constant within the received code block. This has the advantage of simplifying notably the processing required by the estimation algorithm while inherently is the main limitation of the approach itself. However, a possible yet reasonable solution to handle a time-varying phase offset (due, e.g., to phase noise) is shown in [24]. The idea is quite simple and consists in subdividing the entire block in a number of subblocks within which the phase can be considered approximately as constant, and then in applying to each of them the soft-information-based estimation procedure proposed above. Further, yet again for the sake of simplicity, we will consider in the sequel an AWGN channel as well. Hence, putting all these facts together, the baseband received signal $r(t)$ can be written as

$$r(t) = A \sum_{k=0}^{N-1} a_k g(t - kT - \tau) e^{j(2\pi\nu t + \vartheta)} + w(t), \quad (12)$$

where T is the symbol period, $g(t)$ is a unit-energy (e.g., square-root raised-cosine) pulse, and $w(t)$ is complex-valued AWGN with power spectral density $2N_0$ (assumed to be known).

Neglecting irrelevant terms independent of \mathbf{a} and \mathbf{b} , the conditional LLF of (12) is

$$\begin{aligned} \ln p(\mathbf{r}|\mathbf{a}, \tilde{\mathbf{b}}) &= -2\tilde{A} \operatorname{Re} \left\{ \sum_{k=0}^{N-1} a_k^* z_k(\tilde{\nu}, \tilde{\tau}) e^{-j\tilde{\vartheta}} \right\} \\ &\quad + \tilde{A}^2 \sum_{k=0}^{N-1} |a_k|^2, \end{aligned} \quad (13)$$

where

$$\begin{aligned} z_k(\tilde{\gamma}, \tilde{\tau}) &\triangleq \int_{-\infty}^{\infty} r(t) e^{-j2\pi\tilde{\gamma}t} g(t - kT - \tilde{\tau}) dt \\ &= [r(t) e^{-j2\pi\tilde{\gamma}t}] \otimes g(-t)|_{t=kT+\tilde{\tau}} \end{aligned} \quad (14)$$

is obtained by frequency precompensating the received signal by the “trial” value $-\tilde{\gamma}$, then applying the result to the matched filter $g(-t)$, and finally sampling the matched filter output at the “trial” instant $kT + \tilde{\tau}$. Substituting (13) into (8b) and dropping the terms which do not depend on $\tilde{\mathbf{b}}$, we get

$$\begin{aligned} \Lambda(\tilde{\mathbf{b}}, \hat{\mathbf{b}}^{(n-1)}) &= -2\tilde{A} \operatorname{Re} \left\{ \sum_{k=0}^{N-1} \left[\int_{\mathbf{a}} a_k p(\mathbf{a} | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) d\mathbf{a} \right]^* z_k(\tilde{\gamma}, \tilde{\tau}) e^{-j\tilde{\delta}} \right\} \\ &\quad + \tilde{A}^2 \sum_{k=0}^{N-1} \left[\int_{\mathbf{a}} |a_k|^2 p(\mathbf{a} | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) d\mathbf{a} \right]. \end{aligned} \quad (15)$$

We now define $\eta_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ and $\rho_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$, the a posteriori mean and a posteriori mean square value of the channel symbol a_k , respectively, as follows:

$$\begin{aligned} \eta_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) &\triangleq \int_{\mathbf{a}} a_k p(\mathbf{a} | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) d\mathbf{a} \\ &= \sum_{\alpha_m \in A} \alpha_m P(a_k = \alpha_m | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)}), \end{aligned} \quad (16a)$$

$$\begin{aligned} \rho_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) &\triangleq \int_{\mathbf{a}} |a_k|^2 p(\mathbf{a} | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) d\mathbf{a} \\ &= \sum_{\alpha_m \in A} |\alpha_m|^2 P(a_k = \alpha_m | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)}). \end{aligned} \quad (16b)$$

$P(a_k = \alpha_m | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ denotes the marginal a posteriori probability (APP) of the k th channel symbol a_k conditioned on the observation \mathbf{r} and on the estimate $\hat{\mathbf{b}}^{(n-1)}$ at the previous $(n-1)$ th step, and α_m the M possible values taken in the constellation A . Equation (15) can then be rearranged as

$$\begin{aligned} \Lambda(\tilde{\mathbf{b}}, \hat{\mathbf{b}}^{(n-1)}) &= -2\tilde{A} \operatorname{Re} \left\{ \sum_{k=0}^{N-1} \eta_k^*(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) z_k(\tilde{\gamma}, \tilde{\tau}) e^{-j\tilde{\delta}} \right\} \\ &\quad + \tilde{A}^2 \sum_{k=0}^{N-1} \rho_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}). \end{aligned} \quad (17)$$

We emphasize the similarity between (13) and (17): the latter is formally obtained from the former by simply replacing the terms a_k and $|a_k|^2$ by their respective a posteriori expected values $\eta_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ and $\rho_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$.

The new estimate $\hat{\mathbf{b}}^{(n)}$ at the n th step is then determined by applying (8a) and therefore by maximizing $\Lambda(\tilde{\mathbf{b}}, \hat{\mathbf{b}}^{(n-1)})$,

given by (17), with respect to $\tilde{\mathbf{b}}$. The corresponding result is

$$[\hat{\mathbf{v}}^{(n)}, \hat{\tau}^{(n)}] = \operatorname{argmax}_{\tilde{\gamma}, \tilde{\tau}} \left\{ \left| \sum_{k=0}^{N-1} \eta_k^*(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) z_k(\tilde{\gamma}, \tilde{\tau}) \right| \right\}, \quad (18a)$$

$$\hat{\mathbf{v}}^{(n)} = \angle \left\{ \sum_{k=0}^{N-1} \eta_k^*(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) z_k[\hat{\mathbf{v}}^{(n)}, \hat{\tau}^{(n)}] \right\}, \quad (18b)$$

$$\hat{A}^{(n)} = \frac{\left| \sum_{k=0}^{N-1} \eta_k^*(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) z_k[\hat{\mathbf{v}}^{(n)}, \hat{\tau}^{(n)}] \right|}{\sum_{k=0}^{N-1} \rho_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})}. \quad (18c)$$

The obtained solution can be interpreted as an *iterative synchronization* procedure, which can be referred to as *soft-decision-directed* (SDD) synchronization. What we call here *soft decisions* are the a posteriori average values $\eta_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ and $\rho_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ of each channel symbol. They are a sort of “weighted average” over all the constellation points according to the respective symbol APPs. Note that, thanks to (16a) and (16b), these a posteriori average values $\eta_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ and $\rho_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ can be computed from the marginals $P(a_k = \alpha_m | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ only. In other words, due to the particular structure of the digital data-modulated signal, the implementation of the iterative ML estimation algorithm only requires the evaluation of the marginal a posteriori symbol probabilities $P(a_k = \alpha_m | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$.

We now concentrate on the evaluation of the marginal a posteriori symbol probabilities. Whereas for uncoded transmission the usual assumption is that data symbols are independent and equally likely (yielding $P(\mathbf{a} = \boldsymbol{\mu}) = M^{-N}$ for all $\boldsymbol{\mu} \in \mathcal{A}^N$), for a coded transmission with code rate λ , we only have a subset $\mathcal{B} \subset \mathcal{A}^N$ of all possible sequences corresponding to $M^{\lambda N}$ legitimate encoder output sequences. Therefore, taking into account that the APP of the symbol sequence \mathbf{a} is given by

$$P(\mathbf{a} = \boldsymbol{\mu} | \mathbf{r}, \tilde{\mathbf{b}}) = \frac{P(\mathbf{a} = \boldsymbol{\mu}) p(\mathbf{r} | \mathbf{a} = \boldsymbol{\mu}, \tilde{\mathbf{b}})}{\sum_{\boldsymbol{\nu} \in \mathcal{B}} P(\mathbf{a} = \boldsymbol{\nu}) p(\mathbf{r} | \mathbf{a} = \boldsymbol{\nu}, \tilde{\mathbf{b}})}, \quad (19)$$

and assuming that

$$P(\mathbf{a} = \boldsymbol{\mu}) = \begin{cases} M^{-\lambda N}, & \boldsymbol{\mu} \in \mathcal{B}, \\ 0, & \boldsymbol{\mu} \notin \mathcal{B}, \end{cases} \quad (20)$$

we get

$$P(\mathbf{a} = \boldsymbol{\mu} | \mathbf{r}, \tilde{\mathbf{b}}) = \begin{cases} \frac{p(\mathbf{r} | \mathbf{a} = \boldsymbol{\mu}, \tilde{\mathbf{b}})}{\sum_{\boldsymbol{\nu} \in \mathcal{B}} p(\mathbf{r} | \mathbf{a} = \boldsymbol{\nu}, \tilde{\mathbf{b}})}, & \boldsymbol{\mu} \in \mathcal{B}, \\ 0, & \boldsymbol{\mu} \notin \mathcal{B}, \end{cases} \quad (21)$$

which relates the APP of the symbol sequence to the conditional likelihood function. Note that the result for uncoded transmission is obtained from (21) by taking $\mathcal{B} = \mathcal{A}^N$. Finally, the marginal APP related to a symbol a_k is obtained by summing the symbol sequence APPs (21) over all symbols a_i with $i \neq k$.

Evaluation of the APPs according to (21) yields a computational complexity that increases exponentially with the sequence length N , as all possible data sequences must be enumerated. However, in systems where the received signal can be modeled as a Markov process, (i.e., transmission over a frequency selective channel, coded systems, MIMO or CDMA systems, etc.), the marginal symbol APPs $P(a_k = \alpha_m | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ can be efficiently obtained using the BCJR algorithm [25], with a complexity that grows only linearly with the sequence length N . Note however that the computations related to the BCJR algorithm must then be carried out once per iteration of the synchronizer.

3.2. Turbo synchronization

The EM-based synchronization procedure proposed in the previous subsection is intrinsically well suited to iterative (turbo) receivers that perform detection/decoding through extrinsic information exchange between SISO stages. Indeed, one usually assumes that such receivers provide, after convergence of the iterative process, soft information that equals channel symbol APPs. This makes synchronization via the EM algorithm and turbo receivers complementary since the symbol APPs needed by the first one can be provided by the second one.

As shown in the previous subsection, the estimation of the synchronization parameters needs at each EM iteration the knowledge of the marginal APPs $P(a_k = \alpha_m | \mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ in order to compute the a posteriori expected values $\eta_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ and $\rho_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ required for the evaluation of (18a), (18b), and (18c). In a strict implementation, this means that at *each EM iteration* the turbo receiver has to reinitialize the extrinsic information, and then has to iterate until the soft information reaches a steady-state value, in order to yield good approximations of the required symbol APPs. It is clear that the main drawback of this approach is the considerable increase in complexity and latency in comparison with the corresponding ideal synchronized turbo receiver, since the turbo system is required to converge at each EM iteration. To deal with such a trouble, an approximate implementation can be used: the turbo decoder is no longer reinitialized and at each EM iteration only one detection/decoding iteration is performed. In other words, the synchronization iterations (EM algorithm) are merged with the detection/decoding ones (turbo decoder). Note that this approximate “merged” procedure strictly differs from the EM algorithm in that performing only one detection/decoding iteration at each EM iteration (especially in the first ones) leads to poorer estimations of the required symbol APPs. To investigate the potential performance degradation that the proposed simplified algorithm may imply, in [26] the BER performance of both the EM-based synchronizer and its approximate version are evaluated in the context of a BICM (bit-interleaved-coded modulation) 8-PSK transmission scheme. The difference between the two different synchronization methods is that at each EM iteration

in the former we make additionally 5 detection/decoding iterations whereas in the latter only 1 detection/decoding iteration is performed. In spite of this rough simplification, the simulation results surprisingly indicate a negligible performance degradation at EM iteration 10, even though the EM-based method exhibits a faster convergence due to a more reliable symbol APPs estimates in the first iterations.

When applied to the specific case of carrier phase estimation for turbo-coded QAM transmission, the proposed approximate implementation leads to the algorithm introduced earlier in an ad hoc fashion in [11, 27], wherein the symbol APPs computed at each turbo decoding iteration are properly combined with the received samples in order to provide a reliable estimate necessary for coherent demodulation. This leads in this case to a sort of “bootstrapping effect,” wherein decoding helps synchronization that in turn aids decoding and so forth. Therefore, more generally it can be concluded that the proposed mathematical framework provides a theoretical justification to the category of ad hoc algorithms which make use of the available soft decisions in a turbo receiver for the purpose of iteratively estimating the synchronization parameters. Furthermore, if one has to deal with a parameter vector \mathbf{b} for which more than one or two parameters have to be estimated at the same time, it may happen that the turbo receiver must be allowed to proceed for more iterations between the synchronization steps. In this more demanding context, the number of needed detection/decoding iterations has to be selected considering the trade-off between the requirement on providing an accurate estimation of the APPs and the corresponding increase in complexity and latency.

As far as the initial parameter estimate $\hat{\mathbf{b}}^{(0)}$ is concerned, we have to point out that convergence of our iterative, EM-like, synchronization algorithm to the true ML estimate is not unconditional. Due to the highly nonlinear properties shown by the turbo decoding process, a good choice of $\hat{\mathbf{b}}^{(0)}$ certainly affects the system performance and is mandatory in order to enable the convergence of the joint detection and decoding scheme. However, finding a “good” initial value and then refining it through an iterative procedure looks like the acquisition/tracking approach. In our context, the issue of the initial acquisition may be solved in general by making a data-aided preliminary estimate based upon a preamble of pilot symbols. With respect to conventional methods, it is clear that by additionally exploiting the APP information, the length of the pilot sequence may be properly reduced, thereby increasing the spectral efficiency of the transmission system. We will also show in the next section that in some cases (e.g., phase estimation considering turbo-coded QAM transmission) no preamble is required, and acquisition (within a multiple of $\pi/2$) is accomplished as well, provided that the estimate is refined block after block. We will call this approach “time-recursive,” and we will reserve the term “iterative” to successive estimation of a parameter on a *single* data block as described above.

4. SIMULATION RESULTS

Theoretical analysis of the proposed algorithms proved to be extremely difficult. We resorted therefore to simulation to derive performance results of the different iterative SDD turbo synchronization algorithms. As a case study, we consider a turbo-coded QAM-modulated transmission scheme. We focus here on the simple case where the channel gain A and the timing offset τ are known to the receiver, so that only the carrier frequency offset ν and phase offset ϑ , assumed to be constant within the received block, have to be estimated. To be more specific, the corresponding joint SDD phase-frequency recovery procedure is based on (18a)-(18b), that is, assuming the estimates of A and τ replaced by their a priori known values can be written as

$$\hat{\nu}^{(n)} = \underset{\tilde{\nu}}{\operatorname{argmax}} \left\{ \left| \sum_{k=0}^{N-1} \eta_k^*(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) z_k(\tilde{\nu}, \tau) \right| \right\}, \quad (22a)$$

$$\hat{\vartheta}^{(n)} = \angle \left\{ \sum_{k=0}^{N-1} \eta_k^*(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) z_k[\hat{\nu}^{(n)}, \tau] \right\}. \quad (22b)$$

The required a posteriori average values $\eta_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$ given by (16a) are evaluated on the basis of the symbol APPs computed at the output of the turbo decoder (see Section 5 for more details). In the sequel, according to the discussion in Section 3.2, only one decoding iteration is performed at each synchronization iteration, in order to limit the overall complexity and latency. Therefore, at the n th iteration the estimate $\hat{\nu}^{(n)}$ is found according to (22a) and used to reevaluate the samples $z_k(\tilde{\nu}, \tau)$ by *frequency* compensating the received signal by $-\hat{\nu}^{(n)}$ and sampling the matched filter output at the “exact” instant $kT + \tau$. Then, the phase estimate $\hat{\vartheta}^{(n)}$ is computed by applying (22b) and eventually employed for *phase* compensating the matched filter output samples for the next decoding iteration. As initial estimates for the iterative synchronization procedure, we took $(\hat{\nu}^{(0)}, \hat{\vartheta}^{(0)}) = (0, 0)$ in (22a)-(22b).

We consider the simple rate $-\lambda = 3/4$ turbo encoder that encompasses parallel concatenation of two identical binary 16-state rate $-1/2$ recursive systematic convolutional (RSC) encoders with generators $g_1 = (31)_8$ and $g_2 = (33)_8$ [28], via a pseudorandom interleaver with block length $L = 1500$ information bits, and an appropriate puncturing pattern so that the block at the turbo-encoder output comprises 2000 coded bits. This binary turbo code is combined with conventional gray-mapped 16-QAM modulation (giving rise to a transmitted block of 500 symbols) adhering to the so-called suboptimum “pragmatic approach” wherein coding and modulation are performed separately, as illustrated in [29]. Simulation results are provided assuming that the carrier frequency and phase offsets are time-invariant on the transmitted data block. In addition, the above offsets change from one block to the next only in the case of the single-block joint SDD carrier recovery approach, whereas they are considered invariant if the time-recursive algorithm is applied.

The baseband-equivalent architecture of such a turbo-coded transmission system and the encoder schematic are depicted in Figures 1a and 1b, respectively. Note that, in contrast with (14), frequency correction is applied *after* matched filtering. Indeed, in the case of $|\nu T| \ll 1$, this modification causes a negligible performance degradation and, more notably, enables a remarkable reduction in the receiver complexity. At the receiver, consistently to the encoding process, pragmatic disjoint demodulation and binary turbo decoding is performed. As for the latter, to decrease its computational complexity we resort to a suboptimal solution given by the Max-Log-MAP algorithm [30]. Further, the symbol APPs required by the turbo synchronization algorithm can be obtained from the coded bits log-likelihood ratios (LLRs) made available at the output of the binary turbo decoder (see Section 5 for more details).

The proposed synchronization algorithm’s performance will be assessed through evaluation of the mean estimated value (MEV) and the root-mean squared estimation error (RMSEE). We will also investigate the overall BER performance of the coded system with carrier recovery as compared to ideal synchronization, taking as main design parameters the number of decoder iterations I and the energy per bit-to-noise spectral density ratio E_b/N_0 .

4.1. MEV curves

Figure 2 depicts the MEV curves (i.e., the average estimated value $E\{\hat{\vartheta}\}$ as a function of the true phase offset ϑ) for the SDD phase recovery algorithm based on (22b) for different numbers of decoder iterations $I = 8, 10, 12$, assuming a null frequency offset and with $E_b/N_0 = 6$ dB (roughly corresponding to BER = 10^{-4} with ideal carrier recovery). The difference between the MEV curves is not significant for phase errors $|\vartheta| \leq 20^\circ$, whatever the number of iterations, whereas with larger phase errors the bias of the algorithm is negligible only for $I = 10, 12$. For the particular transmission scheme of Figure 1a, the rotational invariance is not destroyed and the usual $\pi/2$ estimation ambiguity due to the four-fold symmetry of the QAM constellation is apparent, as can be found in [11]. Note that, if one can afford an increase in complexity, the above problem can be easily handled by evaluating the average value of the absolute soft output of the decoder for different multiples of $\pi/2$, and choosing the phase offset that provides the highest reliability according to the approach illustrated in [12].

The MEV curves illustrated in Figure 2 suggest using this estimator as a sort of phase error detector in a *time-recursive* recovery scheme. This can be done on a block-by-block recursive basis as follows. We denote with $\hat{\vartheta}_m$ the time-recursive phase estimate related to the m th data block and with $\hat{\varphi}_m^{(I)}$ the phase error estimate after I decoding iterations as described above. After a prerotation of the received samples in the $(m+1)$ th data block by $-\hat{\vartheta}_m$ and a new phase error estimate $\hat{\varphi}_{m+1}^{(I)}$, the phase estimate for the $(m+1)$ th data block is computed as

$$\hat{\vartheta}_{m+1} = \hat{\vartheta}_m + \hat{\varphi}_{m+1}^{(I)} \quad (23)$$

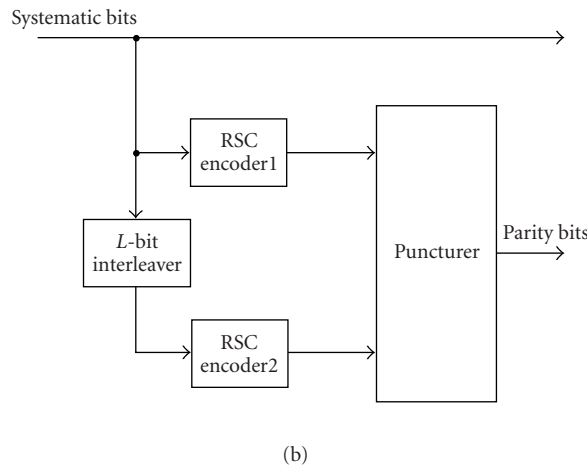
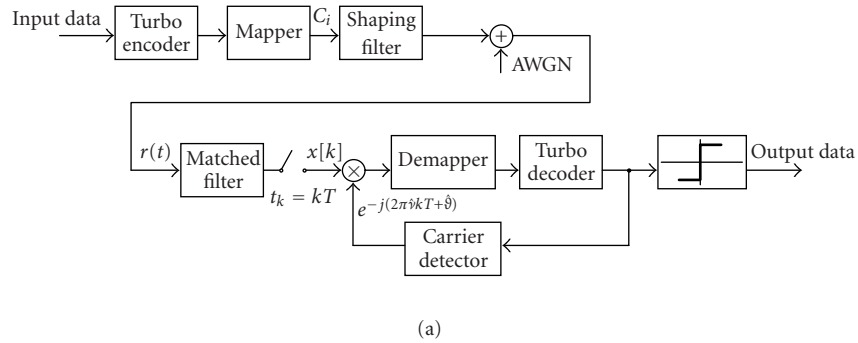


FIGURE 1: (a) Turbo-coded transmission system and (b) turbo-encoder schematic.

assuming as starting condition $\hat{\vartheta}_0 = \hat{\varphi}_0^{(I)}$. We found by simulation that to accomplish an adequate acquisition only 3 blocks are sufficient (i.e., just 3 updates on m in (23)). In doing so, the operating point of the phase error estimator is progressively brought back to the vicinity of the origin, that is, in a negligible-bias zone. Indeed, the results in Figure 3 obtained for $I = 10$ iterations show the improvement of the recursive algorithm with respect to the one based on a single block.

We now tackle the additional issue of carrier frequency recovery. We have to jointly solve (22a) (where the timing offset is considered perfectly known) and (22b). Figure 4 shows the MEV curves for the single-block estimation of the phase offset, for the true values ($\vartheta = 0^\circ, 10^\circ, 20^\circ, 30^\circ$), as a function of the true normalized frequency offset νT . Results are provided for $I = 12$ decoder iterations and $E_b/N_0 = 6$ dB. The joint estimator works fine up to $|\vartheta| \leq 20^\circ$, but the operating interval for frequency recovery is quite narrow, that is, $|\nu T| < 10^{-4}$, if compared with a conventional data-aided method [31]. This can be easily explained if we consider the following fact. For a given block length, the residual frequency offset causes a phase rotation on the received signal samples leading to a considerable performance degradation for the constituent SISO decoders. Clearly, the larger the frequency offset, the larger will also be the phase rotation

on the block samples. Consequently, there exists a threshold value for the frequency offset, such that the overall phase accumulated on a block will be around π , above which the reliability of the decoded bits, even after a few decoding iterations, will stay small. This hinders joint convergence of the (blind) frequency estimator and data decoder.

The time-recursive approach can be used to improve the performance of joint phase-frequency recovery as well. To be more specific, the frequency and phase estimates are used to precorrect the received signal samples in the subsequent block both in frequency and in phase prior to a new iterative estimation. Unfortunately, the improvement for frequency is not as dramatic as for phase estimation, as can be seen from Figure 5. The operating range for the carrier frequency estimator is now $|\nu T| < 3 \cdot 10^{-4}$ for $I = 12$ decoder iterations and $E_b/N_0 = 6$ dB. The conclusion is that some form of “frequency sweeping” is required in order to perform initial frequency acquisition when the offset is larger than the value above. Further enlargement of this range can be alternatively obtained by partitioning the code block into shorter estimation windows, over which we can apply (time-recursive) joint estimation. With shorter windows, a larger frequency operating range is obtained, but the phase estimation accuracy decreases, so that an optimum length will exist for a given E_b/N_0 .

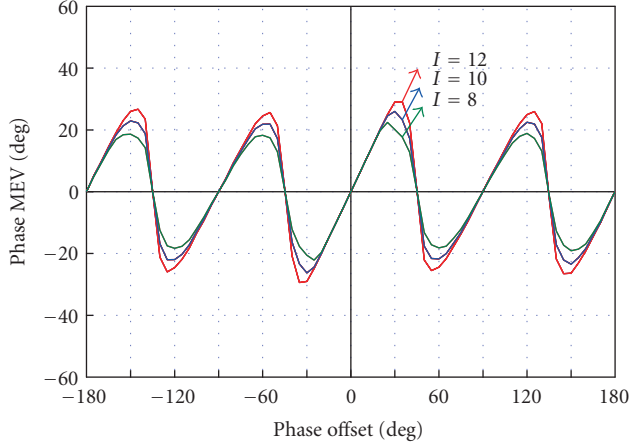


FIGURE 2: MEV curves for single-block phase SDD recovery with different iteration numbers, 16-QAM, $\lambda = 3/4$, $L = 1500$, $E_b/N_0 = 6$ dB.

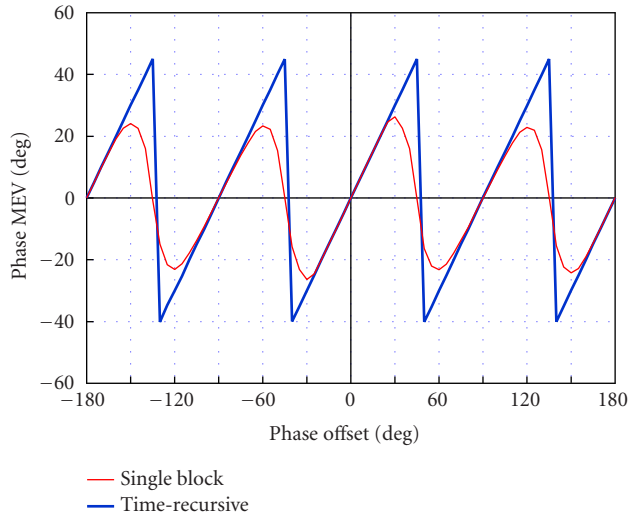


FIGURE 3: MEV curves for time-recursive and single-block SDD phase recovery, 16-QAM, $\lambda = 3/4$, $L = 1500$, $E_b/N_0 = 6$ dB.

4.2. RMSEE curves

Figure 6 shows the curves of RMSEE σ_θ (i.e., $\sqrt{E\{(\hat{\theta} - \vartheta)^2\}}$) of the phase SDD recovery algorithm as a function of E_b/N_0 for various values of the true offset ϑ . The curves are compared to the modified Cramér-Rao bound (MCRB) [31], and with ideal DA estimation that lies exactly on the MCRB. Conversely, the RMSEE performance of SDD gets approximately close to the bound for $E_b/N_0 \geq 6$ dB only, that is, in the interval where soft-data decisions are reliable enough (as will be illustrated in the sequel). It is also noted that the RMSEE curve for conventional hard-decision-directed (HDD) phase estimation, that is, based on the decisions taken at the decoder input, is catastrophic. This is

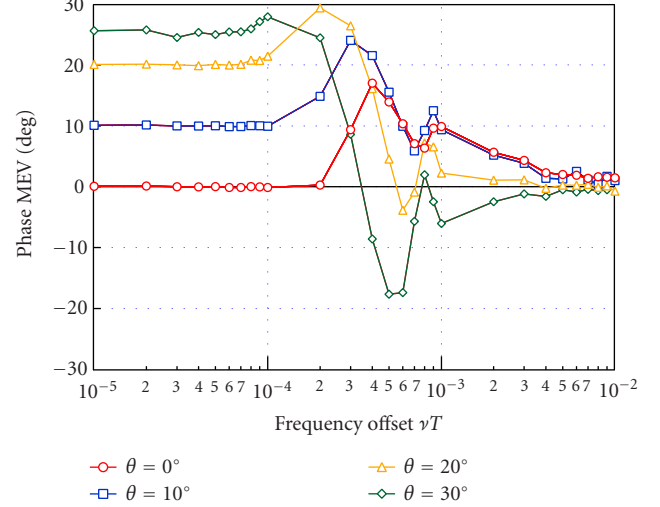


FIGURE 4: MEV phase curves for single-block joint SDD phase-frequency recovery with different true phase, 16-QAM, $\lambda = 3/4$, $L = 1500$, $E_b/N_0 = 6$ dB.

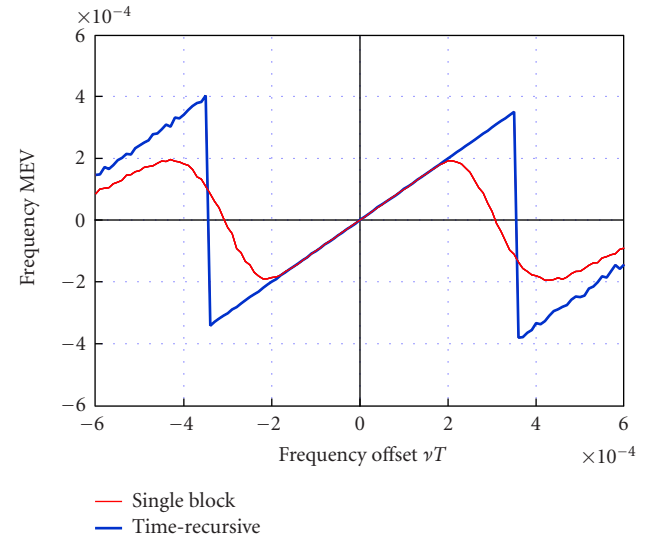


FIGURE 5: MEV curves for time-recursive and single-block joint SDD phase-frequency recovery, 16-QAM, $\lambda = 3/4$, $L = 1500$, $E_b/N_0 = 6$ dB.

easily explained by noting that the BER of hard-detected 16-QAM in our SNR range is definitely poor, leading to an inaccurate phase estimate. On the other side, a different solution is based on applying the proposed iterative estimation algorithm (22b) using the hard-detected QAM symbols taken from the decoder output at each decoding iteration. This kind of scheme can be referred to as iterative hard-decision-directed (IHDD). As illustrated in Figure 6, the performance degradation with respect to SDD of IHDD is small as long as the phase error is $|\vartheta| \leq 10^\circ$, but gets more important for larger values of initial phase offset.

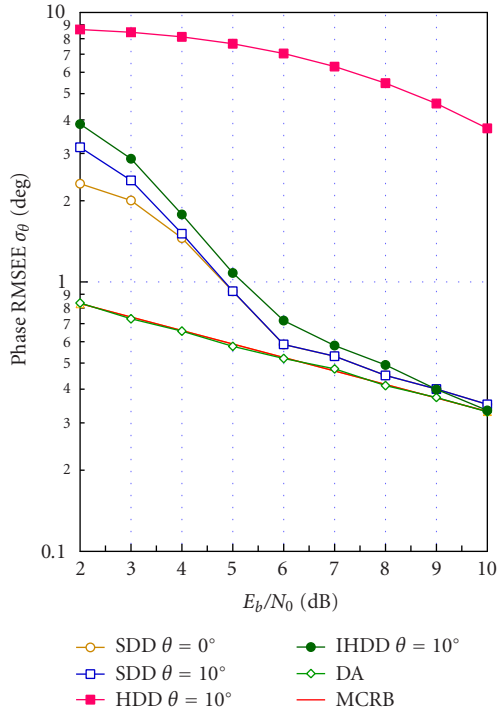


FIGURE 6: RMSEE curves for single-block SDD and IHDD, DA, HDD phase recovery, 16-QAM, $\lambda = 3/4$, $L = 1500$, $I = 12$.

The curves for the frequency RMSEE $\sigma_{\nu T}$ (i.e., $\sqrt{E\{(\hat{\nu} - \nu)^2 T^2\}}$) in Figure 7 follow the same general pattern as those for the phase. As noted for SDD phase recovery, the frequency MCRB bound [31] is attained for $E_b/N_0 \geq 6$ dB, and negligible performance degradation is observed both for the frequency offsets $\nu T = 0$ and $\nu T = 10^{-4}$.

4.3. BER performance

To get a picture about the overall performance of the 16-QAM turbo receiver equipped with the proposed SDD carrier synchronizer, the BER curves can be evaluated as a function of the signal-to-noise ratio E_b/N_0 . For each curve all simulation runs were stopped upon the detection of 100 frame error events. Specifically, Figure 8 shows the BER curves with time-recursive SDD phase recovery and with $I = 10$ iterations. The curves with a phase offset $\vartheta = 20^\circ, 40^\circ$ exhibit a negligible performance degradation with respect to the one with ideal phase recovery. These curves motivate the departure of the RMSEE curves of SDD synchronization from the MCRB. The “knee point” of the RMSEE curves, which roughly corresponds to $E_b/N_0 = 6$ dB, is in fact located in the so-called “waterfall region” (abrupt BER decrease). The associated BER is then sufficiently decreased and the synchronization algorithm performance tends to that of a DA synchronizer. Further, for the sake of completeness, is worthy to point out that a similar behavior is found even with a lower rate, namely 1/2, encoder combined with a 4-QAM modulation format, as shown in the results presented in [11, 27].

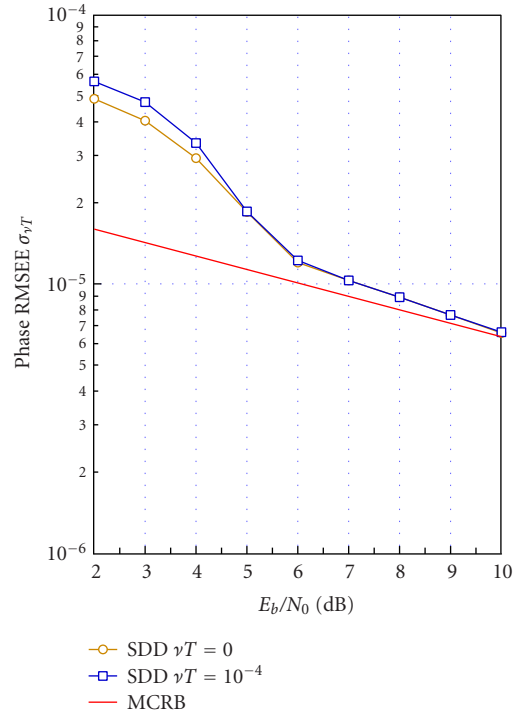


FIGURE 7: RMSEE curves for single-block joint SDD phase-frequency recovery, 16-QAM, $\lambda = 3/4$, $L = 1500$, $I = 12$.

The BER curves for joint SDD phase-frequency recovery are illustrated in Figures 9 and 10 in the case of single-block and time-recursive estimation, respectively. The main result which has to be pointed out is that the performance of single-block-based joint SDD phase-frequency recovery algorithm gets worse for increasing frequency offsets to be estimated, while the time-recursive approach enables to achieve turbo decoding with a negligible degradation with respect to ideal synchronization for a frequency offset up to about $\nu T = 3 \cdot 10^{-4}$. The increased robustness of the time-recursive version of the proposed synchronizer is coherent with what was already observed above in Section 4.1. Indeed, with the iteration of (23) the carrier offset estimation error is progressively reduced despite a nonnegligible initial value due to, for instance, the choice of employing a shorter preamble to achieve a better efficiency.

5. COMPUTATIONAL COMPLEXITY

In this section we focus on the computational complexity of the turbo receiver (whose performance has been evaluated in Section 4) performing soft-decision-based iterative carrier synchronization. In particular, we perform a comparison with the complexity of the turbo receiver with ideal synchronization.

The computational load of both the *iterative* SDD and the *ideal* receiver is dominated by matched filtering, turbo decoding, and carrier synchronization (for the latter only). Depending on the different arrangements for decoding/synchronization, the above functions contribute differently to the overall complexity. For simplicity, we assume that

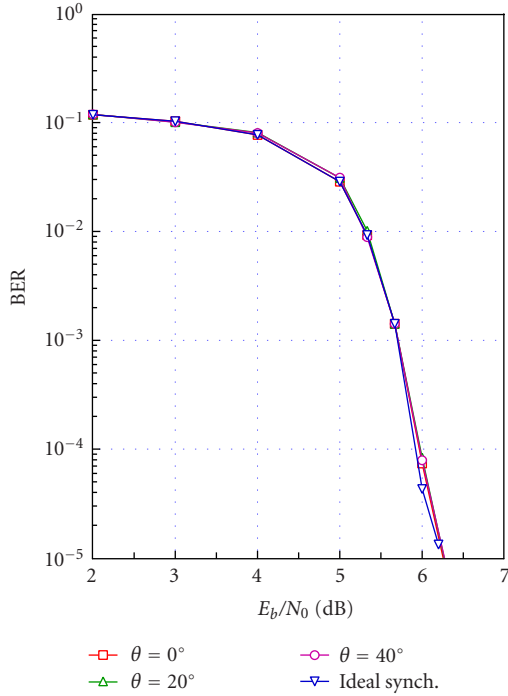


FIGURE 8: BER for time-recursive SDD phase recovery, 16-QAM, $\lambda = 3/4$, $L = 1500$, $I = 10$.

(as adopted in Section 4) the SDD receiver performs at each EM iteration only 1 detection/decoding iteration, whereas for both receivers the matched filtering is carried out only once before applying the decoding and synchronization procedures. Our complexity evaluation is performed on the basis of the number of required floating point (FP) operations, namely additions and multiplications, thereby leaving out (in a first approximation) operations such as comparisons and table lookups.

We denote with L the block length of information bits, with S the number of states of the rate $-1/2$ RSC component decoder and with N the number of transmitted 16-QAM symbols, respectively. The following basic operations have to be performed.

- (OP₀) Matched filtering is based on an FIR filter with an operating frequency equal to $2/T$, where T is the signaling interval. Taking as impulse response a root cosine Nyquist function in the range $(-5T, 5T)$, which corresponds to 20 samples, the relevant computational complexity amounts to $C_0 \cong 80N$.
- (OP₁) Each SISO constituent decoder accomplishes MAP decoding by evaluating the APPs for the systematic bits according to the BCJR algorithm [25]. Specifically, to limit the decoder complexity and avoid multiplications, we adopt the Max-Log-MAP approach illustrated in [30]. This involves the computation of the metrics (related to the states transitions) $\alpha_l(s)$, $\beta_l(s')$, and $\gamma_l(s, s')$ through forward and backward recursions, with s and s' enumerating the trellis states and $1 \leq l \leq L$. As for each decoding iteration two SISO decoders

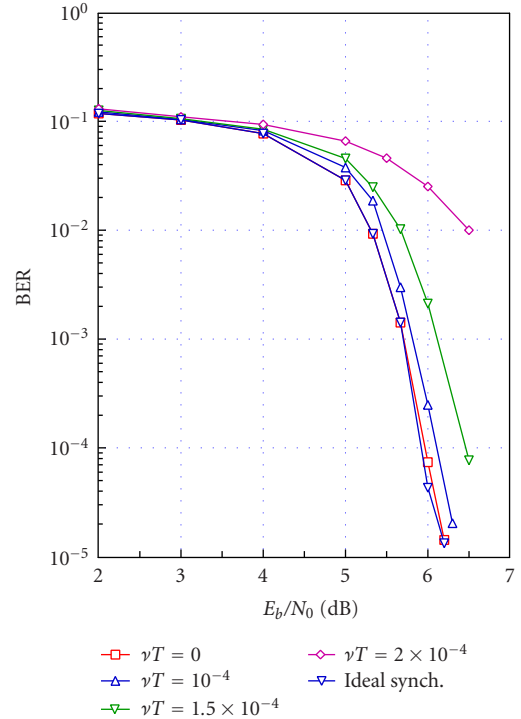


FIGURE 9: BER for single-block joint SDD phase-frequency recovery, 16-QAM, $\lambda = 3/4$, $L = 1500$, $I = 10$.

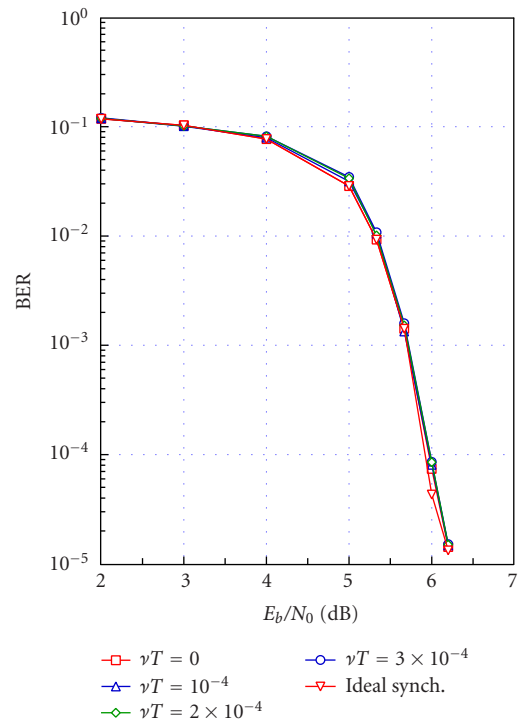


FIGURE 10: BER for time-recursive joint SDD phase-frequency recovery, 16-QAM, $\lambda = 3/4$, $L = 1500$, $I = 10$.

are employed, assuming a total of I iterations, the complexity of this operation is approximately $C_1 = 24S \cdot L \cdot I$.

- (OP₂) In the iterative synchronizer, the APPs for the parity bits are required as well. This means that additional calculations have to be carried out within the two constituent SISO decoders, for a total additional load of roughly $C_2 = 12S \cdot L \cdot I$ operations.
- (OP₃) The matched filter output samples after frequency and phase compensation are exploited to evaluate the initial metrics required by the turbo decoder. Denoting with y_k , $1 \leq k \leq N$, the sequence at the decoder input, the metrics for the four bits associated to the k th transmitted symbol are computed as follows [28]:

$$\begin{aligned} \Gamma_k^{(1)} &= \text{Re}\{y_k\}, & \Gamma_k^{(2)} &= |\text{Re}\{y_k\}| - 2, \\ \Gamma_k^{(3)} &= \text{Im}\{y_k\}, & \Gamma_k^{(4)} &= |\text{Im}\{y_k\}| - 2. \end{aligned} \quad (24)$$

According to (24), the relevant computational complexity can be regarded roughly negligible, that is, $C_3 = 2N \cdot I$.

- (OP₄) The APPs of both the systematic bits and the parity bits provided by the turbo decoder are required in the computation of the a posteriori mean value of the transmitted symbols defined in (16a). After some algebra it can be proved that

$$\begin{aligned} \eta_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)}) &= \tanh \frac{L_n(a_1^{(k)})}{2} \left[2 + \tanh \frac{L_n(a_2^{(k)})}{2} \right] \\ &+ j \tanh \frac{L_n(b_1^{(k)})}{2} \left[2 + \tanh \frac{L_n(b_2^{(k)})}{2} \right], \end{aligned} \quad (25)$$

where $[a_1^{(k)}, a_2^{(k)}, b_1^{(k)}, b_2^{(k)}]$ are the four coded bits associated with the k th transmitted symbol, and $L_n(a_1^{(k)})$, $L_n(a_2^{(k)})$, $L_n(b_1^{(k)})$, and $L_n(b_2^{(k)})$ are the corresponding APPs at the n th decoding iteration, with $1 \leq n \leq I$. Note that in (25) the evaluation of the hyperbolic tangent may be carried out via a proper lookup table, and consequently the complexity of this operation amounts to $C_4 = 4N \cdot I$.

- (OP₅) Using the a posteriori averages $\eta_k(\mathbf{r}, \hat{\mathbf{b}}^{(n-1)})$, the frequency and phase estimates are updated according to (22). The maximization can be carried out through an exhaustive grid search on a small number N_g of trial values (since the operating range of the estimator is narrow), so that the corresponding complexity equals $C_5 = 6N \cdot I \cdot N_g$.

The complexity concerning (OP₀), (OP₃), (OP₄), and (OP₅) is proportional to the block length N , but not to the number of states S . This is the reason why the overall complexity is dominated by (OP₁) and (OP₂). To sum up, the complexity of the SDD receiver is given by

$$C_{\text{SDD}} = \sum_{i=0}^5 C_i \cong 80N + 36S \cdot L \cdot I + 6N \cdot I + 6N \cdot I \cdot N_g. \quad (26)$$

When synchronization is known a priori, (OP₂), (OP₄), and (OP₅) do not apply since they are needed for SDD synchronization. The metrics calculation (OP₃) is carried out only once, and the matched filtering (OP₀) is clearly the same. Therefore, the overall complexity of the ideal receiver is

$$C_{\text{ideal}} = \sum_{i=0}^5 C_i = 80N + 24S \cdot L \cdot I + 2N \cdot I, \quad (27)$$

and the additional complexity introduced by the SDD iterative receiver is

$$\begin{aligned} \Delta C &= \frac{C_{\text{SDD}} - C_{\text{ideal}}}{C_{\text{ideal}}} \cdot 100 \\ &\approx \frac{12S \cdot L \cdot I + 4N \cdot I + 6N \cdot I \cdot N_g}{80N + 24S \cdot L \cdot I + 2N \cdot I} \cdot 100 \approx 50\%. \end{aligned} \quad (28)$$

Summing up, the complexity of the receiver with iterative synchronization is greater, namely around 50%, than that of the receiver with ideal synchronization, because the additional complexity is due mainly to the evaluation of the APPs of the parity bits (OP₂). This price to be paid can be avoided whether one accepts to evaluate them only once at the decoder input. This approximate solution entails a negligible performance degradation in the case of high code rate, that is, when the parity bits are substantially less in number than the information bits, as is shown by the performance results illustrated in Section 4. However, it has to be remarked that for other channel coding schemes also suited to iterative decoding, such as SCCC (serially-concatenated convolutional codes) and LDPC (low-density parity check), at each iteration APPs are available for both systematic and parity bits in a code block, and consequently the incremental complexity due to synchronization is relatively smaller than the complexity of the whole iterative decoder.

6. CONCLUSIONS

The main conclusion of the paper is that ad hoc iterative schemes adopted in the context of joint synchronization and decoding can be justified in a theoretical framework based on the well-known EM algorithm. The resulting estimation procedure can also be easily interpreted as a form of *iterative soft-decision-directed* synchronization, as opposed to conventional *hard-decision-directed* estimation that fails in a condition of low signal-to-noise ratio. Iterative synchronization comes natural in the context of decoding of channels codes suited to iterative detection, such as turbo codes with parallel and serial concatenation, and LDPC codes. This fully justifies the formulation of the so-called *turbo synchronization concept*, that is, soft-decision-directed synchronization and parameter estimation within a turbo (iterative) receiver. As a case study, we demonstrated the application of the proposed mathematical formulation to the particular case of joint carrier phase and frequency offsets estimation in a turbo-coded 16-QAM system. We showed negligible performance degradation with respect to the ideal coherent system down to low signal-to-noise ratios.

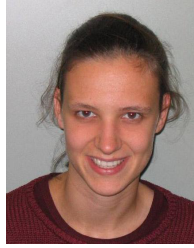
ACKNOWLEDGMENTS

This work has been supported by the Interuniversity Attraction Poles (IAP) Program P5/11, Belgian Science Policy, Federal Office for Scientific, Technical, and Cultural Affairs, Belgium, EU 6FP IST Network of Excellence in Wireless COMmunications (NEWCOM), a project funded by European Community, and FIRB project PRIMO funded by MIUR (Italian Ministry of Education and Research). A. Dejonghe thanks the Belgian "FNRS" for its financial support.

REFERENCES

- [1] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, 1996.
- [2] J. Hagenauer, "The turbo principle: tutorial introduction and state of the art," in *Proc. Int. Symposium on Turbo Codes and Related Topics*, pp. 1–11, Brest, France, September 1997.
- [3] "The turbo principle: from theory to practice. I," *IEEE J. Select. Areas Commun.*, vol. 19, no. 5, 2001.
- [4] "The turbo principle: from theory to practice. II," *IEEE J. Select. Areas Commun.*, vol. 19, no. 9, 2001.
- [5] A. Anastasopoulos and K. M. Chugg, "Adaptive iterative detection for phase tracking in turbo-coded systems," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2135–2144, 2001.
- [6] G. Colavolpe, G. Ferrari, and R. Raheli, "Noncoherent iterative (turbo) decoding," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1488–1498, 2000.
- [7] B. Mielczarek and A. Svensson, "Phase offset estimation using enhanced turbo decoders," in *Proc. IEEE International Conference on Communications (ICC '02)*, vol. 3, pp. 1536–1540, New York, NY, USA, April 2002.
- [8] I. Motedayen-Aval and A. Anastasopoulos, "Polynomial-complexity noncoherent symbol-by-symbol detection with application to adaptive iterative decoding of turbo-like codes," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 197–207, 2003.
- [9] R. Nuriyev and A. Anastasopoulos, "Pilot-symbol-assisted coded transmission over the block-noncoherent AWGN channel," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 953–963, 2003.
- [10] W. Oh and K. Cheun, "Joint decoding and carrier phase recovery algorithm for turbo codes," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 375–377, 2001.
- [11] V. Lottici and M. Luise, "Embedding carrier phase recovery into iterative decoding of turbo-coded linear modulations," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 661–669, 2004.
- [12] A. Burr and L. Zhang, "A novel carrier phase recovery method for turbo-coded QPSK system," in *Proc. European Wireless (EW '02)*, pp. 817–821, Florence, Italy, February 2002.
- [13] C. Morlet, I. Buret, and M.-L. Boucheret, "A carrier phase estimator for multi-media satellite payloads suited to RSC coding schemes," in *Proc. IEEE International Conference on Communications (ICC '00)*, vol. 1, pp. 455–459, New Orleans, La, USA, June 2000.
- [14] C. Langlais and M. Helard, "Phase carrier for turbo codes over a satellite link with the help of tentative decisions," in *2nd International Symposium on Turbo Codes and Related Topics*, vol. 5, pp. 439–442, Brest, France, September 2000.
- [15] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels," *IEEE Trans. Commun.*, vol. 42, no. 7, pp. 2406–2413, 1994.
- [16] R. R. Lopes and J. R. Barry, "Blind iterative channel identification and equalization," in *Proc. IEEE International Conference on Communications (ICC '01)*, vol. 7, pp. 2256–2260, Helsinki, Finland, June 2001.
- [17] C. Brutel, J. Boutros, and P. Mège, "Iterative joint estimation and detection of coded CPM," in *2000 International Zurich Seminar on Broadband Communications*, pp. 287–292, Zurich, Switzerland, February 2000.
- [18] M. Guenach and L. Vandendorpe, "Performance analysis of joint EM/SAGE estimation and multistage detection in UTRA-WCDMA uplink," in *Proc. IEEE International Conference on Communications (ICC '00)*, vol. 1, pp. 638–640, New Orleans, La, USA, June 2000.
- [19] M. Kobayashi, J. Boutros, and G. Caire, "Successive interference cancellation with SISO decoding and EM channel estimation," *IEEE J. Select. Areas Commun.*, vol. 19, no. 8, pp. 1450–1460, 2001.
- [20] C. Lamy, F. Boixadera, and J. Boutros, "Iterative APP decoding and channel estimation for multiple-input multiple-output channels," submitted to *IEEE Trans. on Communications*.
- [21] T. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Mag.*, vol. 13, no. 6, pp. 47–60, 1996.
- [22] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.
- [23] C. Georghiades and J. C. Han, "Sequence estimation in the presence of random parameters via the EM algorithm," *IEEE Trans. Commun.*, vol. 45, no. 3, pp. 300–308, 1997.
- [24] L. Benvenuti, L. Giugno, V. Lottici, and M. Luise, "Code-aware carrier phase noise compensation on turbo-coded spectrally-efficient high-order modulations," in *8th International Workshop on Signal Processing for Space Communications (SPSC '03)*, vol. 1, pp. 177–184, Catania, Italy, September 2003.
- [25] L.R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, 1974.
- [26] N. Noels, C. Herzet, A. Dejonghe, et al., "Turbo-synchronization: an EM interpretation," in *Proc. IEEE International Conference on Communications (ICC '03)*, Anchorage, Alaska, USA, May 2003.
- [27] L. Giugno, V. Lottici, and M. Luise, "Another "turbo-something": Carrier synchronization," in *14th European Signal Processing Conference (EUSIPCO '02)*, vol. 1, pp. 127–131, Toulouse, France, September 2002.
- [28] W. E. Ryan, "A turbo code tutorial," on <http://www.ece.arizona.edu/~ryan>, 1998.
- [29] S. Le Goff, A. Glavieux, and C. Berrou, "Turbo codes and high spectral efficiency modulation," in *Proc. IEEE International Conference on Communications (ICC '94)*, pp. 645–649, New Orleans, La, USA, May 1994.
- [30] P. Robertson, P. Hoeher, and E. Villebrun, "Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo decoding," *European Trans. on Telecomm.*, vol. 8, no. March/April, pp. 119–125, 1997.
- [31] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*, Plenum, New York, NY, USA, 1997.

Nele Noels received the Diploma of Electrical Engineering from Ghent University, Gent, Belgium, in 2001. She is currently a Ph.D. student at the Department of Telecommunications and Information Processing, Ghent University. Her main research interests are in carrier and symbol synchronization. She is the author of several papers in international journals and conference proceedings.



Vincenzo Lottici received the Dr.-Ing. degree (cum laude) in electronic engineering and the Best Thesis SIP Award from the University of Pisa in 1985 and 1986, respectively. From 1987 to 1993 he worked in the design and development of sonar digital signal processing systems. Since 1993 he has been with the Department of Information Engineering at the University of Pisa, where he is currently a Research Fellow and Assistant Professor in telecommunications. His research interests include the area of wireless multicarrier and UWB systems, with particular emphasis on synchronization and channel estimation techniques.



Antoine Dejonghe was born in Ottignies, Belgium, in 1977. He received the Electrical Engineering degree in 2000 from the Université Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium. He is currently pursuing the Ph.D. degree in electrical engineering at the Communications and Remote Sensing Laboratory of UCL. Since 2000, he has been a Research Fellow at the Belgian National Fund for Scientific Research (FNRS). His research interests are in the field of digital communication systems relying on the so-called turbo principle.



Heidi Steendam received the Diploma of Electrical Engineering and the Ph.D. degree in electrical engineering from Ghent University, Gent, Belgium, in 1995 and 2000, respectively. She is Professor at the Department of Telecommunications and Information Processing, Ghent University. Her main research interests are in statistical communication theory, carrier and symbol synchronization, bandwidth-efficient modulation and coding, spread-spectrum (multicarrier spread-spectrum), and satellite and mobile communications. She is the author of more than 50 scientific papers in international journals and conference proceedings.



Marc Moeneclaey received the Diploma of Electrical Engineering and the Ph.D. degree in electrical engineering from Ghent University, Ghent, Belgium, in 1978 and 1983, respectively. He is a Professor at the Department of Telecommunications and Information Processing (TELIN), Ghent University. His main research interests are in statistical communication theory, carrier and symbol synchronization, bandwidth-efficient modulation and coding, spread-spectrum, and satellite and mobile



communications. He is the author of more than 250 scientific papers in international journals and conference proceedings. Together with Professor H. Meyr (RWTH Aachen) and Dr. S. Fechtel (Siemens AG), he coauthors the book *Digital Communication Receivers—Synchronization, Channel Estimation, and Signal Processing* (J. Wiley, 1998). He is the corecipient of the Mannesmann Innovations Prize 2000. During the period 1992–1994, he was the Editor for Synchronization, for the IEEE Transactions on Communications. He served as a Coguest Editor for special issues of the Wireless Personal Communications Journal (on equalization and synchronization in wireless communications) and the IEEE Journal on Selected Areas in Communications (on signal synchronization in digital transmission systems) in 1998 and 2001, respectively.

Marco Luise is a Full Professor of telecommunications at the University of Pisa, Italy. He received the M.S. (cum Laude) and Ph.D. degrees from the University of Pisa, Italy. In the past, he was a Research Fellow of the European Space Agency, a Research Scientist of the Italian National Research Council, and an Associate Professor at the University of Pisa. He chaired the V, VI, VII, and IX editions of the Tyrrhenian International Workshop on Digital Communications, and in 1998 he was the General Chairman of the URSI Symposium ISSSE'98. He was the Technical Cochairman of the 7th International Workshop on DSP Techniques for Space Communications, and of the Conference European Wireless 2002. Professor Luise is a Senior Member of the IEEE, was an Editor of the IEEE Transactions on Communications, and has served as the Coeditor of the Special Issue on Signal Processing in Telecommunications of the European Transactions on Telecommunications (ETT). He has also been the Coeditor of the IEEE JSAC Special Issue on Signal Synchronization in Digital Transmission Systems and Editor for Communication Theory of ETT. His main research interests lie in the area of wireless communications, with particular emphasis on mobile and satellite communication and positioning systems.



Luc Vandendorpe was born in Mouscron, Belgium, in 1962. He received the Electrical Engineering degree (summa cum laude) and the Ph.D. degree from the Université Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, in 1985 and 1991, respectively. Since 1985, Dr. Vandendorpe is with the Communications and Remote Sensing Laboratory of UCL. In 1992, he was a Research Fellow at the Delft Technical University. From 1992 to 1997, Dr. Vandendorpe was a Senior Research Associate of the Belgian NSF at UCL. Presently he is a Professor. He is mainly interested in digital communication systems: equalization, joint detection/synchronization for CDMA, OFDM (multicarrier), MIMO and turbo-based communications systems, and joint source/channel (de-)coding. In 1990, he was a corecipient of the Biennial Alcatel-Bell Award. In 2000, he was a corecipient of the Biennial Siemen. Dr. Vandendorpe is or has been a TPC Member for IEEE VTC Fall 1999, IEEE Globecom 2003 Communications Theory Symposium, the 2003 Turbo Symposium, IEEE VTC Fall 2003, and IEEE SPAWC 2005. He is a Cotechnical Chair (with P. Duhamel) for IEEE ICASSP 2006. He is the Associate Editor of the IEEE Transactions on Wireless Communications, Associate Editor of the IEEE Transactions on Signal Processing, and a Member of the Signal Processing Committee for Communications.



Adaptive Blind Multiuser Detection over Flat Fast Fading Channels Using Particle Filtering

Yufei Huang

*Department of Electrical Engineering, The University of Texas at San Antonio, San Antonio, TX 78249-06615, USA
Email: yhuang@utsa.edu*

Jianqiu (Michelle) Zhang

*Department of Electrical and Computer Engineering, University of New Hampshire, Durham, NH 03824, USA
Email: jianqiu.zhang@unh.edu*

Isabel Tienda Luna

*Departamento de Física Aplicada, Universidad de Granada, Granada 18071, Spain
Email: isabelt@ugr.es*

Petar M. Djurić

*Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY 11794-2350, USA
Email: djuric@ece.sunysb.edu*

Diego Pablo Ruiz Padillo

*Departamento de Física Aplicada, Universidad de Granada, Granada 18071, Spain
Email: druiz@ugr.es*

Received 30 April 2004; Revised 16 September 2004

We propose a method for blind multiuser detection (MUD) in synchronous systems over flat and fast Rayleigh fading channels. We adopt an autoregressive-moving-average (ARMA) process to model the temporal correlation of the channels. Based on the ARMA process, we propose a novel time-observation state-space model (TOSSM) that describes the dynamics of the addressed multiuser system. The TOSSM allows an MUD with natural blending of low-complexity particle filtering (PF) and mixture Kalman filtering (for channel estimation). We further propose to use a more efficient PF algorithm known as the stochastic M -algorithm (SMA), which, although having lower complexity than the generic PF implementation, maintains comparable performance.

Keywords and phrases: multiuser detection, time-observation state-space model, fading channel estimation, particle filtering, mixture Kalman filter.

1. INTRODUCTION

When multiuser detection (MUD) was introduced in the eighties, it has received a great deal of attention due to its ability to reduce multiple access interference (MAI) and potential for increasing the capacity of CDMA systems. Since then, numerous detectors have been proposed in the literature for both synchronous and asynchronous transmission

and some popular ones include the decorrelating detector, the minimum mean square error (MMSE) detector, the multistage detector, and the decision feedback detector [1].

In practice, distortion in signal strength due to time-varying fading channels must be attended while performing MUD. Even though noncoherent detection methods as proposed in [2] are often appealing owing to their simplicity since no inference on fading channels is needed, coherent detection has been proved to deliver better performance [3]. With coherent detection, estimation of channels can be obtained with or without pilot signals. Between them, significant amount of research has been devoted to

This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

schemes without using pilot signals, or blind MUD methods. Blind MUD methods are bandwidth more efficient and the approaches proposed, to name a few, include the recursive least square (RLS) [4, 5], subspace-based [6], expectation-maximization [7], genetic algorithm [8] and Kalman filtering [9, 10, 11, 12, 13, 14]. However, most of the approaches cited above assume slow or quasi-static fading channels.

In this paper, we focus on blind MUD for fast flat Rayleigh fading channels and in synchronous systems. In particular, we assume to know *a priori* the second-order statistics of the underlying channel, based on which a mathematical tractable approximation using autoregressive-moving-average (ARMA) model is adopted. The approximation enables a dynamic state-space modeling (DSSM) of the problem, which lends itself naturally to a Kalman-filtering-related detection solution. The use of Kalman filtering for blind MUD on similar modeling has been seen in [11, 12, 14], where the decision-directed approach was used to estimate the channel variable necessary for the Kalman filtering. One inherent drawback with the decision-directed approach is the error propagation, which greatly limits the performance of such implementation.

Recently, the combined (mixture) Kalman filtering and sequential importance sampling (particle filtering) algorithms have been applied to blind detection of convolutional codes [15], space-time trellis codes [16], and blind MUD [17] over fading channels. The mixture Kalman filtering (MKF) approach is shown to greatly reduce the error propagation of the decision-directed implementations and thus exhibits considerable performance improvement. However, in the proposed use of the MKF to blind MUD in [17], particle filtering (PF) was mainly intended for channel tracking and the embedded MUD at a symbol interval was achieved by an optimum detector, which has exponential complexity with the number of users. Consequently, the proposed MKF algorithm becomes prohibitively complex even for systems with moderate number of users.

In this paper, unlike all existing Kalman filtering detectors, a completely different viewpoint to multiuser systems is taken and we propose a novel time-observation state-space model (TOSSM). Even though the TOSSM is equivalent to the common DSSM, it allows the PF-based multiuser detection to be naturally blended with the mixture Kalman filtering for channel estimation. The new mixture Kalman filtering algorithm samples one user at a time and therefore permits efficient implementation. We further propose to use a more efficient PF algorithm known as the stochastic M -algorithm (SMA), which has shown to attain additional complexity reduction over the generic PF implementation and yet maintain comparable performance.

The rest of the paper is organized as follows. In Section 2, the problem of blind MUD is formulated. In Section 3, a novel TOSSM is described and in Section 4, the optimum solution is discussed. Particle filtering and SMA solutions are proposed in Sections 6 and 7, respectively. The simulation results are presented in Section 8. Section 9 contains some concluding remarks.

2. PROBLEM FORMULATION

Consider a synchronous CDMA system with a processing gain C and K users. Let T denote the symbol duration and $s_k(t)$ the normalized deterministic signature waveform assigned to the k th user. Then, at the n th symbol interval, the received signal $y(t)$ can be expressed as a summation of K antipodally modulated synchronous signature waveforms plus noise, that is,

$$y(t) = \sum_{k=1}^K a_{n,k} b_{n,k} s_k(t) + u(t), \quad t \in [(n-1)T, nT], \quad (1)$$

where $b_{n,k} \in \{-1, +1\}$ is the BPSK modulated bit transmitted by the k th user, $a_{k,n}$ the CSI (fading coefficient) of the k th user, and $u(t)$ the received zero mean additive complex white Gaussian noise with variance σ^2 . The cross-correlation between the signature waveforms of the users is given by the cross-correlation matrix \mathbf{R} , where element r_{k_1, k_2} represents the cross-correlation between the signature waveform of the k_1 th and the k_2 th user and is defined as

$$r_{k_1 k_2} = \langle s_{k_1}, s_{k_2} \rangle = \int_{(n-1)T}^{nT} s_{k_1}(t) s_{k_2}^*(t) dt. \quad (2)$$

The channel for each user is considered as Rayleigh flat fading channel and ARMA processes can be adopted to model its time correlation with satisfaction [11, 15, 18]. Given an ARMA(r_1, r_2) process, the CSI of the k th user at the n th interval $a_{k,n}$ can be represented as

$$\begin{aligned} a_{n,k} &+ \phi_{k,1} a_{n-1,k} \cdots \phi_{k,r_1} a_{n-r_1,k} \\ &= \rho_{k,0} v_{n,k} + \cdots + \rho_{k,r_2} v_{n-r_2,k}, \end{aligned} \quad (3)$$

where $v_{n,k}$ is an i.i.d. random complex Gaussian process that drives the ARMA process, $\{\phi_{k,1}, \dots, \phi_{k,r_1}\}$ and $\{\rho_{k,1}, \dots, \rho_{k,r_2}\}$ are the AR and MA coefficients of the model. We assume that we know *a priori* the second-order statistics of the underlying fading channel, and therefore the coefficients of the ARMA model can be precomputed so that the power spectral density of the ARMA process matches that of the fading channel. For convenience, we assume that $r_1 = r_2 = r$; otherwise zeros can be padded to the coefficients to make the orders equal.

An equivalent form of (1) consists of a set of sufficient statistics represented by the matched filter output,

$$y_{n,k} = \langle y(t), s_k(t) \rangle = \int_{(n-1)T}^{nT} y_n(t) s_k(t) dt. \quad (4)$$

The set of matched filter outputs $\mathbf{y}_n = [y_{n,1}, \dots, y_{n,K}]^T$, where $(\cdot)^T$ stands for matrix transpose, can be represented in vector-matrix form as

$$\mathbf{y}_n = \mathbf{R} \mathbf{A}_n \mathbf{b}_n + \mathbf{u}_n, \quad (5)$$

where $\mathbf{A}_n = \text{diag}\{a_{n,1}, \dots, a_{n,K}\}$ is the diagonal matrix of the channel state information, $\mathbf{b}_n = [b_{n,1}, \dots, b_{n,K}]^T$ is the user data vector, and \mathbf{u}_n is the complex Gaussian noise vector with independent real and imaginary components and with covariance matrix equal to $\sigma^2 \mathbf{R}$. Our objective is to perform sequential symbol detection without knowing the CSI $a_{n,k}$, that is, blind multiuser detection.

3. TIME-OBSERVATION STATE-SPACE SYSTEM MODELING

A succinct mathematical representation of a time-varying system is the dynamic state-space model (DSSM). The state-space representation of CDMA systems in flat fading channels can be found in the existing literatures [11] and it can be expressed as

$$\begin{aligned} \mathbf{h}_{k,n} &= \mathbf{Q}_k \mathbf{h}_{k,n-1} + \mathbf{g} v_{k,n} \quad \forall k, \\ a_{k,n} &= \rho_k^T \mathbf{h}_{k,n} \quad \forall k, \\ \mathbf{y}_n &= \mathbf{R} \mathbf{A}_n \mathbf{b}_n + \mathbf{u}_n, \end{aligned} \quad (6)$$

where $\mathbf{h}_{n,k}^T = [h_{n,k} \cdots h_{n-r,k}]$ is an $(r+1) \times 1$ channel state vector, $\rho_k^T = [\rho_{k,0} \cdots \rho_{k,r}]$,

$$\begin{aligned} \mathbf{Q}_k &= \begin{bmatrix} -\phi_{k,1} & \cdots & -\phi_{k,r} & 0 \\ 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}, \\ \mathbf{g} &= \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \end{aligned} \quad (7)$$

In (6), $\mathbf{h}_{k,n}$ for all k and \mathbf{b}_n are the unknowns to be estimated. Note that the observation \mathbf{y}_n is not linear in $\mathbf{h}_{k,n}$ for all k and \mathbf{b}_n , and therefore the Kalman filter cannot provide the optimum solution. In fact, the optimum solution can be obtained by a so-called splitting Kalman filter, where, at time n , 2^n Kalman filters are required. The complexity of the splitting Kalman filter is exponential with both time and users and thus computational prohibited. Instead, particle filtering can be used to obtain good approximations of the optimum solution with reduced complexity. PF algorithms on (6) incorporated with Kalman filtering were proposed in [17]. However, as mentioned in the introduction, due to the structure of (6), particles of \mathbf{b}_n must be sampled jointly, and the complexity becomes exponential with the number of users. The prohibitive complexity on large user systems implies that this PF algorithm is infeasible for practical applications. To circumvent this difficulty, in the following we introduce a time-observation state-space model (TOSSM) for the system:

$$\begin{aligned} p(\mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK}) &= p(\bar{\mathbf{y}}_{NK} | \mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1}) p(\mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1}) \\ &= p(\bar{\mathbf{y}}_{NK} | \mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1}) \\ &\quad \times p(b_{N,K} | \mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{\mathbf{y}}_{1:NK-1}) \\ &\quad \times p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{\mathbf{y}}_{1:NK-1}) \\ &= p(\bar{\mathbf{y}}_{NK} | \mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1}) p(b_{N,K}) \\ &\quad \times p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{\mathbf{y}}_{1:NK-1}) \\ &= p(\bar{\mathbf{y}}_{NK} | \mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1}) \\ &\quad \times p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{\mathbf{y}}_{1:NK-1}). \end{aligned} \quad (8)$$

In developing the TOSSM, we start with the Cholesky factorization of the cross-correlation matrix \mathbf{R} as

$$\mathbf{R} = \mathbf{F}^T \mathbf{F}, \quad (9)$$

where \mathbf{F} is a uniquely defined $K \times K$ lower triangular matrix. Now, right multiplying $(\mathbf{F}^T)^{-1}$ with the matched filter output, we obtain

$$\bar{\mathbf{y}}_n = (\mathbf{F}^T)^{-1} \mathbf{y}_n = \mathbf{F} \mathbf{A}_n \mathbf{b}_n + \bar{\mathbf{u}}_n \quad (10)$$

or, equivalently,

$$\bar{\mathbf{y}}_n = \mathbf{F} \mathbf{B}_n \mathbf{a}_n + \bar{\mathbf{u}}_n, \quad (11)$$

where $\mathbf{B}_n = \text{diag}\{b_{n,1}, \dots, b_{n,K}\}$ is the diagonal user data matrix, and $\mathbf{a}_n = [a_{n,1}, \dots, a_{n,K}]$ is the $K \times 1$ vector of CSI. Since the covariance matrix of $\bar{\mathbf{u}}_n$ becomes $E[\bar{\mathbf{u}}_n \bar{\mathbf{u}}_n^T] = \sigma^2 \mathbf{F}^{-T} \mathbf{R} \mathbf{F}^{-1} = \sigma^2 \mathbf{I}$, where \mathbf{I} is an identity matrix, $\bar{\mathbf{y}}_n$ is called the whitened matched filter (WMF) output. Next, define a tall channel vector of $K(r+1) \times 1$ as $\mathbf{h}_n = [\mathbf{h}_{1,n}^T \cdots \mathbf{h}_{K,n}^T]^T$ and the channel transition becomes

$$\mathbf{h}_n = \mathbf{Q} \mathbf{h}_{n-1} + \mathbf{G} \mathbf{v}_n, \quad (12)$$

where $\mathbf{v}_n = [v_{1,n}, \dots, v_{K,n}]^T$, $\mathbf{Q} = \text{diag}(\mathbf{Q}_1, \dots, \mathbf{Q}_K)$, and $\mathbf{G} = \text{diag}(\underbrace{\mathbf{g}, \dots, \mathbf{g}}_K)$ are $K(r+1) \times K(r+1)$ and $K(r+1) \times K$ matrices.

We can thus express \mathbf{a}_n by \mathbf{h}_n in a compact form by

$$\mathbf{a}_n = \mathbf{P} \mathbf{h}_n, \quad (13)$$

where $\mathbf{P} = \text{diag}(\rho_1^T, \dots, \rho_K^T)$ is of dimension $K \times K(r+1)$. Now by replacing \mathbf{a}_n in (11) by (13), we have

$$\bar{\mathbf{y}}_n = \mathbf{F} \mathbf{B}_n \mathbf{P} \mathbf{h}_n + \bar{\mathbf{u}}_n. \quad (14)$$

If we denote the k th row of \mathbf{F} by \mathbf{f}_k^T , the k th WMF output \bar{y}_n can be written as

$$\bar{y}_{n,k} = \mathbf{f}_k^T \mathbf{B}_n \mathbf{P} \mathbf{h}_n + \bar{u}_{n,k}, \quad (15)$$

where $\bar{u}_{n,k}$ is the k th element of $\bar{\mathbf{u}}_n$. Now, instead of considering the system evolving only along time, we imagine a system progressing alternately along the path of time and the WMF observations $\bar{y}_{n,k}$. The concept is further illustrated in Figure 1. To describe this new system, we must collapse the time index n and the observation index k into one time-observation index l , where $l = (n-1)K + k$. This conversion is reversible or, in other words, we can also calculate k and n from l by $k = \text{mod}(l, K)$ and $n = (l-k)/K + 1$, where $\text{mod}(k, K)$ is the k modulo K operation. In the following description of the TOSSM indexed by l , all k and n are assumed to be obtained from the corresponding l . Now, we introduce a $K \times K$ auxiliary matrix $\tilde{\mathbf{B}}_l = \text{diag}\{b_{n,1}, \dots, b_{n,k}, 0, \dots, 0\}$. The state-space representation for the new time-observation system indexed by l can be then constructed as

$$\begin{aligned} \mathbf{h}_l &= \begin{cases} \mathbf{Q} \mathbf{h}_{l-1} + \mathbf{G} \mathbf{v}_l & \text{if } k = 1, \\ \mathbf{h}_{l-1} & \text{if } k \neq 1, \end{cases} \\ \bar{y}_l &= \mathbf{f}_k^T \tilde{\mathbf{B}}_l \mathbf{P} \mathbf{h}_l + \bar{u}_l \end{aligned} \quad (16)$$

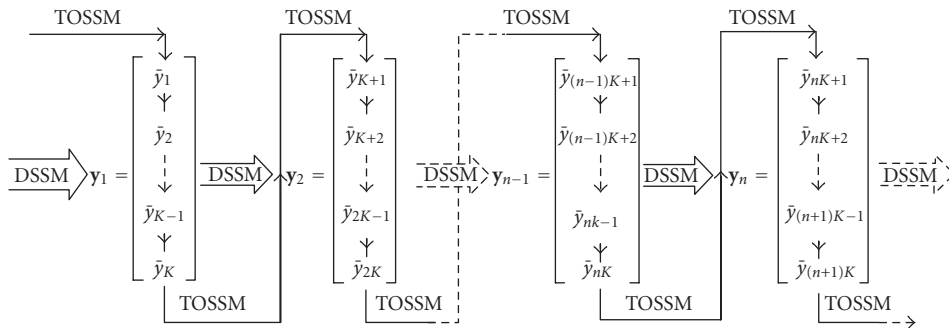


FIGURE 1: Illustrative plot of the TOSSM.

and we call (16) the TOSSM. Note that (16) and (6) describe the same system. There are, however, key differences between the two models. Unlike (6), the state transitions of \mathbf{h}_l in the TOSSM are time (or index) varying, that is, at different l , different transition is applied. Specifically, when $k = 1$ or, equivalently, n increases by 1, \mathbf{h}_l updates according to the ARMA channel model, and otherwise when $k \neq 1$ and n remains unchanged from $l - 1$, \mathbf{h}_l is assumed to be static. Additionally, in the TOSSM, the number of the unknown user bits changes with l and especially, only one new unknown signal $b_{n,k}$ is included each time when l is incremented by one. Therefore, if we assume perfect detection at $l - 1$, that is, $b_{n,1}, \dots, b_{n,k-1}$ are known exactly, then there is only one unknown user bit to be detected. Note that in the conventional DSSM (6), K unknown users bits need to be detected altogether as the system evolves to time n . This is the key of the model that leads to efficient particle filtering solutions. We, however, want to stress that the decision on $b_{n,k}$ (except $k = K$) is not finalized at l . Since the observations from y_{l+1} up to y_{l+r} with $r = K - k$ all contain information about $b_{n,k}$, the final decision is reached only at $l + r$, or in general, when $k = K$.

4. OPTIMUM BAYESIAN BLIND DETECTION

In a Bayesian framework, the optimum decision on \mathbf{b}_N can be obtained by the marginalized posterior mode (MPM) criterion, which is expressed as

$$(\hat{b}_{N,k})_{\text{MPM}} = \text{sgn} \left(\sum_{\mathbf{b}_N \in \{-1,1\}^{NK}} b_{n,k} p(\mathbf{b}_N | \bar{y}_{1:NK}) \right), \quad (17)$$

where $p(\mathbf{b}_N | \bar{y}_{1:NK})$ is the posterior distribution that is essential for computing (17) and the subscript $1:NK$ denotes a collection of the variable indexed from 1 to NK , e.g., $\bar{y}_{1:NK} = \{\bar{y}_1, \dots, \bar{y}_{NK}\}$. Notice that the posterior distribution $p(\mathbf{b}_N | \bar{y}_{1:NK})$ is independent of $\mathbf{b}_{1:(N-1)}$, that is, the bits transmitted prior to time n . Further, the marginalization in (17) suggests that $(\hat{b}_{N,k})_{\text{MPM}}$ is also independent of other users' bits transmitted at n . Therefore, the MPM solution is immune to decision errors on $\mathbf{b}_{1:(N-1)}$ and other users' bits transmitted at n .

Now, to derive $p(\mathbf{b}_N | \bar{y}_{1:NK})$, marginalization on $p(\mathbf{b}_{1:N} | \bar{y}_{1:NK})$ over $\mathbf{b}_{1:(N-1)}$ is needed, that is,

$$\begin{aligned} p(\mathbf{b}_N | \bar{y}_{1:NK}) &= \sum_{\mathbf{b}_{1:N-1}} p(\mathbf{b}_{1:N} | \bar{y}_{1:NK}) \\ &= \frac{\sum_{\mathbf{b}_{1:N-1}} p(\mathbf{b}_{1:N}, \bar{y}_{1:NK})}{\sum_{\mathbf{b}_{1:N}} p(\mathbf{b}_{1:N}, \bar{y}_{1:NK})}. \end{aligned} \quad (18)$$

Considering the TOSSM (16), we found the joint distribution in (8), where the last equation was obtained by assuming the noninformative priors for $\mathbf{b}_{N,K}$, that is, $p(b_{N,K} = 1) = 0.5$. Equation (8) indicates a recursive calculation of $p(\mathbf{b}_{1:N}, \bar{y}_{1:NK})$ from $l = 1$ to NK through multiplying the marginal likelihood $p(\bar{y}_l | \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l})$ at each recursion. These likelihoods $p(\bar{y}_l | \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l})$ for $l = 1, \dots, NK$ are obtained by marginalizing the channel state vector \mathbf{h}_l from $p(\bar{y}_l, \mathbf{h}_l | \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l})$, and we show in the appendix that

$$p(\bar{y}_l | \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l}) = \mathcal{N}(m_l, c_l) \quad (19)$$

and the mean m_l and variance c_l can be calculated sequentially through the Kalman filter. This is equivalent to say that $p(\bar{y}_l | \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1})$ can be calculated from a run of the Kalman filter. Now, revisiting (18), we see that, to calculate $p(\mathbf{b}_N | \bar{y}_{1:NK})$, $p(\bar{y}_l | \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1})$ must be evaluated for 2^{NK} combinations of $\mathbf{b}_{1:N}$, or 2^{NK} Kalman filters are needed, each of which corresponding to one possible combination. As a result, totally 2^{NK} Kalman filters are required for the MPM solution. The expansion of the numbers of the Kalman filters with l presents a tree structure illustrated in Figure 2. The MPM solution has thus a complexity exponentially increasing with both time n and the number of users K . This is apparently a formidable task not possible for real applications. We, therefore, must resort to suboptimum solutions with manageable complexity. One choice is particle filtering.

5. A DECISION-DIRECTED APPROACH TO BLIND MUD

A decision-directed approach to blind MUD was proposed in [11] based on DSSM (6). We describe in the following a corresponding decision-directed approach on the TOSSM (16).

