AFRL-IF-RS-TR-2004-260
Final Technical Report
September 2004

ROBUST ADAPTIVE SPREAD-SPECTRUM RECEIVERS: SMART ANTENNA/SPACE TIMER INTERFERENCE MITIGATION TECHNIQUES

State University of New York at Buffalo

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

STINFO FINAL REPORT

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-IF-RS-TR-2004-260 has been reviewed and is approved for publication

APPROVED: /s/

MICHAEL J. MEDLEY
Project Engineer

FOR THE DIRECTOR: /s/

WARREN H. DEBANY, JR., Technical Advisor
Information Grid Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	SEPTEMBER 2004	Final May 00 – Aug 03

4. TITLE AND SUBTITLE
ROBUST ADAPTIVE SPREAD-SPECTRUM RECEIVERS: SMART ANTENNA/SPACE TIMER INTERFERENCE MITIGATION TECHNIQUES

5. FUNDING NUMBERS
C - F30602-00-1-0520
PE - 62702F
PR - 4519
TA - 42
WU - 05

6. AUTHOR(S)
Stella N. Batalama and
Dimitris A. Pados

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
State University of New York at Buffalo
Department of Electrical Engineering
Buffalo New York 14260

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory/IFGC
525 Brooks Road
Rome New York 13441-4505

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
AFRL-IF-RS-TR-2004-260

11. SUPPLEMENTARY NOTES

AFRL Project Engineer: Michael J. Medley/IFGC/(315) 330-4830/ Michael.Medley@rl.af.mil

12a. DISTRIBUTION / AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

This work focuses on the problem of adaptive receiver design for direct sequence spread-spectrum wireless communications systems with limited data support dictated by the rapidly changing characteristics of the wireless channel. Both linear and non-linear (neutral network) receiver structures are considered, usually equipped with an antenna array (space-time processing).

14. SUBJECT TERMS
Adaptive Signal Processing, Space-Time Signal Processing, Spread-Spectrum Communications, Multiple-Access, Multi-User Detection

15. NUMBER OF PAGES
178

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Table of Contents

List of Figures

List of Tables

Summary

This work focuses on the problem of adaptive receiver design for direct sequence spread-spectrum wireless communications systems with limited data support dictated by the rapidly changing characteristics of the wireless channel. Both linear and non-linear (neural network) receiver structures are considered, usually equipped with an antenna array (space-time processing). A total of 18 publications have been supported by this grant which are listed at the end of this summary. Our developments are presented in Chapters 1 through 11 and are summarized below.

In Chapter 1, based on statistical conditional optimization criteria, we develop an iterative algorithm that starts from the matched filter (or constraint vector) and generates a sequence of filters that converges to the minimum-variance-distortionless-response (MVDR) solution for any positive definite input autocorrelation matrix. Computationally, the algorithm is a simple recursive procedure that avoids explicit matrix inversion, decomposition, or diagonalization operations. When the input autocorrelation matrix is replaced by a conventional sample-average (positive definite) estimate, the algorithm effectively generates a sequence of MVDR filter estimators: The bias converges rapidly to zero and the covariance trace rises slowly and asymptotically to the covariance trace of the familiar sample-matrix-inversion (SMI) estimator. For short data records, early, non-asymptotic, elements of the generated sequence of estimators offer favorable bias/covariance balance and are seen to outperform in mean-square estimation error constraint-LMS, RLS-type, orthogonal multistage decomposition, as well as plain and diagonally loaded SMI estimates.

The problem of selecting the most successful (in some appropriate sense) filter estimator in the sequence for a given data record is also addressed and two data-driven selection criteria are proposed. The first criterion minimizes the cross-validated sample average variance of the filter estimator output. The second criterion maximizes the estimated J-divergence of the filter estimator output conditional distributions. Illustrative interference suppression examples are followed throughout this presentation. The material presented in this chapter has been published in [A1].

In Chapter 2 we investigate the user capacity, throughput, and delay characteristics of a mobile slotted ALOHA direct-sequence code-division-multiple-access (DS-CDMA) link with dedicated signatures under multipath fading and packet-rate adaptive antenna array signal reception. For a given system transmission bit rate, the packet size is designed to be sufficiently small to conform with the coherence time of the channel. Then, on an individual packet-by-packet basis a phase-ambiguous spatial-temporal channel estimate is produced by a blind (unsupervised) eigen-subspace procedure. The space-time channel estimate is phase corrected via a few pilot packet mid-amble bits and used for joint spatial-temporal multiple-access-interference (MAI) suppression according to the principles of auxiliary-vector filtering. Subsequently, packet success probabilities are derived in the presence or absence of forward error correction (FEC) and are used to evaluate the throughput and delay characteristics of the link. The material presented in this chapter has appeared in [A2].

In Chapter 3, we consider a multi-layer perceptron neural network receiver architecture for the recovery of the information bits of a direct-sequence code-division-multiple-access (DS-CDMA) user. We develop a fast converging adaptive training algorithm that minimizes the bit error rate (BER) at the output of the receiver. The adaptive algorithm has three key features: *(i)* it incorporates the BER, i.e. the ultimate performance evaluation measure, directly into the learning process, *(ii)* it utilizes constraints that are derived from the properties of the optimum single-user decision boundary for AWGN multiple-access channels, and *(iii)* it embeds importance sampling principles directly into the receiver optimization process. Simulation studies illustrate the BER performance of the proposed scheme. The material presented in Chapter 3 has been published in [A3], [A4].

In Chapter 4, we investigate the relative output signal-to-interference-plus-noise ratio (SINR) performance of two linear direct-sequence code-division-multiple-access (DS/CDMA) multiuser detectors: the full decorrelator and the partial decorrelator. We derive necessary and sufficient conditions on the system parameters under which the partial decorrelator outperforms the full decorrelator in the output SINR sense. As a side study, we consider a blind implementation of the full decorrelator that is based on eigendecomposition of the interference-plus-noise autocovariance matrix and can be easily modified to provide a partial decorrelator. Simulation studies illustrate the relative SINR and BER performance of the full and partial decorrelator under perfectly known and sample-average-estimated input statistics. The material presented in this chapter has been published in [A5], [A6].

The problem under consideration in Chapter 5 is the adaptive reception of a spread-spectrum (SS) signal in the presence of unknown correlated SS interference and additive impulsive noise. The general receiver structure is composed of an adaptive chip-based non-linear pre-processor followed by an adaptive linear post-processor. We develop two blind algorithms for the adaptation of the parameters of the non-linear pre-processor; the first algorithm couples itself with the adaptation of the linear post-processor, while the second algorithm is a decoupled procedure that is independent of the linear post-processor. The material presented in this chapter has been published in [A7], [A8].

In Chapter 6, we analyze the performance of DS-CDMA single-user antenna-array receivers that utilize either a cascade of spatial and temporal linear filters or one joint space-time linear filter. We consider both MF-type and MMSE/MVDR-type filters, and we use pre-detection signal-to-interference-plus-noise-ratio (SINR) as the figure of merit. Our goal is twofold: *(i)* Quantify the receiver output variance and receiver output SINR performance of disjoint and joint space-time configurations given ideal input statistics; *(ii)* quantify the confidence level in a certain neighborhood of the optimal performance point of the corresponding estimated structures that are evaluated based on limited sample support due to a fast frequency-nonselective fading channel. These developments have been reported in [A9].

The prohibitive -exponential in the number of users- computational complexity of the maximum likelihood (ML) multiuser detector for direct-sequence code-division-multiple-access (DS/CDMA) communications has fueled an extensive research effort for the development of low complexity multiuser detection alternatives. In Chapter 7, we show that we can efficiently and effectively approach the error rate performance of the optimum multiuser detector as follows. We utilize a multiuser zero-forcing or minimum-mean-square-error (MMSE) linear filter as a pre-processor and we estab-

lish that the output magnitudes, when properly scaled, provide a reliability measure for each user bit decision. Then, we prepare an ordered reliability-based error search sequence of length linear in the number of users that returns the most likely user bit vector among all visited options. Numerical and simulation studies for moderately loaded systems that permit exact implementation of the optimum detector indicate that the error rate performance of the optimum and the proposed detector are nearly indistinguishable over the whole pre-detection signal-to-noise ratio (SNR) range of practical interest. Similar studies for higher user loads (that prohibit comparisons with the optimum detector) demonstrate error rate performance gains of orders of magnitude in comparison with straight decorrelating or MMSE multiuser detection. The material presented in this chapter has been published in [A10]-[A12].

In Chapter 8 we consider blind adaptive linear receivers for the demodulation of direct-sequence code-division-multiple-access (DS/CDMA) signals in asynchronous transmissions. The proposed structures are self-synchronized in the sense that adaptive synchronization and demodulation are viewed and treated as an integrated receiver operation. Two computationally efficient *combined* synchronization/demodulation schemes are proposed, developed and analyzed. The first scheme is based on the principles of minimum-variance-distortionless-response (MVDR) processing, while the second scheme follows the principles of auxiliary-vector filtering and exhibits enhanced performance in short data record scenarios. In both cases the resulting receiver is a linear structure of order exactly equal to the system processing gain. Simulation studies included in this chapter demonstrate the coarse synchronization as well as the bit-error-rate performance of the proposed strategies. The material presented in this chapter has appeared in [A13].

Chapter 9 deals with the theoretical finite data-record performance analysis of the developments in Chapter 8. In this chapter we investigate the coarse synchronization performance of blind adaptive linear *self-synchronized* receivers for asynchronous direct-sequence code-division-multiple-access communications under finite data record adaptation. Based on transformation noise modeling techniques, three alternative methods are developed leading to analytic expressions that approximate the probability of coarse synchronization error of matched-filter-type (MF) and minimum-variance-distortionless-response-type (MVDR) receivers. The expressions are explicit functions of the data record size and the filter order and reveal the effect of short-data-record sample-matrix-inversion (SMI) implementations on the coarse synchronization performance. Besides their theoretical interest, the derived expressions provide simple, highly-accurate alternatives to computationally demanding performance evaluation through simulations. The effect of the data record size on the probability of coarse synchronization error is further quantified through the use of a receiver synchronization resolution metric. Numerical and simulation studies examine the accuracy of the theoretical developments and show that the derived expressions approximate closely the actual coarse synchronization performance. The material presented in this chapter has appeared in [A14].

In Chapter 10 we address the problem of navigation data demodulation by an adaptive GPS receiver that utilizes a bank of single-satellite linear-tap-delay filters and employs antenna-array reception. The presence of an antenna array allows the receiver to operate in the spatial domain in addition to the temporal (code) domain. We investigate disjoint-domain as well as joint-domain space-time GPS signal processing techniques and we consider design criteria of conventional matched-filter (MF) type, minimum-variance-distortionless-response (MVDR) type and auxiliary-vector (AV) type. The proposed structures utilize filters that operate at a fraction of the navigation

data bit period (1 msec) and are followed by hard-decision detectors. Hard decisions taken over a navigation data bit period are then combined according to a simple combining rule for further bit-error-rate (BER) performance improvements. Analytic, numerical and simulation comparisons illustrate the relative merits of the investigated design alternatives. The material presented in this chapter has been published in [A15], [A16].

Finally in Chapter 11, we develop a new low complexity algorithm for decoding low-density parity-check (LDPC) codes. The developments are oriented specifically toward the low cost -yet effective- decoding of (high rate) finite geometry LDPC codes. The decoding procedure updates the hard-decision received vector iteratively in search of a valid codeword in the vector space. Only one bit is changed in each iteration and the bit selection criterion combines the number of failed checks and the reliability of the received bits. Prior knowledge of the signal amplitude and noise power is not required. An optional mechanism to avoid infinite loops in the search is also proposed. Our studies show that the algorithm achieves an appealing performance versus complexity trade-off for finite geometry LDPC codes. The material presented in this chapter has appeared in [A17], [A18].

Chapter 1

Short-data-record Adaptive Filtering: The Auxiliary-Vector Algorithm

1.1 Introduction

Minimum-variance-distortionless-response (MVDR) filtering refers to the problem of identifying a linear filter that minimizes the variance at its output, while at the same time the filter maintains a distortionless response toward a specific input vector direction of interest. If $\mathbf{r}$ is a random, zero mean without loss of generality, complex input vector of dimension L, $\mathbf{r} \in \mathbb{C}^L$, that is processed by an L-tap filter $\mathbf{w} \in \mathbb{C}^L$, then the filter output variance is $\mathbf{w}^H \mathbf{R} \mathbf{w}$, where $\mathbf{R} = E\{\mathbf{r}\mathbf{r}^H\}$ is the input autocorrelation matrix ($E\{\cdot\}$ denotes the statistical expectation operation and $\mathbf{x}^H$ denotes the Hermitian that is, transpose conjugate of $\mathbf{x}$). The MVDR filter minimizes $\mathbf{w}^H \mathbf{R} \mathbf{w}$ and simultaneously satisfies $\mathbf{w}^H \mathbf{v} = 1$, or more general $\mathbf{w}^H \mathbf{v} = \rho \in \mathbb{C}$, where $\mathbf{v} \in \mathbb{C}^L$ is the signal vector direction to be protected. In this set-up, MVDR filtering is a standard linear constraint optimization problem and a conventional Lagrange multipliers procedure leads to the well known solution [1], [2]

$$\mathbf{w}_{\mathrm{MVDR}} = \rho^* \frac{\mathbf{R}^{-1}\mathbf{v}}{\mathbf{v}^H \mathbf{R}^{-1} \mathbf{v}} \tag{1.1}$$

where ρ^* denotes the conjugate of the desired response $\mathbf{w}^H \mathbf{v} = \rho$. MVDR filtering has been used extensively in unsupervised signal processing applications where a desired scalar filter output $d \in \mathbb{C}$ cannot be identified or cannot be assumed available for each input $\mathbf{r} \in \mathbb{C}^L$ (for example in radar and array processing problems where the constraint vector $\mathbf{v}$ is usually referred to as the "target" or "look" direction of interest). We may also observe the close relationship between the MVDR filter and the minimum-mean-square-error (MMSE) or Wiener filter. Indeed, if the constraint vector $\mathbf{v}$ is chosen to be the statistical cross-correlation vector between the desired output d and the input vector $\mathbf{r}$, that is if $\mathbf{v} = E\{\mathbf{r}d^*\}$, then the MVDR and MMSE filters become scaled versions of each other, $c\mathbf{R}^{-1}\mathbf{v}$, $c \in \mathbb{C}$. For this reason, in the rest of the chapter we will use the term "MMSE/MVDR filter" to refer to either filter.

In this work, first we present an iterative algorithm for the calculation of the MMSE/MVDR vector $\mathbf{w}_{\mathrm{MMSE/MVDR}}$ in (1.1). The algorithm is a non-invasive procedure where no explicit matrix inversion/eigen-decomposition/diagonalization is attempted. The MMSE/MVDR computation algorithm creates a sequence of filters $\mathbf{w}_n$, $n = 0, 1, 2, \ldots$, that begins from $\mathbf{w}_0 = \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v}$ and converges to the MMSE/MVDR filter ($\mathbf{w}_\infty = \mathbf{w}_{\mathrm{MMSE/MVDR}}$). At each step $n = 1, 2, \ldots$, $\mathbf{w}_n$ is given as a simple, direct function of $\mathbf{R}$, $\mathbf{v}$, and $\mathbf{w}_{n-1}$.

The development of the iterative algorithm (which we call "the auxiliary-vector algorithm" for reasons that will become apparent in the sequel) is founded solely on statistical signal processing principles. The motivation behind its development is *adaptive* signal processing where the input autocorrelation matrix $\mathbf{R}$ is assumed unknown and it is sample-average estimated by a data record of M points, $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_M$:

$$\hat{\mathbf{R}}(M) = \frac{1}{M} \sum_{m=1}^{M} \mathbf{r}_m \mathbf{r}_m^H. \tag{1.2}$$

When $\mathbf{R}$ is substituted by $\hat{\mathbf{R}}(M)$ in the recursively generated sequence of filters $\mathbf{w}_n$, $n = 0, 1, 2, \ldots$, the corresponding filter estimators $\hat{\mathbf{w}}_n(M)$, $n = 0, 1, 2, \ldots$, offer the means for effective control over the filter estimator bias versus (co-)variance trade-off [3]. Starting from the *0*-variance, high-bias (for non-white inputs) $\hat{\mathbf{w}}_0(M) = \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v}$ estimate, we can go all the way up to the unbiased, yet high-variance for small data record sizes M, $\hat{\mathbf{w}}_\infty(M)$ estimate and anywhere in between, $\hat{\mathbf{w}}_n(M)$, $1 \leq n < \infty$. As a result, adaptive filters from this newly developed class can be seen to outperform in expected norm-square estimation error, $E\{\|\hat{\mathbf{w}}(M) - \mathbf{w}_{\text{MMSE/MVDR}}\|^2\}$, (constraint-)LMS [4], sample-matrix-inversion (SMI) [5] with or without diagonal loading [6], RLS-type [7], [8], and orthogonal multistage decomposition [9], [10] adaptive filter implementations. It is worth mentioning that the familiar trial-and-error tuning to problem and data-record-size specifics of the real-valued LMS gain or RLS inverse matrix initialization constant or SMI diagonal loading parameter that plagues field practitioners is now replaced by an integer choice of one of the recursively generated filters.

The problem of selecting the best (in some appropriate sense) filter estimator in the sequence for a given data record is addressed and two data-driven selection criteria are proposed. The first criterion is rather general and is motivated by the asymptotic minimum output variance property of the generated sequence of filter estimators. In particular, for a given data record, we select the filter estimator that has minimum cross-validated average output variance (energy). The second rule is built specifically for binary antipodal (BPSK-type) communication signals and is related to the objective of achieving maximum stochastic distance between the two conditional distributions of the filter estimator output. Under this rule, we choose the filter estimator in the sequence that exhibits maximum estimated J-divergence of the conditional output distributions. We pursue and analyze both supervised and unsupervised (blind) implementations of this criterion. Illustrative case studies drawn from the code-division-multiple-access (CDMA) communications literature are followed throughout this work.

The rest of the chapter is organized as follows. In Section 1.2 we present the basic algorithmic development and analysis results. Filter estimation issues are discussed in Section 1.3. The two data-driven criteria for the selection of a filter estimator from the generated sequence are developed in Section 1.4. In Section 1.5 we examine the quality of the proposed criteria through simulations. A few concluding remarks are given in Section 1.6.

1.2 Algorithmic Developments and Convergence Analysis

For a given constraint vector $\mathbf{v} \in \mathbb{C}^L$ consider the set of filters $\mathcal{D} = \{\mathbf{w} \in \mathbb{C}^L : \mathbf{w} = \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v} + \mathbf{u}, \mathbf{u} \in \mathbb{C}^L \text{ and } \mathbf{v}^H\mathbf{u} = 0\}$. $\mathcal{D}$ is the class of all filters $\mathbf{w}$ in $\mathbb{C}^L$ that are distortionless in $\mathbf{v}$, that is $\mathbf{w}^H\mathbf{v} = \rho$. In this section we develop an iterative algorithm for the computation of the $\mathbf{u}$ component of the MMSE/MVDR filter.

Algorithmic designs that focus on the MMSE/MVDR filter part $\mathbf{u}$ that is orthogonal to the constraint vector, or "look", direction $\mathbf{v}$ have been widely pursued in the array processing literature and have been known as generalized sidelobe cancelers (GSC) [11] or partially adaptive beamformers [12]. Recent developments have been influenced by Principal Component Analysis reduced-rank processing principles [13]. In general, the MMSE/MVDR filter part $\mathbf{u}$ ($\mathbf{u}^H\mathbf{v} = 0$) has been approximated by $\mathbf{u}_{L\times 1} \simeq -\mathbf{B}_{L\times(L-1)}\mathbf{T}_{(L-1)\times p}\mathbf{w}^{\text{GSC}}_{p\times 1}$, where $\mathbf{B}$ is the so called "blocking matrix" that satisfies $\mathbf{B}^H\mathbf{v} = \mathbf{0}_{L-1}$ ($\mathbf{B}$ is a full column-rank matrix that can be derived by Gram-Schmidt orthogonalization of an $L \times L$ orthogonal projection operator such as $\mathbf{I} - \frac{\mathbf{v}\mathbf{v}^H}{\|\mathbf{v}\|^2}$, where $\mathbf{I}$ is the identity matrix), $\mathbf{T}$ is the rank reducing matrix with $1 \leq p < L-1$ columns to be designed, and $\mathbf{w}^{\text{GSC}}$ is the MS-optimum vector of weights of the p columns of $\mathbf{T}$ ($\mathbf{w}^{\text{GSC}} = \frac{\rho^*}{\|\mathbf{v}\|^2}[\mathbf{T}^H\mathbf{B}^H\mathbf{R}\mathbf{B}\mathbf{T}]^{-1}\mathbf{T}^H\mathbf{B}^H\mathbf{R}\mathbf{v}$ [12]). In [14] and [15] the p columns of $\mathbf{T}$ were chosen to be the p maximum eigenvalue eigenvectors of the blocked data autocorrelation matrix $\mathbf{B}^H\mathbf{R}\mathbf{B}$. If, however, the columns of $\mathbf{T}$ *have to be* eigenvectors of $\mathbf{B}^H\mathbf{R}\mathbf{B}$, then the best way to choose them in the minimum output variance p-rank approximation sense was presented in [16]: Select the p eigenvectors $\mathbf{q}_i$ of $\mathbf{B}^H\mathbf{R}\mathbf{B}$, with corresponding eigenvalues λ_i, that maximize $\frac{|\mathbf{v}^H\mathbf{R}\mathbf{q}_i|^2}{\lambda_i}$, $i = 1, \ldots, p$. This design algorithm has also been known as "cross-spectral metric" reduced-rank processing [17]. A different approach from a different point of view is described in this work. A *conditional* statistical optimization procedure is shown to offer the means for *exact* computation of $\mathbf{u}$ as the convergence point of an infinite series of the form $-\sum_{n=1}^{\infty}\mu_n\mathbf{g}_n$, $\mu_n \in \mathcal{R}^+$, $\mathbf{g}_n \in \mathcal{C}^L$ and $\mathbf{g}_n^H\mathbf{v} = 0$, $\forall n = 1, 2, \ldots$.

We begin the algorithmic developments from the conventional matched filter (MF) with desired response

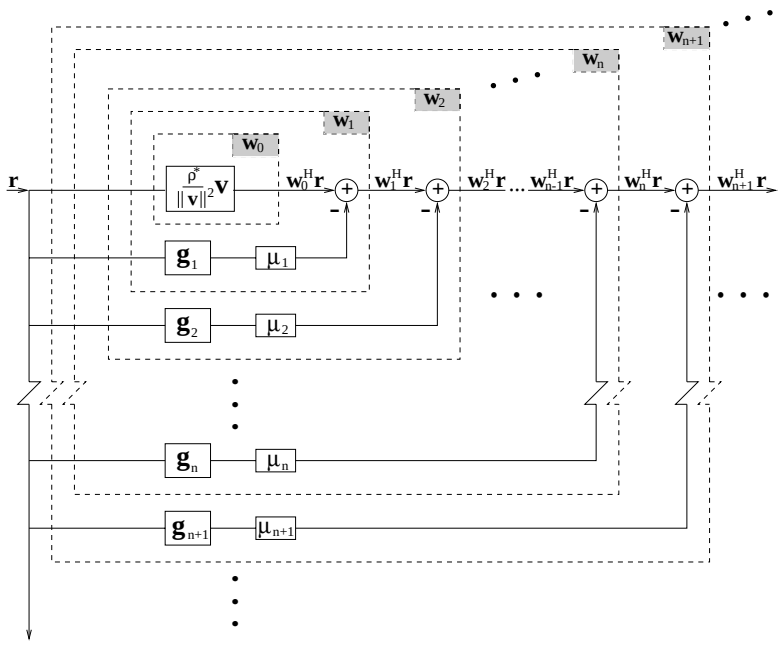

Fig. 1.1: Block diagram representation of the iteratively generated sequence of filters $\mathbf{w}_0, \mathbf{w}_1, \mathbf{w}_2, \ldots$.

$\mathbf{w}^H\mathbf{v} = \rho$

$$\mathbf{w}_0 = \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v}, \tag{1.3}$$

which is MMSE/MVDR optimum for white $\mathbb{C}^L$ vector inputs (when $\mathbf{R} = \sigma^2\mathbf{I}$, $\sigma > 0$). We recall that, w.l.o.g. and for notational simplicity, we assume throughout this presentation that the input vectors $\mathbf{r} \in \mathbb{C}^L$ are zero mean. Next, we incorporate in $\mathbf{w}_0$ an "auxiliary" vector component that is orthogonal to $\mathbf{v}$ and we form (Fig. 1.1)

$$\mathbf{w}_1 = \mathbf{w}_0 - \mu_1\mathbf{g}_1 = \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v} - \mu_1\mathbf{g}_1 \tag{1.4}$$

where $\mathbf{g}_1 \in \mathbb{C}^L - \{\mathbf{0}\}$, $\mu_1 \in \mathbb{C}$, and $\mathbf{g}_1^H\mathbf{v} = 0$. We assume for a moment that the orthogonal auxiliary vector $\mathbf{g}_1$ is arbitrary but non-zero and fixed and we concentrate on the selection of the scalar μ_1. The value of μ_1 that minimizes the variance of the output of the filter $\mathbf{w}_1$ can be found by direct differentiation of $E\{|\mathbf{w}_1^H\mathbf{r}|^2\}$ or simply as the value that minimizes the MS error between $\mathbf{w}_0^H\mathbf{r} = \frac{\rho}{\|\mathbf{v}\|^2}\mathbf{v}^H\mathbf{r}$ and $\mu_1^*\mathbf{g}_1^H\mathbf{r}$. This is essentially a scalar version of the GSC weight determination problem and we present the solution in the form of a proposition [18]:

Proposition 1.1 *The scalar μ_1 that minimizes the variance at the output of $\mathbf{w}_1$ or equivalently minimizes the MS error between $\mathbf{w}_0^H\mathbf{r} = \frac{\rho}{\|\mathbf{v}\|^2}\mathbf{v}^H\mathbf{r}$ and $\mu_1^*\mathbf{g}_1^H\mathbf{r}$ is*

$$\mu_1 = \frac{\mathbf{g}_1^H\mathbf{R}\mathbf{w}_0}{\mathbf{g}_1^H\mathbf{R}\mathbf{g}_1} \tag{1.5}$$

where $\mathbf{R} = E\{\mathbf{r}\mathbf{r}^H\}$ is the input autocorrelation matrix. □

Since $\mathbf{g}_1$ is set to be orthogonal to $\mathbf{v}$, (1.5) shows that if the vector $\mathbf{R}\mathbf{w}_0$ happens to be "on $\mathbf{v}$" (that is $\mathbf{R}\mathbf{w}_0 = \frac{\mathbf{v}^H\mathbf{R}\mathbf{w}_0}{\|\mathbf{v}\|^2}\mathbf{v}$ or equivalently $(\mathbf{I} - \frac{\mathbf{v}\mathbf{v}^H}{\|\mathbf{v}\|^2})\mathbf{R}\mathbf{w}_0 = \mathbf{0}$), then $\mu_1 = 0$. Indeed, if $\mathbf{R}\mathbf{w}_0 = \frac{\mathbf{v}^H\mathbf{R}\mathbf{w}_0}{\|\mathbf{v}\|^2}\mathbf{v}$ then $\mathbf{w}_0$ is *already* the MMSE/MVDR filter. To avoid this trivial case and continue with our developments, we suppose

that $\mathbf{R}\mathbf{w}_0 \neq \frac{\mathbf{v}^H\mathbf{R}\mathbf{w}_0}{\|\mathbf{v}\|^2}\mathbf{v}$. By inspection, we also observe that for the MS-optimum value of μ_1 the product $\mu_1\mathbf{g}_1 = \frac{\mathbf{g}_1^H\mathbf{R}\mathbf{w}_0}{\mathbf{g}_1^H\mathbf{R}\mathbf{g}_1}\mathbf{g}_1$ is independent of the norm of $\mathbf{g}_1$. Hence, so is $\mathbf{w}_1$. At this point, we decide to choose the auxiliary vector $\mathbf{g}_1$ as the normalized vector that maximizes the magnitude of the cross-correlation between $\mathbf{w}_0^H\mathbf{r} = \frac{\rho}{\|\mathbf{v}\|^2}\mathbf{v}^H\mathbf{r}$ and $\mathbf{g}_1^H\mathbf{r}$, under the constraint that $\mathbf{g}_1^H\mathbf{v} = 0$ and $\mathbf{g}_1^H\mathbf{g}_1 = 1$:

$$\mathbf{g}_1 = \arg\max_{\mathbf{g}} \left| E\left\{ \mathbf{w}_0^H\mathbf{r}\left(\mathbf{g}^H\mathbf{r}\right)^* \right\} \right| = \arg\max_{\mathbf{g}} \left| \mathbf{w}_0^H\mathbf{R}\mathbf{g} \right| \quad \text{subject to } \mathbf{g}^H\mathbf{v} = 0 \text{ and } \mathbf{g}^H\mathbf{g} = 1. \tag{1.6}$$

For the sake of mathematical accuracy, we note that both the criterion function $\left|\mathbf{w}_0^H\mathbf{R}\mathbf{g}\right|$ to be maximized as well as the orthogonality constraints are phase invariant. In other words, if $\mathbf{g}_1$ satisfies (1.6) so does $\mathbf{g}_1 e^{j\phi}$ for any phase ϕ. Without loss of generality, to avoid any ambiguity in our presentation and to have a uniquely defined auxiliary vector, we seek the one and only auxiliary vector that satisfies (1.6) and places the cross-correlation value on the positive real line ($\mathbf{w}_0^H\mathbf{R}\mathbf{g} > 0$). This constraint optimization problem was first posed and solved in [19] where the filter $\mathbf{w}_1$ in (1.4) was used for multiple access interference suppression in multipath CDMA communication channels. Intuitively, the maximum magnitude cross-correlation criterion as defined in (1.6) strives to identify the auxiliary vector orthonormal to $\mathbf{v}$ that can capture the most interference present in $\mathbf{w}_0^H\mathbf{r}$. The solution, derived through conventional Lagrange multipliers optimization, is given below.

Proposition 1.2 *Suppose that* $(\mathbf{I} - \frac{\mathbf{v}\mathbf{v}^H}{\|\mathbf{v}\|^2})\mathbf{R}\mathbf{w}_0 \neq \mathbf{0}$ $(\mathbf{w}_0 \neq \mathbf{w}_{\text{MMSE/MVDR}})$. *Then, the auxiliary vector*

$$\mathbf{g}_1 = \frac{\mathbf{R}\mathbf{w}_0 - \frac{\mathbf{v}^H\mathbf{R}\mathbf{w}_0}{\|\mathbf{v}\|^2}\mathbf{v}}{\left\|\mathbf{R}\mathbf{w}_0 - \frac{\mathbf{v}^H\mathbf{R}\mathbf{w}_0}{\|\mathbf{v}\|^2}\mathbf{v}\right\|} \tag{1.7}$$

maximizes the magnitude of the cross-correlation between $\mathbf{w}_0^H\mathbf{r} = \frac{\rho}{\|\mathbf{v}\|^2}\mathbf{v}^H\mathbf{r}$ *and* $\mathbf{g}_1^H\mathbf{r}$, $\left|\mathbf{w}_0^H\mathbf{R}\mathbf{g}_1\right|$, *subject to the constraints* $\mathbf{g}_1^H\mathbf{v} = 0$ *and* $\mathbf{g}_1^H\mathbf{g}_1 = 1$. *In addition,* $\mathbf{w}_0^H\mathbf{R}\mathbf{g}_1$ *is real positive* $(\mathbf{w}_0^H\mathbf{R}\mathbf{g}_1 > 0)$. □

So far we have defined $\mathbf{w}_0$ in (1.3) and $\mathbf{w}_1$ in (1.4) with $\mathbf{g}_1$ and μ_1 given by (1.7) and (1.5), respectively. The iterative algorithm for the generation of an infinite sequence of filters $\mathbf{w}_0, \mathbf{w}_1, \mathbf{w}_2, \ldots$ is already taking shape. Formally, we just need to specify the inductive step. Assuming that the filter $\mathbf{w}_n = \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v} - \sum_{i=1}^{n}\mu_i\mathbf{g}_i$ has been identified for some $n \geq 1$ and $\mathbf{w}_n \neq \mathbf{w}_{\text{MMSE/MVDR}}$, we argue as in Propositions 1 and 2 and we define $\mathbf{w}_{n+1}$ as follows:

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \mu_{n+1}\mathbf{g}_{n+1} \tag{1.8}$$

where

$$\mathbf{g}_{n+1} = \frac{\mathbf{R}\mathbf{w}_n - \frac{\mathbf{v}^H\mathbf{R}\mathbf{w}_n}{\|\mathbf{v}\|^2}\mathbf{v}}{\left\|\mathbf{R}\mathbf{w}_n - \frac{\mathbf{v}^H\mathbf{R}\mathbf{w}_n}{\|\mathbf{v}\|^2}\mathbf{v}\right\|} \tag{1.9}$$

is the orthonormal with respect to $\mathbf{v}$ auxiliary vector that, given $\mathbf{w}_n$, maximizes conditionally the cross-correlation magnitude $\left|E\left\{\mathbf{w}_n^H\mathbf{r}\left(\mathbf{g}_{n+1}^H\mathbf{r}\right)^*\right\}\right| = \left|\mathbf{w}_n^H\mathbf{R}\mathbf{g}_{n+1}\right|$ and

$$\mu_{n+1} = \frac{\mathbf{g}_{n+1}^H\mathbf{R}\mathbf{w}_n}{\mathbf{g}_{n+1}^H\mathbf{R}\mathbf{g}_{n+1}} \tag{1.10}$$

is the scalar that minimizes the MS error between $\mathbf{w}_n^H\mathbf{r}$ and $\mu_{n+1}^*\mathbf{g}_{n+1}^H\mathbf{r}$ (minimizes $E\{|\mathbf{w}_{n+1}^H\mathbf{r}|^2\}$).

It is important to note that, while the generated auxiliary vectors $\mathbf{g}_1, \mathbf{g}_2, \ldots$ are all constrained to be orthogonal to $\mathbf{v}$, orthogonality among the auxiliary vectors is *not* imposed [20], [21]. This is in sharp contrast to previous work that involved filtering with up to $L-1$ orthogonal to each other and to $\mathbf{v}$ vectors [22]-[24], where L is the data input vector dimension. We observe, however, that *successive* auxiliary vectors generated by the above recursive conditional optimization procedure (1.8)-(1.10) *do* come up orthogonal: $\mathbf{g}_n^H\mathbf{g}_{n+1} = 0$, $\forall n = 1, 2, 3, \ldots$ (while $\mathbf{g}_n^H\mathbf{g}_m \neq 0$, $\forall n, m$, $|n-m| \neq 1$). For completeness purposes, below we present this observation in the form of a Lemma.

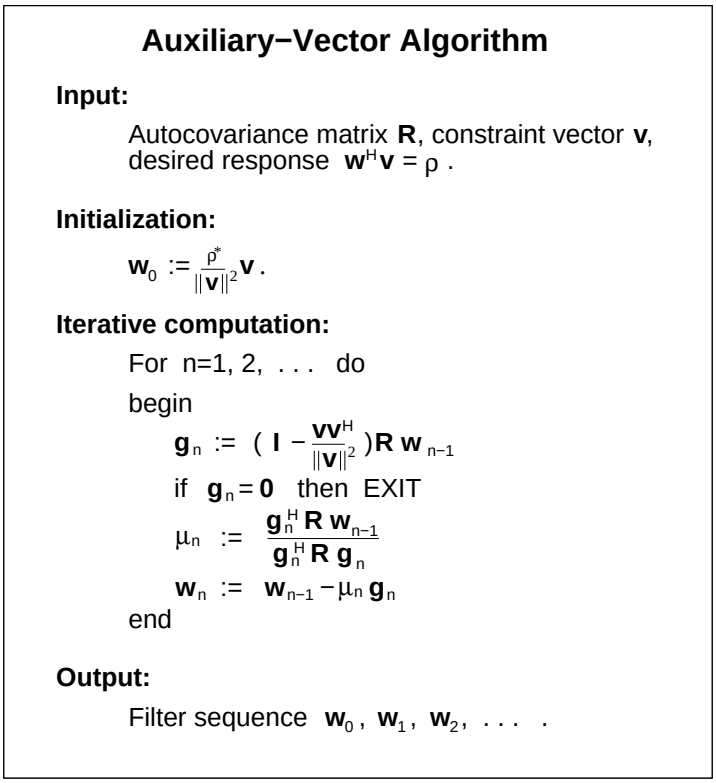

Fig. 1.2: The algorithm for the iterative generation of the filter sequence $\mathbf{w}_0, \mathbf{w}_1, \mathbf{w}_2, \ldots$.

Lemma 1.1 *Successive auxiliary vectors generated through (1.8)-(1.10) are orthogonal:* $\mathbf{g}_n^H \mathbf{g}_{n+1} = 0$, $n = 1, 2, 3, \ldots$. *However,* $\mathbf{g}_n^H \mathbf{g}_m \neq 0$, $\forall n, m$, $|n-m| \neq 1$. □

The algorithm is summarized in Fig. 1.2. The conceptual simplicity of the conditional statistical optimization process led to a computationally simple recursion. In Fig. 1.2 we chose to drop the unnecessary, as previously explained, normalization of the auxiliary vectors and we also factorized their numerator to make the orthogonal projection operator apparent. Formal convergence of the filter sequence $\mathbf{w}_0, \mathbf{w}_1, \mathbf{w}_2, \ldots$ to the MMSE/MVDR filter $\rho^* \frac{\mathbf{R}^{-1}\mathbf{v}}{\mathbf{v}^H \mathbf{R}^{-1}\mathbf{v}}$ is established by the following Theorem. The proof can be found in [21].

Theorem 1.1 *Let* $\mathbf{R}$ *be a Hermitian positive definite matrix. Consider the iterative algorithm of Fig. 1.2.*
(i) *The generated sequence of auxiliary-vector weights* $\{\mu_n\}$, $n = 1, 2, \ldots$, *is real-valued, positive, and bounded:*

$$0 < \frac{1}{\lambda_{\max}} \leq \mu_n \leq \frac{1}{\lambda_{\min}}, \qquad n = 1, 2, \ldots, \tag{1.11}$$

where $\lambda_{\max}$ *and* $\lambda_{\min}$ *are the maximum and minimum, correspondingly, eigenvalues of* $\mathbf{R}$.
(ii) *The sequence of auxiliary vectors* $\{\mathbf{g}_n\}$, $n = 1, 2, \ldots$, *converges to the* $\mathbf{0}$ *vector:*

$$\lim_{n \to \infty} \mathbf{g}_n = \mathbf{0}. \tag{1.12}$$

(iii) *The sequence of auxiliary-vector filters* $\{\mathbf{w}_n\}$, $n = 1, 2, \ldots$, *converges to the MMSE/MVDR filter:*

$$\lim_{n \to \infty} \mathbf{w}_n = \rho^* \frac{\mathbf{R}^{-1}\mathbf{v}}{\mathbf{v}^H \mathbf{R}^{-1}\mathbf{v}}. \tag{1.13}$$

□

We conclude this section with an illustration. We draw a signal model example from the direct-sequence code-division-multiple-access (DS-CDMA) communications literature and we assume a synchronous system where the input signal vector $\mathbf{r} \in \mathcal{R}^L$ is given by

$$\mathbf{r} = \sum_{k=1}^{K} \sqrt{E_k} b_k \mathbf{s}_k + \mathbf{n}. \tag{1.14}$$

In this set-up, K denotes the total number of signals ("users") present and each signal is defined through an L-dimensional, normalized, binary-antipodal vector waveform (or "user signature") $\mathbf{s}_k$, $k = 1, 2, \ldots, K$. The signature vector dimension L is usually referred to as the system "spreading gain". With respect to the k-th user signal, E_k is the received signal energy and $b_k \in \{-1, +1\}$ is the information bit modeled as a random variable with equally probable values and assumed to be statistically independent from all other user bits b_j, $j \neq k$. Additive white Gaussian noise contributions are accounted for by $\mathbf{n}$ with autocorrelation matrix $E\{\mathbf{n}\mathbf{n}^T\} = \sigma^2 \mathbf{I}_{L \times L}$ ($\mathbf{x}^T$ denotes the transpose of $\mathbf{x}$). With this notation and normalized user signatures, the signal-to-noise ratio of the k-th user signal is defined by $\mathrm{SNR}_k \triangleq 10 \log_{10} \frac{E_k}{\sigma^2}$dB, $k = 1, 2, \ldots, K$.

MMSE/MVDR filtering for DS-CDMA type problems has attracted significant interest [25]-[29]. If we wish to recover the information bits of, say, user *1*, then all other signals constitute multiple-access interference and the MMSE/MVDR filter is built with constraint vector $\mathbf{v} = \mathbf{s}_1$, desired response $\mathbf{w}^T \mathbf{s}_1 = 1$, and autocorrelation matrix $\mathbf{R} = \sum_{k=1}^{K} E_k \mathbf{s}_k \mathbf{s}_k^T + \sigma^2 \mathbf{I}$. We choose $L = 32$, $K = 13$, and we draw an arbitrary set of signatures $\mathbf{s}_1$, $\mathbf{s}_2$, ..., $\mathbf{s}_{13}$. For purposes of completeness in presentation, the exact signature assignment is given in the Appendix. We fix the SNR of the user of interest at $\mathrm{SNR}_1 = 12$dB while the "interferers" $k = 2, \ldots, 13$ are at $\mathrm{SNR}_{2-5} = 10$dB, $\mathrm{SNR}_{6-9} = 12$dB and $\mathrm{SNR}_{10-13} = 14$dB. Fig. 1.3 shows how the sequence of filters $\mathbf{w}_0$, $\mathbf{w}_1$, ... generated by the algorithm in Fig. 1.2 converges to the MMSE/MVDR solution. The convergence is captured in terms of the norm-square metric $\|\mathbf{w}_n - \mathbf{w}_{\mathrm{MMSE/MVDR}}\|^2$ as a function of the iteration step (index of the AV filter in the sequence or number of auxiliary vectors used) n.

1.3 Filter Estimation

Consider a constraint vector $\mathbf{v}$ and a Hermitian positive definite autocorrelation matrix $\mathbf{R}$ of an input vector $\mathbf{r} \in \mathbb{C}^L$. Assume that $\mathbf{R}$ is in fact unknown and it is sample-average estimated from a data record of M

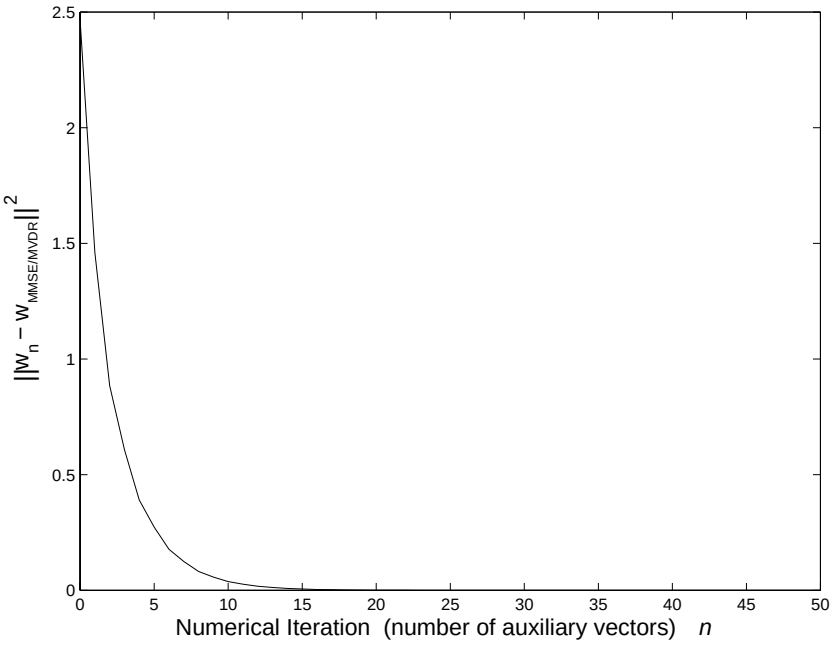

Fig. 1.3: Convergence of the sequence of filters $\mathbf{w}_n$, $n = 0, 1, 2, \ldots$, to the MMSE/MVDR solution for the signal model example in (1.14).

points: $\hat{\mathbf{R}}(M) = \frac{1}{M}\sum_{m=1}^{M} \mathbf{r}_m \mathbf{r}_m^H$. For Gaussian inputs, $\hat{\mathbf{R}}(M)$ is a maximum-likelihood (ML), consistent, unbiased estimator of $\mathbf{R}$ [3], [30]. For a large class of multivariate elliptically contoured input distributions that includes the Gaussian, if $M \geq L$ then $\hat{\mathbf{R}}(M)$ is positive definite (hence invertible) with probability 1 (w.p. 1) [31]-[33]. Then, Theorem 1.1 in Section 1.2 shows that

$$\hat{\mathbf{w}}_n(M) \underset{n\to\infty}{\longrightarrow} \hat{\mathbf{w}}_\infty(M) = \rho^* \frac{\left[\hat{\mathbf{R}}(M)\right]^{-1}\mathbf{v}}{\mathbf{v}^H\left[\hat{\mathbf{R}}(M)\right]^{-1}\mathbf{v}} \tag{1.15}$$

where $\hat{\mathbf{w}}_\infty(M)$ is the widely used MMSE/MVDR filter estimator known as the sample-matrix-inversion (SMI) filter [5].

The output sequence begins from $\hat{\mathbf{w}}_0(M) = \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v}$, which is a *0*-variance, fixed-valued, estimator that may be severely biased ($\hat{\mathbf{w}}_0(M) = \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v} \neq \mathbf{w}_{\mathrm{MMSE/MVDR}}$) unless $\mathbf{R} = \sigma^2\mathbf{I}$, for some $\sigma > 0$. In the latter trivial case, $\hat{\mathbf{w}}_0(M)$ is already the perfect MMSE/MVDR filter. Otherwise, the next filter estimator in the sequence, $\hat{\mathbf{w}}_1(M)$, has a significantly reduced bias due to the optimization procedure employed, at the expense of non-zero estimator (co-)variance. As we move up in the sequence of filter estimators $\hat{\mathbf{w}}_n(M)$, $n = 0, 1, 2, \ldots$, the bias decreases rapidly to zero[1] while the variance rises slowly to the SMI ($\hat{\mathbf{w}}_\infty(M)$) levels (cf. (1.15)). To quantify these remarks, we plot in Fig. 1.4 the norm-square bias $\|E\{\hat{\mathbf{w}}_n(M)\} - \mathbf{w}_{\mathrm{MMSE/MVDR}}\|^2$ and the trace of the covariance matrix $E\left\{[\hat{\mathbf{w}}_n(M) - E\{\hat{\mathbf{w}}_n(M)\}][\hat{\mathbf{w}}_n(M) - E\{\hat{\mathbf{w}}_n(M)\}]^H\right\}$ as a function of the iteration step n, for the signal model example of Fig. 1.3 and data record size $M = 256$. Bias and cov-trace values are calculated from 100000 independent filter estimator realizations for each iteration point n. Formal, theoretical statistical analysis of the generated estimators $\hat{\mathbf{w}}_n(M)$, $n = 0, 1, 2, \ldots$, is beyond the scope of this presentation. We do note, however, that for multivariate elliptically contoured input distributions, an analytic expression for the covariance matrix of the SMI estimator $\hat{\mathbf{w}}_\infty(M)$ can be found

[1] The SMI estimator is unbiased for multivariate elliptically contoured input distributions [33], [34]: $E\{\hat{\mathbf{w}}_\infty(M)\} = \mathbf{w}_{\mathrm{MMSE/MVDR}} = \rho^* \frac{\mathbf{R}^{-1}\mathbf{v}}{\mathbf{v}^H\mathbf{R}^{-1}\mathbf{v}}$.

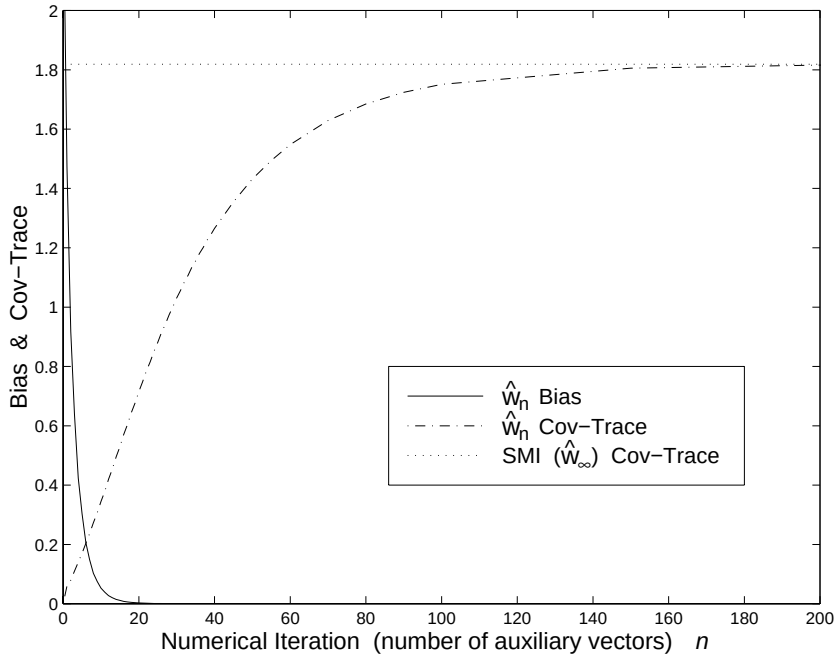

Fig. 1.4: Norm-square bias and covariance trace for the sequence of estimators $\hat{\mathbf{w}}_n(M)$, $n = 0, 1, \ldots$. The signal model is as in Fig. 1.3 and $M = 256$.

in [33]: $E\{[\hat{\mathbf{w}}_\infty(M) - E\{\hat{\mathbf{w}}_\infty(M)\}][\hat{\mathbf{w}}_\infty(M) - E\{\hat{\mathbf{w}}_\infty(M)\}]^H\} = \frac{|\rho|^2}{(\mathbf{v}^H\mathbf{R}^{-1}\mathbf{v})(M-L+1)}\left(\mathbf{R}^{-1} - \frac{\mathbf{R}^{-1}\mathbf{v}\mathbf{v}^H\mathbf{R}^{-1}}{\mathbf{v}^H\mathbf{R}^{-1}\mathbf{v}}\right)$. Since under these input distribution conditions $\hat{\mathbf{w}}_\infty(M)$ is unbiased, the trace of the covariance matrix is the MS filter estimation error. It is important to observe that the covariance matrix and, therefore, the MS filter estimation error depend on the data record size M, the filter length L, as well as the specifics of the signal processing problem at hand ($\mathbf{R}$ and $\mathbf{v}$). It is also important to note that for the CDMA signal model example in (1.14) the input is Gaussian-mixture distributed. Therefore, the analytic result in [33] is not directly applicable and can only be thought of as an approximation (a rather close approximation as we concluded in our studies). In any case, from the results in Fig. 1.4 for $M = 256$, we see that the estimators $\hat{\mathbf{w}}_1(M)$, $\hat{\mathbf{w}}_2(M)$, ..., up to about $\hat{\mathbf{w}}_{20}(M)$ are particularly appealing. In contrast, the estimators $\hat{\mathbf{w}}_n(M)$ for $n > 20$ do not justify their increased cov-trace cost since they have almost nothing to offer in terms of further bias reduction.

The mean-square estimation error expression $E\left\{\|\hat{\mathbf{w}}_n(M) - \mathbf{w}_{\mathrm{MMSE/MVDR}}\|^2\right\}$ captures the bias/variance balance of the individual members of the estimator sequence $\hat{\mathbf{w}}_n(M)$, $n = 0, 1, 2, \ldots$. In Fig. 1.5 we plot the MS estimation error as a function of the iteration step n (index of AV filter in the sequence or number of auxiliary vectors) for the case study in Fig. 1.4, for $M = 256$ (Part (a)) and $M = 2048$ (Part (b)). As a reference, we also include the MS-error of the constraint-LMS estimator [4], [35]

$$\hat{\mathbf{w}}_{\mathrm{LMS}}(m) = \left(\mathbf{I} - \frac{\mathbf{v}\mathbf{v}^H}{\|\mathbf{v}\|^2}\right)\left[\hat{\mathbf{w}}_{\mathrm{LMS}}(m-1) - \mu\mathbf{r}_m\mathbf{r}_m^H\hat{\mathbf{w}}_{\mathrm{LMS}}(m-1)\right] + \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v} \quad , \qquad (1.16)$$
$$m = 1, \ldots, M,$$

with $\hat{\mathbf{w}}_{\mathrm{LMS}}(0) = \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v}$ and some $\mu > 0$, and the RLS estimator [7], [8] with Matrix-Inversion-Lemma-based $\mathbf{R}^{-1}$ estimation:

$$\hat{\mathbf{R}}^{-1}(m) = \hat{\mathbf{R}}^{-1}(m-1) - \frac{\hat{\mathbf{R}}^{-1}(m-1)\mathbf{r}_m\mathbf{r}_m^H\hat{\mathbf{R}}^{-1}(m-1)}{1 + \mathbf{r}_m^H\hat{\mathbf{R}}^{-1}(m-1)\mathbf{r}_m} \quad , \quad m = 1, \ldots, M, \qquad (1.17)$$

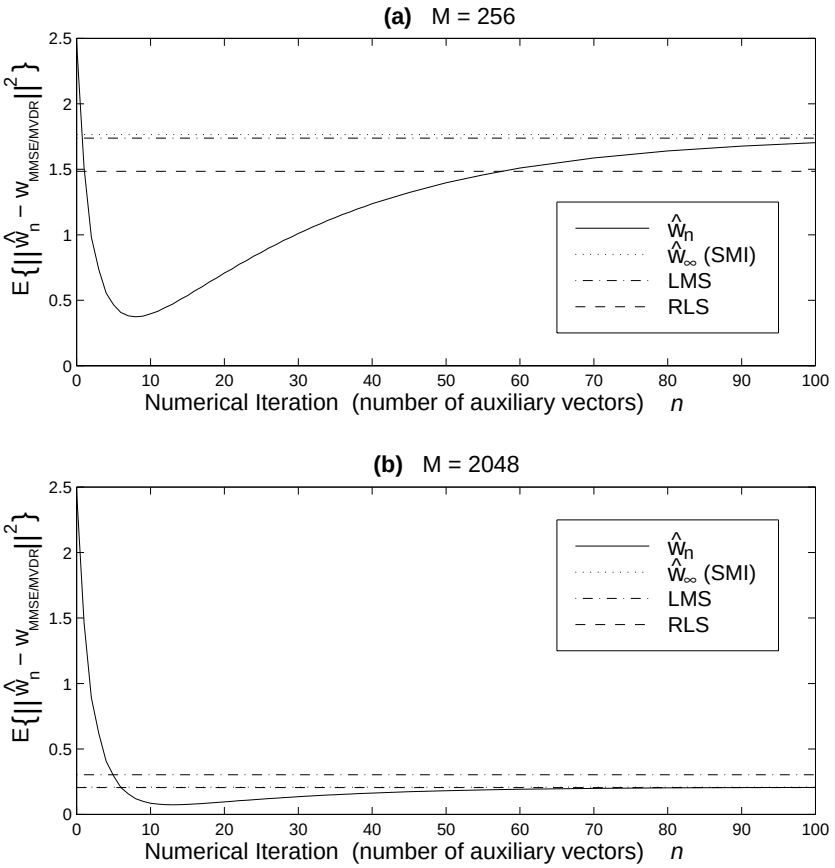

Fig. 1.5: MS estimation error for the sequence of estimators $\hat{\mathbf{w}}_n(M)$, $n = 0, 1, \ldots$. (a) Data record size $M = 256$. (b) $M = 2048$.

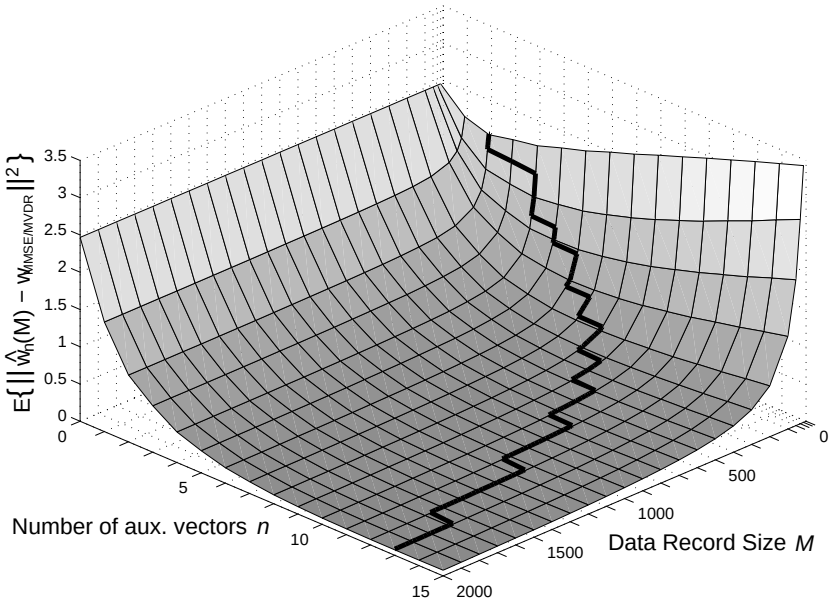

Fig. 1.6: MS estimation error versus number of auxiliary vectors n and sample support M.

with $\hat{\mathbf{R}}^{-1}(0) = \frac{1}{\epsilon_0}\mathbf{I}$ for some $\epsilon_0 > 0$. Theoretically, it is known that the LMS gain parameter $\mu > 0$ [36] has to be less than $\frac{1}{2\cdot\lambda_{\max}^{\text{blocked}}}$, where $\lambda_{\max}^{\text{blocked}}$ is the maximum eigenvalue of the "blocked-data" autocorrelation matrix $\left(\mathbf{I} - \frac{\mathbf{v}\mathbf{v}^H}{\|\mathbf{v}\|^2}\right)\mathbf{R}\left(\mathbf{I} - \frac{\mathbf{v}\mathbf{v}^H}{\|\mathbf{v}\|^2}\right)$. While this is a theoretical upper bound, practitioners are well aware that empirical, data-dependent "optimization" or "tuning" of the LMS gain $\mu > 0$ or the RLS initialization parameter $\epsilon_0 > 0$ [37] is necessary to achieve acceptable performance (in our study we set $\mu = \frac{1}{200\cdot\lambda_{\max}^{\text{blocked}}}$ and $\epsilon_0 = 20$, respectively). This data specific tuning frequently results in misleading, over-optimistic conclusions about the short-data-record performance of the LMS/RLS algorithms. In contrast, when the filter estimators $\hat{\mathbf{w}}_n$ generated by the algorithm of Fig. 1.2 are considered instead, tuning of the real-valued parameters μ and ϵ_0 is virtually replaced by an integer choice among the first several members of the $\{\hat{\mathbf{w}}_n\}$ sequence. Adaptive, data-dependent criteria for the selection of the most appropriate AV filter in the sequence for a given data record are developed in the next section. In Fig. 1.5(a), for $M = 256$ all estimators $\hat{\mathbf{w}}_n$ from $n = 2$ up to about $n = 55$ outperform in MS-error their RLS, LMS, and SMI ($\hat{\mathbf{w}}_\infty$) counterparts. $\hat{\mathbf{w}}_8$ ($n = 8$ auxiliary vectors) has the least MS-error of all (best bias/variance trade-off). When the data record size is increased to $M = 2048$ (Fig. 1.5(b)), we can afford more iterations (more auxiliary vectors) and $\hat{\mathbf{w}}_{13}$ offers the best bias/variance trade-off (lowest MS-error). All filter estimators $\hat{\mathbf{w}}_n$ for $n > 8$ outperform the LMS/RLS/SMI ($\hat{\mathbf{w}}_\infty$) estimators. For such large data record sets ($M = 2048$), the RLS and the SMI ($\hat{\mathbf{w}}_\infty$) MS-error are almost identical. Fig. 1.6 offers a 3-dimensional plot of the MS estimation error as a function of the number of auxiliary vectors n and the sample support M. The dark line that traces the bottom of the MS estimation error surface identifies the best number of auxiliary vectors for any given data record size M.

An alternative bias/variance trading mechanism through real-valued tuning is the diagonally-loaded (DL) SMI estimator [6]

$$\hat{\mathbf{w}}_{\text{DL-SMI}}(\Delta) = \rho^* \frac{\left[\hat{\mathbf{R}}(M) + \Delta\mathbf{I}\right]^{-1}\mathbf{v}}{\mathbf{v}^H\left[\hat{\mathbf{R}}(M) + \Delta\mathbf{I}\right]^{-1}\mathbf{v}} \tag{1.18}$$

where $\Delta \geq 0$ is the diagonal loading parameter. We observe that $\hat{\mathbf{w}}_{\text{DL-SMI}}(\Delta = 0)$ is the regular SMI estimator, while $\lim_{\Delta\to\infty} \hat{\mathbf{w}}_{\text{DL-SMI}}(\Delta) = \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v}$ which is the properly scaled matched filter. In Fig. 1.7(a) we plot the MS estimation error of the DL-SMI estimator as a function of the diagonal loading parameter Δ

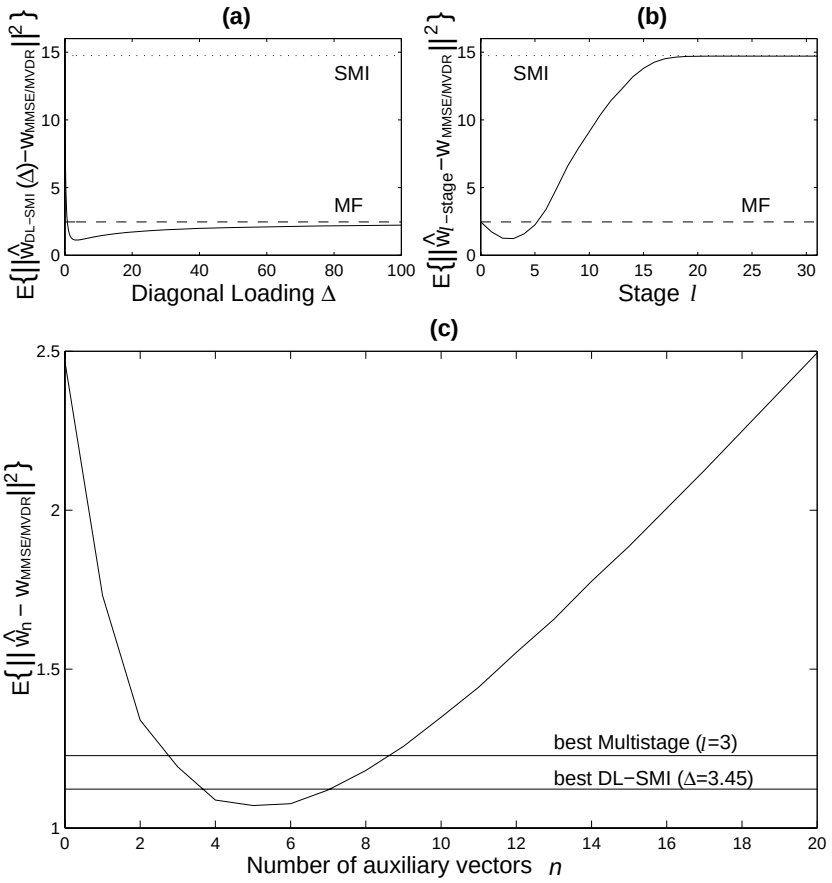

Fig. 1.7: MS estimation error studies for (a) diagonally loaded SMI, (b) multistage, and (c) auxiliary-vector estimators ($M = 60$).

($M = 60$). We identify the *best possible* diagonal loading value $\Delta \simeq 3.45$ (at significant computational cost) and in Fig. 1.7(c) we compare the best DL-SMI estimator against the AV estimator sequence for which *no* diagonal loading is performed. Interestingly, the AV estimators $\hat{\mathbf{w}}_n$ from $n = 4$ to 7 outperform in MS error the best possible DL-SMI estimator ($\Delta \simeq 3.45$).

Finally, a *finite set* of L filter estimators with varying bias/covariance balance can be obtained through the use of the orthogonal "multistage" filter decomposition procedure in [9], [10]. It can be shown theoretically that the l-stage filter, $\mathbf{w}_{l\text{-stage}}$, $0 \leq l \leq L-1$, is equivalent to the following structure. First, change the auxiliary-vector generation recursion in (1.9) or Fig. 1.2 to impose orthogonality not only with respect to the constraint vector $\mathbf{v}$ but also with respect to *all previously defined* auxiliary vectors $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{n-1}$, $n \leq L-1$:

$$\mathbf{y}_n = \left(\mathbf{I} - \frac{\mathbf{v}\mathbf{v}^H}{\|\mathbf{v}\|^2} - \sum_{i=1}^{n-1} \frac{\mathbf{y}_i \mathbf{y}_i^H}{\|\mathbf{y}_i\|^2} \right) \mathbf{R}\mathbf{w}_{n-1}. \tag{1.19}$$

Next, terminate the recursion at $n = l$, $0 \leq l \leq L-1$, and organize the l orthogonal to each other and to $\mathbf{v}$ vectors $\mathbf{y}_1, \ldots, \mathbf{y}_l$ in the form of a blocking matrix $\mathbf{B}_{L\times l} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l]$. Then,

$$\mathbf{w}_{l\text{-stage}} = \frac{\rho^*}{\|\mathbf{v}\|^2}\mathbf{v} - \mathbf{B}_{L\times l}\vec{\boldsymbol{\alpha}}_{l\times 1} \tag{1.20}$$

where

$$\vec{\boldsymbol{\alpha}} = \frac{\rho^*}{\|\mathbf{v}\|^2} \left[\mathbf{B}^H \mathbf{R} \mathbf{B}\right]^{-1} \mathbf{B}^H \mathbf{R}\mathbf{v} \tag{1.21}$$

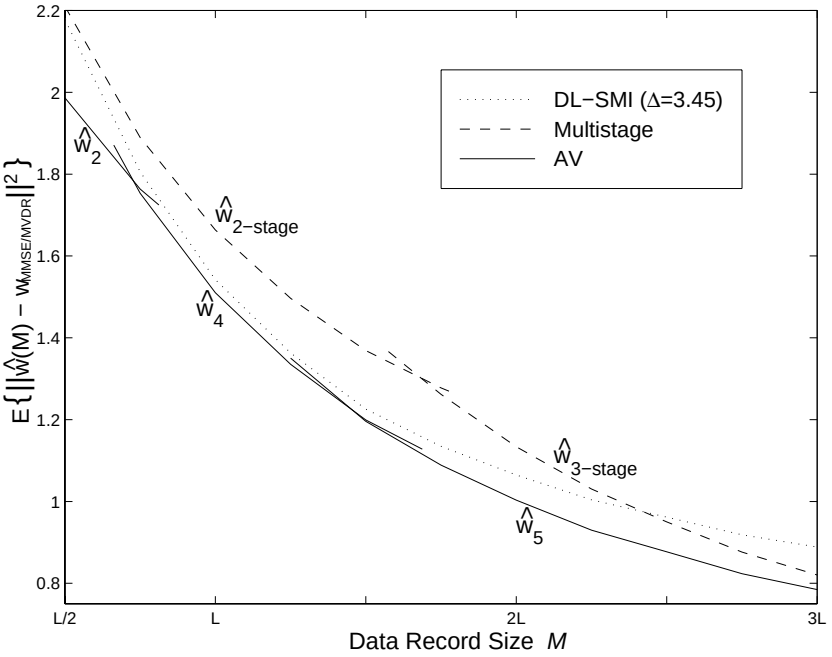

Fig. 1.8: MS estimation error for the *best* multistage and AV estimators over the data support range $M = \frac{L}{2} = 16$ to $M = 3L = 96$. The MS estimation error of the $\Delta = 3.45$ DL-SMI estimator is also included as a reference.

is the MS vector-optimum (unconditionally optimum) set of weights of the vectors $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l$.[2] In the context of MMSE/MVDR filter estimation from a data record of size M, $\hat{\mathbf{w}}_{0\text{-stage}}(M)$ is the matched filter and $\hat{\mathbf{w}}_{(L-1)\text{-stage}}(M)$ is the SMI estimator. In Fig. 1.7(b) we plot the MS estimation error of $\hat{\mathbf{w}}_{l\text{-stage}}(M)$ as a function of l, $0 \leq l \leq L-1 = 31$, $(M = 60)$. We identify the *best* multistage estimator ($l = 3$ stages) and in Fig. 1.7(c) we compare against the AV estimator sequence. We see that all AV estimators $\hat{\mathbf{w}}_n$ from $n = 3$ to 8 outperform in MS-error the best multistage estimator ($l = 3$ stages). Finally, as a last study, in Fig. 1.8 we plot the MS error of the $\Delta = 3.45$ DL-SMI estimator together with the MS error of the *best* multistage and AV estimators over the data support range $M = \frac{L}{2} = 16$ to $M = 3L = 96$.

1.4 How to Choose the Number of Auxiliary Vectors

In this section we present two data-driven rules for the selection of the number of auxiliary vectors n [39]. The first rule selects the AV filter estimator with n auxiliary vectors that has minimum cross-validated average filter output energy. The second selection rule is specific to BPSK communications receivers that employ a sign detector at the output of the linear auxiliary-vector filter. Details are given below.

[2] Therefore, the multistage filter in [9], [10] is identical to the filter '$\mathbf{w_B}$' as it appears in [22]-[24]. The multistage decomposition algorithm is a computationally efficient procedure for the calculation of this filter tailored to the particular structure of $\mathbf{B}^H\mathbf{R}\mathbf{B}$ (tri-diagonal matrix). The same computational savings can be achieved by the general forward calculation algorithm of Liu and Van Veen [38] that returns all intermediate stage filters along the way, up to the stage of interest l (total computational complexity of order $O\left((M+l)\,L^2\right)$). The AV algorithm in Fig. 1.2 has computational complexity $O\left((M+n)\,L^2\right)$ where n is the desired number of auxiliary vectors. Again, all intermediate AV filters are returned. Estimators of practical interest have $l \ll M$ or $n \ll M$. Therefore, the complexity of all such algorithms is dominated by $O\left(ML^2\right)$ which is required for the computation of $\hat{\mathbf{R}}(M)$.

1.4.1 Cross-Validated Minimum-Output-Variance Rule (CV-MOV)

Cross-validation is a well-known statistical method [40]. Here, we use cross-validation to select the filter parameter of interest (number of auxiliary vectors n) that minimizes the output variance which is estimated based on input observations that have not been used in the process of building the filter estimator itself.

A particular case of cross-validation that we use in this work is the "leave-one-out" method. The following criterion defines the CV-MOV AV filter estimator selection process.

Criterion 1.1 *For a given data record of size M, the cross-validated minimum-output-variance AV filter estimator selection rule chooses the AV filter estimator $\widehat{\mathbf{w}}_{n_1}(M)$ that minimizes the cross-validated sample average output variance, i.e.*

$$n_1 = \arg\min_n \left\{ \sum_{m=1}^{M} \widehat{\mathbf{w}}_n^H(M\backslash m)\mathbf{r}_m\mathbf{r}_m^H\widehat{\mathbf{w}}_n(M\backslash m) \right\} \tag{1.22}$$

where $(M\backslash m)$ identifies the AV filter estimator that is evaluated from the available data record after removing the m-th sample. □

It may be important to emphasize the need for invoking the cross-validation technique for the evaluation of the sample-average output variance. If sample-average evaluation using *all* data were attempted, then the selection rule would take the form $\min_n \left\{ \widehat{\mathbf{w}}_n^H(M)\widehat{\mathbf{R}}(M)\widehat{\mathbf{w}}_n(M) \right\}$, where $\widehat{\mathbf{R}}(M) = \frac{1}{M}\sum_{m=1}^{M}\mathbf{r}_m\mathbf{r}_m^H$ and $\widehat{\mathbf{w}}_n(M)$ is given by the algorithm of Fig. 1.2 with $\widehat{\mathbf{R}}(M)$ in place of $\mathbf{R}$. Such minimization, however, would result in $n = \infty$ since we know that $\widehat{\mathbf{w}}_n^H(M)\widehat{\mathbf{R}}(M)\widehat{\mathbf{w}}_n(M) \underset{n\to\infty}{\longrightarrow} \widehat{\mathbf{w}}_\infty^H(M)\widehat{\mathbf{R}}(M)\widehat{\mathbf{w}}_\infty(M)$ and the SMI estimator $\widehat{\mathbf{w}}_\infty(M)$ achieves minimum sample-output-variance $\widehat{\mathbf{w}}_\infty^H(M)\widehat{\mathbf{R}}(M)\widehat{\mathbf{w}}_\infty(M)$ (but not minimum true output variance $\widehat{\mathbf{w}}_\infty^H(M)\mathbf{R}\widehat{\mathbf{w}}_\infty(M)$, of course).

1.4.2 Output J-divergence Rule

For illustration purposes we reconsider the BPSK CDMA signal model example in (1.14). The output J-divergence rule selects the AV filter estimator from the sequence of AV estimators that maximizes the J-divergence of the Gaussian approximated conditional filter-output distributions (conditioned on the transmitted information bit $b_1 = +1$ or $b_1 = -1$). The appropriateness of such a criterion as well as implementation details are presented below.

For a given AV filter estimator $\widehat{\mathbf{w}}_n(M)$, we denote by p_m the real part[3] of the filter output with input the mth data vector $\mathbf{r}_m$, $m = 1, 2, \ldots, M$:

$$\begin{aligned} p_m &\triangleq Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}_m\right] \\ &= \sqrt{E_1}Re\left[b_1(m)\widehat{\mathbf{w}}_n^H(M)\mathbf{s}_1\right] + \sum_{k=2}^{K}\sqrt{E_k}Re\left[b_k(m)\widehat{\mathbf{w}}_n^H(M)\mathbf{s}_k\right] + Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{n}_m\right] \end{aligned} \tag{1.23}$$

where the information bits $b_k(m)$, $k = 1, 2, \ldots, K$, $m = 1, 2, \ldots, M$, are assumed to be independent identically distributed (i.i.d.) with equally probable outcomes and $\mathbf{n}_m$ is a 0-mean complex white Gaussian random vector with autocovariance matrix $\sigma^2\mathbf{I}$. Then, the scalars p_m, $m = 1, 2, \ldots, M$, are i.i.d. with common distribution $f_P(x)$ given by

$$f_P(x) = \frac{1}{2^K\sqrt{2\pi}\sigma}\sum_{i=1}^{2^K}\exp\left\{ \frac{-\left\{x - \sum_{k=1}^{K}\sqrt{E_k}Re\left[b_k^{(i)}\widehat{\mathbf{w}}_n^H(M)\mathbf{s}_k\right]\right\}^2}{2\sigma^2\left\|\widehat{\mathbf{w}}_n(M)\right\|^2} \right\} \tag{1.24}$$

where $b_k^{(i)}$, $i = 1, 2, \ldots, 2^K$, is the bit of user k in the ith bit-combination.

[3] While the signal model in (1.14) is real-valued, we choose to carry out this presentation in the more general context of complex input vectors and filters.

Conditioned on the transmitted information bit of the user of interest, user *1*, the pdf of the filter output is a mixture of 2^{K-1} Gaussian distributions. However, for "effective" interference suppressive filters we can safely approximate the conditional output distribution by a Gaussian distribution as argued in [41] for MMSE/MVDR linear filtering. Under this approximation, the filter output conditional distributions given that $+1$ or -1 is transmitted are $f_{1,n} \sim \mathcal{N}\left[\mu(n), \sigma^2_{I+N}(n)\right]$ and $f_{0,n} \sim \mathcal{N}\left[-\mu(n), \sigma^2_{I+N}(n)\right]$, respectively, where $\mu(n) \triangleq \sqrt{E_1} Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{s}_1\right]$ and $\sigma^2_{I+N}(n) \triangleq \sum_{k=2}^{K} E_k Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{s}_k\right]^2 + \sigma^2 \left\|\widehat{\mathbf{w}}_n(M)\right\|^2$ is the conditional variance due to multiple-access-interference and AWGN (the index "I+N" denotes comprehensively the disturbance contribution). The effect of the above approximation on the performance of the output J-divergence selection rule will be examined in Section V.

The J-divergence distance $J(f_{1,n}, f_{0,n})$ between the distributions $f_{1,n}(\cdot)$ and $f_{0,n}(\cdot)$ is defined as the sum of the Kullback-Leibler (K-L) distances between $f_{1,n}$ and $f_{0,n}$

$$J(f_{1,n}, f_{0,n}) \triangleq D(f_{1,n}, f_{0,n}) + D(f_{0,n}, f_{1,n}) \tag{1.25}$$

where the K-L distance of $f_{1,n}$ from $f_{0,n}$ is defined by $D(f_{1,n}, f_{0,n}) \triangleq \int_{-\infty}^{\infty} f_{1,n}(x) \log \frac{f_{1,n}(x)}{f_{0,n}(x)} dx$ [42]. Since $J(f_{1,n}, f_{0,n})$ is a function of the AV filter-estimator parameter n (number of auxiliary vectors), in the rest of this chapter we will use the notation $J(n)$ to represent the J-divergence distance between $f_{1,n}(x)$ and $f_{0,n}(x)$. For the Gaussian approximated pdf's $f_{1,n}$ and $f_{0,n}$, we have $D(f_{1,n}, f_{0,n}) = D(f_{0,n}, f_{1,n}) = \frac{[2\mu(n)]^2}{2\sigma^2_{I+N}(n)}$ and the J-divergence simplifies to

$$J(n) = \frac{4\mu^2(n)}{\sigma^2_{I+N}(n)}. \tag{1.26}$$

Expression (1.26) justifies our choice of the output J-divergence as one of the underlying rules for the selection of the best AV filter estimator. We recall that under the same Gaussian approximation of the conditional filter-output pdf's the filter output SINR can be expressed as $\frac{J(n)}{4}$ and, consequently, the BER as $Q(\frac{\sqrt{J(n)}}{2})$, where $Q(x) \triangleq \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du$. To this extent, maximization of the output J-divergence in (1.26) implies minimization of the BER. Therefore, we propose to select the estimator from the generated sequence of AV filter estimators that exhibits maximum *estimated* J-divergence.

1) Supervised Output J-divergence Rule

Exploiting the symmetry of $f_{1,n}(\cdot)$ and $f_{0,n}(\cdot)$, we can show in a straightforward manner that

$$J(n) = \frac{4\,E^2\left\{b_1 Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}\right]\right\}}{Var\left\{b_1 Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}\right]\right\}} \tag{1.27}$$

$$= \frac{4\left[\sum\limits_{i=\pm 1} E\left\{i \;\; Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}\right] | b_1 = i\right\} Pr(b_1 = i)\right]^2}{\sum\limits_{i=\pm 1} Var\left\{Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}\right] | b_1 = i\right\} Pr(b_1 = i)} \tag{1.28}$$

where $Var\{\cdot\}$ and $Pr\{\cdot\}$ denote variance and probability, respectively. Assuming availability of a pilot information bit sequence $\{b_1(m)\}_{m=1}^{M}$, we propose to estimate $J(n)$ by estimating statistical expectations and probabilities via sample averaging and frequencies of occurrence, respectively. We note that although (1.27) and (1.28) are ideally equivalent (when all statistical quantities are known), this is not the case in general when estimated measures are considered. So, let $\{p_1^+, p_2^+, p_3^+, \ldots, p_{M_1}^+\} \triangleq \left\{Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}_m\right] : b_1(m) = +1,\right.$ $\left. m = 1, 2, 3 \ldots, M\right\}$ and $\{p_1^-, p_2^-, p_3^-, \ldots, p_{M_2}^-\} \triangleq \left\{Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}_m\right] : b_1(m) = -1,\right.$ $\left. m = 1, 2, 3, \ldots, M\right\}$ be the sets of all filter outputs under $b_1(m) = +1$ and $b_1(m) = -1$, respectively ($M_1, M_2 \neq 0$ and $M_1 + M_2 = M$). First, it can be shown in a straightforward manner that the estimator of the numerator of (1.28) (that is, the weighted average of the sample average mean of the set $\{p_m^+\}_{m=1}^{M_1}$ and the sample average mean of the set $\{p_m^-\}_{m=1}^{M_2}$ weighted by the frequency of each set) is equivalent to the estimator of the numerator of (1.27) $4\widehat{\mu}^2(n)$ where $\widehat{\mu}(n) \triangleq \frac{\sum_{m=1}^{M_1+M_2} b_1(m) p_m}{M_1+M_2}$ (in fact, $\widehat{\mu}(n)$ is the minimum variance unbiased estimator of $\mu(n)$ [43, p. 178]).

Estimators of the denominator of (1.27) and (1.28) are examined in the following proposition. The proof is included in the Appendix.

Proposition 1.3 *Consider the estimator of $\sigma^2_{I+N}(n)$ which is the weighted average of the sample average variance of the set $\{p_m^+\}_{m=1}^{M_1}$ and the sample average variance of the set $\{p_m^-\}_{m=1}^{M_2}$, weighted by the frequency of occurrence of each set:*

$$\widehat{\sigma}_1^2(n) = \frac{M_1}{M_1+M_2}\ \widehat{\sigma}^{2+}(n) + \frac{M_2}{M_1+M_2}\ \widehat{\sigma}^{2-}(n) \tag{1.29}$$

where

$$\widehat{\sigma}^{2+}(n) \triangleq \frac{1}{M_1}\sum_{m=1}^{M_1}(p_m^+ - \widehat{\mu}^+(n))^2, \qquad \widehat{\mu}^+(n) \triangleq \tfrac{1}{M_1}\sum_{m=1}^{M_1} p_m^+,$$

$$\widehat{\sigma}^{2-}(n) \triangleq \frac{1}{M_2}\sum_{m=1}^{M_2}(p_m^+ - \widehat{\mu}^-(n))^2, \qquad \widehat{\mu}^-(n) \triangleq \tfrac{1}{M_2}\sum_{m=1}^{M_2} p_m^-,$$

and $\{p_m^+\}_{m=1}^{M_1}$, $\{p_m^-\}_{m=1}^{M_2}$, M_1 and M_2 are as defined previously. Consider also the direct sample average estimator of $\sigma^2_{I+N}(n)$:

$$\widehat{\sigma}_2^2(n) = \frac{1}{M_1+M_2}\sum_{m=1}^{M_1+M_2}[b_1(m)p_m - \widehat{\mu}(n)]^2 \tag{1.30}$$

where p_m is given by (1.23). The estimators $\widehat{\sigma}_1^2(n)$ and $\widehat{\sigma}_2^2(n)$ exhibit the following properties: (i) They are both biased and (ii) $\widehat{\sigma}_2^2(n)$ exhibits smaller MSE from the true value than $\widehat{\sigma}_1^2(n)$. □

Utilizing Proposition 1.3, estimators for the filter-output J-divergence become readily available. The following theorem identifies their relative merits. The proof is included in the Appendix.

Theorem 1.2 *Define the two supervised estimators of the output J-divergence $\widehat{J}_{S,1}(n) \triangleq \frac{4\hat{\mu}^2(n)}{\hat{\sigma}_1^2(n)}$ and $\widehat{J}_{S,2}(n) \triangleq \frac{4\hat{\mu}^2(n)}{\hat{\sigma}_2^2(n)}$, where the subscript "S" identifies a supervised implementation. Then, for a given information bit pilot sequence of size $M = M_1 + M_2$, where M_1 and M_2 are the cardinalities of the sets $\{b_1(m) = +1\}$ and $\{b_1(m) = -1\}$ respectively, both estimators are biased while the MSE of $\widehat{J}_{S,2}(n)$ is less than the MSE of $\widehat{J}_{S,1}(n)$, i.e.*

$$E\left\{\left[\widehat{J}_{S,2}(n) - \frac{4\mu^2(n)}{\sigma^2_{I+N}(n)}\right]^2\right\} < E\left\{\left[\widehat{J}_{S,1}(n) - \frac{4\mu^2(n)}{\sigma^2_{I+N}(n)}\right]^2\right\} \tag{1.31}$$

where $\frac{4\mu^2(n)}{\sigma^2_{I+N}(n)}$ is the true value of $J(n)$ as given by (1.26). □

Using the preferred estimator $\widehat{J}_{S,2}(n)$, the supervised implementation of the output J-divergence AV filter estimator selection rule takes the final form given by the following criterion.

Criterion 1.2 *For a given information bit pilot sequence of size M, the supervised J-divergence AV filter estimator selection rule chooses the estimator $\widehat{\mathbf{w}}_{n_2}(M)$ with n_2 auxiliary vectors where*

$$n_2 = \arg\max_n\left\{\widehat{J}_{S,2}(n)\right\} = \arg\max_n\left\{\frac{4\left[\frac{1}{M}\sum_{m=1}^{M} b_1(m)Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}_m\right]\right]^2}{\frac{1}{M}\sum_{m=1}^{M}\left[b_1(m)Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}_m\right] - \widehat{\mu}(n)\right]^2}\right\}. \tag{1.32}$$

□

2) Unsupervised (Blind) Output J-divergence Rule

The blind implementation of the rule is obtained by substituting the information bit b_1 in (1.27) by the detected bit $\widehat{b}_1 = \text{sgn}\left[Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}\right]\right]$ (output of the sign detector that follows the linear filter). In particular, using $\widehat{b}_1$ in place of b_1 in (1.27) we obtain the following J-divergence expression:

$$J_B(n) = \frac{4\,E^2\left\{\widehat{b}_1 Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}\right]\right\}}{Var\left\{\widehat{b}_1 Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}\right]\right\}} = \frac{4\,E^2\left\{\left|Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}\right]\right|\right\}}{Var\left\{\left|Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}\right]\right|\right\}} \tag{1.33}$$

where the subscript "B" identifies the blind version of the J-divergence function. The following proposition provides the conditions under which $J_B(n)$ is nearly equal to $J(n)$. The proof is included in the Appendix.

Proposition 1.4 *If* $\frac{\mu(n)}{\sigma_{I+N}(n)} >> 1$, *i.e. the filter output SINR is significantly higher than* $0dB$, *then* $J_B(n) \approx J(n)$. □

To estimate $J_B(n)$ from a data record of finite size, we substitute the statistical expectations in (1.33) by sample averages. The following criterion summarizes the corresponding AV filter estimator selection rule.

