NEAR-FAR RESISTANT LINEAR MULTIUSER DETECTION

by

RUXANDRA LUPAS

M.S. Princeton University, 1985

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Princeton, December 1988

Acknowledgements

First of all I would like to thank my advisor, Professor Sergio Verdú, for having ever so often provided guidance, during my most demanding and long-termed enterprise as yet. I would like to thank him for having always made himself available when I had anything to discuss, and for his personal example of motivation, thoroughness and high-set standards.

I would like to thank the readers of my thesis, Professors John B. Thomas and Stuart C. Schwartz, for their comments. I have also valued the comments of Shmuel Miller, who has carefully read parts of my thesis.

I am indebted to Professors Peter Ramadge, Ken Stieglitz and Sergio Verdú, whose teaching of "Optimization in Hilbert Space", "Theory of Algorithms", and "Queueing Systems", respectively, have been stimulating like no textbook could have been.

I am deeply grateful to Professor Thomas, my academic great-grand-father, for letting me adopt him as my non-academic advisor after my first semester at Princeton, and whose reassuring words, experienced objectivity and concern for me have been invaluable.

In a wider sense I am grateful to live in a society which allows the pursuit of personal goals, and I want to thank all the people and institutions which have helped me achieve mine.

Most of all I am grateful to my family, who has shaped my values, and whose spirit I am carrying on. I have always felt their love and support during the challenge of four long goal-oriented years by oneself and abroad.

My best wishes go to my Princeton companions, many of whom have become good friends: Csaba Peter Gabor, who helped me get to enjoy Prof. Hopkins' "Linear Systems" class, Johanna Froehlich, who shared the cockroaches and the chill of Room 1A with me, my Computer Science friends Ken Salem and Boris Kogan, and my office mates, Dave Brady and Charlie Bouman. With all of them we were "burning our candles at both ends", sharing thoughts, hopes, doubts and fun, and we have grown together.

Finally, I think a becoming motto at the end of a PhD thesis is the line from the song "Silent Sunlight": "... and your work has just begun ...".

NEAR-FAR RESISTANT LINEAR MULTIUSER DETECTION

Ruxandra Lupas

Department of Electrical Engineering
Princeton University, 1988

Abstract

This dissertation studies linear detectors for Code-Division Multiple-Access channels perturbed by additive white Gaussian noise, under the assumption that the ensemble of user code waveforms is known to the receiver. In particular the near-far problem of Direct-Sequence Spread-Spectrum Multiple-Access channels, i.e. the problem of demodulating a weak transmitter in the presence of powerful interferers, is solved with a linear receiver.

Two multiuser performance measures are used to quantify the performance degradation due to the presence of multiuser interference: the asymptotic efficiency, which is equivalent to bit-error-rate in the low background noise region, and the near-far resistance, which measures a detector's robustness to the near-far problem.

Conventional single-user detection in a multiuser channel is not near-far resistant, while the substantially higher performance of the optimum multiuser detector requires exponential complexity in the number of users.

We show that the near-far resistance of optimum multiuser detection is achieved by a linear receiver (whose complexity per demodulated bit is only linear in the number of users). The optimum linear detector for worst-case energies - the *decorrelating detector* - is found, along with existence conditions, which are always satisfied in the models of practical interest. Its implementation does not require knowledge of the received energies, its bit-error-rate is energy-independent, and it achieves optimum near-far resistance.

A one-shot decorrelating detector for asynchronous channels is considered, and shown to preserve a reduced near-far resistance. Examples indicate that the performance loss may be insignificant.

An iterative decision-feedback scheme is proposed, using the decorrelating detector in the first stage, and a feedback set optimized for the operating region. Optimum near-far resistance and a non-decreased asymptotic efficiency compared to the decorrelating detector are achieved, in exchange for knowledge of energies.

Finally, the situation of unknown signature sequences but common chip waveform is considered, and an adaptive algorithm is given which adapts to the main channel impairment, i.e. approaches the decorrelating detector in the low noise region and the conventional detector in the low interference region.

Table of Contents

1. Introduction

1.1 Background

Consider the multiuser communication problem: several users share a common channel to transmit information. The design problem is how to allocate usage of the multiple access channel to the different users, in order to maximize the information flow over the channel, while keeping the bit error probability under a given level. This problem is of great interest, since it is representative of a wide variety of data communication systems where there is more than one source (and any number of destinations); e.g. computer networks, satellite broadcast channels and radio networks. Various approaches to this problem are taken in practice. These can be divided into three broad subclasses.

The first class of approaches divides the channel among the users, such that the multiple-access capability relies on the orthogonality between the assigned signals. Different kinds of orthogonality are exploited in different systems, e.g. time-orthogonality in Time-Division Multiple-Access (each user is allocated a time slot in which he alone can transmit data, utilizing the full channel), frequency-orthogonality in Frequency-Division Multiple-Access (each user transmits on a different frequency band), code-orthogonality, if the transmissions are synchronous, polarization-orthogonality or direction-orthogonality. Existing channel allocation strategies are both static, or, to avoid unused capacity allocation, dynamic, according to the need of each user. Dynamic channel allocation strategies are either centralized, where a central controller uses polling or probing strategies to determine which users have information to send, or decentralized, as in token rings or the Ethernet.

The second class of approaches allows random channel access. The first random access system of this kind was the Aloha system, where the innovative idea was to allow each user to become active whenever he has anything to send. The general feature of this class of channels is that the received waveform can only be demodulated if it consists of one signal at a time, which means that all simultaneous transmissions are lost. A great deal of effort has gone into designing various random access protocols which schedule access to the common channel such that the probability of a simultaneous transmission, called a “collision”, is as low as possible while satisfying constraints on the waiting time distribution. When a collision occurs, it is handled by a collision resolution algorithm, whose task is to reschedule transmission of the collided packets at times which will be nonoverlapping with high probability. Beyond a certain range of channel utilization a common problem of random access algorithms is the problem of stability. Much research has been done recently to find collision resolution algorithms which maximize throughput. Current research focuses on the issues of capture, which models the case when part of the transmissions involved in a collision can be recovered (for example the ones with higher transmission power), and decentralized transmission rate control to ensure stability.

The third approach, the simultaneous transmission philosophy or Code-Division Multiple-Access (CDMA), allows simultaneous, asynchronous access to the common channel, without the penalizing feature of collisions. Here each user is assigned a fixed, distinct signature waveform, which he uses to modulate his digital information sequence, as if he were the only user of the channel. The input to the receiver consists of the superposition of the transmitted waveforms, perturbed by additive noise. The task of the receiver is to demodulate all transmitted sequences, or a proper subset of these.

The fact that each user is assigned a characteristic modulating waveform enables the destination to demodulate his information by correlating the incoming signal with a coherent replica of the desired user's signature waveform. If, aside from comparison to a set of thresholds, no further processing is done after this correlation, the resulting receiver is known as the conventional/single-user receiver. This receiver is widely used in practice due to its simplicity and to the fact that it is the receiver which minimizes the probability of error for reception of a single-user signal on a white Gaussian noise channel. Due to a spurious component of each non-orthogonal interfering waveform in the correlator output of the desired user, the performance of the conventional receiver is adequate only as long as the energies of the interfering users are under a certain level and the crosscorrelations between the signals are low enough. What "low enough" is, depends on the number of users and on the operational energy range. In practice, low crosscorrelations are obtained by assigning the users Spread-Spectrum pseudo-noise sequences with long constraint lengths. Much research has been done on how to design sequences which have good auto- and crosscorrelation properties for all relative shifts (e.g. [Sar 80]). Examples of these are maximal length shift-register sequences, Gold sequences and Kasami sequences.

However, no matter how well the sequences are designed, if one or more of the interferers is sufficiently strong, for example due to much closer proximity to the destination, the probability of error of the conventional receiver is bounded away from zero even in the absence of background noise.

This problem is of great importance in practice due to its ubiquity in communication systems with time-varying or very dissimilar topologies, and is known in the literature as the *near-far problem* ([Scho 77],[Pic 82]). At present it is the main shortcoming of Multiple-Access systems using Direct-Sequence Spread-Spectrum (DS-SS), which is one of the two main spectrum spreading techniques used in practice. (The other being Frequency-Hopping). DS-SS is often used in applications which require anti-jamming capability and immunity from hostile sources. One attempt to combat this problem without changing the receiver structure has been to control the power of the transmitting stations such that the received energies are similar. For example in [Ska 82] each transmitter estimates what power its signal has at the destination by estimating the arriving power of the return signal from the destination, whose power it knows. The disadvantage of this approach is both increased transmitter complexity, which may be undesirable, and the fact that the strong users have to put up with reduced performance for the benefit of the weak ones. Furthermore, the anti-jamming capability of the system is decreased.

However the near-far problem is not an inherent problem of DS-SS systems. If the receiver has knowledge of the interfering waveforms, the performance of CDMA systems can be greatly improved. Then each destination has a correlator and sampler for the signature waveform of each user, thus obtaining a discrete sequence which can be shown to be a sufficient statistic. Verdú [Ver 84c] found and analyzed the maximum-likelihood multiuser receiver for CDMA systems and in particular showed that the optimum receiver is not near-far limited. The optimum multiuser receiver follows the correlating front end by a Viterbi algorithm.

1.2 Previous work

Comparatively little work had been done previously on the demodulation/detection aspects of the multiuser channel. After Viterbi published his well-known maximum-likelihood decoding algorithm, which he had devised for the decoding of convolutional codes, various researchers in the field identified its relevance to related problems, e.g. [Kob 71] to correlative level coding, [For 72] to the intersymbol interference channel and [Ett 76] to M-input M-output dispersive channels with synchronous inputs, an environment which is related to the multiuser channel we are considering, due to the fact that the intersymbol interference introduces memory. The latter is one of a number of works on combatting crosstalk in multi-input multi-output dispersive communication systems using pulse-amplitude modulation, most of them concerned with the structure of optimum linear receivers under various criteria, including the work of [Shn 67] who finds the optimum linear receiver under a zero-forcing constraint, [Ett 75] who finds that the optimum linear filter under the minimum-probability-of-error criterion has the structure of a matched filter front end followed by a tap delay line, and of [Kay 70] who generalize the previous to I-input M-output diversity systems under a mean-squared error criterion. The models of these works are very general and they recognize the fact, first noted by [Shn 67], that intersymbol interference and "crosstalk" (synchronous multiuser interference) can be treated in a unified framework. (We now know that intersymbol interference can be viewed as additional multiuser interference by increasing the dimensionality of the user population, and conversely multiuser interference can either be viewed as periodically time-varying scalar intersymbol interference, or, if synchronization and matched filtering capabilities are assumed at the receiver in order to obtain a discrete-input discrete-output equivalent channel, as a vector generalization of intersymbol interference.) However, due to the generality of their model, the above works contain few specific results and in particular contain very little performance analysis. Earlier work on the multiuser channel of which we are aware considered multiuser receivers for the synchronous channel, specifically the receivers of [Hor 75] and [Sch 79], [Sch 80]. In [Sch 79] it is claimed erroneously (cf. [Ver 86b]) that a memoryless linear transformation on the matched filter front end is optimum in terms of bit error probability. Though this is not the case, the proposed receiver emerges as the solution in the synchronous case to the problem of finding a linear receiver which has desirable near-far performance, which is part of this work. In a short discussion of the asynchronous channel [Sch 79] also suggests that the Viterbi algorithm will provide a suitable solution. His intuition was correct as proved in [Ver 84c]. The performance of the conventional receiver under conditions of multiuser interference has been amply investigated in [Pur 81], [Pur 82], [Ger 82]. In part the relatively small amount of work on the multiuser channel prior to [Ver 84c] (which was in turn motivated by [Poor 80]) is due to the widespread previous belief that a more complex receiver than the conventional one would not yield a worthwhile performance gain ([Pur 81], p. 153). This belief was proved wrong in [Ver 84c], which triggered a new research effort in the multiuser communications area.

1.3 Motivation

The main reason for the quest for new receivers, now that the minimum probability of error and the maximum likelihood receivers for the asynchronous CDMA channel are known, is the exponential computational complexity in the number of users of these decision algorithms.

The computational complexity of the various decision algorithms can be measured and compared by their time complexity per binary decision, TCB, i.e., the limit as the length of the transmitted sequence tends to infinity of the time required by the decision algorithm to select the optimum sequence divided by the number of transmitted bits.

While the TCB of the conventional single-user detector is constant in the number of users, its bit-error rate is bounded away from zero for sufficiently high energy of any interfering non-orthogonal user, that is, the conventional systems can become multiple-access limited even in the absence of additive noise. On the other hand any prespecified error probability was shown [Ver 84c] to be achievable using the optimum detector. Unfortunately it is also shown that the decision algorithm for the optimum multiuser detection problem is NP-hard in the number of users, i.e. has a TCB which is exponential in the number of users, unless NP = P. Therefore the optimum multiuser detector becomes impractical for user populations above, say, 10. It is this trade-off between achievable performance and necessary time complexity per bit which motivates the current research in multiuser detection.

The aim of this thesis is to derive and analyze receiver structures which offer bit-error rates close to that of the optimum detector while maintaining computational feasibility. In particular this work is concerned with remedying the near-far problem with a detector with low computational complexity.

1.4 Outline of the thesis

In Chapter 2 we present the multiuser performance measures used to quantify and compare the performance of multiuser detectors. The first, the *asymptotic efficiency*, is specifically tailored to capture the performance degradation under conditions when the main impairment is the multiple-access interference, rather than the background noise. The second, the *near-far resistance*, measures the detector's robustness to the near-far problem which is our main concern in this work.

Chapter 3 focuses on the synchronous CDMA channel. After deriving the near-far resistance of both the conventional and the optimum detectors, the *decorrelating* multiuser detector is introduced. This detector linearly transforms each vector of matched filter outputs with a generalized inverse of the signal crosscorrelation matrix. It is shown that, somewhat unexpectedly, the near-far resistance of the optimum multiuser detector coincides with that of the decorrelating detector, whose complexity per demodulated bit is only linear in the number of users. In Section 3.5 the optimum linear transformation on the matched filter outputs is found, and conditions on the signal energies and crosscorrelations are given under which, for a certain user, its asymptotic efficiency is equal to that of the optimum multiuser detector. The issue of computing the decorrelating detector is addressed, and an iterative algorithm which converges to the decorrelating detector is given. In Section 3.8 an iterative decision-feedback scheme with the decorrelating detector in the first stage is proposed. This receiver uses the correlation of the noise samples in the matched filter outputs to reduce the variance of the noise components by subtracting a noise estimate based on the past decisions of the other users. The performance in the second stage is analyzed, both for feedback from all the users and for partial feedback, and near-far resistance is shown to be preserved.

Chapter 4 is concerned with the asynchronous CDMA channel. In Section 4.1 it is shown that the near-far resistance of the optimum multiuser detector can be achieved by a linear detector (the decorrelating detector), which is obtained explicitly in Section 4.2, as well as its implementable version as a linear time-invariant system. The dependence of the error probability of this detector on the received delays and phases is discussed and a numerical comparison of the error probability of the decorrelating receiver and the conventional receiver in a scenario of practical interest is given. In Section 4.4 a computationally much simpler one-shot detector is considered, which trades a lower level of near-far resistance than the decorrelating detector in return for lack of memory. A numerical comparison with the decorrelating detector is shown for some of the examples considered earlier.

Finally, in Chapter 5 the situation when the signature sequences of the other users are unknown is considered, in the case of a synchronous channel. An adaptive algorithm is presented which is shown to converge to the decorrelating detector as the level of the background noise vanishes, and to the conventional detector, if the multiuser interference level goes to zero, i.e. to the respectively optimum strategy under the respective (limiting) channel conditions.

1.5 Parallel work in the field

Other possible approaches to the presented issues are to devise suboptimal lower complexity versions of the Viterbi algorithm, as done in e.g. [Due 87] for the intersymbol interference channel, or to find suitable sequential algorithms, with a metric closely related to the optimal one, as in [Rus 88]. However these schemes rely largely on intuition and heuristics, which is why we chose to formulate a computationally favorable class of detectors and to optimize performance over this class. A recent publication which derives an optimal linear receiver under an asymptotic error probability criterion in a hypothesis testing setting is [Gal 88]. Another recent attempt to derive detectors for multiuser channels is [Var 88a], where the decisions of the conventional detector are used to subtract an estimate of the multiuser interference. A similar idea is pursued independently in Section 3.8 of this thesis. Part of the results presented here (also [Lup 86], [Lup 89a]) have been incorporated in [Var 88b], where the decorrelating detector is used instead of the conventional detector to obtain near-far resistant initial decisions. Finally, [Poor 88b] analyzes the form of the optimum single-user detector in a multiuser channel. For a further discussion on research in the field see [Ver 88].

2. Multiuser Performance Measures

2.1 Asymptotic Efficiency

The performance measure of interest in communication systems is the bit-error rate or probability of error achieved by each transmitter. Since the single-user error probability (of the optimum single-user detector) on the Gaussian channel is a one-to-one function of the signal-to-noise ratio (SNR), the same information as the error probability is contained in the *efficiency*, defined for each user as the ratio between his *effective* SNR and his actual SNR, where the effective SNR is the one needed to achieve the same error probability (using the optimum single-user detector) without interference from other users, and the actual SNR is the received energy per bit of the user divided by the power spectral density level of the background noise (not including interference from other users). Since it is the ratio of two energies, the efficiency is nonnegative, and since the effective energy is upper bounded by the actual energy (a user's error probability in a multiuser environment is lower bounded by his single-user error probability), the efficiency is less than or equal to unity. The main performance measure we are interested in is the bit-error-rate in the high signal-to-background noise region. Thus, even though the background thermal noise is not neglected, the main focus will be on the underlying performance degradation due to multiple-access interference. This is a meaningful way to look at the problem, and offers the advantage of tractability.

With this in mind a suitable multiuser performance measure is the *asymptotic efficiency*, introduced in [Ver 84c], [Ver 86a], and defined as the limit of the efficiency as the background noise level goes to zero. Thus the asymptotic efficiency is defined for each user as the limit as the background noise level goes to zero of the ratio between its effective energy and the actual energy it has in the multiuser environment. Therefore it is a measure of the performance loss due to the existence of other active users in the channel. Consider an additive white Gaussian channel with binary antipodal signaling, to which we always refer in the sequel, such that the received waveform upon transmission by a single user is

$$r(t) \;=\; b\,\sqrt{w}\;s(t) \;+\; n(t)$$

where b, w and $s(t)$ are the transmitted bit, the received energy per bit, and the normalized received version of the modulating waveform, and $n(t)$ is a white Gaussian noise process with power spectral density σ^2. It is well known (e.g. [Woz]) that the error probability of the optimum detector for this situation, which is a filter matched to $s(t)$ followed by a sign decision, is[1] $Q(\sqrt{w}/\sigma)$. When K users are transmitting, let the k^{th} user received energy and error probability (achieved by the specific detector under consideration) be equal to w_k and P_k, respectively. Then the effective energy $e_k(\sigma)$ is such that $P_k(\sigma) = Q(\sqrt{e_k(\sigma)}/\sigma)$ and the k^{th} user asymptotic efficiency of this detector can be written as [Ver 86a]

$$\begin{aligned} \eta_k &= \lim_{\sigma \to 0} \frac{e_k(\sigma)}{w_k} \\ &= \sup\left\{0 \le r \le 1;\ \lim_{\sigma \to 0}\ P_k(\sigma)\ /\ Q(\frac{\sqrt{r w_k}}{\sigma}) < +\infty\right\} . \end{aligned} \tag{2.1}$$

We can make the connection between the two above definitions in the following way: Consider the behavior of $P_k(\sigma)$ as σ goes to zero. Since no multiuser detector can outperform the optimum single-user detector in a single-user environment, $P_k(\sigma)$ decays in the limit either as a Q-function, or slower. In the first case the value of r to ensure a finite ratio in (2.1) will be such that the two Q-functions have the same arguments, i.e. it will be the ratio of effective and actual energy. In the second case, (2.1) predicts an asymptotic efficiency of zero, which follows also from the first definition, since for the second case to hold the effective energy must tend to zero.

From the above discussion it becomes clear that the asymptotic efficiency has the following geometric interpretation: the logarithm of the k^{th} user error probability decays asymptotically with the slope corresponding to a single-user with energy $\eta_k w_k$. It also follows that while an irreducible error probability entails a zero asymptotic efficiency, conversely an asymptotic efficiency of zero means that the error probability does not tend to zero exponentially fast with increasing k^{th} user signal to Gaussian background noise ratio. To illustrate how the asymptotic efficiency is obtained, consider a linear detector. The k^{th} user error probability can be shown to be a weighted sum of Q-functions, one for each possible interfering bit-combination; as $\sigma \to 0$ the Q-function with the smallest argument dominates and determines the error probability. The asymptotic efficiency is obtained from this smallest argument.

The importance of the asymptotic efficiency as a multi-user performance measure on the Gaussian channel is that - while probability of error, the actually interesting parameter in any binary communication environment, is highly intractable, which is why other measures like mean-squared

---

(1) $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-v^2/2} dv$

errors are resorted to, - it is a measure which is equivalent to probability of error in the high signal-to-noise region, while offering the advantage of tractability.

2.2 Near-far Resistance

Since we are interested in alleviating the near-far problem, i.e. are interested in detectors whose performance level is high for all received energies, a suitable performance indicator for near-far robustness is the k^{th} user *near-far resistance*, which is defined as the worst-case asymptotic efficiency over all possible energies of the interfering users. Thus for a synchronous channel the near-far resistance of a detector is defined as

$$\overline{\eta_k} = \inf_{\substack{w_j \geq 0 \\ j \neq k}} \eta_k \, . \tag{2.2}$$

The definition for the asynchronous channel is given in Chapter 4. A detector is *near-far resistant* for User k if the near-far resistance of User k is nonzero.

3. Linear multiuser detectors for synchronous CDMA channels

3.1 Preliminaries

Suppose that the k^{th} user is assigned a unit energy signature waveform, $\{s_k(t)\ ,\ t \in [0,T]\}$, and he transmits a string of bits by modulating that waveform antipodally. If the users maintain symbol-synchronism, and they share an ideal white Gaussian multiple-access channel, then the receiver observes

$$r(t) = \sum_{k=1}^{K} b_k(l) \sqrt{w_k(l)}\ s_k(t - lT) + \sigma\, n(t)\ , \qquad t \in [lT,\ lT+T) \tag{3.1}$$

where $n(t)$ is a realization of a unit spectral density white Gaussian process, $\{b_k(l) \in \{-1,1\}\}_l$ and $w_k(l)$ are the k^{th} user information sequence and the possibly time-dependent received energy sequence, respectively. Assuming that all possible information sequences are equally likely, it suffices to restrict attention to a specific symbol interval in (3.1), e.g. $l = 0$. For this reason in the sequel specification of the symbol interval is omitted.

It is easy to check that the likelihood function depends on the observations only through the outputs of a bank of matched filters:

$$y_k = \int_0^T r(t)\ s_k(t)\ dt, \qquad k = 1, \cdots K \tag{3.2}$$

and therefore $\mathbf{y} = (y_1, \ldots, y_K)$ are sufficient statistics for demodulating $\mathbf{b} = (b_1, \ldots, b_K)$. In this section we investigate ways of processing these sufficient statistics, which according to (3.1) and (3.2) depend on the transmitted bits in the following way:

$$\mathbf{y} = \mathbf{R}\ \mathbf{W}\ \mathbf{b} + \mathbf{n} \tag{3.3}$$

where $\mathbf{R}$ is the nonnegative definite Hermitian matrix of crosscorrelations between the assigned waveforms:

$$R_{kj} = \int_0^T s_k(t)\ s_j(t)\ dt \tag{3.4}$$

with diagonal entries $R_{kk} = 1$, $\mathbf{W}$ is diagonal with entries $\sqrt{w_k}$ and $\mathbf{n}$ is a zero-mean Gaussian K-vector with covariance matrix equal to $\sigma^2 \mathbf{R}$.

In this chapter we do not restrict the signal set of the K interfering users to be linearly independent, which means that $\mathbf{R}$ can be singular. Therefore many of the results are formulated in terms of the generalized inverse of $\mathbf{R}$, which obviously will reduce to the usual inverse, if the signal set is linearly independent.

Note that the model of equation (3.3) is not the only one we could work with. Equivalently, we could either choose $r \stackrel{\triangle}{=} \text{rank}(\mathbf{R})$ independent users and discard the other components of $\mathbf{y}$, or use a set of r orthonormal matched filters which are obtained from the waveform ensemble $\{s_k(t), t \in [0,T], k = 1, ..., K\}$ via Gram-Schmidt orthonormalization. Both representations yield sufficient statistics for the demodulation of $\mathbf{b}$. The orthonormalized matched filter set yields a white output noise sequence, therefore it is equivalent to a K-input r-output whitened matched filter. However, both representations yield non-square matrices for $r < K$, and the Gram-Schmidt procedure requires increased computational effort. For these reasons this work adopts the model of (3.3), although if the signal set is linearly dependent, the sufficient statistic $\mathbf{y}$ is redundant.

In order to see where the additional demodulation difficulty comes from when the signal set is linearly dependent, consider the case of singular noiseless demodulation, i.e. the problem of demodulating $\mathbf{b} = (b_1, \ldots, b_K)$ from $\mathbf{y} = (y_1, \ldots, y_K)$,

$$\mathbf{y} = \mathbf{R}\,\mathbf{W}\,\mathbf{b}\,,$$

when $\mathbf{R}$ is singular. Since the collection of all possible hypothesis vectors $\mathbf{b}$ spans $\mathbb{R}^K$, it is apparent that no linear transformation can recover $\mathbf{b}$ from $\mathbf{y}$. In fact it is easy to show that noiseless singular demodulation is NP-complete, because "PARTITION" ([Gar]: given $L = \{l_1, ..., l_n\}$, $l_i \in \mathbb{Z}^+$ and $G \in \mathbb{Z}^+$, decide whether there exists a subset $L' \in L$ such that $\sum_{l_i \in L'} l_i = G + \sum_{l_i \in L-L'} l_i$) can be reduced to a special case of "NOISELESS SINGULAR DEMODULATION", namely the case where the rank of $\mathbf{R}$ is unity.(2) It is not hard to find an algorithm which solves noiseless singular demodulation, i.e. given $\mathbf{y}$ decides whether a solution $\mathbf{b}$ with components in $\{-1, 1\}$ exists, and in the latter case finds it, with a time complexity per bit of

$$T(K) = \frac{1}{K}\, 2^{K-r}\, O(r^2)$$

(2) Let $R_{ij} = l_i l_j$, $y_i = l_i G$, $\mathbf{W} = \mathbf{I}$. Clearly $\mathbf{R}$ is symmetric, nonnegative definite, and rank $\mathbf{R} = 1$. Then since $l_i \neq 0$, $\mathbf{RWb} = \mathbf{y} \leftrightarrow G = \sum l_i b_i$.

where r is the rank of $\mathbf{R}$. To do this, one possibility is to choose r linearly independent columns of $\mathbf{R}$, assign all possible values to the $K-r$ bits corresponding to the other columns, solve for the r remaining bits and accept a solution when these bits turn out to be -1 or +1. The given complexity is then immediate.

Seeing that there is no linear transformation which solves noiseless singular demodulation, an interesting question to ask is whether a subset of $\mathbf{b}$ can be recovered by a linear transformation. The answer is given by Lemma 3.1. The following definition characterizes the dependence which is the cause of the singularity of $\mathbf{R}$. It is easy to show that dependence of modulating waveforms is directly translated into linear dependence of the corresponding columns of $\mathbf{R}$, $\mathbf{r}_i,\ i=1,\ldots,K$.

Definition : Users $u_i,\ i \in I \subseteq \{1,\ldots,K\}$ form a *maximal dependent block* if

$$\forall\ \mathbf{x} \neq 0\ ,\ \text{s.t.}\ \ \mathbf{Rx} = 0\ ,\quad \sum_{i \in I} \mathbf{r}_i\, x_i = 0$$

and no subset I' ⊂ I satisfies the above for all $\mathbf{x}$ in the nullspace of $\mathbf{R}$.

Lemma 3.1: Application of the Moore-Penrose inverse[3] $\mathbf{R}^+$ on the matched filter output vector $\mathbf{y}$ decouples the users into maximal dependent blocks, i.e. if the users are relabeled such that users in the same dependent block have consecutive labels, then $\mathbf{R}^+\mathbf{R}$ is block diagonal. ◊

Proof: Singling out a maximal dependent block I, let the matrices $\mathbf{R}$, $\mathbf{R}^+$ and $\mathbf{R}^+\mathbf{R}$ be partitioned according to the indices in I as

$$\begin{aligned} \mathbf{R} &\rightarrow (\mathbf{M},\ \mathbf{A}\ ,\ \mathbf{N}) \\ \mathbf{R}^+ &\rightarrow (\mathbf{B}\ ,\ \mathbf{C}\ ,\ \mathbf{D}) \\ \mathbf{R}^+\mathbf{R} &\rightarrow (\mathbf{Y}\ ,\ \mathbf{X}\ ,\ \mathbf{Z}) \end{aligned}$$

where the first entry corresponds to the maximal dependent block, the last entry to its complement, and since all the matrices are symmetric the notation $(\mathbf{M},\mathbf{A},\mathbf{A}^{\mathbf{T}},\mathbf{N})$ has been abbreviated to the above form. We want to show $\mathbf{X} = \mathbf{0}$. By definition of a Moore-Penrose inverse,

$$\mathbf{R}\ (\mathbf{R}^+\mathbf{R}) = \mathbf{R} \Rightarrow \begin{bmatrix} \mathbf{M} & \mathbf{A} \\ \mathbf{A}^T & \mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{Y}-\mathbf{I} & \mathbf{X} \\ \mathbf{X}^T & \mathbf{Z}-\mathbf{I} \end{bmatrix} = 0$$

(3) A generalized inverse $\mathbf{A}$ of a matrix $\mathbf{B}$ is any matrix that satisfies 1. $\mathbf{ABA} = \mathbf{A}$ and 2. $\mathbf{BAB} = \mathbf{B}$. The Moore-Penrose generalized inverse, denoted by $\mathbf{B}^+$, is the unique generalized inverse that satisfies 3. $\mathbf{AB}$ and $\mathbf{BA}$ are Hermitian.

$$\Longrightarrow \quad \mathbf{MY} = \mathbf{M}$$
$$\mathbf{A}^T\mathbf{Y} = \mathbf{A}^T$$

where the last implication follows from the definition of a maximal dependent block. Also

$$(\mathbf{R}^+\mathbf{R}) = \mathbf{R}^+\mathbf{R} \Rightarrow \mathbf{Y} = \mathbf{BM} + \mathbf{CA}^T$$

where from, multiplying by $\mathbf{Y}$ and using the previous equalities

$$\mathbf{Y}^2 = \mathbf{BMY} + \mathbf{CA}^T\mathbf{Y} = \mathbf{BM} + \mathbf{CA}^T = \mathbf{Y}$$

but

$$(\mathbf{R}^+\mathbf{R})^2 = (\mathbf{R}^+\mathbf{R}) \Rightarrow \mathbf{Y}^2 + \mathbf{XX}^T = \mathbf{Y}$$

which means that

$$\mathbf{XX}^T = 0 \Rightarrow \mathbf{X} = 0\,.$$

■

This result implies that independent users, i.e. users whose modulating waveform is linearly independent of the others, can be easily demodulated in the noiseless case, since they are decoupled by multiplication of $\mathbf{y}$ with $\mathbf{R}^+$, while for dependent users, a time complexity which is exponential in the size of the dependent block is feasible.

3.2 Single-user detection and optimum multiuser detection

3.2.1 The conventional single-user detector

If it were possible to accurately model the multiple-access interference as a zero-mean white Gaussian random process, then the optimum receiver would be the one which is known to be optimum for detection of a known signal in white Gaussian noise, namely a filter matched to the known signal, followed by a threshold. This strategy is optimal in the absence of multiuser interference, which is why the aforementioned detector is called a single-user detector in this context. However, in the multiuser environments encountered in practical applications, the Gaussian assumption is unfounded, and the colored non-Gaussian nature of the multiuser interference has to be taken into account. Regardless of this fact, due to its simplicity, the conventional single-user detector is the detector which is commonly used in practical situations. We focus now on this detector, along with its performance in a multiuser environment.

Conventional single-user detection for the k^{th} user decides on b_k in the simplest possible way. The k^{th} user receiver consists of a single matched filter, matched to the k^{th} user's signal, instead of the bank of matched filters which are necessary to generate the sufficient statistic $\mathbf{y} = (y_1, \ldots, y_K)$ for demodulation of each user's bit stream. Thus only y_k is generated at the receiver, yielding the following decisions for the k^{th} user:

$$\hat{b}_k^c = \text{sgn } y_k .$$

It is apparent that the conventional receiver requires very little complexity, and that its time-complexity per bit is independent of the number of users. On the other hand,

$$y_k = \sqrt{w_k}\, b_k + \sum_{i \neq k} \sqrt{w_i}\, R_{ki}\, b_i + n_k , \qquad n_k \sim N(0, \sigma^2)$$

so that taking a sign decision on y_k completely ignores the multiple-access interference component $\sum_{i \neq k} \sqrt{w_i}\, R_{ki}\, b_i$ present in y_k. As a result it becomes apparent that a sufficiently high interference energy from any nonorthogonal user will result in an irreducible error probability, even for a vanishing background noise level. A necessary and sufficient condition for this to occur is that the interfering energies are such that $\sqrt{w_k} \leq \sum_{i \neq k} |R_{ki}| \sqrt{w_i}$. This shows that the only way to prevent the conventional receiver from becoming multiple-access limited for sufficiently high interfering energy is to use an orthogonal signal set. The k^{th} user error probability of the conventional single-user detector is:

$$\begin{aligned} P_k^c &= P\left[y_k > 0 \mid b_k = -1\right] \\ &= \sum_{\substack{\mathbf{b} \in \{-1,1\}^K \\ b_k = -1}} P\left[y_k > 0 \mid \mathbf{b}\right] P\left[\mathbf{b} \mid b_k = -1\right] \\ &= 2^{1-K} \sum_{\substack{\mathbf{b} \in \{-1,1\}^K \\ b_k = -1}} Q\left(\frac{\sqrt{w_k} - \sum_{i \neq k} R_{ik} \sqrt{w_i}\, b_i}{\sigma} \right) . \end{aligned} \tag{3.5}$$

In the low background noise region the foregoing summation is dominated by the term corresponding to the least-favorable bits of the interfering users, i.e., $b_i = \text{sgn } (R_{ik})$. Thus, the asymptotic efficiency of the conventional detector is equal to

$$\eta_k^c \;=\; \sup \{0 \;\leq\; r \;\leq\; 1 \,;\;\; \lim_{\sigma \,\to\, 0} P_k^c \;/\; Q(\frac{\sqrt{r w_k}}{\sigma}) \;<\; +\infty\} \tag{3.6}$$

$$= \max^2 \left\{ 0 \;,\; 1 \;-\; \sum_{i \neq k} |R_{ik}| \; \frac{\sqrt{w_i}}{\sqrt{w_k}} \right\} . \tag{3.7}$$

It follows from (3.7) that the conventional k^{th} user detector is not near-far resistant (i.e., its asymptotic efficiency is not bounded away from zero as a function of the interfering users' energies), unless $R_{ik} \;=\; 0$ for all $i \neq k$, i.e., only if the k^{th} user's signal is orthogonal to the subspace spanned by the other signals. Otherwise,

$$\overline{\eta_k^c} \;=\; \inf_{\substack{w_i \geq 0 \\ i \neq k}} \; \eta_k^c \;=\; 0 \,. \tag{3.8}$$

For example for two active users, $\eta_k^c = 0$ for $\sqrt{w_j / w_k} \geq 1/\rho$, where ρ is the correlation coefficient between the two waveforms. Actually, we can make a stronger statement than this. As explained in Section 2.1, an asymptotic efficiency of zero does not imply that the probability of error will be bounded away from zero as the background noise vanishes; it only limits the speed of decay to be slower than exponential. However the error probability of the conventional receiver *does not decay to zero* as $\sigma \to 0$ if its asymptotic efficiency is zero. (This holds for any linear detector, and in general for any detector whose error probability can be represented as a sum of Q-functions). Specifically, if the crosscorrelation coefficients and the user energies are such that for n_1 out of the 2^{K-1} possibilities for $\mathbf{b} \in \{-1, 1\}^K$, for fixed b_k, we have

$$\sum_{i \neq k} b_i \; R_{ik} \; \sqrt{\frac{w_i}{w_k}} \;>\; 1$$

and equality holds for n_2 possibilities, then the limit of the error probability in (3.5) as $\sigma \to 0$ is

$$P_k^c \;\underset{\sigma \to 0}{\longrightarrow}\; 2^{1-K} \; (1 \times n_1 + \frac{1}{2} \times n_2) \;=\; \frac{2n_1 + n_2}{2^K} \,.$$

For example in the two-user case, if $\sqrt{w_2}/\sqrt{w_1} = (1 + \Delta^2)/\rho$, the error probability of the conventional receiver for User 1 tends to 1/4 if $\Delta = 0$ and to 1/2 if $\Delta > 0$, for increasing SNR of User 1.