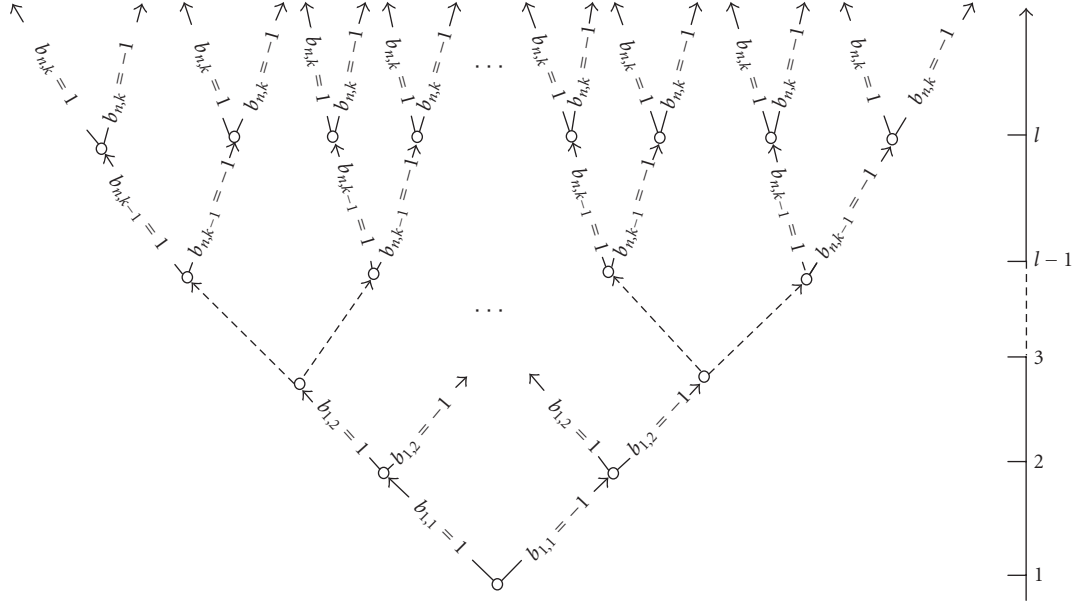


FIGURE 2: The tree structure of the optimum solution. Each path in the tree represents a run of the Kalman filter.

<p>Predictive step: $\tilde{\mathbf{h}}_l = \begin{cases} \mathbf{Q}\hat{\mathbf{h}}_{l-1} + \mathbf{G}\mathbf{u}_l & \text{if } k = 1, \\ \hat{\mathbf{h}}_{l-1} & \text{if } k \neq 1, \end{cases}$</p> <p>$\hat{\Sigma}_l = \begin{cases} \mathbf{Q}\hat{\Sigma}_{l-1}\mathbf{Q}^T + \sigma^2\mathbf{G}\mathbf{G}^T & \text{if } k = 1, \\ \hat{\Sigma}_{l-1} & \text{if } k \neq 1. \end{cases}$</p> <p>Detection step:</p> <p>$\hat{b}_{n,k} = \text{sgn}(z_{n,k});$</p> <p>$z_{n,k} = (\bar{y}_l - \sum_{j=1}^{k-1} f_{k,j} a_{i,j} b_{i,j}) a_{i,k}^*;$</p> <p>$a_{i,k} = \rho_k \tilde{\mathbf{h}}_l.$</p> <p>Update step:</p> <p>$\hat{\mathbf{K}}_l = \hat{\Sigma}_l \hat{\mathbf{C}}_l^H / \hat{c}_l$ with $\hat{c}_l = \hat{\mathbf{C}}_l \hat{\Sigma}_l \hat{\mathbf{C}}_l^H + \sigma^2,$</p> <p>$\hat{\mathbf{h}}_l = \tilde{\mathbf{h}}_l + \hat{\mathbf{K}}_l (\bar{y}_l - \hat{\mathbf{C}}_l \tilde{\mathbf{h}}_l),$</p> <p>$\hat{\Sigma}_l = (\mathbf{I} - \hat{\mathbf{K}}_l \hat{\mathbf{C}}_l) \hat{\Sigma}_l,$</p> <p>where $\hat{\mathbf{C}}_l = \mathbf{f}_k^T \hat{\mathbf{B}}_l \mathbf{P}$ and $\hat{\mathbf{B}}_l = \text{diag}\{\hat{b}_{n,1}, \dots, \hat{b}_{n,k}, 0, \dots, 0\}.$</p>
--

ALGORITHM 1: Decision-directed detector (DD).

One distinct feature of the decision-directed approach on the TOSSM is that the decision on only one user's bit is made at each l . Specifically, let $\hat{b}_{n,k-1}$ and $\hat{\mathbf{h}}_{l-1}$ represent the decisions on $b_{n,k-1}$ and \mathbf{h}_{l-1} at $l-1$, then the decision-directed approach at l can be summarized in Algorithm 1. Clearly, the above decision-directed algorithm is equivalent to one run of the Kalman filter, and therefore it is a lot simpler than the optimum MPM solution. Nevertheless, the user bit is determined based on the prediction of the channel states and the decisions on previous users' bits, and thus it is not optimum. Compared with the algorithm based on DSSM (6), at time k with k from 1 to K , the above algorithm makes

a decision on one user at a time and updates the channel state vector \mathbf{h}_l whenever a decision is reached. The updated \mathbf{h}_l will then influence the decision on $b_{n,k+1}$. Therefore, in both a good and a bad way, decisions at early stages (smaller k) would have more impact on decisions at later stages (larger k) than those made by the algorithm on DSSM. If detection error exists in early stages, they will be propagated into later stages. It is therefore beneficial to rank the users according to the estimated SNR. The performance of the decision-directed algorithm is, however, ultimately limited by error propagation.

6. PARTICLE FILTERING DETECTOR FOR BLIND MUD

Particle filtering belongs to the family of Monte Carlo sampling which aims at using samples to approximate posterior distribution. However, particle filtering distinguishes itself by employing a sequential importance sampling scheme, and in particular, it is designed for nonlinear and non-Gaussian systems described through state-space modeling such as the problem concerned.

In the context of the proposed problem, when \mathbf{y}_N , or equivalently $\bar{\mathbf{y}}_N$, is observed at time N , the objective of particle filtering is to draw, say, J weighted random samples $\{\mathbf{b}_{1:N}^{(j)}, w_{NK}^{(j)}\}_{j=1}^J$ from $p(\mathbf{b}_{1:N} | \bar{\mathbf{y}}_{1:nK})$, where $w_{NK}^{(j)}$ is the weight of the j th sample $\mathbf{b}_{1:N}^{(j)}$. With the samples, $p(\mathbf{b}_N | \bar{\mathbf{y}}_{1:nK})$ can be approximated by

$$p(\mathbf{b}_{1:N} | \bar{\mathbf{y}}_{1:nK}) \approx \sum_{j=1}^J w_{NK}^{(j)} \prod_{l=1}^{NK} \delta(b_{n,k} - b_{n,k}^{(j)}), \quad (20)$$

where $\delta(\cdot)$ is the Dirac delta function, and hence the MPM solution of \mathbf{b} by a simple weighted summation is

$$(\hat{b}_{N,k})_{\text{MPM}} \approx \text{sgn} \left(\sum_{j=1}^J w_{NK}^{(j)} b_{N,k}^{(j)} \right) \quad (21)$$

for $k = 1, \dots, K$. By the law of large numbers, the approximation will converge to the true MPM solution with the increase of the number of samples J . If these samples are taken directly from the posterior distribution, then all the samples have equal weights. However, direct sampling from $p(\mathbf{b}_{1:N} | \bar{\mathbf{y}}_{1:NK})$ is prohibited since all possible combinations of $\mathbf{b}_{1:N}$ must be evaluated on $p(\mathbf{b}_{1:N} | \bar{\mathbf{y}}_{1:NK})$, which again requires 2^{NK} Kalman filters. To circumvent the difficulty, importance sampling is performed where samples are taken from a proposal importance function $\pi(\mathbf{b}_{1:KN} | \bar{\mathbf{y}}_{1:KN})$ and weighted according to

$$w_{KN}^{(j)} = \frac{p(\mathbf{b}_{1:KN}^{(j)} | \bar{\mathbf{y}}_{1:KN})}{\pi(\mathbf{b}_{1:KN}^{(j)} | \bar{\mathbf{y}}_{1:KN})} \quad \forall j. \quad (22)$$

Notice that $\pi(\mathbf{b}_{1:KN} | \bar{\mathbf{y}}_{1:KN})$ is a very high-dimensional distribution and it is burdensome to sample the variables and calculate the weights altogether. Fortunately, the TOSSM allows a Markovian factorization on the posterior distribution as

$$\begin{aligned} p(\mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK}) &\propto p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1}) p(b_{N,K}) \\ &\quad \times p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1} | \bar{\mathbf{y}}_{1:NK-1}) \\ &= p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1}) \\ &\quad \times p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1} | \bar{\mathbf{y}}_{1:NK-1}). \end{aligned} \quad (23)$$

Then, if we choose the importance distribution as

$$\begin{aligned} \pi(\mathbf{b}_{1:N} | \bar{\mathbf{y}}_{1:NK}) &= p(b_{N,k} | \mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{\mathbf{y}}_{1:NK}) \\ &\quad \times \pi(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1} | \bar{\mathbf{y}}_{1:NK-1}), \end{aligned} \quad (24)$$

the weight can be calculated by

$$\begin{aligned} w_{KN}^{(j)} &= \frac{p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1}) p(b_{N,K}^{(j)})}{p(b_{N,K}^{(j)} | \mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{\mathbf{y}}_{1:NK})} \\ &\quad \times \frac{p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1} | \bar{\mathbf{y}}_{1:NK-1})}{\pi(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1} | \bar{\mathbf{y}}_{1:NK-1})} \\ &= \frac{p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1}) p(b_{N,K}^{(j)})}{p(b_{N,K}^{(j)} | \mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{\mathbf{y}}_{1:NK})} w_{KN-1}^{(j)} \\ &\propto p(\bar{y}_{NK} | \mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{\mathbf{y}}_{1:NK-1}) w_{KN-1}^{(j)} \\ &= \sum_{b_{N,K}} p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1}) w_{KN-1}^{(j)} \\ &= \mu_{KN-1}^{(j)} w_{KN-1}^{(j)}, \end{aligned} \quad (25)$$

where $\mu_{KN-1}^{(j)}$ is the weight update factor. Examining (24) and (25), we find that given $w_{KN-1}^{(j)}$ and $p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1} | \bar{\mathbf{y}}_{1:NK-1})$, the importance function (24) and the weights (25) are known exactly as long as $p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1})$ can be derived. In fact, we have indicated in Section 4 that $p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{\mathbf{y}}_{1:NK-1})$ can be calculated through the Kalman filter as

$$\begin{aligned} \lambda_{NK}(i) &= p(\bar{y}_{NK} | b_{N,K} = 2 * i - 3, \mathbf{b}_{N,1:K}^{(j)}, \mathbf{b}_{1:N}^{(j)}, \bar{\mathbf{y}}_{1:NK-1}) \\ &= \mathcal{N}_c(m_{NK}^{(j)}(i), c_{NK}^{(j)}(i)) \end{aligned} \quad (26)$$

for $i = 1, 2$ where $m_i^{(j)}(i)$ and $c_i^{(j)}(i)$ are calculated the same way as shown in the appendix but for a set of $\mathbf{b}_{1:NK}$ given in (26). We can therefore obtain samples and weights using a recursive algorithm. To put the idea in concrete procedure, we assume that at $l-1$, we have obtained from a previous recursion the trajectories (samples) $\{\mathbf{b}_{0:l-1}^{(j)}\}_{j=1}^J$ appropriately weighted with the weights $\{w_{l-1}^{(j)}\}_{j=1}^J$. Using the recent observations \bar{y}_l , we update the trajectories and weights as in Algorithm 2. This process of recursively obtaining particles is called particle filtering. After each recursion, the mean $\boldsymbol{\eta}_l^{(j)}$ and covariance vectors $\boldsymbol{\Xi}_l^{(j)}$ are passed on to the next recursion. From (21), we also see that to calculate all the elements of $\{\mathbf{b}_N\}_{\text{MPM}}$, $w_{NK}^{(j)}$ is required. Therefore the decision on all the elements can only be made after recursion $l = KN$ and the particles for $b_{N,k}$ for $k = 1, 2, \dots, K-1$ must be stored.

In the above derivation of particle filtering, the adopted importance function is known as optimum in the sense that minimizes the variance of the weights. The above particle filtering procedure suffers from particle impoverishment, that is, after several recursions, some weights of the samples become negligible and stop contributing to the overall evaluation. To prevent it, we insert a residue resampling step [15] after every fixed recursion. Particularly, during the resampling at recursion l , the particles for $\mathbf{b}_{n,1:k}^{(j)}$, the mean vectors, and covariance matrices must be treated as a set in the resampling process.

7. STOCHASTIC M -DETECTOR FOR BLIND MUD

Recently, a very efficient particle filtering algorithm called stochastic M -algorithm (SMA) was proposed in [19] for problems with discrete unknowns. SMA can provide similar performance as generic particle filtering but with much reduced complexity. SMA can be considered as a particle filtering algorithm with the discrete delta functions as importance functions. In addition, each trajectory produces two samples (-1 and 1) for the binary case rather than one sample as in the generic PF. A key feature with SMA is that no two trajectories are identical, which is however rarely true with the generic PF. As a result, the SMA can provide more sample diversities with less trajectories than the generic PF. Nonetheless, notice that the number of trajectories doubles after each sampling and therefore a selection step is required

For $j = 1$ to J , do as follows.

(1) Predictive step:

Calculate

$$\boldsymbol{\mu}_l^{(j)} = \begin{cases} \mathbf{Q}\boldsymbol{\eta}_{l-1}^{(j)} & \text{if } k = 1, \\ \boldsymbol{\eta}_{l-1}^{(j)} & \text{if } k \neq 1, \end{cases}$$

and

$$\boldsymbol{\Sigma}_l^{(j)} = \begin{cases} \mathbf{Q}\boldsymbol{\Xi}_{l-1}^j \mathbf{Q}^\top + \sigma^2 \mathbf{G}\mathbf{G}^\top & \text{if } k = 1, \\ \boldsymbol{\Xi}_{l-1}^{(j)} & \text{if } k \neq 1. \end{cases}$$

(2) Sampling step.

(a) For $i = 1$ and -1 , calculate

(i) $m_l^{(j)}(i) = \mathbf{c}_l^{(j)}(i)\boldsymbol{\mu}_l^{(j)}$ and

$$c_l^{(j)}(i) = \mathbf{c}_l(i)\boldsymbol{\Sigma}_l^{(j)}\mathbf{c}_l^{(j)}(i) + \sigma^2,$$

$$\text{where } \mathbf{c}_l^{(j)}(i) = \mathbf{f}_k^\top \mathbf{B}_l^{(j)}(i)\mathbf{P}, \mathbf{B}_l^{(j)}(i) =$$

$$\text{diag}\{b_{n,1}^{(j)}, \dots, b_{n,k-1}^{(j)}, i, 0, \dots, 0\};$$

(ii) $\lambda_l^{(j)}(i) = \mathcal{N}_c(m_l^{(j)}(i), c_l^{(j)}(i))$.

(b) Sample $m \in \{-1, 1\}$ with probability

proportional to $\lambda_l^{(j)}(i) \forall i$.

(c) Set $b_l^{(j)} = m$.

(d) Calculate $\mu_l^{(j)} = \sum_{i \in \{-1, 1\}} \lambda_l^{(j)}(i)$ and the unnormalized weight $\tilde{w}_l^{(j)} = \mu_l^{(j)} w_{l-1}^{(j)}$.

(3) Updating step. Calculate

(i) $\mathbf{K}_l^{(j)} = \boldsymbol{\Sigma}_l^{(j)} \mathbf{c}_l^{(j)}(m) / c_l^{(j)}(m)$;

(ii) $\boldsymbol{\eta}_l^{(j)} = \boldsymbol{\mu}_l^{(j)} + \mathbf{K}_l^{(j)}(\bar{y}_l - \mathbf{c}_l^{(j)}(m)\boldsymbol{\mu}_l^{(j)})$;

(iii) $\boldsymbol{\Xi}_l^{(j)} = (\mathbf{I} - \mathbf{K}_l^{(j)} \mathbf{c}_l^{(j)}(m)) \boldsymbol{\Sigma}_l^{(j)}$.

Form the new trajectories $\mathbf{b}_{0:l}^{(j)} = \{b_l^{(j)}, \mathbf{b}_{0:l-1}^{(j)}\} \forall j$.

Normalize the weight as $w_l^{(j)} = \tilde{w}_l^{(j)} / \sum_{j=1}^J \tilde{w}_l^{(j)}$.

ALGORITHM 2: Particle filtering detector (PFD).

to avoid the exponential increase of trajectories. Here, we use the optimal resampling algorithm [20] since it is a sampling-without-replacement algorithm and does not produce replicates of the same trajectories, the feature that is required by SMA. The SMA for the problem concerned at the l th recursion is outlined as in Algorithm 3.

The structure of the SMA resembles the popular M -algorithm. However, since the SMA is still a PF algorithm, it can provide probability information about the unknowns and thus can be applied to iterative MUD of a coded system.

7.1. Discussion on the MPM, decision-directed, and particle filtering solutions

Comparing the PFD and the SMD with the decision-directed algorithm, we see that the processes along each trajectory is almost as identical as a decision-directed algorithm except that a sampling step is used in the place of the detection step, and they all resemble one run of Kalman filter which corresponds to a path in the tree of Figure 2. There are two paths going out at every node in the tree, and in selecting a path,

Trajectory expansion

(1) For $j = 1$ to J ,

(i) perform the predictive step in the PFD Algorithm;

(ii) perform (2)(a) in Algorithm PFD;

(iii) set $b_l^{(2j-1)} = 1$ and calculate the weight by $\tilde{w}_l^{(2j-1)} = \lambda_l^{(j)}(1)w_{l-1}^{(j)}$;

(iv) set $b_l^{(2j)} = -1$ and calculate the weight by $\tilde{w}_l^{(2j)} = \lambda_l^{(j)}(-1)w_{l-1}^{(j)}$;

(v) form $2J$ new trajectories by setting $\mathbf{b}_l^{(2j-1)} = \{b_l^{(2j-1)}, \mathbf{b}_{0:l}^{(j)}\}$ and $\mathbf{b}_l^{(2j)} = \{b_l^{(2j)}, \mathbf{b}_{0:l}^{(j)}\}$.

(2) Normalize the weights $\tilde{w}_k^{(j)}$ to obtain $w_k^{(j)}$.

(3) Trajectory selection: select J trajectories from $2M$ trajectories using the optimal resampling algorithm.

(4) Updating step: for $j = 1$ to J ;

perform the updating step in the PFD Algorithm.

ALGORITHM 3: Stochastic M detector (SMD).

the decision-directed algorithm uses a deterministic approach, while PFD and SMD adopt a soft measure which is based on probability. What is more, each trajectory is also associated with a weight which indicates the significance of the trajectory in final decision. Although trajectories with small weight do not seem to contribute much to current decision making at the present stage, they, however, might flourish in later recursions and carry significant weights in decision. The soft measure can apparently prevent current decision errors from greatly influencing the future decision, a key advantage over the decision-directed approach.

Comparing the PFD and the SMD with the optimum MPM solution, PFD, especially the SMD, has clear edge in complexity since it only maintains J trajectories or equivalently J Kalman filter at all times, but the required Kalman filter for the MPM grows exponentially with time. Further, the PFD and the SMD achieve every effective and efficient approximation to the true posterior distribution and therefore provide decision performance closer to optimum. Since the two detectors produce soft (probabilistic) results, they are readily applied in turbo MUD.

8. SIMULATION RESULTS

In this section, the bit error rate (BER) performance of the proposed PFDs and SMDs are studied through experiments. In all the experiments, the transmitted signal was differential BPSK modulated. The number of users was 15. For the PFDs, 151 trajectories were maintained, whereas 4 and 32 trajectories were tested for SMDs. Further, an AR model was adopted for the fading process, which was normalized to have a unit power, and thus the signal-to-noise ratio (SNR) was obtained by $10 \log(1/\sigma^2)$.

In Figure 3, we provide the BER versus SNR for the different algorithms on a scenario of $\Omega_d = 0.03$. The genie-aided detector is included as a lower bound. We notice that the PFDs and SMDs with 32 trajectories are of the same

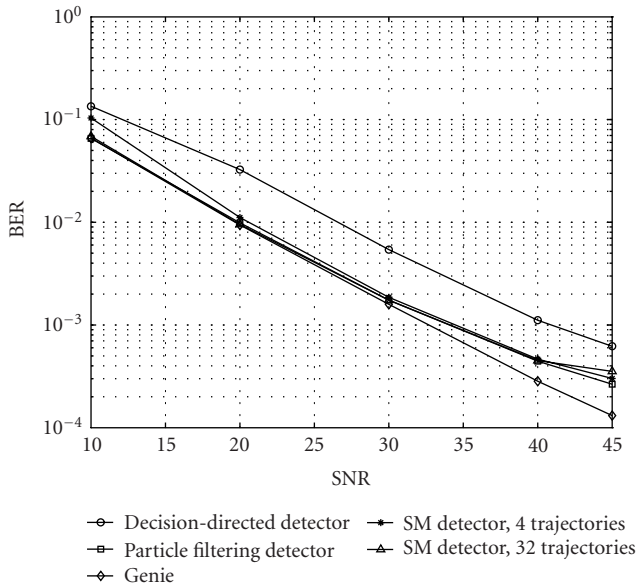


FIGURE 3: BERs versus SNR performance for various detectors. $\Omega = 0.03$.

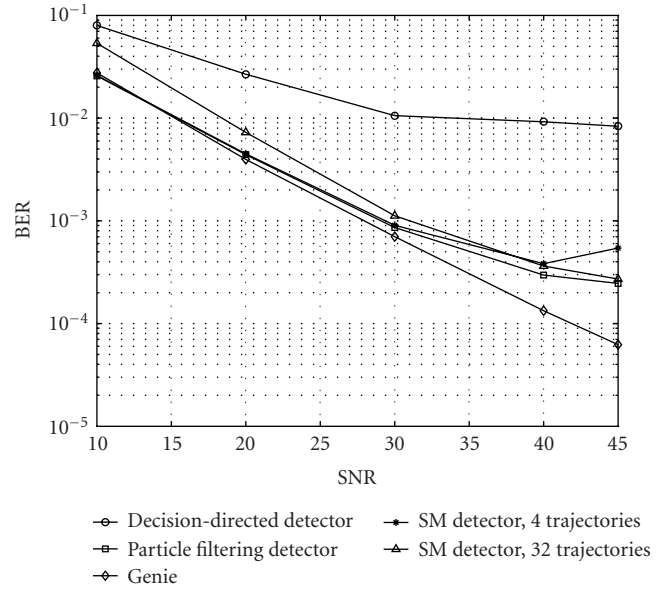


FIGURE 5: BERs versus SNR performance for various detectors for users with different power. $\Omega = 0.03$.

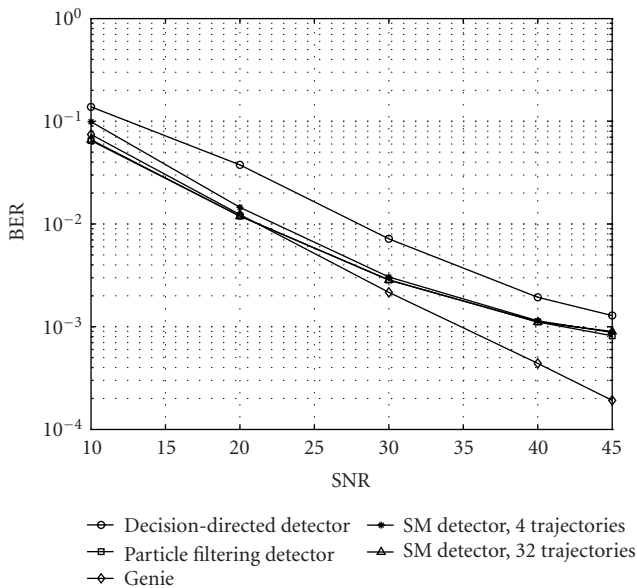


FIGURE 4: BERs versus SNR performance for various detectors. $\Omega = 0.05$.

order of magnitude as that of the genie-aided detector at low SNR (less than 30 dB). On the other hand, the results obtained by the SMDs with 4 and 32 trajectories are very close, especially after 20 dB, and comparable to that of the PFD. The SMD with 4 trajectories is obviously more favorable since it requires only about 1/35 of complexity of the PFD. As a final note, the PFD and SMDs achieve about 7 dB gain over the decision-directed detectors at 10^{-3}

BER. In Figure 4, we provide the BER versus SNR performance for a higher Doppler frequency of $\Omega_d = 0.05$. Similar observations can be drawn as for the previous case even though the overall performance of the detectors is worse, which is reasonable considering that the channels are fading faster.

In Figure 5, we provide the BER versus SNR of the first user for the different algorithms on a scenario of $\Omega_d = 0.03$. In addition, the users have different power. The difference between the power of the first user and that of the last user is 10 dB and the other users' powers are equally spaced in between. The genie-aided detector is also included as a lower bound. In this case, the PFDs and SMDs with 32 trajectories are approximately of the same order of magnitude as that of the genie-aided detector at SNRs of the first user less than 30 dB. As in the case of equal power, the results obtained by the SMDs with 4 and 32 trajectories are very close, especially after 30 dB, and comparable to that of the PFD. Again, the SMD with 4 trajectories is obviously more favorable since it requires only about 1/35 of complexity of the PFD. In this experiment, the performance of the decision-directed detector is much worse compared to the performance of the PFD and SMDs. For example, the latter achieves about 11 dB gain over the former at 10^{-2} BER. In Figure 6, we provide the BER versus SNR performance for a Doppler frequency of $\Omega_d = 0.05$. Since the channels considered are fading faster, the performance of the detectors is worse. However, in general, similar observations to the tested detectors can be drawn. It is important to outline that the performance of the decision-directed detector gets worse in this case, for example, the PFD and SMDs achieve about 20 dB gain over the decision-directed detectors at 10^{-2} BER.

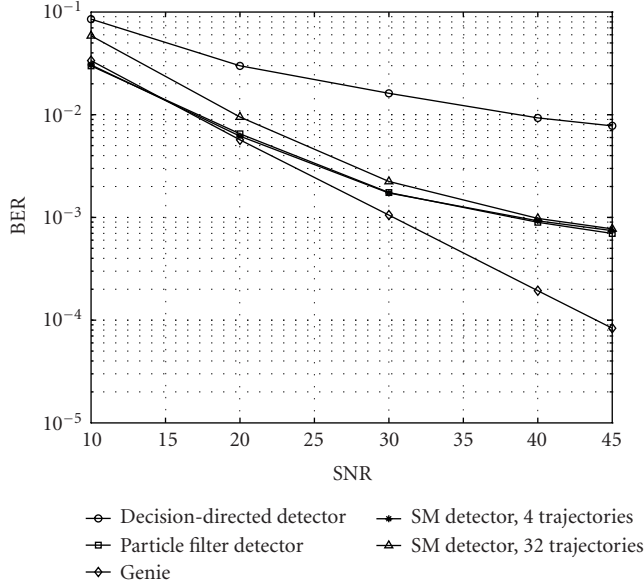


FIGURE 6: BERs versus SNR performance for various detectors for users with different power. $\Omega = 0.05$.

9. CONCLUSION

In this paper, we proposed to solve blind MUD over flat fast fading channels. We constructed a novel time-observation state-space model, based on which efficient particle filtering and stochastic M detectors were proposed. Particularly, the detectors based on the SMA demonstrated greater potential than those using generic PF. The former can provide comparable performance as the latter but with much smaller complexity.

APPENDIX

DERIVATION OF THE LIKELIHOOD $p(\bar{y}_l | \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l})$

The likelihood $p(\bar{y}_l | \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l})$ can be obtained as

$$\begin{aligned} p(\bar{y}_l | \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l-1}) &= \int p(\bar{y}_l, \mathbf{h}_l | \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l}) d\mathbf{h}_l \\ &= \int p(\bar{y}_l | \mathbf{h}_l, \mathbf{b}_{n,1:k}) p(\mathbf{h}_l | \mathbf{b}_{n,1:k-1}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l-1}) d\mathbf{h}_l, \end{aligned} \quad (\text{A.1})$$

where the last equality is arrived by the fact that, given \mathbf{h}_l , and $\mathbf{b}_{n,1:k}$, \bar{y}_l is independent of other variables, and \mathbf{h}_l is independent of $\mathbf{b}_{n,k}$. In (A.1), two distributions are involved in the integral. The first distribution is the likelihood defined by the observation equation which is

$$p(\bar{y}_l | \mathbf{h}_l, \mathbf{b}_{n,1:k}) = \mathcal{N}(\mathbf{C}_l \mathbf{h}_l, \sigma^2), \quad (\text{A.2})$$

where $\mathbf{C}_l = \mathbf{f}_k^T \tilde{\mathbf{B}}_l \mathbf{P}$. The second distribution $p(\mathbf{h}_l | \mathbf{b}_{n,1:k-1}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l-1})$ is the predictive density which can be obtained from the predictive step of the Kalman filter [21, 22], that is,

$$p(\mathbf{h}_l | \mathbf{b}_{n,1:k-1}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l-1}) = \mathcal{N}(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l), \quad (\text{A.3})$$

where

$$\boldsymbol{\mu}_l = \begin{cases} \mathbf{Q}\boldsymbol{\eta}_{l-1} & \text{if } k = 1, \\ \boldsymbol{\eta}_{l-1} & \text{if } k \neq 1, \end{cases} \quad (\text{A.4})$$

and

$$\boldsymbol{\Sigma}_l = \begin{cases} \mathbf{Q}\boldsymbol{\Xi}_{l-1}\mathbf{Q}^T + \sigma^2\mathbf{G}\mathbf{G}^T & \text{if } k = 1, \\ \boldsymbol{\Xi}_{l-1} & \text{if } k \neq 1. \end{cases} \quad (\text{A.5})$$

In (A.4) and (A.5), $\boldsymbol{\eta}_{l-1}$ and $\boldsymbol{\Xi}_{l-1}$ are computed from the update steps of the Kalman filter expressed in terms of l as

$$\boldsymbol{\eta}_l = \boldsymbol{\mu}_l + \mathbf{K}_l(\bar{y}_l - m_l), \quad (\text{A.6})$$

and

$$\boldsymbol{\Xi}_l = (\mathbf{I} - \mathbf{K}_l\mathbf{C}_l)\boldsymbol{\Sigma}_l, \quad (\text{A.7})$$

where $m_l = \mathbf{C}_l\boldsymbol{\mu}_l$ and $\mathbf{K}_l = \boldsymbol{\Sigma}_l\mathbf{C}_l^H/c_l$ with $c_l = \mathbf{C}_l\boldsymbol{\Sigma}_l\mathbf{C}_l^H + \sigma^2$. Now the integration in (A.1) is readily derived as

$$p(\bar{y}_l | \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l}) = \mathcal{N}(m_l, c_l). \quad (\text{A.8})$$

ACKNOWLEDGMENTS

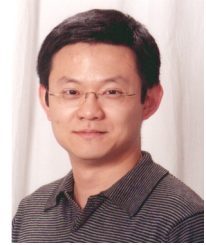
This work was supported by the National Science Foundation under Awards no. CCR-9903120 and no. CCR-0082607 and partially supported by the ‘‘Ministerio de Ciencia y Tecnología,’’ Spain, under Project TIC 2001-2902.

REFERENCES

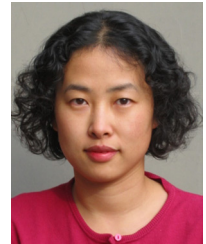
- [1] S. Verdú, *Multiuser Detection*, Cambridge University Press, New York, NY, USA, 1998.
- [2] A. Russ and M. K. Varanasi, ‘‘Noncoherent multiuser detection for nonlinear modulation over the Rayleigh-fading channel,’’ *IEEE Trans. Inform. Theory*, vol. 47, no. 1, pp. 295–307, 2001.
- [3] H.-Y. Wu and A. Duel-Hallen, ‘‘On the performance of coherent and noncoherent multiuser detectors for mobile radio CDMA channels,’’ in *Proc. 5th IEEE International Conference on Universal Personal Communications*, vol. 1, pp. 76–80, Cambridge, Mass, USA, September–October 1996.
- [4] M. Honig, U. Madhow, and S. Verdú, ‘‘Blind adaptive multiuser detection,’’ *IEEE Trans. Inform. Theory*, vol. 41, no. 4, pp. 944–960, 1995.

- [5] R. A. Iltis, "Exact and approximate maximum-likelihood parameter estimation for quasi-synchronous CDMA signals," *IEEE Trans. Commun.*, vol. 48, no. 7, pp. 1208–1216, 2000.
- [6] X. Wang and H. V. Poor, "Blind multiuser detection: a subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 677–690, 1998.
- [7] U. Fawer and B. Aazhang, "A multiuser receiver for code division multiple access communications over multipath channels," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 1556–1565, 1995.
- [8] K. Yen and L. Hanzo, "Genetic algorithm assisted joint multiuser symbol detection and fading channel estimation for synchronous CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 19, no. 6, pp. 985–998, 2001.
- [9] T. J. Lim, L. K. Rasmussen, and H. Sugimoto, "An asynchronous multiuser CDMA detector based on the Kalman filter," *IEEE J. Select. Areas Commun.*, vol. 16, no. 9, pp. 1711–1722, 1998.
- [10] T. J. Lim and Y. Ma, "The Kalman filter as the optimal linear minimum mean-squared error multiuser CDMA detector," *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2561–2566, 2000.
- [11] P. H.-Y. Wu and A. Duel-Hallen, "Multiuser detectors with disjoint Kalman channel estimators for synchronous CDMA mobile radio channels," *IEEE Trans. Commun.*, vol. 48, no. 5, pp. 752–756, 2000.
- [12] S. Vasudevan and M. K. Varanasi, "Achieving near-optimum asymptotic efficiency and fading resistance over the time-varying Rayleigh-faded CDMA channel," *IEEE Trans. Commun.*, vol. 44, no. 9, pp. 1130–1143, 1996.
- [13] K. J. Kim and R. A. Iltis, "Joint detection and channel estimation algorithms for QS-CDMA signals over time-varying channels," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 845–855, 2002.
- [14] B. Flanagan, C. Suprin, S. Kumaresan, and J. Duniak, "Performance of a joint Kalman demodulator for multiuser detection," in *Proc. 56th IEEE Vehicular Technology Conference (VTC '02)*, vol. 3, pp. 1525–1529, Vancouver, Canada, 2002.
- [15] R. Chen, X. Wang, and J. S. Liu, "Adaptive joint detection and decoding in flat-fading channels via mixture Kalman filtering," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 2079–2094, 2000.
- [16] J. Zhang and P. M. Djurić, "Joint estimation and decoding of space-time trellis codes," *EURASIP J. Appl. Signal Process.*, vol. 2002, no. 3, pp. 305–315, 2002.
- [17] E. Punskeya, C. Andrieu, A. Doucet, and W. J. Fitzgerald, "Particle filtering for multiuser detection in fading CDMA channels," in *Proc. 11th IEEE Signal Processing Workshop on Statistical Signal Processing*, pp. 38–41, Singapore, August 2001.
- [18] Y. Huang and P. M. Djurić, "A blind particle filtering detector of signals transmitted over flat fading channels," *IEEE Trans. Signal Processing*, vol. 52, no. 7, pp. 1891–1900, 2004.
- [19] Y. Huang, J. Zhang, and P. M. Djurić, "Bayesian detection for BLAST," *IEEE Transactions on Signal Processing*, vol. 53, no. 3, pp. 1086–1096, March 2005.
- [20] P. Fearnhead, *Sequential Monte Carlo method in filter theory*, Ph.D. thesis, Oxford University, Oxford, UK, 1998, available from <http://www.maths.lancs.ac.uk/~fearnhead/>.
- [21] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, John Wiley & Sons, New York, 1996.
- [22] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Upper Saddle River, NJ, USA, 4th edition, 2002.

Yufei Huang received the B.S. degree in applied electronics from Northwestern Polytechnical University, Xi'an, China, in 1995, and the M.S. and Ph.D. degrees in electrical engineering from the State University of New York at Stony Brook (SUNYSB), Stony Brook, NY, in 1997 and 2001, respectively. He is now an Assistant Professor in the Department of Electrical Engineering, The University of Texas at San Antonio. From 2001 to 2002, he worked as a Postdoctoral Researcher in the Department of Electrical and Computer Engineering, SUNYSB. His current research interests are in Bayesian inference, Monte Carlo methods, signal processing for communications, and bioinformatics.



Jianqiu (Michelle) Zhang received her B.S. and M.S. degrees in electrical engineering in 1992 and 1995, respectively, from Zhejiang University, Hangzhou University, and Zhongshan University, Guangzhou, China. From 1995 to 1997, she worked as a Software Engineer in R&D, Guangdong Nortel, China. She received her Ph.D. degree in electrical engineering from the State University of New York at Stony Brook in 2002.



Currently, she is an Assistant Professor in the Department of Electrical and Computer Engineering at the University of New Hampshire. Her general research interest lies in the fields of wireless communications, information theory, and statistical methods in signal processing. She had been working on topics including CDMA multiuser detection, turbo-coding, particle filtering, space-time coding, and channel capacity.

Isabel Tienda Luna was born in Dona Mencía, Córdoba, Spain, on October 5th 1978. She received her B.S. and M.S. degrees from the University of Granada, Spain, in 1999 and 2001, respectively. Now she is a predoctoral researcher within the Systems, Signals and Waves Research Group, Department of Applied Physics, the University of Granada. Her predoctoral research is funded by the "Ministerio de Educación y Ciencia" of Spain. Her research interests are in the areas of digital communications and signal processing including particle filtering applications, channel estimation, and multiuser detection.



Petar M. Djurić received his B.S. and M.S. degrees in electrical engineering from the University of Belgrade, in 1981 and 1986, respectively, and his Ph.D. degree in electrical engineering from the University of Rhode Island in 1990. From 1981 to 1986 he was a Research Associate with the Institute of Nuclear Sciences, Vinca, Belgrade. Since 1990 he has been with Stony Brook University, where he is a Professor in the Department of Electrical and Computer Engineering.



He works in the area of statistical signal processing, and his primary interests are in the theory of modeling, detection, estimation, and time series analysis and its application to a wide variety of disciplines including wireless communications and biomedicine. Professor Djurić

has served on numerous technical committees for the IEEE and SPIE and has been invited to lecture at universities in the United States and overseas. He is the Area Editor of special issues of the Signal Processing Magazine and Associate Editor of the IEEE Transactions on Signal Processing. He is also Chair of the IEEE Signal Processing Society Committee on Signal Processing—Theory and Methods and is on the Editorial Board of Digital Signal Processing, the EURASIP Journal on Applied Signal Processing, and the EURASIP Journal on Wireless Communications and Networking. Professor Djurić is a Member of the American Statistical Association and the International Society for Bayesian Analysis.

Diego Pablo Ruiz Padillo was born in Cabra, Córdoba, Spain, on July 17th 1968. He received his B.S. and M.S. degrees in electronic physics from the University of Granada, in 1991 and 1993, respectively, and his Ph.D. degree with honors in 1995. He was a granted national researcher from the “Ministerio de Ciencia y Tecnología” of Spain, and an Assistant Professor in the Universities of Malaga and Granada from



1991 to 1998. Now Dr. Diego P. Ruiz is an Associate Professor within the Systems, Signals and Waves Research Group of the Department of Applied Physics in the University of Granada, Spain. His research interests include statistical signal processing and its applications to wireless communications, blind channel identification and equalization, adaptive algorithms, higher-order statistics, and radar signal processing applied to transient electromagnetic problems.

Blind Decoding of Multiple Description Codes over OFDM Systems via Sequential Monte Carlo

Zigang Yang

*Texas Instruments Inc, 12500 TI Boulevard Dallas, MS 8653 Dallas, TX 75243, USA
Email: zigang@ti.com*

Dong Guo

*Department of Electrical Engineering, Columbia University, New York, NY 10027, USA
Email: guodong@ee.columbia.edu*

Xiaodong Wang

*Department of Electrical Engineering, Columbia University, New York, NY 10027, USA
Email: wangx@ee.columbia.edu*

Received 1 May 2004; Revised 20 December 2004

We consider the problem of transmitting a continuous source through an OFDM system. Multiple description scalar quantization (MDSQ) is applied to the source signal, resulting in two correlated source descriptions. The two descriptions are then OFDM modulated and transmitted through two parallel frequency-selective fading channels. At the receiver, a blind turbo receiver is developed for joint OFDM demodulation and MDSQ decoding. Transformation of the extrinsic information of the two descriptions are exchanged between each other to improve system performance. A blind soft-input soft-output OFDM detector is developed, which is based on the techniques of importance sampling and resampling. Such a detector is capable of exchanging the so-called extrinsic information with the other component in the above turbo receiver, and successively improving the overall receiver performance. Finally, we also treat channel-coded systems, and a novel blind turbo receiver is developed for joint demodulation, channel decoding, and MDSQ source decoding.

Keywords and phrases: multiple description codes, OFDM, frequency-selective fading, sequential Monte Carlo, turbo receiver.

1. INTRODUCTION

Multiple description scalar quantization (MDSQ) is a source coding technique that can exploit diversity communication systems to overcome channel impairments. An MDSQ encoder generates multiple descriptions for a source and sends them over different channels provided by the diversity systems. At the receiver, when all descriptions are received correctly, a high-quality reconstruction is possible. In the event of failure of one or more of the channels, the reconstruction would still be of acceptable quality.

The problem of designing multiple description scalar quantizers is addressed in [1, 2], where a theoretical performance bound is derived in [1] and practical design methods are given in [2, 3]. Conventionally, MDSQ has been

investigated only from the perspective of transmission over erasure channels, that is, channels which either transmit noiselessly or fail completely [1, 2, 4]. Recently, it was shown in [5] that an MDSQ can be used effectively for communication over slow-fading channels. In that system, a threshold on the channel fade values is used to determine the acceptability of the received description. The signal received from the bad connection is not utilized at the receiver.

In this paper, we propose an iterative MDSQ decoder for communication over fading channels, where the extrinsic information of the descriptions is exchanged with each other by exploiting the correlation between the two descriptions. Although the MDSQ coding scheme provided in [2] is optimized with the constraint of erasure channels, it provides very nice correlation property between different descriptions. Therefore, the same MDSQ scheme will be applied to the continuous fading environment considered in this paper [6, 7, 8].

This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

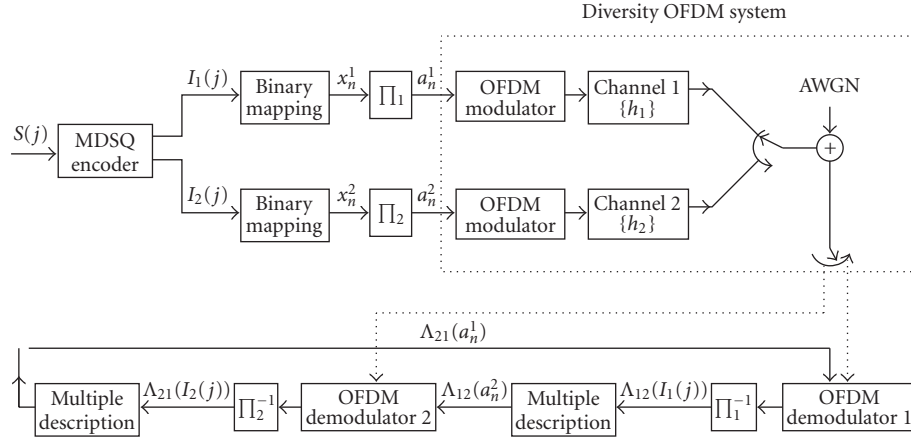


FIGURE 1: Continuous source transmitted through a diversity OFDM system with MDSQ.

Providing high-data-rate transmission is a key objective for modern communication systems. Recently, orthogonal frequency-division multiplexing (OFDM) has received a considerable amount of interests for high-rate wireless communications. Because OFDM increases the symbol duration and transmitting data in parallel, it has become one of the most effective modulation techniques for combating multipath delay spread over mobile wireless channels.

In this paper, we consider the problem of transmitting a continuous source through an OFDM system over parallel frequency-selective fading channels. The source signals are quantized and encoded by an MDSQ, resulting in two correlated descriptions. These two descriptions are then modulated by OFDM and sent through two parallel fading channels. At the receiver, a blind turbo receiver is developed for joint OFDM demodulation and MDSQ decoding. Transformation of the extrinsic information of the two descriptions are exchanged between each other to improve system performance. The transformation is in terms of a transformation matrix which describes the correlation between the two descriptions. Another novelty in this paper is the derivation of a blind detector based on a Bayesian formulation and sequential Monte Carlo (SMC) techniques for the differentially encoded OFDM system. Being soft-input and soft-output in nature, the proposed SMC detector is capable of exchanging the so-called extrinsic information with the other component in the above turbo receiver, successively improving the overall receiver performance.

For a practical communication system, channel coding is usually applied to improve the reliability of the system. In this paper, we also treat a channel-coded OFDM system, where each stream of the source description is channel encoded and then OFDM modulated before being sent to the channel. At the receiver, a novel blind turbo receiver is developed for joint demodulation, channel decoding, and source decoding.

The rest of this paper is organized as follows. In Section 2, the diversity of an OFDM system with an MDSQ encoder is described. In Section 3, the turbo receiver is discussed for

the MDSQ encoded OFDM system. In Section 4, we develop an SMC algorithm for blind symbol detection of OFDM systems. A turbo receiver for a channel-coded OFDM system is derived in Section 5. Simulation results are provided in Section 6, and a brief summary is given in Section 7.

2. SYSTEM DESCRIPTION

We consider transmitting a continuous source through a diversity OFDM system. The diversity of an OFDM system is made up of two N -subcarrier OFDM systems, signalling through two parallel frequency-selective fading channels. Such a parallel channel structure was first introduced in [9]. A block diagram of the system is shown in Figure 1. A sequence of continuous sources $\{S(j)\}$ is encoded by a multiple description scalar quantizer (MDSQ), resulting in two sets of equal-length indices $\{(I_1(j), I_2(j))\}$, where j denotes the sequence order. The detailed MDSQ encoder will be discussed in Section 2.1. These indices can be further described in a binary sequence $\{(x_n^1, x_n^2)\}$ with the order denoted by n . The bit interleavers π_1 and π_2 are used to reduce the influence of error bursts at the input of the MDSQ decoder. After the interleaved bits $\{a_n^1\}$, $\{a_n^2\}$ are modulated by OFDM, we use the parallel concatenated transmission scheme shown in Figure 1; that is, one description of the source is transmitted through one channel and the other description is transmitted through another channel. At the receiver, the OFDM demodulators, which will be discussed in Section 4, generate soft information, which is then exchanged between the two OFDM detectors in the form of *a priori* probabilities of the information symbols. Next, we will focus on the structure of the MDSQ encoder and the diversity OFDM system.

2.1. Multiple description scalar quantizer

2.1.1. Multiple description scalar quantizer for diversity on/off channels

The multiple description scalar quantizer (MDSQ) is a scalar quantizer designed for the channel model illustrated

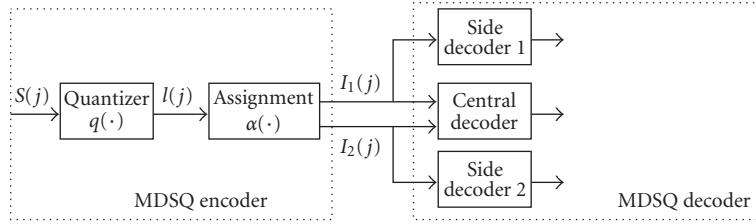


FIGURE 2: Conventional MDSQ in a diversity system.

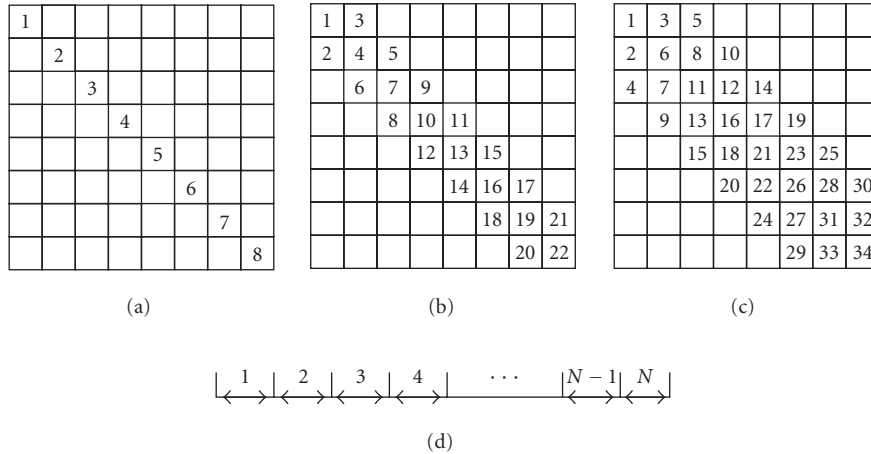


FIGURE 3: MDSQ index assignment for $R = 3$. A quantized source sample $l(j) \in \{1, 2, \dots, N\}$ is mapped to a pair of indices $(I_1(j), I_2(j)) \subset \mathcal{C}$ composed of its associated row and column determined by the assignment $\alpha(\cdot)$. (a) Assignment with $N = 8$. (b) Assignment with $N = 22$. (c) Assignment with $N = 34$. (d) Quantizer.

in Figure 2. The channel model consists of two channels that connect the source to the destination. Either channel may be broken or lossless at any time. The encoder of an MDSQ sends information over each channel at a rate of R bits/sample. Based on the decoder structure shown in Figure 2, the objective is to design an MDSQ encoder so as to minimize the average distortion when both channels are lossless (center distortion), subject to a constraint on the average distortion when only one channel is lossless (side distortion).

Next, we give a brief summary of the MDSQ design presented in [2]. Denote an index set $\mathcal{I} = \{1, 2, \dots, M\}$, where $M = 2^R$. Let $\mathcal{C} \subset \mathcal{I} \times \mathcal{I}$ and $|\mathcal{C}| = N \leq M^2$. The MDSQ encoder consists of an N -level quantizer $q(\cdot) : \mathcal{R} \rightarrow \{1, 2, \dots, N\}$ followed by index assignment $\alpha(\cdot) : \{1, 2, \dots, N\} \rightarrow \mathcal{C}$. Note that N is both the size of \mathcal{C} and the number of the quantization levels. Specifically, a source sample $S(j)$ is mapped to an index $l(j) \in \{1, 2, \dots, N\}$ by the quantizer $q(\cdot)$, which is further mapped to a pair of indices $(I_1(j), I_2(j)) \subset \mathcal{C}$ by the assignment $\alpha(\cdot)$.

Assume a uniform quantizer. The main issue in MDSQ design is the choice of the set \mathcal{C} , and the index assignment $\alpha(\cdot)$. Following [2], an example of good assignment

for $R = 3$ bits/sample is illustrated in Figure 3. We assume that the cells of a quantizer are numbered $1, 2, \dots, N$, in increasing order from left to right as shown in Figure 3d. Intuitively, with a larger set \mathcal{C} , center distortion will be improved at the expense of degraded side distortion. With the same size of the set \mathcal{C} , the center distortion is fixed, and a diagonal-like assignment is preferred to minimize the side distortion.

2.1.2. Multiple description scalar quantizer for diversity fading channels

Although MDSQ was originally designed for diversity erasure channels, it provides a possible solution that combines source coding and channel coding to exploit the diversity provided by communication systems. Next, we consider the application of MDSQ techniques in diversity fading channels.

At the transmitter, we apply the MDSQ encoder as the conventional (cf. Figure 2). For each continuous source $S(j)$, a pair of indices $(I_1(j), I_2(j))$ is generated by the MDSQ, and is further mapped to binary bits $\{x_n^1, x_n^2\}_{n=(j-1)R+1}^{jR}$. Recall that R denotes the bit-length of each description. At the receiver,

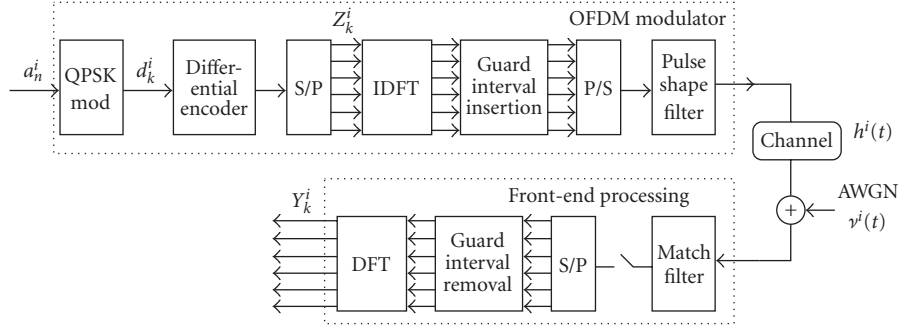


FIGURE 4: Block diagram of a baseband OFDM system.

instead of using the side decoder and central decoder, a soft MDSQ decoder is employed for MDSQ over fading channels. It is assumed that a soft demodulator is available at the receiver, which generates the *a posteriori* symbol probability for each bit x_n^i ,

$$\Lambda_i[n] \triangleq \log \frac{P(x_n^i = 1 | \mathbf{Y})}{P(x_n^i = 0 | \mathbf{Y})}, \quad (1)$$

where \mathbf{Y} denotes the received signal which is given by (3). Based on this posterior information, the soft MDSQ decoding rule is given by

$$(\hat{I}_1(j), \hat{I}_2(j)) = \arg \max_{(l,m) \in \mathcal{C}} P(I_1(j) = l | \{\Lambda_1[n]\}_n) \cdot P(I_2(j) = m | \{\Lambda_2[n]\}_n), \quad (2)$$

which maximizes the posterior probability of the indices subject to a code structure constraint, that is, $(I_1(j), I_2(j)) \in \mathcal{C}$.

2.2. Signal model for diversity OFDM system

Consider an OFDM system with N -subcarriers signaling through a frequency-selective fading channel. The channel response is assumed to be constant during one symbol duration. The block diagram of such a system is shown in Figure 4. The diversity OFDM system is just the parallel concatenation of combination of two such OFDM systems.

The binary information data $\{a_n^i\}_n$ are grouped and mapped into multiphase signals, which take values from a finite alphabet set $\mathcal{A} = \{\beta_1, \dots, \beta_{|\mathcal{A}|}\}$. In this paper, QPSK modulation is employed. The QPSK signals $\{d_k^i\}_{k=0}^{N-2}$ are differentially encoded to resolve the phase ambiguity inherent in any blind receiver, and the output is given by $Z_k^i = Z_{k-1}^i d_k^i$. These differentially encoded symbols are then inverse DFT transformed. A guard interval is inserted to prevent possible interference between OFDM frames. After pulse shaping and parallel-to-serial conversion, the signals are transmitted through a frequency-selective fading

channel. At the receiver end, after matched-filtering and removing the guard interval, the sampled received signals are sent to a DFT block to demultiplex the multicarrier signals.

For the i th OFDM system with proper cyclic extensions and proper sample timing, the demultiplexing sample of the k th subcarrier can be expressed as [10]

$$Y_k^i = Z_k^i H_k^i + V_k^i, \quad k = 0, 1, \dots, N-1; i = 1, 2, \quad (3)$$

where $V_k^i \sim \mathcal{N}_c(0, \sigma^2)$ is the i.i.d. complex Gaussian noise and H_k^i is the channel frequency response at the k th subcarrier. Using the fact that H_k^i can be further expressed as a DFT transformation of the channel time response, the signal model (3) becomes

$$Y_k^i = Z_k^i \mathbf{w}_f^H(k) \mathbf{h}^i + V_k^i, \quad k = 0, 1, \dots, N-1; i = 1, 2, \quad (4)$$

where $\mathbf{h}^i = [h_0^i, h_1^i, \dots, h_{L-1}^i]^T$ contains the time responses of all L taps; $L \triangleq \lceil \tau_m \Delta_f + 1 \rceil$ denotes the maximum number of resolvable taps, with τ_m being the maximum multipath spread and Δ_f being the tone spacing of the carriers; and $\mathbf{w}_f(k) \triangleq [1, e^{-j2\pi k/N}, \dots, e^{-j2\pi k(L-1)/N}]^T$ contains the corresponding DFT coefficients.

3. TURBO RECEIVER

The receiver under consideration is an iterative receiver structure as shown in Figure 5. It consists of two blind Bayesian OFDM detectors, which compute the soft information for the corresponding descriptions. At the output of the blind detector, information about one description is transferred to the other based on the existence of correlation between the two descriptions. Such information transfer is then repeated between the two blind detectors to improve the system performance. Next, we will focus on the operation on the first description to illustrate the iterative procedure.

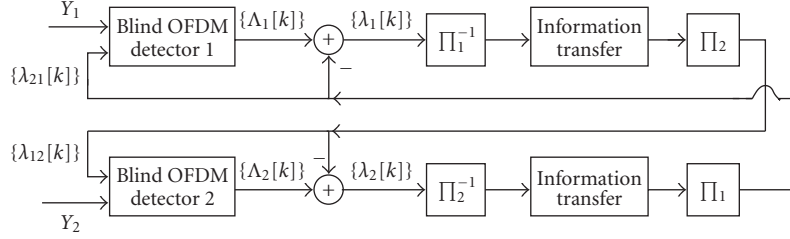


FIGURE 5: Turbo decoding for multiple description over a diversity OFDM system; Π_i and Π_i^{-1} denote the interleaver and deinterleaver, respectively, for the i th description.

3.1. Blind Bayesian OFDM detector

Denote $\mathbf{Y}^1 \triangleq \{Y_0^1, Y_1^1, \dots, Y_{N-1}^1\}$ as the received signals for the first description. The blind Bayesian OFDM detector for the first description computes the *a posteriori* probabilities of the information bits $\{a_n^1\}_n$,

$$\Lambda_1[n] \triangleq \log \frac{P(a_n^1 = 1 | \mathbf{Y}^1)}{P(a_n^1 = 0 | \mathbf{Y}^1)}. \quad (5)$$

The design of such a blind Bayesian detector will be discussed later in Section 4. For now, we assume the Bayesian detector provides us such soft information, and focus on the structure of the turbo receiver.

The *a posteriori* information delivered by the blind detector can be further expressed as

$$\Lambda_1[n] = \underbrace{\log \frac{P[\mathbf{Y}^1 | a_n^1 = 1]}{P[\mathbf{Y}^1 | a_n^1 = 0]}}_{\lambda_1[n]} + \underbrace{\log \frac{P[a_n^1 = 1]}{P[a_n^1 = 0]}}_{\lambda_{21}^p[n]}. \quad (6)$$

The second term in (6), denoted by $\lambda_{21}^p[n]$, represents the *a priori* log-likelihood ratio (LLR) of the bit a_n^1 fed from detector 2. The superscript p indicates the quantity obtained from the previous iteration. The first term in (6), denoted by $\lambda_1[n]$, represents the *extrinsic* information delivered by detector 1, based on the received signals \mathbf{Y}^1 , the structure of signal model (4), and the *a priori* information about all other bits $\{a_l^1\}_{l \neq n}$. The extrinsic information $\{\lambda_1[n]\}$ is transformed into *a priori* information $\{\lambda_{12}^p[n]\}$ for bits $\{a_n^2\}_n$. This information transformation procedure is described next.

3.2. Information transformation

Assume that $\{a_n^i\}_n$ is mapped to $\{x_n^i\}_n$ after passing through the i th deinterleaver Π_i^{-1} , with $x_n^i \triangleq a_{\pi_i(n)}^i$. To transfer the information from detector 1 to detector 2, the following steps are required.

- (1) Compute the bit probability of the deinterleaved bits

$$P(x_n^1 = 1) = \frac{e^{\lambda_1[\pi_1(n)]}}{1 + e^{\lambda_1[\pi_1(n)]}}. \quad (7)$$

- (2) Compute the probability distribution for the first index I_1 based on the deinterleaved bit probabilities

$$P(I_1(j) = l) = \prod_{k=1}^R P(x_{(j-1)R+k}^1 = b_k(l)), \quad l = 1, \dots, |\mathcal{L}|, \quad (8)$$

where $\{b_k(l), k = 1, \dots, R\}$ is the binary representation for the index $l \in \mathcal{L}$. Recall that R denotes the bit length of each description.

- (3) Compute the probability distribution for the second index I_2 according to

$$P(I_2(j) = m) = \sum_{l=1}^{|\mathcal{L}|} P(I_2(j) = m | I_1(j) = l) \cdot P(I_1(j) = l), \quad m = 1, \dots, |\mathcal{L}|. \quad (9)$$

- (4) Compute the bit probability that is associated with index $I_2(j)$,

$$P(x_{(j-1)R+k}^2 = 1) = \sum_{m:b_1(m)=1} P(I_2(j) = m). \quad (10)$$

- (5) Compute the log likelihood of interleaved code bit

$$\lambda_{12}[\pi_2(n)] = \log \frac{P(x_n^2 = 1)}{1 - P(x_n^2 = 1)}. \quad (11)$$

It is important to mention here that the key step is the calculation of the conditional probability $P(I_2(j) = m | I_1(j) = l)$ in (9). Hence, the proposed turbo receiver exploits the correlation between the two descriptions, which is measured by the conditional probabilities in (9). From the discussion in

the previous section, these conditional probabilities can be easily obtained from the index assignment rule $\alpha(\cdot)$ as shown in Figure 3.

4. BLIND BAYESIAN OFDM DETECTOR

4.1. Problem statement

Denote $\mathbf{Y}^i \triangleq \{Y_0^i, Y_1^i, \dots, Y_{N-1}^i\}$. The Bayesian OFDM receiver estimates the *a posteriori* probabilities of the information symbols

$$P(d_k^i = \beta_l | \mathbf{Y}^i), \quad \beta_l \in \mathcal{A}; k = 1, \dots, N-1, \quad (12)$$

based on the received signals \mathbf{Y}^i and the *a priori* symbol probabilities of $\{d_k^i\}_{k=1}^{N-1}$, without knowing the channel response \mathbf{h}^i . Assume the bit a_n^i is mapped to symbol $d_{\kappa(n)}^i$. Based on this symbol *a posteriori* probability, the LLR of the code bit as required in (5) can be computed by

$$\begin{aligned} \Lambda_i[n] &\triangleq \log \frac{P(a_n^i = 1 | \mathbf{Y}^i)}{P(a_n^i = 0 | \mathbf{Y}^i)} \\ &= \log \frac{\sum_{\beta_l \in \mathcal{A}: d_{\kappa(n)}^i = \beta_l, a_n^i = 1} P(d_{\kappa(n)}^i = \beta_l | \mathbf{Y}^i)}{\sum_{\beta_l \in \mathcal{A}: d_{\kappa(n)}^i = \beta_l, a_n^i = 0} P(d_{\kappa(n)}^i = \beta_l | \mathbf{Y}^i)}. \end{aligned} \quad (13)$$

Assume that the unknown quantities \mathbf{h}^i , $\mathbf{Z}^i \triangleq \{Z_k^i\}_{k=1}^{N-1}$ are independent of each other and have *a priori* distribution $p(\mathbf{h}^i)$ and $p(\mathbf{Z}^i)$, respectively. The direct computation of (12) is given by

$$P(d_k^i = a_l | \mathbf{Y}^i) \propto \sum_{\mathbf{Z}^i: d_k^i = a_l} \int p(\mathbf{Y}^i | \mathbf{h}^i, \mathbf{Z}^i) p(\mathbf{h}^i) p(\mathbf{Z}^i) d\mathbf{h}^i, \quad (14)$$

where $p(\mathbf{Y}^i | \mathbf{h}^i, \mathbf{Z}^i)$ is a Gaussian density function [cf. (4)]. Clearly, the computation in (14) involves a very high-dimensional integration which is certainly infeasible in practice. Therefore, we resort to the sequential Monte Carlo method for numerical evaluation of the above multidimensional integration.

4.2. SMC-based blind MAP detector

Sequential Monte Carlo (SMC) is a family of methodologies that use Monte Carlo simulations to efficiently estimate the *a posteriori* distributions of the unknown states in a dynamic system [11, 12, 13]. In [14], an SMC-based blind MAP symbol detection algorithm for OFDM systems is proposed. This algorithm is summarized as follows.

- (0) Initialization. Draw the initial samples of the channel vector from $\mathbf{h}_{-1}^{(j)} \sim N_c(\mathbf{0}, \mathbf{\Sigma}_{-1})$, for $j = 1, \dots, m$. All importance weights are initialized as $w_{-1}^{(j)} = 1$, $j = 1, \dots, m$.

The following steps are implemented at the k th recursion ($k = 0, \dots, N-1$) to update each weighted sample. For $j = 1, \dots, m$, the following hold.

- (1) For each $a_i \in \mathcal{A}$, compute the following quantities:

$$\mu_{k,i}^{(j)} = a_i \mathbf{w}_f^H(k) \mathbf{h}_{k-1}^{(j)},$$

$$\sigma_{k,i}^{2(j)} = \sigma^2 + \mathbf{w}_f^H(k) \mathbf{\Sigma}_{k-1}^{(j)} \mathbf{w}_f(k),$$

$$\alpha_{k,i}^{(j)} = \frac{1}{\pi \sigma_{k,i}^{2(j)}} \exp \left\{ -\frac{\|Y_k - \mu_{k,i}^{(j)}\|^2}{\sigma_{k,i}^{2(j)}} \right\} \cdot P(d_k = a_i Z_{k-1}^{(j)*}). \quad (15)$$

- (2) Impute the symbol Z_k . Draw $Z_k^{(j)}$ from the set \mathcal{A} with probability

$$P(Z_k = a_i | \mathbf{Z}_{k-1}^{(j)}, \mathbf{Y}_k) \propto \alpha_{k,i}^{(j)}, \quad a_i \in \mathcal{A}. \quad (16)$$

- (3) Compute the importance weight:

$$w_k^{(j)} = w_{k-1}^{(j)} \cdot \sum_{a_i \in \mathcal{A}} \alpha_{k,i}^{(j)}. \quad (17)$$

- (4) Update the *a posteriori* mean and covariance of the channel. If the imputed sample $Z_k^{(j)} = a_i$ in step (2), set $\mu_k^{(j)} = \mu_{k,i}^{(j)}$, $\sigma_k^{2(j)} = \sigma_{k,i}^{2(j)}$; and update

$$\mathbf{h}_k^{(j)} = \mathbf{h}_{k-1}^{(j)} + \frac{Y_k - \mu_k^{(j)}}{\sigma_k^{2(j)}} \boldsymbol{\xi}, \quad (18)$$

$$\mathbf{\Sigma}_k^{(j)} = \mathbf{\Sigma}_{k-1}^{(j)} - \frac{1}{\sigma_k^{2(j)}} \boldsymbol{\xi} \boldsymbol{\xi}^T,$$

with

$$\boldsymbol{\xi} \triangleq \mathbf{\Sigma}_{k-1}^{(j)} \mathbf{w}_f(k) Z_k^{(j)*}. \quad (19)$$

- (5) Perform resampling when k is a multiple of k_0 , where k_0 is the resampling interval.

4.3. APP detection

The above sampling procedure generates a set of random samples $\{(Z_k^{(j)}, w_k^{(j)})\}_{j=1}^m$, properly weighted with respect to the distribution $p(\mathbf{Z}_k | \mathbf{Y}_k)$. Based on these samples, an on-line estimation and a delayed-weight estimation can be obtained straightforwardly as

$$P(d_k = \beta_l | \mathbf{Y}_k) \cong \frac{1}{W_k} \sum_{j=1}^m \mathbf{1}(Z_{k+1}^{(j)} Z_k^{(j)*} = \beta_l) w_k^{(j)}, \quad (20)$$

$$P(d_k = \beta_l | \mathbf{Y}_{k+\delta}) \cong \frac{1}{W_{k+\delta}} \sum_{j=1}^m \mathbf{1}(Z_{k+1}^{(j)} Z_k^{(j)*} = \beta_l) w_{k+\delta}^{(j)},$$

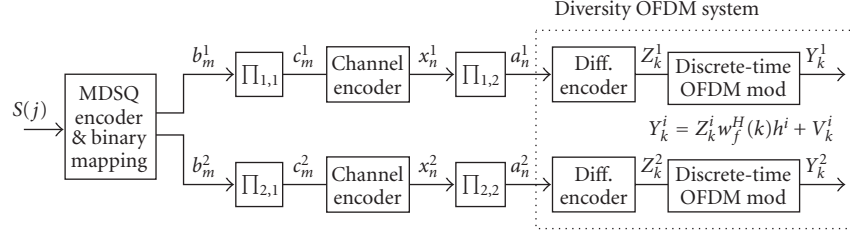


FIGURE 6: MDSQ over a channel-coded diversity OFDM system.

where $W_k \triangleq \sum_j w_k^{(j)}$, and $\mathbf{1}(\cdot)$ denotes the indicator function. Note that both of these two estimates are only approximations to the *a posteriori* symbol probability $P(d_k = \beta_l | \mathbf{Y}_{N-1})$.

We next propose a novel APP estimator, where the channel is estimated as a mixture vector, based on which the symbol APPs are then computed. Specifically, we have

$$p(\mathbf{h} | \mathbf{Y}_{N-1}) = \frac{1}{W_{N-1}} \sum_{j=1}^m \underbrace{p(\mathbf{h} | \mathbf{Y}_{N-1}, \mathbf{Z}_{N-1}^{(j)})}_{\mathcal{N}_c(\mathbf{h}_{N-1}^{(j)}, \Sigma_{N-1}^{(j)})} \cdot w_{N-1}^{(j)}. \quad (21)$$

The symbol *a posteriori* probability is then given by

$$\begin{aligned} & P(d_k = \beta_l | \mathbf{Y}_{N-1}) \\ &= \int P(d_k = \beta_l | \mathbf{Y}_{N-1}, \mathbf{h}) p(\mathbf{h} | \mathbf{Y}_{N-1}) d\mathbf{h} \\ &= \int P(d_k = \beta_l | \mathbf{Y}_{N-1}, \mathbf{h}) \\ &\quad \times \left[\frac{1}{W_{N-1}} \sum_{j=1}^m p(\mathbf{h} | \mathbf{Y}_{N-1}, \mathbf{Z}_{N-1}^{(j)}) \cdot w_{N-1}^{(j)} \right] d\mathbf{h} \\ &= \frac{1}{W_{N-1}} \sum_{j=1}^m w_{N-1}^{(j)} \cdot \left[\int P(d_k = \beta_l | \mathbf{Y}_{N-1}, \mathbf{h}) \right. \\ &\quad \left. \cdot p(\mathbf{h} | \mathbf{Y}_{N-1}, \mathbf{Z}_{N-1}^{(j)}) d\mathbf{h} \right] \\ &\propto \frac{1}{W_{N-1}} \sum_{j=1}^m w_{N-1}^{(j)} \\ &\quad \cdot \left[P(d_k = \beta_l) \sum_{Z_k, Z_{k-1}^* = \beta_l} \int P(\mathbf{Y}_{k-1}^k | \mathbf{Z}_{k-1}^k, \mathbf{h}) \right. \\ &\quad \left. \cdot p(\mathbf{h} | \mathbf{Y}_{N-1}, \mathbf{Z}_{N-1}^{(j)}) d\mathbf{h} \right], \end{aligned} \quad (22)$$

where $\mathbf{Y}_{k-1}^k \triangleq [Y_{k-1}, Y_k]^T$, $\mathbf{Z}_{k-1}^k \triangleq [Z_{k-1}, Z_k]^T$. Note that the integral within (22) is an integral of a Gaussian pdf with respect to another Gaussian pdf. The resulting distribution is

still Gaussian, that is,

$$\begin{aligned} & \int P(\mathbf{Y}_{k-1}^k | \mathbf{Z}_{k-1}^k, \mathbf{h}) \cdot p(\mathbf{h} | \mathbf{Y}_{N-1}, \mathbf{Z}_{N-1}^{(j)}) d\mathbf{h} \\ & \sim \mathcal{N}_c(\boldsymbol{\mu}_{k,j}(\mathbf{Z}_{k-1}^k), \boldsymbol{\Sigma}_{k,j}(\mathbf{Z}_{k-1}^k)), \end{aligned} \quad (23)$$

with mean and variance given, respectively, by

$$\begin{aligned} & \boldsymbol{\mu}_{k,j}(\mathbf{Z}_{k-1}^k) \\ &= \begin{bmatrix} \mu_{k,j}(\mathbf{Z}_k) \\ \mu_{k-1,j}(\mathbf{Z}_{k-1}) \end{bmatrix}, \quad \text{with } \mu_{k,j}(x) \triangleq \mathbf{x} \mathbf{w}_k^H \mathbf{h}_{N-1}^{(j)}, \end{aligned} \quad (24)$$

$$\begin{aligned} & \boldsymbol{\Sigma}_{k,j}(\mathbf{Z}_{k-1}^k) \\ &= \begin{bmatrix} \sigma_{k,j}^2 & 0 \\ 0 & \sigma_{k-1,j}^2 \end{bmatrix}, \quad \text{with } \sigma_{k,j}^2 \triangleq \mathbf{w}_k^H \boldsymbol{\Sigma}_{N-1}^{(j)} \mathbf{w}_k + \sigma^2. \end{aligned} \quad (25)$$

Equations (24) and (25) follow from the fact that conditioned on the channel \mathbf{h} , Y_k and Y_{k+1} are independent. The symbol *a posteriori* probability can then be computed in a close form as

$$\begin{aligned} & P(d_k = \beta_l | \mathbf{Y}_{N-1}) \\ & \approx \sum_{j=1}^m \sum_{Z_k, Z_{k-1}^* = \beta_l} w_N^{(j)} \\ & \quad \cdot \frac{P(d_k = \beta_l)}{\sigma_{k,j}^2 + \sigma_{k-1,j}^2} \exp \left\{ - \frac{|Y_k - \mu_{k,j}(\mathbf{Z}_k)|^2}{\sigma_{k,j}^2} \right. \\ & \quad \left. - \frac{|Y_{k-1} - \mu_{k-1,j}(\mathbf{Z}_{k-1})|^2}{\sigma_{k-1,j}^2} \right\}. \end{aligned} \quad (26)$$

5. CHANNEL-CODED SYSTEMS

Although the MDSQ introduces some redundancy to the system, it has limited capability for error correction. In order to improve the system reliability, we next consider introducing channel coding to the proposed MDSQ system.

A block diagram of an MDSQ system over a channel-coded diversity OFDM system is shown in Figure 6. A stream of source signal $\{S(j)\}_j$ is MDSQ encoded, resulting in two sets of indices $\{I_1(j), I_2(j)\}_j$. Binary descriptions of these

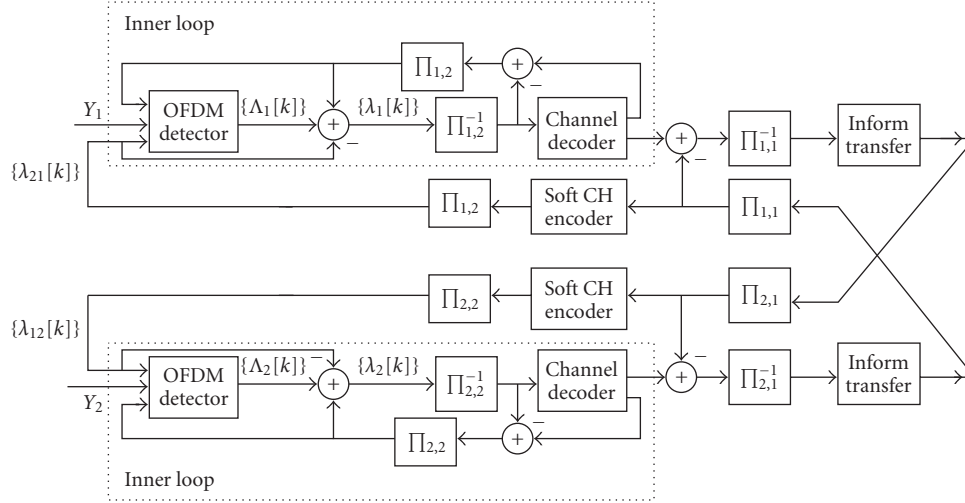


FIGURE 7: Turbo decoding for MDSQ over a channel-coded diversity OFDM system.

indices, $\{b_m^1, b_m^2\}_m$, are then channel encoded and OFDM modulated. There are two sets of bit interleavers in the system: one set, named $\{\Pi_{i,1}\}_{i=1}^2$, is applied between the MDSQ encoder and channel encoder; the other set, named $\{\Pi_{i,2}\}_{i=1}^2$, is applied between the channel encoder and OFDM modulator.

At the receiver, a novel blind iterative receiver is developed for joint demodulation, channel decoding, and MDSQ decoding. The receiver structure, as shown in Figure 7, consists of two loops of iterative operations. For each description, there is an inner loop (iterative procedure) for joint OFDM demodulation and channel decoding. At the outer loop, soft information of the coded bits is exchanged between the two inner loops to exploit the correlations between the two descriptions. Next, we discuss the operation of both the inner loop and the outer loop.

Inner loop: joint OFDM demodulation and channel decoding

We consider a subsystem of the original MDSQ system, which consists of the channel coding and OFDM modulation for only one source description. Since the combination of a differential encoder and OFDM system acts as an inner encoder, the above subsystem is a typical serial concatenated code, and an iterative (turbo) receiver can be designed for such a system, which is denoted as the inner loop part in Figure 7. It consists of two stages: the SMC OFDM detector developed in the previous sections, followed by a MAP channel decoder [15]. The two stages are separated by a deinterleaver and an interleaver. Note that both the SMC OFDM detector and the MAP channel decoder can incorporate the *a priori* probabilities and output *a posteriori* probabilities of the code bits $\{a_n^i\}_n$, that is, they are soft-input and soft-output algorithms. Based on the turbo principle, extrinsic information of the channel-coded bits can be

exchanged iteratively between the SMC OFDM detector and the MAP channel decoder to improve the performance of the subsystem.

Outer loop: exploiting the correlation between the two descriptions

In Section 3, an iterative receiver was proposed for joint MDSQ decoding and OFDM demodulation. Extrinsic information from one description is transformed into the soft information for the other description, and is fed into the OFDM demodulator as the *a priori* information. For channel-coded MDSQ systems, similar approaches can be considered to exploit the correlation between the two descriptions. As shown in Figure 7, the MAP channel decoder incorporates the *a priori* information for the channel-coded bits, and outputs the *a posteriori* probability of both channel-coded bits and uncoded bits. On the other hand, the OFDM detector incorporates and produces as output only the soft information for the channel-coded bits. Taking into account that only uncoded bits will be considered in the MDSQ decoder, the inner loop, when considered as one unit operation, is a SISO algorithm that incorporates the *a priori* information of the channel-coded bits, and produces the output *a posteriori* information of the uncoded bits. Altogether, the two inner loops constitute a turbo structure in parallel, and the transferred soft information provided by the information transformation block (IF-T) can be exchanged iteratively between the two inner loops. This iterative procedure is the outer loop of the system, which aims at further improving the system performance by exploiting the correlation between the two descriptions. It is shown in Section 3 that this correlation can be measured by the probability transformation matrix, and adopted by the IF-T block. For the outer loop, the soft output of the inner loop can be used directly as the *a priori* information for

the IF-T; the soft output of IF-T, however, must be transformed before being fed into the inner loop as *a priori* information. Specifically, a soft channel encoder by the BCJR algorithm [15] is required to transform the soft information of the uncoded bits into the soft information of the coded bits.

6. SIMULATION RESULTS

In this section, we provide computer simulation results to illustrate the performance of the turbo receiver for MDSQ over diversity OFDM systems. In the simulations, the continuous alphabet source is assumed to be uniformly distributed on $(-1, 1)$, and a uniform quantizer is applied. The source range is divided into 8, 22, and 34 intervals. Two indices are assigned to describe the source according to the index assignment $\alpha(\cdot)$ as shown in Figure 3, where each index is described with $R = 3$ bits. Assume the channel bandwidth for each OFDM system is divided into $N = 128$ subchannels. Guard interval is long enough to protect the OFDM blocks from intersymbol interference due to the delay spread. The frequency-selective fading channels are assumed to be uncorrelated. All $L = 5$ taps of the fading channel are Rayleigh distributed with the same variance, normalized such that $E\{\sum_{n=0}^{L-1} \|h_n\|^2\} = 1$, and have delays $\tau_l = l/\Delta_f$, $l = 0, 1, \dots, L - 1$. For channel-coded systems, a rate-1/2 constraint length-5 convolutional code (with generators 23 and 35 in octal notation) is used. The interleavers are generated randomly and fixed for all simulations.

The blind SMC detector implements the algorithm described in Section 4.2. The variance of the noise V_k in (24) is assumed known at the detector with values specified by the given SNR. The SMC algorithm draws $m = 50$ Monte Carlo samples at every recursion with Σ_{-1} set to $1000\mathbf{I}_L$. Two qualities were used in the simulation to measure the performance of the SMC detector: bit error rate (BER) and word error rate (WER). Here, the bit error rate denotes the information bit error rate and word error rate denotes the error rate of the whole data block transferred during one symbol duration. On the other hand, mean square error (MSE) will be used to measure the performance of the whole system.

Performance of the SMC detector

The blind SMC detector, as a SISO algorithm for OFDM demodulation, is an important component of the proposed turbo receiver. Next, we illustrate the performance of the blind SMC detector. In Figure 8, the BER and WER performance is plotted. In the same figure, we also plot the known channel lower bound, where the fading coefficients are assumed to be perfectly known to the receiver and a MAP receiver is employed to compute the *a posteriori* symbol probabilities.