Criterion 1.3 *For a given data record of size* M, *the unsupervised (blind) J-divergence AV filter estimator selection rule chooses the estimator* $\widehat{\mathbf{w}}_{n_3}(M)$ *with* n_3 *auxiliary vectors where*

$$\begin{aligned} n_3 &= \arg\max_n \left\{ \widehat{J}_B(n) \right\} \\ &= \arg\max_n \left\{ \frac{4\left[\frac{1}{M}\sum_{m=1}^{M} \left| Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}_m\right]\right|\right]^2}{\frac{1}{M}\sum_{m=1}^{M} \left|Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}_m\right]\right|^2 - \left[\frac{1}{M}\sum_{m=1}^{M} \left|Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}_m\right]\right|\right]^2} \right\}. \end{aligned} \tag{1.34}$$

□

1.5 Simulation Studies

We examine the performance of the proposed short-data-record AV filter estimator selection rules for a DS-CDMA system with K users, spreading gain L, and multipath fading reception by a narrowband antenna array with N elements. All elements experience identical fading. Let J denote the number of chip-interval-spaced paths per baseband user signal. After conventional carrier demodulation, chip-matched filtering and sampling at the chip rate over a multipath extended symbol interval of $L+J-1$ chips, the $L+J-1$ data samples from the i-th antenna element, $i = 1, 2, \ldots, N$, are organized in the form of a vector $\mathbf{r}_m^{(i)}$ given by

$$\mathbf{r}_m^{(i)} = \sum_{k=1}^{K}\sum_{t=1}^{J} c_{k,t}\sqrt{E_k}\left(b_k(m)\mathbf{s}_{k,t} + b_k^-(m)\mathbf{s}_{k,t}^- + b_k^+(m)\mathbf{s}_{k,t}^+\right)a_{k,t}[i] + \mathbf{n}_m^{(i)}, \tag{1.35}$$

$$m = 1, \ldots, M, \quad i = 1, \ldots, N.$$

In (1.35), with respect to the k-th user signal, E_k is the transmitted energy, $b_k(m)$, $b_k^-(m)$, and $b_k^+(m)$ are the present, the previous, and the following transmitted bits, respectively, and $c_{k,t}$ is the coefficient of the t-th path of the k-th user signal. The channel coefficients are modeled as independent zero-mean complex Gaussian random variables that are assumed to remain constant over the filter adaptation data record of size M. $\mathbf{s}_{k,t}$ represents the $(J-1)$-zero-padded and $(t-1)$-right-shifted version of the signature of the k-th user $\mathbf{s}_k$, $\mathbf{s}_{k,t}^-$ is the 0-filled, L-left-shifted version of $\mathbf{s}_{k,t}$, and $\mathbf{s}_{k,t}^+$ is the 0-filled, L-right-shifted version of $\mathbf{s}_{k,t}$. Finally, $\mathbf{n}_m^{(i)}$ represents additive complex white Gaussian noise and $a_{k,t}[i]$ denotes the i-th coordinate of the array response vector $\mathbf{a}_{k,t}$ that corresponds to the t-th path of the k-th user signal:

$$a_{k,t}[i] = \exp\left\{ j2\pi(i-1)\frac{\sin\theta_{k,t}q}{\lambda} \right\}, \quad i = 1, \ldots, N, \tag{1.36}$$

where $\theta_{k,t}$ is the angle of arrival, λ is the carrier wavelength, and q is the inter-element spacing (in our studies we set $q \triangleq \lambda/2$).

We "vectorize" the $(L+J-1)\times N$ space-time received data matrix $[\mathbf{r}_m^{(1)}, \mathbf{r}_m^{(2)}, \ldots, \mathbf{r}_m^{(N)}]$ to form the joint space-time data vector $\mathbf{r}_m$, which is a $(L+J-1)N$-long column vector:

$$\mathbf{r}_m = Vec\{[\mathbf{r}_m^{(1)}, \mathbf{r}_m^{(2)}, \ldots, \mathbf{r}_m^{(N)}]_{(L+J-1)\times N}\}. \tag{1.37}$$

The joint space-time RAKE filter for user 1 is $\mathbf{v}_1 \triangleq E_{b_1}\{\mathbf{r}_m b_1(m)\} = Vec\{[\mathbf{v}_{1,1}, \mathbf{v}_{1,2}, \ldots, \mathbf{v}_{1,N}]\}$ ($E_{b_1}\{\cdot\}$ denotes statistical expectation with respect to $b_1(m)$), and $\mathbf{v}_{1,i} \triangleq \sum_{t=1}^{J} c_{1,t}\mathbf{s}_{1,t}a_{1,t}[i]$, $i = 1, 2, \ldots, N$. The

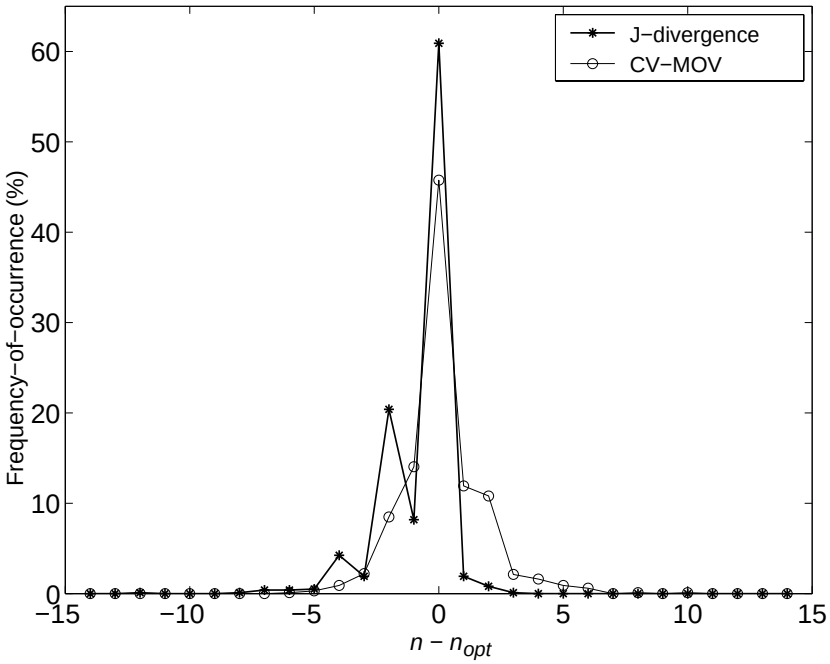

Fig. 1.9: Histogram of the two differences $n_1 - n_{opt}$ and $n_3 - n_{opt}$ where n_1 is the CV-MOV choice and n_3 is the blind J-divergence choice ($M = 230$, $\text{SNR}_1 = 8\text{dB}$).

MMSE/MVDR filter is built with constraint vector $\mathbf{v} = \mathbf{v}_1$, desired response $\mathbf{w}^H \mathbf{v}_1 = 1$, and autocorrelation matrix $\mathbf{R} = E\{\mathbf{r}_m \mathbf{r}_m^H\}$.

We choose $K = 20$, $N = 5$, $J = 3$ paths with independent zero-mean complex Gaussian fading coefficients of variance one (i.e., $E\{|c_{k,t}|^2\} = 1$), and Gold signatures with processing gain $L = 31$. The total SNR's (over the three paths) of the 19 interferers are set at $\text{SNR}_{2-6} = 6\text{dB}$, $\text{SNR}_{7-8} = 7\text{dB}$, $\text{SNR}_{9-13} = 8\text{dB}$, $\text{SNR}_{14-15} = 9\text{dB}$, $\text{SNR}_{16-20} = 10\text{dB}$. The space-time product (filter length) equals $(L + J - 1)N = (31 + 2)5 = 165$. All experimental results that follow are averages over 100 different channel realizations and 10 independent data record generations per channel.

We first examine the performance of the AV filter estimator selection rules under the assumption that no info-bit pilot sequence is available. The data record size is set equal to $M = 230$ while the total SNR of the user of interest is set at $\text{SNR}_1 = 8\text{dB}$. In Fig. 1.9, we plot the empirical pdf of the differences $(n_1 - n_{opt})$ and $(n_3 - n_{opt})$ where n_1 and n_3 denote selections according to Criterion 1.1 and 1.3, respectively, while n_{opt} denotes the "genie" maximum SINR optimum choice of the number of auxiliary vectors. We observe that both criteria provide a reliable estimate of the "genie" assisted optimum number of auxiliary vectors.

The overall short-data-record adaptive filter performance is examined in Figs. 10 and 11. In Fig. 1.10, we plot the BER[4] of the AV filter estimators $\widehat{\mathbf{w}}_{n_1}(M)$ and $\widehat{\mathbf{w}}_{n_3}(M)$ as a function of the SNR of the user of interest for data records of size $M = 230$. The BER curve of the "genie" assisted BER optimum filter choice $\widehat{\mathbf{w}}_{n_{opt}}(M)$ as well as the corresponding curves of the ideal MMSE/MVDR filter $\mathbf{w}_{\text{MMSE/MVDR}}$, the SMI filter estimator $\widehat{\mathbf{w}}_{\infty}(M)$, the S-T RAKE matched-filter (MF) $\widehat{\mathbf{w}}_0(M) = \mathbf{v}_1$, and the multistage filter [9], [10] with the preferred number of stages[5] $l = 7$ are also included for comparison purposes. We observe that both $\widehat{\mathbf{w}}_{n_1}(M)$ and $\widehat{\mathbf{w}}_{n_3}(M)$ are very close to the "genie" optimum AV filter estimator choice and outperform significantly the SMI filter estimator, the multistage filter estimator, and the matched filter. We also observe

[4]The BER of each filter under consideration is approximated by $Q\left(\sqrt{\text{SINR}_{\text{out}}}\right)$ [41], since the computational complexity of the BER expression for this antenna array CDMA system prohibits exact analytic evaluation.

[5]In [44] it is argued that $l = 7$ ($D = 8$ in the notation of [44]) stages is "nearly optimal over a wide range of loads and SNRs".

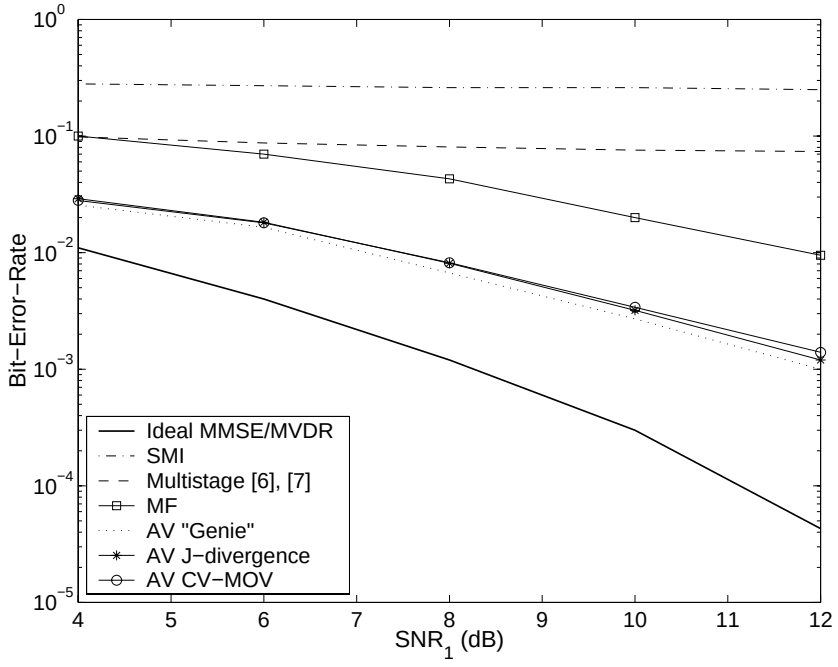

Fig. 1.10: BER versus SNR for the user signal of interest ($M = 230$).

that for moderate to high SNR's of the user of interest, the J-divergence selection rule is slightly superior to the CV-MOV selection rule. The opposite is true in the low SNR range. This is explained by the fact that the J-divergence approximation $J(n) \approx J_B(n)$ used in Proposition 1.4 is less accurate for low filter output SINR values. On the other hand, for high filter output SINR's the discrimination capability of the CV-MOV rule is not as sharp.

Finally, Fig. 1.11 repeats the study of Fig. 1.10 as a function of the data record size. The SNR of the user of interest is fixed at 8dB.

1.6 Concluding Remarks

In this work we relied strictly on statistical *conditional* optimization principles to derive an iterative algorithm that starts from the "white-noise matched filter" and converges to the "MMSE/MVDR filter" solution for any given positive definite input autocorrelation matrix. The conceptual simplicity of the employed conditional optimization criteria led to a computationally simple iteration step. We analyzed basic algorithmic properties and we established formal convergence to the MMSE/MVDR filter.

When the input autocorrelation matrix is substituted by a sample-average (positive definite) estimate, the algorithm generates a sequence of filter estimators that converges to the familiar sample-matrix-inversion (SMI) unbiased estimator. The bias of the generated estimator sequence decreases rapidly to zero while the estimator covariance trace rises slowly from zero (for the initial, fixed-valued, matched-filter estimator) to the asymptotic covariance trace of SMI. Sequences of practical estimators that offer such exceptional control over favorable bias/covariance balance points are always a prime objective in the estimation theory literature. Indeed, for finite data record sets, members of the generated sequence of estimators were seen to outperform in MS estimation error LMS/RLS-type, SMI and diagonally loaded SMI, and orthogonal multistage decomposition filter estimators. In addition, the troublesome, data-dependent tuning of the real-valued LMS learning gain parameter, the RLS initialization parameter or the SMI diagonal loading parameter is replaced by an integer choice among the first several members of the estimator sequence. Two data-driven criteria were proposed for the identification of the best AV filter estimator in the sequence.

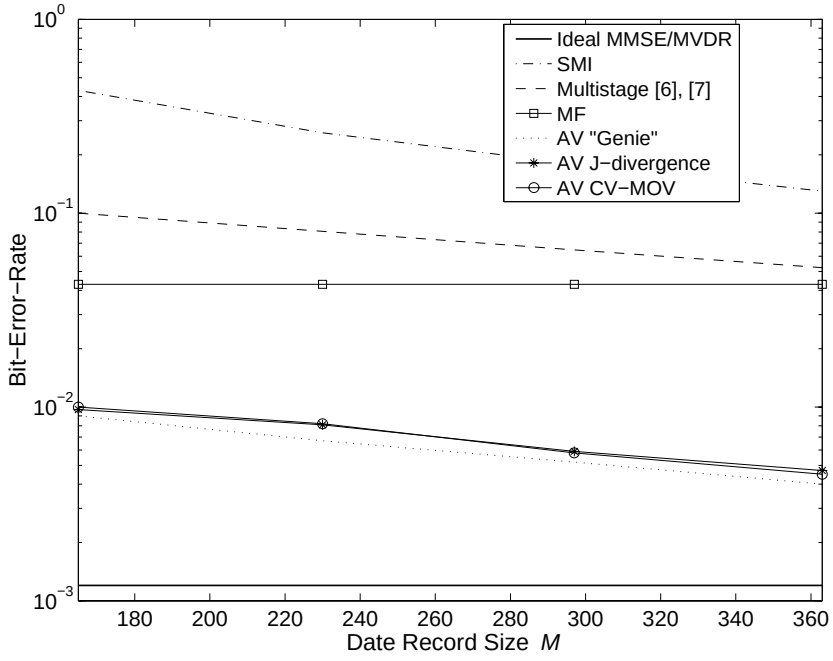

Fig. 1.11: BER versus data record size ($\mathrm{SNR}_1 = 8\mathrm{dB}$).

The first criterion calls for the minimization of the cross-validated filter-estimator output variance. The second criterion calls for the maximization of the J-divergence of the filter-estimator-output conditional distributions. Simulation studies examined and compared the operational characteristics of the proposed selection methods. With respect to the relative merits of the minimum cross-validated output variance and the maximum output J-divergence selection rules, we observed that for moderate to high output SINR's the latter method appears superior to the former (for high SINR's the cross-validated minimum output variance rule is not as sharp in discrimination ability). In contrast, in low output SINR the J-divergence method is somewhat lacking in performance (technically, the approximation in Proposition 1.4 is less accurate for near 0dB or lower output SINR values). As a final general comment, the use of a sufficiently long antenna array in combination with an "effective" interference suppressive filter can result in high output SINR which favors the J-divergence selection rule, even when the *transmitted* energy of the user of interest is much lower than that of the interferers.

The "auxiliary-vector" algorithm in Fig. 1.2 together with Criteria 1, 2, and 3 form a complete toolbox for state-of-the-art estimation of MMSE/MVDR filters. The developments are of particular interest in high-dimensional adaptive signal processing applications that rely on data records of limited size.

1.7 Appendix

Signature Assignment for the DS/CDMA Example of Sections III and IV

The matrix $\mathbf{S}_{32\times 13} = [\mathbf{s}_1\mathbf{s}_2\cdots\mathbf{s}_{13}]$ with columns the signature vectors $\mathbf{s}_1$, $\mathbf{s}_2$, ..., $\mathbf{s}_{13}$ is given below.

$$\mathbf{S} = \frac{1}{\sqrt{32}}\begin{bmatrix}
-1 & +1 & -1 & -1 & -1 & -1 & +1 & -1 & -1 & +1 & +1 & -1 & +1\\
-1 & -1 & -1 & -1 & +1 & +1 & +1 & -1 & -1 & +1 & -1 & -1 & +1\\
+1 & -1 & +1 & +1 & -1 & +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1\\
+1 & -1 & +1 & +1 & -1 & +1 & +1 & +1 & +1 & +1 & -1 & +1 & +1\\
+1 & +1 & +1 & +1 & -1 & -1 & +1 & -1 & +1 & -1 & -1 & -1 & +1\\
-1 & +1 & +1 & +1 & +1 & +1 & -1 & -1 & -1 & +1 & -1 & -1 & -1\\
+1 & -1 & -1 & -1 & -1 & -1 & +1 & +1 & -1 & +1 & -1 & -1 & -1\\
+1 & -1 & +1 & -1 & -1 & -1 & +1 & -1 & +1 & -1 & -1 & -1 & +1\\
+1 & -1 & +1 & +1 & +1 & -1 & +1 & -1 & -1 & -1 & -1 & -1 & -1\\
+1 & -1 & +1 & -1 & -1 & -1 & +1 & +1 & -1 & -1 & -1 & -1 & +1\\
-1 & +1 & -1 & +1 & -1 & +1 & +1 & +1 & -1 & -1 & +1 & -1 & -1\\
-1 & +1 & +1 & +1 & +1 & -1 & -1 & +1 & -1 & +1 & +1 & +1 & +1\\
-1 & -1 & -1 & -1 & -1 & -1 & -1 & +1 & -1 & -1 & +1 & +1 & +1\\
+1 & -1 & -1 & +1 & +1 & -1 & +1 & +1 & +1 & +1 & -1 & +1 & -1\\
-1 & -1 & -1 & -1 & -1 & -1 & +1 & -1 & -1 & -1 & +1 & +1 & +1\\
-1 & +1 & -1 & +1 & -1 & +1 & -1 & +1 & +1 & -1 & +1 & +1 & +1\\
-1 & -1 & -1 & -1 & -1 & +1 & +1 & -1 & +1 & -1 & +1 & +1 & +1\\
+1 & +1 & -1 & -1 & -1 & -1 & +1 & -1 & -1 & -1 & -1 & -1 & +1\\
+1 & -1 & -1 & +1 & -1 & +1 & +1 & +1 & +1 & +1 & -1 & -1 & +1\\
+1 & +1 & +1 & +1 & -1 & +1 & -1 & +1 & -1 & +1 & -1 & +1 & +1\\
-1 & +1 & -1 & +1 & +1 & -1 & +1 & -1 & +1 & +1 & +1 & -1 & -1\\
+1 & +1 & -1 & -1 & -1 & -1 & +1 & -1 & +1 & +1 & -1 & +1 & +1\\
+1 & +1 & -1 & +1 & +1 & +1 & -1 & +1 & +1 & +1 & -1 & +1 & +1\\
-1 & -1 & +1 & +1 & +1 & +1 & -1 & -1 & -1 & +1 & +1 & -1 & +1\\
-1 & +1 & -1 & +1 & +1 & -1 & +1 & -1 & -1 & +1 & -1 & +1 & +1\\
+1 & +1 & +1 & +1 & +1 & -1 & -1 & +1 & +1 & +1 & -1 & +1 & -1\\
+1 & -1 & -1 & -1 & -1 & +1 & -1 & +1 & +1 & +1 & +1 & -1 & -1\\
-1 & +1 & -1 & -1 & -1 & +1 & +1 & +1 & -1 & +1 & +1 & +1 & -1\\
+1 & -1 & +1 & -1 & -1 & -1 & -1 & -1 & -1 & +1 & +1 & +1 & -1\\
-1 & +1 & +1 & -1 & +1 & +1 & -1 & +1 & -1 & -1 & +1 & +1 & -1\\
+1 & -1 & -1 & +1 & -1 & +1 & +1 & -1 & +1 & -1 & -1 & +1 & -1\\
-1 & +1 & +1 & -1 & -1 & -1 & -1 & -1 & -1 & +1 & +1 & +1 & +1
\end{bmatrix}. \tag{1.38}$$

Proof of Proposition 1.3
The quantities $\frac{1}{M_1}\sum_{m=1}^{M_1}\left(p_m^+ - \widehat{\mu}^+(n)\right)^2$ and $\frac{1}{M_2}\sum_{m=1}^{M_2}\left(p_m^- - \widehat{\mu}^-(n)\right)^2$ are the ML estimators for the variance of the filter output conditioned on $b_1 = +1$ and $b_1 = -1$, respectively [43, p. 179]. Both estimators are biased. (In fact, their unbiased counterparts that have multiplying factors $\frac{1}{M_1-1}$ and $\frac{1}{M_2-1}$ instead of $\frac{1}{M_1}$ and $\frac{1}{M_2}$, respectively, exhibit higher MSE). The MSE's of the estimators of interest $\widehat{\sigma}_1^2(n)$ and $\widehat{\sigma}_2^2(n)$ are as follows:

$$\begin{aligned}
\mathrm{MSE}_{\widehat{\sigma}_1^2(n)} =& E\left\{\left(\widehat{\sigma}_1^2(n) - \sigma_{I+N}^2(n)\right)^2\right\}\\
=& E\left\{\left[\left(\frac{\sum_{m=1}^{M_1}(p_m^+ - \widehat{\mu}^+(n))^2}{M_1+M_2} - \frac{M_1\sigma_{I+N}^2(n)}{M_1+M_2}\right)\right.\right.\\
&\left.\left. + \left(\frac{\sum_{m=1}^{M_2}(p_m^- - \widehat{\mu}^-(n))^2}{M_1+M_2} - \frac{M_2\sigma_{I+N}^2(n)}{M_1+M_2}\right)\right]^2\right\}\\
=& \frac{(2M_1-1)\,\sigma_{I+N}^4(n)}{(M_1+M_2)^2} + \frac{(2M_2-1)\,\sigma_{I+N}^4(n)}{(M_1+M_2)^2} + \frac{2\sigma_{I+N}^4(n)}{(M_1+M_2)^2}\\
=& \frac{2\sigma_{I+N}^4(n)}{M_1+M_2}
\end{aligned} \tag{1.39}$$

and

$$\mathrm{MSE}_{\widehat{\sigma}_2^2(n)} = E\left\{\left(\widehat{\sigma}_2^2(n) - \sigma_{I+N}^2(n)\right)^2\right\} = \frac{[2(M_1+M_2)-1]\,\sigma_{I+N}^4(n)}{(M_1+M_2)^2}. \tag{1.40}$$

Thus,

$$\mathrm{MSE}_{\widehat{\sigma}_1^2(n)} - \mathrm{MSE}_{\widehat{\sigma}_2^2(n)} = \frac{\sigma_{I+N}^4(n)}{(M_1+M_2)^2} > 0. \tag{1.41}$$

Proof of Theorem 1.2
$\widehat{\mu}^+(n)$ and $\widehat{\sigma}^{2+}(n)$ are independent random variables with distributions $\widehat{\mu}^+(n) \sim \mathcal{N}\left(\mu(n), \frac{\sigma_{I+N}^2(n)}{M_1}\right)$ and $\frac{M_1}{\sigma_{I+N}^2(n)}\widehat{\sigma}^{2+}(n) \sim \chi_{M_1-1}^2$ [45] ($\chi_{M_1-1}^2$ denotes the chi-square distribution with (M_1-1) degrees of freedom).

Similarly, $\widehat{\mu}^-(n)$ and $\widehat{\sigma}^{2-}(n)$ are independent random variables with distributions $\widehat{\mu}^-(n) \sim \mathcal{N}\left(-\mu(n), \frac{\sigma^2_{I+N}(n)}{M_2}\right)$ and $\frac{M_2}{\sigma^2_{I+N}(n)}\widehat{\sigma}^{2-}(n) \sim \chi^2_{M_2-1}$. Furthermore, $(\widehat{\mu}^+(n)\ ,\ \widehat{\sigma}^{2+}(n))$ and $(\widehat{\mu}^-(n), \widehat{\sigma}^{2-}(n))$ are mutually independent because $(\widehat{\mu}^+(n), \widehat{\sigma}^{2+}(n))$ and $(\widehat{\mu}^-(n), \widehat{\sigma}^{2-}(n))$ are evaluated from two independent training sets.

Since $\widehat{\mu}(n) = \frac{M_1}{M}\widehat{\mu}^+(n) - \frac{M_2}{M}\widehat{\mu}^-(n)$ and $\widehat{\sigma}^2_1(n) = \frac{M_1}{M}\widehat{\sigma}^{2+}(n) + \frac{M_2}{M}\widehat{\sigma}^{2-}(n)$, $\widehat{\mu}(n)$ and $\widehat{\sigma}^2_1(n)$ are independent random variables with distributions $\widehat{\mu}(n) \sim \mathcal{N}\left(\mu(n), \frac{\sigma^2_{I+N}(n)}{M}\right)$ and $\frac{M}{\sigma^2_{I+N}(n)}\widehat{\sigma}^2_1(n) \sim \chi^2_{M-2}$ where $M = M_1 + M_2$ is the size of the given information bit pilot sequence. Therefore,

$$E\{\widehat{J}_{S,1}(n)\} = 4\,E\{\widehat{\mu}^2(n)\}E\left\{\frac{1}{\widehat{\sigma}^2_1(n)}\right\} \tag{1.42}$$

$$= 4\left[\frac{\sigma^2_{I+N}(n)}{M} + \mu^2(n)\right]\frac{M}{\sigma^2_{I+N}(n)\,(M-4)} = \frac{4\mu^2(n)}{\sigma^2_{I+N}(n)}\ \ \frac{M}{M-4} + \frac{4}{M-4},$$

$$E\{\widehat{J}^2_{S,1}(n)\} = 16\,E\{\widehat{\mu}^4(n)\}E\left\{\frac{1}{\widehat{\sigma}^4_1(n)}\right\} \tag{1.43}$$

$$= 16\left[\mu^4(n) + \frac{6\mu^2(n)\sigma^2_{I+N}(n)}{M} + \frac{3\sigma^4_{I+N}(n)}{M^2}\right]\frac{M^2}{\sigma^4_{I+N}(n)\,(M-4)\,(M-6)}.$$

The inverse moments $E\left\{\frac{1}{\widehat{\sigma}^2_1(n)}\right\}$ and $E\left\{\frac{1}{\widehat{\sigma}^4_1(n)}\right\}$ exist for $M > 4$ and $M > 6$, respectively, and are given by [46]

$$E\left\{\frac{1}{\widehat{\sigma}^2_1(n)}\right\} = \frac{M}{\sigma^2_{I+N}(n)\,(M-4)}, E\left\{\frac{1}{\widehat{\sigma}^4_1(n)}\right\} \qquad = \frac{M^2}{\sigma^4_{I+N}(n)\,(M-4)\,(M-6)}. \tag{1.44}$$

Thus, the MSE of $\widehat{J}_{S,1}(n)$ is

$$\mathrm{MSE}_1 \triangleq E\left\{\left[\widehat{J}_{S,1}(n) - \frac{4\mu^2(n)}{\sigma^2_{I+N}(n)}\right]^2\right\} = \frac{16}{(M-4)\,(M-6)}a - \frac{16}{M-4}b + \frac{16\mu^4(n)}{\sigma^4_{I+N}(n)} \tag{1.45}$$

where

$$\begin{aligned} a &\triangleq \left[\mu^4(n) + \frac{6\mu^2(n)\sigma^2_{I+N}(n)}{M} + \frac{3\sigma^4_{I+N}(n)}{M^2}\right]\frac{M^2}{\sigma^4_{I+N}(n)} \quad \text{and} \\ b &\triangleq \left[\frac{\sigma^2_{I+N}(n)}{M} + \mu^2(n)\right]\frac{2M\mu^2(n)}{\sigma^4_{I+N}(n)}. \end{aligned} \tag{1.46}$$

On the other hand, $\widehat{\mu}(n)$ and $\widehat{\sigma}^2_2(n)$ are independent random variables with distributions $\widehat{\mu}(n) \sim \mathcal{N}\left(\mu(n), \frac{\sigma^2_{I+N}(n)}{M}\right)$ and $\frac{M}{\sigma^2_{I+N}(n)}\widehat{\sigma}^2_2(n) \sim \chi^2_{M-1}$, respectively. Therefore,

$$E\{\widehat{J}_{S,2}(n)\} = 4\left[\frac{\sigma^2_{I+N}(n)}{M} + \mu^2(n)\right]\frac{M}{\sigma^2_{I+N}(n)\,(M-3)} = \frac{4\mu^2(n)}{\sigma^2_{I+N}(n)}\ \ \frac{M}{M-3} + \frac{4}{M-3}, \tag{1.47}$$

$$E\{\widehat{J}^2_{S,2}(n)\} = 16\left[\mu^4(n) + \frac{6\mu^2(n)\sigma^2_{I+N}(n)}{M} + \frac{3\sigma^4_{I+N}(n)}{M^2}\right]\frac{M^2}{\sigma^4_{I+N}(n)\,(M-3)\,(M-5)}. \tag{1.48}$$

Thus, the MSE of $\widehat{J}_{S,2}(n)$ is

$$\mathrm{MSE}_2 \triangleq E\left\{\left[\widehat{J}_{S,2}(n) - \frac{4\mu^2(n)}{\sigma^2_{I+N}(n)}\right]^2\right\} = \frac{16}{(M-3)\,(M-5)}a - \frac{16}{M-3}b + \frac{16\mu^4(n)}{\sigma^4_{I+N}(n)} \tag{1.49}$$

where the first and second inverse moments $E\left\{\frac{1}{\widehat{\sigma}^2_2(n)}\right\}$ and $E\left\{\frac{1}{\widehat{\sigma}^4_2(n)}\right\}$ exist for $M > 3$ and $M > 5$, respectively, and a and b are given by (1.46).

From (1.45) and (1.49), assuming $M > 6$ which is the condition for all inverse moments to exist, we obtain

$$\begin{aligned}
\mathrm{MSE}_1 - \mathrm{MSE}_2 &= \frac{16a}{(M-4)(M-6)} - \frac{16a}{(M-3)(M-5)} + \frac{b}{M-3} - \frac{b}{M-4} \\
&= \frac{16}{(M-3)(M-4)} \left[\frac{2M-9}{(M-5)(M-6)} a - b \right] \\
&> \frac{16}{(M-3)(M-4)} \left[\frac{2(M-5)}{(M-5)(M-6)} a - b \right] \\
&= \frac{32\left[6M\mu^4(n) + 5M\mu^2(n)\sigma^2_{I+N}(n) + 6\mu^2(n)\sigma^2_{I+N}(n) + 3\sigma^4_{I+N}(n)\right]}{(M-3)(M-4)(M-6)\sigma^4_{I+N}(n)} \\
&> 0.
\end{aligned} \tag{1.50}$$

Proof of Proposition 1.4
Define $Y \triangleq \left| Re\left[\widehat{\mathbf{w}}_n^H(M)\mathbf{r}\right]\right|$. The pdf of Y is

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_{I+N}(n)} \left[\exp \frac{-(y-\mu(n))^2}{2\sigma^2_{I+N}(n)} + \exp \frac{-(y+\mu(n))^2}{2\sigma^2_{I+N}(n)} \right] U(y) \tag{1.51}$$

where $U(y)$ is the unit step function. The mean of Y is

$$E\{Y\} = \int y f_Y(y) dy = \mu(n) + \frac{2\sigma_{I+N}(n)}{\sqrt{2\pi}} \exp\left(-\frac{\mu^2(n)}{2\sigma^2_{I+N}} \right) - 2\mu(n) Q\left(\frac{\mu(n)}{\sigma_{I+N}(n)} \right). \tag{1.52}$$

The series expansion of $Q(x)$ is [47]

$$Q(x) = \frac{1}{\sqrt{2\pi}x} \exp\left(-\frac{x^2}{2}\right) \left\{ 1 - \frac{1}{x^2} + \frac{1\cdot 3}{x^4} + \cdots + \frac{(-1)^n 1\cdot 3\cdots(2n-1)}{x^{2n}} \right\} + R_n \tag{1.53}$$

where $R_n = (-1)^{n+1} 1\cdot 3\cdots(2n+1) \int_x^\infty \frac{1}{\sqrt{2\pi}t^{2n+2}} \exp\left(-\frac{t^2}{2}\right) dt$ is the remainder which is always less (in absolute value) than the first neglected term. If $\frac{\mu(n)}{\sigma_{I+N}(n)} >> 1$, (1.52) can be written as

$$\begin{aligned}
E\{Y\} =& \mu(n) \left\{ 1 + \frac{2\sigma_{I+N}(n)}{\sqrt{2\pi}\mu(n)} \exp\left(-\frac{\mu^2(n)}{2\sigma^2_{I+N}(n)} \right) \right. \\
& \left. - \frac{2\sigma_{I+N}(n)}{\sqrt{2\pi}\mu(n)} \exp\left(-\frac{\mu^2(n)}{2\sigma^2_{I+N}(n)} \right) \left[1 - O\left[\left(\frac{\mu(n)}{\sigma_{I+N}(n)} \right)^{-2} \right] \right] \right\} \\
=& \mu(n) \left\{ 1 + \frac{2}{\sqrt{2\pi}} \exp\left(-\frac{\mu^2(n)}{2\sigma^2_{I+N}(n)} \right) O\left[\left(\frac{\mu(n)}{\sigma_{I+N}(n)} \right)^{-3} \right] \right\} \\
\approx& \mu(n).
\end{aligned} \tag{1.54}$$

The variance of Y is

$$Var\{Y\} = E\{Y^2\} - E^2\{Y\} = \mu^2(n) + \sigma^2_{I+N}(n) - E^2\{Y\} \tag{1.55}$$

and using (1.54) we may approximate

$$Var\{Y\} \approx \sigma^2_{I+N}(n). \tag{1.56}$$

Finally, from (1.54) and (1.56) we obtain

$$J_B(n) = \frac{4E^2\{Y\}}{Var\{Y\}} \approx \frac{4\mu^2(n)}{\sigma^2_{I+N}(n)} = J(n). \tag{1.57}$$

Bibliography

[1] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings IEEE*, vol. 57, pp. 1408-1418, Aug. 1969.

[2] N. L. Owsley, "A recent trend in adaptive spatial processing for sensor arrays: Constraint adaptation," in *Signal Processing*, J. W. R. Griffiths et al., ed., pp. 591-604. New York, NY: Academic Press, 1973.

[3] N. E. Nahi, *Estimation Theory and Applications*. Huntington, NY: R. E. Krieger Publishing Co., 1976.

[4] B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode. "Adaptive antenna systems," *Proceedings IEEE*, vol. 55, pp. 2143-2158, Dec. 1967.

[5] I. S. Reed, J. D. Mallet, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerospace and Electr. Syst.*, vol. 10, pp. 853-863, Nov. 1974.

[6] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerospace and Electr. Syst.*, vol. 24, pp. 397-401, July 1988.

[7] R. L. Plackett, "Some theorems in least squares," *Biometrika*, vol. 37, p. 149, 1950.

[8] R. A. Wiggins and E. A. Robinson, "Recursive solution to the multichannel filtering problem," *Journ. Geophys. Res.*, vol. 70, pp. 1885-1891, 1965.

[9] J. S. Goldstein, I. S. Reed, P. A. Zulch, and W. L. Melvin, "A multistage STAP CFAR detection technique," in *Proc. IEEE Radar Conf.*, Dallas, TX, May 1998, pp. 111-116.

[10] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2943-2959, Nov. 1998.

[11] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagation*, vol. 30, pp. 27-34, Jan. 1982.

[12] B. D. Van Veen and R. A. Roberts, "Partially adaptive beamformer design via output power minimization," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 35, pp. 1524-1532, Nov. 1987.

[13] L. L. Scharf and D. W. Tufts, "Reduced rank methods for modeling stationary signals," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 35, pp. 350-354, Mar. 1987.

[14] B. D. Van Veen, "Eigenstructure based partially adaptive array design," *IEEE Trans. Antennas Propagation*, vol. 36, pp. 357-362, Mar. 1988.

[15] A. M. Haimovich and Y. Bar-Ness, "An eigenanalysis interference canceler," *IEEE Trans. Signal Proc.*, vol. 39, pp. 76-84, Jan. 1991.

[16] K. A. Byerly and R. A. Roberts, "Output power based partially adaptive array design," in *Proc. Asilomar Conf. Signals, Syst., Computers*, Pacific Grove, CA, 1989, pp. 576-580.

[17] J. S. Goldstein and I. S. Reed, "Reduced-rank adaptive filtering," *IEEE Trans. Signal Processing*, vol. 45, pp. 492-496, Feb. 1997.

[18] D. A. Pados and S. N. Batalama, "Low-complexity blind detection of DS/CDMA signals: Auxiliary-vector receivers," *IEEE Trans. Commun.*, vol. 45, pp. 1586-1594, Dec. 1997.

[19] A. Kansal, S. N. Batalama, and D. A. Pados, "Adaptive maximum SINR RAKE filtering for DS-CDMA multipath fading channels," *IEEE Journ. Select. Areas Commun.*, vol. 16, pp. 1765-1773, Dec. 1998.

[20] S. N. Batalama, M. J. Medley, and D. A. Pados, "Robust adaptive recovery of spread-spectrum signals with short data records," *IEEE Trans. Commun.*, vol. 48, pp. 1725-1731, Oct. 2000.

[21] D. A. Pados and G. N. Karystinos, "An iterative algorithm for the computation of the MVDR filter," *IEEE Trans. Signal Proc.*, vol. 49, pp. 290-300, Feb. 2001.

[22] D. A. Pados and S. N. Batalama, "Joint space-time auxiliary-vector filtering for DS/CDMA systems with antenna arrays," in *Proc. Conf. on Inform. Sc. and Syst.*, Princeton University, Princeton, NJ, Mar. 1998, vol. II, pp. 1007-1013.

[23] D. A. Pados, J. H. Michels, T. Tsao, and M. C. Wicks, "Joint domain space-time adaptive processing with small training data sets," in *Proc. IEEE Radar Conf.*, Dallas, TX, May 1998, pp. 99-104.

[24] D. A. Pados and S. N. Batalama, "Joint space-time auxiliary-vector filtering for DS/CDMA systems with antenna arrays," *IEEE Trans. Commun.*, vol. 47, pp. 1406-1415, Sept. 1999.

[25] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE Journ. Select. Areas Commun.*, vol. 12, pp. 685-697, May 1994.

[26] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, pp. 3178-3188, Dec. 1994.

[27] S. L. Miller, "An adaptive DS-CDMA receiver for multiuser interference rejection," *IEEE Trans. Commun.*, vol. 43, pp. 1746-1755, Feb./Mar./Apr. 1995.

[28] M. L. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944-960, July 1995.

[29] C. N. Pateros and G. J. Saulnier, "An adaptive correlator receiver for direct-sequence spread-spectrum communication," *IEEE Trans. Commun.*, vol. 44, pp. 1543-1552, Nov. 1996.

[30] C. D. Richmond, "Derived PDF of maximum likelihood signal estimator which employs an estimated noise covariance," *IEEE Trans. Signal Proc.*, vol. 44, pp. 305-315, Feb. 1996.

[31] R. L. Dykstra, "Establishing the positive definiteness of the sample covariance matrix," *Ann. Math. Statist.*, vol. 41, no. 6, pp. 2153-2154, 1970.

[32] A. O. Steinhardt, "Adaptive multisensor detection and estimation," in *Adaptive Radar Detection and Estimation*, S. Haykin and A. O. Steinhardt, ed., Ch. 3. New York, NY: Wiley, 1992.

[33] C. D. Richmond, "PDF's, confidence regions, and relevant statistics for a class of sample covariance-based array processors," *IEEE Trans. Signal Proc.*, vol. 44, pp. 1779-1793, July 1996.

[34] A. O. Steinhardt, "The PDF of adaptive beamforming weights," *IEEE Trans. Signal Proc.*, vol. 39, pp. 1232-1235, May 1991.

[35] L. C. Godara and A. Cantoni, "Analysis of constrained LMS algorithm with application to adaptive beamforming using perturbation sequences," *IEEE Trans. Antennas Propagation*, vol. 34, pp. 368-379, Mar. 1986.

[36] V. Solo, "The limiting behavior of LMS," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 37, pp. 1909-1922, Dec. 1989.

[37] J. M. Cioffi and T. Kailath, "Fast recursive-least-squares transversal filters for adaptive filtering," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 32, pp. 304-337, Apr. 1984.

[38] T.-C. Liu and B. Van Veen, "A modular structure for implementation of linearly constrained minimum variance beamformers," *IEEE Trans. Signal Proc.*, vol. 39, pp. 2343-2346, Oct. 1991.

[39] H. Qian and S. N. Batalama, "Data-record-based criteria for the selection of an auxiliary-vector estimator of the MVDR filter," in *Proc. Asilomar Conf. Signals, Syst., Computers*, Pacific Grove, CA, Oct. 2000, pp. 802-807.

[40] C. R. Rao, *Handbook of Statistics 9*. New York, NY: Elsevier, 1993.

[41] H. V. Poor and S. Verdu, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inform. Theory*, vol. 43, pp. 858-871, May 1997.

[42] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY: Wiley, 1991.

[43] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York, NY: Springer-Verlag, 1994.

[44] M. L. Honig and W. Xiao, "Performance of reduced-rank linear interference suppression," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1928-1946, July 2001.

[45] M. H. DeGroot, *Probability and Statistics*. Reading, MA: Addison-Wesley, 1984.

[46] M. C. Jones, "Expressions for inverse moments of positive quadratic forms in normal variables," *Austral. J. Statist.*, vol. 28, pp. 242-250, Feb. 1986.

[47] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York, NY: Dover Publications, 1965.

Chapter 2

Capacity, Throughput, and Delay of Slotted ALOHA DS-CDMA Links with Adaptive Space-Time Auxiliary-Vector Receivers

2.1 Introduction

Mobile communication systems are a key development toward realizing a global personal communications network [1]. In recent years, we saw significant interest and commercial activity in the area of mobile networks which employ a base station transceiver serving a number of mobile terminals (star architecture). Land-mobile satellite systems and cellular communications are manifestations of such networks that can provide a wide range of radio services including interactive sessions, file transfer, voice, facsimile, and position finding. Past commercially available mobile communication systems supported information bit rates of 9.6-32 kb/s. Effective mobile Internet access is one application example that requires the support of higher data rates and user capacity. Third generation (3G) wireless systems promise to deliver significantly improved data rates and capacity [2].

Several 3G wireless system blueprints are based on direct-sequence code-division-multiple-access (DS-CDMA) technology. Two leading proposals are cdma2000 (an extension of IS-95B) and Wideband CDMA (W-CDMA). Both depend on packet data communication services between the mobile terminal and the base station [2]. In the past, random access packet switching has received considerable attention for systems with high peak-to-average traffic ratios [3], [4]. Most of the early research effort was directed toward ALOHA-type systems that operate under the tacit assumption that when two or more transmissions occur at the same time all packets are destroyed due to collision [5], [6]. Various modifications of the ALOHA scheme, including the time slotted version, have been under investigation. Application of code division spread-spectrum concepts [7] permits simultaneous transmission and successful reception of two or more packets. When a medium-access-control (MAC) protocol keeps the number of simultaneous transmissions below a system capacity threshold, each transmitting user may be assigned a unique spreading code for the duration of the transmission [4], [8]-[13]; this system is called "slotted ALOHA DS-CDMA with dedicated signatures" [11], [12], [14] and is free of collision errors and pertinent user identification problems [15]-[17].

In many cases of data transmission, the main measure of link quality is the throughput (either packet throughput or information throughput) which is directly related to the operational physical layer packet-error-rate (PER). Physical layer PER requirements may be determined by the type of service and its sensitivity to delays. For example, real-time voice imposes stringent delay constraints and may require one-shot transmission with guaranteed upper bound on PER of about 10^{-2} [18]. On the other hand, data packets can tolerate reasonable delays but (near) error-free final data delivery may be expected [19]. In this context, a retransmission protocol (ARQ) can soften the physical layer single-transmission PER requirement significantly. Various works, for example [8], [20]-[23], have analyzed the performance of DS-CDMA

transmissions in a random access mode in terms of throughput, delay characteristics and system capacity improvements. In DS-CDMA links with dedicated signatures, the physical layer PER performance is limited only by multiple-access-interference (MAI) and time-varying channel characteristics (such as multipath propagation and fading).

While PER improvements can be achieved by employing forward error correction (FEC) [24], significant underlying "raw" bit-error-rate (BER) -and, hence, PER- improvements are achieved by using adaptive antenna array ("smart antenna") receiver techniques. DS-CDMA systems equipped with adaptive antenna arrays offer the opportunity for jointly effective spatial (array) and temporal (code) MAI and channel noise suppression [25], [26]. Adaptive minimum-mean-square-error (MMSE) filtering algorithms, such as sample-matrix-inversion (SMI) [27], least-mean-squares (LMS) [28], and recursive-least-squares (RLS) [29], [30] outperform significantly *upon convergence* the plain space-time (S-T) RAKE filter. However, the time-varying channel characteristics require fast adaptive joint S-T optimization through small input data sets that can "catch up" with multipath fading communication channels.

In this direction, the Auxiliary-Vector (AV) adaptive filtering theory [26], [31]-[33] offers the opportunity for effective interference suppression with short data records. The AV algorithm generates a sequence of estimators of the MMSE filter. No explicit or implicit matrix inversion, decomposition, or diagonalization operation is required. Most importantly, the early, non-asymptotic, elements of the generated sequence of estimators offer favorable bias/covariance balance and have been seen to outperform LMS, RLS, or SMI estimates in mean-square filter estimation error. Recently, blind criteria for the selection of the most successful AV filter estimator in the generated sequence were developed [34], [35].

In this present work, we focus on mobile packet data networks and we examine the PER, throughput, and delay characteristics of slotted ALOHA DS-CDMA random access links with dedicated signatures equipped with packet-rate S-T adaptive AV-filter receivers. The communication channel is assumed to be multipath Rayleigh fading and the packet size is chosen sufficiently small to conform with the channel coherence time (that is, the channel realization is assumed to remain nearly constant during each packet transmission). For the purposes of adaptive joint S-T processing, a subspace S-T channel estimation procedure is also developed based on the single-antenna channel estimation method of [36] (appropriately modified and extended to S-T CDMA systems). Since second-order blind channel estimation methods suffer from phase ambiguity, mean-square optimum phase recovery is attempted through a short pilot bit sequence that is included as a mid-amble [37] in each packet. As we will observe, AV S-T filtering results in improved channel BER, which in turn translates to higher packet success probability and higher user capacity for a given PER upperbound constraint. Higher packet success probability ensures a significant improvement in terms of packet throughput and delay profile for networks with or without user capacity control. Additional performance improvements through FEC coding can be pursued and FEC-enhanced performance studies are presented.

The rest of this chapter is organized as follows. In Section 2.2 we describe the system and data model under consideration. In Section 2.3 we present the AV filter estimation algorithm. Blind S-T channel estimation together with supervised channel-phase recovery are considered in Section 2.4. In Section 2.5 we focus on the packet data unit and examine the overall packet system performance. A few concluding remarks are drawn in Section 2.6.

2.2 System Model

We consider a slotted ALOHA DS-CDMA link of a packet data network. We assume the presence of K actively transmitting mobile terminals (users) with dedicated distinct signatures and we concentrate on mobile-to-base-station transmissions (uplink connection). The K simultaneous in time and frequency user transmissions take place over multipath fading additive white Gaussian noise (AWGN) channels. The received signal at the base station is collected by a narrowband antenna array. For illustration purposes, we consider uniform linear arrays. Details and notation are given below.

2.2.1 Data-packet structure

The binary data of each user are organized in identically structured packets of size M bits. The kth user data packet,

$$\{b_k(0), b_k(1), \ldots, b_k(M-1)\}, \qquad k = 0, 1, \ldots, K-1, \tag{2.1}$$

contains $M - M_p$ information bits and M_p pilot bits (bits that are known to the receiver). We recall that blind second-order channel estimation methods return phase-ambiguous estimates; the M_p known bits will be utilized for the supervised recovery of the phase of such blind channel estimates. An example of the data packet structure is shown in Fig. 2.1 where the M_p pilot bits appear as a *mid-amble* in the transmitted packet.

Without loss of generality, we assume that each user may transmit one data packet per slot and the slot duration is T_s seconds. Therefore, the data packet size M equals the number of bits transmitted by each user in one time slot and if the duration of each bit transmission is T_b seconds, then $T_s = MT_b$.

2.2.2 Transmitted signal

The bit sequence $\{b_k(0), b_k(1), \ldots, b_k(M-1)\}$ of the kth user data packet, $k = 0, 1, \ldots, K-1$, is modulated by the binary signature assigned to that user

$$\mathbf{d}_k = [d_k(0), d_k(1), \ldots, d_k(L-1)]^T \tag{2.2}$$

where $d_k(l) \in \{\pm 1\}$, $l = 0, 1, \ldots, L-1$, and $(\cdot)^T$ denotes the transpose operation. The signature codelength L is also known as the "processing gain" of the DS-CDMA system. Each active (transmitting) user is assigned a unique dedicated signature that remains fixed for the duration of the transmission. If signature assignment is administered directly by the base station, then the signature itself can be used for user identification purposes. If instead each user draws its own dedicated signature from the available pool without further book-keeping, then user identification information needs to be included in the $M - M_p$ information bits of every packet.[1] In either case, upon completion of the transmission the signature is released back to the system. Adopting a conventional approach, binary Gold user signatures are considered in this work. Recently, quasi-stationary (QS) sequences derived from Gold sequences were seen to be well suited for CDMA slotted ALOHA systems [38].

The contribution of the kth user packet to the transmitted radio signal is denoted by

$$u_k(t) = \sum_{i=0}^{M-1} b_k(i)\sqrt{E_k}s_k(t - iT)e^{j(2\pi f_c t + \phi_k)}. \tag{2.3}$$

The sinusoidal carrier frequency has the same nominal value f_c for all users, while each transmitter has a time invariant carrier phase shift ϕ_k; $s_k(t)$ is the user signature waveform given by

$$s_k(t) = \sum_{l=0}^{L-1} d_k(l)\psi(t - lT_c) \tag{2.4}$$

where $\psi(t)$ is the time-limited chip waveform with duration (period) $T_c = \frac{T_b}{L}$. The signature waveform is assumed to be normalized to unit energy per bit period ($\int_0^{T_b} s_k^2(t)dt = 1$). Therefore, E_k in (2.3) represents transmitted energy per bit for user k.

2.2.3 Channel model and received signal

The transmitted signal $u_k(t)$ of each user k, $k = 0, 1, \ldots, K-1$, propagates over a multipath fading AWGN channel. All user transmitters (as well as the base station receiver) are assumed to be synchronized to the universal slot clock. A guard interval of length several chip periods can be "inserted" between slots to avoid slot spillover. After multipath fading channel "processing," the total received signal *due to all K users* is

$$r(t) = \sum_{k=0}^{K-1}\sum_{n=0}^{N-1} \alpha_{k,n}u_k(t - \frac{n}{B}) + n(t) \tag{2.5}$$

where $n(t)$ is a zero-mean complex white Gaussian noise process with power spectral density $\frac{N_0}{2}$ and N is the number of resolvable multipaths encountered by the packets (w.l.o.g. the number of resolvable multipaths

[1] In DS-CDMA packet data systems with random (non-dedicated) signature selection, the user identification issue becomes significantly more complicated due to the potential collisions [15]-[17].

N is assumed to be the same for all user transmissions). With $r(t)$ bandlimited to $B = \frac{1}{T_c}$, we can assume a tapped-delay line channel model with taps spaced at chip intervals T_c [39]. The path coefficients $\alpha_{k,n}$, $k = 0, 1, \ldots, K-1$, $n = 0, 1, \ldots, N-1$, are independent zero-mean complex Gaussian random variables (Rayleigh distributed amplitude and uniformly distributed phase) that model the fading phenomena and are assumed to remain constant over the entire packet duration. For example [26], the path coefficients for a typical vehicle mobile transmitting data at 19.5 kb/s on a carrier $f_c = 900$MHz remain practically unchanged for about 280 bit periods (channel fading rate 70Hz). Thus, keeping the packet size $M \leq 280$ validates the quasi-static packet assumption.

The base station receiver is equipped with a narrowband linear antenna array of Z elements. The $Z \times 1$ array response vector for the nth path of the kth user signal is defined by

$$\mathbf{a}_{k,n}(z) = e^{j2\pi(z-1)\frac{d}{\lambda}\sin\theta_{k,n}}, \qquad z = 1, 2, \ldots, Z, \tag{2.6}$$

where $\theta_{k,n}$ identifies the angle of arrival of the corresponding path, λ is the carrier wavelength, and d is the element spacing of the antenna array (usually $d = \frac{\lambda}{2}$). The received $Z \times 1$ vector at the antenna array after carrier demodulation is given by

$$\mathbf{r}(t) = \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} b_k(i) \sum_{n=0}^{N-1} c_{k,n} \sqrt{E_k} s_k(t - iT_b - nT_c) \mathbf{a}_{k,n} + \mathbf{n}(t) \tag{2.7}$$

where $c_{k,n} = \alpha_{k,n} e^{-j(2\pi f_c n T_c - \phi_k)}$. Chip-matched filtering and sampling at the chip rate $\frac{1}{T_c}$ (or low-pass filtering, Nyquist sampling, and chip-rate accumulation) over the multipath extended $L + N - 1$ chip period prepares the data for one-shot detection of the ith information bit of interest, say $b_0(i)$. We visualize the S-T data in the form of a $Z \times (L + N - 1)$ matrix:

$$\mathbf{X}_{Z \times (L+N-1)} = [\mathbf{r}(0), \mathbf{r}(T_c), \ldots, \mathbf{r}((L + N - 2)T_c)] \,. \tag{2.8}$$

To avoid cumbersome two-dimensional filtering operations and notation, we "vectorize" $\mathbf{X}_{Z \times (L+N-1)}$ by sequencing all matrix columns in the form of a single vector:

$$\mathbf{x}_{Z(L+N-1) \times 1} = \mathrm{Vec}\left\{\mathbf{X}_{Z \times (L+N-1)}\right\}. \tag{2.9}$$

From now on, $\mathbf{x}$ denotes the joint S-T data in the $\mathbb{C}^{Z(L+N-1)}$ complex vector domain.

The desired information bit of the user of interest is detected by

$$\hat{b}_0 = \mathrm{sgn}\left(\mathrm{Re}\left\{\mathbf{w}^H \mathbf{x}\right\}\right) \tag{2.10}$$

for any generic S-T filter vector $\mathbf{w}$, where $\mathrm{sgn}(\cdot)$ is the ± 1 hard-limiter, $\mathrm{Re}\{\cdot\}$ extracts the real part of a complex number, and $(\cdot)^H$ denotes the Hermitian operation. We are particularly interested in the joint S-T MMSE optimum filter $\mathbf{w}_{\mathrm{MMSE}}$. Section 2.3 is devoted to the development of short-data-record (packet length) estimators of the joint S-T MMSE filter.

2.3 Auxiliary-Vector Adaptive Filtering

The foundation of joint S-T MMSE filtering is the S-T RAKE filter (well known for time-only processing [40]) that we define compactly as the cross-correlation between the received S-T data $\mathbf{x}$ and the desired information bit b_0 :

$$\mathbf{v}_0 \triangleq E_{b_0}\{\mathbf{x} b_0\} \tag{2.11}$$

where the statistical expectation operation $E_{b_0}\{\cdot\}$ is taken with respect to b_0 only. It is straightforward to obtain the MMSE (a.k.a. Wiener) filter solution [41] for our problem if we view the sequence of information bits of the user of interest as the "desired" filter output when the filter input is the S-T data vectors in (2.9). The solution is of the form

$$\mathbf{w}_{\mathrm{MMSE}} = c\mathbf{R}^{-1}\mathbf{v}_0 \tag{2.12}$$

where c is a positive scalar, $\mathbf{v}_0$ is the joint S-T RAKE filter previously defined, and $\mathbf{R} = E\left\{\mathbf{x}\mathbf{x}^H\right\}$ is the S-T input autocorrelation matrix. If we choose $c = \frac{1}{\mathbf{v}_0^H \mathbf{R}^{-1} \mathbf{v}_0}$, then we obtain the well known equivalent (cf. (10)) minimum-variance-distortionless-response (MVDR) filter

$$\mathbf{w}_{\text{MVDR}} = \frac{\mathbf{R}^{-1}\mathbf{v}_0}{\mathbf{v}_0^H \mathbf{R}^{-1}\mathbf{v}_0} \tag{2.13}$$

for which $\mathbf{w}_{\text{MVDR}}^H \mathbf{v}_0 = 1$.

The MMSE/MVDR filter is a function of the true S-T input autocorrelation matrix $\mathbf{R}$ and the true S-T RAKE filter $\mathbf{v}_0$. However, neither $\mathbf{R}$ nor $\mathbf{v}_0$ is known to the receiver. In this section, we present the auxiliary-vector (AV) algorithm which provides us with estimates of the optimum MMSE/MVDR filter when $\mathbf{R}$ is unknown and sample-average estimated from the size M packet data record $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{M-1}$:

$$\hat{\mathbf{R}}(M) = \frac{1}{M} \sum_{i=0}^{M-1} \mathbf{x}_i \mathbf{x}_i^H. \tag{2.14}$$

Throughout this section, $\mathbf{v}_0$ is assumed to be known. A procedure for the estimation of $\mathbf{v}_0$ from the *same* packet data record $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{M-1}$ will be presented in Section 2.4.

The AV algorithm generates an infinite sequence of filters $\{\mathbf{w}_n\}_{n=0}^{\infty}$. The sequence is initialized at the S-T RAKE filter

$$\mathbf{w}_0 = \frac{\mathbf{v}_0}{\|\mathbf{v}_0\|^2}, \tag{2.15}$$

which is here scaled to satisfy $\mathbf{w}_0^H \mathbf{v}_0 = 1$. At each step $k+1$ of the algorithm, $k = 0, 1, 2, \ldots$, we incorporate in $\mathbf{w}_k$ an "auxiliary" vector component $\mathbf{g}_{k+1}$ that is orthogonal to $\mathbf{v}_0$ and weighted by a scalar μ_{k+1} and we form the next filter in the sequence

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mu_{k+1}\mathbf{g}_{k+1}. \tag{2.16}$$

The auxiliary vector $\mathbf{g}_{k+1}$ is chosen to maximize, under fixed norm, the magnitude of the cross-correlation between its output $\mathbf{g}_{k+1}^H \mathbf{x}$ and the previous filter output $\mathbf{w}_k^H \mathbf{x}$ and is given by

$$\mathbf{g}_{k+1} = \mathbf{R}\mathbf{w}_k - \frac{\mathbf{v}_0^H \mathbf{R}\mathbf{w}_k}{\|\mathbf{v}_0\|^2}\mathbf{v}_0. \tag{2.17}$$

The scalar μ_{k+1} is selected such that it minimizes the output variance of the filter $\mathbf{w}_{k+1}$ or equivalently minimizes the MS error between $\mathbf{w}_k^H \mathbf{x}$ and $\mu_{k+1}^* \mathbf{g}_{k+1}^H \mathbf{x}$. The MS-optimum μ_{k+1} is

$$\mu_{k+1} = \frac{\mathbf{g}_{k+1}^H \mathbf{R}\mathbf{w}_k}{\mathbf{g}_{k+1}^H \mathbf{R}\mathbf{g}_{k+1}}. \tag{2.18}$$

The AV filter recursion is completely defined by (2.15)-(2.18). Theoretical analysis of the AV algorithm was pursued in [33]. The results are summarized below in the form of a theorem.

Theorem 2.1 *Let* $\mathbf{R}$ *be a Hermitian positive definite matrix. Consider the iterative algorithm of eqs. (2.15)-(2.18).*
(i) *Successive auxiliary vectors generated through (2.16)-(2.18) are orthogonal:* $\mathbf{g}_i^H \mathbf{g}_{i+1} = 0$, $i = 1, 2, 3, \ldots$ *(however, in general* $\mathbf{g}_i^H \mathbf{g}_j \neq 0$ *for* $|i - j| \neq 1$*).*
(ii) *The generated sequence of auxiliary-vector weights* $\{\mu_n\}$, $n = 1, 2, \ldots$, *is real-valued, positive, and bounded:* $0 < \frac{1}{\lambda_{\max}} \leq \mu_n \leq \frac{1}{\lambda_{\min}}$, $n = 1, 2, \ldots$, *where* $\lambda_{\max}$ *and* $\lambda_{\min}$ *are the maximum and minimum, correspondingly, eigenvalues of* $\mathbf{R}$*.*
(iii) *The sequence of auxiliary vectors* $\{\mathbf{g}_n\}$, $n = 1, 2, \ldots$, *converges to the* $\mathbf{0}$ *vector:* $\lim_{n\to\infty} \mathbf{g}_n = \mathbf{0}$. **(iv)** *The sequence of auxiliary-vector filters* $\{\mathbf{w}_n\}$, $n = 1, 2, \ldots$, *converges to the MVDR filter:* $\lim_{n\to\infty} \mathbf{w}_n = \frac{\mathbf{R}^{-1}\mathbf{v}_0}{\mathbf{v}_0^H \mathbf{R}^{-1}\mathbf{v}_0}$.
□

If $\mathbf{R}$ is unknown and sample-average estimated from a packet data record of M points as in (2.14), then Theorem 2.1 shows that

$$\hat{\mathbf{w}}_n(M) \underset{n\to\infty}{\longrightarrow} \hat{\mathbf{w}}_\infty(M) = \frac{\left[\hat{\mathbf{R}}(M)\right]^{-1}\mathbf{v}_0}{\mathbf{v}_0^H\left[\hat{\mathbf{R}}(M)\right]^{-1}\mathbf{v}_0} \tag{2.19}$$

where $\hat{\mathbf{w}}_\infty(M)$ is the widely used MVDR filter estimator known as the sample-matrix-inversion (SMI) filter [27]. The output sequence begins from $\hat{\mathbf{w}}_0(M) = \frac{\mathbf{v}_0}{\|\mathbf{v}_0\|^2}$, which is a *0*-variance, fixed-valued, estimator that may be severely biased ($\hat{\mathbf{w}}_0(M) = \frac{\mathbf{v}_0}{\|\mathbf{v}_0\|^2} \neq \mathbf{w}_{\rm MVDR}$) unless $\mathbf{R} = \sigma^2\mathbf{I}$ for some $\sigma > 0$. In the latter trivial case, $\hat{\mathbf{w}}_0(M)$ is already the perfect MVDR filter. Otherwise, the next filter estimator in the sequence, $\hat{\mathbf{w}}_1(M)$, has significantly reduced bias due to the optimization procedure employed at the expense of non-zero estimator (co-)variance. As we move up in the sequence of filter estimators $\hat{\mathbf{w}}_n(M)$, $n = 0, 1, 2, \ldots$, the bias decreases rapidly to zero while the variance rises slowly to the SMI ($\hat{\mathbf{w}}_\infty(M)$) levels (cf. (2.19)).

An adaptive data-dependent procedure for the selection of the most appropriate member of the AV filter estimator sequence $\{\hat{\mathbf{w}}_n(M)\}$ for a given data record of size M is presented in [34], [35]. The procedure selects the estimator $\hat{\mathbf{w}}_{n_0}$ from the generated sequence of AV filter estimators that exhibits maximum J-divergence between the filter output conditional distributions given that $+1$ or -1 is transmitted. Under a Gaussian approximation on the conditional filter output distribution, it was shown in [34], [35] that the J-divergence of the filter estimator with n auxiliary vectors is

$$J(n) = \frac{4\,E^2\left\{b_0 \mathrm{Re}\left[\hat{\mathbf{w}}_n^H(M)\mathbf{x}\right]\right\}}{Var\left\{b_0 \mathrm{Re}\left[\hat{\mathbf{w}}_n^H(M)\mathbf{x}\right]\right\}}. \tag{2.20}$$

To estimate the J-divergence $J(n)$ from the data packet of size M, the transmitted information bits b_0 are required to be known. We can obtain a blind *approximate* version of $J(n)$ by substituting the information bit b_0 in (2.20) by the detected bit $\hat{b}_0 = \mathrm{sgn}\left(\mathrm{Re}\left\{\hat{\mathbf{w}}_n^H(M)\mathbf{x}\right\}\right)$ (output of the sign detector that follows the linear filter). In particular, using $\hat{b}_0$ in place of b_0 in (2.20) we obtain the following J-divergence expression:

$$J_B(n) = \frac{4\,E^2\left\{\hat{b}_0 \mathrm{Re}\left[\hat{\mathbf{w}}_n^H(M)\mathbf{x}\right]\right\}}{Var\left\{\hat{b}_0 \mathrm{Re}\left[\hat{\mathbf{w}}_n^H(M)\mathbf{x}\right]\right\}} = \frac{4\,E^2\left\{\left|\mathrm{Re}\left[\hat{\mathbf{w}}_n^H(M)\mathbf{x}\right]\right|\right\}}{Var\left\{\left|\mathrm{Re}\left[\hat{\mathbf{w}}_n^H(M)\mathbf{x}\right]\right|\right\}} \tag{2.21}$$

where the subscript "B" identifies the blind version of the J-divergence function. To estimate $J_B(n)$ from the data packet of size M, we substitute the statistical expectations in (2.21) by sample averages. The following criterion summarizes the corresponding AV filter estimator selection rule.

Criterion 2.1 *For a given data record of size M, the unsupervised (blind) J-divergence AV filter estimator selection rule chooses the estimator $\hat{\mathbf{w}}_{n_0}(M)$ with n_0 auxiliary vectors where*

$$n_0 = \arg\max_n\left\{\hat{J}_B(n)\right\} = \arg\max_n\left\{\frac{4\left[\frac{1}{M}\sum_{i=0}^{M-1}\left|\mathrm{Re}\left[\hat{\mathbf{w}}_n^H(M)\mathbf{x}_i\right]\right|\right]^2}{\frac{1}{M}\sum_{i=0}^{M-1}\left|\mathrm{Re}\left[\hat{\mathbf{w}}_n^H(M)\mathbf{x}_i\right]\right|^2 - \left[\frac{1}{M}\sum_{i=0}^{M-1}\left|\mathrm{Re}\left[\hat{\mathbf{w}}_n^H(M)\mathbf{x}_i\right]\right|\right]^2}\right\}. \tag{2.22}$$

□

Criterion 2.1 completes the design of the joint S-T auxiliary-vector filter estimator. As a brief summary, the sequence of calculations is as follows. The whole packet data record of joint S-T input data vectors $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{M-1}$ is utilized to obtain the estimate $\hat{\mathbf{R}}(M)$ of $\mathbf{R}$ according to (2.14). Then, the sequence of AV filter estimators $\hat{\mathbf{w}}_0(M), \hat{\mathbf{w}}_1(M), \hat{\mathbf{w}}_2(M), \ldots$ is generated by the recursive algorithm of (2.15)-(2.18) with the true autocorrelation matrix $\mathbf{R}$ replaced by its estimate $\hat{\mathbf{R}}(M)$. The most appropriate AV filter estimator $\hat{\mathbf{w}}_{n_0}(M)$ in the sequence is selected according to the "maximum estimated blind J-divergence rule" of (2.22). Finally, the AV filter estimator $\hat{\mathbf{w}}_{n_0}(M)$ is used for bit-by-bit detection of the packet information bit sequence $b_0(0), b_0(1), \ldots, b_0(M-1)$ according to (2.10).

For simplicity in the presentation of the auxiliary-vector filter estimation algorithm in this section, the S-T RAKE filter of the user of interest $\mathbf{v}_0$ was assumed to be known which is certainly not true in practice. Section 2.4 is devoted to the estimation of $\mathbf{v}_0$ from the *same* packet data record $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{M-1}$ which, we recall, includes M_p pilot bits.

2.4 S-T RAKE Filter Estimation

The S-T RAKE filter of the user of interest (user *0*) $\mathbf{v}_0$ defined by (2.11) consists of shifted versions of the S-T matched filter multiplied by the corresponding channel coefficients:[2]

$$\mathbf{v}_0 = \sum_{n=0}^{N-1} c_{0,n} \left[\underbrace{0 \ldots 0}_{n} \;\; \mathbf{d}_0^T \;\; \underbrace{0 \ldots 0}_{N-n-1} \right]^T \odot \mathbf{a}_{0,n} \tag{2.23}$$

where the symbol $\odot$ denotes the Kronecker tensor product. Hence, $\mathbf{v}_0$ is a function of the binary signature vector (spreading sequence) of the user of interest $\mathbf{d}_0$, the channel coefficients $c_{0,0}, c_{0,1}, \ldots, c_{0,N-1}$, and the corresponding angles of arrival $\theta_{0,0}, \theta_{0,1}, \ldots, \theta_{0,N-1}$. While the spreading sequence is assumed to be known to the receiver, the channel coefficients and the angles of arrival are in general unknown.

Channel estimation with one or multiple sensors is a subject that has been researched extensively over the past several years. In [42], the channel coefficients of a DS-CDMA link are estimated in the absence of multipaths based on the cyclostationarity features of the received signals (see also [43] where the receiver is equipped with multiple sensors and maximum-likelihood (ML) channel estimation is attempted by modeling the interference as colored Gaussian noise). In [44], ML estimation of multipath channels is pursued again using multiple sensors at the receiver. Both ML estimation methods of [43] and [44], however, are supervised (they require knowledge of the transmitted information bits). A subspace method for direct estimation of the MMSE filter (without intermediate channel estimation) is presented in [45]. Identifiability results on subspace-based channel estimation are derived in [46]. In [47], a constrained adaptive algorithm is presented that combines all multipath components and minimizes the multiuser interference at the receiver output. However, this algorithm (as well as the methods in [43] and [45]) has satisfactory performance only when large data records are utilized for the training of the estimator. Based on a genetic algorithm, an iterative method for joint channel estimation and symbol detection is presented in [48]. The reduced-rank multistage Wiener filter of [49] is used for channel estimation purposes in [50]. It has been shown, however, that the AV filter estimators presented in the previous section outperform significantly in MS filter estimation error the reduced-rank multistage Wiener filters in short-data-record situations [33]. Finally, least-squares-based supervised channel estimation for long-coded CDMA systems (which is not covered in the present work) is performed in [51].

In this section, we are interested in carrying out blind (non-supervised) multipath S-T channel estimation when only a short input data record of size M is available for the preparation of the estimate. To estimate the channel coefficients $\mathbf{c}_0 \triangleq [c_{0,0}, c_{0,1}, \ldots, c_{0,N-1}]^T$ and the angles of arrival $\boldsymbol{\theta}_0 \triangleq [\theta_{0,0}, \theta_{0,1}, \ldots, \theta_{0,N-1}]^T$ for the user of interest *0* from the S-T packet data record $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{M-1}$, we design a subspace-based estimation procedure. This procedure can be viewed as a generalization of the algorithm in [36] for an antenna-array set-up. An additional modification is introduced that increases the rank of the noise subspace.

The rank r_s of the *signal subspace* of the received data vectors $\mathbf{x}$ can be controlled by considering one-sided or two-sided truncations of $\mathbf{x}$ (the latter eliminates inter-symbol-interference (ISI)). The possible values of r_s, depending on the data formation of choice, are as follows.

(i)	No truncation:	Data dimension $= Z(L+N-1)$,	$2K+1 \le r_s \le 3K$.
(ii)	One-sided truncation:	Data dimension $= ZL$,	$2K \le r_s \le 3K-1$.
(iii)	Two-sided truncation:	Data dimension $= Z(L-N+1)$,	$K \le r_s \le 2K-1$.

To have a guaranteed minimum rank of the *noise subspace* of $Z(L-N+1)-(2K-1)$, we decide to truncate $\mathbf{x}$ from both sides (Case (iii)) as shown in Fig. 2.2 and we form the "truncated" received vector $\mathbf{x}^{\mathrm{tr}}$ of length $Z(L-N+1)$ as follows:

$$\mathbf{x}^{\mathrm{tr}} = \begin{bmatrix} \mathbf{r}((N-1)T_c) \\ \mathbf{r}(NT_c) \\ \vdots \\ \mathbf{r}((L-1)T_c) \end{bmatrix}. \tag{2.24}$$

[2] For the sake of mathematical accuracy, $\mathbf{v}_0 = \frac{\sqrt{E_0}}{L} \sum_{n=0}^{N-1} c_{0,n} \left[\underbrace{0 \ldots 0}_{n} \;\; \mathbf{d}_0^T \;\; \underbrace{0 \ldots 0}_{N-n-1} \right]^T \odot \mathbf{a}_{0,n}$ according to the definition in (11). The positive multiplier $\frac{\sqrt{E_0}}{L}$ is dropped from (23) as inconsequential.

Then, with respect to the ith information bit of user *0*, $\mathbf{x}_i^{\mathrm{tr}}$ can be expressed as

$$\mathbf{x}_i^{\mathrm{tr}} = b_0(i)\frac{\sqrt{E_0}}{L}\mathbf{A}_0\mathbf{B}(\boldsymbol{\theta}_0)\mathbf{c}_0 + \mathrm{MAI}_i + \mathbf{n}_i \tag{2.25}$$

where MAI_i accounts comprehensively for S-T multiple-access-interference of rank $r_s - 1$, $\mathbf{B}(\boldsymbol{\theta}_0)$ is a block diagonal matrix of the form $\mathbf{B}(\boldsymbol{\theta}_0) \triangleq \mathrm{diag}\left(\mathbf{a}_{0,0}, \mathbf{a}_{0,1}, \ldots, \mathbf{a}_{0,N-1}\right)$, and $\mathbf{A}_0 = \mathbf{A}_0^s \odot \mathbf{I}_Z$ where $\mathbf{I}_Z$ is the $Z \times Z$ identity matrix and

$$\mathbf{A}_0^s = \begin{bmatrix} d_0[N-1] & d_0[N-2] & \ldots & d_0[0] \\ d_0[N] & d_0[N-1] & \ldots & d_0[1] \\ \vdots & \vdots & & \vdots \\ d_0[L-1] & d_0[L-2] & \ldots & d_0[L-N] \end{bmatrix}. \tag{2.26}$$

Let $\mathbf{R}_{\mathrm{tr}} = E\left\{\mathbf{x}^{\mathrm{tr}}\mathbf{x}^{\mathrm{tr}H}\right\}$ be the autocorrelation matrix of $\mathbf{x}^{\mathrm{tr}}$. We form a sample-average estimate

$$\hat{\mathbf{R}}_{\mathrm{tr}} = \frac{1}{M}\sum_{i=0}^{M-1}\mathbf{x}_i^{\mathrm{tr}}\mathbf{x}_i^{\mathrm{tr}H} \tag{2.27}$$

based on the truncated M available S-T input vectors $\mathbf{x}_i^{\mathrm{tr}}, i = 0, 1, \ldots, M-1$. If $\hat{\mathbf{R}}_{\mathrm{tr}} = \hat{\mathbf{Q}}\hat{\boldsymbol{\Lambda}}\hat{\mathbf{Q}}^H$ represents the eigen-decomposition of $\hat{\mathbf{R}}_{\mathrm{tr}}$ where the columns of $\hat{\mathbf{Q}}$ are the eigenvectors of $\hat{\mathbf{R}}_{\mathrm{tr}}$ and $\hat{\boldsymbol{\Lambda}}$ is a diagonal matrix consisting of the eigenvalues of $\hat{\mathbf{R}}_{\mathrm{tr}}$, then we use the eigenvectors that correspond to the $Z(L - N + 1) - (2K - 1)$ smallest eigenvalues to define our estimated noise subspace. Let the matrix $\hat{\mathbf{U}}_n$ of size $[Z(L-N+1)] \times [Z(L-N+1) - (2K-1)]$ consist of these "noise eigenvectors."[3] We propose to estimate $\mathbf{c}_0$ and $\boldsymbol{\theta}_0$ indirectly through an estimate of the $ZN \times 1$ vector

$$\mathbf{h}_0 \triangleq \mathbf{B}(\boldsymbol{\theta}_0)\mathbf{c}_0. \tag{2.28}$$

We estimate $\mathbf{h}_0$ as the vector that minimizes the norm of the projection of the signal of the user of interest *0*, $\mathbf{A}_0\mathbf{B}(\boldsymbol{\theta}_0)\mathbf{c}_0 = \mathbf{A}_0\mathbf{h}_0$, onto the estimated noise subspace $\hat{\mathbf{U}}_n$ [36], [52]:

$$\hat{\mathbf{h}}_0 = \arg\min_{\mathbf{h}_0}\left\|(\mathbf{A}_0\mathbf{h}_0)^H\hat{\mathbf{U}}_n\right\| \qquad \text{subject to} \qquad \left\|\hat{\mathbf{h}}_0\right\| = 1. \tag{2.29}$$

The solution to this constrained minimization problem is given by the following proposition. The proof is rather straightforward and, therefore, omitted.

Proposition 2.1 *The vector $\hat{\mathbf{h}}_0$ that minimizes the norm of the projection of the signal of user 0 to the estimated noise subspace $\hat{\mathbf{U}}_n$ subject to the norm constraint $\|\hat{\mathbf{h}}_0\| = 1$ is the eigenvector that corresponds to the minimum eigenvalue of $\mathbf{A}_0^H\hat{\mathbf{U}}_n\hat{\mathbf{U}}_n^H\mathbf{A}_0$.* □

After obtaining $\hat{\mathbf{h}}_0$, we may extract the desired vectors $\hat{\mathbf{c}}_0$ and $\hat{\boldsymbol{\theta}}_0$ by applying least-squares (LS) fitting to $\hat{\mathbf{h}}_0$. Then, the S-T RAKE filter estimate $\hat{\mathbf{v}}_0$ is completely defined by (2.23).

Since the S-T channel estimation method described above in Proposition 2.1 is based on a blind second-order criterion, the phase information is absorbed and the estimate $\hat{\mathbf{v}}_0$ is phase ambiguous. Inherently, adaptive filter estimators that utilize a phase ambiguous RAKE filter estimate are also phase ambiguous. In the rest of this section, we deal with the recovery (correction) of the phase of S-T linear filters when the S-T RAKE vector $\mathbf{v}_0$ is known within a phase ambiguity.

Without loss of generality, let $\tilde{\mathbf{v}}_0$ denote a phase ambiguous version of $\mathbf{v}_0$, i.e.

$$\tilde{\mathbf{v}}_0 e^{j\psi} = \mathbf{v}_0 \tag{2.30}$$

[3]The definition of $\hat{\mathbf{U}}_n$ requires knowledge of the number of active users K. If K is not known, it can be either estimated adaptively from the eigenvalue profile of $\hat{\mathbf{R}}_{\mathrm{tr}}$ or set to the maximum possible value $K = K_{\max}$ allowed by the network protocol for capacity controlled systems. The latter conservative option $K = K_{\max}$ is quite appealing due to its simplicity and effectiveness; system performance is not affected significantly since in practical applications the S-T product $Z(L - N + 1)$ is much larger than $2K_{\max} - 1$.

where ψ is the unknown phase. We consider the class of linear filters $\mathbf{w} \in \mathbb{C}^{Z(L+N-1)}$ that are functions of the S-T RAKE vector $\mathbf{v}_0$ and share the following property:

$$\mathbf{w}(\mathbf{v}_0) = \mathbf{w}(\tilde{\mathbf{v}}_0)e^{j\psi}. \tag{2.31}$$

Such filters include: *(i)* the S-T RAKE filter itself $\mathbf{v}_0$, *(ii)* the S-T MMSE/MVDR filter of (2.12) or (2.13), and *(iii)* the auxiliary-vector sequence of S-T filters $\{\mathbf{w}_n\}$ defined by eqs. (2.15)-(2.18). Given $\tilde{\mathbf{v}}_0$, we attempt to correct the phase of $\mathbf{w}(\tilde{\mathbf{v}}_0)$ directly. The selection criterion for the phase correction parameter ψ that we propose in this work is the minimization of the mean-square-error (MSE) between the phase corrected S-T filter processed data $\left[\mathbf{w}(\tilde{\mathbf{v}}_0)e^{j\psi}\right]^H \mathbf{x}$ and the desired information bit b_0 (Fig. 2.3):

$$\hat{\psi} = \arg\min_{\psi} E\left\{\left|\left[\mathbf{w}(\tilde{\mathbf{v}}_0)e^{j\psi}\right]^H \mathbf{x} - b_0\right|^2\right\}, \qquad \psi \in [-\pi, \pi). \tag{2.32}$$

The following proposition identifies the optimum phase correction value according to our criterion.

Proposition 2.2 *The phase correction value*

$$\hat{\psi} = angle\left\{\mathbf{w}(\tilde{\mathbf{v}}_0)^H E\left\{\mathbf{x}b_0\right\}\right\} \tag{2.33}$$

minimizes the MSE between the phase-corrected S-T filter processed data $\left[\mathbf{w}(\tilde{\mathbf{v}}_0)e^{j\psi}\right]^H \mathbf{x}$ *and the desired information bit* b_0. □

In essence, Proposition 2.2 suggests to project the phase ambiguous filter $\mathbf{w}(\tilde{\mathbf{v}}_0)$ onto the ideal S-T RAKE filter $\mathbf{v}_0 = E\{\mathbf{x}b_0\}$. However, $E\{\mathbf{x}b_0\}$ is certainly not known. Since we have assumed that a pilot information bit sequence of length M_p is included in each packet, the expectation $E\{\mathbf{x}b_0\}$ can be sample-average estimated by $\frac{1}{M_p}\sum_{j=1}^{M_p} \mathbf{x}_j b_0(j)$ where $b_0(j)$, $j = 1, 2, \ldots, M_p$, is the jth pilot information bit and $\mathbf{x}_j$ is the corresponding S-T input data vector. Then, the phase-corrected adaptive filter estimate is given by

$$\mathbf{w}(\hat{\mathbf{v}}_0, \hat{\mathbf{R}})e^{j\hat{\psi}}, \qquad \hat{\psi} = \text{angle}\left\{\mathbf{w}(\hat{\mathbf{v}}_0, \hat{\mathbf{R}})^H \left[\sum_{j=1}^{M_p} \mathbf{x}_j b_0(j)\right]\right\}. \tag{2.34}$$

We note that if the autocorrelation matrix $\mathbf{R}$ were known exactly, the phase correction estimate for the filter $\tilde{\mathbf{w}}(\tilde{\mathbf{v}}_0, \mathbf{R}) = \frac{\mathbf{R}^{-1}\tilde{\mathbf{v}}_0}{\tilde{\mathbf{v}}_0^H \mathbf{R}^{-1}\tilde{\mathbf{v}}_0}$ would become $\hat{\psi} = \text{angle}\left\{\tilde{\mathbf{v}}_0^H \mathbf{R}^{-1}\left[\sum_{j=1}^{M_p} \mathbf{x}_j b_0(j)\right]\right\}$ which is also the ML estimate of ψ when a data record $\{(\mathbf{x}_1, b_0(1)), (\mathbf{x}_2, b_0(2)), \ldots, (\mathbf{x}_{M_p}, b_0(M_p))\}$ is available and the interference-plus-noise contribution to the received vector $\mathbf{x}_j$ is modeled as a colored Gaussian vector. This result -and phase correction proposal- is in sharp contrast to previous work, [39], [53], that ignored the interference or, equivalently, treated the interference-plus-noise contribution as a white Gaussian vector leading to the ML estimate $\hat{\psi} = \text{angle}\left\{\tilde{\mathbf{v}}_0^H \left[\sum_{j=1}^{M_p} \mathbf{x}_j b_0(j)\right]\right\}$.

Since M represents the packet size of the DS-CDMA system and M_p is the number of mid-amble pilot information bits per packet, the ratio $\frac{M_p}{M}$ quantifies bandwidth waste due to the use of the pilot bit sequence. Ideally, $\frac{M_p}{M}$ is to be kept small. As we will see in Section 2.5, a few pilot bits (of the order of 5 bits) are sufficient for effective recovery of the filter phase. As a numerical example, when the packet size is set at $M = 256$ and $M_p = 5$ is chosen, the wasted bandwidth is only $\frac{M_p}{M} \simeq 2\%$.

2.5 Capacity, Throughput, and Delay Analysis

So far, we concentrated on the design/estimation of the receiver filter $\mathbf{w}$ in (2.10). Our filter estimate is based on the AV algorithm combined with subspace-based techniques and supervised phase correction for S-T channel estimation purposes. The proposed overall AV filter estimate is generated adaptively on an individual packet-by-packet basis. All M received vectors of the packet are utilized for the design of $\mathbf{w}$ (estimation of $\mathbf{R}$, $\mathbf{v}_0$, and the number of AV's n_0), while M_p of them are also used for supervised phase correction (estimation of ψ). Then, the $M - M_p$ information bits of *user 0* associated with the remaining $M - M_p$ received vectors are detected by (2.10).

The effectiveness of the filter is characterized statistically by the probability distribution of the number of bit errors in a packet:

$$p(i) \triangleq \Pr\{i \text{ bit errors in the packet}\}, \qquad i = 0, 1, \ldots, M - M_p. \tag{2.35}$$

W.l.o.g. we define the bit-error-rate (BER) as the probability of erroneous detection of $b_0(0)$ (the first bit of *user 0* in the packet):

$$\begin{aligned} \text{BER} &\triangleq \Pr\left\{\hat{b}_0(0) \neq b_0(0)\right\} = \sum_{i=0}^{M-M_p} \Pr\left\{\hat{b}_0(0) \neq b_0(0) \middle| i \text{ bit errors in the packet}\right\} p(i) \\ &= \sum_{i=0}^{M-M_p} \frac{i}{M-M_p} p(i) = \frac{1}{M-M_p} \sum_{i=0}^{M-M_p} i\, p(i). \end{aligned} \tag{2.36}$$

It is interesting to note at this point that if the filter $\mathbf{w}$ were independent of the information bit stream $b_0(0)$, $b_0(1)$, ..., then the BER could have been expressed analytically as a weighted sum of the value of the error function $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \mathrm{e}^{t^2/2} dt$ at 2^{K-1} different points [39]. However, this independence assumption does not hold true in our system since the packet data record $\mathbf{x}_0$, $\mathbf{x}_1$, ..., $\mathbf{x}_{M-1}$ (that includes the information bits $b_0(0)$, $b_0(1)$, ..., $b_0(M-1)$ to be detected) is directly utilized for the calculation of $\mathbf{w}$. Therefore, use of the analytical BER expression is not appropriate in our performance evaluation studies for which we can only rely on (2.36).

A data packet is received successfully within a single time slot if the number of errors in the detection of the $M - M_p$ information bits is less than or equal to the maximum number of (correctable) bit errors allowed by the forward-error-correction (FEC) module. If no FEC is present, then the packet is successfully received when all $M - M_p$ information bits are detected correctly. The packet-error-rate (PER) is defined as the probability of receiving an uncorrectable packet within a single time slot and is given by

$$\text{PER}(h) \triangleq \Pr\{\text{more than } h \text{ bit errors in the packet}\} = \sum_{i=h+1}^{M-M_p} p(i) = 1 - \sum_{i=0}^{h} p(i) \tag{2.37}$$

where h is the maximum number of correctable bit errors per packet. By setting $h = 0$ we obtain the PER of a system without FEC: $\text{PER}(0) = \sum_{i=1}^{M-M_p} p(i) = 1 - p(0)$.