This clearly shows the multi-access limitation of the conventional detector, as well as the fact that in order to obtain an adequate performance within a nominal range of energies stringent requirements have to be put on the crosscorrelations allowable between the modulating signals, without being able to prevent a severe performance degradation if the multiple-access interference exceeds the limit specified in the signal design.

3.2.2 The maximum likelihood detector (optimal receiver)

Due to the fact that there are several users on the channel, optimum detection in the sense of minimizing the probability of error can be conceived in two equally meaningful ways: the goal can be a global one, i.e. maximization of the *joint* posterior density of the transmitted symbols given the received signal, i.e.

$$\hat{b}_k \in \mathbf{u}_k^T \arg \max_{\mathbf{b} \in \{-1,1\}^K} P\left[\mathbf{b} \mid \{r(t), t \in \mathbb{R}\}\right],$$

where $\mathbf{u}_k$ is the k^{th} unit vector, or maximization of the *marginal* posterior density, i.e.

$$\hat{b}_k \in \arg \max_{b \in \{-1,1\}} P\left[b_k = b \mid \{r(t), t \in \mathbb{R}\}\right].$$

We will refer to the first as maximum likelihood detection and to the second as minimum-error-probability detection. Note that the two criteria do indeed lead to different detectors, which is due to the fact that the symbols of the interfering users are no longer independent conditioned on the received signal.

Example 3.1. As an easy example, consider a two-user CDMA detection problem, where the matched filters are matched to the Gram-Schmidt orthogonalized versions of the modulating signals, as discussed in Section 3.1. The matched filter outputs also form a sufficient statistic, with the difference that the noise vector is uncorrelated, which eased the construction of the desired example. The two-user detection problem then is

$$\begin{aligned} y_1 &= b_1 + \rho b_2 + n_1 \\ y_2 &= \sqrt{1-\rho^2}\, b_2 + n_2 . \end{aligned}$$

Now let the crosscorrelation between the two signals be $\rho = 0.6$, the noise variance be $\sigma^2 = 1$ and consider the situation where the received vector $\mathbf{y}$ is $[1, -0.1]$. The posterior probability $P\left[b_1, b_2 | \mathbf{y} = [1, -0.1]\right]$ is shown in Table 1, for the four possibilities, i.e., rowwise, $[b_1, b_2] = [1,1], [1,-1], [-1,1], [-1,-1]$. The maximum likelihood detector decides for the composite hypothesis which is most likely conditioned on the received vector, hence in this case will choose [1,-1]. Thus the maximum likelihood decision on the bit transmitted by the second user will be -1. On the other hand the minimum-error-probability detector for User 2 maximizes

$$P\left(b_2 | \mathbf{y}\right) = \sum_{b_1 \in \{1,-1\}} P\left(b_2, b_1 | \mathbf{y}\right)$$

$P[b_1, b_2 \mid y = \begin{bmatrix} 1 \\ -0.1 \end{bmatrix}]$:

	(1,1)	(1,-1)
	0.37	0.44
	0.17	0.02
	(-1,1)	(-1,-1)

Table 1. Posterior probability of hypotheses for Example 3.1.

i.e. in Table 1 it will choose the symbol corresponding to the column with largest element sum. In this case this is the first column, hence the minimum-error-probability decision on the bit transmitted by the second user will be +1. △

Both the maximum likelihood and the minimum-error-probability receivers for asynchronous CDMA have been found in [Ver 84c], and while they are both dynamic programming algorithms, the first is a forward Viterbi algorithm, while the second is of the backward-forward type, and is in general more complicated. However, [Ver 84c] shows that as the noise level decreases the number of symbols in which the optimum sequences according to both criteria differ goes to zero. Intuitively, this is because for vanishing noise levels the probability mass function $P[\mathbf{b}|\mathbf{y}]$ concentrates increasingly on one element, i.e. a table analogous to Table 1 would have one element close to 1 and the others close to 0, so that both detectors would choose the same element. Since this work deals with the performance of multiuser detectors in the high signal-to-background-noise region, where the limiting factor is the multiuser interference, the aforementioned convergence in the high SNR region of the performances of the two optimal detectors is the reason for which in the sequel we may restrict attention to the maximum likelihood detector and refer to it as the *optimum multiuser detector.*

The optimum multiuser detector selects the most likely hypothesis $\hat{\mathbf{b}}^* = (\hat{b}_1^*, \ldots, \hat{b}_K^*)$ given the observations, which corresponds to the noise realization with minimum energy, i.e.,

$$\hat{\mathbf{b}}^* \in \arg \min_{\mathbf{b} \in \{-1,1\}^K} \int_0^T [\, r(t) - \sum_{k=1}^{K} b_k \, s_k(t) \,]^2 \, dt$$

$$= \arg \min_{\mathbf{b} \in \{-1,1\}^K} \mathbf{b}^T \mathbf{W} \mathbf{R} \mathbf{W} \mathbf{b} - 2 \mathbf{y}^T \mathbf{W} \mathbf{b} . \quad (3.9)$$

The computational complexity of the optimum multiuser detector is radically different from that of the single-user detector. While we have seen that the time-complexity per bit of the single-user detector is independent of the number of users, no algorithm that solves (3.9) in polynomial time in K is known. The reason for this is the NP-completeness of optimum multiuser detection [Ver 85], [Ver 89].

However the performance of both detectors is quite different as well. The k^{th} user error probability of the optimum multiuser receiver is asymptotically (as $\sigma \rightarrow 0$) equivalent to that of a binary test between the two closest hypotheses that differ in the k^{th} bit (see [Ver 86b]). The square of the Euclidean distance between the signals corresponding to these two hypotheses is equal to

$$\min_{\mathbf{b} \in \{-1,1\}^K} \min_{\substack{\mathbf{d} \in \{-1,1\}^K \\ d_k \neq b_k}} \| \sum_{i=1}^{K} b_i \, s_i(t) - \sum_{i=1}^{K} d_i \, s_i(t) \|^2 =$$

$$= 4 \min_{\substack{\epsilon \in \{-1,0,1\}^K \\ \epsilon_k = 1}} \epsilon^T \mathbf{W} \mathbf{R} \mathbf{W} \epsilon . \quad (3.10)$$

Hence, the asymptotic efficiency of the optimum multiuser detector is equal to

$$\eta_k = \frac{1}{w_k} \min_{\substack{\epsilon \in \{-1,0,1\}^K \\ \epsilon_k = 1}} \epsilon^T \mathbf{W} \mathbf{R} \mathbf{W} \epsilon . \quad (3.11)$$

This is the highest asymptotic efficiency attainable by any detector because as $\sigma \rightarrow 0$ the optimum multiuser detector achieves minimum probability of error for each user. In the two-user case, denoting $\rho = R_{12}$, (3.11) reduces to

$$\eta_1 = \min \left\{ 1, \; 1 + \frac{w_2}{w_1} - 2|\rho| \frac{\sqrt{w_2}}{\sqrt{w_1}} \right\} , \quad (3.12)$$

and analogously for User 2. Unfortunately, no explicit expressions are known for (3.11) for an arbitrary number of users. In fact, it is shown in [Ver 85] that the combinatorial optimization problem in (3.11) is also NP-complete, if $\mathbf{R}$ is nonnegative definite. We will give a modified proof which extends the result to positive definite matrices. This extension is nonobvious, because any additional structure introduced (here it is the requirement of nonsingularity) may turn a difficult problem into one solvable in polynomial time.

Proposition 3.1: The following problem is NP-complete.

"MULTIUSER ASYMPTOTIC EFFICIENCY":

Given $K \in I\!N$, $k \in \{1, \ldots, K\}$, and a positive definite matrix $\mathbf{R} \in I\!R^{K \times K}$,

find the k^{th} user asymptotic efficiency $\eta_k \;=\; \frac{1}{R_{kk}} \min_{\substack{\epsilon \in \{-1,0,1\}^K \\ \epsilon_k \,=\, 1}} \epsilon^T \, \mathbf{R} \, \epsilon.$ ◇

Note that we have absorbed the invertible positive diagonal matrices $\mathbf{W}$ into $\mathbf{R}$, since there is a one to one correspondence between the two situations, and positive definiteness is independent of, and preserved by, multiplication by $\mathbf{W}$ on both sides.

Proof : The standard technique in proving NP-completeness of a given problem is to reduce to the problem in question a similar problem, known to be NP-complete. We adopt the approach in [Ver 85] and reduce -1/0/1 KNAPSACK, which is shown therein to be NP-complete, to our problem, using a modified reduction tailored for the positive definite case.

Reduction of '-1/0/1 KNAPSACK' to "MULTIUSER ASYMPTOTIC EFFICIENCY":

Given : $\{l_1, l_2, \ldots, l_L\}$, $l_i \in Z\!\!Z^+$, $i = 1, \ldots, L$ and $G \in Z\!\!Z^+$

find whether or not there exist $\epsilon_i \in \{-1, 0, 1\}$, such that $\sum_{i=1}^{L} \epsilon_i l_i = G$

given that the following problem can be solved for all $K \in Z\!\!Z^+$ and positive definite matrices $\mathbf{R}$:

find $\alpha_k \;\stackrel{\triangle}{=}\; \min_{\substack{\epsilon \in \{-1,0,1\}^K \\ \epsilon_k \,=\, 1}} \epsilon^T \, \mathbf{R} \, \epsilon.$

To reduce the first problem to the second we define

$$K \;=\; L \,+\, 1$$

$$l'_i \;=\; \begin{cases} l_i \,, & i \;=\; 1, \ldots, K-1 \\ G \,, & i \;=\; K \end{cases}$$

$$\mathbf{R} \;:\quad R_{ij} \;=\; l'_i \, l'_j \;+\; \Delta^2 \, \delta_{ij} \,, \qquad i, j \in \{1, \ldots, K\} \,. \tag{3.13}$$

then

$$\begin{aligned} \mathbf{x}^T \, \mathbf{R} \, \mathbf{x} \;&=\; \sum_i \sum_j \, x_i \, x_j \, R_{ij} \\ &=\; (\, \sum_{i=1}^{K} \, x_i \, l'_i \,)^2 \;+\; \Delta^2 \, \| \, \mathbf{x} \, \|^2 \,, \end{aligned} \tag{3.14}$$

hence $\|x\| > 0 \Rightarrow \mathbf{x}^T \mathbf{R} \mathbf{x} > 0$, which means that the matrix $\mathbf{R}$ we have defined is indeed positive definite, if $\Delta \neq 0$. Now, for ϵ as defined in the asymptotic efficiency expression,

$$\epsilon^T \mathbf{R} \, \epsilon \; = \; (\sum_{i=1}^{K} \epsilon_i \, l_i \; + \; G \,)^2 \; + \; \Delta^2 \parallel \epsilon \parallel^2 . \tag{3.15}$$

We can find the minimum of the right hand side (by assumption), and would like to know whether the first term can be zero for some choice of $\{\epsilon_i\}$. Now, since $\|\epsilon\|^2 \leq K$, if we pick $\Delta^2 < 1/K$ the second term on the right hand side will be less than unity. Therefore, since the first term is an integer, the sum is minimized if the first term is minimized. With this in mind, letting $\Delta^2 \; := \; \frac{1}{K+1}$,

$$\alpha_k \; < \; 1 \; \Leftrightarrow \; \text{"YES" instance of -1/0/1 KNAPSACK}$$

$$\alpha_k \; \geq \; 1 \; \Leftrightarrow \; \text{"NO" instance of -1/0/1 KNAPSACK.}$$

We have shown that if we could solve "MULTIUSER ASYMPTOTIC EFFICIENCY", we could equivalently solve "-1/0/1 KNAPSACK", which has been shown to have a formulation as a special case of the former. Thus our problem is at least as hard as "-1/0/1 KNAPSACK", hence at least NP-complete. It is easy to see that it is in NP, which completes the proof. ∎

Nevertheless, it is indeed possible to obtain a closed-form expression for the near-far resistance of the optimum multiuser detector, because the minimization of the asymptotic efficiencies with respect to the energies of the interferers reduces the combinatorial optimization problem in (3.11) to a continuous optimization problem whose solution is given by the following result.

Proposition 3.2: Denote the Moore-Penrose generalized inverse (see [3], p. 15) of the normalized crosscorrelation matrix $\mathbf{R}$, by $\mathbf{R}^+$. If the signal of the k^{th} user is linearly independent, i.e. it does not belong to the subspace spanned by the other signals, then

$$\overline{\eta_k} \; = \; \inf_{\substack{w_i \, \geq \, 0 \\ i \, \neq \, k}} \eta_k \; = \; \frac{1}{R^+_{kk}} \tag{3.16}$$

Otherwise, $\overline{\eta_k} \; = \; 0.$ ◊

Proof: Using expression (3.11) for the asymptotic efficiency of the k^{th} user we obtain

$$\overline{\eta_k} \; = \; \min_{\substack{w_i \, \geq \, 0 \\ i \, \neq \, k}} \; \min_{\substack{\epsilon \, \in \, \{-1,0,1\}^K \\ \epsilon_k \, = \, 1}} \frac{1}{w_k} \, \epsilon^T \, \mathbf{WRW} \, \epsilon$$

$$= \min_{\substack{\mathbf{x} \in \mathbb{R}^K \\ x_k = 1}} \mathbf{x}^T \mathbf{R} \mathbf{x}$$

$$= \min_{\mathbf{z} \in \mathbb{R}^{K-1}} (1 + 2\, \mathbf{z}^T \mathbf{m}_k + \mathbf{z}^T \mathbf{R}_k \mathbf{z}) \tag{3.17}$$

where $\mathbf{R}_k$ is obtained from $\mathbf{R}$ by striking out the k^{th} row and column and $\mathbf{m}_k$ is the k^{th} column of $\mathbf{R}$ with the k^{th} entry removed. Henceforth, we will denote such a partitioning of a symmetric matrix with respect to the k^{th} row and column by $\mathbf{R} = [1,\ \mathbf{m}_k,\ \mathbf{R}_k]$, where the leftmost element in the square brackets is the k^{th} diagonal entry. The minimum in the right-hand side of (3.17) is achieved by any element $\mathbf{z}^*$ such that

$$\mathbf{R}_k\, \mathbf{z}^* = -\mathbf{m}_k\,. \tag{3.18}$$

Because of the Fredholm theorem [Lan 85, p. 115] (the range space of a matrix is orthogonal to the nullspace of its transpose), the solvability of (3.18) is equivalent to $\mathbf{m}_k$ being orthogonal to the nullspace of $\mathbf{R}_k$. But for all $\mathbf{z} \in \mathbb{R}^{K-1}$ the parabola $q(v) = v^2 + 2v\, \mathbf{z}^T \mathbf{m}_k + \mathbf{z}^T \mathbf{R}_k \mathbf{z}$ has at most one zero because it is equal to the quadratic form of the nonnegative definite matrix $\mathbf{R}$ with a vector whose k^{th} coordinate is v and whose other components are equal to $\mathbf{z}$. Therefore, the discriminant of the parabola satisfies $(\mathbf{z}^T \mathbf{m}_k)^2 - \mathbf{z}^T \mathbf{R}_k \mathbf{z} \le 0$; in particular, if $\mathbf{z}$ belongs to the nullspace of $\mathbf{R}_k$, then $\mathbf{z}^T \mathbf{m}_k = 0$. So $\mathbf{m}_k$ is indeed orthogonal to the nullspace of $\mathbf{R}_k$. Substituting (3.18) into (3.17) we obtain

$$\begin{aligned} \overline{\eta_k} &= 1 - \mathbf{z}^{*T}\, \mathbf{R}_k\, \mathbf{z}^* \\ &= 1 - \mathbf{z}^{*T}\, \mathbf{R}_k\, \mathbf{R}_k^+\, \mathbf{R}_k \mathbf{z}^* \\ &= 1 - \mathbf{m}_k^T\, \mathbf{R}_k^+\, \mathbf{m}_k\,. \end{aligned} \tag{3.19}$$

Notice that the k^{th} user is linearly dependent if and only if there exists a linear combination of the columns of $\mathbf{R}$ that includes the k^{th} column and is equal to the zero vector. Therefore, if a user is linearly dependent then we can find $\mathbf{x}$ such that $\mathbf{R}\mathbf{x} = 0$ and $x_k = 1$, in which case the penultimate equation in (3.17) indicates that $\overline{\eta_k} = 0$.

In order to obtain the near-far resistance of a linearly independent user, we will employ the following property, which will also be invoked in the sequel.

Lemma 3.2: If the k^{th} user is linearly independent, then every generalized inverse $\mathbf{R}^I$ of $\mathbf{R}$ satisfies: $(\mathbf{R}^I\mathbf{R})_{kj} = \delta_{kj}$, $(\mathbf{R}\mathbf{R}^I)_{jk} = \delta_{jk}$ for $j = 1, \ldots, K$ and $R^I_{kk} = R^+_{kk}$. ◊

Proof of Lemma 3.2: Let $\mathbf{S} = \mathbf{R}^I\mathbf{R} - \mathbf{I}$. By the definition of generalized inverse, it follows that $\mathbf{RS} = \mathbf{0}$, i.e., every column of $\mathbf{S}$ is in the nullspace of $\mathbf{R}$. But if the k^{th} user is linearly independent, it is necessary that the k^{th} element of each such column be zero. Hence $(\mathbf{R}^I\mathbf{R} - \mathbf{I})_{kj} = 0$ for all $j = 1, \cdots K$.

Similarly, with $\mathbf{S} = \mathbf{R}\mathbf{R}^I - \mathbf{I}$ and $\mathbf{SR} = 0$, we obtain $(\mathbf{R}\mathbf{R}^I)_{jk} = \delta_{jk}$. Equivalently $\mathbf{R}\mathbf{R}^I\mathbf{u}_k = \mathbf{u}_k$, using the k_{th} unit vector $\mathbf{u}_k$. Hence, for any generalized inverses $\mathbf{R}^I_1$, $\mathbf{R}^I_2$, $\mathbf{R}(\mathbf{R}^I_1 - \mathbf{R}^I_2)\mathbf{u}_k = 0$. But since the k^{th} user is linearly independent, it is necessary that the k^{th} element of each vector in the nullspace of $\mathbf{R}$ be zero. Hence $(\mathbf{R}^I_1 - \mathbf{R}^I_2)_{kk} = 0$. ■

Proof of Proposition 3.2 (cont.): Partitioning $\mathbf{R}^+$ with respect to the k^{th} row and column we have, say, $\mathbf{R}^+ = [\gamma, \mathbf{c}, \mathbf{C}]$. Now, computing the submatrices of the partitioned matrix $\mathbf{R}^+\mathbf{R}$ and using Lemma 3.2, it follows that

$$\mathbf{R}_k\mathbf{c} + \gamma\,\mathbf{m}_k = 0 \tag{3.20}$$

and

$$\mathbf{c}^T\mathbf{m}_k + \gamma = 1 \tag{3.21}$$

Notice that $\gamma \neq 0$ for otherwise $\mathbf{c}$ would belong to the nullspace of $\mathbf{R}_k$ and would not be orthogonal to $\mathbf{m}_k$, which, as we saw, is not possible. Finally, substituting (3.20) into (3.19) we obtain

$$\begin{aligned}
\overline{\eta}_k &= 1 - \frac{1}{\gamma^2}\,\mathbf{c}^T\,\mathbf{R}_k\mathbf{R}_k^+\mathbf{R}_k\,\mathbf{c} \\
&= 1 - \frac{1}{\gamma^2}\,\mathbf{c}^T\mathbf{R}_k\mathbf{c} \\
&= 1 + \frac{1}{\gamma}\,\mathbf{c}^T\mathbf{m}_k \\
&= \frac{1}{\gamma} = \frac{1}{R^+_{kk}}
\end{aligned} \tag{3.22}$$

where the second, third and fourth equations follow from the definition of generalized inverse, (3.20) and (3.21), respectively. ■

3.3 The decorrelating detector

In the absence of noise the matched filter output vector is $\mathbf{y} = \mathbf{RWb}$, so if the signal set is linearly independent (i.e. $\mathbf{R}$ invertible), the natural strategy to follow in this hypothetical situation is to premultiply $\mathbf{y}$ by the inverse normalized crosscorrelation matrix $\mathbf{R}^{-1}$. The detector $\hat{b} = \text{sgn } \mathbf{R}^{-1}\mathbf{y}$ was analyzed in [Lup 86], where its performance was quantified in the presence of noise. In [Sch 79] it was erroneously shown (cf.[Ver 86b]) that this detector is optimum in terms of bit-error rate. Note that the noise components in $\mathbf{R}^{-1}\mathbf{y}$ are correlated, and therefore sgn $\mathbf{R}^{-1}\mathbf{y}$ does not result in optimum decisions.

Since here (also [Lup 89a]) the signal set is not constrained to be linearly independent, the above detector need not exist. In general, we will consider the set $I(\mathbf{R})$ of generalized inverses (see [(3)], p. 15) of the normalized crosscorrelation matrix $\mathbf{R}$ and we will analyze the properties of the detector

$$\hat{\mathbf{b}} \;=\; \text{sgn } \;\mathbf{R}^{I}\mathbf{y}\;, \tag{3.23}$$

which we refer to as a *decorrelating* detector. Its name is due to the detector's effect upon input of $\mathbf{y}$ when the signal set is linearly independent: then the output is $\mathbf{Wb} + \mathbf{R}^{-1}\mathbf{n}$, i.e. the matched filter outputs have been "decorrelated".

The k^{th} user asymptotic efficiency achieved by a general linear transformation $\mathbf{T}$ can be obtained similarly to that of the conventional single-user detector $\mathbf{T} = \mathbf{I}$ (Section 3.2.1). The first step is to find the bit error probability of the k^{th} user:

$$\begin{aligned} P_k &= P\,[\hat{b}_k = 1 | b_b = -1] \;=\; P\,[(\mathbf{TRWb} \;+\; \mathbf{Tn})_k \;>\; 0 | b_k = -1] \\ &= P\,[(\mathbf{Tn})_k \;>\; (\mathbf{TRW})_{kk} - \sum_{j \neq k} (\mathbf{TRW})_{kj} b_j] \\ &= 2^{1-K} \sum_{\substack{\mathbf{b} \in \{-1,1\}^K \\ b_k = -1}} P\,[(\mathbf{Tn})_k \;>\; (\mathbf{TR})_{kk}\sqrt{w_k} - \sum_{j \neq k} (\mathbf{TR})_{kj} \sqrt{w_j}\; b_j]\,. \end{aligned} \tag{3.24}$$

Since the random variable $(\mathbf{Tn})_k$ is Gaussian with zero mean and variance equal to $\sigma^2 (\mathbf{TRT}^T)_{kk}$, the sum in (3.24) is dominated as $\sigma \to 0$ by the term

$$2^{1-K}\; Q\,(\min_{\substack{\mathbf{b} \in \{-1,1\}^K \\ b_k = -1}} [(\mathbf{TR})_{kk}\sqrt{w_k} - \sum_{j \neq k} (\mathbf{TR})_{kj}\sqrt{w_j}\; b_j] \;/\; \sigma\sqrt{(\mathbf{TRT}^T)_{kk}})$$

which is equal to

$$2^{1-K}Q(\ ((\mathbf{TR})_{kk}\sqrt{w_k}\ -\ \sum_{j\neq k}|(\mathbf{TR})_{kj}|\sqrt{w_j})/\sigma\sqrt{(\mathbf{TRT}^T)_{kk}}\ )\,. \tag{3.25}$$

Hence according to definition (2.1) the k^{th} user asymptotic efficiency of the linear receiver is equal to zero if $(\mathbf{TR})_{kk}\sqrt{w_k} \leq \sum_{j\neq k}|(\mathbf{TR})_{kj}|\sqrt{w_j}$. Otherwise it is either equal to the square of the ratio of the argument of the foregoing Q-function and the argument corresponding to the single-user probability of error $Q(\sqrt{w_k}/\sigma)$, or equal to 1, whichever is smaller. Therefore

$$\eta_k(\mathbf{T})\ =\ \min\{1,\ \max{}^2\{0,\ \frac{(\mathbf{TR})_{kk}\ -\ \sum_{j\neq k}|(\mathbf{TR})_{kj}|\sqrt{\frac{w_j}{w_k}}}{\sqrt{(\mathbf{TRT}^T)_{kk}}}\}\}\,. \tag{3.26}$$

The min-operation can be seen to be redundant in this case as follows. Clearly the claim is true if $(\mathbf{TR})_{kk} \leq 0$. Otherwise, since all the terms subtracted in the numerator are nonnegative, it is sufficient to show for all $\mathbf{T}$

$$(\mathbf{TR})_{kk}\ \leq\ \sqrt{(\mathbf{TRT}^T)_{kk}}\,. \tag{3.27}$$

Let $\mathbf{v}^\mathbf{T}$ denote the k^{th} row of $\mathbf{T}$ and $\mathbf{r}_k$ the k^{th} column of $\mathbf{R}$. Then we have to show

$$\mathbf{v}^T\mathbf{r}_k\ \leq\ \sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}} \tag{3.28}$$

or, after squaring both sides and collecting terms, it suffices to show that the kernel matrix is nonnegative definite, i.e.

$$\mathbf{r}_k\mathbf{r}_k^T\ -\ \mathbf{R}\ \leq\ 0\,. \tag{3.29}$$

We show this as follows. Letting $\mathbf{u}_k$ be the k^{th} unit vector,

$$(\mathbf{x}-x_k\mathbf{u}_k)^T\ \mathbf{R}\ (\mathbf{x}-x_k\mathbf{u}_k)\ =\ \mathbf{x}^T\mathbf{R}\mathbf{x}\ -\ 2x_k\mathbf{x}^T\mathbf{r}_k\ +\ x_k^2\ \geq\ 0 \tag{3.30}$$

for all pairs $(\mathbf{x}, x_k) \in (\mathbb{R}^K, \mathbb{R})$, since $\mathbf{R}$ is nonnegative definite. But (3.30) can be viewed as a second degree polynomial in x_k, whose discriminant has to be nonpositive in order to guarantee that the polynomial does not change sign. Hence

$$\Delta\ =\ (\mathbf{x}^T\mathbf{r}_k)^2\ -\ \mathbf{x}^T\mathbf{R}\mathbf{x}\ \leq\ 0 \tag{3.31}$$

or equivalently, for all $\mathbf{x}$

$$\mathbf{x}^T(\mathbf{r}_k\mathbf{r}_k^T\ -\ \mathbf{R})\mathbf{x}\ \leq\ 0$$

which completes the proof that the min-operation in (3.26) is superfluous.

Hence the k^{th} user asymptotic efficiency achieved by the linear mapping $\mathbf{T}$ is

$$\eta_k(\mathbf{T}) \;=\; \max{}^2 \left\{ 0, \; \frac{(\mathbf{TR})_{kk} \;-\; \sum_{j \neq k} |\,(\mathbf{TR})_{kj}\,| \sqrt{\frac{w_j}{w_k}}}{\sqrt{(\mathbf{TRT}^T)_{kk}}} \right\} . \tag{3.32}$$

Thus the k^{th} user asymptotic efficiency of a decorrelating detector with matrix $\mathbf{R}^I$ is given by

$$\eta_k(\mathbf{R}^I) \;=\; \max{}^2 \left\{ 0, \; \frac{(\mathbf{R}^I\mathbf{R})_{kk} \;-\; \sum_{j \neq k} |\,(\mathbf{R}^I\mathbf{R})_{kj}\,| \sqrt{\frac{w_j}{w_k}}}{\sqrt{(\mathbf{R}^I\mathbf{R}\mathbf{R}^{I^T})_{kk}}} \right\} . \tag{3.33}$$

Define the highest asymptotic efficiency achievable by a generalized inverse, respectively the optimum linear map by

$$\eta_k^d \;\stackrel{\triangle}{=}\; \sup_{\mathbf{R}^I \in I(\mathbf{R})} \quad \eta_k(\mathbf{R}^I) \tag{3.34}$$

and

$$\eta_k^l \;\stackrel{\triangle}{=}\; \sup_{\mathbf{T} \in \mathbb{R}^{K \times K}} \quad \eta_k(\mathbf{T}) \, . \tag{3.35}$$

Proposition 3.3: If User k is linearly independent, every $\mathbf{R}^I \in I(\mathbf{R})$ satisfies

$$\eta_k \; (\mathbf{R}^I) \;=\; \eta_k^d \;=\; 1/R_{kk}^+ \, , \tag{3.36}$$

where the notation is as in Proposition 3.2. ◇

Thus for independent users the asymptotic efficiency of the decorrelating detector is independent of the energy of other users and of the specific generalized inverse selected.

Proof : If user k is linearly independent we established in Lemma 3.2 that $(\mathbf{R}^I\mathbf{R})_{kj} = \delta_{kj}$. Hence, it follows from (3.33) that

$$\eta_k(\mathbf{R}^I) \;=\; \frac{1}{R_{kk}^I} \tag{3.37}$$

and Proposition 3.3 follows, using the fact that, by Lemma 3.2, $R_{kk}^I = R_{kk}^+$. ∎

If User k is linearly independent it follows from Lemma 3.1 that the Moore-Penrose decorrelating detector, which isolates User k, has a probability of error given by

$$P_k(e) \;=\; Q\,(\frac{\sqrt{w_k}}{\sigma\,\sqrt{R^{+}_{kk}}})\,. \tag{3.38}$$

In Section 3.5.2 it is shown that if user k is linearly dependent, then

$$\eta^d_k \;=\; \eta^l_k$$

i.e. for a dependent user the best decorrelating detector and the best linear detector achieve the same k^{th} user asymptotic efficiency.

Proposition 3.4 : The near-far resistance of the decorrelating detector equals that of the optimum multiuser detector, i.e., for all $\mathbf{R}^I \in I(\mathbf{R})$,

$$\inf_{\substack{w_j \,\geq\, 0 \\ j\neq k}} \;\eta_k(\mathbf{R}^I) \;=\; \inf_{\substack{w_j \,\geq\, 0 \\ j\neq k}} \;\eta_k \;\equiv\; \overline{\eta_k}\,. \tag{3.39}$$

◇

Proof : If User k is linearly independent, then according to Propositions 3.2 and 3.3 the near-far resistance of the optimum detector is equal to the asymptotic efficiency of the decorrelating detector, which is independent of the energy of the other users. If user k is linearly dependent, Proposition 3.2 states that the near-far resistance of the optimum detector is zero, and hence the same is true for any detector. ∎

The result of Proposition 3.4 is of special importance in a near-far environment, where the received signals have different energies, and where the energy ratios may vary continuously over a broad scale if the positions of the users evolve dynamically. In this environment any decorrelating detector, with its linear time-complexity per bit, offers the same near-far resistance as the optimum multiuser detector, whose time-complexity per bit is exponential.

Proposition 3.5: If the signature waveforms are linearly independent, the k^{th} user asymptotic efficiency of the decorrelating detector is lower bounded by:

$$\eta^d_k \;=\; \frac{1}{\mathbf{R}^{-1}_{kk}} \;\geq\; \frac{4\;\lambda_{\max}/\lambda_{\min}}{(\lambda_{\max}/\lambda_{\min}\;+\;1)^2}\,. \tag{3.40}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the largest respectively smallest eigenvalues of $\mathbf{R}$. ◇

This gives a lower bound of $.89, .75, .56, .33$ and $.04$ for a spectral condition number $\lambda_{\max}/\lambda_{\min}$ of $2, 3, 5, 10$ and 100, respectively. As always when dealing with matrix inversion, a small eigenvalue spread is desirable.

Proof: We use Kantorovich's inequality [Horn 85, p.444], which states that, given a positive definite matrix $\mathbf{B}$ with eigenvalues $0 < \lambda_1 \leq \ldots \leq \lambda_n$,

$$\|\mathbf{x}\|_2^4 \geq \frac{4\, \lambda_1 \lambda_n}{(\lambda_1 + \lambda_n)^2} (\mathbf{x}^* \mathbf{B} \mathbf{x})(\mathbf{x}^* \mathbf{B}^{-1} \mathbf{x}) \tag{3.41}$$

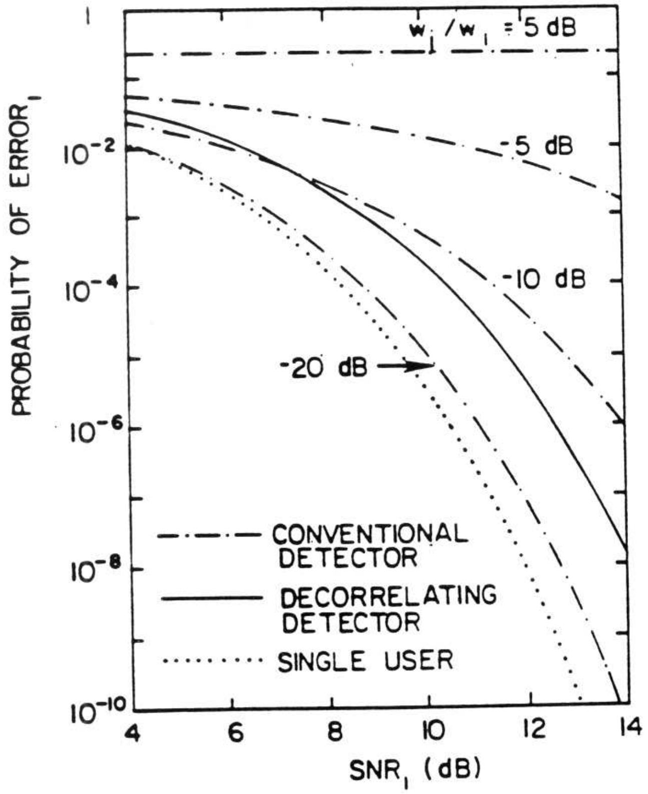
for all $\mathbf{x} \in \mathbb{C}^n$. Moreover, there is a unit-norm vector $\mathbf{x}$ for which equality holds. Proposition 3.5 follows by letting $\mathbf{x}$ be the k^{th} unit vector, and using the fact that $R_{kk} = 1$. ■

Numerical examples

The following examples illustrate the difference between the error probability behavior of the conventional and of the decorrelating detector, for a 3-user and for a 6-user environment. The waveforms used are Spread-Spectrum m-sequences of length 31. The first example, shown in Figure 1, employs the set of 3 sequences reported in [Gar 80, Table 5], to be optimal with respect to a signal-to-multiple-access interference parameter when the conventional detector is used. These sequences have also been used in related works ([Ger 82], [Ver 86a]).