Although the SMC detector generates soft outputs in terms of the symbol *a posteriori* probabilities, only hard decisions are used in an uncoded system. However in a coded system, the channel decoder, such as a MAP decoder, requires

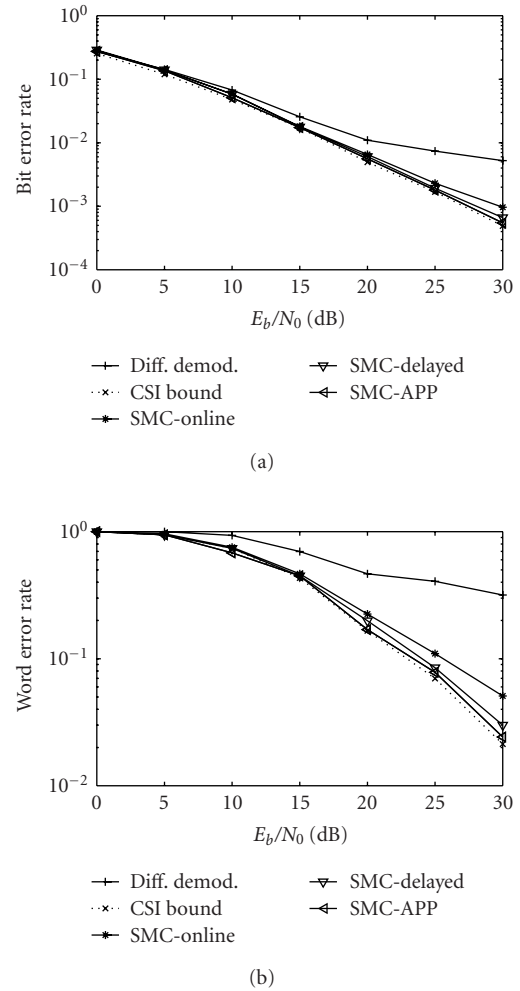
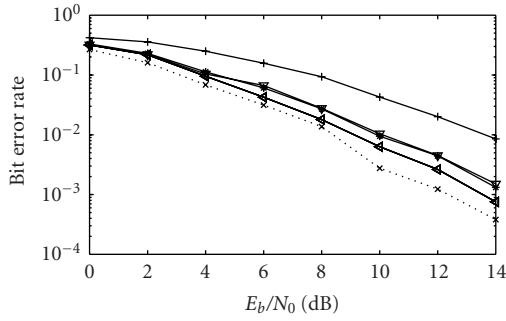


FIGURE 8: The (a) BER and (b) WER performance in an uncoded OFDM system.

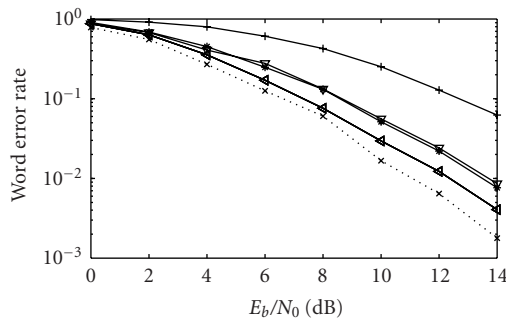
soft information provided by the demodulator. Next, we examine the accurateness of the soft output provided by the SMC detector in a coded OFDM scenario. In Figure 9, the BER and WER performance for the information bits is plotted. In the same figure, the known channel lower bound is also plotted. The MAP convolutional decoder is employed in conjunction with the different detection algorithms. It is seen from Figure 9 that the three SMC detector yield different performance after the MAP decoder because of the different quality of the soft information they provide. Specifically, the APP detector achieves the best performance.

Performance of turbo receiver for MDSQ system

The performance of the turbo receiver is shown in Figures 10, 11, and 12 for MDSQ systems with assignments 8, 22, and 34, respectively, as in Figure 3. The SMC blind detector is employed. In each figure, the BER, WER, and MSE are plotted. In the same figure, the quantization error bound $s^2/12$, where



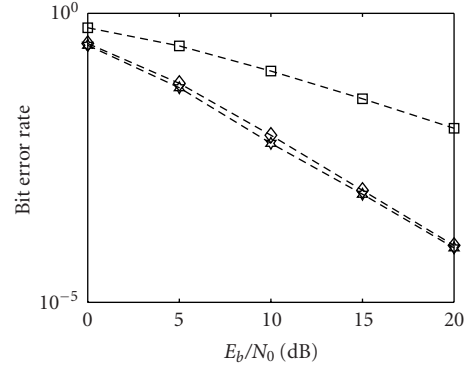
(a)



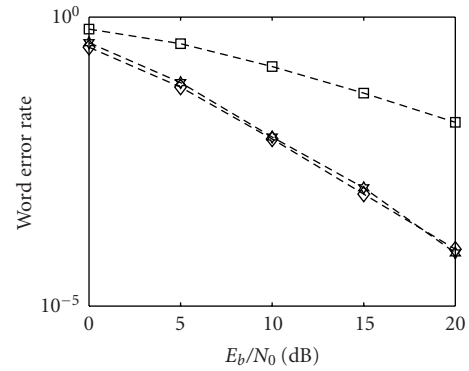
(b)

FIGURE 9: The (a) BER and (b) WER performance in a channel-coded OFDM system.

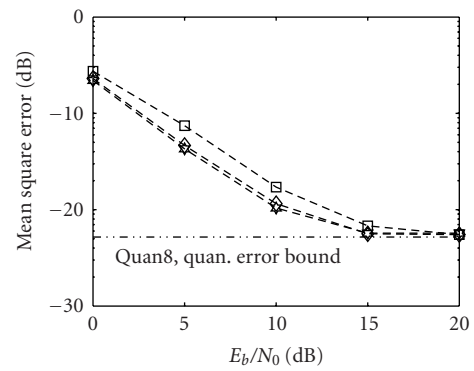
s denote the quantization interval, is also plotted in a dotted line. It is seen that the BER and WER performance is significantly improved at the second iteration, that is, 15 dB better for $N = 8$, 4 dB better for $N = 22$ and 2 dB better for $N = 34$. However, no significant gain is achieved by more iterations. Note that the MSEs of the turbo receivers are very close to the quantization error bound at high SNR. The quantization error bound (5.2×10^{-3}) for $N = 8$ is achieved at about 15 dB. However, much lower quantization error bounds are achieved at higher SNR by the turbo receiver with $N = 22$ and 34, that is, 6.9×10^{-4} for $N = 22$ at SNR = 25 dB and 2.8×10^{-4} for $N = 34$ at SNR = 30 dB. Moreover, due to the different quantization error bounds determined by N and the BER and the WER performance achieved by the turbo receiver, different MDSQ scheme should be chosen at different SNRs to minimize the MSE. For example, the MDSQ with $N = 8$ is superior to other assignments below SNR = 10 dB. However, at SNR = 20 dB, the MDSQ scheme with $N = 22$ is the best choice among the three assignments considered in this paper.



(a)

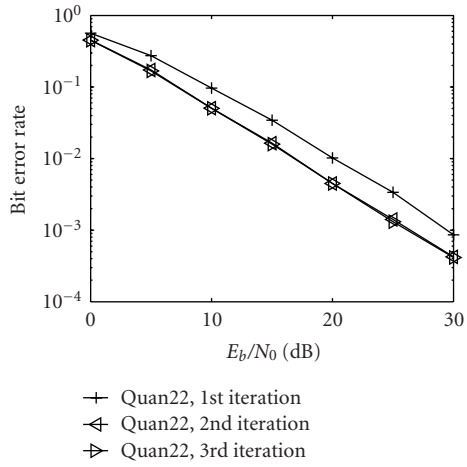


(b)

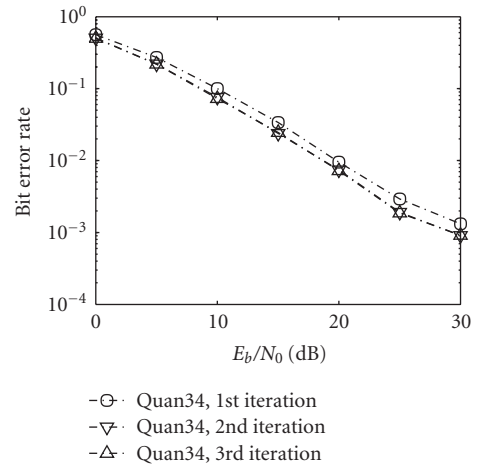


(c)

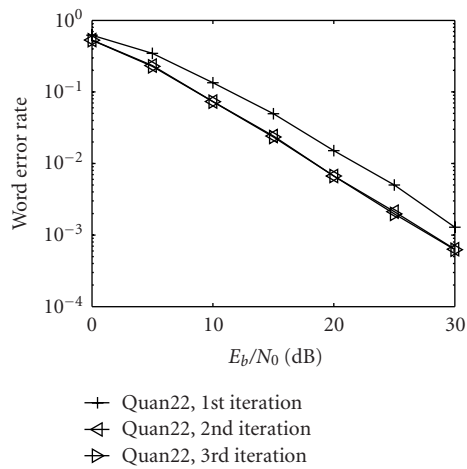
FIGURE 10: Performance of iterative receiver for the MDSQ system with $N = 8$. (a) BER. (b) WER. (c) MSE.



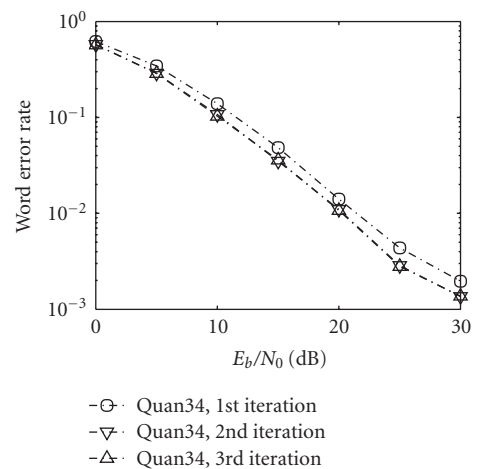
(a)



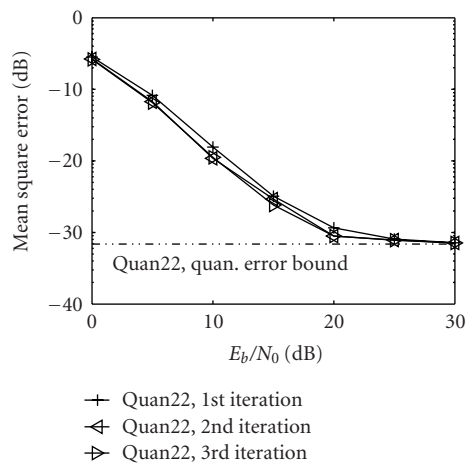
(a)



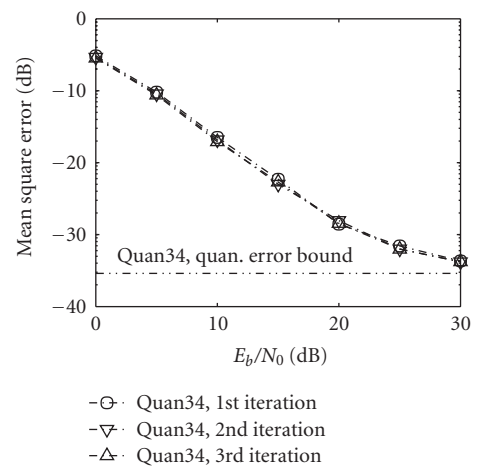
(b)



(b)



(c)



(c)

FIGURE 11: Performance of iterative receiver for the MDSQ system with $N = 22$. (a) BER. (b) WER. (c) MSE.

FIGURE 12: Performance of iterative receiver for the MDSQ system with $N = 34$. (a) BER. (b) WER. (c) MSE.

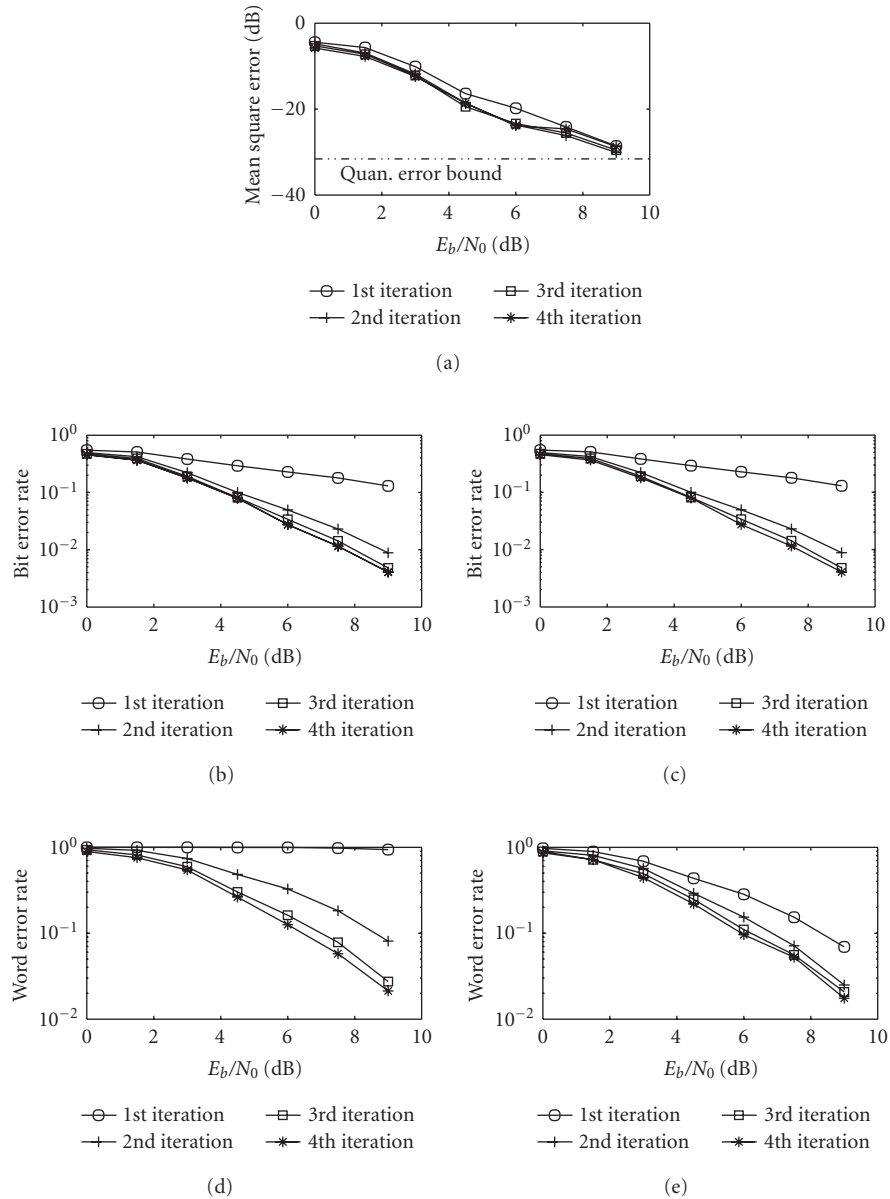


FIGURE 13: Performance of iterative receiver for channel coded MDSQ system, with 1 iteration for inner loop and 4 iterations for outer loop. (a) MSE. (b) BER of coded bits. (c) BER of information bits. (d) WER of coded bits. (e) WER of information bits.

Performance of turbo receiver for channel-coded MDSQ system

Finally, we consider the performance of the channel-coded MDSQ system discussed in Section 5. Performance is compared for systems with different iterative profiles. Specifically, the BER, WER, and MSE performance for the information bits and coded bits are plotted in Figures 13 and 14 for the 4-inner-loop and 1-outer-loop turbo receivers and the 3-inner-loop and 2-outer-loop turbo receivers, respectively. In the simulation, the source range is divided into 22 intervals as shown in Figure 3b. It is seen that the proposed turbo receiver structure can successively improve

the receiver performance through iterative processing. Moreover, the quantization error bounds are achieved at very low SNR, that is, 10 dB.

7. CONCLUSIONS

In this paper, we have proposed a blind turbo receiver for transmitting MDSQ-coded sources over frequency-selective fading channels. Transformation of the extrinsic information of the two descriptions are exchanged between each other to improve the system performance. A novel blind APP OFDM detector, which computes the *a posteriori* symbol

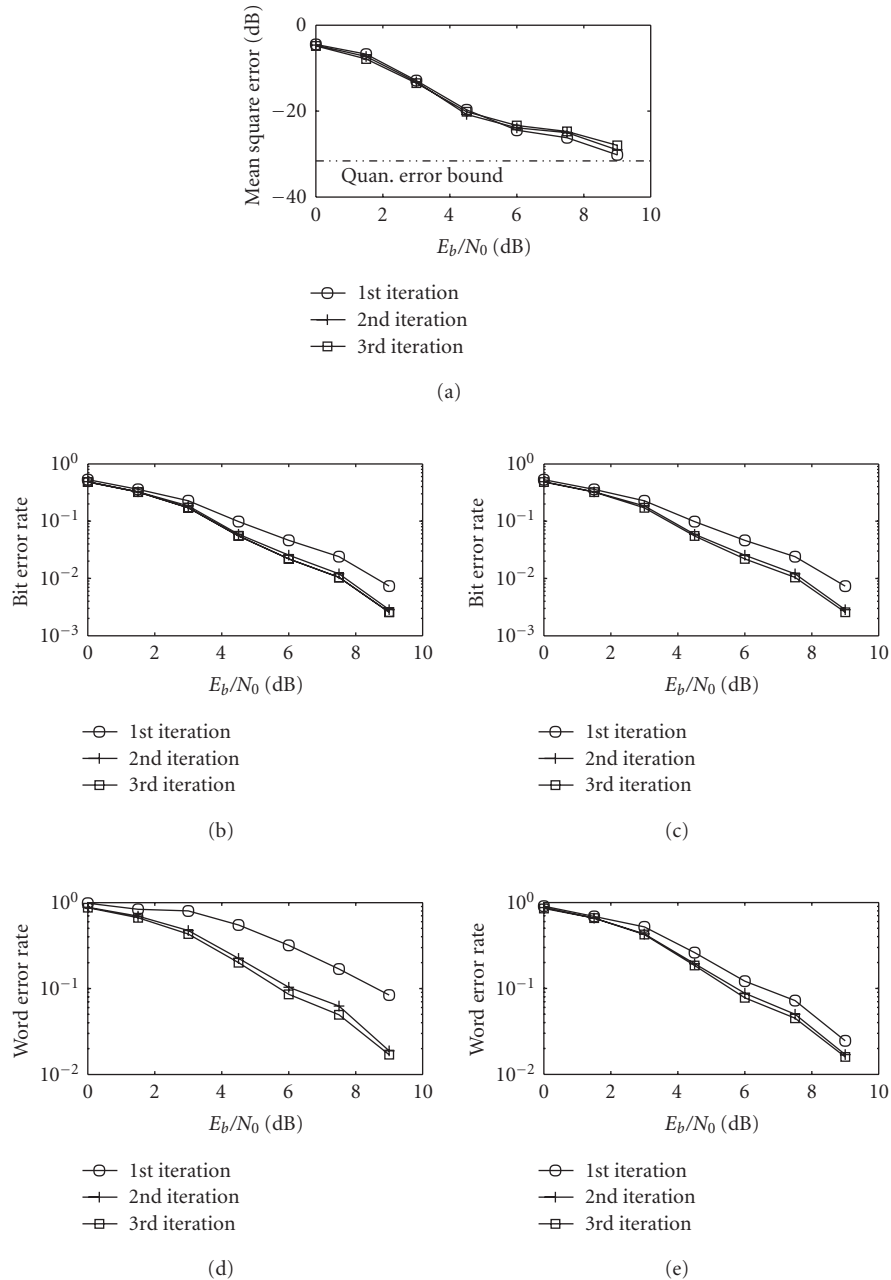


FIGURE 14: Performance of iterative receiver for channel-coded MDSQ system, with 2 iterations for inner loop and 3 iterations for outer loop. (a) MSE. (b) BER of coded bits. (c) BER of information bits. (d) WER of coded bits. (e) WER of information bits.

probabilities, is developed using sequential Monte Carlo (SMC) techniques. Being soft-input and soft-output in nature, the proposed SMC detector is capable of exchanging the so-called extrinsic information with other component in the above turbo receiver, and successively improving the overall receiver performance. Finally, we have also treated channel-coded systems, and a novel blind turbo receiver is developed for joint demodulation, channel decoding, and MDSQ decoding. Simulation results have demonstrated the effectiveness of the proposed techniques.

REFERENCES

- [1] A. E. Gamal and T. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inform. Theory*, vol. 28, no. 6, pp. 851–857, 1982.
- [2] V. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 821–834, 1993.
- [3] Y. Zhang, M. Motani, and H. Garg, "Wireless video transmission using multiple description codes combined with prioritized DCT compression," in *IEEE International Conference on*

Multimedia and Expo (ICME '02), vol. 1, pp. 261–264, Lausanne, Switzerland, August 2002.

- [4] L. Ozarow, "On a source coding problem with two transmitters and three receivers," *Bell Labs Technical Journal*, vol. 59, no. 10, pp. 1909–1921, 1980.
- [5] S.-M. Yang and V. Vaishampayan, "Low-delay communication for Rayleigh fading channels: an application of the multiple description quantizer," *IEEE Trans. Commun.*, vol. 43, no. 11, pp. 2771–2783, 1995.
- [6] J. Barros, J. Hagenauer, and N. Gortz, "Turbo cross decoding of multiple descriptions," in *Proc. IEEE International Conference on Communications (ICC '02)*, vol. 3, pp. 1398–1402, New York, NY, USA, April 2002.
- [7] N. Kamaci, Y. Altunbasak, and R. Mersereau, "Multiple description coding with multiple transmit and receive antennas for wireless channels: the case of digital modulation," in *IEEE Global Telecommunications Conference (GLOBECOM '01)*, pp. 3272–3276, San Antonio, Tex, USA, November 2001.
- [8] D. Sachs, R. Anand, and K. Ramchandran, "Wireless image transmission using multiple-description based concatenated codes," in *Proc. IEEE Data Compression Conference (DCC '00)*, p. 569, Snowbird, Utah, USA, March 2000.
- [9] K. Balachandran and J. Anderson, "Mismatched decoding of intersymbol interference using a parallel concatenated scheme," *IEEE J. Select. Areas Commun.*, vol. 16, no. 2, pp. 255–259, 1998.
- [10] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Börjesson, "On channel estimation in OFDM systems," in *Proc. IEEE Vehicular Technology Conference (VTC '95)*, pp. 815–819, Chicago, Ill, USA, July 1995.
- [11] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo in Practice*, Springer-Verlag, New York, NY, USA, 2001.
- [12] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.
- [13] X. Wang, R. Chen, and J. Liu, "Monte Carlo signal processing for wireless communications," *Journal of VLSI Signal Processing*, vol. 30, no. 1-3, pp. 89–105, 2002.
- [14] Z. Yang and X. Wang, "A sequential Monte Carlo blind receiver for OFDM systems in frequency-selective fading channels," *IEEE Trans. Signal Processing*, vol. 50, no. 2, pp. 271–280, 2002.
- [15] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (Corresp.)," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, 1974.

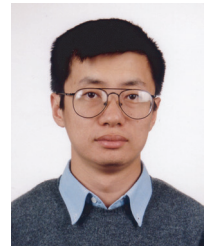
Zigang Yang received the B.S. degree in electrical engineering and applied mathematics in 1995, and the M.S. degree in electrical engineering in 1998, both from Shanghai Jiaotong University (SJTU), Shanghai, China. In 2002, she got the Ph.D. degree in electrical engineering from Texas A&M University, College Station, Texas. From 1999 till 2002, she was a Research Assistant with the Department of Electrical Engineering, Texas A&M University. Currently, she is working as a system engineer at Texas Instrument, Communication R&D Lab. Her research interests are in the area of statistical signal processing and its applications, primarily in digital communications.



Dong Guo received the B.S. degree in geophysics and computer science from China University of Mining and Technology (CUMT), Xuzhou, China, in 1993, and the M.S. degree in geophysics from the Graduate School of Research Institute of Petroleum Exploration and Development (RIPED), Beijing, China, in 1996. In 1999, he received the Ph.D. degree in applied mathematics from Beijing University, Beijing, China. In 2004, he received a second Ph.D. degree in electrical engineering from Columbia University, New York. His research interests are in the area of statistical signal processing and communications.



Xiaodong Wang received the B.S. degree in electrical engineering and applied mathematics (with the highest honors) from Shanghai Jiao Tong University, Shanghai, China, in 1992; the M.S. degree in electrical and computer engineering from Purdue University in 1995; and the Ph.D. degree in electrical engineering from Princeton University in 1998. From July 1998 to December 2001, he was on the faculty of the Department of Electrical Engineering, Texas A&M University. In January 2002, he joined the Department of Electrical Engineering, Columbia University. Dr. Wang's research interests fall in the general areas of computing, signal processing, and communications. Among his publications is a recent book entitled *Wireless Communication Systems: Advanced Techniques for Signal Reception*, published by Prentice Hall. Dr. Wang has received the 1999 NSF CAREER Award. He has also received the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award. He currently serves as an Associate Editor for the IEEE Transactions on Signal Processing, the IEEE Transactions on Communications, the IEEE Transactions on Wireless Communications, and IEEE Transactions on Information Theory.



Adaptive Iterative Soft-Input Soft-Output Parallel Decision-Feedback Detectors for Asynchronous Coded DS-CDMA Systems

Wei Zhang

School of Information Technology and Engineering (SITE), University of Ottawa, 800 King Edward Avenue, Ottawa, ON, Canada K1N 6N5
Email: weizhang@site.uottawa.ca

Claude D'Amours

School of Information Technology and Engineering (SITE), University of Ottawa, 800 King Edward Avenue, Ottawa, ON, Canada K1N 6N5
Email: damours@site.uottawa.ca

Abbas Yongaçoğlu

School of Information Technology and Engineering (SITE), University of Ottawa, 800 King Edward Avenue, Ottawa, Ontario, Canada K1N 6N5
Email: yongacog@site.uottawa.ca

Received 29 April 2004; Revised 4 October 2004

The optimum and many suboptimum iterative soft-input soft-output (SISO) multiuser detectors require a priori information about the multiuser system, such as the users' transmitted signature waveforms, relative delays, as well as the channel impulse response. In this paper, we employ adaptive algorithms in the SISO multiuser detector in order to avoid the need for this a priori information. First, we derive the optimum SISO parallel decision-feedback detector for asynchronous coded DS-CDMA systems. Then, we propose two adaptive versions of this SISO detector, which are based on the normalized least mean square (NLMS) and recursive least squares (RLS) algorithms. Our SISO adaptive detectors effectively exploit the a priori information of coded symbols, whose soft inputs are obtained from a bank of single-user decoders. Furthermore, we consider how to select practical finite feedforward and feedback filter lengths to obtain a good tradeoff between the performance and computational complexity of the receiver.

Keywords and phrases: soft-input soft-output multiuser detection, adaptive multiuser detection, parallel decision-feedback detection, adaptive soft-input soft-output parallel decision-feedback detection, asynchronous coded CDMA systems.

1. INTRODUCTION

Iterative soft-input soft-output (SISO) multiuser receivers for coded multiuser systems have received widespread attention since they can provide near single-user performance in a system with multiple-access interference (MAI) by iteratively combining multiuser detection and single-user decoding. The optimum SISO multiuser detector employs either the cross-entropy minimization [1] or the maximum *a posteriori* (MAP) algorithm [2]. The computational complexity of these techniques is exponentially proportional to the

number of users which can be prohibitive for large systems. Therefore, much work has been done on reduced-complexity suboptimum SISO multiuser detectors.

SISO multiuser detection based on the reduced-complexity MAP algorithms which are applied to the trellis of the multiple-access channel is proposed in [3, 4]. The simplest SISO multiuser detector is the soft interference canceller proposed in [5, 6], which has a linear computational complexity in terms of the number of users. However, it slowly converges to the performance of the single-user system. Linear iterative SISO multiuser detectors, which employ a decorrelator [7] or a minimum mean square error (MMSE) filter [8] on the output of the soft interference cancellation, significantly improve the system performance. Moreover, their computational complexity is only a cubic function of the number

This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

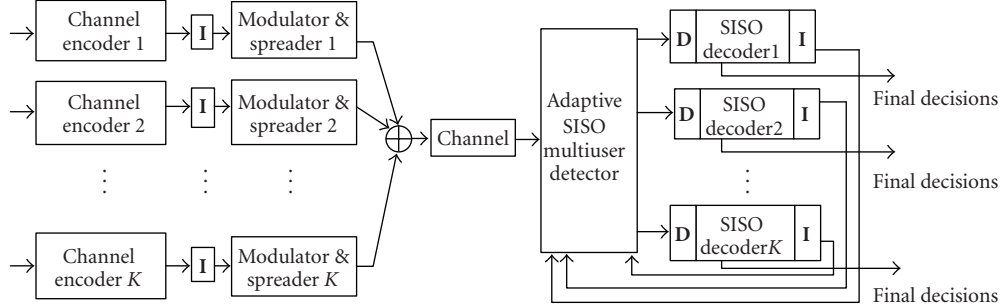


FIGURE 1: A general coded DS-CDMA system with an iterative receiver (I and D denote interleavers and deinterleavers, respectively).

of users. In [9, 10], nonlinear MMSE-based SISO decision-feedback detectors are investigated.

The above optimum and suboptimum SISO multiuser detectors require accurate a priori information about the multiuser system, such as all users' received signature waveforms which are functions of their transmitted signature waveforms, relative delays, and the channel impulse response. In practical situations, this information may not be easily obtainable for time-varying fading channels.

Fortunately, if the system parameters are constant or slowly varying, adaptive detectors (non-SISO) can successfully track these parameters from the received signal [11, 12, 13, 14, 15]. In [16], an adaptive SISO parallel decision-feedback detector for synchronous direct-sequence code-division multiple-access (DS-CDMA) systems with short spreading sequences is presented. By employing an approximate least squares algorithm and soft symbol estimates, the detector exploits the joint statistics of soft symbol estimates and transmitted symbols.

In this paper, we use adaptive algorithms in the iterative SISO parallel decision-feedback detector (PDFD) for asynchronous coded DS-CDMA systems in order to avoid the need for the a priori information about system parameters, such as multiple users' spreading codes and relative delays between users. First, we derive the optimum SISO parallel decision-feedback detector assuming the receiver knows the transmitted signature waveforms and relative delays between all the users. Then, we propose two adaptive versions of this SISO detector, which employ the normalized least mean square (NLMS) and recursive least squares (RLS) algorithms to estimate the filter coefficients of the detector. All users are assumed to employ short spreading codes. A training sequence is required for each user. Our adaptive SISO detectors effectively exploit the a priori information of coded symbols, which is obtained from the soft outputs of a bank of single-user decoders, to further improve their convergence performance.

Furthermore, for adaptive implementation of the SISO PDFD for asynchronous DS-CDMA systems, we select practical finite feedforward and feedback filter lengths to obtain a good tradeoff between the system performance and computational complexity of the receiver. We employ a feedforward filter which covers a two-symbol duration for each user and we consider several options for the feedback filter length.

Monte-Carlo simulation results for these adaptive SISO detectors are presented and compared.

The outline of the rest of this paper is as follows. A system model of asynchronous coded DS-CDMA systems is introduced in Section 2. The optimum SISO PDFD with a general processing window for asynchronous coded DS-CDMA systems is derived in Section 3. Adaptive SISO PDFDs are proposed in Section 4, which are based on the NLMS and RLS algorithms. Monte-Carlo simulation results are presented and compared in Section 5. Finally in Section 6, the conclusions are given.

2. SYSTEM MODEL AND NOTATION

Throughout the paper, matrices and vectors are denoted as boldface uppercases and lowercases, respectively. Notations $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^T$ denote the complex conjugate, Hermitian transpose, and transpose, respectively.

A general coded DS-CDMA system with an iterative receiver is shown in Figure 1. There are K active users in the system. The information bits of each user are first encoded, then interleaved, modulated, and spread before they are transmitted over the channel. The iterative receiver consists of two parts, an adaptive soft-input soft-output multiuser detector and a bank of SISO single-user decoders, which are separated by deinterleavers and interleavers. These two parts cooperate iteratively by transferring updated extrinsic soft information of coded symbols between them.

In our paper, we consider an asynchronous coded DS-CDMA system over the additive white Gaussian noise (AWGN) channel. The equivalent baseband received multiuser signal is

$$r(t) = \sum_{k=1}^K \sum_{i=1}^{N_b} b_k(i) s_k(t - iT - \tau_k) + n(t), \quad (1)$$

where K is the number of active users, N_b is the number of symbols transmitted by each user, $b_k(i)$ is the i th coded symbol of the k th user, $s_k(t)$ is its transmitted signature waveform, τ_k and T are the delay of user k and the symbol interval, respectively, and $n(t)$ is an additive white Gaussian noise process with double-sided power spectral density $N_0/2$. Each user's information bits are encoded and then BPSK modulated, that is, $b_k(i) \in \{+1, -1\}$.

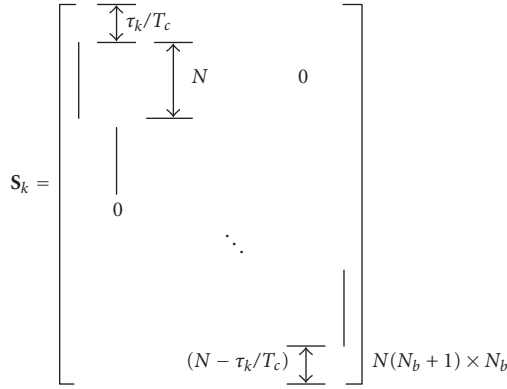


FIGURE 2: System signature matrix \mathbf{S}_k of user k , where the nonzero part of each column is the signature vector \mathbf{s}_k of user k .

For simple implementation, we consider a chip-synchronous and symbol-asynchronous DS-CDMA system. All users' delays are uniformly distributed in $[0, T]$ and are multiples of T_c , which is the chip interval. In the receiver, first we employ a chip-matched filter on the received signal $r(t)$ and then sample its output at frequency $1/T_c$. If the system is chip-asynchronous, we can oversample the output of the chip-matched filter and design a fractionally spaced feedforward filter instead. Without loss of generality and for simplicity of notation, we assume the delays of multiple users satisfy the following inequality:

$$0 \leq \tau_1 \leq \tau_2 \leq \dots \leq \tau_K \leq T. \quad (2)$$

The symbol vector consisting of the transmitted symbols of all users is denoted as

$$\mathbf{b} = [\mathbf{b}_1^T, \dots, \mathbf{b}_k^T, \dots, \mathbf{b}_K^T]_{KN_b \times 1}^T, \quad (3)$$

where

$$\mathbf{b}_k = [b_k(1), b_k(2), \dots, b_k(N_b)]^T. \quad (4)$$

The received signal vector \mathbf{r} at the output of the chip-matched filter during the whole symbol transmission interval can be expressed as follows:

$$\mathbf{r} = \mathbf{S}\mathbf{b} + \mathbf{n}, \quad (5)$$

where \mathbf{S} is the system signature matrix and can be expressed as

$$\mathbf{S} = [\mathbf{S}_1, \dots, \mathbf{S}_k, \dots, \mathbf{S}_K]_{N(N_b+1) \times KN_b}. \quad (6)$$

The construction of \mathbf{S}_k in (6) is shown in Figure 2, where the nonzero part of each column is the signature vector \mathbf{s}_k of user k and N is the number of chips per coded symbol. The vector \mathbf{n} in (5) is an $N(N_b + 1) \times 1$ column vector which represents the output noise component of the chip-matched filter. It has zero mean and covariance matrix $\sigma_n^2 \mathbf{I}$, where σ_n^2 is the variance of the output noise component.

3. OPTIMUM SISO PDFD FOR ASYNCHRONOUS DS-CDMA SYSTEMS

In general, the optimum SISO PDFD filters for asynchronous DS-CDMA systems have infinite lengths [17]. For implementation purposes, we consider finite-length feedforward and feedback filters. Furthermore, these filters are suitable for use in adaptive applications. The use of these filters in our adaptive detectors will be discussed in detail in Section 4.

In the receiver, we assume that the processing window length is N_p , which is measured in chips and is much less than $N_b \times N$. In each processing window, the received signal vector is denoted as $\bar{\mathbf{r}}_{N_p \times 1}$, which consists of N_p rows of \mathbf{r} falling to this processing window. The windowed system signature matrix $\bar{\mathbf{S}}_{N_p \times KN_b}$ and noise vector $\bar{\mathbf{n}}_{N_p \times 1}$ consist of N_p corresponding rows of \mathbf{S} and \mathbf{n} , respectively. Therefore, we have the following equation:

$$\bar{\mathbf{r}} = \bar{\mathbf{S}}\mathbf{b} + \bar{\mathbf{n}}. \quad (7)$$

We can write \mathbf{b} as the following sum:

$$\mathbf{b} = \mathbf{b}_U + \mathbf{b}_D, \quad (8)$$

where \mathbf{b}_U consists of the symbols which are not feedback and its other elements are zeros. The nonzero elements of \mathbf{b}_D consist of the feedback symbols. They have no common elements. In the same way by which we construct \mathbf{b}_U and \mathbf{b}_D , we extract columns of $\bar{\mathbf{S}}$ and construct the corresponding signature matrices $\bar{\mathbf{S}}_U$ and $\bar{\mathbf{S}}_D$. Therefore, the windowed received signal vector $\bar{\mathbf{r}}$ can also be expressed as

$$\bar{\mathbf{r}} = \bar{\mathbf{S}}_U \mathbf{b}_U + \bar{\mathbf{S}}_D \mathbf{b}_D + \bar{\mathbf{n}}. \quad (9)$$

The feedforward filter of user k has N_p taps and is denoted by a column vector $\bar{\mathbf{m}}_{fk}$. The feedback filter $\bar{\mathbf{m}}_{bk}$ of user k has the size $KN_b \times 1$, whose nonzero elements are corresponding to feedback symbols. That is, its effective number of taps is determined by the number of feedback symbols. The optimum filters satisfy the following minimum mean square error (MMSE) criterion:

$$\min_{\bar{\mathbf{m}}_{fk}, \bar{\mathbf{m}}_{bk}} E \left[b_k(i) - \bar{\mathbf{m}}_{fk}^H \cdot \bar{\mathbf{r}} - \bar{\mathbf{m}}_{bk}^H \cdot \hat{\mathbf{b}}_D \right]^2. \quad (10)$$

Nonzero elements of $\hat{\mathbf{b}}_D$ are soft symbol estimates of those elements of \mathbf{b}_D , respectively. We will introduce the soft symbol estimate of each coded symbol in the following.

The soft inputs of a SISO multiuser detector, $\{\lambda_{\text{in}}[b_k(j)], 1 \leq k \leq K, 1 \leq j \leq N_b\}$, are extrinsic log-likelihood ratios (LLRs) of $\{b_k(j)\}$ provided by a bank of K single-user decoders. Based on these inputs, we can obtain the soft symbol estimate of $\{b_k(j)\}$:

$$\hat{b}_k(j) = E[b_k(j) | \lambda_{\text{in}}[b_k(j)]] = \tanh \left(\frac{\lambda_{\text{in}}[b_k(j)]}{2} \right). \quad (11)$$

Furthermore, we have the following a priori statistics (12) for nonzero elements of \mathbf{b}_U and \mathbf{b}_D . For feedback symbols, their mean values are their soft symbol estimates, while nonfeedback symbols have zero mean. Note that $b_k(i)$ in (10) belongs

to nonfeedback symbols. Denote u and v as one of the nonzero elements of \mathbf{b}_U and \mathbf{b}_D , respectively. The soft symbol estimate of v is denoted as \hat{v} . Thus, we have

$$\begin{aligned} E[u] &= 0, \\ E[u^2] &= 1, \\ E[v] &= \hat{v}, \\ E[v^2] &= 1 - (\hat{v})^2. \end{aligned} \quad (12)$$

We also assume that all users' transmitted symbols are independent of one another and of the background noise vector $\bar{\mathbf{n}}$ as well.

Employing the above statistics about the coded symbols, we can get the optimum feedforward and feedback filters of user k which satisfy the MMSE criterion in (10):

$$\bar{\mathbf{m}}_{fk} = (\bar{\mathbf{R}}_U + \bar{\mathbf{R}}_D + \sigma_n^2 \mathbf{I})^{-1} \cdot \mathbf{s}_{b_k(i)}, \quad (13)$$

$$\bar{\mathbf{m}}_{bk} = -\bar{\mathbf{S}}_D^H \cdot \bar{\mathbf{m}}_{fk}, \quad (14)$$

where

$$\begin{aligned} \bar{\mathbf{R}}_U &= \bar{\mathbf{S}}_U \bar{\mathbf{S}}_U^H, \\ \bar{\mathbf{R}}_D &= \bar{\mathbf{S}}_D \left[\mathbf{I} - \text{diag}(\hat{\mathbf{b}}_D \hat{\mathbf{b}}_D^H) \right] \bar{\mathbf{S}}_D^H, \end{aligned} \quad (15)$$

and $\mathbf{s}_{b_k(i)}$ is a one column of $\bar{\mathbf{S}}_U$, whose column index is the same as the row index of $b_k(i)$ in \mathbf{b}_U . The feedforward filter in (13) is actually a linear MMSE filter which suppresses the interference from non-feedback symbols, as well as the residual interference after canceling the feedback symbols and the background Gaussian noise.

From (15), we can see that the optimum feedforward and feedback filters require the knowledge of all users' signature vectors and delays. In order to avoid the need for this information, we can adaptively implement the SISO PDFD, which will be discussed in the next section.

4. ADAPTIVE SISO PDFD FOR ASYNCHRONOUS DS-CDMA SYSTEMS

In this section, we assume that both short spreading codes and delays of all users are unknown to the receiver. We design and employ adaptive SISO PDFDs to track these parameters from the received signal directly.

It is well known that the asynchronous system performance can be improved by using detection filters with an increased number of taps. However, increasing the number of taps increases the computational complexity of the detector. Moreover, this will have an adverse effect on the convergence speed. Therefore, we need to select suitable filter lengths to achieve a good tradeoff among the system performance, detector complexity, and system overhead.

In the parallel decision-feedback detector, the feedforward and feedback filters cooperate to suppress the multiple-access interference. Specifically, the feedback filter tries to cancel some interfering symbols, while the feedforward filter

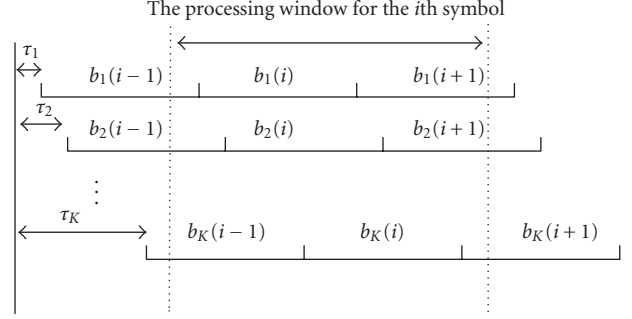


FIGURE 3: An asynchronous system.

suppresses the remaining MAI, as well as the residual interference due to imperfect cancellation by the feedback filter and the background Gaussian noise. Therefore, if the feedback filter effectively cancels most of the interference caused by the interfering symbols, the remaining interference to be suppressed by the feedforward filter is reduced.

On each iteration except for the first one, the SISO PDFD can obtain soft symbol estimates of all symbols from soft inputs. Thus, we have both causal and noncausal soft symbol decisions of interfering symbols for the interested symbol. We may cancel part or all of them by the feedback filter.

In this paper, we employ a feedforward filter which covers a two-symbol duration and consider several options for the feedback filter length. The length of the observation interval is $2T$, which is the minimum length such that one complete symbol of each user falls in this interval regardless of its relative delay. Figure 3 shows the processing window of the detector in the i th signaling interval. The output vector $\mathbf{r}(i)$ of the chip-matched filter in this processing window is

$$\mathbf{r}(i) = [\mathbf{P}^- \quad \mathbf{P}^0 \quad \mathbf{P}^+] \begin{bmatrix} \mathbf{b}(i-1) \\ \mathbf{b}(i) \\ \mathbf{b}(i+1) \end{bmatrix} + \mathbf{n}(i), \quad (16)$$

where $\mathbf{b}(i) = [b_1(i) \ b_2(i) \ \dots \ b_K(i)]^T$ and $\mathbf{n}(i)$ is a Gaussian random vector with zero mean and covariance matrix $\sigma_n^2 \mathbf{I}_{(2N \times 2N)}$. We define the punctured signature vectors of user k as

$$\begin{aligned} \mathbf{p}_k^- &= \left[(\mathbf{s}_k^r)^H \quad \mathbf{0}^H \right]_{(2N \times 1)}^H, \\ \mathbf{p}_k^0 &= \left[\mathbf{0}_{(1 \times N_k)}^H \quad \mathbf{s}_k^H \quad \mathbf{0}_{(1 \times N_k)}^H \right]_{(2N \times 1)}^H, \\ \mathbf{p}_k^+ &= \left[\mathbf{0}^H \quad (\mathbf{s}_k^l)^H \right]_{(2N \times 1)}^H, \end{aligned} \quad (17)$$

where $\mathbf{0}$ is a column vector. \mathbf{s}_k^l and \mathbf{s}_k^r are denoted in Figure 4 and are parts of \mathbf{s}_k :

$$\mathbf{s}_k = \left[(\mathbf{s}_k^l)^H \quad (\mathbf{s}_k^r)^H \right]^H. \quad (18)$$

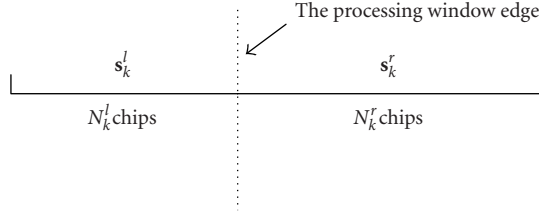


FIGURE 4: Punctured signatures of the k th user in the asynchronous system.

The matrices \mathbf{P}^- , \mathbf{P}^0 , and \mathbf{P}^+ in (16) are constructed as follows:

$$\begin{aligned} \mathbf{P}^- &= [\mathbf{p}_1^- \ \mathbf{p}_2^- \ \cdots \ \mathbf{p}_K^-], \\ \mathbf{P}^0 &= [\mathbf{p}_1^0 \ \mathbf{p}_2^0 \ \cdots \ \mathbf{p}_K^0], \\ \mathbf{P}^+ &= [\mathbf{p}_1^+ \ \mathbf{p}_2^+ \ \cdots \ \mathbf{p}_K^+]. \end{aligned} \quad (19)$$

Thus, when multiple users' delays are unknown to the receiver, for the symbol of interest $b_k(i)$ of user k , it has at most $(3K - 1)$ interfering symbols. For implementation of the adaptive SISO multiuser detector in Figure 1, we consider three adaptive SISO PDFDs with the same feedforward filter length, that is, $2N$ taps. The feedback filter of the first detector (labeled as detector1) has $(K - 1)$ taps which tries to cancel the current $(K - 1)$ interfering symbols for the desired symbol. Detector2 has a feedback filter with $(2K - 1)$ taps which tries to cancel the current $(K - 1)$ and previous K interfering symbols. The feedback filter of detector3 has $(3K - 1)$ taps and tries to cancel all possible previous, current, and future interfering symbols.

In the following, we employ the NLMS and RLS algorithms in adaptive SISO PDFDs to update the feedforward filter $\bar{\mathbf{m}}_{fk}$ and feedback filter $\bar{\mathbf{m}}_{bk}$. Moreover, the a priori information of coded symbols is employed efficiently to improve the performance of the adaptive detector. The adaptive SISO PDFD requires only a training sequence for each user to estimate all filter coefficients.

The adaptive detector employing the NLMS algorithm to resolve the MMSE criterion in (10) updates the feedforward and feedback filters of user k as follows for $m = 0, 1, 2, \dots$:

$$\begin{aligned} \bar{\mathbf{m}}_{fk}(m+1) &= \bar{\mathbf{m}}_{fk}(m) - \frac{\tilde{\mu}_f}{a + \|\bar{\mathbf{r}}(m)\|^2} \tilde{b}_k(m) e_k^*(m) \bar{\mathbf{r}}(m), \\ \bar{\mathbf{m}}_{bk}(m+1) &= \bar{\mathbf{m}}_{bk}(m) - \frac{\tilde{\mu}_b}{a + \|\tilde{\mathbf{b}}_D(m)\|^2} \tilde{b}_k(m) e_k^*(m) \tilde{\mathbf{b}}_D(m), \end{aligned} \quad (20)$$

where m is the recursive index and also the time index, $\tilde{\mu}_f$ and $\tilde{\mu}_b \in (0, 2)$ and are step sizes for the feedforward and feedback filters, respectively. a is a small positive constant. The error signal for the m th recursion is

$$e_k(m) = \tilde{b}_k(m) - \bar{\mathbf{m}}_{fk}^H(m) \cdot \bar{\mathbf{r}}(m) - \bar{\mathbf{m}}_{bk}^H(m) \cdot \tilde{\mathbf{b}}_D(m), \quad (21)$$

where $\tilde{b}_k(m) = b_k(m)$ and $\tilde{\mathbf{b}}_D(m) = \mathbf{b}_D(m)$ in the training mode, $\tilde{b}_k(m) = \hat{b}_k(m)$ and $\tilde{\mathbf{b}}_D(m) = \hat{\mathbf{b}}_D(m)$ in the decision-

directed mode. Furthermore, in the decision-directed mode, $|\hat{b}_k(m)|$ is used as the reliability of the error signal $e_k(m)$ in (20). Both filters are updated per symbol and their initial states are $\bar{\mathbf{m}}_{fk}(0) = \mathbf{0}$ and $\bar{\mathbf{m}}_{bk}(0) = \mathbf{0}$.

When the detector employs the RLS algorithm, we denote $\mathbf{w}_k(m) = [\bar{\mathbf{m}}_{fk}^H(m) \ \bar{\mathbf{m}}_{bk}^H(m)]^H$ and $\mathbf{u}(m) = [\bar{\mathbf{r}}^H(m) \ \tilde{\mathbf{b}}_D^H(m)]^H$. Then the filters are updated for $m = 0, 1, 2, \dots$:

$$\begin{aligned} \mathbf{g}_k(m+1) &= \frac{\lambda^{-1} \mathbf{P}_k(m) \mathbf{u}(m+1)}{1 + \lambda^{-1} \mathbf{u}^H(m+1) \mathbf{P}_k(m) \mathbf{u}(m+1)}, \\ \xi_k(m+1) &= \tilde{b}_k(m+1) - \mathbf{w}_k^H(m) \mathbf{u}(m+1), \end{aligned} \quad (22)$$

$$\mathbf{w}_k(m+1) = \mathbf{w}_k(m) + \mathbf{g}_k(m+1) \tilde{b}_k(m+1) \xi_k^*(m+1),$$

$$\mathbf{P}_k(m+1) = \lambda^{-1} \mathbf{P}_k(m) - \lambda^{-1} \mathbf{g}_k(m+1) \mathbf{u}^H(m+1) \mathbf{P}_k(m).$$

The algorithm is initialized with $\mathbf{P}_k(0) = \delta^{-1} \mathbf{I}$, where δ is a small positive number and $\mathbf{w}_k(0) = \mathbf{0}$.

Both of the adaptive detectors described above try to exploit the joint statistics of the received signal vector $\bar{\mathbf{r}}$, the transmitted symbol b_k or its soft estimate \hat{b}_k , and the soft symbol estimates $\hat{\mathbf{b}}_D$ which are feedback. In the first iteration, since there is no feedback information of coded symbols, we only employ a linear MMSE feedforward filter and set the feedback filter coefficients to zeros for each user.

The output of the adaptive SISO PDFD is

$$y_k(m) = \bar{\mathbf{m}}_{fk}^H(m) \cdot \bar{\mathbf{r}}(m) + \bar{\mathbf{m}}_{bk}^H(m) \cdot \hat{\mathbf{b}}_D(m). \quad (23)$$

Applying the Gaussian assumption to the output in (23), we can calculate the soft outputs of the SISO PDFD. For the m th symbol of the k th user, the output $y_k(m)$ can be expressed as

$$y_k(m) = \mu_k b_k(m) + \eta_k, \quad (24)$$

where μ_k is a constant and η_k is a Gaussian random variable with zero mean and variance $\sigma_{\eta_k}^2$:

$$\begin{aligned} \mu_k &= E[b_k^*(m) y_k(m)], \\ \sigma_{\eta_k}^2 &= E[y_k(m) - \mu_k b_k(m)]^2. \end{aligned} \quad (25)$$

Estimates of (25) can be obtained by the corresponding sample averages in (26), respectively, where we replace $b_k(m)$ by $\tilde{b}_k(m)$ in these equations:

$$\hat{\mu}_k = \frac{1}{N_b} \sum_{m=1}^{N_b} \tilde{b}_k^*(m) y_k(m), \quad (26)$$

$$\hat{\sigma}_{\eta_k}^2 = \frac{1}{N_b} \sum_{m=1}^{N_b} [y_k(m) - \hat{\mu}_k \tilde{b}_k(m)]^2.$$

The soft output, that is, the extrinsic log-likelihood ratio, of $b_k(m)$ is

$$\lambda_k^o(m) = \log \frac{P[y_k(m) | b_k(m) = +1]}{P[y_k(m) | b_k(m) = -1]} = \frac{2\mu_k y_k(m)}{\sigma_{\eta_k}^2}. \quad (27)$$

5. SIMULATION RESULTS

The DS-CDMA system which we simulate in this section has 12 active users. All users employ the same convolutional code with rate 1/2, constraint length 7, and generators [1011011], [1111001]. Each user has a randomly selected short spreading code. The spreading factor is 16 chips per information bit. The system load is 12/16 (K /spreading factor). Multiple users' delays are randomly selected and fixed during simulation.

There are 300 training symbols which are randomly selected and inserted at the beginning of coded symbol frames of each user. SISO single-user decoders are based on the log-MAP algorithm in [18]. Noise random variables at the output of the chip-matched filter are identical independent Gaussian random variables with zero mean and $N_0/2$ variance.

At the first iteration, since there are no soft inputs from single-user decoders, only a feedforward filter is employed for each user. That is, at this time, a linear minimum mean square error filter is used instead. It is initially trained by the training symbols, and then is used for the following transmitted coded symbols. For the later iterations, both the feedforward and feedback filters are employed. After the training mode, they are updated by feedback symbol decisions. In the first two iterations, the filter coefficients are initialized to zeros before the adaptive algorithm is employed. In each of the following iterations, the filter coefficients are set to the values obtained at the end of the previous iteration.

We consider an asynchronous DS-CDMA system over the additive white Gaussian noise (AWGN) channel. It is assumed that the receiver has no knowledge of the short spreading codes used by the users and their delays. Three adaptive SISO PDFDs proposed in Section 4 are simulated. Figures 5 and 6 show average bit error rates of all users in the first, second, and tenth iterations provided by three adaptive detectors based on the NLMS and RLS algorithms, respectively. In (20) of the NLMS algorithm, we use $a = 0.00001$, and step sizes $\bar{\mu}_f = \bar{\mu}_b = 0.2$ in the training mode and $\bar{\mu}_f = \bar{\mu}_b = 0.05$ in the decision-directed mode. Parameters in (22) of the RLS algorithm are $\lambda = 1$ and $\delta = 0.04$. For comparison, we also show the bit error rate performance of the single-user system in these two figures, where the user's spreading code and delay are known to the receiver. In Figures 5 and 6, we observe that after the first iteration, all three detectors have similar performances and their curves appear to overlap. A similar behaviour is observed for the second iteration of detector1 and detector2 in Figure 5 and all three detectors in Figure 6.

We can see that with our adaptive SISO detectors, the system performance is improved with the increased number of iterations. Furthermore, Figure 6 shows that the performance provided by the adaptive RLS receiver approaches the performance of the single-user system after a few iterations at high signal-to-noise ratios. Among the three adaptive SISO PDFDs proposed in Section 4, detector3 provides the best performance, though it has the highest computational

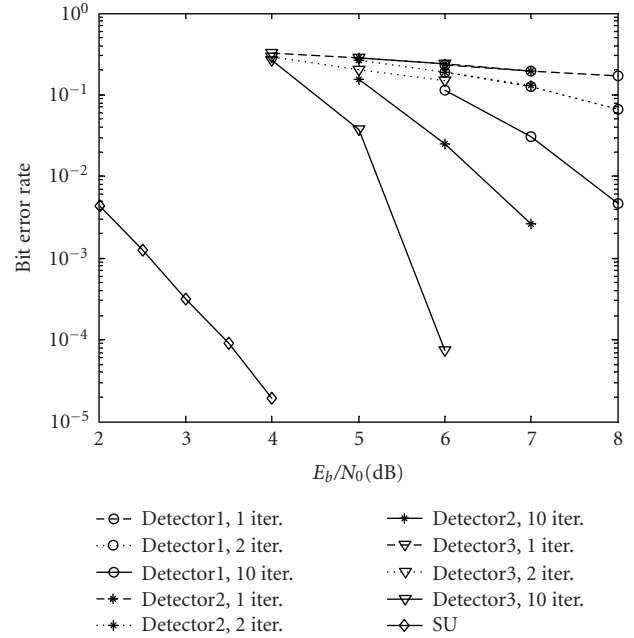


FIGURE 5: Bit error rate performance provided by three NLMS adaptive SISO PDFDs for the asynchronous DS-CDMA system at the first, second, and tenth iterations, and that of the single-user system (SU).

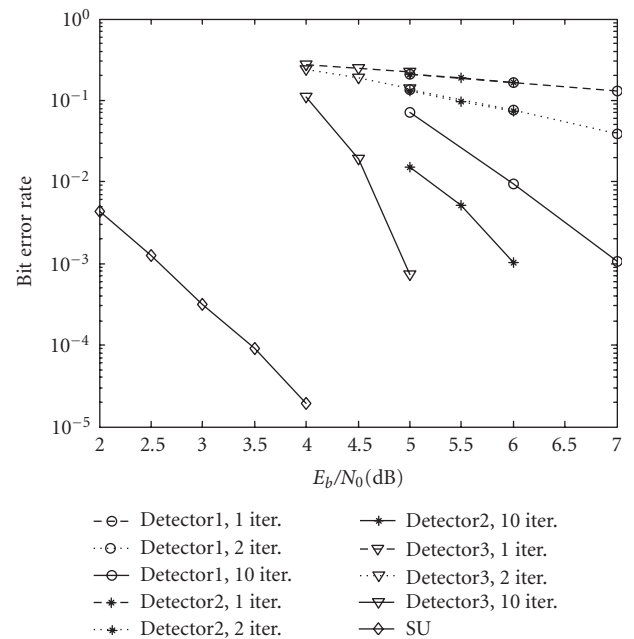


FIGURE 6: Bit error rate performance provided by three RLS adaptive SISO PDFDs for the asynchronous DS-CDMA system at the first, second, and tenth iterations, and that of the single-user system (SU).

complexity, since its feedback filter has the maximum number of taps compared with the other two detectors.

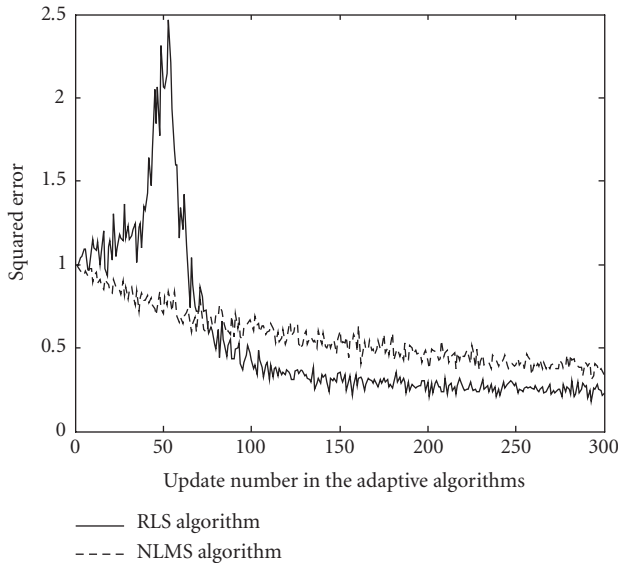


FIGURE 7: Comparison between the experimental learning curves of the adaptive SISO PDFD detector3 based on the NLMS and RLS algorithms after the second iteration during the training mode at SNR = 6 dB.

By comparing average bit error rates of all the users provided by the adaptive detector based on the RLS algorithm in Figure 6 and those obtained by the NLMS algorithm in Figure 5, we can see that the bit error rate performance provided by the adaptive SISO PDFD based on the RLS algorithm is better than the one provided by the detector based on the NLMS algorithm. For example, at a bit error rate 10^{-3} , detector3 based on the RLS algorithm has about 0.7 dB gain with respect to detector3 based on the NLMS algorithm. This is due to the faster convergence property of the RLS algorithm, which is shown by Figure 7. The averaged squared errors $e_k^2(m)$ and $\xi_k^2(m)$ after the second iteration of the adaptive detector3 during the training mode versus the number of updates in the NLMS and RLS algorithms, respectively, are shown and compared in Figure 7. We set the signal-to-noise (SNR) ratio of each user to 6 dB. Each curve of the squared error is averaged over 200 independent trials of the experiment. However, the RLS algorithm has a greater computational complexity. Denote the length of the adaptive filter as L . The computational complexity of the RLS and the NLMS algorithms are $\sim O(L^2)$ and $\sim O(L)$ per update, respectively.

6. CONCLUSIONS

In this paper, first we presented an optimum SISO parallel decision-feedback detector for asynchronous coded DS-CDMA systems, and then proposed an adaptive implementation of it when all users' signature waveforms and relative delays were unknown to the receiver. All users were assumed to employ short spreading codes. A chip-synchronous and symbol-asynchronous DS-CDMA system was considered.

A training sequence was required by each user. We showed that the resulting system performance provided by adaptive SISO PDFDs approaches that of the single-user system after a few iterations at high signal-to-noise ratios. Moreover, the adaptive detector employing the RLS algorithm provides a better bit error rate performance than the adaptive detector based on the NLMS algorithm, though at the expense of higher computational complexity. For asynchronous coded DS-CDMA systems, we further showed that the adaptive detector with more feedback filter taps gives a better bit error rate performance.

REFERENCES

- [1] M. Moher, "An iterative multiuser decoder for near-capacity communications," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 870–880, 1998.
- [2] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (Corresp.)," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, 1974.
- [3] P. D. Alexander, M. C. Reed, J. A. Asenstorfer, and C. B. Schlegel, "Iterative multiuser interference reduction: turbo CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1008–1014, 1999.
- [4] Z. Qin, K. C. Teh, and E. Gunawan, "Iterative multiuser detection for asynchronous CDMA with concatenated convolutional coding," *IEEE J. Select. Areas Commun.*, vol. 19, no. 9, pp. 1784–1792, 2001.
- [5] P. Alexander, A. Grant, and M. Reed, "Iterative detection in code-division multiple-access with error control coding," *European Transactions on Telecommunication*, vol. 9, no. 5, pp. 419–426, 1998.
- [6] Z. Shi and C. Schlegel, "Joint iterative decoding of serially concatenated error control coded CDMA," *IEEE J. Select. Areas Commun.*, vol. 19, no. 8, pp. 1646–1653, 2001.
- [7] W. Zhang and C. D'Amours, "Iterative multiuser detection and decoding for highly correlated narrowband systems and heavily loaded CDMA systems," *Canadian Journal of Electrical and Computer Engineering*, vol. 28, no. 2, pp. 75–80, 2003.
- [8] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046–1061, 1999.
- [9] H. E. Gamal and E. Geraniotis, "Iterative multiuser detection for coded CDMA signals in AWGN and fading channels," *IEEE J. Select. Areas Commun.*, vol. 18, no. 1, pp. 30–41, 2000.
- [10] B. F. Beidas, H. E. Gamal, and S. Kay, "Iterative interference cancellation for high spectral efficiency satellite communications," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 31–36, 2002.
- [11] M. L. Honig and H. V. Poor, "Adaptive interference suppression in wireless communication systems," in *Wireless Communications: Signal Processing Perspectives*, Prentice-Hall, Upper Saddle River, NJ, USA, 1998, chapter 2.
- [12] H. V. Poor, "Adaptivity in multiple-access communications," in *Proc. 34th Conference on Decision and Control*, pp. 835–840, New Orleans, La, USA, December 1995.
- [13] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 12, no. 4, pp. 685–697, 1994.
- [14] P. B. Rapajic and D. K. Borah, "Adaptive MMSE maximum likelihood CDMA multiuser detection," *IEEE J. Select. Areas Commun.*, vol. 17, no. 12, pp. 2110–2122, 1999.
- [15] D. K. Borah and P. B. Rapajic, "Optimal adaptive multiuser detection in unknown multipath channels," *IEEE J. Select. Areas Commun.*, vol. 19, no. 6, pp. 1115–1127, 2001.

- [16] M. L. Honig, G. Woodward, and P. D. Alexander, "Adaptive multiuser parallel-decision-feedback with iterative decoding," in *Proc. IEEE International Symposium on Information Theory (ISIT '00)*, p. 335, Sorrento, Italy, June 2000.
- [17] A. Duel-Hallen, "A family of multiuser decision-feedback detectors for asynchronous code-division multiple-access channels," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 421–434, 1995.
- [18] A. J. Viterbi, "An intuitive justification and a simplified implementation of the MAP decoder for convolutional codes," *IEEE J. Select. Areas Commun.*, vol. 16, no. 2, pp. 260–264, 1998.

Wei Zhang received the B.A.Sc. degree from XiDian University, China, in 1995, and the M.A.Sc. degree from Beijing University of Posts and Telecommunications, China, in 1998. Now she is pursuing a Ph.D. degree at the University of Ottawa, Canada. All of these are in electrical engineering. She worked as a software engineer in CTC Communication Development Ltd, China, in 1998. In 1999, she joined Agilent Technologies, Beijing, China, as a Research Scientist. Her research interest is in signal processing for the physical layer of wireless communications.



Claude D'Amours graduated with the degrees of B.A.Sc., M.A.Sc., and Ph.D. in electrical engineering from the University of Ottawa in 1990, 1992, and 1995, respectively. He was employed briefly at the Communications Research Centre in Ottawa as a Systems Engineer in 1995. From 1995–1999, he was employed as an Assistant Professor in the Department of Electrical and Computer Engineering, the Royal Military College in Kingston, Ontario, Canada. He is presently employed as an Assistant Professor in the School of Information Technology and Engineering, the University of Ottawa.



Abbas Yongaçoğlu received the B.S. degree from Boğaziçi University, Turkey, in 1973, the M. Eng. degree from the University of Toronto, Canada, in 1975, and the Ph.D. degree from the University of Ottawa, Canada, in 1987, all in electrical engineering. He worked as a researcher and a System Engineer at TUBITAK Marmara Research Institute, Turkey, Philips Research Labs, Holland, and Miller Communications Systems, Ottawa. In 1987, he joined the University of Ottawa as an Assistant Professor. He became an Associate Professor in 1992 and a Full Professor in 1996. His area of research is digital communications with emphasis on modulation, coding, equalization, and multiple access for wireless and high-speed wireline communications.



A Low-Complexity KL Expansion-Based Channel Estimator for OFDM Systems

Habib Şenol

Department of Computer Engineering, Kadir Has University, Cibali 34230, Istanbul, Turkey
Email: hsenol@khas.edu.tr

Hakan A. Çırpan

Department of Electrical-Electronics Engineering, Istanbul University, Avcılar 34850, Istanbul, Turkey
Email: hcirpan@istanbul.edu.tr

Erdal Panayırıcı

Department of Electronics Engineering, Işık University, Maslak 80670, Istanbul, Turkey
Email: eepanay@isikun.edu.tr

Received 23 April 2004; Revised 18 October 2004

This paper first proposes a computationally efficient, pilot-aided linear minimum mean square error (MMSE) batch channel estimation algorithm for OFDM systems in unknown wireless fading channels. The proposed approach employs a convenient representation of the discrete multipath fading channel based on the Karhunen-Loeve (KL) orthogonal expansion and finds MMSE estimates of the uncorrelated KL series expansion coefficients. Based on such an expansion, no matrix inversion is required in the proposed MMSE estimator. Moreover, optimal rank reduction is achieved by exploiting the optimal truncation property of the KL expansion resulting in a smaller computational load on the estimation algorithm. The performance of the proposed approach is studied through analytical and experimental results. We then consider the stochastic Cramér-Rao bound and derive the closed-form expression for the random KL coefficients and consequently exploit the performance of the MMSE channel estimator based on the evaluation of minimum Bayesian MSE. We also analyze the effect of a modelling mismatch on the estimator performance. To further reduce the complexity, we extend the batch linear MMSE to the sequential linear MMSE estimator. With the fast convergence property and the simple structure, the sequential linear MMSE estimator provides an attractive alternative to the implementation of channel estimator.

Keywords and phrases: channel estimation, OFDM systems, MMSE estimation.

1. INTRODUCTION

With unprecedented demands on bandwidth due to the explosive growth of broadband wireless services usage, there is an acute need for a high-rate and bandwidth-efficient digital transmission. In response to this need, the research community has been extensively investigating efficient schemes that make efficient utilization of the limited bandwidth and cope with the adverse access environments [1]. These access environments may create different channel impairments and dictate unique sets of advanced signal processing algorithms to combat specific impairments.

Multicarrier (MC) transmission scheme, especially orthogonal frequency-division multiplexing (OFDM), has recently attracted considerable attention since it has been shown to be an effective technique to combat delay spread or frequency-selective fading of wireless or wireline channels, thereby improving the capacity and enhancing the performance of transmission. This approach has been adopted as the standard in several outdoor and indoor high-speed wireless and wireline data applications, including terrestrial digital broadcasting (DAB and DVB) in Europe, and high-speed modems over digital subscriber lines in the US. It has also been implemented for broadband indoor wireless systems including IEEE802.11a, MMAC, and HIPERLAN/2.

An OFDM system operating over a frequency-selective wireless communication channel effectively forms a number of parallel frequency-nonselective fading channels, thereby

This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

reducing intersymbol interference (ISI) and obviating the need for complex equalization, thus greatly simplifying channel estimation/equalization task. Moreover, OFDM is bandwidth efficient since the spectra of the neighboring subchannels overlap, yet channels can still be separated through the use of orthogonality of the carriers. Furthermore, its structure also allows efficient hardware implementations using fast Fourier transform (FFT) and polyphase filtering [2].

Although the structure of OFDM signalling avoids ISI arising due to channel memory, fading multipath channel still introduces random attenuations on each tone. Furthermore, simple frequency-domain equalization, which divides the FFT output by the corresponding channel frequency response, does not assure symbol recovery if the channel has nulls on some subcarriers. Hence, advanced signal processing algorithms have to be used for accurate channel estimation to improve the performance of the OFDM systems. Numerous pilot-aided channel estimation methods for OFDM have been developed [3, 4, 5, 6]. In particular, a low-rank approximation is applied to linear MMSE estimator for the estimation of subcarrier channel attenuations by using the frequency correlation of the channel [3]. Two pilot-aided MLE and MMSE schemes are revisited and compared in terms of computational complexity in [4]. In [5], an MMSE channel estimator, which makes full use of the time and frequency correlation of the time-varying dispersive channel, was proposed. Moreover, low-complexity MMSE doubly channel estimation approaches were presented in [6] based on embedding Kronecker-delta pilot sequences.

Multipath fading channels have been studied extensively, and several models have been developed to describe their variations [7]. In many cases, the channel taps are modelled as general lowpass stochastic processes (e.g., [8]), the statistics depend on mobility parameters. A different approach explicitly models the multipath channel taps by the Karhunen-Loeve (KL) series representation [9]. KL expansion models have also been used previously in modelling the multipath channel within a CDMA scenario [10]. In the case of KL series representation of stochastic process, a convenient choice of orthogonal basis set is one that makes the expansion coefficient random variables uncorrelated. When these orthogonal bases are employed to expand the channel taps of the multipath channel, uncorrelated coefficients indeed represent the multipath channel. Therefore, KL representation allows one to tackle the estimation of correlated multipath parameters as a parameter estimation problem of the uncorrelated coefficients. Exploiting KL expansion, the main contribution of this paper is to propose a computationally efficient, pilot-aided MMSE channel estimation algorithm. Based on such representation, no matrix inversion is required in the proposed batch approach. Moreover, optimal rank reduction is achieved by exploiting the optimal truncation property of the KL expansion resulting in a smaller computational load on the estimation algorithm. The performance of the proposed batch approach is explored based on the evaluation of the stochastic Cramér-Rao bound for the random KL coefficients. We also analyze the effect of a modelling mismatch on the estimator performance. In contrast to [3],

the proposed batch approach employs KL expansion of multipath channel parameters and reduces the complexity of the singular value decomposition (SVD) used in *eigendecomposition* by estimating multipath channel parameters instead of channel attenuations on each tone. In addition, we propose the simple sequential MMSE implementation for the estimation of the KL expansion coefficients, which does not require to perform matrix inversion as well.

The rest of the paper is organized as follows. Section 2 describes a general model for OFDM systems and briefly introduces the channel estimation task. Section 3 derives a basic and simplified approach to MMSE batch channel estimation for OFDM systems. To show its efficiency, the performance bounds are analyzed and the performance degradation due to a mismatch of the estimator to the channel statistics as well as the SNR is demonstrated. The sequential MMSE estimator is introduced in Section 4 and its convergence behavior is also analyzed. Some simulation examples are provided in Section 5. Finally, conclusions are drawn in Section 6.

2. SYSTEM MODEL

In order to eliminate ISI arising due to multipath channel and preserve orthogonality of the subcarrier frequencies (tones), conventional OFDM systems first take the IFFT of data symbols and then insert redundancy in the form of a cyclic prefix (CP) of length L_{CP} larger than the channel order L . CP is discarded at the receiver and the remaining part of the OFDM symbol is FFT processed. Combination of IFFT and CP at the transmitter with the FFT at the receiver divides the frequency-selective channel into several separate flat-fading subchannels. The block diagram in Figure 1 describes the conventional OFDM system. We consider an OFDM system with K subcarriers for the transmission of K parallel data symbols. Thus, the information stream $X(n)$ is parsed into K -long blocks: $\mathbf{X}_i = [X_i(0), X_i(1), \dots, X_i(K-1)]^T$, where $i = 1, 2, \dots$ is the block index and the superscript $(\cdot)^T$ indicates the vector transpose. The $K \times 1$ symbol block is then mapped to a $(K+L) \times 1$ vector by first taking the IFFT of \mathbf{X}_i and then replicating the last L_{CP} elements as

$$\mathbf{s}_i = [s_i(0), s_i(1), \dots, s_i(K+L_{CP}-1)]^T. \quad (1)$$

\mathbf{s}_i is serially transmitted over the channel. At the receiver, the CP of length L_{CP} is removed first and FFT is performed on the remaining $K \times 1$ vector. Therefore, we can write the output of the FFT unit in matrix form as

$$\mathbf{Y}_i = \mathbf{A}_i \mathbf{H}_i + \boldsymbol{\eta}_i, \quad (2)$$

where \mathbf{A}_i is the diagonal matrix $\mathbf{A}_i = \text{diag}(\mathbf{X}_i)$ and \mathbf{H}_i is the channel vector. The elements of \mathbf{H}_i are the values of the channel frequency response evaluated at the subcarriers. Therefore, we can write $\mathbf{H}_i = [H_i(0), H_i \exp(j2\pi/K), \dots, H_i \exp(j2\pi(K-1)/K)]^T$ as $\mathbf{H}_i = \mathcal{F} \mathbf{h}_i$, where \mathcal{F} is the FFT matrix with (m, n) entry $\exp(-j2\pi mn/K)$ and $\mathbf{h}_i = [h_i(0), h_i(1), \dots, h_i(L-1)]^T$. \mathbf{h}_i modelled as a complex Gaussian vector with $\mathbf{h}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{h_i})$ represents the overall

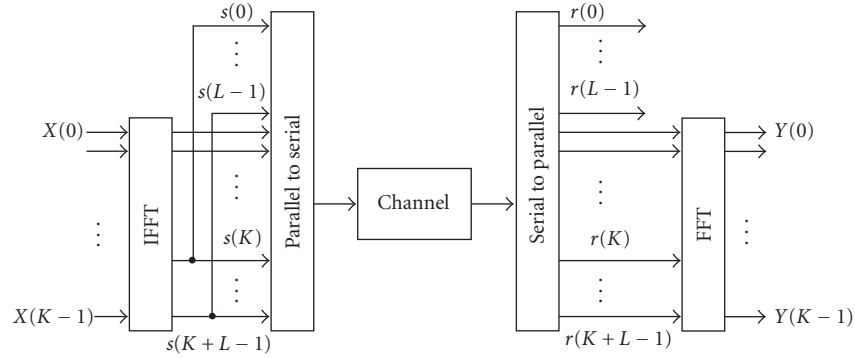


FIGURE 1: OFDM system block diagram.

channel impulse during the i th OFDM block. Finally, $\boldsymbol{\eta}_i$ is a $K \times 1$ zero-mean, i.i.d. complex Gaussian vector that models additive noise in the K subchannels (tones). We have $E[\boldsymbol{\eta}_i \boldsymbol{\eta}_i^\dagger] = \sigma^2 \mathbf{I}_K$ where \mathbf{I}_K represents a $K \times K$ identity matrix, σ^2 is the variance of the additive noise entering the system, and the superscript $(\cdot)^\dagger$ indicates the Hermitian transpose.

Based on model (2), our main objective in this paper is to develop both batch and sequential pilot-aided channel estimation algorithm according to MMSE criterion and then explore the performance of the estimators. A batch approach adapted herein explicitly models the channel parameters by the KL series representation and estimates the uncorrelated expansion coefficients. Furthermore, the computational load of the proposed MMSE estimation technique is further reduced with the application of the KL expansion optimal truncation property [9]. We then introduce batch channel MMSE approach first.

3. MMSE ESTIMATION OF KL COEFFICIENTS: BATCH APPROACH

A low-rank approximation to the frequency-domain linear MMSE channel estimator is provided by [3] to reduce the complexity of the estimator. Optimal rank reduction is achieved in this approach by using the SVD of the channel attenuations covariance matrix \mathbf{C}_H of dimension $K \times K$. In contrast, we adopt the MMSE estimator for the estimation of multipath channel parameter \mathbf{h} that uses covariance matrix of dimension $L \times L$. The proposed approach employs KL expansion of multipath channel parameters and reduces the complexity of the SVD used in *eigendecomposition* since L is usually much less than K . We will now develop MMSE batch estimator for pilot-assisted OFDM system in the sequel.

3.1. MMSE channel estimation

Pilot-symbol-assisted techniques can provide information about an undersampled version of the channel that may be easier to identify. In this paper, we therefore address the problem of estimating multipath channel parameters by exploiting the distributed training symbols. Considering (2), and in order for the pilot symbols to be included in the output vector for our estimation purposes, we focus on an

undersampled signal model. Assuming that K_p pilot symbols are uniformly inserted at known locations of the i th OFDM block, the $K_p \times 1$ vector corresponding to the FFT output at the pilot locations becomes

$$\mathbf{Y} = \mathbf{A}\mathbf{F}\mathbf{h} + \boldsymbol{\eta}, \quad (3)$$

where $\mathbf{A} = \text{diag}[\mathbf{A}_i(0), \mathbf{A}_i(\Delta), \dots, \mathbf{A}_i((K_p - 1)\Delta)]$ is a diagonal matrix with pilot-symbol entries, Δ is pilot spacing interval, \mathbf{F} is a $K_p \times L$ FFT matrix generated based on pilot indices, and similarly $\boldsymbol{\eta}$ is the undersampled noise vector.

For the estimation of \mathbf{h} , the new linear signal model can be formed by premultiplying both sides of (3) by \mathbf{A}^\dagger and assuming that pilot symbols are taken from a PSK constellation $\mathbf{A}^\dagger \mathbf{A} = \mathbf{I}_{K_p}$, then the new form of (3) becomes

$$\begin{aligned} \mathbf{A}^\dagger \mathbf{Y} &= \mathbf{F}\mathbf{h} + \mathbf{A}^\dagger \boldsymbol{\eta}, \\ \tilde{\mathbf{Y}} &= \mathbf{F}\mathbf{h} + \tilde{\boldsymbol{\eta}}, \end{aligned} \quad (4)$$

where $\tilde{\mathbf{Y}}$ and $\tilde{\boldsymbol{\eta}}$ are related to \mathbf{Y} and $\boldsymbol{\eta}$ by the linear transformation, respectively. Furthermore, $\tilde{\boldsymbol{\eta}}$ is statistically equivalent to $\boldsymbol{\eta}$.