To examine the PER performance of the DS-CDMA system described in Section 2.2 equipped with the proposed packet-rate adaptive S-T AV filter receiver, we proceed with an illustration. We consider dedicated Gold user signature assignments of length $L = 31$. We fix the packet size at $M = 256$ bits and we use $M_p = 5$ of them as pilot mid-amble bits. Each user signal experiences $N = 3$ independent Rayleigh fading paths with equal average received energy per path and independent angles of arrival uniformly distributed in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. In all of our studies we consider averages over 20,000 independently drawn multipath Rayleigh fading S-T channels. The receiver antenna array consists of $Z = 4$ elements. With these numbers, the multipath extended S-T product (or length of the adaptive filter) is $Z(L + N - 1) = 132$. For simplicity in our studies, we assume that the packets of all users arrive at the same *average* total power level (affected by independent fading per packet, user, and path). The average total received pre-detection SNR of each user (the sum of the average received SNR's over all paths) is set to 11dB where, following the notation in (2.7), the average total received SNR for user k, $k = 0, 1, \ldots$, is defined by $2E_k \sum_{n=0}^{N-1} E\left\{|c_{k,n}|^2\right\} / N_0$ (we recall that $\frac{N_0}{2}$ is the AWGN power spectral density assumed to be identical for every spatial channel/antenna element).

In Fig. 2.4(a) we set the number of users to $K = 15$ and we plot $p(i)$ for various receivers: *(i)* S-T RAKE, *(ii)* SMI [27], *(iii)* Auxiliary-Vector, and for reference purposes *(iv)* the reduced-rank orthogonal multistage Wiener filter in [49] (with the preferred number of stages $D = 8$ [54]). All filter estimates utilize exactly the same data and for all estimates phase recovery with $M_p = 5$ pilot bits is carried out according to Section 2.4. The AV and S-T RAKE filters are seen to be significantly less vulnerable to bit errors than their multistage and SMI counterparts. We also observe that the probability mass function $p(i)$ is practically zero for the AV filter when $i \geq 6$. In Fig. 2.4(b), we concentrate only on the AV filter and we plot $p(i)$ for the cases of $K = 15$ and $K = 30$ active users. For $K = 30$, the probability mass is shifted gently to the right.

In Fig. 2.5, we study the BER performance of all these receivers as a function of the number of active users K in the system. We also add to this study the multistage Wiener filter [49] with $D = 2$ stages which we found empirically to be the best number of stages for this specific problem. For each receiver and every value of K, the BER is evaluated through (2.36). We immediately observe the disappointing performance of the SMI and multistage ($D = 8$) Wiener filters. We also note the disturbing behavior of the best multistage ($D = 2$) Wiener filter when $K < 4$ due to the implicit matrix inversion implemented through the backward recursion of the multistage algorithm (for $K < 4$ the corresponding reduced-rank data autocorrelation matrix estimate becomes ill-conditioned and its implicit inversion in the multistage algorithm leads to this phenomenon). The superiority of the AV filtering algorithm over all other schemes and over the whole range of the number of users K is apparent.

In Fig. 2.6(a), we plot the PER as a function of the number of active users K using (2.37). No FEC is assumed ($h = 0$). As expected, the relative PER performance results parallel the BER findings in Fig. 2.5. Joint S-T processing of the antenna-array received, multipath combined data has "increased" the effective system processing gain from $L = 31$ to $Z(L + N - 1) = 132$; indeed, it is satisfying to observe that the PER of the S-T RAKE, best multistage Wiener filter ($D = 2$), and AV receivers approaches 1 for user numbers much greater than 31. In a rather more interesting study, Fig. 2.6(b) shows the PER under $h = 4$ FEC.

System BER/PER performance improvements due to the use of the AV receiver under joint S-T data processing allow us to accommodate more users for a set PER constraint (or reduce the transmitting power of the hand-set or increase the range/coverage of the base-station transceiver for a pre-set maximum number of active users). We may express the PER as a function of both the FEC capabilities h and the number of active users K, $\mathrm{PER}(h, K)$. Then, we define the *user capacity* $C(h)$ as the maximum number of active users (packets in a slot) under which a given PER constraint is satisfied:

$$C(h) \triangleq \max\{K : \ \mathrm{PER}(h, K) \leq \lambda\} \tag{2.38}$$

where λ is the PER constraint threshold. The choice of the threshold value λ depends on the specific user application requirements [9], [10]. For example, real-time voice data cannot benefit -in general- from packet retransmission and may require a λ value close to 10^{-2}. Upholding of the PER performance constraint requires that the network controls the number of active users in the system K and enforces $K \leq C(h)$ at all times.

We return to the previous illustration to examine the user capacity of the system for the S-T adaptive receivers under consideration. We set, for example, the PER constraint threshold to $\lambda = 10^{-2}$. From Fig. 2.6(a) we conclude that in the absence of FEC the user capacity of the AV system is 10 and the capacity of the S-T RAKE system is 2. Significant capacity improvement is achieved with 4-bit FEC (Fig. 2.6(b)). The AV system supports 23 users and the S-T RAKE 6. Neither the $D = 8$ multistage Wiener filter nor the SMI receiver can meet the PER constraint for any number of users $K \geq 1$, with or without FEC. The best multistage Wiener filter ($D = 2$) receiver can support 4 to 6 users when no FEC is considered and 3 to 14 users with 4-bit FEC. However, it cannot meet the PER requirement when $K \leq 2$ and it has to be withdrawn as an option for the packet-rate adaptive S-T system under consideration. We conclude that we have two viable solutions: the AV and plain S-T RAKE receiver systems. The capacity gain of the proposed AV system is certainly clear. The AV system allows up to $23/31 \simeq 74\%$ loading for this Gold coded system with signature length 31 and 4-bit FEC (with $Z = 4$ antenna elements and multipath fading reception with 11dB average total pre-detection SNR per user).

Let us now assume that if a data packet is received with only correctable errors a positive ACK is sent back to the user over a different downlink channel (FDD). Once an uncorrectable error is detected in the packet, a NAK is sent back to the mobile which then retransmits after waiting for a random number of time slots. We denote the probability of having K packets transmitted during a time slot interval T_s by $A(K)$. $A(K)$ depends on the traffic model. For example, if potential users (packet contributors) come from an infinite population and the number of users that attempt to transmit a packet during a time slot is Poisson distributed with mean G (offered packet traffic load) [10], [11], then $A(K) = \frac{G^K e^{-G}}{K!}$, $K = 0, 1, 2, \ldots$. If, on the other hand, potential users come from a finite population of size N and each user attempts independently to transmit a packet with probability $\frac{G}{N}$ (G represents again the offered traffic load per slot), then $A(K) = \binom{N}{K} \left(\frac{G}{N}\right)^K \left(1 - \frac{G}{N}\right)^{N-K}$, $0 \leq K \leq N$, [9], [10], [12], [13]. Of course, for large values of N the hard-to-evaluate binomial distribution approaches quickly the Poisson distribution. In either case, when the network system enforces a user capacity cap $K \leq C(h)$ the distribution of the number of packets in a slot

$A(K)$ remains unchanged for $K = 0, 1, \ldots, C(h)-1$, while $A(K = C(h)) = 1 - \sum_{k=0}^{C(h)-1} A(k)$ and $A(K) = 0$ for $K > C(h)$.

Packet throughput is defined as the average number of successful packet transmissions per time slot. Let $\mathrm{PSR}(h, K)$ be the physical layer one-shot packet success rate (that is the probability that a packet is received successfully within a single time slot) with h-bit FEC in the presence of K users. Then, $\mathrm{PSR}(h, K) = 1 - \mathrm{PER}(h, K)$ and the packet throughput of the system S can be expressed as

$$S(h) = \sum_{K=1}^{\infty} K\, A(K)\, \mathrm{PSR}(h, K) = \sum_{K=1}^{\infty} K\, A(K)\, \left(1 - \mathrm{PER}(h, K)\right). \tag{2.39}$$

In Fig. 2.7, we plot the packet throughput $S(h)$ as a function of the offered packet traffic load G for binomial packet arrivals ($N = 33$) without user capacity control. The cases of no FEC ($h = 0$) and 4-bit FEC ($h = 4$) are examined separately in Part (a) and (b), respectively. We see that at the maximum traffic load of 33 packets, $\frac{25.6}{33}100\% \simeq 77.6\%$ of the packets in the AV system go through successfully on the average when $h = 0$. This number increases to $\frac{31.9}{33}100\% \simeq 96.7\%$ when $h = 4$. In Fig. 2.8 we repeat the exact same study of Fig. 2.7 for a network system that enforces user capacity control (number of packets K in each system slot is less than or equal to $C(h)$ to guarantee $\mathrm{PER}(h, K)$ quality $\lambda = 10^{-2}$ or better at all times). As expected (cf. (2.39)), the packet throughput curves converge quickly to $C(h)\mathrm{PSR}(h, C(h))$ as G increases. When $h = 0$, the limit throughput value is 9.9 packets per slot for the AV system and 2.0 packets per slot for the S-T RAKE system. When $h = 4$, the limit throughput values become 22.8 and 5.9 packets per slot, correspondingly. Fig. 2.8 quantifies the potential contribution of AV receivers in networks with user capacity control. The study of the important trade-off between the PER quality constraint λ and the limit throughput $C(h)\mathrm{PSR}(h, C(h))$ is beyond the scope of this work.

As a final study, we consider the network delay metric $D(h)$ defined as the average number of slots from the time a packet is presented to the network until it is successfully received. Markov-chain analysis techniques lead us to the following expression for the network delay of a system with user capacity cap $C(h)$.[4] The proof is given in the Appendix.

$$D(h) = \delta_0 + \frac{\delta_1 \left(\sum_{K=1}^{C(h)} A(K)\mathrm{PER}(h, K) + \sum_{K=C(h)+1}^{\infty} A(K) \frac{C(h)}{K} \mathrm{PER}(h, C(h)) \right) + \delta_2 \sum_{K=C(h)+1}^{\infty} A(K) \left(1 - \frac{C(h)}{K}\right)}{1 - \sum_{K=1}^{C(h)} A(K)\mathrm{PER}(h, K) - \sum_{K=C(h)+1}^{\infty} A(K) \frac{C(h)}{K} \mathrm{PER}(h, C(h)) - \sum_{K=C(h)+1}^{\infty} A(K) \left(1 - \frac{C(h)}{K}\right)} \tag{2.40}$$

where δ_0, δ_1, and δ_2 denote, correspondingly, the mean delay in slot units for the execution of *(i)* a single transmission with successful reception, *(ii)* a retransmission due to bit errors, and *(iii)* a retransmission due to rejection by the capacity controller (when more than $C(h)$ packets are presented for potential transmission during a slot). It is easy to verify that if we let $C(h) \to \infty$ in our delay expression in (2.40) we obtain the familiar delay expression for links without user capacity control [9], [10], [55]:

$$D(h) = \delta_0 + \frac{\delta_1 \sum_{K=1}^{\infty} A(K)\mathrm{PER}(h, K)}{1 - \sum_{K=1}^{\infty} A(K)\mathrm{PER}(h, K)} = \delta_0 + \delta_1 \left(\frac{1}{E\left\{\mathrm{PSR}(h, K)\right\}} - 1 \right). \tag{2.41}$$

Fig. 2.9 presents the network delay $D(h)$ as a function of the throughput $S(h)$ for a system without user capacity control and binomial ($N = 33$) arrivals. The mean retransmission delay is set to $\delta_1 = 3$ slots and the mean successful transmission delay is $\delta_0 = 1.5$ slots.[5] No FEC ($h = 0$) is examined in Fig. 2.9(a) and 4-bit FEC is examined in Fig. 2.9(b). As desired for packetized voice applications, the network delay plots indicate that the system maintains a relatively constant delay as the throughput approaches its maximum

[4] This delay analysis assumes a simple stop-and-wait ARQ protocol. Certainly, for large loads the transmission delay may be reduced by means of a more efficient selective retransmission protocol which enables continuous transmission.

[5] It is assumed that the minimum delay for slotted systems is 1.5 slots [10] which includes the average time interval between the presentation of a packet to the transmitter and the beginning of the next slot.

value (at $G = N = 33$). In Fig. 2.10, we repeat the study of Fig. 2.9 for a system that enforces capacity control with mean "rejection delay" $\delta_2 = 1$ slot. The plots capture again the possible throughput-delay operating points and help us visualize how the delay $D(h)$ in (2.40) increases when the capacity-controlled throughput $S(h)$ converges to $C(h)\text{PSR}(h, C(h))$ for high traffic loads G.

2.6 Summary and Conclusions

We considered a random-access wireless DS-CDMA multipath fading AWGN link with dedicated signature assignments. We suggested packet length design that is sufficiently short to conform with the coherence time of the link (quasi-static fading over the duration of a packet). Then, we developed a complete adaptive antenna array CDMA linear filter receiver design that adapts itself and detects the transmitted information bits on an individual packet-by-packet basis. Effective short-data-record (packet-rate) filter adaptation/estimation and bit detection was made possible due to the proposed auxiliary-vector (AV) filtering algorithm. The latter incorporated seamlessly packet-rate blind subspace-based space-time channel estimation and supervised recovery of the space-time channel phase through the use of a few packet mid-amble pilot bits. Our studies showed that very limited mid-amble pilot signaling (of the order of $5/256 \simeq 2\%$) can be sufficient for phase recovery and effective adaptive receiver design. Therefore, differential modulation to overcome the phase ambiguity problem is not a necessity.

Bit-error-rate, packet-error-rate, and user capacity studies and comparisons demonstrated state-of-the-art performance of the proposed linear adaptive AV receiver system. Through the development of the probability mass distribution of the bit errors in a packet we were able to translate these findings to packet throughput results (with or without user capacity control). We found for example that for a multipath Rayleigh fading AWGN link, 11dB average total received pre-detection SNR per user, four antenna elements, and 4-bit FEC the proposed adaptive AV receiver system offers $\frac{22.8-5.9}{5.9}100\% \simeq 286\%$ improvement in maximum packet throughput with user capacity control over the corresponding system that employs conventional space-time RAKE reception. Network delay derivations and studies showed a slow, well-behaved increase of the AV system delay for throughput up to its maximum value. In this context, the packet-rate adaptive receiver design developed using the AV filtering principles coupled with FEC techniques seems to provide an effective tool to improve the performance of random access DS-CDMA mobile communication links.

2.7 Appendix

We denote by R_1 and R_2 the total number of retransmissions of a packet until successful reception due to bit errors and due to rejection by the capacity controller (when more than $C(h)$ packets are presented for potential transmission during a slot), respectively. The joint probability mass function (pmf) of R_1 and R_2 is given by

$$\Pr\{R_1 = i,\ R_2 = j\} = \binom{i+j}{j} \rho_1^i \rho_2^j (1 - \rho_1 - \rho_2), \qquad i, j = 0, 1, 2, \ldots, \tag{2.42}$$

where ρ_1 and ρ_2 denote the probability that a retransmission is needed due to bit errors and due to rejection by the capacity controller, respectively:

$$\rho_1 = \sum_{K=1}^{C(h)} A(K)\text{PER}(h, K) + \sum_{K=C(h)+1}^{\infty} A(K)\frac{C(h)}{K}\text{PER}(h, C(h)) \tag{2.43}$$

and

$$\rho_2 = \sum_{K=C(h)+1}^{\infty} A(K)\left(1 - \frac{C(h)}{K}\right). \tag{2.44}$$

From (2.42), the marginal pmf of R_1 is

$$\Pr\{R_1 = i\} = \sum_{j=0}^{\infty} \Pr\{R_1 = i,\ R_2 = j\} = \rho_1^i (1 - \rho_1 - \rho_2) \sum_{j=0}^{\infty} \binom{i+j}{j} \rho_2^j. \tag{2.45}$$

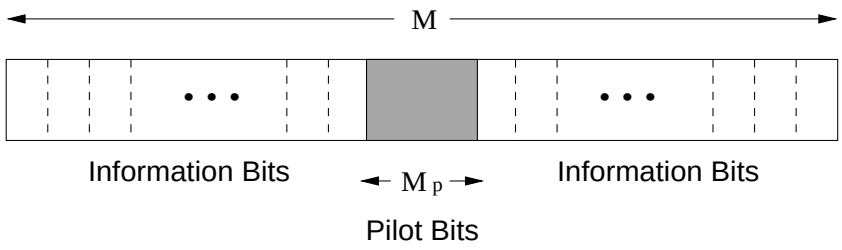

Fig. 2.1. Data packet structure of total length M bits that contains a mid-amble of M_p pilot bits.

For any $\rho_2 \in (0,1)$, $\sum_{j=0}^{\infty} \rho_2^j = \frac{1}{1-\rho_2}$; we differentiate both sides of this equality i times, $i = 0, 1, 2, \ldots$, to obtain $\sum_{j=0}^{\infty} \frac{(i+j)!}{j!} \rho_2^j = \frac{i!}{(1-\rho_2)^{i+1}}$ or

$$\sum_{j=0}^{\infty} \binom{i+j}{j} \rho_2^j = \frac{1}{(1-\rho_2)^{i+1}}, \qquad i = 0, 1, 2, \ldots . \tag{2.46}$$

From (2.45) and (2.46) we conclude that

$$\Pr\{R_1 = i\} = \rho_1^i (1 - \rho_1 - \rho_2) \frac{1}{(1-\rho_2)^{i+1}} = \left(\frac{\rho_1}{1-\rho_2}\right)^i \left(1 - \frac{\rho_1}{1-\rho_2}\right), \qquad i = 0, 1, 2, \ldots . \tag{2.47}$$

Therefore, R_1 is geometrically distributed with parameter $\frac{\rho_1}{1-\rho_2} \in (0,1)$ and mean

$$E\{R_1\} = \frac{\frac{\rho_1}{1-\rho_2}}{1 - \frac{\rho_1}{1-\rho_2}} = \frac{\rho_1}{1 - \rho_1 - \rho_2}. \tag{2.48}$$

Similarly, following the same derivation steps for the random variable R_2 we conclude that

$$E\{R_2\} = \frac{\rho_2}{1 - \rho_1 - \rho_2}. \tag{2.49}$$

Then, the network delay is evaluated as follows:

$$D(h) = \delta_0 + \delta_1 E\{R_1\} + \delta_2 E\{R_2\} = \delta_0 + \delta_1 \frac{\rho_1}{1 - \rho_1 - \rho_2} + \delta_2 \frac{\rho_2}{1 - \rho_1 - \rho_2} = \delta_0 + \frac{\delta_1 \rho_1 + \delta_2 \rho_2}{1 - \rho_1 - \rho_2} \tag{2.50}$$

which coincides with (2.40) (ρ_1 and ρ_2 are given by (2.43) and (2.44), respectively). □

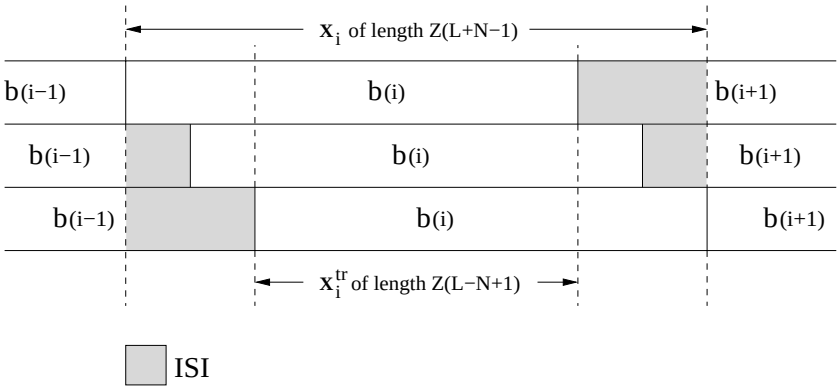

Fig. 2.2. Data collection and ISI trimming.

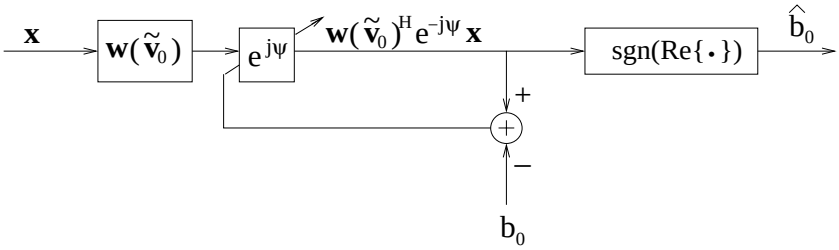

Fig. 2.3. Phase correction for the S-T linear filter $\mathbf{w}(\tilde{\mathbf{v}}_0)$.

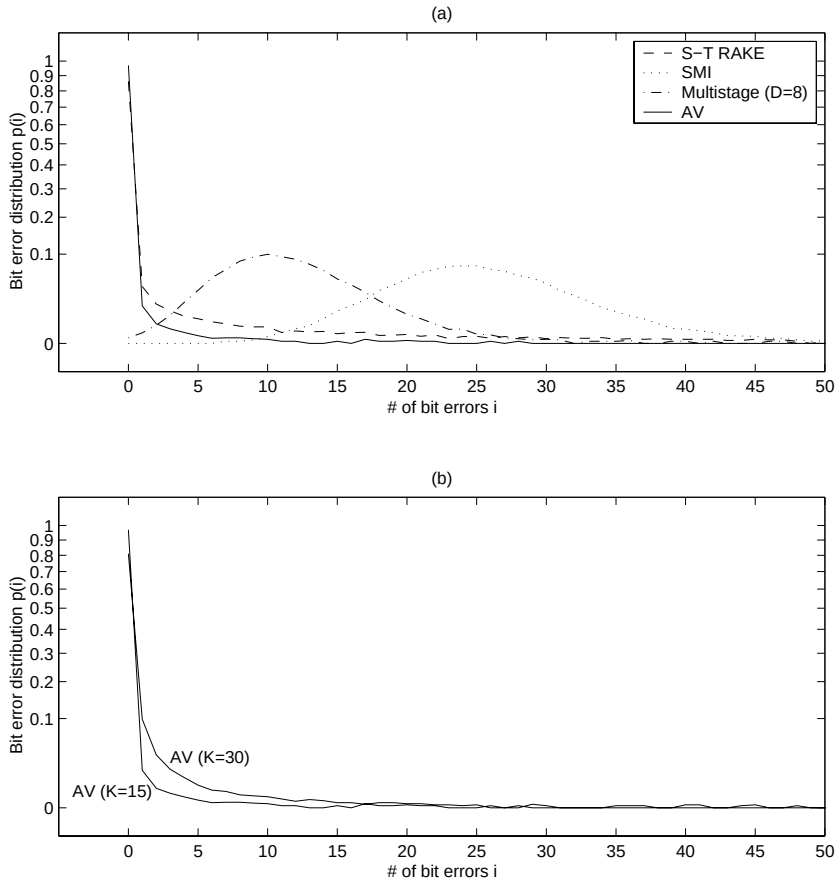

Fig. 2.4. Distribution $p(i)$ of number of bit errors per packet for *(a)* the S-T RAKE, SMI, multistage Wiener filter ($D = 8$), and AV receivers in the presence of $K = 15$ users and *(b)* the AV receiver when $K = 15$ or $K = 30$.

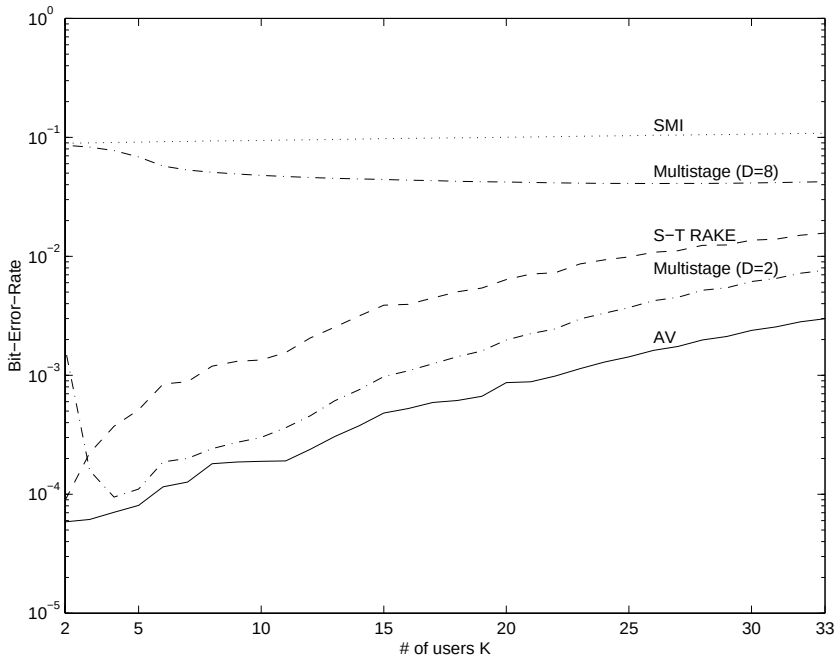

Fig. 2.5. BER versus number of active users K for the SMI, multistage Wiener filter with the preferred number of stages ($D = 8$), S-T RAKE, best multistage Wiener filter ($D = 2$), and AV receivers.

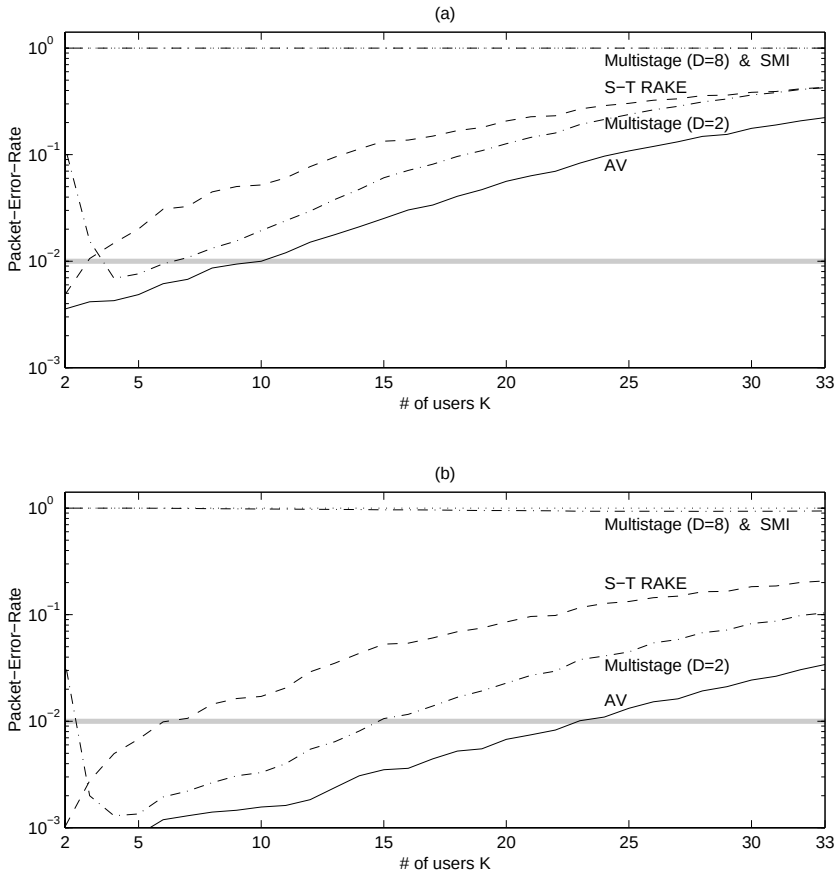

Fig. 2.6. PER versus number of active users K for a system with *(a)* no FEC and *(b)* 4-bit FEC.

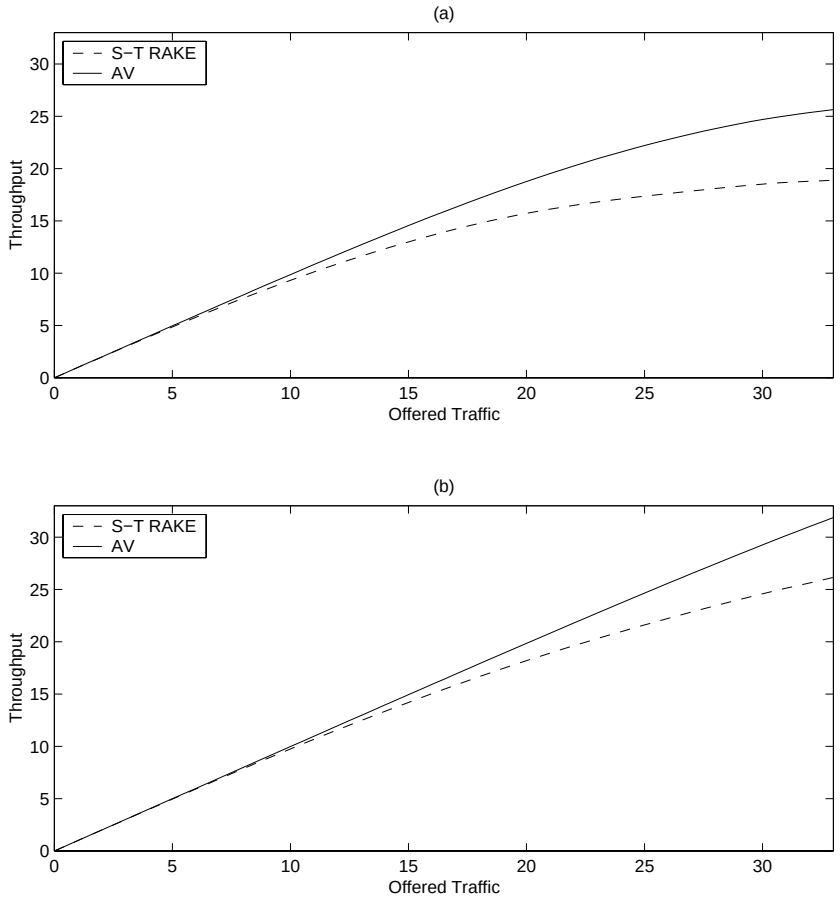

Fig. 2.7. Packet throughput versus offered packet traffic load without capacity control: *(a)* no FEC, *(b)* 4-bit FEC.

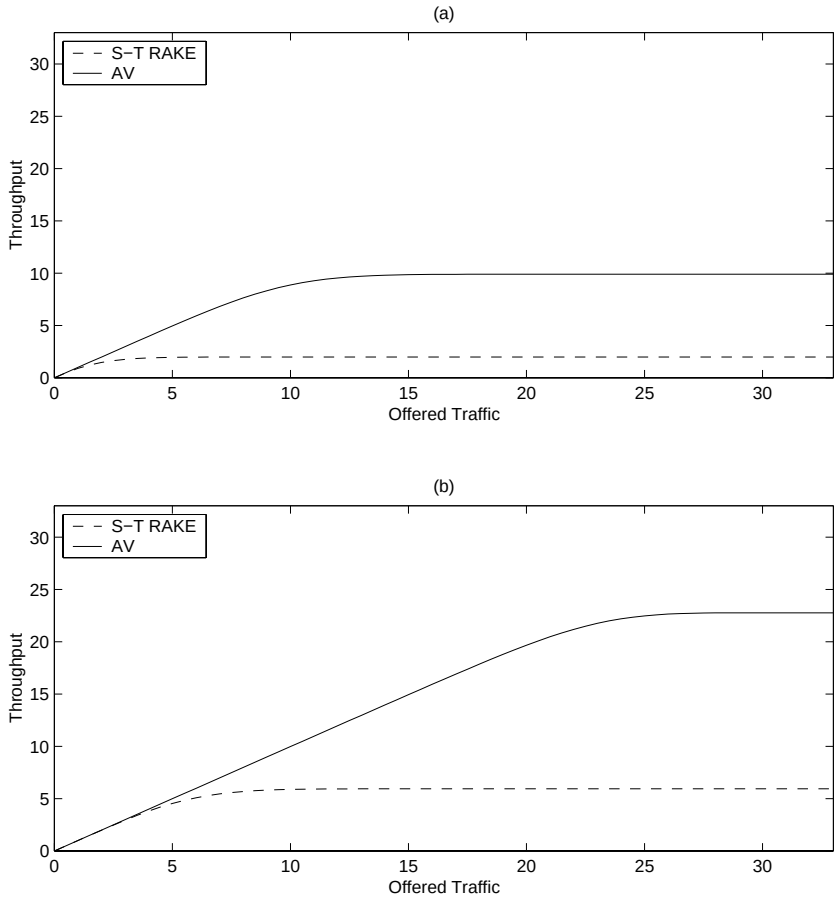

Fig. 2.8. Packet throughput versus offered traffic with capacity control ($K \leq C(h)$): *(a)* no FEC, *(b)* 4-bit FEC.

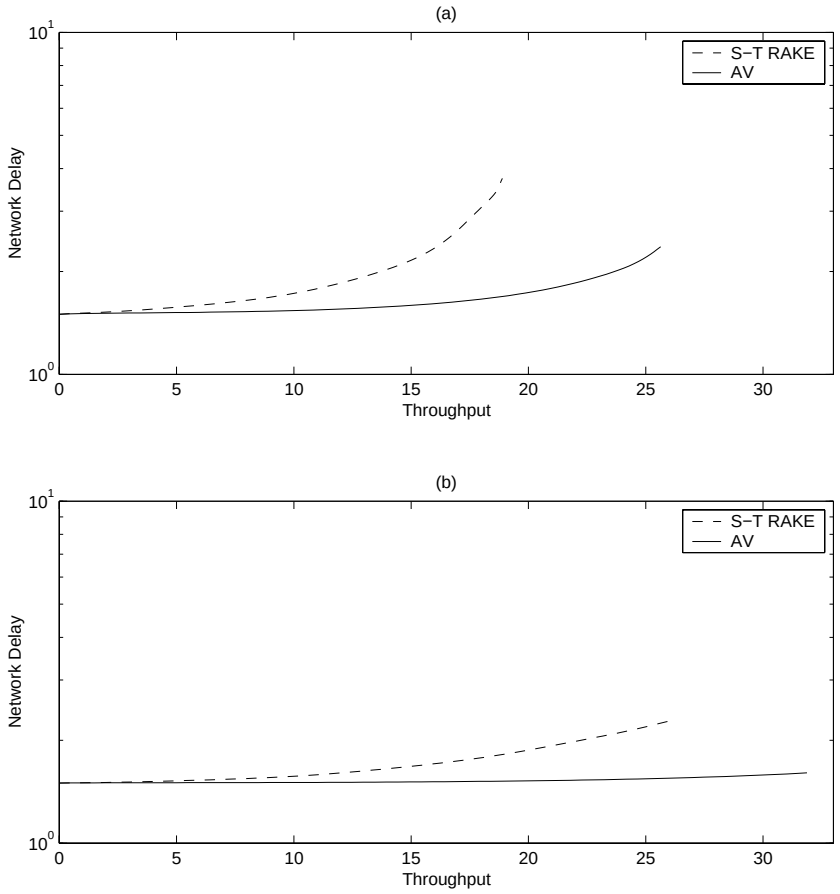

Fig. 2.9. Network delay (in slots) versus throughput for a system without capacity control: *(a)* no FEC, *(b)* 4-bit FEC.

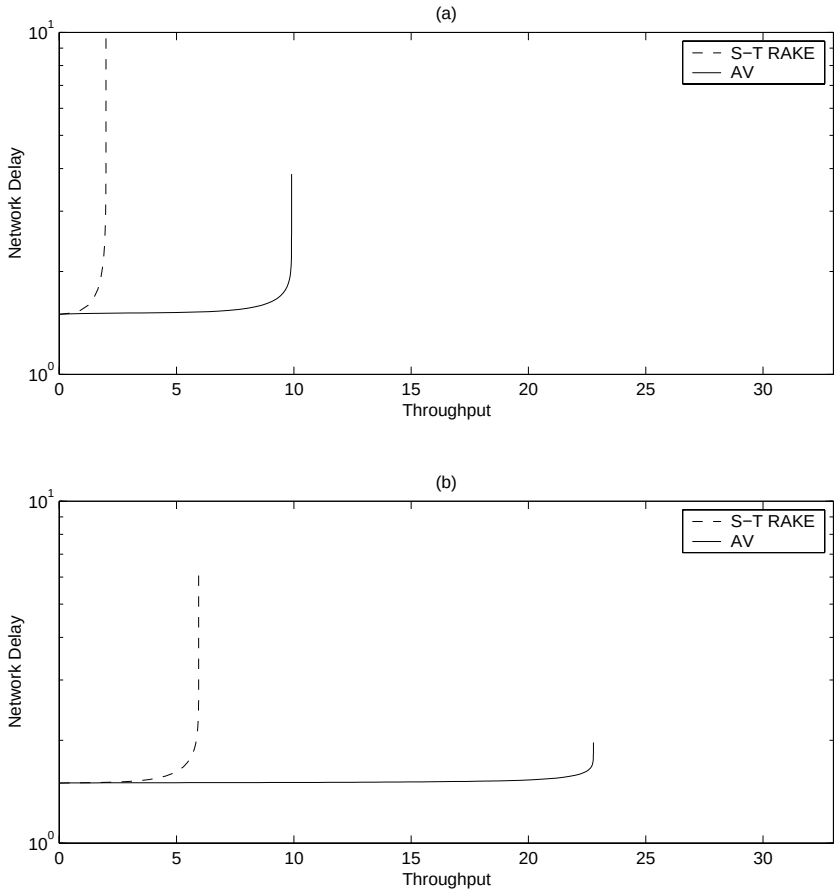

Fig. 2.10. Network delay (in slots) versus throughput for a system with capacity control: *(a)* no FEC, *(b)* 4-bit FEC.

Bibliography

[1] J. H. Lodge, "Mobile satellite-communications systems: Toward global personal communications," *IEEE Comm. Mag.*, vol. 29, pp. 24-30, Nov. 1991.

[2] B. Sarikaya, "Packet mode in wireless networks: Overview of transition to third generation," *IEEE Comm. Mag.*, vol. 38, pp. 164-172, Sept. 2000.

[3] N. Abramson, "The ALOHA system–Another alternative for computer communications," in *Proc. Fall Joint Computer Conf. of American Federation of Info. Processing Soc. (AFIPS)*, vol. 37, Montvale, NJ, Nov. 1970, pp. 281-285.

[4] D. Raychaudhuri, "Performance analysis of random access packet-switched code division multiple access systems," *IEEE Trans. Commun.*, vol. 29, pp. 895-901, June 1981.

[5] A. B. Carleial and M. E. Hellman, "Bistable behavior of ALOHA-type systems," *IEEE Trans. Commun.*, vol. 23, pp. 401-410, Apr. 1975.

[6] L. Kleinrock and S. S. Lam, "Packet switching in a multiaccess broadcast channel: Performance evaluation," *IEEE Trans. Commun.*, vol. 23, pp. 410-423, Apr. 1975.

[7] K. G. Johanssen, "Code division multiple access channel capacity in mobile satellite communication," *IEEE Trans. Vehic. Tech.*, vol. 39, pp. 17-26, Feb. 1990.

[8] R. K. Morrow, Jr., and J. Lehnert, "Packet throughput in slotted ALOHA DS/SSMA radio systems with random signature sequences," *IEEE Trans. Commun.*, vol. 40, pp. 1223-1230, July 1992.

[9] C. A. F. J. Wijffels, H. S. Misser, and R. Prasad, "A micro-cellular CDMA system over slow and fast Rician fading radio channels with forward error correcting coding and diversity," *IEEE Trans. Vehic. Tech.*, vol. 42, pp. 570-580, Nov. 1993.

[10] R. Prasad, *CDMA for Wireless Personal Communications*. Boston, MA: Artech House, 1996.

[11] M. Saito, H. Okada, T. Sato, T. Yamazato, M. Katayama, and A. Ogawa, "Throughput improvement of CDMA slotted ALOHA systems," *IEICE Trans. Commun.*, vol. E80B, pp. 74-80, Jan. 1997.

[12] O. Sallent and R. Agustí, "A proposal for an adaptive S-ALOHA access system for a mobile CDMA environment," *IEEE Trans. Vehic. Tech.*, vol. 47, pp. 977-986, Aug. 1998.

[13] O. Sallent and R. Agustí, "Adaptive S-ALOHA CDMA as an alternative way of integrating services in mobile environments," *IEEE Trans. Vehic. Tech.*, vol. 49, pp. 936-947, May 2000.

[14] T.-W. Ban, J.-C. Lee, S.-M. Lee, H.-W. Jeon, Y.-Z. Cho, and J.-S. Song, "A fair access algorithm for packet data service in DS/CDMA-based slotted-ALOHA system," in *Proc. IEEE Vehic. Tech. Conf.*, Tokyo, Japan, May 2000, vol. 2, pp. 1160-1164.

[15] K. W. Halford and M. Brandt-Pearce, "New-user identification in a CDMA system," *IEEE Trans. Commun.*, vol. 46, pp. 144-155, Jan. 1998.

[16] A. Yener and R. D. Yates, "Multiuser access detection for CDMA systems," in *Proc. Conf. on Inform. Sc. and Syst.*, Princeton, NJ, Mar. 1998, vol. 1, pp. 17-22.

[17] A. Yener and R. D. Yates, "Multiuser access capability of packet switched CDMA systems," in *Proc. IEEE Vehic. Tech. Conf.*, Houston, TX, May 1999, vol. 3, pp. 1846-1850.

[18] T. K. Liu and J. A. Silvester, "Joint admission congestion control for wireless CDMA systems supporting integrated services," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 845-857, Aug. 1998.

[19] H. Bischl and E. Lutz, "Packet error rate in the non-interleaved Rayleigh channel," *IEEE Trans. Commun.*, vol. 43, pp. 1375-1382, Apr. 1995.

[20] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, and L. A. Weaver, "Increased capacity using CDMA for mobile satellite communication," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 503-514, May 1990.

[21] R. Prasad, R. D. J. van Nee, and R. N. van Wolfswinkel, "Performance analysis of multiple access technique for land-mobile satellite communication," in *Proc. IEEE Globecom 1994*, San Francisco, CA, Nov. 1994, pp. 740-744.

[22] D. Makrakis and K. M. S. Murthy, "Spread slotted ALOHA techniques for mobile and personal satellite communication-systems," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 985-1002, Aug. 1992.

[23] P. B. Rapajic and B. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 685-697, May 1994.

[24] R. D. J. van Nee, R. N. van Wolfswinkel, and R. Prasad, "Slotted ALOHA and code-division multiple-access techniques for land-mobile satellite personal communications," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 382-388, Feb. 1995.

[25] X. Wu and A. Haimovich, "Space-time processing for CDMA communications," in *Proc. Conf. on Inform. Sc. and Syst.*, Baltimore, MD, Mar. 1995, pp. 371-376.

[26] D. A. Pados and S. N. Batalama, "Joint space-time auxiliary-vector filtering for DS/CDMA systems with antenna arrays," *IEEE Trans. Commun.*, vol. 47, pp. 1406-1415, Sept. 1999.

[27] I. S. Reed, J. D. Mallet, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerospace and Electr. Syst.*, vol. 10, pp. 853-863, Nov. 1974.

[28] B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode, "Adaptive antenna systems," *Proceedings IEEE*, vol. 55, pp. 2143-2158, Dec. 1967.

[29] R. L. Plackett, "Some theorems in least squares," *Biometrika*, vol. 37, p. 149, 1950.

[30] R. A. Wiggins and E. A. Robinson, "Recursive solution to the multichannel filtering problem," *Journ. Geophys. Res.*, vol. 70, pp. 1885-1891, 1965.

[31] D. A. Pados and S. N. Batalama, "Low-complexity blind detection of DS/CDMA signals: Auxiliary-vector receivers," *IEEE Trans. Commun.*, vol. 45, pp. 1586-1594, Dec. 1997.

[32] S. N. Batalama, M. J. Medley, and D. A. Pados, "Robust adaptive recovery of spread-spectrum signals with short data records," *IEEE Trans. Commun.*, vol. 48, pp. 1725-1731, Oct. 2000.

[33] D. A. Pados and G. N. Karystinos, "An iterative algorithm for the computation of the MVDR filter," *IEEE Trans. Signal Proc.*, vol. 49, pp. 290-300, Feb. 2001.

[34] H. Qian and S. N. Batalama, "Data-record-based criteria for the selection of an auxiliary-vector estimator of the MVDR filter," in *Proc. Asilomar Conf. Signals, Syst., Computers*, Pacific Grove, CA, Oct. 2000, pp. 802-807.

[35] H. Qian and S. N. Batalama, "Data-record-based criteria for the selection of an auxiliary-vector estimator of the MMSE/MVDR filter," *IEEE Trans. Commun.*, vol. 51, pp. 1700-1708, Oct. 2003.

[36] S. E. Bensley and B. Aazhang, "Subspace-based channel estimation for code division multiple access communication systems," *IEEE Trans. Commun.*, vol. 44, pp. 1009-1020, Aug. 1996.

[37] P. Chaudhury, W. Mohr, and S. Onoe, "The 3GPP proposal for IMT-2000," *IEEE Comm. Mag.*, vol. 37, pp. 72-81, Dec. 1999.

[38] M. Saito, T. Yamazato, M. Katayama, and A. Ogawa, "New quasi-synchronous sequences for CDMA slotted ALOHA systems," *IEICE Trans. Fund. Elec. Comm. and Comp.*, vol. 11, pp. 2274-2280, Nov. 1998.

[39] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 3rd ed., 1995.

[40] A. Kansal, S. N. Batalama, and D. A. Pados, "Adaptive maximum SINR RAKE filtering for DS-CDMA multipath fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1765 -1773, Dec. 1998.

[41] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[42] A. Napolitano and M. Tanda, "Blind parameter estimation in multiple-access systems," *IEEE Trans. Commun.*, vol. 49, pp. 688-698, Apr. 2001.

[43] Z.-S. Liu, J. Li, and S. L. Miller, "An efficient code-timing estimator for receiver diversity DS-CDMA systems," *IEEE Trans. Commun.*, vol. 46, pp. 826-835, June 1998.

[44] C. Sengupta, J. R. Cavallaro, and B. Aazhang, "On multipath channel estimation for CDMA systems using multiple sensors," *IEEE Trans. Commun.*, vol. 49, pp. 543-553, Mar. 2001.

[45] D. Gesbert, J. Sorelius, P. Stoica, and A. Paulraj, "Blind multiuser MMSE detector for CDMA signals in ISI channels," *IEEE Comm. Letters*, vol. 3, pp. 233-235, Aug. 1999.

[46] P. Loubaton and E. Moulines, "On blind multiuser forward link channel estimation by the subspace method: Identifiability results," *IEEE Trans. Signal Proc.*, vol. 48, pp. 2366-2376, Aug. 2000.

[47] H. Liu and K. Li, "A decorrelating RAKE receiver for CDMA communications over frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1036-1045, July 1999.

[48] 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, pp. 985-998, June 2001.

[49] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2943-2959, Nov. 1998.

[50] O. Ozdemir and M. Torlak, "Reduced-rank advanced RAKE receivers for dispersive CDMA signals," in *Proc. IEEE Vehic. Tech. Conf.*, Atlantic City, NJ, Oct. 2001, vol. 1, pp. 409-413.

[51] S. Buzzi and H. V. Poor, "Channel estimation and multiuser detection in long-code DS/CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1476-1487, Aug. 2001.

[52] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 2nd ed., 1991.

[53] K. J. Quirk and L. B. Milstein, "The effect of low SNR in phase estimation in wideband CDMA," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 107-120, Jan. 2001.

[54] M. L. Honig and W. Xiao, "Performance of reduced-rank linear interference suppression," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1928-1946, July 2001.

[55] L. Kleinrock and F. A. Tobagi, "Packet switching in radio channels: Part 1–Carrier-sense multiple-access modes and their throughput-delay characteristics," *IEEE Trans. Commun.*, vol. 23, pp. 1400-1416, Dec. 1975.

Chapter 3

Fast converging minimum probability of error neural network receivers for DS-CDMA communications

3.1 Introduction

Direct-sequence code-division-multiple-access $\stackrel{\triangle}{=}$(DS-CDMA) systems have received considerable interest in response to an ever-increasing demand for better utilization of the available resources in communications environments. In DS-CDMA communication systems, the target operation of the receiver is the detection of the transmitted information bits of one (mobile-end) or more (base station) users.

For the case of single-user detection in DS-CDMA systems, the optimum decision boundary for the additive white Gaussian noise (AWGN) multiple-access channel is nonlinear [1]. As such, multi-layer perceptron (MLP) neural networks (NNs) are particularly suitable for the detection of DS-CDMA signals since they are capable of approximating arbitrary hypersurfaces in the input space. MLPs as well as other neural network structures for single or multi-user detection were considered in [2]-[11]. Independently of the *design* optimality criteria that are used to determine the parameters of the receiver structure, the ultimate *performance* measure of interest is the probability of error in detecting the transmitted information bit of each user (also referred to as bit-error-rate or BER). This is exactly the motivation for this present work.

In recent literature, minimum BER (MBER) optimization in the context of *linear* receiver structures was considered in [12]-[14] where stochastic gradient adaptive algorithms were proposed that approximate the linear minimum BER single-user DS-CDMA receiver. The work in [14] was extended in [15] to a stochastic gradient adaptive "least error rate" algorithm for training two-class radial basis function (RBF) network classifiers. A Newton-type recursive minimum-BER *linear* detector that requires knowledge of all system users and noise statistics was reported in [16].

Motivated by the inherent nonlinearity of the single user optimum decision boundary for AWGN multiple access channels, in this chapter we focus on MLP neural networks and their potential application to single-user detection. We assume a synchronous DS-CDMA system and we consider the problem of single-user detection in the presence of *unknown* multiuser interference and additive channel noise. Our receiver structure of choice consists of an adaptive MLP-NN filter front-end followed by a sign detector. There are three key contributions in this chapter that are summarized below. *(i)* The NN parameters (weights and biases) are adapted online in a way that minimizes *directly* the induced probability of error. Although an analytic closed-form NN filter solution for the adopted minimum-probability-of-error criterion is certainly not attainable in general, this mathematical intractability is bypassed using results from Stochastic Approximation theory that allow recursive probability of error minimization. Relative to other previously proposed interference suppression approaches, the major advantage of a minimum probability of error scheme is the incorporation of the performance evaluation measure of choice *directly* into the learning process. *(ii)* Based on two key symmetry properties exhibited by the optimum decision boundary of the user of interest we identify two principal characteristics of the MLP structure under consideration and translate them into constraints that

are incorporated in the MBER design procedure of the network weights and biases. The end result is a significant reduction of the number of the network parameters to be trained which implies faster network convergence. We note that the NN constraints are independent of the specifics of MBER optimization and can be adopted by *any* MLP training algorithm to accelerate convergence (e.g. [2]-[11]). *(iii)* We embed importance sampling (IS) principles directly into the receiver optimization process. This is in sharp contrast to the traditional use of IS for system performance evaluation only. In addition and contrary to the traditional IS assumption of known input statistics, the proposed scheme does not require such knowledge. The proposed method utilizes only a short pilot sequence and in effect creates virtual high BER receiver conditions that accelerate the adaptation (training) process.

The rest of the chapter is organized as follows. In Section 3.2 we present briefly the DS-CDMA signal model and we discuss the properties of the optimum single-user decision boundary. Constrained neural network optimization and, in particular, constrained minimum BER optimization are the subjects of Sections 3.3 and 3.3, respectively. Importance-sampling-accelerated MBER neural network optimization is discussed in Section 3.5. Simulation studies that illustrate the merits of the proposed scheme are included in Section 3.6. Finally a few concluding remarks are drawn in Section 3.7.

3.2 Optimum Single-User Detection

We consider a K-user synchronous DS-CDMA communication system with processing gain L. Extension to asynchronous systems is straightforward[1]. We assume that each user is assigned a normalized signature vector $\mathbf{s}_k$, $k = 0,1,\ldots,K\text{-}1$. The binary antipodal ($+1$ or -1) information bits are assumed to be independent, identically distributed and equiprobable. After conventional chip-matched filtering and chip-rate sampling, the baseband received signal vector over one information bit period takes the form

$$\mathbf{r} = \sum_{k=0}^{K-1} \sqrt{E_k} d_k \mathbf{s}_k + \mathbf{n} \tag{3.1}$$

where E_k and d_k denote the bit energy and the information bit of user k, respectively. The vector $\mathbf{n}$ represents AWGN with mean $\mathbf{0}$ and autocovariance matrix $\sigma^2 \mathbf{I}_L$ where σ^2 is the channel noise variance and $\mathbf{I}_L$ is the $L \times L$ identity matrix.

Using the total probability theorem we can show that the probability density function (pdf) $f(\mathbf{r})$ of the random vector $\mathbf{r}$ in (3.1) is a mixture of 2^K multivariate Gaussian distributions, i.e.

$$f(\mathbf{r}) = \frac{1}{2^K} \frac{1}{\sqrt{2\pi}\sigma} \sum_{i=1}^{2^K} \exp\left(-\frac{1}{2\sigma^2} \left\| \mathbf{r} - \sum_{k=0}^{K-1} \sqrt{E_k} d_k^{(i)} \mathbf{s}_k \right\|^2 \right) \tag{3.2}$$

where $d_k^{(i)}$ is the bit of the kth user in the ith *possible bit combination*, $i = 1,\ldots,2^K$ [1]. The minimum BER detector of the user of interest (*user 0*) decides in favor of the information bit d_0 that maximizes the conditional pdf of the received vector $\mathbf{r}$ given that d_0 was transmitted [1]:

$$\hat{d}_0 = \arg \max_{d_0 \in \{\pm 1\}} f_{\mathbf{r}|d_0}(\mathbf{r}). \tag{3.3}$$

By considering the particular form of $f(\mathbf{r})$ in (3.2) we may rewrite the optimum DS-CDMA detector of (3.3) in the following way:

$$\sum_{i=1}^{2^{K-1}} \exp\left(-\frac{\left\| \mathbf{r} - \left(\sqrt{E_0}\mathbf{s}_0 + \sum_{k=1}^{K-1} d_k^{(i)} \sqrt{E_k} \mathbf{s}_k \right) \right\|^2}{2\sigma^2} \right) \underset{\text{decide } -1}{\overset{\text{decide } \triangleq +1}{\gtrless}} \sum_{i=1}^{2^{K-1}} \exp\left(-\frac{\left\| \mathbf{r} - \left(-\sqrt{E_0}\mathbf{s}_0 + \sum_{k=1}^{K-1} d_k^{(i)} \sqrt{E_k} \mathbf{s}_k \right) \right\|^2}{2\sigma^2} \right). \tag{3.4}$$

[1]For the case of one-shot detection, an asynchronous DS-CDMA system with K users can be modeled as a synchronous system with $2K-1$ virtual users [17].

The optimum (minimum BER) single-user detector in (3.4) requires knowledge of the energy terms E_1, E_2, ..., $E_{K\text{-}1}$ and signatures $\mathbf{s}_1$, $\mathbf{s}_2$, ..., $\mathbf{s}_{K\text{-}1}$ of all interfering users as well as knowledge of the noise power σ^2 (E_0 and $\mathbf{s}_0$ are also supposed to be known). The computational complexity of the decision rule in (3.4) is, in general, exponential in the number of users K.

As a general comment, the optimum single-user detector in (3.4) is nonlinear and the optimum decision boundary is a hypersurface whose characteristics are determined by all user signatures and energies as well as the additive noise power. The single-user optimum decision boundary does have, however, some properties that *are not system-specific* and can be useful in the design of MLP neural networks that attempt to approximate this boundary [18]. For the sake of completeness, we repeat below Part (i) of Proposition 3.1 in [18] which will serve as the basis for the developments of this work.

Proposition 3.1 *Let $\mathcal{B}$ be the set of all vectors in $\mathbb{R}^L$ that define the optimum single-user decision boundary:*

$$\mathcal{B} \triangleq \{\mathbf{r} \in \mathbb{R}^L : \ f_{\mathbf{r}|d_0=1}(\mathbf{r}) = f_{\mathbf{r}|d_0=-1}(\mathbf{r})\}. \tag{3.5}$$

Then,
(i) $\mathbf{0} \in \mathcal{B}$ and
(ii) the elements of $\mathcal{B}$ are pairwise opposite, i.e.

$$\mathbf{r} \in \mathcal{B} \quad \text{if and only if} \quad -\mathbf{r} \in \mathcal{B}. \tag{3.6}$$

□

Proposition 3.1 essentially implies that the optimum decision boundary passes through the origin $\mathbf{0}$ and is also symmetric with respect to the origin. This property is independent of the system setup and can be useful in practice when the system parameters are unknown and the optimum decision boundary has to be approximated (estimated). In the next section, we exploit this property in the form of a set of constraints that will be incorporated in the receiver adaptation process.

3.3 Constrained NN Optimization

We consider a 3-layer (input layer, hidden layer, and output layer) feedforward fully connected MLP-NN with an odd number of hidden neurons and a single output neuron followed by a sign detector (Fig. 3.1). The input and hidden layers consist of L and $2H+1$ neurons, respectively. With respect to the hidden layer, $w_{i,j}$ denotes the weight associated with the connection between the jth input neuron and the ith hidden neuron, $j = 1, 2, \ldots, L$, $i = -H, (-H+1), \ldots, 0, 1, \ldots, H$, while b_i denotes the bias associated with the ith hidden neuron, $i = -H, (-H+1), \ldots, 0, 1, \ldots, H$. With respect to the output layer (single output neuron), v_i is the weight associated with the connection between the ith hidden neuron and the output neuron, $i = -H, (-H+1), \ldots, 0, 1, \ldots, H$, while b_{out} is the bias of the output neuron. We assume that all (hidden and output) neurons employ the same (symmetric) activation function φ that exhibits the following properties: *(i)* $\varphi(x) > 0$ for $x > 0$ and *(ii)* $\varphi(-x) = -\varphi(x)$ for any x. For notational simplicity we combine all weights associated with the connections between all input neurons and a single hidden neuron i to form the weight vector $\mathbf{w}_i = [w_{i,1} \ w_{i,2} \ \ldots \ w_{i,L}]^T$, $i = -H, (-H+1), \ldots, 0, 1, \ldots, H$.

Let $\mathbf{z} \triangleq \left[\mathbf{w}_{\text{-}H}^T, \mathbf{w}_{\text{-}H+1}^T, \ldots, \mathbf{w}_H^T, b_{\text{-}H}, b_{\text{-}H+1}, \ldots, b_H, v_{\text{-}H}, v_{\text{-}H+1}, \ldots, v_H, b_{out}\right]^T$ be the vectorized form of all network parameters. Let also $g(\mathbf{r};\mathbf{z})$ denote the (pre-detection) output of the NN as a function of the network input $\mathbf{r}$ and parameterized by $\mathbf{z}$. Let $\mathbf{z}_{opt}$ be the value of the vector $\mathbf{z}$ that minimizes a given average (expected) cost, i.e.

$$\mathbf{z}_{opt} = \arg\min_{\mathbf{z}} \ \mathrm{E}\{\mathcal{C}[g(\mathbf{r};\mathbf{z})]\} \tag{3.7}$$

where $\mathcal{C}$ denotes the cost as a function of the NN output and E denotes the expectation operation. Examples of average cost functions include the probability of error evaluated at the output of the sign detector that follows the NN structure or the mean square error (MSE) evaluated at the output of the NN itself.

The optimization problem of (3.7) is defined over a hypersurface parameterized with respect to $2H(L+2)+L+3$ parameters in all. According to Proposition 3.1, the optimum decision boundary is symmetric with respect

to the origin $\mathbf{0}$ and passes through it. If we consider these characteristics in forming the NN optimization problem, then we can reduce significantly the size of the parameter vector.

First, we require the hyperplanes associated with the hidden neurons to be pairwise parallel. Consequently, we require that the neurons of each pair exhibit antipodal shifts from the origin $\mathbf{0}$ (i.e. same absolute value and opposite sign). Let us denote the elements of such an arbitrary pair as "*neuron* i" and "*neuron* $-i$" ($i \neq 0$). Then, the network decision boundary will maintain the properties of the optimum decision boundary identified in Proposition 3.1 if the neurons of each pair satisfy the following constraints: *(a)* $\mathbf{w}_{-i} = \mathbf{w}_i$ (the hyperplanes are parallel); *(b)* $b_{-i} = -b_i$ (the hyperplanes exhibit opposite shifts); *(c)* $b_0 = 0$ (the middle neuron hyperplane passes through the origin); *(d)* $v_{-i} = v_i$, $i \neq 0$ (the output weights of the same-pair neurons are equal to maintain overall decision-boundary symmetry); *(e)* $b_{out} = 0$ (the neural-network decision boundary passes through the origin $\mathbf{0}$). The above constraints for the L—$(2H{+}1)$—1 neural network single-user DS-CDMA receiver are summarized in Table I.

Taking into account the above constraints we define the reduced size parameter vector $\mathbf{z}' \stackrel{\triangle}{=} \left[\mathbf{w}_0^T\ \mathbf{w}_1^T, \ldots, \mathbf{w}_H^T,\ b_1, \ldots, b_H,\ v_0,\ v_1,\ \ldots,\ v_H\right]^T$. Then, we propose the following *constrained NN optimization*:

$$\mathbf{z}'_{opt} = \arg\min_{\mathbf{z}'} \mathrm{E}\{\mathcal{C}[g(\mathbf{r};\mathbf{z}')]\} \tag{3.8}$$

subject to the constraints of Table I.

We note that each hidden neuron $i \in \{\text{-}H, \text{-}H+1, \ldots, H\}$ of the network (or, equivalently, each hidden neuron weight-bias pair $(\mathbf{w}_i,\ b_i)$) determines a hyperplane $\mathcal{P}_i$ in the observation space that is orthogonal to $\mathbf{w}_i$. The shift of $\mathcal{P}_i$ from the origin $\mathbf{0}$ equals the bias b_i. The $2H{+}1$ hyperplanes $\mathcal{P}_{\text{-}H}$, $\mathcal{P}_{\text{-}H+1}$, $\ldots$, $\mathcal{P}_H$ are combined through the output weights $v_{\text{-}H}$, $v_{\text{-}H+1}$, $\ldots$, v_H and bias b_{out} to form the nonlinear network decision boundary. The optimization problem in (3.8) appears in its most general form; it applies to any given neural network $g(\mathbf{r};\mathbf{z}')$ with input $\mathbf{r}$ and reduced set of parameters $\mathbf{z}'$. We emphasize that the constraints in Table I can be used in conjunction with *any* NN training method/criterion that may be used to approximate the optimum decision boundary. Incorporation of such constraints leads to significant performance enhancement when compared to unconstrained training, which is primarily due to the effective reduction of the number of system parameters to be optimized. For an L—$(2H{+}1)$—1 neural network, this number is reduced from $2H(L{+}2){+}L{+}3$ parameters when unconstrained optimization is pursued to $H(L{+}2){+}L{+}1$ when constrained optimization of the type of (3.8) is invoked. For large values of H this reduction amounts to nearly a 50% decrease of the original parameter vector size. Additionally, a welcome side effect of constrained NN optimization is reduced risk for "redundancy", since, for example, neuron i can no longer become equivalent to neuron $-i$ due to the constraint $b_{-i} = -b_i$.

3.4 Constrained Minimum BER NN Training

We now focus on the training procedure for solving the optimization problem in (3.8) when the average cost function is the probability of error at the output of the sign detector that follows the neural network structure. We recall that the NN receiver structure is represented by the function $g(\mathbf{r};\mathbf{z}')$, operates on input $\mathbf{r}$, and is parameterized by vector $\mathbf{z}'$. The final bit decision $\hat{d}_0$ is the sign of the NN output $g(\mathbf{r};\mathbf{z}')$:

$$\hat{d}_0 = \mathrm{sgn}[g(\mathbf{r};\mathbf{z}')]. \tag{3.9}$$

The probability of error at the final output of the receiver is

$$P_e \;=\; P[\hat{d}_0 \neq d_0] \;=\; \pi_0 P[\hat{d}_0 = 1 \,|\, d_0 = -1] \;+\; \pi_1 P[\hat{d}_0 = -1 \,|\, d_0 = 1] \tag{3.10}$$

where $\pi_0 = P[d_0 = -1]$ and $\pi_1 = P[d_0 = 1]$ are the *a priori* probabilities of transmitting -1 or $+1$, respectively. For the purposes of this presentation, we assume that $\pi_0 = \pi_1 = 1/2$. To emphasize the dependency that the induced probability of error has on the parameter vector $\mathbf{z}'$, we will denote the error probability expression as $P_e(\mathbf{z}')$. Our goal is to adapt the NN weight and bias parameter vector $\mathbf{z}'$ such that the induced probability of error becomes minimum. We will refer to the resulting receiver as the constrained minimum bit-error-rate NN (MBER-NN) detector.

Let

$$u(\mathbf{r},\mathbf{z}') = \mathrm{sgn}[g(\mathbf{r};\mathbf{z}')] \tag{3.11}$$

be the output of the receiver. We define the *single-letter distortion measure* $\zeta(\cdot)$ as follows

$$\zeta(\mathbf{r}_0, \mathbf{r}_1; \mathbf{z}') \stackrel{\triangle}{=} \frac{1}{2}\{\pi_0[1 + u(\mathbf{r}_0, \mathbf{z}')] + \pi_1[1 - u(\mathbf{r}_1, \mathbf{z}')]\} \tag{3.12}$$

where $\mathbf{r}_0$ and $\mathbf{r}_1$ are assumed to be received data vectors from hypothesis H_0 (i.e. with $d_0 = -1$) and hypothesis H_1 (i.e. with $d_0 = +1$), respectively. This is the modification for the antipodal ± 1 case of the distortion measure that was first defined in [19]. Intuitively, $\zeta(\cdot)$ measures the distortion at the output of the receiver. The terms $[1 + u(\mathbf{r}_0, \mathbf{z}')]$ and $[1 - u(\mathbf{r}_1, \mathbf{z}')]$ are both zero when the receiver makes the correct decision and strictly positive otherwise. The two terms are weighted by the *a priori* probabilities of the two hypotheses H_0 and H_1, so failure of the receiver to detect the most likely hypothesis will result in a proportionally large penalty. As a result, the expected value of this single-letter distortion measure ζ is as follows [19]:

$$\begin{aligned} &E[\zeta(\mathbf{r}_0, \mathbf{r}_1; \mathbf{z}')] = \\ &E\left[\frac{1}{2}[\pi_0(1 + u(\mathbf{r}_0, \mathbf{z}')) + \pi_1(1 - u(\mathbf{r}_1, \mathbf{z}'))]\right] = \\ &\pi_0 P[u(\mathbf{r}_0, \mathbf{z}') = 1] + \pi_1 P[u(\mathbf{r}_1, \mathbf{z}') = -1] = P_e(\mathbf{z}'). \end{aligned} \tag{3.13}$$

Equation (3.13) allows us to reduce the problem of finding the parameter vector $\mathbf{z}'$ that minimizes the probability of error $P_e(\mathbf{z}')$ to that of finding the parameter vector $\mathbf{z}'$ that minimizes the expected value of $\zeta(\cdot\,, \cdot\,; \mathbf{z}')$.

Let in addition

$$\mathbf{x}_n(\mathbf{z}') \stackrel{\triangle}{=} [x_{n,1}(\mathbf{z}'), \ldots, x_{n,M'}(\mathbf{z}')]^T \tag{3.14}$$

be a vector of the same length as $\mathbf{z}'$, with its jth element $x_{n,j}(\mathbf{z}')$, $j = 1, 2, \ldots, H(L+2)+L+1$, defined by

$$x_{n,j}(\mathbf{z}') \stackrel{\triangle}{=} \frac{1}{2c_n}[\zeta(\mathbf{r}_{0,n}, \mathbf{r}_{1,n}; \mathbf{z}' + c_n\mathbf{e}_j) - \zeta(\mathbf{r}_{0,n}, \mathbf{r}_{1,n}; \mathbf{z}' - c_n\mathbf{e}_j)] \tag{3.15}$$

where $c_n = cn^{-1/4}$ for some $c > 0$ and $\mathbf{e}_j$ denotes the jth coordinate unit vector. The sequences $\{\mathbf{r}_{0,n}\}$ and $\{\mathbf{r}_{1,n}\}$, $n = 1, \ldots, N$, are assumed to be pilot (training) sequences of received vectors with $d_0 = -1$ and $d_0 = +1$, respectively. We note that in this setup, $\{\mathbf{r}_{0,n}\}$ and $\{\mathbf{r}_{1,n}\}$, $n = 1, \ldots, N$, consist of independent identically distributed vectors. The variance of all components of $\mathbf{x}_n(\mathbf{z}')$ is finite for all $\mathbf{z}'$. The regression function $P_e(\mathbf{z}')$ is twice continuously differentiable with bounded second derivative at least as long as $\mathbf{n}$ in (3.1) is white Gaussian [19]. Finally, the above selection of the gain sequence $\{c_n\}$ guarantees that $c_n \to 0$. Then, for a monotonically decreasing sequence of positive numbers $\{\alpha_n\}$ such that $\sum \alpha_n = \infty$, $\sum \alpha_n c_n < \infty$, and $\sum \alpha_n^2 c_n^{-2} < \infty$, the recursion

$$\mathbf{z}'_{n+1} = \mathbf{z}'_n - \alpha_n \mathbf{x}_n(\mathbf{z}'_n) \;\; n = 1, \ldots, N \tag{3.16}$$

converges with probability one (w.p. 1) to the parameter vector $\mathbf{z}'$, say $\boldsymbol{\theta}$, that minimizes $P_e(\mathbf{z}')$ provided that for every $\epsilon > 0$ there exists a positive number $J(\epsilon)$ such that

$$\begin{aligned} &||\mathbf{z}' - \boldsymbol{\theta}|| \geq \epsilon \qquad \text{implies} \\ &\qquad\qquad P_e(\mathbf{z}') \geq P_e(\boldsymbol{\theta}) \text{ and } ||\nabla P_e(\mathbf{z}')|| \geq J(\epsilon). \end{aligned} \tag{3.17}$$

The foundation of the recursive algorithm in (3.16) is provided by [20] which is an extension to multivariate regressions of the well-known Kiefer-Wolfowitz stochastic approximation method [21] for finding the extrema of a regression function. At every stage of recursion (3.16) each component of the gradient of the regression function $P_e(\mathbf{z}')$ is evaluated/estimated simultaneously by the two-sided difference approximation in (3.15). Common selections for the sequence c_n in (3.15) and the learning gain α_n in (3.16) are $c_n = cn^{-1/4}$, $c > 0$, and $\alpha_n = \alpha/n$, $\alpha > 0$, respectively. These selections, however, are arbitrary and any sequences that satisfy $c_n \to 0$, $\alpha_n \to 0$, $\sum \alpha_n = \infty$, $\sum \alpha_n c_n < \infty$, $\sum \alpha_n^2 c_n^{-2} < \infty$ can be used.

3.5 Importance-Sampling-accelerated MBER-NN optimization

The constrained MBER training algorithm described in the previous section (eq. (3.16)) exhibits slow convergence as the NN approaches a low BER operating region. This is caused by the fact that $\mathbf{z}'_n$ is updated only when one of the distortion measures $\zeta(\mathbf{r}_{0,n}, \mathbf{r}_{1,n}; \mathbf{z}' + c_n\mathbf{e}_j)$ or $\zeta(\mathbf{r}_{0,n}, \mathbf{r}_{1,n}; \mathbf{z}' - c_n\mathbf{e}_j)$ is non-zero, i.e. when an "error" occurs. This event may become extremely rare when the BER of the NN is low and the value of c_n is small. To overcome this limitation we modify the MBER recursion in (3.16) to achieve a higher parameter-update rate (and, thus, a higher convergence rate) by creating virtual high BER receiver conditions. This is accomplished by altering the received vector statistics using importance sampling (IS) principles. We note that importance sampling is traditionally used for system performance evaluation; in contrast, what we propose here for the first time is to use IS directly for receiver design optimization.

We recall that the MBER training algorithm described in the previous section operates on data vectors $\mathbf{r}_0$ and $\mathbf{r}_1$ whose conditional pdfs are

$$f_0(\mathbf{r}) = f(\mathbf{r}|d_0 = -1) = \frac{1}{2^{K-1}} \frac{1}{\sqrt{2\pi}\sigma} \sum_{i=1}^{2^{K-1}} \exp\left(-\frac{1}{2\sigma^2} \left\| \mathbf{r} + \sqrt{E_0}\mathbf{s}_0 + \sum_{k=1}^{K-1} \sqrt{E_k} d_k^{(i)} \mathbf{s}_k \right\|^2 \right) \tag{3.18}$$

and

$$f_1(\mathbf{r}) = f(\mathbf{r}|d_0 = +1) = \frac{1}{2^{K-1}} \frac{1}{\sqrt{2\pi}\sigma} \sum_{i=1}^{2^{K-1}} \exp\left(-\frac{1}{2\sigma^2} \left\| \mathbf{r} - \sqrt{E_0}\mathbf{s}_0 + \sum_{k=1}^{K-1} \sqrt{E_k} d_k^{(i)} \mathbf{s}_k \right\|^2 \right) \tag{3.19}$$

respectively.

Following the principles of IS, the proposed IS-based constrained MBER-NN training algorithm operates on vectors $\mathbf{y}_0$ and $\mathbf{y}_1$ drawn from distributions h_0 and h_1, respectively, that are different than f_0 and f_1 in (3.18) and (3.19). In general, the densities $h_0(\cdot)$ and $h_1(\cdot)$ should be chosen so that the BER of the receiver (operating on received vectors drawn from h_0 and h_1) is high enough to induce plenty of "errors" during the training process. Thus, the choice of h_0 and h_1 is application dependent. For the problem of DS-CDMA receiver design, we use parametrically described densities $h_{0,\beta}(\mathbf{y})$ and $h_{1,\beta}(\mathbf{y})$ parameterized by a scalar $\beta \in (0, 1]$ that are given by

$$h_{0,\beta}(\mathbf{y}) = \frac{1}{2^{K-1}} \frac{1}{\sqrt{2\pi}\sigma} \sum_{i=1}^{2^{K-1}} \exp\left(-\frac{1}{2\sigma^2} \left\| \mathbf{y} + \beta\sqrt{E_0}\mathbf{s}_0 + \sum_{k=1}^{K-1} \sqrt{E_k} d_k^{(i)} \mathbf{s}_k \right\|^2 \right) \tag{3.20}$$

and

$$h_{1,\beta}(\mathbf{y}) = \frac{1}{2^{K-1}} \frac{1}{\sqrt{2\pi}\sigma} \sum_{i=1}^{2^{K-1}} \exp\left(-\frac{1}{2\sigma^2} \left\| \mathbf{y} - \beta\sqrt{E_0}\mathbf{s}_0 + \sum_{k=1}^{K-1} \sqrt{E_k} d_k^{(i)} \mathbf{s}_k \right\|^2 \right). \tag{3.21}$$

In other words, $h_{0,\beta}(\mathbf{y})$ and $h_{1,\beta}(\mathbf{y})$ are densities of virtual received vectors where the desired signal is attenuated by a factor β relative to the received vectors distributed according to (3.18) and (3.19), respectively. For $\beta = 0$ we have complete attenuation of the desired signal while for $\beta = 1$ we have no attenuation at all. The advantage of the specific choices for $h_{0,\beta}$ and $h_{1,\beta}$ in (3.20) and (3.21) is that we can easily draw vectors from this distribution by subtracting a part of the desired signal from the actual received vector as follows:

$$\mathbf{y}_0 = \mathbf{r}_0 + (1-\beta)\sqrt{E_0}\mathbf{s}_0 \tag{3.22}$$

$$\mathbf{y}_1 = \mathbf{r}_1 - (1-\beta)\sqrt{E_0}\mathbf{s}_0. \tag{3.23}$$

Using f_0, $h_{0,\beta}$ and f_1, $h_{1,\beta}$ we define the following weighting functions $W_0(\mathbf{y})$ and $W_1(\mathbf{y})$:

$$W_0(\mathbf{y}) \triangleq f_0(\mathbf{y})/h_{0,\beta}(\mathbf{y}) \tag{3.24}$$

$$W_1(\mathbf{y}) \triangleq f_1(\mathbf{y})/h_{1,\beta}(\mathbf{y}). \tag{3.25}$$

Let also $\zeta'(\mathbf{y}_0, \mathbf{y}_1; \mathbf{z}')$ be a new modified distortion measure with equal a priori probabilities defined as

$$\zeta'(\mathbf{y}_0, \mathbf{y}_1; \mathbf{z}') \triangleq \frac{1}{4}\left\{ W_0(\mathbf{y}_0)[1 + \mathrm{sgn}[g(\mathbf{y}_0; \mathbf{z}')] + W_1(\mathbf{y}_1)[1 - \mathrm{sgn}[g(\mathbf{y}_1; \mathbf{z}')] \right\}. \tag{3.26}$$

Then, it is straightforward to show that

$$E\{\zeta'(\mathbf{y}_0, \mathbf{y}_1; \mathbf{z}')\} = E\{\zeta(\mathbf{r}_0, \mathbf{r}_1; \mathbf{z}')\} = P_e(\mathbf{z}'). \tag{3.27}$$

In other words, the expected value of the distortion measure ζ' is equal to the true BER of the NN parameterized by $\mathbf{z}'$ when the latter operates on received vectors $\mathbf{r}$ distributed according to $f(\cdot)$ given by (3.2).

Proceeding similarly to Section IV, let

$$\mathbf{x}'_n(\mathbf{z}') \triangleq [x'_{n,1}(\mathbf{z}'), \ldots, x'_{n,M'}(\mathbf{z}')]^T \tag{3.28}$$

be a vector with its jth element $x'_{n,j}(\mathbf{z}')$, $j = 1, 2, \ldots, H(L+2)+L+1$, defined by

$$x'_{n,j}(\mathbf{z}') \triangleq \frac{1}{2c_n}[\zeta'(\mathbf{y}_{0,n}, \mathbf{y}_{1,n}; \mathbf{z}' + c_n\mathbf{e}_j) - \zeta'(\mathbf{y}_{0,n}, \mathbf{y}_{1,n}; \mathbf{z}' - c_n\mathbf{e}_j)]. \tag{3.29}$$

In (3.29), $\{\mathbf{y}_{0,n}\}$ and $\{\mathbf{y}_{1,n}\}$, $n = 1, \ldots, N$, are sequences obtained from the pilot (training) sequences $\{\mathbf{r}_{0,n}\}$, $\{\mathbf{r}_{1,n}\}$ by (3.22) and (3.23), respectively. Then, from (3.27) we see that for sequences of positive numbers a_n and c_n such that $c_n \to 0$, $a_n \to 0$, $\sum a_n = \infty$, $\sum a_n c_n < \infty$, $\sum a_n^2 c_n^{-2} < \infty$ the recursion

$$\mathbf{z}'_{n+1} = \mathbf{z}'_n - a_n\mathbf{x}'_n(\mathbf{z}'_n), \ n = 1, \ldots, N \tag{3.30}$$

converges with probability one (w.p. 1) to the value of $\mathbf{z}'$ that minimizes $P_e(\mathbf{z}')$. In comparison with (3.16), the convergence rate of the recursion in (3.30) is much improved since in this latter case the network during training operates always in a virtual high BER region due to the attenuated desired signal in the training vectors $\mathbf{y}_{0,n}$ and $\mathbf{y}_{1,n}$.

We must note at this point that theoretically the weighting functions $W_0(\mathbf{y})$ and $W_1(\mathbf{y})$ in (3.24), (3.25) require exact knowledge of the pdfs $h_{i,\beta}(\cdot)$ and $f_i(\cdot)$, $i = 0,1$. For a practical solution, we propose to approximate the above pdfs by colored Gaussian densities as follows. Let

$$\mathbf{r}_{I+N} \triangleq \sum_{k=1}^{K-1} \sqrt{E_k} d_k \mathbf{s}_k + \mathbf{n} \tag{3.31}$$

be the interference-plus-noise component of $\mathbf{r}$. The vector $\mathbf{r}_{I+N}$ is distributed according to a multivariate Gaussian mixture of 2^{K-1} components. We approximate this mixture by a multivariate colored-Gaussian distribution with mean $\mathbf{0}_{L\times 1}$ and covariance matrix $\mathbf{R}_{I+N} = E\{\mathbf{r}_{I+N}\mathbf{r}_{I+N}^T\} = \sum_{k=1}^{K-1} E_k\mathbf{s}_k\mathbf{s}_k^T + \sigma^2\mathbf{I}$. The approximated pdfs, that we distinguish from the exact $h_{i,\beta}(\cdot)$ and $f_i(\cdot)$, $i = 0,1$ by using the symbol $\sim$, are given below:

$$\widetilde{h}_{0,\beta}(\mathbf{y}) = \frac{1}{2(2\pi)^{L/2}\left|\mathbf{R}_{I+N}\right|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y} + \beta\sqrt{E_0}\mathbf{s}_0)^T\mathbf{R}_{I+N}^{-1}(\mathbf{y} + \beta\sqrt{E_0}\mathbf{s}_0)\right) \tag{3.32}$$

$$\widetilde{h}_{1,\beta}(\mathbf{y}) = \frac{1}{2(2\pi)^{L/2}\left|\mathbf{R}_{I+N}\right|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y} - \beta\sqrt{E_0}\mathbf{s}_0)^T\mathbf{R}_{I+N}^{-1}(\mathbf{y} - \beta\sqrt{E_0}\mathbf{s}_0)\right) \tag{3.33}$$

$$\widetilde{f}_0(\mathbf{r}) = \frac{1}{2(2\pi)^{L/2}\left|\mathbf{R}_{I+N}\right|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{r} + \sqrt{E_0}\mathbf{s}_0)^T\mathbf{R}_{I+N}^{-1}(\mathbf{r} + \sqrt{E_0}\mathbf{s}_0)\right) \tag{3.34}$$

$$\widetilde{f}_1(\mathbf{r}) = \frac{1}{2(2\pi)^{L/2}\left|\mathbf{R}_{I+N}\right|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{r} - \sqrt{E_0}\mathbf{s}_0)^T\mathbf{R}_{I+N}^{-1}(\mathbf{r} - \sqrt{E_0}\mathbf{s}_0)\right). \tag{3.35}$$

We observe that $\widetilde{h}_{i,\beta}(\mathbf{y})$, $i = 0,1$ is an explicit function of the second order statistics of $\mathbf{r}_{I+N}$ which are easy to estimate. The estimation proceeds as follows. First, we observe that $E\{d_0(n)\mathbf{r}_n\} = \sqrt{E_0}\mathbf{s}_0$. Thus, we may estimate the *product* $\sqrt{E_0}\mathbf{s}_0$ by

$$\widehat{\sqrt{E_0}\mathbf{s}_0} = \frac{1}{T}\sum_{n=1}^{T} d_0(n)\mathbf{r}_n \tag{3.36}$$

where T is the number of available samples. We can estimate the interference-plus-noise component of the vectors $\mathbf{r}_n$ by

$$\widehat{\mathbf{r}}_{I+N,n} = \mathbf{r}_n - d_0(n)\widehat{\sqrt{E_0}\mathbf{s}_0}. \tag{3.37}$$

Then, the interference-plus-noise covariance matrix estimate can be evaluated as $\widehat{\mathbf{R}}_{I+N} = \frac{1}{T}\sum_{n=1}^{T}\widehat{\mathbf{r}}_{I+N,n}\widehat{\mathbf{r}}_{I+N,n}^T$. Substituting the above estimates in (3.33) we obtain the estimates

$$\widehat{\widehat{h}}_{0,\beta}(\mathbf{y}) = \frac{1}{2(2\pi)^{L/2}\left|\widehat{\mathbf{R}}_{I+N}\right|^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{y}+\beta\widehat{\sqrt{E_0}\mathbf{s}_0})^T\widehat{\mathbf{R}}_{I+N}^{-1}(\mathbf{y}+\beta\widehat{\sqrt{E_0}\mathbf{s}_0})\right) \tag{3.38}$$

$$\widehat{\widehat{h}}_{1,\beta}(\mathbf{y}) = \frac{1}{2(2\pi)^{L/2}\left|\widehat{\mathbf{R}}_{I+N}\right|^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{y}-\beta\widehat{\sqrt{E_0}\mathbf{s}_0})^T\widehat{\mathbf{R}}_{I+N}^{-1}(\mathbf{y}-\beta\widehat{\sqrt{E_0}\mathbf{s}_0})\right) \tag{3.39}$$

$$\widehat{\widetilde{f}_0}(\mathbf{r}) = \frac{1}{2(2\pi)^{L/2}\left|\widehat{\mathbf{R}}_{I+N}\right|^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{r}+\widehat{\sqrt{E_0}\mathbf{s}_0})^T\widehat{\mathbf{R}}_{I+N}^{-1}(\mathbf{r}+\widehat{\sqrt{E_0}\mathbf{s}_0})\right) \tag{3.40}$$

$$\widehat{\widetilde{f}_1}(\mathbf{r}) = \frac{1}{2(2\pi)^{L/2}\left|\widehat{\mathbf{R}}_{I+N}\right|^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{r}-\widehat{\sqrt{E_0}\mathbf{s}_0})^T\widehat{\mathbf{R}}_{I+N}^{-1}(\mathbf{r}-\widehat{\sqrt{E_0}\mathbf{s}_0})\right). \tag{3.41}$$

Finally, an estimate of the two weighting functions is given by

$$\widehat{W}_0(\widehat{\mathbf{y}}_n) = \widehat{\widetilde{f}}_0(\widehat{\mathbf{y}}_n)/\widehat{\widehat{h}}_{0,\beta}(\widehat{\mathbf{y}}_n) \tag{3.42}$$

$$\widehat{W}_1(\widehat{\mathbf{y}}_n) = \widehat{\widetilde{f}}_1(\widehat{\mathbf{y}}_n)/\widehat{\widehat{h}}_{1,\beta}(\widehat{\mathbf{y}}_n) \tag{3.43}$$

where

$$\widehat{\mathbf{y}}_n \triangleq \mathbf{r}_n - d_0(n)(1-\beta)\widehat{\sqrt{E_0}\mathbf{s}_0}. \tag{3.44}$$

3.6 Simulation Studies

We consider a synchronous DS-CDMA system populated by $K = 6$ users, each with a length $L = 16$ signature. The signature cross-correlations between users are in the range $[0.125, 0.375]$. In each and every study below, the BER results that we present are averaged over 100 independent neural-network receiver training realizations (each realization corresponds to a single epoch consisting of 4000 data).

We examine two feedforward MLP neural network structures for $H = 1$ and $H = 3$, i.e. one of size 16—3—1 (Fig. 3.2) and the other of size 16—7—1 (Figs. 3.3-3.5). The MLP NN of the desired user is trained by either the plain MBER algorithm (defined similarly to (3.16) with $\mathbf{z}$ in place of $\mathbf{z}'$), or the constrained-MBER algorithm in (3.16), or the IS-constrained-MBER algorithm in (3.30). For each case, the BER performance is shown as a function of the number of new training data arriving at the receiver. As a comparative study we also consider the standard backpropagation (BP) training algorithm[2]. We recall that the BP algorithm is a recursive procedure that aims at the minimization of $\mathrm{E}\{[d_0 - g(\mathbf{r};\mathbf{z})]^2\}$, i.e. minimizes the mean-square-error (MSE) between the output of the MLP $g(\mathbf{r};\mathbf{z})$ and the desired signal d_0. According to stochastic approximation principles, the BP algorithm originates from the recursion

$$\mathbf{z}_{n+1} = \mathbf{z}_n - \mu_n\nabla_{\mathbf{z}_n}\left[d_0(n) - g(\mathbf{r}_n;\mathbf{z}_n)\right]^2 \tag{3.45}$$

for any monotonically decreasing sequence of positive numbers $\{\mu_n\}_n$ such that $\sum_{n=0}^{\infty}\mu_n = \infty$ and $\sum_{n=0}^{\infty}\mu_n^2 < \infty$ ($\{\mathbf{r}_n\}_n$ is a pilot sequence of input data vectors).

[2]The results in this work apply to systems/neural network structures with any value of K, L and H. We note, however, that higher values of L and H may require prohibitive amount of time for evaluation of BER performance of *any* neural network algorithm through Monte-Carlo simulations.