The second example, Figure 2, uses the set of auto-optimal m-sequences of length 31 found in [Pur 79, Fig. A.1] to be optimal with respect to certain peak and mean-square correlation parameters which play an important role in the error probability analysis of the conventional detector. The figures show the error probability of User 1 in a baseband environment (where the crosscorrelation values are highest) with equal energy interferers, whose energy ratio to User 1 is the parameter which indexes the different error probability curves for the conventional receiver. Also shown are the error probability of the decorrelating detector for user 1 and, for comparison purposes, the error probability of the single user channel. Note that the former is independent of the energy of the interferers. The matrix $\mathbf{R}$ is given for interest, as well as the decorrelating detector asymptotic efficiencies of all the users.

Both figures illustrate the strong dependence of the performance of the conventional receiver on the relative energies of the active users, and the fact that the error probability of the conventional



crosscorrelation matrix R :

1.00	0.55	0.61
0.55	1.00	0.81
0.61	0.81	1.00

decorrelating asymptotic efficiencies of all 3 users

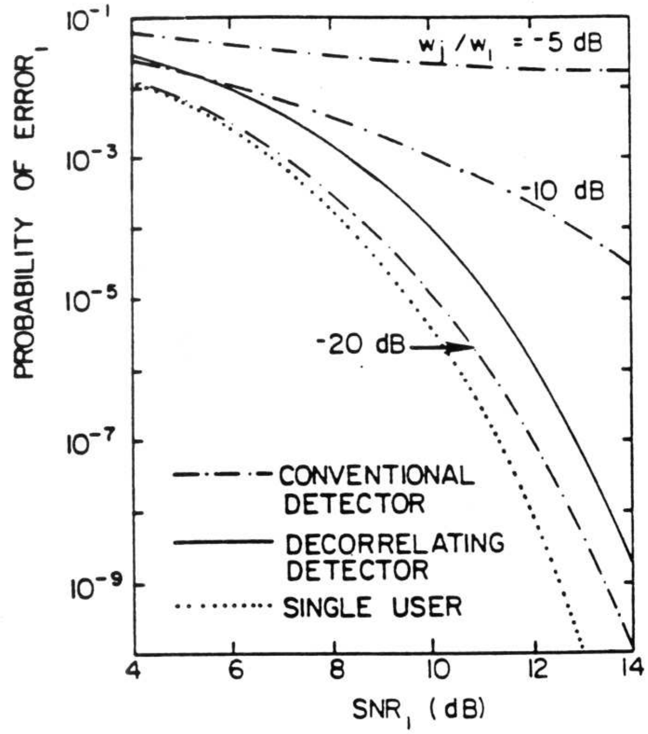
0.78 0.59 0.55

Fig. 1. Error probability of User 1 with 2 active equal energy interferers, each of energy w_j, averaged over the interfering bit sequences, for the decorrelating and conventional receiver versus the SNR of User 1, for m-sequences of length 31 and different interference levels.

receiver becomes irreducible even for vanishing background noise levels if the interference energy is high enough. For 6 users the latter is seen to happen if each of the interferers has more than 1/3 the power of User 1. Only if the multiple-access interference level plays a negligible role compared to the background noise does the conventional detector outperform the decorrelating detector, which pays a penalty for combatting the interference instead of ignoring it.

The same sets of sequences are used to illustrate the error probability constellation in the asynchronous case (Section 4.3, Fig. 26, 27). The single error probabilities can be seen to be lower in the asynchronous case, though the qualitative relations stay the same.

Figure 3 shows the asymptotic efficiency of User 1 achieved by the conventional detector, the optimum multiuser detector and the decorrelating detector, for two users with crosscorrelation



crosscorrelation matrix **R** :

1.00	0.35	0.23	0.35	0.35	0.48
0.35	1.00	0.48	0.48	0.61	0.35
0.23	0.48	1.00	0.74	0.48	0.35
0.35	0.48	0.74	1.00	0.74	0.35
0.35	0.61	0.48	0.74	1.00	0.35
0.48	0.35	0.35	0.35	0.35	1.00

decorrelating asymptotic efficiencies of all 6 users

0.83	0.73	0.62	0.49	0.58	0.83

Fig. 2. Same as Fig.1, with 5 active equal energy interferers.

coefficient $R_{12} = \rho$, versus the square root of the energy ratio of the two users. The figure shows the good performance of the optimum detector, who asymptotically performs as well as in the absence of a second user, if this user is powerful enough; the decay to zero of the asymptotic efficiency of the conventional detector for relatively low interference power, and the energy independence of the asymptotic efficiency of the decorrelating detector, which is much superior to the conventional detector except for very low interference.

3.4 Optimality criteria leading to the decorrelating detector

In this chapter we assume that **R** is invertible and present a number of optimality criteria which lead to the decorrelating detector, thus providing further justification for its study. We have already mentioned (Section 3.3) that the decorrelating detector is the optimal strategy in the absence of noise.

Proposition 3.6: The decorrelating detector is the maximum likelihood detector in the case when the energies are not known to the receiver. $\diamond$

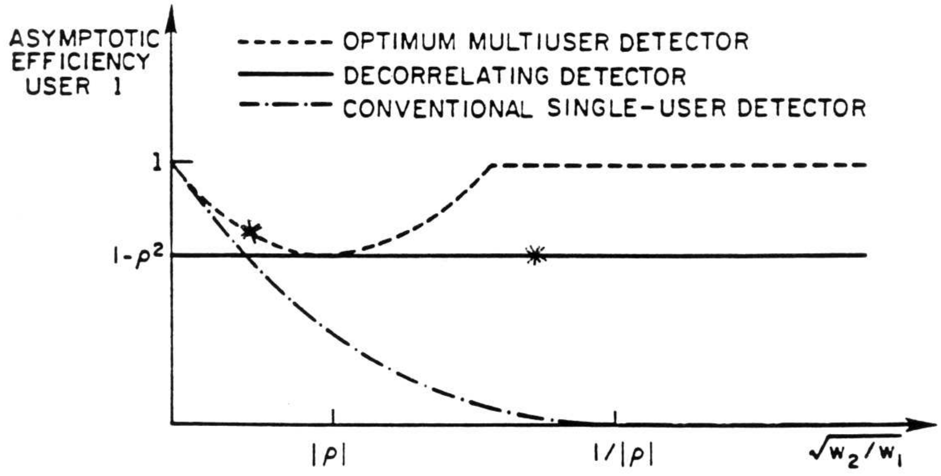


Fig. 3. Asymptotic efficiencies in the 2-user case ($\rho = 0.6$)

The * indicates the asymptotic efficiency of the best linear detector.

Proof : The maximum likelihood receiver selects the decisions that maximize the maximum of the likelihood function over the unknown parameters ([Poor, Ch.2], [Hel, p. 291]), i.e.

$$\begin{aligned}\hat{\mathbf{b}} &\in \arg \min_{\mathbf{b}\in\{-1,1\}^K} \min_{\substack{w_i>0\\ i=1,\ldots,K}} \int_0^T [\, r(t) - \sum_{k=1}^{K} b_k \sqrt{w_k}\, s_k(t)\,]^2\, dt \\ &= \arg \min_{\mathbf{b}\in\{-1,1\}^K} \min_{\substack{w_i>0\\ i=1,\ldots,K}} \int_0^T r^2(t)\, dt - 2\, \mathbf{y}^T \mathbf{W}\, \mathbf{b} + \mathbf{b}^T \mathbf{W}\, \mathbf{R}\, \mathbf{W}\, \mathbf{b} \\ &= \operatorname{sgn} \left(\arg \min_{\mathbf{x}\in \mathbb{R}^K} \mathbf{x}^T \mathbf{R}\, \mathbf{x} - 2\, \mathbf{x}^T \mathbf{y} \right) = \operatorname{sgn} \mathbf{R}^{-1}\, \mathbf{y} \qquad (3.42)\end{aligned}$$

■

Proposition 3.7: The decorrelating detector is the limit as the Gaussian noise level tends to zero ($\sigma \to 0$) of the minimum-variance linear estimate of $\mathbf{b}$ given $\mathbf{y}$, followed by a sign decision. ♢

Proof : The minimum-variance linear estimate (e.g. [Lue, p. 87]) of $\mathbf{b}$ given $\mathbf{y}$ is $\tilde{\mathbf{b}} = \mathbf{T}^*\mathbf{y}$ where

$$\mathbf{T}^* = \arg \min_{\mathbf{T}\in \mathbb{R}^{K\times K}} E\, \|\mathbf{T}\mathbf{y} - \mathbf{b}\|^2 \qquad (3.43)$$

The expectation is with respect to the noise and to the transmitted information vector, the two of which are independent, and $\|\cdot\|$ is the Euclidean norm. From the Projection theorem, the optimum estimate is obtained when each component of the estimation error is orthogonal to each component of the measurement vector $\mathbf{y}$. Thus

$$E\left[(\mathbf{T}^*\mathbf{y} - \mathbf{b})\,\mathbf{y}^T\right] = 0$$

whence,

$$\mathbf{T}^* = E\left[\mathbf{b}\,\mathbf{y}^T\right]\left[E\left[\mathbf{y}\mathbf{y}^T\right]\right]^{-1}. \tag{3.44}$$

Since the bits transmitted by different users are independent of each other and of the background noise, the above expectations are

$$E\left[\mathbf{b}\,\mathbf{y}^T\right] = E\left[\mathbf{b}\,(\mathbf{b}^T\mathbf{W}\mathbf{R} + \mathbf{n}^T)\right] = \mathbf{W}\,\mathbf{R} \tag{3.45}$$

and

$$E\left[\mathbf{y}\mathbf{y}^T\right] = E\left[(\mathbf{R}\mathbf{W}\mathbf{b} + \mathbf{n})\,(\mathbf{b}^T\mathbf{W}\mathbf{R} + \mathbf{n}^T)\right] = \mathbf{R}\,\mathbf{W}^2\,\mathbf{R} + \sigma^2\mathbf{R}, \tag{3.46}$$

so that the minimum-variance linear estimate is given by

$$\mathbf{T}^* = \mathbf{W}\,\mathbf{R}\,(\mathbf{R}\,\mathbf{W}^2\,\mathbf{R} + \sigma^2\mathbf{R})^{-1} = \mathbf{W}\,(\mathbf{R}\,\mathbf{W}^2 + \sigma^2\mathbf{I})^{-1} \tag{3.47}$$

and

$$\lim_{\sigma\to 0}\mathbf{T}^* = \mathbf{W}^{-1}\,\mathbf{R}^{-1}. \tag{3.48}$$

Finally, since in order to use the additional information that $\mathbf{b}$ is binary data, a sign decision is taken on $\tilde{\mathbf{b}}$, multiplication by the diagonal matrix with positive entries $\mathbf{W}^{-1}$ does not affect the resulting decision $\hat{\mathbf{b}}$, and can thus be omitted. ■

(3.8): The decorrelating detector is the analogue in multiuser communication of the zero-forcing solution to the problem of minimizing peak distortion in automatic equalization. ◊

In his pioneering paper [Luc 65], Lucky considers the automatic equalization problem of recovering the term a_o from

$$y_o = h_o\left[a_o + \frac{1}{h_o}\sum_{n\neq 0} a_n\,h_{-n}\right] \tag{3.49}$$

where the second term constitutes intersymbol interference and the h_n depend linearly on the set of N parameters $c_j, j \in K_N$ (called tap gains), which the system designer is free to choose, via

$$h_n = \sum_{j\in K_N} c_j\,x_{n-j}. \tag{3.50}$$

The criterion Lucky chooses to minimize, since it is the maximum value the intersymbol interference term can assume, is the so-called peak distortion

$$D \triangleq \frac{1}{h_o} \sum_{n \neq 0} | h_n | . \tag{3.51}$$

Under the condition that the level of the initial distortion $D_o \triangleq \sum_{n \neq 0} | x_n |$ is less than 1, Lucky shows that the zero-forcing solution is optimal, namely to choose the tap gains which simultaneously cause $h_n = 0$, for all $n \in K_N, n \neq 0$.

Setting our problem up along these terms, we have (for User k)

$$\tilde{b}_k = \mathbf{v}^T \mathbf{y} = \mathbf{r}_k^T \mathbf{v} \sqrt{w_k} b_k + \sum_{j \neq k} \mathbf{r}_j^T \mathbf{v} \sqrt{w_j} b_j + n'_k \tag{3.52}$$

therefore in our case the peak distortion is

$$D(\mathbf{v}) = \frac{1}{\mathbf{r}_k^T \mathbf{v}} \sum_{j \neq k} | \mathbf{r}_j^T \mathbf{v} | \sqrt{\frac{w_j}{w_k}} . \tag{3.53}$$

From here it is apparent that the decorrelating detector would be optimal if we wanted to minimize the peak distortion, a trivial result in this case, since in our case we have the same number of tap weights as of interfering samples, which means that we can force the peak distortion to zero. Note that the asymptotic efficiency can be expressed in terms of the peak distortion D as

$$\eta_k(\mathbf{v}) = \frac{\mathbf{r}_k^T \mathbf{v} \ (1 - D(\mathbf{v}))}{\sqrt{\mathbf{v}^T \ \mathbf{R} \ \mathbf{v}}} \tag{3.54}$$

and we show in Section 3.5 that though in general the zero-forcing solution (i.e. the decorrelating detector) is not the optimal linear rule, there exists a region of energies where it is. In his formulation of the problem Lucky neglects additive background noise, and motivates his choice of the peak distortion criterion with the words “it is a minimax criterion in that we seek to maximize the customer’s minimum margin against noise over all data sequences”. A more appropriate procedure might be to also take into account the noise, since the noise variance is also affected by the equalizer. In this case one would maximize a functional equivalent to the asymptotic efficiency.

3.5 The optimum linear multiuser detector

We now turn to the question of finding the optimum linear detector. We have seen that this is a fruitful approach, since a particular type of linear detector, the decorrelating detector, offered a substantial improvement in asymptotic efficiency compared to the single-user detector, while its near-far resistance equaled that of the optimum multiuser detector. While we now know that no detector, either linear or nonlinear, can outperform the decorrelating detector with respect to near-far resistance, for fixed energies it is indeed possible to obtain linear detectors that have a higher asymptotic efficiency than the one achieved by the decorrelating detector.

We find the linear detector which maximizes the asymptotic efficiency (or equivalently minimizes the probability of bit error in the low-noise region) and compare the achieved asymptotic efficiency to the ones achieved by the conventional and optimal detectors. Thus we ask which mapping $\mathbf{T}\colon \mathbb{R}^K \to \mathbb{R}^K$ maximizes the asymptotic efficiency of the decision scheme

$$\hat{\mathbf{b}} \;=\; \mathrm{sgn}\;(\;\mathbf{Ty}\;) \;=\; \mathrm{sgn}\;(\;\mathbf{TRWb} + \mathbf{Tn}\;)\;. \tag{3.55}$$

The interpretation of this optimization problem in terms of decision regions is to find the optimal partition of the K -dimensional hypotheses space into K decision-cones with vertices at the origin. The surfaces of these cones determine the columns of the inverse $\mathbf{T}^{-1}$ of the desired mapping. Application of $\mathbf{T}$ on the cone configuration will map the cones on quadrants, after which a sign detector is used.

Letting $\mathbf{v}^{\mathbf{T}}$ denote the k^{th} row of $\mathbf{T}$, the k^{th} user asymptotic efficiency of a general linear detector, as given by (3.55) was derived in (3.32):

$$\eta_k(\mathbf{T}) \;=\; \max{}^2\;\{0,\; \frac{\mathbf{r}_k^T\mathbf{v} - \sum_{j\neq k} \mid \mathbf{r}_j^T\mathbf{v} \mid \sqrt{\frac{w_j}{w_k}}}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}}\;\}\;. \tag{3.56}$$

The best linear detector has the asymptotic efficiency

$$\eta_k^l \;=\; \sup_{\mathbf{v}\in\mathbb{R}^K} \;\eta_k(\mathbf{v})\;. \tag{3.57}$$

Hence the asymptotic efficiency of the best linear detector is equal to

$$\eta_k^l = \sup_{\mathbf{v}\in \mathbb{R}^K} \max{}^2 \{0, \frac{\mathbf{r}_k^T\mathbf{v} - \sum_{j\neq k} \mid \mathbf{r}_j^T\mathbf{v} \mid \sqrt{\frac{w_j}{w_k}}}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}}\} \tag{3.58}$$

$$= \max{}^2 \{0, \sup_{\mathbf{v}\in \mathbb{R}^K} \eta_k(\mathbf{v})\} \qquad \text{with } \eta_k(\mathbf{v}) = \frac{\mathbf{r}_k^T\mathbf{v} - \sum_{j\neq k} \mid \mathbf{r}_j^T\mathbf{v} \mid \sqrt{\frac{w_j}{w_k}}}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}} . \tag{3.59}$$

In order to minimize the probability of bit-error, P_k, we have to maximize the smallest argument in the sum of Q-functions, and equivalently maximize the asymptotic efficiency $\eta_k(\mathbf{v})$, with respect to the components of the vector $\mathbf{v}$. Since the map applied on the matched filter outputs is linear, the asymptotic efficiencies of all the users can be simultaneously maximized, each such maximization yielding the corresponding row of the map to be applied.

For the sake of clarity we first consider the two-user case, for which explicit expressions for the maximum linear asymptotic efficiency can be obtained.

3.5.1 The two-user case

Throughout this subsection we denote the normalized crosscorrelation between both signals by $\rho = R_{12}$. Without loss of generality, let $k = 1$. We first give an explicit expression for the optimum linear detector:

Proposition 3.9 : The 1^{st} user optimal linear transformation $\mathbf{T}_1(\mathbf{y}) = \mathbf{v}^T\mathbf{y}$ on the matched filter outputs prior to threshold detection is given by

$$\mathbf{v}^T = [\ 1\ ;\ -\text{sgn}\rho \ \ \min\left\{ \frac{\sqrt{w_2}}{\sqrt{w_1}},\ |\rho| \right\}], \tag{3.60}$$

$$= \begin{cases} [1 \quad -\text{sgn}\rho\ \sqrt{w_2/w_1}]\ , & \text{if } \sqrt{w_2/w_1} \leq |\rho| \\ [1 \quad -\rho]\ , & \text{otherwise .} \end{cases} \tag{3.61}$$

◇

Note that $[1 \quad -\rho]$ is the 1^{st} row of the decorrelating detector.

Proof : We have

$$\mathbf{R} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \qquad \mathbf{v}^T = [1\ ;\ v_2] \tag{3.62}$$

$$\eta_1(\mathbf{v}) = \frac{\mathbf{r}_1^T\mathbf{v} - |\ \mathbf{r}_2^T\mathbf{v}\ |\sqrt{w_j/w_k}}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}} = \frac{1 + \rho\, v_2 - |\ \rho + v_2\ |\sqrt{w_j/w_k}}{\sqrt{1 + 2\rho\ v_2 + v_2^2}} \tag{3.63}$$

and the objective is to maximize the right-hand side of (3.63) with respect to v_2. We consider the case $|\rho| = 1$ separately. Equation (3.63) depends on the user energies only through the ratio $r \triangleq \sqrt{w_2/w_1}$. With this substitution

$$\eta_1(v_2) = \frac{1 + \rho\, v_2 - |\ \rho + v_2\ |\ r}{\sqrt{1 + 2\rho\ v_2 + v_2^2}}\ . \tag{3.64}$$

a) Case $|\rho| \neq 1$: Introduce an indicator function for the absolute value term, as follows.

$$I = \begin{cases} 1\ , & \rho + v_2 > 0 \\ -1\ , & \rho + v_2 < 0 \\ 0\ , & \text{else}\ . \end{cases} \tag{3.65}$$

Then

$$\frac{d\eta_1}{dv_2} = -\frac{(\ 1 - \rho^2\ )\ (\ I\ r\ +\ v_2\ )}{(\ 1 + 2\rho\ v_2 +\ v_2^2\ )^{3/2}}\ . \tag{3.66}$$

Therefore we should take $v_2 = -I\ r$ when this is consistent with the definition of I as a function of v_2. Thus,

$$\begin{aligned} v_2 &= r \quad \text{if } I = -1 \Longleftrightarrow 0 < r < -\rho \\ v_2 &= -r \quad \text{if } I = 1 \Longleftrightarrow 0 < r < \rho\ . \end{aligned} \tag{3.67}$$

As can easily be seen, both values correspond to maxima. If neither of these conditions is met, the derivative does not have a zero. The optimal value for v_2 can be determined from a closer look at the behavior of $d\eta_1/dv_2$ of (3.66), shown in Figure 4 for both $I = 1$ and $I = -1$. Looking at the curves we see the following. For both $I = 1$ and $I = -1$, the derivative of η_1 is positive for v_2 smaller than the abscissa of the zero of the derivative (which is equal to $-I\ r$), and negative afterwards. Due to the nonlinearity of η_1 the derivative has the form corresponding to $I = -1$ for $v_2 < -\rho$ and the form corresponding to $I = 1$ afterwards. The dashed lines show possible positions of $-\rho$ on the v_2 axis. Depending on where $-\rho$ is located relative to $-r$ and r, the resulting derivative will have a zero (this happens when $-\rho < -r$ or $-\rho > r$), or not (otherwise). In the latter case, since the second branch (for $I = 1$) turns negative before the first one, we have to take

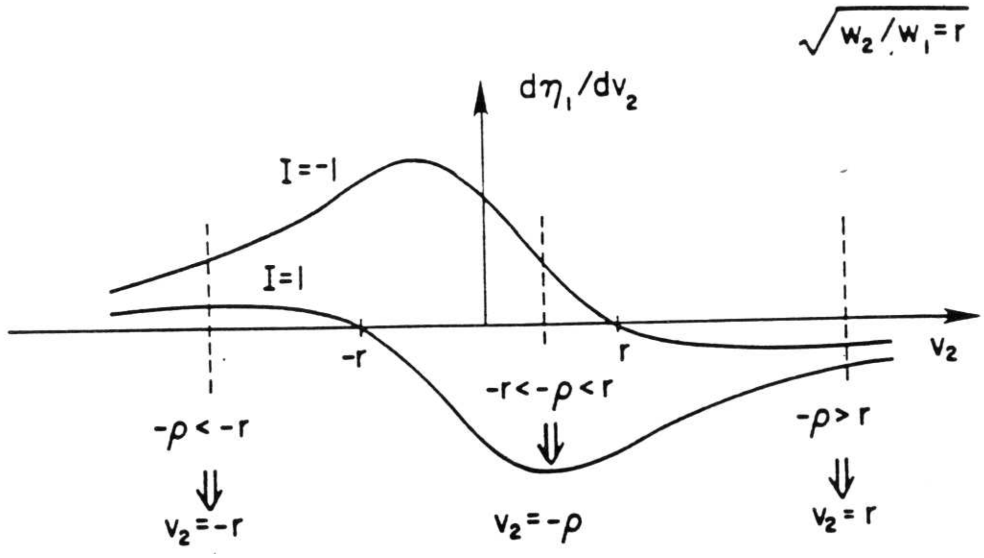


Fig. 4. Behavior of the derivative in (3.66).

the largest value of v_2 yielding a positive derivative on the first branch. It can easily be seen that in the "no-zero" case, $-r < -\rho < r$, this is the point of discontinuity, i.e. $v_2 = -\rho$. Note that for $\rho = 0$ we get $\mathbf{v}^T = [1 \quad 0]$, the identity transformation, as expected, since then the users are decoupled and a single-user detector is optimal. By taking the inverse of $\mathbf{R}$ we also see that in the "no-zero" case the optimal transformation vector is exactly the corresponding row of the inverse correlation matrix.

b) Case $|\rho| = 1$: Equation (3.64) becomes

$$\begin{aligned}\eta_1(v_2) &= \frac{1 + \mathrm{sgn}\rho\; v_2 \;-\; |1 + \mathrm{sgn}\rho\; v_2|\; r}{|1 + \mathrm{sgn}\rho\; v_2|} \\ &= \mathrm{sgn}\,(1 + \mathrm{sgn}\rho\; v_2) \;-\; r\,. \end{aligned} \tag{3.68}$$

We see that for $r < 1$, any v_2 satisfying $v_2\ \mathrm{sgn}\rho > -1$ is optimal, in particular the one given in (3.61). Otherwise the asymptotic efficiency of the best linear transformation is 0, hence all linear transformations are equivalent. Substituting the result of Proposition 3.9 into the asymptotic efficiency of (3.64), we obtain the following.

Proposition 3.10 : The k^{th} user asymptotic efficiency of the optimal linear two-user detector equals :

$$\eta_k^l = \begin{cases} 1 - 2|\rho|(w_i/w_k)^{1/2} + w_i/w_k \ , & \text{if} \ \ (w_i/w_k)^{1/2} \leq |\rho| \\ 1 - \rho^2 \ , & \text{otherwise} \end{cases} \tag{3.69}$$

for $(i, k) \in \{(1, 2), (2, 1)\}$. ♢

The k^{th} user asymptotic efficiency obtained in the range $(w_i/w_k)^{1/2} < |\rho|$ equals the optimum asymptotic efficiency, obtained in (3.12). Even where it equals the decorrelating detector, outside the region of optimality, the best linear detector shows a far better performance than the conventional single user detector (see Figure 3), since if $w_i/w_k > \rho^2$, then η_1^l is independent of w_i/w_k, whereas according to (3.7) the asymptotic efficiency of the conventional detector is equal to zero for $w_i/w_k \geq 1/\rho^2$.

There is an intuitive interpretation of the dual behavior of the best linear detector and of the boundary point $r = |\rho|$. The input to the threshold device corresponding to the first user, $z_1 = \mathbf{v}^T\mathbf{y}$, has three components:

$$\begin{aligned} z_1 \ &= \ \sqrt{w_1} \ [1 - \rho^2 + \rho \ (\rho + v_2)] \ b_1 \\ &+ \sqrt{w_1} \ [\qquad r \ (\rho + v_2)] \ b_2 \ \ + \ \tilde{n} \ , \qquad \tilde{n} \sim N(0, \sigma^2[1 - \rho^2 + (\rho + v_2)^2]). \end{aligned} \tag{3.70}$$

For $r > |\rho|$, the second term outweighs the second part of the first term, so the best one can do is to eliminate it, by choosing $v_2 = -\rho$ (the decorrelating detector). Since this minimizes the noise variance at the same time, it is the best strategy in this region. If, however, $r < |\rho|$, and if additionally v_2 is such that the term $\rho(\rho + v_2)$ is positive, it is a better policy to allow interference from User 2, which is compensated by the second part in the first term, and use the residual positive contribution in the first term to increase the SNR compared to the decorrelating case. We have seen that this strategy leads to the same performance as the more complex maximum likelihood detector.

Note that in the two-user case the signal energies and crosscorellations cannot be picked such as to allow *both* users optimal performance at the same time: for User 1 we need $r < |\rho| < 1$, whereas for User 2 we need $r > \frac{1}{|\rho|} > 1$.

3.5.2 The K-user case

Unlike Propositions 3.3 and 3.10, in the general K-user case it is not feasible to obtain an explicit expression for the asymptotic efficiency achieved by the best linear detector.

Proposition 3.11: The k^{th} user asymptotic efficiency of the best linear detector equals:

$$\eta_k^l = \max{}^2 \{0, \max_{\substack{e_j \in \{-1,1\} \\ j \neq k}} \eta(\mathbf{e})\} \qquad \text{with} \quad \eta(\mathbf{e}) = \sup_{\substack{\mathbf{v} \in R^K \\ \mathbf{v}^T \mathbf{R} \mathbf{v} = 1 \\ e_j \mathbf{r}_j^T \mathbf{v} \geq 0 \\ j \neq k}} \mathbf{v}_o^T \mathbf{R} \mathbf{v} \tag{3.71}$$

where the i^{th} component of $\mathbf{v}_o$ is equal to $(\mathbf{v}_o)_i = \begin{cases} -e_i \sqrt{w_j/w_k}, & i \neq k \\ 1, & i = k \end{cases}$

Then the maximum $\eta(\mathbf{e})$ is achieved for $\tilde{\mathbf{v}}$ such that

$$\tilde{\mathbf{v}} = \frac{\mathbf{v}_o + \sum_{j \neq k} \lambda_j e_j \mathbf{u}_j}{(\mathbf{v}_o^T \mathbf{R} \mathbf{v}_o + \mathbf{v}_o^T \mathbf{R} \sum_{j \neq k} \lambda_j e_j \mathbf{u}_j)^{1/2}}, \qquad (\mathbf{u}_j)_i = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \tag{3.72}$$

$$e_j \mathbf{r}_j^T \tilde{\mathbf{v}} \geq 0 \quad \text{for} \quad j \neq k \tag{3.73}$$

$$\mathbf{r}_j^T \tilde{\mathbf{v}} \neq 0 \quad => \quad \lambda_j = 0 \tag{3.74}$$

$$\lambda_j \geq 0 \quad j \neq k \tag{3.75}$$

◇

Proof : Let

$$S_j^+ = \{\mathbf{x} \in I\!\!R^K \ : \ \mathbf{r}_j^T \mathbf{x} \geq 0 \} \tag{3.76}$$

$$S_j^- = \{\mathbf{x} \in I\!\!R^K \ : \ \mathbf{r}_j^T \mathbf{x} \leq 0 \}$$

From (3.59) we seek

$$\sup_{\mathbf{v}\in\mathbb{R}^K} \frac{1}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}}(\mathbf{r}_k^T\mathbf{v} - \sum_{j\neq k} |\mathbf{r}_j^T\mathbf{v}| \sqrt{\frac{w_j}{w_k}})$$

$$= \max_{\substack{e_j\in\{-1,1\}\\ j\neq k}} \sup_{\mathbf{v}\in\bigcap_j S_j^{e_j}} \frac{1}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}}(\mathbf{r}_k^T\mathbf{v} - \sum_{j\neq k} |\mathbf{r}_j^T\mathbf{v}| \sqrt{\frac{w_j}{w_k}})) \tag{3.77}$$

$$= \max_{\substack{e_j\in\{-1,1\}\\ j\neq k}} \eta(\mathbf{e}) \ , \quad \text{with} \quad \eta(\mathbf{e}) = \sup_{\mathbf{v}\in\bigcap_j S_j^{e_j}} \frac{1}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}}(\mathbf{r}_k^T\mathbf{v} - \sum_{j\neq k} |\mathbf{r}_j^T\mathbf{v}| \sqrt{\frac{w_j}{w_k}}) \ . \tag{3.78}$$

From the definition of $\mathbf{v}_o$ we see that the term in parentheses equals $\mathbf{v}_o^T\mathbf{R}\mathbf{v}$. Now $\mathbf{v} \in \bigcap_j S_j^{e_j} \iff e_j\mathbf{r}_j^T\mathbf{v} \geq 0, \ j \neq k$ and, since η_k^l is invariant to scaling of $\mathbf{v}$, maximization of the given functional over $\mathbb{R}^K$ is equivalent to maximization over the ellipsoid $\mathbf{v}^T\mathbf{R}\mathbf{v} = 1$.

This proves the first part of Proposition 3.11. We now have a sequence of two maximizations to perform, where the second one has the explicit form of an exhaustive search. We turn our attention to the inner maximization in (3.78). We first show that it is possible to replace the feasible set therein by an equivalent convex set, i.e., the asymptotic efficiency is unchanged if we replace

$$\eta(\mathbf{e}) = \sup_{\substack{\mathbf{v}\in\mathbb{R}^K\\ \mathbf{v}^T\mathbf{R}\mathbf{v}=1\\ e_j\mathbf{r}_j^T\mathbf{v}\geq 0\\ j\neq k}} \mathbf{v}_o^T\mathbf{R}\mathbf{v} \qquad \text{by} \qquad \eta(\mathbf{e}) = \sup_{\substack{\mathbf{v}\in\mathbb{R}^K\\ \mathbf{v}^T\mathbf{R}\mathbf{v}\leq 1\\ e_j\mathbf{r}_j^T\mathbf{v}\geq 0\\ j\neq k}} \mathbf{v}_o^T\mathbf{R}\mathbf{v} \ . \tag{3.79}$$

In order to show (3.79), let $\mathbf{y} = \mathbf{R}^{1/2}\mathbf{v}$, $\mathbf{z}_j^T = j^{th}$ row of $\mathbf{R}^{1/2}$. Then it follows that $\mathbf{r}_j^T\mathbf{v} = \mathbf{z}_j^T\mathbf{y}$, $\mathbf{v}_o^T\mathbf{R}^{1/2} = \mathbf{y}_o^T$, $\mathbf{v}^T\mathbf{R}\mathbf{v} = \mathbf{y}^T\mathbf{y} = \|\mathbf{y}\|^2$, and

$$\eta(\mathbf{e}) = \sup_{\substack{\mathbf{y}\in\mathbb{R}^K\\ \|\mathbf{y}\|=1\\ e_j\mathbf{z}_j^T\mathbf{y}\geq 0\\ j\neq k}} \mathbf{y}_o^T\mathbf{y} = \sup_{\substack{\mathbf{y}\in\mathbb{R}^K\\ \|\mathbf{y}\|=1\\ e_j\mathbf{z}_j^T\mathbf{y}\geq 0\\ j\neq k}} \|\mathbf{y}_o\|\|\mathbf{y}\| \cos\ \alpha \tag{3.80}$$

where α is the angle between the vectors $\mathbf{y}_o$ and $\mathbf{y}$. Since the inequality constraints are linear and partition the space into convex cones with vertex at the origin, the optimal angle α is independent of $\|\mathbf{y}\|$. Either the optimal cos α is nonnegative, in which case $\eta(\mathbf{e})$ is maximized for $\|\mathbf{y}\|$ maximal in both versions, or it is negative, in which case $\eta(\mathbf{e}) < 0$. In either case the value of η_k^l, which involves comparison with 0, is unchanged if the maximization is performed over the interior of the ellipsoid, which completes the proof of the claim.