Equation (4) offers a Bayesian linear model representation. Based on this representation, the minimum variance estimator for the time-domain channel vector \mathbf{h} for the i th OFDM block, that is, conditional mean of \mathbf{h} given $\tilde{\mathbf{Y}}$, can be obtained using MMSE estimator. We should clearly make the assumptions that $\mathbf{h} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_h)$, $\tilde{\boldsymbol{\eta}} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\tilde{\boldsymbol{\eta}}})$, and \mathbf{h} is uncorrelated with $\tilde{\boldsymbol{\eta}}$. Therefore, MMSE estimate of \mathbf{h} is given by [11]

$$\hat{\mathbf{h}} = (\mathbf{F}^\dagger \mathbf{C}_{\tilde{\boldsymbol{\eta}}}^{-1} \mathbf{F} + \mathbf{C}_h^{-1})^{-1} \mathbf{F}^\dagger \mathbf{C}_{\tilde{\boldsymbol{\eta}}}^{-1} \tilde{\mathbf{Y}}. \quad (5)$$

Due to PSK pilot-symbol assumption together with the result $\mathbf{C}_{\tilde{\boldsymbol{\eta}}} = E[\tilde{\boldsymbol{\eta}} \tilde{\boldsymbol{\eta}}^\dagger] = \sigma^2 \mathbf{I}_{K_p}$, we can therefore express (5) by

$$\hat{\mathbf{h}} = (\mathbf{F}^\dagger \mathbf{F} + \sigma^2 \mathbf{C}_h^{-1})^{-1} \mathbf{F}^\dagger \tilde{\mathbf{Y}}. \quad (6)$$

Under the assumption that uniformly spaced pilot symbols are inserted with pilot spacing interval Δ and $K = \Delta \times K_p$, correspondingly, $\mathbf{F}^\dagger \mathbf{F}$ reduces to $\mathbf{F}^\dagger \mathbf{F} = K_p \mathbf{I}_L$. Then according to (6) and $\mathbf{F}^\dagger \mathbf{F} = K_p \mathbf{I}_L$, we arrive at the expression

$$\hat{\mathbf{h}} = (K_p \mathbf{I}_L + \sigma^2 \mathbf{C}_h^{-1})^{-1} \mathbf{F}^\dagger \tilde{\mathbf{Y}}. \quad (7)$$

Since MMSE estimation still requires the inversion of \mathbf{C}_h^{-1} , it therefore suffers from a high computational complexity. However, it is possible to reduce the complexity of the MMSE algorithm by diagonalizing channel covariance matrix with a KL expansion.

3.2. KL expansion

Channel impulse response \mathbf{h} is a zero-mean Gaussian process with covariance matrix \mathbf{C}_h . The KL transformation is therefore employed here to rotate the vector \mathbf{h} so that all its components are uncorrelated. The vector \mathbf{h} , representing the channel impulse response during i th OFDM block, can be expressed as a linear combination of the orthonormal basis vectors as follows:

$$\mathbf{h} = \sum_{l=0}^{L-1} g_l \boldsymbol{\psi}_l = \boldsymbol{\Psi} \mathbf{g}, \quad (8)$$

where $\boldsymbol{\Psi} = [\boldsymbol{\psi}_0, \boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_{L-1}]$, $\boldsymbol{\psi}_i$'s are the orthonormal basis vectors, $\mathbf{g} = [g_0, g_1, \dots, g_{L-1}]^T$, and g_i 's are the weights of the expansion. If we form the covariance matrix \mathbf{C}_h as

$$\mathbf{C}_h = \boldsymbol{\Psi} \boldsymbol{\Lambda}_g \boldsymbol{\Psi}^\dagger, \quad (9)$$

where $\boldsymbol{\Lambda}_g = E\{\mathbf{g}\mathbf{g}^\dagger\}$, the KL expansion is the one in which $\boldsymbol{\Lambda}_g$ of \mathbf{C}_h is a diagonal matrix (i.e., the coefficients are uncorrelated). If $\boldsymbol{\Lambda}_g$ is diagonal, then the form $\boldsymbol{\Psi} \boldsymbol{\Lambda}_g \boldsymbol{\Psi}^\dagger$ is called an *eigendecomposition* of \mathbf{C}_h . The fact that only the eigenvectors diagonalize \mathbf{C}_h leads to the desirable property that the KL coefficients are uncorrelated. Furthermore, in the Gaussian case, the uncorrelatedness of the coefficients renders them independent as well, providing additional simplicity.

Thus, the channel estimation problem in this application is equivalent to estimating the i.i.d. complex Gaussian vector \mathbf{g} KL expansion coefficients.

3.3. Estimation of KL coefficients

In contrast to (4) in which only \mathbf{h} is to be estimated, we now assume that the KL coefficient vector \mathbf{g} is unknown. Thus the data model (4) is rewritten for each OFDM block as

$$\tilde{\mathbf{Y}} = \mathbf{F} \boldsymbol{\Psi} \mathbf{g} + \tilde{\boldsymbol{\eta}} \quad (10)$$

which is also recognized as a Bayesian linear model, and recall that $\mathbf{g} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}_g)$. As a result, the MMSE estimator of \mathbf{g} is

$$\begin{aligned} \hat{\mathbf{g}} &= \boldsymbol{\Lambda}_g (K_p \boldsymbol{\Lambda}_g + \sigma^2 \mathbf{I}_L)^{-1} \boldsymbol{\Psi}^\dagger \mathbf{F}^\dagger \tilde{\mathbf{Y}} \\ &= \boldsymbol{\Gamma} \boldsymbol{\Psi}^\dagger \mathbf{F}^\dagger \tilde{\mathbf{Y}}, \end{aligned} \quad (11)$$

where

$$\begin{aligned} \boldsymbol{\Gamma} &= \boldsymbol{\Lambda}_g (K_p \boldsymbol{\Lambda}_g + \sigma^2 \mathbf{I}_L)^{-1} \\ &= \text{diag} \left\{ \frac{\lambda_{g_0}}{\lambda_{g_0} K_p + \sigma^2}, \dots, \frac{\lambda_{g_{L-1}}}{\lambda_{g_{L-1}} K_p + \sigma^2} \right\} \end{aligned} \quad (12)$$

and $\lambda_{g_0}, \lambda_{g_1}, \dots, \lambda_{g_{L-1}}$ are the singular values of $\boldsymbol{\Lambda}_g$.

It is clear that the complexity of the MMSE estimator in (7) is reduced by the application of KL expansion. However, the complexity of $\hat{\mathbf{g}}$ can be further reduced by exploiting the optimal truncation property of the KL expansion [9]. MMSE estimator of \mathbf{g} requires $4L^2 + 4LK_p + 2L$ real multiplications. From the results presented in [4], ML estimator of \mathbf{g} is obtained as follows:

$$\hat{\mathbf{g}} = \frac{1}{K_p} \boldsymbol{\Psi}^\dagger \mathbf{F}^\dagger \tilde{\mathbf{Y}}. \quad (13)$$

Note that, according to (13), the ML estimator of \mathbf{g} requires $4L^2 + 4LK_p$ real multiplications.

3.4. Truncated KL expansion

A truncated expansion \mathbf{g}_r can be formed by selecting r orthonormal basis vectors among all basis vectors that satisfy $\mathbf{C}_h \boldsymbol{\Psi} = \boldsymbol{\Psi} \boldsymbol{\Lambda}_g$. The optimal one that yields the smallest average mean squared truncation error $(1/L)E[\boldsymbol{\epsilon}_r^\dagger \boldsymbol{\epsilon}_r]$ is the one expanded with the orthonormal basis vectors associated with the first largest r eigenvalues as given by

$$\frac{1}{L} E[\boldsymbol{\epsilon}_r^\dagger \boldsymbol{\epsilon}_r] = \frac{1}{L} \sum_{i=r}^{L-1} \lambda_{g_i}, \quad (14)$$

where $\boldsymbol{\epsilon}_r = \mathbf{g} - \mathbf{g}_r$. For the problem at hand, truncation property of the KL expansion results in a low-rank approximation as well. Thus, a rank- r approximation to $\boldsymbol{\Lambda}_g$ is defined as

$$\boldsymbol{\Lambda}_g = \text{diag} \{ \lambda_{g_0}, \lambda_{g_1}, \dots, \lambda_{g_{r-1}}, 0, \dots, 0 \}. \quad (15)$$

Since the trailing $L - r$ variances $\{\lambda_{g_i}\}_{i=r}^{L-1}$ are small compared to the leading r variances $\{\lambda_{g_i}\}_{i=0}^{r-1}$, then the trailing $L - r$ variances are set to zero to produce the approximation. However, typically the pattern of eigenvalues for $\boldsymbol{\Lambda}_g$ splits the eigenvectors into dominant and subdominant sets. Then the choice of r is more or less obvious. The optimal truncated KL (rank- r) estimator of (11) now becomes

$$\hat{\mathbf{g}}_r = \boldsymbol{\Gamma}_r \boldsymbol{\Psi}^\dagger \mathbf{F}^\dagger \tilde{\mathbf{Y}}, \quad (16)$$

where

$$\begin{aligned} \boldsymbol{\Gamma}_r &= \boldsymbol{\Lambda}_g (K_p \boldsymbol{\Lambda}_g + \sigma^2 \mathbf{I}_L)^{-1} \\ &= \text{diag} \left\{ \frac{\lambda_{g_0}}{\lambda_{g_0} K_p + \sigma^2}, \dots, \frac{\lambda_{g_{r-1}}}{\lambda_{g_{r-1}} K_p + \sigma^2}, 0, \dots, 0 \right\}. \end{aligned} \quad (17)$$

Since our ultimate goal is to obtain MMSE estimator for the channel frequency response \mathbf{H} , from the invariance property of the MMSE estimator, it follows that if $\hat{\mathbf{g}}$ is the estimate of \mathbf{g} , then the corresponding estimate of \mathbf{H} can be obtained for the i th OFDM block as

$$\hat{\mathbf{H}} = \mathcal{F} \boldsymbol{\Psi} \hat{\mathbf{g}}. \quad (18)$$

Thus, from (16) and (17), the truncated MMSE estimator of \mathbf{g} requires $4Lr + 4LK_p + 2r$ real multiplications.

3.5. Performance analysis

We turn our attention to analytical performance results of the batch MMSE approach. We first consider the CRB and derive the closed-form expression for the random KL coefficients, and then exploit the performance of the MMSE channel estimator based on the evaluation of minimum Bayesian MSE.

3.5.1. Cramér-Rao bound for random KL coefficients

The mean squared estimation error for unbiased estimation of a nonrandom parameter has a lower bound, the *Cramér-Rao bound* (CRB), which defines the ultimate accuracy of unbiased estimation procedure. Suppose $\hat{\mathbf{g}}$ is an unbiased estimator of a vector of unknown parameters \mathbf{g} (i.e., $E\{\hat{\mathbf{g}}\} = \mathbf{g}$) then the mean squared error matrix is lower bounded by the inverse of the Fisher information matrix (FIM):

$$E\{(\mathbf{g} - \hat{\mathbf{g}})(\mathbf{g} - \hat{\mathbf{g}})^\dagger\} \geq \mathbf{J}^{-1}(\mathbf{g}). \quad (19)$$

Since the estimation of unknown random parameters \mathbf{g} via MMSE approach is considered in this paper, the modified FIM needs to be taken into account in the derivation of stochastic CRB [12]. Fortunately, the modified FIM can be obtained by a straightforward modification of (19) as

$$\mathbf{J}_M(\mathbf{g}) \triangleq \mathbf{J}(\mathbf{g}) + \mathbf{J}_P(\mathbf{g}), \quad (20)$$

where $\mathbf{J}_P(\mathbf{g})$ represents the *a priori* information.

Under the assumption that \mathbf{g} and $\tilde{\boldsymbol{\eta}}$ are independent of each other and $\tilde{\boldsymbol{\eta}}$ is a zero mean, from [12] and (10), the conditional PDF is given by

$$p(\tilde{\mathbf{Y}}|\mathbf{g}) = \frac{1}{\pi^{K_p} |\mathbf{C}_{\tilde{\boldsymbol{\eta}}}|} \exp\{-\tilde{\mathbf{Y}} - \mathbf{F}\boldsymbol{\Psi}\mathbf{g}\}^\dagger \mathbf{C}_{\tilde{\boldsymbol{\eta}}}^{-1} \{\tilde{\mathbf{Y}} - \mathbf{F}\boldsymbol{\Psi}\mathbf{g}\} \quad (21)$$

from which the derivatives follow as

$$\begin{aligned} \frac{\partial \ln p(\tilde{\mathbf{Y}}|\mathbf{g})}{\partial \mathbf{g}^T} &= (\tilde{\mathbf{Y}} - \mathbf{F}\boldsymbol{\Psi}\mathbf{g})^\dagger \mathbf{C}_{\tilde{\boldsymbol{\eta}}}^{-1} \mathbf{F}\boldsymbol{\Psi}, \\ \frac{\partial^2 \ln p(\tilde{\mathbf{Y}}|\mathbf{g})}{\partial \mathbf{g}^* \partial \mathbf{g}^T} &= -\boldsymbol{\Psi}^\dagger \mathbf{F}^\dagger \mathbf{C}_{\tilde{\boldsymbol{\eta}}}^{-1} \mathbf{F}\boldsymbol{\Psi}, \end{aligned} \quad (22)$$

where the superscript $(\cdot)^*$ indicates the conjugation operation.

Using $\mathbf{C}_{\tilde{\boldsymbol{\eta}}} = \sigma^2 \mathbf{I}_{K_p}$, $\boldsymbol{\Psi}^\dagger \boldsymbol{\Psi} = \mathbf{I}_L$, and $\mathbf{F}^\dagger \mathbf{F} = K_p \mathbf{I}_L$, and taking the expected value yield the following simple form:

$$\begin{aligned} \mathbf{J}(\mathbf{g}) &= -E\left[\frac{\partial^2 \ln p(\tilde{\mathbf{Y}}|\mathbf{g})}{\partial \mathbf{g}^* \partial \mathbf{g}^T}\right] \\ &= -E\left[-\frac{K_p}{\sigma^2} \mathbf{I}_L\right] = \frac{K_p}{\sigma^2} \mathbf{I}_L. \end{aligned} \quad (23)$$

The second term in (20) is easily obtained as follows. Consider the prior PDF of \mathbf{g} as

$$p(\mathbf{g}) = \frac{1}{\pi^L |\boldsymbol{\Lambda}_{\mathbf{g}}|} \exp\{-\mathbf{g}^\dagger \boldsymbol{\Lambda}_{\mathbf{g}}^{-1} \mathbf{g}\}. \quad (24)$$

The respective derivatives are found as

$$\begin{aligned} \frac{\partial \ln p(\mathbf{g})}{\partial \mathbf{g}^T} &= -\mathbf{g}^\dagger \boldsymbol{\Lambda}_{\mathbf{g}}^{-1}, \\ \frac{\partial^2 \ln p(\mathbf{g})}{\partial \mathbf{g}^* \partial \mathbf{g}^T} &= -\boldsymbol{\Lambda}_{\mathbf{g}}^{-1}. \end{aligned} \quad (25)$$

Upon taking the negative expectations, the second term in (20) becomes

$$\begin{aligned} \mathbf{J}_P(\mathbf{g}) &= -E\left[\frac{\partial^2 \ln p(\mathbf{g})}{\partial \mathbf{g}^* \partial \mathbf{g}^T}\right] \\ &= -E[-\boldsymbol{\Lambda}_{\mathbf{g}}^{-1}] \\ &= \boldsymbol{\Lambda}_{\mathbf{g}}^{-1}. \end{aligned} \quad (26)$$

Substituting (23) and (26) in (20) produces for the modified FIM the following:

$$\begin{aligned} \mathbf{J}_M(\mathbf{g}) &= \mathbf{J}(\mathbf{g}) + \mathbf{J}_P(\mathbf{g}) \\ &= \frac{K_p}{\sigma^2} \mathbf{I}_L + \boldsymbol{\Lambda}_{\mathbf{g}}^{-1} \\ &= \frac{1}{\sigma^2} (K_p \mathbf{I}_L + \sigma^2 \boldsymbol{\Lambda}_{\mathbf{g}}^{-1}) \\ &= \frac{1}{\sigma^2} \boldsymbol{\Gamma}^{-1}. \end{aligned} \quad (27)$$

Inverting the matrix $\mathbf{J}_M(\mathbf{g})$ yields

$$\text{CRB}(\hat{\mathbf{g}}) = \mathbf{J}_M^{-1}(\mathbf{g}) = \sigma^2 \boldsymbol{\Gamma}. \quad (28)$$

3.5.2. Bayesian MSE

For the MMSE estimator $\hat{\mathbf{g}}$, the error is

$$\boldsymbol{\epsilon} = \mathbf{g} - \hat{\mathbf{g}}. \quad (29)$$

Since the diagonal entries of the covariance matrix of the error represent the minimum Bayesian MSE, we now derive covariance matrix $\mathbf{C}_{\boldsymbol{\epsilon}}$ of the error vector. From *the performance of the MMSE estimator for the Bayesian linear model theorem* [11], the error covariance matrix is obtained as

$$\begin{aligned} \mathbf{C}_{\boldsymbol{\epsilon}} &= (\boldsymbol{\Lambda}_{\mathbf{g}}^{-1} + (\mathbf{F}\boldsymbol{\Psi})^\dagger \mathbf{C}_{\tilde{\boldsymbol{\eta}}}^{-1} (\mathbf{F}\boldsymbol{\Psi}))^{-1} \\ &= \sigma^2 (K_p \mathbf{I}_L + \sigma^2 \boldsymbol{\Lambda}_{\mathbf{g}}^{-1})^{-1} \\ &= \sigma^2 \boldsymbol{\Gamma} \end{aligned} \quad (30)$$

and then the minimum Bayesian MSE of the full rank estimator becomes (see Appendix A)

$$\begin{aligned} \mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}) &= \frac{1}{L} \text{tr}(\mathbf{C}_{\boldsymbol{\epsilon}}) \\ &= \frac{1}{L} \text{tr}(\sigma^2 \boldsymbol{\Gamma}) = \frac{1}{L} \sum_{i=0}^{L-1} \frac{\lambda_{g_i}}{1 + K_p \lambda_{g_i} \text{SNR}}, \end{aligned} \quad (31)$$

where $\text{SNR} = 1/\sigma^2$ and tr denotes trace operator on matrices.

Comparing (28) with (30), the error covariance matrix of the MMSE estimator coincides with the stochastic CRB of the random vector estimator. Thus, $\hat{\mathbf{g}}$ achieves the stochastic CRB.

As the details are given in Appendix A, $\mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}})$ given in (31) can also be computed for the truncated (low-rank) case as follows:

$$\mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}_r) = \frac{1}{L} \sum_{i=0}^{r-1} \frac{\lambda_{g_i}}{1 + K_p \lambda_{g_i} \text{SNR}} + \frac{1}{L} \sum_{i=r}^{L-1} \lambda_{g_i}. \quad (32)$$

Notice that the second term in (32) is the sum of the powers in the KL transform coefficients not used in the truncated estimator. Thus, the truncated $\mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}_r)$ can be lower bounded by $(1/L) \sum_{i=r}^{L-1} \lambda_{g_i}$ which will cause an irreducible error floor in the SER results.

3.6. Mismatch analysis

Once the true frequency-domain correlation, characterizing the channel statistics and the SNR, is known, the optimal channel estimator can be designed as indicated in Section 4.

However, in mobile wireless communications, the channel statistics depend on the particular environment, for example, indoor or outdoor, urban or suburban, and change with time. Hence, it is important to analyze the performance degradation due to a mismatch of the estimator to the channel statistics as well as the SNR, and to study the choice of the channel correlation and SNR for this estimator so that it is robust to variations in the channel statistics. As a performance measure, we use uncoded symbol error rate (SER) for QPSK signaling. The SER expression for this case is given in [13] as a function of the SNR and the average $\mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}})$ as follows:

$$\text{SER}_{\text{QPSK}} = \frac{3}{4} - \frac{\mu}{2} - \frac{\mu}{\pi} \arctan(\mu), \quad (33)$$

where

$$\mu = \frac{\Omega_g}{\sqrt{(\Omega_g + \mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}))(1 + 2/\text{SNR})}}, \quad (34)$$

and Ω_g represents the normalized variance of the channel gains ($\Omega_g = \sum_{i=0}^{L-1} \lambda_{g_i} = 1$) and $\text{SNR} = 1/\sigma^2$. In practice, the true channel correlations and SNR are not known. If the MMSE channel estimator is designed to match the correlation of a multipath channel impulse response \mathbf{C}_h and SNR, but the true channel parameter $\tilde{\mathbf{h}}$ has the correlation $\mathbf{C}_{\tilde{\mathbf{h}}}$ and the true $\widetilde{\text{SNR}}$, then average Bayesian MSE for the designed channel estimator is obtained as (see Appendices A and B)

(i) SNR mismatch:

$$\mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}) = \frac{1}{L} \sum_{i=0}^{L-1} \frac{\lambda_{g_i}}{(1 + K_p \lambda_{g_i} \text{SNR})^2} \left[1 + K_p \lambda_{g_i} \frac{\text{SNR}^2}{\widetilde{\text{SNR}}} \right]; \quad (35)$$

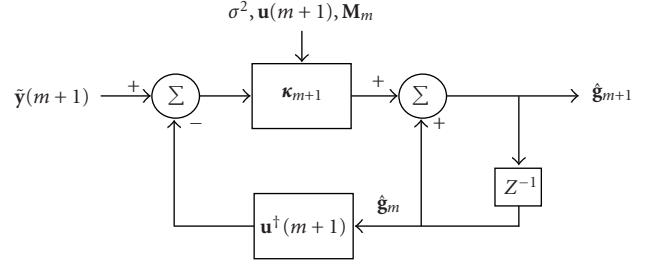


FIGURE 2: Block diagram of sequential MMSE estimator.

(ii) correlation mismatch:

$$\mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}) = \frac{1}{L} \sum_{i=0}^{L-1} \frac{\tilde{\lambda}_{g_i} + K_p \text{SNR} \lambda_{g_i} (\tilde{\lambda}_{g_i} + \lambda_{g_i} - 2\beta_i)}{1 + K_p \text{SNR} \lambda_{g_i}}, \quad (36)$$

where $\tilde{\lambda}_{g_i}$ is the i th diagonal element of $\tilde{\Lambda}_{\tilde{\mathbf{g}}} = \Psi^\dagger \mathbf{C}_{\tilde{\mathbf{h}}} \Psi$, and β_i is i th diagonal element of the real part of the cross-correlation matrix between $\tilde{\mathbf{g}}$ and \mathbf{g} .

4. MMSE ESTIMATION OF KL COEFFICIENTS: SEQUENTIAL APPROACH

We now turn our attention to the derivation of the sequential MMSE algorithm with simple structure. The sequential MMSE approach is proposed in this paper to follow the channel variations by exploiting only channel correlations in frequency. The block diagram for this is shown in Figure 2.

To begin with the algebraic derivation, we use (10) to write m th component of $\tilde{\mathbf{Y}}$ as

$$\tilde{\mathbf{Y}}[m] = \mathbf{u}^\dagger(m) \mathbf{g} + \tilde{\boldsymbol{\eta}}[m], \quad (37)$$

where $\mathbf{u}^\dagger(m)$ is the m th row of $\mathbf{F}\Psi$ and $\tilde{\boldsymbol{\eta}}[m]$ is the m th element of the noise vector $\tilde{\boldsymbol{\eta}}$.

If an MMSE estimator of $\tilde{\mathbf{Y}}[m+1]$ can be found based on $\tilde{\mathbf{Y}}[m]$, denoted by $\hat{\tilde{\mathbf{Y}}}_{m+1|m}$, the prediction error $f_{m+1} = \tilde{\mathbf{Y}}[m+1] - \hat{\tilde{\mathbf{Y}}}_{m+1|m}$ will be orthogonal to $\tilde{\mathbf{Y}}[m]$. We can therefore project \mathbf{g} onto each vector separately and add the results, so that

$$\begin{aligned} \hat{\mathbf{g}}_{m+1} &= \hat{\mathbf{g}}_m + \boldsymbol{\kappa}_{m+1} f_{m+1} \\ &= \hat{\mathbf{g}}_m + \boldsymbol{\kappa}_{m+1} (\tilde{\mathbf{Y}}[m+1] - \mathbf{u}^\dagger(m+1) \hat{\mathbf{g}}_m), \end{aligned} \quad (38)$$

where $\hat{\mathbf{g}}_{m+1}$ is the $(m+1)$ th estimate of \mathbf{g} , and $\boldsymbol{\kappa}_{m+1}$ is the gain factor given as

$$\boldsymbol{\kappa}_{m+1} = \frac{\mathbf{M}_m \mathbf{u}(m+1)}{\mathbf{u}^\dagger(m+1) \mathbf{M}_m \mathbf{u}(m+1) + \sigma^2}. \quad (39)$$

It can be seen that $\mathbf{M}_m = E[(\mathbf{g} - \hat{\mathbf{g}}_m)(\mathbf{g} - \hat{\mathbf{g}}_m)^\dagger]$ is needed in (39), hence update equation for the minimum MSE matrix should also be given. If we substitute (38) in $\mathbf{M}_{m+1} = E[(\mathbf{g} - \hat{\mathbf{g}}_{m+1})(\mathbf{g} - \hat{\mathbf{g}}_{m+1})^\dagger]$, we obtain an update equation

for \mathbf{M}_{m+1} as

$$\mathbf{M}_{m+1} = [\mathbf{I} - \kappa_{m+1} \mathbf{u}^\dagger(m+1)] \mathbf{M}_m. \quad (40)$$

Based on these results, the steps of the sequential MMSE estimator for \mathbf{g} can be summarized as follows.

Initialization. Set the parameters to some initial value $\hat{\mathbf{g}}_0 = \mathbf{0}$, $\mathbf{M}_0 = \Lambda_{\mathbf{g}}$.

- (1) Compute the gain κ_{m+1} from (39).
- (2) Update the estimate of \mathbf{g} from (38).
- (3) Update the minimum MSE matrix from (40).
- (4) Repeat step (1)–step (3) until $m = K_p - 1$.

Some remarks and observations are now in order.

- (i) No matrix inversions are required.
- (ii) Since the MMSE estimator (11) requires $\mathbf{F}^\dagger \mathbf{F}$ to be equal to $K_p \mathbf{I}$ which is satisfied only when $\Delta = K/K_p$ is an integer, however, the sequential version of (11) works as long as $\Delta \leq K/L$.

We now analyze the complexity of the sequential MMSE algorithm. It follows from (39) in step (1) that one needs $4L^2 + 5L$ real multiplications to compute the gain. Similarly, from (38) in step (2), it requires $5L$ real multiplications for the estimator update. Finally, in step (3), we need $8L^2$ real multiplications for the MMSE matrix update. Therefore, the total sequential MMSE algorithm requires $12L^2 + 10L$ real multiplications for one iteration.

4.1. Performance analysis

We turn our attention now to the performance analysis of the adaptive algorithm. We will try to evaluate its convergence properties in terms of mean square error.

From (39) and (40), we conclude that

$$\begin{aligned} \kappa_{m+1} \sigma^2 &= (\mathbf{I} - \kappa_{m+1} \mathbf{u}^\dagger(m+1)) \mathbf{M}_m \mathbf{u}(m+1) \\ &= \mathbf{M}_{m+1} \mathbf{u}(m+1). \end{aligned} \quad (41)$$

Substituting (41) in (39), we have

$$\left(\mathbf{M}_{m+1} - \frac{\sigma^2}{\mathbf{u}^\dagger(m+1) \mathbf{M}_m \mathbf{u}(m+1) + \sigma^2} \mathbf{M}_m \right) \mathbf{u}(m+1) = \mathbf{0}_{L \times 1}. \quad (42)$$

Based on (42), the following recursion is obtained:

$$\begin{aligned} \mathbf{M}_{m+1} &= \frac{\sigma^2}{\mathbf{u}^\dagger(m+1) \mathbf{M}_m \mathbf{u}(m+1) + \sigma^2} \mathbf{M}_m \\ &= \delta_{m+1|m} \mathbf{M}_m. \end{aligned} \quad (43)$$

Due to positive definite property of error covariance matrix \mathbf{M}_m , it follows that $\mathbf{u}^\dagger(m+1) \mathbf{M}_m \mathbf{u}(m+1) > 0$. As a result, $0 < \delta_{m+1|m} < 1$.

Define average MSE at the m th step as $\text{MSE}_m = (1/L) \text{tr}(\mathbf{M}_m)$, then it follows from (43) that

$$\text{MSE}_{m+1} = \delta_{m+1|m} \text{MSE}_m. \quad (44)$$

Thus, we observe that as $m \rightarrow \infty$, $\text{MSE}_m \rightarrow 0$ which means that $\hat{\mathbf{g}}_m$ converges to \mathbf{g} in the mean square.

5. SIMULATIONS

In this section, the merits of our channel estimators are illustrated through simulations. We choose average mean square error (MSE) and symbol error rate (SER) as our figure of merits. We consider the fading multipath channel with L paths given by (45) with an exponentially decaying power delay profile $\theta(\tau_l) = C e^{-\tau_l/\tau_{\text{RMS}}}$ with delays τ_l that are uniformly and independently distributed over the length L_{CP} . Note that \mathbf{h} is chosen as complex Gaussian leading to a Rayleigh fading channel with root mean square (RMS) width τ_{RMS} and normalizing constant C . In [3], it is shown that the normalized exponential discrete channel correlation for different subcarriers is

$$r_f(k) = \frac{1 - \exp(-L(1/\tau_{\text{RMS}} + 2\pi jk/K))}{\tau_{\text{RMS}}(1 - \exp(-L/\tau_{\text{RMS}}))(1/\tau_{\text{RMS}} + 2\pi jk/K)}. \quad (45)$$

The scenario for our simulation study consists of a wireless QPSK-OFDM system employing the pulse shape as a unit-energy Nyquist-root raised-cosine shape with rolloff $\alpha = 0.2$, with a symbol period (T_s) of 0.120 microsecond, corresponding to an uncoded symbol rate of 8.33 Mbps. Transmission bandwidth (5 MHz) is divided into 1024 tones. We assume that the fading multipath channel has $L = 40$ paths with an exponentially decaying power delay profile (45) with $\tau_{\text{RMS}} = 5$ sample (0.6 microsecond) long.

5.1. Batch MMSE approach

A QPSK-OFDM sequence passes through channel taps and is corrupted by AWGN (0 dB, 5 dB, 10 dB, 15 dB, 20 dB, 25 dB, and 30 dB, respectively). We use a pilot symbol for every twenty ($\Delta = 20$) symbols. The MSE at each SNR point is averaged over 1000 realizations. We compare the experimental MSE performance and its theoretical Bayesian MSE of the proposed full-rank MMSE estimator with maximum-likelihood (ML) estimator and its corresponding Cramér-Rao bound (CRB). Figure 3 confirms that MMSE estimator performs better than ML estimator at low SNR. However, the 2 approaches have comparable performance at high SNRs. To observe the performance, we also present the MMSE and ML estimated channel SER results together with theoretical SER in Figure 4. Due to the fact that spaces between the pilot symbols are not chosen as a factor of the number of subcarriers, an error floor is observed in Figures 3 and 4. In the case of choosing the pilot space as a factor of number of subcarriers, the error floor vanishes because of the fact that the orthogonality condition between the subcarriers at pilot locations is satisfied. In other words, the curves labeled as simulation results for MMSE estimator and ML estimator fit to the theoretical curve at high SNRs. It also shows that the MMSE estimated channel SER results are better than ML estimated channel SER especially at low SNRs.

5.1.1. SNR design mismatch

In order to evaluate the performance of the proposed full-rank MMSE estimator to mismatch only in SNR design, the

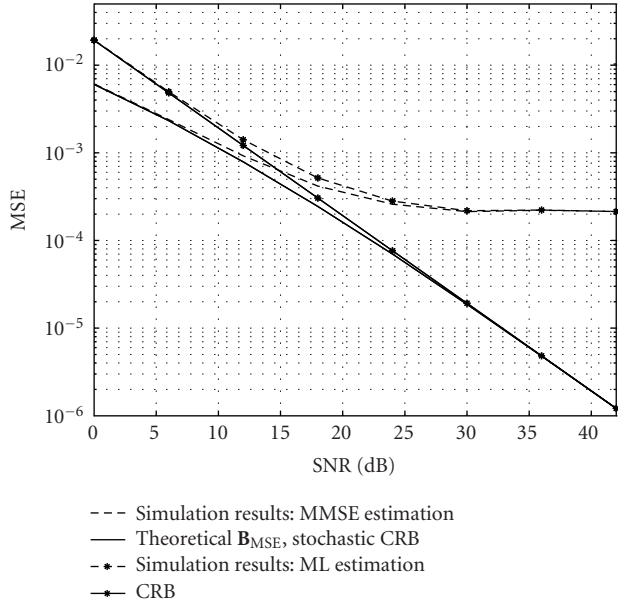


FIGURE 3: Performance of proposed MMSE and MLE together with \mathbf{B}_{MSE} and CRB.

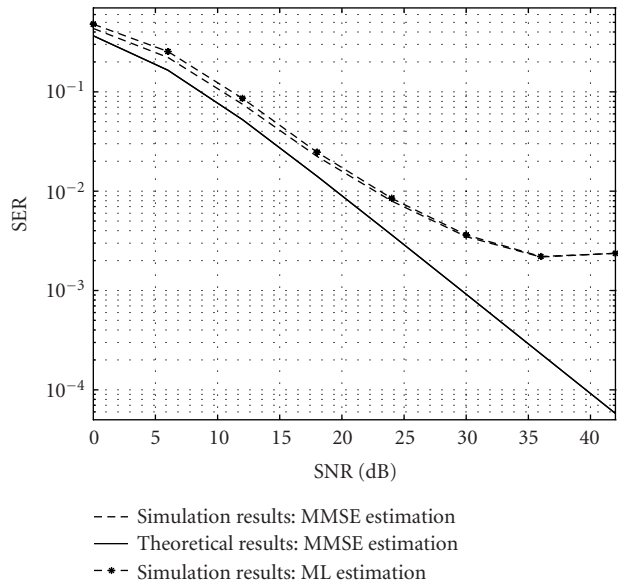


FIGURE 4: Symbol error rate results.

estimator is tested when SNRs of 10 and 30 dB are used in the design. The SER curves for a design SNR of 10, 30 dB are shown in Figure 5. The performance of the MMSE estimator for high SNR (30 dB) design is better than low SNR (10 dB) design across a range of SNR values (0–30 dB). This results confirm that channel estimation error is concealed in noise for low SNR whereas it tends to dominate for high SNR. Thus, the system performance degrades especially for low SNR design.

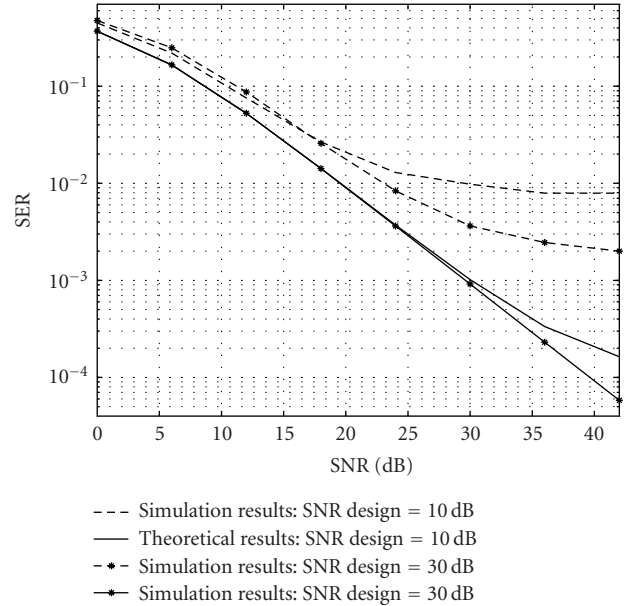


FIGURE 5: Effects of SNR design mismatch on SER.

5.1.2. Correlation mismatch

To further analyze full-rank MMSE estimator's performance, we need to study sensitivity of the estimator to design errors, that is, correlation mismatch. We therefore designed the estimator for a uniform channel correlation which gives the worst MSE performance among all channels [3, 5] and evaluated it for an exponentially decaying power delay profile. The uniform channel correlation between the attenuations can be obtained by letting $\tau_{\text{RMS}} \rightarrow \infty$ in (45), resulting in

$$r_f(k) = \frac{1 - \exp(2\pi jLk/K)}{2\pi jk/K}. \quad (46)$$

Figures 6 and 7 demonstrate the estimator's sensitivity to the channel statistics in terms of average MSE and SER performance measures, respectively. As it can be seen from Figures 6 and 7, only small performance loss is observed for low SNRs when the estimator is designed for mismatched channel statistics. This justifies the result that a design for worst correlation is robust to mismatch.

5.1.3. Performance of the truncated estimator

The truncated estimator performance is also studied as a function of the number of KL coefficients. Figure 8 presents the MSE result of the truncated MMSE estimator for SNR = 10, 20, and 30 dB. If only a few expansion coefficients are employed to reduce the complexity of the proposed estimator, then the MSE between channel parameters becomes large. However, if the number of parameters in the expansion is increased, the irreducible error floor still occurs.

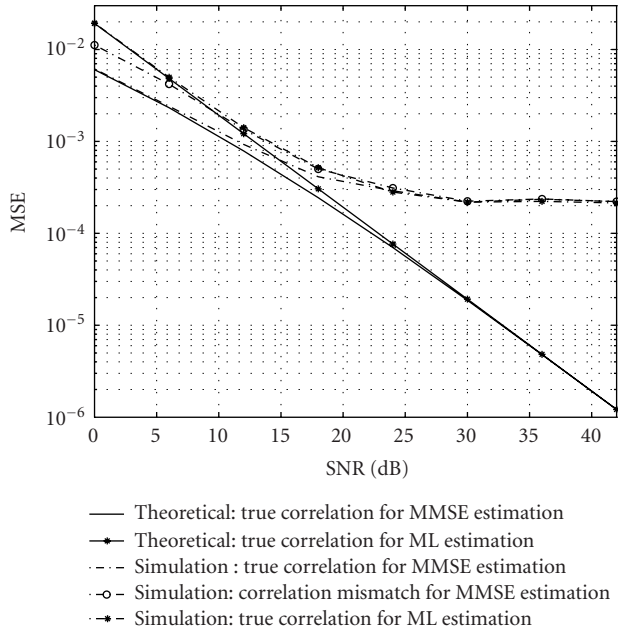


FIGURE 6: Effects of correlation mismatch on MSE.

5.2. Sequential MMSE approach

The MSE results of the sequential full-rank MMSE algorithm are obtained and presented as shown in Figure 9. In order to better evaluate the performance of the proposed sequential MMSE estimation algorithm, we compare it with previously developed least mean square (LMS) and recursive least squares (RLS) algorithms. It can be seen from simulations that recursive MMSE estimator yields better performance than LMS and RLS approaches and achieves Bayesian MSE especially for low SNR.

For the convergence of the proposed adaptive algorithm, MSE versus iteration is plotted for SNR = 10, 20, 30, and 40 dB in Figure 10. As expected, the proposed sequential algorithm converges faster for high SNR values.

Finally, we wish to evaluate the performance of the algorithm for different values of pilot spacing 10, 20, 30, 40, and 50 by plotting the MSEs and SERs with respect to SNR in Figures 11 and 12, respectively. For the values pilot spacing Δ larger than K/L , the SER and MSE performances decrease as Δ increases.

6. CONCLUSION

We consider the design of low-complexity MMSE channel estimators for OFDM systems in unknown wireless dispersive fading channels. We first derive the batch MMSE estimator based on the stochastic orthogonal expansion representation of the channel via KL transform. Based on such representation, we show that no matrix inversion is needed in the MMSE algorithm. Therefore, the computational cost for implementing the proposed MMSE estimator is low and computation is numerically stable. Moreover, the performance of our proposed batch method was first studied through the derivation of stochastic CRB for Bayesian approach.

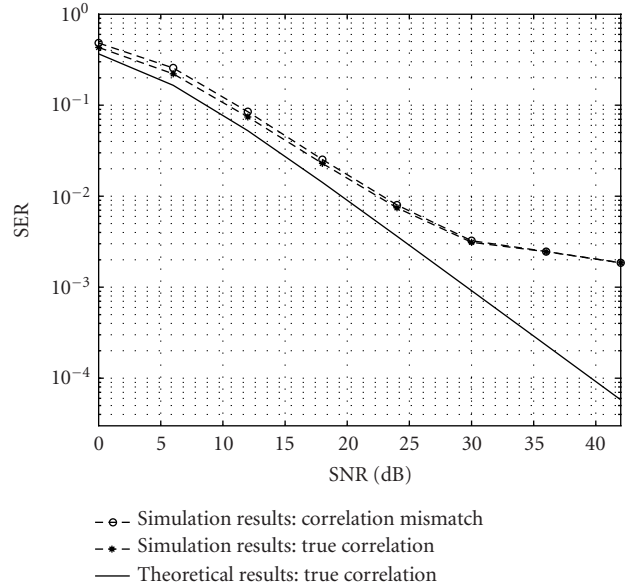


FIGURE 7: Effects of correlation mismatch on SER.

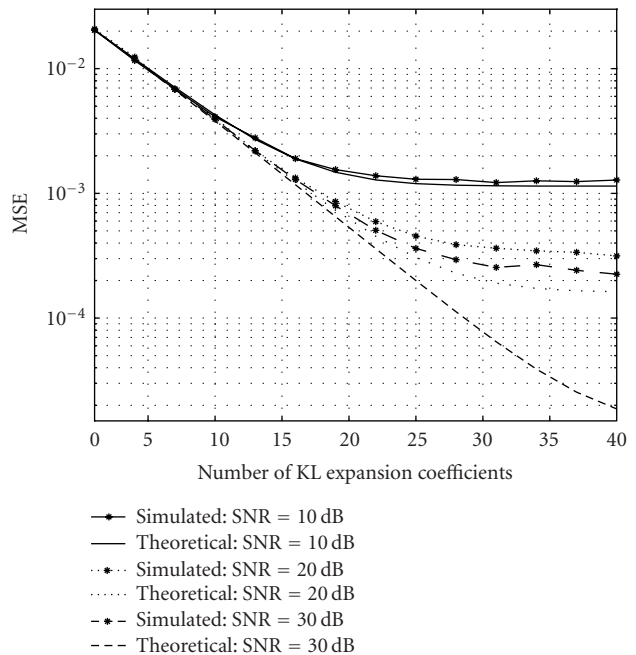


FIGURE 8: MSE as a function of KL expansion coefficients.

Since the actual channel statistics and SNR may vary within OFDM block, we have also analyzed the effect of modelling mismatch on the estimator performance and shown both analytically and through simulations that the performance degradation due to such mismatch is negligible for low SNR values. The MMSE estimator is then extended to sequential implementation which enjoys the elegance of the simple structure and fast convergence.

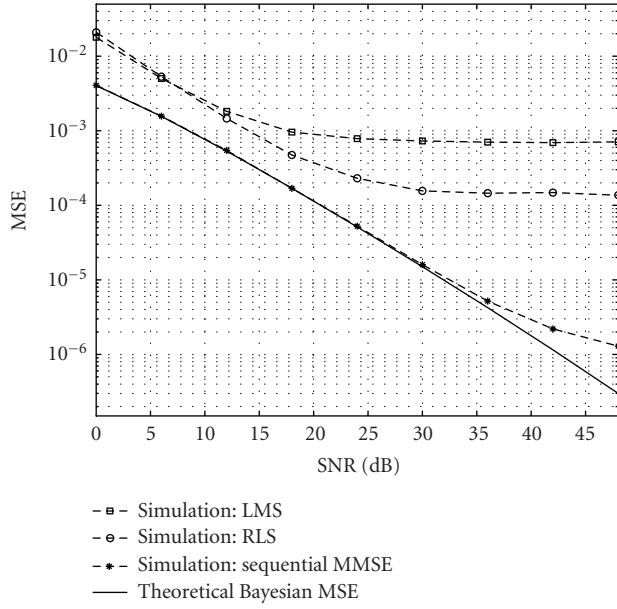


FIGURE 9: Sequential MMSE performance.

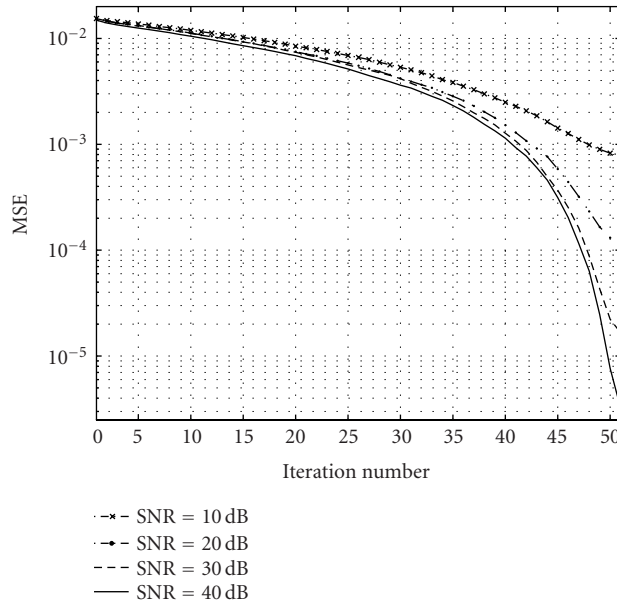


FIGURE 10: Convergence of the sequential MMSE estimator.

APPENDICES

A. BAYESIAN MSE FOR TRUNCATED MMSE KL ESTIMATOR UNDER SNR MISMATCH

Substituting (10) in (16), the truncated MMSE KL estimator now becomes

$$\hat{\mathbf{g}}_r = K_p \Gamma_r \mathbf{g} + \Gamma_r \Psi^\dagger \mathbf{F}^\dagger \tilde{\boldsymbol{\eta}}. \quad (\text{A.1})$$

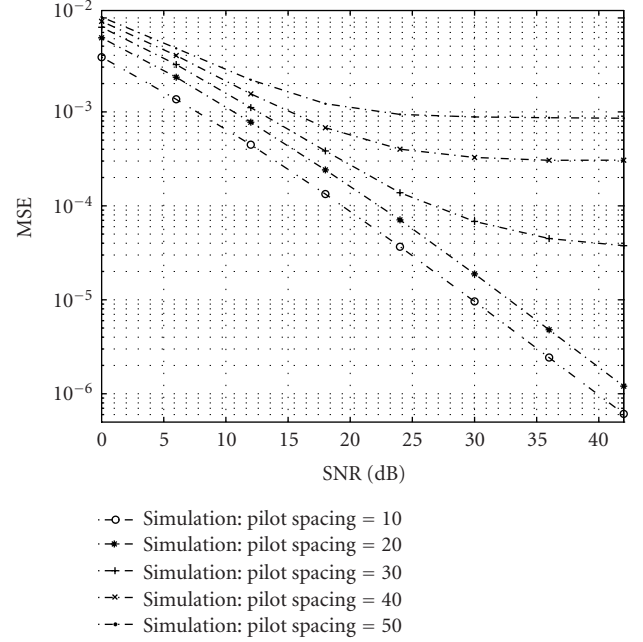


FIGURE 11: Performance of the sequential MMSE for different pilot spacings.

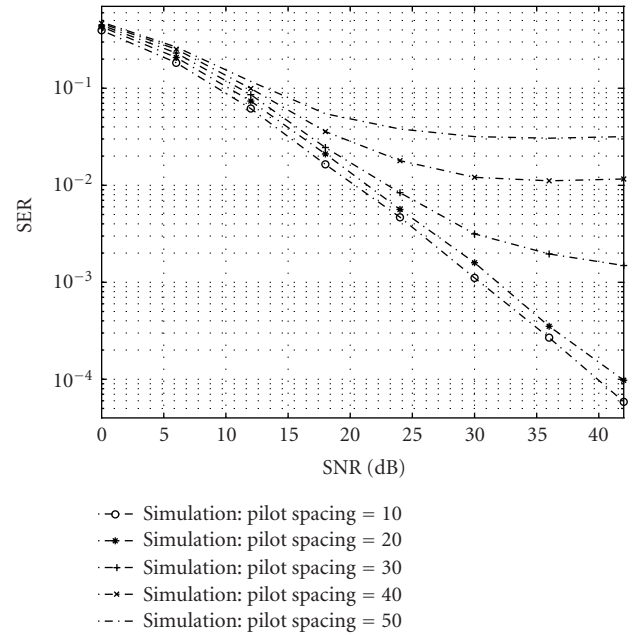


FIGURE 12: Symbol error rate of the sequential MMSE for different pilot spacings.

The estimation error

$$\begin{aligned} \hat{\boldsymbol{\epsilon}}_r &= \mathbf{g} - \hat{\mathbf{g}}_r \\ &= \mathbf{g} - (K_p \Gamma_r \mathbf{g} + \Gamma_r \Psi^\dagger \mathbf{F}^\dagger \tilde{\boldsymbol{\eta}}) \\ &= (\mathbf{I}_L - K_p \Gamma_r) \mathbf{g} - \Gamma_r \Psi^\dagger \mathbf{F}^\dagger \tilde{\boldsymbol{\eta}} \end{aligned} \quad (\text{A.2})$$

and then the average Bayesian MSE is

$$\begin{aligned}
 \mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}_r) &= \frac{1}{L} \text{tr}(\mathbf{C}_{\hat{\mathbf{e}}_r}) \\
 &= \frac{1}{L} \text{tr}(\mathbf{\Lambda}_{\mathbf{g}}(\mathbf{I}_L - K_p \mathbf{\Gamma}_r)^2 + K_p \tilde{\sigma}^2 \mathbf{\Gamma}_r^2) \\
 &= \frac{1}{L} \sum_{i=0}^{r-1} \left[\lambda_{g_i} \left(1 - K_p \frac{\lambda_{g_i}}{\lambda_{g_i} K_p + \sigma^2} \right)^2 \right. \\
 &\quad \left. + K_p \tilde{\sigma}^2 \left(\frac{\lambda_{g_i}}{\lambda_{g_i} K_p + \sigma^2} \right)^2 \right] + \frac{1}{L} \sum_{i=r}^{L-1} \lambda_{g_i} \\
 &= \frac{1}{L} \sum_{i=0}^{r-1} \lambda_{g_i} \frac{\tilde{\sigma}^2 K_p \lambda_{g_i} + \sigma^4}{(K_p \lambda_{g_i} + \sigma^2)^2} + \frac{1}{L} \sum_{i=r}^{L-1} \lambda_{g_i} \\
 &\quad \text{where } \sigma^2 = \frac{1}{\text{SNR}}, \tilde{\sigma}^2 = \frac{1}{\text{SNR}} \\
 &= \frac{1}{L} \sum_{i=0}^{r-1} \frac{\lambda_{g_i}}{(1 + K_p \lambda_{g_i} \text{SNR})^2} \left[1 + K_p \lambda_{g_i} \frac{\text{SNR}^2}{\text{SNR}} \right] \\
 &\quad + \frac{1}{L} \sum_{i=r}^{L-1} \lambda_{g_i}. \tag{A.3}
 \end{aligned}$$

Based on the result obtained in (A.3), Bayesian estimator performance can be further elaborated for the following scenarios.

- (i) By taking $\widetilde{\text{SNR}} = \text{SNR}$, the performance result for the case of no SNR mismatch is

$$\mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}_r) = \frac{1}{L} \sum_{i=0}^{r-1} \frac{\lambda_{g_i}}{1 + K_p \lambda_{g_i} \text{SNR}} + \frac{1}{L} \sum_{i=r}^{L-1} \lambda_{g_i}. \tag{A.4}$$

- (ii) As $r \rightarrow L$ in (A.3), $\mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}})$ under SNR mismatch results in the following Bayesian MSE:

$$\mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}) = \frac{1}{L} \sum_{i=0}^{L-1} \frac{\lambda_{g_i}}{(1 + K_p \lambda_{g_i} \text{SNR})^2} \left[1 + K_p \lambda_{g_i} \frac{\text{SNR}^2}{\widetilde{\text{SNR}}} \right]. \tag{A.5}$$

- (iii) Finally, the Bayesian MSE in the case of no SNR mismatch can also be obtained as

$$\mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}) = \frac{1}{L} \sum_{i=0}^{L-1} \frac{\lambda_{g_i}}{1 + K_p \lambda_{g_i} \text{SNR}}. \tag{A.6}$$

B. BAYESIAN MSE FOR TRUNCATED MMSE KL ESTIMATOR UNDER CORRELATION MISMATCH

In this appendix, we derive the Bayesian MSE of the truncated MMSE KL estimator under correlation mismatch. Although the real multipath channel \mathbf{h} has the expansion correlation $\mathbf{C}_{\mathbf{h}}$, we designed the estimator for the multipath channel $\mathbf{h} = \mathbf{\Psi} \mathbf{g}$ with correlation $\mathbf{C}_{\mathbf{h}}$. To evaluate the estimation error $\tilde{\mathbf{g}} - \hat{\mathbf{g}}_r$ in the same space, we expand $\tilde{\mathbf{h}}$ onto the eigenspace of \mathbf{h} as $\tilde{\mathbf{h}} = \mathbf{\Psi} \tilde{\mathbf{g}}$ resulting in correlated expansion coefficients.

For the real channel, data model in (10) can be rewritten as

$$\tilde{\mathbf{Y}} = \mathbf{F} \mathbf{\Psi} \tilde{\mathbf{g}} + \tilde{\boldsymbol{\eta}} \tag{B.1}$$

and substituting in (16), the truncated MMSE KL estimator now becomes

$$\hat{\mathbf{g}}_r = K_p \mathbf{\Gamma}_r \mathbf{g} + \mathbf{\Gamma}_r \mathbf{\Psi}^\dagger \mathbf{F}^\dagger \tilde{\boldsymbol{\eta}}. \tag{B.2}$$

For the truncated MMSE estimator, the error is

$$\begin{aligned}
 \hat{\mathbf{e}}_r &= \tilde{\mathbf{g}} - \hat{\mathbf{g}}_r \\
 &= \tilde{\mathbf{g}} - K_p \mathbf{\Gamma}_r \mathbf{g} - \mathbf{\Gamma}_r \mathbf{\Psi}^\dagger \mathbf{F}^\dagger \tilde{\boldsymbol{\eta}}. \tag{B.3}
 \end{aligned}$$

As a result, the average Bayesian MSE is

$$\begin{aligned}
 \mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}_r) &= \frac{1}{L} \text{tr}(\mathbf{C}_{\hat{\mathbf{e}}_r}) \\
 &= \frac{1}{L} \text{tr}(\mathbf{\Lambda}_{\tilde{\mathbf{g}}} + K_p^2 \mathbf{\Gamma}_r^2 \mathbf{\Lambda}_{\mathbf{g}} + \sigma^2 K_p \mathbf{\Gamma}_r^2 - 2K_p \mathbf{\Gamma}_r \boldsymbol{\beta}) \\
 &= \frac{1}{L} \sum_{i=0}^{r-1} \left[\tilde{\lambda}_{g_i} + \frac{K_p \lambda_{g_i} (\lambda_{g_i} - 2\beta_i)}{K_p \lambda_{g_i} + \sigma^2} \right] \\
 &\quad + \frac{1}{L} \sum_{i=r}^{L-1} \tilde{\lambda}_{g_i}, \quad \sigma^2 = \frac{1}{\text{SNR}} \\
 &= \frac{1}{L} \sum_{i=0}^{r-1} \left[\tilde{\lambda}_{g_i} + \frac{K_p \text{SNR} \lambda_{g_i} (\lambda_{g_i} - 2\beta_i)}{1 + K_p \text{SNR} \lambda_{g_i}} \right] + \frac{1}{L} \sum_{i=r}^{L-1} \tilde{\lambda}_{g_i} \\
 &= \frac{1}{L} \sum_{i=0}^{r-1} \frac{\tilde{\lambda}_{g_i} + K_p \text{SNR} \lambda_{g_i} (\tilde{\lambda}_{g_i} + \lambda_{g_i} - 2\beta_i)}{1 + K_p \text{SNR} \lambda_{g_i}} \\
 &\quad + \frac{1}{L} \sum_{i=r}^{L-1} \tilde{\lambda}_{g_i}, \tag{B.4}
 \end{aligned}$$

where $\boldsymbol{\beta}$ is the real part of $E[\tilde{\mathbf{g}} \mathbf{g}^\dagger]$ and β_i 's are the diagonal elements of $\boldsymbol{\beta}$. With this result, we will now highlight some special cases.

- (i) Letting $\beta_i = \lambda_{g_i} = \tilde{\lambda}_{g_i}$ in (B.4) for the case of no mismatch in the correlation of KL expansion coefficients, the truncated Bayesian MSE is identical to that obtained in (A.4).
- (ii) As $r \rightarrow L$ in (B.4), Bayesian MSE under correlation mismatch is obtained to yield

$$\mathbf{B}_{\text{MSE}}(\hat{\mathbf{g}}) = \frac{1}{L} \sum_{i=0}^{L-1} \frac{\tilde{\lambda}_{g_i} + K_p \text{SNR} \lambda_{g_i} (\tilde{\lambda}_{g_i} + \lambda_{g_i} - 2\beta_i)}{1 + K_p \text{SNR} \lambda_{g_i}}. \tag{B.5}$$

- (iii) Under no correlation mismatch in (B.5) where $\beta_i = \lambda_{g_i} = \tilde{\lambda}_{g_i}$, Bayesian MSE obtained from (B.5) is identical to that in (A.6).
- (iv) Also note that as $\text{SNR} \rightarrow \infty$, (B.4) reduces to $\text{MSE}(\tilde{\mathbf{g}} - \mathbf{g}_r)$.

ACKNOWLEDGMENTS

This paper has been produced as part of the NEWCOM Network of Excellence, a project funded by the European Commission's 6th Framework Programme. This work was supported by the Research Fund of the University of Istanbul, Project numbers UDP-362/04082004 and 220/29042004. Part of the results of this paper was presented at the Sixth Baiona Workshop on Signal Processing in Communications, Baiona, Spain, September 8–10, 2003.

REFERENCES

- [1] R. Van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*, Artech House Publishers, Boston, Mass, USA, 2000.
- [2] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, no. 2, pp. 100–109, 1995.
- [3] O. Edfors, M. Sandell, J.-J. van de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, 1998.
- [4] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Processing*, vol. 49, no. 12, pp. 3065–3073, 2001.
- [5] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, 1998.
- [6] P. Schniter, "Low-complexity estimation of doubly-selective channels," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC'03)*, pp. 200–204, Rome, Italy, June 2003.
- [7] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.
- [8] W. C. Jakes Jr., *Microwave Mobile Communications*, John Wiley & Sons, New York, NY, USA, 1974.
- [9] K. W. Yip and T.-S. Ng, "Karhunen-Loeve expansion of the WSSUS channel output and its application to efficient simulation," *IEEE J. Select. Areas Commun.*, vol. 15, no. 4, pp. 640–646, 1997.
- [10] M. Siala and D. Duponteil, "Maximum a posteriori multipath fading channel estimation for CDMA systems," in *IEEE 49th Vehicular Technology Conference (VTC '99)*, vol. 2, pp. 1121–1125, Houston, Tex, USA, May 1999.
- [11] S. M. Kay, *Fundamentals of Statistical Signal Processing. Estimation Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.
- [12] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, Wiley-Interscience, New York, NY, USA, 2001.
- [13] J. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 1983.

Habib Şenol received the B.S. and M.S. degrees from the University of Istanbul in 1993 and in 1999, respectively. He is currently a Ph.D. student in the Department of Electronics Engineering, Işık University. From 1996 to 1999, he was a Research Assistant at the University of Istanbul. In 1999, as a Lecturer, he joined the faculty of the Department of Computer Engineering,

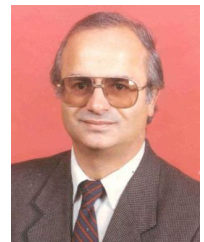


Kadir Has University. His general research interests cover communication theory, estimation theory, statistical signal processing, and information theory. His current research activities are focused on wireless communication concepts with specific attention to channel estimation algorithms for multicarrier (OFDM) systems.

Hakan A. Çırpan received the B.S. degree in 1989 from Uludag University, Bursa, Turkey, the M.S. degree in 1992 from Istanbul University, Istanbul, Turkey, and the Ph.D. degree in 1997 from the Stevens Institute of Technology, Hoboken, New Jersey, USA, all in electrical engineering. From 1995 to 1997, he was a Research Assistant at the Stevens Institute of Technology, working on signal processing algorithms for wireless communication systems. In 1997, he joined the faculty of the Department of Electrical-Electronics Engineering, Istanbul University. His general research interests cover wireless communications, statistical signal and array processing, system identification, and estimation theory. His current research activities are focused on signal processing and communication concepts with specific attention to channel estimation and equalization algorithms for space-time coding and multicarrier (OFDM) systems. Dr. Çırpan received Piskin Award from Stevens Institute of Technology as well as Professor Nazim Terzioğlu Award from the Research Fund of The University of Istanbul. He is a Member of IEEE and Member of Sigma Xi.



Erdal Panayırıcı received the Diploma Engineering degree in electrical engineering from Istanbul Technical University, Istanbul, Turkey, in 1964, and the Ph.D. degree in electrical engineering and system science from Michigan State University, East Lansing Michigan, USA, in 1970. Between 1970–2000, he was with the Faculty of Electrical and Electronics Engineering at the Istanbul Technical University, where he was a Professor and Head of the Telecommunications Chair. Currently, he is a Professor and Head of the Electronics Engineering Department at Işık University, Istanbul, Turkey. He is engaged in research and teaching in digital communications and wireless systems, equalization and channel estimation in multicarrier (OFDM) communication systems, and efficient modulation and coding techniques (TCM and turbo coding). He spent two years (1979–1981) with the Department of Computer Science, Michigan State University, as a Fulbright-Hays Fellow and a NATO Senior Scientist. Between 1983–1986 he served as a NATO Advisory Committee Member for the Special Panel on Sensory Systems for Robotic Control. From August 1990 to December 1991, he was with the Center for Communications and Signal Processing, New Jersey Institute of Technology, as a Visiting Professor, and took part in the research project on interference cancellation by array processing. Between 1998–2000, he was a Visiting Professor in the Department of Electrical Engineering, Texas A&M University, and took part in research on developing efficient synchronization algorithms for OFDM systems. Between 1995–1999, Professor Panayırıcı was an Editor for the IEEE Transactions on Communications in the fields of synchronization and equalizations. He is a Fellow of the IEEE and Member of Sigma Xi.



Impact of Channel Estimation Errors on Multiuser Detection via the Replica Method

Husheng Li

*Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA
Email: hushengl@princeton.edu*

H. Vincent Poor

*Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA
Email: poor@princeton.edu*

Received 26 January 2005

For practical wireless DS-CDMA systems, channel estimation is imperfect due to noise and interference. In this paper, the impact of channel estimation errors on multiuser detection (MUD) is analyzed under the framework of the replica method. System performance is obtained in the large system limit for optimal MUD, linear MUD, and turbo MUD, and is validated by numerical results for finite systems.

Keywords and phrases: CDMA, multiuser detection, replica method, channel estimation.

1. INTRODUCTION

Multiuser detection (MUD) [1] can be used to mitigate multiple access interference (MAI) in direct-sequence code division multiple access (DS-CDMA) systems, thereby substantially improving the system performance compared with the conventional matched filter (MF) reception. The maximum likelihood (ML)-based optimal MUD, introduced in [2], is exponentially complex in the number of users, thus being difficult to implement in practical systems. Consequently, various suboptimal MUD algorithms have been proposed to effect a tradeoff between performance and computational cost. For example, linear processing can be applied, based on zero-forcing or minimum mean square error (MMSE) criteria, thus resulting in the decorrelator [1] and the MMSE detector [3]. For nonlinear processing, a well-known approach is decision-feedback-based interference cancellation (IC) [1], which can be implemented in a parallel fashion (PIC) or successive fashion (SIC). It should be noted that the above algorithms are suitable for systems without channel codes. For channel-coded CDMA systems, the turbo principle can be introduced to improve the performance iteratively using the decision feedback from channel decoders, resulting in

turbo MUD [4], which can also be simplified using PIC [5]. The decisions of channel decoders can also be fed back in the fashion of SIC, and it has been shown that SIC combined with MMSE MUD achieves the sum channel capacity [6].

It is difficult to obtain explicit expressions for the performance of most MUD algorithms in finite systems (here, “finite” means that the number of users and spreading gain are finite). In recent years, asymptotic analysis has been applied to obtain the performance of such systems in the large system limit, which means that the system size tends to infinity while keeping the system load constant. The explicit expressions obtained from asymptotic analysis can provide more insight than simulation results and can be used as approximations for finite systems. The theory of large random matrices [1, 7] has been applied to the asymptotic analysis of MMSE MUD, resulting in the Tse-Hanly equation [8], which quantifies implicitly multiuser efficiency. However, this method is valid for only linear MUD and cannot be used for the analysis of nonlinear algorithms. For ML optimal MUD, the performance is determined by the sum of many exponential terms, which is difficult to tackle with matrices. Recently, attention has been paid to the analogy between optimal MUD and free energy in statistical mechanics [9], which has motivated researchers to apply mathematical tools developed in statistical mechanics to the analysis of MUD. In [10, 11], the replica method, which was developed in the context of spin glasses

theory, has been applied as a unified framework to both optimal and linear MUD, resulting in explicit asymptotic expressions for the corresponding bit error rates and spectral efficiencies. These results have been extended to turbo MUD in [12]. It should be noted that the replica method is based on some assumptions which still require rigorous mathematical proof. However, the corresponding conclusions match simulation results and some known theoretical conclusions well.

In practical wireless communication systems, the transmitted signals experience fading. In the above MUD algorithms, the channel state information (CSI) is assumed to be known to the receiver. However, this is not a reasonable assumption since channel estimation is imperfect due to the existence of noise and interference. Therefore, it is of interest to analyze the performance of MUD with imperfect channel estimates. For linear MUD, the impact of channel estimation error on detection has been studied in [13, 14, 15] using the theory of large random matrices. In this paper, we will apply the replica method to analyze the corresponding impact on optimal MUD, and then extend the results to linear or turbo MUD, under some assumptions on the channel estimation error. The results can be used to determine the number of training symbols needed for channel estimation.

The remainder of this paper is organized as follows. The signal model is explained in Section 2 and the replica method is briefly introduced in Section 3. Optimal MUD with imperfect channel estimation is discussed in Section 4 and the results are extended to linear and turbo MUD in Section 5. Simulation results and conclusions are given in Sections 6 and 7, respectively.

2. SIGNAL MODEL

2.1. Signal model

We consider a synchronous uplink DS-CDMA system, which operates over a frequency selective fading channel of order P (i.e., P is the delay spread in chip intervals). Let K denote the number of active users, N the spreading gain, and $\beta \triangleq K/N$ the system load. In this paper, our analysis is based on the large system limit, where $K, N, P \rightarrow \infty$ while keeping K/N and P/N constant.

We model the frequency selective fading channels as discrete finite-impulse-response (FIR) filters. For simplicity, we assume that the channel coefficients are real. The z -transform of the channel response of user k is given by

$$h_k(z) = \sum_{p=0}^{P-1} g_k(p)z^p, \quad (1)$$

where $\{g_k(p)\}_{p=0, \dots, P-1}$ are the corresponding independent and identically distributed (i.i.d.) (with respect to both k and p) channel coefficients having variance $1/P$. For simplicity, we consider only the case in which $P/N \ll 1$. Thus we can ignore the intersymbol interference (ISI) and deal with only the portion uncontaminated by ISI.

The chip matched filter output at the l th chip period in a fixed symbol period can be written as

$$r(l) = \frac{1}{\sqrt{N}} \sum_{k=1}^K b_k h_k(l) + n(l), \quad l = P, P+1, \dots, N, \quad (2)$$

where b_k denotes the binary phase shift keying (BPSK) modulated channel symbol of user k with normalized power 1, $\{n(l)\}$ is additive white Gaussian noise (AWGN), which satisfies $E\{|n(l)|^2\} = \sigma_n^2$,¹ and $\{h_k(l)\}$ is the convolution of the spreading codes and channel coefficients:

$$h_k(l) = s_k(l) \star g_k(l), \quad (3)$$

where $s_k(l)$ is the l th chip of the original spreading code of user k , which is i.i.d. with respect to both k and l and takes values 1 and -1 equiprobably. We call the $(N+P-1) \times 1$ vector² $\mathbf{h}_k = (h_k(0), \dots, h_k(N+P-2))^T$ the *equivalent spreading code* of user k . Due to the assumption that $P/N \ll 1$, we can approximate $N-P+1$ by N for notational simplicity. Then the received signal in the fixed symbol period can be written in a vector form

$$\mathbf{r} = \frac{1}{\sqrt{N}} H \mathbf{b} + \mathbf{n}, \quad (4)$$

where $\mathbf{r} = (r(P), \dots, r(N))^T$, $H = (\mathbf{h}_1, \dots, \mathbf{h}_K)$, and $\mathbf{b} = (b_1, \dots, b_K)^T$. It is easy to show that $(1/N) \|\mathbf{h}_k\|^2 \rightarrow 1$, as $P \rightarrow \infty$. Thus, we can ignore the performance loss incurred by the fluctuations of received power in the fading channels and consider only the impact of channel estimation error.

2.2. Channel estimation error

In practical wireless communication systems, the channel coefficients $\{g_k(l)\}$ are unknown to the receiver, and the corresponding channel estimates $\{\hat{g}_k(l)\}$ are imprecise due to the existence of noise and interference. We assume that training symbol-based channel estimation [16] is applied to provide the channel estimates. On denoting the channel estimation error by $\delta g_k(l) \triangleq g_k(l) - \hat{g}_k(l)$, $\{\delta g_k(l)\}$ are jointly Gaussian-distributed and mutually independent for sufficiently large numbers of training symbols [16]. Therefore, it is reasonable to assume that $\{\delta g_k(l)\}$ is independent for different values of k and l . In this paper, we consider only the following two types of channel estimation.

- (i) *ML channel estimation.* It is well known that ML estimation is asymptotically unbiased under some regulation conditions. Thus, we can assume that the estimation error $\delta g_k(l)$ has zero expectation conditioned on $g_k(l)$, and is therefore correlated with $\hat{g}_k(l)$.

¹Note that σ_n^2 is the noise variance, normalized to represent the inverse signal-to-noise ratio.

²Superscript T denotes transposition and superscript H denotes conjugate transposition.

(ii) *MMSE channel estimation.* An important property of the MMSE estimate, namely the conditional expectation $E\{g_k(l)|Y\}$, where Y is the observation, is that the estimation error $\delta g_k(l)$ is uncorrelated with $\hat{g}_k(l)$, and thus is biased.

We assume that the receiver uses the imperfect channel estimates to construct the corresponding equivalent spreading code, namely $\hat{\mathbf{h}}_k$. Thus, the error of the i th chip of $\hat{\mathbf{h}}_k$ is given by

$$\begin{aligned} \delta h_k(i) &\triangleq h_k(i) - \hat{h}_k(i) \\ &= \sum_{l=0}^{P-1} s_k(i-l) \delta g_k(l), \end{aligned} \quad (5)$$

from which it follows that the variance of $\delta h_k(i)$ is given by $\Delta_n^2 = P \text{Var}\{\delta g_k(l)\}$.

Fixing $\{\delta g_k(l)\}$ and considering $\{\delta s_k(l)\}$ as random variables, it is easy to show that $\delta h_k(l)$ is asymptotically Gaussian as $P \rightarrow \infty$ by applying the central limit theorem to (5). Due to the assumption that $P/N \ll 1$, for any l , $\delta h_k(l)$ is independent of most $\{\delta h_k(m)\}_{m \neq l}$ since for any $|l-m| > P$, $\delta h_k(l)$ and $\delta h_k(m)$ are mutually independent. Thus, it is reasonable to assume that the elements in $\delta \mathbf{h}_k$ are Gaussian and mutually independent, which substantially simplifies the analysis and will be validated with simulation results in Section 6. Similarly, we can assume that the elements of \mathbf{h}_k are mutually independent as well.

3. BRIEF REVIEW OF REPLICA METHOD

In this section, we give a brief introduction to the replica method, on which the asymptotic analysis in this paper is based. The details can be found in [9, 10, 11, 17].

On assuming $P(b_k = 1) = P(b_k = -1)$, we consider the following ratio:

$$\begin{aligned} &\frac{P(b_k = 1|\mathbf{r})}{P(b_k = -1|\mathbf{r})} \\ &= \frac{\sum_{\{\mathbf{b}|b_k=1\}} \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{r} - (1/\sqrt{N})\mathbf{H}\mathbf{b}\|^2\right)}{\sum_{\{\mathbf{b}|b_k=-1\}} \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{r} - (1/\sqrt{N})\mathbf{H}\mathbf{b}\|^2\right)}, \end{aligned} \quad (6)$$

where σ^2 is a control parameter. Various MUD algorithms can be obtained using this ratio. In particular, we can obtain individually optimal (IO), or maximum a posteriori probability (MAP), MUD ($\sigma^2 = \sigma_n^2$), jointly optimal (JO), or ML, MUD ($\sigma^2 = 0$) and the MF ($\sigma^2 = \infty$).

The key point of the replica method is the computation of the free energy, which is given by

$$\begin{aligned} \mathcal{F}_K(\mathbf{r}, H) &\triangleq K^{-1} \log Z(\mathbf{r}, H) \\ &= \lim_{K \rightarrow \infty} \int_{\mathbb{R}^N} \overline{P(\mathbf{r}|H) \log Z(\mathbf{r}, H)} d\mathbf{r}, \end{aligned} \quad (7)$$

where

$$Z(\mathbf{r}, H) \triangleq \sum_{\{\mathbf{b}\}} P(\mathbf{b}) \exp\left(-\frac{1}{2\sigma^2} \left\| \mathbf{r} - \frac{1}{\sqrt{N}} \mathbf{H}\mathbf{b} \right\|^2\right), \quad (8)$$

and the overbar denotes the average over the randomness of the equivalent spreading codes. It should be noted that the second equation is based on the self-averaging assumption [11].

To evaluate the free energy, we can use the replica method, by which we have

$$\mathcal{F}_K(\mathbf{r}, H) = \lim_{K \rightarrow \infty} \left(\lim_{n_r \rightarrow 0} \frac{\log \Xi_{n_r}}{K} \right), \quad (9)$$

where

$$\Xi_{n_r} = \int_{\mathbf{b}_0, \dots, \mathbf{b}_{n_r}} \prod_{a=0}^{n_r} P(\mathbf{b}_a) \left\{ \frac{1}{\sqrt{2\pi\sigma_n^2}} \int_{\mathbb{R}} \exp\left[-\frac{1}{2\sigma_n^2} \left(r - \frac{1}{\sqrt{N}} \sum_{k=1}^K h_k b_{0k}\right)^2\right] \prod_{a=1}^{n_r} \exp\left[-\frac{1}{2\sigma^2} \left(r - \frac{1}{\sqrt{N}} \sum_{k=1}^K h_k b_{ak}\right)^2\right] dr \right\}^{n_r}, \quad (10)$$

where \mathbf{b}_0 is the same as the \mathbf{b} in (4). However, it is difficult to find an exact physical meaning for $\{\mathbf{b}_a\}_{a=1, \dots, n_r}$. We can roughly consider \mathbf{b}_a to be the a th estimates of the received binary symbols \mathbf{b} .

An assumption, which still lacks rigorous mathematical proof, is proposed in [11], which states that Ξ_{n_r} around $n_r = 0$ can be evaluated by directly using the expression of Ξ_{n_r} obtained for positive integers n_r . With this assumption, we can regard n_r as an integer when evaluating Ξ_{n_r} , and $\{\mathbf{x}_a\}$ as n_r replicas of \mathbf{x} .

To exploit the asymptotic normality of

$$\frac{1}{\sqrt{N}} \sum_{k=1}^K h_k b_{ak}, \quad a = 0, \dots, n_r, \quad (11)$$

we define variables $\{v_a\}_{a=0, \dots, n_r}$ as

$$\begin{aligned} v_0 &= \frac{1}{\sqrt{K}} \sum_{k=1}^K h_k b_{0k}, \\ v_a &= \frac{1}{\sqrt{K}} \sum_{k=1}^K h_k b_{ak}, \quad a = 1, \dots, n_r. \end{aligned} \quad (12)$$

The cross-correlations of $\{v_a\}$ are denoted by parameters $\{Q_{ab}\}$, where $Q_{ab} \triangleq \overline{v_a v_b}$. With these definitions, we can obtain

$$\Xi_{n_r} = \int_{\mathbb{R}} \exp(K\beta^{-1}\mathcal{G}\{Q\}) \mu_K\{Q\} \prod_{a<b} dQ_{ab}, \quad (13)$$

where³

$$\mu_K\{Q\} = \sum_{\mathbf{b}_0, \dots, \mathbf{b}_{n_r}} \prod_{a=0}^{n_r} P(\mathbf{b}_a) \prod_{a<b} \delta(\mathbf{b}_a^H \mathbf{b}_b - KQ_{ab}), \quad (14)$$

and

$$\exp(\mathcal{G}\{Q\}) = \frac{1}{2\pi\sigma_n^2} \int_{\mathbb{R}} \exp\left[-\frac{\beta}{2\sigma_n^2} \left(\frac{r}{\sqrt{\beta}} - v_0\{Q\}\right)^2\right] \times \prod_{a=1}^{n_r} \exp\left[-\frac{\beta}{2\sigma^2} \left(\frac{r}{\sqrt{\beta}} - v_a\{Q\}\right)^2\right] dr + O(K^{-1}). \quad (15)$$

By applying Varadhan's large deviations theorem [18], Ξ_{n_r} converges to the following expression as $K \rightarrow \infty$:

$$\lim_{K \rightarrow \infty} K^{-1} \log \Xi_{n_r} = \sup_{\{Q\}} (\beta^{-1}\mathcal{G}\{Q\} - \mathcal{I}\{Q\}), \quad (16)$$

where $\mathcal{I}\{Q\}$ is the rate function of $\mu_K\{Q\}$, which is based on an optimization over a set of parameters $\{\tilde{Q}_{ab}\}_{a<b}$.

Thus, the evaluation of the free energy $\mathcal{F}_K(\mathbf{r}, H)$ depends on the optimization of (16) over the parameters $\{Q_{ab}\}$ and $\{\tilde{Q}_{ab}\}$, which is computationally prohibitive. This problem is tackled by the assumption of *replica symmetry*; that is, $Q_{0a} = m$, $\tilde{Q}_{0a} = E$, for all $a \neq 0$ and $Q_{ab} = q$, $\tilde{Q}_{ab} = F$, for all $a < b$, $a \neq 0$. Then the optimization of (16) is performed on the parameter set $\{m, q, E, F\}$. The optimal $\{m, q, E, F\}$ are given by solving the following implicit expressions:

$$\begin{aligned} m &= \int_{\mathbb{R}} \tanh(\sqrt{F}z + E) Dz, \\ q &= \int_{\mathbb{R}} \tanh^2(\sqrt{F}z + E) Dz, \\ E &= \frac{\beta^{-1}B}{1 + B(1 - q)}, \\ F &= \frac{\beta^{-1}B^2(B_0^{-1} + 1 - 2m + q)}{(1 + B(1 - q))^2}, \end{aligned} \quad (17)$$

where $Dz = (1/\sqrt{2\pi})e^{-z^2/2}dz$, $B_0 = \beta/\sigma_n^2$, and $B = \beta/\sigma^2$. Then, the performance of MUD can be derived from the free energy, which is determined by m, q, E, F . It is shown in [11] that the bit error rate of MUD is given by

$$P_e = Q\left(\frac{E}{\sqrt{F}}\right), \quad (18)$$

where $Q(z) = \int_z^\infty Dt$ is the complementary Gaussian cumulative distribution function. Thus the multiple access system

is equivalent to a single-user system operating over an AWGN channel with an equivalent signal-to-noise ratio (SNR) E^2/F . The parameters m and q are the first and second moments, respectively, of the soft output, $\hat{b}_k = P(b_k = 1) - P(b_k = -1)$. When $B = B_0$ ($\sigma^2 = \sigma_n^2$), it is easy to check that $m = q$ and $E = F$ using (17).

4. OPTIMAL MUD

In this section, we discuss two types of receivers distinguished by whether or not the receiver considers the distribution of the channel estimation error. We denote the case of *directly* using the channel estimates for MUD by a prefix D, and the case of considering the distribution of the channel estimation error to *compensate* the corresponding impact by a prefix C.