Table 3.1. Neural-Network Symmetry Constraints

Hidden Layer
$\mathbf{w}_{-i} = \mathbf{w}_i \ (i = 1, 2, \ldots, H)$
$b_{-i} = -b_i (i = 1, 2, \ldots, H)$
$b_0 = 0$
Output Layer
$v_{-i} = v_i \ \ (i = 1, 2, \ldots, H)$
$b_{out} = 0$

All algorithms are initialized with the same parameter values drawn randomly and uniformly in the interval $[-0.25, 0.25]$. For Figs. 3.2 and 3.3, the energy of the user of interest is set at 14dB while the energies of the interferers range from 8-11dB. We see that the IS-constrained-MBER algorithm converges noticeably faster in both studies of Fig. 3.2 and Fig. 3.3. In both figures, all various forms of the MBER training algorithm outperform the standard BP training algorithm. In Fig. 3.4, we repeat the same study as in Fig. 3.3 where now the user signal of interest is at 10dB and all five interferers are at 12dB. In this figure we also include the performance curve of the MLP-NN single user detection algorithm proposed in [2]. We note that the algorithm in [2] requires *noiseless* training data as well as knowledge of the interfering information bits. The performance of the network when trained by the algorithm in [2] is sensitive to the number of hidden neurons of the structure (it has been reported that a large number of hidden neurons - exponential in the number of users - is necessary to achieve acceptable performance levels [4]). Finally, in Fig. 3.5 we compare the performance of the training algorithms of Fig. 3.4 when used on a larger MLP-NN (16—7—1) with smaller user population ($K = 3$). Again, we see that the IS-constrained-MBER algorithm converges noticeably faster.

3.7 Conclusions

Motivated by the ultimate performance metric of a digital communications link, namely BER, a fast converging adaptive minimum BER MLP-NN-based receiver architecture for DS-CDMA communications was developed. The adaptive receiver incorporates the BER optimization criterion directly into the NN learning process. Furthermore, properties of the optimum single-user decision boundary for AWGN multiple-access channels were translated into a set of constraints that were incorporated in the minimum BER training procedure, thus, significantly reducing the number of network parameters (weights and biases) to be optimized. Coupled with importance sampling (IS) principles embedded directly into the receiver optimization process, the end result is faster network convergence and lower vulnerability to architectural NN redundancy. The proposed unconventional use of IS does not assume knowledge of the input statistics and requires only a short pilot sequence to create virtual high BER receiver conditions that accelerate the NN parameter optimization.

Simulation results supported these theoretical developments and illustrated the merits of the minimum BER training approach as well as the significant convergence rate gains obtained as a result of the proposed IS-constrained method.

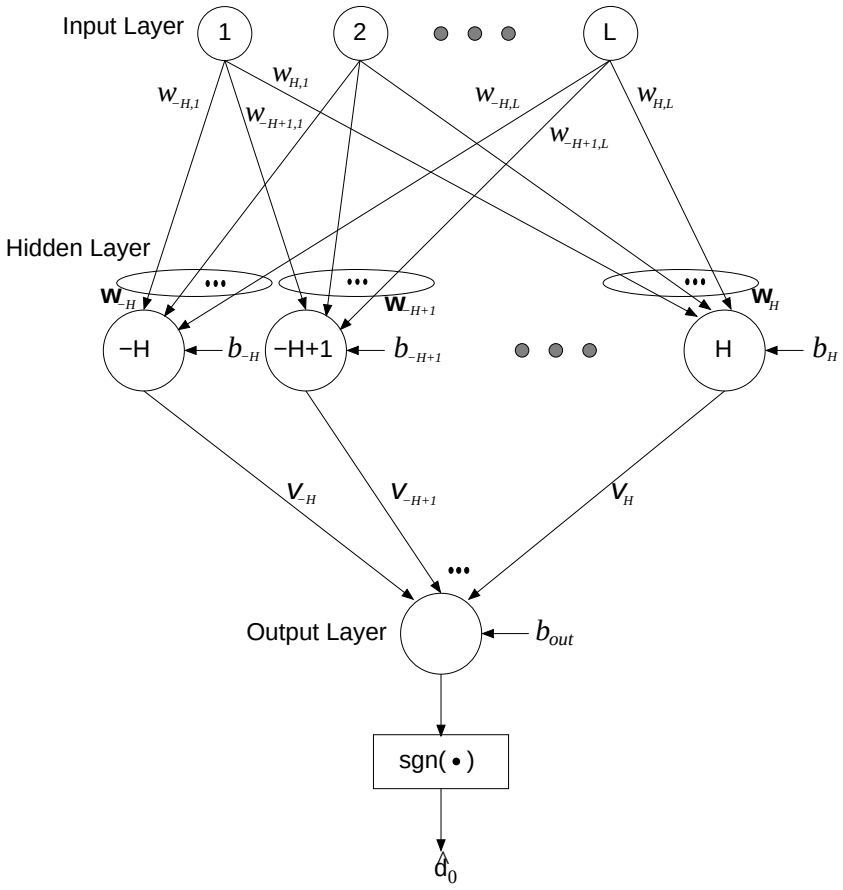

Fig. 3.1. Neural network receiver architecture under consideration.

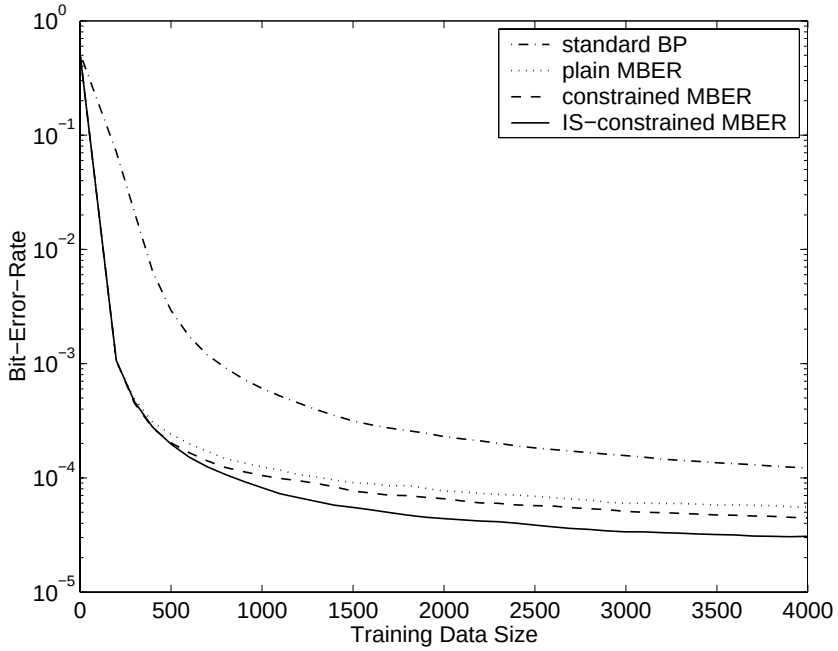

Fig. 3.2. BER performance versus number of training data for a 16-3-1 MLP NN (L=16, H=1) trained by either the standard BP algorithm in (3.45), or the plain MBER algorithm (defined similarly to (3.16) with $\mathbf{z}$ in place of $\mathbf{z}'$), or the constrained-MBER algorithm in (3.16), or the IS-constrained-MBER algorithm in (3.30).

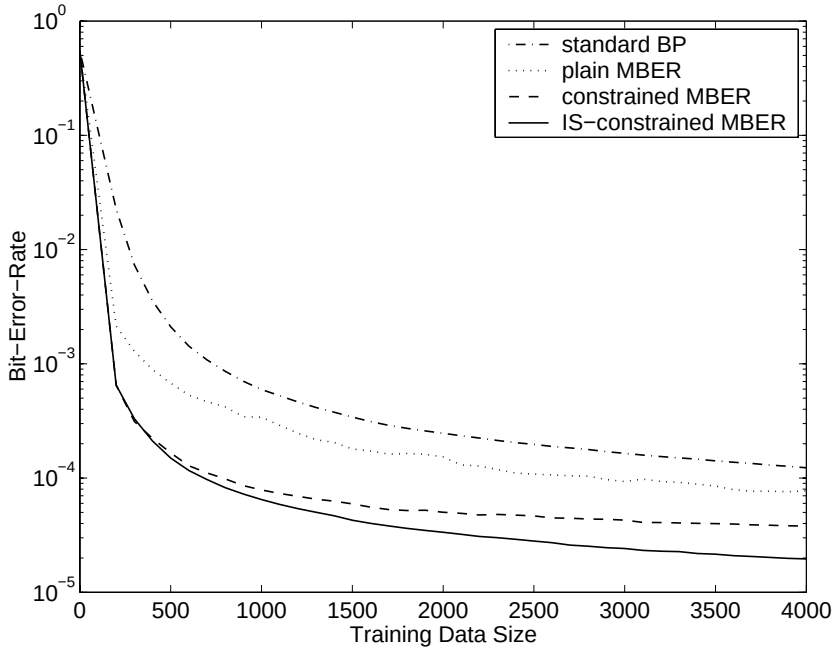

Fig. 3.3. BER performance versus number of training data for a 16-7-1 MLP NN (L=16, H=3) trained by either the standard BP algorithm in (3.45), or the plain MBER algorithm (defined similarly to (3.16) with $\mathbf{z}$ in place of $\mathbf{z}'$), or the constrained-MBER algorithm in (3.16), or the IS-constrained-MBER algorithm in (3.30).

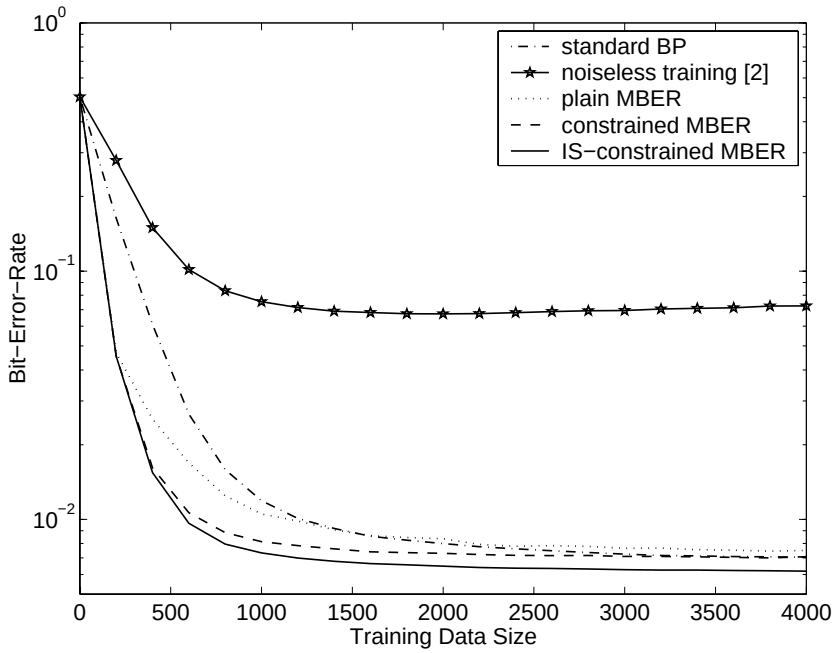

Fig. 3.4. Same study as in Fig. 3.3 with the user signal of interest at 10dB and the five interferers at 12dB. The 16-7-1 MLP-NN (L=16, H=3) is trained by either the standard BP algorithm in (3.45), or the *noiseless* training scheme in [2], or the plain MBER algorithm (defined similarly to (3.16) with $\mathbf{z}$ in place of $\mathbf{z}'$), or the constrained-MBER algorithm in (3.16), or the IS-constrained-MBER algorithm in (3.30).

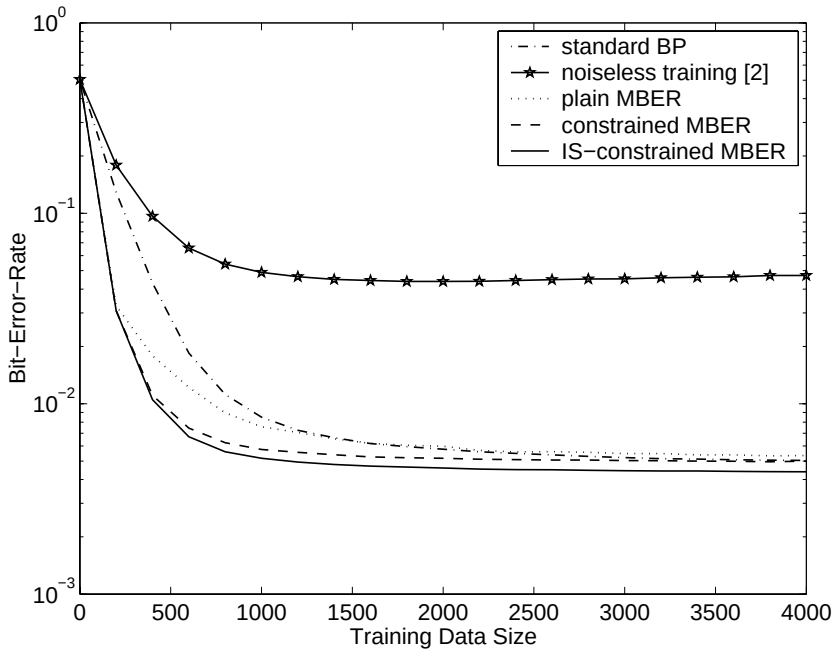

Fig. 3.5. BER performance versus number of training data for a 16-7-1 MLP NN (L=16, H=3) with $K = 3$ (with the user signal of interest at 10dB and the two interferers at 12dB). The neural network is trained by either the standard BP algorithm in (3.45), or the *noiseless* training scheme in [2], or the plain MBER algorithm (defined similarly to (3.16) with $\mathbf{z}$ in place of $\mathbf{z}'$), or the constrained-MBER algorithm in (3.16), or the IS-constrained-MBER algorithm in (3.30).

Bibliography

[1] H. V. Poor and S. Verdu, "Single-user detectors for multiuser channels," *IEEE Trans. Commun.*, vol. 36, no. 1, pp. 50-60, Jan. 1988.

[2] B. Aazhang, B. P. Paris, and G. C. Orsak, "Neural networks for multiuser detection in code-division multiple-access communications," *IEEE Trans. Commun.*, vol. 40, pp. 1212-1222, July 1992.

[3] R. K. Madyastha and B. Aazhang, "An algorithm for training multilayer perceptrons for data classification and function interpolation," *IEEE Trans. Circ. Syst. I: Fundamental Theories and Applications*, vol. 41, pp. 866-875, Dec. 1994.

[4] U. Mitra and H. V. Poor, "Adaptive receiver algorithms for near-far resistant CDMA," *IEEE Trans. Commun.*, vol. 43, pp. 1713-1724, Feb./Mar./Apr. 1995.

[5] U. Mitra and H. V. Poor, "Neural network techniques for adaptive multiuser demodulation," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 1460-1470, Dec. 1994.

[6] B. Mulgrew, E. S. Warner, and P. M. Grant, "A Bayesian receiver for asynchronous code division multiple access communications," in *Proc. PIMRC*, Toronto, Canada, Sept. 1995, pp. 985-989.

[7] K. Das and S. D. Morgera, "Adaptive interference cancellation for DS-CDMA systems using neural network techniques," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1774-1784, Dec. 1998.

[8] T. C. Chuah, B. S. Sharif, and O. R. Hinton, "Robust adaptive spread-spectrum receiver with neural-net preprocessing in non-Gaussian noise," *IEEE Trans. Neural Networks*, vol. 12, pp. 546-558, May 2001.

[9] G. I. Kechriotis and E. S. Manolakos, "Hopfield neural network implementation of the optimal CDMA multiuser detector," *IEEE Trans. Neural Networks*, vol. 7, pp. 131-141, Jan. 1996.

[10] G. I. Kechriotis and E. S. Manolakos, "A hybrid digital signal processing-neural network CDMA multiuser detection scheme," *IEEE Trans. Circ. Syst. II: Analog and Digital Signal Processing*, vol. 43, pp. 96-104, Feb. 1996.

[11] D. C. Chen and B. J. Sheu, "A compact neural-network-based CDMA receiver," *IEEE Trans. Circ. Syst. II: Analog and Digital Signal Processing*, vol. 45, pp. 384-387, Mar. 1998.

[12] I. N. Psaromiligkos, S. N. Batalama, and D. A. Pados, "On adaptive minimum probability of error linear-filter receivers for DS-CDMA channels," *IEEE Trans. Commun.*, vol. 47, pp. 1092-1102, July 1999.

[13] C. C. Yeh, R. R. Lopes, and J. R. Barry, "Approximate minimum bit-error rate multiuser detection," in *Proc. IEEE Globecom*, Sydney, Australia, Nov. 1998, pp. 3590-3595.

[14] S. Chen, A. K. Samingan, B. Mulgrew, and L. Hanzo, "Adaptive minimum-BER linear multiuser detection for DS-CDMA signals in multipath channels," *IEEE Trans. on Signal Processing*, vol. 49, pp. 1240-1247, June 2001.

[15] S. Chen, B. Mulgrew, and L. Hanzo, "Adaptive least error rate algorithm for neural network classifiers," in *Proc. 2001 IEEE Signal Processing Society Workshop: Neural Networks for Signal Processing XI*, N. Falmouth, MA, Sept. 2001, pp. 223-232.

[16] X. F. Wang, W. S. Lu, and A. Antoniou, "Constrained minimum-BER multiuser detection," *IEEE Trans. on Signal Processing*, vol. 48, pp. 2903-2909, Oct. 2000.

[17] S. Verdu, *Multiuser Detection*, Cambridge, UK: Cambridge University Press, 1998.

[18] J. D. Matyjas, G. N. Karystinos, and S. N. Batalama, "On the training of DS-CDMA neural-network receivers," in *Proc. 2002 Int. Conf. on Acoustics Speech and Signal Processing (ICASSP)*, Orlando, FL, 2002, pp. 1017-1020.

[19] D. A. Pados and P. Papantoni-Kazakos, "New nonleast-squares neural network learning algorithms for hypothesis testing," *IEEE Trans. Neural Networks*, vol. 6, pp. 596-609, May 1995.

[20] J. Blum, "Multidimensional stochastic approximation methods," *Ann. Math. Stat.*, vol. 25, pp. 737-744, 1954.

[21] J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Stat.*, vol. 23, pp. 462-466, 1952.

Chapter 4

On the Relative Output SINR of Full and Partial Decorrelators

4.1 Introduction

Linear DS/CDMA multiuser detectors have emerged as low complexity alternatives to the optimum multiuser detector [1]. Within the class of linear detectors, minimum-mean-square-error (MMSE) and decorrelating-type receivers strike an appealing balance between complexity and performance.[1] In this work, we focus solely on linear decorrelating-type receivers [3]-[8].

In the presence of additive white Gaussian noise (AWGN), full decorrelation of the multiple access interference (MAI) results in enhanced noise variance at the decorrelator output. Partial decorrelation occasionally strikes a better balance between interference suppression and noise enhancement as it was observed in [9]. A partial decorrelator (PDEC) chooses purposefully to decorrelate only a part of the MAI as opposed to the full decorrelator (FDEC) that decorrelates completely the MAI.[2]

In this work, motivated primarily by scientific curiosity rather than searching for practical solutions, we identify necessary and sufficient conditions on the system parameters under which a PDEC outperforms the FDEC in the output SINR sense. Our developments provide further insight into the operational characteristics and limitations of decorrelating receivers and show why decorrelation of MAI that meets certain conditions should be purposefully avoided, if at all possible. As a side study, we consider a blind implementation of the FDEC that is based on eigendecomposition of the interference-plus-noise autocovariance matrix and can be readily modified to provide a PDEC as explained in the sequel.

4.2 Signal Model

We consider a K-user synchronous[3] DS/CDMA communication system that utilizes short spreading codes of length L. In this system, user-k, $k = 0, \ldots, K-1$, is assigned a normalized signature vector $\mathbf{s}_k$, where $\mathbf{s}_k[l] = \{\pm 1/\sqrt{L}\}$, $l = 1, 2, \cdots, L$. The signature matrix is denoted by $\mathbf{S}_{L\times K} \triangleq [\mathbf{s}_0 \ \mathbf{s}_1 \cdots \mathbf{s}_{K-1}]$ and is assumed to be full rank (rank K). In the rest of the work, we focus on single-user detection and without loss of generality we assume that the user of interest is *user 0*. After chip-matched filtering and chip-rate

[1] In extreme, non-practical interference models the matched-filter (MF) scheme may outperform the decorrelating or MMSE detector in terms of bit-error-rate (BER) [2].

[2] This is in contrast to unintentional operation under partial decorrelating conditions when unknown interference is present in the system (e.g. inter-cell interference [10]).

[3] Generalization to asynchronous transmissions over multipath channels is straightforward if we formulate the asynchronous channel as a synchronous one with a higher virtual active user population and perform "one-shot" decorrelation [1].

sampling, the baseband received signal vector over one bit period takes the form

$$\mathbf{r} = \sum_{k=0}^{K-1} \sqrt{E_k} b_k \mathbf{s}_k + \mathbf{n} = \sqrt{E_0} b_0 \mathbf{s}_0 + \mathbf{i} + \mathbf{n}, \tag{4.1}$$

where E_k and b_k denote the bit energy and the information bit of user-k, respectively, (the binary information bits are assumed to take the values ± 1 independently and equiprobably), $\mathbf{i} \triangleq \sum_{k=1}^{K-1} \sqrt{E_k} b_k \mathbf{s}_k$ identifies comprehensively the MAI, and $\mathbf{n}$ represents the AWGN vector with mean $\mathbf{0}$ and autocovariance matrix $\sigma^2 \mathbf{I}_L$, where $\mathbf{I}_L$ is the $L \times L$ identity matrix.

4.3 Full Decorrelator (FDEC)

Under the assumption of perfectly known signature matrix $\mathbf{S}$, the FDEC for *user 0* is[4]

$$\mathbf{d_s} \triangleq \mathbf{S}\mathbf{X}^{-1}\mathbf{e}_1 \tag{4.2}$$

where $\mathbf{X} \triangleq \mathbf{S}^T\mathbf{S}$ is the signature cross-correlation matrix and $\mathbf{e}_1 \triangleq [1\ 0\ \cdots\ 0]^T$. We note that with simple algebraic manipulations we can show that $\mathbf{d_s}$ in (4.2) is a scalar multiple of the filter $(\mathbf{I}_L - \mathbf{S}_{\backslash 0}\mathbf{S}_{\backslash 0}^{+})\mathbf{s}_0$ where $\mathbf{S}_{\backslash 0} \triangleq [\mathbf{s}_1, \ldots, \mathbf{s}_{K-1}]$ and the superscript $+$ denotes the Moore-Penrose generalized inverse [1].

The FDEC can be also formed by means of ED of the desired-signal-present received-vector autocovariance matrix $\mathbf{R} \triangleq E\{\mathbf{r}\mathbf{r}^T\} = \sum_{k=0}^{K-1} E_k \mathbf{s}_k \mathbf{s}_k^T + \sigma^2 \mathbf{I}_L$, as follows. Let the ED of $\mathbf{R}$ be $\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$, where $\mathbf{\Lambda} \triangleq diag(\lambda_0, \lambda_1, \cdots, \lambda_{L-1})$, $\lambda_0 \geq \cdots \geq \lambda_{K-1} > \lambda_K = \cdots = \lambda_{L-1} = \sigma^2$, are the eigenvalues of $\mathbf{R}$ in descending order, and $\mathbf{Q} \triangleq [\mathbf{q}_0, \ldots, \mathbf{q}_{L-1}]$ denotes the matrix of the corresponding eigenvectors. Let $\mathbf{Q}_S \triangleq [\mathbf{q}_0\ \ \mathbf{q}_1 \cdots \mathbf{q}_{K-1}]$, where $\mathbf{q}_0, \mathbf{q}_1, \cdots, \mathbf{q}_{K-1}$ span the *signal subspace*, and $\mathbf{\Lambda}_S \triangleq diag(\lambda_0, \lambda_1, \cdots, \lambda_{K-1})$. Then, the FDEC for *user 0* takes the form:

$$\mathbf{d}_{ED1} = \mathbf{Q}_S(\mathbf{\Lambda}_S - \sigma^2\mathbf{I}_K)^{-1}\mathbf{Q}_S^T\mathbf{s}_0 = \sum_{k=0}^{K-1} \frac{\mathbf{q}_k\mathbf{q}_k^T}{\lambda_k - \sigma^2}\mathbf{s}_0. \tag{4.3}$$

The FDEC $\mathbf{d}_{ED1}$ in (4.3) was proposed in [7].

In the following, we derive an alternative subspace-type FDEC, $\mathbf{d}_{ED2}$, that instead of $\mathbf{R}$, utilizes the desired-signal-absent (interference-plus-noise only) received-vector autocovariance matrix, $\mathbf{R}_{I+N} \triangleq E\{(\mathbf{i}+\mathbf{n})(\mathbf{i}+\mathbf{n})^T\} = \mathbf{R} - E_0\mathbf{s}_0\mathbf{s}_0^T$. Let the ED of $\mathbf{R}_{I+N}$ be $\mathbf{R}_{I+N} = \mathbf{U}\mathbf{D}\mathbf{U}^T = [\mathbf{U}_I\ \ \mathbf{U}_{I^\perp}]\mathbf{D}[\mathbf{U}_I\ \ \mathbf{U}_{I^\perp}]^T$, where $\mathbf{D} \triangleq diag(d_1, d_2, \cdots, d_L)$ denotes the diagonal eigenvalue matrix, and $\mathbf{U}_I \triangleq [\mathbf{u}_1\ \ \mathbf{u}_2 \cdots \mathbf{u}_{K-1}]$ (i.e. the columns of $\mathbf{U}_I$ span the MAI-subspace). The filter $\mathbf{d}_{ED2}$ is given by

$$\mathbf{d}_{ED2} = \left(\mathbf{I} - \mathbf{U}_I\mathbf{U}_I^T\right)\mathbf{s}_0, \tag{4.4}$$

i.e. $\mathbf{d}_{ED2}$ is the component of $\mathbf{s}_0$ that is orthogonal to the MAI subspace. We note that the decorrelating filter $\mathbf{d}_{ED2}$ is a CDMA version of the "eigencanceler" [11]-[13]. Under perfect knowledge of $\mathbf{S}$, $\mathbf{R}$, and $\mathbf{R}_{I+N}$, FDECs $\mathbf{d_s}$, $\mathbf{d}_{ED1}$, and $\mathbf{d}_{ED2}$ are all equivalent[5] (scalar multiples of each other) which implies that the corresponding output SINRs $\gamma_\mathbf{s}$, γ_{ED1}, and γ_{ED2} are equal and given by

$$\gamma_\mathbf{s} = \gamma_{ED1} = \gamma_{ED2} = \frac{(\mathbf{d_s}^T\mathbf{s}_0)^2 E_0}{\|\mathbf{d_s}\|^2\sigma^2} = \frac{E_0}{\sigma^2\{\mathbf{X}^{-1}\}_{1,1}} \tag{4.5}$$

where $\{\mathbf{X}^{-1}\}_{i,j}$ identifies the (i,j)th element of matrix $\mathbf{X}^{-1}$.

[4]The bold subscript "s" of vector $\mathbf{d}$ is used to distinguish the signature-matrix-based ($\mathbf{S}$-type) decorrelator $\mathbf{d_s}$ from the eigendecomposition (ED)-based structures (identified by the subscript "ED") that will be presented later in this section. To avoid any confusion, we clarify that the non-bold subscript "S" of a matrix will indicate operation in the complete signal subspace while "$S^\perp$" will indicate operation in the corresponding orthogonal subspace.

[5]Any filter $\mathbf{w}$ that satisfies the constraints *(i)* $\mathbf{w}^T\mathbf{s}_0 > 0$, *(ii)* $\mathbf{w}^T\mathbf{s}_k = 0$, $k = 1, \cdots, K-1$, and *(iii)* $\mathbf{w}$ lies in the range of $\mathbf{S}$, yields the same bit decisions as $\mathbf{d_s}$.

However, when $\mathbf{R}$ and $\mathbf{R}_{I+N}$ are unknown and are sample-average estimated the corresponding data-adaptive FDECs differ and so do their SINR and BER performance. The sample-average estimate of $\mathbf{R}$ is $\widehat{\mathbf{R}} = \frac{1}{M}\sum_{m=1}^{M}\mathbf{r}_m\mathbf{r}_m^T$ where M is the size of the available data record. $\mathbf{R}_{I+N}$ can be estimated similarly during a silent period of the user of interest.[6] By performing eigenvalue decomposition on $\widehat{\mathbf{R}}$ and $\widehat{\mathbf{R}}_{I+N}$ we obtain the estimates of $\mathbf{d}_{ED1}$ and $\mathbf{d}_{ED2}$:

$$\widehat{\mathbf{d}}_{ED1} = \widehat{\mathbf{Q}}_S(\widehat{\mathbf{\Lambda}}_S - \widehat{\sigma}^2\mathbf{I}_K)^{-1}\widehat{\mathbf{Q}}_S^T\mathbf{s}_0 \tag{4.6}$$

$$\widehat{\mathbf{d}}_{ED2} = \mathbf{s}_0 - \widehat{\mathbf{U}}_I\widehat{\mathbf{U}}_I^T\mathbf{s}_0. \tag{4.7}$$

Comparing expression (4.6) and (4.7) we observe that the estimate $\widehat{\mathbf{d}}_{ED2}$ is a function of the eigenvector estimates $\widehat{\mathbf{U}}_I$ only, while $\widehat{\mathbf{d}}_{ED1}$ is a function of both the eigenvector estimates $\widehat{\mathbf{Q}}_S$ *and* the eigenvalue estimates $\widehat{\mathbf{\Lambda}}_S$. Most importantly, evaluation of (4.6) requires inversion of the modified eigenvalue matrix $(\widehat{\mathbf{\Lambda}}_S - \widehat{\sigma}^2\mathbf{I}_K)$, an operation that yields increased estimation variance particularly when estimation is based on a short data record. Therefore, it is expected that $\widehat{\mathbf{d}}_{ED2}$ outperforms $\widehat{\mathbf{d}}_{ED1}$ when the number of available data is limited (and the same is true with the corresponding PDECs developed in the next section).

4.4 Partial Decorrelator (PDEC)

While the FDEC decorrelates the complete MAI, a PDEC aims at decorrelating only a part of it by excluding one or more user-signatures or eigenvectors from the corresponding implementation method. In this section, we investigate the conditions under which the PDEC outperforms the FDEC in the output SINR sense.

Let $\mathbf{d}_{\mathbf{s}\setminus j}$, $\mathbf{d}_{ED1\setminus j}$, and $\mathbf{d}_{ED2\setminus j}$ denote the PDEC for *user 0* that decorrelates all interfering users but user j, $j = 1,\ldots,K-1$, eigenvector-direction j of $\mathbf{R}$, $j = 0,\ldots,K-1$, or eigenvector-direction j of $\mathbf{R}_{I+N}$, $j = 1,\ldots,K-1$, respectively[7].Similarly, we define

$$\begin{aligned}
&\mathbf{S}_{\setminus j} \triangleq [\mathbf{s}_0\cdots\mathbf{s}_{j-1}\ \ \mathbf{s}_{j+1}\cdots\mathbf{s}_{K-1}],\\
&\mathbf{X}_{\setminus j} \triangleq \mathbf{S}_{\setminus j}^T\mathbf{S}_{\setminus j}, \qquad \boldsymbol{\psi}_j \triangleq \mathbf{S}_{\setminus j}^T\mathbf{s}_j,\\
&\mathbf{Q}_{S\setminus j} \triangleq [\mathbf{q}_0\cdots\mathbf{q}_{j-1}\ \ \mathbf{q}_{j+1}\cdots\mathbf{q}_{K-1}],\\
&\mathbf{\Lambda}_{S\setminus j} \triangleq \mathrm{diag}[\lambda_0\cdots\lambda_{j-1}\ \ \lambda_{j+1}\cdots\lambda_{K-1}],\\
&\mathbf{U}_{I\setminus j} \triangleq [\mathbf{u}_1\cdots\mathbf{u}_{j-1}\ \ \mathbf{u}_{j+1}\cdots\mathbf{u}_{K-1}].
\end{aligned}$$

Then, the PDECs that correspond to the FDECs presented in the previous section, as well as their output SINRs are given below.

$$\mathbf{d}_{\mathbf{s}\setminus j} = \mathbf{S}_{\setminus j}\mathbf{X}_{\setminus j}^{-1}\mathbf{e}_1, \tag{4.8}$$

$$\begin{aligned}
\gamma_{\mathbf{s}\setminus j} &= \frac{(\mathbf{d}_{\mathbf{s}\setminus j}^T\mathbf{s}_0)^2E_0}{\|\mathbf{d}_{\mathbf{s}\setminus j}\|^2\sigma^2 + E_j(\mathbf{d}_{\mathbf{s}\setminus j}^T\mathbf{s}_j)^2}\\
&= \frac{E_0}{\sigma^2\{\mathbf{X}_{\setminus j}^{-1}\}_{1,1} + E_j(\mathbf{e}_1^T\mathbf{X}_{\setminus j}^{-1}\boldsymbol{\psi}_j)^2},
\end{aligned} \tag{4.9}$$

$$\begin{aligned}
\mathbf{d}_{ED1\setminus j} &= \mathbf{Q}_{S\setminus j}\left(\mathbf{\Lambda}_{S\setminus j} - \sigma^2\mathbf{I}_{K-1}\right)^{-1}\mathbf{Q}_{S\setminus j}^T\mathbf{s}_0\\
&= \mathbf{d}_{ED1} - \frac{\mathbf{q}_j\mathbf{q}_j^T}{\lambda_j - \sigma^2}\mathbf{s}_0,
\end{aligned} \tag{4.10}$$

[6] Alternatively, $\mathbf{R}_{I+N}$ can be estimated in a decision-directed way, $\widehat{\mathbf{R}}_{I+N} = \frac{1}{M}\sum_{m=1}^{M}\left[\mathbf{r}_m - \widehat{b}_0(m)\widehat{\sqrt{E_0}\mathbf{s}_0}\right]\left[\mathbf{r}_m - \widehat{b}_0(m)\ \widehat{\sqrt{E_0}\mathbf{s}_0}\right]^T$. The bit estimates $\widehat{b}_0(m)$ may initially be the outputs of the matched filter detector, then switched to the outputs of the decorrelator detector when enough samples are accumulated. A simple estimator of the product $\sqrt{E_0}\mathbf{s}_0$ is $\widehat{\sqrt{E_0}\mathbf{s}_0} = \frac{1}{M}\sum_{m=1}^{M}\widehat{b}_0(m)\mathbf{r}_m$ [14]-[15].

[7] The construction procedures and conditions under which PDECs outperform the corresponding FDECs described in this section can be generalized in a straightforward manner to the case where more than one signature or eigenvector direction is eliminated.

$$\gamma_{ED1\backslash j} = \frac{(\mathbf{d}_{ED1\backslash j}^T \mathbf{s}_0)^2 E_0}{\mathbf{d}_{ED1\backslash j}^T \mathbf{R} \mathbf{d}_{ED1\backslash j} - (\mathbf{d}_{ED1\backslash j}^T \mathbf{s}_0)^2 E_0}, \tag{4.11}$$

$$\mathbf{d}_{ED2\backslash j} = \mathbf{s}_0 - \mathbf{U}_{I\backslash j} \mathbf{U}_{I\backslash j}^T \mathbf{s}_0 = \mathbf{d}_{ED2} + \mathbf{u}_j \mathbf{u}_j^T \mathbf{s}_0, \tag{4.12}$$

$$\begin{aligned} \gamma_{ED2\backslash j} &= \frac{(\mathbf{d}_{ED2\backslash j}^T \mathbf{s}_0)^2 E_0}{\|\mathbf{d}_{ED2\backslash j}\|^2 \sigma^2 + (d_j - \sigma^2)(\mathbf{s}_0^T \mathbf{u}_j)^2} \\ &= \frac{E_0}{\sigma^2} \frac{[\ \|\mathbf{U}_{I^\perp}^T \mathbf{s}_0\|^2 + (\mathbf{s}_0^T \mathbf{u}_j)^2\ ]^2}{\|\mathbf{U}_{I^\perp}^T \mathbf{s}_0\|^2 + d_j/\sigma^2 (\mathbf{s}_0^T \mathbf{u}_j)^2}. \end{aligned} \tag{4.13}$$

The following proposition identifies conditions on the interference SNR and the signature cross-correlations under which the PDEC outperforms the FDEC in terms of output SINR.

Proposition 4.1 **(i)** $\gamma_{\mathbf{s}\backslash j} > \gamma_{\mathbf{s}}$ *if and only if* $\frac{E_j}{\sigma^2} < \left(1 - \boldsymbol{\psi}_j^T \mathbf{X}_{\backslash j}^{-1} \boldsymbol{\psi}_j\right)^{-1}$, **(ii)** $\gamma_{ED1\backslash j} > \gamma_{ED1}$ *if and only if* $\frac{\gamma_{ED1}}{1+\gamma_{ED1}} > E_0 \left[\frac{2(\lambda_j - \sigma^2)}{\lambda_j} \sum_{k=0}^{K-1} \frac{(\mathbf{q}_k^T \mathbf{s}_0)^2}{\lambda_k - \sigma^2} - \frac{(\mathbf{q}_j^T \mathbf{s}_0)^2}{\lambda_j}\right]$, **(iii)** $\gamma_{ED2\backslash j} > \gamma_{ED2}$ *if and only if* $\frac{d_j}{\sigma^2} < 2 + \frac{(\mathbf{u}_j^T \mathbf{s}_0)^2}{\|\mathbf{U}_{I^\perp}^T \mathbf{s}_0\|^2}$.

Proof: **(i)** *Let* $\widetilde{\mathbf{S}} \triangleq [\mathbf{S}_{\backslash j}\ \ \mathbf{s}_j]$ *and* $\widetilde{\mathbf{X}} \triangleq \widetilde{\mathbf{S}}^T \widetilde{\mathbf{S}}$. *There exists a permutation matrix* $\mathbf{P}_{K \times K}$ *such that* $\widetilde{\mathbf{S}} = \mathbf{S}\mathbf{P}$, *thus,* $\widetilde{\mathbf{X}} = (\mathbf{S}\mathbf{P})^T (\mathbf{S}\mathbf{P}) = \mathbf{P}^T \mathbf{X} \mathbf{P}$ *and* $\widetilde{\mathbf{X}}^{-1} = \mathbf{P}^T \mathbf{X}^{-1} \mathbf{P}$. *In addition,* $\{\widetilde{\mathbf{X}}^{-1}\}_{1,1} = \{\mathbf{X}^{-1}\}_{1,1}$,

$$\widetilde{\mathbf{X}}^{-1} = \begin{bmatrix} \mathbf{X}_{\backslash j} & \boldsymbol{\psi}_j \\ \boldsymbol{\psi}_j^T & 1 \end{bmatrix}^{-1} = \tag{4.14}$$

$$\begin{bmatrix} \left[\mathbf{X}_{\backslash j} - \boldsymbol{\psi}_j \boldsymbol{\psi}_j^T\right]^{-1} & \mathbf{X}_{\backslash j}^{-1} \boldsymbol{\psi}_j \left[\boldsymbol{\psi}_j^T \mathbf{X}_{\backslash j}^{-1} \boldsymbol{\psi}_j - 1\right]^{-1} \\ \left[\boldsymbol{\psi}_j^T \mathbf{X}_{\backslash j}^{-1} \boldsymbol{\psi}_j - 1\right]^{-1} \boldsymbol{\psi}_j^T \mathbf{X}_{\backslash j}^{-1} & \left[1 - \boldsymbol{\psi}_j^T \mathbf{X}_{\backslash j}^{-1} \boldsymbol{\psi}_j\right]^{-1} \end{bmatrix},$$

and

$$\left[\mathbf{X}_{\backslash j} - \boldsymbol{\psi}_j \boldsymbol{\psi}_j^T\right]^{-1} = \mathbf{X}_{\backslash j}^{-1} + \mathbf{X}_{\backslash j}^{-1} \boldsymbol{\psi}_j \boldsymbol{\psi}_j^T \mathbf{X}_{\backslash j}^{-1} / [1 - \boldsymbol{\psi}_j^T \mathbf{X}_{\backslash j}^{-1} \boldsymbol{\psi}_j]. \tag{4.15}$$

Then, (4.5) and (4.9) lead to the following expression:

$$\begin{aligned} & E_0 \left(\frac{1}{\gamma_{\mathbf{s}\backslash j}} - \frac{1}{\gamma_{\mathbf{s}}}\right) \\ = & \left[\sigma^2 \{\mathbf{X}_{\backslash j}^{-1}\}_{1,1} + E_j (\mathbf{e}_1^T \mathbf{X}_{\backslash j}^{-1} \mathbf{S}_{\backslash j}^T \mathbf{s}_j)^2\right] - \sigma^2 \{\mathbf{X}^{-1}\}_{1,1} \\ = & E_j (\mathbf{e}_1^T \mathbf{X}_{\backslash j}^{-1} \boldsymbol{\psi}_j)^2 - \sigma^2 \left[\{\widetilde{\mathbf{X}}^{-1}\}_{1,1} - \{\mathbf{X}_{\backslash j}^{-1}\}_{1,1}\right] \\ = & (\mathbf{e}_1^T \mathbf{X}_{\backslash j}^{-1} \boldsymbol{\psi}_j)^2 \left[E_j - \sigma^2 / (1 - \boldsymbol{\psi}_j^T \mathbf{X}_{\backslash j}^{-1} \boldsymbol{\psi}_j)\right]. \end{aligned} \tag{4.16}$$

Therefore, $\gamma_{\mathbf{s}\backslash j} > \gamma_{\mathbf{s}}$ *if and only if* $\frac{E_j}{\sigma^2} < \left(1 - \boldsymbol{\psi}_j^T \mathbf{X}_{\backslash j}^{-1} \boldsymbol{\psi}_j\right)^{-1}$.

(ii) *Instead of a direct comparison of* γ_{ED1} *and* $\gamma_{ED1\backslash j}$, *we evaluate (and compare) the function* $f(\gamma) = \frac{\gamma}{1+\gamma}$ *for* $\gamma = \gamma_{ED1}$ *and* $\gamma = \gamma_{ED1\backslash j}$ *(we note that* $f(\gamma)$ *is monotonically increasing for* $\gamma \geq 0$*).*

$$\begin{aligned} & f(\gamma_{ED1\backslash j}) = \frac{E_0 (\mathbf{d}_{ED1\backslash j}^T \mathbf{s}_0)^2}{\mathbf{d}_{ED1\backslash j}^T \mathbf{R} \mathbf{d}_{ED1\backslash j}} \\ = & \frac{E_0 \left[\mathbf{d}_{ED1}^T \mathbf{s}_0 - \frac{(\mathbf{q}_j^T \mathbf{s}_0)^2}{\lambda_j - \sigma^2}\right]^2}{\mathbf{d}_{ED1}^T \mathbf{R} \mathbf{d}_{ED1} - 2 \frac{\mathbf{q}_j^T \mathbf{s}_0}{\lambda_j - \sigma^2} \mathbf{d}_{ED1}^T \mathbf{R} \mathbf{q}_j + \frac{(\mathbf{q}_j^T \mathbf{s}_0)^2}{(\lambda_j - \sigma^2)^2} \mathbf{q}_j^T \mathbf{R} \mathbf{q}_j} \\ = & \frac{E_0 \left[(\mathbf{d}_{ED1}^T \mathbf{s}_0)^2 - 2 \frac{(\mathbf{q}_j^T \mathbf{s}_0)^2}{\lambda_j - \sigma^2} \mathbf{d}_{ED1}^T \mathbf{s}_0 + \frac{(\mathbf{q}_j^T \mathbf{s}_0)^4}{(\lambda_j - \sigma^2)^2}\right]}{\mathbf{d}_{ED1}^T \mathbf{R} \mathbf{d}_{ED1} - \frac{(\mathbf{q}_j^T \mathbf{s}_0)^2}{(\lambda_j - \sigma^2)^2} \lambda_j}, \end{aligned}$$

$$f(\gamma_{ED1}) > E_0 \left[\frac{2(\lambda_j - \sigma^2)}{\lambda_j} \mathbf{d}_{ED1}^T \mathbf{s}_0 - \frac{(\mathbf{q}_j^T \mathbf{s}_0)^2}{\lambda_j} \right]$$

$$\iff \quad f(\gamma_{ED1}) < f(\gamma_{ED1\backslash j}) \iff \gamma_{ED1} < \gamma_{ED1\backslash j} \tag{4.17}$$

Thus, $\gamma_{ED1} < \gamma_{ED1\backslash j}$ *if and only if* $f(\gamma_{ED1}) > E_0 \left[\frac{2(\lambda_j - \sigma^2)}{\lambda_j} \mathbf{d}_{ED1}^T \mathbf{s}_0 - \frac{(\mathbf{q}_j^T \mathbf{s}_0)^2}{\lambda_j} \right]$.

(iii) *The output SINR of* $\mathbf{d}_{ED2}$ *is* $\gamma_{ED2} = \frac{(\mathbf{d}_{ED2}^T \mathbf{s}_0)^2 E_0}{\|\mathbf{d}_{ED2}\|^2 \sigma^2} = \frac{E_0}{\sigma^2} \|\mathbf{U}_{I^\perp}^T \mathbf{s}_0\|^2$. *Then,*

$$\begin{aligned} & \frac{\sigma^2}{E_0} \left(\gamma_{ED2} - \gamma_{ED2\backslash j} \right) \\ = \quad & \|\mathbf{U}_{I^\perp}^T \mathbf{s}_0\|^2 - \frac{[\, \|\mathbf{U}_{I^\perp}^T \mathbf{s}_0\|^2 + (\mathbf{s}_0^T \mathbf{u}_j)^2 \,]^2}{\|\mathbf{U}_{I^\perp}^T \mathbf{s}_0\|^2 + d_j/\sigma^2 (\mathbf{s}_0^T \mathbf{u}_j)^2} \\ = \quad & c' \left[\left(\frac{d_j}{\sigma^2} - 2 \right) \|\mathbf{U}_{I^\perp}^T \mathbf{s}_0\|^2 - (\mathbf{s}_0^T \mathbf{u}_j)^2 \right] \end{aligned} \tag{4.18}$$

where c' *is a positive scalar. Therefore, whenever* $\frac{d_j}{\sigma^2} < 2 + \frac{(\mathbf{s}_0^T \mathbf{u}_j)^2}{\|\mathbf{U}_{I^\perp}^T \mathbf{s}_0\|^2}$, *we have* $\gamma_{ED2} < \gamma_{ED2\backslash j}$. □

4.5 Simulation Results

We consider a DS/CDMA system with $K = 10$ active users and processing gain $L = 31$. First, we examine the synchronous single-path case. The signature cross-correlation between the interfering users and *user 0* is in the range (0.03, 0.29). The input SNR of *user 0* is fixed at 12 dB while the input SNR of the interferers are fixed between 8 and 14 dB.

Fig. 4.1(a) shows the output SINR of the *ideal* PDECs as a function of the index of the signature or eigenvector that is eliminated from the FDEC structure when forming the corresponding PDEC. The output SINR of the ideal MF and FDEC are also included as reference points. Fig. 4.1(b) shows the corresponding BER values. We see that for this specific system setup, the PDEC that ignores the signature or eigenvector with index j=9 has higher SINR and lower BER than the FDEC.

Fig. 4.2(a) shows the output SINR of *estimated* full and partial (j=9) decorrelators as a function of the data record size; Fig. 4.2(b) presents the corresponding BER values. The output SINR and BER values are averages over 400 independent decorrelator estimates (data record realizations). $\mathbf{R}_{I+N}$ is estimated during a silent period of the user of interest. For comparison purposes, we also included in our study the blind decorrelating algorithm of [8].

In Fig. 4.3, we assume that each user experiences a 3-path Rayleigh fading channel. The propagation delays of the users are independent and uniformly distributed over a bit interval. We assume that the receiver is synchronized with respect to the first path of the user of interest. To illustrate acceptable BER levels, we consider a 3 dB total input SNR increase for all users in comparison with the corresponding SNR values used in Figs. 1 and 2. We evaluate the output SINR and BER performance of the estimated FDECs, the blind decorrelating algorithm of [8], as well as the performance of the *best* PDEC of type ED1 and ED2.[8] The results that we present are averages over 10000 runs (1000 independent asynchronous multi-path fading channel realizations and 10 independent received data record generations per channel). Figs. 2 and 3 demonstrate that in environments where the data support that is available for adaptation and redesign is limited, ED2-type decorrelators outperform ED1-type, both under full and partial decorrelation.

[8]The best PDEC is the PDEC that exhibits maximum output SINR and is found by exhaustive search over all PDECs. The SINR of each candidate filter is evaluated using the ideal autocorrelation matrix $\mathbf{R}_{I+N}$. As a result, the SINR (BER) performance curves of Figs. 2-3 represent upper (lower) bounds of the SINR (BER) performance that would be attained if PDEC selection were based on SINR estimates. Although Proposition 4.1 can form the basis for the development of efficient PDEC selection methods, such investigation is beyond the scope of this work.

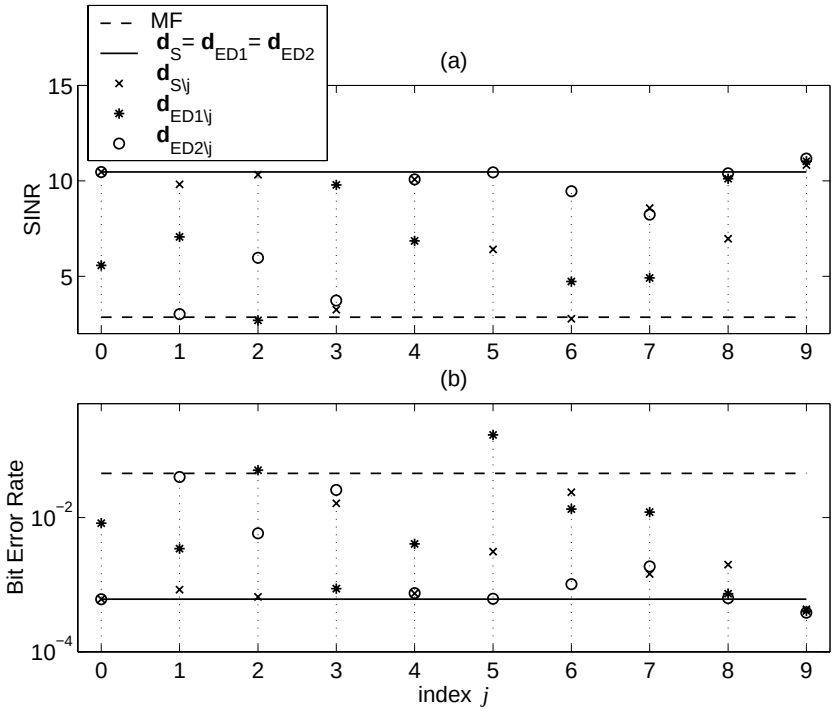

Fig. 4.1. Output SINR and BER of PDECs as a function of the index of the signature or eigenvector that is eliminated from the corresponding FDEC (synchronous system).

4.6 Conclusions

In this work, we investigated the relative output SINR performance of full and partial decorrelators and we derived necessary and sufficient conditions under which the latter outperform the former in the output SINR sense. We showed that decorrelation of MAI that satisfies certain conditions should be purposefully avoided. As a side study we considered a blind implementation of the decorrelator that is based on eigendecomposition of the desired-signal-absent autocovariance matrix. When estimation of the decorrelator is based on limited data support, the latter implementation exhibits significantly better SINR and BER performance than the decorrelator that is based on eigendecomposition of the desired-signal-present autocovariance matrix.

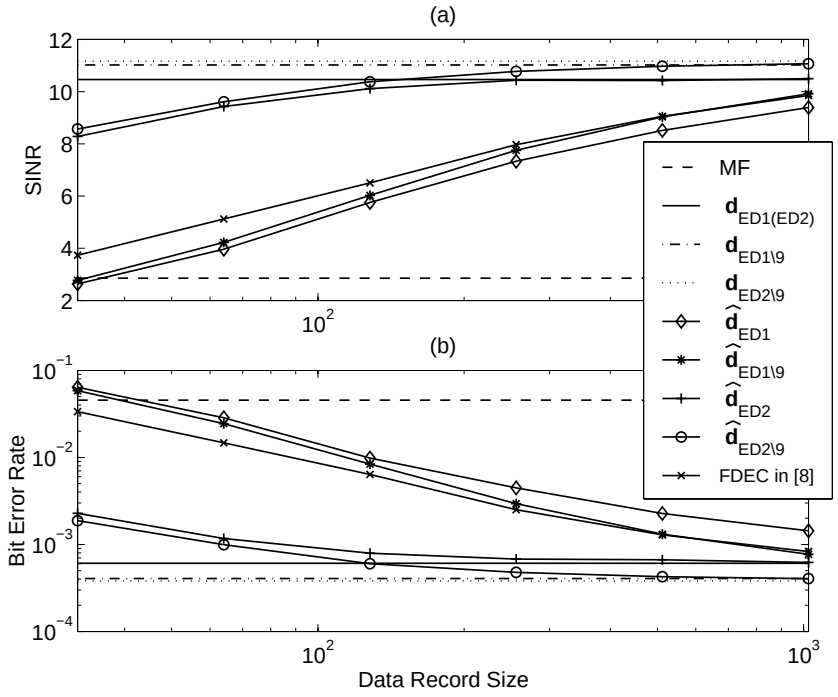

Fig. 4.2. SINR and BER of estimated FDECs and PDECs as a function of the data record size (synchronous system).

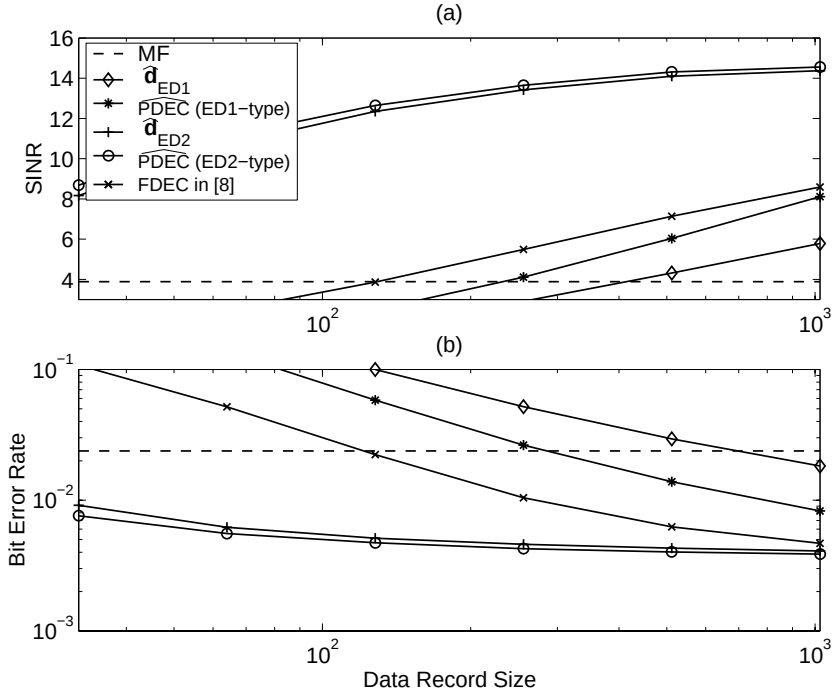

Fig. 4.3. SINR and BER of estimated FDECs and PDECs (best single eigenvector elimination per realization) as a function of the data record size (asynchronous multipath-fading channel).

Bibliography

[1] S. Verdú, *Multiuser Detection*. Cambridge, U. K.: Cambridge Univ. Press, 1998.

[2] G. V. Moustakides and H. V. Poor, "On the relative error probabilities of linear multiuser detectors," *IEEE Trans. Inform. Theory*, vol. 47, pp. 450-456, Jan. 2001.

[3] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 123-136, Jan. 1989.

[4] U. Mitra and H. V. Poor, "Analysis of an adaptive decorrelating detector for synchronous CDMA channels," *IEEE Trans. Commun.*, vol. 44, pp. 257-268, Feb. 1996.

[5] U. Mitra and H. V. Poor, "Adaptive decorrelating detector for CDMA systems," *Wireless Personal Commun.*, vol. 2, pp. 265-290, Dec. 1995.

[6] M. J. Juntti, "Linear multiuser detector update in synchronous dynamic CDMA systems," in *Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Commun.*, Toronto, Canada, Sept. 1995, vol. 3, pp. 980-984.

[7] X. Wang and H. V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, pp. 677-690, Mar. 1998.

[8] S. Ulukus and R. Yates, "A blind adaptive decorrelating detector for CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1530-1541, Oct. 1998.

[9] S. J. Baines, T. C. Tozer, and A. G. Burr, "Performance limits for multi-user decorrelating detectors in DS-CDMA cellular radio systems," in *Proc. IEEE Int. Symp. Spread Spectrum Techniques and Applications*, Mainz, Germany, Sept. 1996, vol. 2, pp. 486-491.

[10] M. Saquib and R. Yates, "Analysis of a partial decorrelator in a multi-cell DS-CDMA system," in *Proc. IEEE Globecom*, Rio de Janeiro, Brazil, Dec. 1999, vol. 4, pp. 2193-2197.

[11] I. P. Kirsteins and D. W. Tufts, "Adaptive detection using low rank approximation to a data matrix," *IEEE Trans. Aerosp. and Electr. Syst.*, vol. 30, pp. 55-67, Jan. 1994.

[12] A. Haimovich, "The eigencanceler: Adaptive radar by eigenanalysis methods," *IEEE Trans. Aerosp. and Electr. Syst.*, vol. 32, pp. 532-542, Apr. 1996.

[13] A. Haimovich, A. Shah, and X. Wu, "Reduced-rank array processing for wireless communications with applications to IS-54/IS-136," *IEEE Trans. Commun.*, vol. 48, pp. 743-747, May 2000.

[14] I. N. Psaromiligkos and S. N. Batalama, "Interference-plus-noise covariance matrix estimation for adaptive space-time processing of DS/CDMA signals," in *Proc. IEEE Vehic. Tech. Conf.*, Boston, MA, Sept. 2000, vol. 5, pp. 2197-2204.

[15] I. N. Psaromiligkos and S. N. Batalama, "Recursive AV and MVDR filter estimation for maximum SINR adaptive space-time processing," *IEEE Trans. Commun.*, accepted for publication.

[16] J. M. Farrell, I. N. Psaromiligkos, and S. N. Batalama, "Adaptive MMSE/MVDR filter estimators," in *Proc. SPIE, Digital Wireless Comm. Conf.*, Orlando, FL, Apr. 2003, vol. 5100.

[17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U. K.: Cambridge Univ. Press, 1985.

Chapter 5

All-Blind Adaptive Non-linear Spread-Spectrum Receivers

5.1 Introduction

Most of the research chapter on spread-spectrum (SS) signal detection in the presence of multiple access interference (MAI) and additive noise is based on the additive Gaussian noise assumption. Although it is often inaccurate to model the underlying noise of many physical channels as Gaussian [1], such an assumption is favored primarily because of its mathematical simplicity and tractability. It is well understood that the presence of non-Gaussian noise can degrade severely the performance of conventional signal processors which are based on the Gaussian assumption. In contrast, when the non-Gaussian nature of a physical channel is correctly modeled and accounted for during receiver design, the performance of the receiver may benefit significantly since the Gaussian noise is the most entropic among all noise models with full real/complex support probability density function (pdf) and finite variance.

Signal detection in non-Gaussian (impulsive) noise has attracted research interest since the early 1960's (see for example the list of references in [2]). In particular, SS signal detection in the presence of non-Gaussian noise was studied in [2]-[6]. In [2], the analysis of an approximately optimum single-user receiver motivated a general adaptive receiver structure that is composed of an adaptive front-end chip-based nonlinear processor followed by an adaptive post-processing linear filter. This general structure combines the merits of both linear and nonlinear signal processing; the nonlinear front-processor suppresses the impulsive noise and the adaptive linear post-processor suppresses the MAI. The chapter in [2] was further extended in [6] where the emphasis of the problem under consideration (SS signal detection in the presence of MAI and additive non-Gaussian noise) was shifted to short-data-record optimized receiver design. While in both [2] and [6] the parameters of the front-end nonlinear processor were optimized through supervised training, in this chapter we propose new blind parameter optimization criteria and we develop an all-blind robust adaptive SS receiver.

5.2 Signal Model

For simplicity in presentation, we assume a synchronous DS-CDMA communications system. The developments in this chapter can be extended in a straightforward manner to accommodate asynchronous multipath fading transmissions as well. Let L and K denote the system processing gain and the number of active users in the system, respectively. We assume that each user is assigned a normalized signature vector $\mathbf{S}_k$, $k = 0, 1, \cdots, K-1$, where $\mathbf{S}_k[l] = \{\pm 1/\sqrt{L}\}$, $l = 1, 2, \cdots, L$. The binary antipodal $(+1, -1)$ information bits are assumed to be independent identically distributed and equally probable. After conventional chip-matched filtering and chip-rate sampling over one information bit period T, the baseband received signal vector takes the form

$$\mathbf{r} = \sum_{k=0}^{K-1} \sqrt{E_k} b_k \mathbf{S}_k + \mathbf{n} = \sqrt{E_0} b_0 \mathbf{S}_0 + \mathbf{I} + \mathbf{n} \tag{5.1}$$

where E_k and b_k denote the power and the information bit of user k, $\mathbf{I} \triangleq \sum_{k=1}^{K} \sqrt{E_k} b_k \mathbf{S}_k$ identifies MAI and $\mathbf{n}$ represents additive (non-Gaussian) noise. Our goal is to detect the information bit of the user of interest, say *user 0*.

5.3 Receiver Architecture

For the signal detection problem under consideration, the general receiver structure is shown in Fig. 5.1(a) [2], [6]. The receiver is composed of a front-end chip-based nonlinear pre-processor that is followed by a linear post-processing tap-weight filter.

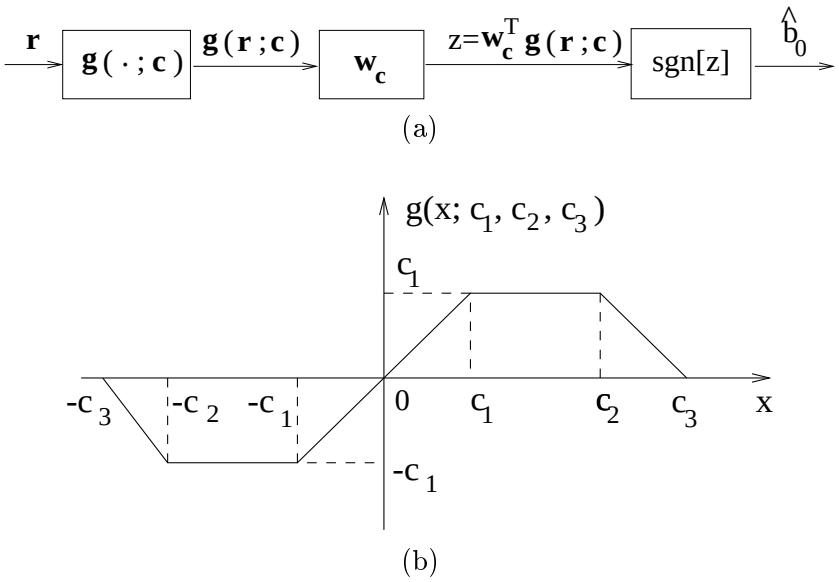

Fig. 5.1. (a) General receiver structure. (b) Hampel nonlinearity.

The information bit decision is $\hat{b}_0 = sgn[z]$, where $z = \mathbf{w}^T \mathbf{g}(\mathbf{r}; \mathbf{c})$ is the linear tap-weight filter output, superscript T denotes transpose, and $sgn[\cdot]$ is the sign operator. The nonlinear pre-processor, $\mathbf{g}(\mathbf{r}; \mathbf{c}) = [g(r_1; \mathbf{c}) \cdots g(r_L; \mathbf{c})]^T$, considered in this chapter is of Hampel-type:

$$
\begin{aligned}
g(x; \mathbf{c}) &= g(x; c_1, c_2, c_3) \\
&\triangleq \begin{cases} x, & \text{if } |x| < c_1 \\ c_1 \frac{x}{|x|}, & \text{if } c_1 \le |x| < c_2 \\ \frac{c_3 - |x|}{c_3 - c_2} c_1 \frac{x}{|x|}, & \text{if } c_2 \le |x| < c_3 \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
\tag{5.2}
$$

where $\mathbf{c} = [c_1, c_2, c_3]^T$ is the cutoff parameter vector to be obtained adaptively ($0 < c_1 \le c_2 \le c_3$). In (5.2), x is a complex number and $|x|$ denotes the magnitude of x. The Hampel nonlinearity preserves the received measurements whose magnitude is less than the threshold value c_1 and either de-emphasizes or discards higher magnitude measurements when $c_1 \le |x| \le c_3$ or $|x| > c_3$, respectively. The real valued version of (5.2) is shown in Fig. 5.1(b). We note that if the cutoff parameters are such that $0 < c_1 = c_2 = c_3$, then the Hampel nonlinear processor degenerates to the puncher. As $c_2 \to \infty$ the Hampel operator converges to the clipper. With respect to the linear filter post-processor $\mathbf{w}$, we suggest as candidates and we investigate both the minimum-variance-distortionless-response (MVDR) and the auxiliary-vector (AV) filters. We recall that the MVDR filter, $\mathbf{w}_{MVDR}$, that is distortionless in the vector direction of the user signal of interest $\mathbf{S}_0$ is given

by

$$\mathbf{w}_{MVDR} = \frac{\mathbf{R}_{\mathbf{c}}^{-1}\mathbf{S}_0}{\mathbf{S}_0^T\mathbf{R}_{\mathbf{c}}^{-1}\mathbf{S}_0} \tag{5.3}$$

where $\mathbf{R}_{\mathbf{c}} \triangleq E\{\mathbf{g}(\mathbf{r};\mathbf{c})\mathbf{g}^T(\mathbf{r};\mathbf{c})\}$ is the autocovariance matrix of the Hampel nonlinearly processed input data ($E\{\cdot\}$ denotes the statistical expectation operation). On the other hand, the AV algorithm generates a sequence of AV filters, $\mathbf{w}_{AV}^{(n)}$, $n = 0, 1, \cdots$, that are distortionless in the $\mathbf{S}_0$ direction as follows:

$$\begin{aligned}
\mathbf{w}_{AV}^{(0)} &= \mathbf{S}_0. && (5.4)\\
\text{For } n &= 1, 2, \cdots \\
\mathbf{G}^{(n)} &= \mathbf{R}_{\mathbf{c}}\mathbf{w}_{AV}^{(n-1)} - \mathbf{S}_0^T\mathbf{R}_{\mathbf{c}}\mathbf{w}_{AV}^{(n-1)}\mathbf{S}_0, && (5.5)\\
\mu^{(n)} &= \frac{\mathbf{G}^{(n)T}\mathbf{R}_{\mathbf{c}}\mathbf{w}_{AV}^{(n-1)}}{\mathbf{G}^{(n)T}\mathbf{R}_{\mathbf{c}}\mathbf{G}^{(n)}}, && (5.6)\\
\mathbf{w}_{AV}^{(n)} &= \mathbf{w}_{AV}^{(n-1)} - \mu^{(n)}\mathbf{G}^{(n)}. && (5.7)
\end{aligned}$$

We review briefly the basic properties of the sequence of AV filters $\mathbf{w}_{AV}^{(1)}, \mathbf{w}_{AV}^{(2)}, \cdots, \mathbf{w}_{AV}^{(n)}, \cdots$ [7]. Under perfectly known input statistics (matrix $\mathbf{R}_{\mathbf{c}}$), the sequence converges to the ideal MVDR solution $\mathbf{w}_{MVDR}$ in (5.3). When the autocovariance matrix $\mathbf{R}_{\mathbf{c}}$ is estimated from a finite data record through sample averaging, $\widehat{\mathbf{R}}_{\mathbf{c}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{g}(\mathbf{r}_n;\mathbf{c})\mathbf{g}^T(\mathbf{r}_n;\mathbf{c})$, then the early non-asymptotic elements of the sequence of AV filter estimates $\widehat{\mathbf{w}}_{AV}^{(1)}, \widehat{\mathbf{w}}_{AV}^{(2)}, \cdots$ offer favorable bias/variance balance and outperform in mean-square (MS) filter estimation error the sample-matrix-inversion (SMI) MVDR filter estimate.[1] Therefore, we need to emphasize that the sequence $\widehat{\mathbf{w}}_{AV}^{(1)}, \widehat{\mathbf{w}}_{AV}^{(2)}, \cdots$ can be viewed not only as a sequence of estimates of the ideal AV filters but, most importantly, as a sequence of estimates of the ideal MVDR filter. Data-record-based criteria for the selection of the best AV estimate of the MVDR filter were recently reported in [8].

5.4 Algorithmic Developments

In [2], [6], the cutoff parameter vector $\mathbf{c}$ of the front-end nonlinearity $\mathbf{g}(\mathbf{r};\mathbf{c})$ was obtained through supervised training according to either a mean-square-error (MSE) or a bit-error-rate (BER) optimization criterion. In this chapter we develop and analyze blind techniques for the estimation of the cutoff parameter vector of the nonlinear processor. [2]

We propose to use as a criterion for the selection of the parameter vector $\mathbf{c}$ the Fisher discriminant function [9] evaluated at the output of the linear post-processor filter. The Fisher discriminant function $\mathcal{F}(\mathbf{c},\mathbf{w})$ for the SS signal model under consideration takes the form

$$\mathcal{F}(\mathbf{c},\mathbf{w}) = \frac{2\mu^2(\mathbf{c},\mathbf{w})}{\sigma_{I+n}^2(\mathbf{c},\mathbf{w})} \tag{5.8}$$

where $\mathbf{c}$ denotes the cutoff parameter vector, $\mathbf{w}$ denotes a given linear post-processor filter, $\mu(\mathbf{c},\mathbf{w})$ denotes the conditional mean of the linear filter output conditioned on $b_0 = +1$, and $\sigma_{I+n}^2(\mathbf{c},\mathbf{w})$ is the corresponding output variance due to MAI and noise.[3] We note that if $\mathbf{w}$ is an adaptive filter such as $\mathbf{w}_{MVDR}$ or $\mathbf{w}_{AV}$, then it depends on the parameter $\mathbf{c}$ through the input autocovariance matrix $\mathbf{R}_{\mathbf{c}}$. Maximization of the Fisher discriminant function with respect to $\mathbf{c}$ aims at creating two well separated clusters of filter outputs, one corresponding to $b_0 = +1$ and the other corresponding to $b_0 = -1$.

[1] $\widehat{\mathbf{w}}_{AV}^{(n)}, n = 1, 2, \cdots$, are obtained from (5.4)-(5.7) with $\widehat{\mathbf{R}}_{\mathbf{c}}$ in place of $\mathbf{R}_{\mathbf{c}}$. Similarly, $\widehat{\mathbf{w}}_{SMI-MVDR} = \frac{\widehat{\mathbf{R}}_{\mathbf{c}}^{-1}\mathbf{S}_0}{\mathbf{S}_0^T\widehat{\mathbf{R}}_{\mathbf{c}}^{-1}\mathbf{S}_0}$ from (5.3). For a fixed data record $\lim_{n\to\infty}\widehat{\mathbf{w}}_{AV}^{(n)} = \widehat{\mathbf{w}}_{SMI-MVDR}$ [7].

[2] We recall that the evaluation of $\widehat{\mathbf{w}}_{SMI-MVDR}$ or $\widehat{\mathbf{w}}_{AV}$ is inherently unsupervised (blind).

[3] The linear filter output conditioned on $b_0 = -1$ has mean $-\mu(\mathbf{c},\mathbf{w})$.

Exploiting the symmetry of the conditional pdf's of the linear filter output given $b_0 = +1$ and $b_0 = -1$ (recall that the Hampel nonlinearity is also symmetric), we can show in a straightforward manner that

$$\mathcal{F}(\mathbf{c}, \mathbf{w_c}) = \frac{2E^2\left\{b_0 \mathbf{w_c}^T \mathbf{g}(\mathbf{r};\mathbf{c})\right\}}{Var\left\{b_0 \mathbf{w_c}^T \mathbf{g}(\mathbf{r};\mathbf{c})\right\}} \tag{5.9}$$

where $Var(\cdot)$ denotes variance and the filter subscript "$\mathbf{c}$" serves as a reminder that the filter is a function of $\mathbf{c}$. The Fisher discriminant function in (5.9) is of supervised nature, which is exactly what we would like to avoid in our developments. We propose to design a blind (or decision directed) version of (5.9) by substituting b_0 by $sgn\left[\mathbf{w_c}^T \mathbf{g}(\mathbf{r};\mathbf{c})\right]$ (output of the sign detector that follows the linear filter). Then,

$$\mathcal{F}_B(\mathbf{c}, \mathbf{w_c}) = \frac{2E^2\{sgn[\mathbf{w_c}^T \mathbf{g}(\mathbf{r};\mathbf{c})]\, \mathbf{w_c}^T \mathbf{g}(\mathbf{r};\mathbf{c})\}}{Var\{sgn[\mathbf{w_c}^T \mathbf{g}(\mathbf{r};\mathbf{c})]\, \mathbf{w_c}^T \mathbf{g}(\mathbf{r};\mathbf{c})\}} \tag{5.10}$$

$$= \frac{2E^2\{|\mathbf{w_c}^T \mathbf{g}(\mathbf{r};\mathbf{c})|\}}{Var\{|\mathbf{w_c}^T \mathbf{g}(\mathbf{r};\mathbf{c})|\}} \tag{5.11}$$

where the subscript "B" identifies the blind version of the Fisher discriminant function. The following proposition identifies conditions under which $\mathcal{F}_B(\mathbf{c}, \mathbf{w_c})$ is nearly equal to $\mathcal{F}(\mathbf{c}, \mathbf{w_c})$.

Proposition 5.1 *To the extent that the linear filter output conditional distributions (conditioned on the transmitted information bit of the user of interest) can be approximated by Gaussian pdf's, if* $\frac{\mu(\mathbf{c},\mathbf{w_c})}{\sigma_{I+n}(\mathbf{c},\mathbf{w_c})} \gg 1$, *i.e. the filter output SINR is significantly higher than 0dB, then* $\mathcal{F}_B(\mathbf{c}, \mathbf{w_c}) \approx \mathcal{F}(\mathbf{c}, \mathbf{w_c})$. □

To estimate $\mathcal{F}_B(\mathbf{c}, \mathbf{w_c})$ from a data record of finite size we substitute the statistical expectations in (5.11) by sample averages and we denote the estimate by $\widehat{\mathcal{F}}_B(\mathbf{c}, \widehat{\mathbf{w}}_\mathbf{c})$. The following criterion summarizes the procedure for the selection of the cutoff parameter $\mathbf{c}$.

Criterion 5.1 *For a given data record of size N, the Fisher discriminant selection rule chooses the cutoff parameter* $\mathbf{c}$ *such that*

$$\mathbf{c}_o = \arg\max_{\mathbf{c}} \widehat{\mathcal{F}}_B(\mathbf{c}, \widehat{\mathbf{w}}_\mathbf{c}) \tag{5.12}$$

□

Some critical observations now follow. We understand that Criterion 5.1 introduces a coupled maximization problem that can be solved by utilizing a gradient type adaptation for the parameter $\mathbf{c}$. In each iteration step the adaptation of $\mathbf{c}$ is followed by the evaluation of $\widehat{\mathbf{w}}_\mathbf{c}$ (be it $\widehat{\mathbf{w}}_{SMI-MVDR,\mathbf{c}}$ or $\widehat{\mathbf{w}}_{AV,\mathbf{c}}$). We also emphasize that the Fisher discriminant function is used for the selection of the cutoff parameter $\mathbf{c}$ only, while the linear filter is determined strictly according to the MVDR or AV filtering principles. The use of the Fisher discriminant criterion $\mathcal{F}_B(\cdot,\cdot)$ to select both $\mathbf{c}$ and $\mathbf{w}$ would return phase ambiguous filters that may be locked on a strong SS interferer instead of having a coherent lock on the signal of interest. Finally, we note that the Fisher discriminant function for the case of conditionally Gaussian linear filter outputs coincides with the J-divergence distance [4] between the corresponding pdf's [8].

To reduce the computational complexity of the above coupled optimization approach due to the adaptive evaluation of the linear post-processor at each iteration step, we examine the following decoupled optimization scheme. In the first step, we select the parameter $\mathbf{c}$ that maximizes $\widehat{\mathcal{F}}_B(\mathbf{c}, \mathbf{w}_{MF})$ (i.e. the linear filter is *fixed* at the user of interest signature-matched filter $\mathbf{w}_{MF} = \mathbf{S}_0$). In the second step, given the value of the parameter $\mathbf{c}$ returned by the first step, say $\mathbf{c}_1$, we calculate the adaptive filter $\widehat{\mathbf{w}}_{SMI-MVDR,\mathbf{c}_1}$ or $\widehat{\mathbf{w}}_{AV,\mathbf{c}_1}$ based on the input autocovariance matrix $\widehat{\mathbf{R}}_{\mathbf{c}_1}$. These two simple steps complete the decoupled scheme. There are two issues associated with the above two-step decoupled procedure that need to be addressed. The first issue involves the direction at which both the coupled and the decoupled gradient method for the evaluation of $\mathbf{c}$ is moving. This issue is treated in the following proposition. The second issue involves possible performance loss due to the decoupled optimization and it is examined through simulations in the next section.

[4] The J-divergence distance $J(f_1, f_2)$ between two distributions $f_1(\cdot)$ and $f_2(\cdot)$ is defined as the sum of the Kullback-Leibler (K-L) distances between f_1 and f_2: $J(f_1, f_2) \triangleq D(f_1, f_2) + D(f_2, f_1)$ where $D(f_1, f_2) \triangleq \int_{-\infty}^{\infty} f_1(x) \log \frac{f_1(x)}{f_2(x)} dx$ is the K-L distance from $f_2(\cdot)$ to $f_1(\cdot)$.

Proposition 5.2 *To the extent that the nonlinear pre-processor suppresses only the impulsive noise component while it passes undistorted the user of interest as well as the MAI component, both steepest descent recursive methods (coupled and decoupled) for the determination of* $\mathbf{c}$ *increase the value of the Fisher discriminant function at each iteration step.*

Proof: We can show in a straightforward manner that if $\mathbf{c}_1$ and $\mathbf{c}_2$ are two cutoff parameter vectors for which $\mathcal{F}_B(\mathbf{c}_1, \mathbf{w}_{MF}) > \mathcal{F}_B(\mathbf{c}_2, \mathbf{w}_{MF})$, then

$$\sigma^2_{I+n,g}(\mathbf{g}(\mathbf{r};\mathbf{c}_1)) < \sigma^2_{I+n,g}(\mathbf{g}(\mathbf{r};\mathbf{c}_2)), \tag{5.13}$$

$$\sigma^2_{n,g}(\mathbf{g}(\mathbf{r};\mathbf{c}_1)) < \sigma^2_{n,g}(\mathbf{g}(\mathbf{r};\mathbf{c}_2)), \tag{5.14}$$

$$\begin{aligned}\mathcal{F}_B(\mathbf{c}_1,\widehat{\mathbf{w}}_{SMI-MVDR,\mathbf{c}_1}) &> \mathcal{F}_B(\mathbf{c}_1,\widehat{\mathbf{w}}_{SMI-MVDR,\mathbf{c}_2}) \\ &> \mathcal{F}_B(\mathbf{c}_2,\widehat{\mathbf{w}}_{SMI-MVDR,\mathbf{c}_2}),\end{aligned} \tag{5.15}$$

$$\mathcal{F}_B(\mathbf{c}_1, \widehat{\mathbf{w}}_{AV,\mathbf{c}_1}) > \mathcal{F}_B(\mathbf{c}_1, \widehat{\mathbf{w}}_{AV,\mathbf{c}_2}) > \mathcal{F}_B(\mathbf{c}_2, \widehat{\mathbf{w}}_{AV,\mathbf{c}_2}) \tag{5.16}$$

where $\sigma^2_{I+n,g}$ and $\sigma^2_{n,g}$ denote the squared norm of a vector with elements the variances of the chip-nonlinearly processed samples of interference and noise $(I+n)$ or just noise (n), respectively. □

5.5 Simulation Results

We consider a system of 6 synchronous SS signals with processing gain 31. The normalized signature cross-correlations between the interfering signals and the signal of interest are in the range of 0.22 to 0.35 while the interfering signal signatures are approximately orthogonal to each other. The power of the interfering users (users 1-5) is fixed at 9, 10, 11, 12 and 13 dB, respectively. We also assume the presence of additive non-Gaussian noise under the familiar ϵ-contamination model

$$f_\epsilon(x) = (1-\epsilon)f_0(x) + \epsilon f_1(x) \tag{5.17}$$

where $f_0(x)$ denotes the ambient background Gaussian noise pdf. The contamination parameter $\epsilon \in [0,1]$ indicates the probability that the chip noise samples are $f_1(x)$ distributed. We assume that $f_0(x)$ and $f_1(x)$ are zero-mean Gaussian with variance $\sigma_0^2 = 1$ and $\sigma_1^2 = \gamma^2\sigma_0^2$ $(\gamma^2 = 1000)$, respectively, and we set ϵ equal to 0.2.

Fig. 5.2 examines the BER performance as a function of the data record size of the conventional MF and SMI-MVDR receivers (without nonlinear pre-processing) and the Hampel pre-processor followed by either the SMI-MVDR or the AV-filter estimator $\widehat{\mathbf{w}}_{AV}$. The Hampel-MF performance curve is also included for reference purposes. The power of the signal of interest is fixed at 12 dB. The solid lines indicate the performance of the coupled procedures while the dotted lines show the performance of the corresponding decoupled procedures. Fig. 5.3 reports the same studies as in Fig as a function of the power of the signal of interest. The estimation of the cutoff parameter $\mathbf{c}$ and the estimation of the adaptive filter ($\widehat{\mathbf{w}}_{SMI-MVDR,\mathbf{c}}$ or $\widehat{\mathbf{w}}_{AV,\mathbf{c}}$) is based on the same record of 128 data samples. The results in both Figs. 2 and 3 are averages over 150 independent Monte-Carlo runs. We observe that the decoupled procedures do not suffer substantial performance loss while they maintain significantly lower computational complexity when compared to the coupled procedures. We also note the performance gains that the Hampel AV-filter structure exhibits in low sample support situations when compared to the Hampel SMI-MVDR filter.

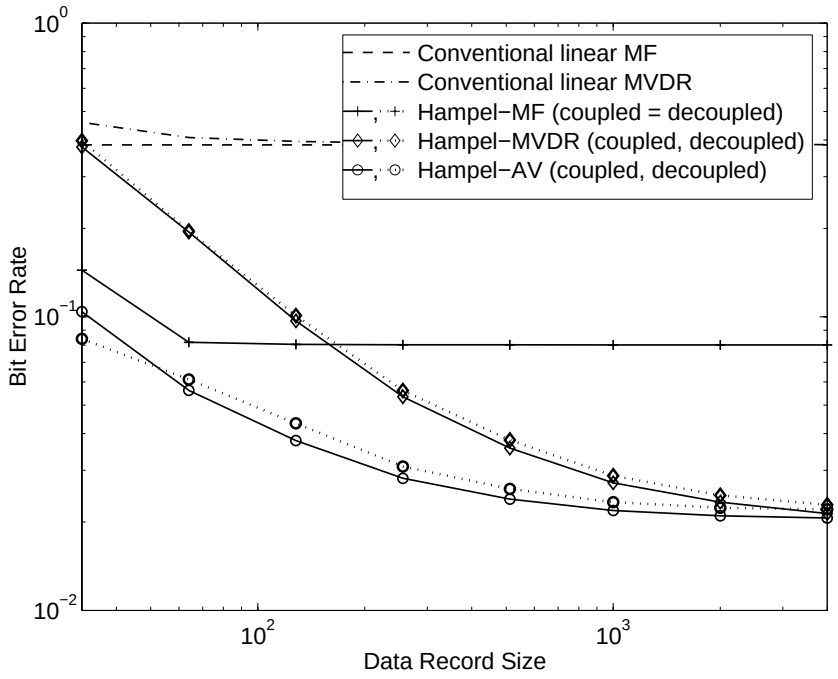

Fig. 5.2. BER as a function of the data record size.

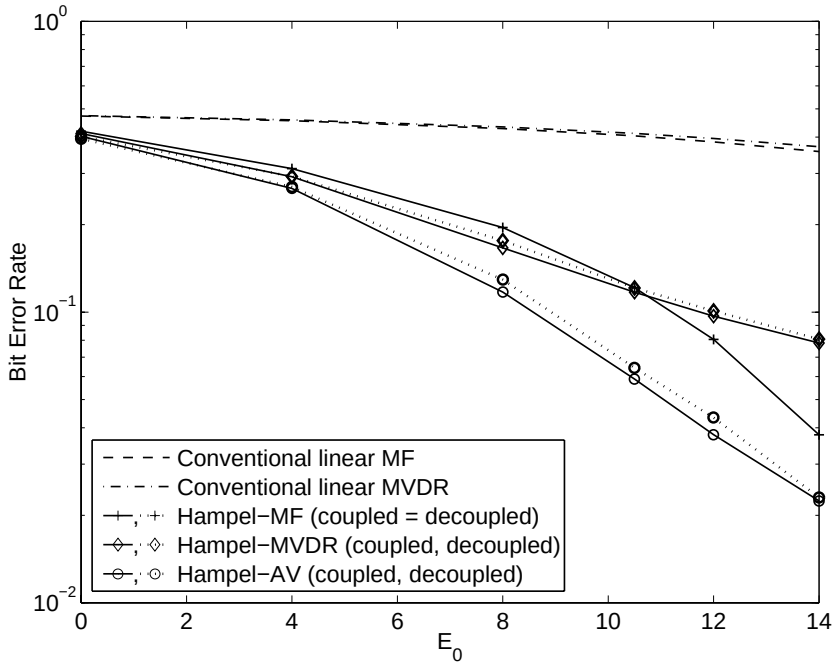

Fig. 5.3. BER as a function of the power of the user of interest.

Bibliography

[1] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: New methods and results for Class A and Class B noise models," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1129-1149, May 1999.

[2] S. N. Batalama, M. J. Medley, and I. N. Psaromiligkos, "Adaptive robust spread-spectrum receivers," *IEEE Trans. Commun.*, vol. 47, pp. 905-917, June 1999.

[3] B. Aazhang and H. V. Poor, "Performance of DS/SSMA communications in impulsive channels- Part II: Hard-limiting correlation receivers," *IEEE Trans. Commun.*, vol. 36, pp. 88-97, Jan. 1988.

[4] B. Aazhang and H. V. Poor, "An analysis of nonlinear direct-sequence correlators," *IEEE Trans. Commun.*, vol. 37, pp. 723-731, July 1989.

[5] X. Wang and H. V. Poor, "Robust multiuser detection in non-Gaussian channels," *IEEE Trans. Signal Processing*, vol. 47, pp. 289-305, Feb. 1999.

[6] S. N. Batalama, M. J. Medley, and D. A. Pados, "Robust adaptive recovery of spread-spectrum signals with short data records," *IEEE Trans. Commun.*, vol. 48, pp. 1725-1731, Oct. 2000.

[7] D. A. Pados and G. N. Karystinos, "An Iterative Algorithm for the Computation of the MVDR Filter," *IEEE Trans. Signal Processing*, vol. 49, pp. 290-300, Feb. 2001.

[8] H. Qian and S. N. Batalama, "Data-record-based criteria for the selection of an auxiliary-vector estimator of the MVDR filter," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, USA, 2000, vol. 1, pp. 802-807.