We now have to consider the following problem:

$$\eta(\mathbf{e}) \;=\; \inf_{\substack{\mathbf{v}\in R^K \\ \mathbf{v}^T\mathbf{R}\mathbf{v}-1\le 0 \\ -e_j\mathbf{r}_j^T\mathbf{v}\le 0 \\ j\neq k}} \; -\mathbf{v}_o^T\mathbf{R}\mathbf{v}\ . \tag{3.81}$$

Since this is a minimization problem of a continuous real function on a compact set, it achieves a minimum on the set [Rud, Thm. 4.16]. Since both the cost function and the feasible set are convex, any local minimum is a global minimum. Let $\tilde{\mathbf{v}}$ be a minimizing $\mathbf{v}$, unique up to addition of a vector in the nullspace of $\mathbf{R}$. Since all the functions are differentiable, we can apply the Kuhn-Tucker conditions [(4)] , e.g. [Bro], to get, from condition (1),

$$-\,\mathbf{R}\mathbf{v}_o + \lambda_o\, 2\,\mathbf{R}\tilde{\mathbf{v}} - \sum_{j\neq k} \lambda_j e_j \mathbf{r}_j \;=\; 0$$

hence, since $\mathbf{r} = \mathbf{R}\mathbf{u}_j$,

$$\tilde{\mathbf{v}} \;=\; \frac{1}{2\lambda_o}(\ \mathbf{v}_o + \sum_{j\neq k}\lambda_j e_j \mathbf{u}_j\ ) \tag{3.82}$$

with $\mathbf{u}_j$ the j^{th} unit vector, as defined above. Equations (3.73) and (3.74) result from the Kuhn-Tucker conditions (2) and (3), condition (3.75) expresses the nonnegativity requirement for the λ_i. There is one more constraint to satisfy, which is $\tilde{\mathbf{v}}^T\mathbf{R}\tilde{\mathbf{v}} \;=\; 1$ (we know from (3.79) that the bound is achieved) :

$$1 \;=\; \tilde{\mathbf{v}}^T\mathbf{R}\tilde{\mathbf{v}} \;=\; \frac{1}{2\ \lambda_o}(\mathbf{v}_o^T\mathbf{R}\tilde{\mathbf{v}} + \sum_{j\neq k}\lambda_j e_j \mathbf{r}_j^T\tilde{\mathbf{v}}) \;=\; \frac{\mathbf{v}_o^T\mathbf{R}\tilde{\mathbf{v}}}{2\ \lambda_o} \tag{3.83}$$

We used condition (3.74) to get the last equality. So

$$2\ \lambda_o \;=\; \mathbf{v}_o^T\mathbf{R}\tilde{\mathbf{v}} \;=\; \eta(\mathbf{e}) \tag{3.84}$$

and since

$$\mathbf{v}_o^T\mathbf{R}\tilde{\mathbf{v}} \;=\; \frac{1}{2\ \lambda_o}(\ \mathbf{v}_o^T\mathbf{R}\mathbf{v}_o + \mathbf{v}_o^T\mathbf{R}\ \sum_{j\neq k}\lambda_j e_j \mathbf{u}_j\ )$$

we get

$$2\ \lambda_o \;=\; (\mathbf{v}_o^T\mathbf{R}\mathbf{v}_o + \mathbf{v}_o^T\mathbf{R}\ \sum_{j\neq k}\lambda_j e_j \mathbf{u}_j\ )^{1/2} \tag{3.85}$$

Together with equation (3.82) this completes the proof of Proposition 3.11. ∎

[(4)] Kuhn-Tucker conditions for minimum of differentiable convex function $F(x)$, subject to the set of differentiable convex constraints $f_i(x) \le 0$, $i = 1, \ldots K$: x is a minimum of $F(x)$ if and only if there exist nonnegative λ_i, $i = 1, \ldots K$ such that (1) $\nabla F(x) + \sum \lambda_i \nabla f_i(x) = 0$, (2) $f_i(x) \le 0$, all i, (x feasible), (3) $f_i(x) \neq 0$, $\Rightarrow \lambda_i = 0$.

We would now like to have an explicit procedure to find the maximizing vector $\tilde{\mathbf{v}}$ given implicitly by Proposition 3.11. Next we give an algorithm which solves this problem. The idea is the following: condition (3.74) states that if the maximizing vector $\tilde{\mathbf{v}}$ lies in the intersection of a subset of the delimiting hyperplanes with equations $\mathbf{r}_j^T \tilde{\mathbf{v}} = 0,\ j \in S$, with S the index set of the specific hyperplanes, only the $\lambda_j\ ,\ j \in S$ are possibly nonzero and enter into the expression defining $\tilde{\mathbf{v}}$. Thus we have $|S|$ equations with $|S|$ unknowns, which we can solve to get the λ_i, and then $\tilde{\mathbf{v}}$.

In order to state (and prove the correctness of) an algorithm that finds the optimum linear transformation, the following terminology will be used.

Definition 1 : Let S be an index set $\{j_{1,}j_2, \ldots, j_n\}$, $0 \leq n \leq K-1$, with $j_1,\ldots,j_n \in \{1,..,K\}-\{k\}$, labeled in increasing order. Define

$$D_S(j) \quad = \quad \det \begin{vmatrix} \mathbf{r}_j^T \mathbf{v}_o & R_{jj_1} & . & . & . & R_{jj_n} \\ \mathbf{r}_{j_1}^T \mathbf{v}_o & R_{j_1 j_1} & . & . & . & R_{j_1 j_n} \\ \vdots & \vdots & & \ddots & & \vdots \\ \mathbf{r}_{j_n}^T \mathbf{v}_o & R_{j_n j_1} & . & . & . & R_{j_n j_n} \end{vmatrix} \tag{3.86}$$

Definition 2 : We introduce an indicator for the second Kuhn-Tucker condition:

$$\text{If} \qquad e_j\ D_S(j) > 0 \quad \text{then}\ \ C_S(j) \ = \ \text{yes}\ \ ,\ \ \text{else}\ \ C_S(j) \ = \ \text{no} \tag{3.87}$$

Definition 3 : An n-tuple S of $\{1,..,K\}-\{k\}$ is *matched* if for all i $\in S$:
$C_{S-\{i\}}\ (i) = \text{no}$.

Definition 4 : An n-tuple S contains a basis B if $\{\mathbf{r}_j | j \in B\}$ is a basis for $\{\mathbf{r}_j | j \in S\}$.

Proposition 3.12 : The following algorithm finds a vector $\tilde{\mathbf{v}}$ achieving the maximum in Proposition 3.11 :

[A] Search for the index set with least cardinality $S \subseteq \{1,...,K\}-\{k\}$, for which λ_i, $i \in S$, are possibly nonzero

n := 0

all n-tuples := untried; S_o := matched

while n $\leq$ K-2

while there is still an untried n-tuple containing a matched basis B

select untried matched n-tuple := S_n, contained matched basis := B

if $\forall\, j \notin S_n, j \neq k \quad C_B\ (j) \;=$ yes, **return** S_n, B, **stop**

else S_n := tried

return

n := n+1

return

print "decorrelating detector is optimal", output $\{1,..,K\}-\{k\}$,

stop

[B] Computation of the λ_i :

$i \notin B$: $\lambda_i \;=\; 0$

$i \in B$: λ_i are the solutions of the |B| equations in |B| unknowns $\mathbf{r}_i^T \mathbf{v} \;=\; 0,\; i \in B$, where

$\mathbf{v} \;=\; \mathbf{v}_o + \sum_{i \in B} \lambda_i e_i \mathbf{u}_i$

[C]

$$\tilde{\mathbf{v}} \;=\; \frac{\mathbf{v}_o + \sum_{i \in B} \lambda_i e_i \mathbf{u}_i}{(\mathbf{v}_o^T \mathbf{R} \mathbf{v}_o + \mathbf{v}_o^T \mathbf{R} \sum_{i \in B} \lambda_i e_i \mathbf{u}_i)^{1/2}} \quad .$$

$\diamondsuit$

Comment : Recall that this procedure has to be repeated for all the different e_j in search of the maximal $\eta(\mathbf{e})$ value, until either the efficiency $\eta(\mathbf{e})$ reaches the upper bound given by the optimal detector, or all 2^K possibilities have been exhausted. Prior to running the algorithm, the sufficient conditions given in Propositions 3.13 and 3.14 should be checked.

Proof: Conditions (3.72) and (3.74) are obviously satisfied by construction of $\tilde{\mathbf{v}}$ in [C], and the requirement $\mathbf{r}_j^T\tilde{\mathbf{v}} = 0$ for the possibly nonzero λ_i in [B]. To prove conditions (3.73) and (3.75), consider the system of $|B|$ linear equations in $|B|$ unknowns of [B]. From [A] the set B is matched, and satisfies $C_B\ (j) =$ yes for all $j \neq k, j \notin S_n$. We have to show:

a) $\lambda_i \geq 0$, for all i $= 1, 2, \ldots, K$.

b) $C_B\ (j) =$ yes for all $j \neq k,\ j \notin S_n$ is equivalent to condition (3.73).

a): $\lambda_i = 0,\ i \notin B$, by construction of the index set S_n and [B]. For $i \in B$:

In step [B] we solve $\mathbf{r}_{j_i}{}^T\tilde{\mathbf{v}} = 0$, all i = 1,2,...,$|B|$: (let $|B| =$ n)

$$\begin{aligned} \mathbf{r}_{j_1}^T\mathbf{v}_o + \lambda_{j_1}e_{j_1}R_{j_1j_1} + \cdots + \lambda_{j_n}e_{j_n}R_{j_1j_n} &= 0 \\ \mathbf{r}_{j_2}^T\mathbf{v}_o + \lambda_{j_1}e_{j_1}R_{j_2j_1} + \cdots + \lambda_{j_n}e_{j_n}R_{j_2j_n} &= 0 \\ \cdots \\ \mathbf{r}_{j_n}^T\mathbf{v}_o + \lambda_{j_1}e_{j_1}R_{j_nj_1} + \cdots + \lambda_{j_n}e_{j_n}R_{j_nj_n} &= 0 \, . \end{aligned} \tag{3.88}$$

Denote by D_B the determinant of the coefficient matrix of the $\lambda_{j_i}e_{j_i}$. This coefficient matrix is the reduction of $\mathbf{R}$ to rows and columns indexed by elements of B. Therefore it is nonnegative definite, which implies $D_B \geq 0$. However, because B is a basis, the signal set restricted to indices in B is linearly independent [(5)] , therefore D_B is strictly positive. Then, by Cramer's rule,

$$\lambda_{j_i} = -\, e_{j_i} \frac{D_{B-\{j_i\}}\ (j_i)}{D_B} \, . \tag{3.89}$$

(5)

$$\begin{aligned} \text{B is a basis} &\Longleftrightarrow (\ \forall \alpha \in \boldsymbol{R}^{|B|},\ \exists j \in \{1,\ldots,K\}\ \ s.t.\ \sum_{j\in B} \alpha_i R_{ij} \neq 0\ ) \\ &\Longleftrightarrow (\ \forall \alpha \in \boldsymbol{R}^{|B|},\ \exists j \in \{1,\ldots,K\}\ \ s.t.\ \int_0^T [\sum_{j\in B} \alpha_i \tilde{s}_i(t)]\tilde{s}_j(t)dt \neq 0) \\ &\Longrightarrow \{\ \tilde{s}_i(t), i \in B\ \}\text{linearly independent} \end{aligned}$$

(The converse is also true, as is easily seen).

The numerator is obtained by i row flips and i column flips in order to get j_i into position (1,1). Since the set B is matched, the numerator is nonnegative. As obtained above, the denominator is positive, hence $\lambda_i \geq 0$ for all i $\in B$. This completes the proof of a).

b): Since $\mathbf{r}_j^T\tilde{\mathbf{v}} = 0, j \in B$ and B is a basis of S_n, $\mathbf{r}_j^T\mathbf{v} = 0,\ j \in S_n$. For $j \notin S_n,\ j \neq k$:

With the values obtained for λ compute the "feasibility" expressions (omitting the positive denominator) :

$$\begin{aligned} e_j\mathbf{r}_j^T\tilde{\mathbf{v}} &= e_j\mathbf{r}_j^T\ (\ \mathbf{v}_o + \sum_{i\in B} \lambda_i e_i \mathbf{u}_i\ ) \\ &= \frac{e_j}{D_B}\ (\ D_B\mathbf{r}_j^T\mathbf{v}_o + \sum_{i\in B} -D_{B-\{i\}}(i)\ R_{ji}\ ) \\ &= \frac{1}{D_B}\ e_j\ D_B(j) \qquad > 0\ , \quad \text{since}\ \ C_B(j) = \text{yes},\ \forall j {\notin} S_n, j{\neq}k\ . \end{aligned} \tag{3.90}$$

The last equality is obtained by expanding along the first row of $D_B\ (j)$. This completes the proof of b). By construction the algorithm terminates after at most K-2 steps. ■

In Part [A] of the algorithm notice that $n = 0$ corresponds to a solution in the interior of the feasible cone, with all λ equal to zero, and $\tilde{\mathbf{v}} = \mathbf{v}_o/\sqrt{\mathbf{v}_o^T\mathbf{R}\mathbf{v}_o}$. The corresponding asymptotic efficiency $\eta^2(\mathbf{e}) = \mathbf{v}_o^T\mathbf{R}\mathbf{v}_o$, which is equal to the asymptotic efficiency of the maximum likelihood detector as given by (3.11). Call this case "the optimality case". On the other hand, $n = 1$ corresponds to a solution on exactly one of the delimiting hyperplanes, with exactly one λ nonzero (let it be λ_j). Then λ_j is found in [B] by setting

$$\begin{aligned} \mathbf{r}_j^T\tilde{\mathbf{v}} = 0 \ &\Rightarrow\ \mathbf{r}_j^T\ (\mathbf{v}_o\ +\ \lambda_j e_j \mathbf{u}_j)\ =\ 0 \\ &\Rightarrow\ \lambda_j e_j\ =\ -\mathbf{r}_j^T\mathbf{v}_o\ /\ R_{jj} \end{aligned}$$

Therefore,

$$\tilde{\mathbf{v}} = \frac{1}{\eta(\mathbf{e})}(\mathbf{v}_o - \frac{\mathbf{r}_j^T\mathbf{v}_o}{R_{jj}}\ \mathbf{u}_j) \tag{3.91}$$

and

$$\eta^2(\mathbf{e}) = \mathbf{v}_o^T\mathbf{R}\mathbf{v}_o - \frac{(\mathbf{r}_j^T\mathbf{v}_o)^2}{R_{jj}}\ . \tag{3.92}$$

The asymptotic efficiency achieved in this case is bounded above by the one for $n = 0$, since the second term is nonnegative. If the matrix $\mathbf{R}$ does not have a lot of structure, which is to be expected in practical applications, this is the most probable case. For increasing n the computational effort grows fast, but in most cases the algorithm will terminate for very small n.

We also have an explicit solution for the "terminal case", $n{=}K{-}1$, which corresponds to the decorrelating detector case. Then without loss of generality $\tilde{\mathbf{v}} = \mathbf{r}_k^I/\sqrt{\mathbf{R}_{kk}^I}$, a scaled version of the k^{th} column of any generalized inverse of $\mathbf{R}$, in particular of $\mathbf{R}^+$, and $\eta(\mathbf{e}) = 1/R_{kk}^+$, which is equal to the k^{th} user asymptotic efficiency of the decorrelating detector. This can be showed as follows: In the terminal case $\mathbf{r}_j^T\tilde{\mathbf{v}} = 0$, for all $j \neq k$. Hence

$$\eta(\mathbf{e}) = \max_{\substack{\mathbf{v}\in R \\ \mathbf{vRv}=1 \\ \mathbf{r}_j^T\mathbf{v}=0 \\ j\neq k}} \mathbf{v}_o^T\mathbf{R}\mathbf{v} = \max_{\substack{\mathbf{v}\in R \\ \mathbf{r}_j^T\mathbf{v}=0 \\ j\neq k \\ v_k\mathbf{r}_k^T\mathbf{v}=1}} \mathbf{r}_k^T\mathbf{v} = \max_{\substack{\mathbf{v}\in R \\ \mathbf{Rv}=\frac{1}{v_k}\mathbf{u}_k}} \frac{1}{v_k}. \tag{3.93}$$

If User k is dependent, $\mathbf{r}_j^T\tilde{\mathbf{v}} = 0, \forall j{\neq}k$ implies $\mathbf{r}_k^T\tilde{\mathbf{v}} = 0$, hence the feasible set F in (3.93), $F = \{\mathbf{v}|\mathbf{Rv} = \frac{1}{v_k}\,\mathbf{u}_k\}$, is empty and $\eta(\mathbf{e}) = 0$. Since this was the best choice of $\tilde{\mathbf{v}}$, we can without loss of generality replace $\mathbf{v}$ by the k^{th} row of any generalized inverse, because the resulting asymptotic efficiency cannot become negative. If User k is independent, Lemma 3.2 implies $\mathbf{R}\mathbf{R}^I\mathbf{u}_k = \mathbf{u}_k$, and for all $\mathbf{v}$ in the feasible set,

$$\mathbf{Rv} = \frac{1}{v_k}\mathbf{u}_k \iff \mathbf{R}^I\mathbf{Rv} = \frac{1}{v_k}\mathbf{R}^I\mathbf{u}_k$$

hence, using Lemma 3.2 to obtain the k^{th} elements of the vectors on both sides,

$$v_k = \frac{1}{v_k}R_{kk}^I \Rightarrow v_k = \sqrt{R_{kk}^I} = \sqrt{R_{kk}^+}\,.$$

The last equality was also obtained in Lemma 3.2. If User k is independent the feasible set F in (3.93) is nonempty, (e.g. it contains the set $\frac{1}{v_k}\,\mathbf{r}_k^{I}$, $\mathbf{R}^I \in$ I($\mathbf{R}$), since $\mathbf{R}\mathbf{r}_k^I = \mathbf{u}_k$), from Lemma 3.2, and for all $\mathbf{v} \in$ F, $v_k = \sqrt{R_{kk}^+}$. Hence $\eta(\mathbf{e}) = 1/\sqrt{R_{kk}^+}$, which is the energy independent asymptotic efficiency of the decorrelating detector for independent users. ∎

We showed that there is an energy region for which the best linear detector is equivalent to the optimum multiuser detector ("optimality case"), and an energy region where it is equivalent to its lower bound, the decorrelating detector ("terminal case"). In the following results we give sufficient conditions for these two boundary cases.

Proposition 3.13 : The following are sufficient conditions on the signal energies and crosscorrelations for the best linear detector to achieve optimal k^{th} user asymptotic efficiency:

$$\sqrt{w_k} > \max_{j=1,\dots,K} \left(\frac{1}{|R_{kj}|} \sum_{i \neq k} \sqrt{w_i}\, |R_{ij}| \right) . \tag{3.94}$$

$\diamondsuit$

Proof : In the optimality case, $e_j \mathbf{r}_j^T \mathbf{v}_o > 0$ for all $j \neq k$. If we introduce $e_k = 1$ this has to hold also for $j = k$, otherwise we get negative asymptotic efficiency. Letting $\mathbf{D}$ be the diagonal matrix with i^{th} diagonal element equal to e_i and noting that

$$\mathbf{v}_o = \mathbf{D} \begin{bmatrix} -\sqrt{w_1/w_k} \\ \vdots \\ 1 \\ \vdots \\ -\sqrt{w_K/w_k} \end{bmatrix} \leftarrow \mathrm{k^{th}position} \tag{3.95}$$

an equivalent requirement is that each component of the vector

$$\mathbf{DRv}_o = \begin{bmatrix} R_{11} & e_1 e_2 R_{12} & \dots & e_1 R_{1k} & \dots & e_1 e_K R_{1K} \\ e_1 e_2 R_{21} & R_{22} & \dots & e_2 R_{2k} & \dots & e_2 e_K R_{2K} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ e_K e_1 R_{K1} & e_K e_2 R_{K2} & \dots & e_K R_{Kk} & \dots & R_{KK} \end{bmatrix} \begin{bmatrix} -\sqrt{w_1/w_k} \\ -\sqrt{w_2/w_k} \\ \vdots \\ 1 \\ \vdots \\ -\sqrt{w_K/w_k} \end{bmatrix} \tag{3.96}$$

be positive. We now see that a sufficient condition for this to be satisfied for some $e_1, \dots, e_K$ is

$$|R_{jk}| > \sum_{i \neq k} |R_{ji}| \frac{\sqrt{w_i}}{\sqrt{w_k}} \quad , \quad j = 1, \dots, K. \tag{3.97}$$

The corresponding e_j are $e_j = \operatorname{sgn} R_{jk}$. ■

Note that the above condition can be satisfied by only one user, because then

$$\sqrt{w_k} > \sqrt{w_j}/|R_{kj}| > \sqrt{w_j} \ , \quad \text{for all} \quad j \ . \tag{3.98}$$

Proposition 3.14 : If User k is linearly independent, the following condition is sufficient for the k^{th} row of the decorrelating detector $\mathbf{R}^I \in I(\mathbf{R})$ to be the best k^{th} user linear detector for a given set of signal energies and crosscorrelations :

$$|R^I_{jk}| \leq R^I_{kk} \frac{\sqrt{w_j}}{\sqrt{w_k}} \quad , \quad \text{for all} \quad j \neq k. \tag{3.99}$$

$\diamondsuit$

Proof : We showed that in the terminal case if User k is linearly independent, any $\tilde{\mathbf{v}} = \mathbf{r}_k^I/\sqrt{R_{kk}^I}$ is a maximizing vector for $\mathbf{v}_o^T\mathbf{R}\mathbf{v}$. From (3.84) and Lemma 3.2 ($\mathbf{R}\mathbf{R}^I\mathbf{u}_k = \mathbf{u}_k$)

$$2\lambda_o = \mathbf{v}_o^T\mathbf{R}\tilde{\mathbf{v}} = \mathbf{v}_o^T\frac{\mathbf{R}\mathbf{r}_k^I}{\sqrt{R_{kk}^I}} = \frac{\mathbf{v}_o^T\mathbf{u}_k}{\sqrt{R_{kk}^I}} = \frac{1}{\sqrt{R_{kk}^I}} . \tag{3.100}$$

As a consequence, there are nonnegative Kuhn-Tucker multipliers λ_i, such that, with (3.82),

$$\tilde{\mathbf{v}} = \frac{\mathbf{r}_k^I}{\sqrt{R_{kk}^I}} = \sqrt{R_{kk}^I}\ \ (\mathbf{v}_o + \sum_{j\neq k}\lambda_j e_j \mathbf{u}_j) \tag{3.101}$$

or

$$\frac{1}{R_{kk}^I}\ \mathbf{r}_k^I = \left[(\lambda_1 - \sqrt{\frac{w_1}{w_k}})e_1, \ldots, 1, \ldots, (\lambda_K - \sqrt{\frac{w_K}{w_k}})e_K\right]^T \tag{3.102}$$

so

$$\lambda_j = \frac{\sqrt{w_j}}{\sqrt{w_k}} + e_j R_{jk}^I/R_{kk}^I , \qquad j \neq k . \tag{3.103}$$

Hence (3.99) is sufficient to ensure $\lambda_j \geq 0$ regardless of $\{e_i\ ,\ i \neq k\}$. ∎

Note that in the two-user case, Proposition 3.10 implies that the sufficient conditions found in Propositions 3.13 and 3.14 are also necessary.

Proposition 3.15 : If User k is linearly dependent, then

$$\eta_k^d \ \stackrel{\triangle}{=} \sup_{\mathbf{R}^I \in I(\mathbf{R})} \eta_k(\mathbf{R}^I) = \sup_{\mathbf{T}\in \mathbf{R}^{K\times K}} \eta_k(\mathbf{T}) \ \stackrel{\triangle}{=}\ \eta_k^l \tag{3.104}$$

i.e. for a dependent user the best decorrelating detector has the same asymptotic efficiency as the best linear detector. $\diamondsuit$

Proof : Using (3.33), we can write:

$$\eta_k^d = \max{}^2\ \{0, \sup_{\mathbf{R}^I \in I(\mathbf{R})} \frac{(\mathbf{R}^I\mathbf{R})_{kk} - \sum_{j\neq k} |(\mathbf{R}^I\mathbf{R})_{kj}|\ \sqrt{\frac{w_j}{w_k}}}{\sqrt{(\mathbf{R}^I\mathbf{R}\mathbf{R}^{I^T})_{kk}}}\} . \tag{3.105}$$

Since $\mathbf{R}$ is nonnegative definite of rank r, it can be represented using its orthonormal eigenvector matrix $\mathbf{T}$, and the $r * r$ diagonal matrix $\mathbf{\Lambda}$ of nonzero eigenvalues of $\mathbf{R}$, as

$$\mathbf{R} = \mathbf{T}\begin{bmatrix}\mathbf{\Lambda} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\mathbf{T}^T . \tag{3.106}$$

Then (cf.[Bou]), $\mathbf{R}^I$ is a generalized inverse of $\mathbf{R}$ if and only if, for some matrices $\mathbf{U}$ and $\mathbf{V}$ of appropriate dimensions, it can be represented as:

$$\mathbf{R}^I = \mathbf{T} \begin{bmatrix} \mathbf{\Lambda}^{-1} & \mathbf{V} \\ \mathbf{U} & \mathbf{U\Lambda V} \end{bmatrix} \mathbf{T}^T . \tag{3.107}$$

Hence, using the corresponding partition of $\mathbf{T}$, we can write:

$$(\mathbf{R}^I\mathbf{R})_{kj} = \mathbf{u}_k^T [\mathbf{T}_1 \quad \mathbf{T}_2] \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{U\Lambda} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{T}_1^T \\ \mathbf{T}_2^T \end{bmatrix} \mathbf{u}_j = \mathbf{u}_k^T (\mathbf{T}_1\mathbf{T}_1^T + \mathbf{T}_2\mathbf{U\Lambda T}_1^T)\, \mathbf{u}_j \tag{3.108}$$

$$\begin{aligned}(\mathbf{R}^I\mathbf{R}\mathbf{R}^{I^T})_{kk} &= \mathbf{u}_k^T [\mathbf{T}_1 \quad \mathbf{T}_2] \begin{bmatrix} \mathbf{\Lambda}^{-1} & \mathbf{U}^T \\ \mathbf{U} & \mathbf{U\Lambda U}^T \end{bmatrix} \begin{bmatrix} \mathbf{T}_1^T \\ \mathbf{T}_2^T \end{bmatrix} \mathbf{u}_k \\ &= \mathbf{u}_k^T (\mathbf{T}_1\mathbf{\Lambda}^{-1}\mathbf{T}_1^T + \mathbf{T}_2\mathbf{U}\mathbf{T}_1^T + \mathbf{T}_1\mathbf{U}\mathbf{T}_2^T + \mathbf{T}_2\mathbf{U\Lambda U}^T\mathbf{T}_2^T)\, \mathbf{u}_k \end{aligned} \tag{3.109}$$

and

$$\eta_k^d = \max{}^2 \{0, \sup_{\mathbf{U}\in \mathbb{R}^{(K-r)*r}} \frac{\mathbf{u}_k^T(\mathbf{T}_1 + \mathbf{T}_2\mathbf{U\Lambda})\mathbf{T}_1^T\mathbf{u}_k - \sum_{j\neq k} |\mathbf{u}_k^T(\mathbf{T}_1 + \mathbf{T}_2\mathbf{U\Lambda})\mathbf{T}_1^T\mathbf{u}_j| \sqrt{\frac{w_j}{w_k}}}{\sqrt{\mathbf{u}_k^T(\mathbf{T}_1 + \mathbf{T}_2\mathbf{U\Lambda})\mathbf{\Lambda}^{-1}(\mathbf{T}_1 + \mathbf{T}_2\mathbf{U\Lambda})^T\mathbf{u}_k}} \} . \tag{3.110}$$

Since User k is dependent, $\mathbf{u}_k^T\mathbf{T}_2$, whose components are the k^{th} components of the eigenvectors to eigenvalue zero, is nonzero. (Otherwise for all $\mathbf{x}$ with $\mathbf{Rx} = 0$, x_k would be zero, which implies that the k^{th} user is linearly independent of the other users.) Therefore, and since $\mathbf{\Lambda}$ is invertible, we can make the change of variables

$$\mathbf{x} = (\mathbf{T}_1 + \mathbf{T}_2\mathbf{U\Lambda})^T\mathbf{u}_k \tag{3.111}$$

to get

$$\eta_k^d = \max{}^2 \{ 0, \sup_{\mathbf{x}} \frac{\mathbf{x}^T\mathbf{T}_1^T\mathbf{u}_k - \sum_{j\neq k} |\mathbf{x}^T\mathbf{T}_1^T\mathbf{u}_j| \sqrt{\frac{w_j}{w_k}}}{\sqrt{\mathbf{x}^T\mathbf{\Lambda}^{-1}\mathbf{x}}} \} . \tag{3.112}$$

From here, with the same reasoning as in the proof of Proposition 3.11 for the best linear detector, we obtain

$$\eta_k^d = \max{}^2 \{ 0, \max_{\substack{e_j\in\{-1,1\} \\ j\neq k}} \sup_{\substack{\mathbf{x}\in R^r \\ \mathbf{x}^T\mathbf{\Lambda}^{-1}\mathbf{x} = 1 \\ e_j\mathbf{x}^T\mathbf{T}_1^T\mathbf{u}_j \geq 0 \\ j\neq k}} \mathbf{v}_o^T\mathbf{T}_1\mathbf{x} \} \tag{3.113}$$

where the i^{th} component of $\mathbf{v}_o$ is equal to $(\mathbf{v}_o)_i = \begin{cases} -e_i \sqrt{w_j/w_k} , & i \neq k \\ 1 , & i = k \end{cases}$

$$= \max{}^2 \{0, \max_{\substack{e_j\in\{-1,1\} \\ j\neq k}} \eta_k^d (\mathbf{e})\} \quad \text{with} \quad \eta_k^d (\mathbf{e}) = \sup_{\substack{\mathbf{x}\in R^r \\ \mathbf{x}^T\mathbf{\Lambda}^{-1}\mathbf{x} = 1 \\ e_j\mathbf{x}^T\mathbf{T}_1^T\mathbf{u}_j \geq 0 \\ j\neq k}} \mathbf{v}_o^T\mathbf{T}_1\mathbf{x} \tag{3.114}$$

whereas the k^{th} user asymptotic efficiency of the best linear detector equals, cf. (3.71)

$$\eta_k^l \;=\; \max{}^2 \{0, \max_{\substack{e_j \in \{-1,1\} \\ j \neq k}} \; \eta_k^l\ (\mathbf{e})\} \qquad \text{with} \;\; \eta_k^l\ (\mathbf{e}) \;=\; \sup_{\substack{\mathbf{v} \in R^K \\ \mathbf{v}^T\mathbf{R}\mathbf{v} = 1 \\ e_j r_j^T \mathbf{v} \geq 0 \\ j \neq k}} \quad \mathbf{v}_o^T \mathbf{R} \mathbf{v}\ .$$

Let

$$\mathbf{v}^* \;\in\; \arg\ \eta_k^l(\mathbf{e}) \;\in\; \arg \max_{\substack{\mathbf{v} \in R^K \\ \mathbf{v}^T\mathbf{R}\mathbf{v} = 1 \\ e_j r_j^T \mathbf{v} \geq 0 \\ j \neq k}} \quad \mathbf{v}_o^T \mathbf{R} \mathbf{v}\ . \tag{3.115}$$

We show that $\mathbf{x}^* = \mathbf{\Lambda} \mathbf{T}_1^T \mathbf{v}^*$ is feasible in (3.114), and $\mathbf{v}_o^T \mathbf{T}_1 \mathbf{x}^* = \eta_k^l(\mathbf{e})$:

$$e_j {\mathbf{x}^*}^T \mathbf{T}_1^T \mathbf{u}_j \;=\; e_j {\mathbf{v}^*}^T \mathbf{T}_1 \mathbf{\Lambda} \mathbf{T}_1^T \mathbf{u}_j \;=\; e_j {\mathbf{v}^*}^T \mathbf{R} \mathbf{u}_j \tag{3.116}$$

$$= \; e_j {\mathbf{v}^*}^T \mathbf{r}_j \quad \geq \; 0 \tag{3.117}$$

since $\mathbf{v}^*$ feasible. Also

$${\mathbf{x}^*}^T \mathbf{\Lambda}^{-1} \mathbf{x}^* \;=\; {\mathbf{v}^*}^T \mathbf{T}_1 \mathbf{\Lambda} \mathbf{\Lambda}^{-1} \mathbf{\Lambda} \mathbf{T}_1^T \mathbf{v}^* \;=\; {\mathbf{v}^*}^T \mathbf{R} \mathbf{v}^* \;=\; 1\ . \tag{3.118}$$

Hence $\mathbf{x}^*$ is feasible, and

$$\mathbf{v}_o^T \mathbf{T}_1 \mathbf{x}^* \;=\; \mathbf{v}_o^T \mathbf{T}_1 \mathbf{\Lambda} \mathbf{T}_1^T \mathbf{v}^* \;=\; \mathbf{v}_o^T \mathbf{R} \mathbf{v}^* \;=\; \eta_k^l(\mathbf{e})\ . \tag{3.119}$$

We know that $\eta_k^d \leq \eta_k^l$, since the decorrelating detector belongs to the class of linear detectors. We exhibited for each $\mathbf{e}$ a feasible vector $\mathbf{x}^*$, which satisfied $\mathbf{v}_o^T \mathbf{T}_1 \mathbf{x}^* = \eta_k^l(\mathbf{e})$.

Since from (3.114) $\eta_k^d(\mathbf{e}) \geq \mathbf{v}_o^T \mathbf{T}_1 \mathbf{x}$ for all feasible $\mathbf{x}$, we have, for all $\mathbf{e}$, $\eta_k^d(\mathbf{e}) \geq \eta_k^l(\mathbf{e})$. Hence $\eta_k^d \geq \eta_k^l$, which establishes (3.104). ∎

Since the k^{th} user asymptotic efficiency depends only on the k^{th} row of the applied linear transformation, optimization of $\eta_k(\mathbf{R}^I)$ over the class of generalized inverses for each dependent user k, yields different rows, each belonging to a different generalized inverse. Consequently, the collection of the K optimal rows need not be a generalized inverse.

Finally notice that the near-far resistance of the optimum linear detector is equal to that of the optimum detector, since it is shown in Proposition 3.4 that a particular type of linear detector, i.e. the decorrelating detector, achieves optimum near-far resistance.

3.6 A geometric interpretation

This section gives a geometric interpretation of the problem of constructing a linear decision scheme for K-user synchronous CDMA, in the case where the signal set is linearly independent, using the familiar formulation in terms of a hypothesis testing problem. Every linear detector for User k can be viewed as a hyperplane in K-dimensional Euclidean space which separates the region where the detector decides that User k has transmitted $+1$ from the region where it decides -1. As a result the k^{th} user asymptotic efficiency of a linear detector is determined by the smallest distance among all those hypotheses which coincide in the k^{th} component to the separating hyperplane, which implies that the best linear receiver problem for User k consists in finding the hyperplane with maximal minimum distance to the multiuser hypotheses which coincide in the k^{th} component. Also an explanation of the equality between the near-far resistance of the maximum likelihood detector and the asymptotic efficiency of the decorrelating detector is given, as well as a geometric derivation of the best linear detector in the two-user case.