4.1. D-optimal MUD

In this subsection, we discuss the D-optimal MUD, where the receiver applies the channel estimates directly to MUD and does not consider the distribution of the channel estimation error. When the equivalent spreading codes contain errors incurred by the channel estimation error, the corresponding free energy is given by

$$\mathcal{F}_K(\mathbf{r}, \hat{H}) = K^{-1} \log Z(\mathbf{r}, \hat{H}), \quad (19)$$

where \hat{H} is the estimation of channel coefficients H and

$$Z(\mathbf{r}, \hat{H}) \triangleq \sum_{\{\mathbf{b}\}} P(\mathbf{b}) \exp\left(-\frac{1}{2\sigma^2} \left\| \mathbf{r} - \frac{1}{\sqrt{N}} \hat{H} \mathbf{b} \right\|^2\right). \quad (20)$$

We assume that the self-averaging assumption is also valid for $\delta H \triangleq H - \hat{H}$, and thus (7) still holds with the corresponding Ξ_n given by

³ $\delta(x)$ is the Dirac delta function.

$$\Xi_n = \int_{\mathbf{b}_0, \dots, \mathbf{b}_{n_r}} \prod_{a=0}^{n_r} P(\mathbf{b}_a) \times \left\{ \frac{1}{\sqrt{2\pi\sigma_n^2}} \int_{\mathbb{R}} \exp \left[-\frac{1}{2\sigma_n^2} \left(r - \frac{1}{\sqrt{N}} \sum_{k=1}^K h_k b_{0k} \right)^2 \right] \prod_{a=1}^{n_r} \exp \left[-\frac{1}{2\sigma^2} \left(r - \frac{1}{\sqrt{N}} \sum_{k=1}^K \hat{h}_k b_{ak} \right)^2 \right] dr \right\}^N. \quad (21)$$

We can apply the same methodology as in Section 3 to the evaluation of the free energy with imperfect channel estimation. The only difference is that we need to take into account the distribution of the channel estimation error. In a way similar to (12), we define

$$v_a = \frac{1}{\sqrt{K}} \sum_{k=1}^K \hat{h}_k b_{ak}, \quad a = 1, \dots, n_r. \quad (22)$$

For ML channel estimation, δh_k is uncorrelated with h_k , thus resulting in $E\{h_k \hat{h}_k\} = 1$ and $E\{\hat{h}_k \hat{h}_k\} = 1 + \Delta_h^2$. Then we have

$$\begin{aligned} \overline{v_0 v_a} &= \frac{1}{K} \sum_{k=1}^K b_{0k} b_{ak} \quad \forall a > 0, \\ \overline{v_a v_b} &= \frac{1 + \Delta_h^2}{K} \sum_{k=1}^K b_{ak} b_{bk} \quad \forall a, b > 0. \end{aligned} \quad (23)$$

For MMSE channel estimation, δh_k is uncorrelated with \hat{h}_k , thus resulting in $E\{h_k \hat{h}_k\} = E\{\hat{h}_k^2\} = 1 - \Delta_h^2$. Then we have

$$\begin{aligned} \overline{v_0 v_a} &= \frac{1 - \Delta_h^2}{K} \sum_{k=1}^K b_{0k} b_{ak} \quad \forall a > 0, \\ \overline{v_a v_b} &= \frac{1 - \Delta_h^2}{K} \sum_{k=1}^K b_{ak} b_{bk} \quad \forall a, b > 0. \end{aligned} \quad (24)$$

Thus, the free energy with imprecise channel estimation still depends on the same parameter set $\{m, q, E, F\}$ as in Section 3. An important observation is that the existence of $\{\delta h_k\}$ affects only the term $\mathcal{G}\{Q\}$ in (13), and $\mu_K\{Q\}$ remains unchanged, which implies that the expressions for m and q are identical to those in (17). Hence, we can focus on only the computation of $\mathcal{G}\{Q\}$. By supposing that the assumption of replica symmetry is still valid, the asymptotically Gaussian random variables v_0 and v_a can be constructed using expressions similar to those in [11]. For ML channel estimation, we have

$$\begin{aligned} v_0 &= u \sqrt{1 - \frac{m^2}{(1 + \Delta_h^2)q}} - t \frac{m}{\sqrt{(1 + \Delta_h^2)q}}, \\ v_a &= \sqrt{1 + \Delta_h^2} (z_a \sqrt{1 - q} - t \sqrt{q}), \quad a = 1, \dots, n_r, \end{aligned} \quad (25)$$

where u , t , and $\{z_a\}$ are mutually independent Gaussian random variables with zero mean and unit variance.

With the same definitions of u , t , and $\{z_a\}$, for MMSE channel estimation, we have

$$v_0 = u \sqrt{1 - \frac{(1 - \Delta_h^2)m^2}{q}} - t \frac{m \sqrt{1 - \Delta_h^2}}{\sqrt{q}}, \quad (26)$$

$$v_a = \sqrt{1 - \Delta_h^2} (z_a \sqrt{1 - q} - t \sqrt{q}), \quad a = 1, \dots, n_r.$$

Substituting the above expressions into (13), we can obtain the following conclusions using some calculus similar to that of [11]. For ML channel estimation, the free energy is given by

$$\begin{aligned} \mathcal{F}_K(\mathbf{r}, \hat{H}) &= \int_{\mathbb{R}} \log(\cosh(\sqrt{F}z + E)) Dz - Em - \frac{F(1-q)}{2} \\ &\quad - \frac{1}{2\beta} \left(\log(1 + (1 + \Delta_h^2)(1-q)B) \right. \\ &\quad \left. + \frac{B(B_0^{-1} + 1 - 2m + (1 + \Delta_h^2)q)}{1 + B(1-q)(1 + \Delta_h^2)} \right). \end{aligned} \quad (27)$$

The corresponding E and F are given by

$$\begin{aligned} E &= \frac{\beta^{-1}B}{1 + B(1-q)(1 + \Delta_h^2)}, \\ F &= \frac{(1 + \Delta_h^2)\beta^{-1}B^2(B_0^{-1} + 1 - 2m + (1 + \Delta_h^2)q)}{(1 + B(1-q)(1 + \Delta_h^2))^2}. \end{aligned} \quad (28)$$

For MMSE channel estimation, we can obtain

$$\begin{aligned} \mathcal{F}_K(\mathbf{r}, \hat{H}) &= \int_{\mathbb{R}} \log(\cosh(\sqrt{F}z + E)) Dz - Em - \frac{F(1-q)}{2} \\ &\quad - \frac{1}{2\beta} \left(\log(1 + (1 - \Delta_h^2)(1-q)B) \right. \\ &\quad \left. + \frac{B(B_0^{-1} + 1 - (1 - \Delta_h^2)(2m - q))}{1 + B(1-q)(1 - \Delta_h^2)} \right), \end{aligned} \quad (29)$$

and the corresponding E and F are given by

$$\begin{aligned} E &= \frac{\beta^{-1}B(1 - \Delta_h^2)}{1 + B(1-q)(1 - \Delta_h^2)}, \\ F &= \frac{\beta^{-1}B^2(1 - \Delta_h^2)(B_0^{-1} + 1 - (1 - \Delta_h^2)(2m - q))}{(1 + B(1-q)(1 - \Delta_h^2))^2}. \end{aligned} \quad (30)$$

The corresponding output signal-to-interference-plus-noise-ratios (SINRs) of the ML and MMSE channel estimation are given by the following expressions, respectively:

$$\text{SINR}_{\text{ML}} = \frac{1}{(1 + \Delta_h^2)} \frac{1}{(\sigma_n^2 + \beta(1 - 2m + (1 + \Delta_h^2)q))}, \quad (31)$$

$$\text{SINR}_{\text{MMSE}} = \frac{1 - \Delta_h^2}{(\sigma_n^2 + \beta(1 - (1 - \Delta_h^2)(2m - q)))}. \quad (32)$$

Thus, we can summarize the impact of the channel estimation error on the D-optimal MUD as follows.

- (i) The factors $1/(1 + \Delta_h^2)$ in (31) and $1 - \Delta_h^2$ in the numerator of (32) represent the impact of the error of the desired user's equivalent spreading codes, which is equivalent to increasing the noise level.
- (ii) The imperfect channel estimation also increases the variance of the residual MAI, which equals $\beta(1 - 2m + (1 + \Delta_h^2)q)$ for ML channel estimation-based systems and $\beta(1 - (1 - \Delta_h^2)(2m - q))$ for MMSE channel estimation-based systems.
- (iii) The equations $m = q$ and $E = F$ are no longer valid when $\sigma^2 = \sigma_n^2$. Thus, there are no simple analytical expressions for obtaining the multiuser efficiency in a way similar to the Tse-Hanly equation [8].

4.2. C-optimal MUD

In this subsection, we consider the C-optimal MUD, where the distribution of the channel estimation error is exploited to compensate for the imperfection of channel estimation. For simplicity, we consider only the IO MUD (C-IO MUD).

4.2.1. ML channel estimation

When deriving the expressions of C-IO MUD, we consider a fixed chip period and drop the index of the chip period for simplicity. The conditional probability $P(\{h_k\} | \{\hat{h}_k\})$ should be taken into account to attain the optimal detection. Thus, the a posteriori probability of the received signal r at this chip period, conditioned on the channel estimates $\{\hat{h}_k\}$ and the transmitted symbols $\{b_k\}$, is given by

$$P(r | \{\hat{h}_k\}, \{b_k\}) \propto \int_{\mathbb{R}^K} P(r | \{h_k\}, \{b_k\}) P(\{h_k\} | \{\hat{h}_k\}) \prod_{k=1}^K dh_k, \quad (33)$$

where

$$P(\{h_k\} | \{\hat{h}_k\}) = \prod_{k=1}^K P(h_k | \hat{h}_k), \quad (34)$$

$$P(h_k | \hat{h}_k) \propto \exp\left(-\frac{(h_k - \hat{h}_k)^2}{2\Delta_h^2}\right) \exp\left(-\frac{h_k^2}{2}\right).$$

It should be noted that the above two expressions are based on the assumption of normality and mutual independence of $\{\delta h_k\}$ in Section 2.2. Then we have

$$P(r | \{\hat{h}_k\}, \{b_k\}) \propto \int_{\mathbb{R}^K} \exp\left(-\frac{(r - (1/\sqrt{N}) \sum_{k=1}^K h_k b_k)^2}{2\sigma_n^2}\right) \times \prod_{k=1}^K p(h_k | \hat{h}_k) dh_k. \quad (35)$$

Let $r_1 = r - (1/\sqrt{N}) \sum_{k=2}^K h_k b_k$, then the integral with respect to h_1 is given by

$$\int_{\mathbb{R}} \exp\left(-\frac{(r_1 - (1/\sqrt{N})h_1 b_1)^2}{2\sigma_n^2}\right) \exp\left(-\frac{(h_1 - \hat{h}_1)^2}{2\Delta_h^2}\right) \times \exp\left(-\frac{h_1^2}{2}\right) dh_1 \propto \exp\left(-\frac{(r_1 - b_1 \hat{h}_1 / \sqrt{N} (1 + \Delta_h^2))^2}{2(\sigma_n^2 + \Delta_h^2 / (1 + \Delta_h^2) N)}\right), \quad (36)$$

where the factors common for different $\{b_k\}$ are ignored for simplicity.

Applying the same procedure for h_2, \dots, h_K , we obtain that

$$P(r | \{\hat{h}_k\}, \{b_k\}) \propto \exp\left(-\frac{(r - (1/\sqrt{N}(1 + \Delta_h^2)) \sum_{k=1}^K b_k \hat{h}_k)^2}{2(\sigma_n^2 + \beta \Delta_h^2 / (1 + \Delta_h^2))}\right). \quad (37)$$

Thus the LR of IO MUD is given by

$$\frac{P(b_k = 1 | \mathbf{r})}{P(b_k = -1 | \mathbf{r})} = \frac{\sum_{\{\mathbf{b}|b_k=1\}} \exp\left(-\frac{(1/2\sigma^2) \|\mathbf{r} - (1/\sqrt{N}(1 + \Delta_h^2)) \hat{H} \mathbf{b}\|^2}{2}\right)}{\sum_{\{\mathbf{b}|b_k=-1\}} \exp\left(-\frac{(1/2\sigma^2) \|\mathbf{r} - (1/\sqrt{N}(1 + \Delta_h^2)) \hat{H} \mathbf{b}\|^2}{2}\right)}, \quad (38)$$

where $\sigma^2 = \sigma_n^2 + \beta \Delta_h^2 / (1 + \Delta_h^2)$. Therefore, the channel estimation error is compensated for merely by changing the equivalent noise variance and scaling the channel estimate with a factor of $1/(1 + \Delta_h^2)$.

Similarly to the analysis in Section 4.1, we can define

$$v_0 = u \sqrt{1 - \frac{m^2}{(1 + \Delta_h^2)q}} - t \frac{m}{\sqrt{(1 + \Delta_h^2)q}}, \quad (39)$$

$$v_a = \frac{1}{\sqrt{1 + \Delta_h^2}} (z_a \sqrt{1 - q} - t \sqrt{q}), \quad a = 1, \dots, n_r.$$

Then we can obtain the free energy, which is given by

$$\begin{aligned} \mathcal{F}_K(\mathbf{r}, \hat{H}) &= \int_{\mathbb{R}} \log(\cosh(\sqrt{F}z + E)) Dz - Em - \frac{F(1-q)}{2} \\ &\quad - \frac{1}{2\beta} \left(\log \left(1 + \frac{B(1-q)}{(1 + \Delta_h^2)} \right) \right. \\ &\quad \left. + \frac{B((B_0^{-1} + 1)(1 + \Delta_h^2) - 2m + q)}{1 + \Delta_h^2 + B(1-q)} \right), \end{aligned} \quad (40)$$

where $B = \beta/(\sigma_n^2 + \beta\Delta_h^2/(1 + \Delta_h^2))$. The corresponding E and F are given by

$$\begin{aligned} E &= \frac{\beta^{-1}B_0}{1 + \Delta_h^2 + B_0(1 + \Delta_h^2 - q)}, \\ F &= \frac{\beta^{-1}B_0^2((B_0^{-1} + 1)(1 + \Delta_h^2) - 2m + q)}{(1 + \Delta_h^2 + B_0(1 + \Delta_h^2 - q))^2}. \end{aligned} \quad (41)$$

An interesting observation is that the equations $m = q$ and $E = F$ are recovered in this case. Also we can obtain the equivalent SINR, which is given by

$$\text{SINR}_{\text{ML}} = \frac{1}{\sigma_n^2(1 + \Delta_h^2) + \beta\Delta_h^2 + \beta(1 - q)}. \quad (42)$$

The corresponding multiuser efficiency η is given by solving the following Tse-Hanly style equation:

$$\frac{1}{\eta} + \frac{\beta}{\sigma_n^2} \int_{\mathbb{R}} \tanh^2 \left(\sqrt{\frac{\eta}{\sigma_n^2}} z + \frac{\eta}{\sigma_n^2} \right) Dz = (1 + \Delta_h^2) \left(1 + \frac{\beta}{\sigma_n^2} \right). \quad (43)$$

From (42), we can see that the impact of channel estimation error consists of three aspects, which are represented by the three terms in the denominator of the expression (42). The term $\sigma_n^2(1 + \Delta_h^2)$ embodies the negative impact of the channel estimation error on the user being detected, which causes uncertainty in the equivalent spreading codes of this user and is equivalent to scaling the noise by a factor of $(1 + \Delta_h^2)$. Besides implicitly affecting the parameter q in the third term, the channel estimation error of the interfering users also results in the term of $\beta\Delta_h^2$; an intuitive explanation for this is that, since the output of IO MUD can be regarded as the output of an interference canceller using the conditional mean estimates of all other users [10], the channel estimation error causes imperfection in the reconstruction of the signals of the other users and the variance of residual interference equals $\beta\Delta_h^2$ when the decision feedback is free of errors. The corresponding equivalent channel model is illustrated in Figure 1.

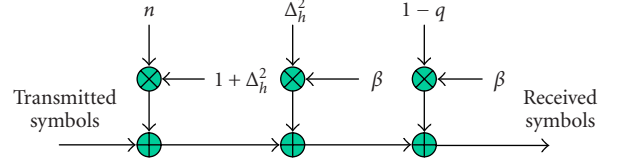


FIGURE 1: Bit error rate of D-IO MUD as a function of channel estimation error variance.

4.2.2. MMSE channel estimation

For MMSE channel estimation, the channel estimation error δh_k is uncorrelated with the estimate \hat{h}_k . Thus, we have

$$\begin{aligned} P(h_k | \hat{h}_k) &= P(\delta h_k + \hat{h}_k | \hat{h}_k) \\ &\propto \exp \left(- \frac{(h_k - \hat{h}_k)^2}{2\Delta_h^2} \right). \end{aligned} \quad (44)$$

Applying the same procedure as ML channel estimation, we can obtain the LR of IO MUD, which is given by

$$\begin{aligned} \frac{P(b_k = 1 | \mathbf{r})}{P(b_k = -1 | \mathbf{r})} &= \frac{\sum_{\{\mathbf{b}|b_k=1\}} \exp \left(- (1/2\sigma^2) \|\mathbf{r} - (1/\sqrt{N})\hat{H}\mathbf{b}\|^2 \right)}{\sum_{\{\mathbf{b}|b_k=-1\}} \exp \left(- (1/2\sigma^2) \|\mathbf{r} - (1/\sqrt{N})\hat{H}\mathbf{b}\|^2 \right)}, \end{aligned} \quad (45)$$

where the control parameter, or equivalent noise power, $\sigma^2 = \sigma_n^2 + \beta\Delta_h^2$. Substituting $B = \beta/(\sigma_n^2 + \beta\Delta_h^2)$ into (30), we have

$$\begin{aligned} E &= \frac{\beta^{-1}B_0(1 - \Delta_h^2)}{1 + B_0(1 - (1 - \Delta_h^2)q)}, \\ F &= \frac{\beta^{-1}B_0^2(1 - \Delta_h^2)(B_0^{-1} - (2m - q)(1 - \Delta_h^2))}{(1 + B_0(1 - (1 - \Delta_h^2)q))^2}. \end{aligned} \quad (46)$$

Similarly to the case of ML channel estimation, the equations $m = q$ and $E = F$ are recovered as well. The equivalent output SINR is given by

$$\text{SINR}_{\text{MMSE}} = \frac{1 - \Delta_h^2}{\sigma_n^2 + \beta(1 - (1 - \Delta_h^2)q)}, \quad (47)$$

and the corresponding multiuser efficiency is given by solving the following equation:

$$\frac{1}{\eta} + \frac{\beta}{\sigma_n^2} \int_{\mathbb{R}} \tanh^2 \left(\sqrt{\frac{\eta}{\sigma_n^2}} z + \frac{\eta}{\sigma_n^2} \right) Dz = \frac{1 + \beta/\sigma_n^2}{1 - \Delta_h^2}. \quad (48)$$

The intuition behind (47) is similar to that of ML channel estimation. On comparing (43) and (48), an immediate conclusion is that the C-IO MUD is more susceptible to the error incurred by MMSE channel estimation than that incurred by ML estimation, when Δ_h^2 is identical for both estimators.

5. LINEAR MUD AND TURBO MUD

We now turn to the consideration of linear and turbo multiuser detection. For simplicity, we discuss only ML channel estimation-based systems in this section. MMSE channel estimation-based systems can be analyzed in a similar way.

5.1. Linear MUD

The analysis of linear MUD can be incorporated into the framework of the replica method (for MMSE MUD, $\sigma^2 = \sigma_n^2$; for the decorrelator, $\sigma^2 \rightarrow 0$) by merely regarding the channel symbols as Gaussian-distributed random variables. The system performance is determined by the parameter set $\{m, q, p, E, F, G\}$ and a group of saddle-point equations [11].

Particularly, when $\sigma^2 = \sigma_n^2$ (MMSE MUD), the parameters can be simplified to $\{q, E\}$, which satisfy $q = E/(1 + E)$ and $E = \beta^{-1}B_0/(1 + B_0(1 - q))$. The multiuser efficiency is determined by the Tse-Hanly equation [8].

5.1.1. D-MMSE MUD

Since the channel estimation error does not affect $\mathcal{I}\{Q\}$, the parameters m, q , and p are unchanged. With the same manipulation on $\mathcal{G}\{Q\}$ as in Section 4, we can obtain the parameters E, F , and G as follows:

$$\begin{aligned} E &= \frac{\beta^{-1}B}{1 + B(p - q)(1 + \Delta_h^2)}, \\ F &= \frac{(1 + \Delta_h^2)\beta^{-1}B^2(B_0^{-1} + 1 - 2m + (1 + \Delta_h^2)q)}{(1 + B(p - q)(1 + \Delta_h^2))^2}, \\ G &= F - (1 + \Delta_h^2)E. \end{aligned} \quad (49)$$

5.1.2. C-MMSE MUD

Similarly to Section 4, the MMSE detector considering the distribution of the channel estimation error is given by merely scaling \hat{H} with a factor of $1/(1 + \Delta_h^2)$ and changing σ^2 to $\sigma_n^2 + \beta\Delta_h^2/(1 + \Delta_h^2)$. Then, we have $E = F, G = 0, m = q$, and $p = 0$. The corresponding multiuser efficiency is given implicitly by

$$\left(1 + \Delta_h^2 + \frac{\beta\Delta_h^2}{\sigma_n^2}\right)\eta + \frac{\beta\eta}{\sigma_n^2 + \eta} = 1. \quad (50)$$

5.2. Turbo MUD

5.2.1. Optimal turbo MUD

For optimal turbo MUD [4], since the channel estimation error does not affect $\mathcal{I}\{Q\}$ when evaluating the free energy, the impact of channel estimation error is similar to the optimal MUD in Section 4, namely, the corresponding saddle-point equations remain the same as in [12] except that the parameters E and F are changed in the same way as in (28) and (41).

5.2.2. MMSE filter-based PIC

However, greater complications arise in the case of MMSE filter-based PIC [4], where the MAI is cancelled with the decision feedback from channel decoders and the residual MAI

is further suppressed with an MMSE filter. The corresponding MMSE filter is constructed with the estimated equivalent spreading codes $\{\hat{\mathbf{h}}_k\}$ and the estimated power of the residual interference. In an unconditional MMSE filter, the power estimate is given by $\Delta_b^2 \triangleq E\{(b_k - \hat{b}_k)^2\}$, where \hat{b}_k is the soft decision feedback; and in a conditional MMSE filter, the power estimate is given by $1 - \hat{b}_k^2$. However, this power estimate for user k is different from the true value $|b_k - \hat{b}_k|^2$ since b_k is unknown to the receiver, thus making the filter unmatched for the MAI. Hence, the analysis in [12] may overestimate the system performance since such power estimation errors are not considered there. Thus we need to take into account the corresponding power mismatch. For simplicity, we consider only unbiased power estimation. Note that this scenario can be applied to general cases where the received signal power is not perfectly estimated.

For the MMSE filter-based PIC, the powers of the residual interference are different for different users. Similarly to the analysis of unequal-power systems in [17], we can divide the users into a finite number (L) of equal-power groups, with power $\{P_l\}_{l=1,\dots,L}$, estimated power $\{\hat{P}_l\}_{l=1,\dots,L}$, and the corresponding proportion $\{\alpha_l\}_{l=1,\dots,L}$, and obtain the results for any arbitrary user power distribution by letting $L \rightarrow \infty$. Confining our discussion to unbiased MAI power estimation, we normalize the MAI power such that $\sum_{l=1}^L \alpha_l P_l = 1$ and $\sum_{l=1}^L \alpha_l \hat{P}_l = 1$. The equivalent noise variance is given by $\sigma^2 = \sigma_n^2/\Delta_b^2$. Thus, the bit error rate of MUD is given by $Q(E/\sqrt{F\Delta_b^2})$ since the power of the desired user is unity.

Similarly to the previous analysis, we define

$$\begin{aligned} v_0 &= \frac{1}{\sqrt{K}} \sum_{l=1}^L \sqrt{P_k} \sum_{k \in C_l} h_k b_{0k}, \\ v_a &= \frac{1}{\sqrt{K}} \sum_{l=1}^L \sqrt{\hat{P}_k} \sum_{k \in C_l} \hat{h}_k b_{ak}, \quad a = 1, \dots, n_r, \end{aligned} \quad (51)$$

where C_l represents the set of users with power P_l . We can see that the uneven and mismatched power distribution does not affect the analysis of $\exp(\mathcal{G}\{Q\})$, which incorporates the impact of channel estimation error. However, the rate function $\mathcal{I}\{Q\}$ is changed to

$$\mathcal{I}\{Q\} = \sup_{\{\tilde{Q}\}} \left(\sum_{a \leq b} \tilde{Q}_{ab} Q_{ab} - \sum_{l=1}^L \alpha_l \log M_{\{l\}}^G(\tilde{Q}) \right), \quad (52)$$

where

$$\begin{aligned} M_{\{l\}}^G(\tilde{Q}) &= \frac{1}{2} \int_{\mathbb{R}^{n_r}} \exp \left(\sqrt{P_l \hat{P}_l} E b_0 \sum_{a=1}^{n_r} b_a + \hat{P}_l F \sum_{a < b} b_a b_b \right. \\ &\quad \left. + \frac{G \hat{P}_l}{2} \sum_{a=1}^{n_r} b_a^2 \right) \prod_{a=1}^n D b_a, \end{aligned} \quad (53)$$

in which $\{b_a\}_{a=1,\dots,n_r}$ are Gaussian random variables. Similarly to [17], after some algebra, we can obtain the free

energy, which is given by

$$\begin{aligned} \mathcal{F}_K(\mathbf{r}, \hat{H}) &= \frac{1}{2} \sum_{l=1}^L \alpha_l \left(\log(1 + (F - G)\hat{P}_l) - \frac{\hat{P}_l F + P_l \hat{P}_l E^2}{1 + (F - G)\hat{P}_l} \right) \\ &+ Em - \frac{1}{2} Fq + \frac{1}{2} Gp \\ &- \frac{1}{2\beta} \left(\log(1 + (1 + \Delta_h^2)(p - q)B) \right. \\ &\quad \left. + \frac{B(B_0^{-1} + 1 - 2m + (1 + \Delta_h^2)q)}{1 + B(p - q)(1 + \Delta_h^2)} \right). \end{aligned} \quad (54)$$

Letting $L \rightarrow \infty$, we can obtain that

$$\begin{aligned} m &= E \left\{ \frac{P\hat{P}E}{1 + \hat{P}(F - G)} \right\}, \\ q &= E \left\{ \frac{\hat{P}^2(PE^2 + F)}{(1 + \hat{P}(F - G))^2} \right\}, \\ p &= E \left\{ \frac{\hat{P}(\hat{P}PE^2 + 2\hat{P}F + 1 - \hat{P}G)}{(1 + \hat{P}(F - G))^2} \right\}, \end{aligned} \quad (55)$$

where the expectation is with respect to the joint distribution of P and \hat{P} .

For the unconditional MMSE filter, the expressions for m , q , and p can be simplified to the following expressions, since $\hat{P} = E\{P\} = \Delta_b^2$:

$$\begin{aligned} m &= \frac{(\Delta_b^2)^2 E}{1 + \Delta_b^2(F - G)}, \\ q &= \frac{(\Delta_b^2)^2 (\Delta_b^2 E^2 + F)}{(1 + \Delta_b^2(F - G))^2}, \\ p &= \frac{\Delta_b^2 \left((\Delta_b^2)^2 E^2 + 2\Delta_b^2 F + 1 - \Delta_b^2 G \right)}{(1 + \Delta_b^2(F - G))^2}. \end{aligned} \quad (56)$$

This implies the interesting conclusion that if the MMSE MUD based receiver regards the received powers of different users as being equal to the average received power, the multiuser efficiency will be identical to that of the corresponding equal-power system. It should be noted that the corresponding bit error rates are different although the multiuser efficiencies are the same. Thus, the analysis of the unconditional MMSE filter-based PIC in [12] yields correct results. It should be noted that, for IO MUD with binary channel symbols, this conclusion does not hold since the expressions for m , q , and p are nonlinear in P .

This conclusion can also be applied to frequency-flat fading channels. When the received power is perfectly known, the multiuser efficiency of MMSE MUD is given by

$$\eta + E \left\{ \frac{\beta P \eta}{\sigma_n^2 + P \eta} \right\} = 1, \quad (57)$$

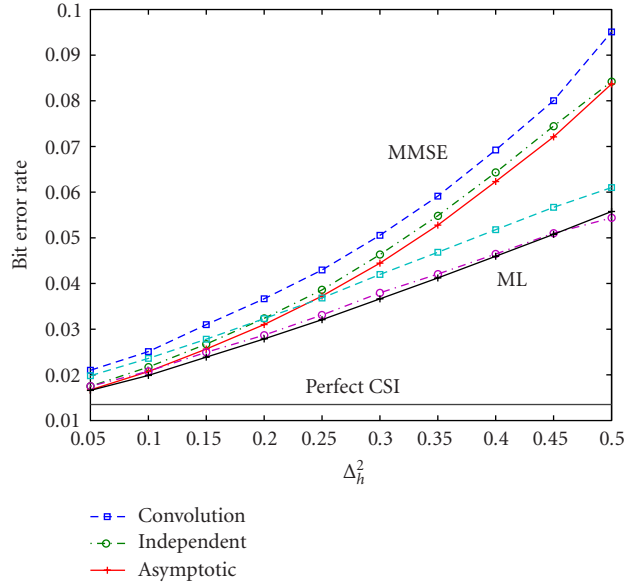


FIGURE 2: Bit error rate of D-IO MUD as a function of channel estimation error variance.

where the random variable P is the received power and the expectation is with respect to the distribution of P . When the receiver is unaware of the fading and uses equal-power MMSE MUD, the multiuser efficiency of this power-mismatched MMSE MUD is given by that of an equal-power system:

$$\eta + \frac{\beta E\{P\} \eta}{\sigma_n^2 + E\{P\} \eta} = 1. \quad (58)$$

Comparing (57) and (58) and applying the fact that, for any positive random variable x , $E\{x/(1+x)\} \leq E\{x\}/(1+E\{x\})$, we can see that this power mismatch incurs a loss in multiuser efficiency.

6. SIMULATION RESULTS

In this section, we provide simulation results to verify and illustrate the analysis of the preceding sections.

Figure 2 shows the bit error rates versus the variance of the channel estimation error for a D-IO MUD system with $K = 10$, $N = 150$, $P = 50$, and $\sigma_n^2 = 0.2$. In this figure, “independent” represents the case of equivalent spreading codes with mutually independent elements and “convolution” represents the case in which the equivalent spreading codes are the convolutions of binary spreading codes and channel gains. From this figure, we can see that the assumption of independent elements in the equivalent spreading codes appears to be valid and the asymptotic results can predict the performance of finite systems fairly well. This figure also shows that D-IO MUD is more susceptible to the error of MMSE channel estimation than that of ML estimation.

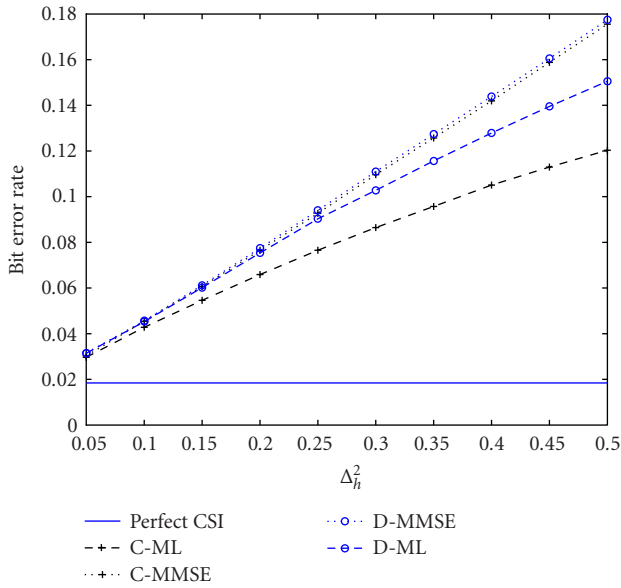


FIGURE 3: Bit error rate of C-IO MUD as a function of channel estimation error variance.

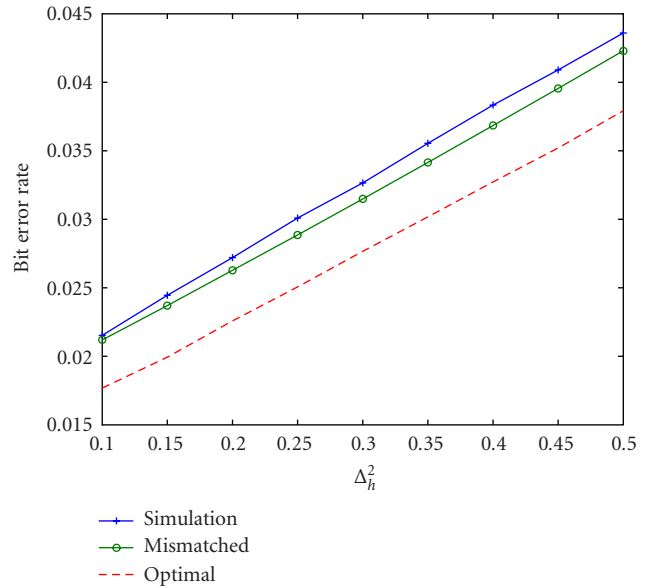


FIGURE 5: Bit error rate of MMSE filter-based PIC as a function of channel estimation error variance.

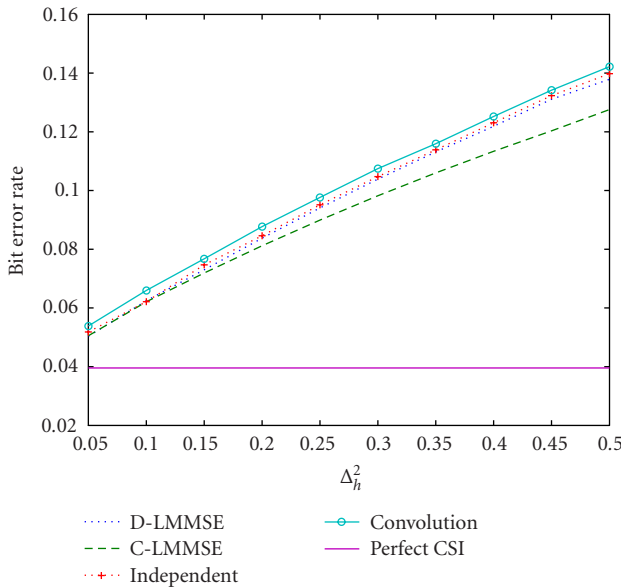


FIGURE 4: Bit error rate of MMSE MUD as a function of channel estimation error variance.

Figure 3 compares the bit error rates in D-IO and C-IO MUD systems with $\beta = 0.5$ and $\sigma_h^2 = 0.2$. For ML channel estimation, the C-IO MUD achieves considerably better performance than the D-IO MUD. For MMSE channel estimation, the two IO MUD schemes attain almost the same performance.

Figure 4 shows the bit error rates for MMSE MUD systems with the same configuration as in Figure 3. Both the numerical simulations (for both independent and convolution models of the equivalent spreading codes) and asymptotic

results are given for D-MMSE MUD, and match fairly well. Note that C-MMSE MUD achieves marginally better performance than D-MMSE MUD.

Figure 5 shows the bit error rates of MMSE filter-based PIC systems with the same configurations as in Figure 4. The decision feedback is from the channel decoder of a convolutional code $(23, 33, 37)_8$ when the input SINR is 3 dB. In this figure, the theoretical and simulation results for the unconditional MMSE filter are represented with “mismatched” and “simulation,” respectively; the results with the assumption that the residual interference power is known are represented by “optimal.” We can observe that the optimal scheme, which assumes that the decision feedback error is known, achieves only marginally better performance.

For Rayleigh flat-fading channels, the multiuser efficiency, obtained by numerical simulations, versus SNR is given in Figure 6. In this figure, “equal power” means the case of equal received power. For the case of Rayleigh-distributed received power, the results of mismatched (regarding the received power as being equal) MMSE MUD and optimal (the received powers are known) MMSE MUD are represented by “Rayleigh-mismatch” and “Rayleigh,” respectively. We can see that the numerical results verify our conclusion about the power-mismatched MMSE MUD in Section 5.2. Also, the knowledge of received power provides marginal improvement in multiuser efficiency.

In Figure 7, we apply the results for C-MMSE MUD to obtain the optimal proportion α of training symbols, versus the coherence time M (measured in symbol periods) and system load β , to maximize the spectral efficiency given by $(1 - \alpha) \log(1 + \eta \text{SNR})$, where $\text{SNR} = 5$ dB, η is determined by (50), and $\Delta_h^2 = \sigma_h^2 / \alpha M$. We can see that the required proportion of training data increases with the system load and decreases with the coherence time.

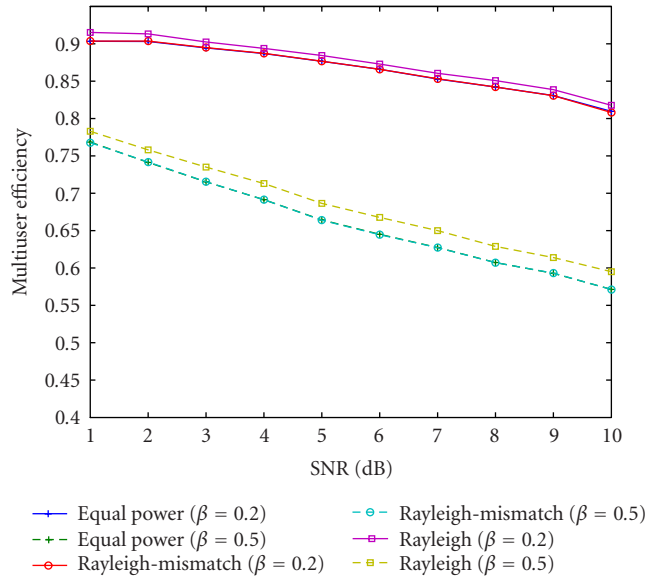


FIGURE 6: Multiuser efficiency versus SNR for nonfading and Rayleigh fading systems.

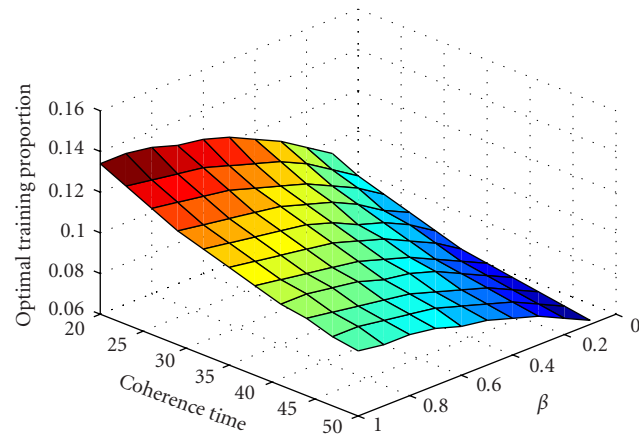


FIGURE 7: Optimal proportion of training symbols versus coherence times and system load.

7. CONCLUSIONS

In this paper, we have discussed the impact of channel estimation error on various types of MUD algorithms in DS-CDMA systems by obtaining asymptotic expressions for the system performance in terms of the channel estimation error variance. The analysis is unified under the framework of the replica method. The following conclusions are of particular interest.

- (i) The performance of MUD is more susceptible to MMSE channel estimation errors than ML ones.
- (ii) The MUD schemes that consider the distribution of channel estimation errors can improve the system performance, considerably for ML channel estimation errors and marginally for MMSE channel estimation errors.

- (iii) When the MMSE MUD treats different users as being received with equal power, it attains the same multiuser efficiency as the corresponding equal-power systems.

ACKNOWLEDGMENTS

This research was supported in part by the Office of Naval Research under Grant N00014-03-1-0102 and in part by the New Jersey Center for Wireless Telecommunications. This paper was presented in part at the 2004 IEEE Global Telecommunications Conference, Dallas, Tex, November 29–December 3, 2004.

REFERENCES

- [1] S. Verdú, *Multiuser Detection*, Cambridge University Press, Cambridge, UK, 1998.
- [2] S. Verdú, *Optimal Multi-user Signal Detection*, Ph.D. thesis, University of Illinois at Urbana-Champaign, Urbana, USA, August 1984.
- [3] R. Lupas and S. Verdú, “Linear multiuser detectors for synchronous code-division multiple-access channels,” *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 123–136, 1989.
- [4] X. Wang and H. V. Poor, “Iterative (turbo) soft interference cancellation and decoding for coded CDMA,” *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046–1061, 1999.
- [5] P. Alexander and A. Grant, “Iterative detection in code-division multiple-access with error control coding,” *European Trans. Telecommunications*, vol. 9, no. 5, pp. 419–426, 1998.
- [6] S. Verdú and S. Shamai, “Spectral efficiency of CDMA with random spreading,” *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 622–640, 1999.
- [7] D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free Random Variables*, vol. 1 of *CRM Monograph Series*, American Mathematical Society, Providence, RI, USA, 1992.
- [8] D. N. C. Tse and S. Hanly, “Linear multiuser receivers: effective interference, effective bandwidth and user capacity,” *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 641–657, 1999.
- [9] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing*, Oxford University Press, Oxford, UK, 2001.
- [10] D. Guo and S. Verdú, “Spectral efficiency of large-system CDMA via statistical physics,” in *Proc. Conference on Information Sciences and Systems*, The Johns Hopkins University, Baltimore, Md, USA, March 2003.
- [11] T. Tanaka, “A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors,” *IEEE Trans. Inform. Theory*, vol. 48, no. 11, pp. 2888–2910, 2002.
- [12] G. Caire, R. Müller, and T. Tanaka, “Iterative multiuser joint decoding: optimal power allocation and low-complexity implementation,” *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1950–1973, 2004.
- [13] J. Evans and D. N. C. Tse, “Large system performance of linear multiuser receivers in multipath fading channels,” *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 2059–2078, 2000.
- [14] H. Li and H. V. Poor, “Impact of imperfect channel estimation on turbo multiuser detection in DS-CDMA systems,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC ’04)*, pp. 30–35, Atlanta, Ga, USA, March 2004.
- [15] Z. Xu, “Effects of imperfect blind channel estimation on performance of linear CDMA receivers,” *IEEE Trans. Signal Processing*, vol. 52, no. 10, pp. 2873–2884, 2004.

- [16] H. Li and H. V. Poor, "Performance of channel estimation in long code DS-CDMA with and without decision feedback," in *Proc. Conference on Information Sciences and Systems*, The Johns Hopkins University, Baltimore, Md, USA, March 2003.
- [17] D. Guo and S. Verdú, "Multiuser detection and statistical mechanics," in *Communications, Information and Network Security*, V. Bhargava, H. V. Poor, V. Tarokh, and S. Yoon, Eds., chapter 13, pp. 229–277, Kluwer Academic Publishers, Norwell, Mass, USA, 2002.
- [18] F. den Hollander, *Large Deviations*, vol. 14 of *Fields Institute Monographs*, American Mathematical Society, Providence, RI, USA, 2000.

Husheng Li received the B.S. and M.S. degrees in electronics engineering from Tsinghua University, Beijing, China, in 1998 and 2000, respectively, and the Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ, in 2005. In 2005, he joined Qualcomm, San Diego, Calif. His research interests include statistical signal processing, wireless communication, and information theory.



H. Vincent Poor received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the faculty at Princeton, where he is the George Van Ness Lothrop Professor in engineering. He has also held visiting appointments at a number of universities, including recently Imperial College, Stanford, and Harvard.



Dr. Poor's research interests are in the areas of wireless networks, advanced signal processing, and related fields. Among his publications in these areas is the recent book *Wireless Networks: Multiuser Detection in Cross-Layer Design*, Springer, 2005. Dr. Poor is a Member of the National Academy of Engineering, and is a Fellow of the IEEE, the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. He is a past President of the IEEE Information Theory Society, and is the current Editor-in-Chief of the IEEE Transactions on Information Theory. Recent recognition of his work includes the Joint Paper Award of the IEEE Communications and Information Theory Societies (2001), the NSF Director's Award for Distinguished Teaching Scholars (2002), a Guggenheim Fellowship (2002–2003), and the IEEE Education Medal (2005).

Factor-Graph-Based Soft Self-Iterative Equalizer for Multipath Channels

Ben Lu

*Silicon Laboratories, Inc., Austin, TX 78735, USA
Email: ben.lu@silabs.com*

Guosen Yue

*NEC Laboratories America, Inc., Princeton, NJ 08540, USA
Email: yueg@nec-labs.com*

Xiaodong Wang

*Department of Electrical Engineering, Columbia University, New York, NY 10027, USA
Email: wangx@ee.columbia.edu*

Mohammad Madihian

*NEC Laboratories America, Inc., Princeton, NJ 08540, USA
Email: madihian@nec-labs.com*

Received 30 April 2004; Revised 23 August 2004

We consider factor-graph-based soft self-iterative equalization in wireless multipath channels. Since factor graphs are able to characterize multipath channels to per-path level, the corresponding soft self-iterative equalizer possesses reduced computational complexity in sparse multipath channels. The performance of the considered self-iterative equalizer is analyzed in both single-antenna and multiple-antenna multipath channels. When factor graphs of multipath channels have no cycles or mild cycle conditions, the considered self-iterative equalizer can converge to optimum performance after a few iterations; but it may suffer local convergence in channels with severe cycle conditions.

Keywords and phrases: factor graph, equalizer, iterative processing, multipath fading, MIMO.

1. INTRODUCTION

A multipath fading channel, which can be mathematically described by a convolution of transmitted signals and linear channel response, is one of many typical channel models occurring in digital communications. In general, an equalizer that makes detection based on a number of adjacent received symbols is necessary to achieve optimal or near-optimal performance in multipath channels. In classical communication theory, different representations of multipath channels have led to equalizers with different designs. By representing multipath channels as trellis structures, the optimum sequence detector can be computed by the Viterbi algorithm [1], and the optimum symbol detector can be computed by BCJR algorithm [2]. Starting from the transfer function

representation of linear multipath systems, people proposed various low-complexity designs such as linear zero-forcing (ZF) equalizer, linear minimum mean-square-error (MMSE) equalizer, nonlinear zero-forcing decision feedback equalizer (ZF-DFE), non-linear MMSE-DFE, and so forth. [3]. In this work, the multipath channels are represented by factor graphs, and soft self-iterative equalizers that execute belief propagation algorithm on factor graphs are studied. (Please refer to [4] for an excellent tutorial on factor graph and its applications.)

One question might rise regarding the motivation of this work, since we have already had both Viterbi algorithm and BCJR algorithm as exact optimum equalizers. The answer to this question lies in the flexibility of factor graph in characterizing multipath channels to per-path level. As a well-known fact, the computational complexity of Viterbi and BCJR algorithms are exponential in the total number of multipaths L . In practice, there exist cases when only L' out of L paths (with $L' < L$) have significant channel gains and

This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

moreover the location of these significant L' paths can be slowly changing in time, for example, rural wireless channels. Then, a reduced-complexity equalizer that avoids or reduces the computations spent on those zero multipath taps is desirable. Some efforts along this direction have been made in earlier works, for example, parallel Viterbi and parallel BCJR algorithms in [5, 6], which however may require specifically designed control logic for a different multipath scenario. In the considered factor-graph-based soft iterative equalizer, the log-likelihood probabilities are passed as messages in factor graphs between channel nodes and information nodes only along the edges that correspond to paths with significant gain, thus it inherently results in a complexity reduction owing to the sparseness of multipath channels. In particular, we consider three schemes to compute the messages passed from channel nodes to information nodes, namely the scheme based on the a posteriori probability (APP) algorithm, the one based on the linear-MMSE-soft-interference-cancellation (LMMSE-SIC), and the one based on match-filter-soft-interference-cancellation (MF-SIC); and we analyze their performance and applicabilities in practical multipath channels.

One main focus of this paper is the effect of cycles that existed in factor graph on the equalization performance. As compared to the Viterbi and BCJR algorithms which themselves are belief propagation algorithms operating in trellis trees of multipath channels and guarantee the optimum performance, the belief propagation algorithm operating in factor graphs guarantee global optimality only if the underlying factor graph is a tree. Although the condition of factor graph being a tree (i.e., without cycles) is not always met in practice, the factor-graph-based belief propagation algorithm has achieved great success in decoding cycle-contained linear turbo codes and low-density parity-check (LDPC) codes. For the considered self-iterative equalizer, we quantitatively analyze the cycle effect in single-input single-output (SISO), multiple-input single-output (MISO), and multiple-input multiple-output (MIMO) wireless systems; and discuss an alternative representation of factor graphs that ameliorates the performance degradation due to cycle effects.

While it bears similarities to various iterative receivers developed earlier, for example, [7, 8, 9, 10], we highlight that the soft self-iterative equalizer is a *self-iterative* device which successively improves the equalization performance by taking advantage of the constraints in received signals due to multipaths, instead of other constraints for instance imposed by error-control coding. Moreover, since the considered equalizer inputs prior and outputs a posteriori probabilities of information symbols, it can easily concatenate with other receiver modules to achieve the turbo receiver processing gains [8].

The rest of this paper is organized as follows. In Section 2, the system model and factor graph representation of multipath channels are described. In Section 3, the factor-graph-based soft iterative equalizer is derived. In Section 4, the performance of the soft self-iterative equalizer is analyzed by numerical simulations for both single-antenna and multiple-antenna systems. Finally, Section 5 contains the conclusions.

2. SYSTEM MODEL AND FACTOR GRAPH REPRESENTATION

Assume match-filtering and symbol-rate sampling, the received signals of multipath channels are normally described by the following time-domain equation [11]:

$$y_t = \sum_{l=0}^{L-1} h_{t,l} x_{t-l} + n_t, \quad t = 1, 2, \dots, T, \quad (1)$$

where $y_t \in \mathcal{C}$ and $x_t \in \Omega$ are the receive and transmit signals at time t , respectively; Ω is the modulation set; $h_{t,l} \in \mathcal{C}$ is the channel impulse response with delay of l times the symbol rate at time t ; $n_t \in \mathcal{C} \sim \mathcal{N}(0, \sigma^2)$ is the zero-mean σ -variance circularly symmetrical Gaussian ambient noise that has been properly whitened and is independent of data; L is the total number of multipaths; T is the frame length. In this paper, we are concerned with block signal processing, and assume that zero prefix is inserted in each signal frame, that is, $x_t = 0$, $t = -L + 1, \dots, -1$. For ease of comparison, we also assume that channel gain is properly normalized: in static channels, $\sum_{l=0}^{L-1} |h_{t,l}|^2 = 1$; and in fading channels, $\sum_{l=0}^{L-1} E(|h_{t,l}|^2) = 1$, where $E(\cdot)$ denotes the expectation over random variables $h_{t,l}$, for all l . As mentioned earlier, we only consider uncoded systems in this work, thus x_t have equal prior probabilities and are assumed to be independent for different t .

Equivalently, (1) can be written in a matrix form as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_t \\ \vdots \\ y_T \end{bmatrix} = \underbrace{\begin{bmatrix} h_{1,L-1} & h_{1,L-2} & \cdots & h_{1,0} & & & \\ & \ddots & \ddots & \ddots & \ddots & & \\ & & h_{t,L-1} & h_{t,L-2} & \cdots & h_{t,0} & \\ & & & \ddots & \ddots & \ddots & \\ & & & & h_{T,L-1} & h_{T,L-2} & \cdots & h_{T,0} \end{bmatrix}}_{\mathbf{H}} \times \begin{bmatrix} x_{-L+2} \\ \vdots \\ x_{t-L+2} \\ \vdots \\ x_T \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_t \\ \vdots \\ n_T \end{bmatrix}, \quad (2)$$

where \mathbf{H} is a $T \times (T + L - 1)$ Toeplitz matrix. Throughout this paper, we assume that \mathbf{H} is perfectly known to the receiver, and $h_{t,l}$, for all t, l , can be either time invariant or time variant within each signal frame. In addition, we define $\mathbf{I}_{\mathbf{H}}$ as the incidence matrix of \mathbf{H} , such that $\{\mathbf{I}_{\mathbf{H}}\}_{i,j} = 1$, if $\{|\mathbf{H}\}_{i,j}|^2 > 0$; $\{\mathbf{I}_{\mathbf{H}}\}_{i,j} = 0$, otherwise. $\mathbf{I}_{\mathbf{H}}$ will later be used to help explain the cycle effects of the factor-graph-based soft equalizer.

In the above, we described single-input single-output (SISO) multipath systems. Without much difficulty, (1)

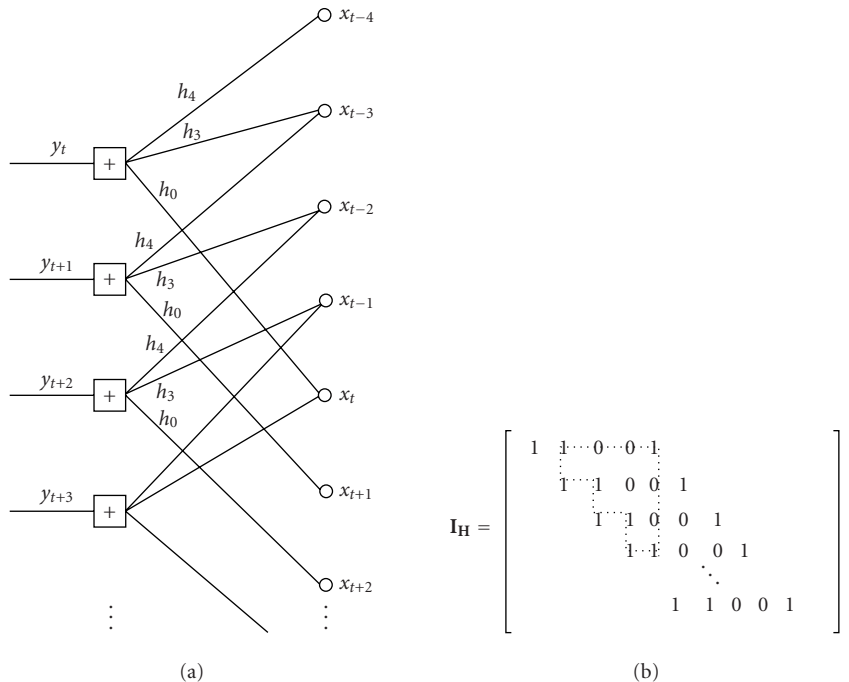


FIGURE 1: (a) The factor graph representation and (b) the incidence matrix of a single-antenna multipath channel: $y_t = h_0x_t + h_3x_{t-3} + h_4x_{t-4} + n_t$.

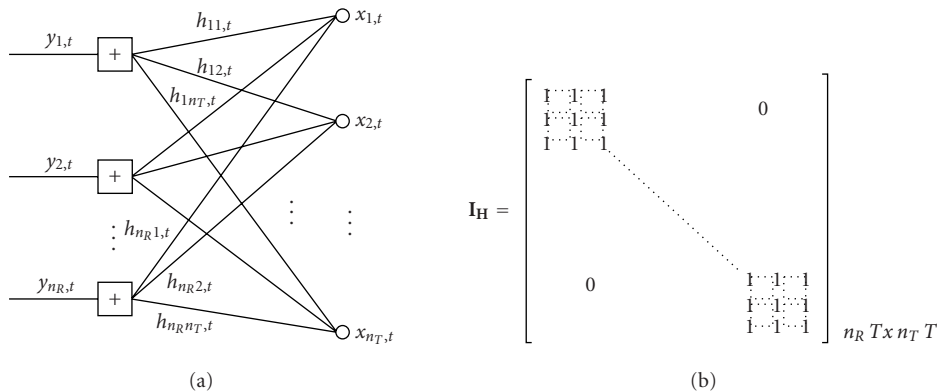


FIGURE 2: (a) The factor graph representation and (b) the incidence matrix of an $n_T \times n_R$ single-path MIMO channel.

and (2) as well as \mathbf{I}_H can be extended to multiple-input single-output (MISO) and multiple-output multiple-output (MIMO) cases, by simply replacing $y_t, h_{t,l}, x_t, n_t$ with their matrix/vector counterparts $\mathbf{y}_t, \mathbf{h}_{t,l}, \mathbf{x}_t, \mathbf{n}_t$. As a result, \mathbf{H} and \mathbf{I}_H now become $N_r T \times N_t \cdot (T + L - 1)$ matrices, where N_r and N_t are the number of receive and transmit antennas, respectively.

The above multipath channels in (1) and (2) can also be depicted by factor graphs. The example of the factor graph representations of SISO multipath and MIMO single-path channels are given in Figures 1 and 2. There are two types of nodes in the factor graph: the channel nodes for y_t , for all t , and the information nodes x_t , for all t . An edge connects channel node t and information node t' , only if the channel gain is significant, that is, $|h_{t-t',l}|^2 > 0$. We remark that by no

means the factor graphs shown in the figures are unique representation of the corresponding multipath channels; indeed, different representations of the same multipath channel lead to different designs of the factor-graph-based soft iterative equalizer, which we will discuss in Section 4.3.

3. SOFT SELF-ITERATIVE EQUALIZER BASED ON FACTOR GRAPH

The considered soft self-iterative equalizer computes the marginal probabilities of information symbol $\{x_t\}_{t=0}^T$ based on prior probabilities of the receive signals $\{y_t\}_{t=0}^T$ and $\{x_t\}_{t=0}^T$, by executing belief propagations in factor graphs. (As a comparison, both Viterbi algorithm and BCJR algorithm execute belief propagation in trellis trees.)

The messages, defined as the log-likelihood ratio (LLR) of information symbols, are iteratively passed among the nodes in factor graphs, such as to compute the marginal probabilities of information symbols. For BPSK modulation, the message is 1-tuple. In this paper, we will mainly study the complex modulation schemes such as MPSK and MQAM for which the message is $\log_2 |\Omega|$ -tuple. Let $m_{ci}^{(p)}$ be the message passed from the channel node c to the information node i at the p th iteration, $m_{ci}^{(p)} \triangleq (m_{ci,0}^{(p)}, m_{ci,1}^{(p)}, \dots, m_{ci,\log_2 |\Omega|-1}^{(p)})$; and it is updated as

$$m_{ci,k}^{(p)} \triangleq F_{ci,k}(y_c, m_{i'c}^{(p)}, \forall i' \in \mathcal{U}_c) \quad \forall k, \quad (3)$$

$$= \log \frac{\Pr[b_{i,k}=0 | y_c, m_{i'c}^{(p-1)}, \forall i' \in \mathcal{U}_c \setminus \{i\}, m_{ic,k'}^{(p-1)}, \forall k' \neq k]}{\Pr[b_{i,k}=1 | y_c, m_{i'c}^{(p-1)}, \forall i' \in \mathcal{U}_c \setminus \{i\}, m_{ic,k'}^{(p-1)}, \forall k' \neq k]}, \quad \forall k, \quad (4)$$

where the mapping function from $\log_2 |\Omega|$ -tuple $(b_{i,0}, \dots, b_{i,\log_2 |\Omega|-1})$ to complex symbol x_i is usually referred to as modulation format; $m_{i'c}$ is the message sent from information node i' to channel node c , as explained next; \mathcal{U}_c denotes the set of all information nodes incident to channel node c , $\mathcal{U}_c \setminus \{i\}$ denotes \mathcal{U}_c excluding information node i ; and y_c is the received signal at time c . The message update rule in (3) follows the general principle of a belief propagation algorithm, that is, the component message $m_{ci,k}^{(p)}$ sent from channel node c to information node i is updated based on received signal y_c and all incident messages to chan-

nel node c except for the same incident component message $m_{ci,k}^{(p-1)}$. Similarly, we let $m_{ic}^{(p)}$ be the message passed from the information node i to the channel node c at the p th iteration, $m_{ic}^{(p)} \triangleq (m_{ic,0}^{(p)}, m_{ic,1}^{(p)}, \dots, m_{ic,\log_2 |\Omega|-1}^{(p)})$; and it is updated as

$$m_{ic,k}^{(p)} \triangleq G_{ic,k}(m_i^{(0)}, m_{c'i}^{(p)}, \forall c' \in \mathcal{V}_i) = \log \frac{\Pr[b_{i,k}=0 | m_i^{(0)}, m_{c'i}^{(p-1)}, \forall c' \in \mathcal{V}_i \setminus \{c\}, m_{ci,k'}^{(p-1)}, \forall k' \neq k]}{\Pr[b_{i,k}=1 | m_i^{(0)}, m_{c'i}^{(p-1)}, \forall c' \in \mathcal{V}_i \setminus \{c\}, m_{ci,k'}^{(p-1)}, \forall k' \neq k]}, \quad \forall k, \quad (4)$$

where $m_i^{(0)}$ denotes the prior probabilities of the i th information symbol, input from other receiver modules (e.g., a channel decoder); \mathcal{V}_i denotes the set of all channel nodes incident to information node i .

In (4), assume that the messages $m_i^{(0)}$ and $m_{c'i}^{(p-1)}$, for all c' are independent random variables, then we have

$$G_{ic,k}(m_{i,k}^{(0)}, m_{c'i,k}^{(p)}, \forall c' \in \mathcal{V}_i) = m_{i,k}^{(0)} + \sum_{c' \in \mathcal{V}_i \setminus \{c\}} m_{c'i,k}^{(p)}. \quad (5)$$

On the other hand, we have the following three different approaches, that is, *a-posteriori*-probability- (APP-) based scheme, linear-MMSE-soft-interference-cancellation- (LMMSE-SIC-)based scheme, and match-filter-soft-interference-cancellation- (MF-SIC-)based scheme, to compute (3), that is,

$$F_{ci,k}(y_c, m_{i'c}^{(p)}, \forall i' \in \mathcal{U}_c) = \begin{cases} \log \frac{\sum_{x_i \in \mathcal{Q}_{i,k}^+} \exp\left(-|y_c - \sum_{i' \in \mathcal{U}_c} h_{c,c-i'} x_{i'}|^2 / \sigma^2 + \sum_{k=0}^{\log_2 |\Omega|-1} b_{i',k} \cdot m_{i'c,k}^{(p-1)} / 2\right)}{\sum_{x_i \in \mathcal{Q}_{i,k}^-} \exp\left(-|y_c - \sum_{i' \in \mathcal{U}_c} h_{c,c-i'} x_{i'}|^2 / \sigma^2 + \sum_{k=0}^{\log_2 |\Omega|-1} b_{i',k} \cdot m_{i'c,k}^{(p-1)} / 2\right)} - m_{ic,k}^{(p-1)}, & \text{for APP,} \\ \log \frac{\sum_{x_i \in \mathcal{S}_{i,k}^+} \exp\left(-|w_{c,i}^* (y_c - \tilde{y}_c) - \mu_{c,i} x_i|^2 / \nu_{c,i}^2 + \sum_{k=0}^{\log_2 |\Omega|-1} b_{i,k} \cdot m_{ic,k}^{(p-1)} / 2\right)}{\sum_{x_i \in \mathcal{S}_{i,k}^-} \exp\left(-|w_{c,i}^* (y_c - \tilde{y}_c) - \mu_{c,i} x_i|^2 / \nu_{c,i}^2 + \sum_{k=0}^{\log_2 |\Omega|-1} b_{i,k} \cdot m_{ic,k}^{(p-1)} / 2\right)} - m_{ic,k}^{(p-1)}, & \text{for LMMSE-SIC, MF-SIC,} \end{cases} \quad (6)$$

and for LMMSE-SIC,

$$w_{c,i}^* = \frac{h_{c,c-i}^*}{\sum_{i' \in \mathcal{U}_c \setminus \{i\}} |h_{c,c-i'}|^2 (1 - |\tilde{x}_{c-i'}|^2) + |h_{c,c-i}|^2 + \sigma^2}, \quad \mu_{c,i} = w_{c,i}^* h_{c,c-i}, \quad \nu_{c,i}^2 = \mu_{c,i} - \mu_{c,i}^2, \quad (7)$$

and for MF-SIC,

$$w_{c,i}^* = \frac{h_{c,c-i}^*}{|h_{c,c-i}|^2}, \quad \mu_{c,i} = 1, \quad \nu_{c,i}^2 = \frac{\sum_{i' \in \mathcal{U}_c \setminus \{i\}} |h_{c,c-i'}|^2 (1 - |\tilde{x}_{c-i'}|^2) + \sigma^2}{|h_{c,c-i}|^2}, \quad (8)$$

Initialize: for all edges $m_{ic}^{(0)} = 0$ for all edges $m_{ci}^{(0)} = F_{ci}(y_c, m_{i'c}^{(0)}, \forall i' \in \mathcal{U}_c)$
Self-iterative equalize: for $p = 1$ to P /* compute messages from channel nodes to information nodes */ for all edges $m_{ci}^{(p)} = F_{ci}(y_c, m_{i'c}^{(p)}, \forall i' \in \mathcal{U}_c)$ /* compute messages from information nodes to channel nodes */ for all edges $m_{ic}^{(p)} = G_{ic}(m_i^{(0)}, m_{c'i}^{(p)}, \forall c' \in \mathcal{V}_i)$ end
Output: /* compute information symbols' a posteriori probabilities $m_i^{(p)}$ */ for $i = 0$ to T $m_i^{(p)} = \sum_{c' \in \mathcal{V}_i} m_{c'i}^{(p)}$ end

ALGORITHM 1: Algorithm description of the factor-graph-based soft self-iterative equalizer.

with $\tilde{y}_c = \sum_{i' \in \mathcal{U}_c \setminus \{i\}} h_{c,c-i} \tilde{x}_{c-i}$,

$$\tilde{x}_i = \sum_{x_i \in \Omega} x_i \prod_{k=0}^{\log_2 |\Omega| - 1} \frac{b_{i,k} m_{ic,k}^{(p-1)}}{1 + b_{i,k} m_{ic,k}^{(p-1)}}, \quad (9)$$

where $\mathcal{S}_{i,k}^+$ is the set defined as $\{x_i \in \Omega \mid b_{i,k} = 0\}$, and similarly is $\mathcal{S}_{i,k}^-$; $\mathcal{Q}_{i,k}^+$ is the union of $\{x_{i'} \in \Omega \mid \text{for all } i' \in \mathcal{U}_c \setminus \{i\} \text{ and } \mathcal{S}_{i',k}^+\}$ and $\mathcal{S}_{i,k}^+$, and similarly is $\mathcal{Q}_{i,k}^-$. The detailed derivation of (6) is shown in the appendix.

Finally, the whole steps of the proposed equalizer are given in Algorithm 1.

4. NUMERICAL SIMULATIONS AND ANALYSIS

In this section, we analyze the factor-graph-based soft self-iterative equalizer in sparse wireless multipath channels through numerical simulations. For simplicity, we assume that channel gains remain constant in one frame and change independently from one to the other. The modulator uses the QPSK constellation with Gray mapping. Each frame contains 128 QPSK symbols per transmit antenna; proper zero prefix information symbols are inserted in each frame. The soft equalizer is a self-iterative device; and we only study the uncoded system. The performance is evaluated in terms of frame error rate (FER) versus the signal-to-noise ratio (SNR).

4.1. SISO multipath fading channels

First, consider a sparse 4-path fading channel: $y_t = h_0 x_t + h_3 x_{t-3} + n_t$, with $E\{|h_0|^2\} = 0.8$, $E\{|h_3|^2\} = 0.2$; thus, $L = 4$ and $L' = 2$. In Figure 3, the performance of three different approaches, (i.e., APP, LMMSE-SIC, and MF-SIC), to computing the extrinsic messages passed from channel nodes to

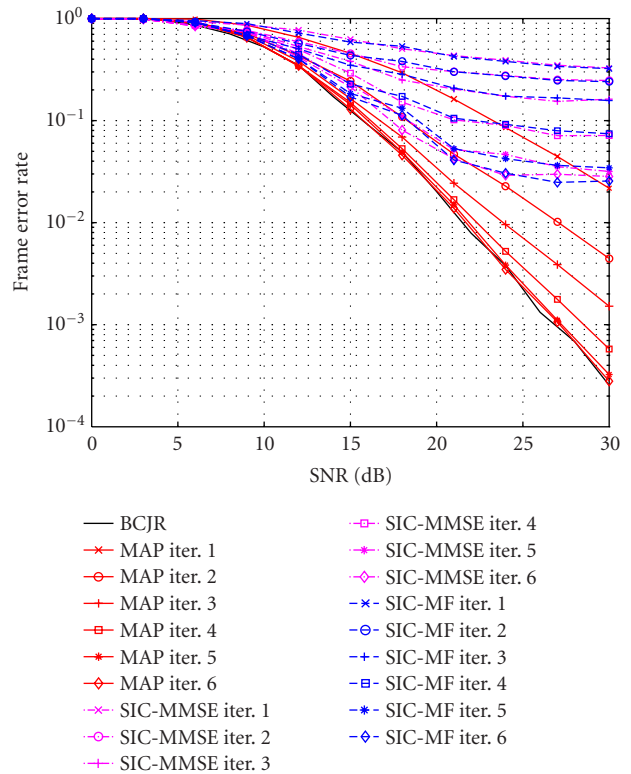


FIGURE 3: FER performance of the factor-graph-based soft iterative equalizer in SISO multipath fading channels ($n_T = 1$, $n_R = 1$, $L = 4$, $L' = 2$).

information nodes is presented. For each scheme, total six iterations, that is, $P = 6$, are conducted in the self-iterative equalizer. Serving as a benchmark, the performance of the optimum maximum likelihood equalizer based on BCJR algorithm is also included in the figure. Since the factor graph of this channel is cycle free, the belief propagation algorithm theoretically is able to achieve optimum performance. Indeed, the soft iterative equalizer using APP-based message update scheme achieves the optimum performance after a few iterations. On the contrary, two low-complexity schemes, LMMSE-SIC and MF-SIC, suffer error floors at high SNRs. We remark that the prior probability input from other receiver modules (e.g., channel decoder) can lower but never eradicate such error floors; henceforth we will only consider the APP-based scheme for channel node message updating.

Now, consider a sparse 5-path fading channel: $y_t = h_0 x_t + h_3 x_{t-3} + h_4 x_{t-4} + n_t$, where $E\{|h_0|^2\} = 0.7$, $E\{|h_3|^2\} = 0.2$, and $E\{|h_4|^2\} = 0.1$; thus, $L = 5$ and $L' = 3$. As seen in Figure 1, there exist a number of cycles with length 8 in the factor graph, where a “cycle” is defined as a close loop in the graph and its “length” is defined as the number of edges traversed by that loop. This cycle condition accounts for the marginal gap between the factor-graph-based equalization and the optimum performance, as shown in Figure 4.

4.2. MISO multipath fading channels

Equalization of MISO multipath channels falls into the group of “underdetermined” problems: at each time instance a mix-

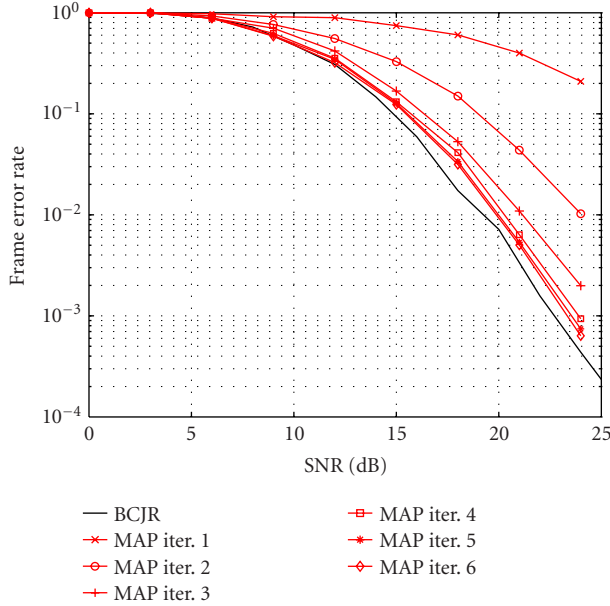


FIGURE 4: FER performance of the factor-graph-based soft iterative equalizer in SISO multipath fading channels ($n_T = 1$, $n_R = 1$, $L = 5$, $L' = 3$).

ture of plural information symbols that transmitted with different delays and from different antennas is to be detected from a single-receiver observation. Conventional linear equalization or decision-feedback-cancellation equalization schemes would lead to unsatisfactory performance, whereas an optimal equalizer has complexity exponential in $(L - 1) \cdot n_T$. When MISO multipath channels exhibit sparseness, the factor-graph-based soft equalizer becomes potentially attractive, as it can reduce the complexity exponent to $(L' - 1) \cdot n_T$.

We consider a two-transmit-one-receive-antenna (2×1) MISO system in a sparse 3-path fading. Every transmit-receive antenna pair follows the same multipath profile, that is, $E\{|h_0|^2\} = 0.8$, and $E\{|h_2|^2\} = 0.2$; fading coefficients for different paths and different antenna pairs are assumed to be mutually independent. The performance is illustrated in Figure 5. It is seen that after a few iterations the considered factor-graph-based equalizer performs slightly more than one dB away from the optimum equalizer. Again, this performance gap is due to the existence of length-4 cycles in the factor graphs. It is worth to remark that the complexity of optimum BCJR equalizer soon becomes prohibitive for (2×1) MISO systems with QPSK modulation and $L > 3$ multipaths; in comparison, the complexity exponent of factor-graph-based equalizer is proportional to L' , hence in sparse channels it is strictly lower than the original L .

4.3. MIMO multipath fading channels

Recently, there has been increasing interest in developing MIMO equalization schemes in multipath channels. We analyze the performance of the factor-graph-based equalizer as

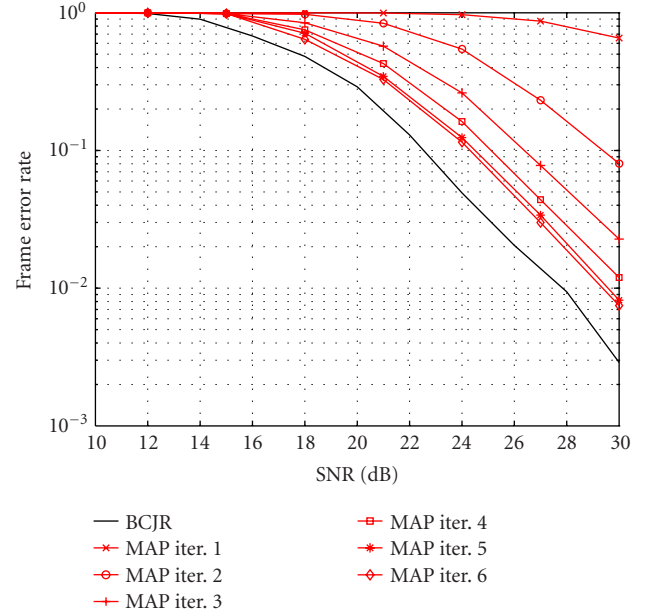


FIGURE 5: FER performance of the factor-graph-based soft iterative equalizer in MISO multipath fading channels ($n_T = 2$, $n_R = 1$, $L = 3$).

below. First, we consider $(n_T \times n_R)$ MIMO systems in single path fading channels. It is easily seen from Figure 2 that the incidence matrix \mathbf{I}_H contains length-4 cycles everywhere; and the cycle condition worsens as more antennas are employed. To the best of our knowledge, little efforts have been made to rigorously quantify the cycle condition of factor graphs. Empirically, the cycle condition is better, if the length of cycles is increased, or given the cycle length, the number of cycles is reduced, or the cycles have a larger number of edges connecting to rest of the graph. However, by and large, the combined effect of these empirical assertions is unclear; we then have to resort to numerical simulations. It is seen from Figures 6 and 7 that the considered self-iterative equalizer approaches optimum demodulation performance in (2×2) MIMO channels, but it suffers considerable performance loss in 4×4 MIMO channels. Especially from the (4×4) MIMO case, we conclude that the direct application of the factor-graph-based equalizer may not be a good option for MIMO channels. It is seen from Figure 8 that the above observation also holds for MIMO multipath channels—as much as 2.5 dB performance loss is seen in a (2×2) MIMO with 3 multipaths.

Alternative factor graph representation for MIMO multipath fading channels

The previous simulation results and analysis has identified the difficulty in directly applying the factor-graph-based equalizer in MIMO channels. An alternative way to ameliorate this problem is to reconstruct the underlying factor graphs. Shown in Figure 9 the idea is to glue all channel nodes in the original graph $\{y_{1,t}, \dots, y_{n_R,t}\}$ that corresponds to different receiver antennas at the same time instance t into

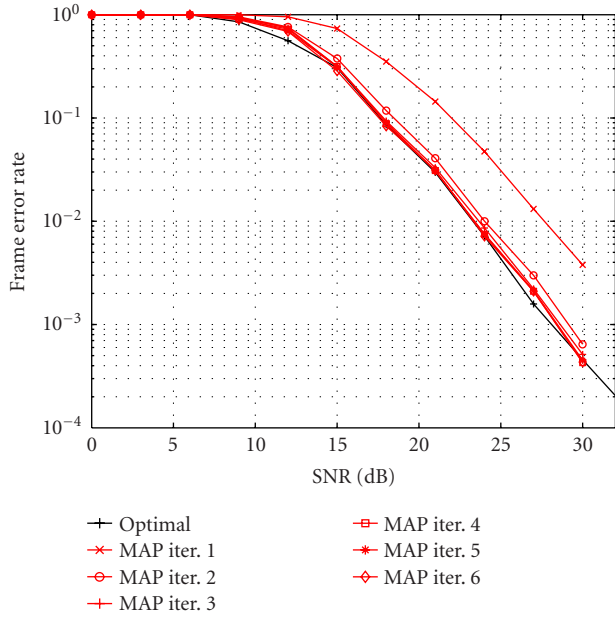


FIGURE 6: FER performance of the factor-graph-based soft iterative equalizer in MIMO multipath fading channels ($n_T = 2, n_R = 2, L = 1$).

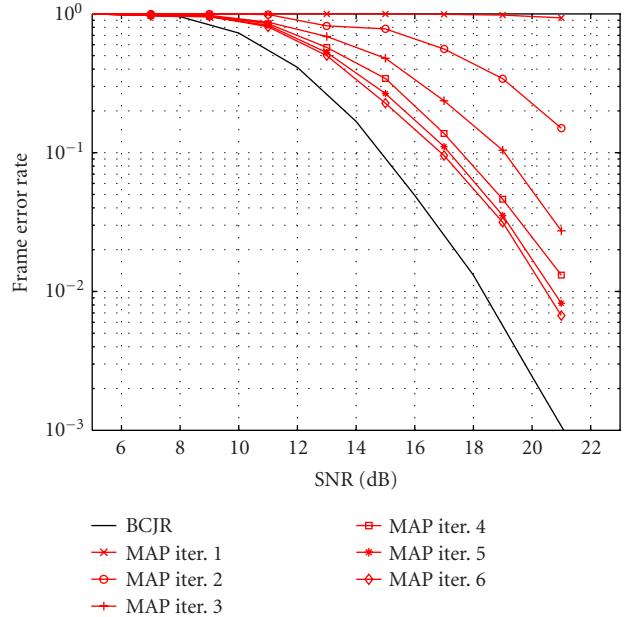


FIGURE 8: FER performance of the factor-graph-based soft iterative equalizer in MIMO multipath fading channels ($n_T = 2, n_R = 2, L = 3, L' = 2$).

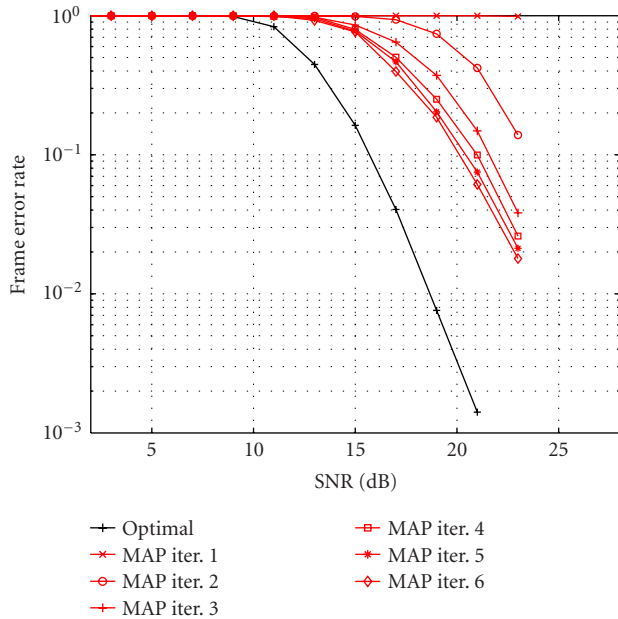


FIGURE 7: FER performance of the factor-graph-based soft iterative equalizer in MIMO multipath fading channels ($n_T = 4, n_R = 4, L = 1$).

a new channel node $\mathbf{y}_t \triangleq [y_{1,t}, \dots, y_{n_R,t}]^T$; the channel coefficient on each edge is now an $(n_R \times 1)$ vector instead of a scalar. In doing so, the alternative factor graph still represents the same MIMO multipath systems, but the extensive short cycles due to multiple receive antennas are systematically avoided. The belief propagation algorithm can be accordingly rederived; and in single-path channels, it converges

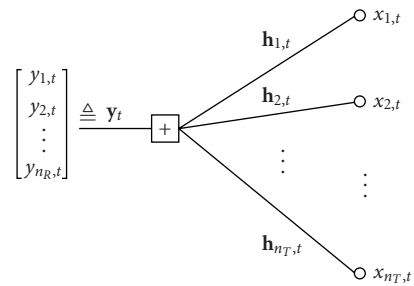


FIGURE 9: The alternative factor graph representation of an $n_T \times n_R$ single-path MIMO channel. Compared to Figure 2, here all channel nodes $\{y_{1,t}, y_{2,t}, \dots, y_{n_R,t}\}$ that correspond to different receiver antennas at the same time instance t are glued to form a new channel node \mathbf{y}_t .

in one iteration and coincides with the optimal APP MIMO demodulator [12]. With this alternative factor graph representation, we can continue to apply the self-iterative equalizer for MIMO multipath fading channels to improve the performance. We now consider the case of (2×2) MIMO with 3 multipaths as an example. The FER curves are shown in Figure 10. It is seen that the resulting performance is significantly improved and approaches the performance from the optimum demodulation.