[9] R. J. Schalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches*. New York, NY: Wiley, 1992.

Chapter 6

Performance Analysis of Joint and Disjoint Space/Time Receivers in Fast Frequency-Nonselective Fading DS-CDMA Channels

6.1 Introduction

In this chapter we analyze the performance of DS-CDMA single-user antenna-array receivers that utilize either a cascade of spatial and temporal linear filters or one joint space-time linear filter. Receiver filters are estimated from limited sample support due to a fast frequency-nonselective fading channel considered. Linear filters are preferred due to their conceptual simplicity and low computational cost. Among them, the conventional matched filter (MF) receiver has the lowest complexity but exhibits poor performance in strong interference environments or when the user of interest is highly correlated with the interferers. On the other hand, the linear minimum-mean-square-error (MMSE) and the minimum-variance-distortionless-response (MVDR) receivers strike a balance between complexity and performance [1]. We consider both MF-type and MMSE/MVDR-type filters, and we use pre-detection signal-to-interference-plus-noise-ratio (SINR) as the figure of merit. To simplify the analysis, we confine ourselves within the class of filters that are distortionless in the signal direction of the user of interest. Our goal is twofold: *(i)* Quantify the receiver output variance and receiver output SINR performance of disjoint and joint space-time configurations given ideal input statistics; *(ii)* quantify the confidence level in a certain neighborhood of the optimal performance point of the corresponding estimated structures that are evaluated based on limited sample support.

6.2 Signal Model

We consider a synchronous DS-CDMA communication system. Let L and K denote the system processing gain and the number of active users in the system, respectively. We assume that each user is assigned a normalized signature vector $\mathbf{s}_k$, $k = 1, 2, \cdots K$, where $\mathbf{s}_k[l] = \{\pm 1/\sqrt{L}\}$, $l = 1, 2, \cdots, L$. The receiver employs a uniform linear antenna array with M antenna elements, spaced at half-the-wavelength apart. If the angle-of-arrival (AoA) of user-k is denoted by θ_k, then the array response vector (spatial signature) of user-k is defined as $\mathbf{a}_k \triangleq \left[1 \;\; e^{-j\pi \sin\theta_k} \;\; \ldots \;\; e^{-j(M-1)\pi \sin\theta_k}\right]^T$.

Without loss of generality, we assume chip synchronization with the user of interest (*user-1*) at the reference antenna element ($m = 1$). After conventional chip matched filtering and chip rate sampling, we can visualize the space-time data samples in the form of a complex matrix $\mathbf{X} \in \mathbb{C}^{M \times L}$

$$\mathbf{X}_{M\times L}(i) = [\mathbf{x}_{M\times 1}(iL+1) \;\; \mathbf{x}_{M\times 1}(iL+2) \;\; \ldots \;\; \mathbf{x}_{M\times 1}(iL+L)] \tag{6.1}$$

where the column vector $\mathbf{x}_{M\times 1}(iL+l)$ associated with the lth chip sample within the ith information bit

period is

$$\mathbf{x}_{M\times 1}(iL+l)=\sum_{k=1}^{K} b_k(i)\sqrt{E_k}c_k(l)\mathbf{a}_k+\mathbf{n}_{M\times 1}(iL+l),\ l=1,\ldots,L. \tag{6.2}$$

In (6.2), E_k, $b_k(i)$ and $c_k(l)$ denote the energy, the ith *binary* information bit and the lth signature chip of user-k, respectively. $\mathbf{n}_{M\times 1}(iL+l)\sim\mathcal{N}(\mathbf{0}_{M\times 1},\sigma^2\mathbf{I}_M)$ is an AWGN vector ($\mathbf{I}_M$ denotes the $M\times M$ identity matrix).

Let $\tilde{\mathbf{s}}_k\triangleq\mathbf{s}_k\otimes\mathbf{a}_k$ be the joint space-time signature of user-k, where $\tilde{\mathbf{s}}_k\in\mathbb{C}^{ML}$ and $\otimes$ denotes the Kronecker product. Let also

$$\begin{aligned}
&\mathbf{A}_{M\times K}\triangleq[\mathbf{a}_1\ \ \mathbf{a}_2\ \ \ldots\mathbf{a}_K],\quad \mathbf{S}_{L\times K}\triangleq[\mathbf{s}_1\ \ \mathbf{s}_2\ \ \ldots\mathbf{s}_K],\\
&\tilde{\mathbf{S}}_{ML\times K}\triangleq[\tilde{\mathbf{s}}_1\ \tilde{\mathbf{s}}_2\ \ldots\tilde{\mathbf{s}}_K],\quad \mathbf{E}_{K\times K}\triangleq\mathrm{diag}\,(E_1,\ E_2,\ \ldots,E_K),\\
&\mathbf{B}(i)_{K\times K}\triangleq\mathrm{diag}\,(b_1(i),\ \ b_2(i),\ \ \ldots,b_K(i)),\\
&\mathbf{N}(i)_{M\times L}\triangleq[\mathbf{n}(iL+1)\ \ \ldots\ \ \mathbf{n}(iL+L)],\\
&\mathbf{b}(i)_{K\times 1}\triangleq[b_1(i),\ \ b_2(i),\ldots,\ \ b_K(i)]^T.
\end{aligned} \tag{6.3}$$

Then, (6.1) can be written compactly as

$$\mathbf{X}(i)=\mathbf{A}\mathbf{E}^{1/2}\mathbf{B}(i)\mathbf{S}^T+\mathbf{N}(i). \tag{6.4}$$

We note that bold upper-case letters denote matrices while bold lower-case letters denote vectors. Superscripts T, H and $*$ denote transpose, Hermitian and conjugation operation, respectively.

In the following section, we summarize the receiver configurations that we consider in this chapter for one-shot detection of the information bit of the user of of interest (the index i in (6.4) is dropped for simplicity in notation).

6.3 Joint and Disjoint Space-Time Filter Configurations

While in disjoint domain processing we operate on the data matrix $\mathbf{X}$ directly, to facilitate joint domain processing, we vectorize the matrix $\mathbf{X}_{M\times L}$ by stacking all columns to form a vector as follows:

$$\mathcal{X}_{ML\times 1}\triangleq\mathrm{Vec}\{\mathbf{X}_{M\times L}\}=\tilde{\mathbf{S}}\mathbf{E}^{1/2}\mathbf{b}+\tilde{\mathbf{n}}, \tag{6.5}$$

where $\tilde{\mathbf{n}}=\mathrm{Vec}\{\mathbf{N}_{M\times L}\}$.

Let $\mathbf{w}_J\in\mathbb{C}^{ML}$, $\mathbf{w}_s\in\mathbb{C}^{M}$, and $\mathbf{w}_t\in\mathbb{C}^{L}$ denote an arbitrary joint space-time, spatial and temporal linear filter, respectively. Information bit decisions are made by taking the sign of the real part of the final output of the corresponding receiver configuration, i.e. information bits are estimated by $\widehat{b}_1=sgn\left[Re\{\mathbf{w}_J^H\mathcal{X}\}\right]$ in joint domain filtering, or by $\widehat{b}_1=sgn\left[Re\{\mathbf{w}_s^H\mathbf{X}\mathbf{w}_t\}\right]$ in disjoint domain filtering. We also assume that the filters $\mathbf{w}_J$, $\mathbf{w}_s$ and $\mathbf{w}_t$ are distortionless in the joint space-time, spatial and temporal direction of the user of interest, respectively, i.e. $\mathbf{w}_J^H\tilde{\mathbf{s}}_1=1$, $\mathbf{w}_s^H\mathbf{a}_1=1$, and $\mathbf{w}_t^H\mathbf{s}_1=1$. Let $\Gamma(\Omega)$ and $P(\Omega)$ denote output SINR and output variance, respectively, of an arbitrarily chosen configuration Ω. Since $\Gamma(\Omega)=E_1/\left[P(\Omega)-E_1\right]$, then lower receiver output variance implies higher receiver output SINR.

The joint space-time MF $\mathbf{w}_{JMF}$ and the joint MVDR filter $\mathbf{w}_{JMV}$ [2] for *user-1* are given by

$$\mathbf{w}_{JMF}=\tilde{\mathbf{s}}_1/M,\qquad \mathbf{w}_{JMV}=\frac{\mathbf{R}^{-1}\tilde{\mathbf{s}}_1}{\tilde{\mathbf{s}}_1^H\mathbf{R}^{-1}\tilde{\mathbf{s}}_1}, \tag{6.6}$$

where $\mathbf{R}\triangleq E\{\mathcal{X}\mathcal{X}^H\}=\tilde{\mathbf{S}}\mathbf{E}\tilde{\mathbf{S}}^H+\sigma^2\mathbf{I}$ is the auto-covariance matrix of the space-time received vector, $E\{\cdot\}$ denotes statistical expectation operation. The output variance of the joint matched filter is $P_{JMF}=\tilde{\mathbf{s}}_1^H\mathbf{R}\tilde{\mathbf{s}}_1/M^2$, while the output variance of the joint MVDR filter is $P_{JMV}=\left(\tilde{\mathbf{s}}_1^H\mathbf{R}^{-1}\tilde{\mathbf{s}}_1\right)^{-1}$.

On the other hand, disjoint domain processing can take the form of spatial filtering followed by temporal filtering or temporal filtering followed by spatial filtering. Clearly spatial MF combined with temporal

MF in any order is equivalent to joint MF (JMF). Below we list all other possible configurations (in the processing order), and evaluate the corresponding filters as well as the output variance of each configuration, as a function of M, L, $\mathbf{E}$, σ^2, $\mathbf{A}$ and $\mathbf{S}$ (in each configuration the left operation is the type of processing performed first). In the following, the letter s and t when used as part of the subscript of a vector or matrix indicates spatial and temporal filtering, respectively, while MF and MV refers to MF and MVDR type processing. The symbol $|$ is used as a "condition" indicator (e.g. $_{tMV|sMF}$ implies that temporal MVDR processing follows spatial MF processing).

(i) Spatial MF – temporal MVDR

$$\mathbf{w}_s = \mathbf{w}_{_{sMF}} \triangleq \mathbf{a}_1/M, \tag{6.7}$$

$$\mathbf{w}_t = \mathbf{w}_{_{tMV|sMF}} \triangleq \mathbf{R}^{-1}_{_{t2|sMF}}\mathbf{s}_1/\left(\mathbf{s}_1^H\mathbf{R}^{-1}_{_{t2|sMF}}\mathbf{s}_1\right), \tag{6.8}$$

$$P_{_{tMV|sMF}} = \left[\mathbf{s}_1^H\mathbf{R}^{-1}_{_{t2|sMF}}\mathbf{s}_1\right]^{-1}, \tag{6.9}$$

where $\mathbf{R}_{_{t2|sMF}} \triangleq E\left\{(\mathbf{w}^H_{_{sMF}}\mathbf{X})^H(\mathbf{w}^H_{_{sMF}}\mathbf{X})\right\}$

$$= \mathbf{S}^*\mathbf{E}^{1/2}\mathrm{diag}\left(\frac{\mathbf{A}^H\mathbf{a}_1\mathbf{a}_1^H\mathbf{A}}{M^2}\right)\mathbf{E}^{1/2}\mathbf{S}^T + \frac{\sigma^2}{M}\mathbf{I}.$$

(ii) Spatial MVDR – temporal MVDR

$$\mathbf{w}_s = \mathbf{w}_{_{sMV}} \triangleq \mathbf{R}^{-1}_{s1}\mathbf{a}_1/\left(\mathbf{a}_1^H\mathbf{R}^{-1}_{s1}\mathbf{a}_1\right), \tag{6.10}$$

$$\mathbf{w}_t = \mathbf{w}_{_{tMV|sMV}} \triangleq \mathbf{R}^{-1}_{_{t2|sMV}}\mathbf{s}_1/\left(\mathbf{s}_1^H\mathbf{R}^{-1}_{_{t2|sMV}}\mathbf{s}_1\right), \tag{6.11}$$

$$P_{_{tMV|sMV}} = \left[\mathbf{s}_1^H\mathbf{R}^{-1}_{_{t2|sMV}}\mathbf{s}_1\right]^{-1}, \tag{6.12}$$

where $\mathbf{R}_{s1} \triangleq E\left\{\mathbf{x}\mathbf{x}^H\right\} = \frac{1}{L}\mathbf{A}\mathbf{E}\mathbf{A}^H + \sigma^2\mathbf{I}$,

$\mathbf{R}_{_{t2|sMV}} \triangleq E\left\{(\mathbf{w}^H_{_{sMV}}\mathbf{X})^H(\mathbf{w}^H_{_{sMV}}\mathbf{X})\right\}$

$= \mathbf{S}^*\mathbf{E}^{1/2}\mathrm{diag}\left(\mathbf{A}^H\mathbf{w}_{_{sMV}}\mathbf{w}^H_{_{sMV}}\mathbf{A}\right)\mathbf{E}^{1/2}\mathbf{S}^T$

$+ \|\mathbf{w}_{_{sMV}}\|^2\sigma^2\mathbf{I}$.

(iii) Spatial MVDR – temporal MF

$$\mathbf{w}_s = \mathbf{w}_{_{sMV}}, \quad \mathbf{w}_t = \mathbf{w}_{_{tMF}} = \mathbf{s}_1,$$
$$P_{_{tMF|sMV}} = \mathbf{s}_1^H\mathbf{R}_{_{t2|sMV}}\mathbf{s}_1. \tag{6.13}$$

(iv) Temporal MF – spatial MVDR

$$\mathbf{w}_t = \mathbf{s}_1,$$

$$\mathbf{w}_s = \mathbf{w}_{_{sMV|tMF}} \triangleq \mathbf{R}^{-1}_{_{s2|tMF}}\mathbf{a}_1/\left(\mathbf{a}_1^H\mathbf{R}^{-1}_{_{s2|tMF}}\mathbf{a}_1\right), \tag{6.14}$$

$$P_{_{sMV|tMF}} = \left[\mathbf{a}_1^H\mathbf{R}^{-1}_{_{s2|tMF}}\mathbf{a}_1\right]^{-1}, \tag{6.15}$$

where $\mathbf{R}_{_{s2|tMF}} \triangleq E\left\{(\mathbf{X}\mathbf{w}_{_{tMF}})(\mathbf{X}\mathbf{w}_{_{tMF}})^H\right\}$

$$= \mathbf{A}\mathbf{E}^{1/2}\mathrm{diag}\left(\mathbf{S}^T\mathbf{s}_1\mathbf{s}_1^H\mathbf{S}^*\right)\mathbf{E}^{1/2}\mathbf{A}^H + \sigma^2\mathbf{I}. \tag{6.16}$$

(v) Temporal MVDR – spatial MVDR

$$\mathbf{w}_t = \mathbf{w}_{_{tMV}} \triangleq \mathbf{R}^{-1}_{t1}\mathbf{s}_1/\left(\mathbf{s}_1^H\mathbf{R}^{-1}_{t1}\mathbf{s}_1\right), \tag{6.17}$$

$$\mathbf{w}_s = \mathbf{w}_{_{sMV|tMV}} \triangleq \mathbf{R}^{-1}_{_{s2|tMV}}\mathbf{a}_1/\left(\mathbf{a}_1^H\mathbf{R}^{-1}_{_{s2|tMV}}\mathbf{a}_1\right), \tag{6.18}$$

$$P_{_{sMV|tMV}} = \left[\mathbf{a}_1^H\mathbf{R}^{-1}_{_{s2|tMV}}\mathbf{a}_1\right]^{-1}, \tag{6.19}$$

where $\mathbf{R}_{t1} \triangleq E\left\{\mathbf{X}^H\mathbf{X}\right\}/M = \mathbf{S}^*\mathbf{E}\mathbf{S}^T + \sigma^2\mathbf{I}$,

$\mathbf{R}_{_{s2|tMV}} \triangleq E\left\{(\mathbf{X}\mathbf{w}_{_{tMV}})(\mathbf{X}\mathbf{w}_{_{tMV}})^H\right\}$

$$= \mathbf{A}\mathbf{E}^{1/2}\text{diag}\left(\mathbf{S}^T\mathbf{w}_{tMV}\mathbf{w}_{tMV}^H\mathbf{S}^*\right)\mathbf{E}^{1/2}\mathbf{A}^H$$
$$+ \|\mathbf{w}_{tMV}\|^2\sigma^2\mathbf{I}.$$

(vi) Temporal MVDR – spatial MF

$$\mathbf{w}_t = \mathbf{w}_{tMV}, \quad \mathbf{w}_s = \mathbf{a}_1/M,$$
$$P_{sMF|tMV} = \mathbf{a}_1^H\mathbf{R}_{s2|tMV}\mathbf{a}_1/M^2. \tag{6.20}$$

Let us define the relative performance measure $\mu_1(\Omega) \triangleq P(JMV)/P(\Omega)$, where Ω denotes an arbitrary receiver configuration and JMV denotes the joint space-time MVDR receiver configuration. Then, $0 < \mu_1(\Omega) \leq \mu_1(JMV) = 1$, while high $\mu_1(\Omega)$ indicates low output variance $P(\Omega)$ (or, equivalently, high output SINR). In the following we will use $\mu_1(\Omega)$ to rank the above mentioned configurations in terms of total output SINR.

6.4 Receiver Performance under Limited Sample Support

In practice, the input statistics $\mathbf{R}$ as well as $\mathbf{R}_s$ and $\mathbf{R}_t$ are not known, but estimated from received data vectors, e.g. $\widehat{\mathbf{R}} = \frac{1}{N}\sum_{i=1}^{N}\mathcal{X}_i\mathcal{X}_i^H$, where N denotes the size of the data record. The estimated MVDR filter that utilizes the inverse of a sample-average estimate of the autocovariance matrix is known as the sample-matrix-inversion (SMI) filter. In the following, $\widehat{\mathbf{w}}$ and $\widehat{P}(\cdot)$ refer to an estimate of $\mathbf{w}$ and the output variance of a configuration that utilizes filter estimates, respectively.

Let $\mu_2(\Omega) \triangleq P(\Omega)/\widehat{P}(\Omega)$, and $\mu(\Omega) \triangleq \mu_1(\Omega)\mu_2(\Omega)$. Then, $\mu(\Omega) = P(JMV)/\widehat{P}(\Omega)$ characterizes the output variance of the "estimated" configuration Ω relative to the ideal JMV configuration. Evidently, higher $\mu(\Omega)$ implies higher output SINR exhibited by the estimated configuration.

In [3] it was shown that under the assumption that the received data vectors are independent and identically distributed according to a multivariate complex Gaussian distribution, then, for any given $\zeta \in (0,1)$ and $N \geq ML$,

$$Pr[\mu_2(JMV) > 1-\zeta] \approx 1 - Q\left(\frac{N\zeta - ML + 2}{\sqrt{N\zeta(1-\zeta)}}\right), \tag{6.21}$$

where $Q(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt$ is the complementary Gaussian cumulative distribution function.

In the following theorem, we extend the result in (6.21) for all disjoint domain receiver configurations presented in the previous section.

Theorem 6.1 *For any given $\zeta \in (0,1)$ and $N \geq max(M,L)$,*

$$Pr\left[\mu_2(tMV|sMF) > 1-\zeta\right] \approx 1 - Q\left(\frac{N\zeta - L + 2}{\sqrt{N\zeta(1-\zeta)}}\right) \tag{6.22}$$

$$Pr\left[\mu_2(tMF|sMV) > 1-\zeta\right] \approx 1 - Q\left(\frac{LN\zeta - M + 2}{\sqrt{LN\zeta(1-\zeta)}}\right) \tag{6.23}$$

$$Pr\left[\mu_2(sMV|tMF) > 1-\zeta\right] \approx 1 - Q\left(\frac{N\zeta - M + 2}{\sqrt{N\zeta(1-\zeta)}}\right) \tag{6.24}$$

$$Pr\left[\mu_2(sMF|tMV) > 1-\zeta\right] \approx 1 - Q\left(\frac{MN\zeta - L + 2}{\sqrt{MN\zeta(1-\zeta)}}\right) \tag{6.25}$$

$$Pr\left[\mu_2(tMV|sMV) > 1-\zeta\right] \approx \left[1 - Q\left(\frac{LN\zeta - M + 2}{\sqrt{LN\zeta(1-\zeta)}}\right)\right]$$
$$\left[1 - Q\left(\frac{N\zeta - L + 2}{\sqrt{N\zeta(1-\zeta)}}\right)\right] \tag{6.26}$$

$$Pr\left[\mu_2(sMV|tMV) > 1-\zeta\right] \approx \left[1 - Q\left(\frac{MN\zeta - L + 2}{\sqrt{MN\zeta(1-\zeta)}}\right)\right] \left[1 - Q\left(\frac{N\zeta - M + 2}{\sqrt{N\zeta(1-\zeta)}}\right)\right] \tag{6.27}$$

□

Using Theorem 6.1 we can evaluate the data record size that is necessary to achieve a certain confidence level ϵ by the estimated configuration in a neighborhood ζ of the ideal configuration, which can be expressed by $Pr\left[\mu_2(\Omega) > 1-\zeta\right] > \epsilon$ for any given $\zeta \in (0,1)$ and $\epsilon \in (0,1)$. The results are summarized in Theorem 6.2 below where the last statement (6.33) is due to [3] (the proof is omitted). For convenience in presentation we define the following operation:

$$\ominus \triangleq \begin{cases} -, & \text{if } \epsilon < 0.5 \\ +, & \text{if } \epsilon \geq 0.5. \end{cases} \tag{6.28}$$

Theorem 6.2 *For any given $\zeta \in (0,1)$ and $\epsilon \in (0,1)$, the least number of samples that guarantees $Pr\left[\mu_2 > 1-\zeta\right] > \epsilon$ can be approximated by*

$$N_{tMV|sMF} = max\left\{L, \left\lceil \frac{(1-\zeta)\psi^2 + 2(L-2) \ominus \sqrt{\psi^2(1-\zeta)[4L - 8 + (1-\zeta)\psi^2]}}{2\zeta} \right\rceil\right\} \tag{6.29}$$

$$N_{tMF|sMV} = max\left\{\frac{M}{L}, \left\lceil \frac{(1-\zeta)\psi^2 + 2(M-2) \ominus \sqrt{\psi^2(1-\zeta)[4M - 8 + (1-\zeta)\psi^2]}}{2L\zeta} \right\rceil\right\} \tag{6.30}$$

$$N_{sMV|tMF} = max\left\{M, \left\lceil \frac{(1-\zeta)\psi^2 + 2(M-2) \ominus \sqrt{\psi^2(1-\zeta)[4M - 8 + (1-\zeta)\psi^2]}}{2\zeta} \right\rceil\right\} \tag{6.31}$$

$$N_{sMF|tMV} = max\left\{\frac{L}{M}, \left\lceil \frac{(1-\zeta)\psi^2 + 2(L-2) \ominus \sqrt{\psi^2(1-\zeta)[4L - 8 + (1-\zeta)\psi^2]}}{2M\zeta} \right\rceil\right\} \tag{6.32}$$

$$N_{JMV} = max\left\{ML, \left\lceil \frac{(1-\zeta)\psi^2 + 2(ML-2) \ominus \sqrt{\psi^2(1-\zeta)[4ML - 8 + (1-\zeta)\psi^2]}}{2\zeta} \right\rceil\right\} \tag{6.33}$$

where $\psi \triangleq Q^{-1}(1-\epsilon)$, $\lceil x \rceil$ denotes the smallest integer larger than or equal to x. □

We note that

$$\begin{aligned} & Pr\left[\mu(\Omega) > 1-\zeta\right] = Pr\left[\mu_2(\Omega) > \frac{1-\zeta}{\mu_1(\Omega)}\right] \\ = \quad & Pr\left[\mu_2(\Omega) > 1 - \frac{\mu_1(\Omega) + \zeta - 1}{\mu_1(\Omega)}\right]. \end{aligned} \tag{6.34}$$

Using Theorem 6.2 we can evaluate N'_Ω such that $Pr\left[\mu(\Omega) > 1-\zeta\right] > \epsilon$, i.e. evaluate the data record size N'_Ω that is necessary to achieve a certain performance confidence level ϵ by the estimated configuration in the neighborhood ζ of the optimum (JMV) configuration. This can be accomplished by substituting $\zeta'(\Omega) = 1 - \frac{1-\zeta}{\mu_1(\Omega)}$ into eqs. (6.29)–(6.32).

6.5 Simulation Results

We consider a synchronous DS-CDMA system with $K = 9$ users and processing gain $L = 15$. The temporal signature cross-correlations are in the range of [0.07, 0.33]. We consider *user-1* as the user of interest. We assume that the AoAs of all users are equally spaced in $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and $\theta_1 = 0$. We consider two scenarios with respect to different power profile.

Scenario (i): Perfect power control
All users have the same input SNR fixed at 13dB, including the antenna array gain evaluated as $10\log(M)$. The results are tabulated in Table 6.1 and Table 6.2.

Table 6.1 indicates that for a given L, if M is much smaller than L, then the tMV|sMF configuration outperforms all other disjoint configurations with high probability. On the other hand, when M is much larger than L then the sMV|tMF configuration performs the best among all disjoint configurations with high probability. Similar performance characteristics were also observed in [5].

Table 6.1. Output SINR performance under ideal statistics

configuration	$M = 2$	$M = 5$	$M = 10$	$M = 18$
JMV	15.10	19.65	19.82	19.94
tMV|sMF	10.12	15.85	17.81	19.74
tMF|sMV	3.82	14.67	18.81	19.86
tMV|sMV	10.12	15.58	18.84	19.86
sMV|tMF	3.82	15.80	18.91	19.88
sMF|tMV	6.78	13.28	16.24	18.56
sMV|tMV	6.79	13.34	16.70	18.62
JMF	3.82	15.21	17.50	19.74

Table 6.2 provides the minimum number of data samples needed by each configuration to achieve, with confidence level ϵ, receiver output SINR in a neighborhood ζ of the optimum JMV performance, for different values of ϵ and ζ. An empty entry indicates that the corresponding configuration can never reach, with confidence ϵ, an SINR performance level that lies within distance ζ from the JMV performance. We observe that when a disjoint configuration *can* achieve output SINR with confidence level ϵ in a neighborhood ζ of the JMV performance, it requires much less data support than the JMV configuration does.

Table 6.2. Minimum number of data samples

N'	$M = 10$		$M = 18$	
	ζ=0.3 ϵ=0.9	ζ=0.1 ϵ=0.9	ζ=0.3 ϵ=0.9	ζ=0.1 ϵ=0.9
JMV	539	1636	954	2887
tMV|sMF	81	—	60	201
tMF|sMV	3	16	5	15
sMV|tMF	45	219	71	224
sMF|tMV	9	—	5	37

Scenario (ii): Unequal power users
We use the same setup as in Scenario (i) except that now the SNR of users 6-9 is fixed at 23dB (while the SNR of users 1-5 is kept at 13dB as before). In other words, users 6-9 have a 10dB power advantage relative to the user of interest. The simulation results are tabulated in Table 6.3 and Table 6.4.

Table 6.3. Output SINR performance under ideal statistics

configuration	$M = 2$	$M = 5$	$M = 10$	$M = 18$
JMV	14.99	19.64	19.82	19.94
tMV|sMF	8.91	13.06	15.31	18.94
tMF|sMV	1.20	9.43	18.51	19.83
tMV|sMV	6.32	11.58	18.53	19.83
sMV|tMF	1.25	9.98	18.59	19.85
sMF|tMV	5.68	10.40	12.93	15.20
sMV|tMV	5.83	10.57	13.39	15.26
JMF	0.87	8.00	11.03	18.81

Table 6.4. Minimum number of data samples

N'	$M = 10$		$M = 18$	
	ζ=0.3 ϵ=0.9	ζ=0.1 ϵ=0.9	ζ=0.3 ϵ=0.9	ζ=0.1 ϵ=0.9
JMV	539	1636	954	2887
tMV|sMF	194	—	67	349
tMF|sMV	4	23	5	16
sMV|tMF	47	308	71	227
sMF|tMV	—	—	15	—

Bibliography

[1] S. Verdú, *Multiuser Detection*. Cambridge, U. K.: Cambridge Univ. Press, 1998.

[2] D. A. Pados and S. N. Batalama, "Joint space-time auxiliary-vector filtering for DS-CDMA systems with antenna arrays," *IEEE Trans. Commun.*, vol. 47, pp. 1406-1415, Sept. 1999.

[3] I. N. Psaromiligkos and S. N. Batalama, "Data record size requirements for adaptive space-time DS/CDMA signal detection and direction-of-arrival estimation," *IEEE Trans. Commun.*, accepted for publication, pending revision.

[4] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. New York: Dover, 1974.

[5] P. Xiong, M. J. Medley and S. N. Batalama, "Spatial and temporal processing for global navigation satellite systems: The GPS receiver paradigm," *IEEE Trans. Aerosp. Electr. Syst.*, vol. 39, pp. 1471-1484, Oct. 2003.

Chapter 7

Near-ML Multiuser Detection with Linear Filters and Reliability-based Processing

7.1 Introduction

In direct-sequence code-division-multiple-access (DS-CDMA) systems, multiple users transmit information bit sequences over a common channel utilizing distinct individually assigned signature waveforms. In the presence of strong interferers and/or high cross-correlation signature waveforms, the performance of the conventional single-user signature matched-filter detector degrades significantly. On the other hand, if the user signatures, energies, and the channel noise power are known at the receiver end, use of the maximum-likelihood (ML) multiuser detector [1] guarantees minimum probability of error system performance. ML optimum multiuser detection, however, comes with computational complexity that grows exponentially with the number of users.

In search of a satisfactory trade-off between performance and complexity, numerous sub-optimum multiuser detectors have been proposed. Two linear multiuser detectors that are of particular interest to our work are the familiar zero-forcing decorrelating detector [2] and the minimum-mean-square-error (MMSE) detector [3]-[6]. Both are known to be near-far resistant [7]. The decorrelator is the maximum likelihood estimator of the energy modulated user bits for additive white Gaussian noise (AWGN) multiple-access channels. The MMSE receiver is the linear filter solution that maximizes the pre-detection signal-to-interference-plus-noise ratio (SINR) of each user.

Recently, there has been an effort to directly approximate the ML multiuser detector in an efficient and effective manner. Examples of such work are the "greedy" detector in [8] which partitions the likelihood metric function and proceeds with "part-by-part" maximization and the neighbor searching mechanisms in [9, 10]. The "MK-face" detector in [9] searches within a pre-calculated set of neighboring candidates. The gradient search algorithm in [10] happens to coincide with the "K-face" detector of [9].

As a combinatorial optimization problem, ML multiuser detection has certain similarities with ML decoding of binary linear block codes. One of the earliest, well known low-complexity algorithms for sub-ML block-code decoding is Chase's algorithm [11]. The Chase procedure selects a small fixed number of bit decision perturbations based on reliability measurements which are supposed to be part of the demodulator output (usually directly related to the sufficient decision statistic). In this chapter, we follow an analogous approach to tackle the ML multiuser detection problem. First, we use a linear multiuser detector (such as the decorrelator or MMSE detector) to obtain an initial multiuser bit decision. The soft output vector, before being discarded, is processed to derive the corresponding reliability measurement vector. Next, as dictated by the reliability vector, a sequence of "error patterns" is generated to "correct" the initial decision vector and produce multiuser bit combination candidates that include the maximum likelihood solution with high probability. Final decision is made by simply choosing the bit combination with the highest likelihood among all candidates. When the length of the error pattern sequence is in the order of the number of users,

the proposed algorithm exhibits near-ML performance at remarkably reduced computational complexity compared to the ML-optimum multiuser detector.[1]

We organize the rest of the chapter as follows. In Section 7.2, we present our signal model and notation. The near-ML multiuser detection algorithm is developed in Section 7.3. Simulation studies are presented in Section 7.4 and a few final conclusions are drawn in Section 7.5.

7.2 Signal Model and Notation

For the sake of clarity in presentation, we consider a synchronous DS/CDMA system where K users share a multiple-access AWGN channel with power spectral density σ^2. Our developments can be generalized to cover near-ML block multiuser detection for asynchronous systems by creating error pattern sequences of length proportional to the total number of bits in a block. The general block diagram representation of the receiver that we develop in this work is given in Fig. 7.1. The received signal is processed by a front end that consists of a block of K signature-matched filters. We denote the ith bit period output of the matched filter bank by $\mathbf{y}(i) \triangleq [y_1(i),\ \ y_2(i), \cdots, y_K(i)]^T$ (T denotes the transpose operation) and we write

$$\mathbf{y}(i) = \mathbf{R}\mathbf{E}(i)\mathbf{b}(i) + \mathbf{n}(i), \qquad i = 1,\ 2,\ \ldots\ , \tag{7.1}$$

where $\mathbf{b}(i) \triangleq [b_1(i), b_2(i), \cdots, b_K(i)]^T \in \{\pm 1\}^K$ is the ith period information bit vector, $\mathbf{E}(i) \triangleq \text{diag}\left(\left[\sqrt{E_1(i)},\ \sqrt{E_2(i)}, \cdots, \sqrt{E_K(i)}\right]\right)$ is the $K \times K$ user signal amplitude matrix, $\mathbf{R}$ is the $K \times K$ signature cross-correlation matrix and $\mathbf{n}(i) \triangleq [n_1(i), n_2(i), \cdots, n_K(i)]^T$ is the filtered noise vector. The ML optimum multiuser detector [1] is

$$\hat{\mathbf{b}}_{\text{ML}} = \arg \max_{\mathbf{b} \in \{\pm 1\}^K} \left(\Omega(\mathbf{b}, \mathbf{y}) \triangleq 2\mathbf{b}^T\mathbf{E}\mathbf{y} - \mathbf{b}^T\mathbf{E}\mathbf{R}\mathbf{E}\mathbf{b}\right) \tag{7.2}$$

where for notational simplicity we dropped the index i with the understanding that the decision rule in (7.2) refers to a given information bit interval.

7.3 Near-ML Multiuser Detection

We switch now our attention to the linear operator $\mathbf{L}$ and the "Search & Decision" blocks in Fig. 7.1. The linear operator can be viewed as a multiuser detector that operates on the output of the matched-filter bank $\mathbf{y}$ to produce an initial decision $\hat{\mathbf{b}}^{(0)}$ for the transmitted bit vector $\mathbf{b}$. Alongside the initial decision, we also extract a "reliability" measurement vector $\boldsymbol{\alpha}$. During the final "Search & Decision" stage, an ordered sequence of D distinct error patterns $\mathbf{e}^{(1)}$, $\mathbf{e}^{(2)}$, ..., $\mathbf{e}^{(D)} \in \{0,1\}^K$ is generated from the reliability vector $\boldsymbol{\alpha}$ where the positive integer D is a system parameter to be adjusted according to our performance versus complexity requirements. Corresponding to each error pattern $\mathbf{e}^{(d)}$, $d = 1,\ 2,\ \ldots,\ D$, a new tentative decision $\hat{\mathbf{b}}^{(d)}$ is produced. The final decision output becomes

$$\hat{\mathbf{b}}_{\text{Near-ML}} = \arg \max_{\mathbf{b} \in \{\hat{\mathbf{b}}^{(1)},\ \hat{\mathbf{b}}^{(2)},\ \ldots,\ \hat{\mathbf{b}}^{(D)}\}} \Omega(\mathbf{b}, \mathbf{y}). \tag{7.3}$$

As we will see later in Section 7.4, when $\mathbf{L}$ is the decorrelating or MMSE operator, a small D (for instance of the same order as the number of users/bits to be detected) results in bit-error-rate system performance that is nearly indistinguishable from the maximum likelihood detector in operational environments of practical interest.

7.3.1 Initial Decision and Reliability Measurement

We use a linear operator (multiuser detector) $\mathbf{L}_{K \times K}$ to obtain the initial bit decision vector $\hat{\mathbf{b}}^{(0)} = \text{sgn}(\mathbf{L}\mathbf{y})$. To exploit the reliability information contained in the linear operator/filter output, we define for each initial

[1] Another coding theory inspired multiuser detector was presented recently in [12]. In contrast to the ML approximations in [8]-[10] and the work herein, the multiuser detector of [12] is a complete exponential-complexity depth-first tree search implementation of the full ML decision rule based on the lattice decoder developed in [13].

bit decision $\hat{b}_j^{(0)}$ the corresponding soft output

$$\tilde{\alpha_j} \triangleq \boldsymbol{l}_j^T \mathbf{y}, \quad j = 1,\ 2,\ \ldots,\ K, \tag{7.4}$$

where $\boldsymbol{l}_j^T$ is the jth row of the matrix $\mathbf{L}$. For convenience, we also define $\mathbf{w}_j^T \triangleq \boldsymbol{l}_j^T \mathbf{RE}$ and substitute (7.1) in (7.4) to calculate

$$\tilde{\alpha_j} = \mathbf{w}_j^T \mathbf{b} + \boldsymbol{l}_j^T \mathbf{n} = w_{jj} b_j + \sum_{i=1, i\neq j}^{K} w_{ji} b_i + \boldsymbol{l}_j^T \mathbf{n}, \quad j = 1,\ 2,\ \ldots,\ K. \tag{7.5}$$

The first term on the right-hand side of (7.5) is the scaled jth transmitted bit of interest. The last term, $\boldsymbol{l}_j^T \mathbf{n} \triangleq n_j$, is zero-mean filtered Gaussian noise with variance $E\{n_j^2\} = \sigma^2 \boldsymbol{l}_j^T \mathbf{R} \boldsymbol{l}_j \triangleq \sigma_{n_j}^2$. The middle term $\sum_{i=1, i\neq j}^{K} w_{ji} b_i \triangleq z_j$ contains the information bits transmitted by the other users (multiple-access-interference or MAI) and is a zero-mean random variable with variance $E\{z_j^2\} = \sum_{i=1, i\neq j}^{K} w_{ji}^2 = \|\mathbf{w}_j\|^2 - w_{jj}^2 \triangleq \sigma_{z_j}^2$. We define the *reliability value* α_j of the initial bit decision $\hat{b}_j^{(0)}$ as the log-likelihood ratio $\alpha_j \triangleq \frac{1}{2} \ln \frac{f\left(\tilde{\alpha_j} | b_j = \hat{b}_j^{(0)}\right)}{f\left(\tilde{\alpha_j} | b_j = -\hat{b}_j^{(0)}\right)}$ where $f\left(\tilde{\alpha_j} | b_j\right)$ denotes the probability density function of $\tilde{\alpha_j}$ conditioned on b_j. If the filtered MAI variance $E\{z_j^2\}$ is relatively small, the soft output $\tilde{\alpha_j}$ is well approximated [14, 15] by a Gaussian random variable with mean $E\left\{\tilde{\alpha_j}\right\} = w_{jj} b_j$ and variance $E\left\{\tilde{\alpha_j}^2\right\} = \sigma_{z_j}^2 + \sigma_{n_j}^2$. Then, the reliability value α_j simplifies to a scaled version of $|\tilde{\alpha_j}|$:

$$\alpha_j \triangleq \frac{1}{2} \ln \frac{f\left(\tilde{\alpha_j} | b_j = \hat{b}_j^{(0)}\right)}{f\left(\tilde{\alpha_j} | b_j = -\hat{b}_j^{(0)}\right)} \cong \frac{\tilde{\alpha_j} w_{jj} \hat{b}_j^{(0)}}{\sigma_{z_j}^2 + \sigma_{n_j}^2} = \frac{w_{jj}}{\sigma_{z_j}^2 + \sigma_{n_j}^2} \left|\tilde{\alpha_j}\right|. \tag{7.6}$$

The generated reliability vector $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_K]^T$ is passed to the "Search & Decision" stage along with the hard-limited initial "guess" $\hat{\mathbf{b}}^{(0)}$.

7.3.2 Generation of the Error Pattern Sequence

When $\mathbf{L}$ is the decorrelating or MMSE filter operator, we expect intuitively that a small perturbation of the initial vector $\hat{\mathbf{b}}^{(0)}$ may produce the maximum likelihood decision $\hat{\mathbf{b}}_{\mathrm{ML}}$. With this motivation, we generate a sequence of error patterns $\mathbf{e}^{(d)} \in \{0,1\}^K$, $d = 1,\ 2,\ \ldots,$ where $e_j^{(d)} = 1$ stands for an error in bit position j. We apply the error sequence $\mathbf{e}^{(d)}$ on the initial bit vector $\hat{\mathbf{b}}^{(0)}$ to create the bit vector sequence $\hat{\mathbf{b}}^{(d)}$ as follows:

$$\hat{b}_j^{(d)} = \hat{b}_j^{(0)} \oplus e_j^{(d)}, \quad j = 1,\ 2,\ \ldots,\ K, \tag{7.7}$$

where $\oplus$ denotes the error correction operation $b \oplus e = b \cdot (-1)^e$, $b \in \{\pm 1\}$, $e \in \{0,1\}$. Naturally, the error sequence generation criterion will be the likelihood of $\hat{\mathbf{b}} = \hat{\mathbf{b}}^{(0)} \oplus \mathbf{e}$ with observation the soft-output vector $\tilde{\boldsymbol{\alpha}} = \mathbf{Ly}$, $f\left(\tilde{\boldsymbol{\alpha}} | \mathbf{b} = \hat{\mathbf{b}}\right)$. If, approximately, we treat the soft outputs $\tilde{\alpha_j}$ as conditionally independent random variables, the likelihood function becomes $\prod_{j=1}^{K} f\left(\tilde{\alpha_j} | b_j = \hat{b}_j\right)$ and the log-likelihood ratio is

$$\Lambda(\hat{\mathbf{b}}) \cong \frac{1}{2} \sum_{j=1}^{K} \ln \frac{f\left(\tilde{\alpha_j} | b_j = \hat{b}_j\right)}{f\left(\tilde{\alpha_j} | b_j = -\hat{b}_j\right)} = \sum_{j:\hat{b}_j = \hat{b}_j^{(0)}} \alpha_j - \sum_{j:\hat{b}_j \neq \hat{b}_j^{(0)}} \alpha_j = \sum_{j=1}^{K} \alpha_j - 2 \sum_{j:e_j=1} \alpha_j. \tag{7.8}$$

We immediately observe that the first term in (7.8), $\sum_{j=1}^{K} \alpha_j$, is independent of the choice of the error pattern $\mathbf{e}$. Therefore, the error pattern sequence should be designed in ascending order of the other term $\phi(\mathbf{e}, \boldsymbol{\alpha}) \triangleq \sum_{j:e_j=1} \alpha_j$. For our purposes of near-ML multiuser detection, if all 2^K error patterns $\left\{\mathbf{e}^{(d)}\right\}_{j=1}^{2^K}$ are in ascending order under the $\phi(\cdot, \boldsymbol{\alpha})$ key

$$\phi\left(\mathbf{e}^{(d)}, \boldsymbol{\alpha}\right) \leq \phi\left(\mathbf{e}^{(d+1)}, \boldsymbol{\alpha}\right), \quad d = 1,\ 2,\ \ldots,\ 2^K - 1, \tag{7.9}$$

we will consider and process only the first D elements of the sequence.

We now derive an algorithm that produces the first D $\phi(\cdot,\cdot)$-ordered error patterns with computational complexity linear in the number of bits, $O(DK)$. The origins of our algorithm are in the "subset-sum problem" where the objective is to find all subsets of a set of numbers that have sum of elements less than or equal to a given number [16, 17]. Suppose that the reliability values α_j, $j = 1,\ 2,\ \ldots,\ K$, are such that $\alpha_{h_1} \leq \alpha_{h_2} \leq \cdots \leq \alpha_{h_K}$. Define the error pattern sequences $\mathcal{E}_p \triangleq \left\{\mathbf{e}^{(p_t)}\right\}$, $p = 0,\ 1,\ 2,\ \ldots,\ K$, $t = 1,\ 2,\ \ldots,\ 2^p$, where $\left\{\mathbf{e}^{(p_t)}\right\}$ is ordered under the key $\phi(\cdot,\boldsymbol{\alpha})$

$$\phi\left(\mathbf{e}^{(p_t)},\boldsymbol{\alpha}\right) \leq \phi\left(\mathbf{e}^{(p_{t+1})},\boldsymbol{\alpha}\right), \qquad t = 1,\ 2,\ \ldots,\ 2^p - 1, \tag{7.10}$$

and satisfies

$$e_h^{(p_t)} = 0 \quad \text{for every } \ h \in \{h_{p+1},\ h_{p+2},\ \cdots,\ h_K\} \ \text{ and } \ 1 \leq t \leq 2^p. \tag{7.11}$$

Here, $e_h^{(\cdot)}$ is the hth bit in the error pattern $\mathbf{e}^{(\cdot)}$ and (7.11) states that $\mathbf{e}^{(p_t)}$ may include errors only in the p lowest reliability positions of the initial bit decision vector $\hat{\mathbf{b}}^{(0)}$. Define also the error pattern sequences $\mathcal{E}'_p \triangleq \left\{\mathbf{e}'^{(p_t)}\right\}$, $p = 0,\ 1,\ 2,\ \ldots,\ K-1$, $t = 1,\ 2,\ \ldots,\ 2^p$, where $\left\{\mathbf{e}'^{(p_t)}\right\}$ is such that

$$e'^{(p_t)}_h = \begin{cases} 1, & \text{if } h = h_{p+1} \\ e_h^{(p_t)}, & \text{otherwise.} \end{cases} \tag{7.12}$$

Therefore, each $\mathbf{e}'^{(p_t)}$ contains *exactly* one more error in bit position h_{p+1} than the corresponding error pattern $\mathbf{e}^{(p_t)}$. It follows that $\phi\left(\mathbf{e}'^{(p_t)},\boldsymbol{\alpha}\right) = \phi\left(\mathbf{e}^{(p_t)},\boldsymbol{\alpha}\right) + \alpha_{h_{p+1}}$ and $\left\{\mathbf{e}'^{(p_t)}\right\}$ is also ordered under the key $\phi(\cdot,\boldsymbol{\alpha})$.

At this point, it is important to observe the following. If we merge the two ordered sequences $\mathcal{E}_p = \{\mathbf{e}^{(\cdot)}\}$ and $\mathcal{E}'_p = \{\mathbf{e}'^{(\cdot)}\}$ in increasing order of the same key $\phi\left(\cdot,\boldsymbol{\alpha}\right)$, then we will obtain $\mathcal{E}_{p+1}$. Therefore, iterative generation of $\mathcal{E}'_p$ from $\mathcal{E}_p$ and merging to form $\mathcal{E}_{p+1}$ for $p = 0,\ 1,\ \ldots,\ K-1$ leads to $\mathcal{E}_K$ where $\mathcal{E}_K$ is the whole set of 2^K error patterns ordered under the key $\phi(\cdot,\boldsymbol{\alpha})$. Since only the first D error patterns in $\mathcal{E}_K$ are needed, we can simplify this process greatly. We observe that the first D elements in $\mathcal{E}_K$ come from the first D elements of $\mathcal{E}_{K-1}$ and the first D elements[2] of $\mathcal{E}'_{K-1}$. Therefore, we can truncate the length of $\mathcal{E}_{K-1}$ to D (which also limits the length of $\mathcal{E}'_{K-1}$ to D). By induction, we conclude that every sequence $\mathcal{E}_0$, $\mathcal{E}_1$, $\ldots$, $\mathcal{E}_K$ can be shortened to length D. Considering the involved computational complexity, we need up to D additions to calculate $\phi(\cdot,\boldsymbol{\alpha})$ and another D comparisons to merge $\mathcal{E}_p$ and $\mathcal{E}'_p$ in each step. Therefore, the computational complexity is $O(DK)$.

In summary, the proposed multiuser detection system in Fig. 7.1 operates as follows:

(i) Apply the linear operator/filter $\mathbf{L}$ to the matched-filter bank output $\mathbf{y}$ to obtain the initial decision $\hat{\mathbf{b}}^{(0)}$ and the reliability measures $\{\alpha_j\}$ in (7.6).

(ii) Sort $\{\alpha_j\}$ in increasing order.

(iii) Initialize $p \leftarrow 0$, $\mathcal{E}_0 = \left\{\mathbf{e}^{(1)} = [0\ 0\ \cdots\ 0]\right\}$, $\phi\left(\mathbf{e}^{(1)},\boldsymbol{\alpha}\right) = 0$.

(iv) Create $\mathcal{E}'_p = \left\{\mathbf{e}'^{(i)} = \mathbf{e}^{(i)} + [\underbrace{0\ \cdots\ 0}_{h_{p+1}-1}\ 1\ \underbrace{0\ \cdots\ 0}_{K-h_{p+1}}] : \mathbf{e}^{(i)} \in \mathcal{E}_p\right\}$. Calculate $\phi\left(\mathbf{e}'^{(i)},\boldsymbol{\alpha}\right) = \phi\left(\mathbf{e}^{(i)},\boldsymbol{\alpha}\right) + \alpha_{h_{p+1}}$.

(v) Merge $\mathcal{E}_p$ and $\mathcal{E}'_p$ to form $\mathcal{E}_{p+1}$ in ascending order under key $\phi(\cdot,\boldsymbol{\alpha})$. Truncate the length of $\mathcal{E}_{p+1}$ to D if longer than D.

(vi) If $p < K-1$, then set $p \leftarrow p+1$ and go back to Step 4. Else go to Step 7.

(vii) For $d = 1,\ 2,\ \ldots,\ D$ apply $\mathbf{e}^{(d)}$ to the initial decision $\hat{\mathbf{b}}^{(0)}$ as in (7.7). Produce the final decision $\hat{\mathbf{b}}_{\text{Near-ML}}$ by (7.3).

[2] In fact, only the first $\lfloor \frac{D}{2} \rfloor$ elements of $\mathcal{E}'_{K-1}$ can contribute to the first D elements of $\mathcal{E}_K$ but this does not affect our conclusion on the rough order of the involved computational complexity and may complicate our presentation unnecessarily.

The overall system computational complexity can be calculated as follows. The complexity of applying the linear operator $\mathbf{L}$ to the matched-filter output $\mathbf{y}$ is $O\left(K^2\right)$. Sorting the reliability values costs $O\left(K\log_2 K\right)$. As stated earlier, the complexity of generating the ordered sequence $\left\{\mathbf{e}^{(d)}\right\}_{d=1}^{D}$ is $O(DK)$. Finally, $\Omega(\cdot,\cdot)$ is calculated D times with complexity, again, $O(DK)$. Therefore, the complexity of the whole scheme is approximately $O\Big(K(K+2D+\log_2 K)\Big)$ per bit period. The cost of the "Near-ML Search & Decide" block in Fig. 7.1 is barely more than twice the cost of the $\mathbf{Ly}$ operation if D is chosen equal to K (that is, if D is equal to the number of user bits to be detected). As we will see later in Section 7.4, when the linear operator $\mathbf{L}$ is chosen appropriately a value of D equal to K or -more general- a few times K can be sufficient for the algorithm to maintain near-ML performance for all practical purposes.

7.3.3 Selection of the Linear Operator

So far, we have intentionally treated the linear operator $\mathbf{L}$ as arbitrary. Yet, it is important to choose an operator/filter that produces initial decisions with low probability of error such that only a small perturbation is needed to approach the ML optimum decision vector. Conversely, if the initial decisions are poor, then to achieve the same bit-error-rate performance level we will need a larger search length parameter D and that will thereby increase the computational complexity.

Besides the initial performance considerations, the condition that the variance of the MAI term $z_j = \sum_{i=1,i\neq j}^{K} w_{ji}b_i$ in (7.5) is relatively small should also be fulfilled. This directly affects the accuracy of the Gaussian approximation of the soft output $\tilde{\alpha_j}$ which in turn affects the quality of the reliability measure α_j in (7.6). There are two well-researched linear filters that can satisfy the quality of initial decision and small MAI variance requirements: the decorrelating and MMSE filters. The decorrelating filter $\mathbf{L}_{\rm DEC} = \mathbf{R}^{-1}$, leading to zero errors in the absence of noise, has perfect near-far resistance and is the maximum likelihood solution for real-valued "symbol" vectors. If we use $\mathbf{L}_{\rm DEC} = \mathbf{R}^{-1}$ in (7.4), (7.5), we obtain $w_{jj} = \mathbf{E}_{jj}$, $\sigma_{n_j}^2 = \sigma^2(\mathbf{R}^{-1})_{jj}$, and $\sigma_{z_j}^2 = 0$. Because MAI is completely cancelled, the Gaussian approximation is not needed anymore. The reliability measures in (7.6) take the form

$$\alpha_j = \frac{\mathbf{E}_{jj}}{\sigma^2\left(\mathbf{R}^{-1}\right)_{jj}}|\tilde{\alpha_j}|, \qquad j = 1,\ 2,\ \ldots,\ K, \tag{7.13}$$

with *no approximation*. The MMSE multiuser filter $\mathbf{L}_{\rm MMSE} = (\mathbf{R} + \sigma^2\mathbf{E}^{-2})^{-1}$ minimizes the mean norm square error between its output and the true transmitted information bit vector and maximizes the output SINR per user. However, the filtered MAI term in (7.5) is not zero: $z_j = \sum_{i=1,i\neq j}^{K} w_{ji}b_i \neq 0$. Still, the MMSE filtered MAI is approximately Gaussian [14],[15] and the reliability measure in (7.6) maintains its accuracy.

As a final comment, it is clear that technically the soft outputs $\tilde{\alpha_j}$, $j = 1,\ 2,\ \ldots,\ K$, are not independent and their degree of dependency can affect the quality of the log-likelihood ratio approximation in (7.8) and, in turn, the quality of the generated sequence of error patterns. When the linear operator $\mathbf{L}$ is the decorrelator, it is easy to calculate the correlation coefficient ρ_{ij} between $\tilde{\alpha_i}$ and $\tilde{\alpha_j}$, $i \neq j$: $\rho_{ij} = \frac{E\{\tilde{\alpha_i}\tilde{\alpha_j}\}}{\sqrt{E\{\tilde{\alpha_i}^2\}E\{\tilde{\alpha_j}^2\}}} = \frac{\mathbf{L}_{ij}}{\sqrt{(\mathrm{SNR}_i+\mathbf{L}_{ii})(\mathrm{SNR}_j+\mathbf{L}_{jj})}}$ where $\mathrm{SNR}_i \triangleq \frac{E_i}{\sigma^2}$. In most cases of practical interest, $|\rho_{ij}|$ is less than 10^{-2} and the independence approximation is well justified. In such cases, we found (see for example Section 7.4) that $D = K$ is sufficient to approximate very closely the performance of the true ML detector. In extreme situations, such as pre-detection SNR values less than 0dB and high signature cross-correlations (greater than 0.5), $|\rho_{ij}|$ can be greater than 0.1 and thus non-negligible. In these cases, a greater D should be used to approach ML performance. For MMSE preprocessors, we can calculate $|\rho_{ij}|$ numerically and, again, adjust D accordingly.

7.4 Simulation Studies

We consider the DS/CDMA signal model of Section 7.2 with spreading gain 31 and $K = 15$ users with Gold signature code assignments. In Fig. 7.2, we plot the bit-error-rate (BER) of the user of interest (user 1) under different multiuser detection schemes as a function of its SNR over the $4 - 16$dB range. The SNRs

of the 14 interferers are distributed in the $8.5 - 15$dB range with a 0.5dB increment. We use the multiuser MMSE filter as the pre-processor in our proposed system and we fix the length of the error pattern sequence to $D = K = 15$ (hence, the computational complexity per bit interval is $O(DK) = O(K^2)$). For purposes of comparison, we include *(i)* the ML optimum multiuser detector [1] (which serves as a BER lower bound), *(ii)* the stand-alone multiuser MMSE detector[3] [3]-[6], *(iii)* the "greedy" multiuser detector in [8] with different memory parameter settings $L = 1, 2, 3$ and computational complexity per bit interval $O(LK^2)$, *(iv)* the "MK-face" detector [9] with parameter setting $MK = 2 \cdot 15 = 30$, computational complexity per bit interval $O(MK^2)$ and space complexity $O(2^K MK)$, and *(v)* the "gradient search" detector [10] that coincides with the "K-face" algorithm of [9] and has computational complexity $O(K^2)$ per iteration. We observe that in the relatively low SNR region $(4 - 12$dB) the greedy, the MK-face, and the proposed detector have performance very close to the optimum. We recall that the greedy algorithm decides the value of the user information bits sequentially and the order favors high energy bits/users. As the SNR of the "user of interest" increases relative to the other users, bit decisions are made early with the adverse effect of an increase in BER as seen in Fig. 7.2. For the MK-face detector, the pre-calculated neighboring candidates fail to perturb high SNR user bits and the performance curve degrades to the pre-detector (MMSE in this case). The proposed system maintains "near-ML" performance throughout the range of this study. In Fig. 7.3, we repeat the study of Fig. 7.2 and we examine the BER of the proposed detector as a function of the length of the error pattern sequence D at $\text{SNR}_1 = 12$dB and 14dB. The proposed algorithm approaches effectively the ML bound for modest values of D (the number of bits to be detected is $K = 15$; $D = 1$ corresponds to the stand-alone MMSE detector and $D = 2^{15}$ corresponds to the true ML detector).

Finally, in Fig. 7.4 we increase the number of active users from $K = 15$ to $K = 25$. The SNRs of the 24 interferers are distributed over the $8.5 - 20$dB range with a 0.5dB increment. The performance gain of the proposed "near-ML" algorithm with $D = 4K = 100$ over the MMSE detector is remarkable. True ML detection has complexity proportional to 2^{25} and we cannot afford this computational cost at present. For this reason, we replace the true ML bound that appears in the previous Figs. 7.2 and 7.3 by the *single-user bound*. We also cannot afford execution of the "MK-face" algorithm that requires exponential $(2^{25}MK)$ storage space. Instead, we present in Fig. 7.4 the performance curves of the "greedy" $(L = 4)$ and "gradient search" ("K-face") algorithms [8]-[10].

7.5 Conclusions

We described a new multiuser detection algorithm. A decorrelating or MMSE multiuser filter is used as a pre-processor that provides initial decisions and reliability measurements based on which an ordered error pattern sequence of variable length is formed. The error pattern sequence is followed to its end and the most likely bit vector among all visited options is returned. When the length of the pattern sequence is of the order of the number of bits to be detected, the additional imposed computational cost compared to straight decorrelating or MMSE detection is rather insignificant. Still, in extensive simulation studies for both synchronous and asynchronous links (not reported herein) we saw that the proposed multiuser detection algorithm is able to maintain near-ML bit-error-rate performance over the whole studied SNR range of interest. There is strong resemblance between this scheme and "efficient" decoding algorithms for binary linear block codes [11, 17].

[3] In this study we choose a Gold code assignment that makes the normalized signature cross-correlation between user 1 and the other users equal to $\frac{7}{31}$ or $-\frac{9}{31}$. Notice the remarkable performance difference between the MMSE and the ML detector.

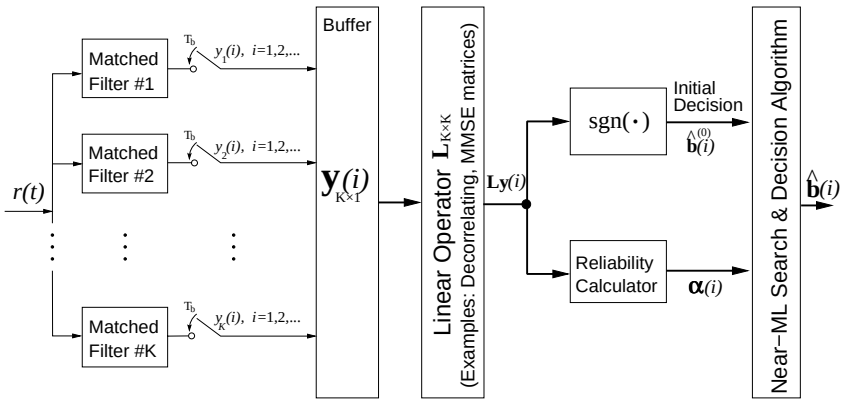

Fig. 7.1. The proposed "near-ML" multiuser detection architecture.

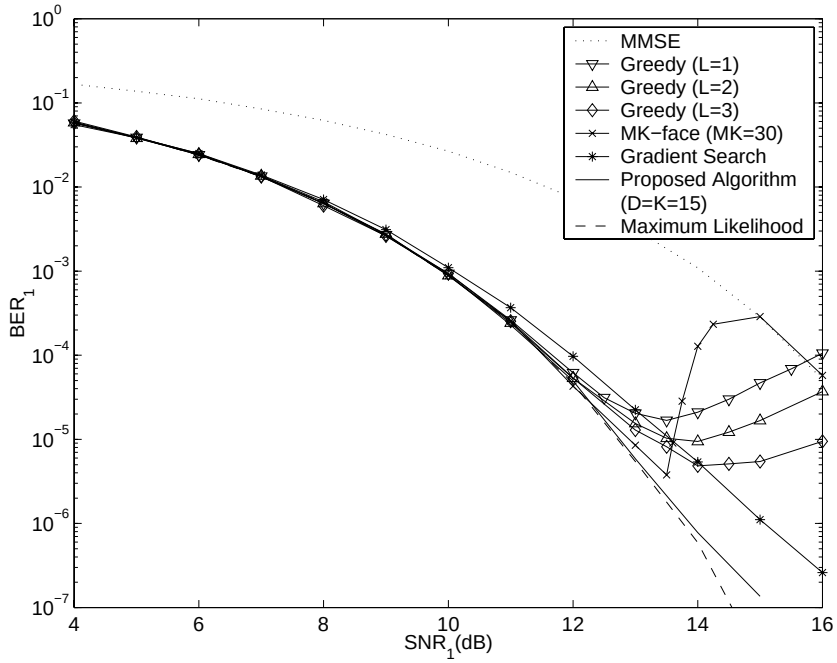

Fig. 7.2. Bit-error-rate versus SNR for various multiuser detectors ($K = 15$, SNR_i in 8.5 to 15dB range, $i = 2, 3, \cdots, 15$).

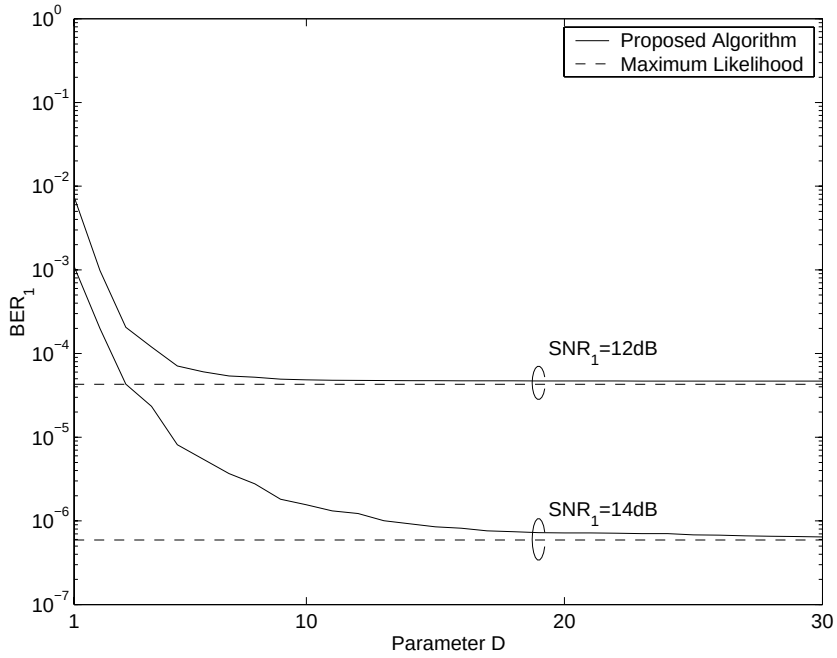

Fig. 7.3. Bit-error-rate versus length of the error pattern sequence D for the system of Fig. 7.2 (detection of $K = 15$ user bits).

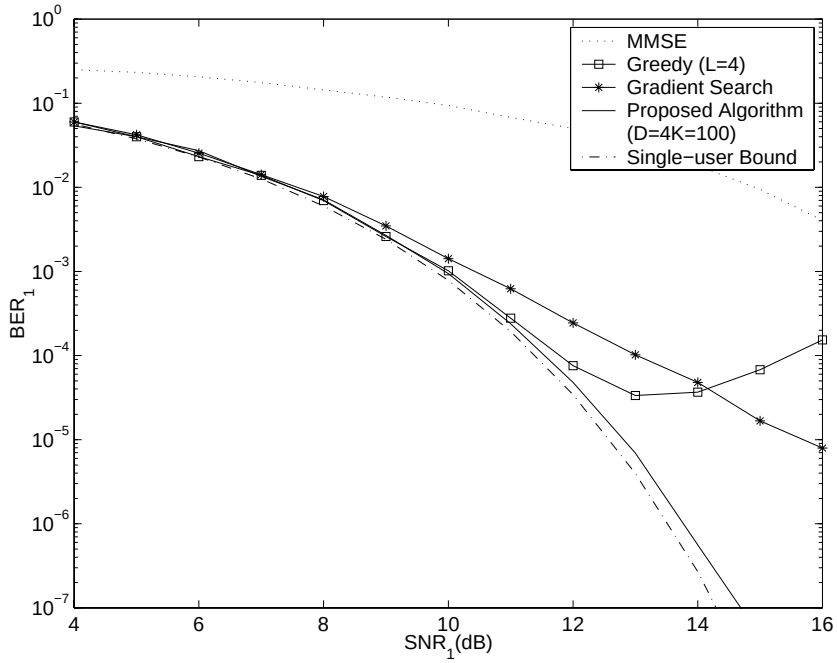

Fig. 7.4. Bit-error-rate versus SNR performance in the presence of severe MAI ($K = 25$, SNR_i in 8.5 to 20dB range, $i = 2, 3, \cdots, 25$).

Bibliography

[1] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 32, pp. 85-96, Jan. 1986.

[2] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 123-136, Jan. 1989.

[3] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 685-697, May 1994.

[4] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, pp. 3178-3188, Dec. 1994.

[5] S. L. Miller, "An adaptive direct-sequence code-division multiple-access receiver for multiuser interference rejection," *IEEE Trans. Commun.*, vol. 43, pp. 1746-1755, Feb./Mar./Apr. 1995.

[6] C. N. Pateros and G. J. Saulnier, "An adaptive correlator receiver for direct-sequence spread-spectrum communication," *IEEE Trans. Commun.*, vol. 44, pp. 1543-1552, Nov. 1996.

[7] R. Lupas and S. Verdú, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Commun.*, vol. 38, pp. 496-508, Apr. 1990.

[8] A. AlRustamani and B. R. Vojcic, "A new approach to greedy multiuser detection," *IEEE Trans. Commun.*, vol. 50, pp. 1326-1336, Aug. 2002.

[9] Q. Li and C. N. Georghiades, "On a geometric view of multiuser detection for synchronous DS/CDMA channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2723-2731, Nov. 2000.

[10] J. Hu and R. S. Blum, "A gradient guided search algorithm for multiuser detection," *IEEE Commun. Letters*, vol. 4, pp. 340-342, Nov. 2000.

[11] D. Chase, "A class of algorithms for decoding block codes with channel measurement information," *IEEE Trans. Inform. Theory*, vol. 18, pp. 170-182, Jan. 1972.

[12] L. Brunel and J. Boutros, "Lattice decoding for joint detection in direct sequence CDMA systems," *IEEE Trans. Inform. Theory*, submitted Jan. 2000.

[13] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1639-1642, July 1999.

[14] H. V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Trans. Commun.*, vol. 43, pp. 858-871, May 1997.

[15] J. Zhang, E. K. P. Chong, and D. N. C. Tse, "Output MAI distributions of linear MMSE multiuser receivers in DS-CDMA systems," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1128-1144, Mar. 2001.

[16] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*. New York: McGraw-Hill, 1990.

[17] Y. Wu and D. A. Pados, "An adaptive two-stage algorithm for ML and sub-ML decoding of binary linear block codes," *IEEE Trans. Inform. Theory*, vol. 49, pp. 261-269, Jan. 2003.

Chapter 8

Rapid Combined Synchronization/Demodulation Structures for DS/CDMA Systems, Part I: Algorithmic Developments

8.1 Introduction

The exponential complexity of the optimum direct-sequence code-division-multiple-access (DS/CDMA) demodulator [1] in conjunction with the severe performance degradation of the conventional matched-filter (MF) receiver in the presence of high power multiple access interferers (MAI) triggered a substantial research effort toward the development of demodulators with reasonable complexity that provide reliable decisions on the transmitted information bits. As a result, many suboptimal schemes have appeared in the literature such as the decorrelating demodulator [2],[3] which is the zero-forcing solution for multiuser interference rejection in noiseless channels, multistage architectures [4],[5], and decision feedback detectors [6]. Supervised back-propagation [7],[8], and unsupervised neural network structures [9], minimum-mean-square-error (MMSE) [10]-[13], minimum-variance-distortionless-response (MVDR) [14], and auxiliary-vector receivers [15]-[19] are additional examples.

Similarly to demodulation algorithms, synchronization schemes can be classified into supervised [20]-[27] or unsupervised (blind) [28]-[32] as well as multi-user [28], [29], [32] or single-user techniques [20]-[28],[31]-[33]. In particular, the successful operation of supervised synchronization algorithms relies on the periodic transmission of a known training sequence of information bits dedicated to the user of interest. This is done at the expense of reduced channel utilization relative to unsupervised techniques where no such training takes place. In addition, when supervised synchronization algorithms allow the active users to use correlated training sequences (for example the "all-1" sequence for all users) then a significant amount of coordination is required to ensure that no more than one user is in "training" mode (otherwise the receiver is trained for the correlated information bit transmissions of all users in "training" mode). Alternatively, supervised synchronization using pseudo-random training sequences requires that symbol-level synchronization has already been achieved before the corresponding training takes place. On the other hand, multi-user synchronization algorithms require knowledge of the signatures of all active users while single-user techniques assume known only the signature of the desired user (and possibly the number of the active users in the system [24],[28],[33]). Just like their demodulator counterparts, multi-user synchronization techniques may be employed only at the base station where the required information is available, while single-user techniques can be employed at either end of the communication link. As a final remark, we note that channel estimation algorithms [33]-[37] implicitly estimate the timing of the desired transmission and, thus, may be used in place of a dedicated synchronization algorithm.

We recall that timing information pertinent to the desired user(s) is the underlying assumption of all demodulator proposals in [1]-[19]. So, in practice, demodulators must be preceded by a synchronizer [38].

Another approach is to integrate the demodulator into a reliable synchronization scheme. In the first case the receiver design is a two-stage process, while in the second case only one *combined* stage is involved that integrates both timing estimation and demodulation of the desired user signal. In this context, the synchronization algorithms developed in [20]-[32] are demodulator-independent and as such they can precede any structure proposed in [1]-[19]. However, conventional matched-filter-type coarse synchronization techniques [20] are found to be interference-limited at least as much as their demodulator counterparts [21] and as such they are unable to provide reliable timing estimates in a multiuser environment. On the other hand, the synchronization schemes developed in [24]-[41] exhibit the desired near-far resistance that the conventional approaches fail to provide but they do so at the expense of significant increase in computational complexity. A viable alternative that balances well computational complexity, system simplicity, and MAI resistance is the *combined* synchronization and demodulation approach where both timing estimation and demodulation are performed at the same time. We note that the *combined* approach should not be confused with *joint* timing estimation and information bit detection strategies as in [39]-[41]. Combined synchronization and demodulation proposals include the work in [42],[43] where an MVDR/MMSE sample-matrix-inversion (SMI)-type receiver is developed of order (length) equal to twice the system processing gain (to ensure that the receiver input vector always contains a complete information symbol period). However, as we shall demonstrate, the choice of a high order filter combined with adaptive SMI implementations degrades severely the performance of the algorithms under small data support scenarios.

In this chapter we focus on rapid synchronization and combined demodulation based on adaptive short data record operations that enables tracking of rapidly changing communication environments. We propose two blind single-user combined synchronization/demodulation algorithms that result in linear structures of order L, where L is the system processing gain. First, we develop an MVDR-type *L-order* algorithm. Then, motivated by the limitations of the MVDR structure under short-data-record operation, we design an auxiliary-vector (AV)-type *L-order* scheme. Simulation studies demonstrate the effects of the choice of the filter order and the choice of the design algorithm on coarse synchronization and bit-error-rate (BER) receiver performance.

The chapter is organized as follows. In Section 8.2 we present the signal model. Design preliminaries are covered in Section 8.3. In Section 8.4 we develop the proposed combined synchronization/demodulation algorithms. Implementation issues are discussed in Section 8.5. Numerical and simulation comparisons are included in Section 8.6 and some conclusions are drawn in Section 8.7.

8.2 Signal Model

We consider an asynchronous DS/CDMA system populated by K active users transmitting over a common additive white Gaussian noise (AWGN) channel. The received signal $r(t)$ is the superposition of the K transmissions corrupted by channel noise, i.e.

$$r(t) = \sum_{k=0}^{K-1} r_k(t-\tau_k) + \nu(t). \tag{8.1}$$

In the above expression, τ_k is the delay of the kth user relative to the receiver's time reference point and $\nu(t)$ is AWGN. The kth user's contribution to the received signal $r(t)$ is given by

$$r_k(t) = \sum_{i=-\infty}^{\infty} \sqrt{E_k} b_k(i) S_k(t-iT), \quad k = 0, \ldots, K-1, \tag{8.2}$$

where $b_k(i) \in \{-1, +1\}$ is the ith information bit, E_k is transmitted energy, and T is the information bit period. The signature waveform $S_k(t)$ assigned to the kth user has the form

$$S_k(t) = \sum_{l=0}^{L-1} s_k(l) P_{T_c}(t-lT_c), \quad k = 0, \ldots, K-1, \tag{8.3}$$

where T_c is the chip duration, P_{T_c} is a rectangular pulse with support $[0, T_c]$, $L = T/T_c$ is the system processing gain, and $s_k(l)$ is the lth element of the normalized bipolar signature or spreading vector $\mathbf{s}_k \stackrel{\triangle}{=}$

$[s_k(0), s_k(1), \ldots, s_k(L-1)]^T$ that uniquely identifies the kth user. We assume that the signature waveforms are normalized to unit energy over a bit period, i.e. $\int_0^T S_k^2(t)dt = 1, \quad k = 0, \ldots, K-1$.

Without loss of generality, we let $\tau_k \in [0, T)$, $k = 0, \ldots, K-1$, in (8.1). Thus, we may write the delay τ_k as a sum of an integer multiple of chips plus a fraction of a chip, i.e. $\tau_k = (n_k + \delta_k)T_c$ where $n_k \in \{0, \ldots, L-1\}$ and $\delta_k \in [0, 1)$. After chip-matched filtering and chip-rate sampling, the continuous-time signal $r(t)$ in (8.1) is reduced to a discrete-time sequence $\{r(n)\}_{n=-\infty}^{\infty}$ where

$$r(n) = \int_{nT_c}^{(n+1)T_c} r(t)dt. \tag{8.4}$$

We note that since the chip timing of the desired user is unknown to the receiver the integrate-and-dump operation in (8.4) will result in SNR loss due to the averaging of adjacent chips. This SNR loss does not exceed 3dB [24] and can be reduced by oversampling the output of the chip-matched filter at the expense of increased receiver complexity. By substituting (8.1), (8.2), and (8.3) into (8.4), we obtain:

$$\begin{aligned} r(n) &= \sum_{k=0}^{K-1} \left\{ \sqrt{E_k} b_k \left(\left\lfloor \frac{n - n_k}{L} \right\rfloor \right) (1-\delta_k) s_k \left((n - n_k) \div L \right) \right. \\ &\quad \left. + \sqrt{E_k} b_k \left(\left\lfloor \frac{n - n_k - 1}{L} \right\rfloor \right) \delta_k s_k \left((n - n_k - 1) \div L \right) \right\} + \nu(n) \end{aligned} \tag{8.5}$$

where $\nu(n)$ is chip-matched filtered and sampled AWGN with variance σ^2, $\lfloor x \rfloor$ denotes the largest integer smaller than or equal to x and $\div L$ denotes the modulo-L operation.

The problem we consider in this work is the development of a linear receiver for combined synchronization and demodulation of the information bit of the user of interest, *user 0*. The only known quantity is the signature of the user of interest $\mathbf{s}_0$. All other system parameters, such as energies and delays of all users as well as signature vectors of the interfering users, are considered unknown.

8.3 Design Preliminaries and Background

Let

$$\begin{aligned} \mathbf{s}_k^{(n_k)} &\triangleq (1-\delta_k)[\underbrace{0, \ldots, 0}_{n_k}, s_k(0), \ldots, s_k(L - n_k - 1)]^T + \\ &\quad \delta_k[\underbrace{0, \ldots, 0}_{n_k+1}, s_k(0), \ldots, s_k(L - n_k - 2)]^T, \end{aligned} \tag{8.6}$$

and

$$\begin{aligned} \mathbf{s}_k^{(n_k-)} &\triangleq (1-\delta_k)[s_k(L - n_k), \ldots, s_k(L-1), \underbrace{0, \ldots, 0}_{L-n_k}]^T + \\ &\quad \delta_k[s_k(L - n_k - 1), \ldots, s_k(L-1), \underbrace{0, \ldots, 0}_{L-n_k-1}]^T, \end{aligned} \tag{8.7}$$

for $k = 0, \ldots, K-1$. Then, under the assumption of a linear filter receiver of length L, the corresponding $L \times 1$ received input vector $\mathbf{r}_i \triangleq [r(iL), \ldots, r((i+1)L - 1)]^T \in \Re^L$ can be written as follows:

$$\begin{aligned} \mathbf{r}_i &= \sqrt{E_0} \left\{ b_0(i) \mathbf{s}_0^{(n_0)} + b_0(i-1) \mathbf{s}_0^{(n_0-)} \right\} + \\ &\quad \sum_{k=1}^{K-1} \sqrt{E_k} \left\{ b_k(i) \mathbf{s}_k^{(n_k)} + b_k(i-1) \mathbf{s}_k^{(n_k-)} \right\} + \boldsymbol{\nu} \end{aligned} \tag{8.8}$$

where $\boldsymbol{\nu}$ is an AWGN vector with autocorrelation matrix $\sigma^2 \mathbf{I}$. The decision on the bit $b_0(i)$ that modulates $\mathbf{s}_0^{(n_0)}$ in the vector $\mathbf{r}_i$ is set to be

$$\hat{b}_0(i) = \text{sgn}(\mathbf{w}^T \mathbf{r}_i) \tag{8.9}$$

where $\mathbf{w} \in \Re^L$ is the filter tap-weight vector.

In this chapter, we develop combined synchronization/demodulation algorithms that are based on either the MVDR or the AV filtering approach. We recall that an MVDR-type filter [44] is a tap-weight vector $\mathbf{w}$ that minimizes the output variance/energy $E\{|\mathbf{w}^T\mathbf{r}_i|^2\}$ ($E\{\cdot\}$ denotes the statistical expectation operation) and at the same time is "distortionless" in a *reference* direction $\mathbf{d}$, i.e. $\mathbf{w}^T\mathbf{d} = 1$. If $\mathbf{R} \stackrel{\triangle}{=} E\{\mathbf{r}_i\mathbf{r}_i^T\}$ denotes the autocorrelation matrix of the received vector $\mathbf{r}_i$, then the MVDR tap-weight vector is given by

$$\mathbf{w}_{\mathbf{d},MVDR} = \frac{\mathbf{R}^{-1}\mathbf{d}}{\mathbf{d}^T\mathbf{R}^{-1}\mathbf{d}}. \tag{8.10}$$

In the rest of this chapter, the first subscript of the filter variable $\mathbf{w}$ will identify the reference vector while the second subscript will identify the receiver type (MVDR or AV). When the latter is not present, either type is applicable.

To obtain the solution in (8.10) for a given reference direction $\mathbf{d}$, it suffices to decompose $\mathbf{w}$ in the form $\mathbf{w} = \frac{\mathbf{d}}{\|\mathbf{d}\|^2} + \mathbf{y}$ where $\mathbf{d}^T\mathbf{y} = 0$ and search for the vector $\mathbf{y}$ that minimizes the variance at the output of $\mathbf{w}$. Instead of optimizing the orthogonal direction $\mathbf{y}$ as a whole, we further decompose the problem and optimize with respect to a sequence of weighted *auxiliary* vectors $\mathbf{g}_m$ and their corresponding scalar weights μ_m, $m = 1, 2, \ldots$, separately. The general method is termed auxiliary-vector (AV) filtering. The concept of AV filtering as introduced in [15]-[17] pertains to a particular choice for $\mathbf{g}_m$ and the use of a *conditionally MS-optimum* scalar μ_m for a given $\mathbf{g}_m$. For a receiver design with M auxiliary vectors, the linear tap weight vector $\mathbf{w}_{\mathbf{d},AV(M)}$ can be expressed as follows:

$$\mathbf{w}_{\mathbf{d},AV(M)} = \frac{\mathbf{d}}{\|\mathbf{d}\|^2} - \sum_{m=1}^{M} \mu_m \mathbf{g}_m, \quad M = 1, 2, 3, \ldots \ . \tag{8.11}$$

In contrast to [17] that proposes auxiliary vectors $\mathbf{g}_m$, $m = 1, \ldots, M$, orthogonal to one another (and thus limits the maximum number of auxiliary vectors M to $L-1$), in this chapter we relax the latter constraint and we follow the proposal of [18] and [19]. That is, we select the auxiliary vectors $\mathbf{g}_m$ and the steering scalars μ_m, $m = 1, \ldots, M$, recursively through conditional optimization according to the following two criteria: *(i)* The mth auxiliary vector $\mathbf{g}_m$ is set to maximize the magnitude of the cross-correlation between its output and the output of the previous step filter, $\mathbf{w}_{\mathbf{d},AV(m-1)} = \mathbf{d} - \sum_{j=1}^{m-1} \mu_j \mathbf{g}_j$, subject to the constraint that $\mathbf{g}_m$ is orthonormal to the reference vector $\mathbf{d}$. *(ii)* The steering scalar μ_m is set to minimize the MSE between the outputs $\mathbf{w}_{\mathbf{d},AV(m-1)}^T\mathbf{r}_i$ and $\mathbf{g}_m^T\mathbf{r}_i$. The whole algorithm is summarized below in the following simple recursive form [19]:

$$\mathbf{w}_{\mathbf{d},AV(m)} = \mathbf{w}_{\mathbf{d},AV(m-1)} - \mu_m \mathbf{g}_m, \ m = 1, 2, \ldots, \ \mathbf{w}_{\mathbf{d},AV(0)} = \frac{\mathbf{d}}{\|\mathbf{d}\|^2}, \tag{8.12}$$

$$\mathbf{g}_m = \frac{\mathbf{R}\mathbf{w}_{\mathbf{d},AV(m-1)} - (\mathbf{d}^T\mathbf{R}\mathbf{w}_{\mathbf{d},AV(m-1)})\mathbf{d}}{\|\mathbf{R}\mathbf{w}_{\mathbf{d},AV(m-1)} - (\mathbf{d}^T\mathbf{R}\mathbf{w}_{\mathbf{d},AV(m-1)})\mathbf{d}\|}, \tag{8.13}$$

$$\mu_m = \frac{\mathbf{g}_m^T\mathbf{R}\mathbf{w}_{\mathbf{d},AV(m-1)}}{\mathbf{g}_m^T\mathbf{R}\mathbf{g}_m}. \tag{8.14}$$

By inspection, we observe that for the MS optimum value of μ_m the product $\mu_m\mathbf{g}_m$ is independent of the norm of $\mathbf{g}_m$. Therefore, we may drop the unnecessary normalization operation in (8.13). Then, the auxiliary vector generation procedure may stop when $\mathbf{g}_{m+1} = \mathbf{0}$ and in that case $\mathbf{w}_{\mathbf{d},AV(m)}$ is exactly equal to $\mathbf{w}_{\mathbf{d},MVDR}$ (formal theoretical analysis and convergence proof can be found in [19]).

The auxiliary-vector filter $\mathbf{w}_{\mathbf{d},AV(m)}$ has two advantages in comparison with the $\mathbf{w}_{\mathbf{d},MVDR}$ filter in (8.10). First, no matrix inversion operation (neither explicit nor implicit) is required for the auxiliary-vector filter. The second and most important advantage has to do with the short data record behavior of the filter estimators $\widehat{\mathbf{w}}_{\mathbf{d},AV(0)}, \widehat{\mathbf{w}}_{\mathbf{d},AV(1)}, \widehat{\mathbf{w}}_{\mathbf{d},AV(2)}, \ldots$ and $\widehat{\mathbf{w}}_{\mathbf{d},MVDR}$ that are based on an N-point estimate of the autocorrelation matrix $\widehat{\mathbf{R}} = \frac{1}{N}\sum_{i=1}^{N-1}\mathbf{r}_i\mathbf{r}_i^T$[1]. As illustrated in [19], for a fixed finite data-record-size N the sequence $\{\widehat{\mathbf{w}}_{\mathbf{d},AV(m)}\}_m$ provides an infinite number of filter estimates with varying bias versus covariance

[1]The sequence $\{\widehat{\mathbf{w}}_{\mathbf{d},AV(m)}\}_m$ can be viewed not only as a sequence of estimates of the ideal AV-filter sequence in (8.12)-(8.14) but also, and most importantly, as a sequence of estimates of the MVDR filter in (8.10).

characteristics. For short data records N, the early, non-asymptotic, elements of the generated sequence of AV estimators offer favorable bias/covariance balance and are seen to outperform significantly in mean-square estimation error the $\widehat{\mathbf{w}}_{\mathbf{d},MVDR}$ estimator. This translates to rapid acquisition and superior short data record BER performance as demonstrated in the following section. Data-record-based criteria for the selection of the best AV-filter in the sequence have been recently reported in [45], [46].

The use of the reference vector $\mathbf{d}$ as described above in the context of MVDR or AV filtering allows the design of DS/CDMA filters that are distortionless in the vector direction of the user signal of interest. For example, in perfectly synchronous systems the vector direction of the signal of interest is $\mathbf{s}_0$. In coarsely synchronized systems ($n_0 = 0$) the vector direction $\mathbf{s}_0^*$ of the signal of interest is a convex combination of $\mathbf{s}_0$ and $\mathbf{s}_0^{(+1)} \triangleq [0, s_0(0), \ldots, s_0(L-2)]^T$ given by the following expression:

$$\mathbf{s}_0^* = (1-\delta_0)[s_0(0), \ldots, s_0(L-1)]^T + \delta_0[0, s_0(0), \ldots, s_0(L-2)]^T \quad (8.15)$$

$$= (1-\delta_0)\mathbf{s}_0 + \delta_0 \mathbf{s}_0^{(+1)}. \quad (8.16)$$

By setting the reference vector direction $\mathbf{d}$ equal to $\mathbf{s}_0^*$ we guarantee that the energy of the desired transmission is protected and the reduction of the filter output energy will be due to interference and noise suppression only. However, in non-synchronized systems the actual vector direction of the desired signal $\mathbf{s}_0^{(n_0)}$ is given by (8.6) for some unknown $n_0 \neq 0$ and it is different from $\mathbf{s}_0^*$. In this case, a receiver distortionless in the $\mathbf{s}_0^*$ direction will treat the desired transmission as interference (since it is no longer protected by the constraint $\mathbf{w}_{\mathbf{s}_0^*}^T \mathbf{s}_0^* = 1$) and it will suppress a portion of it. The lower the cross-correlation between $\mathbf{s}_0^*$ and $\mathbf{s}_0^{(n_0)}$ is, the more severe the suppression of the desired signal can be. Therefore, a minimum output variance receiver distortionless in a direction different than the actual vector direction of interest exhibits lower output energy than the output energy of the minimum output variance receiver that is distortionless in the actual direction of the user of interest. This observation forms the basis of the proposed combined synchronization/demodulation algorithm which is discussed in more detail in the following section.