It is advantageous for the geometric intuition to view the detection problem in a domain where the noise is spherically symmetric. Therefore set $\mathbf{z} \stackrel{\triangle}{=} \mathbf{R}^{-1/2}\mathbf{y}$, where $\mathbf{R}^{1/2}$ is the unique positive definite square root of $\mathbf{R}$. Then

$$\mathbf{z} = \mathbf{R}^{1/2}\mathbf{W}\mathbf{b} + \mathbf{n}' , \qquad \mathbf{n}' \sim N(0, \sigma^2\mathbf{I}) . \tag{3.120}$$

and in this domain the noise is spherically symmetric. The k^{th} user receiver has to decide on the basis of $\mathbf{z}$ whether $b_k = 1$ or $b_k = -1$. Letting

$$\begin{aligned} S_1 &\stackrel{\triangle}{=} \{\mathbf{R}^{1/2}\mathbf{W}\mathbf{b} \mid \mathbf{b} \in \{-1,1\}^K,\ b_k = 1\} \\ S_{-1} &\stackrel{\triangle}{=} \{\mathbf{R}^{1/2}\mathbf{W}\mathbf{b} \mid \mathbf{b} \in \{-1,1\}^K,\ b_k = -1\} , \end{aligned} \tag{3.121}$$

this is a two-hypotheses problem with

$$\begin{aligned} H^0 &: \mathbf{z} \in S_1 + \mathbf{n}' \quad \to \text{decide for } b_k = 1 \\ H^1 &: \mathbf{z} \in S_{-1} + \mathbf{n}' \quad \to \text{decide for } b_k = -1 . \end{aligned} \tag{3.122}$$

Any decision rule corresponds to a partition of the signal space span (S_1, S_{-1}) into a decision region for H_0 and one for H_1. The maximum likelihood detector decides -since the noise is spherically symmetric- for the hypothesis which is closest in Euclidean distance to $\mathbf{z}$. In the two-user case Figure

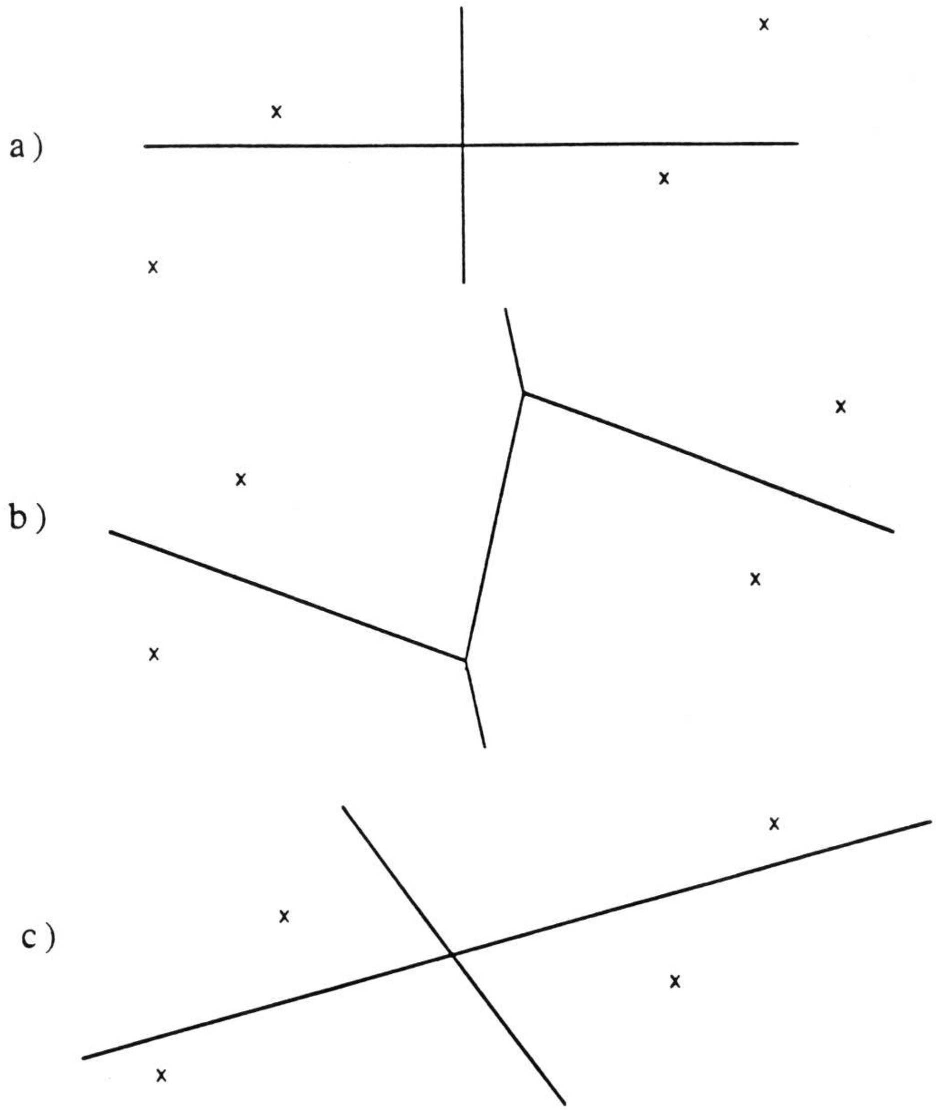


Fig. 5. Decision regions of a) the conventional, b) the maximum-likelihood and c) a linear detector, for spherically symmetric noise.

5 illustrates the decision regions for the conventional, the maximum-likelihood and a general linear detector.

On the other hand a linear detector decides according to

$$\hat{b}_k = \text{sgn}\, \mathbf{v}^T \mathbf{y} = \text{sgn}\, < \mathbf{u}, \mathbf{z} > \tag{3.123}$$

where for the second equality we have set $\mathbf{u} = \mathbf{R}^{1/2}\mathbf{v}/\|\mathbf{R}^{1/2}\mathbf{v}\|$. The normalizing factor is introduced for convenience, and does not affect the sign decision. In other words, since $\mathbf{u}$ is a unit length vector,

$$\hat{b}_k \;=\; \text{sgn}\, P_{\mathbf{u}}\, \mathbf{z} \tag{3.124}$$

where $P_{\mathbf{u}}$ denotes the projection operator onto $\mathbf{u}$. Therefore the decision boundary for the k^{th} user is the hyperplane $L_{\mathbf{u}}$ perpendicular on $\mathbf{u}$, since this is where $P_{\mathbf{u}}$ changes sign, i.e.,

$$L_{\mathbf{u}} \;=\; \{\mathbf{x} \mid P_{\mathbf{u}}\, \mathbf{x} \;=\; 0\}\,. \tag{3.125}$$

Hence a linear detector for the k^{th} user is specified equivalently by $\mathbf{u}$ or by the hyperplane $L_{\mathbf{u}}$ perpendicular on $\mathbf{u}$. (We know that it is sufficient to consider hyperplanes, i.e., we do not need to introduce a possible offset from the origin, because of the symmetry of the hypotheses with respect to a sign change in all coordinates). In order to have zero error probability in the absence of noise it is apparent that the hyperplane corresponding to a reasonable linear detector has to separate S_1 and S_{-1}. We will refer to such a hyperplane as a *separating hyperplane.* So this is how the decision regions depends on the chosen $\mathbf{u}$. On the other hand the asymptotic efficiency depends on $\mathbf{u}$ as follows. From (3.56)

$$\eta_k(\mathbf{v}) \;=\; \max{}^2\, \{0,\; \frac{\mathbf{r}_k^T\mathbf{v} - \sum_{j\neq k} \mid \mathbf{r}_j^T\mathbf{v} \mid \sqrt{\frac{w_j}{w_k}}}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}}\;\}$$

and recall that the absolute values resulted from the worst-case set of $b_j, j \neq k$. Therefore

$$\begin{aligned}
\eta_k(\mathbf{v}) \;&=\; \frac{1}{w_k}\max{}^2\,\{0,\; \min_{\substack{\mathbf{b}\in\{-1,1\}^K \\ b_k=1}} \frac{<\mathbf{v}, \mathbf{RWb}>}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}}\}\\
&=\; \frac{1}{w_k}\max{}^2\,\{0,\; \min_{\substack{\mathbf{b}\in\{-1,1\}^K \\ b_k=1}} <\frac{\mathbf{R}^{1/2}\mathbf{v}}{\|\mathbf{R}^{1/2}\mathbf{v}\|},\; \mathbf{R}^{1/2}\mathbf{Wb}>\}\\
&=\; \frac{1}{w_k}\max{}^2\,\{0,\; \min_{\mathbf{s}\in S_1} <\mathbf{u}, \mathbf{s}>\}\\
&=\; \frac{1}{w_k}\max{}^2\,\{0,\; \min_{\mathbf{s}\in S_1} P_{\mathbf{u}}\,\mathbf{s}>\}\,.
\end{aligned} \tag{3.126}$$

The first equality uses the nonnegative definiteness of $\mathbf{R}$, while the last two equalities make use of the definition of $\mathbf{u}$ and S_1, respectively the fact that $\mathbf{u}$ has unit length. Hence for a linear detector for the k^{th} user, specified by the separating hyperplane $L_{\mathbf{u}}$, the asymptotic efficiency is given by the minimum distance of a hypothesis to $L_{\mathbf{u}}$. The problem of selecting the best linear detector is therefore that of selecting a separating hyperplane with maximal minimum distance from the hypotheses (since S_1 and S_{-1} are symmetric with respect to the origin it suffices to maximize the minimum distance to S_1). That this is a desirable goal is intuitively clear, because the decision on the transmitted bit is based entirely on whether the received vector $\mathbf{z}$ falls on one side of the hyperplane $L_{\mathbf{u}}$ or on the other. Since in the absence of noise $\mathbf{z}$ has to be one of the hypotheses in S_1 or S_{-1}, this means that the white noise resistance of this binary decision is determined by the hypothesis which is closest to the decision boundary (i.e., the term corresponding to this hypothesis dominates the error probability in the high SNR region).

Explanation of Proposition 3.4 for a linearly independent signal set

Since the signal set is linearly independent, $\mathbf{R}$ is nonsingular and $\eta_k^d = \eta_k(\mathbf{R}^{-1})$ is energy-independent for all users. We want to explain the equality of $\bar{\eta}_k$ and η_k^d. First consider the two-user case, and assume we want to decide on the transmitted bit of User 1.

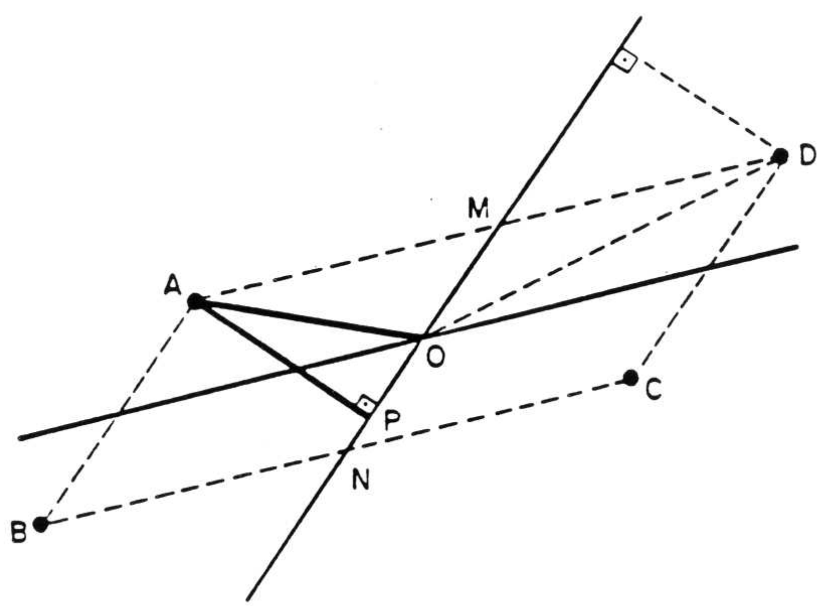


Fig. 6. Geometric illustration of Proposition 3.4.

Figure 6 shows the four hypotheses, where $S_{-1} = \{A, B\}$ and $S_1 = \{C, D\}$. Since in this domain the matched filter output noise is spherically symmetric and Gaussian, the decision regions of the maximum likelihood detector, determined by the minimum Euclidean distance rule, are given by the perpendicular bisectors of the segments between the different hypotheses, and the k^{th} user asymptotic efficiency corresponds to the square of half the minimum distance between distinct hypotheses differing in the k^{th} bit $(Q(d_{\min}/2\sigma) = Q(\sqrt{\eta w_k}/\sigma))$. Thus, in Figure 6, up to the factor w_k which will be ignored because it multiplies the k^{th} user asymptotic efficiency for linear detectors too, $\sqrt{\eta_1}$ is the length of the shortest of the segments AM,AO and BO.

The decision regions of the decorrelating detector for User 1 are determined by a straight line through the origin, such that application of $\mathbf{R}^{-1/2}$ maps it to the y-axis (since a sign decision is then taken). This means that the separating line passes through the points $\pm\mathbf{R}^{1/2}\mathbf{u}_2$, with $\mathbf{u}_2$ is the unit vector in y direction, i.e., $[0\ 1]^T$. These points are at the centers of the sides AD and BC of the parallelogram formed by the hypotheses, because the unit vector is collinear to half the sum of adjacent hypotheses differing in the first bit. Hence the decorrelating detector decision boundary is parallel to the parallelogram sides AB and CD, which follows because it passes through the centers of the sides AD and BC. As a consequence, *all the hypotheses have equal distance to the decision boundary of the decorrelating detector.* This is intuitively clear, because the k^{th} bit error probability of the decorrelating detector reduces to a single Q-function. Now the first user asymptotic efficiency of the decorrelating detector is equal to the square of the distance of any hypothesis to the decision boundary, e.g. in Figure 6 $\sqrt{\eta_1^d}$ is the length of AP.

The result of Proposition 3.4 can now be interpreted as follows: since η appears as the hypotenuse and η^d as the leg of a right angled triangle, η is lower bounded by the energy independent η^d. However, since the triangle angles vary with increasing energy of the interfering user, there is a particular energy ratio for which the triangle degenerates into a line segment. This is the point when η reaches its minimum, $\bar{\eta}$, which is geometrically identical with η^d. For the parallelogram formed by the hypotheses this is the case when a diagonal is perpendicular to a side (eg. AO perpendicular to CD).

For more than two users the explanation is analogous. The set of 2^K hypotheses will be the corners of a parallelepiped in $\mathbb{R}^K$, since the nonsingular linear map $\mathbf{R}$ maps hyperplanes into hyperplanes, hence the rectangular parallelepiped with corners $\mathbf{Wb}$, $\mathbf{b} \in \{-1, 1\}^K$ into the one we are considering. Call the sides of the parallelepiped corresponding to S_1 and S_{-1} *significant* sides.

Any linear detector corresponds to a hyperplane separating the significant sides. Since the probability of error of the decorrelating detector is for each user a single Q-function, the corresponding separating hyperplane is equidistant to the hypotheses, i.e., is parallel to the significant sides, and the asymptotic efficiency is determined by their distance (see Figure 7).

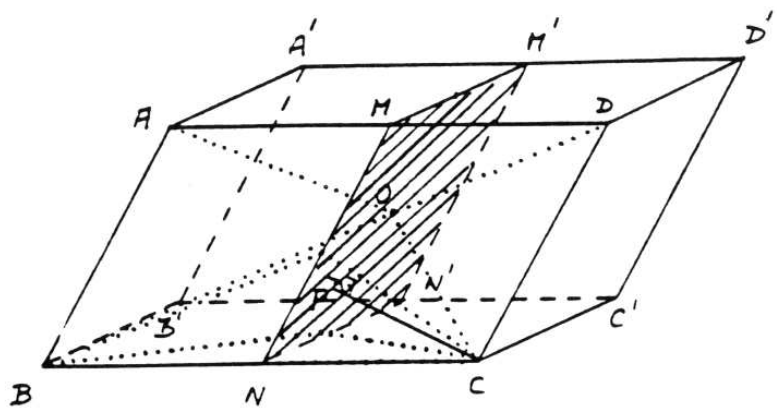


Fig. 7. Same as Figure 6, in the case of three users.

As the energies vary, there will be particular energy constellations (because the deformations of the parallelepiped are continuous) where the vector from the origin to some hypothesis, say $\mathbf{R}^{1/2}\mathbf{W}\mathbf{b}^*$, is perpendicular on the corresponding significant side, i.e., also on the decorrelating hyperplane. This is the case when equality occurs between the asymptotic efficiencies of the decorrelating and the maximum likelihood detector. To see this note that in this constellation the asymptotic efficiency of the maximum likelihood detector must simultaneously be higher than that of the decorrelating detector (since being the optimum multiuser detector it constitutes an upper bound on the performance of any detector) and smaller, since from (3.10) it equals the minimum of $\|\mathbf{R}^{1/2}\mathbf{W}\epsilon\|^2$ for $\epsilon \in \{-1,0,1\}^K$, $\epsilon_k = 1$, hence is upper bounded by $\|\mathbf{R}^{1/2}\mathbf{W}\mathbf{b}^*\|^2$, which in this case is the asymptotic efficiency of the decorrelating detector.

Again, consider User 1. The hypotheses in S_1 are $h_{++} \triangleq \mathbf{R}^{1/2}\mathbf{W}[1 \;\; 1]^T$ and $h_{+-} \triangleq \mathbf{R}^{1/2}\mathbf{W}[1 \;\; -1]^T$. To find the best linear detector we have seen that we have to find the separating line L_* through the origin which has maximal minimum distance to h_{++} and h_{+-}. If we only consider h_{++}, the line having maximal distance from it is the perpendicular through the origin on the vector from the origin to h_{++}. Call this line L_{++}. Similarly for h_{+-}, which defines L_{+-}. If we start with an arbitrary separating line, it is easy to see that a rotation of this line in the direction of L_{++} increases its distance to h_{++}, while a rotation in the direction of L_{+-} increases its distance to h_{+-}. Call the decision boundary corresponding to the decorrelating detector L_d. We have seen that L_d is equidistant from h_{++} and h_{+-}. Now given that we have three lines through the origin, L_{++}, L_{+-} and L_d, one can have two inherently different situations. In one case L_d lies between L_{++} and L_{+-}, in the second both L_{++} and L_{+-} are on the same side of L_d. The two cases are illustrated in Figure 8 a) and b).

The difference between them is that, while in the first case attempting to rotate L_* away from L_d will decrease its distance from either h_{++} or h_{+-} (though increasing the other), so that the optimal solution is to have both distances equal, i.e., $L_* = L_d$, in the second case rotating L_* in the coinciding direction of L_{++} and L_{+-} will increase the distance to both hypotheses, as long as L_{++} and L_{+-} are still on the same side of the line we are rotating. Clearly this can be done until the first of L_{++}, L_{+-} is encountered and this will be the optimal solution L_*. Note that this is the linear continuation of the (piecewise linear) decision boundary of the maximum likelihood detector, which is given by the perpendicular bisector of the segments between hypotheses differing in the 1^{st} bit. This can be seen also from Proposition 3.9, where if $w_2/w_1 \leq \rho^2$ the optimum linear detector is collinear to $[1 \;\; -\text{sgn}\rho\sqrt{w_2/w_1}]$, i.e., is collinear to one of the hypotheses $\mathbf{Wb}$, i.e., is still collinear to one of the hypotheses after application of the map $\mathbf{R}^{1/2}$. This means that the corresponding delimiting line is perpendicular on this hypothesis, i.e., is a relevant boundary for the minimum Euclidean distance detector, which is just what the maximum likelihood detector is in this case.

This explains why there are two different cases for the optimum linear detector in the two user case, why the decorrelating detector is one of them, and also why in one of the cases optimum asymptotic efficiency is achieved.

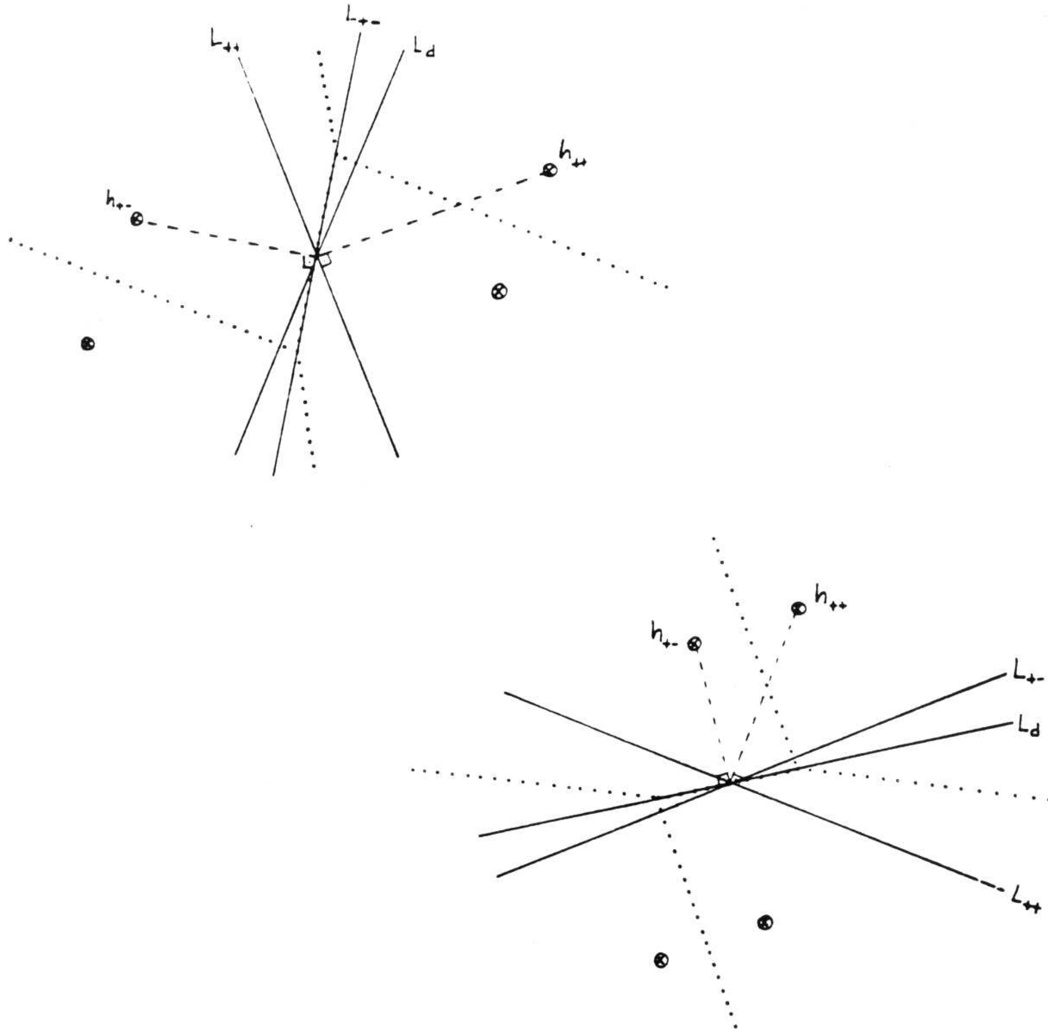


Fig. 8. Geometric illustration of Proposition 3.9.

3.7 Computation of the decorrelating detector

The decorrelating detector for a synchronous communication system with a linearly independent signal set is the inverse of a positive definite symmetric matrix. The straightforward computation via the Cholesky factorization or Gauss elimination requires on the order of K^3 multiplications and the same number of additions. Thereafter, as long as the active user configuration does not change, the demodulation of $\mathbf{b}$ requires a TCB of K multiplications and additions. In the following an iterative scheme is given, which for each $\mathbf{z}$ converges to the solution $\mathbf{x}^{\infty} = \mathbf{R}^{-1}\,\mathbf{z}$ upon starting with an arbitrary initial vector $\mathbf{x}^{o}$.

Proposition 3.16: The iterative scheme

$$\mathbf{x}^{n+1} = (\mathbf{I} - \beta\mathbf{R})\mathbf{x}^{n} + \beta\mathbf{z} \tag{3.127}$$

converges to

$$\mathbf{x}^{\infty} = \mathbf{R}^{-1}\mathbf{z} \tag{3.128}$$

for any initial vector $\mathbf{x}^{o}$ as long as $0 < \beta < 2/\lambda_{\text{max}}$, where λ_{max} is the largest eigenvalue of $\mathbf{R}$. Moreover the fastest convergence is achieved for

$$\beta^{\text{opt}} = \frac{2}{\lambda_{\text{min}} + \lambda_{\text{max}}} . \tag{3.129}$$

$\diamondsuit$

Proof and Discussion: It is easy to prove convergence by defining the error vector $\mathbf{e}^{n} = \mathbf{x}^{n} - \mathbf{R}^{-1}\mathbf{z}$ and deriving from (3.127) that

$$\mathbf{e}^{n} = (\mathbf{I} - \beta\mathbf{R})^{n}\,\mathbf{e}^{o} \tag{3.130}$$

where from convergence follows if the spectral radius of $\mathbf{I} - \beta\mathbf{R}$ is less than unity, which results in the given condition on β. Alternatively, (3.128) can be interpreted as a gradient solution to the problem

$$\min_{\mathbf{x}\in\boldsymbol{R}} \; (\mathbf{x}^{T}\mathbf{R}\mathbf{x} - 2\mathbf{x}^{T}\mathbf{z}) \tag{3.131}$$

whose exact solution is $\mathbf{x}^{\infty} = \mathbf{R}^{-1}\mathbf{z}$ and for which the gradient solution is given by

$$\mathbf{x}^{n+1} = \mathbf{x}^{n} - \frac{\beta}{2}\,\nabla\,|_{\mathbf{x}=\mathbf{x}^{n}}\,(\mathbf{x}^{T}\mathbf{R}\mathbf{x} - 2\mathbf{x}^{T}\mathbf{z})\,. \tag{3.132}$$

Iterative gradient solutions of this kind have been well studied and used in practice , among others in adaptive filtering applications (e.g. [Hon], [Ben]). Denoting by ρ the spectral radius of the matrix $\mathbf{I} - \beta\mathbf{R}$, the Euclidean norm of the error vector satisfies

$$\|\mathbf{e}^{n}\| \leq \|\mathbf{R}\|_{2} = \rho^{n}\,\|\mathbf{e}^{o}\|\,, \tag{3.133}$$

where the equality follows since, $\mathbf{R}$ being Hermitian, its spectral norm equals its spectral radius.[6] It can be shown (e.g. [Ben, p. 375]) that the smallest spectral radius is achieved by the choice of β given in the proposition, and equals

$$\rho_{\text{opt}} = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \tag{3.134}$$

which is a function of the eigenvalue spread of the matrix $\mathbf{R}$. As always, a small eigenvalue spread is desirable.

Note that if $\mathbf{z} = \mathbf{u}_k$, i.e., the k^{th} unit vector, $\mathbf{x}^\infty$ is the k^{th} column of $\mathbf{R}^{-1}$, which is sufficient for decentralized detection of the information sent by User k. The iterations would be run before the beginning of the detection process in order to find the decorrelating detector. K different iterations have to be run beforehand, in order to get the matrix $\mathbf{R}^{-1}$.

Given an initial error norm and a value for a sufficient final error, equation (3.133) can be used to upper bound the required number of iterations, for purposes of comparison with the method of direct matrix inversion. Choosing $\mathbf{x}^o = 0$, the initial error vector has a norm upper bounded by

$$\|\mathbf{e}^n\| \;\le\; \rho^n \,\|\mathbf{R}^{-1}\mathbf{u}_k\| \;\le\; \rho^n \,\|\mathbf{R}^{-1}\| \;=\; \rho^n \,\frac{1}{\lambda_{\min}} \;\le\; \rho_{\text{opt}}^n \,\frac{1}{\lambda_{\min}} \;.$$

Therefore an upper bound on the necessary number of iterations for obtaining the k^{th} column of $\mathbf{R}^{-1}$ with precision ϵ is

$$n^* \;\le\; \left\lceil \frac{\log\,(\epsilon\,\lambda_{\min})}{\log\,(\rho_{\text{opt}})} \right\rceil . \tag{3.135}$$

Each iteration involves K multiplications and additions, therefore the computational complexity is $K * n^*$ for each column of $\mathbf{R}^{-1}$. Whether it is preferable to use the direct inversion scheme or the iteration algorithm, depends on which of the two complexities is smaller, K^3 or $K * P * n^*$ (for demodulation of P users). For example, for $\lambda_{\max} = 3$, $\lambda_{\min} = 0.5$, $\epsilon = 10^{-12}$ and demodulation of one user, the necessary number of iterations is, from (3.135), at most 85, which means that the iterative scheme is more efficient than the direct inversion scheme for $K > 9$.

---

(6)

$$\begin{aligned}
\|\mathbf{A}\|_2 &\stackrel{\triangle}{=} \max\,\{\sqrt{\lambda}:\; \lambda \text{ is an eigenvalue of } \mathbf{A}^*\mathbf{A}\} \\
&\stackrel{(\mathbf{A}=\mathbf{A}^*)}{=} \max\,\{\sqrt{\lambda}:\; \lambda \text{ is an eigenvalue of } \mathbf{A}^2\} \\
&= \max\,\{\sqrt{\xi^2}:\; \xi \text{ is an eigenvalue of } \mathbf{A}\} \\
&= \max\,\{|\lambda|:\; \lambda \text{ is an eigenvalue of } \mathbf{A}\} \stackrel{\triangle}{=} \rho(\mathbf{A})
\end{aligned}$$

3.8 On a simple decision-feedback detection scheme for synchronous CDMA

In synchronous CDMA the receiver has to make an estimate of the transmitted information vector $\mathbf{b}$ of the K synchronous users in one symbol interval, or on the k^{th} component thereof, based in the sufficient statistic $\mathbf{y}$, which depends on $\mathbf{b}$ according to

$$\mathbf{y} \; = \; \mathbf{RWb} \, + \, \mathbf{n}' \, , \qquad \mathbf{n}' \, \sim \, N \, (0, \sigma^2 \, \mathbf{R}) \; .$$

In the following it is assumed for simplicity that $\mathbf{R}$ is invertible (otherwise generalized inverses will be used, along the lines of the previous sections). Then the sufficient statistic is equivalently

$$\begin{aligned} \mathbf{z} \; &= \; \mathbf{R}^{-1} \, \mathbf{y} \\ &= \; \mathbf{W} \, \mathbf{b} \, + \, \mathbf{n} \, , \qquad \mathbf{n} \, \sim \, N \, (0, \sigma^2 \, \mathbf{R}^{-1}) \; . \end{aligned} \tag{3.136}$$

If the noise vector $\mathbf{n}$ were uncorrelated, a sign decision at this point, which is what the decorrelating detector does, would be optimal. Also, if the variance of the noise samples were σ^2, i.e., $R^{-1}_{kk} = 1$, the probability of error of a sign decision for User k would be the same as the probability of error of a single user channel with the same noise level. On the other hand when $R^{-1}_{kk} > 1$ (from (3.27) $R^{-1}_{kk} \geq 1$ is obtained by setting $\mathbf{T} = \mathbf{R}^{-1}$), the noise sample variance, and hence the error probability of a sign decision is higher. Therefore a natural question to ask, is how to use the fact that the noise samples in (3.136) are correlated, to reduce the variance of the noise sample affecting User k. One can reason as follows. The k^{th} noise sample has an a priori expected value of zero. However, its expected value, given the symbols transmitted by the other users, is nonzero in general, because of the correlation of the noise components. Subtracting this expected value from z_k reduces the variance of the additive noise component contained therein. It turns out that the variance is reduced to unity, therefore decreasing the probability of error to that of the single-user channel, as expected, since the bits transmitted by the other users have been given. Of course the symbols transmitted by the other users are not available, but if the detector used has a small error probability, then there exist estimates of these symbols which are correct most of the time. Therefore the above scheme can be expected to still work well if decisions on the transmitted symbols are used, instead of the real values, as long as these decisions are reliable enough. The aim of this section is to investigate under what conditions feedback in the aforementioned form improves performance, and by how much.

3.8.1 Detector with full feedback of decisions

Given that $z_i = \sqrt{w_i}\, b_i + n_i$, upon decision $\hat{b}_i$ for the transmitted bit, an estimate of the noise is obtained as a byproduct, namely

$$\hat{n}_i = z_i - \sqrt{w_i}\, \hat{b}_i \, . \tag{3.137}$$

Such an estimate is obtained for all the components of the noise vector. Note that knowledge of the received energies is required, in contrast to the decorrelating detector.

Denote by $(n_k, \tilde{\mathbf{n}})$ and $(b_k, \tilde{\mathbf{b}})$ the partitions of the noise and information vector, respectively, according to User k (the tilde indicates that the entry corresponding to User k has been deleted), and let $\hat{b}_k^{(2)}$ denote the second stage decision on the bit transmitted by User k, obtained by using the first stage decisions $\hat{b}_j,\ j \neq k$, of the other users. The question of interest is to investigate the performance of the decision scheme

$$\begin{aligned} \hat{b}_k^{(2)} &= \operatorname{sgn}\left(z_k - E\,[n_k \mid \tilde{\mathbf{b}}\,]\big|_{\tilde{\mathbf{b}}=\hat{\tilde{\mathbf{b}}}} \right) \\ &= \operatorname{sgn}\left(z_k - E\,[n_k \mid \tilde{\mathbf{n}}\,]\big|_{\tilde{\mathbf{n}}=\hat{\tilde{\mathbf{n}}}} \right) . \end{aligned} \tag{3.138}$$

$E[n_k \mid \tilde{\mathbf{n}}]$ is the minimum mean-square estimate (MMSE) of n_k given $\tilde{\mathbf{n}}$, and since the n_i are jointly Gaussian, this MMSE is linear, therefore it equals the projection of n_k on the space spanned by the components of $\tilde{\mathbf{n}}$. Hence

$$E\,[n_k \mid \tilde{\mathbf{n}}] = \sum_{i \neq k} c_i n_i \, , \tag{3.139}$$

and by the Projection theorem

$$E\,[(\, n_k - \sum_{j \neq k} c_i n_i \,)\, n_j] = 0 \, , \qquad \forall j \neq k \, , \tag{3.140}$$

i.e., since the covariance matrix of the n_i is $\mathbf{R}^{-1}$,

$$R_{kj}^{-1} - \sum_{i \neq k} c_i R_{ij}^{-1} = 0 \, , \qquad \forall j \neq k \, . \tag{3.141}$$

Equivalently, the vector $[-c_1, \ldots, -c_{k-1},\ 1\ , -c_{k+1}, \ldots, -c_K]^T$ is perpendicular on the j^{th} column of $\mathbf{R}^{-1}$, for all $j \neq k$, i.e., is the k^{th} row of $\mathbf{R}$. Therefore, letting $[1, \mathbf{m}_k, \mathbf{R}_k]$ denote the partition of the matrix $\mathbf{R}$ with respect to the k^{th} row and column, we have

$$E\,[n_k \mid \tilde{\mathbf{n}}] = -\mathbf{m}_k^T\, \tilde{\mathbf{n}} \tag{3.142}$$

and the decision scheme of (3.138) becomes

$$\hat{b}_k^{(2)} = \text{sgn} (z_k + \mathbf{m}_k^T \hat{\mathbf{n}}) . \tag{3.143}$$

Using (3.137)

$$\hat{b}_k^{(2)} = \text{sgn} \left(z_k + \mathbf{m}_k^T (\tilde{\mathbf{z}} - \tilde{\mathbf{W}} \hat{\tilde{\mathbf{b}}}) \right) = \text{sgn} \left(\mathbf{r}_k^T \mathbf{z} - \mathbf{m}_k^T \tilde{\mathbf{W}} \hat{\tilde{\mathbf{b}}} \right) \tag{3.144}$$

is the obtained feedback detector, where $\hat{\tilde{\mathbf{b}}}$ are the initial decisions of the detector in the forward branch. We can rewrite the last expression as

$$\hat{b}_k^{(2)} = \text{sgn} (y_k - \mathbf{m}_k^T \tilde{\mathbf{W}} \hat{\tilde{\mathbf{b}}}) . \tag{3.145}$$

Interestingly, the detector we obtained subtracts the available estimate of the cross-interference, which is intuitively pleasing, although not obvious from the starting point of reducing the Gaussian noise variance. Subtracting an estimate of the multi-user interference has been independently proposed in [Var 88a], using the conventional detector to obtain initial decisions. However a weakness of that work is the absence of the desirable property of near-far resistance of the original estimates. This criticism has been incorporated recently in [Var 88b], where the decorrelating detector is used in the first stage. However, the error probability analysis in [Var 88b] is numerical, in contrast to the asymptotic efficiency results obtained here. Also the approach via noise variance reduction rather than interference cancellation is easily generalized to a partial feedback scheme which is near-far resistant, a property which partial interference removal obviously lacks. For all users together the decision scheme is

$$\hat{\mathbf{b}}^{(2)} = \text{sgn} (\mathbf{y} - (\mathbf{R} - \mathbf{I})\mathbf{W}\hat{\mathbf{b}}) . \tag{3.146}$$

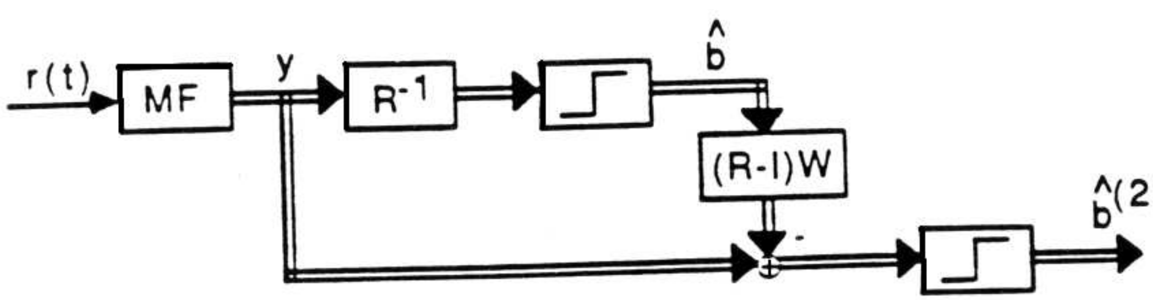


Fig. 9. Structure of the decision-feedback detector of (3.146).