5. CONCLUSIONS

Since a factor graph is able to characterize multipath channels to per-path level, the factor-graph-based soft self-iterative equalizer with reduced computational complexity is a potential candidate for sparse multipath channel

equalization. By numerical simulations, we have shown that the cycles in factor graphs are crucial to the convergence property of the considered soft self-iterative equalization. While being able to achieve near-optimum performance in single-input single-output (SISO) and multiple-input single-output (MISO) sparse multipath channels with mild cycle conditions, a factor-graph-based soft self-iterative equalizer may suffer noticeable performance loss in multiple-input

multiple-output (MIMO) multipath channels, unless proper means is taken to ameliorate the cycle conditions in factor graphs.

APPENDIX

DERIVATION OF (6)

(i) For APP detection, we have

$$\begin{aligned}
& F_{ci,k}(y_c, m_{i'c}^{(p)}, \forall i' \in \mathcal{U}_c) \\
&= \log \frac{\sum_{x_{i'} \in \mathcal{Q}_{i,k}^+} P(x_{ci} = x_{i'} | y_c)}{\sum_{x_{i'} \in \mathcal{Q}_{i,k}^-} P(x_{ci} = x_{i'} | y_c)} - \underbrace{\log \frac{P(b_{i,k} = +1)}{P(b_{i,k} = -1)}}_{m_{ic,k}^{(p-1)}} \\
&= \log \frac{\sum_{x_{i'} \in \mathcal{Q}_{i,k}^+} P(y_c | x_{ci} = x_{i'}) P(x_{ci} = x_{i'})}{\sum_{x_{i'} \in \mathcal{Q}_{i,k}^-} P(y_c | x_{ci} = x_{i'}) P(x_{ci} = x_{i'})} - m_{ic,k}^{(p-1)} \tag{A.1} \\
&= \log \frac{\sum_{x_{i'} \in \mathcal{Q}_{i,k}^+} \exp\left(-|y_c - \sum_{i' \in \mathcal{U}_c} h_{c,c-i'} x_{i'}|^2 / \sigma^2\right) \prod_{x_{i'} \in \mathcal{Q}_{i,k}^+} P(b_{i',k}^{(p-1)})}{\sum_{x_{i'} \in \mathcal{Q}_{i,k}^-} \exp\left(-|y_c - \sum_{i' \in \mathcal{U}_c} h_{c,c-i'} x_{i'}|^2 / \sigma^2\right) \prod_{x_{i'} \in \mathcal{Q}_{i,k}^-} P(b_{i',k}^{(p-1)})} - m_{ic,k}^{(p-1)} \\
&= \log \frac{\sum_{x_{i'} \in \mathcal{Q}_{i,k}^+} \exp\left(-|y_c - \sum_{i' \in \mathcal{U}_c} h_{c,c-i'} x_{i'}|^2 / \sigma^2 + \sum_{k=0}^{\log_2 |\Omega| - 1} b_{i',k} \cdot m_{i'c,k}^{(p-1)} / 2\right)}{\sum_{x_{i'} \in \mathcal{Q}_{i,k}^-} \exp\left(-|y_c - \sum_{i' \in \mathcal{U}_c} h_{c,c-i'} x_{i'}|^2 / \sigma^2 + \sum_{k=0}^{\log_2 |\Omega| - 1} b_{i',k} \cdot m_{i'c,k}^{(p-1)} / 2\right)} - m_{ic,k}^{(p-1)}.
\end{aligned}$$

(ii) For LMMSE-SIC detection, we first obtain the MMSE filtering output, given by

$$z_{c,i} = w_{c,i}^* (y_c - \tilde{y}_c). \tag{A.2}$$

Based on Gaussian approximation of $z_{c,i}$, the extrinsic messages can be computed by

$$\begin{aligned}
& F_{ci,k}(y_c, m_{i'c}^{(p)}, \forall i' \in \mathcal{U}_c) \\
&= \log \frac{\sum_{x_i \in \mathcal{S}_{i,k}^+} \exp\left(-|z_{c,i} - \mu_{c,i} x_i|^2 / \nu_{c,i}^2\right) \prod_{x_i \in \mathcal{S}_{i,k}^+} P(b_{i,k}^{(p-1)})}{\sum_{x_i \in \mathcal{S}_{i,k}^-} \exp\left(-|z_{c,i} - \mu_{c,i} x_i|^2 / \nu_{c,i}^2\right) \prod_{x_i \in \mathcal{S}_{i,k}^-} P(b_{i,k}^{(p-1)})} - m_{ic,k}^{(p-1)} \tag{A.3} \\
&= \log \frac{\sum_{x_i \in \mathcal{S}_{i,k}^+} \exp\left(-|w_{c,i}^* (y_c - \tilde{y}_c) - \mu_{c,i} x_i|^2 / \nu_{c,i}^2 + \sum_{k=0}^{\log_2 |\Omega| - 1} b_{i,k} m_{ic,k}^{(p-1)} / 2\right)}{\sum_{x_i \in \mathcal{S}_{i,k}^-} \exp\left(-|w_{c,i}^* (y_c - \tilde{y}_c) - \mu_{c,i} x_i|^2 / \nu_{c,i}^2 + \sum_{k=0}^{\log_2 |\Omega| - 1} b_{i,k} m_{ic,k}^{(p-1)} / 2\right)} - m_{ic,k}^{(p-1)},
\end{aligned}$$

where

$$\begin{aligned}
w_{c,i}^* &= \frac{h_{c,c-i}^*}{\sum_{i' \in \mathcal{U}_c \setminus \{i\}} |h_{c,c-i'}|^2 (1 - |\tilde{x}_{c-i'}|^2) + |h_{c,c-i}|^2 + \sigma^2}, \\
\mu_{c,i} &= w_{c,i}^* h_{c,c-i}, \quad \nu_{c,i}^2 = \mu_{c,i} - \mu_{c,i}^2. \tag{A.4}
\end{aligned}$$

The details for obtaining $w_{c,i}^*$, $\mu_{c,i}$, and $\nu_{c,i}^2$ can be found in [7].

(iii) For MF-SIC, we simply apply the match filter to the soft interference canceled output, that is,

$$z_{c,i} = w_{c,i}^* (y_c - \tilde{y}_c), \quad w_{c,i}^* = \frac{h_{c,c-i}^*}{|h_{c,c-i}|^2}. \tag{A.5}$$

We then approximate the MF-SIC output as Gaussian distributed, and compute extrinsic message in the same form in

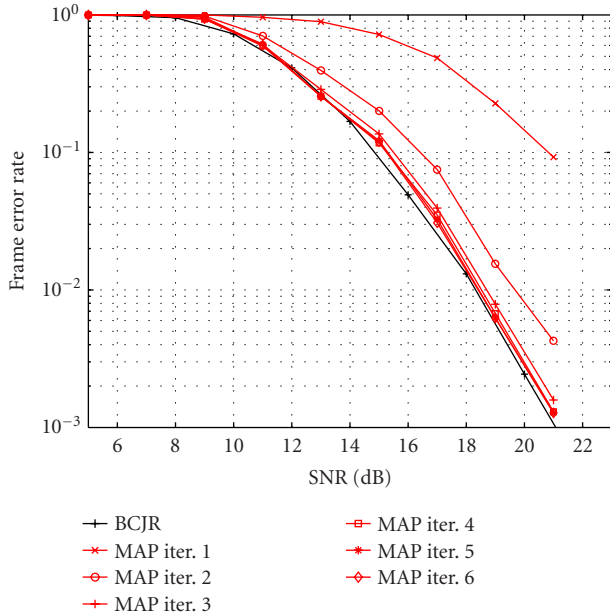


FIGURE 10: FER performance of the soft iterative equalizer based on alternative factor graph representation in MIMO multipath fading channels ($n_T = 2$, $n_R = 2$, $L = 3$, $L' = 2$).

(6) with mean and variance given by

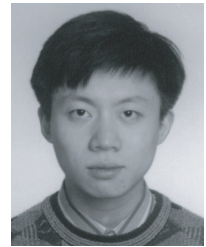
$$\mu_{c,i} = 1, \quad \nu_{c,i}^2 = \frac{\sum_{i' \in \mathcal{U}_c \setminus \{i\}} |h_{c,c-i'}|^2 (1 - |\tilde{x}_{c-i'}|^2) + \sigma^2}{|h_{c,c-i}|^2}. \quad (\text{A.6})$$

REFERENCES

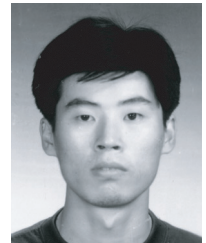
- [1] D. Forney, G., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 18, no. 3, pp. 363–378, 1972.
- [2] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (corresp.)," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, 1974.
- [3] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 3rd edition, 1995.
- [4] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [5] N. C. McGinty, R. A. Kennedy, and P. Hoher, "Parallel trellis Viterbi algorithm for sparse channels," *IEEE Commun. Lett.*, vol. 2, no. 5, pp. 143–145, 1998.
- [6] F. K. Lee and P. J. McLane, "Iterative parallel-trellis MAP equalizers with nonuniformly-spaced prefilters for sparse multipath channels," in *Proc. 56th IEEE Vehicular Technology Conference (VTC '02)*, vol. 4, pp. 2201–2205, Vancouver, BC, Canada, September 2002.
- [7] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046–1061, 1999.
- [8] J. Hagenauer, "The turbo principle: Tutorial introduction and state of the art," in *Proc. International Symposium on Turbo Codes and Related Topics*, pp. 1–11, Brest, France, September 1997.

- [9] J. Boutros and G. Caire, "Iterative multiuser joint decoding: unified framework and asymptotic analysis," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1772–1793, 2002.
- [10] P. Li, L. Liu, and W. K. Leung, "A simple approach to near-optimal multiuser detection: interleaved-division multiple-access," in *Proc. IEEE Wireless Communications and Networking (WCNC '03)*, vol. 1, pp. 391–396, New Orleans, La, USA, March 2003.
- [11] K. A. Hamied and G. L. Stuber, "Performance of trellis-coded modulation for equalized multipath fading ISI channels," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 50–58, 1995.
- [12] G. Bauch, "Concatenation of space-time block codes and turbo-TCM," in *Proc. IEEE International Conference on Communications, (ICC '99)*, vol. 2, pp. 1202–1206, Vancouver, BC, Canada, June 1999.

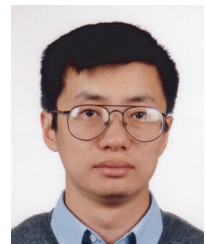
Ben Lu received the B.S. and M.S. degrees in electrical engineering from Southeast University, Nanjing, China, in 1994 and 1997, and the Ph.D. degree from Texas A&M University, in 2002. From 1994 to 1997, he was a Research Assistant with National Mobile Communications Research Laboratory at Southeast University, China. From 1997 to 1998, he was with the CDMA Research Department of Zhongxing Telecommunication Equipment Co., Shanghai, China. From 2002 to 2004, he worked for the project of high-speed wireless packet data transmission (4G prototype) at NEC Laboratories America, Inc., Princeton, New Jersey. He is now with Silicon Laboratories. His research interests include the signal processing and error-control coding for mobile and wireless communication systems.



Guosen Yue received the B.S. degree in physics and the M.S. degree in electrical engineering from Nanjing University, Nanjing, China, in 1994 and 1997, and the Ph.D. degree from Texas A&M University, College Station, Texas, in 2004. Since August 2004, he has been with NEC Laboratories America, Inc., Princeton, New Jersey, conducting research on broadband wireless systems and mobile networks. His research interests are in the area of advanced modulation and channel coding techniques for wireless communications.



Xiaodong Wang received the B.S. degree in electrical engineering and applied mathematics (with the highest honors) from Shanghai Jiao Tong University, Shanghai, China, in 1992; the M.S. degree in electrical and computer engineering from Purdue University, in 1995; and the Ph.D. degree in electrical engineering from Princeton University, in 1998. From July 1998 to December 2001, he was an Assistant Professor in the Department of Electrical Engineering, Texas A&M University. In January 2002, he joined the faculty of the Department of Electrical Engineering, Columbia University. Dr. Wang's research interests fall in the general areas of computing, signal processing, and communications. He has worked in the areas of digital communications, digital signal processing, parallel and distributed computing, nanoelectronics, and bioinformatics, and has published



extensively in these areas. Among his publications is a recent book entitled *Wireless Communication Systems: Advanced Techniques for Signal Reception*, published by Prentice Hall, Upper Saddle River, in 2003. His current research interests include wireless communications, Monte-Carlo-based statistical signal processing, and genomic signal processing. Dr. Wang received the 1999 NSF CAREER Award, and the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award. He currently serves as an Associate Editor for the IEEE Transactions on Communications, the IEEE Transactions on Wireless Communications, the IEEE Transactions on Signal Processing, and the IEEE Transactions on Information Theory.

Mohammad Madihian received the Ph.D. degree in electronic engineering from Shizuoka University, Japan, in 1983. He joined NEC Central Research Laboratories, Kawasaki, Japan, where he worked on research and development of Si and GaAs device-based digital as well as microwave and millimeter-wave monolithic ICs. In 1999, he moved to NEC Laboratories America, Inc., Princeton, New Jersey, and is presently the Department Head of Microwave and Signal Processing and Chief Patent Officer. He conducts PHY/MAC layer signal processing activities for high-speed wireless networks and personal communication applications. He has authored or coauthored more than 130 scientific publications including 20 invited talks, and holds 35 Japan/US patents. Dr. Madihian has received the IEEE MTT-S Best Paper Microwave Prize in 1988, and the IEEE Fellow Award in 1998. He holds 8 NEC Distinguished R&D Achievement Awards. He has served as a Guest Editor for the IEEE Journal of Solid-State Circuits, Japan IEICE Transactions on Electronics, and IEEE Transactions on Microwave Theory and Techniques. He is presently serving on the IEEE Speaker's Bureau, IEEE Compound Semiconductor IC Symposium (CSICS) Executive Committee, IEEE Radio and Wireless Conference Steering Committee, IEEE International Microwave Symposium (IMS) Technical Program Committee, IEEE MTT-6 Subcommittee, IEEE MTT Editorial Board, and Technical Program Committee of the International Conference on Solid State Devices and Materials (SSDM). Dr. Madihian is an Adjunct Professor at the Electrical and Computer Engineering Department, Drexel University, Philadelphia, Pennsylvania.



Estimation of Directions of Arrival by Matching Pursuit (EDAMP)

Güneş Z. Karabulut

*School of Information Technology and Engineering, University of Ottawa, ON, Canada K1N 6N5
Email: gkarabul@site.uottawa.ca*

Tolga Kurt

*School of Information Technology and Engineering, University of Ottawa, ON, Canada K1N 6N5
Email: tkurt@site.uottawa.ca*

Abbas Yongaçoğlu

*School of Information Technology and Engineering, University of Ottawa, ON, Canada K1N 6N5
Email: yongacog@site.uottawa.ca*

Received 30 April 2004; Revised 7 October 2004

We propose a novel system architecture that employs a matching pursuit-based basis selection algorithm for directions of arrival estimation. The proposed system does not require a priori knowledge of the number of angles to be resolved and uses very small number of snapshots for convergence. The performance of the algorithm is not affected by correlation in the input signals. The algorithm is compared with well-known directions of arrival estimation methods with different branch-SNR levels, correlation levels, and different angles of arrival separations.

Keywords and phrases: directions of arrival estimation, adaptive antennas, matching pursuit algorithm, spatial resolution.

1. INTRODUCTION

In recent years, the impact of adaptive antennas and array processing to the system performance of wireless communication systems has gained intense attention. Adaptive (or smart) antennas consist of an antenna array combined with space and time processing. The processing of different antennas helps to improve system performance in terms of both capacity and quality, in particular by decreasing cochannel interference. A detailed overview of adaptive antennas can be found in [1, 2].

One of the most important problems for adaptive antenna systems in order to perform well is to have reliable reference inputs. These references include array element positions and characteristics, directions of arrivals, planar properties and dimensionality of the incoming signals. In this paper we investigate one of the most critical problems of adaptive antenna systems, namely directions of arrival (DOA) estimation.

For an adaptive system to be effective, it must have very accurate estimations of the DOA for the signal and the interferers. Once the directions are estimated accurately then processing in spatial, time, or other domains can be accomplished in order to improve the system performance.

There are many different approaches and algorithms for estimating DOA with various complexities and resolution properties such as ML [3], Bartlett [4], MVDR [1], MUSIC [5], and ESPRIT [6]. Variations to these models can also be found in the recent literature, some of which will be referred to in the following section.

For estimation of DOA, we consider a high-resolution basis selection algorithm, the flexible tree-search-based orthogonal matching pursuit (FTB-OMP) algorithm that is proposed in [7]. The FTB-OMP algorithm heuristically converges to the maximum likelihood solution. The algorithm selects a basis for signal decomposition by determining a small, possibly the smallest, subset of vectors chosen from a large redundant set of vectors to match the given data. This problem has various applications such as time/frequency representations [8], speech coding [9], and spectral estimation [10]. For the case of DOA, this set of vectors are modeled as possible outputs of the antenna array elements when the signal is arriving from a certain direction. The problem

of selecting correct linear combination of these elements is equivalent to the problem of selecting correct DOA.

In DOA estimation, typically only a small number of directions contain the signal. Hence, the solution to the DOA estimation problem will be sparse. In this paper, we propose to use the FTB-OMP algorithm for DOA estimation, by exploiting the sparsity property of the DOA. The proposed technique is named as estimation of directions of arrival by matching pursuit (EDAMP). The main advantages of EDAMP are the flexibility and increased resolution at low signal-to-noise ratio (SNR) levels. It also does not require the a priori knowledge of the number of signals to be resolved, and it is not affected by the correlation of the signals arriving from different directions. The output of the algorithm is directly the angles of arrivals and their corresponding amplitudes; hence it does not require any postprocessing of output amplitudes at different angles as would be required in the case of conventional DOA estimators.

In the next section, the problem statement for the DOA estimation will be presented. In Section 3, the FTB-OMP algorithm employed in EDAMP structure will be summarized. In Section 4, the system model for estimating directions will be given. In Section 5, the simulation results will be presented for different scenarios. Finally in Section 6, the conclusions will be given.

2. PROBLEM STATEMENT

Consider an antenna array consisting of N elements. The output of these elements is a vector \mathbf{x} of size $N \times 1$. Generally \mathbf{x} corresponds to a linear combination of signals from different directions. If we consider i th and j th elements of \mathbf{x} , depending on DOA and the distance between them, x_i and x_j contain the same signals with different phase shifts. The problem is to identify each signal's DOA from \mathbf{x} which is a weighted sum of the signals plus noise.

In the literature, different methods for achieving this goal are presented.

- (i) The first one is the maximum likelihood (ML) approach [3]. Although it is the best one in terms of performance, it has formidable complexity. So other suboptimum algorithms which generally converge to ML performance at high SNR are proposed.
- (ii) The second approach is finding the array response in the spectral domain for different angles, and recovering the local maximas as DOA [1, 4].
- (iii) The third one is the eigenstructure method. In this method the space spanned by the eigenvectors is partitioned into signal subspace and noise subspace, hence they are referred to as subspace algorithms. After partitioning, signal subspace is investigated to recover DOA. The most popular subspace algorithms are ESPRIT [6] and MUSIC [5]. These algorithms are more complex than spectral domain algorithms since they require eigenvalue decomposition. However they have performances in between ML algorithm and spectral domain algorithms. On the other hand, they have poor performances in the low-SNR regions [1, 2].

Many different techniques, including independent component analysis [11], and many modified versions of these algorithms have been proposed in addition to the main ones mentioned above [1, 12, 13, 14].

In this paper we propose to use the EDAMP algorithm as a solution to the DOA estimation problem in order to achieve high resolution with low complexity. In EDAMP, we propose to use a high-resolution basis selection algorithm FTB-OMP. In the next section, the FTB-OMP algorithm will be described in detail.

3. BASIS SELECTION ALGORITHMS

The basis selection problem can be stated over \mathbb{C} as follows. Let $\mathcal{D} = \{a_k\}_{k=1}^n$ be a set/dictionary of vectors which is highly redundant (i.e., $a_k \in \mathbb{C}^m$ and $m \ll n$ with $\mathbb{C}^m = \text{Span}(\mathcal{D})$).

The basis selection problem can be viewed as finding the most sparse solution to a linear system of equations. More precisely, if we form a matrix A from the columns of the dictionary \mathcal{D} , $A = [a_1, a_2, \dots, a_n]$, the problem can be stated as finding an $\bar{\mathbf{x}}$, with at most r nonzero entries such that

$$\|\bar{\mathbf{x}} - \mathbf{x}\| \leq \epsilon \quad (1)$$

for $\epsilon \geq 0$, and $r > 1$.

Even though it would give the ML solution, finding the most sparse solution to (1) in an overcomplete dictionary using an exhaustive search is infeasible for large dimensions. In order to solve this problem, suboptimal methods based on sequential and parallel basis selection have been proposed. Due to high-complexity requirements of the parallel basis selection algorithms [15], sequential basis selection (SBS) methods are more frequently used for practical purposes [10, 16, 17].

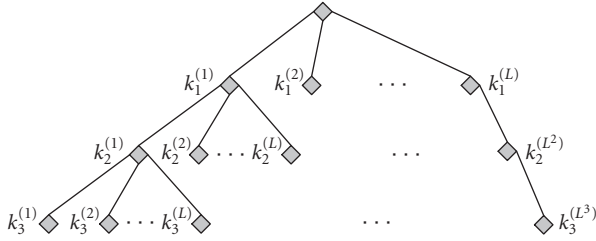
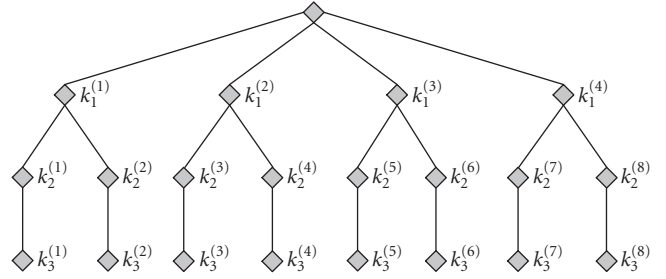
In the following sections, we describe the orthogonal matching pursuit (OMP), and the tree-search-based OMP algorithms. There are several other decomposition algorithms such as best orthogonal basis [18] and method of frames [19], which are not considered here due to their low resolution and poor sparsity properties.

The algorithms are explained based on the notation in [20]. As mentioned before, basis selection in OMP algorithms is performed sequentially, that is, one at a time.

Let the residual vector after the p th iteration be denoted by b_p , with $b_0 = \mathbf{x}$. P_{S_p} denotes the orthogonal projection matrix onto the range space of S_p , and $P_{S_p}^\perp = I - P_{S_p}$ denotes its orthogonal complement with $P_{S_0} = 0$ and $P_{S_0}^\perp = I$. The projection matrix on the space spanned by a_k , with $\|a_k\| = 1$, is $P_{a_k} = a_k a_k^T$. The algorithm terminates after r iterations.

3.1. Orthogonal matching pursuit algorithm

The orthogonal matching pursuit (OMP) algorithm is proposed in [20, 21], independently. OMP is also called modified matching pursuit algorithm [20].


 FIGURE 1: L -branch search tree.

 FIGURE 2: $L = 4, d = 2$ search tree for $r = 4$.

The OMP selects k_p in the p th iteration by finding the vector best aligned with the residual obtained by projecting b onto the orthogonal complement of the range space S_{p-1} , that is,

$$\begin{aligned} k_p &= \arg \max_l |a_l^T P_{S_{p-1}}^\perp b| \\ &= \arg \max_l |a_l^T b_{p-1}|, \quad l \notin I_{p-1}. \end{aligned} \quad (2)$$

With the initial values, $\hat{a}_{k_p}^0 = a_{k_p}$, $q_0 = 0$, we can write

$$P_{S_p} = P_{S_{p-1}} + q_p q_p^T, \quad (3)$$

where

$$\begin{aligned} \hat{a}_{k_p}^l &= \hat{a}_{k_p}^{l-1} - (q_{l-1}^T \hat{a}_{k_p}^{l-1}) q_{l-1}, \quad l = 1, 2, \dots, p, \\ q_p &= \frac{\hat{a}_{k_p}^p}{\|\hat{a}_{k_p}^p\|}. \end{aligned} \quad (4)$$

The residual b_p is updated as follows:

$$b_p = P_{S_p}^\perp b_{p-1} = b_{p-1} - (q_p^T b_{p-1}) q_p. \quad (5)$$

The coefficients c_i change with each iteration and can be evaluated by taking the orthogonal projection of \mathbf{x} onto S_p . The algorithm terminates when either $p = r$, or $\|b_p\| \leq \epsilon$.

3.2. Tree-search-based orthogonal matching pursuit algorithm

Matching pursuit algorithms with tree-based search are proposed in [22]. We focus on TB-OMP algorithm.

In this algorithm, the best matching vector indices, $\{k_p^{(1)}, k_p^{(2)}, \dots, k_p^{(L)}\}$ at the p th iteration are selected according to

$$\begin{aligned} k_p^{(i)} &= \arg \max_l |a_l^T P_{S_{p-1}}^\perp b|, \\ l &\neq \{k_p^{(1)}, k_p^{(2)}, \dots, k_p^{(i-1)}\}, \quad i = 1, \dots, L. \end{aligned} \quad (6)$$

At the end of r iterations, the search grows exponentially to a tree with L^r leaves as shown in Figure 1. The leaf corresponding to the smallest residual error vector yields the solution.

3.3. Flexible tree-search-based orthogonal matching pursuit

In [22], it is concluded that OMP algorithm offers a good compromise between performance and running time among the tree-search techniques, namely the matching pursuit and the order recursive matching pursuit algorithms.

In this section, we summarize the efficient tree-search-based OMP algorithms with branch pruning, the flexible tree-search-based OMP (FTB-OMP), that has been recently proposed in [7]. A maximum of L branches are searched at each partial solution. Thus, the resolution is adaptive, since it changes for different values of L in the algorithm. Note that TB-OMP (proposed in [22]) also has this adaptive nature, but has a prohibitive running time since it does not employ tree-pruning.

Our objective is to prune the tree branches that are heuristically believed to be unnecessary. Our heuristic is only to keep branches among $k_p^{(1)}, k_p^{(2)}, \dots, k_p^{(L)}$ which are closely ‘‘aligned’’ with the OMP first choice branch $k_p^{(1)}$. We measure this alignment by the correlation between vectors which is defined as

$$\rho_{ij} = \frac{\langle a_i, a_j \rangle}{\|a_i\| \|a_j\|}. \quad (7)$$

In the algorithm, an input design parameter correlation threshold ξ is given. A branch is assumed to be unnecessary when the candidate vector is not aligned with $k_p^{(1)}$, that is, $|\rho_{k_p^{(1)}, k_p^{(i)}}| < \xi$.

In flexible tree-search-based OMP (FTB-OMP), the branching factor L is of variable size. In the first iteration $L = M$, where M is a parameter of the algorithm. At the i th iteration L is set to $\lceil M/d^i \rceil$, where $\lceil \cdot \rceil$ represents the ceiling function. The parameter $d > 0$, represents the speed of the decay in the branching factor of the search tree. The idea in this algorithm is to start the search with a large number of branches at the initial iteration, where an erroneous selection is more likely to appear, and to reduce the branching factor as the number of iterations increases. A search tree for $L = 4$, $d = 2$ is shown in Figure 2. For the special case $d = 1$, the algorithm keeps L as the branching factor.

Note that FTB-OMP is a generalization of both OMP and TB-OMP algorithms. By choosing $\xi = 1$, we require full alignment so that only $k_p^{(1)}$ is kept, reproducing OMP. By choosing $\xi = 0$, and $d = 1$, we place no restriction on

```

FTB-OMP ( $d, p, r, L, \xi, \epsilon$ )
Global  $K = [k_1, k_2, \dots]$ , Best_res, Best_k
Calculate  $b_{p-1}$  as in (5)
If  $\|b_{p-1}\| < \text{Best\_res}$ 
    Best_k =  $[k_1, \dots, k_{p-1}]$ 
    Best_res  $\leftarrow \|b_{p-1}\|$ 
end
If  $p > r$  or  $\|b_{p-1}\| < \epsilon$ , then return
Calculate  $\{k_p^{(1)}, k_p^{(2)}, \dots, k_p^{(L)}\}$  as in (6)
For each  $i = 1-L$  do
    If  $|\rho_{k_p^{(1)}, k_p^{(i)}}| \geq \xi$ 
         $k_p = k_p^{(2)}$ 
        FTB-OMP ( $d, p + 1, r, \lceil L/d \rceil, \xi, \epsilon$ )
    end
end
end

```

ALGORITHM 1: Pseudocode for FTB-OMP.

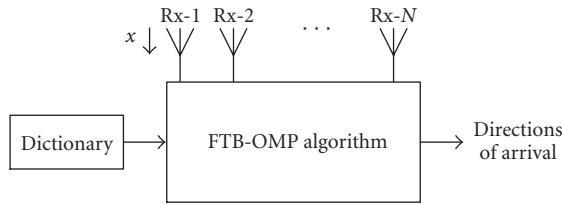


FIGURE 3: EDAMP estimation of DOA.

alignment, reproducing TB-OMP. A value $0 < \xi < 1$ represents a compromise between the number of nodes for OMP (r nodes), and for TB-OMP $((L^{r+1} - 1)/(L - 1))$. Further reduction on the tree-size is achieved by using decay parameter d . This reduction makes the algorithm more competitive even without tree-pruning ($\xi = 0$). A pseudocode for FTB-OMP is given in Algorithm 1.

4. SYSTEM MODEL

In our system model for DOA estimation, we consider an adaptive antenna array of N elements as in Figure 3. The input signal is assumed to be a plane wave or equivalently it can be decomposed into plane waves.

Let \mathbf{x} be the received vector formed by the received signal at each antenna element. For a uniform linear array the dictionary \mathcal{D} can be obtained as

$$\mathcal{D} = \begin{Bmatrix} 1 & 1 & \dots & 1 \\ e^{j\psi_1} & e^{j\psi_2} & \dots & e^{j\psi_M} \\ \vdots & \vdots & \vdots & \vdots \\ e^{j(N-1)\psi_1} & e^{j(N-1)\psi_2} & \dots & e^{j(N-1)\psi_M} \end{Bmatrix}, \quad (8)$$

where ψ_i is the phase difference between elements of array, when the signal arrives from angle θ_i . The relation between ψ_i and θ_i is given as $\psi_i = (2\pi l/\lambda) \cos(\theta_i)$, where λ is the wavelength and l is the array spacing between the antenna elements. For the case in (8), the possible range of DOA is divided into M sections. These sections form the dictionary \mathcal{D} . Also for presentation purposes, we stick to the notation of [2] and define $u = \cos(\theta_i)$.

Depending on the DOA, the received signal vector of size $N \times 1$ will be a linear combination of the columns of \mathcal{D} plus noise. Hence, detecting the DOA problem will reduce to finding correct linear combination of the columns of \mathcal{D} .

When the signal arrives from an individual angle only, the problem is straightforward and algorithm chooses the column of \mathcal{D} , which has the maximum inner product with the received vector \mathbf{x} . However when the signal arrives from more than one angle, \mathbf{x} is a linear combination of columns of \mathcal{D} and trying every possible linear combination would give the ML solution. On the other hand, this would bring formidable complexity to the system. By employing the FTB-OMP algorithm presented in the previous section, we propose a heuristic approximation to ML solution.

FTB-OMP algorithm selects the columns of \mathcal{D} which are estimated to form \mathbf{x} , and these columns correspond to the DOA. FTB-OMP also returns to the coefficients of these columns, which represent the amplitude of the corresponding DOA.

There are three main advantages of the application of FTB-OMP.

- (i) It does not require the number of directions to be estimated. By comparing the amplitude in \mathbf{x} and amplitude of the resolved signals defined by the space spanned by the columns of \mathcal{D} , which have already been chosen by the algorithm, it is capable of deciding whether all the components are resolved or not. Considering that most of the spectral and subspace algorithms require the number of directions as an input, this is a very important advantage.
- (ii) The algorithm allows flexibility between complexity and resolution property. By increasing the search depth, a closer solution to ML can be achieved, by decreasing the search depth algorithm running time can be decreased. But for both cases, it is computationally advantageous to the subspace-based algorithms, since it works on spectral domain and does not require eigenvalue decomposition.
- (iii) In EDAMP, not the signal subspaces but the amplitudes of the received signals are used. As a result, system performance is robust to correlation between the inputs from different angles.

In the next section we support these advantages by simulation results.

5. SIMULATION RESULTS

In the simulations we consider a 10-element uniform linear array (ULA) that has element separation of $\lambda/2$ as shown in Figure 4. The SNR values correspond to the signal-to-noise ratios at the input of each antenna element and they are assumed to be the same. However the noise at each element is assumed to be independent identically distributed (i.i.d.) additive white Gaussian noise (AWGN). The system SNR is much higher than the SNR at each element. Hence, low-SNR results presented in the paper are of practical interest as well.

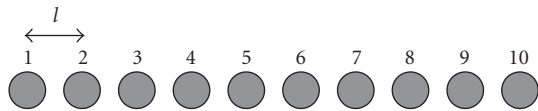


FIGURE 4: Array structure of ULA.

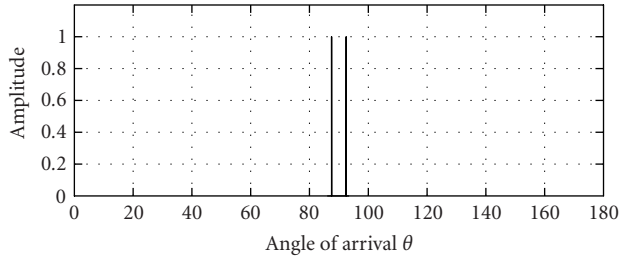


FIGURE 5: Arrival angles $\theta_1 = 87.52^\circ$, $\theta_2 = 92.48^\circ$.

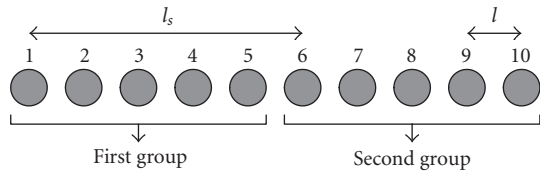


FIGURE 6: Subarrays for ESPRIT: first five elements of the original array form the first subarray, and last five elements of the original array form the second subarray.

Unless stated otherwise, two different signal directions with $u_1 = 0.0433$ and $u_2 = -0.0433$ (the minimum distance that can be resolved for a 10-element ULA [2]) are considered. The amplitudes in both directions are assumed to be the same. These u values correspond to 87.52° and 92.48° . As shown in Figure 5, the range of estimation is between 0° and 180° .

In the subspace-based algorithms, for the convergence of the eigenvalues, 100 independent snapshots are used. The results are averaged over 1000 Monte Carlo simulations.

Other than the proposed EDAMP algorithm as described in the previous section, Bartlett [4], MVDR [2], MUSIC [5], and ESPRIT [6] algorithms have also been considered. These algorithms have been simulated with the parameters defined above, and all of the results presented in this work about these algorithms have been calibrated with the results on their performances presented in the literature prior to this work [1, 2].

Bartlett algorithm is generated as a traditional beamformer with 10 elements, steered along different angles and acquiring the maximum amplitude points. Application of MVDR is simply using MVDR beamformer coefficients instead of uniform coefficients of Bartlett. For MUSIC, the parameters described in [2, 5] are employed for 10 antenna elements.

For the ESPRIT algorithm, the antenna array is divided into two subarrays, one being the shifted version of the other in space. The constant phase shift between two subarrays is employed for the resolution. For simulations, 5-element-shifted ESPRIT is considered as shown in Figure 6.

TABLE 1: Parameters of FTB-OMP algorithm used in EDAMP simulations.

Parameter	Value
Tree-pruning (ξ)	0.25
Number of branches (L)	100
Decaying parameter (d)	10
Maximum iteration (r)	3

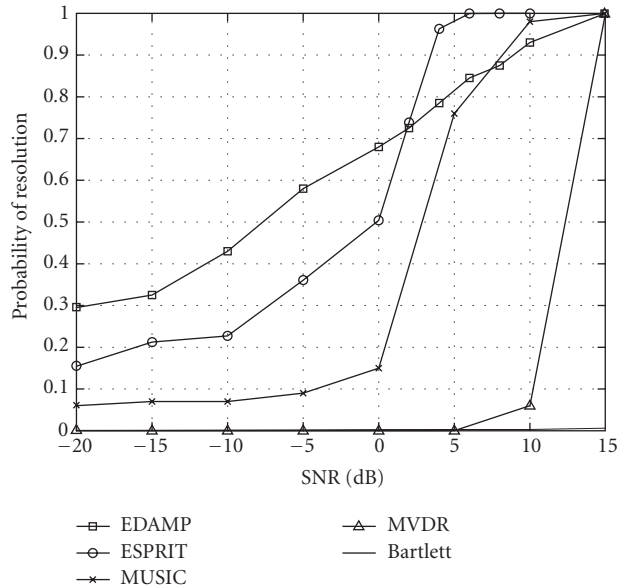


FIGURE 7: Probability of resolution versus SNR for uncorrelated inputs.

In Table 1, the parameters used for FTB-OMP algorithm employed in the simulations are given. With these parameters, EDAMP requires much less computational time when compared to ESPRIT and MUSIC. In terms of floating point operations in MATLAB simulation platform, EDAMP requires approximately half the number of flops required by ESPRIT, and one fourth the number of flops required by MUSIC.

5.1. Uncorrelated inputs

We first look at the case when the signals arriving from different angles are uncorrelated. In Figure 7, the novel EDAMP algorithm is compared with all four algorithms mentioned above. As can be seen in Figure 7, EDAMP performs well especially in the low-SNR region and the probability of resolution increases linearly with SNR. For uncorrelated channels at low SNR, EDAMP outperforms every other algorithm, and at high SNR, ESPRIT performs the best.

In Figure 8, root mean square error (RMSE) in the estimated angles is shown. RMSE is normalized by the null-to-null beamwidth (BW_{NN}) of the 10-element antenna array. As it is seen in Figure 8, at low SNR EDAMP outperforms ESPRIT and at high SNR, ESPRIT is better in terms of RMSE performance.

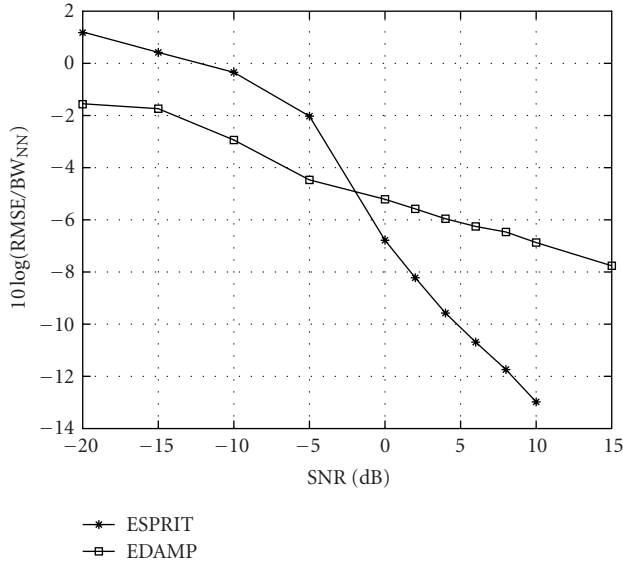


FIGURE 8: RMSE of DOA normalized by null-to-null beamwidth for uncorrelated inputs.

Next, the effect of angular separation on the probability of resolution is investigated. In Figure 9, it is depicted that for $\text{SNR} = 3$ dB, EDAMP can resolve more closely separated signals when compared to ESPRIT. Also in Figure 9, we can see another limitation of ESPRIT. In ESPRIT algorithm, the antenna array is divided into two symmetric subarrays. The resolution property is highly dependent on the distance between the first element of the first array and first element of the second array, which is denoted by l_s [2]. The ESPRIT scheme that we employ in our simulations is the one with highest resolution available for a 10-element antenna array [2]. However, in ESPRIT algorithm, the resolvable angles are limited by the relation

$$-\frac{1}{l_s} < u < \frac{1}{l_s}. \quad (9)$$

For the scheme employed which is shown in Figure 6, $l_s = 5$. Since

$$-\frac{1}{5} < u < \frac{1}{5}, \quad (10)$$

the largest value of Δu , for resolution is $1/5 + 1/5 = 0.4$. It is clearly seen that for $u > 0.4$, the performance of ESPRIT degrades very fast. On the other hand, EDAMP has no such limitation. One could select an ESPRIT scheme with smaller l_s hence increasing the resolvable range, but this would result in lower probability of resolution and worse RMSE in the resolvable range [2, 6].

5.2. Correlated inputs

Above we considered the case when two signals arriving from different angles were uncorrelated. Here, we investigate the effect of correlation on the system performance. The perfor-

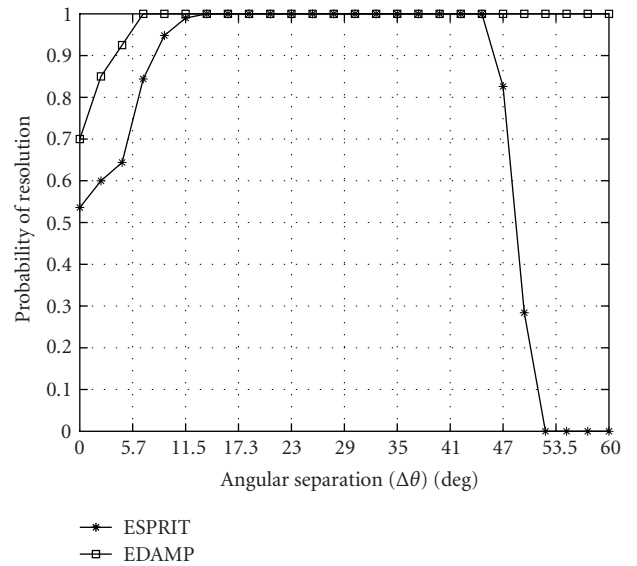


FIGURE 9: Probability of resolution versus angular separation for uncorrelated inputs for $\text{SNR} = 3$ dB.

mance of subspace algorithms, namely MUSIC and ESPRIT are highly dependent on the correlation between input signals arriving from different angles [1, 2, 5, 6]. This is a natural outcome of subspace algorithms making use of eigenspace decomposition in order to separate noise, signal, and interference.

On the other hand, the performance of EDAMP is independent of correlation in the signals, since its resolving power depends solely on the amplitudes in different directions. This is supported by the results of Figures 10 and 11. Even for 90% correlation, the performance of EDAMP is the same as its performance with uncorrelated channels. However, as shown in Figures 10 and 11, the performances of MUSIC and ESPRIT are severely degraded with increased correlation.

It is seen that for highly correlated signals EDAMP resolution performance is much better than subspace algorithms such as MUSIC and ESPRIT.

5.3. Effect of number of snapshots

In wireless communications, especially for real-time applications, delays in the system are very critical. In DOA estimation, a number of snapshots is required for the estimation to be accurate [1]. When the number of snapshots increases, the delay in the system increases. It is well known that with insufficient number of snapshots, traditional DOA algorithms perform poorly. In EDAMP, snapshots are only utilized for running the algorithm again and averaging the estimations. For known signals, the snapshots can be utilized to decrease the SNR by averaging the signals from different snapshots. The number of snapshots, therefore, is not very critical as in the case of subspace algorithms. Here we investigate the effect of number of snapshots by decreasing it from 100 to 10, and the effect of number of snapshots when the SNR is 15 dB.

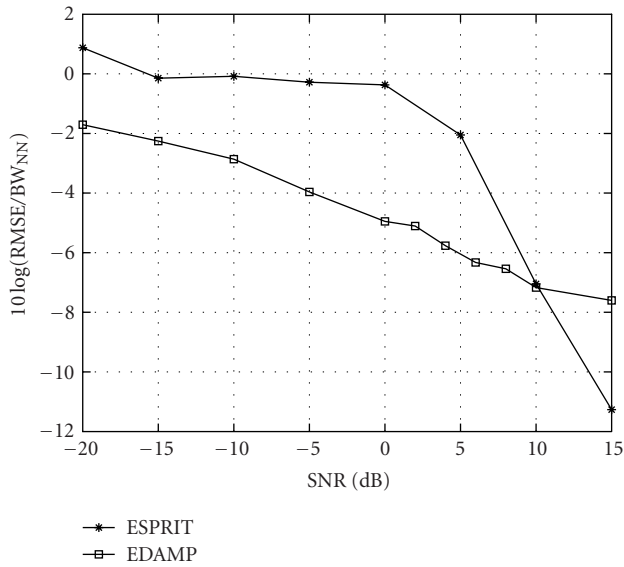


FIGURE 10: RMSE of DOA normalized by null-to-null beamwidth for 90%-correlated inputs.

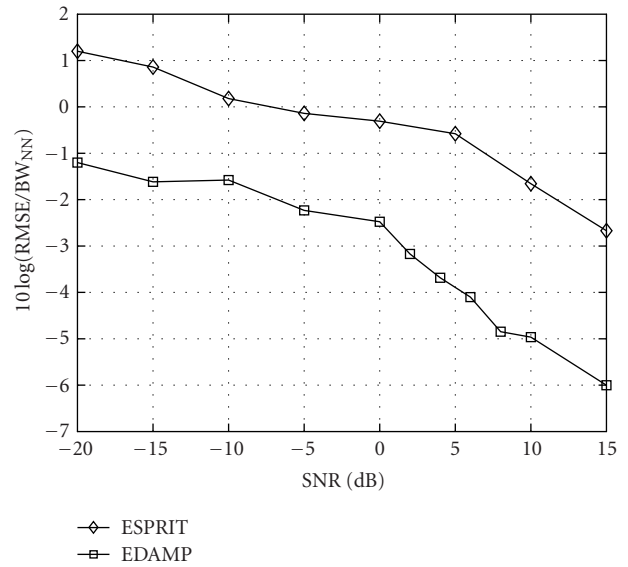


FIGURE 12: RMSE of DOA normalized by null-to-null beamwidth for 90%-correlated inputs with 10 snapshots.

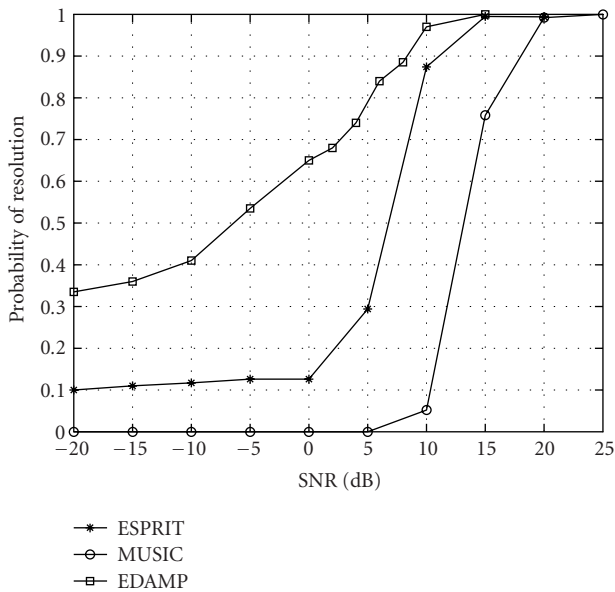


FIGURE 11: Probability of resolution versus SNR for 90%-correlated inputs.

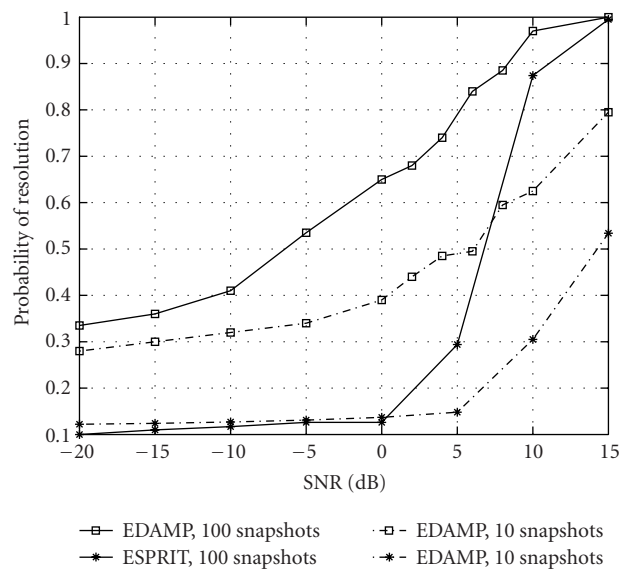


FIGURE 13: Comparison of probabilities of resolution of 90%-correlated inputs for 10 and 100 snapshots.

In Figures 12, 13, and 14 it is clearly depicted that EDAMP performs much better for low number of snapshots. Even at 10 snapshots, EDAMP shows acceptable performance, which makes EDAMP even more valuable for applications requiring short delays.

6. CONCLUSIONS

In this paper, we have presented a novel DOA estimator, EDAMP, which employs a based basis selection algorithm,

namely FTB-OMP. Many advantages of EDAMP when compared to the traditional algorithms are presented, which can be summarized as follows.

The EDAMP algorithm gives directions of arrival and their corresponding amplitudes as output, so it does not require postprocessing to detect amplitudes after detecting directions. On the other hand, the algorithm does not need preprocessing since it does not require the number of DOA as input.

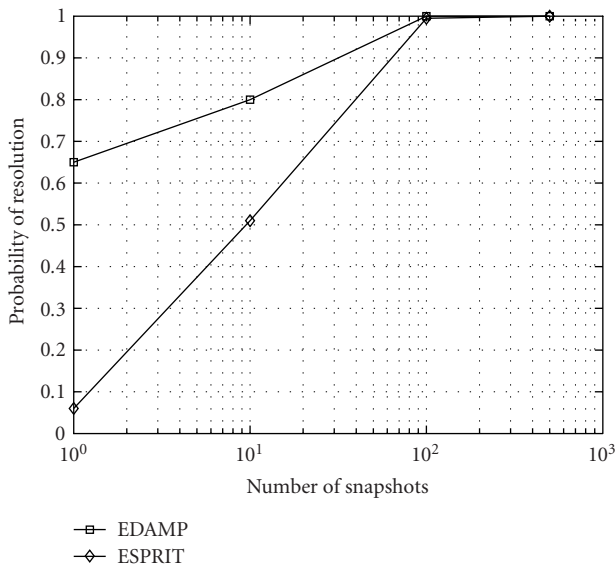


FIGURE 14: Probability of resolution versus number of snapshots for 90%-correlated inputs with SNR = 15 dB.

EDAMP is not affected by the correlations in the signals from different DOA, hence it is expected to perform better in multipath situations when compared to traditional techniques.

Since it is a heuristic approach to ML solution, it gives good resolution properties even at low-SNR situations. It also requires very few snapshots, when compared to subspace algorithms, thus decreasing processing time.

Many different variations of basis selection algorithms can be utilized for DOA estimation or similar estimation problems employing overcomplete sets and sparse solutions. Hence the idea presented in this paper promises many possible future research areas in several areas of signal processing, other than DOA estimation.

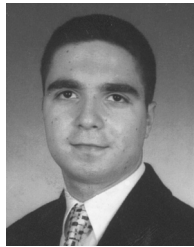
REFERENCES

- [1] L. C. Godara, "Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations," *Proc. IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.
- [2] H. L. Van Trees, *Optimum Array Processing*, Wiley, New York, NY, USA, 2002.
- [3] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 38, no. 7, pp. 1132–1143, 1990.
- [4] V. A. N. Barroso, M. J. Rendas, and J. P. Gomes, "Impact of array processing techniques on the design of mobile communication systems," in *Proc. 7th IEEE Mediterranean Electrotechnical Conference*, vol. 3, pp. 1291–1294, Antalya, Turkey, April 1994.
- [5] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. 34, no. 3, pp. 276–280, 1986.
- [6] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.
- [7] G. Z. Karabulut, L. Moura, D. Panario, and A. Yongacoglu, "Efficient tree search based orthogonal matching pursuit algorithm with adaptive resolution," Internal report, University of Ottawa, Ottawa, Ontario, Canada, May 2004.
- [8] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [9] A. M. Kondoz, *Digital Speech*, Wiley, New York, NY, USA, 1996.
- [10] S. S. Chen and D. L. Donoho, "Application of basis pursuit in spectrum estimation," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '98)*, vol. 3, pp. 1865–1868, Seattle, Wash, USA, May 1998.
- [11] H. Sawada, R. Mukai, and S. Makino, "Direction of arrival estimation for multiple source signals using independent component analysis," in *Proc. 7th IEEE International Symposium on Signal Processing and Its Applications (ISSPA '03)*, vol. 2, pp. 411–414, Paris, France, July 2003.
- [12] M. Buhren, M. Pesavento, and J. E. Bohme, "A new approach to array interpolation by generation of artificial shift invariances: interpolated ESPRIT," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, vol. 5, pp. 205–208, Hong Kong, China, April 2003.
- [13] J. Xin and A. Sano, "Computationally efficient subspace-based method for direction-of-arrival estimation without eigendecomposition," *IEEE Trans. Signal Processing*, vol. 52, no. 4, pp. 876–893, 2004.
- [14] P. Charge, Y. Wang, and J. Saillard, "An extended cyclic MUSIC algorithm," *IEEE Trans. Signal Processing*, vol. 51, no. 7, pp. 1695–1701, 2003.
- [15] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm," *IEEE Trans. Signal Processing*, vol. 45, no. 3, pp. 600–616, 1997.
- [16] Y. H. Chan, "An efficient weight optimization algorithm for image representation using nonorthogonal basis images," *IEEE Signal Processing Lett.*, vol. 5, no. 8, pp. 193–195, 1998.
- [17] R. Gribonval, E. Bacry, S. Mallat, P. Depalle, and X. Rodet, "Analysis of sound signals with high resolution matching pursuit," in *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pp. 125–128, Paris, France, June 1996.
- [18] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 713–718, 1992.
- [19] I. Daubechies, "Time-frequency localization operators: a geometric phase space approach," *IEEE Trans. Inform. Theory*, vol. 34, no. 4, pp. 605–612, 1988.
- [20] J. Adler, B. D. Rao, and K. Kreutz-Delgado, "Comparison of basis selection methods," in *Proc. 30th IEEE Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 252–257, Pacific Grove, Calif, USA, November 1996.
- [21] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proc. 27th IEEE Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 40–44, Pacific Grove, Calif, USA, November 1993.
- [22] S. F. Cotter and B. D. Rao, "Application of tree-based searches to matching pursuit," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '01)*, vol. 6, pp. 3933–3936, Salt Lake City, Utah, USA, May 2001.

Güneş Z. Karabulut received the B.S. degree in electronics and electrical engineering from Boğaziçi University, Istanbul, Turkey, in 2000. She received her M.A.Sc. degree in electrical engineering from the University of Ottawa, Ontario, Canada. Currently she is working towards her Ph.D. degree at the University of Ottawa, Ontario, Canada. From 1999 to 2000 she was working at Boğaziçi University Signal and Image Processing Laboratory, where she worked on motion estimation algorithms. She is presently employed as a Research Assistant at CASP Group, University of Ottawa. Her research interests include coding theory, basis selection algorithms, sparse signal representations, and adaptive time/frequency decompositions. Ms. Karabulut is a Member of the IEEE Information Theory Society.



Tolga Kurt received his B.S. and M.S. degrees from Boğaziçi University, Istanbul, Turkey, in 2000 and 2002, respectively. Currently he is working towards his Ph.D. degree at the University of Ottawa, Ontario, Canada. From 2000 to 2002, he was working at Turkcell Telecommunication Ltd., Istanbul, Turkey. He worked as a Research Assistant at CASP Group, University of Ottawa, between 2002 and 2004. He is now with Marconi Wireless R&D, Ottawa, Canada. His research interests include OFDM systems, smart antennas, and radio over fiber systems.



Abbas Yongaçoğlu received the B.S. degree from Boğaziçi University, Turkey, in 1973, the M. Eng. degree from the University of Toronto, Canada, in 1975, and the Ph.D. degree from the University of Ottawa, Canada, in 1987, all in electrical engineering. He worked as a researcher and a System Engineer at TUBITAK Marmara Research Institute, Turkey, Philips Research Labs, Holland, and Miller Communications Systems, Ottawa. In 1987 he joined the University of Ottawa as an Assistant Professor. He became an Associate Professor in 1992 and a Full Professor in 1996. His area of research is digital communications with emphasis on modulation, coding, equalization, and multiple access for wireless and high-speed wireline communications.



Blind Multiuser Detection for Long-Code CDMA Systems with Transmission-Induced Cyclostationarity

Tongtong Li

Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA
Email: tongli@egr.msu.edu

Weiguo Liang

Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA
Email: liangwg@egr.msu.edu

Zhi Ding

Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, USA
Email: zding@ece.ucdavis.edu

Jitendra K. Tugnait

Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849, USA
Email: tugnajk@eng.auburn.edu

Received 30 April 2004; Revised 5 August 2004

We consider blind channel identification and signal separation in long-code CDMA systems. First, by modeling the received signals as cyclostationary processes with modulation-induced cyclostationarity, long-code CDMA system is characterized using a time-invariant system model. Secondly, based on the time-invariant model, multistep linear prediction method is used to reduce the intersymbol interference introduced by multipath propagation, and channel estimation then follows by utilizing the nonconstant modulus precoding technique with or without the matrix-pencil approach. The channel estimation algorithm without the matrix-pencil approach relies on the Fourier transform and requires additional constraint on the code sequences other than being a nonconstant modulus. It is found that by introducing a random linear transform, the matrix-pencil approach can remove (with probability one) the extra constraint on the code sequences. Thirdly, after channel estimation, equalization is carried out using a cyclic Wiener filter. Finally, since chip-level equalization is performed, the proposed approach can readily be extended to multirate cases, either with multicode or variable spreading factor. Simulation results show that compared with the approach using the Fourier transform, the matrix-pencil-based approach can significantly improve the accuracy of channel estimation, therefore the overall system performance.

Keywords and phrases: long-code CDMA, multiuser detection, cyclostationarity.

1. INTRODUCTION

In addition to intersymbol and interchip interference, one of the key obstacles to signal detection and separation in CDMA systems is the detrimental effect of multiuser interference (MUI) on the performance of the receivers and the overall communication system. Compared to the conventional single-user detectors where interfering users are modeled as

noise, significant improvement can be obtained with multiuser detectors where MUI is explicitly part of the signal model [1].

In literature [2], if the spreading sequences are periodic and repeat every information symbol, the system is referred to as short-code CDMA, and if the spreading sequences are aperiodic or essentially pseudorandom, it is known as long-code CDMA. Since multiuser detection relies on the cyclostationarity of the received signal, which is significantly complicated by the time-varying nature of the long-code system, research on multiuser detection has largely been limited to short-code CDMA for some time, see, for

This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

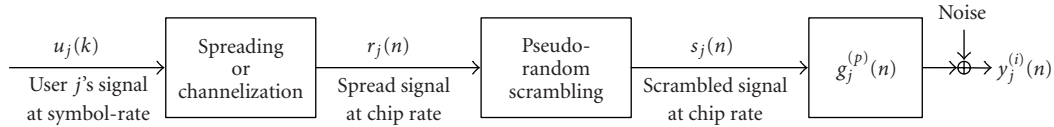


FIGURE 1: Block diagram of a long-code DS-CDMA system.

example, [3, 4, 5, 6, 7] and the references therein. On the other hand, due to its robustness and performance stability in frequency fading environment [2], long code is widely used in virtually all operational and commercially proposed CDMA systems, as shown in Figure 1. Actually, each user's signal is first spread using a code sequence spanning over just one symbol or multiple symbols. The spread signal is then further scrambled using a long-periodicity pseudorandom sequence. This is equivalent to the use of an aperiodic (long) coding sequence as in *long-code CDMA system*, and the chip-rate sampled signal and MUIs are generally modeled as time-varying vector processes [8]. The time-varying nature of the received signal model in the long-code case severely complicates the equalizer development approaches, since consistent estimation of the needed signal statistics cannot be achieved by time-averaging over the received data record.

More recently, both training-based (e.g., [9, 10, 11]) and blind (e.g., [8, 12, 13, 14, 15, 16, 17, 18, 19]) multiuser detection methods targeted at the long-code CDMA systems have been proposed. In this paper, we will focus on blind channel estimation and user separation for long-code CDMA systems. Based on the channel model, most existing blind algorithms can roughly be divided into three classes.

- (i) *Symbol-by-symbol approaches.* As in long-code systems, each user's spreading code changes for every information symbol, symbol-by-symbol approaches (see [8, 17, 18, 19], e.g.) process each received symbol individually based on the assumption that channel is invariant in each symbol. In [8, 17, 18], channel estimation and equalization is carried out for each individual received symbol by taking instantaneous estimates of signal statistics based on the sample values of each symbol. In [19], based on the BCJR algorithm, an iterative turbo multiuser detector was proposed.
- (ii) *Frame-by-frame approaches.* Algorithms in this category (see [15, 20], e.g.) stack the total received signal corresponding to a whole frame or slot into a long vector, and formulate a deterministic channel model. In [15], computational complexity is reduced by breaking the big matrix into small blocks and implementing the inversion "locally." As can be seen, the "localization" is similar to the process of the symbol-by-symbol approach. And the work is extended to fast fading channels in [20].
- (iii) *Chip-level equalization.* By taking chip-rate information as input, the time-varying effect of the pseudo-random sequence is absorbed into the input sequence.

With the observation that channels remain approximately stationary over each time slot, the underlying channel, therefore, can be modelled as a time-invariant system, and at the receiver, chip-level equalization is performed. Please refer to [14, 21, 22, 23] and the references therein.

In all these three categories, one way or another, the time-varying channel is "converted" or "decomposed" into *time-invariant* channels.

In this paper, the long-code CDMA system is characterized as a time-invariant MIMO system as in [14, 23]. Actually, the received signals and MUIs can be modeled as cyclostationary processes with modulation-induced cyclostationarity, and we consider blind channel estimation and signal separation for long-code CDMA systems using multistep linear predictors. Linear prediction-based approach for MIMO model was first proposed by Slock in [24], and developed by others in [25, 26, 27, 28]. It has been reported [26, 28] that compared with subspace methods, linear prediction methods can deliver more accurate channel estimates and are more robust to overmodeling in channel order estimate. In this paper, multistep linear prediction method is used to separate the intersymbol interference introduced by multipath channel, and channel estimation is then performed using non-constant modulus precoding technique both with and without the matrix-pencil approach [29, 30]. The channel estimation algorithm without the matrix-pencil approach relies on the Fourier transform, and requires additional constraint on the code sequences other than being nonconstant modulus. It is found that by introducing a random linear transform, the matrix-pencil approach can remove (with probability one) the extra constraint on the code sequences. After channel estimation, equalization is carried out using a cyclic Wiener filter. Finally, since chip-level equalization is performed, the proposed approach can readily be extended to multirate cases, either with multicode or variable spreading factor. Simulation results show that compared with the approach using the Fourier transform, the matrix-pencil-based approach can significantly improve the accuracy of channel estimation, therefore the overall system performance.

2. SYSTEM MODEL

Consider a DS-CDMA system with M users and K receive antennas, as shown in Figure 2. Assume the processing gain is N , that is, there are N chips per symbol. Let $u_j(k)$ ($j = 1, \dots, M$) denote user j 's k th symbol. Assume that the code sequence extends over L_c symbols. Let $\mathbf{c}_j =$

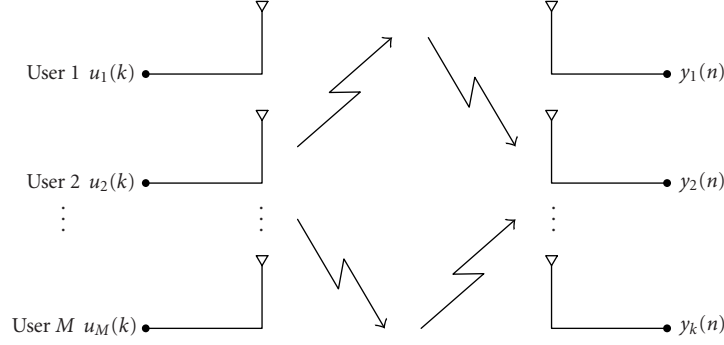


FIGURE 2: Block diagram of a MIMO system.

$[c_j(0), c_j(1), \dots, c_j(N-1), c_j(N), \dots, c_j(L_c N - 1)]$ denote user j 's spreading code sequence. For notations used for each individual user, please refer to Figure 1. When k is a multiple of L_c , the spread signal (at chip rate) with respect to the signal block $[u_j(k), \dots, u_j(k + L_c - 1)]$ is

$$\begin{aligned} & [r_j(kN), \dots, r_j((k + L_c)N - 1)] \\ & = [u_j(k)c_j(0), \dots, u_j(k)c_j(N-1), \dots, \\ & \quad u_j(k + L_c - 1)c_j((L_c - 1)N), \dots, \\ & \quad u_j(k + L_c - 1)c_j(L_c N - 1)]. \end{aligned} \quad (1)$$

The successive scrambling process is achieved by

$$\begin{aligned} & [s_j(kN), \dots, s_j((k + L_c)N - 1)] \\ & = [r_j(kN), \dots, r_j((k + L_c)N - 1)] \\ & \quad \cdot [d_j(kN), d_j(kN + 1), \dots, d_j((k + L_c)N - 1)], \end{aligned} \quad (2)$$

where “ \cdot ” stands for point-wise multiplication, and $[d_j(kN), d_j(kN + 1), \dots, d_j(kN + N - 1)]$ denotes the chip-rate scrambling sequence with respect to symbol $u_j(k)$. Defining

$$\begin{aligned} & [v_j(kN), \dots, v_j((k + L_c)N - 1)] \\ & \triangleq [u_j(k)d_j(kN), \dots, u_j(k)d_j(kN + N - 1), \dots, \\ & \quad u_j(k + L_c - 1)d_j((k + L_c - 1)N), \dots, \\ & \quad u_j(k + L_c - 1)d_j((k + L_c)N - 1)], \end{aligned} \quad (3)$$

we get

$$\begin{aligned} & [s_j(kN), s_j(kN + 1), \dots, s_j((k + L_c)N - 1)] \\ & = [v_j(kN), v_j(kN + 1), \dots, v_j((k + L_c)N - 1)] \\ & \quad \cdot [c_j(0), c_j(1), \dots, c_j(L_c N - 1)]. \end{aligned} \quad (4)$$

If we regard the chip rate $v_j(n)$ as the input signal of user j , then $s_j(n)$ is the precoded transmit signal corresponding to the j th user and

$$s_j(n) = v_j(n)c_j(n), \quad n \in \mathbb{Z}, \quad j = 1, 2, \dots, M, \quad (5)$$

where $c_j(n) = c_j(n + L_c N)$ serves as a periodic precoding sequence with period $L_c N$. We note that this form of periodic precoding has been suggested by Serpedin and Giannakis in [31] to introduce cyclostationarity in the transmit signal, thereby making blind channel identification based on second-order statistics in symbol-rate-sampled single-carrier system possible. More general idea of transmitter-induced cyclostationarity has been suggested previously in [32, 33]. In [34], nonconstant precoding technique has been applied to blind channel identification and equalization in OFDM-based multiantenna systems.

Based on Figures 1 and 2, the received chip-rate signal at the p th antenna ($p = 1, 2, \dots, K$) can be expressed as

$$y_p(n) = \sum_{j=1}^M \sum_{l=0}^{L-1} g_j^{(p)}(l) s_j(n-l) + w_p(n), \quad (6)$$

where $L - 1$ is the maximum multipath delay spread in chips, $\{g_j^{(p)}(l)\}_{l=0}^{L-1}$ denotes the channel impulse response from j th transmit antenna to p th receive antenna, and $w_p(n)$ is the p th antenna additive white noise. Let $\mathbf{s}(n) = [s_1(n), s_2(n), \dots, s_M(n)]^T$ be the precoded signal vector. Collect the samples at each receive antenna and stack them into a $K \times 1$ vector, we get the following *time-invariant* MIMO system model:

$$\mathbf{y}(n) = [y_1(n), y_2(n), \dots, y_K(n)]^T = \sum_{l=0}^{L-1} \mathbf{H}(l) \mathbf{s}(n-l) + \mathbf{w}(n), \quad (7)$$

where

$$\mathbf{H}(l) = \begin{bmatrix} g_1^{(1)}(l) & g_2^{(1)}(l) & \cdots & g_M^{(1)}(l) \\ g_1^{(2)}(l) & g_2^{(2)}(l) & \cdots & g_M^{(2)}(l) \\ \vdots & \vdots & \ddots & \vdots \\ g_1^{(K)}(l) & g_2^{(K)}(l) & \cdots & g_M^{(K)}(l) \end{bmatrix}_{K \times M} \quad (8)$$

$$\text{and } \mathbf{w}(n) = [w_1(n), w_2(n), \dots, w_K(n)]^T.$$

Defining $\mathcal{H}(z) = \sum_{l=0}^{L-1} \mathbf{H}(l)z^{-l}$, it then follows that

$$\mathbf{y}(n) = \mathcal{H}(z)\mathbf{s}(n) + \mathbf{w}(n) \triangleq \mathbf{y}_s(n) + \mathbf{w}(n). \quad (9)$$

In the following section, channels are estimated based on the desired user's code sequence and the following assumptions.

- (A1) The multiuser sequences $\{u_j(k)\}_{j=1}^M$ are zero mean, mutually independent, and i.i.d. Take $E\{|u_j(k)|^2\} = 1$ by absorbing any nonidentity variance of $u_j(k)$ into the channel.
- (A2) The scrambling sequences $\{d_j(k)\}_{j=1}^M$ are mutually independent i.i.d. BPSK sequences, independent of the information sequences.
- (A3) The noise is zero mean Gaussian, independent of the information sequences, with $E\{\mathbf{w}(k+l)\mathbf{w}^H(k)\} = \sigma_w^2 \mathbf{I}_K \delta(l)$ where \mathbf{I}_K is the $K \times K$ identity matrix.
- (A4) $\mathcal{H}(z)$ is irreducible when regarded as a polynomial matrix of z^{-1} , that is, $\text{Rank}\{\mathcal{H}(z)\} = M$ for all complex z except $z = 0$.

3. BLIND CHANNEL IDENTIFICATION BASED ON MULTISTEP LINEAR PREDICTORS

In this section, first, multistep linear prediction method is used to resolve the intersymbol interference introduced by multipath channel. Secondly, based on the ISI-free MIMO model, two channel estimation approaches are proposed by exploiting the advantage of nonconstant modulus precoding: one uses the Fourier analysis, and the other is based on the matrix-pencil technique.

3.1. ISI reduction and separation based on multistep linear predictors

Based on the results in [6, 28, 35], it can be shown that under (A1), (A2), (A3), and (A4), finite length predictors exist for the noise-free channel observations

$$\mathbf{y}_s(n) = \mathcal{H}(z)\mathbf{s}(n) = \sum_{l=0}^{L-1} \mathbf{H}(l)\mathbf{s}(n-l) \quad (10)$$

such that it has the following canonical representation:

$$\mathbf{y}_s(n) = \sum_{i=l}^{L_l} A_{n,i}^{(l)} \mathbf{y}_s(n-i) + \mathbf{e}(n|n-l), \quad l = 1, 2, \dots, \quad (11)$$

for some $L_l \leq M(L-1) + l - 1$, where the l -step ahead linear prediction error $\mathbf{e}(n|n-l)$ is given by

$$\mathbf{e}(n|n-l) = \sum_{i=0}^{l-1} \mathbf{H}(i)\mathbf{s}(n-i) \quad (12)$$

satisfying

$$E\{\mathbf{e}(n|n-l)\mathbf{y}_s^H(n-m)\} = 0 \quad \forall m \geq l. \quad (13)$$

Therefore, based on (11) and (13), the coefficient matrices $A_{n,i}^{(l)}$ s can be determined from

$$E\{\mathbf{y}_s(n)\mathbf{y}_s^H(n-m)\} = \sum_{i=l}^{L_l} A_{n,i}^{(l)} E\{\mathbf{y}_s(n-i)\mathbf{y}_s^H(n-m)\} \quad \forall m \geq l. \quad (14)$$

Actually, consider

$$\begin{aligned} \mathbf{R}_s(n, k) &\triangleq E\{\mathbf{s}(n)\mathbf{s}^H(n-k)\} \\ &= \text{diag}\left[|c_1(n)|^2, \dots, |c_M(n)|^2\right] \delta(k). \end{aligned} \quad (15)$$

It follows that $\mathbf{R}_s(n, k)$ is periodic with respect to n :

$$\mathbf{R}_s(n, k) = \mathbf{R}_s(n + L_c N, k) \quad (16)$$

(where N is the processing gain) since $c_j(n) = c_j(n + L_c N)$ for $j = 1, 2, \dots, M$. Note that $\mathbf{R}_s(n, k) = 0$ for any $k \neq 0$. Defining $\mathbf{R}_s(n) \triangleq \mathbf{R}_s(n, 0)$, then

$$\mathbf{R}_s(n) = \mathbf{R}_s(n + L_c N). \quad (17)$$

It follows that the $K \times K$ autocorrelation matrix of the noise-free channel output

$$\begin{aligned} \mathbf{R}_{y_s}(n, k) &\triangleq E\{\mathbf{y}_s(n)\mathbf{y}_s^H(n-k)\} \\ &= \sum_{l=0}^{L-1} \mathbf{H}(l)\mathbf{R}_s(n-l)\mathbf{H}^H(l-k) \end{aligned} \quad (18)$$

is also periodic with period $L_c N$ in this circumstance. In (14), letting $m = l, l+1, \dots, L_l$, we have

$$\begin{aligned} &[A_{n,l}^{(l)}, A_{n,l+1}^{(l)}, \dots, A_{n,L_l}^{(l)}] \\ &= [\mathbf{R}_{y_s}(n, l), \dots, \mathbf{R}_{y_s}(n, L_l)] \mathcal{R}^\#(n, l, L_l), \end{aligned} \quad (19)$$

where $\#$ stands for pseudoinverse and $\mathcal{R}(n, l, L_l)$ is a $(L_l - l + 1)K \times (L_l - l + 1)K$ matrix with its (i, j) th $K \times K$ block element as $\mathbf{R}_{y_s}(n-l-i+1, j-i) = E\{\mathbf{y}_s(n-l-i+1)\mathbf{y}_s^H(n-l-j+1)\}$ for $i, j = 1, \dots, L_l - l + 1$. And $\mathbf{R}_{y_s}(n, k)$ can be estimated from

$$\mathbf{R}_{y_s}(n, k) \triangleq E\{\mathbf{y}(n)\mathbf{y}^H(n-k)\} = \mathbf{R}_{y_s}(n, k) + \sigma_n^2 \mathbf{I}_K \delta(k) \quad (20)$$

through noise variance estimation, please see [6, 28] for more details.

Now define $\bar{\mathbf{e}}_l(n) \triangleq \mathbf{e}(n|n-l) - \mathbf{e}(n|n-l+1)$ and let

$$\mathbf{E}(n) \triangleq \begin{bmatrix} \bar{\mathbf{e}}_{d+1}(n+d) \\ \bar{\mathbf{e}}_d(n+d-1) \\ \vdots \\ \bar{\mathbf{e}}_2(n+1) \\ \bar{\mathbf{e}}_1(n-1) \end{bmatrix}. \quad (21)$$

It then follows from (12) that

$$\mathbf{E}(n) = \begin{bmatrix} \mathbf{H}(d) \\ \mathbf{H}(d-1) \\ \vdots \\ \mathbf{H}(0) \end{bmatrix} \mathbf{s}(n) \triangleq \tilde{\mathbf{H}}\mathbf{s}(n), \quad (22)$$

where

$$\tilde{\mathbf{H}} \triangleq \begin{bmatrix} \mathbf{H}(d) \\ \mathbf{H}(d-1) \\ \vdots \\ \mathbf{H}(0) \end{bmatrix}. \quad (23)$$

Thus, we obtained an ISI-free MIMO model (22).

3.2. Channel estimation through the Fourier analysis

Consider the correlation matrix of $\mathbf{E}(n)$,

$$\begin{aligned} \mathbf{R}_{\mathbf{E}}(n) &\triangleq E\{\mathbf{E}(n)\mathbf{E}^H(n)\} = \tilde{\mathbf{H}}\mathbf{R}_s(n)\tilde{\mathbf{H}}^H \\ &= \tilde{\mathbf{H}} \text{diag} \{ |c_1(n)|^2, |c_2(n)|^2, \dots, |c_M(n)|^2 \} \tilde{\mathbf{H}}^H. \end{aligned} \quad (24)$$

Note that $c_j(n) = c_j(n + L_cN)$, $j = 1, 2, \dots, M$, so $\mathbf{R}_{\mathbf{E}}(n)$ is periodic with period L_cN . The Fourier series of $\mathbf{R}_{\mathbf{E}}(n)$ is

$$\begin{aligned} S_E(m) &= \sum_{n=0}^{L_cN-1} \mathbf{R}_{\mathbf{E}}(n) e^{-i(2\pi mn/L_cN)} \\ &= \tilde{\mathbf{H}}\mathbf{C}_s(m)\tilde{\mathbf{H}}^H, \end{aligned} \quad (25)$$

where

$$\begin{aligned} \mathbf{C}_s(m) &\triangleq \text{diag} \left(\sum_{n=0}^{L_cN-1} |c_1(n)|^2 e^{-i(2\pi mn/L_cN)}, \dots, \right. \\ &\quad \left. \sum_{n=0}^{L_cN-1} |c_M(n)|^2 e^{-i(2\pi mn/L_cN)} \right) \\ &= \text{diag} (C_{s_1}(m), \dots, C_{s_M}(m)). \end{aligned} \quad (26)$$

The basic idea of this channel estimation algorithm is to design precoding code sequences $\{c_j(n)\}_{n=0}^{L_cN-1}$ ($j = 1, 2, \dots, M$) such that for a given cycle $m = m_j$, $C_{s_j}(m_j) \neq 0$ and $C_{s_k}(m_j) = 0$ for all $k \neq j$. That is, all but one entries in $\mathbf{C}_s(m)$ are zero. Choosing a different cycle m_j for each user (obviously, we need $L_cN > M$), blind identification of each individual channel can then be achieved through (25).

In fact, if for $m = m_j$, $C_{s_j}(m_j) \neq 0$, but $C_{s_k}(m_j) = 0$, for all $k \neq j$, then

$$S_E(m_j) = \tilde{\mathbf{H}} \text{diag} (0, \dots, 0, C_{s_j}(m_j), 0, \dots, 0) \tilde{\mathbf{H}}^H. \quad (27)$$

It then follows from (8), (23), and (27) that

$$\mathbf{g}_j = [g_j^{(1)}(d), \dots, g_j^{(K)}(d), \dots, g_j^{(1)}(0), \dots, g_j^{(K)}(0)]^T \quad (28)$$

can be determined up to a complex scalar from the $K(d+1) \times K(d+1)$ Hermitian matrix $\mathbf{g}_j\mathbf{g}_j^H$. In other words, the channel responses from user j to each receive antenna $p = 1, 2, \dots, K$ can be identified up to a complex scalar. This ambiguity can be removed either by using one training symbol or using differential encoding.

3.3. Channel estimation using the matrix-pencil approach

Noting that $\mathbf{R}_{\mathbf{E}}(n) = \mathbf{R}_{\mathbf{E}}(n + L_cN)$, we form a matrix pencil $\{S_1, S_2\}$ based on linear combination of $\{\mathbf{R}_{\mathbf{E}}(n)\}_{n=0}^{L_cN-1}$ with random weighting. Let $\alpha_i(n)$ be uniformly distributed in interval $(0,1)$, where $i = 1, 2$. Define

$$\begin{aligned} S_i &= \sum_{n=0}^{L_cN-1} \alpha_i(n) \mathbf{R}_{\mathbf{E}}(n) \\ &= \tilde{\mathbf{H}} \text{diag} \left(\sum_{n=0}^{L_cN-1} \alpha_i(n) |c_1(n)|^2, \dots, \sum_{n=0}^{L_cN-1} \alpha_i(n) |c_M(n)|^2 \right) \tilde{\mathbf{H}}^H \\ &\triangleq \tilde{\mathbf{H}}\Gamma_i\tilde{\mathbf{H}}^H \quad \text{for } i = 1, 2. \end{aligned} \quad (29)$$

According to the definition,

$$\begin{aligned} \Gamma_i &= \text{diag} \left(\sum_{n=0}^{L_cN-1} \alpha_i(n) |c_1(n)|^2, \dots, \right. \\ &\quad \left. \sum_{n=0}^{L_cN-1} \alpha_i(n) |c_M(n)|^2 \right), \quad i = 1, 2, \end{aligned} \quad (30)$$

are two positively-definited matrices.