8.4 Order-L Combined Synchronization/Demodulation

Effective demodulation of DS/CDMA signals requires a received vector that contains at least one complete information symbol. Chip-rate sampling of the chip-MF output at the receiver front-end implies that when we use a $2L$-long [42] or longer received vector, the latter condition is always satisfied regardless of the delay of the desired transmission. This requires the use of a linear filter of order at least $2L$. However, the longer the employed filter is, the higher the optimization and computational complexity. In addition, as we shall show in Chapter 9, adaptive SMI implementations of receivers employing long linear filters result in severe performance degradation in small data support situations. For these reasons, practical receiver designs should be based on small filter order choices. On the other hand, the shortest input vector that is possible to contain a complete information symbol is of length L, where L is the system processing gain. This implies that the length of an effective combined synchronization/demodulation structure is lower limited by L. Thus, our first objective toward the development of an effective receiver is to employ filters of order L. Our second objective is to integrate synchronization and demodulation into a "combined" scheme so that re-alignment of the received data and re-evaluation of a demodulation filter is no longer needed. To have a receiver readily available by the time synchronization is achieved, the proposed L-order combined synchronization/demodulation algorithms consist of the following two general steps. In the first step, the $[0, T)$ timing uncertainty interval is quantized to a finite number of hypotheses. For each hypothesis, the received samples are grouped into vectors of length L and a pair of linear receivers (distortionless in two different fractional shifts of the direction of interest) is evaluated based on the statistics of the corresponding input vector sequence. For each distortionless direction, the receiver with the highest output energy over all hypotheses provides a coarsely synchronized receiver. In the second step, we obtain a refined structure by maximizing the output energy of a linear combination of the two receivers that have already been evaluated in the first step. In the rest of the chapter, we refer to the first step of the proposed algorithm as the *coarse synchronization/demodulation* step and to the second step as *refined synchronization/demodulation*. Finally, our third objective is to achieve rapid synchronization and superior BER performance under short data record adaptation. In this section we develop two structures: An L-order MVDR-type and an L-order

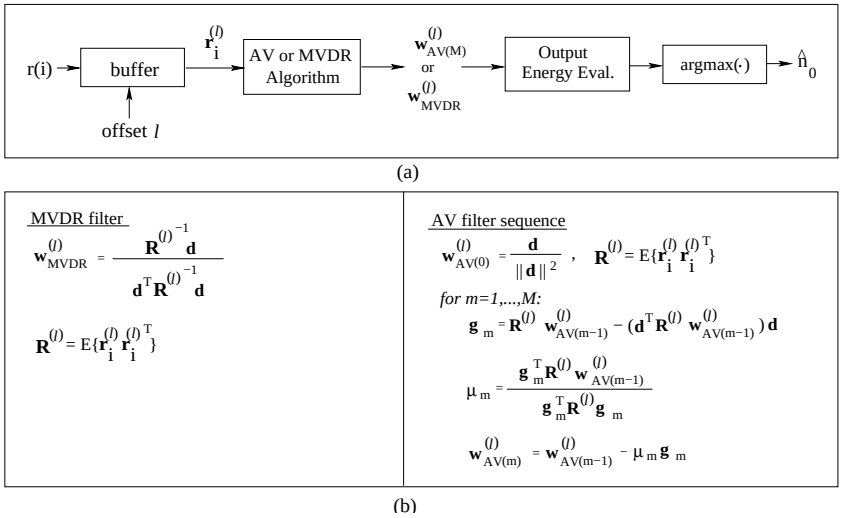

Fig. 8.1. Coarse synchronization based on filters of order L: (a) Block diagram. (b) Algorithmic description of the MVDR and AV filters.

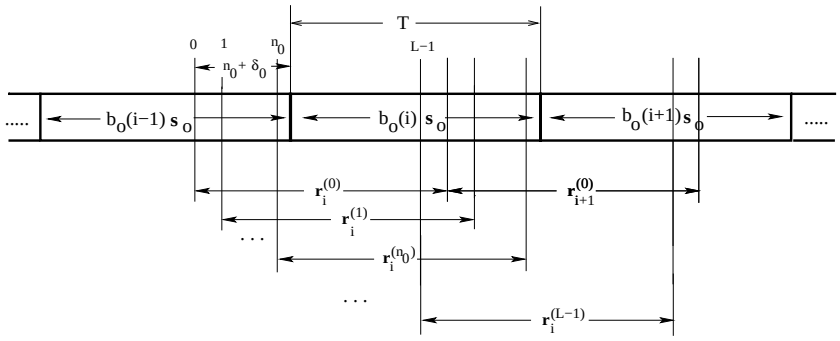

Fig. 8.2. Construction of the sequences $\mathbf{r}^{(l)}, l = 0, \ldots, L-1$.

AV-type linear structure. Both schemes meet the first two objectives and exhibit superior BER performance when compared to their $2L$ counterparts; the latter scheme is the most favorable under short data record adaptation.

In the rest of the section we present the details of the proposed schemes. Fig. 8.1 shows the block diagram description of an MVDR-type or AV-type coarse synchronizer/demodulator that operates on L-long received vectors. The buffer follows a chip-rate sampler and groups the received samples into L sequences of vectors, one for each of the L possible values of n_0, as depicted in Fig. 8.2.

The ith vector of the l-vector sequence is given by

$$\mathbf{r}_i^{(l)} = [r(iL+l), r(iL+l+1), \ldots, r((i+1)L+l-1)]^T, \quad l = 0, \ldots, L-1. \tag{8.17}$$

The subscript i of $\mathbf{r}_i^{(l)}$, $i = 0, \ldots, N-1$, denotes the order of the received sample vector in any of the above sequences in a data record of size N. For $l = n_0$, the received vectors $\mathbf{r}_i^{(n_0)}$, $i = 0, 1, \ldots, N-1$, contain a *complete* information symbol that is modulated by the vector $\mathbf{s}_0^*$ given by (8.16). To efficiently demodulate the desired user transmission, a linear receiver should operate exactly on the $\{\mathbf{r}_i^{(n_0)}\}$ received vector sequence. Thus, efficient and effective estimation of n_0 is needed and this is performed implicitly by the first step of the proposed algorithm. The discussion in the previous section *suggested* that the output energy of a receiver that is distortionless in the $\mathbf{s}_0^*$ direction is maximized when the offset l in Fig. 8.2 is equal to n_0. In other

words,

$$n_0 = \arg\max_l \{\mathbf{w}_{\mathbf{s}_0^*}^{(l)^T} \mathbf{R}^{(l)} \mathbf{w}_{\mathbf{s}_0^*}^{(l)} \mid l = 0, \ldots, L-1\} \tag{8.18}$$

where

$$\mathbf{R}^{(l)} = E\{\mathbf{r}_i^{(l)} \mathbf{r}_i^{(l)^T}\} \tag{8.19}$$

is the autocorrelation matrix of the received vectors of the lth sequence, $l = 0, \ldots, L-1$, and $\mathbf{w}_{\mathbf{s}_0^*}^{(l)}$ is an AV or MVDR-type filter distortionless in $\mathbf{s}_0^*$ and evaluated based on the statistics of the filter input sequence $\{\mathbf{r}_i^{(l)}\}$ (for example $\mathbf{w}_{\mathbf{s}_0^*, MVDR}^{(l)} = \mathbf{R}^{(l)^{-1}} \mathbf{s}_0^* / \left(\mathbf{s}_0^{*T} \mathbf{R}^{(l)^{-1}} \mathbf{s}_0^*\right)$). Indeed, the following proposition shows that (8.18) *holds true* for the case of a power-controlled system employing an MVDR receiver, provided that the SNR of the received signals is sufficiently large. The proof is included in the Appendix.

Proposition 8.1 *In power-controlled systems and for sufficiently large user SNRs, the quantity*

$$\mathbf{w}_{\mathbf{s}_0^*, MVDR}^{(l)^T} \mathbf{R}^{(l)} \mathbf{w}_{\mathbf{s}_0^*, MVDR}^{(l)}, \quad l = 0, \ldots, L-1, \tag{8.20}$$

is maximized for $l = n_0$, where $\mathbf{w}_{\mathbf{s}_0^*, MVDR}^{(l)} = \frac{\mathbf{R}^{(l)^{-1}} \mathbf{s}_0^*}{\mathbf{s}_0^{*T} \mathbf{R}^{(l)^{-1}} \mathbf{s}_0^*}$. □

Therefore, to achieve coarse synchronization with the desired transmission, it suffices to find the value of l that maximizes the output energy of a receiver distortionless in the direction of the vector $\mathbf{s}_0^*$. To overcome the fact that at this stage the parameter δ_0 of the effective signature $\mathbf{s}_0^*$ is unknown, we use two normalized reference vectors $\mathbf{d}_0$ and $\mathbf{d}_1$ given by (8.16) for $\delta_0 = 0$ and $\delta_0 = 0.5$, respectively, i.e.

$$\mathbf{d}_0 = \mathbf{s}_0, \tag{8.21}$$

$$\mathbf{d}_1 = (0.5\mathbf{s}_0 + 0.5\mathbf{s}_0^{(+1)}) / \|0.5\mathbf{s}_0 + 0.5\mathbf{s}_0^{(+1)}\|. \tag{8.22}$$

The coarse synchronization algorithm is summarized below.

L-order coarse synchronization algorithm

(i) For $l = 0, \ldots, L-1$, calculate the lth autocorrelation matrix

$$\mathbf{R}^{(l)} = E\{\mathbf{r}_i^{(l)} \mathbf{r}_i^{(l)^T}\} \tag{8.23}$$

where the vector $\mathbf{r}_i^{(l)}$ is given by (8.17).

(ii) For each autocorrelation matrix calculate the AV (or MVDR) filters $\mathbf{w}_{\mathbf{d}_0}^{(l)}$ and $\mathbf{w}_{\mathbf{d}_1}^{(l)}$ that are distortionless in the $\mathbf{d}_0$ and $\mathbf{d}_1$ vector direction, respectively.

(iii) Calculate the output energy of the filters:

$$OE^{(2l)} \triangleq \mathbf{w}_{\mathbf{d}_0}^{(l)^T} \mathbf{R}^{(l)} \mathbf{w}_{\mathbf{d}_0}^{(l)} \quad \text{and} \quad OE^{(2l+1)} \triangleq \mathbf{w}_{\mathbf{d}_1}^{(l)^T} \mathbf{R}^{(l)} \mathbf{w}_{\mathbf{d}_1}^{(l)}, \quad l = 0, \ldots, L-1. \tag{8.24}$$

(iv) Let

$$j_{max} = \arg\max_j \{OE^{(j)} | j = 0, \ldots, 2L-1\}. \tag{8.25}$$

Then,

$$\hat{n}_0 = \lfloor j_{max}/2 \rfloor. \tag{8.26}$$

The effective signature in $\mathbf{r}_i^{(\hat{n}_0)}$ is the vector $\mathbf{s}_0^*$.

Receiver refinement can be performed as follows. In the case of the MVDR receiver, the ideal (perfectly refined) receiver is given by

$$\mathbf{w}_{\mathbf{s}_{\|0\|}^*(\delta_0), MVDR} = \frac{\mathbf{R}^{(\hat{n}_0)^{-1}} \mathbf{s}_{\|0\|}^*(\delta_0)}{\mathbf{s}_{\|0\|}^{*T}(\delta_0) \mathbf{R}^{(\hat{n}_0)^{-1}} \mathbf{s}_{\|0\|}^*(\delta_0)} \tag{8.27}$$

where $\mathbf{s}^*_{\|0\|}(\delta_0)$ is the normalized effective signature

$$\mathbf{s}^*_{\|0\|}(\delta) \stackrel{\triangle}{=} \frac{(1-\delta)\mathbf{s}_0 + \delta\mathbf{s}_0^{(+1)}}{||(1-\delta)\mathbf{s}_0 + \delta\mathbf{s}_0^{(+1)}||} \tag{8.28}$$

evaluated at $\delta = \delta_0$. Since $(1-\delta_0)\mathbf{s}_0+\delta_0\mathbf{s}_0^{(+1)} = (1-2\delta_0)\mathbf{d}_0+2\delta_0\|0.5\mathbf{s}_0+0.5\mathbf{s}_0^{(+1)}\|\mathbf{d}_1$, the filter in (8.27) can be expressed as a linear combination of the two already evaluated MVDR filters $\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_0,MVDR}$ and $\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_1,MVDR}$ (or as a linear combination of $\mathbf{w}^{(\hat{n}_0-1)}_{\mathbf{d}_0,MVDR}$ and $\mathbf{w}^{(\hat{n}_0-1)}_{\mathbf{d}_1,MVDR}$ when j_{max} is even and $OE(j_{max}-1) > OE(j_{max}+1)$). Based on the above observation we may refine our initial filter estimates $\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_0,MVDR}$ and $\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_1,MVDR}$ by linearly combining them to produce a filter $\widetilde{\mathbf{w}}_{MVDR}$ with maximum output energy. The combination rule is given by the following proposition. The proof is included in the Appendix.

Proposition 8.2 *(i) Let $\mathcal{W}$ be the set of linear combinations of $\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_0,MVDR}$ and $\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_1,MVDR}$ of the form*

$$\mathbf{w} = \frac{\alpha_0(\delta)\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_0,MVDR} + \alpha_1(\delta)\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_1,MVDR}}{\left(\alpha_0(\delta)\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_0,MVDR} + \alpha_1(\delta)\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_1,MVDR}\right)^T \mathbf{s}^*_{\|0\|}(\delta)} \tag{8.29}$$

with

$$\alpha_0(\delta) = (1-2\delta)\mathbf{s}_0^T{\mathbf{R}^{(\hat{n}_0)}}^{-1}\mathbf{s}_0, \tag{8.30}$$

$$\alpha_1(\delta) = \frac{2\delta(0.5\mathbf{s}_0 + 0.5\mathbf{s}_0^{(+1)})^T{\mathbf{R}^{(\hat{n}_0)}}^{-1}(0.5\mathbf{s}_0 + 0.5\mathbf{s}_0^{(+1)})}{\|0.5\mathbf{s}_0 + 0.5\mathbf{s}_0^{(+1)}\|}, \text{and } \delta \in [0,1). \tag{8.31}$$

Then, the refined filter $\widetilde{\mathbf{w}}_{MVDR}$ is the solution to the following optimization problem:

$$\widetilde{\mathbf{w}}_{MVDR} = \arg\max_{\mathbf{w}\in\mathcal{W}} \mathbf{w}^T\mathbf{R}^{(\hat{n}_0)}\mathbf{w}. \tag{8.32}$$

(ii) Equivalently, the filter $\widetilde{\mathbf{w}}_{MVDR}$ is given by

$$\widetilde{\mathbf{w}}_{MVDR} = \frac{\alpha_0(\delta_{max})\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_0,MVDR} + \alpha_1(\delta_{max})\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_1,MVDR}}{\left(\alpha_0(\delta_{max})\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_0,MVDR} + \alpha_1(\delta_{max})\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_1,MVDR}\right)^T \mathbf{s}^*_{\|0\|}(\delta_{max})} \tag{8.33}$$

where δ_{max} is the value of $\delta \in [0,1)$ that maximizes the output energy of an MVDR receiver, i.e.

$$\delta_{max} = \arg\max_{\delta\in[0,1)} \left[\mathbf{s}^{*T}_{\|0\|}(\delta){\mathbf{R}^{(\hat{n}_0)}}^{-1}\mathbf{s}^*_{\|0\|}(\delta)\right]^{-1}. \tag{8.34}$$

□

The case of the AV receiver is treated similarly. To obtain a combined synchronization/demodulation scheme, we form (as in the MVDR case) the final refined receiver $\widetilde{\mathbf{w}}_{AV}$, as a linear combination of two coarsely synchronized AV receivers, namely $\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_0,AV}$ and $\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_1,AV}$, that were already obtained during the coarse synchronization process (or as a linear combination of $\mathbf{w}^{(\hat{n}_0-1)}_{\mathbf{d}_0,AV}$ and $\mathbf{w}^{(\hat{n}_0-1)}_{\mathbf{d}_1,AV}$ when j_{max} is even and $OE(j_{max}-1) > OE(j_{max}+1)$), i.e.

$$\widetilde{\mathbf{w}}_{AV} = \alpha_0(\delta)\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_0,AV} + \alpha_1(\delta)\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_1,AV}. \tag{8.35}$$

To ensure that $\widetilde{\mathbf{w}}_{AV}$ is of the general form

$$\widetilde{\mathbf{w}}_{AV} = \mathbf{s}^*_{\|0\|}(\delta) + \mathbf{y} \tag{8.36}$$

for some $\mathbf{y}$ orthogonal to $\mathbf{s}^*_{\|0\|}(\delta)$, the scalar weights $\alpha_0(\delta)$ and $\alpha_1(\delta)$ in (8.35) are first set equal to

$$\alpha_0(\delta) = (1-2\delta)/\|(1-\delta)\mathbf{s}_0 + \delta\mathbf{s}_0^{(+1)}\| \quad \text{and} \tag{8.37}$$

$$\alpha_1(\delta) \quad = \quad 2\delta\|0.5\mathbf{s}_0 + 0.5\mathbf{s}_0^{(+1)}\|/\|(1-\delta)\mathbf{s}_0 + \delta\mathbf{s}_0^{(+1)}\|. \tag{8.38}$$

If we combine (8.35), (8.37), and (8.38) we obtain

$$\widetilde{\mathbf{w}}_{AV} = \mathbf{s}^*_{\|0\|}(\delta) + \alpha_0(\delta)\mathbf{y}_0 + \alpha_1(\delta)\mathbf{y}_1 \tag{8.39}$$

where $\mathbf{y}_0 = -\sum_{i=1}^{M} \mu_{\mathbf{d}_0,i}\mathbf{g}_{\mathbf{d}_0,i}$ and $\mathbf{y}_1 = -\sum_{i=1}^{M} \mu_{\mathbf{d}_1,i}\mathbf{g}_{\mathbf{d}_1,i}$ (the subscripts $\mathbf{d}_0$ and $\mathbf{d}_1$ of μ_i and $\mathbf{g}_i$ indicate the corresponding AV filters obtained during coarse synchronization, $\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_0,AV}$ and $\mathbf{w}^{(\hat{n}_0)}_{\mathbf{d}_1,AV}$ respectively). We observe that (8.39) is not necessarily of the form of (8.36) since the vector $\alpha_0(\delta)\mathbf{y}_0 + \alpha_1(\delta)\mathbf{y}_1$ is not necessarily orthogonal to $\mathbf{s}^*_{\|0\|}(\delta)$. All we need is to use the orthogonal projections $\mathbf{y}_0^{\perp}$ and $\mathbf{y}_1^{\perp}$ onto the subspace spanned by $\mathbf{s}^*_{\|0\|}(\delta)$ for $\delta \in [0,1)$ instead of $\mathbf{y}_0$ and $\mathbf{y}_1$, respectively. Then, $\mathbf{y} = \alpha_0(\delta)\mathbf{y}_0^{\perp} + \alpha_1(\delta)\mathbf{y}_1^{\perp}$ is orthogonal to $\mathbf{s}^*_{\|0\|}(\delta)$ for any $\delta \in [0,1)$ and does qualify as the vector $\mathbf{y}$ in (8.36). We note that the vector set $\{\mathbf{s}^*_{\|0\|}(\delta) \mid \delta \in [0,1)\}$ has dimension 2, therefore $\mathbf{y}$ is given as follows:

$$\mathbf{y} = \alpha_0(\delta)(\mathbf{I} - \mathbf{u}_0\mathbf{u}_0^T - \mathbf{u}_1\mathbf{u}_1^T)\mathbf{y}_0 + \alpha_1(\delta)(\mathbf{I} - \mathbf{u}_0\mathbf{u}_0^T - \mathbf{u}_1\mathbf{u}_1^T)\mathbf{y}_1, \tag{8.40}$$

where $\{\mathbf{u}_0, \mathbf{u}_1\}$ is an orthonormal basis of the space spanned by $\{\mathbf{s}_0, \mathbf{s}_0^{(+1)}\}$. Combining (8.39) and (8.40) we obtain an expression for the AV-type receiver as a function of δ:

$$\widetilde{\mathbf{w}}_{\mathbf{s}^*_{\|0\|}(\delta),AV} = \mathbf{s}^*_{\|0\|}(\delta) + \alpha_0(\delta)(\mathbf{I} - \mathbf{u}_0\mathbf{u}_0^T - \mathbf{u}_1\mathbf{u}_1^T)\mathbf{y}_0 + \alpha_1(\delta)(\mathbf{I} - \mathbf{u}_0\mathbf{u}_0^T - \mathbf{u}_1\mathbf{u}_1^T)\mathbf{y}_1. \tag{8.41}$$

Finally, the combined synchronization/demodulation AV-type receiver is given by $\widetilde{\mathbf{w}}_{\mathbf{s}^*_{\|0\|}(\delta_{max}),AV}$ where

$$\delta_{max} = \arg\max_{\delta\in[0,1)} E(\delta) = \arg\max_{\delta\in[0,1)} \{\widetilde{\mathbf{w}}^T_{\mathbf{s}^*_{\|0\|}(\delta),AV}\mathbf{R}^{(\hat{n}_0)}\widetilde{\mathbf{w}}_{\mathbf{s}^*_{\|0\|}(\delta),AV}\}. \tag{8.42}$$

In practice, the autocorrelation matrix $\mathbf{R}^{(l)}, l = 0, \ldots, L-1$, is sample-average estimated:

$$\widehat{\mathbf{R}}^{(l)} = \frac{1}{N}\sum_{i=0}^{N-1} \mathbf{r}_i^{(l)}\mathbf{r}_i^{(l)^T} \tag{8.43}$$

where N is the data record size. For values of N less than the filter order L the estimate in (8.43) is not full-rank, thus the sample matrix inversion (SMI) estimate of the MVDR filter does not exist.

8.5 Implementation Issues

Evaluation of the sample average estimate of the input autocorrelation matrices $\mathbf{R}^{(l)}, l = 0, \ldots, L-1$, constitutes a major part of the computational burden of both proposed algorithms. Straightforward application of (8.43) involves a total number of floating point operations of order $O(L^3)$. However, the special structure of the matrix $\mathbf{R}^{(l)}$ allows evaluation with a total number of floating point operations of order $O(L^2)$. This can be achieved by exploiting the "overlapping" property of the autocorrelation matrix. Indeed, we may observe that the matrices $\widehat{\mathbf{R}}^{(l)}$ are submatrices of $\widehat{\overline{\mathbf{R}}} = \frac{1}{N}\sum_{i=1}^{N} \bar{\mathbf{r}}_i\bar{\mathbf{r}}_i^T$, where the $2L$-long vectors $\bar{\mathbf{r}}_i$, $i = 0, \ldots, N-1$, are formed by a sliding window such that successive vectors share L common data samples:

$$\bar{\mathbf{r}}_i \quad = \quad [r[iL], r[iL+1], \ldots, r[(i+2)L-1]]^T, \tag{8.44}$$

$$\bar{\mathbf{r}}_{i+1} \quad = \quad [r[(i+1)L], r[(i+1)L+1], \ldots, r[(i+3)L-1]]^T. \tag{8.45}$$

More precisely,

$$\widehat{\mathbf{R}}^{(l)} = \begin{pmatrix} \widehat{\overline{R}}(l,l) & \ldots & \widehat{\overline{R}}(l,l+L-1) \\ \vdots & \ddots & \vdots \\ \widehat{\overline{R}}(l+L-1,l) & \ldots & \widehat{\overline{R}}(l+L-1,l+L-1) \end{pmatrix}, \; l = 0, \ldots, L-1, \tag{8.46}$$

Table 8.1. Computational Complexity of Covariance Matrix Estimation

L-order	
Additions:	$\frac{3}{2}L(L+1)(N-1)$
Multiplications:	$\frac{3}{2}L(L+1)N$
$2L$-order	
Additions:	$L(2L+1)(N-1)$
Multiplications:	$L(2L+1)N$

where $\widehat{\overline{R}}(i,j), i=0,\ldots,2L-1,\ j=0,\ldots,2L-1$, denotes the (i,j)th element of $\widehat{\overline{\mathbf{R}}}$. In Table 8.1 we show the total number of multiplications and additions needed for the evaluation of $\widehat{\mathbf{R}}^{(l)}$, $l=0,\ldots,L-1$, (proposed L-order algorithms) and $\widehat{\overline{\mathbf{R}}}$ ($2L$-order algorithms [42]) based on a data record of size N.

On the other hand, the evaluation of the inverse estimates $\left[\widehat{\mathbf{R}}^{(l)}\right]^{-1}$ required by the SMI-MVDR-type algorithm involves computations of order $O(NL^3)$. Inverse matrix estimation can be achieved recursively described in the following proposition. The proof is included in the Appendix.

Proposition 8.3 *Let the matrix* $\widehat{\mathbf{R}}^{(l+1)}$ *be partitioned as follows:*

$$\widehat{\mathbf{R}}^{(l+1)} = \begin{pmatrix} \mathbf{C} & \mathbf{b}_{l+1} \\ \mathbf{b}_{l+1}^T & a_{l+1} \end{pmatrix} \tag{8.47}$$

where $\mathbf{C}$ *is an* $(L-1)\times(L-1)$ *matrix,* $\mathbf{b}_{l+1}$ *is an* $(L-1)\times 1$ *vector, and* a_{l+1} *is a scalar. Then*

$$\widehat{\mathbf{R}}^{(l+1)^{-1}} = \begin{pmatrix} \mathbf{C}^{-1} + \frac{\mathbf{C}^{-1}\mathbf{b}_{l+1}\mathbf{b}_{l+1}^T\mathbf{C}^{-1}}{a_{l+1}-\mathbf{b}_{l+1}^T\mathbf{C}^{-1}\mathbf{b}_{l+1}} & (\mathbf{b}_{l+1}^T\mathbf{C}^{-1}\mathbf{b}_{l+1}-a_{l+1})^{-1}\mathbf{C}^{-1}\mathbf{b}_{l+1} \\ (\mathbf{b}_{l+1}^T\mathbf{C}^{-1}\mathbf{b}_{l+1}-a_{l+1})^{-1}\mathbf{b}_{l+1}^T\mathbf{C}^{-1} & (a_{l+1}-\mathbf{b}_{l+1}^T\mathbf{C}^{-1}\mathbf{b}_{l+1})^{-1} \end{pmatrix}. \tag{8.48}$$

The matrix $\mathbf{C}^{-1}$ *is given by*

$$\mathbf{C}^{-1} = \mathbf{\Gamma} - \frac{1}{\alpha}\boldsymbol{\beta}\boldsymbol{\beta}^T \tag{8.49}$$

where the scalar α*, the vector* $\boldsymbol{\beta}_{(L-1)\times 1}$*, and the matrix* $\mathbf{\Gamma}_{(L-1)\times(L-1)}$ *form the partition of* $\widehat{\mathbf{R}}^{(l)^{-1}}$

$$\widehat{\mathbf{R}}^{(l)^{-1}} = \begin{pmatrix} \alpha & \boldsymbol{\beta}^T \\ \boldsymbol{\beta} & \mathbf{\Gamma} \end{pmatrix}. \tag{8.50}$$

□

In Table 8.2 we show the total number of additions, multiplications, and divisions needed by the proposed L-order coarse synchronization algorithms (SMI-MVDR and AV-type) and the $2L$-order SMI-MVDR-type coarse synchronization algorithm [42]. The AV scheme utilizes M auxiliary vectors. It is interesting to note that the sum of the total number of floating point multiplications and divisions required by the L-order SMI-MVDR algorithm is at least 45% less than the corresponding number required by the $2L$-order SMI-MVDR algorithm. At the same time the L-order algorithms outperform the $2L$-order algorithms in coarse synchronization error and bit-error-rate, as we will see in the next Section.

8.6 Simulation Results and Comparisons

We consider a 11-user asynchronous DS-CDMA system that utilizes Gold sequences of length $L=31$. We compare the proposed L-order coarse synchronization algorithms with the $2L$-order scheme in [42]. We also develop the $2L$-order AV-type algorithm that we include together with the L and $2L$-order MF-type receivers as reference points. The probability of coarse synchronization error or "coarse-synchronization-error-rate"

Table 8.2. Computational Complexity of Coarse Synchronization Algorithms

L-order	
SMI-MVDR	
Additions:	$(2L-1)(L-1)^2+\frac{1}{4}(L-1)L(L-1/2)+\frac{1}{4}(L-1)L+2L(L^2-1)$
Multiplications:	$2(L-1)^2(L+1)+\frac{7}{12}L(L-1)(L-1/2)+\frac{5}{4}L(L-1)+2L^2(L+1)$
Divisions:	$5L-2$
AV	
Additions:	$2L(L^2-1)+2ML[6L-3+L(L-1)]$
Multiplications:	$(12M+2)L^2+2(M+1)L^3$
Divisions:	$2ML$

$2L$-order	
SMI-MVDR	
Additions:	$\frac{1}{2}L(2L-1)(2L-1/2)+\frac{3}{2}L(2L-1)+2L^3$
Multiplications:	$\frac{7}{6}L(2L-1)(2L-1/2)+\frac{5}{2}L(2L-1)+2L^3+2L(L+1)$
Divisions:	$4L$

(CSER) is defined as the probability that the delay estimate provided by the coarse synchronization algorithm is off by more than $0.5T_c$ from the actual delay. The coarse-synchronization-error-rate is evaluated by averaging over 5,000 independent simulation experiments. The delays of all users are chosen randomly and are kept constant over 5 independent experiments. The autocorrelation matrix of the received signal is estimated by sample averaging for both the SMI-MVDR and the AV scheme. Finally, the number of auxiliary vectors utilized by the AV receiver is optimized ("genie" assisted) with respect to the corresponding performance metric (CSER or BER). Data record based criteria for the selection of the best number of auxiliary vectors (i.e. best AV filter in the sequence) have been recently reported in [45], [46].

In Fig. 8.3 the coarse-synchronization-error-rate performance is depicted as the data support size ranges from 15 to 200 samples. The SNRs[2] of the interferers are fixed at 10, 12, 14, 16, 18, 20, 22, 24, 26, 28dB. The SNR of the user of interest is fixed at 22dB. We observe that the AV order-L and the AV order-$2L$ curves are close to each other which suggests that the AV-type algorithms are not very sensitive with respect to the filter length. Both AV-type algorithms outperform significantly the L-order and $2L$-order MVDR algorithms (in the small data-support-size region the L-order MVDR algorithm achieves the same CSER as the $2L$-order MVDR algorithm approximately L samples faster).

To quantify the (in-)sensitivity of the proposed receivers to synchronization errors we compare the BER performance of the AV and the SMI-MVDR receiver to the performance of their perfectly synchronized counterparts. Fig. 8.4 depicts the performance of the proposed receivers as a function of the data record size for receivers of order L and $2L$. The user SNR setup in Fig. 8.4 is identical to the setup in Fig. 8.3. We observe that the proposed receivers perform close to the perfect synchronization performance levels. As in Fig. 8.3, the best performance is attained by the AV-type schemes while the L-order SMI-MVDR outperforms the $2L$-order SMI-MVDR over the whole data-record-size range.

In Fig. 8.5 we show the BER performance of the L and $2L$-order *diagonally loaded* SMI-MVDR receivers. The $2L$-order SMI-MVDR receiver utilizes a sample-average autocorrelation matrix that is diagonally loaded by the amount $\alpha\,\mathrm{trace}(\widehat{\mathbf{R}})\,\mathbf{I}$, where $\alpha=10^{-4}$ as suggested in [42]. For a fair comparison, the same diagonal loading value is used in the implementation of the L-order algorithm. We understand, however, that any algorithm that utilizes diagonal loading is sensitive to the specific choice of the real-valued diagonal-loading parameter. The results in Fig. 8.5 are quite similar to those of Fig. 8.4 and the main benefit is operation under sample support less than the filter order.

In Figs. 8.6 and 8.7, the coarse-synchronization-error-rate performance is shown as a function of the

[2]The SNR of user k, $k=0,\ldots,K-1$, is defined as E_k/σ^2 where σ^2 is the variance of the chip-matched filtered and sampled AWGN.

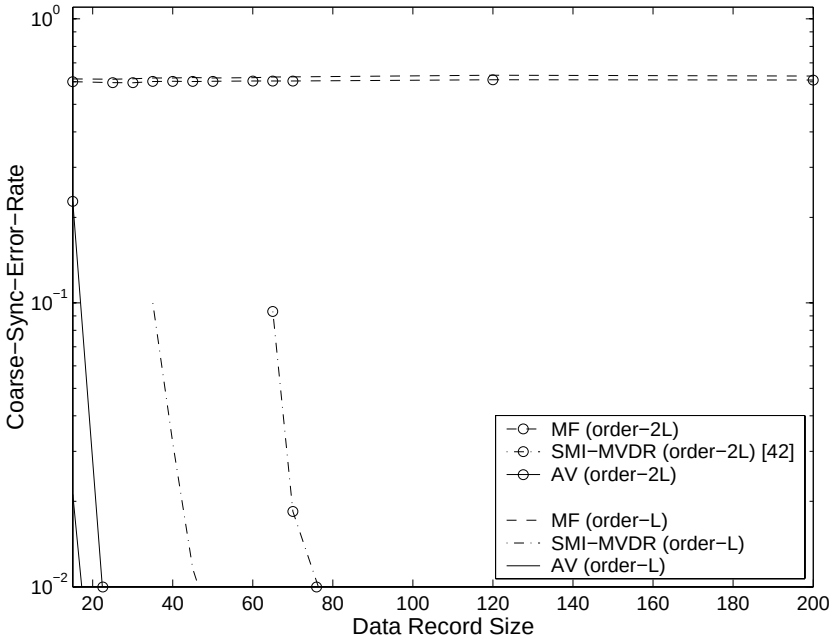

Fig. 8.3. Coarse-synchronization-error-rate as a function of the data-record-size for linear receivers of order-L and order-$2L$.

SNR of the user of interest for receivers of order L and $2L$, respectively. The SNRs of the interferers are the same as in Fig. 8.3. To illustrate the performance of the receivers in short data record scenarios, we choose the total number of samples used for the estimation of the autocorrelation matrix to be slightly larger than the employed filter order. Thus, for the L-order algorithms 55 samples are used, while for the $2L$-order algorithms 85 samples are used. It is interesting to note that the L-order SMI-MVDR receiver performs close to its $2L$-order counterpart, even though the former utilizes a smaller data record size. As expected, the AV receivers (of order L or $2L$) offer the lowest synchronization error.

Figs. 8.8 and 8.9 depict the BER as a function of the SNR of the user of interest for receivers of order L and $2L$, respectively. The interfering users SNR setup in Figs. 8.8 and 8.9 is identical to that in Fig. 8.3. As in Figs.8.6 and 8.7, 55 samples are used to estimate the autocorrelation matrix of the L-order algorithms and 85 samples for matrix estimation of the $2L$-order algorithms. The performance results in Figs. 8.8 and 8.9 parallel the CSER findings in Figs. 8.6 and 8.7, while – as in Fig. 8.4 – the proposed receivers perform close to the perfect synchronization performance levels.

8.7 Conclusions

We considered the receiver design problem for asynchronous DS/CDMA systems. We focused on blind combined synchronization/demodulation receiver designs that lead to simple structures and eliminate the need for an extra stage that performs timing extraction of the desired transmission. Along these lines, first we developed an MVDR-type structure of order L with enhanced performance over higher order structures of the same type. To overcome the basic limitations of MVDR-type filtering under short-data-record operations, we developed an L-order linear structure based on auxiliary-vector filtering principles. Computationally efficient implementations of the proposed schemes were discussed and simulation studies were performed that demonstrated the superiority of the proposed L-order schemes when the objective is rapid synchronization and reliable demodulation with limited data support.

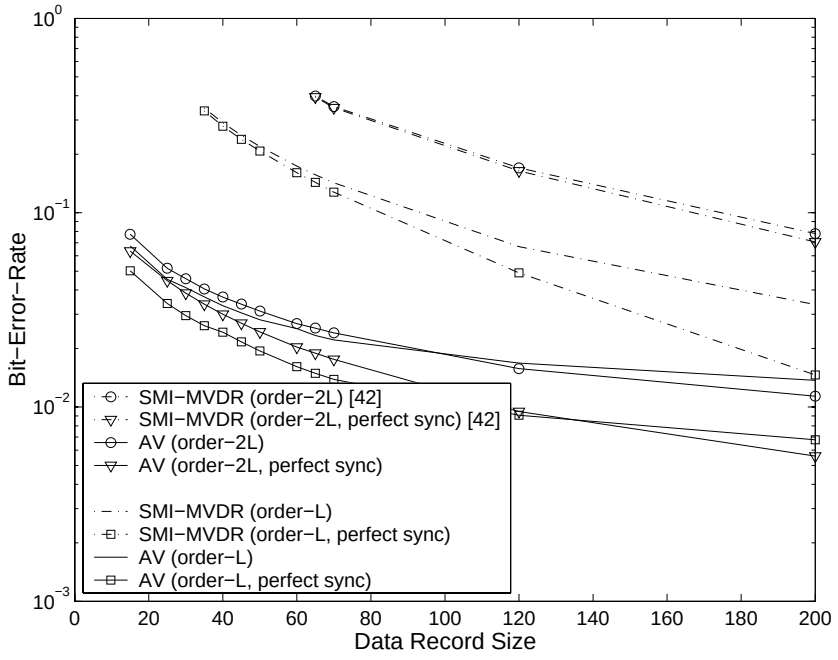

Fig. 8.4. Bit-error-rate as a function of the data record size for linear receivers of order-L and order-$2L$ (the BER performance of their perfectly synchronized counterparts is also included for reference).

8.8 Appendix

Proof of Proposition 8.1

In a power controlled system, $E_i = E$, $i = 0, \ldots, K-1$. If $\mathbf{w}^{(l)}_{\mathbf{s}_0^*, MVDR} = \frac{\mathbf{R}^{(l)^{-1}}\mathbf{s}_0^*}{\mathbf{s}_0^{*T}\mathbf{R}^{(l)^{-1}}\mathbf{s}_0^*}$, $l = 0, \ldots, L-1$, then the output energy is given by $(\mathbf{s}_0^{*T}\mathbf{R}^{(l)^{-1}}\mathbf{s}_0^*)^{-1}$. For the system under consideration, the autocorrelation matrix $\mathbf{R}^{(l)}$ is given by

$$\mathbf{R}^{(l)} = E\boldsymbol{\Sigma}^{(l)}\boldsymbol{\Sigma}^{(l)^T} + \sigma^2\mathbf{I} \tag{8.51}$$

where

$$\boldsymbol{\Sigma}^{(l)} \triangleq \left(\boldsymbol{\Sigma}_0^{(l)}, \boldsymbol{\Sigma}_1^{(l)}, \ldots, \boldsymbol{\Sigma}_{K-1}^{(l)}\right) \tag{8.52}$$

and the $L \times 2$ matrix $\boldsymbol{\Sigma}_k^{(l)}$, $k = 0, \ldots, K-1$, is constructed from the user signatures as follows:

$$\boldsymbol{\Sigma}_k^{(l)} \triangleq \begin{pmatrix} (1-\delta_k')[\underbrace{0, \ldots, 0}_{n_k'}, s_k(0), \ldots, s_k(L-n_k'-1)]^T + \\ \delta_k'[\underbrace{0, \ldots, 0}_{n_k'-1}, s_k(0), \ldots, s_k(L-n_k')]^T \\ (1-\delta_k')[s_k(L-n_k'), \ldots, s_k(L-1), \underbrace{0, \ldots, 0}_{L-n_k'}]^T + \\ \delta_k'[s_k(L-n_k'+1), \ldots, s_k(L-1), \underbrace{0, \ldots, 0}_{L-n_k'-1}]^T \end{pmatrix} \tag{8.53}$$

with

$$n_k' = \begin{cases} l - n_k & \text{if } l \le n_k \\ L + (l - n_k) & \text{if } l > n_k \end{cases} \quad \text{and} \quad \delta_k' = \begin{cases} \delta_k & \text{if } l \le n_k \\ 1 - \delta_k & \text{if } l > n_k \end{cases} . \tag{8.54}$$

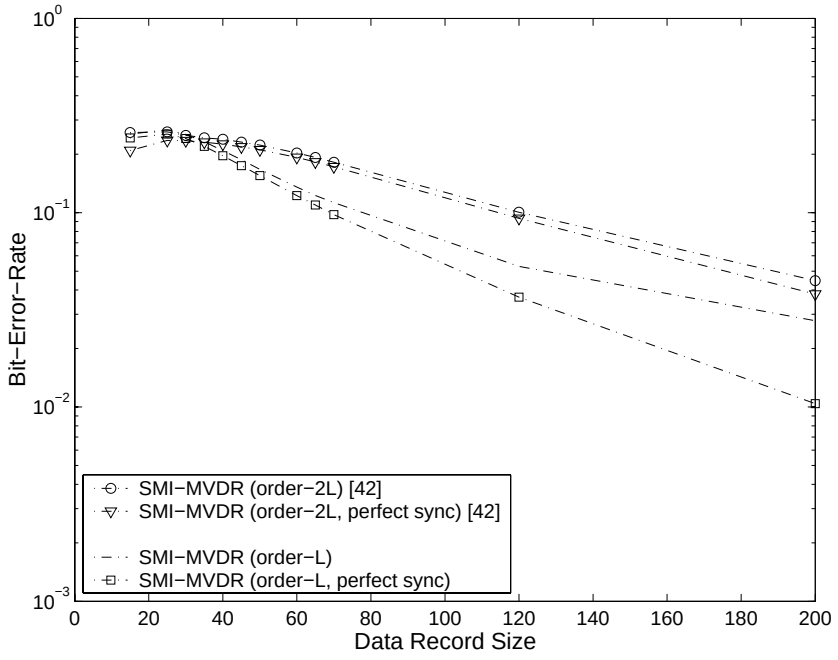

Fig. 8.5. Bit-error-rate as a function of the data record size for diagonally loaded SMI-MVDR receivers of order-L and order-$2L$ (the BER performance of their perfectly synchronized counterparts is also included for reference).

Let $\lambda_1^{(l)}, \ldots, \lambda_\rho^{(l)}$ be the non-zero eigenvalues of $\mathbf{\Sigma}^{(l)}\mathbf{\Sigma}^{(l)^T}$ and $\mathbf{v}_1^{(l)}, \ldots, \mathbf{v}_\rho^{(l)}$ their corresponding orthonormal eigenvectors. Then, we may express $\mathbf{R}^{(l)}$ as follows:

$$\mathbf{R}^{(l)} = \sigma^2\mathbf{I} + E\mathbf{V}^{(l)}\mathbf{\Lambda}^{(l)}\mathbf{V}^{(l)^T} \tag{8.55}$$

where $\mathbf{\Lambda}^{(l)} = diag[\lambda_1^{(l)}, \ldots, \lambda_\rho^{(l)}]$ and $\mathbf{V}^{(l)} = [\mathbf{v}_1^{(l)}, \ldots, \mathbf{v}_\rho^{(l)}]$. The inverse of $\mathbf{R}^{(l)}$ is given below [47]:

$$\mathbf{R}^{(l)^{-1}} = \frac{1}{\sigma^2}\left[\mathbf{I} - \mathbf{V}^{(l)}\left(\frac{\sigma^2}{E}\mathbf{\Lambda}^{(l)^{-1}} + \mathbf{I}\right)^{-1}\mathbf{V}^{(l)^T}\right]. \tag{8.56}$$

Assuming without loss of generality that $\|\mathbf{s}_0^*\|^2 = 1$ we have

$$\mathbf{s}_0^{*^T}\mathbf{R}^{(l)^{-1}}\mathbf{s}_0^* = \frac{1}{\sigma^2}\left[1 - \mathbf{s}_0^{*^T}\mathbf{V}^{(l)}\left(\frac{\sigma^2}{E}\mathbf{\Lambda}^{(l)^{-1}} + \mathbf{I}\right)^{-1}\mathbf{V}^{(l)^T}\mathbf{s}_0^*\right]. \tag{8.57}$$

The above expression implies that minimization of $\mathbf{s}_0^{*^T}\mathbf{R}^{(l)^{-1}}\mathbf{s}_0^*$ is equivalent to maximization of $\mathbf{s}_0^{*^T}\mathbf{V}^{(l)}(\frac{\sigma^2}{E}\mathbf{\Lambda}^{(l)^{-1}} + 1)^{-1}\mathbf{V}^{(l)^T}\mathbf{s}_0^*$. But,

$$\lim_{E\to\infty}\left[\mathbf{s}_0^{*^T}\mathbf{V}^{(l)}(\frac{\sigma^2}{E}\mathbf{\Lambda}^{(l)^{-1}} + 1)^{-1}\mathbf{V}^{(l)^T}\mathbf{s}_0^*\right] = \mathbf{s}_0^{*^T}\mathbf{V}^{(l)}\mathbf{V}^{(l)^T}\mathbf{s}_0^* = ||\mathbf{V}^{(l)^T}\mathbf{s}_0^*||^2 \tag{8.58}$$

and

$$||\mathbf{V}^{(l)^T}\mathbf{s}_0^*||^2 \leq ||\mathbf{s}_0^*||^2 = 1 = ||\mathbf{V}^{(n_0)^T}\mathbf{s}_0^*||^2. \tag{8.59}$$

Thus, there exists an $E^* > 0$ such that for $E > E^*$ $\arg\max_l \mathbf{s}_0^{*^T}\mathbf{R}^{(l)^{-1}}\mathbf{s}_0^* = n_0$ if and only if $||\mathbf{V}^{(l)^T}\mathbf{s}_0^*||^2 < 1$ for $l \neq n_0$. Indeed, the latter condition is satisfied if and only if $\mathbf{s}_0^*$ does not lie in the subspace spanned by the columns of $\mathbf{\Sigma}^{(l)}$ for $l \neq n_0$. □

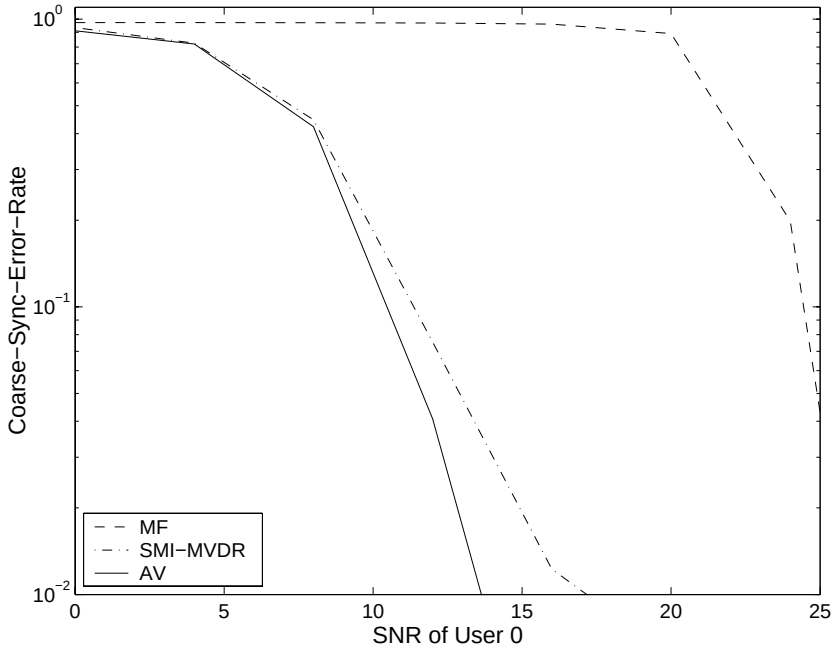

Fig. 8.6. Coarse-synchronization-error-rate as a function of the SNR of the user of interest for linear receivers of order-L (sample support $N = 55$).

Proof of Proposition 8.2

(i) We observe that

$$\alpha_0(\delta)\mathbf{w}_{\mathbf{d}_0,MVDR}^{(\hat{n}_0)} = (1-2\delta)\mathbf{R}^{(\hat{n}_0)^{-1}}\mathbf{s}_0, \tag{8.60}$$

$$\alpha_1(\delta)\mathbf{w}_{\mathbf{d}_1,MVDR}^{(\hat{n}_0)} = 2\delta\mathbf{R}^{(\hat{n}_0)^{-1}}(0.5\mathbf{s}_0 + 0.5\mathbf{s}_0^{(+1)}). \tag{8.61}$$

Then,

$$\alpha_0(\delta)\mathbf{w}_{\mathbf{d}_0,MVDR}^{(\hat{n}_0)} + \alpha_1(\delta)\mathbf{w}_{\mathbf{d}_1,MVDR}^{(\hat{n}_0)} = \mathbf{R}^{(\hat{n}_0)^{-1}}\mathbf{s}_{\|0\|}^*(\delta). \tag{8.62}$$

So, the filters $\mathbf{w} \in \mathcal{W}$ are of the form

$$\mathbf{w} = \frac{\mathbf{R}^{(\hat{n}_0)^{-1}}\mathbf{s}_{\|0\|}^*(\delta)}{\mathbf{s}_{\|0\|}^{*T}(\delta)\mathbf{R}^{(\hat{n}_0)^{-1}}\mathbf{s}_{\|0\|}^*(\delta)}. \tag{8.63}$$

(ii) (8.63) implies that the maximization in (8.34) over all $\delta \in [0,1)$ is equivalent to maximization of $\mathbf{w}^T\mathbf{R}^{(\hat{n}_0)}\mathbf{w}$ over all $\mathbf{w} \in \mathcal{W}$. □

Proof of Proposition 8.3

From the partition matrix inversion formula [47] we have

$$\widehat{\mathbf{R}}^{(l+1)^{-1}} = \begin{pmatrix} \left(\mathbf{C} - \frac{1}{a_{l+1}}\mathbf{b}_{l+1}\mathbf{b}_{l+1}^T\right)^{-1} & (\mathbf{b}_{l+1}^T\mathbf{C}^{-1}\mathbf{b}_{l+1} - a_{l+1})^{-1}\mathbf{C}^{-1}\mathbf{b}_{l+1} \\ (\mathbf{b}_{l+1}^T\mathbf{C}^{-1}\mathbf{b}_{l+1} - a_{l+1})^{-1}\mathbf{b}_{l+1}^T\mathbf{C}^{-1} & (a_{l+1} - \mathbf{b}_{l+1}^T\mathbf{C}^{-1}\mathbf{b}_{l+1})^{-1} \end{pmatrix}, \tag{8.64}$$

while from the matrix inversion lemma we have

$$\left(\mathbf{C} - \frac{1}{a_{l+1}}\mathbf{b}_{l+1}\mathbf{b}_{l+1}^T\right)^{-1} = \mathbf{C}^{-1} + \frac{\mathbf{C}^{-1}\mathbf{b}_{l+1}\mathbf{b}_{l+1}^T\mathbf{C}^{-1}}{a_{l+1} - \mathbf{b}_{l+1}^T\mathbf{C}^{-1}\mathbf{b}_{l+1}}. \tag{8.65}$$

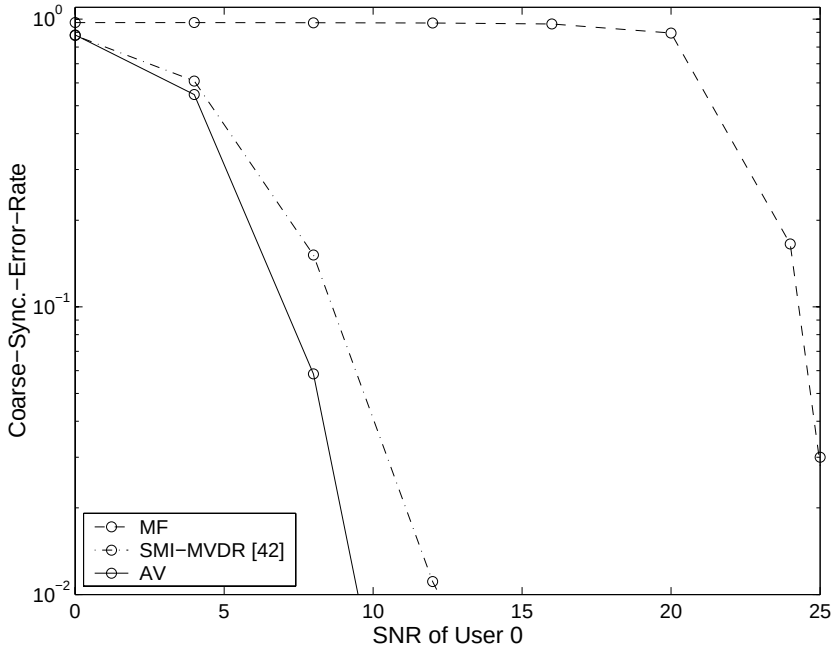

Fig. 8.7. Coarse-synchronization-error-rate as a function of the SNR of the user of interest for linear receivers of order-$2L$ (sample support $N = 85$).

Partition $\widehat{\mathbf{R}}^{(l)}$ as follows:

$$\widehat{\mathbf{R}}^{(l)} = \begin{pmatrix} a_l & \mathbf{b}_l^T \\ \mathbf{b}_l & \mathbf{C} \end{pmatrix} \tag{8.66}$$

where a_l is a scalar and $\mathbf{b}_l$ is an $(L-1) \times 1$ vector. Then, application of the partition matrix inversion formula and the matrix inversion lemma leads to the following expressions for $\mathbf{\Gamma}$, α, and $\boldsymbol{\beta}$ in (8.50):

$$\mathbf{\Gamma} = \mathbf{C}^{-1} + \frac{\mathbf{C}^{-1}\mathbf{b}_l\mathbf{b}_l^T\mathbf{C}^{-1}}{a_l - \mathbf{b}_l^T\mathbf{C}^{-1}\mathbf{b}_l}, \tag{8.67}$$

$$\alpha = \left(a_l - \mathbf{b}_l^T\mathbf{C}^{-1}\mathbf{b}_l\right)^{-1}, \tag{8.68}$$

$$\boldsymbol{\beta} = -\frac{1}{a_l}\left(1 + \frac{\mathbf{b}_l^T\mathbf{C}^{-1}\mathbf{b}_l}{a_l - \mathbf{b}_l^T\mathbf{C}^{-1}\mathbf{b}_l}\right)\mathbf{C}^{-1}\mathbf{b}_l. \tag{8.69}$$

But, $\mathbf{b}_l^T\mathbf{C}^{-1}\mathbf{b}_l = a_l - \alpha^{-1}$. Hence, $\beta = -\alpha\mathbf{C}^{-1}\mathbf{b}_l$ which implies

$$\mathbf{C}^{-1}\mathbf{b}_l = -\frac{\boldsymbol{\beta}}{\alpha}. \tag{8.70}$$

From (8.67) and (8.70) we obtain

$$\mathbf{C}^{-1} = \mathbf{\Gamma} - \frac{1}{\alpha}\boldsymbol{\beta}\beta^T. \tag{8.71}$$

□

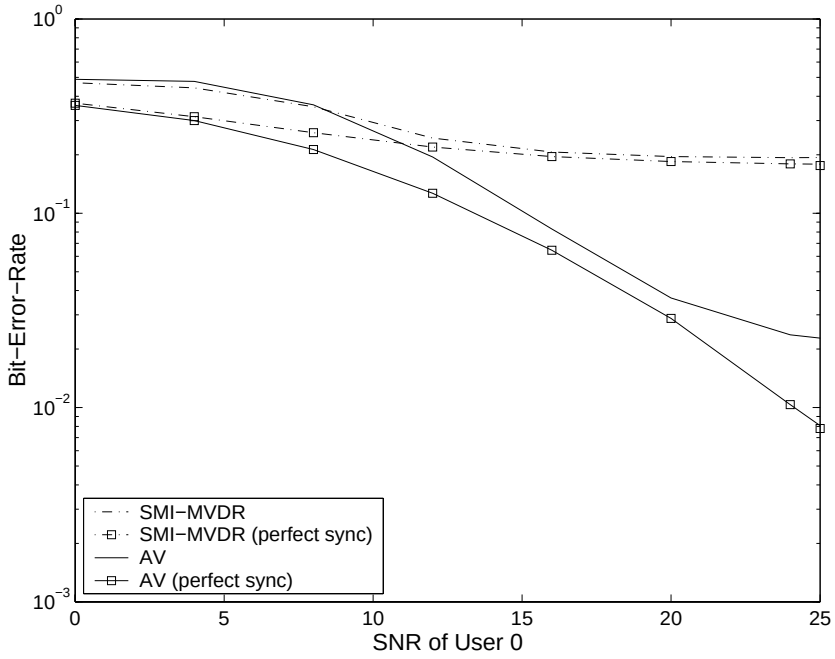

Fig. 8.8. Bit-error-rate as a function of the SNR of the user of interest for linear receivers of order-L (the BER performance of their perfectly synchronized counterparts is also included for reference). The sample support $N = 55$.

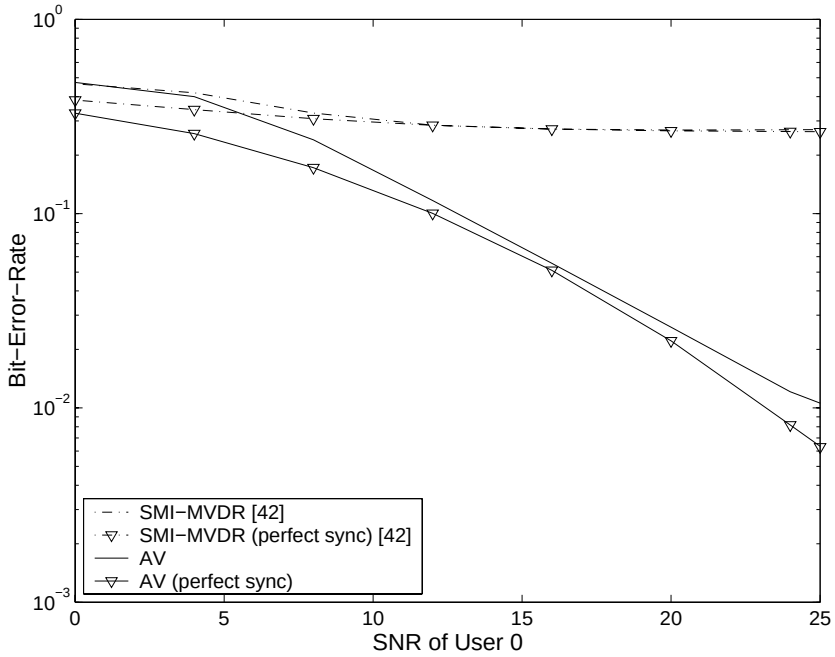

Fig. 8.9. Bit-error-rate as a function of the SNR of the user of interest for linear receivers of order-$2L$ (the BER performance of their perfectly synchronized counterparts is also included for reference). The sample support is $N = 85$.

Bibliography

[1] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 32, pp. 85-96, Jan. 1986.

[2] K. S. Schneider, "Optimum detection of code division multiplexed signals," *IEEE Trans. Aerospace Electron. Syst.*, vol. 15, pp. 181-185, Jan. 1979.

[3] R. Lupas and S. Verdu, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 123-136, Jan. 1989.

[4] M. K. Varanasi and B. Aazhang, "Multistage detection in asynchronous code-division multiple-access communications," *IEEE Trans. Commun.*, vol. 38, pp. 509-519, Apr. 1990.

[5] M. K. Varanasi and B. Aazhang, "Near-optimum detection in synchronous code-division multiple-access systems," *IEEE Trans. Commun.*, vol. 39, pp. 725-736, May 1991.

[6] A. Duel-Hallen, "A family of multiuser decision-feedback detectors for asynchronous code-division multiple access channels," *IEEE Trans. Commun.*, vol. 43, pp. 421-434, Feb./Mar./Apr. 1995.

[7] B. Aazhang, B. P. Paris, and G. Orsak, "Neural networks for multiuser detection in CDMA communication," *IEEE Trans. Commun.*, vol. 40, pp. 1212-1222, July 1992.

[8] U. Mitra and H. V. Poor, "Adaptive receiver algorithms for near-far resistant CDMA," *IEEE Trans. Commun.*, vol. 43, pp. 1713-1724, Feb./Mar./Apr. 1995.

[9] U. Mitra and H. V. Poor, "Neural network techniques for adaptive multiuser demodulation," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 1460-1470, Dec. 1994.

[10] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, pp. 3178-3188, Dec. 1994.

[11] Z. Xie, R. T. Short, and C. K. Rushforth. "A family of suboptimum detectors for coherent multiuser communications", *IEEE J. Select. Areas Commun.*, vol. 8, pp. 683-690, May 1990.

[12] C. Pateros and G. Saulnier, "An adaptive correlator receiver for direct-sequence spread-spectrum communications," *IEEE Trans. Commun.*, vol. 44, pp. 1543-1552, Nov. 1996.

[13] S. L. Miller, "Adaptive direct-sequence code-division multiple-access receiver for multiuser interference rejection," *IEEE Transactions on Communications*, vol. 43, pp. 1746-1755, Feb./Mar./Apr. 1995.

[14] M. L. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944-960, July 1995.

[15] D. A. Pados and S. N. Batalama, "Low-complexity blind detection of DS/CDMA signals: Auxiliary vector receivers," *IEEE Trans. Commun.*, vol. 45, pp. 1586–1594, Dec. 1997.

[16] A. Kansal, S. N. Batalama, and D. A. Pados, "Adaptive maximum SINR RAKE filtering for DS-CDMA multipath fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1765-1773, Dec. 1998.

[17] D. A. Pados and S. N. Batalama, "Joint space-time auxiliary-vector filtering for DS/CDMA systems with antenna arrays," *IEEE Trans. Commun.*, vol. 47, pp. 1406-1415, Sept. 1999.

[18] S. N. Batalama, M. Medley, and D. A. Pados, "Robust adaptive recovery of spread-spectrum signals with short data records," *IEEE Trans. Commun.*, vol. 48, pp. 1725–1731, Oct. 2000.

[19] D. A. Pados and G. N. Karystinos, "An iterative algorithm for the computation of the MVDR filter," *IEEE Trans. Signal Proc.*, vol. 49, pp. 290-300, Feb. 2001.

[20] P. M. Hopkins, "A unified analysis of pseudonoise synchronization by envelope correlation," *IEEE Trans. Commun.*, vol. 25, pp. 770-778, Aug. 1977.

[21] U. Madhow and M. B. Pursley, "Acquisition in direct-sequence spread-spectrum communication networks: An asymptotic analysis," *IEEE Trans. Inform. Theory*, vol. 39, pp. 903-912, May 1993.

[22] J. S. Yamamoto and R. Kohno, " Dynamic digital matched filter acquisition of DS receiver," in *Proc. IEEE Intern. Symposium on Spread Spectrum Techniques & Applications*, Mainz, Germany, Sept. 1996, vol. 2, pp. 751-755.

[23] L. B. Milstein, and R. Rick, "Parallel acquisition in mobile DS-CDMA systems," *IEEE Trans. Commun.*, vol. 45, pp. 1466-1476, Nov. 1997.

[24] S. E. Bensley and B. Aazhang, "Maximum-likelihood synchronization of a single user for code-division multiple-access communication systems," *IEEE Trans. Commun.*, vol. 46, pp. 392-399, Mar. 1998.

[25] R. F. Smith and S. L. Miller, "Acquisition performance of an adaptive receiver for a DS/CDMA system," *IEEE Trans. Commun.*, vol. 49, pp. 1416-1424, Sept. 1998.

[26] Z. Liu, J. Li, and S. L. Miller, "An efficient code-timing estimator for receiver diversity DS-CDMA Systems," *IEEE Trans. Commun.*, vol. 46, pp. 826-835, Jun. 1996.

[27] U. Madhow and M. B. Pursley, "Mathematical modeling and performance analysis for a two-stage acquisition scheme for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 43, pp. 2511-2520, Sept. 1995.

[28] E. G. Strom, S. Parkvall, S. L. Miller, and B. Ottersten, "Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems," *IEEE Trans. Commun.*, vol. 44, pp. 84-93, Jan. 1996.

[29] R. A. Iltis, "Demodulation and code acquisition using decorrelator detectors for QS-CDMA," *IEEE Trans. Commun.*, vol. 44, pp. 1553-1560, Nov. 1996.

[30] R. De Gaudenzi, F. Giannetti, and M. Luise, "Signal recognition and signature code acquisition in CDMA mobile packet communications," *IEEE Trans. Vehic. Tech.*, vol. 47, pp. 196-207, Feb. 1998.

[31] M. Latva-Aho, J. Lilleberg, J. Iinatti, and M. Juntti, "CDMA downlink code acquisition performance in frequency-selective fading channels," in *Proc. IEEE Inter. Symposium on Personal, Indoor and Mobile Radio Commun. (PIMRC'98)*, Boston, MA, Sept. 1998, pp. 1476-1479.

[32] M. Latva-Aho and J. Lilleberg, "Delay trackers for multiuser CDMA receivers," in *1996 Annual Intern. Conf. on Universal Personal Commun.*, Boston, MA, Sept. 1996, pp. 326-330.

[33] S. E. Bensley and B. Aazhang, "Subspace-based channel estimation for code division multiple access communication systems," *IEEE Trans. Commun.*, vol. 44, pp. 1009-1020, Aug. 1996.

[34] B. Steiner and P. Jung, "Uplink channel estimation in synchronous CDMA mobile radio systems with joint detection," in *Proc. IEEE Inter. Symposium on Personal, Indoor and Mobile Radio Commun. (PIMRC'93)*, Yokohama, Japan, Sept. 1993, pp. 123-127.

[35] B. Steiner and P. Jung, "Optimum and suboptimum channel estimation for the uplink of CDMA mobile radio systems with joint detection," *European Trans. on Telecomm.*, vol. 5, pp. 39-50, Jan./Feb. 1994.

[36] M. Torlak and G. Xu, "Blind multiuser channel estimation in asynchronous CDMA systems," *IEEE Trans. Signal Proc.*, vol. 45, pp. 137-147, Jan. 1997.

[37] M. Missiroli, Y. J. Guo, and S. K. Barton, "Near-far resistant channel estimation for CDMA systems using the linear decorrelating detector," *IEEE Trans. Commun.*, vol. 48, pp. 514-524, Mar. 2000.

[38] S. Parkvall, E. G. Strom, L. B. Milstein, and B. E. Ottersten, "Asynchronous near-far resistant DS-CDMA receivers without a priori synchronization," *IEEE Trans. Commun.*, vol. 47, pp. 78-88, Jan. 1999.

[39] R. A. Iltis, and L. Mailender, "An adaptive multiuser detector with joint amplitude and delay estimation," *IEEE J. Select. Areas. Commun.*, vol. 12, pp.774-785, Jun. 1994.

[40] T. J. Lim, and L. K. Rasmussen, "Adaptive symbol and parameter estimation in asynchronous multiuser CDMA detectors," *IEEE Trans. Commun.*, vol. 45, pp. 213-220, Feb. 1997.

[41] A. Radovic, and B. Aazhang, "Iterative algorithms for joint data detection and delay estimation for code division multiple access communication systems," in *Proc. 31st Annual Allerton Conf. on Commun., Control, and Computing*, Monticello, IL, Sept. 1993.

[42] U. Madhow, "Blind adaptive interference suppression for the near-far resistant acquisition and demodulation of direct-sequence CDMA signals," *IEEE Trans. Signal Proc.*, vol. 45, pp. 124-136, Jan. 1997.

[43] U. Madhow, "MMSE interference suppression for timing acquisition and demodulation in direct-sequence CDMA systems," *IEEE Trans. Commun.*, vol. 46, pp. 1065-1075, Aug. 1998.

[44] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 1996.

[45] H. Qian and S. N. Batalama, "Data-record-based criteria for the selection of an auxiliary-vector estimator of the MVDR filter," in *Proc. Thirty-fourth Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2000, pp. 802-807.

[46] H. Qian and S. N. Batalama, "Data-record-based criteria for the selection of an auxiliary-vector estimator of the MVDR filter," *IEEE Trans. Commun.*, submitted.

[47] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 5th ed. New York: Cambridge Univ. Press, 1991.

Chapter 9

Rapid Combined Synchronization/Demodulation Structures for DS/CDMA Systems, Part II: Finite Data Record Performance Analysis

9.1 Introduction

The effectiveness of a receiver designed for rapidly changing wireless direct-sequence code-division-multiple-access (DS/CDMA) communication environments depends on the following design attributes: *(i)* low computational complexity, *(ii)* multiple-access-interference (MAI) near-far resistance, and *(iii)* system adaptivity with superior performance under limited data support. Adaptive short-data-record designs appear as the natural next step [1]-[4] to a matured discipline that has extensively addressed the first two design objectives in ideal setups (perfectly known or asymptotically estimated statistical properties) [5], [6]. System adaptivity based on short data records is necessary for the development of practical receivers that exhibit superior bit-error-rate performance when they operate in rapidly changing communication environments that limit substantially the available data support.

In Chapter 8 we considered self-synchronized receivers (integrated synchronizers/demodulators) and we presented three schemes along the classical receiver evolution path. We started with a matched-filter-type (MF) structure, we continued with a minimum-variance-distortionless-response-type (MVDR) scheme and, finally, we developed an auxiliary-vector-type (AV) alternative, all of order either twice or equal to the system spreading gain L. Simulation studies illustrated the performance of the proposed structures in terms of bit-error-rate (BER) versus data support for a given signal-to-noise ratio (SNR) of the user of interest or in terms of BER versus SNR for a given data record size.

In this chapter, we develop three methods for the analysis of the finite data record behavior of self-synchronized linear receivers and in particular the effect of the filter order and the data-record-size on the coarse-synchronization-error-rate (CSER). We derive analytic expressions that approximate closely the probability of coarse synchronization error of MF-type and MVDR-type schemes. Probability of coarse synchronization error analysis for AV-type schemes based on a finite data-record-size is prohibitively complex and thus not attempted at this time.

The chapter is organized as follows. In Section 9.2, we present the system model and we summarize briefly the combined synchronization/demodulation algorithms developed in Chapter 8. In Section 9.3, we develop analytic expressions that approximate the coarse synchronization error probability (P_{cse}) while in Section 9.4 we derive a sequence of increasingly tight lower bounds on P_{cse}. In Section 9.5 we discuss the effect of the filter order and the data-record-size on the coarse synchronization error rate. The accuracy of

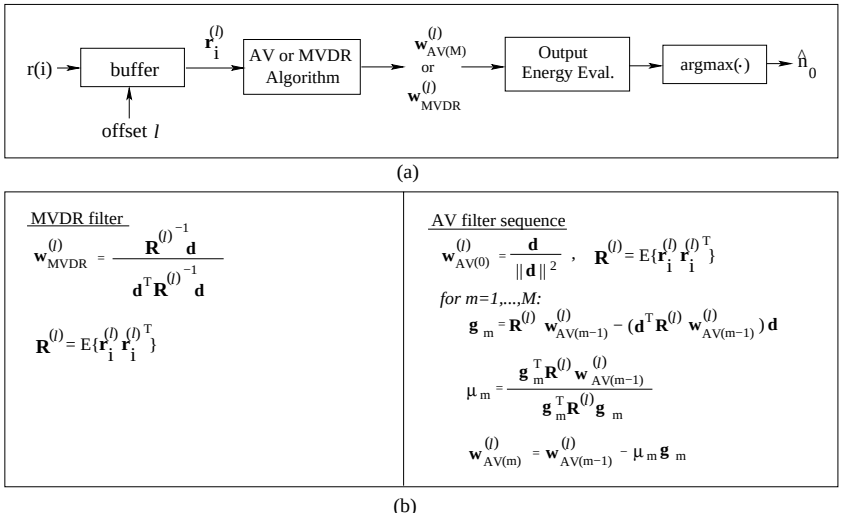

Fig. 9.1. Coarse synchronization based on filters of order L.

the analytic approximations of Sections 9.3 and 9.4 is examined through simulations in Section 9.6. Some final conclusions are drawn in Section 9.7.

9.2 System Model and Algorithmic Description

We consider an asynchronous DS/CDMA system populated by K active users transmitting over a common additive white Gaussian noise (AWGN) channel. The received signal $r(t)$ is the superposition of the K transmissions corrupted by channel noise, i.e.

$$r(t) = \sum_{k=0}^{K-1} \sum_{i=-\infty}^{\infty} \sqrt{E_k} b_k(i) S_k(t - iT - \tau_k) + n(t). \tag{9.1}$$

In the above expression, with respect to the kth user τ_k is the signal delay relative to the receiver's time reference point, $b_k(i) \in \{-1, +1\}$ is the ith information bit, E_k is the transmitted energy, T is the information bit period, and $n(t)$ is AWGN. The normalized signature waveform $S_k(t)$ assigned to the kth user has the form

$$S_k(t) = \sum_{l=0}^{L-1} s_k(l) P_{Tc}(t - lT_c), \quad k = 0, \ldots, K-1, \tag{9.2}$$

where T_c is the chip duration, $P_{Tc}(\cdot)$ is a rectangular pulse with support $[0, T_c]$, $L = T/T_c$ is the system spreading gain, and $s_k(l)$ is the lth element of the normalized bipolar signature or spreading vector $\mathbf{s}_k \stackrel{\triangle}{=} [s_k(0), s_k(1), \ldots, s_k(L-1)]^T$ that uniquely identifies the kth user. Without loss of generality, we assume that $\tau_k \in [0, T),\ k = 0, \ldots, K-1$. Thus, we may write the delay τ_k as a sum of an integer multiple of chips plus a fraction of a chip: $\tau_k = (n_k + \delta_k)T_c$ where $n_k \in \{0, \ldots, L-1\}$ and $\delta_k \in [0, 1)$. After chip-matched filtering and chip-rate sampling, the continuous-time signal $r(t)$ in (9.1) is converted to a discrete-time sequence $\{r(n)\}_{n=-\infty}^{\infty}$ where $r(n) = \int_{nT_c}^{(n+1)T_c} r(t)dt$.

In this work we analyze the coarse synchronization performance of L and $2L$-order algorithms of MF and sample-matrix-inversion (SMI) MVDR type. A brief description of the coarse synchronization step of these algorithms follows (details can be found in Chapter 8).

Fig. 9.1 shows the block diagram representation of an L-order coarse synchronizer/demodulator. The buffer follows a chip-rate sampler and groups the received samples into L sequences of vectors, $\{\mathbf{r}_i^{(l)}\}_i \in \Re^L,\ l = 0, \ldots, L-1$, as depicted in Fig. 9.2 where

$$\mathbf{r}_i^{(l)} = [r(iL + l), r(iL + l + 1), \ldots, r((i+1)L + l - 1)]^T, \quad l = 0, \ldots, L-1. \tag{9.3}$$

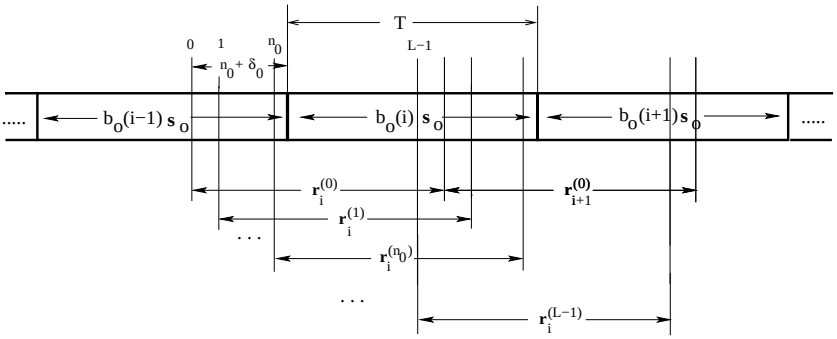

Fig. 9.2. Construction of the sequences $\mathbf{r}^{(l)}, l = 0, \ldots, L-1$.

In the case of $2L$-order algorithms, the received samples $\{r(n)\}_n$ are grouped into a sequence of overlapping vectors, $\bar{\mathbf{r}}_i \in \Re^{2L}$, obtained by a sliding window such that successive vectors share L common data samples:

$$\bar{\mathbf{r}}_i = [r(iL), r(iL+1), \ldots, r((i+2)L-1)]^T. \tag{9.4}$$

To facilitate our presentation overlined variables always refer to $2L$-order processing to distinguish themselves from the corresponding variables used by L-order algorithms. Table 9.1 provides the definition of both the ideal and estimated decision statistic (filter output energy) of each filter under consideration, namely SMI-MVDR and MF of order L or $2L$. Table 9.2 provides additional definitions needed in the expressions of Table 9.1.

For the L-order case of Table 9.1, $\mathbf{R}^{(l)} \triangleq E\{\mathbf{r}_i^{(l)} \mathbf{r}_i^{(l)^T}\}$ is the input covariance matrix and $\widehat{\mathbf{R}}^{(l)} \triangleq \frac{1}{N}\sum_{i=0}^{N-1} \mathbf{r}_i^{(l)} \mathbf{r}_i^{(l)^T}$ is its sample average estimate (N is the total number of samples or data record size). Similarly, for the $2L$-order case, $\overline{\mathbf{R}} \triangleq E\{\bar{\mathbf{r}}_i \bar{\mathbf{r}}_i^T\}$ and $\widehat{\overline{\mathbf{R}}} \triangleq \frac{1}{N}\sum_{i=0}^{N-1} \bar{\mathbf{r}}_i \bar{\mathbf{r}}_i^T$. In Table 9.2, $\mathbf{s}_0^{(+1)} \triangleq [0, s_0(0), \ldots, s_0(L-2)]^T$ denotes the one-chip right-shifted zero-filled version of $\mathbf{s}_0$, while $\bar{\mathbf{s}}_0^{(l)}$, $l = 0, \ldots, L-1$, is the l-shifted version of the zero-padded $2L$-long vector $[\mathbf{s}_0^T, 0, \ldots, 0]^T$.

An SMI-MVDR-type or MF-type coarse estimate of the delay τ_0 can be determined as follows:

$$\hat{\tau}_0 = \begin{cases} (\lfloor j_{\max}/2 \rfloor) T_c, & \text{if } j_{\max} \text{ is even} \\ (\lfloor j_{\max}/2 \rfloor + 0.5) T_c, & \text{if } j_{\max} \text{ is odd} \end{cases} \tag{9.5}$$

where

$$j_{\max} = \arg\max_j \left\{ \widehat{U}_j \;\middle|\; j = 0, \ldots, 2L-1 \right\} \tag{9.6}$$

and $\widehat{U}_j$ is obtained from Table 9.1 for each filter respectively.

9.3 Finite Data Record Performance Analysis

The probability of coarse synchronization error is defined as the probability that the coarse synchronization algorithm fails to provide an estimate within the "pull-in" range. In this chapter, we adopt the commonly used assumption that the pull-in range is equal to $\frac{1}{2}T_c$ around the correct delay τ_0. However, the performance analysis presented in this section can be carried out with straightforward modifications for any given pull-in range value. The probability of coarse synchronization error is given by

$$P_{cse} \triangleq 1 - Pr\left[|\hat{\tau}_0 - \tau_0| < \frac{1}{2} T_c \right], \tag{9.7}$$

where the coarse estimate of $\hat{\tau}_0$ is given by (9.5), (9.6). We recall that $j_{\max}$ in (9.5), (9.6) is the index of the filter with the highest output energy. If we define $\mathcal{H}(\tau_0)$ as the set of the filter indices that yield a timing estimate $\hat{\tau}_0$ within $\frac{1}{2}T_c$ about τ_0, then the probability of coarse synchronization error will be equal to

$$P_{cse} = 1 - Pr\left[j_{\max} \in \mathcal{H}(\tau_0) \right]. \tag{9.8}$$

Table 9.1. Decision Statistic

L-order	
SMI-MVDR	
$U_j = \left(\mathbf{d}_{\|0\|}^{(j)^T} \mathbf{R}^{(\lfloor j/2 \rfloor)^{-1}} \mathbf{d}_{\|0\|}^{(j)}\right)^{-1}$	$\widehat{U}_j = \left(\mathbf{d}_{\|0\|}^{(j)^T} \widehat{\mathbf{R}}^{(\lfloor j/2 \rfloor)^{-1}} \mathbf{d}_{\|0\|}^{(j)}\right)^{-1}$
MF	
$U_j = \mathbf{d}_{\|0\|}^{(j)^T} \mathbf{R}^{(\lfloor j/2 \rfloor)} \mathbf{d}_{\|0\|}^{(j)}$	$\widehat{U}_j = \mathbf{d}_{\|0\|}^{(j)^T} \widehat{\mathbf{R}}^{(\lfloor j/2 \rfloor)} \mathbf{d}_{\|0\|}^{(j)}$
$2L$-order	
SMI-MVDR	
$U_j = \left(\bar{\mathbf{d}}_{\|0\|}^{(j)^T} \bar{\mathbf{R}}^{-1} \bar{\mathbf{d}}_{\|0\|}^{(j)}\right)^{-1}$	$\widehat{U}_j = \left(\bar{\mathbf{d}}_{\|0\|}^{(j)^T} \widehat{\bar{\mathbf{R}}}^{-1} \bar{\mathbf{d}}_{\|0\|}^{(j)}\right)^{-1}$
MF	
$U_j = \bar{\mathbf{d}}_{\|0\|}^{(j)^T} \bar{\mathbf{R}} \bar{\mathbf{d}}_{\|0\|}^{(j)}$	$\widehat{U}_j = \bar{\mathbf{d}}_{\|0\|}^{(j)^T} \widehat{\bar{\mathbf{R}}} \bar{\mathbf{d}}_{\|0\|}^{(j)}$

Table 9.2. Effective signature

L-order	
$\mathbf{d}_{\|0\|}^{(j)} \triangleq \begin{cases} \mathbf{s}_{\|0\|}(0), & \text{if } j \text{ is even} \\ \mathbf{s}_{\|0\|}(0.5) & \text{if } j \text{ is odd,} \end{cases}$	$\mathbf{s}_{\|0\|}(\delta) \triangleq \frac{(1-\delta)\mathbf{s}_0 + \delta \mathbf{s}_0^{(+1)}}{\|\|(1-\delta)\mathbf{s}_0 + \delta \mathbf{s}_0^{(+1)}\|\|}$
$2L$-order	
$\bar{\mathbf{d}}_{\|0\|}^{(j)} \triangleq \begin{cases} \bar{\mathbf{s}}_{\|0\|}^{(\lfloor j/2 \rfloor)}(0), & \text{if } j \text{ is even} \\ \bar{\mathbf{s}}_{\|0\|}^{(\lfloor j/2 \rfloor)}(0.5) & \text{if } j \text{ is odd,} \end{cases}$	$\bar{\mathbf{s}}_{\|0\|}^{(l)}(\delta) \triangleq \frac{(1-\delta)\bar{\mathbf{s}}_0^{(l)} + \delta \bar{\mathbf{s}}_0^{(l+1)}}{\|\|(1-\delta)\bar{\mathbf{s}}_0^{(l)} + \delta \bar{\mathbf{s}}_0^{(l+1)}\|\|}$

Different choices of the "pull-in" range can be accommodated by modifying appropriately the definition of $\mathcal{H}(\tau_0)$. In the rest of this section we derive close approximations of the coarse synchronization performance of the four algorithms under consideration.

9.3.1 MF-type receivers ($2L$-order)

We recall that in this case $U_j = \overline{\mathbf{d}}_{\|0\|}^{(j)^T} \, \overline{\mathbf{R}} \, \overline{\mathbf{d}}_{\|0\|}^{(j)}$ and $\widehat{U}_j = \overline{\mathbf{d}}_{\|0\|}^{(j)^T} \, \widehat{\overline{\mathbf{R}}} \, \overline{\mathbf{d}}_{\|0\|}^{(j)}$, $j = 0, 1, \ldots, 2L-1$. Then, the coarse synchronization error rate $P_{cse,MF}^{(2L)}$ of the $2L$-order MF algorithm can be expressed as the probability that the index of the largest decision variable $\widehat{U}_j$, $j = 0, 1, \ldots, 2L-1$, is not contained in $\mathcal{H}(\tau_0)$; that is,

$$P_{cse,MF}^{(2L)} = 1 - Pr\left[\arg\max_j \left\{ \widehat{U}_j \;\middle|\; j = 0, \ldots, 2L-1 \right\} \in \mathcal{H}(\tau_0)\right] \tag{9.9}$$

$$= 1 - Pr\left[\max\left\{ \widehat{U}_k \;\middle|\; k \in \mathcal{H}(\tau_0)\right\} > \max\left\{ \widehat{U}_j \;\middle|\; j = 0, \ldots, 2L-1, j \notin \mathcal{H}(\tau_0)\right\}\right]. \tag{9.10}$$

The latter form of $P_{cse,MF}^{(2L)}$ is particularly easy to handle as we will see in the sequel. We observe that the evaluation of the joint probability density function (pdf) of the random quantities $\widehat{U}_j$, $j = 0, \ldots, 2L-1$, requires knowledge of the pdf of the random matrix $\widehat{\overline{\mathbf{R}}}$. We recall that $\widehat{\overline{\mathbf{R}}}$ is the sample average estimate of the autocorrelation matrix $\overline{\mathbf{R}}$ of the received vectors $\overline{\mathbf{r}}_i$ which are distributed according to a mixture of 2^{3K} Gaussian distributions and they are, by construction, statistically dependent. Unfortunately, the use of exact closed form expressions for the pdf of the sample covariance matrix that is formed by vector samples drawn from a Gaussian mixture distribution [7] leads to mathematically intractable expressions for the probability of synchronization error. An additional complicating factor for the analysis of the finite data support behavior of the MF-type (and SMI-MVDR-type) receivers is the underlying dependence of the vector samples. To that respect, we make the following simplifying assumptions. We assume that the received vectors are uncorrelated and identically Gaussian $\mathcal{N}(\mathbf{0}, \overline{\mathbf{R}})$ distributed (thus i.i.d.) [8],[9]. The latter Gaussianity postulation can be considered as a wishful approximation of the true distribution, while the former independence postulation can be justified if we consider "guard" bands of size L between the received vectors so that successive input vectors do not share common information bits. The effect of the above assumptions will be examined in the simulations section through comparisons with the exact P_{cse}.

It is known [10] that if $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{N-1} \in \Re^p$ are i.i.d. $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ vectors, then the distribution of the matrix $\mathbf{X} \stackrel{\triangle}{=} \sum_{i=0}^{N-1} \mathbf{x}_i \mathbf{x}_i^T$ is

$$f(\mathbf{X}) = \frac{|\mathbf{X}|^{(N-p-1)/2} \exp\{-\frac{1}{2}\mathrm{tr}\mathbf{\Sigma}^{-1}\mathbf{X}\}}{2^{pN/2}\pi^{p(p-1)/4}|\mathbf{\Sigma}|^{N/2}\prod_{i=1}^{p}\Gamma\left[(N+1-i)/2\right]}, \quad p > N, \tag{9.11}$$

where $\Gamma(\cdot)$ is the Gamma function. The distribution in (9.11) is called Wishart distribution with N degrees of freedom and is denoted by $W_p(\mathbf{\Sigma}; N)$. A random matrix distributed according to (9.11) is called Wishart matrix. In our case, $\widehat{\overline{\mathbf{R}}}$ is approximately Wishart with distribution $W_{2L}(\overline{\mathbf{R}}\,/N; N)$. If we define the matrix

$$\overline{\mathbf{D}} \stackrel{\triangle}{=} (\overline{\mathbf{d}}_{\|0\|}^{(0)}, \overline{\mathbf{d}}_{\|0\|}^{(1)}, \ldots, \overline{\mathbf{d}}_{\|0\|}^{(2L-1)}), \tag{9.12}$$

then the random variables $\widehat{U}_j$, $j = 0, \ldots, 2L-1$, form the main diagonal of $\overline{\mathbf{D}}^T \, \widehat{\overline{\mathbf{R}}} \, \overline{\mathbf{D}}$. For a full rank matrix $\overline{\mathbf{D}}$, $\overline{\mathbf{D}}^T \, \widehat{\overline{\mathbf{R}}} \, \overline{\mathbf{D}}$ is Wishart $W_{2L}(\frac{1}{N}\,\overline{\mathbf{D}}^T\overline{\mathbf{R}}\overline{\mathbf{D}}; N)$ [11] and the main diagonal elements $\widehat{U}_j$, $j = 0, \ldots, 2L-1$, are jointly distributed according to a multivariate Gamma distribution [12]. However, $\overline{\mathbf{D}}$ is not full rank in general and even if it were, it would have been of little help since closed form expressions of the multivariate Gamma pdf do not exist except for some special cases [13]. Thus, appropriate approximations of the joint pdf of the diagonal elements of $\overline{\mathbf{D}}^T \, \widehat{\overline{\mathbf{R}}} \, \overline{\mathbf{D}}$ with tractable multivariate integrals are needed.