Figure 9 shows the resulting detector. Since equation (3.145) can be rewritten as

$$\hat{b}_k^{(2)} = \text{sgn}\left(\sqrt{w_k} b_k + \mathbf{m}_k^T \tilde{\mathbf{W}} (\tilde{\mathbf{b}} - \hat{\tilde{\mathbf{b}}}) + n'_k\right), \qquad \mathbf{n}' \sim N(0, \sigma^2 \mathbf{R}) \tag{3.147}$$

it becomes apparent that if the first-stage estimates $\hat{\mathbf{b}}$ are correct, the probability of error of the above scheme is the same as that of the single-user channel, since the n'_k has variance σ^2. Obviously, if the second-stage decisions $\hat{\mathbf{b}}^{(2)}$ are better than the original estimates, the procedure can be repeated using the new decisions instead of $\hat{\mathbf{b}}$, and the iteration process can be repeated until the decisions of two consecutive iterations are the same. In this light it is to be expected that the performance of the above feedback scheme depends critically on the robustness of the first-stage estimates, and has the potential of significant performance improvement over the decorrelating detector, if these estimates are good enough. This potential is analyzed in the sequel. The question, in a more general form than will be answered, is the following.

Problem statement

Given

$$\mathbf{y} = \mathbf{RWb} + \mathbf{n}, \qquad \mathbf{n} \sim N(0, \sigma^2 \mathbf{R}) \tag{3.148}$$

and an estimate $\hat{\mathbf{b}}$, with error probability $P_{\hat{\mathbf{b}},k}(e)$. When is the performance of User k improved by doing

$$\hat{b}_k^{(2)} = \text{sgn}\left(y_k - \mathbf{m}_k^T \tilde{\mathbf{W}} \hat{\tilde{\mathbf{b}}}\right) \tag{3.149}$$

Our measure of performance is the k^{th} user asymptotic efficiency.

It would be interesting to be able to answer this question for an arbitrary first stage detector, though probably intractable. The results when using the decorrelating detector for the first stage are summarized in the following propositions. The reason which speaks for using the decorrelating detector to obtain initial decisions is that it is the only near-far resistant linear detector, which moreover achieves the near-far resistance of optimum multi-user detection, thereby ensuring a nondegrading performance level regardless of the received energies.

Proposition 3.17 : The k^{th} user second stage asymptotic efficiency η_k^f of the iterated decorrelating detector with full feedback, which decides for

$$\hat{\mathbf{b}}^{(2)} = \text{sgn}\,(\mathbf{y} - (\mathbf{R} - \mathbf{I})\mathbf{W}\hat{\mathbf{b}}) \tag{3.150}$$

where $\hat{\mathbf{b}}$ are the decisions of the decorrelating detector, satisfies

$$\eta_k^f \geq \min_{i \neq k} f\,(a_i, c_i)\;, \tag{3.151}$$

where

$$a_i = \frac{\sqrt{w_i}}{\sqrt{w_k}} \sqrt{\eta_i^d}\;, \tag{3.152}$$

$$c_i = 2 \sum_{\substack{n \neq k \\ n s.t. a_n \leq a_i}} |R_{kn}| \frac{\sqrt{w_n}}{\sqrt{w_k}}\;, \tag{3.153}$$

and

$$f\,(a,c) = \begin{cases} 1\,, & \text{for } c \geq 1,\ a \geq 1 \\ & \text{and } c < 1,\ (1-c)^2 + a^2 \geq 1 \\ a^2\,, & \text{for } c \geq 1,\ a < 1 \\ (1-c)^2 + a^2\,, & \text{else}\ . \end{cases} \tag{3.154}$$

$\diamond$

The function $f(a,c)$ is shown in Figure 10.

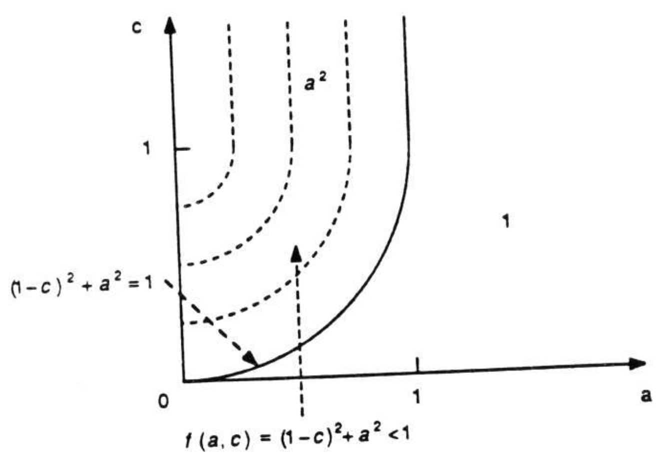


Fig. 10. Function f(a,c) of Proposition 3.17.

The dotted lines indicate curves of constant f, while the labels 1 and a^2 show the values of the function in the corresponding region. Note that f is nondecreasing in a and nonincreasing in c. The following two corollaries give sufficient conditions for unit asymptotic efficiency of the feedback detector, and for a performance improvement over the decorrelating detector used for the first-stage decisions, respectively.

Corollary 3.1: A sufficient condition for unit k^{th} user asymptotic efficiency is that

$$\forall\, i \neq k \qquad (a_i, c_i) \in R1 \,, \tag{3.155}$$

where R1 is the shaded region in Figure 11a, and a_i, c_i are defined in Proposition 3.17. The condition

$$\frac{w_i}{w_k} \geq R_{ii}^{-1} \,, \qquad \forall\, i \neq k \tag{3.156}$$

is sufficient for (3.155). $\diamondsuit$

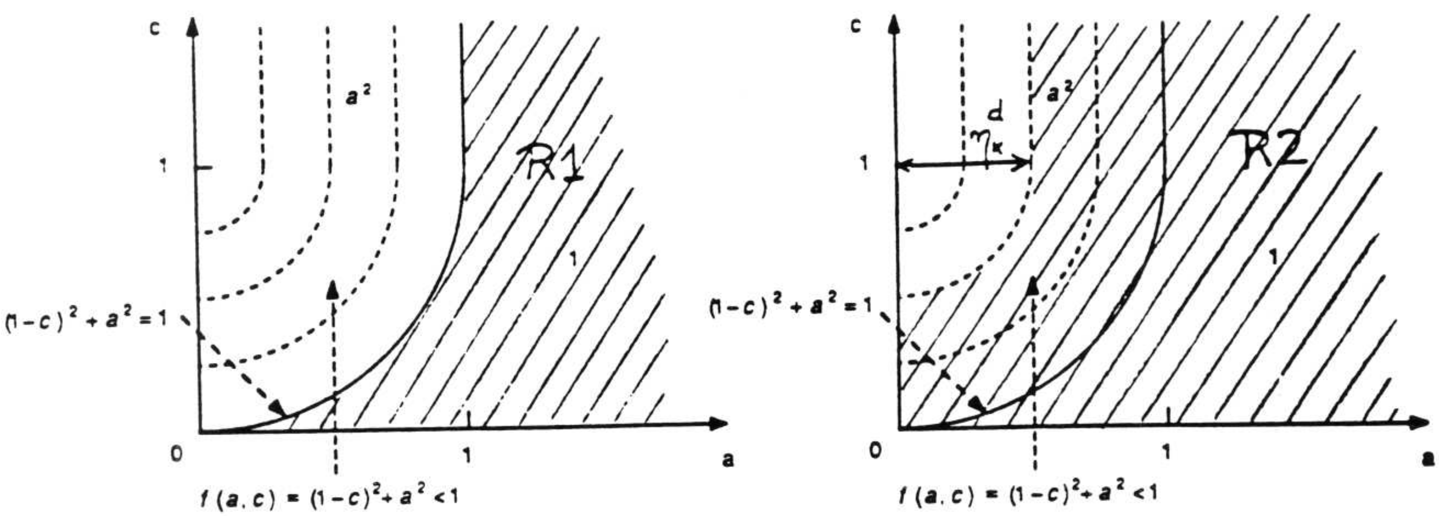


Fig. 11. a) R1 of Corollary 3.1, b) R2 of Corollary 3.2.

Corollary 3.2: A sufficient condition for achieving an improvement over decorrelating detector k^{th} user asymptotic efficiency is that

$$\forall\, i \neq k \qquad (a_i, c_i) \in R2 , \tag{3.157}$$

where R2 is the shaded region in Figure 11b, and a_i, c_i are defined in Proposition 3.17. The condition

$$\frac{w_i}{w_k} \geq \frac{R^{-1}_{ii}}{R^{-1}_{kk}} , \qquad \forall\, i \neq k \tag{3.158}$$

is sufficient for (3.157). ♢

Note that since $R^{-1}_{kk} \geq 1$, condition (3.158) is less stringent than the one in Corollary 3.1.

Proof of Corollaries : Conditions (3.155), (3.157) follow immediately from Proposition 3.17. For conditions (3.156), (3.158) notice that

$$f(a, c) \geq \min\,(a^2, 1) \tag{3.159}$$

and require $a_j^2 \geq 1,\ \forall j \neq k \Leftrightarrow w_j/w_k \geq 1/\eta_j^d$ for Corollary 3.1,

and $a_j^2 \geq \eta_k^d,\ \forall j \neq k \Leftrightarrow w_j/w_k \geq \eta_k^d/\eta_j^d$ for Corollary 3.2. ■

The following proposition is important, because it shows that by using feedback the near-far resistance of the decorrelating detector is still positive.

Proposition 3.18: For a linearly independent signal set the feedback detector is near-far resistant. Moreover

$$\overline{\eta_k^f} \triangleq \min_{\substack{w_i \geq 0 \\ i \neq k}} \eta_k^f \geq \frac{1}{1 + \xi_k^2} , \tag{3.160}$$

where

$$\xi_k = 2 \sum_{i \neq k} |R_{ki}| \frac{1}{\sqrt{\eta_i^d}} . \tag{3.161}$$

♢

Particularizing (3.160) to the two-user case, the near-far resistance there is lower bounded by

$$\overline{\eta_k^f} \geq \frac{1}{1 + (\frac{2|\rho|}{\sqrt{1-\rho^2}})^2} = \frac{1-\rho^2}{1+3\rho^2} = 1 - \frac{4}{3+\rho^{-2}} , \tag{3.162}$$

which is monotonically decreasing on $[1, 0]$ for $\rho \in [0, 1]$. The proofs of the propositions are given in a more general setting in Appendix 3.1. To particularize to the full-feedback case, F has to be set to $\{1,...,K\} \backslash \{k\}$.

Example 3.2. In the two-user case, the lower bound given by the right-hand side of (3.151) on the asymptotic efficiency of User 1, when feeding back the decorrelating detector estimates of the transmitted bit of User 2, is shown in Figure 12.

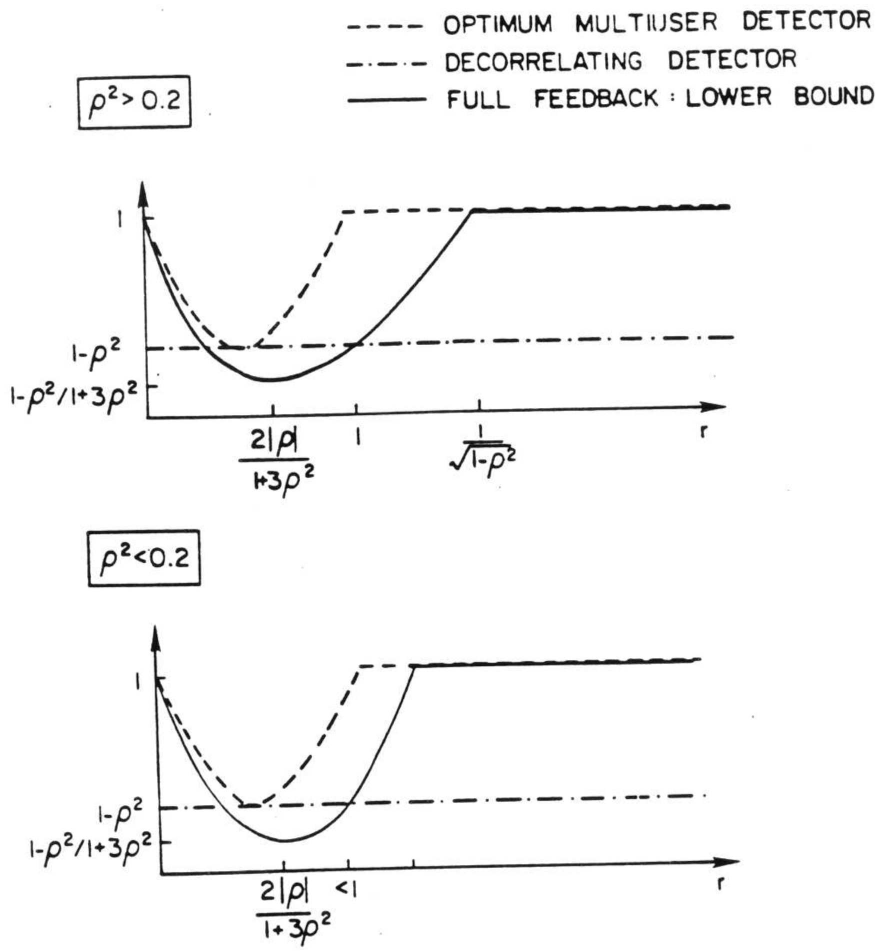


Fig. 12. Lower bound on the asymptotic efficiency of the full feedback detector for 2 users. Also shown are the asymptotic efficiency of the optimum multiuser detector and of the decorrelating detector.

Also shown are the asymptotic efficiencies of the decorrelating detector and of the optimum multiuser detector, the latter as an upper bound on the asymptotic efficiency of any detector. We see that the asymptotic efficiency of the decorrelating detector with feedback is no longer energy independent (we can infer this from the behavior of the lower bound, because otherwise the asymptotic efficiency would have to always be one, i.e., larger than that of the optimal multiuser detector, which is a contradiction), which comes about because the correctness of the fed-back decisions on the symbols of User 2 is a function of his energy w_2. Notice that a performance improvement is obtained over the decorrelating detector if either User 2 is very weak, or is stronger than User 1. In the latter case it is always better to use decision feedback for User 1, with a performance gain which is monotonically increasing as User 2 gets stronger with respect to User 1. When User 2 is strong enough, e.g., 1.15 times stronger than User 1 for a correlation of $\rho = 1/2$ between the normalized signals of the two users, the asymptotic efficiency of User 1 is unity, because the upper and lower bounds coincide. This means that for sufficiently strong interference the decorrelating detector with feedback performs as well as the optimum multiuser detector, and both achieve single-user performance, if the SNR with respect to the background noise is high enough. In this region the performance improvement results from the fact that the decorrelating detector decisions of User 2 have a very small error probability, due to the fact that User 2 is strong. The fact that there is also a performance gain if User 2 is weak, has a different explanation. In this case the detector for User 1 should ignore user 2, and not eliminate it at the expense of increasing the background noise, as the decorrelating detector does. This we have seen from Figure 4, when the conventional detector performs better than the decorrelating detector as long as w_2/w_1 $< (1 - \sqrt{1-\rho^2})/\rho$, e.g., for $\rho = 0.5$ the power of User 2 is under one fourth of that of User 1. In this region the performance gain of the fedback detector results not from the feedback itself (since User 2 is weak his decisions are not overly reliable) but because the -insignificant- interference is not eliminated any longer at the expense of the background noise. Finally, there is an intermediate energy region, where the feedback strategy may signify a performance loss over the decorrelating detector.

Intuitively there are two critical factors which may affect detrimentally the performance when using feedback: feeding back decisions of users with low η_i^d, and feeding back decisions of weak users, such that their error probability is high (in spite of possibly a high η_i^d), but who are not weak enough that their contributions are negligible.

To remedy this problem, in the following we analyze partial feedback, where only the decisions on the symbols of a subset of the interferers -considered reliable- are used.

3.8.2 Partial feedback of decisions from users in index-set F

Recall that in the full feedback case we subtracted from $z_k = \sqrt{w_k} b_k + n_k$ an estimate of the noise n_k based on estimates of the (correlated) noise components of the other users, which equaled $\hat{n}_i = z_i - \sqrt{w_i}\hat{b}_i$. In the partial feedback case only a subset F of noise estimates of this form is used. The other estimates, for $i \in \bar{F}$, are considered unreliable, and are replaced by their expectation in the absence of information, namely zero. The conditional expectation

$$E\,[n_k \mid \tilde{\mathbf{n}}] \;=\; -\mathbf{m}_k^T\,\tilde{\mathbf{n}}$$

is now replaced by

$$E\,[n_k \mid \tilde{\mathbf{n}},\, F] \;=\; -\mathbf{m}_k^T\,\mathbf{I}_F\,\tilde{\mathbf{n}} \qquad (I_F)_{ij} \;=\; \begin{cases} \delta_{ij}\,, & i \in F \\ 0\,, & \text{else}\,. \end{cases} \tag{3.163}$$

The partial feedback detector decides for:

$$\begin{aligned} \hat{b}_k^{(2)} &= \operatorname{sgn}\,(\,z_k \;+\; \mathbf{m}_k^T\,\mathbf{I}_F\,\hat{\tilde{\mathbf{n}}}\,) \\ &= \operatorname{sgn}\,\Big(\,z_k \;+\; \mathbf{m}_k^T\,\mathbf{I}_F\,(\,\tilde{\mathbf{z}} \;-\; \tilde{\mathbf{W}}\,\hat{\tilde{\mathbf{b}}}\,)\Big)\,, \end{aligned} \tag{3.164}$$

where as before $[1, \mathbf{m}_k, \mathbf{R}_k]$ denotes the k^{th} user partition of the matrix $\mathbf{R}$ and the tilde indicates that the entries corresponding to the k^{th} user have been deleted. Inserting $\mathbf{z} = \mathbf{W}\mathbf{b} + \mathbf{n}$, the dependence of $\hat{b}_k^{(2)}$ on the transmitted bits can be made explicit, namely

$$\hat{b}_k^{(2)} \;=\; \operatorname{sgn}\,(\sqrt{w_k}b_k \;+\; \mathbf{m}_k^T\,\mathbf{I}_F\,\tilde{\mathbf{W}}(\tilde{\mathbf{b}} - \hat{\tilde{\mathbf{b}}}) \;+\; n_k'')\,, \qquad n'' \sim N\,(0, \sigma^2\,(1 + \mathbf{m}_k^T\,\mathbf{I}_{\bar{F}}\,\mathbf{N}_k\,\mathbf{I}_{\bar{F}}\,\mathbf{m}_k)) \tag{3.165}$$

where $[1/\eta_k^d, *, \mathbf{N}_k]$ denotes the k^{th} user partition of the matrix $\mathbf{R}^{-1}$. Notice that the two limiting cases $F = \emptyset$ and $F = \{1, ..., K\}\backslash\{k\}$ reduce to the decorrelating and to the full feedback case, respectively. In the following the case $F = \emptyset$ is excluded, since no additional insight is gained by considering it. If the decisions that are fedback are correct

$$\eta_k^F \;=\; \frac{1}{1 \;+\; \mathbf{m}_k^T\,\mathbf{I}_{\bar{F}}\,\mathbf{N}_k\,\mathbf{I}_{\bar{F}}\,\mathbf{m}_k}\,, \tag{3.166}$$

which is less than 1 since $\mathbf{N}_k$ is positive definite (≥ 0 if $\mathbf{R} \geq 0$), unless full feedback is used or $\mathbf{m}_k = 0$. Therefore, in this optimistic case, a performance gain over the decorrelating detector is obtained as long as $\mathbf{m}_k^T\mathbf{I}_{\bar{F}}\mathbf{N}_k\mathbf{I}_{\bar{F}}\mathbf{m}_k < \mathbf{m}_k^T\mathbf{N}_k\mathbf{m}_k$. The following propositions generalize the results obtained in the full feedback case to include partial feedback.

Proposition 3.19: The k^{th} user asymptotic efficiency of the decorrelating detector with partial feedback set $F \neq \emptyset$ satisfies

$$\eta_k^F \;\geq\; \frac{1}{v^2(F)} \min_{i \in F} f\,(a_i v(F), c_i(F)) \tag{3.167}$$

where

$$a_i \;=\; \frac{\sqrt{w_i}}{\sqrt{w_k}} \sqrt{\eta_i^d}\,, \tag{3.168}$$

$$c_i(F) \;=\; 2 \sum_{\substack{n \in F \\ n\ s.t.\, a_n \leq a_i}} |R_{kn}| \frac{\sqrt{w_n}}{\sqrt{w_k}}\,, \tag{3.169}$$

$$v^2(F) \;=\; 1 + \mathbf{m}_k^T\, \mathbf{I}_{\bar{F}}\, \mathbf{N}_k\, \mathbf{I}_{\bar{F}}\, \mathbf{m}_k\,, \tag{3.170}$$

and

$$f\,(x, c) \;=\; \begin{cases} 1\,, & \text{for } c \geq 1,\ x \geq 1 \\ & \text{and } c < 1,\ (1-c)^2 + x^2 \geq 1 \\ x^2\,, & \text{for } c \geq 1,\ x < 1 \\ (1-c)^2 + x^2\,, & \text{else}\ . \end{cases} \tag{3.171}$$

with

$$\mathbf{R} \;=\; \begin{bmatrix} 1 & \mathbf{m}_k^T \\ \mathbf{m}_k & \mathbf{R}_k \end{bmatrix} \qquad \mathbf{R}^{-1} \;=\; \begin{bmatrix} 1/\eta_k^d & *^T \\ * & \mathbf{N}_k \end{bmatrix}\,. \tag{3.172}$$

$\diamondsuit$

The proof is given in Appendix 3.1. Note that the a_i and the function $f(.,.)$ are the same as in Proposition 3.17 and do not depend on the feedback set F. Obviously, if $F \neq \{1,...,K\} \backslash \{k\}$, $v^2(F) > 1$. Therefore we cannot achieve unit asymptotic efficiency using strictly partial feedback, regardless of the energy region. However, counterbalancing the detrimental effect of division by $v^2 > 1$, there are three favorable effects due to $v > 1$:

- $a_i\, v \;>\; a_i$, hence higher values of f,
- c_i possibly decreased, since summing over fewer terms, hence higher values of f,
- taking the $\min_{i \in F} f(.,.)$ over fewer terms.

Whether an improvement over feedback-free detection is achieved or not, depends on which influence dominates; hence it depends on the operating energy region, since the first two favorable effects are energy dependent. One thing is obvious: we need not consider feedback sets F, such that $1/v^2(F) \;\leq\; \eta_k^d$.

Proposition 3.20: If $\eta_i^d \neq 0 \ \forall i \in F$, the feedback detector with feedback set F is near-far resistant. Moreover

$$\overline{\eta_k^F} \triangleq \min_{\substack{w_i \geq 0 \\ i \in F}} \eta_k^F \geq \frac{1}{1 + \mathbf{m}_k^T \, \mathbf{I}_{\bar{F}} \, \mathbf{N}_k \, \mathbf{I}_{\bar{F}} \, \mathbf{m}_k + (2 \sum_{n \in F} |R_{kn}| \frac{1}{\sqrt{\eta_n^d}})^2} \ . \tag{3.173}$$

◇

The proof is given in Appendix 3.2.

Example 3.3. First consider the following example of a 3-user environment. Figure 13 shows the lower bound given by the right-hand side of equation (3.167) on the asymptotic efficiency of User 1, under feedback-free, partial feedback and full feedback conditions, assuming that users 1 and 2 are equally strong, with the energy of User 3 as a parameter.

Under feedback-free conditions, ($F = \emptyset$), the lower bound is achieved, and equals the asymptotic efficiency of the decorrelating detector. The asymptotic efficiency of the optimum multiuser detector, which constitutes an upper bound, is also shown. The signals of the 3 users are the ones shown in Figure 13, with crosscorrelations $\rho_{12} = -0.5$, $\rho_{13} = 0.87$ and $\rho_{23} = -0.65$, hence $\eta_1^d = 0.24$, $\eta_2^d = 0.56$ and $\eta_3^d = 0.19$. Since $v^2(\{2\}) \approx 5$, i.e., $1/v^2(\{2\}) < \eta_1^d$, the feedback sets to consider are $\{2,3\}$, $\{3\}$ and $\emptyset$. Considering Figure 13, it is apparent that significant performance improvement can be achieved in certain operating energy regions by using feedback of the decisions of other users, and that partial feedback can be superior to full feedback. Also both feedback schemes are near-far resistant, in accordance with Proposition 3.20, though with possibly smaller worst-case asymptotic efficiency than the decorrelating detector (recall that only lower bounds are tractable).

As can be seen from this 3-user example, different strategies are superior for different energy regions. While it is not desirable to use feedback regardless of the operating energy region - since taken by itself the feedback detector does not preserve optimum near-far resistance - substantial improvement can be achieved if feedback is used in certain energy regions. Thus in the given example, where users 1 and 2 have equal energy, if User 3 is four times as powerful, the asymptotic efficiency can be more than doubled by full feedback as compared to the decorrelating detector, and tripled by using only partial feedback of User 3.

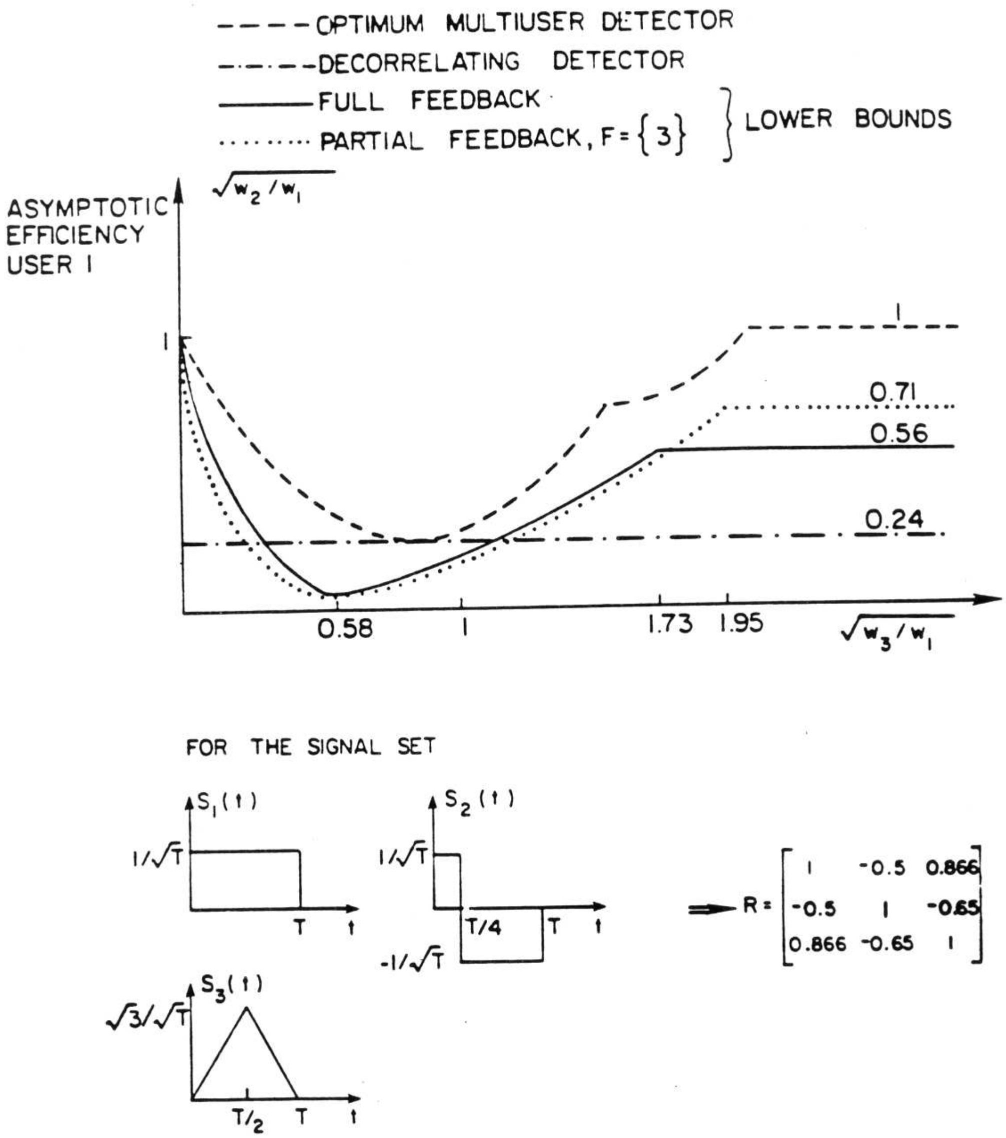


Fig. 13. Lower bound on the asymptotic efficiency of User 3 for full and partial feedback for a 3-user example with equally powerful interferers.

(3.21): By adapting the feedback set F to be the optimum set for the corresponding energy region the asymptotic efficiency is increased to

$$\eta_k^a \geq \max \{ \eta_k^d, \max_F \frac{1}{v^2(F)} \min_{i \in F} f(a_i v(F), c_i(F)) \} . \tag{3.174}$$

The following algorithm finds the optimum feedback set F:

[A] compute $v^2(F)$ for all sets $F \subseteq \{1, ..., K\} \backslash \{k\}$.

discard all F with $1\ v^2(F)\ <\ \eta_k^d$.

all remaining sets := admissible

order the admissible sets according to increasing $v^2(F)$.

note: full set will be first in list, with $v^2 = 1$

$F = \emptyset$ is admissible, since $v^2 = 1\ /\eta_k^d$.

[B] $\eta_{max}\ :=\ \eta_k^d$

$F_{opt}\ :=\ \emptyset$

while not at end of admissible sets

$F\ :=$ next admissible set

$v\ :=\ \sqrt{v^2(F)}$

compute the $a_i, c_i, \forall i \in F$

$\eta\ =\ \frac{1}{v^2} \min_{i \in F}\ f(a_i v, c_i)$

if $(\eta_{max}\ <\ \eta)\ F_{opt} = F,\ \eta_{max} = \eta$

if $(v^2 \eta\ =\ 1)$ **return** F_{opt}, η_{max}

end.

return

return F_{opt}, η_{max}

end. ◇

Part [A] of the procedure is independent of energies, hence is done only once. If the energies are not constant the optimum set F will change in time, hence part [B] has to be executed periodically to update F. If the energies are known exactly as a function of time, the adaptive feedback detector is seen to have optimum near-far resistance. If the energies are not known, and have to be estimated at the receiver, we have the following cases.

a) The energies vary slowly in time, such that for long transmission lengths they can be considered constant. In this case estimates of the energies can be obtained up to any desired precision. One obvious possibility is to estimate

$$\hat{w}_k = \frac{1}{N} \sum_{i=1}^{N} (z_k)_i^2 - \sigma^2 R_{kk}^{-1} ,$$

where $(z_k)_i$ is the input sequence to the decorrelating detector. Since the information bits and the noise are uncorrelated

$$E[z_k^2] = E[w_k + 2\sqrt{w_k} b_k n_k + n_k^2] = w_k + \sigma^2 R_{kk}^{-1} ,$$

hence, by the weak law of large numbers $\hat{w}_k$ converges to w_k as $N \to \infty$. More elaborate and effective schemes can be found in the literature. Since the energies can be estimated with arbitrary precision the energy-adaptive feedback detector can be used. It can be robustified with respect to small errors in the energy estimates by accepting a nonempty feedback set F only if the achieved performance improvement exceeds a certain threshold. Moreover, Proposition 3.20 shows that feedback from users with nonzero asymptotic efficiency is near-far resistant, hence near-far resistance is not lost even for completely erroneous energy estimates. However optimum near-far resistance is then lost, which is why large errors in the estimates of the user energies are detrimental.

b) A subset of users has slowly varying energies, while the rest has energies that vary rapidly in time. In this case only feedback from the slowly moving users should be considered, otherwise identical to a).

c) The energies of all the users change rapidly. In this case no feedback can be used. The decorrelating detector is well suited for this environment since it is energy independent and has optimum near-far resistance.

Also note that an inherent drawback of feedback schemes compared to the decorrelating detector is that the former require knowledge of the received energies. Finally, it can be shown that, if $\mathbf{R}$ is not invertible, the results obtained still apply if F excludes all linearly dependent users.