Consider the generalized eigenvalue problem

$$S_1\mathbf{x} = \lambda S_2\mathbf{x} \iff \tilde{\mathbf{H}}(\Gamma_1 - \lambda\Gamma_2)\tilde{\mathbf{H}}^H\mathbf{x} = 0. \quad (31)$$

If $\tilde{\mathbf{H}}$ is of full column rank (which is ensured by assumption (A4)), then (31) reduces to

$$(\Gamma_1 - \lambda\Gamma_2)\tilde{\mathbf{H}}^H\mathbf{x} = 0. \quad (32)$$

By using random weighting, all the generalized eigenvalues corresponding to (32),

$$\lambda_j = \frac{\sum_{n=0}^{L_cN-1} \alpha_1(n) |c_j(n)|^2}{\sum_{n=0}^{L_cN-1} \alpha_2(n) |c_j(n)|^2}, \quad j = 1, 2, \dots, M, \quad (33)$$

are distinct eigenvalues with probability 1. In this case, since Γ_1 and Γ_2 are both diagonal, the generalized eigenvector \mathbf{x}_j corresponding to λ_j should satisfy

$$\tilde{\mathbf{H}}^H\mathbf{x}_j = \beta_j I_j, \quad (34)$$

where β_j is an unknown scalar, and $I_j = [0, \dots, 1, \dots, 0]^T$ with 1 in the j th entry is the j th column of the $M \times M$ identity matrix \mathbf{I} [29].

It then follows from (31) and (34) that

$$S_1 \mathbf{x}_j = \tilde{\mathbf{H}} \Gamma_1 \tilde{\mathbf{H}}^H \mathbf{x}_j = \beta_j \sum_{n=0}^{L_e N - 1} \alpha_1(n) |c_j(n)|^2 \mathbf{g}_j, \quad (35)$$

where \mathbf{g}_j is as in (28). And β_j can be determined up to a scalar once the generalized eigenvector \mathbf{x}_j is obtained.

Remark 1. It should be noticed that the channel estimation algorithm based on the Fourier analysis requires an additional condition on the coding sequences, which actually implies that for a given cycle, all antennas, except one, are nulled out. More specifically, this constraint on the code sequences implies that for each user, there exists at least one narrow frequency band over which no other user is transmitting. When using the matrix-pencil approach, on the other hand, random weights, hence a random linear transform, is introduced instead of the Fourier transform, resulting in that the condition on the code sequences can be relaxed to any nonconstant modulus sequences which make λ_j 's in (33) be distinct from each other for $j = 1, 2, \dots, M$.

4. CHANNEL EQUALIZATION USING CYCLIC WIENER FILTER

After the channel estimation, in this section, equalization/desired user extraction is carried out using an MMSE cyclic Wiener filter. Without loss of generality, assume user 1 is the desired user. We want to design a chip-level $K \times 1$ MMSE equalizer $\{\mathbf{f}_d(n, i)\}_{i=0}^{L_e-1}$ of length L_e ($L_e \geq L$) which satisfies

$$\mathbf{f}_d(n, i) = \mathbf{f}_d(n + L_e N, i), \quad i = 0, 1, \dots, L_e - 1. \quad (36)$$

The equalizer output can be expressed as

$$\hat{v}_1(n - d) = \sum_{i=0}^{L_e-1} \mathbf{f}_d^H(n, i) \mathbf{y}(n - i). \quad (37)$$

With the above equalizer, the MSE between the input signal and the equalizer output is

$$E\{|e(n)|^2\} = E\left\{\left|\sum_{i=0}^{L_e-1} \mathbf{f}_d^H(n, i) \mathbf{y}(n - i) - v_1(n - d)\right|^2\right\}. \quad (38)$$

Applying the orthogonality principle, we obtain

$$E\left\{\left[\sum_{i=0}^{L_e-1} \mathbf{f}_d^H(n, i) \mathbf{y}(n - i) - v_1(n - d)\right] \mathbf{y}^H(n - k)\right\} = 0 \quad (39)$$

for $k = 0, 1, \dots, L_e - 1$.

Recall that (see (5)) if we define

$$\mathbf{C}(n) \triangleq \text{diag}\{c_1(n), c_2(n), \dots, c_M(n)\}, \quad (40)$$

$$\mathbf{v}(n) \triangleq [v_1(n), v_2(n), \dots, v_M(n)]^T,$$

then

$$\mathbf{s}(n) = [s_1(n), s_2(n), \dots, s_M(n)]^T = \mathbf{C}(n) \mathbf{v}(n). \quad (41)$$

It then follows from (7) that

$$\mathbf{y}(n) = \sum_{l=0}^{L-1} \mathbf{H}(l) \mathbf{C}(n - l) \mathbf{v}(n - l) + \mathbf{w}(n). \quad (42)$$

Stacking L_e successive $\mathbf{y}(n)$ together to form the $KL_e \times 1$ vector

$$Y(n) = \begin{bmatrix} \mathbf{y}(n) \\ \mathbf{y}(n - 1) \\ \vdots \\ \mathbf{y}(n - L_e + 1) \end{bmatrix} \triangleq \mathcal{H}_{C,n} V(n) + W(n), \quad (43)$$

where

$$\mathcal{H}_{C,n} = \begin{bmatrix} \mathbf{H}(0) \mathbf{C}(n) & \cdots & \mathbf{H}(L - 1) \mathbf{C}(n - L + 1) & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{H}(0) \mathbf{C}(n - L_e + 1) & \cdots & \mathbf{H}(L - 1) \mathbf{C}(n - L_e - L + 2) \end{bmatrix} \quad (44)$$

is a $KL_e \times [(L + L_e - 1)M]$ matrix, $V(n) = [\mathbf{v}^T(n), \mathbf{v}^T(n - 1), \dots, \mathbf{v}^T(n - L_e - L + 2)]^T$ and $W(n)$ is defined in the same manner as $Y(n)$. It follows from (A1), (A2), and (A3) that

$$\begin{aligned} \mathbf{R}_Y(n) &\triangleq E\{Y(n)Y^H(n)\} = \mathcal{H}_{C,n} \mathcal{H}_{C,n}^H + \sigma_w^2 \mathbf{I}_{KL_e}, \\ \mathbf{R}_{v_1 Y}(n, d) &\triangleq E\{v_1(n - d)Y^H(n)\} = I_d^H \mathcal{H}_{C,n}^H, \end{aligned} \quad (45)$$

where $I_d = [\mathbf{0}, \dots, \mathbf{0}, \underbrace{\mathbf{1}, \mathbf{0}, \dots, \mathbf{0}}_{(d+1)'s M \times 1 \text{ block}}, \dots, \mathbf{0}]^H$ is the $(Md + 1)$ th column of the $M(L + L_e - 1) \times M(L + L_e - 1)$ identity matrix. Define

$$\tilde{\mathbf{f}}_d(n) \triangleq [\mathbf{f}_d^H(n, 0), \mathbf{f}_d^H(n, 1), \dots, \mathbf{f}_d^H(n, L_e - 1)]^H \quad (46)$$

as the $KL_e \times 1$ equalizer coefficients vector. Then (39) can be rewritten as

$$\mathbf{R}_Y(n) \tilde{\mathbf{f}}_d(n) = \mathcal{H}_{C,n} I_d. \quad (47)$$

It then follows that for $n = 0, \dots, L_c N - 1$,

$$\tilde{\mathbf{f}}_d(n) = \mathbf{R}_Y^\#(n) \mathcal{H}_{C,n} I_d, \quad (48)$$

where # denotes pseudoinverse.

5. EXTENSION TO MULTIRATE CDMA SYSTEMS

To support multimedia services with different quality of services requirements, multirate scheme is implemented in 3G CDMA systems by using *multicode* (MC) or *variable spreading factor* (VSF). In MC systems, the symbols of a high-rate user are subsampled to obtain several symbol streams, and each stream is regarded as the signal from a low-rate virtual user and is spread using a specific signature sequence. In VSF systems, users requiring different rates are assigned signature sequences of different lengths. Thus in the same period, more symbols of high-rate users can be transmitted.

Since chip-level channel modeling and equalization are performed, the proposed approach can readily be extended to multirate case. As an MC system with high-rate users is equivalent to a single-rate system with more users, extension of the proposed approaches to MC multirate CDMA systems is therefore trivial. For VSF systems, let N be the smallest processing gain and let $L_{c,j}N$ denote the length of the j th user's spreading code. Defining

$$L_c = LCM(L_{c,1}, \dots, L_{c,M}) \quad (49)$$

as the least common multiple of $\{L_{c,1}, \dots, L_{c,M}\}$, the generalization of the proposed algorithm to VSF systems is then straightforward.

6. SIMULATION EXAMPLES

We consider the case of two users and four receive antennas. Each user transmits QPSK signals. The spreading gain is chosen to be $N = 8$ or $N = 16$, and three cases are considered. (1) Both users have spreading gain $N = 8$. (2) Both users have spreading gain $N = 16$. (3) Two users have different data rates, the spreading gain for the low-rate user is $N = 16$, and for the high-rate user is $N = 8$.

The nonconstant modulus channelization codes spread over 32 chips (i.e., 2 to 4 symbols depending on the user's spreading gain). Both randomly generated codes which are uniformly distributed within the interval $[0.8, 1.2]$ and codes that satisfy the additional constraint (as described in Section 3.2) are considered. In the simulation, "codes with

constraint" are chosen to be

$$\begin{aligned} \mathbf{c}_1 = & [0.6857, 0.7145, 0.6356, 0.6849, 0.8433, 0.8036, 0.7597, \\ & 0.5856, 0.7488, 0.5641, 0.7300, 0.7542, 0.7482, 0.5870, \\ & 0.7902, 0.6172, 0.5409, 0.5474, 0.6425, 0.7834, 0.7520, \\ & 0.6743, 0.6904, 0.8114, 0.5829, 0.6913, 0.5939, 0.7339, \\ & 0.8608, 0.6380, 0.8207, 0.8808], \\ \mathbf{c}_2 = & [0.6670, 0.7275, 0.8540, 0.6100, 0.7518, 0.6363, 0.5545, \\ & 0.6887, 0.7092, 0.6143, 0.6313, 0.7625, 0.5210, 0.8036, \\ & 0.7582, 0.6979, 0.8136, 0.6944, 0.6902, 0.6660, 0.6536, \\ & 0.6908, 0.6010, 0.8078, 0.7622, 0.5486, 0.6005, 0.6395, \\ & 0.6176, 0.8070, 0.6382, 0.8265]. \end{aligned} \quad (50)$$

The multipath channels have three rays and the multipath amplitudes are Gaussian with zero mean and identical variance. The transmission delays are uniformly spread over 6 chip intervals. Complex zero mean white Gaussian noise was added to the received signals. The normalized mean-square-error of channel estimation (CHMSE) for the desired user is defined as

$$\text{CHMSE} = \frac{1}{KIL} \sum_{i=1}^I \sum_{p=1}^K \frac{\|\hat{\mathbf{g}}_1^{(p)} - \mathbf{g}_1^{(p)}\|^2}{\|\mathbf{g}_1^{(p)}\|^2}, \quad (51)$$

where I stands for the number of Monte-Carlo runs, and K is the number of receive antennas. And SNR refers to the signal-to-noise ratio with respect to the desired user and is chosen to be the same at each receiver. The result is averaged over $I = 100$ Monte-Carlo runs. The channel is generated randomly in each run, and is estimated based on a record of 256 symbols. In the case of multirate, we mean 256 lower-rate symbols. The equalizer with length $L_e = 6$ is constructed according to the estimated channel, and is applied to a set of 1024 independent symbols in order to calculate the symbol MSE and BER for each Monte-Carlo run. Blind channel estimation based on nonconstant modulus precoding is carried out both with and without the matrix-pencil approach. Without the matrix-pencil approach, channel estimation is obtained directly through the second-order statistics of $\mathbf{E}(n)$ (see (22)) based on the nonconstant precoding technique and the Fourier transform, as presented in Section 3.2. Simulation results show that by introducing a random linear transform, the matrix-pencil approach delivers significantly better results for both single-rate and multirate systems. Figures 3 and 4 correspond to the single-rate cases, where both users have spreading gain $N = 8$ or $N = 16$, and the codes in (50) are used. In the figures, "MP" stands for "matrix pencil". Figures 5 and 6 compare the performances of the matrix-pencil-based approach when different codes are used. In the figures, "codes with constraint" denote the codes in (50), and we choose $N = 8$ for the high-rate user and $N = 16$ for the low rate user. Optimal spreading code design and random linear transform design will be investigated in future work.

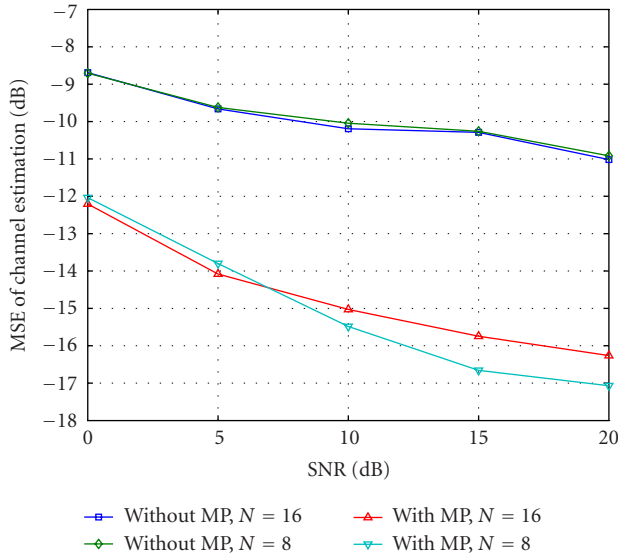


FIGURE 3: Normalized MSE of channel estimation versus SNR, single-rate cases with $N = 8$ and $N = 16$, respectively.

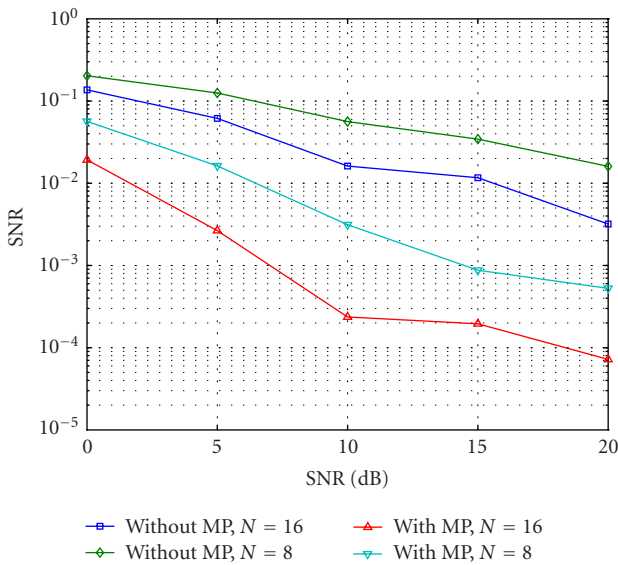


FIGURE 4: Comparison of BER versus SNR, single-rate cases with $N = 8$ and $N = 16$, respectively.

7. CONCLUSIONS

In this paper, blind channel identification and signal separation for long-code CDMA systems are revisited. Long-code CDMA system is characterized using a time-invariant system model by modeling the received signals and MUIs as cyclostationary processes with modulation-induced cyclostationarity. Then, multistep linear prediction method is used to reduce the intersymbol interference introduced by multipath propagation, and channel estimation is performed by exploiting the nonconstant modulus precoding technique with

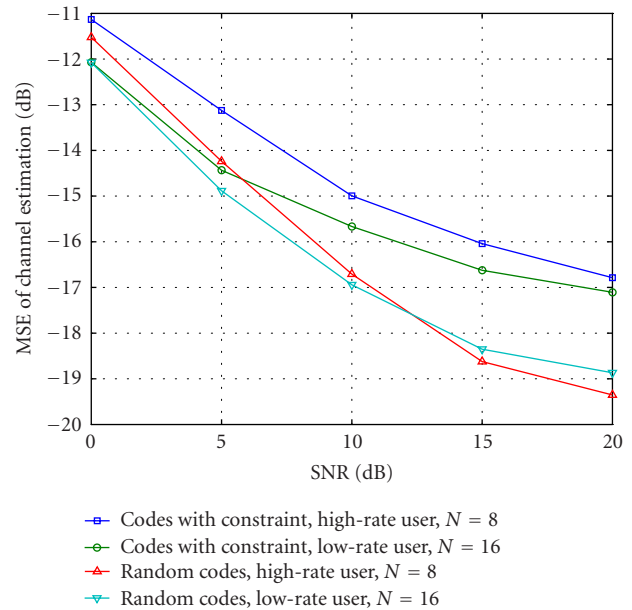


FIGURE 5: Normalized MSE of channel estimation versus SNR for matrix-pencil-based approach with different codes, multirate configuration with $N = 8$ for the high-rate user and $N = 16$ for the low-rate user, respectively.

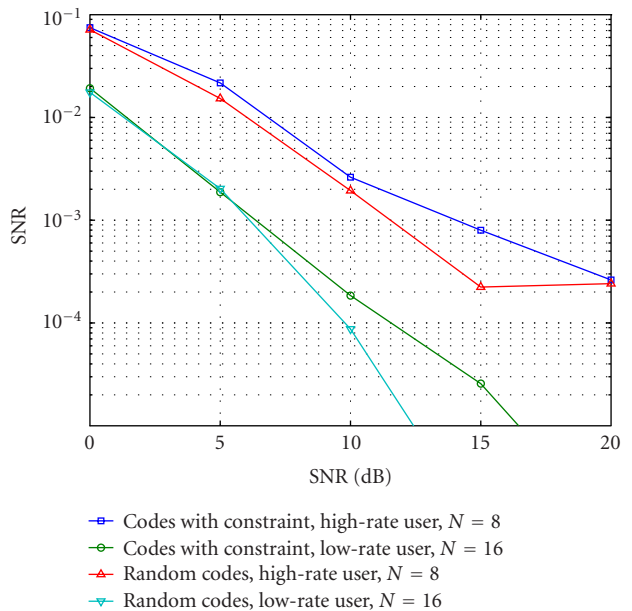


FIGURE 6: Comparison of BER versus SNR for matrix-pencil-based approach with different codes, multirate configuration with $N = 8$ for the high-rate user and $N = 16$ for the low-rate user, respectively.

and without the matrix-pencil approach. It is found that by introducing a random linear transform, the matrix-pencil-based approach delivers a much better result than the one relying on the Fourier transform. As chip-level channel modeling and equalization are performed, the proposed approach can be extended to multirate CDMA systems in a straight forward manner.

ACKNOWLEDGMENT

This paper is supported in part by MSU IRGP 91-4005 and NSF Grants CCR-0196364 and ECS-0121469.

REFERENCES

- [1] S. Verdú, *Multuser Detection*, Cambridge University Press, Cambridge, UK, 1998.
- [2] S. Parkvall, "Variability of user performance in cellular DS-CDMA-long versus short spreading sequences," *IEEE Trans. Commun.*, vol. 48, no. 7, pp. 1178–1187, 2000.
- [3] S. E. Bensley and B. Aazhang, "Subspace-based channel estimation for code division multiple access communication systems," *IEEE Trans. Commun.*, vol. 44, no. 8, pp. 1009–1020, 1996.
- [4] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, no. 4, pp. 944–960, 1995.
- [5] M. Torlak and G. Xu, "Blind multiuser channel estimation in asynchronous CDMA systems," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 137–147, 1997.
- [6] J. K. Tugnait and T. Li, "A multistep linear prediction approach to blind asynchronous CDMA channel estimation and equalization," *IEEE J. Select. Areas Commun.*, vol. 19, no. 6, pp. 1090–1102, 2001.
- [7] X. Wang and H. V. Poor, "Blind adaptive multiuser detection in multipath CDMA channels based on subspace tracking," *IEEE Trans. Signal Processing*, vol. 46, no. 11, pp. 3030–3044, 1998.
- [8] A. J. Weiss and B. Friedlander, "Channel estimation for DS-CDMA downlink with aperiodic spreading codes," *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1561–1569, 1999.
- [9] S. Bhashyam and B. Aazhang, "Multiuser channel estimation and tracking for long-code CDMA systems," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1081–1090, 2002.
- [10] S. Buzzi and H. V. Poor, "Channel estimation and multiuser detection in long-code DS/CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 19, no. 8, pp. 1476–1487, 2001.
- [11] S. Buzzi and H. V. Poor, "A multipass approach to joint data and channel estimation in long-code CDMA systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 2, pp. 612–626, 2004.
- [12] Y.-F. Chen, M. D. Zoltowski, J. Ramos, C. Chatterjee, and V. P. Roychowdhury, "Reduced-dimension blind space-time 2-D RAKE receivers for DS-CDMA communication systems," *IEEE Trans. Signal Processing*, vol. 48, no. 6, pp. 1521–1536, 2000.
- [13] C. J. Escudero, U. Mitra, and D. T. M. Slock, "A toeplitz displacement method for blind multipath estimation for long code DS/CDMA signals," *IEEE Trans. Signal Processing*, vol. 49, no. 3, pp. 654–665, 2001.
- [14] H. Liu and M. D. Zoltowski, "Blind equalization in antenna array CDMA systems," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 161–172, 1997.
- [15] L. Tong, A.-J. van der Veen, P. Dewilde, and Y. Sung, "Blind decorrelating RAKE receivers for long-code WCDMA," *IEEE Trans. Signal Processing*, vol. 51, no. 6, pp. 1642–1655, 2003.
- [16] M. Torlak, B. L. Evans, and G. Xu, "Blind estimation of FIR channels in CDMA systems with aperiodic spreading sequences," in *Proc. the 31st Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 495–499, Pacific Grove, Calif, USA, 1997.
- [17] Z. Xu, "Low-complexity multiuser channel estimation with aperiodic spreading codes," *IEEE Trans. Signal Processing*, vol. 49, no. 11, pp. 2813–2822, 2001.
- [18] Z. Xu and M. K. Tsatsanis, "Blind channel estimation for long code multiuser CDMA systems," *IEEE Trans. Signal Processing*, vol. 48, no. 4, pp. 988–1001, 2000.
- [19] Z. Yang and X. Wang, "Blind turbo multiuser detection for long-code multipath CDMA," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 112–125, 2002.
- [20] Y. Sung and L. Tong, "Tracking of fast-fading channels in long-code CDMA," *IEEE Trans. Signal Processing*, vol. 52, no. 3, pp. 786–795, 2004.
- [21] C. D. Frank, E. Visotsky, and U. Madhow, "Adaptive interference suppression for the downlink of a direct sequence cdma system with long spreading sequences," *The Journal of VLSI Signal Processing*, vol. 30, pp. 273–291, March 2002.
- [22] T. P. Krauss, W. J. Hillery, and M. D. Zoltowski, "Downlink specific linear equalization for frequency selective CDMA cellular systems," *The Journal of VLSI Signal Processing*, vol. 30, pp. 143–161, January 2002.
- [23] T. Li, J. K. Tugnait, and Z. Ding, "Channel estimation of long-code CDMA systems utilizing transmission induced cyclostationarity," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, vol. 4, pp. 105–108, 2003.
- [24] D. T. M. Slock, "Blind joint equalization of multiple synchronous mobile users using oversampling and/or multiple antennas," in *Proc. the 28th Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1154–1158, Pacific Grove, Calif, USA, 1994.
- [25] N. Delfosse and P. Loubaton, "Adaptive blind separation of convolutive mixtures," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '96)*, vol. 5, pp. 2940–2943, Atlanta, Ga, USA, 1996.
- [26] Z. Ding, "Matrix outer-product decomposition method for blind multiple channel identification," *IEEE Trans. Signal Processing*, vol. 45, no. 12, pp. 3053–3061, 1997.
- [27] A. Gorokhov, P. Loubaton, and E. Moulines, "Second order blind equalization in multiple input multiple output FIR systems: a weighted least squares approach," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '96)*, vol. 5, pp. 2415–2418, Atlanta, Ga, USA, 1996.
- [28] J. K. Tugnait and B. Huang, "Multistep linear predictors-based blind identification and equalization of multiple-input multiple-output channels," *IEEE Trans. Signal Processing*, vol. 48, no. 1, pp. 26–38, 2000.
- [29] C. Chang, Z. Ding, S. F. Yau, and F. H. Y. Chan, "A matrix-pencil approach to blind separation of colored nonstationary signals," *IEEE Trans. Signal Processing*, vol. 48, no. 3, pp. 900–907, 2000.
- [30] J. Liang and Z. Ding, "Nonminimum-phase FIR channel estimation using cumulant matrix pencils," *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2310–2320, 2003.
- [31] E. Serpedin and G. B. Giannakis, "Blind channel identification and equalization with modulation-induced cyclostationarity," *IEEE Trans. Signal Processing*, vol. 46, no. 7, pp. 1930–1944, 1998.
- [32] G. B. Giannakis, "Filterbanks for blind channel identification and equalization," *IEEE Signal Processing Lett.*, vol. 4, no. 6, pp. 184–187, 1997.
- [33] M. K. Tsatsanis and G. B. Giannakis, "Transmitter induced cyclostationarity for blind channel equalization," *IEEE Trans. Signal Processing*, vol. 45, no. 7, pp. 1785–1794, 1997.

- [34] H. Bolcskei, R. W. Heath Jr., and A. J. Paulraj, "Blind channel identification and equalization in OFDM-based multiantenna systems," *IEEE Trans. Signal Processing*, vol. 50, no. 1, pp. 96–109, 2002.
- [35] J. K. Tugnait and W. Luo, "Linear prediction error method for blind identification of periodically time-varying channels," *IEEE Trans. Signal Processing*, vol. 50, no. 12, pp. 3070–3082, 2002.

Tongtong Li received her Ph.D. degree in electrical engineering in 2000 from Auburn University. From 2000 to 2002, she was with Bell Labs, and has been working on the design and implementation of wireless communication systems, including 3GPP UMTS and IEEE 802.11a. She joined the faculty of Michigan State University in 2002, and currently is an Assistant Professor at the Department of ECE. Her research interests fall into the areas of wireless and wirelined communication systems, multiuser detection and separation over time-varying wireless channels, wireless networking and network security, and digital signal processing with applications in wireless communications. She is serving as an Editorial Board Member for *EURASIP Journal on Wireless Communications and Networking*.



Weiguo Liang was born in Hebei province, China, January 1975. He received the B.E. degree in biomedical engineering from Tsinghua University, Beijing, China, and the M.S. degree in electrical engineering from the Chinese Academy of Sciences, Beijing, China, in 1998 and 2001, respectively. He is currently pursuing the Ph.D. degree at the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, Mich. Since 2001, he has been a Research Assistant at this department. His research interests include blind equalization, multiuser detection, space-time coding, and wireless sensor network.



Zhi Ding is Professor at the University of California, Davis. He received his Ph.D. degree in electrical engineering from Cornell University in 1990. From 1990 to 2000, he was a faculty member of Auburn University and later, University of Iowa. He has held visiting positions in the Australian National University, Hong Kong University of Science and Technology, NASA Lewis Research Center, and USAF Wright Laboratory. He has active collaboration with researchers from several countries including Australia, China, Japan, Canada, Taiwan, Korea, Singapore, and Hong Kong. He is also a Visiting Professor at the Southeast University, Nanjing, China. He is a Fellow of IEEE and has been an active Member of IEEE, serving on technical programs of several workshops and conferences. He was an Associate Editor for *IEEE Transactions on Signal Processing* from 1994–1997, 2001–2004. He is currently an Associate Editor of the *IEEE Signal Processing Letters*. He was a member of technical committee on statistical signal and array processing and member of technical committee on signal processing for communications. Currently, he is a member of the CAS technical committee on blind signal processing.



Jitendra K. Tugnait received the B.S. (honors) degree in electronics and electrical communication engineering from the Punjab Engineering College, Chandigarh, India, in 1971, the M.S. and E.E. degrees from Syracuse University, Syracuse, NY, and the Ph.D. degree from the University of Illinois at Urbana-Champaign, in 1973, 1974, and 1978, respectively, all in electrical engineering. From 1978 to 1982 he was an Assistant Professor of electrical and computer engineering at the University of Iowa, Iowa City, Iowa. He was with the Long Range Research Division of the Exxon Production Research Company, Houston, Tex, from June 1982 to September 1989. He joined the Department of Electrical and Computer Engineering, Auburn University, Auburn, Ala, in September 1989 as a Professor. He currently holds the title of James B. Davis and Alumni Professor. His current research interests are in statistical signal processing, wireless and wireline digital communications, and stochastic systems analysis. He is a past Associate Editor of the *IEEE Transactions on Automatic Control* and of the *IEEE Transactions on Signal Processing*. He is currently an Editor of the *IEEE Transactions on Wireless Communications*. He was on elected Fellow of the IEEE in 1994.



Adaptive Space-Time-Spreading-Assisted Wideband CDMA Systems Communicating over Dispersive Nakagami- m Fading Channels

Lie-Liang Yang

*School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK
Email: lly@ecs.soton.ac.uk*

Lajos Hanzo

*School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK
Email: lh@ecs.soton.ac.uk*

Received 23 May 2004; Revised 18 December 2004

In this contribution, the performance of wideband code-division multiple-access (W-CDMA) systems using space-time-spreading- (STS-) based transmit diversity is investigated, when frequency-selective Nakagami- m fading channels, multiuser interference, and background noise are considered. The analysis and numerical results suggest that the achievable diversity order is the product of the frequency-selective diversity order and the transmit diversity order. Furthermore, both the transmit diversity and the frequency-selective diversity have the same order of importance. Since W-CDMA signals are subjected to frequency-selective fading, the number of resolvable paths at the receiver may vary over a wide range depending on the transmission environment encountered. It can be shown that, for wireless channels where the frequency selectivity is sufficiently high, transmit diversity may be not necessitated. Under this case, multiple transmission antennas can be leveraged into an increased bitrate. Therefore, an adaptive STS-based transmission scheme is then proposed for improving the throughput of W-CDMA systems. Our numerical results demonstrate that this adaptive STS-based transmission scheme is capable of significantly improving the effective throughput of W-CDMA systems. Specifically, the studied W-CDMA system's bitrate can be increased by a factor of three at the modest cost of requiring an extra 0.4 dB or 1.2 dB transmitted power in the context of the investigated urban or suburban areas, respectively.

Keywords and phrases: CDMA, space-time spreading, Nakagami- m fading, transmit diversity.

1. BACKGROUND ON LINK ADAPTATION

It is widely recognised that the channel quality of wireless systems fluctuates over a wide range and hence it is unrealistic to expect that conventional nonadaptive systems might be able to provide a time-invariant grade of service. Hence in recent years various near-instantaneously adaptive-coding-and-modulation- (ACM-) assisted arrangements have been proposed [1, 2], which have found their way also into the high-speed downlink packet access (HS-DPA) mode of the third-generation wireless systems [3] and in other adaptively reconfigurable multicarrier orthogonal

frequency division multiplex (OFDM) systems [4] as well as into single-carrier and multi-carrier DS-CDMA schemes [5]. The family of multi-carrier systems is now widely considered to be the most potent candidate for the next-generation systems of wireless communications. The taxonomy of ACM schemes and a plethora of open research problems was detailed in [5, Chapter 1], hence here we refrain from detailing these issues. The philosophy of these ACM schemes is that instead of dropping a wireless call, they temporarily drop their throughput [3], when the instantaneous channel quality quantified in terms of the signal to interference-plus-noise ratio (SINR) [5] is too low and hence the resultant bit error ratio (BER) happens to be excessive. In this contribution, we will focus our attention on a less well-documented area of link adaptivity, namely, on the effects on multipath-induced dispersion-controlled adaptivity [5]. Achieving these ambitious objectives requires efficient

This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

cross-layer design,¹ which supports the agile and prompt liaison of the OSI layers concerned, potentially requiring an interaction between the physical, network, and service layers, as it was exemplified in [3, 5]. More explicitly, in order to be able to pass on the benefits of the increased system throughput of these cross-layer optimised ACM-aided transceivers to the service layer in terms of improved video or speech quality, near-instantaneously adaptive speech codecs [6] and video codecs [7] are required. These speech and video codecs must have the ability to reconfigure themselves under the control of the near-instantaneous channel quality, such as the advanced multirate (AMR) speech codec or the H.26L multimedia source codec [8]. The interactions and performance benefits of cross-layer-optimised third-generation wireless systems employing adaptive beamforming were quantified in [3], while a host of further cross-layer optimisation issues were treated in [9, 10, 11, 12, 13, 14].

Against this background, in this contribution we focus our attention on a specific channel-quality controlled link adaptation algorithm, which allows the system to increase its effective throughput, as a function of the instantaneous channel quality with the aid of a novel combination of multiple-antenna-assisted transmitter and receiver diversity schemes. The capacity and the achievable data rate of wireless communication systems is limited by the time-varying characteristics of the channels. An efficient technique of combating the time-varying effects of wireless channels is employing diversity. In recent years, space-time coding has received much attention as an effective transmit diversity technique used for combating fading in wireless communications [15, 16, 17, 18]. Space-time-block-coding-assisted [16] transmit diversity has now been adapted as an optional diversity mode in the third-generation (3G) wireless systems known as IMT2000 using wideband code-division multiple-access (W-CDMA) [19, 20]. Inspired by space-time codes, in [21], an attractive transmit diversity scheme based on space-time spreading (STS) has been proposed by Hochwald et al. for employment in CDMA systems. The simple spreading philosophy of this scheme is portrayed in the schematic of Figure 1 and exemplified with the aid of the signal waveforms seen in Figure 2, both of which will be discussed in detail during our further discourse. An STS scheme designed for supporting two transmission antennas and one receiver antenna has also been included in the cdma2000 W-CDMA standard [20]. In [21], the performance of CDMA systems using STS has been investigated by Hochwald et al., when the channel is modelled either as a flat or as a frequency-selective Rayleigh fading channel in the absence of multiuser

interference. It was argued that the proposed STS scheme is capable of attaining the maximal achievable transmit diversity gain without using extra spreading codes and without an increased transmit power. Furthermore, the results recorded for transmission over frequency-selective Rayleigh fading channels by Hochwald et al. [21, Figure 4] show that when there is a sufficiently high number of resolvable paths, a CDMA system using a single transmit antenna and a conventional RAKE receiver is capable of achieving an adequate diversity gain.

Wideband CDMA channels are typically frequency-selective fading channels, having a number of resolvable paths. Therefore, in this contribution, first we investigate the performance of W-CDMA systems using STS-based transmit diversity, when encountering multipath Nakagami- m fading channels, multiuser interference, and background noise. A BER expression is derived, when Gaussian approximation [22, 23] of the multiuser interference and that of the multipath interference is invoked. This BER expression implies that the diversity order achieved is the product of the transmit diversity order and the frequency selective diversity order. Furthermore, the analysis and the numerical results show that both the STS and the frequency selectivity of the channel appear to have the same order of importance, especially when the power decay factor of the multipath intensity profile (MIP) [24] is low.

The frequency-selective frequency-domain transfer function of W-CDMA wireless channels may vary slowly, but often over a wide dynamic range when roaming in urban and suburban areas [25]. Therefore, the number of resolvable paths at the receiver can be modelled as a random variable distributed over a certain range, depending on the location of the receiver, where the number of resolvable paths varies slowly, as the receiver moves. Consequently, STS schemes designed on the basis of a low number of resolvable paths or based on the premise of encountering a constant number of resolvable paths may not achieve the maximum communication efficiency in terms of the effective throughput.

Motivated by the above arguments, in the second part of this contribution an adaptive STS-based transmission scheme is proposed and investigated, which adapts the mode of operation of its STS scheme and its corresponding data rate according to the near-instantaneous frequency selectivity information fed back from the receiver to the transmitter. Our numerical results show that this adaptive STS scheme is capable of efficiently exploiting the diversity potential provided by the channel's frequency selectivity, hence significantly improving the effective throughput of W-CDMA systems.

The remainder of this paper is organized as follows. In the next section, the W-CDMA system's model using STS and the channel model are described. Section 3 considers the detection of STS-based W-CDMA signals. In Section 4, we derive the corresponding BER expression and summarize our numerical results, while in Section 5 we describe the proposed adaptive STS scheme and investigate its BER performance. Finally, our conclusions are offered in Section 6.

¹Cross-layer design constitutes a novel area of wireless system research, which is motivated by the fact that some elements of wireless systems, such as handovers and power control, do not fit into the classic seven-layer open system interconnection (OSI) architecture and hence an improved system performance may be achieved by jointly optimising several layers. In this contribution, the service layer, namely, the achievable data rate or video quality and voice quality, would be improved by the increased bitrate attained by the proposed system.

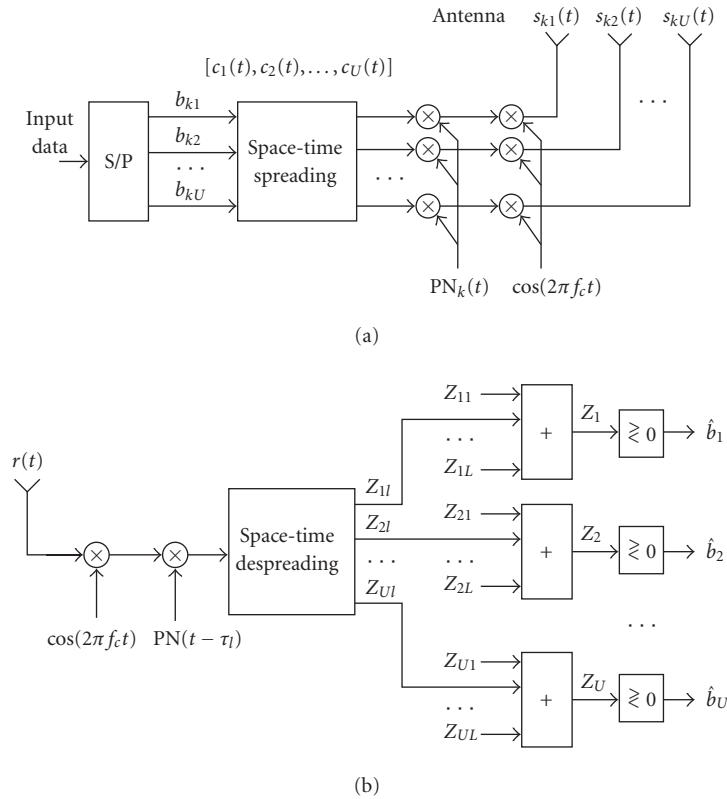


FIGURE 1: (a) Transmitter and (b) receiver block diagram of the W-CDMA system using space-time spreading.

2. SYSTEM MODEL

2.1. Transmitted signal

The W-CDMA system considered in this paper consists of U transmitter antennas and one receiver antenna. The transmitter schematic of the k th user and the receiver schematic of the reference user are shown in Figure 1, where real-valued data symbols using BPSK modulation and real-valued spreading [21] were assumed. Note that the analysis in this contribution can be extended to W-CDMA systems using U transmitter antennas and more than one receiver antenna, or to W-CDMA systems using complex-valued data symbols as well as complex-valued spreading. As shown in Figure 1a, at the transmitter side the binary input data stream having a bit duration of T_b is serial-to-parallel (S/P) converted to U parallel substreams. The new bit duration of each parallel substream, in other words the symbol duration, becomes $T_s = UT_b$. After S/P conversion, the U number of parallel bits are direct-sequence spread using the STS schemes proposed by Hochwald et al. [21] with the aid of U number of orthogonal spreading sequences—for example, Walsh codes—having a period of UG , where $G = T_b/T_c$ represents the number of chips per bit and T_c is the chip duration of the orthogonal spreading sequences. The STS scheme will be further discussed in detail during our forthcoming discourse in this section. As seen in Figure 1a, following STS, the U parallel signals to be mapped to the U transmission antennas are scrambled using the k th user's pseudonoise (PN)

sequence $PN_k(t)$, in order that the transmitted signals become randomised, and to ensure that the orthogonal spreading sequences employed within the STS block of Figure 1 can be reused by the other users. Finally, after the PN-sequence-based scrambling, the U number of parallel signals are carrier modulated and transmitted by the corresponding U number of antennas.

As described above, we have assumed that the number of parallel data substreams, the number of orthogonal spreading sequences used by the STS block of Figure 1, and the number of transmission antennas is the same, namely U . This specific STS scheme constitutes a specific subclass of the generic family of STS schemes, where the number of parallel data substreams, the number of orthogonal spreading sequences required by STS block, and the number of transmission antennas may take different values. The impressive study conducted by Hochwald et al. [21] has shown that the number of orthogonal spreading sequences required by STS is usually higher than the number of parallel substreams. The STS scheme having an equal number of parallel substreams, orthogonal STS-related spreading sequences, as well as transmission antennas constitutes an attractive scheme, since this STS scheme is capable of providing maximal transmit diversity without requiring extra STS spreading codes. Note that for the specific values of $U = 2, 4$ the above-mentioned attractive STS schemes have been specified by Hochwald et al. [21]. In this contribution, we only investigate these attractive STS schemes.

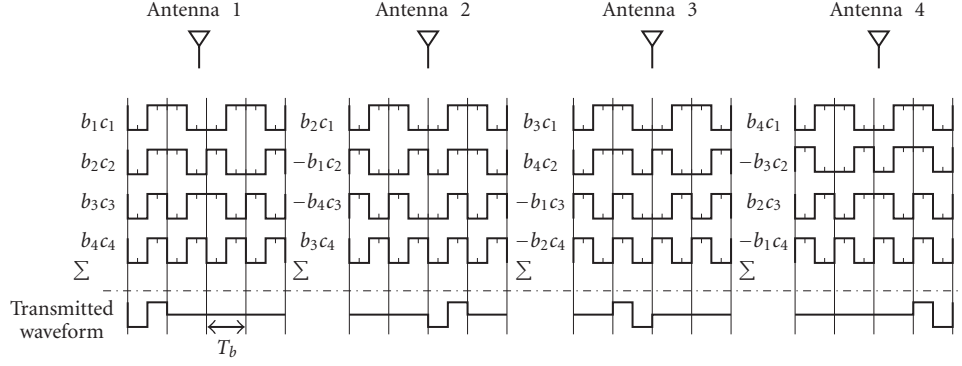


FIGURE 2: Illustration of STS using four transmission antennas transmitting 4 bits within $4T_b$ duration, where $b_1 = b_2 = b_3 = b_4 = +1$ are assumed. Furthermore, c_1, c_2, c_3, c_4 are four STS-related orthogonal codes having a period of $4T_b$. In this example, the STS-codes were chosen as follows: $c_1 = -1 -1 +1 +1 +1 +1 -1 -1 -1 -1 +1 +1 +1 +1 -1 -1$, $c_2 = -1 -1 +1 +1 +1 +1 -1 -1 +1 +1 -1 -1 -1 -1 +1 +1$, $c_3 = -1 -1 +1 +1 -1 -1 +1 +1 +1 +1 -1 -1 +1 +1 -1 -1$, $c_4 = -1 -1 +1 +1 -1 -1 +1 +1 -1 -1 +1 +1 -1 -1 +1 +1$. We note however that the codes used in Figure 3 could be also employed after repeating them four times without the loss of orthogonality.

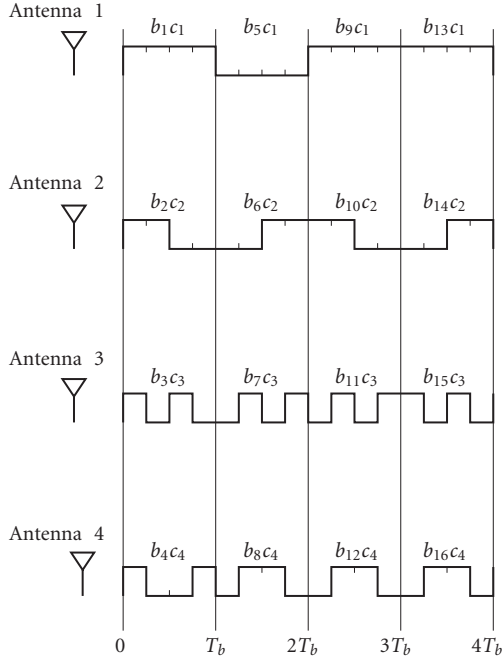


FIGURE 3: Illustration of the transmitted waveforms of the transmission scheme without using STS, that is, the four transmission antennas transmit their data independently. In this figure, we assumed that $b_1 = b_2 = b_3 = b_4 = +1$, $b_5 = b_6 = b_7 = b_8 = -1$, $b_9 = b_{10} = +1$, $b_{11} = b_{12} = -1$, $b_{13} = +1$, $b_{14} = +1$, $b_{15} = +1$, $b_{16} = -1$. Furthermore, c_1, c_2, c_3, c_4 are four STS-related orthogonal codes that have a reduced period of T_b , rather than $4T_b$ as it was in Figure 2 or $2T_b$ as in Figure 4. In this example, the STS-codes were chosen as follows: $c_1 = +1 +1 +1 +1$, $c_2 = +1 +1 -1 -1$, $c_3 = +1 -1 +1 -1$, $c_4 = +1 -1 -1 +1$.

Based on the philosophy of STS as discussed in [21] and referring to Figure 1a, the transmitted signal of the k th user can be expressed as

$$\mathbf{s}_k(t) = \sqrt{\frac{2P}{U^2}} \mathbf{c}(t) \mathbf{B}_U(t) \times \text{PN}_k(t) \cos(2\pi f_c t), \quad (1)$$

where P represents each user's transmitted power, which is constant for all users, $\mathbf{s}_k(t) = [s_{k1}(t) \ s_{k2}(t) \ \cdots \ s_{kU}(t)]$ represents the transmitted signal vector of the U transmission antennas, while $\text{PN}_k(t)$ and f_c represent the DS-scrambling-based spreading waveform and the carrier frequency, respectively. The scrambling sequence waveform is given by $\text{PN}_k(t) = \sum_{j=-\infty}^{\infty} p_{kj} P_{T_c}(t - jT_c)$, where p_{kj} assumes values of $+1$ or -1 with equal probability, while $P_{T_c}(t)$ is the rectangular chip waveform, which is defined over the interval $[0, T_c)$. In (1), the vector $\mathbf{c}(t) = [c_1(t) \ c_2(t) \ \cdots \ c_U(t)]$ is constituted by the U number of orthogonal signals assigned for the STS, $c_i(t) = \sum_{j=-\infty}^{\infty} c_{ij} P_{T_c}(t - jT_c)$, $i = 1, 2, \dots, U$, denotes the individual components of the STS-based orthogonal spread signals, where $\{c_{ij}\}$ is an orthogonal sequence of period UG for each index i ; $\mathbf{B}_U(t)$ represents the $U \times U$ -dimensional transmitted data matrix created by mapping U input data bits to the U parallel substreams according to the specific design rules outlined by Hochwald et al. [21], so that the maximum possible transmit diversity is achieved, while using relatively low-complexity signal detection algorithms. Specifically, $\mathbf{B}_U(t)$ can be expressed as

$$\mathbf{B}_U(t) = \begin{pmatrix} a_{11}b_{k,11} & a_{12}b_{k,12} & \cdots & a_{1U}b_{k,1U} \\ a_{21}b_{k,21} & a_{22}b_{k,22} & \cdots & a_{2U}b_{k,2U} \\ \vdots & \vdots & \ddots & \vdots \\ a_{U1}b_{k,U1} & a_{U2}b_{k,U2} & \cdots & a_{UU}b_{k,UU} \end{pmatrix} (t), \quad (2)$$

where the time dependence of the (i, j) th element is indicated at the right-hand side of the matrix for simplicity. In (2), a_{ij} represents the sign of the element at the i th row and the j th column, which is determined by the STS design rule, while $b_{k,i,j}$ is the data bit assigned to the (i, j) th element, which is one of the U input data bits $\{b_{k1}, b_{k2}, \dots, b_{kU}\}$ of user k . Each input data bit of $\{b_{k1}, b_{k2}, \dots, b_{kU}\}$ appears only once in any given row and in any given column. For $U = 2, 4$, $\mathbf{B}_2(t)$, and

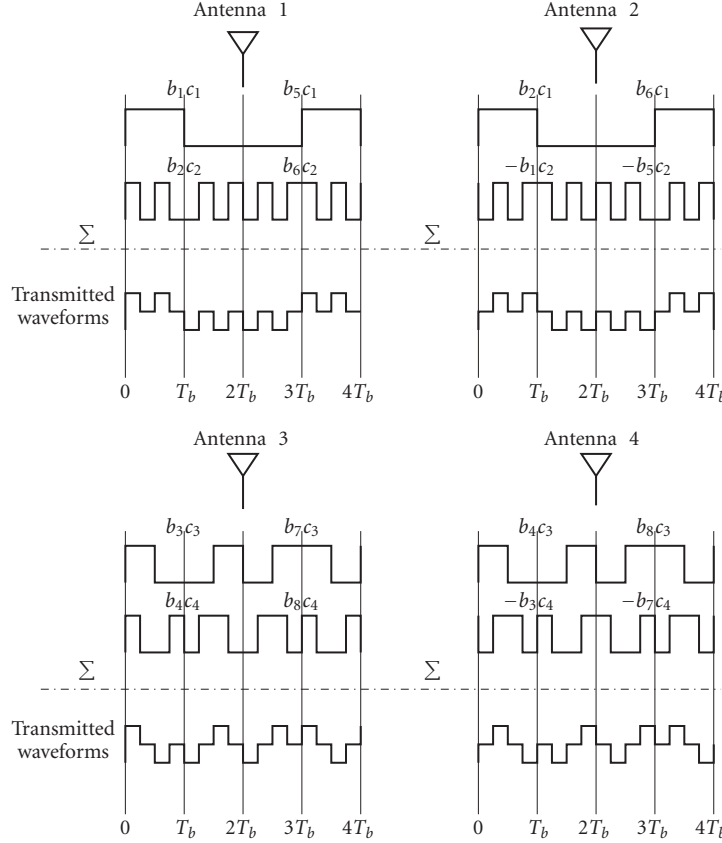


FIGURE 4: Illustration of STS using two transmission antennas transmitting 2 bits within $2T_b$ duration. Hence, four transmission antennas transmit 8 bits within $4T_b$ duration, where $b_1 = b_2 = b_3 = b_4 = +1$ and $b_5 = b_6 = b_7 = b_8 = -1$ were assumed. Furthermore, c_1, c_2, c_3, c_4 are four STS-related orthogonal codes that have a reduced period of $2T_b$, rather than $4T_b$ as it was in Figure 2. In this example, the STS codes were chosen as follows: $c_1 = +1 + 1 + 1 + 1 - 1 - 1 - 1 - 1$, $c_2 = +1 - 1 + 1 - 1 - 1 + 1 - 1 + 1$, $c_3 = +1 + 1 - 1 - 1 - 1 - 1 + 1 + 1$, $c_4 = +1 - 1 - 1 + 1 - 1 + 1 + 1 - 1$. We note however that the codes used in Figure 3 could be also employed after repeating them twice without the loss of orthogonality.

$\mathbf{B}_4(t)$ are given by [21]

$$\mathbf{B}_2(t) = \begin{pmatrix} b_{k1} & b_{k2} \\ b_{k2} & -b_{k1} \end{pmatrix} (t),$$

$$\mathbf{B}_4(t) = \begin{pmatrix} b_{k1} & b_{k2} & b_{k3} & b_{k4} \\ b_{k2} & -b_{k1} & b_{k4} & -b_{k3} \\ b_{k3} & -b_{k4} & -b_{k1} & b_{k2} \\ b_{k4} & b_{k3} & -b_{k2} & -b_{k1} \end{pmatrix} (t). \quad (3)$$

Based on (1) and (2) the signal transmitted by the u th antenna to the k th user can be explicitly expressed as

$$s_{ku}(t) = \sqrt{\frac{2P}{U^2}} [c_1(t)a_{1u}b_{k,1u}(t) + c_2(t)a_{2u}b_{k,2u}(t) + \dots + c_U(t)a_{Uu}b_{k,Uu}(t)] \times \text{PN}_k(t) \cos(2\pi f_c t), \quad u = 1, 2, \dots, U. \quad (4)$$

2.2. Channel model

The U number of parallel subsignals

$$\mathbf{s}_k(t) = [s_{k1}(t) \ s_{k2}(t) \ \dots \ s_{kU}(t)] \quad (5)$$

is transmitted by the U number of antennas over frequency-selective fading channels, where each parallel subsignal experiences independent frequency-selective Nakagami- m fading. The complex lowpass equivalent representation of the impulse response experienced by the u th parallel subsignal of user k is given by [24]

$$h_k^u(t) = \sum_{l=1}^L h_{kl}^u \delta(t - \tau_{kl}) \exp(j\psi_{kl}^u), \quad (6)$$

where h_{kl}^u , τ_{kl} , and ψ_{kl}^u represent the attenuation factor, delay and phase shift of the l th multipath component of the channel, respectively, while L is the total number of resolvable multipath components and $\delta(t)$ is the Kronecker delta function. We assume that the phases $\{\psi_{kl}^u\}$ in (6) are independent identically distributed (i.i.d.) random variables uniformly distributed in the interval $[0, 2\pi)$, while the L

multipath attenuations $\{h_{kl}^u\}$ in (6) are independent Nakagami random variables with a probability density function (PDF) of [22, 23, 24, 25, 26, 27]

$$p(h_{kl}^u) = M(h_{kl}^u, m_{kl}^{(u)}, \Omega_{kl}^u),$$

$$M(R, m, \Omega) = \frac{2m^m R^{2m-1}}{\Gamma(m)\Omega^m} e^{(-m/\Omega)R^2}, \quad (7)$$

where $\Gamma(\cdot)$ is the gamma function [24], and $m_{kl}^{(u)}$ is the Nakagami- m fading parameter, which characterises the severity of the fading over the l th resolvable path [28] between the u th transmission antenna and user k . Furthermore, the parameter Ω_{kl}^u in (7) is defined as $\Omega_{kl}^u = E[(\alpha_{kl}^u)^2]$, which is assumed to be a negative exponentially decaying multipath intensity profile (MIP) given by $\Omega_{kl}^u = \Omega_{kl}^u e^{-\eta(l-1)}$, $\eta \geq 0$, where Ω_{kl}^u is the average signal strength corresponding to the first resolvable path and η is the rate of average power decay, while $(\alpha_{kl}^u)^2$ represents the individual coefficients of the MIP.

When supporting K asynchronous CDMA users and assuming perfect power control, the received complex lowpass equivalent signal can be expressed as

$$R(t) = \sum_{k=1}^K \sum_{l=1}^L \sqrt{\frac{2P}{U^2}} \mathbf{c}(t - \tau_{kl}) \mathbf{B}_U(t - \tau_{kl}) \mathbf{h}_{kl}$$

$$\times \text{PN}_k(t - \tau_{kl}) + N(t), \quad (8)$$

where $N(t)$ is the complex-valued lowpass-equivalent additive white Gaussian noise (AWGN) having a double-sided spectral density of N_0 , while

$$\mathbf{h}_{kl} = \begin{pmatrix} h_{kl}^1 \exp(j\psi_{kl}^1) \\ h_{kl}^2 \exp(j\psi_{kl}^2) \\ \vdots \\ h_{kl}^U \exp(j\psi_{kl}^U) \end{pmatrix}, \quad k = 1, 2, \dots, K, \quad l = 1, 2, \dots, L, \quad (9)$$

represents the channel's complex impulse response in the context of the k th user and the l th resolvable path, where $\psi_{kl}^u = \phi_{kl}^u - 2\pi f_c \tau_{kl}$. Furthermore, in (8) we assumed that the signals transmitted by the U number of transmission antennas arrive at the receiver antenna after experiencing the same set of delays. This assumption is justified by the fact that in the frequency band of cellular system the propagation delay differences among the transmission antenna elements are on the order of nanoseconds, while the multipath delays are on the order of microseconds [21], provided that U is a relatively low number.

2.3. Receiver model

Let the first user be the user of interest and consider a receiver using space-time despreading as well as diversity combining, as shown in Figure 1b, where the subscript of the reference user's signal has been omitted for notational convenience. The receiver of Figure 1b carries out the inverse processing

of Figure 1a, in addition to multipath diversity combining. In Figure 1b, the received signal is first down-converted using the carrier frequency f_c , and then descrambled using the DS scrambling sequence of $\text{PN}(t - \tau_l)$ in the context of the l th resolvable path, where we assumed that the receiver is capable of achieving near-perfect multipath-delay estimation for the reference user. The descrambled signal associated with the l th resolvable path is space-time despread using the approach of [21]—which will be further discussed in Section 3, in order to obtain U separate variables, $\{Z_{1l}, Z_{2l}, \dots, Z_{Ul}\}$, corresponding to the U parallel data bits $\{b_1, b_2, \dots, b_U\}$, respectively. Following space-time despreading, a decision variable is formed for each parallel transmitted data bit of $\{b_1, b_2, \dots, b_U\}$ by combining the corresponding variables associated with the L number of resolvable paths, which can be expressed as

$$Z_u = \sum_{l=1}^L Z_{ul}, \quad u = 1, 2, \dots, U. \quad (10)$$

Finally, the U number of transmitted data bits $\{b_1, b_2, \dots, b_U\}$ can be decided based on the decision variables $\{Z_u\}_{u=1}^U$ using the conventional decision rule of a BPSK scheme.

Above we have described the transmitter model, the channel model, as well as the receiver model of W-CDMA using STS. We will now describe the detection procedure of the W-CDMA scheme using STS.

3. DETECTION OF SPACE-TIME SPREAD W-CDMA SIGNALS

Let $\mathbf{d}_l = [d_{l1} \ d_{l2} \ \dots \ d_{lU}]^T$, $l = 1, 2, \dots, L$, where T denotes vector transpose, represent the correlator's output variable vector in the context of the l th ($l = 1, 2, \dots, L$) resolvable path, where

$$d_{ul} = \int_{\tau_l}^{UT_b + \tau_l} R(t) c_u(t - \tau_l) \text{PN}(t - \tau_l) dt. \quad (11)$$

When substituting (8) into (11), it can be shown that

$$d_{ul} = \sqrt{2P} T_b \left[a_{u1} b_{u1} h_l^1 \exp(j\psi_l^1) + a_{u2} b_{u2} h_l^2 \exp(j\psi_l^2) \right. \\ \left. + \dots + a_{uU} b_{uU} h_l^U \exp(j\psi_l^U) \right] \\ + J_u(l), \quad u = 1, 2, \dots, U, \quad (12)$$

where

$$J_u(l) = J_{Su}(l) + J_{Mu}(l) + N_u(l), \quad u = 1, 2, \dots, U, \quad (13)$$

and $J_{Su}(l)$ is due to the multipath-induced self-interference of the signal of interest inflicted upon the l th path signal, where $J_{Su}(l)$ can be expressed as

$$J_{Su}(l) = \sum_{j=1, j \neq l}^L \sqrt{\frac{2P}{U^2}} \int_{\tau_l}^{UT_b + \tau_l} \mathbf{c}(t - \tau_j) \mathbf{B}_U(t - \tau_j) \mathbf{h}_j \text{PN}(t - \tau_j) \\ \times c_u(t - \tau_l) \text{PN}(t - \tau_l) dt, \quad (14)$$

$J_{Mu}(l)$ represents the multiuser interference due to the signals transmitted simultaneously by the other users, which can be expressed as

$$J_{Mu}(l) = \sum_{k=2}^K \sum_{j=1}^L \sqrt{\frac{2P}{U^2}} \int_{\tau_l}^{UT_b+\tau_l} \mathbf{c}(t - \tau_{kj}) \mathbf{B}_U(t - \tau_{kj}) \times \mathbf{h}_{kj} \text{PN}_k(t - \tau_{kj}) c_u(t - \tau_l) \text{PN}(t - \tau_l) dt, \quad (15)$$

and finally $N_u(l)$ is due to the AWGN, which can be written as

$$N_u(l) = \int_{\tau_l}^{UT_b+\tau_l} N(t) c_u(t - \tau_l) \text{PN}(t - \tau_l) dt, \quad (16)$$

which is a Gaussian distributed variable having zero mean and a variance of $2UN_0T_b$.

Let $\mathbf{J}(l) = [J_1(l) \ J_2(l) \ \cdots \ J_U(l)]^T$. Then, the correlator's output variable vector \mathbf{d}_l can be expressed as

$$\mathbf{d}_l = \sqrt{2PT_b} \mathbf{B}_U \mathbf{h}_l + \mathbf{J}(l), \quad l = 1, 2, \dots, L, \quad (17)$$

where \mathbf{B}_U is the reference user's $U \times U$ -dimensional transmitted data matrix, which is given by (2), but ignoring the time dependence, while \mathbf{h}_l is the channel's complex impulse

response between the base station and the reference user, as shown in (9) in the context of the reference user.

The attractive STS schemes of Hochwald et al. have the property [21] of $\mathbf{B}_U \mathbf{h}_l = \mathbf{H}_U \mathbf{b}$, that is, (17) can be written as

$$\mathbf{d}_l = \sqrt{2PT_b} \mathbf{H}_U \mathbf{b} + \mathbf{J}(l), \quad (18)$$

where $\mathbf{b} = [b_1 \ b_2 \ \cdots \ b_U]^T$ represents the U number of transmitted data bits, while \mathbf{H}_U is a $U \times U$ -dimensional matrix with elements from \mathbf{h}_l . Each element of \mathbf{h}_l appears once and only once in a given row and also in a given column of the matrix \mathbf{H}_U [21]. The matrix \mathbf{H}_U can be expressed as

$$\mathbf{H}_U(l) = \begin{pmatrix} \alpha_{11}(l) & \alpha_{12}(l) & \cdots & \alpha_{1U}(l) \\ \alpha_{21}(l) & \alpha_{22}(l) & \cdots & \alpha_{2U}(l) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{U1}(l) & \alpha_{U2}(l) & \cdots & \alpha_{UU}(l) \end{pmatrix}, \quad (19)$$

where $\alpha_{ij}(l)$ takes the form of $d_{ij} h_l^m \exp(j\psi_l^m)$, and $d_{ij} \in \{+1, -1\}$ represents the sign of the (i, j) th element of \mathbf{H}_U , while $h_l^m \exp(j\psi_l^m)$ belongs to the m th element of \mathbf{h}_l . For $U = 2, 4$, with the aid of [21], it can be shown that

$$\mathbf{H}_2(l) = \begin{pmatrix} h_l^1 \exp(j\psi_l^1) & h_l^2 \exp(j\psi_l^2) \\ -h_l^2 \exp(j\psi_l^2) & h_l^1 \exp(j\psi_l^1) \end{pmatrix},$$

$$\mathbf{H}_4(l) = \begin{pmatrix} h_l^1 \exp(j\psi_l^1) & h_l^2 \exp(j\psi_l^2) & h_l^3 \exp(j\psi_l^3) & h_l^4 \exp(j\psi_l^4) \\ -h_l^2 \exp(j\psi_l^2) & h_l^1 \exp(j\psi_l^1) & -h_l^4 \exp(j\psi_l^4) & h_l^3 \exp(j\psi_l^3) \\ -h_l^3 \exp(j\psi_l^3) & h_l^4 \exp(j\psi_l^4) & h_l^1 \exp(j\psi_l^1) & -h_l^2 \exp(j\psi_l^2) \\ -h_l^4 \exp(j\psi_l^4) & -h_l^3 \exp(j\psi_l^3) & h_l^2 \exp(j\psi_l^2) & h_l^1 \exp(j\psi_l^1) \end{pmatrix}. \quad (20)$$

With the aid of the analysis in [21], it can be shown that the matrix $\mathbf{H}_U(l)$ has the property of $\text{Re}\{\mathbf{H}_U^\dagger(l) \mathbf{H}_U(l)\} = \mathbf{h}_l^\dagger \mathbf{h}_l \cdot \mathbf{I}$, where \dagger denotes complex conjugate transpose and \mathbf{I} represents a $U \times U$ -dimensional unity matrix. Letting $\mathbf{h}_u(l)$ denote the u th column of $\mathbf{H}_U(l)$, the variable Z_{ul} in (10) can be expressed as [21]

$$Z_{ul} = \text{Re}\{\mathbf{h}_u^\dagger(l) \mathbf{d}_l\} = \sqrt{2PT_b} b_u \sum_{u=1}^U |h_l^u|^2 + \text{Re}\{\mathbf{h}_u^\dagger(l) \mathbf{J}(l)\},$$

$$u = 1, 2, \dots, U. \quad (21)$$

Finally, according to (10) the decision variables associated

with the U parallel transmitted data bits $\{b_1, b_2, \dots, b_U\}$ of the reference user can be expressed as

$$Z_u = \sqrt{2PT_b} b_u \sum_{l=1}^L \sum_{u=1}^U |h_l^u|^2 + \sum_{l=1}^L \text{Re}\{\mathbf{h}_u^\dagger(l) \mathbf{J}(l)\},$$

$$u = 1, 2, \dots, U, \quad (22)$$

which shows that the receiver is capable of achieving a diversity order of UL , as indicated by the related sums of the first term.

Above we have analysed the detection procedure applicable to W-CDMA signals generated using STS. We will now derive the corresponding BER expression.

4. BER PERFORMANCE

4.1. BER analysis

In this section, we derive the BER expression of the STS-assisted W-CDMA system by first analysing the statistics of the variable Z_u , $u = 1, 2, \dots, U$, with the aid of the Gaussian approximation [23]. According to (22), for a given set of complex channel transfer factor estimates $\{h_l^u\}$, Z_u can be approximated as a Gaussian variable having a mean given by

$$E[Z_u] = \sqrt{2P}T_b b_u \sum_{l=1}^L \sum_{u=1}^U |h_l^u|^2. \quad (23)$$

Based on the assumption that the interferences imposed by the different users, by the different paths, as well as by the AWGN constitute independent random variables, the variance of Z_u can be expressed as

$$\begin{aligned} \text{Var}[Z_u] &= E \left[\left(\sum_{l=1}^L \text{Re} \{ \mathbf{h}_u^\dagger(l) \mathbf{J}(l) \} \right)^2 \right] \\ &= \sum_{l=1}^L E \left[\left(\text{Re} \{ \mathbf{h}_u^\dagger(l) \mathbf{J}(l) \} \right)^2 \right] \\ &= \frac{1}{2} \sum_{l=1}^L E \left[\left(\mathbf{h}_u^\dagger(l) \mathbf{J}(l) \right)^2 \right]. \end{aligned} \quad (24)$$

Substituting $\mathbf{h}_u(l)$, which is the u th column of $\mathbf{H}_u(l)$ in (19), and $\mathbf{J}(l)$ having elements given by (13) into the above equation, it can be shown that for a given set of channel estimates $\{h_l^u\}$, (24) can be simplified as

$$\begin{aligned} \text{Var}[Z_u] &= \frac{1}{2} \sum_{l=1}^L \sum_{u=1}^U |h_l^u|^2 E \left[(J_u(l))^2 \right] \\ &= \frac{1}{2} \sum_{l=1}^L \sum_{u=1}^U |h_l^u|^2 \text{Var}[J_u(l)], \end{aligned} \quad (25)$$

where $J_u(l)$ is given by (13). In deriving (25) we exploited the assumption of $\text{Var}[J_1(l)] = \text{Var}[J_2(l)] = \dots = \text{Var}[J_U(l)]$.

As shown by Hochwald et al. in (13), $J_u(l)$ consists of three terms, namely the AWGN $N_u(l)$ having a variance of $2UN_0T_b$, $J_{Su}(l)$, which is the multipath-induced self-interference inflicted upon the l th path of the user of interest, and $J_{Mu}(l)$ imposed by the $(K-1)$ interfering users. By careful observation of (14), it can be shown that $J_{Su}(l)$ consists of U^2 terms and each term takes the form of $\sum_{j=1, j \neq l}^L \sqrt{2P/U^2} \int_{\tau_l}^{UT_b + \tau_l} c_m(t - \tau_j) a_{mn} b_{mn}(t - \tau_j) h_j^n \exp(j\psi_j^n) \text{PN}(t - \tau_j) \times c_u(t - \tau_l) \text{PN}(t - \tau_l) dt$. Assuming that $E[(h_j^n)^2] = \Omega_1 e^{-\eta(j-1)}$, that is, that $E[(h_j^n)^2]$ is independent of the index of the transmission antenna, and following the analysis in [22], it can be shown that the above term has a variance of $2\Omega_1 E_b T_b [q(L, \eta) - 1]/(GU)$, where $q(L, \eta) = (1 - e^{-L\eta})/(1 - e^{-\eta})$, if $\eta \neq 0$ and $q(L, \eta) = L$, if $\eta = 0$. Consequently, we have $\text{Var}[J_{Su}(l)] =$

$U^2 \times 2\Omega_1 E_b T_b [q(L, \eta) - 1]/(GU) = 2U\Omega_1 E_b T_b [q(L, \eta) - 1]/G$. Similarly, the multiuser interference term $J_{Mu}(l)$ of (15) also consists of U^2 terms, and each term has the form of $\sum_{k=2}^K \sum_{j=1}^L \sqrt{2P/U^2} \int_{\tau_l}^{UT_b + \tau_l} c_m(t - \tau_{kj}) a_{mn} b_{mn}(t - \tau_{kj}) h_{kj}^n \exp(j\psi_{kj}^n) \text{PN}_k(t - \tau_{kj}) c_u(t - \tau_l) \text{PN}(t - \tau_l) dt$. Again, with the aid of the analysis in [22], it can be shown that this term has the variance of $(K-1)4\Omega_1 E_b T_b q(L, \eta)/(3GU)$, and consequently the variance of $J_{Mu}(l)$ is given by $\text{Var}[J_{Mu}(l)] = (K-1)4U\Omega_1 E_b T_b q(L, \eta)/(3G)$. Therefore, the variance of $J_u(l)$ can be expressed as

$$\begin{aligned} \text{Var}[J_u(l)] &= 2N_0 U T_b + \frac{2U\Omega_1 E_b T_b [q(L, \eta) - 1]}{G} \\ &\quad + \frac{(K-1)4U\Omega_1 E_b T_b q(L, \eta)}{3G}, \end{aligned} \quad (26)$$

and the variance of Z_u for a given set of channel estimates $\{h_l^u\}$ can be expressed as

$$\begin{aligned} \text{Var}[Z_u] &= \sum_{l=1}^L \sum_{u=1}^U |h_l^u|^2 \left[N_0 U T_b + \frac{U\Omega_1 E_b T_b [q(L, \eta) - 1]}{G} \right. \\ &\quad \left. + \frac{(K-1)2U\Omega_1 E_b T_b q(L, \eta)}{3G} \right]. \end{aligned} \quad (27)$$

Based on (23) and (27), the BER conditioned on h_l^u for $u = 1, 2, \dots, U$ and $l = 1, 2, \dots, L$ can be written as

$$P_b(E|\{h_l^u\}) = Q \left(\sqrt{\frac{E^2[Z_u]}{\text{Var}[Z_u]}} \right) = Q \left(\sqrt{2 \cdot \sum_{l=1}^L \sum_{u=1}^U \gamma_{lu}} \right), \quad (28)$$

where $Q(x)$ represents the Gaussian Q -function, which can also be represented in its less conventional form as $Q(x) = (1/\pi) \int_0^{\pi/2} \exp(-x^2/2 \sin^2 \theta) d\theta$, where $x \geq 0$ [28, 29]. Furthermore, γ_{lu} in (28) is given by

$$\gamma_{lu} = \bar{\gamma}_c \cdot \frac{(h_l^u)^2}{\Omega_1}, \quad (29)$$

$$\bar{\gamma}_c = \frac{1}{U} \left[\frac{(2K+1)q(L, \eta) - 3}{3G} + \left(\frac{\Omega_1 E_b}{N_0} \right)^{-1} \right]^{-1}.$$

The average BER, $P_b(E)$, can be obtained by averaging the conditional BER of (28) over the joint PDF of the instantaneous SNR values corresponding to the L multipath components and to the U transmit antennas $\{\gamma_{lu} : l = 1, 2, \dots, L; u = 1, 2, \dots, U\}$. Since the random variables $\{\gamma_{lu} : l = 1, 2, \dots, L; u = 1, 2, \dots, U\}$ are assumed to be statistically independent, the average BER can be expressed as [30, (23)]

$$P_b(E) = \frac{1}{\pi} \int_0^{\pi/2} \prod_{l=1}^L \prod_{u=1}^U I_{lu}(\bar{\gamma}_{lu}, \theta) d\theta, \quad (30)$$

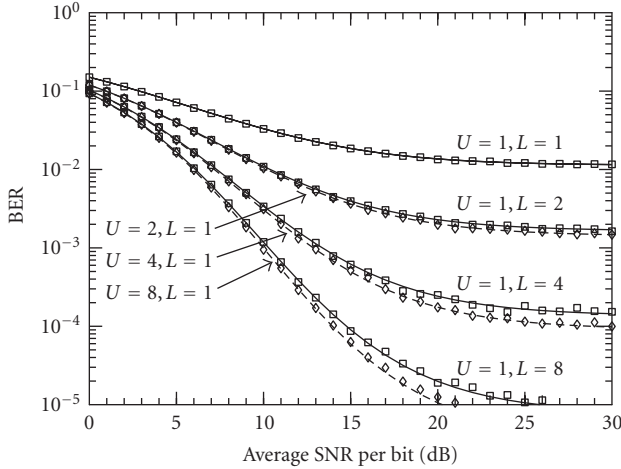


FIGURE 5: BER versus the SNR per bit, E_b/N_0 , performance comparison between the space-time-spreading-based transmit diversity scheme and the conventional RAKE receiver arrangement using only one transmission antenna when communicating over flat-fading (for space-time spreading) and multipath (for RAKE) Rayleigh fading ($m_l = m_c = 1$) channels evaluated from (35) by assuming that the average power decay rate was $\eta = 0$. The solid line indicates the BER of the receiver-diversity-aided schemes, while the dashed line that of the transmit-diversity-assisted schemes ($G = 128, K = 10$).

where

$$I_{lu}(\bar{\gamma}_{lu}, \theta) = \int_0^{\infty} \exp\left(-\frac{\gamma_{lu}}{\sin^2 \theta}\right) p_{\gamma_{lu}}(\gamma_{lu}) d\gamma_{lu}. \quad (31)$$

Since $\gamma_{lu} = \bar{\gamma}_c \cdot ((h_l^u)^2 / \Omega_1)$ and h_l^u obeys the Nakagami- m distribution characterised by (7), it can be shown that the PDF of γ_{lu} can be expressed as

$$p_{\gamma_{lu}}(\gamma_{lu}) = \left(\frac{m_l^{(u)}}{\bar{\gamma}_{lu}}\right)^{m_l^{(u)}} \frac{\gamma_{lu}^{m_l^{(u)}-1} \exp\left(-\frac{m_l^{(u)} \gamma_{lu}}{\bar{\gamma}_{lu}}\right)}{\Gamma(m_l^{(u)})}, \quad \gamma_{lu} \geq 0, \quad (32)$$

where $\bar{\gamma}_{lu} = \bar{\gamma}_c e^{-\eta(l-1)}$ for $l = 1, 2, \dots, L$.

Upon substituting (32) into (31) it can be shown that [28]

$$I_{lu}(\bar{\gamma}_{lu}, \theta) = \left(\frac{m_l^{(u)} \sin^2 \theta}{\bar{\gamma}_{lu} + m_l^{(u)} \sin^2 \theta}\right)^{m_l^{(u)}}. \quad (33)$$

Finally, upon substituting (33) into (30), the average BER of the STS-assisted W-CDMA system using U transmission antennas can be expressed as

$$P_b(E) = \frac{1}{\pi} \int_0^{\pi/2} \prod_{l=1}^L \prod_{u=1}^U \left(\frac{m_l^{(u)} \sin^2 \theta}{\bar{\gamma}_{lu} + m_l^{(u)} \sin^2 \theta}\right)^{m_l^{(u)}} d\theta, \quad (34)$$

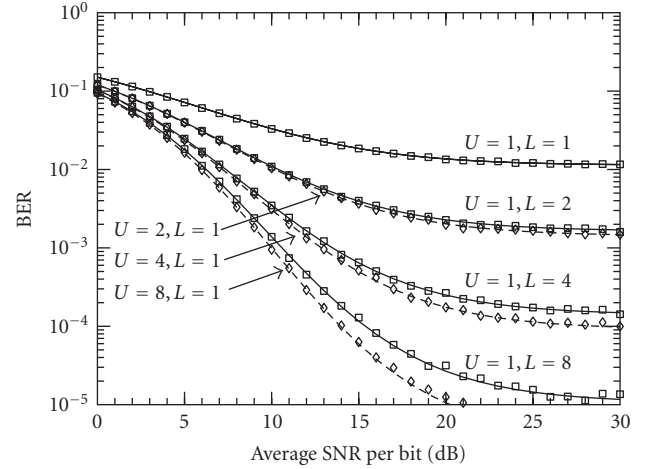


FIGURE 6: BER versus the SNR per bit, E_b/N_0 , performance comparison between the space-time-spreading-based transmit diversity scheme and the conventional RAKE receiver arrangement using only one transmission antenna when communicating over flat-fading (for space-time spreading) and multipath (for RAKE) Rayleigh fading ($m_l = m_c = 1$) channels evaluated from (35) by assuming that the average power decay rate was $\eta = 0.2$. The solid line indicates the BER of the receiver-diversity-aided schemes, while the dashed line that of the transmit-diversity-assisted schemes ($G = 128, K = 10$).

which shows that the diversity order achieved is LU —the product of the transmit diversity order and the frequency-selective diversity order. Furthermore, if we assume that $m_l^{(u)}$ is independent of u , that is, that all of the parallel transmitted subsignals experience an identical Nakagami fading, then (34) can be expressed as

$$P_b(E) = \frac{1}{\pi} \int_0^{\pi/2} \prod_{l=1}^L \left(\frac{m_l \sin^2 \theta}{\bar{\gamma}_{lu} + m_l \sin^2 \theta}\right)^{Um_l} d\theta. \quad (35)$$

4.2. Numerical results and discussions

In Figures 5, 6, 7, 8, and 9 we compare the BER performance of the STS-assisted W-CDMA system transmitting over flat-fading channels and that of the conventional RAKE receiver using only one transmission antenna, but communicating over frequency-selective fading channels. The results in these figures were all evaluated from (35) by assuming appropriate parameters, which are explicitly shown in the corresponding figures. In Figures 5, 6, and 7 the BER was drawn against the SNR/bit, namely E_b/N_0 , while in Figures 8 and 9 the BER was drawn against the number of users, K , supported by the system. From the results we observe that for transmission over Rayleigh fading channels ($m_l = 1$), as characterised by Figures 5, 6, and 8, both the STS-based transmit diversity scheme transmitting over the frequency-nonspecific Rayleigh fading channel and the conventional RAKE receiver scheme communicating over frequency-selective Rayleigh fading channels

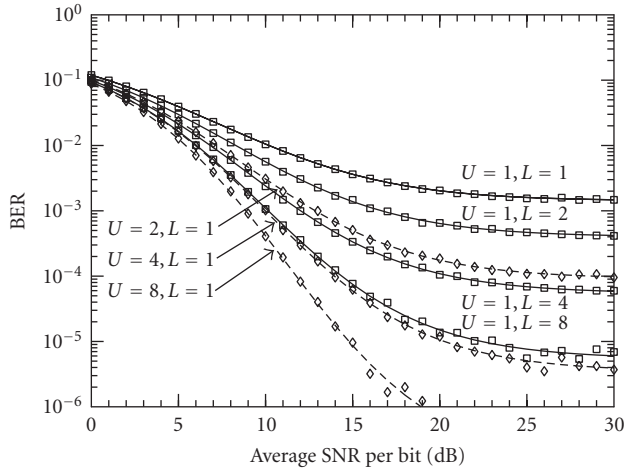


FIGURE 7: BER versus the SNR per bit, E_b/N_0 , performance comparison between the space-time-spreading-based transmit diversity scheme and the conventional RAKE receiver arrangement using only one transmission antenna when communicating over flat-fading (for space-time spreading) and multipath (for RAKE) Nakagami- m fading channels evaluated from (35) by assuming that the average power decay rate was $\eta = 0.2$, where $m_1 = 2$ indicates that the first resolvable path constitutes a moderately fading path, while the other resolvable paths experience more severe Rayleigh fading ($m_c = 1$). The solid line indicates the BER of the receiver-diversity-aided schemes, while the dashed line that of the transmit-diversity-assisted schemes ($G = 128, K = 10$).

having the same number of resolvable paths as the number of transmission antennas in the STS-assisted scheme achieved a similar BER performance, with the STS scheme slightly outperforming the conventional RAKE scheme. For transmission over general Nakagami- m fading channels, if the first resolvable path is less severely faded, than the other resolvable paths, such as in Figures 7 and 9 where $m_1 = 2$ and $m_2 = m_3 = \dots = m_c = 1$, the STS-based transmit diversity scheme communicating over the frequency-nonselective Rayleigh fading channel may significantly outperform the corresponding conventional RAKE-receiver-assisted scheme communicating over frequency-selective Rayleigh fading channels. This is because the STS-based transmit diversity scheme communicated over a single nondispersive path, which benefited from having a path experiencing moderate fading. However, if the number of resolvable paths is sufficiently high, the conventional RAKE receiver scheme is also capable of achieving a satisfactory BER performance.