Motivated by the transformation noise modeling techniques used in [14], we approximate the joint pdf of the MF decision variables $\widehat{U}_j$, $j = 0, \ldots, 2L-1$, by the joint pdf of a nonlinear transformation of Gaussian random variables Y_j, $j = 0, \ldots, 2L-1$. The nonlinear transformation and the covariance of Y_j,

$j = 0, \ldots, 2L-1$, are chosen such that the outputs of the nonlinear transformations have identical covariances with the corresponding $\widehat{U}_j$, $j = 0, \ldots, 2L-1$, and marginal pdfs that are close approximations to the pdfs of $\widehat{U}_j$, $j = 0, \ldots, 2L-1$. The following lemma identifies this non-linear transformation and the statistics of the random variables Y_j, $j = 0, \ldots, 2L-1$. The proof is included in the Appendix.

Lemma 9.1 *The pdf of $\widehat{U}_j$ can be approximated by the pdf of $U_j\, Y_j^3$, $j = 0, \ldots, 2L-1$, where the random variables Y_j, $j = 0, \ldots, 2L-1$, are uncorrelated Gaussian with mean $1 - \frac{2}{9N}$ and variance $\frac{2}{9N}$.* □

The probability of coarse synchronization error can now be evaluated as follows. Let the random variables Z_j, $j = 0, \ldots, 2L-1$ be defined as

$$Z_j = \frac{9N}{9N-2} U_j^{1/3} Y_j, \quad j = 0, \ldots, 2L-1. \tag{9.13}$$

It is straightforward to show that:

$$\begin{aligned}
&Pr\left[\max\left\{\widehat{U}_k \;\middle|\; k \in \mathcal{H}(\tau_0)\right\} > \max\left\{\widehat{U}_j \;\middle|\; j = 0, \ldots, 2L-1, j \notin \mathcal{H}(\tau_0)\right\}\right] \simeq \\
&Pr\left[\max\left\{U_k Y_k^3 \;\middle|\; k \in \mathcal{H}(\tau_0)\right\} > \max\left\{U_j Y_j^3 \;\middle|\; j = 0, \ldots, 2L-1, j \notin \mathcal{H}(\tau_0)\right\}\right] = \\
&Pr\left[\max\left\{Z_k \mid k \in \mathcal{H}(\tau_0)\right\} > \max\left\{Z_j \mid j = 0, \ldots, 2L-1, j \notin \mathcal{H}(\tau_0)\right\}\right] = \\
&Pr\left[\max\left\{Z_j \mid j = 0, \ldots, 2L-1, j \notin \mathcal{H}(\tau_0)\right\} - \max\left\{Z_k \mid k \in \mathcal{H}(\tau_0)\right\} < 0\right].
\end{aligned} \tag{9.14}$$

The random variables Z_j, $j = 0, \ldots, 2L-1$, are uncorrelated Gaussian (thus, independent) with mean $\mu_j = U_j^{1/3}$ and variance $\sigma_j^2 = \frac{18N}{(9N-2)^2} U_j^{2/3}$. The pdf of Z_j, $j = 0, \ldots, 2L-1$, is

$$f_j(z) = \frac{1}{\sigma_j} g\left(\frac{z-\mu_j}{\sigma_j}\right) \tag{9.15}$$

where $g(z) \triangleq \frac{1}{\sqrt{2\pi}} \exp(-z^2/2)$ is the pdf of a Gaussian random variable with mean 0 and variance 1, while the cumulative distribution function (cdf) of Z_j is

$$\mathcal{F}_j(z) = \mathcal{G}\left(\frac{z-\mu_j}{\sigma_j}\right) \tag{9.16}$$

where $\mathcal{G}(z)$ is the cdf of a Gaussian random variable with mean 0 and variance 1. The cdf of $\max\{Z_j \mid j = 0, \ldots, 2L-1, j \notin \mathcal{H}(\tau_0)\}$ is given by [16]

$$\mathcal{F}_{\max, \setminus \mathcal{H}(\tau_0)}(z) \triangleq \prod_{j=0, j \notin \mathcal{H}(\tau_0)}^{2L-1} \mathcal{F}_j(z), \tag{9.17}$$

where the index "$\setminus \mathcal{H}(\tau_0)$" indicates that the maximum is taken over a set that does not include the elements of $\mathcal{H}(\tau_0)$. Therefore, the corresponding pdf is given by

$$f_{\max, \setminus \mathcal{H}(\tau_0)}(z) \triangleq \sum_{l=0, l \notin \mathcal{H}(\tau_0)}^{2L-1} \left(f_l(z) \prod_{j=0, j \notin \mathcal{H}(\tau_0), j \neq l}^{2L-1} \mathcal{F}_j(z) \right). \tag{9.18}$$

Similarly, the cdf and pdf of $\max\{Z_k \mid k \in \mathcal{H}(\tau_0)\}$ are given by

$$\mathcal{F}_{\max, \mathcal{H}(\tau_0)}(z) \triangleq \prod_{j \in \mathcal{H}(\tau_0)} \mathcal{F}_j(z) \tag{9.19}$$

and

$$f_{\max, \mathcal{H}(\tau_0)}(z) \triangleq \sum_{l \in \mathcal{H}(\tau_0)} \left(f_l(z) \prod_{j \in \mathcal{H}(\tau_0), j \neq l} \mathcal{F}_j(z) \right), \tag{9.20}$$

respectively. Since Z_j, $j=0,\ldots,2L-1$, are independent, the two random variables $\max\{Z_j \,|\, j=0,\ldots,2L-1, j\notin\mathcal{H}(\tau_0)\}$ and $\max\{Z_k \,|\, k\in\mathcal{H}(\tau_0)\}$ are also independent, therefore the pdf of their difference $\max\{Z_j \,|\, j=0,\ldots,2L-1, j\notin\mathcal{H}(\tau_0)\} - \max\{Z_k \,|\, k\in\mathcal{H}(\tau_0)\}$ is given by

$$f_{\text{diff}}(z) \triangleq \int_{-\infty}^{+\infty} f_{\max,\setminus\mathcal{H}(\tau_0)}(z-y) f_{\max,\mathcal{H}(\tau_0)}(-y) dy. \tag{9.21}$$

We conclude that the probability of coarse synchronization error for the $2L$-long MF-type coarse synchronization algorithm is

$$\begin{aligned} P_{cse,MF}^{(2L)} &\simeq 1 - \int_{-\infty}^{0} f_{\text{diff}}(z) dz & (9.22)\\ &= 1 - \int_{-\infty}^{0}\int_{-\infty}^{+\infty} f_{\max,\setminus\mathcal{H}(\tau_0)}(z-y) f_{\max,\mathcal{H}(\tau_0)}(-y) dy dz & (9.23)\\ &= 1 - \int_{-\infty}^{0}\int_{-\infty}^{+\infty} \left[\sum_{l=0,l\notin\mathcal{H}(\tau_0)}^{2L-1} \left(f_l(z-y) \prod_{j=0,j\neq l,j\notin\mathcal{H}(\tau_0)}^{2L-1} \mathcal{F}_j(z-y)\right)\right] \\ &\qquad \cdot \left[\sum_{l\in\mathcal{H}(\tau_0)} \left(f_l(-y) \prod_{j\in\mathcal{H}(\tau_0),j\neq l} \mathcal{F}_j(-y)\right)\right] dy\, dz. & (9.24) \end{aligned}$$

Expression (9.24) approximates the probability of coarse synchronization error as a double integral.

9.3.2 MF-type receivers (L-order)

In the case of the MF-type coarse synchronization algorithm that utilizes filters of order L the probability of coarse synchronization error $P_{cse,MF}^{(L)}$ can be evaluated by expression (9.24) where $f_j(\cdot)$ and $F_j(\cdot)$ denote the pdf and cdf, respectively, of Z_j, $j=0,1,...,2L-1$ (Z_j is defined as in eq. (9.13) with the difference that U_j is now given by $U_j = \mathbf{d}_{\|0\|}^{(j)^T} \mathbf{R}^{(\lfloor j/2\rfloor)} \mathbf{d}_{\|0\|}^{(j)}$, $j=0,1,\ldots,2L-1$). The proof follows similar reasoning as in Section 9.3.A with the difference that $\widehat{U}_j = \mathbf{d}_{\|0\|}^{(j)^T} \widehat{\mathbf{R}}^{(\lfloor j/2\rfloor)} \mathbf{d}_{\|0\|}^{(j)}$, $j=0,1,\ldots,2L-1$ and $\widehat{\mathbf{R}}^{(\lfloor j/2\rfloor)}$ is Wishart $W_L(\frac{1}{N}\mathbf{R}^{(\lfloor j/2\rfloor)}; N)$ with N degrees of freedom.

9.3.3 SMI-MVDR-type receivers ($2L$-order)

The MVDR-type SMI-implemented coarse synchronization algorithm that utilizes filters of order $2L$ can be analyzed in a similar manner. In this case, U_j is given by $U_j = \left(\overline{\mathbf{d}}_{\|0\|}^{(j)^T} \overline{\mathbf{R}}^{-1} \overline{\mathbf{d}}_{\|0\|}^{(j)}\right)^{-1}$, $j=0,1,\ldots,2L-1$. If we define the random variables $Z_j = \frac{9(N-2L)}{9(N-2L)-2} U_j^{1/3} Y_j$, $j=0,1,\ldots,2L-1$, where the random variables Y_j, $j=0,1,\ldots,2L-1$, are uncorrelated Gaussian with mean $1-\frac{2}{9(N-2L)}$ and variance $\frac{2}{9(N-2L)}$, then Z_j, $j=0,1,\ldots,2L-1$ are uncorrelated Gaussian random variables with mean $\mu_j = U_j^{1/3}$ and variance $\sigma_j^2 = U_j^{2/3} \frac{18(N-2L)}{[9(N-2L)-2]^2}$. The probability of coarse synchronization error $P_{cse,MVDR}^{(2L)}$ is given by expression (9.24) where $\mathcal{F}_j(\cdot)$ and $f_j(\cdot)$ denote the cdf and pdf of the random variable Z_j, $j=0,1,\ldots,2L-1$. The proof follows again similar reasoning as in Section 9.3.A with the difference that $\widehat{U}_j = \left(\overline{\mathbf{d}}_{\|0\|}^{(j)^T} \widehat{\overline{\mathbf{R}}}^{-1} \overline{\mathbf{d}}_{\|0\|}^{(j)}\right)^{-1}$, $j=0,1,\ldots,2L-1$ and $\widehat{\overline{\mathbf{R}}}$ is Wishart $W_{2L}(\frac{1}{N}\overline{\mathbf{R}}; N)$ with N degrees of freedom. The latter implies that $\widehat{U}_j$ can be expressed as a multiple of a χ^2 random variable with $N-2L$ degrees of freedom.

9.3.4 SMI-MVDR-type receivers (L-order)

For the MVDR-type SMI-implemented coarse synchronization algorithm that utilizes filters of order L we have $U_j = \left(\mathbf{d}_{\|0\|}^{(j)^T} \mathbf{R}^{(\lfloor j/2\rfloor)^{-1}} \mathbf{d}_{\|0\|}^{(j)}\right)^{-1}$, $j=0,1,\ldots,2L-1$. Then, the probabilty of coarse synchronization

error $P_{cse,MVDR}^{(L)}$ is still given by expression (9.24) where $\mathcal{F}_j(\cdot)$ and $f_j(\cdot)$ denote the cdf and pdf of the random variable $Z_j = \frac{9(N-L)}{9(N-L)-2} U_j^{1/3} Y_j$, $j = 0, \ldots, 2L-1$. The random variables Y_j, $j = 0, 1, \ldots, 2L-1$, are uncorrelated Gaussian with mean $1 - \frac{2}{9(N-L)}$ and variance $\frac{2}{9(N-L)}$ therefore, the random variables Z_j, $j = 0, 1, \ldots, 2L-1$ are uncorrelated Gaussian random variables with mean $\mu_j = U_j^{1/3}$ and variance $\sigma_j^2 = U_j^{2/3} \frac{18(N-L)}{[9(N-L)-2]^2}$. The proof follows similar reasoning as in Section 9.3.A with the only difference that $\widehat{U}_j$ is now given by $\widehat{U}_j = \left(\mathbf{d}_{\|0\|}^{(j)^T} \widehat{\mathbf{R}}^{(\lfloor j/2 \rfloor)^{-1}} \mathbf{d}_{\|0\|}^{(j)} \right)^{-1}$ and $\widehat{\mathbf{R}}^{(\lfloor j/2 \rfloor)^{-1}}$ is Wishart $W_L(\frac{1}{N}\mathbf{R}^{(\lfloor j/2 \rfloor)^{-1}}; N)$ with N degrees of freedom. The latter implies that $\widehat{U}_j$ can be expressed as a multiple of a χ^2 random variable with $N-L$ degrees of freedom.

9.4 A Recursive Method for the Approximation of the Probability of Coarse Synchronization Error

In this section, we develop a recursive procedure for the evaluation of the probability of coarse synchronization error, P_{cse}. The method evaluates P_{cse} as a sequence of increasingly tight lower bounds on the approximate expression (9.24) and is computationally simpler than the latter expression that evaluates P_{cse} as a convolution of pdfs. We recall that (9.24) with appropriate definitions of $f_j(\cdot)$ and $\mathcal{F}_j(\cdot)$ can be used for the evaluation of $P_{cse,MF}^{(2L)}$, $P_{cse,MF}^{(L)}$, $P_{cse,MVDR}^{(2L)}$, and $P_{cse,MVDR}^{(L)}$ as discussed in Section 9.3.

We begin with the observation that the cardinality of the set $\mathcal{H}(\tau_0)$ is almost always[1] equal to 2. Thus, without loss of generality we assume that $\mathcal{H}(\tau_0) = \{0, 1\}$. In addition, we assume that the random variables Z_j, $j = 2, \ldots, 2L-1$ are ordered in terms of decreasing variance[2], i.e. $Var\{Z_j\} \geq Var\{Z_{j+1}\}$ for $j = 2, \ldots, 2L-2$. Let

$$P_{cse}^{(k)} \triangleq 1 - Pr\left[\left(\arg\max_j \{ Z_j \mid j = 0, 1, \ldots, k+1 \} \right) \in \mathcal{H}(\tau_0) \right] \tag{9.25}$$

$$= 1 - Pr[\max\{Z_0, Z_1\} > \max\{Z_2, Z_3, \ldots, Z_{k+2}\}] \tag{9.26}$$

where $k = 0, \ldots, 2L-3$. It is straightforward to check that for any $k = 0, \ldots, 2L-4$, we have $P_{cse}^{(k)} < P_{cse}^{(k+1)}$. Indeed, including the extra element Z_{k+3} to $\{Z_2, \ldots, Z_{k+2}\}$ can only decrease the probability that the index of the maximum element of $\{Z_2, \ldots, Z_{k+2}, Z_{k+3}\}$ will be larger than $\max\{Z_0, Z_1\}$. Sorting the elements by decreasing variance implies that the k elements with the highest variance (and mean) will produce the tightest lower bound $P_{cse}^{(k)}$ of all other choices of k elements from $\{Z_2, \ldots, Z_{k+2}\}$. For $k = 2L-3$, all variables Z_j, $j = 0, \ldots, 2L-1$, are included in the evaluation of the kth bound, so

$$P_{cse}^{(2L-3)} = 1 - Pr[\max\{Z_0, Z_1\} > \max\{Z_1, Z_2, \ldots, Z_{2L-1}\}]. \tag{9.27}$$

Thus (cf. (9.14)),

$$P_{cse}^{(0)} < P_{cse}^{(1)} < \ldots < P_{cse}^{(2L-3)} = Pr\left[\max\{ Z_j \mid j = 0, \ldots, 2L-1, j \notin \mathcal{H}(\tau_0) \} - \max\{ Z_k \mid k \in \mathcal{H}(\tau_0) \} < 0 \right]. \tag{9.28}$$

The probability bounds $P_{cse}^{(k)}$, $k = 0, \ldots, 2L-3$, can be evaluated as follows. The pdf of $\max\{Z_0, Z_1\}$ is

$$f_{\max\{Z_0, Z_1\}}(z) = f_0(z) F_1(z) + f_1(z) F_0(z), \tag{9.29}$$

while the pdf of $\max\{Z_2, \ldots, Z_{k+2}\}$ is

$$f_{\max\{Z_2, \ldots, Z_{k+2}\}}(z) \triangleq \sum_{l=2}^{k+2} \left(f_l(z) \prod_{j=2, j \neq l}^{k+2} \mathcal{F}_j(z) \right). \tag{9.30}$$

[1] The cardinality is 1 if and only if $\tau_0 = nT_c$ for some $n = 0, \ldots, L-1$. However, the probability of these events is zero.

[2] Sorting the variables Z_j in terms of decreasing variance is equivalent to sorting them in terms of decreasing mean.

Then, the probability $P_{cse}^{(k)}$ is equal to

$$P_{cse}^{(k)} = 1 - \int_{-\infty}^{0} \int_{-\infty}^{+\infty} f_{\max\{Z_0,Z_1\}}(z-y) f_{\max\{Z_2,\ldots,Z_{k+2}\}}(-y) dy dz. \tag{9.31}$$

Alternatively, a simpler recursive procedure for the evaluation of the probability bounds $P_{cse}^{(k)}$, $k = 0,\ldots,2L-3$, is proposed below. The probability $P_{cse}^{(k)}$, $k = 0,\ldots,2L-3$, is equal to

$$P_{cse}^{(k)} = P\left[Z_{\max}^{(k)} - Z_{\max}^{\{0,1\}} < 0\right] \tag{9.32}$$

where

$$Z_{\max}^{(k)} \triangleq \max\{Z_2,\ldots,Z_{k+2}\} \tag{9.33}$$

and

$$Z_{\max}^{\{0,1\}} \triangleq \max\{Z_0, Z_1\}. \tag{9.34}$$

Assuming that $Z_{\max}^{(k)}$ and $Z_{\max}^{\{0,1\}}$ are normally distributed we may calculate $P_{cse}^{(k)}$ as follows

$$P_{cse}^{(k)} \simeq 1 - \mathcal{G}\left(\frac{E\left\{Z_{\max}^{(k)}\right\} - E\left\{Z_{\max}^{\{0,1\}}\right\}}{\sqrt{Var\left\{Z_{\max}^{(k)}\right\} + Var\left\{Z_{\max}^{\{0,1\}}\right\}}}\right). \tag{9.35}$$

The mean and variance of $Z_{\max}^{(k)}$ can be evaluated recursively as described by the following lemma.

Lemma 9.2 *For $k = 1,\ldots,2L-3$,*

$$\begin{aligned}
E\left\{Z_{\max}^{(k)}\right\} \triangleq\ & E\left\{Z_{\max}^{(k-1)}\right\} \mathcal{G}\left(\frac{E\left\{Z_{\max}^{(k-1)}\right\} - E\{Z_{k+2}\}}{\sqrt{Var\left\{Z_{\max}^{(k-1)}\right\} + Var\{Z_{k+2}\}}}\right) \\
& + E\{Z_{k+2}\}\, \mathcal{G}\left(-\frac{E\left\{Z_{\max}^{(k-1)}\right\} - E\{Z_{k+2}\}}{\sqrt{Var\left\{Z_{\max}^{(k-1)}\right\} + Var\{Z_{k+2}\}}}\right) \\
& + \sqrt{Var\left\{Z_{\max}^{(k-1)}\right\} + Var\{Z_{k+2}\}}\ \cdot\ g\left(\frac{E\left\{Z_{\max}^{(k-1)}\right\} - E\{Z_{k+2}\}}{\sqrt{Var\left\{Z_{\max}^{(k-1)}\right\} + Var\{Z_{k+2}\}}}\right) \quad \text{and} \qquad (9.36)
\end{aligned}$$

$$\begin{aligned}
Var\left\{Z_{\max}^{(1)}\right\} \triangleq\ & Var\left\{Z_{\max}^{(k-1)}\right\} \mathcal{G}\left(\frac{E\left\{Z_{\max}^{(k-1)}\right\} - E\{Z_{k+2}\}}{\sqrt{Var\left\{Z_{\max}^{(k-1)}\right\} + Var\{Z_{k+2}\}}}\right) \\
& + Var\{Z_{k+2}\}\, \mathcal{G}\left(-\frac{E\left\{Z_{\max}^{(k-1)}\right\} - E\{Z_{k+2}\}}{\sqrt{Var\left\{Z_{\max}^{(k-1)}\right\} + Var\{Z_{k+2}\}}}\right) \\
& + \left[\left(E\left\{Z_{\max}^{(k-1)}\right\} - E\{Z_{k+2}\}\right) \mathcal{G}\left(\frac{E\left\{Z_{\max}^{(k-1)}\right\} - E\{Z_{k+2}\}}{\sqrt{Var\left\{Z_{\max}^{(k-1)}\right\} + Var\{Z_{k+2}\}}}\right)\right. \\
& \qquad \left. + \sqrt{Var\left\{Z_{\max}^{(k-1)}\right\} + Var\{Z_{k+2}\}}\ \cdot\ g\left(\frac{E\left\{Z_{\max}^{(k-1)}\right\} - E\{Z_{k+2}\}}{\sqrt{Var\left\{Z_{\max}^{(k-1)}\right\} + Var\{Z_{k+2}\}}}\right)\right]
\end{aligned}$$

$$\cdot \left[\left(E\left\{ Z_{\max}^{(k-1)} \right\} - E\left\{ Z_{k+2} \right\} \right) \mathcal{G}\left(- \frac{E\left\{ Z_{\max}^{(k-1)} \right\} - E\left\{ Z_{k+2} \right\}}{\sqrt{Var\left\{ Z_{\max}^{(k-1)} \right\} + Var\left\{ Z_{k+2} \right\}}} \right) \right.$$
$$\left. - \sqrt{Var\left\{ Z_{\max}^{(k-1)} \right\} + Var\left\{ Z_{k+2} \right\}} \cdot g\left(\frac{E\left\{ Z_{\max}^{(k-1)} \right\} - E\left\{ Z_{k+2} \right\}}{\sqrt{Var\left\{ Z_{\max}^{(k-1)} \right\} + Var\left\{ Z_{k+2} \right\}}} \right) \right] \tag{9.37}$$

with $E\left\{ Z_{\max}^{(0)} \right\} = E\{Z_2\}$ *and* $Var\left\{ Z_{\max}^{(0)} \right\} = Var\{Z_2\}$.

Proof: Since

$$Z_{\max}^{(k)} = \max\{Z_2, \ldots, Z_{k+2}\} \;=\; \max\{Z_{\max}^{(k-1)}, Z_{k+2}\} \tag{9.38}$$

(9.36) and (9.37) provide, respectively, the mean and variance of the maximum of two uncorrelated Gaussian random variables [17]. □

The mean and variance of $Z_{\max}^{\{0,1\}}$ can be evaluated in exactly the same way. As a final comment, we recall that the expression (9.24) of Section 9.3 was derived based on the assumption that the pdf of the received vectors can be closely approximated by the pdf of i.i.d. Gaussian random vectors. The recursive method for the evaluation of P_{cse} developed in this section is based on the additional assumption that the maximum of two independent Gaussian random variables is approximately Gaussian. The effect of both assumptions is examined in the simulations section through comparisons with the exact P_{cse}.

9.5 Discussion

Summarizing the developments so far, the probability of coarse synchronization error was approximated as follows:

$$P_{cse} \simeq 1 - Pr[\arg\max_j\{ Z_j \mid j = 0, \ldots, 2L-1\} \in \mathcal{H}(\tau_0)] \tag{9.39}$$

where $\mathcal{H}(\tau_0)$ is the set that contains the indices of the filters that yield a timing estimate within the pull-in range. The random variables $Z_0, \ldots, Z_{2L-1}$ are uncorrelated Gaussian random variables with means and variances given by

$$E\{Z_j\} = U_j^{1/3}, \quad Var\{Z_j\} = C(N)U_j^{2/3} \tag{9.40}$$

where the quantities U_j, $j = 0, \ldots, 2L-1$, and the functions $C(N)$ are presented in Table 9.1 and 9.3, respectively, for each filter under consideration (SMI-MVDR or MF of order L or $2L$).

Thus, the probability of coarse synchronization error was calculated as the probability that the index of the largest Z_j, $j = 0, \ldots, 2L-1$, is not contained in $\mathcal{H}(\tau_0)$. This approach led to the P_{cse} expressions for the order-L and order-$2L$ MF-type or SMI-MVDR-type receivers presented in Section 9.3, as well as to the evaluation of P_{cse} through the sequence of lower bounds presented in Section 9.4. An intuitively simpler but computationally more complex method for the calculation of P_{cse} is presented in the Appendix. Expression (9.40) and Table 9.3 imply that the mean and variance of Z_j, $j = 0, \ldots, 2L-1$, and, consequently, the performance of the SMI-MVDR-type algorithms depend explicitly on the filter order. Moreover, the data record size N and the filter order appear only in the terms $(N-2L)$ and $(N-L)$ of $C(N)$ for the filters of order $2L$ and L, respectively. Although the filter order does not affect the asymptotic (as $N \to \infty$) performance of the SMI-MVDR-type algorithms, it does cause a performance drop in the small data-record-size operating region since the filter order term is subtracted from the data-record-size term. In other words, the data record size is "effectively" reduced by a number of data vector samples equal to the filter order. The longer the employed filter is, the greater the "effective" data-record-size reduction is. For the SMI-MVDR-type coarse synchronization algorithms, this translates to poorer performance under short-data-support conditions as the filter order increases.

An alternative way to quantify the effect of the data-record-size on the probability of coarse synchronization error is to examine the short-data-record synchronization resolution of the SMI-MVDR-type algorithms

Table 9.3. Definition of $C(N)$

L-order	
SMI-MVDR	MF
$C(N) = \frac{18(N-L)}{[9(N-L)-2]^2}$	$C(N) = \frac{18N}{(9N-2)^2}$
$2L$-order	
SMI-MVDR	MF
$C(N) = \frac{18(N-2L)}{[9(N-2L)-2]^2}$	$C(N) = \frac{18N}{(9N-2)^2}$

that we define as follows. For the pair of decision variables (Z_l, Z_m), $l = 0, \ldots, 2L-1$, $m = 0, \ldots, 2L-1$, with $E\{Z_l\} \neq E\{Z_m\}$, the resolution metric is defined by

$$\rho_{l,m} \triangleq \begin{cases} Pr[Z_l > Z_m], & \text{if } E\{Z_l\} > E\{Z_m\} \\ Pr[Z_l < Z_m], & \text{otherwise.} \end{cases} \tag{9.41}$$

That is, the resolution $\rho_{l,m}$ is the probability that the coarse synchronization algorithm will decide in favor of the asymptotically largest decision variable. In other words, the synchronization resolution is a measure of how close the performance of the finite data record SMI-based algorithm is to the asymptotic performance. It is straightforward to show that the resolution of the SMI-MVDR algorithms is given by

$$\rho_{l,m} = \mathcal{G}\left(C(N)^{-1/2} \frac{\left| U_l^{1/3} - U_m^{1/3} \right|}{\sqrt{U_l^{2/3} + U_m^{2/3}}} \right) \tag{9.42}$$

where $\mathcal{G}(x)$ is the cdf of a normal random variable with mean 0 and variance 1. The following proposition compares the resolution performance of the L-order and $2L$-order algorithms for a given pair of hypotheses. The proof is straightforward and thus omitted.

Proposition 9.1 *For a given pair of hypotheses* (l, m), $l = 0, \ldots, 2L-1$, $m = 0, \ldots, 2L-1$, *we have*

$$\rho_{l,m}^{(L)} > \rho_{l,m}^{(2L)} \tag{9.43}$$

if and only if

$$\frac{9(N-2L)-2}{9(N-L)-2} \frac{\sqrt{N-L}}{\sqrt{N-2L}} \leq S_{l,m} \tag{9.44}$$

where

$$\begin{aligned} S_{l,m} \triangleq & \frac{\left| (\mathbf{d}_{\|0\|}^{(l)^T} \mathbf{R}^{(\lfloor l/2 \rfloor)^{-1}} \mathbf{d}_{\|0\|}^{(l)})^{-1/3} - (\mathbf{d}_{\|0\|}^{(m)^T} \mathbf{R}^{(\lfloor m/2 \rfloor)^{-1}} \mathbf{d}_{\|0\|}^{(m)})^{-1/3} \right|}{\left| (\bar{\mathbf{d}}_{\|0\|}^{(l)^T} \bar{\mathbf{R}}^{-1} \bar{\mathbf{d}}_{\|0\|}^{(l)})^{-1/3} - (\bar{\mathbf{d}}_{\|0\|}^{(m)^T} \bar{\mathbf{R}}^{-1} \bar{\mathbf{d}}_{\|0\|}^{(m)})^{-1/3} \right|} \\ & \cdot \frac{\sqrt{(\bar{\mathbf{d}}_{\|0\|}^{(l)^T} \bar{\mathbf{R}}^{-1} \bar{\mathbf{d}}_{\|0\|}^{(l)})^{-1/3} + (\bar{\mathbf{d}}_{\|0\|}^{(m)^T} \bar{\mathbf{R}}^{-1} \bar{\mathbf{d}}_{\|0\|}^{(m)})^{-2/3}}}{\sqrt{(\mathbf{d}_{\|0\|}^{(l)^T} \mathbf{R}^{(\lfloor l/2 \rfloor)^{-1}} \mathbf{d}_{\|0\|}^{(l)})^{-2/3} - (\mathbf{d}_{\|0\|}^{(m)^T} \mathbf{R}^{(\lfloor m/2 \rfloor)^{-1}} \mathbf{d}_{\|0\|}^{(m)})^{-2/3}}}. \end{aligned} \tag{9.45}$$

□

Let $S_{min} \triangleq \min\{S_{l,m} \,|\, l,m = 0,\ldots,2L-1\}$. Then, the resolution performance of the L-order algorithm will be better than the performance of the $2L$-order algorithm (in the sense that the L-order algorithm performance will be closer to the corresponding asymptotic performance) for *any* pair of hypotheses, provided that

$$\frac{9(N-2L)-2}{9(N-L)-2}\frac{\sqrt{N-L}}{\sqrt{N-2L}} \leq S_{min}. \tag{9.46}$$

(9.46) identifies a condition on the system parameters and the data record size under which SMI-MVDR-type L-order algorithms outperform the $2L$-order algorithms.

On the other hand, in the case of MF-type algorithms the mean and variance of Z_j do not depend explicitly on the filter order. Theoretically, this conclusion is a direct consequence of the fact that for any Wishart $W_p(\Sigma;N)$ matrix $\mathbf{X}$ and any non-zero vector $\mathbf{v} \in \Re^p$, the quantity $\mathbf{v}^T\mathbf{X}\mathbf{v}$ is a χ^2 random variable that maintains N degrees of freedom (while $(\mathbf{v}^T\mathbf{X}^{-1}\mathbf{v})^{-1}$ is χ^2 with $N-p$ degrees of freedom). Therefore, algorithms that do not require the inversion of a Wishart matrix do not suffer from "effective" data-record-size reduction. Consequently, it is expected that the MF-type and AV-type coarse synchronization algorithms are fairly insensitive to the choice of filter length.

As a final note, we recall that an alternative to direct inversion of $\widehat{\mathbf{R}}^{(l)}$ is the use of a matrix-inversion-lemma based RLS algorithm [15]. In reality, the RLS algorithm computes the inverse of a diagonally loaded version of $\widehat{\mathbf{R}}^{(l)}$, $(\widehat{\mathbf{R}}^{(l)} + \frac{\epsilon}{N}\mathbf{I})^{-1}$ where $\frac{1}{\epsilon}\mathbf{I}$ is the initialization matrix of RLS. In fact, we can show that although RLS provides an asymptotically ($N \to \infty$) unbiased estimate of $(\mathbf{R}^{(l)})^{-1}$, for a finite data record size N it provides a biased estimate of $(\mathbf{R}^{(l)})^{-1}$ with a bias that depends on ϵ/N. Since in practice ϵ is small, the effect of the diagonal loading quickly becomes negligible and, consequently, the performance is close to the SMI performance analyzed in this work.

9.6 Simulation Results and Comparisons

We consider a 10-user asynchronous DS-CDMA system that utilizes Gold sequences of length $L = 31$. We compare the coarse-synchronization-error-rate evaluated using the expressions derived in Sections III and IV with the exact coarse-synchronization-error-rate of MF and SMI-MVDR-type algorithms. The exact coarse-synchronization-error-rate is evaluated by averaging over 10,000 independent runs. The delays of all users are chosen randomly and kept constant for the duration of the experiment.

In Figs. 9.3, 9.4, the SNR[3] of the user of interest is 10dB and the data support size ranges from 35 to 80 samples. The SNRs of the interferers are fixed at 6, 8, 10, 12, 15, 16, 18, 20, and 22dB.

In Fig. 9.3, we plot as a function of k the sequence of bounds $P_{cse}^{(k)}$ developed in Section 9.4 for the order-L and order-$2L$ SMI-MVDR-type algorithms (given by (9.26) and (9.35)-(9.37)).

In Fig. 9.4, we compare the exact coarse synchronization performance of order-L and order-$2L$ algorithms to the derived approximations of Section 9.3 and 9.4, that is the direct approximation through (9.24) and the recursive approximation through (9.26), (9.35)-(9.37). Regardless of the filter order, we see that both approximations offer close coarse synchronization performance estimates. We observe that the SMI-MVDR L-order algorithm exhibits improved performance when compared to the SMI-MVDR $2L$-order scheme for data record sizes above $L = 31$. In fact, in the small data support region the L-order algorithm can achieve the same CSER as the $2L$-order algorithm about L samples faster.

In Figs. 9.5 and 9.6 we repeat the studies of Figs. 9.3 and 9.4 for different values of the SNR of the user of interest. The data support for the estimation of the order-L and order-$2L$ SMI-MVDR filters is $N = 55$ and 80 samples, respectively. The results parallel the findings in Figs. 9.3 and 9.4.

The insensitivity of the non-inversion-based coarse synchronization algorithms to the filter order is illustrated by the simulation study of Fig. 9.7 where the probability of coarse synchronization error of the L-order algorithms is plotted as a function of the system processing gain L (the figure also includes the AV-type coarse synchronization scheme presented in Chapter 8). The user SNR values are identical to those in Fig. 9.3 and the data support for filter estimation is $N = 125$ samples. The signature vectors of the users are constructed by concatenating Gold codes of length 31. We observe that the SMI-MVDR algorithm is severely affected by the value of the spreading gain (filter length) as L approaches N, while the MF-type

[3]The SNR of user k, $k = 0,\ldots,K-1$, is defined as E_k/σ^2 where σ^2 is the variance of the chip-matched filtered and sampled AWGN.

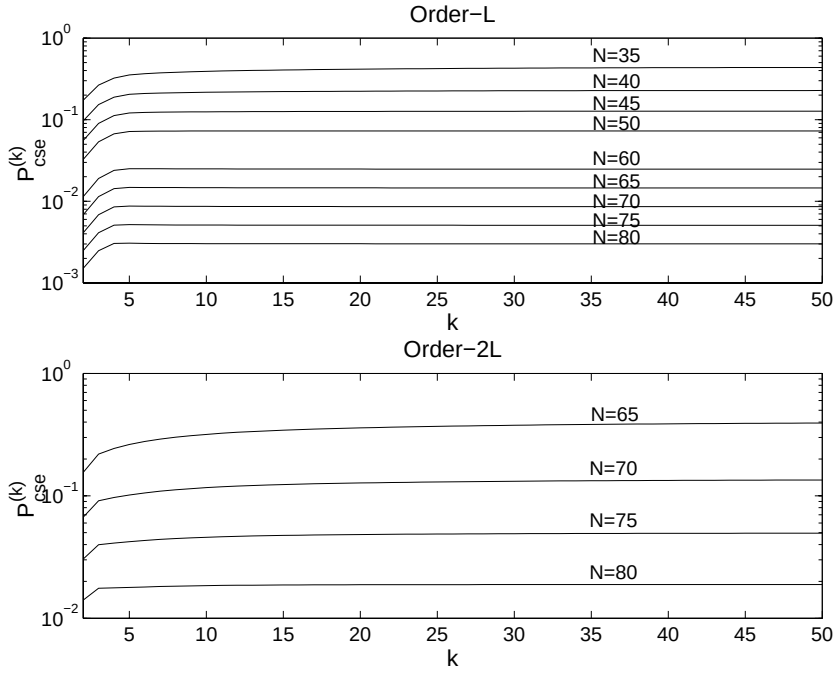

Fig. 9.3. Recursive approximation of the coarse-synchronization-error-rate of SMI-MVDR-type receivers of order-L and order-$2L$ given by (9.35) - (9.37) ($SNR_0 = 10dB$).

algorithm is not affected (but has unacceptable synchronization performance as also seen in Figs. 9.4 and 9.6). The AV-type synchronizer does not suffer from significant performance degradation as the system spreading gain L increases and offers by far the most reliable synchronization.

9.7 Conclusions

In this chapter, we investigated the finite data record coarse synchronization performance of blind adaptive combined synchronizers/demodulators. Using transformation noise modeling techniques we derived analytic expressions that approximate closely the probability of coarse synchronization error. Three alternative methods were developed. The first method approximates the probability of coarse synchronization error as a double integral. The second approximation method, which is the least computationally complex, provides a sequence of increasingly tight lower bounds on the approximate expressions of the first method. Finally, the third and most intuitive method approximates the probability of coarse synchronization error as a $2L$-order integral (the expression is given in the Appendix). The analytic expressions provide simple, highly-accurate alternatives to computationally demanding performance evaluation through simulations. We showed that the coarse synchronization performance of the MF-type receiver is a function of the data record size while the performance of MVDR-type SMI or RLS-based filter estimators is a function of the difference between the data record size and the employed filter order. The latter translates to an "effective" reduction of the data record size with more evident negative effect on the performance as the available data record becomes smaller. MF-type and AV-type (which provide active interference suppression) filter estimators do not suffer such data reduction and thus provide an attractive solution when the environment dictates short data record synchronization.

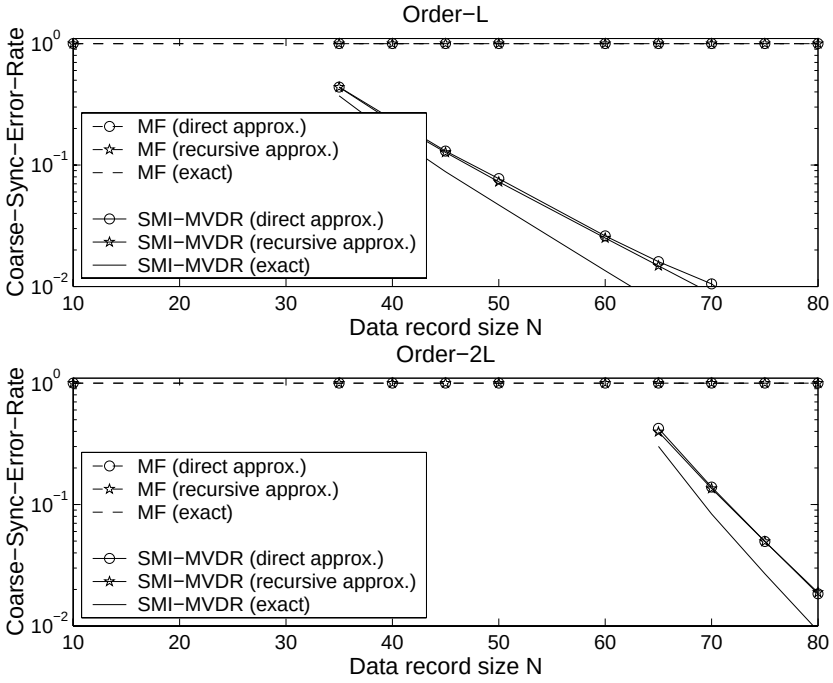

Fig. 9.4. Coarse-synchronization-error-rate of linear receivers of order L and $2L$ as a function of the data record size N $(SNR_0 = 10dB)$.

9.8 Appendix

Proof of Lemma 9.1

The marginal pdf of $\widehat{U}_j$, $j = 0, \ldots, 2L-1$, is

$$h_j(u) = \frac{x^{N/2-1} \exp\{-u/U_j\}}{(\frac{1}{N}U_j)^{N/2}\Gamma(N/2)}. \tag{9.47}$$

Expression (9.47) implies that the random variable $\widehat{U}_j$, $j = 0, \ldots, 2L-1$, is a multiple of a chi-square (χ^2) random variable. More precisely, $\widehat{U}_j$ can be written as

$$\widehat{U}_j \;=\; (\frac{1}{N}U_j)\, X_j, \qquad j = 0, \ldots, 2L-1, \tag{9.48}$$

where X_j is a χ^2 random variable with N degrees of freedom. Since $(X_j/N)^{1/3}$ is approximately Gaussian with mean $\left(1 - \frac{2}{9N}\right)$ and variance $\frac{2}{9N}$ [18], we may approximate the pdf of $\widehat{U}_j$ by the pdf of

$$V_j \;\triangleq\; N(\frac{1}{N}U_j)Y_j^3 \;=\; U_j\, Y_j^3, \quad j = 0, \ldots, 2L-1. \tag{9.49}$$

In the above equation, Y_j is a Gaussian variable with mean $\left(1 - \frac{2}{9N}\right)$ and variance $\frac{2}{9N}$. The joint pdf of $Y_0, \ldots, Y_{2L-1}$, will be completely defined if we specify their covariance matrix that we denote by Ψ. The diagonal elements of Ψ are equal to $\frac{2}{9N}$. The non-diagonal elements of Ψ are chosen such that $V_0, \ldots, V_{2L-1}$ have covariances identical to the covariances of the corresponding $\widehat{U}_0, \ldots, \widehat{U}_{2L-1}$ variables. The latter covariances are [19]:

$$Cov\left[\widehat{U}_j,\ \widehat{U}_k\right] = \frac{2}{N}\left(\overline{\mathbf{d}}_{\|0\|}^{(j)^T}\, \overline{\mathbf{R}}\, \overline{\mathbf{d}}_{\|0\|}^{(k)}\right)^2. \tag{9.50}$$

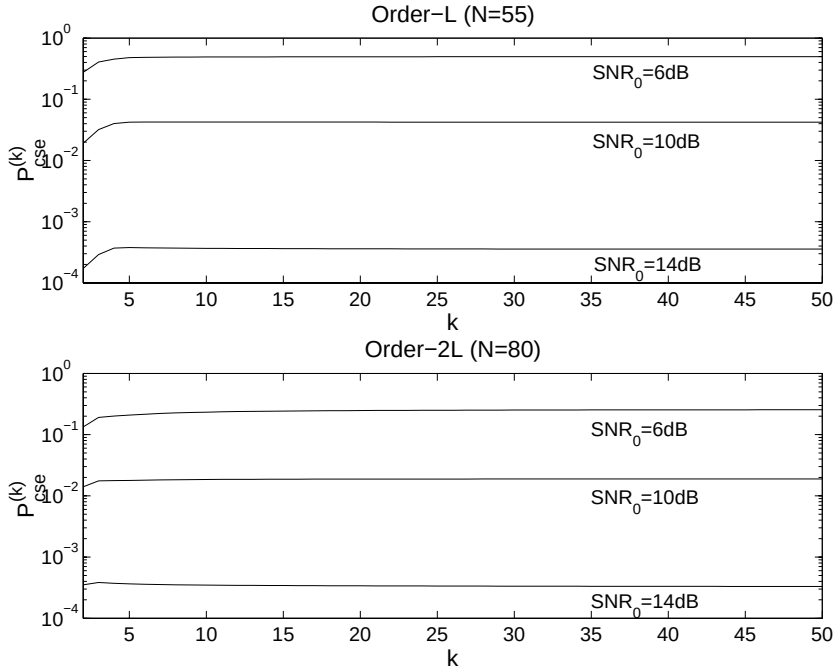

Fig. 9.5. Recursive approximation of the coarse-synchronization-error-rate of SMI-MVDR-type receivers of order-L and order-$2L$ given by (9.35)-(9.37) ($N = 55$ and 80, respectively).

To find the covariances ψ_{jk} of Y_j and Y_k, $j \neq k$, that result in

$$Cov\,[V_j,\, V_k] = \frac{2}{N}\left(\overline{\mathbf{d}}_{\|0\|}^{(j)^T}\,\overline{\mathbf{R}}\,\overline{\mathbf{d}}_{\|0\|}^{(k)}\right)^2, \tag{9.51}$$

we substitute (9.49) in (9.51) and make use of the moment theorem to evaluate $E\{Y_j^3 Y_k^3\}$. Then, we find that the (j,k)th non-diagonal element of Ψ, ψ_{jk}, must satisfy

$$6\psi_{jk}^3 + 18(1-\frac{2}{9N})^2\psi_{jk}^2 + 9\frac{(4-18N-81N^2)^2}{6561N^4}\psi_{jk} = \frac{2}{N}\frac{(\overline{\mathbf{d}}_{\|0\|}^{(j)^T}\,\overline{\mathbf{R}}\overline{\mathbf{d}}_{\|0\|}^{(k)})^2}{(\overline{\mathbf{d}}_{\|0\|}^{(j)^T}\,\overline{\mathbf{R}}\overline{\mathbf{d}}_{\|0\|}^{(j)})(\overline{\mathbf{d}}_{\|0\|}^{(k)^T}\,\overline{\mathbf{R}}\overline{\mathbf{d}}_{\|0\|}^{(k)})} \tag{9.52}$$

or equivalently

$$\psi_{jk}^3 + 3(1-\frac{2}{9N})^2\psi_{jk}^2 + 9\frac{(4-18N-81N^2)^2}{6\times 6561N^4}\psi_{jk} - \frac{1}{3N}\alpha = 0 \tag{9.53}$$

where $\alpha \triangleq \frac{(\overline{\mathbf{d}}_{\|0\|}^{(j)^T}\,\overline{\mathbf{R}}\overline{\mathbf{d}}_{\|0\|}^{(k)})^2}{(\overline{\mathbf{d}}_{\|0\|}^{(j)^T}\,\overline{\mathbf{R}}\overline{\mathbf{d}}_{\|0\|}^{(j)})(\overline{\mathbf{d}}_{\|0\|}^{(k)^T}\,\overline{\mathbf{R}}\overline{\mathbf{d}}_{\|0\|}^{(k)})}$. By the Schwarz inequality $\alpha \in (0,1)$. Equation (9.53) has three real roots, two of which are always negative, provided that $N > 4$.

The third root is plotted in Fig. 9.8 as a function of N and α. We see that the third root is always positive while it drops rapidly to zero as N increases. In fact, for $N > 40$ we see that ψ_{jk} is less than 0.0045. Thus, we may assume that $\psi_{jk} = 0,\ j \neq k,\ j,k = 0,\ldots,2L-1$. In other words, we can safely assume that the Gaussian random variables Y_j, $j = 0,\ldots,2L-1$, are uncorrelated and, therefore, independent. □

An alternative approximation of P_{cse}

The probability of coarse synchronization error can be approximated by (cf. (9.14))

$$P_{cse} \simeq 1 - Pr[\arg\max\{\,Z_j \mid j = 0,\ldots,2L-1\} \in \mathcal{H}(\tau_0)] \tag{9.54}$$

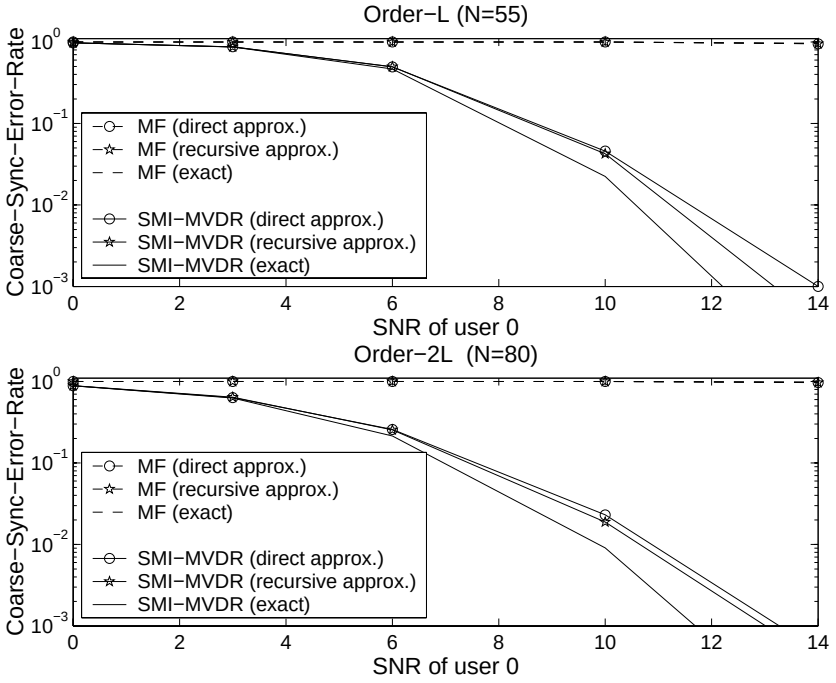

Fig. 9.6. Coarse-synchronization-error-rate of linear receivers of order L and $2L$ as a function of the SNR of the user of interest ($N = 55$ and 80, respectively).

where $\mathcal{H}(\tau_0)$ is the set that contains the indices of the filters that yield a timing estimate within the pull-in range. The random variables Z_j, $j = 0, \ldots, 2L-1$, are independent Gaussian with mean $U_j^{1/3}$ and variance $C(N)U_j^{2/3}$ where the function $C(N)$ is given in Table 9.3 for each of the receivers under consideration. The pdf of the random vector $\mathbf{Z} \triangleq [Z_0, \ldots, Z_{2L-1}]^T$ is

$$f_{\mathbf{Z}}(\mathbf{z}) \triangleq \frac{1}{(2\pi)^L \prod_{j=0}^{2L-1} \sqrt{C(N)U_j^{2/3}}} \exp\left\{ -\sum_{j=0}^{2L-1} \frac{\left(z_j - U_j^{1/3}\right)^2}{2C(N)U_j^{2/3}} \right\} \tag{9.55}$$

where $\mathbf{z} = [z_0, \ldots, z_{2L-1}]^T$. Thus,

$$P_{cse} \simeq 1 - \int_{\mathcal{D}_0} f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \tag{9.56}$$

where

$$\mathcal{D}_0 = \bigcup_{k \in \mathcal{H}(\tau_0)} \left\{ (d_0, d_1, \ldots, d_{2L-1}) \in \Re^{2L} \text{ with } d_k > d_j,\ j = 0, \ldots, 2L-1,\ j \neq k \right\}. \tag{9.57}$$

□

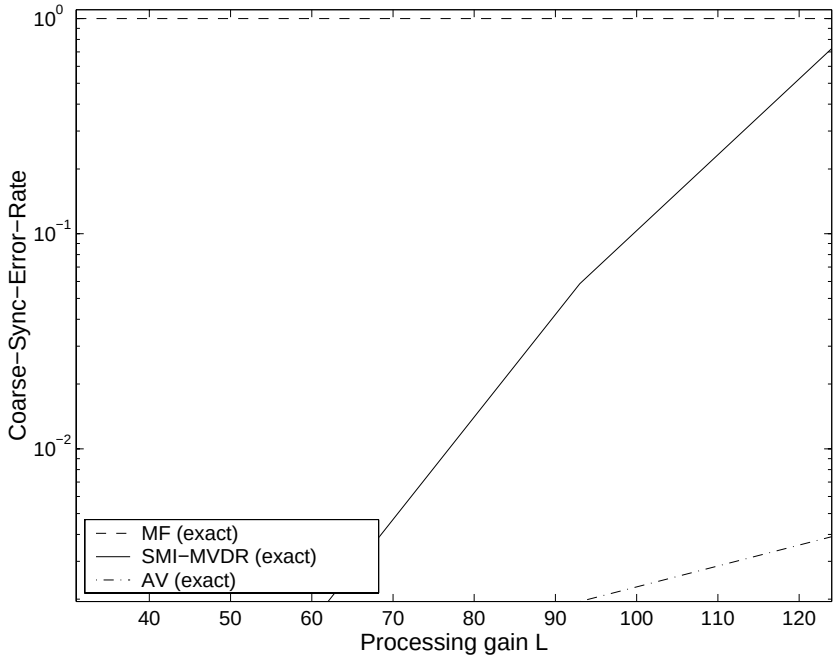

Fig. 9.7. Coarse-synchronization-error-rate of linear receivers of order L as a function of the processing gain L.

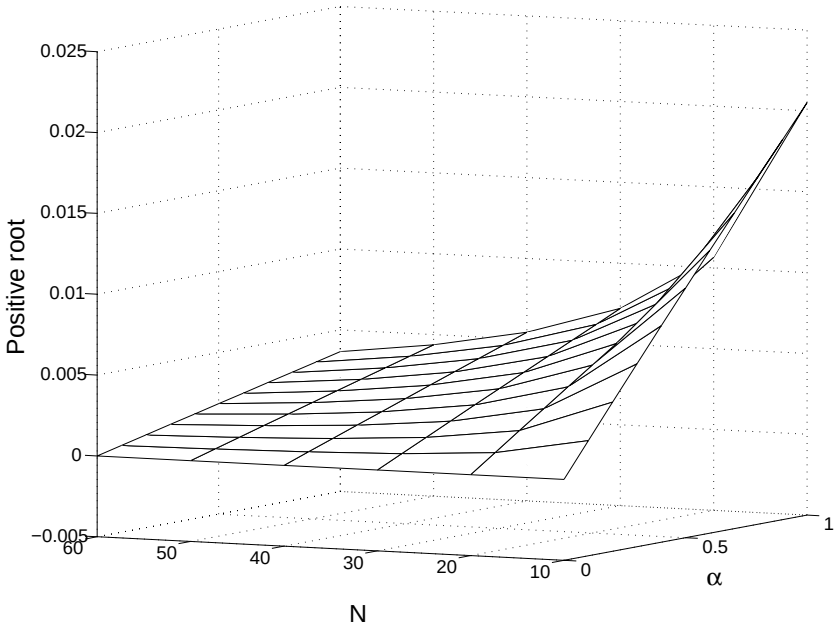

Fig. 9.8. Positive root of (9.53) as a function of α and data-record-size N.

Bibliography

[1] D. A. Pados and S. N. Batalama, "Low-complexity blind detection of DS/CDMA signals: Auxiliary vector receivers," *IEEE Trans. Commun.*, vol. 45, pp. 1586–1594, Dec. 1997.

[2] A. Kansal, S. N. Batalama, and D. A. Pados, "Adaptive maximum SINR RAKE filtering for DS-CDMA multipath fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1765-1773, Dec. 1998.

[3] D. A. Pados and S. N. Batalama, "Joint space-time auxiliary-vector filtering for DS/CDMA systems with antenna arrays," *IEEE Trans. Commun.*, vol. 47, pp. 1406-1415, Sept. 1999.

[4] D. A. Pados and G. N. Karystinos, "An iterative algorithm for the computation of the MVDR filter," *IEEE Trans. Signal Proc.*, vol. 49, pp. 290-300, Feb. 2001.

[5] S. Verdu, "Adaptive multiuser detection," in *Proc. IEEE Intern. Symposium on Spread Spectrum Techniques & Applications,* vol. 1, pp. 43-50, 1994.

[6] R. Kohno, P. B. Rapajic, and B. S. Vucetic, "Overview of adaptive techniques for interference minimization in CDMA systems," *Wireless Personal Communications*, vol. 1, pp. 3-21, 1994.

[7] W. Y. Tan, "On the distribution of the sample covariance matrix from a mixture of normal densities," *South African Stat. J.*, vol. 12, pp. 47-55, 1978.

[8] S. E. Bensley and B. Aazhang, "Maximum-likelihood synchronization of a single user for code-division multiple-access communication systems," *IEEE Trans. Commun.*, vol. 46, pp. 392-399, Mar. 1998.

[9] Z. Liu, J. Li, and S. L. Miller, "An efficient code-timing estimator for receiver diversity DS-CDMA Systems," *IEEE Trans. Commun.*, vol. 46, pp. 826-835, June 1996.

[10] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis.* Wiley: New York, 1958.

[11] R. J. Muirhead, *Aspects of Multivariate Statistical Theory.* Wiley: New York, 1982.

[12] P. R. Krishnaiah and M. M. Rao,"Remarks on a multivariate Gamma distribution," *Amer. Math. Monthly,* vol. 68, pp. 342-346, 1961.

[13] T. Royen, "Expansions for the multivariate chi-square distribution," *J. Multivar. Anal.*, vol. 38, pp. 213-232, 1991.

[14] S. N. Batalama, M. J. Medley, and I. N. Psaromiligkos, "Adaptive robust spread-spectrum receivers," *IEEE Trans. Commun.*, vol. 47, pp. 905-917, June 1999.

[15] L. C. Godara, "Application of antenna arrays to mobile communications, Part II: Beam-forming and direction-of-arrival considerations," *Proceedings of the IEEE*, vol. 85, pp. 1195-1245, Aug. 1997.

[16] H. A. David, *Order Statistics.* Wiley: New York, 1981.

[17] O. Kella, "On the distribution of the maximum of bivariate normal random variables with general means and variances," *Comm. Statist. A - Theory Methods*, vol. 15, pp. 3265–3276, 1986.

[18] N. C. Severo and M. Zelen, "Normal approximation to the chi-square and non-central F probability functions," *Biometrika*, vol. 47, pp. 411-416, 1960.

[19] L. R. Haff, "An identity for the Wishart distribution with applications," *J. Multivar. Anal.*, vol. 9, pp. 531-544, 1979.

Chapter 10

Spatial and Temporal Processing for Global Navigation Satellite Systems: The GPS Receiver Paradigm

10.1 Introduction

The Global Positioning System (GPS), originally developed for military use, has received a lot of attention recently for use in civilian applications such as aviation, agriculture, land-vehicle navigation, surveying and mapping, to name a few [1]-[5]. The GPS satellite navigation system comprises the space segment, the control segment and the user segment. The space segment refers to the constellation of 24 GPS satellites out of which six to eight are usually in simultaneous view by a GPS receiver. The control segment is responsible for the management and monitoring of the GPS system operations. It involves monitor stations and ground antennas that generate and upload navigation data to each GPS satellite. The user segment refers to receiver related activities such as estimation of the receiver position, velocity and/or acceleration.

The GPS system employs direct-sequence spread-spectrum (DS-SS) signaling [6]-[8] and utilizes two carriers, namely, the L_1 carrier at 1575.42 MHz and the L_2 carrier at 1227.6 MHz. Each satellite is assigned a coarse acquisition (C/A) code and a precision (P) code. Both types of codes belong to the general class of product codes, i.e., they are formed by the product of two equal period pseudorandom noise (PN) codes. The C/A-code is a Gold sequence with chipping rate at 1.023 Mchips/sec and period 1 msec (i.e. code-length 1023 chips), while the P-code is a PN code with chipping rate at 10.23 Mchips/sec and truncated period of one week. The L_1 carrier involves both C/A-signals and P-signals while the L_2 carrier involves only P-signals. The use of P-signals is limited to authorized DoD users. On the other hand, C/A-signals are available for civilian applications and are critical in military applications as well, since their demodulation facilitates the acquisition of P-signals. The C/A and P code are modulated by binary navigation data (at 50 bps) and then multiplexed in phase quadrature to form the satellite signal. In this chapter we focus on the C/A component of the transmitted GPS signal (or as commonly stated, we assume perfect separation of the C/A and P-signal component at the receiver).

The working principle of the GPS system is simple: The position of a GPS receiver is modeled as a four-dimensional vector (three coordinates correspond to the spatial position of the receiver and one is related to receiver timing). Estimation of the position can be achieved by utilizing the signals of a minimum of four satellites (or equivalently, by solving a system of four equations with respect to four unknowns, namely the three spatial position coordinates and the receiver clock bias). If additional receiver parameters such as velocity and/or acceleration are of interest then a higher dimension vector is required to model the system.

The signal captured by a GPS receiver is the aggregate of the GPS signals of the satellites that are currently in view, their multipaths, additive white Gaussian noise (AWGN) due to front-end receiver electronics, possible intelligent hostile spread-spectrum interference (spoofing), other wideband and narrowband interference (we note that the only type of wideband interference considered in this work is SS interference while narrowband interference *is not* considered at all; extensive treatment of the problem of narrowband

interference suppression can be found in [9]-[14]). The component of the received signal that is due to the GPS signals of the satellites currently in view is the superposition of very low correlated SS signals due to the fact that Gold sequences are utilized [15], [16]. We recall that the cross-correlation of the C/A-codes assumes three values, namely $\frac{1}{1023}$, $\frac{63}{1023}$, and $\frac{65}{1023}$, all related to shifts at integer multiples of the chip duration. However, an intelligent SS interferer/spoofer who knows the satellite C/A-code can synchronize to the satellite signal of interest (at its own location), mimic the satellite transmission and thus severely corrupt the signal of interest by contributing highly correlated (with the signal of interest) SS interference. Indeed, simple geometric calculations can show that if the difference of the direction-of-arrival (DoA) of the jamming signal and that of the satellite signal is less than 0.2π, then as long as the spoofer is within 1500m from the GPS receiver, the jamming SS signal lags the satellite signal of interest by a fraction of a chip period. The latter condition can result in high signature correlation between the jamming SS signal and the SS satellite signal of interest at the receiver (we recall that the auto-correlation of integrate-and-dump Gold SS signals for lags less than a chip period can be high)[1].

Demodulation of the navigation data of the satellites currently in view can be achieved either by utilizing a bank of single-satellite detectors, or by using a fixed decision device designed to simultaneously detect the data of a number of *specific* satellites [1], [17], [18]. The former receiver architecture appears better suited to the GPS environment since the number of satellites in view changes with time.

In the presence of highly-correlated SS interference, adaptive temporal processing may not be sufficient by itself to achieve satisfactory interference suppression. The use of an antenna array allows the receiver to operate in the spatial domain in addition to the temporal (code) domain and hence exhibit improved interference suppression capabilities. On the other hand, there are cases where interference suppression can be achieved through adaptive spatial processing only (followed or preceded by temporal matched filtering). However, when cost and size requirements limit the degrees of freedom provided by the antenna array, then adaptive spatial processing alone might not be sufficient for effective interference suppression. In such cases adaptive temporal processing in addition to spatial processing is required (either jointly or disjointly optimized). In this chapter, we address the problem of navigation data demodulation by an adaptive GPS receiver that utilizes a bank of single-satellite linear tapped-delay filters and employs antenna-array reception [19], [20]; we investigate *disjoint-domain* as well as *joint-domain* space-time GPS signal processing techniques.

The receiver processors considered in this work are in the form of conventional matched-filters (MF) [21], [22], as well as interference suppressing minimum-variance-distortionless-response (MVDR) filters [23], [24] and auxiliary-vector (AV) filters [25], [26]. The rationale behind the selection of the above processors is as follows. MVDR filtering is a generalization of minimum-mean-square-error (MMSE) filtering (the former reduces to the latter when the constraint vector is chosen to be the statistical cross-correlation between the filter input and the desired response). Thus, the MVDR treatment covers MMSE filtering as a special case. On the other hand, AV filters provide a viable solution when filter estimation with short data records is performed. We recall that the AV algorithm provides a sequence of filters that converges to the linear MMSE/MVDR optimum solution for any positive definite input auto-correlation matrix. When filter estimation with short data records is performed, the early, non-asymptotic elements of the sequence of AV filter estimators provide favorable bias/variance balance and outperform in mean-square filter estimation error (constraint) LMS, RLS, sample-matrix-inversion (SMI), diagonally loaded SMI (DL-SMI), and multi-stage nested Wiener filter estimators [27]. Maximum-likelihood (ML) optimum receivers are not considered in our studies because they require information that is unknown to the receiver (i.e. signatures of *all* active SS interfering signals and all energy terms) and they exhibit unacceptable computational complexity (exponential in the number of active SS signals). Space-time processing in the context of direct-sequence code-division-multiple-access (DS-CDMA) communications with symbol-rate periodic spreading has been extensively studied in the literature (a representative survey chapter is [28]). In this present work, the proposed GPS receiver structures utilize filters that operate at a fraction of the navigation data bit period (1 msec) and are followed by hard-decision detectors. Soft pre-detection measurements or hard post-detection decisions taken over a navigation data bit period are then combined according to a simple combining rule for further bit-error-rate (BER) performance improvements.

The focus of this chapter is on interference resistant GPS receivers. The performance metrics associated with interference resistance characteristics are receiver output SINR and BER. Accurate data demodulation

[1]We note that if the distance parameter value is less than 1500m then even higher DoA-separation parameter values can result in highly correlated jammers with the GPS signal of interest.

is fundamental to the estimation of the location of the GPS receiver since both satellite position information and time stamps are extracted from the received navigation message. A GPS message is transmitted in data frames. Each data frame consists of five subframes while each subframe contains ten 30-bit words. In each data frame, the first subframe provides information for clock correction while the second and third subframes provide ephemeris data which are used for the calculation of the satellite position. The probability that either the satellite position estimate or the timing estimate is in error, P_r, can be assessed via the word error rate in the first three subframes. Accounting for the effects of block coding ((32, 26) Hamming block codes are used by current GPS systems), then the 30-bit word error rate, P_w, can be evaluated (approximately) by $P_w \approx 1-(1-P_e)^{30}-30P_e(1-P_e)^{29}$, where P_e denotes BER. Then, the probability that the range estimate is in error is given by $P_r = 1-(1-P_w)^{30}$. This implies that when $P_e = 0.01$, $P_r \approx 0.67$, while when $P_e = 0.1$, $P_r \approx 1$. Clearly, P_r is a global metric that indicates the effect of BER on the error event of the range estimation process. To further quantify this error, detailed analysis is needed that associates the receiver output SINR with some performance metric (e.g. variance) of the range estimator. Such analysis is beyond the scope of our work.

The rest of this chapter is organized as follows. The signal model is given in Section 10.2. Disjoint space-time processing algorithms are presented in Section 10.3 and are followed by joint domain processing algorithms in Section 10.4. Combining techniques and multipath fading phenomena are addressed in Section 10.5 and Section 10.6, respectively. In Section 10.7 filter estimation issues are considered. Numerical and simulation results are given in Section 10.8 and some concluding remarks are drawn in Section 10.9.

10.2 Signal Model

The C/A and the P-signal component that modulate the L_1 carrier are assumed to be perfectly separated at the receiver. Then, the signal model related to the aggregate C/A component of the received signal can be viewed as an asynchronous DS-SS system with K SS signals in the presence of AWGN. Since the satellite navigation information bit rate is 50 bps, it is equivalent to say that the C/A code of each GPS signal repeats itself 20 times per bit and during this period it is modulated by the same data bit. In this context, the contribution of the kth transmitted GPS signal, $k = 1, 2, \cdots, K$, over $T = 1$ msec is given by

$$u_k(t) = \sum_i b_k(i)\sqrt{E_k}s_k(t-iT)e^{j(2\pi f_c t+\phi_k)} \tag{10.1}$$

where $b_k(i) \in \{-1, +1\}$ is the ith transmitted data bit such that $b_k(i) = b_k(i+m)$, $m = 1, 2, \cdots, 19$, and $b_k(i)$ is independent of $b_k(j)$ for all $|i-j| > 20$. Also, E_k and ϕ_k denote the energy and the carrier phase, respectively, of the kth signal, while s_k is the normalized kth C/A code given by

$$s_k(t) = \sum_{l=0}^{L-1} d_k(l)\psi(t-lT_c). \tag{10.2}$$

In (10.2), $d_k(l) \in \{\pm 1/\sqrt{L}\}$, $l = 0, 1, \cdots, L-1$, are the signature chips of the kth SS signal and $\psi(t)$ is the chip waveform of duration $T_c = T/L$.

The aggregate signal received at the input of a narrowband uniform linear array of M antenna elements is given by

$$\mathbf{x}_c(t) = \sum_{k=1}^{K} u_k(t-\tau_k)\mathbf{a}_k + \mathbf{n}(t) \tag{10.3}$$

where $\mathbf{n}(t)$ is M-dimensional complex AWGN, τ_k is the user time delay and $\mathbf{a}_k$ is the array response vector of the kth SS signal with elements

$$\mathbf{a}_k[m] \triangleq e^{j2\pi(m-1)\frac{\sin\theta_k d}{\lambda}}, \qquad m = 1, 2, \cdots, M. \tag{10.4}$$

In (10.4), θ_k denotes the angle of arrival of the kth signal, λ is the carrier wavelength and d is the inter-element spacing (usually $d = \lambda/2$). For simplicity of presentation, multipath considerations and related signal model generalizations are deferred to Section 10.6.

After carrier demodulation the received signal is given by

$$\mathbf{x}(t) = \sum_{i} \sum_{k=1}^{K} b_k(i)\sqrt{E_k} s_k(t-iT-\tau_k) e^{-j2\pi f_c \tau_k} \mathbf{a}_k + \mathbf{n}(t). \tag{10.5}$$

Without loss of generality, we assume chip synchronization at the reference antenna element (*m=1*) with the SS signal of interest, say *signal 1*, and we also assume $0 \leq \tau_k < T, k = 2, 3, \cdots, K$. After conventional chip matched filtering and sampling at the chip rate $1/T_c$, we can visualize the space-time data samples associated with $b_1(i)$ in the form of an $M \times L$ matrix $\mathbf{X}$

$$\mathbf{X}_{M\times L}(i) = [\mathbf{x}_{M\times 1}(iL) \;\; \mathbf{x}_{M\times 1}(iL+1) \cdots \mathbf{x}_{M\times 1}(iL+L-1)] \tag{10.6}$$

where the column vector $\mathbf{x}_{M\times 1}(iL + j)$ associated with the jth chip-sample within the ith information bit period of signal 1 is given by

$$\begin{aligned} \mathbf{x}_{M\times 1}(iL+j) = \sum_{k=1}^{K} b_k(i)\sqrt{E_k} d_k(j-\tau_k/T_c) e^{-j2\pi f_c \tau_k} \mathbf{a}_k + \\ \mathbf{n}_{M\times 1}(iL+j), \quad j = 0, 1, \cdots, L-1 \end{aligned} \tag{10.7}$$

or equivalently,

$$\begin{aligned} \mathbf{x}_{M\times 1}(iL+j) = b_1(i)\sqrt{E_1} d_1(j)\mathbf{a}_1 + \mathbf{I}_{M\times 1}(iL+j) + \\ \mathbf{n}_{M\times 1}(iL+j), \quad j = 0, 1, \cdots, L-1 \end{aligned} \tag{10.8}$$

where $\mathbf{I}_{M\times 1}(iL + j) = \sum_{k=2}^{K} b_k(i)\sqrt{E_k} d_k(j - \tau_k/T_c)\; e^{-j2\pi f_c \tau_k} \mathbf{a}_k$ denotes comprehensively multiple access interference (MAI) and $\mathbf{n}_{M\times 1}(iL+j)$ is an AWGN noise vector with covariance matrix $\sigma^2 \mathcal{I}_{M\times M}$ ($\mathcal{I}_{M\times M}$ denotes the $M \times M$ identity matrix). In the following, we pursue one-shot detection of the bit of interest $b_1(i)$ and we drop the index i for simplicity in notation.

We note that under the signal model characteristics presented in this section, as well as the GPS receiver architecture adopted in this work (a bank of single-satellite detectors), the focus is on the detection of the navigation data of a single-satellite. In this context and for the single-satellite detector under consideration, the GPS signals of other satellites currently in view as well as possible intelligent SS "GPS-looking" jamming signals can be comprehensively incorporated into the SS interference term **I**. Besides the power advantage of spoofing signals over satellite signals, the difference between the former and the latter SS signals lies in their corresponding signature cross-correlation level with the satellite signal of interest (high and low correlated SS interference, respectively). To justify the above observation it suffices to recall that Gold sequences have low cross-correlation while the auto-correlation of integrate-and-dump Gold SS signals for lags less than one chip period can be high. As an example, suppose now that the distance between the satellite and the GPS receiver is h =20162 km, the distance between the jammer and the GPS receiver is r, and the angle between the DoA of the jamming signal and the DoA of the satellite signal is $\Delta\theta$. Then, by simple geometric calculations, we can evaluate the distance between the satellite and the jammer by $l = \sqrt{h^2 + r^2 - 2hr\cos(\Delta\theta)}$. If $\Delta\theta < 0.2\pi$, then $l + r - h < 300$ m as long as $r \leq 1500$m, which implies that the jamming signal arrives at the GPS receiver within T_c time after the direct satellite signal. Thus, the jammer contributes strong interference from as far away as 1500m from the GPS receiver.

10.3 Disjoint Domain Filtering

In this section we investigate several disjoint domain filtering configurations for GPS signal processing that are formed by the cascade of a spatial filter followed by a temporal filter, or the cascade of a temporal filter followed by a spatial filter. In this context, disjoint domain receiver design is a two stage process with the second stage being conditioned on the design of the first stage. The design optimization criterion imposed at either stage (regardless of the spatial or temporal nature of the corresponding filter) that is of interest in this work is of MF-type, MVDR-type or AV-type. The design optimization criterion for each stage can be selected independently and is usually dictated by considerations of simplicity in implementation,

computational complexity and performance. On the other hand, the space-time or time-space order by which a signal is processed may be influenced by the objective of a particular application. We may expect, for example, that in cases where DoA estimation is critical, a temporal-domain processed received signal is less interference-corrupted than the non-processed received signal and thus it can provide a better input to a DoA estimation algorithm (whose output may, in turn, be needed for subsequent spatial-domain processing).

10.3.1 Space-Time (S-T) Configuration

The general structure of the space-time cascade receiver configuration is shown in Fig. 10.1. For an arbitrary linear spatial processor $\mathbf{w}_s$ and an arbitrary linear temporal processor $\mathbf{w}_t$ the decision on the information bit of the signal of interest b_1 is given by the following expression,

$$\widehat{b}_1 = sgn\left(Re\{\mathbf{w}_t^H(\mathbf{w}_s^H\mathbf{X})^T\}\right) = sgn\left(Re\{\mathbf{w}_s^H\mathbf{X}\mathbf{w}_t^*\}\right) \tag{10.9}$$

where $sgn(\cdot)$ identifies the sign operation (0-threshold hard limiter), superscript H denotes the Hermitian operator, and $Re\{\cdot\}$ extracts the real part of a complex number. We recall that $\mathbf{X}$ is the complex data matrix of size $M \times L$. In the following we present the different forms that $\mathbf{w}_s$ and $\mathbf{w}_t$ may assume.

Spatial Matched-Filtering (sMF)

The spatial matched filter is the filter matched to the array response vector, $\mathbf{a}_1$, of the signal of interest, that is

$$\mathbf{w}_{sMF} = \mathbf{a}_1/M. \tag{10.10}$$

We observe that $E_{b_1}\{\mathbf{X}b_1\mathbf{d}_1\} = \sqrt{E_1}\mathbf{a}_1$, where $\mathbf{d}_1$ is the signature of the signal of interest and the statistical expectation operation $E\{\cdot\}$ is taken with respect to b_1 only. Although estimation of the array response vector *is not* the primary subject of this work, we notice that the above expectation implemented in the form of sample averaging can serve as a supervised estimator of the array response vector of the signal of interest:

$$\widehat{\mathbf{a}}_1 = \frac{1}{J}\sum_{j=1}^{J}\mathbf{X}_j b_1(j)\mathbf{d}_1 \tag{10.11}$$

where $\{\mathbf{X}_j\}_{j=1}^J$ is a sequence of J $M \times L$ matrices of received data. We recall that any subsequent linear filtering followed by 0-threshold hard limiter is positive scalar invariant and thus any scalar that multiplies $\mathbf{a}_1$ does not affect the final output performance. We also note that the scalar $1/M$ in (10.10) preserves the energy of the signal of interest at the output. The latter is merely a convenience consideration for the performance analysis and comparisons presented in subsequent sections.

The output of the filter $\mathbf{w}_{sMF}$ is given by

$$\mathbf{y}_{MF}^T = \mathbf{w}_{sMF}^H\mathbf{X}_{M\times L} = [y_{MF,1} \quad y_{MF,2} \quad \cdots \quad y_{MF,L}] \tag{10.12}$$

$$\begin{aligned} y_{MF,j} &= \mathbf{w}_{sMF}^H\mathbf{x}_{M\times 1}(j) = \sqrt{E_1}b_1 d_1(j) + \mathbf{a}_1^H\mathbf{I}(j)/M \\ &\quad + \mathbf{a}_1^H\mathbf{n}(j)/M, \quad j = 1, 2, \cdots, L. \end{aligned} \tag{10.13}$$

□

Spatial MVDR Filtering (sMVDR)

The MVDR filter is designed to minimize the filter output variance subject to the constraint that the filter remains distortionless in the vector direction of the signal of interest $\mathbf{a}_1$ [23]. The filter is given by

$$\mathbf{w}_{sMV} = \frac{\mathbf{R}_s^{-1}\mathbf{a}_1}{\mathbf{a}_1^H\mathbf{R}_s^{-1}\mathbf{a}_1} \tag{10.14}$$

where $\mathbf{R}_s$ denotes the $M \times M$ covariance matrix of the columns of $\mathbf{X}_{M\times L}$ (i.e. the correlation of the spatial input data). The output of the MVDR filter is given by

$$\begin{aligned} \mathbf{y}_{MV}^T &= \mathbf{w}_{sMV}^H\mathbf{X}_{M\times L} \\ &= \begin{bmatrix} y_{MV,1} & y_{MV,2} & \cdots & y_{MV,L} \end{bmatrix} \end{aligned} \tag{10.15}$$

where

$$y_{MV,j} = \mathbf{w}_{sMV}^H \mathbf{x}_{M\times 1}(j) = \sqrt{E_1} b_1 d_1(j) + \frac{\mathbf{a}_1^H \mathbf{R}_s^{-1} \mathbf{I}(j)}{\mathbf{a}_1^H \mathbf{R}_s^{-1} \mathbf{a}_1} + \frac{\mathbf{a}_1^H \mathbf{R}_s^{-1} \mathbf{n}(j)}{\mathbf{a}_1^H \mathbf{R}_s^{-1} \mathbf{a}_1},\ j = 1, 2, \cdots, L. \tag{10.16}$$

□

Spatial Auxiliary-Vector Filtering (sAV)

For a given spatial covariance matrix $\mathbf{R}_s$, the theory of AV filtering [26], [27] can be applied to the spatial domain and provide an infinite sequence of spatial linear filters that are distortionless in the direction of the signal of interest and can be obtained by the following recursion:

$$\mathbf{w}_{sAV}(0) = \frac{\mathbf{a}_1}{||\mathbf{a}_1||^2} \tag{10.17}$$

For n=1, 2, $\cdots$

$$\mathbf{g}_s(n) = \mathbf{R}_s \mathbf{w}_{sAV}(n-1) - \frac{\mathbf{a}_1^H \mathbf{R}_s \mathbf{w}_{sAV}(n-1)\mathbf{a}_1}{||\mathbf{a}_1||^2} \tag{10.18}$$

$$\mu_s(n) = \frac{\mathbf{g}_s^H(n) \mathbf{R}_s \mathbf{w}_{sAV}(n-1)}{\mathbf{g}_s^H(n) \mathbf{R}_s \mathbf{g}_s(n)} \tag{10.19}$$

$$\mathbf{w}_{sAV}(n) = \mathbf{w}_{sAV}(0) - \sum_{i=1}^{n} \mu_s(i) \mathbf{g}_s(i) \tag{10.20}$$

In the above recursive procedure, the sequence of auxiliary vectors $\mathbf{g}_s(n)$ is selected to maximize the cross-correlation between the $\mathbf{w}_{sAV}(n-1)$ processed data and the $\mathbf{g}_s(n)$ processed data while μ_n is chosen to minimize the variance at the output of the filter $\mathbf{w}_{sAV}(n)$. It is important to note that all auxiliary vectors $\mathbf{g}_s(1)$, $\mathbf{g}_s(2)$, $\cdots$ are constrained to be orthogonal to $\mathbf{a}_1$ but orthogonality among them is not imposed. As shown in [27], the sequence of auxiliary-vector filters converges to $\mathbf{w}_{sMV}$ ($\mathbf{w}_{sAV}(n) \rightarrow \mathbf{w}_{sMV}$ as $n \rightarrow \infty$). The AV filter output when n auxiliary vectors are utilized is given by

$$\mathbf{y}_{AV}^T = \mathbf{w}_{sAV}^H(n) \mathbf{X}_{M\times L}. \tag{10.21}$$

□

Finally, the temporal processor of the space-time configuration is a linear filter of dimension $L \times 1$ that takes as input the spatial-processor output (given by (10.12) or (10.15) or (10.21)). Similar to the optimization criteria used in the design of the spatial processor, the temporal filter can be of MF-type, MVDR-type or AV-type.

Temporal Matched-Filtering (tMF)

The temporal MF is given by

$$\mathbf{w}_{tMF} = \mathbf{d}_1 \tag{10.22}$$

or equivalently, $\mathbf{w}_{tMF} = E_{b_1}\{\mathbf{y} b_1\}$, where the temporal filter input $\mathbf{y}$ is either $\mathbf{y} = \mathbf{y}_{MF}$ or $\mathbf{y} = \mathbf{y}_{MV}$ or $\mathbf{y} = \mathbf{y}_{AV}$. □

Temporal MVDR-Filtering (tMVDR)

The temporal MVDR filter can be shown to be equal to

$$\mathbf{w}_{tMV} = \frac{\mathbf{R}_t^{-1} \mathbf{d}_1}{\mathbf{d}_1^H \mathbf{R}_t^{-1} \mathbf{d}_1} \tag{10.23}$$

where $\mathbf{R}_t$ is the $L \times L$ covariance matrix of the spatial filter output $\mathbf{y}$ and thus it depends on the type of first-stage spatial processing. □

Temporal Auxiliary-Vector Filtering (tAV)
The sequence of auxiliary vector filters in the temporal domain can be obtained as follows,

$$\mathbf{w}_{tAV}(0) = \mathbf{d}_1 \tag{10.24}$$

For $n = 1, 2, \cdots$

$$\mathbf{g}_t(n) = \mathbf{R}_t \mathbf{w}_{tAV}(n-1) - \mathbf{d}_1^H \mathbf{R}_t \mathbf{w}_{tAV}(n-1)\mathbf{d}_1 \tag{10.25}$$

$$\mu_t(n) = \frac{\mathbf{g}_t^H(n)\mathbf{R}_t \mathbf{w}_{tAV}(n-1)}{\mathbf{g}_t^H(n)\mathbf{R}_t \mathbf{g}_t(n)} \tag{10.26}$$

$$\mathbf{w}_{tAV}(n) = \mathbf{w}_{tAV}(0) - \sum_{i=1}^{n} \mu_t(i)\mathbf{g}_t(i) \tag{10.27}$$

and $\mathbf{R}_t$ denotes the covariance matrix of the spatial filter output $\mathbf{y}$ (which again depends on the type of the first-stage spatial processor). □

10.3.2 Time-Space (T-S) Configuration

The time-space (T-S) cascade configuration is the dual alternative to the space-time cascade configuration examined in the previous section. The general T-S cascade structure is depicted in Fig. 10.2 where $\mathbf{z}_{M\times 1}$ and y denote the output of the temporal and spatial filter $\mathbf{w}_t$ and $\mathbf{w}_s$, respectively. Both filters can be designed according to MF, MVDR or AV processing principles in a dual manner to the S-T configuration of the previous section. It is straightforward to show that the S-T configuration sMF/tMF is equivalent to the T-S configuration tMF/sMF. This is not, however, the case when MVDR or AV filters are involved.

With all these design alternatives available for both disjoint S-T and T-S configurations, a natural question arises as to what their relative merits are. We investigate this issue in the following section.

10.3.3 Performance Analysis of Disjoint Processing Schemes

For convenience in presentation we study a hypothetical system consisting of only two spread spectrum signals, say one GPS signal of interest and one "GPS-looking" jamming signal. Our objective is to draw theoretical conclusions about the relative merits of all different design alternatives presented in sections 10.3.1 and 10.3.2. The performance criterion that we consider for a given design is the signal-to-interference-plus-noise-ratio (SINR) at the output of the overall configuration. The results can be generalized to account for the presence of more GPS and jamming signals. However, the mathematical expressions involved are prohibitively complicated with the latter usually at the expense of clarity in presentation. The analysis that follows involves performance comparisons for MF-type and MVDR-type filter combinations only. Disjoint auxiliary-vector filtering analysis renders itself quite complicated and thus it is not considered in this work. The comparisons are summarized in the following theorem. The proof is included in the appendix.