Appendix 3.1: Performance analysis of partial feedback

From (3.165) the partial feedback detector decides for:

$$\hat{b}_k^{(2)} = \text{sgn}\,(\,\sqrt{w_k} b_k \;+\; \mathbf{m}_k^T\, \mathbf{I}_F\, \tilde{\mathbf{W}}(\tilde{\mathbf{b}} - \hat{\tilde{\mathbf{b}}}) \;+\; n_k''\,)\,, \qquad \mathbf{n}'' \sim N\,(0,\, \sigma^2 v^2)\,, \tag{3.175}$$

where we have set $1 + \mathbf{m}_k^T\, \mathbf{I}_{\bar{F}}\, \mathbf{N}_k\, \mathbf{I}_{\bar{F}}\, \mathbf{m}_k := v^2$. The probability of error for this decision scheme is, by symmetry

$$\begin{aligned}
P_{\hat{\mathbf{b}}^{(2)},k}(e) &= P\,(\hat{b}_k^{(2)} = 1 \,\wedge\, b_k = -1) \;+\; P\,(\hat{b}_k^{(2)} = -1 \,\wedge\, b_k = 1) \\
&= P\,(\hat{b}_k^{(2)} = -1 \mid b_k = 1) \\
&= \sum_{\substack{(b_j, \hat{b}_j) \\ j\in F}} P\,(\hat{b}_k^{(2)} = -1\,,\; b_j, \hat{b}_j, j{\in}F \mid b_k{=}1) \\
&= \sum_{\substack{(b_j, \hat{b}_j) \\ j\in F}} P\,(\hat{b}_k^{(2)} = -1 \mid b_j, \hat{b}_j, j{\in}F\,,\; b_k{=}1)\; P\,(\hat{b}_j, j{\in}F \mid b_j, j{\in}F\,,\; b_k{=}1)\;\; P\,(b_j, j{\in}F \mid b_k{=}1) \\
&= \frac{1}{2^{|F|}} \sum_{\substack{(b_j, \hat{b}_j) \\ j\in F}} P\,(n_k' < \mathbf{m}_k^T\, \mathbf{I}_F\, \tilde{\mathbf{W}}(\tilde{\mathbf{b}} - \hat{\tilde{\mathbf{b}}}) - \sqrt{w_k} \mid b_j, \hat{b}_j, j{\in}F)\; P\,(\hat{b}_j, j{\in}F \mid b_j, j{\in}F\,,\; b_k{=}1) \\
&= \frac{1}{2^{|F|}} \sum_{\substack{(b_j, \hat{b}_j) \\ j\in F}} Q\left(\frac{\sqrt{w_k} \;-\; \sum_{j\in F} R_{kj}\sqrt{w_j}\,(b_j - \hat{b}_j)}{\sigma v}\right) P\,(\hat{b}_j, j{\in}F \mid b_j, j{\in}F\,,\; b_k{=}1)\,.
\end{aligned} \tag{3.176}$$

The probability $P\,(\hat{b}_j, j{\in}F \mid b_j, j{\in}F\,,\; b_k{=}1)$ is hard to specify. However, noticing that the joint event $\hat{b}_j = \mathbf{x}, j{\in}F$ is contained in the marginal event $\hat{b}_i = x_i$, for each $i \in F$, we can find an upper bound, which is exact in the two-user case.

$$P\,(\hat{b}_j, j{\in}F \mid b_j, j{\in}F\,,\; b_k{=}1) \;\le\; \min_{i\in F} P\,(\hat{b}_i \mid b_j, j{\in}F\,,\; b_k{=}1)\,. \tag{3.177}$$

Next we have to use the precise form of the detector used to obtain the $\hat{b}_i$. For the decorrelating detector the decision on the symbol of each user is independent of the interfering transmissions of the other users, i.e.,

$$\begin{aligned}
P\,(\hat{b}_i \mid b_j, j{\in}F\,,\; b_k{=}1) &= P\,(\hat{b}_i \mid b_i) \\
&= Q_i \;+\; \delta_{b_i \hat{b}_i}(1 - 2Q_i)\,,
\end{aligned} \tag{3.178}$$

where, for notational simplicity, we have abbreviated by Q_i the probability of error of the decorrelating detector, i.e., $Q_i = Q(\sqrt{w_i/R_{ii}^{-1}}/\sigma)$, and δ_{ij} is as usually zero for $i \neq j$ and one else. So we now have

$$P\,(\hat{b}_j, j{\in}F \mid b_j, j{\in}F\ ,\ b_k{=}1) \ \leq\ \min_{i\in F}\,[\,Q_i \ +\ \delta_{b_i\hat{b}_i}(1-2Q_i)\,]\,. \tag{3.179}$$

Let the index set $\{i{\in}F|\delta_{b_i\hat{b}_i} = 0\}$ of users whose first-stage decisions are erroneous, be denoted by S. Also let $i_1, ..., i_{|F|}$ be the permutation of the users in F according to increasing decorrelating detector error probability Q_i (i.e., we order the users according to their expected reliability), and let $r(i)$ return the rank of User i in this permutation, i.e., $r(i_k) = k$. Let $i(S)$ be the index of the user in S with the lowest Q_i, i.e., $i(S) = i_{\min\{j|i_j\in S\}}$, which is the index of the most reliable user among those who did make a first-stage decoding error. (3.179) now becomes

$$\begin{aligned} P\,(\hat{b}_j, j{\in}F \mid b_j, j{\in}F\ ,\ b_k{=}1) \ &\leq\ \begin{cases} \min\limits_{i\in S}\ Q_i\ , & S \neq \emptyset \\ 1\ -\ \max\limits_{i\in F}\ Q_i\ , & S = \emptyset \end{cases} \\ &=\ \begin{cases} Q_{i(S)}\ , & S \neq \emptyset \\ 1\ -\ Q_{i(F)}\ , & S = \emptyset\,. \end{cases} \end{aligned} \tag{3.180}$$

Thus, combining (3.176) and (3.180), we obtain

$$P_{\hat{\mathbf{b}}^{(2)},k}(e) \ \leq\ P^{UB}_{\hat{\mathbf{b}}^{(2)},k}(e) \ \triangleq\ \frac{1}{2^{|F|}} \sum_{\substack{(b_j,\hat{b}_j)\in\{\pm1\}^2\\ j\in F}} \left[Q\left(\frac{\sqrt{w_k}\ -\ \sum_{j\in F} R_{kj}\sqrt{w_j}\,(b_j-\hat{b}_j)}{\sigma v}\right) \begin{cases} Q_{i(S)}\ , & S \neq \emptyset \\ 1\ -\ Q_{i(F)}\ , & S = \emptyset \end{cases} \right]\,. \tag{3.181}$$

Now, $b_j - \hat{b}_j$ can have the values 0, 2 sgnR_{kj} or -2 sgnR_{kj}, and the "+sgn" case makes a higher contribution to $P(e)$, while having the same $\delta_{b_j\hat{b}_j}$ pattern, so that the other Q-term affecting the error probability is the same. Hence the leading terms for $\sigma \to 0$ are among those for which $(b_j, \hat{b}_j) \in \{(1,1), (-1,-1), (\text{sgn}R_{kj}, -\text{sgn}R_{kj})\}$. Denote by A the subset of pairs $(\mathbf{b}, \hat{\mathbf{b}})$ whose j^{th} components, $j{\in}F$, are in the above set. Then

$$P^{UB}_{\hat{\mathbf{b}}^{(2)},k}(e) \ \underset{\sigma\to0}{\longrightarrow}\ \frac{1}{2^{|F|}} \sum_{(\mathbf{b},\hat{\mathbf{b}})\in A} \left[Q\left(\frac{\sqrt{w_k}\ -\ \sum_{j\in F} |R_{kj}|\sqrt{w_j}\,(1-\delta_{b_j-\hat{b}_j})}{\sigma v}\right) \begin{cases} Q_{i(S)}\ , & S \neq \emptyset \\ 1\ -\ Q_{i(F)}\ , & S = \emptyset \end{cases} \right]\,. \tag{3.182}$$

The above expression depends on $(b_j, \hat{b}_j)$ only through $\delta_{b_j\hat{b}_j}$. Hence we can group the summands whose $\delta_{b_j\hat{b}_j}$ pattern, $j \in F$, coincides in all components, as follows. There are $2^{|F|}$ sets $\{\delta_{b_j\hat{b}_j}, j \in F\}$. Fix a $j \in F$ and consider the corresponding component pair $(b_j, \hat{b}_j)$. If $\delta_{b_j\hat{b}_j} = 0$, i.e., the corresponding components are different, there is only one possibility for the components, since $(\mathbf{b}, \hat{\mathbf{b}}) \in A$. Otherwise, there are two choices which result in $\delta_{b_j\hat{b}_j} = 1$. This implies that the number of vector pairs $(\mathbf{b}, \hat{\mathbf{b}}) \in A$ which result in the same $\delta_{b_j\hat{b}_j}$ pattern for $j \in F$ is

$$2^{\sum_{j\in F} \delta_{b_j\hat{b}_j}} = 2^{|F|-|S|} . \tag{3.183}$$

Therefore the right hand side in (3.182) equals

$$\frac{1}{2^{|F|}} \sum_{\substack{\{\delta_{b_j\hat{b}_j} \in \{0,1\},\\ j\in F\}}} 2^{|F|-|S|} \left[Q\left(\frac{\sqrt{w_k} - \sum_{j\in S} |R_{kj}|\sqrt{w_j}}{\sigma v} \right) \begin{cases} Q_{i(S)} , & S \neq \emptyset \\ 1 - Q_{i(F)} , & S = \emptyset \end{cases} \right] . \tag{3.184}$$

where, recall, $S = \{i \in F | \delta_{b_i\hat{b}_i} = 0\}$ is the index set of the users whose first-stage decisions are erroneous. $i(S)$ is the index of the user in S with the lowest Q_i, i.e., with the lowest error probability when deciding for $\hat{b}_i$ with the decorrelating detector, among those users in F who made wrong first-stage decisions on their bits. Among all sets S with the same $i(S)$, the set which will asymptotically dominate the error probability as $\sigma \to 0$ is the one which includes all users in F with $Q_i \geq Q_{i(S)}$. This set has $|F| - r(i(S)) + 1$ elements. Hence letting $j := r(i(S))$ run from 1 to $|F|$ we have

$$P^{UB}_{\hat{\mathbf{b}}^{(2)},k}(e) \underset{\sigma\to 0}{\longrightarrow}$$
$$\frac{1}{2^{|F|}} \sum_{j=1}^{|F|} \left[2^{j-1}\, Q \left(\frac{\sqrt{w_k} - \sum_{\substack{n\in F\\ Q_n \geq Q_{i_j}}} |R_{kn}|\sqrt{w_n}}{\sigma v} \right) Q_{i_j} \right] + Q(\frac{\sqrt{w_k}}{\sigma v})\ (1 - Q_{i_{|F|}}) \tag{3.185}$$

and finally, neglecting $Q_{i_{|F|}}$ versus 1 (since it becomes arbitrarily small for $\sigma \to 0$), and after the change of variable $i := i_j$,

$$P^{UB}_{\hat{\mathbf{b}}^{(2)},k}(e) \underset{\sigma\to 0}{\longrightarrow}$$
$$Q\left(\frac{\sqrt{w_k}}{\sigma v(F)}\right) + \frac{1}{2^{|F|}} \sum_{i\in F} 2^{r(i)-1}\, Q\left(\frac{\sqrt{w_i/R^{-1}_{ii}}}{\sigma}\right) Q\left(\frac{\sqrt{w_k} - \sum_{\substack{n\in F\\ Q_n\geq Q_i}} |R_{kn}|\sqrt{w_n}}{\sigma v(F)} \right)$$

(3.186)

where we have made explicit the dependence of the noise variance v on the feedback set F. If $F = \emptyset$ the above expression reduces to the first term, and since it is easily shown that $v^2(\emptyset) = 1 + \mathbf{m}^T\mathbf{N}\mathbf{m} = \eta_k^d$ (the asymptotic efficiency of the decorrelating detector), the upper bound in (3.186) is achieved if F is empty. In the sequel we consider the case $F \neq \emptyset$. Then (3.186) has the form

$$P_k^{UB}(e) \underset{\sigma\to 0}{\longrightarrow} \sum_{i\in F} \left[k_o^2\, Q\left(\frac{x}{v}\right) + k_i^2\, Q\left(xa_i\right) Q\left(\frac{x}{v}(1-c_i)\right)\right] , \tag{3.187}$$

$$\text{where} \quad x = \sqrt{w_k}/\sigma\ , \quad k_o^2 = 1/|F|\ , \quad k_j^2 = 2^{r(i)-|F|-1} \tag{3.188}$$

$$a_i = \frac{\sqrt{w_i}}{\sqrt{w_k}} \sqrt{\eta_i^d} > 0 \tag{3.189}$$

$$\text{and} \quad c_i = 2 \sum_{\substack{n\in F \\ s.t. \\ a_n\leq a_i}} |R_{kn}| \frac{\sqrt{w_n}}{\sqrt{w_k}} > 0\,. \tag{3.190}$$

In investigating the associated asymptotic efficiency, we will make use of the following properties, which are immediate if the asymptotic efficiency is thought of as the slope with which the logarithm of the error probability decays in the high SNR region :

1) The asymptotic efficiency obtained from an upper bound on the actual error probability is a lower bound on the actual asymptotic efficiency.

2) The asymptotic efficiency obtained from a sum of terms is lower bounded by the minimum of the asymptotic efficiencies obtained from the single terms.

Then, if η_k^i denotes the asymptotic efficiency resulting from the i^{th} summand, the k^{th} user asymptotic efficiency of the decision-feedback detector is by 1) and 2) lower bounded by

$$\eta_k^F \geq \min_{i\in F} \eta_k^i\,. \tag{3.191}$$

Each η_k^i is the asymptotic efficiency associated with an error probability of the form

$$P = k_o^2 Q\left(\frac{x}{v}\right) + k_i^2\, Q\left(xa\right) Q\left(\frac{x}{v}(1-c)\right) , \qquad a, c > 0 \tag{3.192}$$

hence is a function of a and c. We distinguish the following cases :

A $c \geq 1 \Rightarrow$ in the high SNR region

$$P \to k_o^2\, Q\left(\frac{x}{v}\right) + \frac{k_i^2}{2}\, Q\left(xa\right) \tag{3.193}$$

$\Rightarrow \eta \geq 1/v^2 \ \min(a^2v^2, 1)$.

B $c < 1$. We use the inequality $Q(\alpha) \leq 1/2 \exp(-\alpha^2/2)$ for $\alpha \geq 0$ (e.g. [Woz]) to upper bound P. We get

$$P \leq \frac{k_o^2}{2} e^{-\frac{x^2}{2}\frac{1}{v^2}} + \frac{k_1^2}{4} e^{-\frac{x^2}{2}a^2} e^{-\frac{x^2}{2}\frac{(1-c)^2}{v^2}} \tag{3.194}$$

B1 $(1-c)^2 + a^2v^2 \geq 1$: then in the high SNR region the upper bound will approach the first term, and by 1) $\eta \geq 1/v^2$.

B2 $(1-c)^2 + a^2v^2 < 1$: then in the high SNR region the upper bound will approach the second term, and by 1) $\eta \geq ((1-c)^2 + a^2v^2)/v^2$.

Finally, combining with (3.191), we obtain

$$\eta_k^F \geq \frac{1}{v^2(F)} \min_{i \in F} f\ (a_i v(F), c_i(F)) \tag{3.195}$$

where

$$f\ (x, c) = \begin{cases} 1\ , & \text{for } c \geq 1,\ x \geq 1 \\ & \text{and } c < 1,\ (1-c)^2 + x^2 \geq 1 \\ x^2\ , & \text{for } c \geq 1,\ x < 1 \\ (1-c)^2 + x^2\ , & \text{else}\ . \end{cases} \tag{3.196}$$

■

Appendix 3.2 : Near-far resistance of partial feedback

Let r_i be a short hand notation for $\sqrt{w_i/w_k}$. We want to show that the lower bound on k^{th} user asymptotic efficiency given in Proposition 3.19 is near-far resistant for each feedback set F, i.e.,

$$\overline{\eta_k^F} \triangleq \min_{\substack{r_i \geq 0 \\ i \in F}} \eta_k^F \geq 0\ . \tag{3.197}$$

By definition (3.168) $a_i = \sqrt{\eta_i^d} r_i$. We have the following upper bound on c_i, which will be useful since f is nonincreasing in c.

$$c_i(F) = 2 \sum_{\substack{n \in F \\ a_n \leq a_i}} |R_{kn}|\, r_n = \sum_{\substack{n \in F \\ a_n \leq a_i}} |R_{kn}| \frac{a_n}{\sqrt{\eta_n^d}} \tag{3.198}$$

$$\leq a_i\, 2 \sum_{\substack{n \in F \\ a_n \leq a_i}} |R_{kn}| \frac{1}{\sqrt{\eta_n^d}} \leq a_i\ 2 \sum_{n \in F} |R_{kn}| \frac{1}{\sqrt{\eta_n^d}} \tag{3.199}$$

$$\triangleq a_i \xi_k = \sqrt{\eta_i^d} \xi_k r_i \triangleq k_c r_i\ , \qquad k_c > 0\ . \tag{3.200}$$

Now the function f has three regions, given by the three cases in its definition (3.171). In the second region

$$1 \leq c_i \leq k_c r_i \quad \text{and} \quad k_c r_i v \leq 1 \ \Rightarrow \ \frac{1}{k_c} \leq r_i \leq \frac{1}{k_c v} \tag{3.201}$$

therefore f is lower bounded by

$$f(a_i v, c_i) \ = \ a_i^2 v^2 \ = \ \eta_i^d v^2 r_i^2 \ \geq \ \eta_i^d v^2 \frac{1}{k_c^2} \,. \tag{3.202}$$

In the third region

$$0 \leq c_i \leq \min\{k_c r_i, 1\} \ \Rightarrow \ 1 \geq 1 - c_i \geq 1 - \min\{k_c r_i, 1\} \geq 0 \tag{3.203}$$

therefore f is lower bounded by

$$f(a_i v, c_i) \ = \ (1-c_i)^2 + v^2 a_i^2 \geq \min_{\substack{r_i \\ k_c r_i \leq 1}} [(1-k_c r_i)^2 + \eta_i^d r_i^2 v^2] \ = \ \frac{\eta_i^d v^2}{k_c^2 + \eta_i^d v^2} \,. \tag{3.204}$$

Finally, combining the three cases,

$$f(a_i v, c_i) \ \geq \ \min\{1, \ \frac{\eta_i^d v^2}{k_c^2}, \ \frac{\eta_i^d v^2}{k_c^2 + \eta_i^d v^2}\} \ = \ \frac{\eta_i^d v^2}{k_c^2 + \eta_i^d v^2} \ = \ \frac{v^2}{\xi_k^2 + v^2} \tag{3.205}$$

which is independent of i. From here, using (3.195), $\eta_k^F \geq 1/(\xi_k^2 + v^2)$, which is energy-independent. Using the definition (3.200) of ξ_k, and the definition of v, (3.173) follows, together with a strictly positive near-far resistance, since $\mathbf{N}_k$ is positive definite. ■

4. Near-far resistance of multiuser detectors in asynchronous channels

4.1 Optimum near-far resistance of linear detection

4.1.1 Multiuser communication model

Let the receiver input signal be

$$r(t) = S(t, \mathbf{b}) + n(t) \ , \tag{4.1}$$

where $n(t)$ is zero-mean white Gaussian noise with power spectral density σ^2 and

$$S(t, \mathbf{b}) = \sum_{i=-M}^{M} \sum_{k=1}^{K} b_k(i) \sqrt{w_k(i)}\ s_k(t - iT - \tau_k) \tag{4.2}$$

is the element of $\mathcal{L}_2$ (the Hilbert space of square-integrable functions) which contains the information sequence $\mathbf{b} = \{\mathbf{b}(i) = [b_1(i), ..., b_K(i)],\ b_k(i) \in \{-1, 1\}, k = 1, ..., K; i = -M, ...M\}$, $s_k(t)$ is the normalized signature waveform of User k and is zero outside the interval [0,T], and $w_k(i)$ is the received energy of User k in the i^{th} time slot. Let $N = 2M + 1$ be the length of the transmitted sequence. Without loss of generality it is assumed that the users are numbered such that their delays satisfy $0 \leq \tau_1 \leq ... \leq \tau_K < T$. The normalized signal $\tilde{S}(t, \mathbf{b})$ is the receiver input signal corresponding to unit energies.

Define the vector space $L = \{\mathbf{x} = [\ \mathbf{x}(-M), ..., \mathbf{x}(M)\ ] = [\ [\ x_1(-M), ..., x_K(-M)\ ]\ , \ \ldots\ ,$ $[\ x_1(M), ..., x_K(M)\ ]\ ]^T,\ x_k(i) \in I\!R,\ k = 1, ..., K,\ i = -M, ..., M\}$, (each element of which can be equivalently viewed as a sequence of N $(K*1)$-vectors or as one single $(NK*1)$-vector), and define the $(k,i)^{th}$ unit vector $\mathbf{u}^{k,i}$ in L as $u_j^{k,i}(l) = \delta_{kj}\delta_{li}$. Let $< \cdot, \cdot >$ denote the usual inner product on $\mathcal{L}_2$, i.e. the integral of the product over the region of support, with induced norm $\| \cdot \|$. Henceforth, we make the following assumption on $\tilde{S}(t, \mathbf{b})$:

Linear Independence Assumption (LIA) :

$$\forall\, \mathbf{x} \in L\ ,\ \mathbf{x} \neq 0 \quad \Rightarrow \quad \|\tilde{S}(t, \mathbf{x})\| \ \neq \ 0\ . \tag{4.3}$$

In other words, no matter what the user energies are, the received signal does not vanish everywhere if at least one of the users has transmitted a symbol. This condition fails to hold only in pathological non-practical cases with very heavy crosscorrelation between the signals, such as the two-user example in Figure 14. There, if the delay between the users is $T/2$, the received signal can be identically zero although transmissions have been made (this happens if, for all i, $b_2(i) = -b_1(i)$). It is shown in Appendix 4.1 that such a situation will arise with probability zero if the a priori unknown delays are uniformly distributed, which is the case in an asynchronous unslotted channel used by non-cooperating users. Basically, in order to violate the LIA, a subset of the users must be effectively synchronous and the modulating signals of this subset have to be heavily correlated.

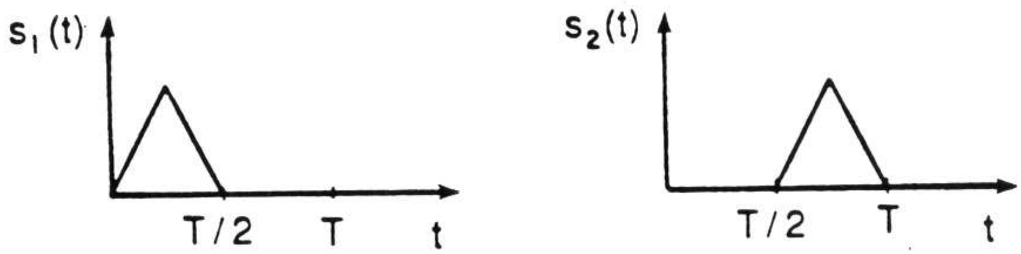


Fig. 14. Example of signature waveforms which can violate the LIA.

For simplicity the LIA will be in effect in the rest of this section. It has been shown to be a mild condition. If it is removed all the given results can be generalized in a manner analogous to the treatment of the synchronous transmission case. The changes that have to be made are given in Appendix 4.2.

The sampled output of the normalized matched filter for the i^{th} bit of the k^{th} user, $i = -M, ..., M$, is

$$y_k(i) = \int_{iT+\tau_k}^{iT+T+\tau_k} r(t) s_k(t - iT - \tau_k) dt \tag{4.4}$$

$$= \int_{-\infty}^{\infty} S(t, \mathbf{b}) s_k(t - iT - \tau_k) dt + \int_{-\infty}^{\infty} n(t) s_k(t - iT - \tau_k) dt \, , \tag{4.5}$$

where the second equality is valid since the signals are zero outside [0,T]. It is well established (e.g. [Ver 86a]) that the whole sequence $\mathbf{y}$ of outputs of the bank of K matched filters, with components $y_k(i)$ given by (4.4), for $k = 1, ..., K,\ i = -M, ..., M$, is a sufficient statistic for decision on the most likely transmitted information sequence $\mathbf{b}$. The multiuser demodulation problem at the receiver of User k now is to recover the sequence $\{b_k(i)\}$, transmitted by User k, from the sequence $\mathbf{y} \in L$. Due to the uncertainty introduced by the noise any detector will have a nonzero probability of making an error. The different detectors are characterized by their probability of error versus computational complexity tradeoff. Motivated by the state of the art - where the choice lies between the optimum multiuser detector, which is of exponential complexity and the ad hoc single user detector whose performance degrades to zero for sufficiently high interference energy - we define a class of simple detectors and optimize performance within this class, to obtain an acceptable error probability versus complexity tradeoff.

A *linear detector* for bit i of User k is characterized by $\mathbf{v}^{k,i} \in L$. The decision of the detector is given by the polarity of the inner product of $\mathbf{v}^{k,i}$ and the vector $\mathbf{y}$ of matched filter outputs, which is equal to

$$\sum_{l=-M}^{M} \sum_{j=1}^{K} v_j^{k,i}(l)\ y_j(l) = \int_{-\infty}^{\infty} \tilde{S}(t, \mathbf{wb})\tilde{S}(t,\ \mathbf{v}^{k,i})dt\ +\ n_{k,i} \tag{4.6}$$

$$= < \tilde{S}(t, \mathbf{wb}), \tilde{S}(t,\ \mathbf{v}^{k,i}) > + \ n_{k,i}\ , \tag{4.7}$$

where for any sequence $\mathbf{b}$ of information bits, $\mathbf{wb}$ will denote the sequence of amplitudes $\mathbf{wb} = \{[\sqrt{w_1(i)}b_1(i), ..., \sqrt{w_K(i)}b_K(i)],\ i = -M, ...M\}$. $n_{k,i}$ is the noise component at the output of the cascade of matched filter, sampler and detector, hence is a Gaussian zero-mean random variable with variance given by

$$E[n_{k,i}^2] = \sum_{k,l,i,j} v_k(l)v_j(i) \int_{-\infty}^{\infty} \sigma^2 s_k(t - lT - \tau_k)s_j(t - iT - \tau_j)dt\ =\ \sigma^2\|\tilde{S}(t,\ \mathbf{v}^{k,i})\|^2\ . \tag{4.8}$$

The receiver decides on the i^{th} bit of the k^{th} user according to the rule

$$\hat{b}_k(i) = \text{sgn} \sum_{l=-M}^{M} \sum_{j=1}^{K} v_j^{k,i}(l)\ y_j(l) \tag{4.9}$$

$$= \text{sgn}\ (< \tilde{S}(t, \mathbf{wb}), \tilde{S}(t,\ \mathbf{v}^{k,i}) > + \ n_{k,i})\ . \tag{4.10}$$

Wherever it is clear from the context, the superscripts k, i will be omitted.

Matrix Notation: It is convenient to introduce the following compact notation. Define the $K * K$ normalized signal crosscorelation matrices $\mathbf{R}(l)$ whose entries are given by

$$R_{kj}(l) = \int_{-\infty}^{\infty} s_k(t-\tau_k)\ s_j(t+lT-\tau_j)\ dt\ . \tag{4.11}$$

Then, since the modulating signals are zero outside $[0,T]$

$$\mathbf{R}(l) = 0 \quad \forall\ |l| > 1\ , \tag{4.12}$$

$$\mathbf{R}(-l)\ =\ \mathbf{R}^T(l)\ , \tag{4.13}$$

and, if the users are numbered according to increasing delays, $\mathbf{R}(1)$ is an upper triangular matrix with zero diagonal. Also let $\mathbf{W}(l) = \text{diag}\ ([\sqrt{w_1(l)}, ..., \sqrt{w_K(l)}])$. With this notation the matched filter outputs for $l = \{-M, ..., M\}$ can be written in vector form as (cf. [Ver 83b])

$$\mathbf{y}(l) = \mathbf{R}(-1)\mathbf{W}(l+1)\ \mathbf{b}(l+1) + \mathbf{R}(0)\mathbf{W}(l)\ \mathbf{b}(l) + \mathbf{R}(1)\mathbf{W}(l-1)\ \mathbf{b}(l-1) + \mathbf{n}(l)\ , \tag{4.14}$$

as can be seen for each component by inserting (4.1) into (4.4). We make the convention that $\mathbf{b}(-M-1) = \mathbf{b}(M+1) = 0$. $\mathbf{n}(l)$ is the matched filter output noise vector, with autocorrelation matrix given by

$$E\ [\mathbf{n}(i)\mathbf{n}^T(j)]\ =\ \sigma^2\ \mathbf{R}(i-j)\ . \tag{4.15}$$

The entries of the matrices $\mathbf{R}(i),\ i = -1, 0, 1$ are obtained at the receiver by crosscorrelating appropriately delayed replicas of the normalized signature waveforms according to (4.11). Note that no additional complexity is hereby required of the receiver, since knowledge of the normalized signature waveforms and the capability to lock onto the respective delays are necessary for matched filtering and sampling at the instant of maximal signal-to-noise ratio.

In contrast to (4.5) the asynchronous nature of the problem is clearly transparent in (4.14). To make this notation more compact we define unifying variables, the $NK * NK$ symmetric block-Toeplitz matrix $\mathcal{R}$ and the $NK * NK$ diagonal matrix $\mathcal{W}$, as follows:

$$\mathcal{R} = \begin{pmatrix} \mathbf{R}(0) & \mathbf{R}(-1) & 0 & \dots & 0 \\ \mathbf{R}(1) & \mathbf{R}(0) & \mathbf{R}(-1) & & \vdots \\ 0 & \mathbf{R}(1) & \mathbf{R}(0) & \ddots & 0 \\ \vdots & & \ddots & \ddots & \mathbf{R}(-1) \\ 0 & \dots & 0 & \mathbf{R}(1) & \mathbf{R}(0) \end{pmatrix}, \tag{4.16}$$

$$\mathcal{W} \;=\; \mathrm{diag}\,([\sqrt{w_1(-M)},...,\sqrt{w_K(-M)},\ldots,\sqrt{w_1(M)},...,\sqrt{w_K(M)}])\,. \qquad (4.17)$$

In this notation the matched filter output vector $\mathbf{y}$ depends on $\mathbf{b}$ via, from (4.14)

$$\mathbf{y} \;=\; \mathcal{R}\mathcal{W}\mathbf{b} \;+\; \mathbf{n}\,. \qquad (4.19)$$

The matrix $\mathcal{R}$ can be interpreted as the cross-correlation matrix for an equivalent synchronous problem where the whole transmitted sequence is considered to result from $N * K$ users, labeled as shown in Figure 15, during one transmission interval of duration $T_e = N * T + \tau_K - \tau_1$.

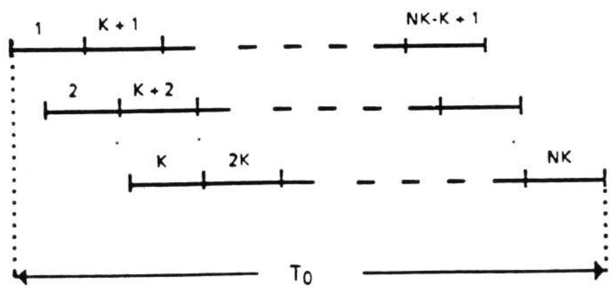


Fig. 15. Equivalent Synchronous Transmitted Sequence.

Then the results presented here for finite transmission length can be derived via analysis of synchronous multiuser communication, as done in Chapter 3. However, the approach taken here is more general and gives more insight into the nature of the problem. The limit $N \to \infty$ is considered in Section 4.2.2.

The decision made on the i^{th} bit of the k^{th} user at the output of the detector $\mathbf{v}$ is:

$$\hat{b}_k(i) \;=\; \mathrm{sgn}\,\mathbf{v}^T\mathbf{y} \;=\; \mathrm{sgn}\,\mathbf{v}^T(\,\mathcal{R}\mathcal{W}\mathbf{b} \;+\; \mathbf{n}\,)\,. \qquad (4.20)$$

As for the inner product, for all $\mathbf{x}$, $\mathbf{y}$ in L

$$<\tilde{S}(t,\mathbf{x})\,,\;\tilde{S}(t,\mathbf{y})> \;=\; \mathbf{x}^T\,\mathcal{R}\,\mathbf{y} \;=<\mathbf{x}\,,\,\mathbf{y}>_{\mathcal{R}}\,. \qquad (4.21)$$

It can be seen from (4.21) and from (4.3) that $\mathcal{R}$ is positive definite.

Definition : We will refer to the $(k,i)^{th}$ row (or column) of a matrix of the dimension of $\mathcal{R}$ when we want to name the k^{th} row (or column) within the i^{th} block in vertical (horizontal) direction.

4.1.2 Near-far resistance

We first assume N to be finite, as is the case in all communication environments, and prove existence of a linear filter which achieves the near-far resistance of optimum multiuser detection. This filter is nonstationary for finite N. The limit $N \to \infty$ is then considered, yielding a stationary noncausal limiting filter, and hence, after appropriate truncation of the noncausal part, an approximation of the near-far optimal linear filter which can be implemented easily.

As shown in [Ver 86b] the asymptotic efficiency of the optimal multiuser detector η is

$$\eta_{k,i} = \frac{1}{w_k(i)} \min_{\epsilon \in Z_k} \|S(t,\epsilon)\|^2 \tag{4.22}$$

$$= \frac{1}{w_k(i)} \min_{\epsilon \in Z_k} \|\tilde{S}(t,\mathbf{w}\epsilon)\|^2 \tag{4.23}$$

where Z_k is the set of error-sequences $\epsilon = \{\epsilon(i) \in \{-1,0,1\}^K, i = -M, ..., M,\ \epsilon_k(i) = 1\}$ that affect the i^{th} bit of the k^{th} user. The NP-completeness of this problem for a positive definite matrix $\mathbf{R}$ was established in Section 3.2.2.

In an environment where the transmission energies change in time, e.g. if the transmitters are mobile, a performance measure of interest for any detector is its k^{th} *user near-far resistance*, $\overline{\eta_{k,i}}$, which is defined for each detector as its worst-case asymptotic efficiency for bit i of User k over all possible energies of the other (interfering and non-interfering) bits, i.e.

$$\overline{\eta_{k,i}} = \inf_{\substack{w_j(l) \geq 0 \\ (j,l) \neq (k,i)}} \eta_{k,i} \,. \tag{4.24}$$

In our definition of near-far resistance we model the most general case, where the energies of the users are allowed to be time-dependent. This captures the worst-case operating conditions of the detector, which are, for example, encountered in mobile radio communication, due to positioning and tracking variations. In the case where the energies are constrained to be arbitrary but non-varying the present near-far resistance is a lower bound. That case is not amenable to closed-form analysis, since one has to deal with a combinatorial optimization problem.

For illustration consider the two-user case. If the user energies are constant over time, i.e. $w_1(i) = w_1$, $w_2(i) = w_2$, the asymptotic efficiency of the optimal multiuser detector given by (4.23) reduces to [Ver 86b]:

$$\eta_1 \;=\; \min\left\{ 1,\; 1 + \frac{w_2}{w_1} - 2\max\{|\rho_{12}|, |\rho_{21}|\}\frac{\sqrt{w_2}}{\sqrt{w_1}},\; 1 + 2\frac{w_2}{w_1} - 2(|\rho_{12}| + |\rho_{21}|)\frac{\sqrt{w_2}}{\sqrt{w_1}} \right\}$$

and hence

$$\eta_{\min} \;\stackrel{\triangle}{=}\; \min_{\substack{w_2 \\ w_1\ const.}} \eta_1 \;=\; \min\left\{ 1 - \rho_{12}^2,\; 1 - \rho_{21}^2,\; 1 - \rho_{12}^2 - \rho_{21}^2 + \frac{(|\rho_{12}| - |\rho_{21}|)^2}{2} \right\}, \qquad (4.25)$$

and analogously for User 2, where $\rho_{12} = R_{12}(0)$ and $\rho_{21} = R_{12}(1)$. The dependence of η_1 for constant energies on the energy ratio is shown in Figure 16.