Above we assumed that the number of resolvable paths was one, if the STS using more than one antenna was considered. By contrast, the number of resolvable paths was equal to the number of transmit antennas of the corresponding STS-based system, when the conventional RAKE receiver was considered. However, in practical W-CDMA systems the number of resolvable paths of each antenna's transmitted signal depends on its transmission environment. The number of resolvable paths dynamically changes, as the mobile

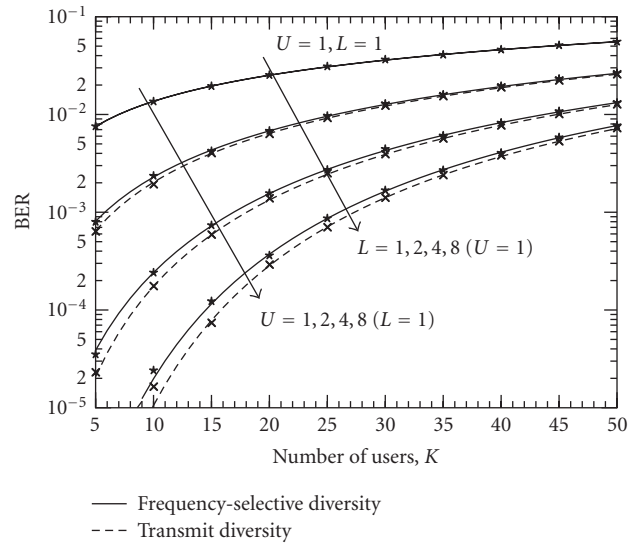


FIGURE 8: BER versus the number of users, K , performance comparison between the space-time-spreading-based transmit diversity scheme and the conventional RAKE receiver arrangement using only one transmission antenna when communicating over flat-fading (for space-time spreading) and multipath (for RAKE) Rayleigh fading channels evaluated from (35) by assuming that the average power decay rate was $\eta = 0$ ($G = 128, E_b/N_0 = 20$ dB, $m_1 = m_c = 1$).

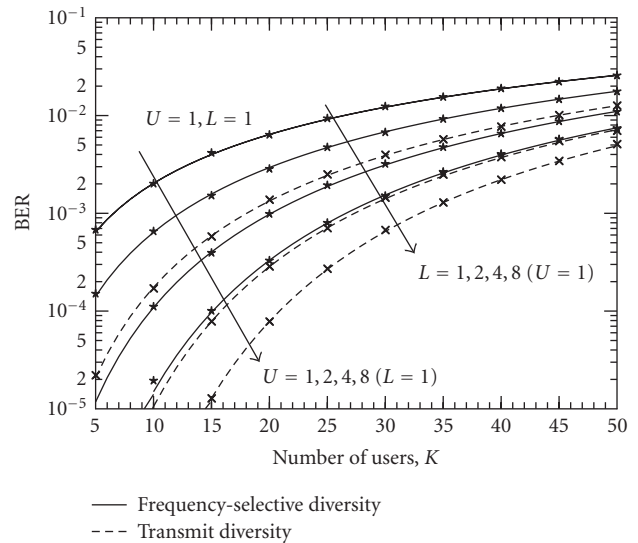


FIGURE 9: BER versus the number of users, K , performance comparison between the space-time-spreading-based transmit diversity scheme and the conventional RAKE receiver arrangement using only one transmission antenna when communicating over flat-fading (for space-time spreading) and multipath (for RAKE) Nakagami- m fading channels evaluated from (35) by assuming that the average power decay rate was $\eta = 0.2$, where $m_1 = 2$ indicates that the first resolvable path constitutes a moderately fading path, while the other resolvable paths experience more severe Rayleigh fading ($m_c = 1$); $G = 128, E_b/N_0 = 20$ dB.

traverses through different transmission environments. Specifically, in some scenarios the number of paths may be as low as $L = 1$, and in other scenarios it may be as high as $L > 10$. When the number of resolvable paths is as low as $L = 1$ or 2, employing STS-based transmit diversity is particularly valuable. However, when the number of resolvable paths is reasonably high, for example, $L > 4$, the employment of STS-based transmit diversity may not be necessary. An attractive approach is to adapt the mode of operation of the STS scheme, which is discussed in the following section.

5. DISPERSION-CONTROLLED ADAPTIVE SPACE-TIME SPREADING

The main philosophy behind the proposed channel-induced dispersion-controlled adaptive STS scheme is the real-time balancing of the link budget through the adaptive control of the STS-based transmission scheme, in order that the system achieves its maximum throughput, while maintaining the required target BER performance. More specifically, in this treatise we will aim for maintaining a target BER of 10^{-4} , regardless of the instantaneous channel quality experienced and exploit the improved channel quality provided by a higher number of resolvable multipath components experienced in scattering rich outdoor channels for increasing the system's effective throughput, ultimately leading to a potentially better speech [6] or video [7] quality for the users of the system.

In the context of the STS-assisted W-CDMA system, the delay spread of the wireless channels, and hence the number of resolvable paths, varies slowly over a range spanning from one to dozens of paths. The STS scheme designed based on a low number of resolvable paths, or even based on a relatively high but constant number of resolvable paths, cannot maximise the achievable throughput. For example, if the STS scheme is designed based on a low number of resolvable paths, in order to guarantee a required quality of service (QoS), the practically achieved QoS may be excessive, when the number of resolvable paths is high, provided that these resolvable paths are efficiently combined. However, if only a low but constant number of resolvable paths is combined, the diversity potential provided by the high number of resolvable paths is inevitably wasted. A high-efficiency STS-based communication scheme must be capable of combining the transmitted energy, which was scattered over an arbitrary number of resolvable paths, and the mode of operation of the STS scheme can be adaptively controlled according to the receiver's detection performance.

When the number of resolvable paths is low and hence the resultant BER is higher than the required BER, then a low throughput STS-assisted transmitter mode is activated, which exhibits a high transmit diversity gain, as it will be demonstrated below with the aid of an example. By contrast, when the number of resolvable paths is high and hence the resultant BER is lower than the required BER, then

a higher throughput STS-assisted transmitter mode is activated, which has a lower transmit diversity gain.²

Specifically, the principle of implementing channel-dispersion-controlled adaptive rate transmission using adaptive STS may be readily interpreted by referring to the following example. Let the transmitter employ a total of four transmission antennas. If the number of resolvable paths experienced by the receiver is low, the transmitter is instructed by the receiver to employ an STS scheme based on four transmit antennas, using the STS scheme described as [21]

$$\mathbf{S} = [c_1 \ c_2 \ c_3 \ c_4] \begin{pmatrix} b_1 & b_2 & b_3 & b_4 \\ b_2 & -b_1 & b_4 & -b_3 \\ b_3 & -b_4 & -b_1 & b_2 \\ b_4 & b_3 & -b_2 & -b_1 \end{pmatrix}, \quad (36)$$

where c_1, c_2, c_3, c_4 are four STS-related orthogonal codes having a period of $4T_b$. The above STS scheme transmits $U = 4$ parallel data bits during the interval of $4T_b$, and hence the effective transmission rate becomes $R_b = 4 \times 1/4T_b = 1/T_b$, as seen in Figure 2. By contrast, when the number of resolvable paths increases, the transmitter is instructed by the receiver to employ four separate STS schemes, each based on two transmit antennas, as seen in Figure 4, which can be formulated as

$$\mathbf{S} = \begin{pmatrix} [c_1 \ c_2] \begin{pmatrix} b_1 & b_2 \\ b_2 & -b_1 \end{pmatrix} & [c_3 \ c_4] \begin{pmatrix} b_3 & b_4 \\ b_4 & -b_3 \end{pmatrix} \\ [c_1 \ c_2] \begin{pmatrix} b_5 & b_6 \\ b_6 & -b_5 \end{pmatrix} & [c_3 \ c_4] \begin{pmatrix} b_7 & b_8 \\ b_8 & -b_7 \end{pmatrix} \end{pmatrix}, \quad (37)$$

which, again, constitutes the four independent two-antenna-based STS schemes $\mathbf{B}_2(t)$ of (3), where c_1, c_2, c_3, c_4 are the $U = 4$ STS-related orthogonal codes having a period of $2T_b$. Based on the above four two-antenna-assisted STS schemes, $U = 4$ parallel data bits are transmitted during the first $2T_b$ -duration interval using the STS scheme $\mathbf{B}_2(t)$ of (3). Specifically, antennas 1 and 2 are activated with the aid of c_1, c_2 , while activating antennas 3 and 4 using c_3, c_4 , as portrayed in Figure 4. During the following $2T_b$ -duration slot another

²The transmitter does not necessarily have to have the explicit knowledge of the number of resolvable paths, there is a range of other criteria, which may be used for controlling the activation of the different antenna configurations. Firstly, since most existing systems employ explicit training for estimating the channel's impulse response (CIR), the significant-energy CIR taps explicitly quantify the number of resolvable multipath components. Another practical metric that may be used for activating the required antenna configuration is the bit error ratio (BER) estimated, for example, by the channel decoder's soft metrics. When the estimated BER is higher than the target BER, the transmitter is instructed to increase its spreading gain and hence reduce its throughput, as well as vice versa. The activation regime has to be conservative for the sake of maintaining the target BER even if the BER was underestimated. Finally, the Doppler frequency does not dramatically affect the system's performance, since the *amount of dispersion*, that is, the CIR duration, changes only, when traversing from an indoor-type nondispersive environment to an outdoor scenario and then to a rural scenario, which may require 10 minutes for the dispersion to change substantially. However, the system's increased throughput is achieved at the cost of an increased complexity.

four data bits are transmitted using the same scheme as outlined above. Consequently, the above four two-antenna-based STS schemes transmit a total of eight data bits during two consecutive $2T_b$ -duration time slots having a total duration of $4T_b$, and the effective transmission rate is now doubled to $2R_b$. Furthermore, if the number of resolvable paths is sufficiently high, which results in requiring no transmit diversity at all, then the four transmission antennas can transmit their information independently, as demonstrated in Figure 3 and the corresponding transmission mode can be described as

$$\mathbf{S} = \begin{pmatrix} c_1b_1 & c_2b_2 & c_3b_3 & c_4b_4 \\ c_1b_5 & c_2b_6 & c_3b_7 & c_4b_8 \\ c_1b_9 & c_2b_{10} & c_3b_{11} & c_4b_{12} \\ c_1b_{13} & c_2b_{14} & c_3b_{15} & c_4b_{16} \end{pmatrix}, \quad (38)$$

which implies that each of the 16 bits is transmitted independently using an antenna within a duration T_b , where c_1, c_2, c_3, c_4 are four orthogonal codes having a period of T_b , each mapped to one antenna. Explicitly, this scheme is capable of transmitting a total of 16 data bits during an interval of $4T_b$, and hence we achieve a transmission rate of $4R_b$, as exemplified in Figure 3.

The PDF of the delay spread in a wireless communication channel can be approximated by a negative exponential distribution given by [31]

$$f(\tau) = \frac{1}{T_m} \exp\left(-\frac{\tau - \tau_0}{T_m}\right), \quad \tau \geq \tau_0, \quad (39)$$

where the minimum delay τ_0 is the time required for the signal to propagate directly following the line of sight from the transmitter to the receiver, and T_m represents the mean square of the distribution, which is also the average value of the delay spread. Some typical examples of T_m in different environments are [25] $T_m < 0.1$ microseconds for an indoor environment, $T_m < 0.2$ microseconds for an open rural area, $T_m \approx 0.5$ microseconds for a suburban area, and $T_m \approx 3$ microseconds for a typical urban area. In (39), we let $\tau_r = (\tau - \tau_0)/T_c$. Then the PDF of τ_r can be expressed as

$$f(\tau_r) = \frac{1}{T_m/T_c} \exp\left(-\frac{\tau_r}{T_m/T_c}\right), \quad \tau_r \geq 0, \quad (40)$$

where T_m/T_c represents the average delay spread to chip-duration ratio, and $\lfloor x \rfloor + 1$ —where $\lfloor x \rfloor$ represents the largest integer not exceeding x —is the average number of resolvable paths, which has been widely used in the performance analysis of DS-CDMA systems transmitting over multipath fading channels.

Let the number of resolvable paths associated with the reference signal be L_r . For DS-CDMA signals having a chip duration of T_c , the number of near-instantaneous resolvable paths $L_r = \lfloor (\tau - \tau_0)/T_c \rfloor + 1$ can be modelled as a discrete random variable, which varies slowly depending on the communication environment encountered. For a given BER, let the maximum throughput conditioned on the number of resolvable paths L_r be $B(L_r)$. Ideally, assuming that the receiver

is capable of combining an arbitrary number of resolvable paths and that the transmitter has the perfect knowledge of the number of resolvable paths with the aid of a feedback channel, and that the feedback delay is negligible, the unconditional throughput, B , using adaptive STS can be written as

$$B = \sum_{L_r=1}^{\infty} P(L_r) \cdot B(L_r), \quad (41)$$

where $P(L_r)$ is the probability that there are L_r resolvable paths at the receiver. With the aid of (40), this probability can be approximated as

$$\begin{aligned} P(L_r) &= \int_{\max\{0, L_r - 1 - 0.5\}}^{L_r - 1 + 0.5} f(\tau_r) d\tau_r \\ &= \int_{\max\{0, L_r - 1 - 0.5\}}^{L_r - 1 + 0.5} \frac{1}{T_m/T_c} \exp\left(-\frac{\tau_r}{T_m/T_c}\right) d\tau_r \\ &= \exp\left(-\frac{\max\{0, L_r - 1 - 0.5\}}{T_m/T_c}\right) - \exp\left(-\frac{L_r - 1 + 0.5}{T_m/T_c}\right), \end{aligned} \quad (42)$$

where $[L_r - 1 - 0.5, L_r - 1 + 0.5]$ is the normalised delay spread range having L_r resolvable paths. In (41), $B(L_r)$ represents the maximum possible throughput conditioned on having L_r number of resolvable paths. For example, for the proposed adaptive STS scheme using four-antenna-based STS, two-antenna-based STS, as well as conventional single-antenna-based transmission, as characterised in (36), (37), and (38), $B(L_r)$ may achieve values of R_b , $2R_b$, or $4R_b$, respectively, depending on the specific number of resolvable paths encountered.

Figures 10 and 11 show the throughput versus SNR/bit performance of the STS-assisted W-CDMA system using a maximum of four antennas. The maximum dispersion of the propagation environment was $T_m = 0.1, 0.2, 0.5$, and 3 microseconds. The corresponding number of resolvable multipath components at the 3.84 Mchip/s chip rate of the third-generation systems [3] became $(T_m/T_c) + 1 = 1, 2, 3$, and 16, respectively. Depending on the number of resolvable paths at the receiver and on the corresponding achievable BER performance, the transmitter may activate one of the transmission schemes described by (36), (37), and (38). In our related investigations, the target BER was set to 0.01. Specifically, if a sufficiently high number of resolvable paths is encountered by the receiver, which results in a BER of less than 0.01 for the scheme described by (38), then the transmitter supports a bitrate of $4R_b$. If the number of resolvable paths is in a range, where the BER using the scheme described by (38) is higher than 0.01, but that of the STS scheme described by (37) is lower than 0.01, then the transmitter transmits at a rate of $2R_b$. Finally, if the number of resolvable paths is in a range, where the BER using the STS scheme described by (37) is higher than 0.01, but that described by (36) is lower than 0.01, then the transmitter transmits at a rate of R_b . Otherwise, if the number of resolvable paths is too low, which results in $\text{BER} > 0.01$ for the STS scheme described by (36), then the transmitter simply disables transmissions.

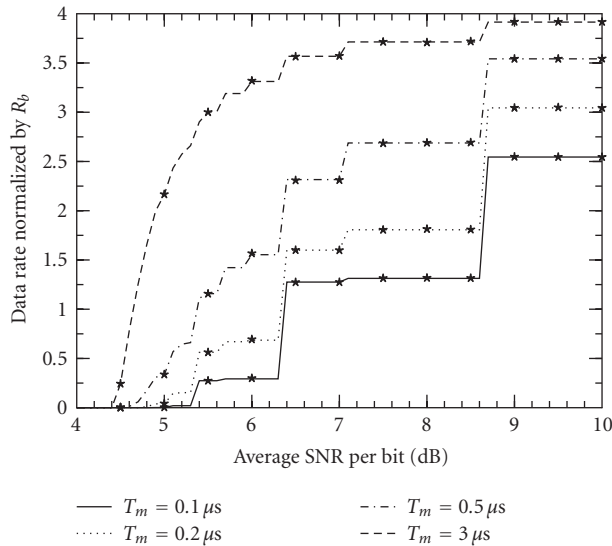


FIGURE 10: Normalized throughput versus the SNR per bit, E_b/N_0 , performance of the adaptive space-time-spreading-assisted W-CDMA system using four-antenna-based STS of (36), the two-antenna-aided STS of (37), and the conventional single-antenna scheme for transmission over four typical wireless channels experiencing Rayleigh fading ($m = 1$). The target BER of the reference user is 0.01 and there are no interference users, that is, $K = 1$ ($G = 128, \eta = 0, R_{\text{chip}} = 3.686$ Mcps/s).

In the context of Figure 10 we assumed that the number of users was $K = 1$, and that the fading associated with each resolvable path obeyed the Rayleigh distribution ($m = 1$). By contrast, in Figure 11 we assumed that the number of users was $K = 10$, and that the fading associated with the first resolvable path obeyed the Nakagami- m distribution in conjunction with $m = 2$, while the fading of the other resolvable paths obeyed the Rayleigh distribution ($m_c = 1$).

From the results of Figures 10 and 11 we observe that with the aid of the adaptive STS scheme, the system's effective throughput is significantly increased, if the average delay spread of the channel is sufficiently high or, in other words, if the number of resolvable paths varies over a sufficiently wide range. We will highlight the significance of this observation in more detail. Using $T_m = 0.5$ microseconds and 3 microseconds as examples and by observing Figure 10 we find that the SNR/bit required for transmitting at the data rate of R_b is about 5.2 dB for $T_m = 0.5$ microseconds and 4.6 dB for $T_m = 3$ microseconds. Similarly, the SNR/bit required for supporting the data rate of $3R_b$ is about 6.4 dB for $T_m = 0.5$ microseconds and 5 dB for $T_m = 3$ microseconds. Hence, the adaptive STS-assisted W-CDMA system increased the achievable transmission rate by a factor of three, while requiring only a modest transmitted power increase of about 1.2 dB for $T_m = 0.5$ microseconds and 0.4 dB for $T_m = 3$ microseconds. Similar results can also be observed in Figure 11, where an extra 0.4 dB or 1.2 dB transmitted power is required for achieving a data rate of $3R_b$ instead of R_b . However, if

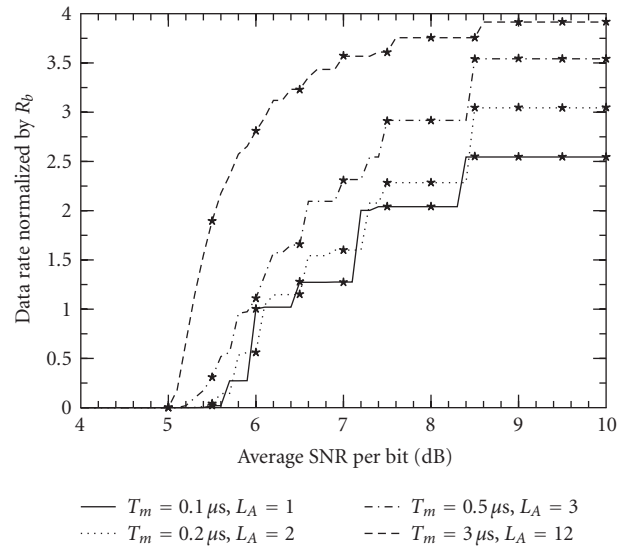


FIGURE 11: Normalized throughput versus the SNR per bit, E_b/N_0 , performance of the adaptive space-time-spreading-assisted W-CDMA system using the four-antenna-based STS of (36), the two-antenna-aided STS of (37), and the conventional single-antenna scheme for transmission over four typical wireless channels obeying the Nakagami- m distribution ($m_1 = 2, m_c = 1$). The target BER of the reference user is 0.01, while the interfering users communicate using the four-antenna-based STS of (36) and each interfering signal has an average of L_A number of resolvable paths ($G = 128, K = 10, \eta = 0, R_{\text{chip}} = 3.686$ Mcps/s).

the number of resolvable paths varies over a relatively low range, the required increase of the transmitted power becomes higher. For example, for the case of $T_m = 0.1$ microseconds in Figures 10 and 11 an extra 2.2 dB (Figure 10) or 1.2 dB (Figure 11) transmitted power must be invested, in order to achieve a data rate of $2R_b$ instead of R_b . In this scenario, due to the associated extra complexity of the adaptive STS-assisted scheme required by the channel dispersion estimation and feedback, and due to the control channel requirement of the dispersion feedback, the adaptive STS-aided scheme might not constitute a more attractive alternative. The system's increased effective throughput ultimately leads to a potentially better speech [6] or video [7] service quality for the users of the system.

6. CONCLUSIONS

In this contribution, we have investigated the performance of STS-assisted W-CDMA systems, when multipath Nakagami- m fading, multiuser interference, and background noise-induced impairments are considered. Our analysis and numerical results demonstrated that the achievable diversity order is the product of the frequency selective diversity order and the transmit diversity order. Furthermore, both the transmit diversity and the frequency selective diversity have a similar influence on the BER performance of the W-CDMA systems considered. Since W-CDMA signals typically experience high-dynamic frequency-selective fading in both urban

and suburban areas, the proposed adaptive transmit diversity scheme will result into an increased throughput and ultimately in a potentially better speech [6] or video [7] service quality for the users of the system. Based on the above scenarios, we proposed an adaptive STS transmission scheme, which adapts its STS configuration using (36), (37), and (38) according to the frequency selectivity information fed back from the receivers. The numerical results show that by efficiently exploiting the channel's frequency selectivity, the proposed adaptive STS scheme is capable of significantly improving the throughput of W-CDMA systems. For W-CDMA systems transmitting at a data rate of $3R_b$ instead of R_b , only an extra of 0.4 dB and 1.2 dB transmitted power is required in the urban and suburban areas considered, respectively, which results in a substantially increased speech [6] or video [7] service quality. Alternatively, a potentially higher number of users may be supported within the same bandwidth, as a benefit of cross-layer optimisation.

A number of related open research problems may be identified, such as the design of more sophisticated STS-aided multicarrier CDMA transceivers. The design of new STS codes is also a promising research area. A particularly promising research topic is designing large area synchronous (LAS) STS schemes, which exhibit a so-called interference-free window (IFW). Provided that the interfering signals arrive within this IFW, no multiuser interference is inflicted. Finally, quantifying the achievable network-layer benefits [3] of STS-aided CDMA systems is an important open problem.

ACKNOWLEDGMENTS

This work has been partly funded in the framework of the IST Project PHOENIX, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues. The financial support of the EPSRC, UK, is also acknowledged.

REFERENCES

- [1] L. Hanzo, C. H. Wong, and M. S. Yee, *Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalized and Space-Time Coded TDMA, CDMA, and OFDM Systems*, John Wiley, New York, NY, USA, 2002.
- [2] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels*, John Wiley, New York, NY, USA, 2002.
- [3] J. S. Blogh and L. Hanzo, *Third-Generation Systems and Intelligent Wireless Networking: Smart Antennas and Adaptive Modulation*, John Wiley, New York, NY, USA, 2002.
- [4] L. Hanzo, M. Münster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-user Communications, WLANs and Broadcasting*, Wiley-IEEE Press, New York, NY, USA, 2003.
- [5] L. Hanzo, L.-L. Yang, E.-L. Kuan, and K. Yen, *Single and Multi-Carrier DS-SS: Multi-User Detection, Space-Time Spreading, Synchronisation, Standards and Networking*, Wiley-IEEE Press, New York, NY, USA, 2003.
- [6] L. Hanzo, F. C. A. Somerville, and J. P. Woodard, *Voice Compression and Communications: Principles and Applications for Fixed and Wireless Channels*, Wiley-IEEE Press, New York, NY, USA, 2001, <http://www-mobile.ecs.soton.ac.uk>.
- [7] L. Hanzo, P. Cherriman, and J. Streit, *Wireless Video Communications: Second to Third Generation and Beyond*, IEEE Press, New York, NY, USA, 2001.
- [8] ITU-T Rec. H.26L/ISO/IEC 11496-10, "Advanced video coding," September 2002.
- [9] E. A. Brewer, R. H. Katz, Y. Chawathe, et al., "A network architecture for heterogeneous mobile computing," *IEEE Pers. Commun.*, vol. 5, no. 5, pp. 8–24, 1998.
- [10] L. Tong, Q. Zhao, and G. Mergen, "Multipacket reception in random access wireless networks: from signal processing to optimal medium access control," *IEEE Commun. Mag.*, vol. 39, no. 11, pp. 108–112, 2001.
- [11] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Communications Magazine*, vol. 9, no. 4, pp. 8–27, 2002.
- [12] C. Comaniciu and H.V. Poor, "Jointly optimal power and admission control for delay sensitive traffic in CDMA networks with LMMSE receivers," *IEEE Trans. Signal Processing*, vol. 51, no. 8, pp. 2031–2042, 2003.
- [13] S. Shakkottai, T. S. Rappaport, and P. C. Karlsson, "Cross-layer design for wireless networks," *IEEE Commun. Mag.*, vol. 41, no. 10, pp. 74–80, 2003.
- [14] A. Maharshi, T. Tong, and A. Swami, "Cross-layer designs of multichannel reservation MAC under Rayleigh fading," *IEEE Trans. Signal Processing*, vol. 51, no. 8, pp. 2054–2067, 2003.
- [15] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, 1998.
- [16] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE J. Select. Areas Commun.*, vol. 17, no. 3, pp. 451–460, 1999.
- [17] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "A space-time coding modem for high-data-rate wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1459–1478, 1998.
- [18] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [19] Proposed TDOC: 662/98 to ETSI SMG2 UMTS Standards, *Space-time block coded transmit antenna diversity for WCDMA*, December 1998.
- [20] Telcomm. Industry Association (TIA), *TIA/EIA Interim Standard: Physical Layer Standard for cdma2000 Standards for Spread Spectrum Systems*, 2000.
- [21] B. Hochwald, T. L. Marzetta, and C. B. Papadias, "A transmitter diversity scheme for wideband CDMA systems based on space-time spreading," *IEEE J. Select. Areas Commun.*, vol. 19, no. 1, pp. 48–60, 2001.
- [22] T. Eng and L. B. Milstein, "Coherent DS-SS performance in Nakagami multipath fading," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 1134–1143, 1995.
- [23] M. B. Pursley, "Performance evaluation for phase-coded spread-spectrum multiple-access communication-Part I: System analysis," *IEEE Trans. Commun.*, vol. 25, no. 8, pp. 795–799, 1977.
- [24] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 3rd edition, 1995.
- [25] W. C. Y. Lee, *Mobile Communications Engineering*, McGraw-Hill, New York, NY, USA, 2nd edition, 1998.
- [26] N. Nakagami, "The m -distribution, a general formula for intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*, W. G. Hoffman, Ed., Pergamon, Oxford, England, 1960.

- [27] V. Aalo, O. Ugweje, and R. Sudhakar, "Performance analysis of a DS/CDMA system with noncoherent M -ary orthogonal modulation in Nakagami fading," *IEEE Trans. Veh. Technol.*, vol. 47, no. 1, pp. 20–29, 1998.
- [28] M.-S. Alouini and A. J. Goldsmith, "A unified approach for calculating error rates of linearly modulated signals over generalized fading channels," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1324–1334, 1999.
- [29] M. K. Simon and M.-S. Alouini, "A unified approach to the probability of error for noncoherent and differentially coherent modulations over generalized fading channels," *IEEE Trans. Commun.*, vol. 46, no. 12, pp. 1625–1638, 1998.
- [30] M. K. Simon and M. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. IEEE*, vol. 86, no. 9, pp. 1860–1877, 1998.
- [31] L. E. Miller and J. S. Lee, *CDMA Systems Engineering Handbook*, Artech House, Boston, Mass, USA, 1998.

conference sessions, presented overview lectures, and has been awarded a number of distinctions. Currently he is managing an academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC), UK, the European IST Programme, and the Mobile Virtual Centre of Excellence (VCE), UK. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. Dr. Hanzo is also an IEEE Distinguished Lecturer of both the Communications Society and the Vehicular Technology Society as well as a Fellow of both the IEEE and IEE. For further information on research in progress and associated publications, please refer to <http://www-mobile.ecs.soton.ac.uk>.

Lie-Liang Yang received his M.Eng and Ph.D. degrees in communications and electronics from Northern Jiaotong University, Beijing, China, in 1991 and 1997, respectively, and his B.Eng. degree in communications engineering from Shanghai Tiedao University, Shanghai, China, in 1988. Since December 1997, he has been with the Communications Research Group at the Department of Electronics and Computer Science, University of Southampton, UK, where he held various research posts as a Visiting Postdoctoral Research Fellow, Research Fellow, and Senior Research Fellow. He currently holds an academic post as a Lecturer. From June 1997 to December 1997 he was a Visiting Scientist of the IREE, The Academy of Sciences of the Czech Republic. He has been involved in a number of projects funded by the National Science Foundation of China, the Grant Agency of the Czech Republic, the Engineering and Physical Sciences Research Council (EPSRC) of UK, and the European Union. His research covers a wide range of areas in communications, which include data network and security, intelligent wireless networking, error control coding, modulation and demodulation, spread-spectrum communications and multiuser detection, pseudonoise (PN) code synchronisation, smart antennas, adaptive wireless systems, as well as wideband, broadband, and ultra-wideband code-division multiple access (CDMA) for advanced wireless mobile communication systems. He has published over 90 papers in various journals and conference proceedings. He is a Senior Member of the IEEE.



Lajos Hanzo, a Fellow of the Royal Academy of Engineering (FREng), received his Master's degree in electronics in 1976 and his Doctorate in 1983. In 2004, he was awarded the Doctor of Sciences (D.S.) degree from the University of Southampton, UK. During his 28-year career in telecommunications he has held various research and academic posts in Hungary, Germany, and the UK. Since 1986 he has been with



the Department of Electronics and Computer Science, University of Southampton, UK, where he holds the Chair in telecommunications. He has coauthored 11 John Wiley/IEEE Press books totalling about 8000 pages on mobile radio communications, published in excess of 500 research papers, organised and chaired

Opportunistic Carrier Sensing for Energy-Efficient Information Retrieval in Sensor Networks

Qing Zhao

*Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, USA
Email: qzhao@ece.ucdavis.edu*

Lang Tong

*School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA
Email: ltong@ece.cornell.edu*

Received 26 January 2005

We consider distributed information retrieval for sensor networks with cluster heads or mobile access points. The performance metric used in the design is energy efficiency defined as the ratio of the average number of bits reliably retrieved by the access point to the total amount of energy consumed. A distributed opportunistic transmission protocol is proposed using a combination of carrier sensing and backoff strategy that incorporates channel state information (CSI) of individual sensors. By selecting a set of sensors with the best channel states to transmit, the proposed protocol achieves the upper bound on energy efficiency when the signal propagation delay is negligible. For networks with substantial propagation delays, a backoff function optimized for energy efficiency is proposed. The design of this backoff function utilizes properties of extreme statistics and is shown to have mild performance loss in practical scenarios. We also demonstrate that opportunistic strategies that use CSI may not be optimal when channel acquisition at individual sensors consumes substantial energy. We show further that there is an optimal sensor density for which the opportunistic information retrieval is the most energy efficient. This observation leads to the design of the optimal sensor duty cycle.

Keywords and phrases: sensor networks, distributed information retrieval, opportunistic transmission, energy efficiency.

1. INTRODUCTION

A key component in the design of sensor networks is the process by which information is retrieved from sensors. In an ad hoc sensor network with cluster heads/gateway nodes, sensors send their packets to their cluster heads using a certain transmission protocol [1, 2, 3]. For sensor networks with mobile access [4, 5], data are collected directly by the mobile access points (see Figure 1). In both cases, a population of sensors (those in the same coverage area of an access point) must share a common wireless channel. Thus, an information retrieval protocol that determines which sensors should transmit and the rates of transmissions needs to be designed for efficient channel utilization.

Distributed information retrieval allows each sensor, by itself, to determine whether it should transmit and the rate of transmission. One such example is ALOHA in which each sensor flips a coin (possibly biased by its channel state) to

determine whether it should transmit [6, 7]. Another example is a fixed TDMA schedule by which each sensor transmits in a predetermined time slot. A centralized protocol, in contrast, requires the scheduling by the access point. A particularly relevant technique is the so-called opportunistic scheduling [8, 9] by which the access point determines which sensor should transmit according to the channel states of the sensors. In this paper, we are interested in distributed information retrieval which, in the context of sensor networks, has many advantages: less overhead, more robust against node failures, and possibly more energy efficient.

1.1. Energy-efficient opportunistic transmission

By opportunistic transmission we mean that the information retrieval protocol utilizes the channel state information (CSI). Specifically, suppose that the channel states of a set of activated sensors are obtained. An opportunistic transmission protocol chooses, according to some criterion, a subset of activated sensors to transmit and determines their transmission rates. Knopp and Humblet [8] showed that, to maximize the sum capacity under the average power constraint, the opportunistic transmission that allows a single user with

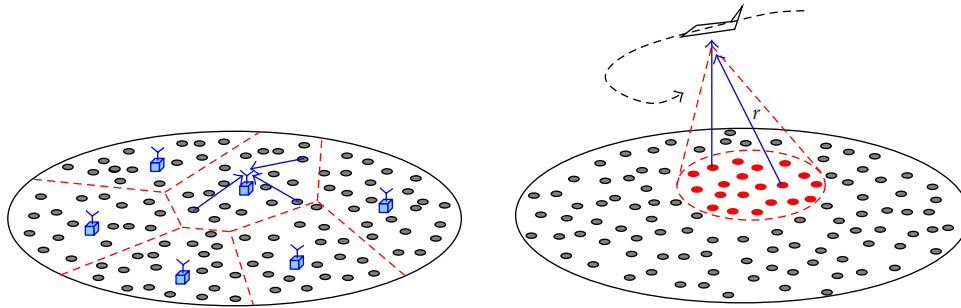


FIGURE 1: Information retrieval in sensor networks.

the best channel to transmit is optimal. Other opportunistic schemes include [6, 7, 9, 10, 11, 12, 13] and the references therein.

The idea of opportunistic information retrieval, at the first glance, is appealing for sensor networks where energy consumption is of primary concern. If the channel realization of a sensor is favorable, the sensor can transmit at a lower power level for the same rate or at a higher rate using the same power. If the sensor has a poor channel, on the other hand, it is better that the sensor saves the energy by not transmitting (and not creating interference to others). What is missing in this line of argument, however, is the cost of obtaining channel states and the cost of determining opportunistic scheduling. If it takes a considerable amount of energy to estimate the channel at each sensor and if determining the set of sensors with the best channels requires additional communications among sensors, it is no longer obvious that an opportunistic information retrieval is more energy efficient than a strategy—for example, using a predetermined schedule—that does not require the channel state information.

It is necessary at this point to specify the performance metric used in the design of information retrieval protocols. For sensor networks, we use energy efficiency (bits/Joule) defined by the ratio of the expected total number of bits reliably received at the access point and the total energy consumed. Here we will include both the energy radiated at the transmitting antenna and the energy consumed in listening, computation, and channel acquisition (when an opportunistic strategy is used). For sensor networks, it has been widely recognized that energy consumption beyond transmission can be substantial [3, 4, 14].

Using energy efficiency as the metric, we aim to address the following questions. If channel acquisition consumes energy, is opportunistic transmission strategy optimal? What would be an energy-efficient *distributed* opportunistic information retrieval? What network parameters affect the energy efficiency? Can these parameters be designed optimally?

While it is debatable whether the information theoretic metric of energy efficiency is appropriate for sensor networks, our goal is to gain insights into the above fundamental questions. It should also be emphasized that the distributed opportunistic protocol developed in this paper applies also

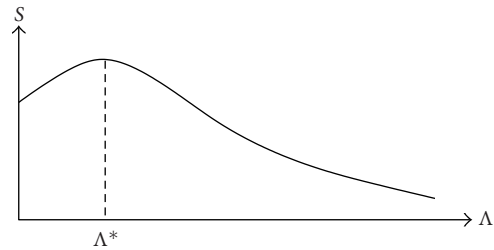


FIGURE 2: Energy-efficiency characteristics.

to noninformation theoretic metrics such as throughput and throughput per unit cost.

1.2. Summary of results

The contribution of this paper is twofold. First, we demonstrate that when the cost of channel acquisition is small as compared to the energy consumed in transmission, the opportunistic transmission is optimal. However, when the average number of activated sensors exceeds a certain threshold, the opportunistic strategy loses its optimality; its energy efficiency approaches zero as the average number of activated sensors approaches infinity. Figure 2 illustrates the generic characteristics of the energy efficiency of the opportunistic transmission where Λ denotes the average number of activated sensors. When Λ is small, the gain in sum capacity due to the use of the best channel dominates the increase in energy consumption. As Λ increases beyond a certain value, the energy cost for acquiring the channel state of every activated sensor overrides the improvement in sum capacity. It is thus critical that the average number Λ of activated sensors be optimized. In Section 5, we study possible schemes of controlling Λ by the design of the sensor duty cycle.

Second, we propose opportunistic carrier sensing—a distributed protocol that achieves a performance upper bound assumed by the centralized opportunistic transmission. The key idea is to incorporate local CSI into the backoff strategy of carrier sensing. Specifically, a decreasing function is used to map the channel state to the backoff time. Each sensor, after measuring its channel, generates the backoff time according to this backoff function. When the propagation delay is negligible, the decreasing property of the backoff function ensures that the sensor with the best channel state

seizes the channel. To minimize the performance loss caused by propagation delay, the backoff function is constructed to balance the energy consumed in carrier sensing and the energy wasted in collision. This protocol also provides a distributed solution to the general problem of finding the maximum/minimum.

1.3. Related work

The metric of energy efficiency considered in this paper can be traced back to capacity per unit cost [15, 16]. For sensor networks, such a metric captures important design trade-offs. However, the literature on using this metric for sensor networks is scarce. Our results explicitly include energy consumed in channel acquisition and listening.

The idea of using CSI was sparked by the work of Knopp and Humblet [8]. Exploiting CSI induces multiuser diversity as the performance increases with the number of users [9, 10]. Throughput optimal scheduling for downlink over time-varying channels by a central controller has been considered in [17, 18], all assuming the knowledge of the channel states at no cost. Decentralized power allocation based on channel states was investigated by Telatar and Shamai under the metric of sum capacity [12]. Viswanath et al. [19] have shown the asymptotic optimality of a decentralized power control scheme for a multiaccess fading channel that uses CDMA with an optimal receiver. The effect of decentralized power control on the sum capacity of CDMA with linear receivers and single-user decoders was studied by Shamai and Verdú in [20]. All the work along this line uses rate, not the energy efficiency, as the performance metric. Using channel state information in random access has been considered in [6, 7, 21]. Qin and Berry, in particular, aimed to schedule the sensor with the best channel to transmit by a distributed protocol—channel-aware ALOHA [7]. The throughput of channel-aware ALOHA, however, is limited by the efficiency of the conventional ALOHA protocol.

1.4. Organization of the paper

In Section 2, we state the network model. The performance of the opportunistic transmission is addressed in Section 3 where we obtain a performance upper bound and characterize the optimal number of transmitting sensors in the opportunistic transmission. In Section 4, we propose opportunistic carrier sensing. A backoff function is constructed and its robustness to propagation delay is demonstrated. In Section 5, we focus on the optimality of the opportunistic transmission. Optimal sensor activation schemes are discussed. Section 6 concludes the paper.

2. THE NETWORK MODEL

2.1. The sensor network

We assume that the sensor nodes form a two-dimensional Poisson field¹ with mean λ . The number M of active sensors

that share the wireless channel to an access point is thus a Poisson random variable with mean $\Lambda = a\lambda$ where a denotes the coverage area of the mobile access point or the size of the cluster, that is,

$$P[M = m] = \frac{e^{-\Lambda} \Lambda^m}{m!}. \quad (1)$$

For a sensor network with mobile access, we consider a single access point. For a sensor network under the structure of clusters, we focus on the information retrieval within one cluster. We assume that there is no interference among adjacent clusters (which can be achieved by, for example, assigning different frequencies to adjacent clusters) and the sensors within the cluster transmit directly to the cluster head as considered in [3]. Thus, information retrieval for a sensor network with mobile access or cluster heads can be modeled as a many-to-one communication problem. Aiming at providing insights to fundamental questions on opportunistic transmission, we further assume that sensors within the coverage area of the mobile access point or the same cluster can hear each other's transmission.

2.2. The wireless fading channel

The physical channel between an active sensor and the access point is subject to flat Rayleigh fading with a block length of T seconds, which is also the length of transmission slot. The channel is thus constant within each slot and varies independently from slot to slot.

Consider the first slot where n nodes transmit simultaneously. The received signal $y(t)$ at the access point can be written as

$$y(t) = \sum_{i=1}^n h_i x_i(t) + n(t), \quad 0 \leq t \leq T, \quad (2)$$

where h_i is the channel fading process experienced by sensor i , $n(t)$ the white Gaussian noise with power spectrum density $N_0/2$, and $x_i(t)$ the transmitted signal with fixed power P_{out} . We point out that the power constraint used here is different from the long-term average power constraint considered in [8]. We assume that sensors can only transmit at a fixed power level P_{out} and do not have the capability of allocating power over time. Define

$$\rho \triangleq \frac{P_{\text{out}}}{WN_0}. \quad (3)$$

Let

$$\gamma_i \triangleq |h_i|^2 \sim \exp(\bar{\gamma}_i) \quad (4)$$

denote the channel gain from sensor i to the access point. Under independent Rayleigh fading, γ_i is exponentially distributed with mean $\bar{\gamma}_i$. The average received SNR of sensor i is thus given by $\rho\bar{\gamma}_i$.

¹As shown in [22], the difference (in terms of network connectivity) between a Poisson field and a uniformly distributed random field is negligible when the number of nodes is large. For the simplicity of the analysis, we assume a Poisson distributed sensor network.

2.3. The energy consumption model

In each slot, energy consumed by active sensors may come from three operations: transmission, reception, and scheduling.

Let E_r and E_t denote, respectively, total energy consumed in receiving and transmitting in one slot. We have [14]

$$E_r = \mathbb{E} \left[P_{\text{rx}} \sum_{i=1}^M T_{\text{rx}}(i) \right], \quad (5)$$

$$E_t = \mathbb{E} \left[P_{\text{tx}} \sum_{i=1}^M T_{\text{tx}}(i) \right], \quad (6)$$

where the expectation is with respect to M , $T_{\text{rx}}(i)$, and $T_{\text{tx}}(i)$ are the average reception and transmission time of node i , P_{rx} is the sensor's receiver circuitry power, P_{tx} is the power consumed in transmission which consists of transmitter circuitry power and antenna output power P_{out} .

In the distributed opportunistic transmission, active sensors perform synchronization and channel acquisition using a beacon signal broadcast by the access point² and determine who should transmit and at what rate. The expected total cost E_c of scheduling transmissions based on the channel states of the active sensors is lower bounded by

$$E_c \geq \Lambda e_c, \quad (7)$$

where e_c is the amount of energy consumed by one sensor in estimating its channel state from the beacon signal. This lower bound holds for both centralized and distributed implementations of the opportunistic transmission. It is achieved when the active sensors, each with access only to its own channel state, can determine the set of transmitting sensors at no cost. We show in Section 4 that when the propagation delay among active sensors is negligible, the scheduling cost of the proposed opportunistic protocol achieves the lower bound given in (7).

3. OPPORTUNISTIC TRANSMISSION FOR ENERGY EFFICIENCY

In this section, we address the performance of the opportunistic transmission under the metric of energy efficiency. As a performance measure, energy efficiency is first defined and the underlying coding scheme specified. We then obtain an upper bound on the performance of the opportunistic transmission and characterize the optimal number of transmitting sensors.

3.1. Sum capacity and coding scheme

Given that the channel fading process h_i is independent among sensors, and strictly stationary and ergodic, the sum

capacity achieved by an information retrieval protocol which enables n sensors in each slot is given by [23]

$$R = W \mathbb{E} \left[\log \left(1 + \rho \sum_{i=1}^n \gamma_i \right) \right], \quad (8)$$

where W is the transmission bandwidth and the expectation is over the fading process γ_i (see (4)). To achieve this rate, the CSI is used in decoding. The information rate is constant over time and each codeword sees a large number of channel realizations.

An alternative coding scheme is to use different transmission rates according to the channel states of the transmitting sensors. In this case, each codeword experiences only one channel realization, resulting in a smaller coding delay. When the block length T is sufficiently large, the achievable sum rate averaged over time can be approximated by (8). Note that using a variable information rate in each slot requires the CSI in both encoding and decoding. If more than one sensor is enabled for transmission, each transmitting sensor must know not only its own channel state, but also the channel states of other simultaneously transmitting sensors in order to determine the rate of transmission. In Section 4, we show that with the proposed opportunistic carrier sensing, each transmitting sensor obtains the channel states of other sensors at no extra cost. The proposed protocol is thus applicable to both coding schemes. Without loss of generality, we assume, for the rest of the paper, this alternative coding scheme which uses variable information rate. We point out that under this coding scheme, (8) is only an approximation to the achievable sum rate. A more rigorous formulation is to use error exponents [15].

3.2. n -TDMA

As a benchmark, we first give an expression of energy efficiency for a predetermined scheduling where n sensors are scheduled for transmission in each slot. At the beginning of each slot, n sensors wake up, measure their channel states, and transmit. Referred to as n -TDMA, this scheme with optimal n has the energy efficiency

$$S_{\text{TDMA}} = \max_n \frac{WT \mathbb{E} [\log (1 + \rho \sum_{i=1}^n \gamma_i)]}{n e_c + n T P_{\text{tx}}}, \quad (9)$$

where expectation³ is over M and $\{\gamma_i\}_{i=1}^n$. Since $n \ll \Lambda$ in general, we have ignored the rare event of $M < n$. The above optimization can be obtained numerically.

3.3. Opportunistic transmission

3.3.1. A performance upper bound

With the opportunistic strategy, n sensors with the best channels are enabled for transmission in each slot. Let $\gamma_M^{(i)}$ denote

²We assume reciprocity. The channel gain from a sensor to the access point is the same as that from the access point to the sensor.

³To be precise, the numerator of (9) should be written as $WT \mathbb{E}_M \{ \mathbb{E}_{\gamma^{(i)}} [\log (1 + \rho \sum_{i=1}^{\min\{n, m\}} \gamma_m^{(i)}) | M = m] \}$.

the i th best channel gain among M sensors. The energy efficiency of the opportunistic strategy with optimal n is

$$S_{\text{opt}} = \max_n \frac{WT\mathbb{E}\left[\log\left(1 + \rho \sum_{i=1}^n \gamma_M^{(i)}\right)\right]}{E_c + nTP_{\text{tx}}}, \quad (10)$$

where expectation is over M and $\{\gamma_M^{(i)}\}_{i=1}^n$. Using the lower bound on E_c given in (7), we obtain a performance upper bound for the opportunistic strategy:

$$S_{\text{opt}} \leq \max_n \frac{WT\mathbb{E}\left[\log\left(1 + \rho \sum_{i=1}^n \gamma_M^{(i)}\right)\right]}{\Lambda e_c + nTP_{\text{tx}}}. \quad (11)$$

3.3.2. The optimal number of transmitting sensors

Since the performance upper bound given in (11) is achieved by the opportunistic carrier sensing proposed in Section 4, we can use this upper bound to study the optimal number n^* of transmitting sensors and the optimality of the opportunistic transmission.

It has been shown by Knopp and Humblet [8] that the optimal transmission scheme for maximizing sum capacity under a long-term average power constraint is to enable only one sensor (the one with the best channel) to transmit. Under the metric of energy efficiency with a fixed transmission power, however, allowing more than one transmission may be optimal when the cost in channel acquisition becomes substantial.

Proposition 1. *For a fixed slot length T , transmission power P_{tx} , and the channel acquisition cost e_c , the optimal number n^* of transmitting sensors for the opportunistic transmission is given by*

$$n^* = 1 \quad \text{if } \Lambda < \frac{TP_{\text{tx}}(2C_1 - C_2)}{e_c(C_2 - C_1)}, \quad (12)$$

$$n^* > 1 \quad \text{otherwise,}$$

where $C_n = WT\mathbb{E}[\log(1 + \rho \sum_{i=1}^n \gamma_M^{(i)})]$.

For the proof of Proposition 1, see Appendix A.

In Figure 3, we plot the energy efficiency of the opportunistic transmission for different numbers n of transmitting sensors. In Figure 3a, the average number Λ of active sensors is 500 while, in Figure 3b, it is set to 5 000. We can see that n^* increases from 1 to 2 when Λ increases. The intuition behind this is that the cost in channel acquisition dominates when $\Lambda = 5000$; allowing one more transmission improves the sum rate without inducing significant increase in energy consumption. The performance of n -TDMA is also plotted in Figure 3 for comparison. For this simulation setup, the optimal number of transmitting sensors for n -TDMA equals 1. We observe that the opportunistic transmission is inferior to the simple predetermined scheduling at $\Lambda = 5000$. Indeed, we show in Section 5 that the opportunistic transmission strategy loses its optimality when Λ exceeds a threshold.

4. OPPORTUNISTIC CARRIER SENSING

In this section, we propose opportunistic carrier sensing, a distributed protocol whose performance approaches to the upper bound of the opportunistic strategy given in (11). We first present the basic idea of the opportunistic carrier sensing under the assumption of negligible propagation delay among active sensors. In Section 4.2, we study the design of the backoff function to minimize the performance loss caused by propagation delay.

4.1. The basic idea

We now present the basic idea of the opportunistic carrier sensing by considering an idealistic scenario. We assume that the transmission of one sensor is immediately detected by other active sensors. In the next subsection, we discuss how to circumvent the propagation delay among active sensors.

The key idea of opportunistic carrier sensing is to exploit CSI in the backoff strategy of carrier sensing. First consider $n^* = 1$, that is, in each slot, only the sensor with the best channel transmits. After each active sensor measures its channel gain γ_i using the beacon of the access point, it chooses a backoff τ based on a predetermined function $f(\gamma)$ which maps the channel state to a backoff time and then listens to the channel. A sensor will transmit with its chosen backoff delay if and only if no one transmits before its backoff time expires. If $f(\gamma)$ is chosen to be a strictly decreasing function of γ as shown in Figure 4, this opportunistic carrier sensing will ensure that only the sensor with the best channel transmits. Under the idealistic scenario where the transmission of one sensor is immediately detected by other active sensors, $f(\gamma)$ can be any decreasing function with range $[0, \tau_{\text{max}}]$, where τ_{max} is the maximum backoff. Since τ_{max} can be chosen as any positive number, the time required for each sensor listening to the channel can be arbitrarily short. Hence, energy consumed in each slot comes only from each sensor estimating its own channel state (the lower bound on E_c given in (7)) and the transmission by one sensor; opportunistic carrier sensing thus achieves the performance upper bound of the opportunistic strategy.

We now consider $n^* > 1$. If the energy detector of each sensor is sensitive enough to distinguish the number of simultaneous transmissions, the opportunistic carrier sensing protocol stated above can be directly applied—a sensor transmits with its chosen backoff if and only if the number of transmissions at that time instant is smaller than n^* . Note that by observing the time instant τ at which the number of simultaneous transmissions increases (energy-level jumps) and mapping this time instant back to the channel gain using $\gamma = f^{-1}(\tau)$, a sensor obtains the channel states of other transmitting sensors and can thus determine its transmission rate. Note that the channel gain of a transmitting sensor is learned by measuring the backoff of the transmission, not the signal strength.

If, however, sensors can not obtain the number of simultaneous transmissions, we generalize the protocol as follows. We partition each slot into two segments: carrier sensing and information transmission (see Figure 5). During the carrier

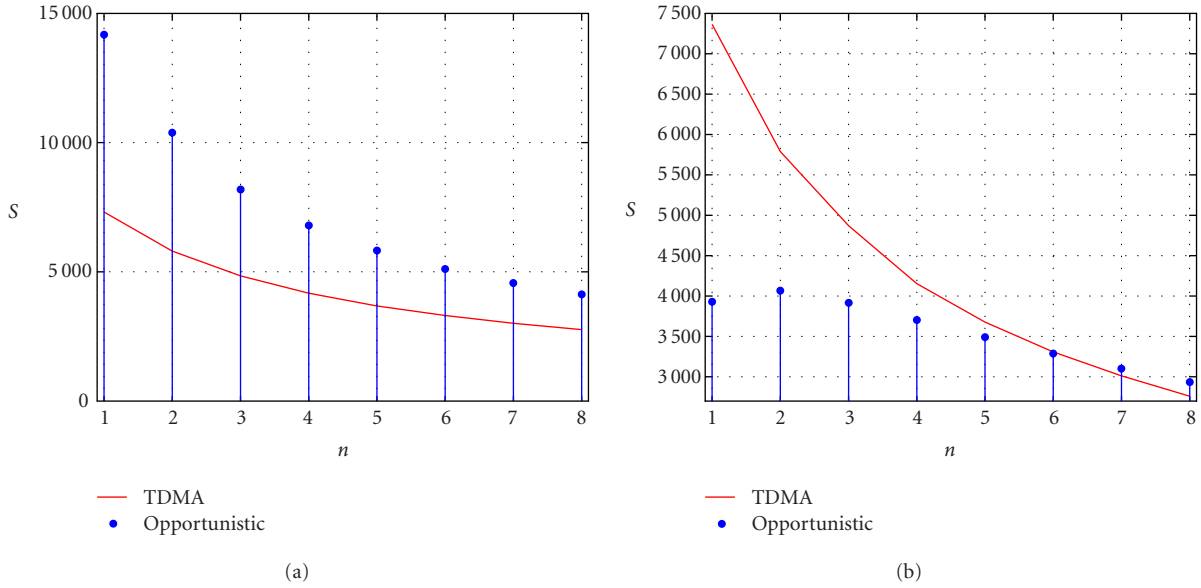


FIGURE 3: The optimal number n^* of transmitting sensors ($W = 1$ kHz, $\rho\bar{\gamma}_i = 3$ dB, $T = 0.01$ second, $P_{tx} = 0.181$ W, $e_c = 1.8$ nJ): (a) $\Lambda = 500$ and (b) $\Lambda = 5000$.

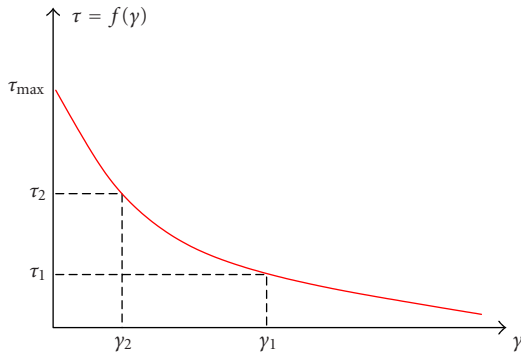


FIGURE 4: Opportunistic carrier sensing.

sensing period, sensors transmit, with backoff delay determined by $f(\gamma)$, a beacon signal with short duration. A sensor transmits a beacon if and only if the number of received beacon signals is smaller than n^* . By measuring the time instant at which each beacon signal is transmitted, those n^* sensors with the best channels can also obtain all n^* channel states from $f^{-1}(\tau)$ and thus encode their messages accordingly. Shown in Figure 5 is an example with $n^* = 2$. During the carrier sensing segment $[0, \tau_{max}]$, two beacon signals are transmitted at τ_1 and τ_2 by two sensors with the best channel gains. Based on τ_1 , τ_2 , and $f^{-1}(\tau)$, these two sensors obtain each other's channel state (see Figure 4). They then encode their messages for transmissions in the second segment of the slot. One possible encoding scheme, as shown in Figure 5, is based on the idea of successive decoding. The sensor with the higher channel gain γ_1 encodes its message at rate $W \log(1 + \rho\gamma_1)$ as if it was the only transmitting node.

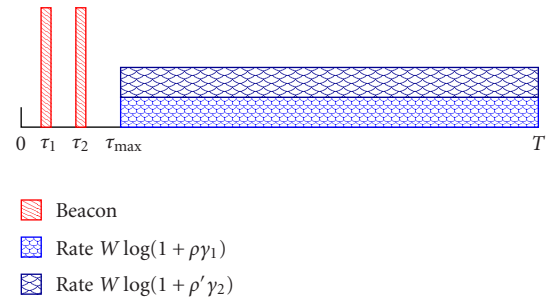


FIGURE 5: Opportunistic carrier sensing for $n^* = 2$.

The other sensor with channel gain γ_2 encodes its message by treating the transmission from the sensor with channel γ_1 as noise. It transmits at rate $W \log(1 + \rho'\gamma_1)$ where

$$\rho' = \frac{P_{out}}{N_0 W + P_{out}\gamma_1}. \tag{13}$$

We point out that the idea of opportunistic carrier sensing provides a distributed solution to the general problem of finding maximum/minimum. By substituting the channel gain γ with, for example, the temperature measured by each sensor, the distance of each sensor to a particular location, or the residual energy of each sensor, we can retrieve information of interest (the highest/lowest temperature, the measurement closest/farthest to a location) from sensors of interest (those with the highest energy level or those with the best channel gain) in a distributed and energy-efficient fashion.

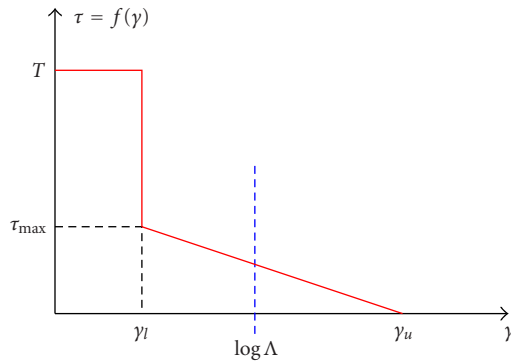


FIGURE 6: Backoff function under significant propagation delay.

4.2. Backoff design under significant delay

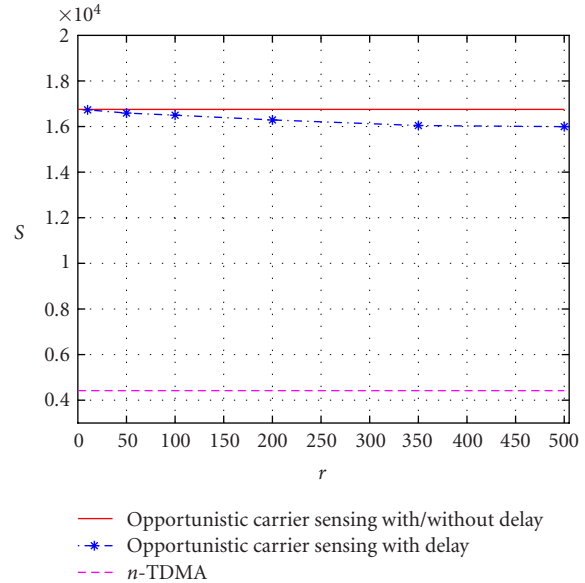
We now generalize the basic idea of opportunistic carrier sensing to scenarios with significant delay which may include both the propagation delay and the time spent in the detection of transmissions. Without loss of generality, we focus on the case of $n^* = 1$.

In the idealistic case considered in the previous subsection, energy consumed in carrier sensing is negligible due to the arbitrarily small carrier sensing time τ_{\max} . Furthermore, using any decreasing function as the backoff function $f(\gamma)$ avoids collision, an event where several nodes transmit simultaneously while no information is received at the access point. When there is substantial delay, however, collision and energy consumed by carrier sensing⁴ are inevitable. To maintain the optimal performance achieved under the idealistic scenario, $f(\gamma)$ needs to be designed judiciously to minimize both the occurrence of collision and the energy consumed in carrier sensing. Unfortunately, these are two conflicting objectives. On one hand, choosing a larger τ_{\max} makes it more likely to map channel gains to well-separated backoff times, thus reducing collisions. On the other hand, a larger τ_{\max} results in less transmission time and more energy consumption of carrier sensing.

To balance the tradeoff between collision and energy consumption of carrier sensing, we propose $f(\gamma)$ as illustrated in Figure 6. This backoff scheme is a linear function on a finite interval $[\gamma_l, \gamma_u)$ where the channel gain is mapped to a backoff time in $(0, \tau_{\max}]$. Sensors with channel gains greater than γ_u transmit without backoff ($\tau = 0$) while sensors with channel gains smaller than γ_l turn off their radios until next slot ($\tau = T$), without even participating in the carrier sensing process.

The proposed backoff function is completely determined by γ_l , γ_u , and τ_{\max} . The choice of a finite γ_u allows better resolution among highly likely channel realizations. The option of a nonzero γ_l avoids the listening cost of sensors whose channels are unlikely to be the best. For a relatively large Λ , a large percentage of active sensors can be freed of carrier

⁴Listening to the channel requires the receiver being turned on, which consumes energy as given in (5).

FIGURE 7: Performance of opportunistic carrier sensing under significant delay ($\Lambda = 100$, $W = 1$ kHz, $\rho\bar{\gamma}_i = 3$ dB, $T = 0.01$ second, $P_{\text{tx}} = 0.181$ W, $P_{\text{rx}} = 0.18$ W, $e_c = 1.8$ nJ).

sensing cost with a carefully chosen γ_l . The maximum backoff time τ_{\max} is chosen to balance collision and energy consumption of carrier sensing. It is jointly optimized with γ_l and γ_u to maximize energy efficiency:

$$\{\gamma_l^*, \gamma_u^*, \tau_{\max}^*\} = \arg \max S(\gamma_l, \gamma_u, \tau_{\max}). \quad (14)$$

The optimal $\{\gamma_l^*, \gamma_u^*, \tau_{\max}^*\}$ can be obtained via numerical evaluation or simulations. To narrow the search range of γ_l and γ_u , asymptotic extreme-order statistics given in Lemma 1 (see Section 5.1) can be exploited. For a relatively large Λ , the best channel gain $\gamma^{(1)}$ is on the order of $\log \Lambda$.

We now consider a simulation example to evaluate the performance of opportunistic carrier sensing with the backoff function $f(\gamma)$ given in Figure 6 using numerically optimized parameters $\{\gamma_l^*, \gamma_u^*, \tau_{\max}^*\}$. We focus on information retrieval by a mobile access point and model the coverage area of the mobile access point as a disk with radius r (see Figure 1). The maximum propagation delay β is then given by

$$\beta = \frac{2r}{v_l}, \quad (15)$$

where v_l is the speed of light.⁵ Shown in Figure 7 is the energy efficiency of opportunistic carrier sensing as a function of the radius r of the coverage area which determines the maximum propagation delay. Compared with the performance in the ideal scenario (no propagation delay), the performance of opportunistic carrier sensing degrades gracefully with

⁵We have ignored the delay in the detection of transmission at sensor nodes. It can be easily accommodated by adding a constant to the propagation delay.

propagation delay. Even with a coverage radius of 500 meters, the performance degradation due to propagation delay is less than 5%.

5. OPTIMAL SENSOR ACTIVATION

In this section, we demonstrate that the energy efficiency of the opportunistic transmission vanishes as the number Λ of active sensors approaches infinity. Possible schemes for optimizing the number of active sensors are discussed.

5.1. Tradeoff between sum capacity and energy consumption

Since the extreme value of i.i.d. samples increases with the sample size, it is easy to show that the sum capacity achieved by n sensors with the best channels increases with Λ . Unfortunately, larger Λ also leads to higher energy consumption in channel acquisition (see (7)). Proposition 2 shows that the gain in sum capacity does not always justify the cost in obtaining the channel states.

Proposition 2. For a fixed slot length T , transmission power P_{tx} , and the channel acquisition cost $e_c > 0$,

$$\lim_{\Lambda \rightarrow \infty} S_{\text{opt}} = 0. \quad (16)$$

A direct consequence of Proposition 2 is that, as summarized in Corollary 1, the opportunistic strategy loses its optimality when Λ exceeds a threshold.

Corollary 1. There exists $\Lambda_0 < \infty$ such that $S_{\text{opt}} < S_{\text{TDMA}}$ when $\Lambda > \Lambda_0$.

The proof (see Appendix B) of Proposition 2 is based on the following result on asymptotic extreme-order statistics [24].

Lemma 1. Let X_1, X_2, \dots be i.i.d. random variables with continuous distribution function $F(x)$. Let x_0 denote the upper boundary, possibly $+\infty$, of the distribution: $x_0 \triangleq \sup\{x : F(x) < 1\}$. If there exists a function $R(t)$ such that for all x ,

$$\lim_{t \rightarrow x_0} \frac{1 - F(t + xR(t))}{1 - F(t)} = e^{-x}, \quad (17)$$

then

$$\frac{X_m^{(1)} - a_m}{b_m} \xrightarrow{d} \exp\{-e^{-x}\}, \quad (18)$$

where $X_m^{(1)} = \max_{i \leq m} X_i$, $1 - F(a_m) = 1/m$, $b_m = R(a_m)$, and \xrightarrow{d} denotes convergence in distribution.

Common fading distributions such as Rayleigh and Ricean satisfy the assumptions of Lemma 1. For Rayleigh fading considered in this paper, we have $a_m = \log m$ and $b_m = 1$, that is,

$$X_m^{(1)} - \log m \xrightarrow{d} \exp\{-e^{-x}\}. \quad (19)$$

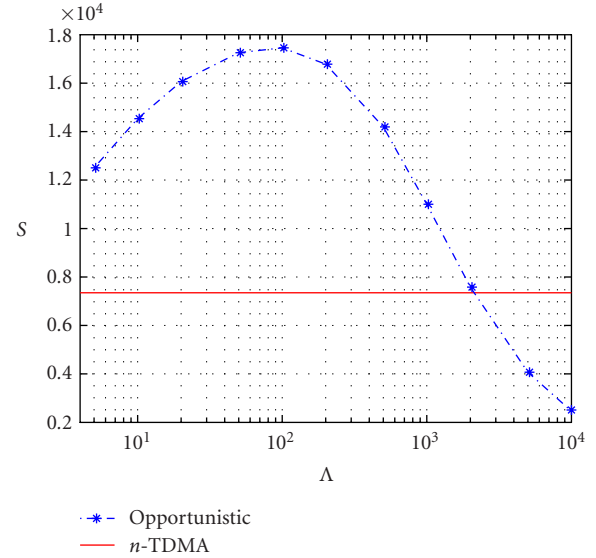


FIGURE 8: Tradeoff between sum capacity and energy consumption ($W = 1$ kHz, $\rho\bar{y}_i = 3$ dB, $T = 0.01$ second, $P_{\text{tx}} = 0.181$ W, $e_c = 1.8$ nJ).

Shown in Figure 8 are simulation results on the energy efficiency of the opportunistic transmission as compared to the predetermined scheduling. Since both the sum rate and the energy consumption of n -TDMA are independent of Λ , the energy efficiency is constant over Λ . For the opportunistic strategy, the energy efficiency increases with Λ when Λ is relatively small. In this region, the energy consumption is dominated by transmission; the increase in the cost of channel acquisition does not significantly affect the total energy expenditure. The energy efficiency thus improves as the sum capacity increases with Λ . When Λ increases beyond 100 where the cost in channel acquisition contributes more than 10% of the total energy expenditure, the increase in energy consumption overrides the improvement in sum rate; the energy efficiency starts to decrease. Eventually, the gain in sum capacity achieved by exploiting CSI can no longer justify the cost in obtaining CSI, and the opportunistic strategy is inferior to the predetermined scheduling.

5.2. The optimal number of active sensors

As shown in Figure 8, the performance of the opportunistic transmission depends on the average number of active sensors's. To achieve the best performance of the opportunistic strategy, the average number Λ of active sensors should be carefully chosen.

The average number of active sensors can be controlled via the sensor duty cycle or the size of the coverage area of the mobile access point (or the cluster). Assume that each sensor with probability p wakes up independently to detect the beacon signal of the access point. For a coverage area of size a , the average number of active sensors is given by $\Lambda = ap\lambda$, where λ is the node density defined in Section 2. The average number of active sensors can thus be controlled by varying either a or the duty cycle p .

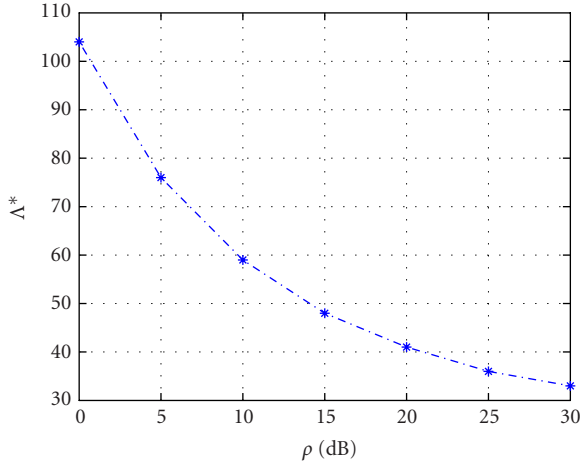


FIGURE 9: The optimal number of active sensors ($W = 1$ kHz, $T = 0.01$ second, $P_{\text{tx}} = 0.181$ W, $e_c = 1.8$ nJ).

In Figure 9, we plot the optimal average number Λ^* of the active sensors as a function of the average SNR. Without loss of generality, we normalize $\bar{\gamma}_i$ to 1. The average received SNR is thus given by ρ . We observe that Λ^* is a decreasing function of ρ . The reason for this is that the larger the average SNR, the smaller the impact of $\gamma^{(1)}$ on the sum rate (see (10)). Thus, the threshold beyond which the channel acquisition cost overrides the gain in sum rate decreases with ρ , resulting in decreasing Λ^* .

6. CONCLUSION

In this paper, we focus on distributed information retrieval in wireless sensor networks. Energy efficiency is introduced as the performance metric. Measured in bits per Joule, this metric captures a major design constraint—energy—of sensor networks.

We examine the performance of the opportunistic transmission which exploits CSI for transmission scheduling. Taking into account energy consumed in channel acquisition, we demonstrate that sum-rate improvement achieved by opportunistic transmission does not always justify the cost in channel acquisition; there exists a threshold of the average number of activated sensor nodes beyond which the opportunistic strategy loses its optimality. Sensor activation schemes are discussed to optimize the energy efficiency of the opportunistic transmission.

We propose a distributed opportunistic transmission protocol that achieves the performance upper bound assumed by the centralized opportunistic scheduler. Referred to as opportunistic carrier sensing, the proposed protocol incorporates CSI into the backoff strategy of carrier sensing. A backoff function which maps channel state to backoff time is constructed for scenarios with substantial propagation delay. The performance of opportunistic carrier sensing with the proposed backoff function degrades gracefully with propagation delay. The proposed protocol also provides a distributed solution to the general problem of finding the maximum/minimum.

A number of issues are not addressed in this paper. We have used the information theoretic metric of energy efficiency that implicitly assumes that data from different sensors are independent. For applications in which data are highly correlated, distributed compression techniques may be necessary [25]. Fairness in transmission is another issue that needs to be considered in practice. For sensor networks with mobile access points or networks with randomly rotated cluster heads, the probability of transmission can be made uniform. For networks with fixed cluster heads, sensors closer to the cluster head tend to have stronger channel, thus transmit more often. This, however, can be easily equalized by using the normalized channel gain in the backoff strategy.

APPENDICES

A. PROOF OF PROPOSITION 1

Let S_n denote the energy efficiency of the opportunistic strategy which enables n sensors with the best channels in each slot. We have

$$S_n = \frac{C_n}{\Lambda e_c + nTP_{\text{tx}}}. \quad (\text{A.1})$$

To prove Proposition 1, we need to show that for $\Lambda < TP_{\text{tx}}(2C_1 - C_2)/e_c(C_2 - C_1)$, $S_1 \geq S_n$ for all n . Since

$$\begin{aligned} \frac{C_1}{\Lambda e_c + TP_{\text{tx}}} &\geq \frac{C_n}{\Lambda e_c + nTP_{\text{tx}}} \\ &\Rightarrow \Lambda e_c(C_n - C_1) \\ &\leq TP_{\text{tx}}(nC_1 - C_n), \end{aligned} \quad (\text{A.2})$$

we only need to show that there exists $\Lambda > 0$ that satisfies (A.2). This reduces to the positiveness of $nC_1 - C_n$ which follows directly from the concavity of the logarithm function.

B. PROOF OF PROPOSITION 2

Let $S_{\text{opt}}(m)$ denote the energy efficiency of the opportunistic transmission where exactly m sensors are active in each slot. We first show, based on Lemma 1, that $\lim_{m \rightarrow \infty} S_{\text{opt}}(m) = 0$:

$$\lim_{m \rightarrow \infty} S_{\text{opt}}(m) = \lim_{m \rightarrow \infty} \max_{1 \leq n \leq m} \frac{\mathbb{E} \left[WT \log \left(1 + \rho \sum_{i=1}^n \gamma_m^{(i)} \right) \right]}{me_c + nTP_{\text{tx}}} \quad (\text{B.1})$$

$$\leq \lim_{m \rightarrow \infty} \frac{\mathbb{E} \left[WT \log \left(1 + m\rho\gamma_m^{(1)} \right) \right]}{me_c} \quad (\text{B.2})$$

$$\leq \lim_{m \rightarrow \infty} \frac{WT \log \left(1 + m\rho\mathbb{E} \left[\gamma_m^{(1)} \right] \right)}{me_c} \quad (\text{B.3})$$

$$\leq \lim_{m \rightarrow \infty} \frac{WT \log \left(1 + m^2\rho \right)}{me_c} \quad (\text{B.4})$$

$$= 0, \quad (\text{B.5})$$

where $\gamma_m^{(i)}$ denotes the i th-order statistics over m samples; the expectations in (B.1) and (B.2) are with respect to $\{\gamma_m^{(i)}\}_{i=1}^n$ and $\gamma_m^{(1)}$, respectively. Jensen's inequality is used to obtain (B.3), and Lemma 1, which shows that $\gamma_m^{(1)} \sim \log(m) < m$, for large m , is used to obtain (B.4). Combining (B.5) and the fact that $S_{\text{opt}}(m) > 0$ for all m , we conclude that $\lim_{m \rightarrow \infty} S_{\text{opt}}(m) = 0$. Thus,

$$\forall \epsilon > 0, \quad \exists M_0 > 0, \quad \text{s.t. } S_{\text{opt}}(m) < \epsilon \quad \forall m > M_0. \quad (\text{B.6})$$

That $S_{\text{opt}}(m)$ vanishes with m also implies that

$$\exists \bar{S} < \infty, \quad \text{s.t. } S_{\text{opt}}(m) < \bar{S} \quad \forall m. \quad (\text{B.7})$$

It is easy to show that for Poisson distributed random variable M ,

$$\lim_{\Lambda \rightarrow \infty} P[M \leq M_0] = \lim_{\Lambda \rightarrow \infty} \frac{\sum_{i=1}^{M_0} (\Lambda)^i / i!}{e^\Lambda} = 0. \quad (\text{B.8})$$

Thus, for ϵ and M_0 given in (B.6), we have

$$\exists M_1 > 0, \quad \text{s.t. } P[M \leq M_0] < \epsilon \quad \forall \Lambda > M_1. \quad (\text{B.9})$$

Combining (B.6), (B.7), and (B.9), we have, for $\Lambda > M_1$,

$$\begin{aligned} S_{\text{opt}} &= \sum_{m=1}^{\infty} P[M = m] S_{\text{opt}}(m) \\ &= \sum_{m=1}^{M_0} P[M = m] S_{\text{opt}}(m) + \sum_{m=M_0+1}^{\infty} P[M = m] S_{\text{opt}}(m) \\ &< \epsilon \bar{S} + \epsilon. \end{aligned} \quad (\text{B.10})$$

We thus obtain Proposition 2 from the arbitrariness of ϵ .

ACKNOWLEDGMENT

This work was supported in part by the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564 and the Army Research Laboratory CTA on Communication and Networks under Grant DAAD19-01-2-0011.

REFERENCES

- [1] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next century challenges: scalable coordination in sensor networks," in *Proc. 5th annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM '99)*, pp. 263–270, Seattle, Wash, USA, August 1999.
- [2] G. Pottie and W. Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, 2000.
- [3] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.
- [4] L. Tong, Q. Zhao, and S. Adireddy, "Sensor networks with mobile agents," in *Proc. IEEE Military Communications Conference (MILCOM '03)*, vol. 1, pp. 688–693, Boston, Mass, USA, October 2003.
- [5] G. Mergen, Q. Zhao, and L. Tong, "Sensor networks with mobile access: energy and capacity considerations," to appear in *IEEE Trans. Commun.*
- [6] P. Venkatasubramanian, S. Adireddy, and L. Tong, "Opportunistic ALOHA and cross layer design for sensor networks," in *Proc. IEEE Military Communications Conference (MILCOM '03)*, vol. 1, pp. 705–710, Boston, Mass, USA, October 2003.
- [7] X. Qin and R. Berry, "Exploiting multiuser diversity for medium access control in wireless networks," in *IEEE Twenty-Second Annual Conference of the IEEE Computer and Communications Societies (INFOCOM '03)*, vol. 2, pp. 1084–1094, San Francisco, Calif, USA, March–April 2003.
- [8] R. Knopp and P. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE International Conference on Communications (ICC '95)*, vol. 1, pp. 331–335, Seattle, Wash, USA, June 1995.
- [9] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1277–1294, 2002.
- [10] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2796–2815, 1998.
- [11] X. Liu, E. Chong, and N. Shroff, "Opportunistic transmission scheduling with resource-sharing constraints in wireless networks," *IEEE J. Select. Areas Commun.*, vol. 19, no. 10, pp. 2053–2064, 2001.
- [12] I. E. Telatar and S. Shamai, "Some information theoretic aspects of decentralized power control in multiple access fading channels," in *Proc. Information Theory and Networking Workshop*, p. 23, Piscataway, NJ, USA, 1999.
- [13] S. Adireddy and L. Tong, "Exploiting decentralized channel state information for random access," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 537–561, 2005.
- [14] E. Shih, S.-H. Cho, N. Ickes, et al., "Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks," in *Proc. 7th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM '01)*, pp. 272–287, Rome, Italy, July 2001.
- [15] R. Gallager, "Energy limited channels: coding, multiaccess, and spread spectrum," Tech. Rep. LIDS-P-1714, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Mass, USA, November 1987.
- [16] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inform. Theory*, vol. 36, no. 5, pp. 1019–1030, 1990.
- [17] S. Shakkottai and A. Stolyar, *Scheduling for multiple flows sharing a time-varying channel: The exponential rule*, vol. 207 of *Translations of the American Mathematical Society, A volume in memory of F. Karpelevich*, American Mathematical Society, Providence, RI, USA, 2001.
- [18] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 466–478, 1993.
- [19] P. Viswanath, D. N. C. Tse, and V. Anantharam, "Asymptotically optimal water-filling in vector multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 1, pp. 241–267, 2001.

- [20] S. Shamai and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1302–1327, 2001.
- [21] S. Adireddy and L. Tong, "On the use of channel state information for slotted ALOHA in CDMA networks with linear MMSE multi-user receivers," in *Proc. IEEE International Symposium on Information Theory (ISIT '03)*, Yokohama, Japan, June–July 2003.
- [22] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," in *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming*, pp. 547–566, Birkhauser, Boston, Mass, USA, 1998.
- [23] S. Shamai and A. D. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels. i," *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 1877–1894, 1997.
- [24] T. Ferguson, *A Course in Large Sample Theory*, Chapman & Hall, London, UK, 1996.
- [25] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Mag.*, vol. 19, no. 2, pp. 51–60, 2002.

Qing Zhao received the B.S. degree in 1994 from Sichuan University, Chengdu, China, the M.S. degree in 1997 from Fudan University, Shanghai, China, and the Ph.D. degree in 2001 from Cornell University, Ithaca, NY, all in electrical engineering. From 2001 to 2003, she was a Communication System Engineer with Aware, Inc., Bedford, Mass. She returned to academy in 2003 as a Postdoctoral Research Associate with the School of Electrical and Computer Engineering, Cornell University. In 2004, she joined the Department of Electrical and Computer Engineering, UC Davis, where she is currently an Assistant Professor. Her research interests are in the general area of signal processing, communication systems, wireless networking, and information theory. Specific topics include adaptive signal processing for communications, design and analysis of wireless and mobile networks, fundamental limits on the performance of large-scale ad hoc and sensor networks, and energy-constrained signal processing and networking techniques. She received the IEEE Signal Processing Society Young Author Best Paper Award.

Lang Tong is a Professor in the School of Electrical and Computer Engineering, Cornell University, Ithaca, New York. He received the B.E. degree from Tsinghua University, and the M.S. and Ph.D. degrees from the University of Notre Dame. He was a Postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University. Prior to joining Cornell University, he was on the faculty at the University of Connecticut and the West Virginia University. He was also the 2001 Cor Wit Visiting Professor at the Delft University of Technology. He received the Young Investigator Award from the Office of Naval Research in 1996, and the Outstanding Young Author Award from the IEEE Circuits and Systems Society. His areas of interest include statistical signal processing, wireless communications, communication networks and sensor networks, and information theory.