Theorem 10.1 *Let $\mathbf{a}_i$ and $\mathbf{d}_i$ be the spatial and temporal signature of signal i, $i = 1, 2$, of length M and L, respectively. Let ρ and η be defined as $\rho \triangleq |\mathbf{d}_1^H \mathbf{d}_2|$, and $\eta \triangleq \left|\frac{\mathbf{a}_1^H \mathbf{a}_2}{M}\right|$, respectively. Let also $\Gamma_i \neq 0$ denote the input signal-to-noise-ratio (SNR) of signal i, and SINR$_{(sXX/tYY)}$ or SINR$_{(tYY/sXX)}$ denote the output SINR of an S-T or T-S configuration, respectively, that utilizes an XX-type spatial filter and a YY-type temporal filter. Then,*

A. Space-Time configuration

(i) SINR$_{(sMF/tMF)} \leq$ SINR$_{(sMF/tMV)}$, with equality if and only if $\rho \in \{0, 1\}$ or $\eta = 0$,

(ii) SINR$_{(sMV/tMF)} \leq$ SINR$_{(sMV/tMV)}$, with equality if and only if $\rho \in \{0, 1\}$ or $\eta = 0$,

(iii) SINR$_{(sMF/tMF)} \geq$ SINR$_{(sMV/tMF)}$ if and only if

$$\eta^2(1-\eta^2)\left[1 - M\Gamma_2\rho^2(1-\eta^2) - 2L\rho^2\right] \geq 0, \tag{10.28}$$

(iv) A loose sufficient condition for SINR$_{(sMV/tMV)}$ < SINR$_{(sMF/tMV)}$ is

$$1 + M\Gamma_2\eta^2(1-\rho^2) \geq$$

$$\left[1+\frac{M\Gamma_2\eta^2(1-\rho^2)}{1+\frac{M\Gamma_2}{L}(2+\frac{M\Gamma_2}{L})(1-\eta^2)}\right]^2\left[1+\frac{M\Gamma_2}{L}(1-\eta^2)\right]^2, \tag{10.29}$$

(v) $SINR(_{sMF/tMV}) \geq SINR(_{sMV/tMF})$ *if and only if*

$$L^2\eta^2\rho^2(1-\rho^2) \geq (1-\eta^2)\left[M\Gamma_2\rho^2(1-\eta^2)+2L\rho^2-1-\eta^2(1-\rho^2)\right]. \tag{10.30}$$

B. Time-Space configuration
(i) $SINR(_{tMF/sMF}) \leq SINR(_{tMF/sMV})$, *with equality if and only if* $\eta \in \{0,1\}$ *or* $\rho = 0$,
(ii) $SINR(_{tMV/sMF}) \leq SINR(_{tMV/sMV})$, *with equality if and only if* $\eta \in \{0,1\}$ *or* $\rho = 0$,
(iii) $SINR(_{tMF/sMF}) \geq SINR(_{tMV/sMF})$ *if and only if*

$$\rho^2(1-\rho^2)\left[1-M\Gamma_2\eta^2(1-\rho^2)-2M\eta^2\right] \geq 0, \tag{10.31}$$

(iv) A loose sufficient condition for $SINR(_{tMV/sMV}) < SINR(_{tMF/sMV})$ *is*

$$1+M\Gamma_2\rho^2(1-\eta^2) \geq \left[1+\frac{M\Gamma_2\rho^2(1-\eta^2))}{1+\Gamma_2(2+\Gamma_2)(1-\rho^2)}\right]^2\left[1+\Gamma_2(1-\rho^2)\right]^2, \tag{10.32}$$

(v) $SINR(_{tMF/sMV}) \geq SINR(_{tMV/sMF})$ *if and only if*

$$M^2\rho^2\eta^2(1-\eta^2) \geq (1-\rho^2)\left[M\Gamma_2\eta^2(1-\rho^2)+2M\eta^2-1-M\Gamma_2\rho^2(1-\eta^2)\right]. \tag{10.33}$$

C. Space-Time versus Time-Space configuration
(i) $SINR(_{sMF/tMF}) = SINR(_{tMF/sMF})$, *for any* ρ, η,
(ii) $SINR(_{sMV/tMF}) \leq SINR(_{tMF/sMV})$, *with equality if and only if* $\eta \in \{0,1\}$ *or* $\rho = 1/\sqrt{L}$,
(iii) $SINR(_{tMV/sMF}) \leq SINR(_{sMF/tMV})$, *with equality if and only if* $\rho \in \{0,1\}$ *or* $\eta = 1/\sqrt{M}$.

□

A few observations based on the above theorem now follow: **(a)** Under the assumption of AWGN and interference that is independent of the signal of interest, the output interference-plus-noise variance of any cascade configuration of linear filters above can be expressed as the sum of two terms, one due to the interference only and the other due to noise only. **(b)** S-T and T-S are dual configurations (hence results in Part B are simply the dual of results in Part A, so in the following we address Part A and Part C only). **(c)** When $\rho = 0$ or 1, $\mathbf{w}_{tMV}$ degenerates to $\mathbf{w}_{tMF}$. Similarly, when $\eta = 0$ or 1, $\mathbf{w}_{sMV}$ degenerates to $\mathbf{w}_{sMF}$. **(d)** The output SINR increases as the number of antenna elements increases. **(e)** An MVDR structure at the second stage suppresses the overall output interference-plus-noise variance while an MVDR structure at the first stage followed by an MF-structure at the second stage suppresses interference-plus-noise variance of that first stage only. **(f)** We understand that Item (iii) of Part A may appear somewhat unexpected at first glance. To add an intuitive justification/explanation of the validity of this statement, we consider a simple example where $\rho = 0$ and $0 < \eta < 1$. In this case, the interference is completely suppressed after temporal MF of the second stage in both $_{sMF/tMF}$ and $_{sMV/tMF}$ configurations. However, the variance component due to AWGN only is higher at the output of the spatial $_{sMV}$ than $_{sMF}$ of the first stage. This filtered noise term which participates in the input to $_{tMF}$ in both configurations leads to the output SINR relationship of A(iii). Similar justification can be provided for A(iv). We note that the output SINR of the MVDR filter depends on the input SINR as well as on the input interference-to-noise ratio (INR). In other words, for a given set of SS signal signatures and a given input SINR, the output SINR of the MVDR filter may be higher in the case of high input INR than in the case of low input INR. A complete numerical study with respect to this issue is carried out in Fig.10.3 for a GPS system with the signal of interest at 15 dB and one "GPS looking" jamming signal at 30 dB. Perfect knowledge of the pertinent auto-covariance matrices is assumed. We observe, for example, that if $\eta > 0.25$ then $\text{SINR}(_{sMV/tMV}) < \text{SINR}(_{sMF/tMV})$ for $\rho < 0.7$.

10.4 Joint Domain Filtering

For joint domain processing, to avoid cumbersome 2-D operations and notation, we vectorize the matrix $\mathbf{X}_{M\times L}$ by stacking all columns in the form of a vector

$$\mathcal{X}_{ML\times 1} \stackrel{\triangle}{=} Vec\{\mathbf{X}_{M\times L}\}. \tag{10.34}$$

In the following, $\mathcal{X}$ denotes the joint space-time data in the C^{ML} complex vector space that constitutes the input to a joint space-time linear filter $\mathbf{w}$ to be designed according to MF, MVDR or AV processing principles. The general structure of the joint space-time receiver is depicted in Fig. 10.4.

Joint Domain Matched-Filtering (JMF)

The joint space-time matched filter for the signal of interest (*signal 1*) is the joint space-time signature of the signal of interest, that is, the Kronecker product $\mathbf{v}_1 = (\mathbf{d}_1 \otimes \mathbf{a}_1)/M$, where $\mathbf{d}_1$ and $\mathbf{a}_1$ are the temporal code and spatial code (steering vector) of the signal of interest, respectively. Similar to the developments of Section 10.3, $\mathbf{v}_1$ is also given as the cross-correlation between the received space-time vector $\mathcal{X}$ and the desired information bit b_1, that is,

$$\mathbf{w}_{JMF} = \mathbf{v}_1 = E_{b_1}\{\mathcal{X}b_1\}/M = (\mathbf{d}_1 \otimes \mathbf{a}_1)/M \tag{10.35}$$

where the scaling by 1/M is done without loss of generality for consistency purposes in subsequent theoretical and simulation comparisons.

With $\mathbf{w}_{JMF}$ defined by (10.35), it is straightforward to see that the JMF filter is equivalent to the disjoint MF/MF configurations. We also note that JMF (and the equivalent cascade MF/MF designs for that purpose) is mean-square optimum only when the multiple access interference and the external disturbances are white Gaussian, which is not the case for most practical SS communication systems. Indeed, non-orthogonal multiple access interferers as well as highly correlated (with the signal of interest) hostile jammers may render the JMF receiver ineffective. A remedy for the latter situation is to proceed with the design of the near-far resistant MVDR filter receiver or the AV filter receiver that are presented below.

Joint Domain MVDR Filtering (JMVDR)

The joint domain MVDR filter is designed to minimize the variance at its output and simultaneously be distortionless toward $\mathbf{v}_1$, the joint space-time signature of the signal of interest. The filter is given by the following expression:

$$\mathbf{w}_{JMV} = \frac{\mathbf{R}^{-1}\mathbf{v}_1}{\mathbf{v}_1^H\mathbf{R}^{-1}\mathbf{v}_1} \tag{10.36}$$

where $\mathbf{R} = E\{\mathcal{X}\mathcal{X}^H\}$ is the covariance matrix of the space-time data input vector.

Joint Domain Auxiliary-Vector Filtering (JAV)

Joint domain auxiliary-vector filter design provides a sequence of joint domain auxiliary-vector filters that are distortionless toward the joint space-time signature of the signal of interest $\mathbf{v}_1$ and can be obtained by the following recursion:

$$\mathbf{w}_{JAV}(0) = \frac{\mathbf{v}_1}{||\mathbf{v}_1||^2} \tag{10.37}$$

For $n = 1, 2, \cdots$

$$\mathbf{g}(n) = \mathbf{R}\mathbf{w}_{JAV}(n-1) - \frac{\mathbf{v}_1^H\mathbf{R}\mathbf{w}_{JAV}(n-1)\mathbf{v}_1}{||\mathbf{v}_1||^2} \tag{10.38}$$

$$\mu(n) = \frac{\mathbf{g}^H(n)\mathbf{R}\mathbf{w}_{JAV}(n-1)}{\mathbf{g}^H(n)\mathbf{R}\mathbf{g}(n)} \tag{10.39}$$

$$\mathbf{w}_{JAV}(n) = \mathbf{w}_{JAV}(0) - \sum_{i=1}^{n}\mu(i)\mathbf{g}(i) \tag{10.40}$$

Similar to the corresponding auxiliary-vector filtering developments for the disjoint configuration of Section 10.3, the joint-domain auxiliary vector $\mathbf{g}(n)$ is chosen to maximize the cross-correlation $|\mathbf{w}_{JAV}^H(n-1)\mathbf{R}\mathbf{g}(n)|$ while the scalar $\mu(n)$ is chosen to minimize the output filter variance $|\mathbf{w}_{JAV}^H(n)\mathbf{R}\mathbf{w}_{JAV}(n)|$, $n = 1, 2, \cdots$. Convergence analysis of the auxiliary-vector filter sequence can be found

in [27]. Data-record-based criteria for the selection of the best AV estimate of the MVDR filter were recently reported in [29].

As a final comment for this section, we note that while disjoint MF processing in space and time is always equivalent to joint MF processing, adaptive joint-domain (MVDR or AV) processing involves $(ML-1)$ degrees of freedom while disjoint processing involves $(M+L-2)$ [30]. As a result, both joint MVDR and joint AV processing are expected to outperform the corresponding disjoint configurations that involve MVDR processing or AV processing in both space and time, given ideal pertinent statistics. In real world applications, however, finite data-record-size puts the "curse of dimensionality" on joint-domain processing, hence we may find disjoint-domain configurations outperform the joint-domain counterpart under short data support.

10.5 GPS Filter Output Combining

Combining techniques have been extensively studied in the communications literature as a means to provide some form of diversity (a representative example is RAKE filtering for multipath combining [22], [31]). In this section we take advantage of the redundancy introduced by spreading the information bit into 20 C/A code periods and implement combining methods to further improve the receiver BER performance. In fact, while in initial acquisition and demodulation mode this feature cannot be utilized, once bit synchronization has been acquired and the information bit boundary has been detected (usually achieved by a threshold testing of the output of a sliding correlator), we can combine the soft or hard GPS filter outputs by using either *(i) selective combining* (SC) or *(ii) equal gain combining* (EGC) or *(iii) maximum ratio combining* (MRC). In SC, we combine those GPS processor outputs whose soft (pre-detection) outputs are above a pre-specified threshold. Depending on the variations of the environment and the velocity of the receiver, the threshold needs to be set dynamically, which increases significantly the complexity of the method. On the other hand, EGC is the simplest method where all 20 GPS filter outputs are just added together with equal gain of 1. In MRC, the GPS filter (soft or hard) outputs are first weighted and then summed such that each weight is proportional to the corresponding magnitude of the GPS filter soft output. Adaptive implementations of the MRC method that appear necessary under time varying conditions of the environment may contribute to additional complexity. As mentioned above, all these methods can be applied to either the soft or the hard filter outputs with the understanding that pre-detection combining is better than post-detection combining in terms of output SINR.

The advantages of filter output combining as well as the relative SINR merits of the corresponding methods described above appear to be more evident under fading and low SNR conditions. To illustrate this and as a precursor to our simulation studies, we may consider a simple scenario where due to fading and in the absence of SS interference the signal component at all but one (out of 20) filter outputs over a navigation data bit period are attenuated by the multiplicative constant 0.1. Then in presence of AWGN of variance σ^2 a SC method that selects only the strongest output exhibits output SNR equal to $SNR_{SC}=\frac{E}{\sigma^2}$, while the EGC method and the MRC method exhibit output SNR, respectively, $SNR_{EGC}=\frac{(\sqrt{E}+19\cdot 0.1\sqrt{E})^2}{20\sigma^2}=0.42SNR_{SC}$ and $SNR_{MRC}=\frac{(\sqrt{E}+19\cdot 0.1^2\sqrt{E})^2}{\sigma^2+19\cdot 0.1^2\sigma^2}=1.19SNR_{SC}$, where E is the energy of the signal component at the filter output under no attenuation conditions. Alternatively, if we assume that due to fading half of the filter outputs are attenuated by the multiplicative constant 0.5, then $SNR_{SC}=\frac{E}{\sigma^2}$, $SNR_{EGC}=11.25SNR_{SC}$ and $SNR_{MRC}=12.5SNR_{SC}$.

10.6 Multipath Fading Considerations

Fading phenomena associated with mobile satellite communications at the L-band can be categorized into long-term fading and short-term fading. Long-term fading effects in signal envelope variations are modeled as multiplicative log-normal random variables while short-term fading effects are modeled as multiplicative complex Gaussian random variables that introduce a Rayleigh distributed amplitude. In this chapter we consider only short-term channel variations that appear more difficult to accommodate.

Short-term multipath fading channels involve independent fading effects across all resolvable paths. The presence of a direct path (with no amplitude distortion) identifies a Rician channel while its absence

characterizes the channel as Rayleigh. After multipath fading "channel processing" the aggregate spread spectrum signal at the input of a narrowband uniform linear array of M elements is given by

$$\mathbf{x}_c(t) = \sum_{k=1}^{K} \sum_{n=1}^{N} \alpha_{k,n} u_k(t - \tau_{k,n}) \mathbf{a}_{k,n} + \mathbf{n}(t) \tag{10.41}$$

where N is the total number of resolvable multipaths, $\tau_{k,n}$ and $\mathbf{a}_{k,n}$ are respectively the delay and the $M \times 1$ array response vector for the nth path of the kth SS signal, and $\alpha_{k,n}, k = 1, \cdots, K$, are independent zero-mean complex Gaussian random variables that model the fading phenomena and are assumed to remain constant over several C/A code intervals. We recall that, the fading rate is the maximum Doppler shift in frequency and is given by $f_m = v f_0/c$ Hz where v, f_0 and c denotes the velocity of the receiver (measured in m/sec), the carrier frequency (measured in Hz) and the propagation speed (measured in m/sec), respectively. As an example, the fading rate for a mobile GPS receiver moving at 35 mph is about 70 Hz, which implies that the channel coherence time (the reciprocal of fading rate) is approximately 15 msec while the information bit period is 20 msec. If, on the other hand, the GPS receiver is moving almost perpendicular to the line of sight (e.g. at an angle 89.5^o) then the channel varies at a frequency $f_m \cos(89.5) \approx 0.1$ Hz.

The developments presented in the previous sections can be generalized to cover multipath fading channels. In particular, the filter processors (for any of the disjoint or joint-domain configurations) retain the same structure (e.g., expressions (10.10), (10.14), (10.20), (10.22), (10.23), (10.27), (10.35)-(10.40)) with the understanding that the spatial "code" $\mathbf{a}_1$, the temporal code $\mathbf{d}_1$, and the space-time code $\mathbf{d}_1 \otimes \mathbf{a}_1$ are substituted by their multipath fading "channel processed" versions, respectively. Similarly, corresponding covariance matrices are appropriately adjusted to reflect the above multipath fading "channel processing". As an example, the multipath version of the joint space-time MF now takes the well known form of the RAKE filter, i.e. the "channel processed" space-time effective signature of the signal of interest,

$$\mathbf{w}_{RAKE} = E_{b_1}\{\mathbf{X} b_1\}/M = \mathbf{d}_1 \otimes \mathbf{a}_1 \otimes \boldsymbol{\alpha}_1 \tag{10.42}$$

where $\mathbf{X}$ denotes now the $M \times (LN)$ space-time data matrix created by sampling at the chip rate the multipath fading received signal in (10.41), and $\boldsymbol{\alpha}_1$ identifies the $N \times 1$ vector of the multipath fading channel coefficients.

10.7 Filter Estimation Considerations

In this section we assume availability of the "channel-processed" spatial, temporal or space-time code needed to implement a particular receiver configuration. We note, however, that a sample average implementation of the expectation operation in (10.11), (10.35) or (10.42), for example, can serve as a supervised estimator of the corresponding channel processed signature. For instance, for joint-domain processing, $\widehat{\mathbf{w}}_{RAKE} = \frac{1}{J}\sum_{j=1}^{J} \mathbf{X}_j b_1(j)$, where $\{\mathbf{X}_j\}_{j=1}^{J}$ is a sequence of J space-time received data vectors. We recognize that if sample averaging extends over multiple multipath channel realizations ($J \to \infty$) then $\widehat{\mathbf{w}}_{RAKE}$ converges either to the direct path space-time signature of the signal of interest when Rician fading is assumed or to the $\mathbf{0}$ vector when no direct path exists, i.e., Rayleigh fading is assumed. In either case, in order to capture the prevailing full space-time effective signature, $\widehat{\mathbf{w}}_{RAKE}$ needs to be continuously reinitialized at a rate that is consistent with the fading rate or a corresponding forgetting factor needs to be embedded in the sample average procedure. We are careful to note, however, that estimation of the channel-processed signature through supervised or blind estimation procedures *is not* the primary subject of this work. What we are concerned with in this chapter is to minimize the induced BER of the overall system with a given channel-processed signature vector and a few data samples $\mathbf{X}_j$, $j = 1, 2, \cdots, J$. Similar considerations hold for the disjoint configurations.

In the example mentioned in the previous section, the fading rate for a mobile GPS receiver moving (along the line of sight) at 35 mph is about 70 Hz, while when the receiver is moving approximately perpendicular to the line of sight, then the channel varies at a frequency 0.1 Hz. The former case implies that the channel can vary faster than the information bit rate, while for the latter case only 5ML data are available. As a result, adaptive signal processing is constrained by short data record operational characteristics. This is, in fact, the basic motivation for the development of the auxiliary-vector GPS processor as explained below.

Besides the fixed MF-type structure, all other filter alternatives discussed in the previous section (that is MVDR or AV type) are adaptive in nature and have been presented/formulated under ideal conditions, that is, under the assumption that the involved covariance matrix $\mathbf{R}$ is known. Here $\mathbf{R}$ denotes either the spatial or temporal or space-time receiver input covariance matrix. In practice, however, $\mathbf{R}$ is unknown and it is sample-average estimated by a data record of finite size. When $\mathbf{R}$ is substituted by the sample average estimate $\widehat{\mathbf{R}}$ then the ideal receiver expressions in (10.14), (10.20), (10.23), (10.27), (10.36), (10.40) assume their estimated versions. In this context, the auxiliary-vector algorithm, in addition to its own merit in the context of MS-optimum signal processing theory, produces a sequence of MVDR filter estimators of the form $\widehat{\mathbf{w}}(0)$, $\widehat{\mathbf{w}}(1), \cdots$. This sequence was extensively studied in [27] and was shown to offer the means for effective control over the filter estimator bias versus covariance trade-off. As a result, adaptive filter estimators from this class have been seen to easily outperform in expected mean-square estimation error (constraint) LMS, sample-matrix-inversion (SMI) and RLS-type adaptive filter implementations. These operational characteristics of the AV filter estimators place them favorably in terms of GPS receiver implementation when adaptive interference suppression with short data records is the objective. Simulation comparisons in the following section illustrate how the above observations translate into superior BER performance.

10.8 Numerical and Simulation Comparisons

We consider the GPS signal model in (10.5) for a system with $M = 2$ antenna elements and spreading gain L =1023. In all cases we assume that 4 satellite signals with fixed C/A Gold codes are in view. In addition, we assume the presence of one or two high power spread-spectrum jammers (spoofers) that exhibit code cross-correlation with the signal of interest (*signal 1*) approximately 0.2 and 0.3, respectively. As explained in Section 10.1, high code cross-correlation models the jamming effect of a "GPS-looking" hostile signal. The angles of arrival of the satellite signals and jammers are randomly generated according to a uniform distribution in $(-\pi/2, \pi/2)$[2].

The simulation/numerical studies in this section evaluate the BER performance of the GPS receiver as a function of either the signal-to-noise-ratio (SNR) of the signal of interest or the data record size. All BER expressions are *analytically* evaluated[3]. The results for the one and two jammer case are averages over 100 and 500, respectively, independent space-time channels. For the BER versus SNR studies, the SNR of the signal of interest varies from 0 dB (weak signal due to foliage or terrain attenuation) to 15 dB (normal strength signal[4]). The SNRs of other satellite signals are fixed at 15 dB while the jammer SNR is fixed at 30 dB. The sampled AWGN is assumed to be *i.i.d.* across spatial channels (antenna elements).

In Figs. 10.5-10.8 different receiver configurations are compared in terms of their BER performance in the absence of multipath fading. Fig. 10.5 plots the BER versus the SNR of *signal 1* for the estimated disjoint S-T cascade and joint configurations with data support equal to 8ML. One jammer is present. We observe that the *sMF/tMV* configuration performs better than the *sMV/tMV* configuration. This behavior is in agreement with the theoretical studies in Section 10.3.3, since $Pr\{\eta > 0.25\} \approx 0.86$ for this simulation setup (the latter probability was evaluated numerically). In agreement with the findings in [25]–[27], AV designs outperform their MVDR counterparts. For this study both the JAV filter and tAV filter utilize 5 auxiliary vectors.

Fig. 10.6 plots the BER versus the data record size of estimated disjoint S-T configurations. Fig. 10.7 replicates the studies of Fig. 10.6 for the corresponding disjoint T-S configurations. Fig. 10.8 plots the BER

[2]Table 10.1 in the appendix presents the average spatial cross-correlation $\overline{\eta}$ (evaluated numerically) between two signals with DoA-separation angles in $[0, 0.2\pi]$, $(0.2\pi, \pi]$, and $[0, \pi]$ when an antenna array of $M = 2$, 3, 4, 8, or 10 elements is utilized. Uniform DoA distribution ($\Delta\theta \in [0, \pi]$) appears as a reasonable maximum-entropy assumption that represents/reflects the correlation characteristics of the jamming environment for various $\Delta\theta$ values.

[3]We recall that the received vectors are distributed according to a Gaussian mixture pdf; thus the BER is analytically evaluated as a sum of $Q(\cdot)$ functions over all possible bit combinations of the interferers weighted by the probability of the corresponding bit combination (equally probable bit combinations in our case).

[4]Under normal operating conditions the nominal thermal noise power spectral density is -205 dBW/Hz [1], which results in -142 dBW noise power in 2 MHz bandwidth. Also, the minimum specified C/A signal power at the receiver end is -160 dBW while the actual received power corresponding to a conservative link budget is about -157 dBW. The latter leads to -157 dBW$-$ (-142 dBW) $= -15$ dB SNR per code chip which, after despreading (that is 30 dB gain), results in -15 dB + 30 dB = 15 dB pre-detection SNR.

versus the data record size for the estimated joint-domain configurations. The SNR of *signal 1* in Figs. 10.6-10.8 is fixed at 15 dB, two jammers are present, the sAV filters utilize one auxiliary vector, while JAV and tAV filters utilize 5 auxiliary vectors.

Fig. 10.9 shows the BER performance as a function of the data record size for the joint-domain configurations and a few selected disjoint configurations in the presence of multipath fading (all AV filters utilize 5 auxiliary vectors). One jammer is present and each signal experiences $N = 3$ independent paths that are $5T_c$ apart with angles of arrival uniformly distributed in $(-\pi/2, \pi/2)$. The SNR of the SS jammer is fixed at 30 dB while the total SNR of the GPS signal of interest (the sum of the received SNR over all paths) is fixed at 15 dB. Following the notation in (10.41) the total received SNR for an SS signal k, $k = 1, 2, \cdots, K$, is defined by $E_k \sum_{n=1}^{N} E\{|\alpha_{k,n}|^2\}/\sigma^2$, where σ^2 is the variance of the sampled AWGN.

Figs. 10.5-10.9 illustrate the performance gains when non-MF-type signal processing is performed by the GPS receiver and do not consider combining. Additional performance gains obtained through EGC combining are illustrated in Fig. 10.10 where we plot the BER as a function of the data record size for selected receiver configurations of the previous figures.

10.9 Conclusions

Disjoint space and time domain as well as joint space-time signal processing algorithms have been investigated for GPS navigation data demodulation in the presence of multiple-access satellite interference, intelligent SS interference highly correlated with the signal of interest and AWGN. In particular, linear GPS filters were derived based on MF, MVDR and AV processing principles. The relative merits of different disjoint domain configurations were analytically evaluated and summarized in Theorem 10.1. Simulation results illustrated the performance of GPS filter estimators of both disjoint and joint domain configurations under different interference conditions, channel conditions, and data record sizes. It was demonstrated that joint domain adaptive processing configurations have advantages over disjoint domain schemes. However, these advantages quickly diminish under limited data support where disjoint (or joint) auxiliary-vector-receivers were seen to offer effective SS interference suppression and superior BER performance. Finally, further BER performance improvements under limited data support were sought through output filter combining methods that exploited the redundancy inherent in the GPS signal due to the 20-times repetition of the *C/A* code in a single information bit period.

10.10 Appendix

10.10.1 Proof of Theorem 10.1

Let E_i and $\Gamma_i \triangleq E_i/\sigma^2$ denote the input energy and input SNR, respectively, of signal i, *i=1, 2* (signal *1* is the signal of interest), where σ^2 is the input noise variance. Let also $\varrho \triangleq \mathbf{d}_1^H \mathbf{d}_2$ and $\zeta \triangleq \frac{\mathbf{a}_1^H \mathbf{a}_2}{M}$ denote the temporal and (complex) spatial cross-correlation, respectively (we note that $\rho = |\varrho|$ and $\eta = |\zeta|$).

Part A We first evaluate the spatial and temporal filter as well as the overall system output interference-plus-noise variance P_{IN} as a function of Γ_1, Γ_2, ρ and η. Since all design alternatives involve filters that preserve the energy of the signal of interest (*signal 1*), the output variance P_{IN} is inversely proportional to output SINR.

sMF/tMF

$$\mathbf{w}_{sMF} = \mathbf{a}_1/M \tag{10.43}$$

$$\mathbf{w}_{tMF} = \mathbf{d}_1 \tag{10.44}$$

$$\frac{P_{IN}(sMF/tMF)}{\sigma^2} = \Gamma_2 \eta^2 \rho^2 + \frac{1}{M}. \tag{10.45}$$

sMF/tMVDR

$$\mathbf{w}_{sMF} = \mathbf{a}_1/M \tag{10.46}$$

$$\mathbf{w}_{tMV} = \frac{1/\Gamma_2'+1}{1/\Gamma_2'+1-\rho^2}\left(\mathbf{d}_1 - \frac{\varrho}{1/\Gamma_2'+1}\mathbf{d}_2\right) \tag{10.47}$$

$$\frac{P_{IN}(sMF/tMV)}{\sigma^2} = \frac{\Gamma_2\eta^2\rho^2}{[1+\Gamma_2'(1-\rho^2)]^2} + \frac{1}{M}\left[1 + \frac{\rho^2(1-\rho^2)}{(1/\Gamma_2'+1-\rho^2)^2}\right] \tag{10.48}$$

where

$$\Gamma_2' \triangleq M\Gamma_2\eta^2. \tag{10.49}$$

sMVDR/tMVDR

$$\mathbf{w}_{sMV} = \frac{1/\Gamma_2''+1}{1/\Gamma_2''+1-\eta^2}\left(\mathbf{a}_1 - \frac{\zeta^*}{1/\Gamma_2''+1}\mathbf{a}_2\right)/M \tag{10.50}$$

$$\mathbf{w}_{tMV} = \frac{1/\Gamma_2'''+1}{1/\Gamma_2'''+1-\rho^2}\left(\mathbf{d}_1 - \frac{\varrho}{1/\Gamma_2'''+1}\mathbf{d}_2\right) \tag{10.51}$$

$$\frac{P_{IN}(sMV/tMV)}{\sigma^2} = \frac{\Gamma_2\eta^2\rho^2}{[1+\Gamma_2''(1-\eta^2)]^2[1+\Gamma_2'''(1-\rho^2)]^2} + \frac{1}{M}\left[1 + \frac{\eta^2(1-\eta^2)}{(1/\Gamma_2''+1-\eta^2)^2}\right]\left[1 + \frac{\rho^2(1-\rho^2)}{(1/\Gamma_2'''+1-\rho^2)^2}\right] \tag{10.52}$$

where

$$\Gamma_2'' \triangleq \frac{M\Gamma_2}{L} \tag{10.53}$$

$$\Gamma_2''' \triangleq \frac{M\Gamma_2\eta^2}{(1+\Gamma_2'')^2 - \Gamma_2''(2+\Gamma_2'')\eta^2}. \tag{10.54}$$

sMVDR/tMF

The $\mathbf{w}_{sMV}$ and $\mathbf{w}_{tMF}$ filters are given by (10.50) and (10.44) respectively. Then,

$$\frac{P_{IN}(sMV/tMF)}{\sigma^2} = \frac{\Gamma_2\eta^2\rho^2}{[1+\Gamma_2''(1-\eta^2)]^2} + \frac{1}{M}\left[1 + \frac{\eta^2(1-\eta^2)}{(1/\Gamma_2''+1-\eta^2)^2}\right] \tag{10.55}$$

where Γ_2'' is given by (10.53). The relations (i) and (ii) follow by direct comparison of (10.45), (10.48), (10.52) and (10.55). To show *(iv)* we observe that the right hand side of (10.48) is equal to

$$\frac{\Gamma_2\eta^2\rho^2}{1+\Gamma_2'(1-\rho^2)} + \frac{1}{M}. \tag{10.56}$$

The second term (noise power) of the right hand side of (10.52) is always greater than the second term of the right hand side of (10.56) and the sufficient condition can be derived by comparing the corresponding first terms.

Part B Part B of Theorem 10.1 can be proved in a similar manner after deriving the dual expressions of (10.43)-(10.55) by direct substitution of ρ^2 by η^2, ϱ by ζ^*, $\mathbf{a}_1$ by $\mathbf{d}_1$, $\mathbf{a}_2$ by $\mathbf{d}_2$, and vice versa.

Part C Part C can be proved by direct comparison of the involved expressions (i.e. (10.55) and (10.48)) obtained in Parts A and B above. □

10.10.2 Average Spatial Cross-correlation versus Number of Antenna Elements

Table 10.1 below shows the average spatial cross-correlation $\overline{\eta}$ between two signals whose DoAs are $\Delta\theta$ apart, as a function of the number of antenna elements M.

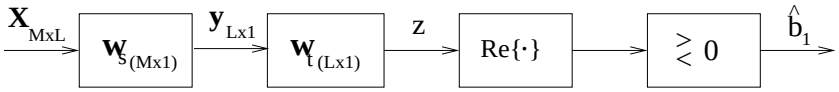

Fig. 10.1. Space-time cascade receiver structure.

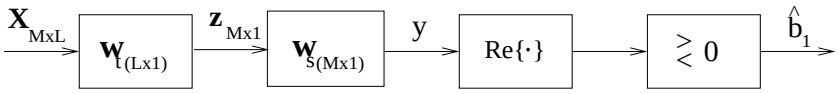

Fig. 10.2. Time-space cascade receiver structure.

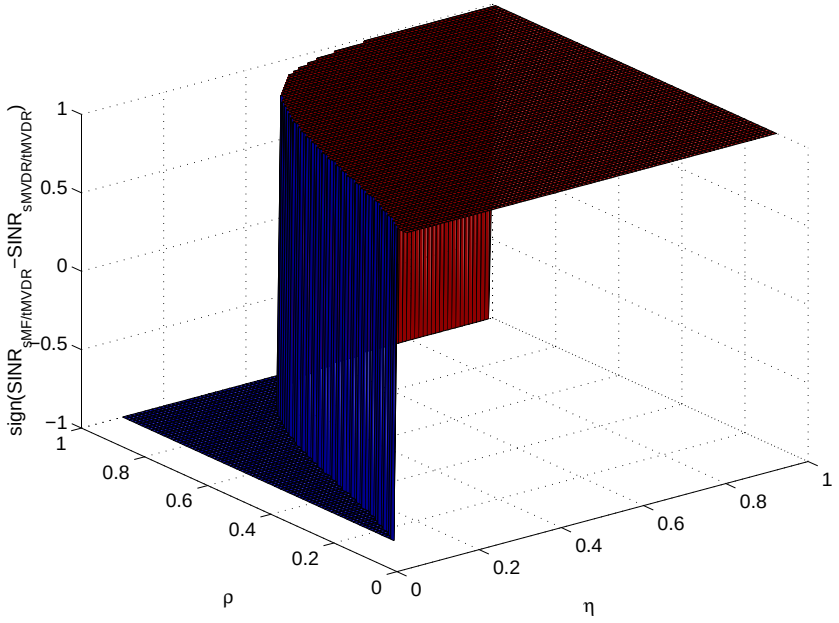

Fig. 10.3. Output SINR comparisons for the disjoint $sMF/tMVDR$ and $sMVDR/tMVDR$ configurations as a function of the spatial cross-correlation η and the temporal cross-correlation ρ.

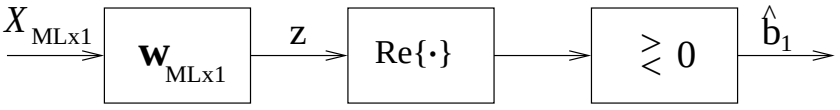

Fig. 10.4. Joint space-time receiver structure.

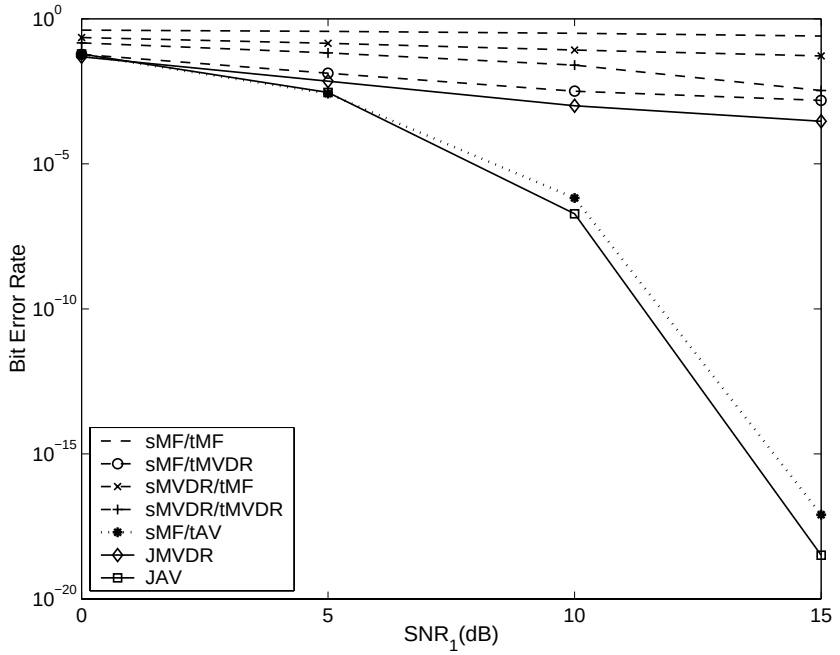

Fig. 10.5. Bit-Error-Rate as a function of the SNR of the signal of interest for disjoint S-T and joint S-T configurations in the presence of one jammer (data record size =8ML).

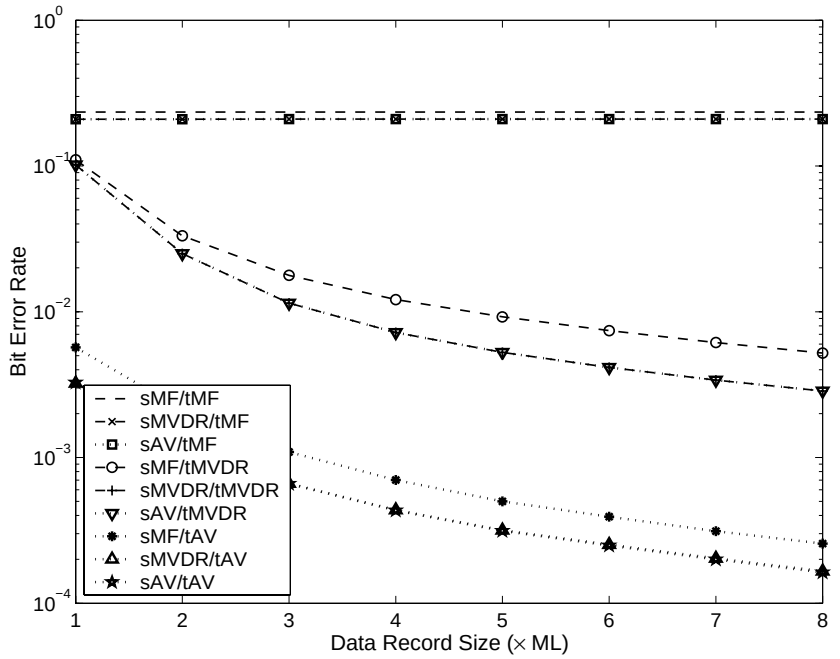

Fig. 10.6. Bit-Error-Rate as a function of the data record size for disjoint S-T configurations (SNR_1=15 dB, two jammers present).

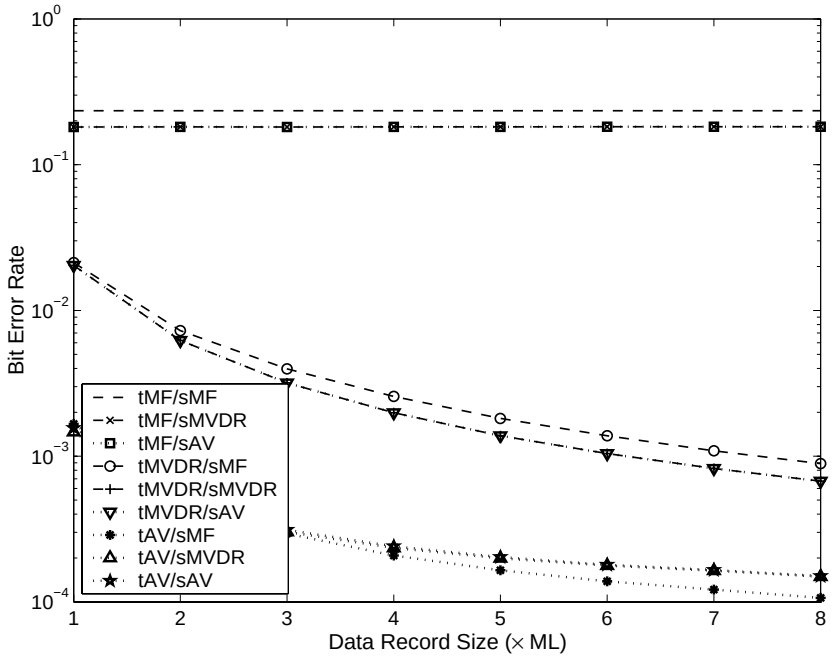

Fig. 10.7. Bit-Error-Rate as a function of the data record size for disjoint T-S configurations (SNR_1=15 dB, two jammers present).

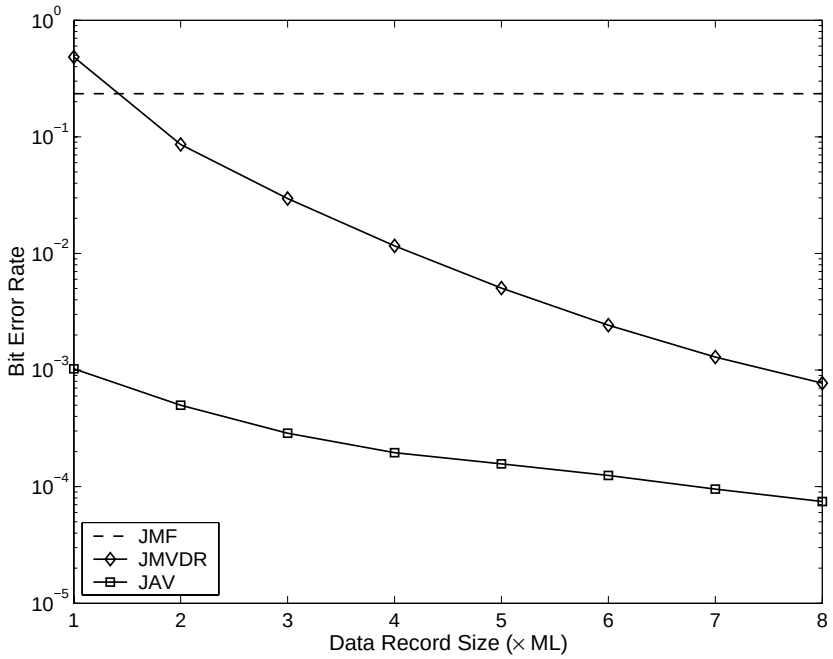

Fig. 10.8. Bit-Error-Rate as a function of the data record size for joint space-time configurations (SNR_1=15 dB, two jammers present).

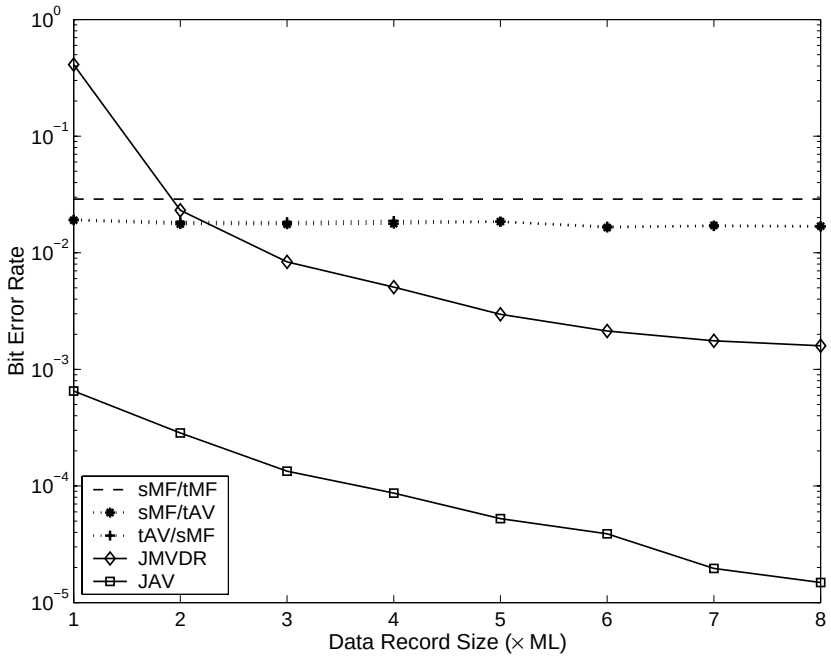

Fig. 10.9. Bit-Error-Rate as a function of the data record size under multipath fading (total SNR_1=15 dB, one jammer present).

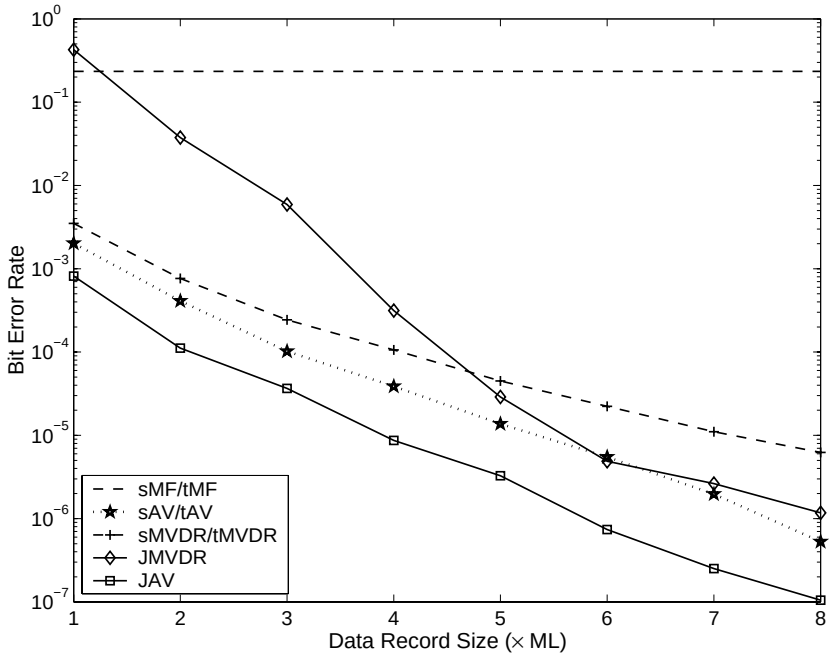

Fig. 10.10. Bit-Error-Rate as a function of the data record size after EGC combining (SNR_1=15 dB, two jammers present).

Table 10.1. $\overline{\eta}$ vs. M

M	2	3	4	5	8	10
$0 \leq \Delta\theta \leq 0.2\pi$	0.91	0.78	0.67	0.59	0.44	0.38
$\Delta\theta = 0.2\pi$	0.70	0.35	0.26	0.23	0.13	0.09
$0.2\pi < \Delta\theta \leq \pi$	0.54	0.37	0.30	0.25	0.17	0.14
$0 \leq \Delta\theta \leq \pi$	0.67	0.52	0.44	0.38	0.27	0.23

Bibliography

[1] B. W. Parkinson, J. J. Spilker, P. Axelrad, and P. Enge Editors, *Global Positioning System: Theory and Applications I*. American Institute of Aeronautics and Astronautics, 1995.

[2] P. Enge and P. Misra, "Scanning the issue/technology," *Proceedings of the IEEE*, vol. 87, pp. 3-15, Jan. 1999.

[3] P. Enge, "Local area augmentation of GPS for the precision approach of aircraft," *Proceedings of the IEEE*, vol. 87, pp. 111-132, Jan. 1999.

[4] E. Abbott and D. Powell, "Land-vehicle navigation using GPS," *Proceedings of the IEEE*, vol. 87, pp. 145-162, Jan. 1999.

[5] P. Axelrad and C. P. Behre, "Satellite attitude determination based on GPS signal-to-noise ratio," *Proceedings of the IEEE*, vol. 87, pp. 133-144, Jan. 1999.

[6] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communication*. Reading, MA: Addison-Wesley, 1995.

[7] M. K. Simon, J. K. Omura, R. A. Sholtz, and B. K. Levitt, *Spread Spectrum Communications Handbook*, 2nd ed. New York: McGraw-Hill, 1994.

[8] R. L. Peterson, R. E. Ziemer, and D. E. Borth, *CDMA: Introduction to Spread Spectrum Communications*. Upper Saddle River, NJ: Prentice-Hall, 1995.

[9] R. A. Iltis, "An EKF-based joint estimator for interference, multipath, and code delay in a DS spread-spectrum receiver," *IEEE Trans. Commun.*, vol. 42, pp. 1288-1299, Feb./Mar./Apr. 1994.

[10] R. A. Iltis and G. A. Hanson, "C/A code tracking and acquisition with interference rejection using extended Kalman filter," in *Proc. 1999 ION National Technical Meeting & 19th Biannual Guidance Test Symposium*, pp. 881-889.

[11] R. L. Fante and J. J. Vacarro, "Cancellation of jammers and jammer multipath in a GPS receiver," *IEEE AES Systems Magazine*, pp. 25-28, Nov. 1998.

[12] N. A. White, P. S. Maybeck, and S. L. DeVilbiss, "Detection of interference/jamming and spoofing in a DGPS-aided inertial system," *IEEE Trans. Aerospace and Electronics Sys.*, pp. 1208-1217, Oct. 1998.

[13] S. Bartlett, R. LeGrand, and E. Staton, "Analysis and simulation of discrete Kalman filtering applied to a GPS receiver," *Naval Engineers Journal*, pp. 29-36, Jan. 1996.

[14] A. M. Haimovich and A. Shah, "The performance of space-time processing for suppressing narrowband interference in CDMA communications," *Wireless Personal Commun.*, pp. 233-255, 1998.

[15] R. Gold, "Optimal binary sequences for spread sprectrum multiplexing," *IEEE Trans. Inform. Theory*, vol. 13, pp. 619-621, Oct. 1967.

[16] D. Sarwate and M. B. Pursley, "Cross-correlation properties of pseudorandom and related sequences," *Proceedings of the IEEE*, pp. 593-619, May 1980.

[17] W. Zhuang and J. Tranquilla, "Digital baseband processor for the GPS receiver modeling and simulations," *IEEE Trans. Aerospace and Electronics Sys.*, pp. 1343-1349, Oct. 1993.

[18] M. S. Braasch and A. J. Van Dierendonck, "GPS receiver architectures and measurements," *Proceedings of the IEEE*, vol. 87, Jan. 1999.

[19] S. C. Fisher and K. Ghassemi, "GPS IIF–The next generation," *Proceedings of the IEEE*, vol. 87, pp. 24-47, Jan. 1999.

[20] C. C. Counselman, III, "Multipath-rejecting GPS antennas," *Proceedings of the IEEE*, vol. 87, pp. 86-91, Jan. 1999.

[21] P. Misra, B. P. Burke, and M. M. Pratt, "GPS performance in navigation," *Proceedings of the IEEE*, vol. 87, pp. 65-85, Jan. 1999.

[22] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[23] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[24] S. Verdú, *Multiuser Detection*. Cambridge, U. K.: Cambridge Univ. Press, 1998.

[25] A. Kansal, S. N. Batalama, and D. A. Pados, "Adaptive maximum SINR RAKE filtering for DS-CDMA multipath fading channels," *IEEE J. Select. Areas Commun.*, pp. 1965-1973, Dec. 1998.

[26] D. A. Pados and S. N. Batalama, "Joint space-time auxiliary-vector filtering for DS-CDMA systems with antenna arrays," *IEEE Trans. Commun.*, vol. 47, pp. 1406-1415, Sept. 1999.

[27] D. A. Pados and G. N. Karystinos, "An iterative algorithm for the computation of the MVDR filter," *IEEE Trans. Signal Processing*, vol. 49, pp. 290-300, Feb. 2001.

[28] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Magazine*, vol. 14, pp. 49-83, Nov. 1997.

[29] H. Qian and S. N. Batalama, "Data-record-based criteria for the selection of an auxiliary-vector estimator of the MVDR filter," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, USA, 2000, vol. 1, pp. 802-807.

[30] X. Bernstein and A. M. Haimovich, "Space-time optimum combining for CDMA communications," *Wireless Personal Commun.*, pp. 73-89, 1996.

[31] R. Price and P. E. Green, " A communication technique for multipath channels," *Proc. IRE*, pp. 555-570, Mar. 1958.

Chapter 11

Low Complexity Decoding of Finite Geometry LDPC Codes

11.1 Introduction

Low-density parity-check (LDPC) codes, originally introduced by Gallager [1, 2] and brought into prominence by MacKay and Neal [3, 4], are currently attracting a great deal of research interest. Indeed, practical decoding algorithms operating on certain long LDPC codes were seen to offer exceptional error protection. On the other hand, although various code design techniques are available, most of the "good" LDPC codes known so far are generated randomly by computer and are lacking a certain algebraic structure. This impairs our ability to carry out theoretical analysis or develop cost effective encoder implementations.

A notable exception to the above is the algebraically constructed finite geometry LDPC (FG-LDPC) codes [5]. Most of the FG-LDPC codes are cyclic or quasi-cyclic; therefore, their encoding can be easily implemented by linear shift registers with feedback connections. Compared to the computer generated LDPC codes with the same code length and dimension, FG-LDPC codes have a more balanced structure and better performance, especially in the range of short to moderate code lengths ($10^2 - 10^4$). The parity-check matrices of FG-LDPC codes employ some redundant checks (the number of rows of the parity-check matrix is greater than the difference between the code length and the code dimension) and row/column weights are also somewhat higher than usual LDPC codes. This implies higher decoding complexity than the computer generated LDPC codes, which may be the price for the improved performance.

In this chapter, we propose a new decoding algorithm for high-rate FG-LDPC codes. LDPC codes draw their error correction strength from their relatively long length in comparison with conventional linear block codes. Undetected errors at the decoder are exceedingly rare. The codewords are distributed sparsely in the N-dimensional space over $GF(2)$ and in almost all cases we either find the correct codeword or cannot find any codeword at all. In other words, when we search around the received vector -according to a properly defined distance metric- we expect that the first valid codeword to encounter will be the correct one with exceedingly high probability. In this work, we begin indeed with the hard-decision received vector of length N and we search around it iteratively. In each iteration, we flip only one bit of the length-N vector. The bit to be flipped is chosen according to a selection metric that is supposed to conform well with the characteristics of FG-LDPC codes (many cycles of length six, redundant checks and somewhat higher than usual row/column weights). The metric combines the number of failed checks that bits participate in and their reliability. After each iteration, the metric values are updated as needed. Calculation of the bit selection metric requires no knowledge of the signal energy or noise power characteristics of additive white Gaussian noise (AWGN) channels. If a valid codeword is found, the search is terminated and successful decoding is claimed. If no codeword is found after a certain number of iterations (search upper bound), then we declare a decoding failure. Last but not least, during decoding we can record a window of searched vectors along the search path to identify and break closed loops. This is an optional modification that can improve the decoding performance without much complexity increase.

11.2 The Decoding Algorithm

11.2.1 Notation and Basic Definitions

Consider a binary (N, K) LDPC code with length N and dimension K. We choose an arbitrary codeword $\mathbf{c} = [c_1, c_2, \cdots, c_N]$. The codeword is transmitted over a symmetric binary-input memoryless discrete-time channel and the received vector is denoted by $\mathbf{r} = [r_1, r_2, \cdots, r_N]$. Here, by "symmetric channel" we mean that $f(r|c = 1) = f(-r|c = 0)$ where $f(\cdot|\cdot)$ is the conditional probability distribution function of the channel output. We define the log-likelihood value for each channel output r_n, $n = 1,\ 2,\ \ldots,\ N$,

$$y_n \triangleq \log \frac{f(r_n|c_n = 1)}{f(r_n|c_n = 0)} \tag{11.1}$$

and the initial hard-limited bit decision

$$b_n^{(0)} \triangleq \mathrm{sgn}(y_n) \tag{11.2}$$

where $\mathrm{sgn}(y) = 1$ if $y \geq 0$ and $\mathrm{sgn}(y) = 0$ if $y < 0$. The absolute value of y_n is called the *reliability* of the initial decision $b_n^{(0)}$. The vectors $[|y_1|, |y_2|, \cdots, |y_N|]$ and $\mathbf{b}^{(0)} \triangleq [b_1^{(0)}, b_2^{(0)}, \cdots, b_N^{(0)}]$ are the inputs to the decoder.

A binary LDPC code is completely described by its sparse binary parity-check matrix $\mathbf{H}$. For an (N, K) LDPC code, $\mathbf{H}$ has N columns and $M \geq N - K$ rows. For any tentative bit decision vector $\mathbf{b}$, the set of check sums -or syndrome- is the vector $\mathbf{s} = \mathbf{b}\mathbf{H}^T$. For a *regular* LDPC code, $\mathbf{H}$ has a constant column weight t_c and a constant row weight t_r. Following exactly the notation of [4], we define the set $\mathcal{N}(m)$ that contains the indices of the t_r bits that participate in check m and the set $\mathcal{M}(n)$ that contains the indices of the t_c checks in which bit n participates:

$$\mathcal{N}(m) \triangleq \{n : \mathbf{H}_{mn} = 1\}, \quad \mathcal{M}(n) \triangleq \{m : \mathbf{H}_{mn} = 1\}. \tag{11.3}$$

For each *check* $1 \leq m \leq M$, we define the "*lower check reliability*" value l_m and the "*upper check reliability*" value u_m as follows:

$$l_m \triangleq \min_{n \in \mathcal{N}(m)} |y_n|, \quad u_m \triangleq \max_{n \in \mathcal{N}(m)} |y_n|. \tag{11.4}$$

11.2.2 Algorithm

The proposed algorithm starts from the hard-limited received data vector $\mathbf{b}^{(0)}$ and searches iteratively the N-dimensional vector space over $GF(2)$ for a valid codeword vector. Assume that prior to the kth iteration, $k \geq 1$, the bit vector under consideration is $\mathbf{b}^{(k-1)}$ with corresponding syndrome vector $\mathbf{s}^{(k-1)} = \mathbf{b}^{(k-1)}\mathbf{H}^T$. The objective is to perturb $\mathbf{b}^{(k-1)}$ by one bit and create a new candidate bit vector $\mathbf{b}^{(k)}$. To choose which bit of $\mathbf{b}^{(k-1)}$ to flip we define for each bit $n = 1,\ 2,\ \ldots,\ N$ the cumulative metric over all checks in $\mathcal{M}(n)$

$$\phi_n^{(k)} \triangleq \sum_{m \in \mathcal{M}(n)} \phi_{n,m}^{(k)}, \quad 1 \leq n \leq N, \tag{11.5}$$

where for each check $m \in \mathcal{M}(n)$

$$\phi_{n,m}^{(k)} \triangleq \begin{cases} |y_n| - l_m/2, & \text{if } s_m^{(k-1)} = 0 \\ |y_n| - (u_m + l_m/2), & \text{if } s_m^{(k-1)} = 1.^1 \end{cases} \tag{11.6}$$

Then, we flip the bit in $\mathbf{b}^{(k-1)}$ with minimum $\phi_n^{(k)}$ value to obtain $\mathbf{b}^{(k)}$:

$$n^{(k)} = \arg \min_{1 \leq n \leq N} \phi_n^{(k)} \quad \text{and} \quad \mathbf{b}^{(k)} = \mathbf{b}^{(k-1)} \oplus \mathbf{e}^{(n^{(k)})} \tag{11.7}$$

where $\oplus$ denotes addition over GF(2) and $\mathbf{e}^{(n^{(k)})}$ is the N-dimensional unit vector with 1 at the $n^{(k)}$th position and 0 everywhere else. We update the syndrome vector to $\mathbf{s}^{(k)}$ from the previous syndrome $\mathbf{s}^{(k-1)}$

($\mathbf{b}^{(k)}$ and $\mathbf{b}^{(k-1)}$ differ only in one bit) and compare with the all-zero vector. If $\mathbf{s}^{(k)} = \mathbf{0}$, then we have a valid codeword and we stop the search. If not, we calculate $\phi_n^{(k+1)}$, $n = 1, 2, \ldots, N$, and continue.

The basic properties of the selection metric in (11.5), (11.6) are as follows. *(i)* If bit n passes check $m \in \mathcal{M}(n)$, then $\phi_{n,m} \in [l_m/2, u_m - l_m/2]$. *(ii)* If bit n fails check $m \in \mathcal{M}(n)$, then $\phi_{n,m} \in [-(u_m - l_m/2), -l_m/2]$. *(iii)* Among all bits that participate in check m, the selection rule favors the least reliable ((11.6) incorporates $|y_n|$ as a fixed term). *(iv)* Over all bits and all checks, the selection rule favors bits that participate in failed checks (because $\phi_{n,m} > 0$ if $s_m = 0$ and $\phi_{n,m} < 0$ if $s_m = 1$). In essence, the selection rule combines "the number of failed checks" and "the reliability" into a new criterion function ϕ_n. The criterion function ϕ_n maintains linearity with respect to $|y_n|$, therefore for AWGN channels we can replace $|y_n|$ by $|r_n|$ and still have an equivalent decoder that requires no knowledge of signal energy/noise power.

11.3 Improvement: Loop Detection

In the algorithm that we described in the previous section, the selection of the bit to flip depends only on the current bit vector $\mathbf{b}^{(k-1)}$ and the reliability vector $[|y_1|, \cdots, |y_N|]$:

$$n^{(k)} = \arg\min_n \phi_n(\mathbf{b}^{(k-1)}, [|y_1|, \cdots, |y_N|]). \tag{11.8}$$

If in the kth iteration the new bit vector $\mathbf{b}^{(k)}$ coincides with a previously considered vector $\mathbf{b}^{(k_0)}$, $k_0 < k$, then $n^{(k+1)} = n^{(k_0+1)}$ and $\mathbf{b}^{(k+1)} = \mathbf{b}^{(k_0+1)}$. By induction, $\mathbf{b}^{(k+\rho)} = \mathbf{b}^{(k_0+\rho)}$ for any $\rho > 0$ and $\mathbf{b}^{(i)}$ is seen to be periodic. Since there is no valid codeword vector in $\left\{\mathbf{b}^{(k_0)}, \mathbf{b}^{(k_0+1)}, \cdots, \mathbf{b}^{(k-1)}\right\}$, we will not find a valid codeword vector if we continue the search. The decoding process is trapped into an infinite loop and a decoding failure will be reported when the maximum allowable iteration number is reached.

To detect and break such infinite loops when they appear, we can record all searched vectors; if in the kth iteration we attempt to generate a previously searched vector $\mathbf{b}^{(k_0)}$, $k_0 < k$, we will discard this bit-flip choice and select instead the bit immediately next in ϕ_n value. At first sight, it appears that to record and compare with all searched vectors has complexity proportional to the code length N (which may be prohibitive for long LDPC codes). The cost, however, can be greatly reduced as we see in the sequel.

For convenience, let us first define the vector sum $\mathbf{E}(x) \triangleq \sum_{i=x+1}^{k} \mathbf{e}^{(n^{(i)})}$. From (11.7) and by induction we have $\mathbf{b}^{(k)} = \mathbf{b}^{(k_0)} \oplus \sum_{i=k_0+1}^{k} \mathbf{e}^{(n^{(i)})} = \mathbf{b}^{(k_0)} \oplus \mathbf{E}(k_0)$ for any $0 \le k_0 < k$. Naturally, we reach the important conclusion that

$$\mathbf{b}^{(k)} = \mathbf{b}^{(k_0)} \quad \text{if and only if} \quad \mathbf{E}(k_0) = \mathbf{0}. \tag{11.9}$$

This means that, to detect and prevent infinite loops, we need only the vectors $\mathbf{E}(k-1)$, $\mathbf{E}(k-2)$, $\cdots$, $\mathbf{E}(0)$. If any of these vectors is zero, an infinite loop is detected and the bit selection $n^{(k)}$ should be discarded. We notice also that since $\mathbf{E}(k-1), \cdots, \mathbf{E}(0)$ depend only on $\left\{n^{(1)}, n^{(2)}, \cdots, n^{(k)}\right\}$, there is no need to record the searched bit vectors $\mathbf{b}^{(k-1)}$, $\cdots$, $\mathbf{b}^{(0)}$ (which otherwise would impose a significant storage requirement for long codes).

In more detail, we can calculate $\mathbf{E}(l)$ iteratively:

$$\mathbf{E}(l-1) = \begin{cases} \mathbf{e}^{(n^{(k)})}, & \text{if} \quad l = k \\ \mathbf{E}(l) \oplus \mathbf{e}^{(n^{(l)})}, & \text{otherwise.} \end{cases} \tag{11.10}$$

To compare $\mathbf{E}(l)$, $l = 0, 1, \cdots, k-1$, with the all-zero vector, we denote its *Hamming weight* by $w(l)$. Then, $w(l-1)$ can be easily derived from $w(l)$ without counting the number of 1's in $\mathbf{E}(l)$:

$$w(l-1) = \begin{cases} 1, & \text{if } \ l = k \\ w(l) - 1, & \text{if } \ n^{(l)}\text{th bit of } \mathbf{E}(l) = 1 \\ w(l) + 1, & \text{if } \ n^{(l)}\text{th bit of } \mathbf{E}(l) = 0. \end{cases} \tag{11.11}$$

Of course, $\mathbf{E}(l) = \mathbf{0}$ if and only if $w(l) = 0$.

The proposed loop detection procedure does not depend on our specific criterion function ϕ_n and can be used in conjunction with other bit-flipping-type decoding algorithms as well.

[1] $l_m/2$ and $u_m + l_m/2$, $m = 1, 2, \ldots, M$, need to be calculated only once before the first iteration of the algorithm.

Table 11.1. Decoding Complexity per Iteration (Unit One Real Operation)

Algorithm	Multiplications	Divisions	Additions
BP algorithm	$11Nt_c - 6(N+M)$	$N(t_c+1)$	$N(3t_c+1)$
Normalized BP-based [7]	0	Nt_c	$N(4t_c-3)+M(\lceil\log_2 t_r\rceil-2)$
Normalized APP-based [8]	0	Nt_c	$2Nt_c+M(\lceil\log_2 t_r\rceil-2)$
UMP BP-based (min-sum) [6, 9]	0	0	$N(4t_c-3)+M\lceil\log_2 t_r\rceil-2)$
Weighted BF [5]	0	0	$N-1+t_ct_r$
Proposed algorithm	0	0	$N-1+t_ct_r$

In summary, the complete decoding algorithm of Section 11.2 that incorporates loop detection and prevention is as follows:

(i) Initialization: Iteration counter $k \leftarrow 0$; "Exclusion list" $B \leftarrow \emptyset$; $l_m/2$, $u_m + l_m/2$, $m = 1, 2, \ldots, M$, calculated by (11.4); $\mathbf{b}^{(0)}$ calculated by (11.2).

(ii) Calculate syndrome $\mathbf{s}^{(k)}$; if $\mathbf{s}^{(k)} = \mathbf{0}$, then return $\mathbf{b}^{(k)}$ and STOP.

(iii) $k \leftarrow k+1$; if $k > k_{\max}$ (user specified maximum iteration number), then declare decoding failure and STOP.

(iv) Calculate $\phi_n^{(k)}$, $n = 1, 2, \ldots, N$, by (11.5), (11.6).

(v) $n^{(k)} \leftarrow \arg\min_{n \notin B} \phi_n^{(k)}$.

(vi) Calculate $\mathbf{E}(l)$, $w(l)$ by (11.10), (11.11), $l = k-1, k-2, \cdots, 0$; if $w(l) = 0$ for any $l = k-1, \cdots, 0$, then $B \leftarrow B \cup \{n^{(k)}\}$ and go back to Step 5.

(vii) Calculate $\mathbf{b}^{(k)}$ by (11.7); $B \leftarrow \emptyset$; go to Step 2.

11.4 Computational Complexity

The computational complexity of the decoding algorithm described above can be calculated as follows. During each iteration, one bit of the N-dimensional bit vector $\mathbf{b}$ is flipped. To select this bit according to the metric ϕ_n, $N-1$ real comparisons are needed. After bit n is selected and flipped, t_c syndrome bits that correspond to the t_c checks in which bit n participated should be updated (flipped):

$$s_m \leftarrow 1 - s_m, \quad m \in \mathcal{M}(n). \tag{11.12}$$

Then, for each $m \in \mathcal{M}(n)$, we update ϕ_n for all $n \in \mathcal{N}(m)$:

$$\phi_n \leftarrow \begin{cases} \phi_n + u_m, & \text{if } s_m = 0 \\ \phi_n - u_m, & \text{if } s_m = 1 \end{cases}, \quad n \in \mathcal{N}(m). \tag{11.13}$$

Since comparisons can be considered as additions, we need altogether $N - 1 + |\mathcal{N}||\mathcal{M}| = N - 1 + t_ct_r$ real additions per iteration. Both the comparison and the update of ϕ_n can be parallelized and accomplished in $\lceil\log_2 N\rceil$ and t_r time units, respectively.

Loop detection, as described in Section 11.3, requires only bit operations and integer counter increments/decrements. Most importantly, the total number of loop detection operations is independent of the block length N and depends only on the total number of executed iterations. Therefore the cost of loop detection in comparison with the $N - 1 + t_ct_r$ real additions required by the core decoding procedure can be safely ignored.

In Table 11.1, we summarize the decoding complexity in real number operations per iteration for several algorithms: Belief propagation (BP), three simplified versions of BP, namely the "normalized BP-based" [7], "normalized APP-based" [8], and "uniformly most powerful (UMP) BP-based" [6] algorithms, the "weighted

bit-flipping (BF)" algorithm [5], and, as a last entry, our proposed algorithm. The UMP BP-based algorithm coincides with the "min-sum" algorithm [9, Algorithm B] although their derivations are different. UMP BP-based (min-sum), weighted BF, and the proposed algorithm all require no multiplication/division operations at all. It is also important to note that for certain LDPC codes the normalized BP-based and the normalized APP-based divisions may be with respect to a power of 2, in which case they degenerate to trivial binary shift operations [7, 8].

Analytic statistical evaluation of the total number of iterations that are required per codeword is not an easy task for either of the above algorithms. For the proposed algorithm in particular, since only one bit is flipped in each iteration, we may suggest intuitively and validate experimentally a statistical model under which the average number of iterations is proportional to the mean number of bits in error in the hard-decision received vector. The mean number of iterations for the proposed algorithm is expected to be higher than BP; however, since the operations per iteration are much fewer, we can count on significant complexity savings even compared with simplified versions of BP. We should also point out that parallel implementations of BP decoders still need approximately $10t_c + 5t_r$ time units per iteration assuming availability of $\max\{M, N\}$ processors. With $\lceil \log_2 N \rceil + t_r$ times units per iteration and $N/2$ processors, our proposed algorithm parallelized would maintain speed advantage for reasonably long codes at half the hardware complexity.

11.5 Simulation Results

In this section, we evaluate the BER of various LDPC decoders as a function of the signal-to-noise ratio (SNR). To allow performance comparisons across different rate codes we use the normalized SNR per information bit $\frac{\gamma_b}{N_0} \triangleq \frac{E_c}{N_0}\frac{K}{N}$. We study the following decoders: *(i)* BP, *(ii)* UMP BP-based [6] (min-sum), *(iii)* normalized BP-based [7], *(iv)* normalized APP-based [8], *(v)* Gallager's original bit-flipping (BF) algorithm [2], *(vi)* weighted BF introduced in [5], and *(vii)* the algorithm proposed herein. The maximum iteration number is set equal to 200 for all algorithms. In Fig. 11.1, we present our findings for the Type-I 2-D (1023, 781) EG-LDPC code with rate 0.763 [5]. The normalization factor for the two normalized decoders is set equal to 4.0 based on the analysis in [7]. We see that the proposed algorithm, with loop detection and prevention as described in Section 11.3, offers an arguably substantial gain of 0.48dB at BER of 10^{-5} over the similar complexity weighted BF algorithm (the gain over UMP BP-based and plain BF is 0.57dB and 0.97dB, respectively). At the same time, the proposed algorithm is just about 0.63dB away from straight BP decoding, 0.61dB from normalized APP-based decoding, and 0.56dB from normalized BP-based decoding. With these performance studies in mind, we move on to algorithmic complexity comparisons. We plot the mean number of real operations needed[2] (real operations per iteration $\times$ mean iteration number) for each algorithm versus γ_b/N_0 (Fig. 11.2). In view of the results in Figs. 11.1 and 11.2, we argue that the proposed algorithm and the normalized APP-based algorithm of [7] offer the two best trade-off points between performance and computational cost (a "low-end" and a "high-end" trade-off point, respectively). As a form of visual aid that may help us gain a better understanding of the underlying algorithmic behavior, in Fig. 11.3 we plot the empirical "distribution" of the number of iterations needed by each algorithm to decode one received vector. We fix γ_b/N_0 at 3.85dB and we run each algorithm 800,000 times. The x-axis represents the iterations needed to decode a received vector and the y-axis represents the density of appearance of an iteration number as observed in the simulation. For effective visualization, we use a log scale for the y-axis. We see that for the BP, UMP BP-based, normalized BP-based and APP-based algorithms, most successful decodings terminate during the first 10 to 15 iterations at the price of high complexity per iteration. Weighted BF and the proposed algorithm with or without loop detection have frequency of occurrence peaks at about 25 iterations. The algorithm without loop detection has almost all successful decodings concentrated in the first 65 or so iterations. The remaining decodings fail as seen by the spike at iteration 200. Apparently, a good number of decoding failures can be attributed to infinite loop traps as seen by the operation of the version of the algorithm that incorporates loop detection and prevention (bottom entry in Fig. 11.3).

To verify that our findings and conclusions continue to hold true for much longer FG-LDPC codes than

[2] In Figs. 11.2 and 11.5, the operations for the BP algorithm are multiplications+divisions+additions. For all other algorithms, the operations are additions only (multiplications/divisions are not required at all).

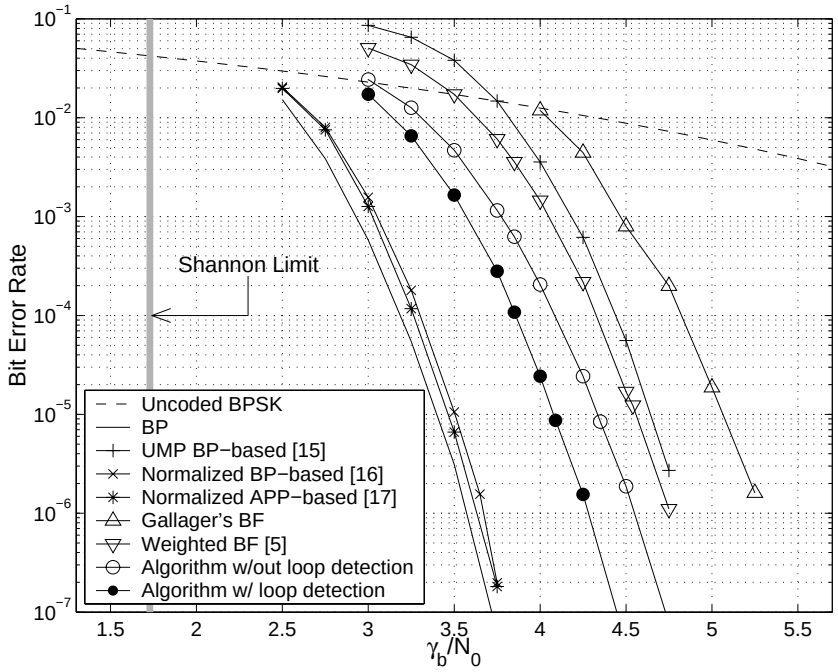

Fig. 11.1. Bit-error-rate versus rate normalized SNR for the Type-I 2-D (1023, 781) EG-LDPC code.

the (1023, 781) code of Figs. 11.1-11.3, in Fig. 11.4 we consider the Type-II 3-D (4599, 4227) EG-LDPC code with rate 0.919 (the normalization factor is 4.0 again). At BER of 10^{-5}, the proposed low cost algorithm is again about 0.6dB away from BP, less than 0.4dB from normalized BP-based decoding, and only 1.7dB away from the Shannon limit[3]. Compared to weighted BF, the algorithm offers a performance improvement of about 0.8dB. Fig. 11.5 establishes the proposed algorithm as the lowest complexity scheme among all BP, BP-based, APP-based, and weighted BF decoders.

[3]The minimum distance of this code is only 9. In our study of Fig. 11.4, we observed that, except for the normalized BP-based decoder, all other decoders, including BP, generated a considerable percentage of *undetected* errors.

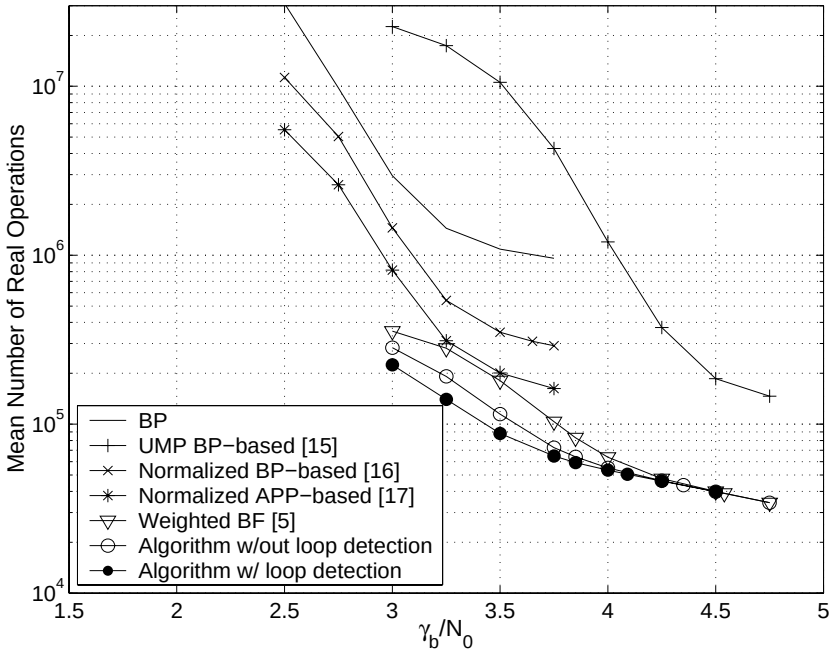

Fig. 11.2. Real operations needed versus rate normalized SNR for the code of Fig. 11.1.

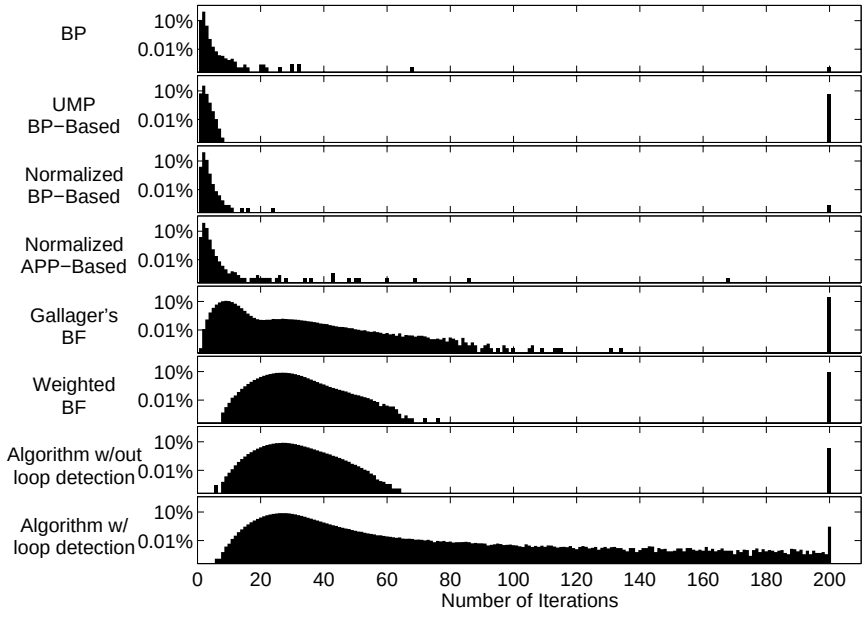

Fig. 11.3. Frequency of occurrence of executed iterations for the code studies in Figs. 11.1 and 11.2 (γ_b/N_0=3.85dB).

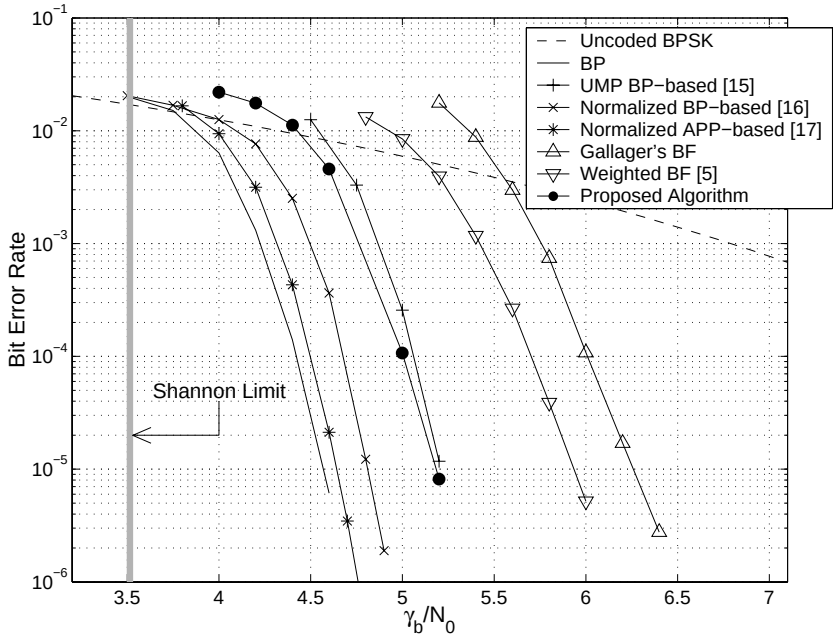

Fig. 11.4. Bit-error-rate versus rate normalized SNR for the Type-II 3-D (4599, 4227) EG-LDPC code.

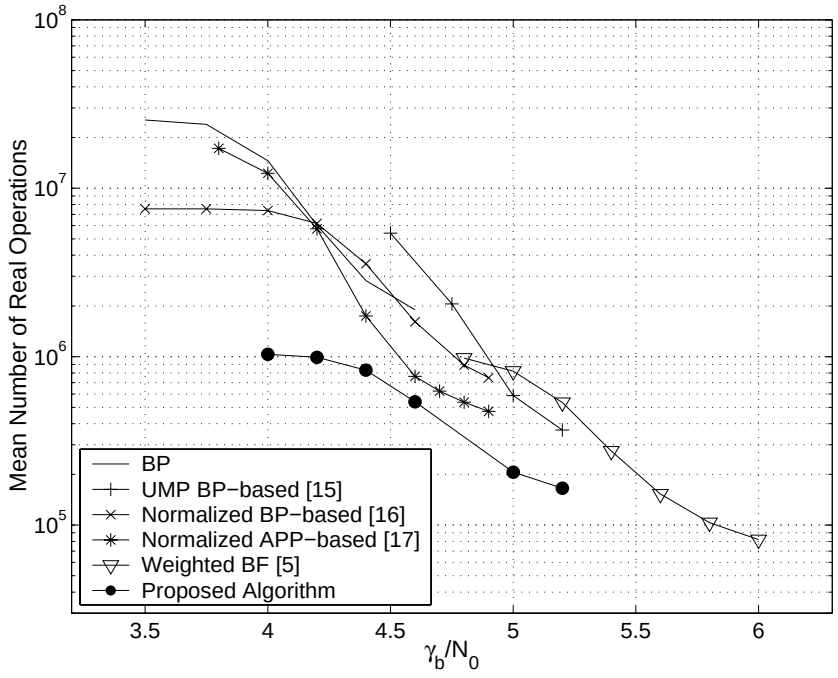

Fig. 11.5. Real operations needed versus rate normalized SNR for the code of Fig. 11.4.

Bibliography

[1] R. G. Gallager, "Low-density parity-check codes," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 21-28, Jan. 1962.

[2] R. G. Gallager, *Low Density Parity Check Codes*. Cambridge, MA: MIT Press, 1963.

[3] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low-density parity-check codes," *Electron. Lett.*, vol. 32, pp. 1645-1646, Aug. 1996.

[4] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, vol. 45, pp. 399-432, Mar. 1999.

[5] Y. Kou, S. Lin, and M. P. C. Fossorier, "Low density parity-check codes based on finite geometries: A rediscovery and new results," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2711-2736, Nov. 2001.

[6] M. P. C. Fossorier, M. Mihaljevic, and H. Imai, "Reduced complexity iterative decoding of low-density parity check codes based on belief propagation," *IEEE Trans. Commun.*, vol. 47, pp. 673-680, May 1999.

[7] J. Chen and M. P. C. Fossorier, "Near optimum universal belief propagation based decoding of low-density parity check codes," *IEEE Trans. Commun.*, vol. 50, pp. 406-414, March 2002.

[8] J. Chen and M. P. C Fossorier, "Decoding low-density parity check codes with normalized APP-based algorithm," in *Proc. IEEE Globecom '01*, vol. 2, San Antonio, TX, Nov. 2001, pp. 1026-1030.

[9] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inform. Theory*, vol. 27, pp. 533-547, Sept. 1981.

List of publications supported by this grant

[A1] G. N. Karystinos, H. Qian, M. J. Medley, and S. N. Batalama, "Short-data-record adaptive filtering: The auxiliary-vector algorithm," in *Proc. 2001/2002 Workshop on Defense Appl. of Signal Process.*, Adelaide, Australia, June 2002.

[A2] S. Gopalan, G. N. Karystinos, and D. A. Pados, "Capacity, throughput, and delay of slotted ALOHA DS-CDMA links with adaptive space-time auxiliary-vector receivers," *IEEE Trans. Wireless Commun.*, submitted Feb. 2002, revised Jan. 2003.

[A3] J. D. Matyjas, I. N. Psaromiligkos, S. N. Batalama, and M. J. Medley, "Fast converging minimum probability of error neural network receivers for DS-CDMA communications,"*IEEE Trans. Neural Networks*, to appear.

[A4] J. D. Matyjas, I. N. Psaromiligkos, S. N. Batalama, and M. J. Medley, "Fast converging adaptive training of DS-CDMA neural network receivers," in *Proc. SPIE's 17th Annual Int. Symp., Digital Wireless Commun. Conference V*, Orlando, FL, Apr. 2003, vol. 5100, pp. 238-248

[A5] P. Xiong, I. N. Psaromiligkos, and S. N. Batalama, "On the relative output SINR of full and partial decorrelators," *IEEE Trans. Commun.*, vol. 51, pp. 1633-1637, Oct. 2003.

[A6] P. Xiong, I. N. Psaromiligkos, and S. N. Batalama, "On the relative output SINR of full and partial decorrelators," in *Proc. IEEE Milcom'01*, McLean, VA, USA, 2001, vol. 2, pp. 1423-1428.

[A7] P. Xiong and S. N. Batalama, "All-blind adaptive non-linear spread-spectrum receivers," in *Proc. Conf. on Inform. Sci. and Syst.*, Baltimore, MD, USA, Mar. 2001, pp. 183-186.

[A8] P. Xiong and S. N. Batalama, "Adaptive non-linear spread-spectrum receivers without pilot signaling," in *Proc. Int. Conf. on Telecomm.*, Bucharest, Romania, Jun. 2001, vol. 3, pp. 245-248.

[A9] P. Xiong and S. N. Batalama, "Performance analysis of joint and disjoint space/time receivers in fast frequency-nonselective fading DS-CDMA channels," in *Proc. Conf. Inform. Sci. and Syst.*, Baltimore, MD, USA, Mar. 2003.

[A10] Z. Liu and D. A. Pados, "Near-ML multiuser detection with linear filters and reliability-based processing," *IEEE Trans. Commun.*, vol. 51, pp. 1446-1450, Sept. 2003.

[A11] Z. Liu and D. A. Pados, "Reliability-based near maximum-likelihood multiuser detection," in *Proc. MILCOM 2001*, vol. 1, Fairfax, VA,Oct. 2001, pp. 576-580.

[A12] Z. Liu and D. A. Pados, "Near-ML multiuser detection," in *Proc. SPIE*, Vol. 4395, Orlando, FL,Apr. 2001, pp. 65-74.

[A13] I. N. Psaromiligkos, S. N. Batalama, and M. J. Medley, "Rapid combined synchronization/demodulation structures for DS-CDMA systems—Part I: Algorithmic developments," *IEEE Trans. Commun.*, vol. 51, pp. 983-994, Jun. 2003.

[A14] I. N. Psaromiligkos and S. N. Batalama, "Rapid combined synchronization/demodulation structures for DS-CDMA systems—Part II: Finite data-record performance analysis," *IEEE Trans. Commun.*, vol. 51, pp. 1162-1172, Jul. 2003.

[A15] P. Xiong, M. J. Medley and S. N. Batalama, "Spatial and temporal processing of GPS signals," in *Proc. IEEE Workshop on Statistical Signal and Array Processing*, Pocono Manor, PA, USA, Aug. 2000, pp. 524-528.

[A16] P. Xiong, M. J. Medley and S. N. Batalama, "Spatial and temporal processing for global navigation satellite systems: The GPS receiver paradigm," *IEEE Trans. Aerosp. Electr. Syst.*, vol. 39, pp. 1471-1484, Oct. 2003.

[A17] Z. Liu and D. A. Pados, "A decoding algorithm for finite geometry LDPC codes," *IEEE Trans. Commun.*, submitted July 3, 2002.

[A18] Z. Liu and D. A. Pados, "A decoding algorithm for finite geometry LDPC codes," in *Proc. 2003 IEEE Int. Conf. on Commun.*, vol. 4, Anchorage, AK, May 2003, pp. 2713-2717.