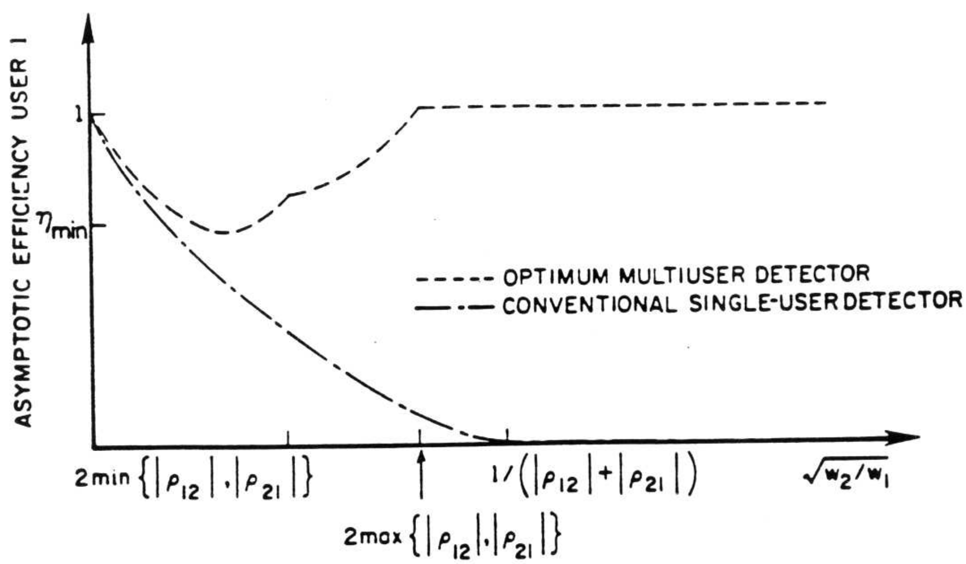


Fig. 16. Asymptotic efficiencies in the 2-user case for infinite transmitted sequence length, when the user energies are constant over time (here we chose $|\rho_{12}|, |\rho_{21}|$=0.3,0.5).

Note that the optimal multiuser detector is near-far resistant, and in fact has an asymptotic efficiency of unity for sufficiently powerful interference ([Ver 86b]). Intuitively this is because in that case the interferers can be "perfectly" recovered and their contributions can be subtracted from the decision statistic. Note also that in this case three different error-sequences minimize (4.23) for different values of w_2/w_1, as can be seen from the discontinuity points of the derivative of η. The minimum of η over constant energies, η_{min}, is an upper bound on the near-far resistance of optimum multiuser detection $\overline{\eta}$, which is the minimum asymptotic efficiency over unconstrained energies.

The near-far resistance of the optimal multi-user detector is important since it is the least upper bound on the near-far resistance of any detector, and a measure of the relative performance of any suboptimal detector. From (4.23) and the definition of near-far resistance it is equal to

$$\overline{\eta_{k,i}} = \inf_{\substack{w_j(l)\geq 0 \\ w_k(i)\ const.}} \frac{1}{w_k(i)} \min_{\epsilon \in Z_k} \|\tilde{S}(t, \mathbf{w}\epsilon)\|^2 \tag{4.26}$$

$$= \inf_{\substack{w_j(l)\geq 0 \\ w_k(i)\ const.}} \min_{\epsilon \in Z_k} \| \tilde{S}(t, \frac{1}{\sqrt{w_k(i)}}\mathbf{w}\epsilon)\|^2 \tag{4.27}$$

$$= \inf_{\substack{\mathbf{y}\in L \\ y_k(i)=1}} \|\tilde{S}(t, \mathbf{y})\|^2 \; . \tag{4.28}$$

In Section 4.2 a closed form expression for (4.28) is obtained as the reciprocal of the $(k,i)^{th}$ diagonal element of the inverse of $\mathcal{R}$ (see footnote 7). Hence, though nonobvious because an *inf* and not a *min* is considered, the near-far resistance of optimum multiuser resistance is guaranteed to be nonzero because of the linear independence assumption of (4.3), which ensures that $\mathcal{R}$ is invertible.

We now turn to the performance analysis of the linear detectors introduced above. The probability of error at decision upon $b_k(i)$ of the linear detector $\mathbf{v}$ is, from (4.10):

$$P_k(i) = P(\hat{b}_k(i) \neq b_k(i)) \tag{4.29}$$

$$= P(< \tilde{S}(t, \mathbf{wb})\, , \, \tilde{S}(t, \mathbf{v}) > + \, n_{k,i} < 0 \,|\, b_k(i) = 1) \; . \tag{4.30}$$

The equality follows since the hypotheses $+1, -1$ are assumed equally likely. Let B be the set of possible transmitted sequences. From (4.8) $n_{k,i}$ is a zero-mean Gaussian random variable with variance $\sigma^2\|\tilde{S}(t,\, \mathbf{v})\|^2$, hence the probability of error in (4.30) is a sum of Q-functions, one for each possible interfering bit-combination. For $\sigma \to 0$ the Q-function with the smallest argument dominates the error probability, hence from (2.1), since the expression below is shown to be upper bounded by 1 via the synchronous equivalent (Section 4.1.1) and the proof in the synchronous case (Section 3.3), the asymptotic efficiency achieved by the linear detector $\mathbf{v}$ for the i^{th} bit of the k^{th} user is

$$\eta_{k,i}(\mathbf{v}) = \frac{1}{w_k(i)} \max{}^2 \left\{ 0, \min_{\substack{\mathbf{b} \in B \\ b_k(i)=1}} \frac{< \tilde{S}(t, \mathbf{wb}), \tilde{S}(t, \mathbf{v}) >}{\|\tilde{S}(t, \mathbf{v})\|} \right\} . \tag{4.31}$$

Knowledge of the asymptotic efficiency of a linear detector is equivalent to knowledge of the worst-case probability of error over the bit sequences of the interfering users, since this error probability, which is a Q-function, is set equal to $Q(\sqrt{\eta_{k,i}(\mathbf{v}) w_k(i)}/\sigma)$ to obtain (4.31). Note that when we say worst-case error probability in this chapter, we mean with respect to interfering bit sequences or, when explicitly stated, with respect to energies, and not also with respect to delays and phases, as e.g. in [Pur 82].

For illustration consider the conventional single-user detector in the two-user case. We have $\mathbf{v} = \mathbf{u}^{k,i}$ (recall that $\mathbf{u}^{k,i}$ is the $(k,i)^{th}$ unit vector in the space L of linear detectors). If the user energies are constant over time, i.e. $w_1(i) = w_1$, $w_2(i) = w_2$, the asymptotic efficiency of the conventional single-user detector is found from (4.31) to be:

$$\eta_1^c = max^2 \left\{ 0, 1 - (|\rho_{12}| + |\rho_{21}|) \frac{\sqrt{w_2}}{\sqrt{w_1}} \right\} , \tag{4.32}$$

and analogously for User 2, where $\rho_{12} = \mathbf{R}_{12}(0)$ and $\rho_{21} = \mathbf{R}_{12}(1)$. The dependence of η_1^c for constant energies on the energy ratio is shown in Figure 16. Note that the asymptotic efficiency of the conventional single-user detector is zero for sufficiently high interference energy ($\sqrt{w_2}/\sqrt{w_1} > 1/(|\rho_{12}| + |\rho_{21}|)$). This implies that its near-far resistance is zero, which is what we want to remedy.

There are three quantities of interest in this communication environment. They are the transmitted bit-sequence, the set of energies (these depend only on the transmitters and determine the *operating points* for the receiver) and the data-processing linear detector $\mathbf{v}$. In determining what linear detector to choose at the receiver a useful procedure is the **minimax** approach, in which the design goal is to optimize the worst-case performance of the receiver over the class of operating points. Thus we are interested in finding the *maximin linear detector*, whose worst-case performance over all allowable input sequences is the highest in the class of linear detectors. The following result quantifies the performance of the maximin detector, in the sequel denoted by $\mathbf{v}^*$.

Proposition 4.1 : There exists a linear detector (which is independent of the received energies) that achieves optimum near-far resistance. $\diamondsuit$

In other words, there is a linear detector which achieves the near-far resistance of the optimum multiuser detector.

Proof: From (3.10) the asymptotic efficiency of the linear detector $\mathbf{v}$ is

$$\begin{aligned}
\eta_{k,i}(\mathbf{v}) &= \frac{1}{w_k(i)} \max{}^2 \left\{ 0, \min_{\substack{\mathbf{b}\in B \\ b_k(i)=1}} \frac{< \tilde{S}(t,\mathbf{wb}), \tilde{S}(t,\mathbf{v}) >}{\|\tilde{S}(t,\mathbf{v})\|} \right\} && (4.33) \\
&= \min_{\substack{\mathbf{b}\in B \\ b_k(i)=1}} \frac{1}{w_k(i)} \max{}^2 \left\{ 0, \frac{< \tilde{S}(t,\mathbf{wb}), \tilde{S}(t,\mathbf{v}) >}{\|\tilde{S}(t,\mathbf{v})\|} \right\} && (4.34) \\
&= \min_{\substack{\mathbf{b}\in B \\ b_k(i)=1}} \frac{1}{w_k(i)} \max{}^2 \left\{ 0, \frac{\mathbf{b}^T \mathcal{W}\mathcal{R}\mathbf{v}}{\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}} \right\} , && (4.35)
\end{aligned}$$

where in the last equality we have used the compact matrix notation of (4.21) for simplicity. We are interested in the energy-independent linear detector with the highest worst-case asymptotic efficiency, i.e. whose near-far resistance is

$$\begin{aligned}
\overline{\eta_{k,i}(\mathbf{v}^*)} &= \sup_{\substack{\mathbf{v}\in L \\ \|\tilde{S}(t,\mathbf{v})\|\neq 0}} \inf_{\substack{w_j(l)\geq 0 \\ w_k(i)\ const.}} \eta_{k,i}(\mathbf{v}) && (4.36) \\
&= \sup_{\substack{\mathbf{v}\in L \\ \mathbf{v}^T\mathcal{R}\mathbf{v}\neq 0}} \underbrace{\inf_{\substack{w_j(l)\geq 0 \\ w_k(i)\ const.}} \min_{\substack{\mathbf{b}\in B \\ b_k(i)=1}}} \frac{1}{w_k(i)} \max{}^2 \left\{ 0, \frac{\mathbf{b}^T \mathcal{W}\mathcal{R}\mathbf{v}}{\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}} \right\} && (4.37) \\
&= \sup_{\substack{\mathbf{v}\in L \\ \mathbf{v}^T\mathcal{R}\mathbf{v}\neq 0}} \inf_{\substack{\mathbf{y}\in L \\ y_k(i)=1}} \max{}^2 \left\{ 0, \frac{\mathbf{y}^T\mathcal{R}\mathbf{v}}{\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}} \right\} && (4.38) \\
&= \max{}^2 \left\{ 0, \sup_{\substack{\mathbf{v}\in L \\ \mathbf{v}^T\mathcal{R}\mathbf{v}\neq 0}} \inf_{\substack{\mathbf{y}\in L \\ y_k(i)=1}} \frac{\mathbf{y}^T\mathcal{R}\mathbf{v}}{\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}} \right\} , && (4.39)
\end{aligned}$$

where we have set $y_j(l) = b_j(l)\sqrt{w_j(l)}/\sqrt{w_k(i)}$ for the third equality. Let $M(\mathbf{v},\mathbf{y})$ denote the penalty function $\mathbf{y}^T\mathcal{R}\mathbf{v}/\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}$, where the first argument is from the set H of detectors and the second from the set Q of operating points. We now show that $M(\mathbf{v},\mathbf{y})$ has a saddle point, i.e.

$$\sup_{\substack{\mathbf{v}\in L\\ \mathbf{v}^T\mathcal{R}\mathbf{v}\neq 0}} \inf_{\substack{\mathbf{y}\in L\\ y_k(i)=1}} \frac{\mathbf{y}^T\mathcal{R}\mathbf{v}}{\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}} = \inf_{\substack{\mathbf{y}\in L\\ y_k(i)=1}} \sup_{\substack{\mathbf{v}\in L\\ \mathbf{v}^T\mathcal{R}\mathbf{v}\neq 0}} \frac{\mathbf{y}^T\mathcal{R}\mathbf{v}}{\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}}\,, \tag{4.40}$$

which means that the sequence of *sup* and *inf* in (4.39) can be interchanged.

Though the penalty function of (4.39) looks similar to the signal-to-noise ratio functional encountered in the robust matched filtering problem [Ver 83a], $| < h, s > |^2 / < h, \Sigma h >$, the problem is different here because the numerator can be negative. Thus we have to establish the result "from scratch". In order to show that $M(\mathbf{v}, \mathbf{y})$ has a saddle point, i.e. satisfies (4.40), we show that it satisfies the requirements of the following theorem:

Theorem 4.1 [Ver 84a, Thm. 2.1]: Suppose Q is a convex set and $M(\mathbf{v}, \cdot)$ is convex on Q for every $\mathbf{v} \in H$. Then if $(\mathbf{v}_L, \mathbf{y}_L)$ is a *regular pair*(7) for (H, Q, M), the following are equivalent:

a) $\mathbf{y}_L \in arg \min_{\mathbf{y}\in Q} \sup_{\mathbf{v}\in H} M(\mathbf{v}, \mathbf{y})$,

b) $(\mathbf{v}_L, \mathbf{y}_L)$ is a saddle point solution for (H, Q, M).

This theorem establishes that if we exhibit a regular pair whose second argument satisfies a), the game (H,Q,M) has a saddle point, which means that the sequence of *max* and *min* in (4.39) can be interchanged. In the following we find a suitable regular pair, thereby proving (4.40).

Clearly the convexity conditions are satisfied. We need to find a candidate regular pair. Note that the value of the *inf* term in (4.39) is $-\infty$ (which gives a near-far resistance of zero) unless $\mathbf{v}$ is picked such that $\mathcal{R}\mathbf{v} = \mathbf{u}^{k,i}$($\eta$ is invariant with respect to scaling of $\mathbf{v}$). $\mathbf{u}^{k,i}$ is the $(k,i)^{th}$ unit vector in the Hilbert space L, defined as $u_j^{k,i}(l) = \delta_{kj}\delta_{li}$. This gives us a candidate for an optimal detector $\mathbf{v}_L$: $\mathbf{d}$, with $\mathcal{R}\mathbf{d} = \mathbf{u}^{k,i}$. (If this detector is indeed optimal, which follows if the candidate pair is regular and satisfies a), it coincides with $\mathbf{v}^*$).

Definition: A **decorrelating detector** $\mathbf{d}^{i,k}$ for the i^{th} bit of the k^{th} user is an element of L for which the following relation holds:

$$\forall\, \mathbf{x} \in L \;\; : \;\; < \tilde{S}(t, \mathbf{d}^{i,k}),\; \tilde{S}(t, \mathbf{x}) > \; = \; x_k(i)\,. \tag{4.41}$$

(7) $(\mathbf{v}_L, \mathbf{y}_L) \in H \times Q$ is a regular pair for (H, Q, M) if, for every $\mathbf{y} \in Q$ such that $\mathbf{y}_\alpha = (1-\alpha)\mathbf{y}_L + \alpha\mathbf{y} \in Q$ for $\alpha \in [0,1]$, we have

$$\sup_{\mathbf{v}\in H} M(\mathbf{v}, \mathbf{y}_\alpha) - M(\mathbf{v}_L, \mathbf{y}_\alpha) = o(\alpha).$$

Equivalently, $\mathcal{R}\mathbf{d} = \mathbf{u}^{i,k}$ and the candidate $\mathbf{v}_L = \mathbf{d}$ (superscripts are omitted). The explanation of the name of "decorrelating" is similar to the one given in the synchronous case in Section 3.3. From (4.10), using $\mathbf{v}^{i,k} = \mathbf{d}^{i,k}$, the detector decides for $sgn(\sqrt{w_k(i)}\, b_k(i) + n_k(i)$, i.e. $b_k(i)$ has been decorrelated. As to existence of the postulated filter, we show in Section 4.2., (4.63), that a decorrelating detector exists for any set of transmitter signals and delays for which the LIA is satisfied.

Next we find a $\mathbf{y}_L$ which meets the requirement of point a) of Theorem 4.1. Using the Cauchy-Schwarz inequality, we find that

$$\sup_{\mathbf{v}\in H} M(\mathbf{v},\mathbf{y}) = \sup_{\substack{\mathbf{v}\in L \\ \mathbf{v}^T\mathcal{R}\mathbf{v}\neq 0}} \frac{\mathbf{y}^T\mathcal{R}\mathbf{v}}{\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}} = \sqrt{\mathbf{y}^T\mathcal{R}\mathbf{y}}\,, \tag{4.42}$$

where the inner product is maximized for $\mathbf{v} = k\mathbf{y} + \{\mathbf{x} \in L \,:\, \mathcal{R}\mathbf{x} = 0\}$.

We now need to solve the Hilbert space optimization problem

$$\inf \quad \mathbf{y}^T\mathcal{R}\mathbf{y} \tag{4.43}$$

$$\text{subject to} \quad y_k(i) = 1\,.$$

Using (4.21) and the definition of $\mathbf{d}$ we can rewrite the minimization problem under consideration as:

$$\inf \quad \|\mathbf{y}\|_R \tag{4.44}$$

$$\text{subject to} \quad <\mathbf{d},\mathbf{y}>_R = 1\,.$$

$\|\cdot\|_R$ is a norm since $\mathcal{R}$ is positive definite. We have obtained a minimum-norm optimization problem in Hilbert space. To prove existence of a solution we need to show the constraint set to be closed, which holds since the Hilbert space is finite dimensional. (Even for $N \to \infty$, when we have an infinite dimensional optimization problem, we could use the fact that the codimension is finite. The problem there is that the signals are no longer square integrable.) The constraint, $\mathbf{y}_k(i) = 1$, is equivalent to $\mathbf{y} = \mathbf{u}^{k,i} + \{\mathbf{x} \,:<\mathbf{x},\mathbf{d}>_R = 0\}$. $\mathcal{A} = [\,\mathbf{d}\,]$, the subspace generated by $\mathbf{d}$, is a closed subspace of dimension 1. Hence the constraint set $\{\mathbf{x} \,:<\mathbf{x},\mathbf{d}>_R = 0\} = \mathcal{A}^\perp$ is closed. We now have a minimum-norm optimization problem in Hilbert space over a closed subspace. Hence the *Projection Theorem*, [Lue], guarantees existence (so we can replace the *inf* by a *min*, as required in a)) and uniqueness of a minimizing equivalence class $\mathbf{y}^*$, with

$$\mathbf{y}^* \in \{\mathcal{A}^\perp + \mathbf{u}^{k,i}\} \cap \mathcal{A}^{\perp\perp} = \{\mathcal{A}^\perp + \mathbf{u}^{k,i}\} \cap \mathcal{A}\,, \tag{4.45}$$

where equality holds since A is closed. Hence $y_k^*(i) = 1$ and $\mathbf{y}^* = k\ \mathbf{d}$, which implies

$$\mathbf{y}^* = \frac{1}{d_k(i)}\ \mathbf{d}\ . \tag{4.46}$$

We now have a candidate regular pair which satisfies a): $(\mathbf{v}_L, \mathbf{y}_L) = (\mathbf{d},\ (d_k(i))^{-1}\mathbf{d})$. From (4.42) and the definition of regularity we have to check the dependence on α of

$$\sqrt{\mathbf{y}_\alpha^T \mathcal{R} \mathbf{y}_\alpha} - \frac{\mathbf{y}_\alpha^T \mathcal{R} \mathbf{v}_L}{\sqrt{\mathbf{v}_L^T \mathcal{R} \mathbf{v}_L}}$$

$$= \sqrt{\mathbf{d}^T\mathcal{R}\mathbf{d} + 2\alpha(\mathbf{y}-\mathbf{d})^T\mathcal{R}\mathbf{d} + \alpha^2(\mathbf{y}-\mathbf{d})^T\mathcal{R}(\mathbf{y}-\mathbf{d})} - \sqrt{\frac{1}{d_k(i)}}$$

$$= \sqrt{\frac{1}{d_k(i)} + \alpha^2(\mathbf{y}-\mathbf{d})^T\mathcal{R}(\mathbf{y}-\mathbf{d})} - \sqrt{\frac{1}{d_k(i)}}\ . \tag{4.47}$$

We have repeatedly used the decorrelating property of $\mathbf{d}$. Since $\sqrt{1+x} \leq 1 + 1/2\ x$, the above quantity lies in the interval $[0\ ,\ (\mathbf{y}-\mathbf{d})^T\mathcal{R}(\mathbf{y}-\mathbf{d})\sqrt{d_k(i)}/2\ \ \alpha^2]$, hence divided by α goes to 0 when $\alpha \downarrow 0$. Thus $(\mathbf{d},\ (d_k(i))^{-1}\mathbf{d})$ is a regular pair which satisfies point a) of the theorem.

Hence it follows from the theorem that the penalty function $\mathbf{y}^T\mathcal{R}\mathbf{v}/\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}$ has a saddle point, i.e.:

$$\sup_{\substack{\mathbf{v}\in L \\ \mathbf{v}^T\mathcal{R}\mathbf{v}\neq 1}} \inf_{\substack{\mathbf{y}\in L \\ y_k(i)=1}} \frac{\mathbf{y}^T\mathcal{R}\mathbf{v}}{\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}} = \inf_{\substack{\mathbf{y}\in L \\ y_k(i)=1}} \sup_{\substack{\mathbf{v}\in L \\ \mathbf{v}^T\mathcal{R}\mathbf{v}\neq 1}} \frac{\mathbf{y}^T\mathcal{R}\mathbf{v}}{\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}}\ . \tag{4.48}$$

which establishes existence of $\mathbf{v}^*$ and hence

$$\overline{\eta_{k,i}(\mathbf{v}^*)} = \max{}^2\ \{0,\ \inf_{\substack{\mathbf{y}\in L \\ y_k(i)=1}} \sup_{\substack{\mathbf{v}\in L \\ \mathbf{v}^T\mathcal{R}\mathbf{v}\neq 0}} \frac{\mathbf{y}^T\mathcal{R}\mathbf{v}}{\sqrt{\mathbf{v}^T\mathcal{R}\mathbf{v}}}\} \tag{4.49}$$

$$= \max{}^2\ \{0,\ \inf_{\substack{\mathbf{y}\in L \\ y_k(i)=1}} \sqrt{\mathbf{y}^T\mathcal{R}\mathbf{y}}\} \tag{4.50}$$

$$= \inf_{\substack{\mathbf{y}\in L \\ y_k(i)=1}} \|\tilde{S}(t,\mathbf{y})\|^2 \tag{4.51}$$

$$= \overline{\eta_{k,i}}\ , \tag{4.52}$$

where the second equality is obtained in (4.42), the third line follows since $\mathcal{R}$ is nonnegative definite and the last equality was obtained in (4.28).

We have proved that there is a linear detector which achieves the near-far resistance of optimum multiuser detection.

$\diamond$

The reason why the near-far optimum linear receiver achieves the same near-far resistance as the optimum receiver can be understood as follows. Let Ω be the set of multiuser signals modulated by all possible amplitudes, i.e. $\Omega = \{\tilde{S}(t, \mathbf{y}), \mathbf{y} \in L\}$ and let S denote the subset of Ω such that the amplitude of the i^{th} symbol of the k^{th} user is fixed to 1, i.e. $S = \{\tilde{S}(t, \mathbf{y}), \mathbf{y} \in L, y_k(i) = 1\}$ (note that S is a convex set, and because of the LIA it does not include the origin). Since the penalty function in (4.39) is invariant to scaling of $\mathbf{v}$ and the operator $\mathcal{R}$ is positive definite, (4.39) can be rewritten as

$$\overline{\eta_{k,i}(\mathbf{v}^*)} = \max{}^2 \left\{ 0, \sup_{\substack{\mathbf{v}\in L \\ \|\tilde{S}(t,\mathbf{v})\|=1}} \inf_{\substack{\mathbf{y}\in L \\ y_k(i)=1}} < \tilde{S}(t, \mathbf{y}), \tilde{S}(t, \mathbf{v}) > \right\} \tag{4.53}$$

$$= \max{}^2 \left\{ 0, \sup_{\substack{v\in\Omega \\ \|v\|=1}} \inf_{y\in S} < y, v > \right\} . \tag{4.54}$$

Therefore the k^{th} user decorrelating detector corresponds to the unit-norm multiuser waveform whose minimum inner product with the elements of S is highest. But since S is a convex set, that signal is a scaled version of the closest vector in S to the origin (Figure 17), and its near-far resistance (c.f. (4.51)) is the norm squared of that vector. But, as (4.28) indicates, the square of the distance from S to the origin is precisely the near-far resistance of the optimum detector.

Equation (4.28) leads to a nice intuitive interpretation of near-far resistance. Rewrite this equation, using the definition of $\tilde{S}(t, .)$, as

$$\overline{\eta_{k,i}} = \inf_{\substack{y_j(l)\in R \\ (j,l)\neq(k,i)}} \| s_k(t - iT - \tau_k) + \sum_{(j,l)\neq(k,i)} y_j(l)\, s_j(t - lT - \tau_j) \|^2 . \tag{4.55}$$

Letting $\{y_j(l)\}$ vary over the admissible set, the second term above generates all points of a linear subspace which includes the origin, therefore the infimum in (3.26) is the distance of $\tilde{s}_k(t - iT - \tau_k)$ to this space, i.e.

$$\overline{\eta_{k,i}} = d^2 \; (\, \tilde{s}_k(t - iT - \tau_k) \, , \, \text{span} \, \{ \, \tilde{s}_j(t - lT - \tau_j), (j, l) \neq (k, i) \, \} \,) \, , \tag{4.56}$$

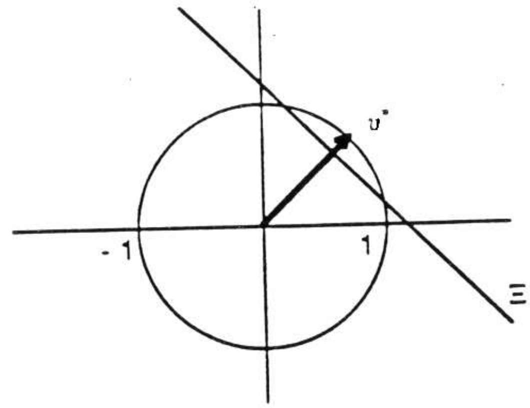


Fig. 17. Interpretation of near-far resistance. Vector in boldface corresponds to decorrelating filter.

where $d(a, b)$ denotes the Euclidean distance between the $\mathcal{L}_2$ elements a and b. In the synchronous case, because the time-support is disjoint, the infimum in (3.26) is achieved when $y_j(l) = 0,\ l \neq i$, and (3.27) reduces to

$$\overline{\eta_k} \; = \; d^2 \; (\tilde{s}_k(t) \; , \; \text{span} \; \{ \; \tilde{s}_j(t), \; j \neq k \; \} \;) \; , \tag{4.57}$$

i.e., the k^{th} user near-far resistance in a synchronous channel is the square of the *innovation of the k^{th} user signal with respect to the space spanned by the signals of the interfering users.* Viewing the asynchronous problem in terms of the equivalent synchronous system with $N * K$ users and period NT, the near-far resistance of asynchronous communication allows for the same interpretation. Note, however, that the shifted versions $s_k(t - lT - \tau_k), i \neq l$ of the k^{th} user signal affect the near-far resistance of the i^{th} symbol of User k.

The following section characterizes a linear detector that achieves the optimum near-far resistance anticipated by Proposition 4.1.

4.2 The decorrelating detector

4.2.1 The finite sequence length case

Definition: A *decorrelating detector* $\mathbf{d}^{k,i}$ for the i^{th} bit of the k^{th} user is a linear detector for which

$$\mathcal{R}\mathbf{d}^{k,i} = \mathbf{u}^{k,i} \tag{4.58}$$

or equivalently, from (4.21), $< \tilde{S}(t, \mathbf{d}^{k,i}),\ \tilde{S}(t, \mathbf{x}) > = x_k(i)$, for all $\mathbf{x}$ in L (cf. (4.41)).

Existence: By the LIA, statement (4.59) below holds for all k, i. Hence the following equivalences show the existence of the decorrelating detectors for each bit of each user.

$$\forall\, \mathbf{x} \in L \text{ with } x_k(i) \neq 0 : \ \|\tilde{S}(t, \mathbf{x})\| \neq 0 \tag{4.59}$$

$$\Longleftrightarrow \quad \forall\, \mathbf{x} \in L \text{ with } x_k(i) \neq 0 : \ \mathbf{x}^T \mathcal{R}\mathbf{x} \neq 0 \tag{4.60}$$

$$\Longleftrightarrow \quad \nexists\, \mathbf{x} \in L \text{ with } x_k(i) \neq 0 \text{ s.t. } \mathcal{R}\mathbf{x} = 0 \tag{4.61}$$

$$\Longleftrightarrow \quad \text{the } (k,i)^{th} \text{ column}^{(7)} \text{ of } \mathcal{R} \text{ is linearly independent of the others} \tag{4.62}$$

$$\Longleftrightarrow \quad \exists\, \mathbf{d} \text{ s.t. } \mathcal{R}\mathbf{d} = \mathbf{u}^{k,i} \tag{4.63}$$

Properties:

i) The decorrelating detector for each bit of each user is invariant with respect to received energies and does not require knowledge thereof.

> *Proof:* Since the elements of the matrix $\mathcal{R}$ are normalized crosscorrelation coefficients, the defining equation (4.58) is energy independent.

ii) The decorrelating detector eliminates the multiuser interference present in the respective matched filter output. (Hence its name).

---

(7) We refer to the $(k,i)^{th}$ row (or column) of a matrix of the dimension of $\mathcal{R}$ when we want to name the k^{th} row (or column) within the i^{th} block in vertical (horizontal) direction. This notation was adopted since $\mathcal{R}$ is block-Toeplitz.

Proof: From (4.20) the decision made on the i^{th} bit of the k^{th} user at the output of the decorrelating filter **d** is,

$$\begin{aligned} \hat{b}_k(i) &= \text{sgn}\ \left(\mathbf{d}^T \mathcal{R}\mathcal{W}\mathbf{b} + \mathbf{d}^T\mathbf{n}\right) \\ &= \text{sgn}\ \left(\sqrt{w_k(i)}\, b_k(i) + \mathbf{d}^T\mathbf{n}\right). \end{aligned} \tag{4.64}$$

Interestingly, this natural strategy, though not necessarily optimal for specific user-energies, is optimal with respect to the worst possible distribution of energies. The *sgn* decision in (4.64) is based on the assumption that the respective user is known to be active.

iii) The k^{th}-user bit-error-rate of the decorrelating detector is independent of the energies of the interfering users $w_j(i), j \neq k, i = -M, ..., M$.

Proof: It follows from (4.64) that the decision statistic that is compared to a zero threshold is independent of the energies of the interfering users.

iv) The efficiency of the decorrelating detector is independent of the energies and is given by

$$\eta^d_{k,i} = \max{}^2 \left\{ 0,\ \min_{\substack{\mathbf{b}\in B \\ b_k(i)=1}} \frac{1}{\sqrt{w_k(i)}} \frac{< \tilde{S}(t, \mathbf{W}\mathbf{b}), \tilde{S}(t, \mathbf{d}) >}{\|\tilde{S}(t, \mathbf{d})\|} \right\} \tag{4.65}$$

$$= \max{}^2 \left\{ 0,\ \min_{\substack{\mathbf{b}\in B \\ b_k(i)=1}} \frac{1}{\sqrt{w_k(i)}} \frac{\sqrt{w_k(i)}\, b_k(i)}{\sqrt{d_k(i)}} \right\} \tag{4.66}$$

$$= \frac{1}{d_k(i)}, \tag{4.67}$$

which by i) is energy-independent.

v) The decorrelating detector achieves the highest near-far resistance of any linear detector. In addition, it achieves the near-far resistance of optimum multiuser detection.

Proof: The proof of Proposition 4.1 is constructive, hence the first part of v) was obtained as a byproduct during the proof. Here is a shorter proof, using the following fact: Any single (i.e., energy-independent) linear strategy which is not decorrelating has a near-far resistance of zero. This is shown as follows: The near-far resistance of an energy-independent linear filter is (cf. (4.39)):

$$\overline{\eta_{k,i}(\mathbf{v})} = \max{}^2 \left\{ 0,\ \inf_{\substack{\mathbf{y}\in L \\ y_k(i)=1}} \frac{\mathbf{y}^T \mathcal{R}\mathbf{v}}{\sqrt{\mathbf{v}^T \mathcal{R}\mathbf{v}}} \right\}. \tag{4.68}$$

Unless $\mathcal{R}\mathbf{v} = \mathbf{u}^{k,i}$ (note invariance of η to scaling of $\mathbf{v}$) the value of the inf-term is $-\infty$. Hence any linear filter which is not decorrelating has a near-far resistance $\overline{\eta} = 0$. This fact together with the nonzero asymptotic efficiency (4.67) of the decorrelating detector establish optimality of the decorrelating detector within the class of energy-independent linear filters. Therefore the second part of v) results from Proposition 4.1.

Note that since the asymptotic efficiency of the decorrelating detector is independent of energies (Property iv) it equals the near-far resistance. This gives us an explicit solution for the Hilbert space optimization problem we obtained for the near-far resistance of optimal multi-user detection in (4.28), namely

$$\overline{\eta_{k,i}} = \eta^d_{k,i} = \frac{1}{d_k(i)} \tag{4.69}$$

and outlines an alternative proof for Proposition 4.1: one could have explicitly solved the optimization problem (4.28) by proceeding as in (4.43 ff), postulated the decorrelating detector by reasoning as in Fact under v), and shown that the asymptotic efficiency of the decorrelating detector and the near-far resistance of optimal multi-user detection are equal (see [Lup 89a]). However, the game theoretic proof provides more insight into the nature of the solution.

Property iii) is of special importance. By this property the decorrelating detector does not become multiple-access limited, no matter how strong the multiple-access interference is. Also the decorrelating detector demodulates the data perfectly in the absence of noise, as can be seen from (4.64).

Characterization

We would now like to find an explicit expression for the decorrelating detector which we have up to now defined implicitly. It follows immediately from (4.58) and the uniqueness of the inverse of an invertible matrix that the decorrelating detector for the i^{th} bit of User k is the $(k,i)^{th}$ row of the inverse of $\mathcal{R}$.

From the above and (4.67) the asymptotic efficiency of the decorrelating detector for the i^{th} bit of User k is given by the $(k,i)^{th}$ diagonal element of the inverse of $\mathcal{R}$:

$$\eta^d_{k,i} = \frac{1}{\mathcal{R}^{-1}_{(k,i),(k,i)}} \,. \tag{4.70}$$

For the values of N encountered in practical applications, inverting a $NK * NK$ matrix is not possible. This issue is addressed in Section 4.2.2, where we represent the decorrelating detector as a K-input K-output time-varying linear filter, and then show that in the limit as N tends to infinity the filter becomes time-invariant.

4.2.2 The limiting case $N \to \infty$

Proposition 4.2 : As the length of the transmitted sequence increases ($N \to \infty$) the decorrelating detector approaches the K-input K-output linear time-invariant filter with transfer function

$$\mathbf{G}(z) \;=\; [\,\mathbf{R}^T(1)z \;+\; \mathbf{R}(0) \;+\; \mathbf{R}(1)z^{-1}\,]^{-1}\,. \qquad (4.71)$$

◊

Proof : From (4.14) and (4.13) the matched filter outputs for $l = \{-M, ..., M\}$ are

$$\mathbf{y}(l) = \mathbf{R}^T(1)\mathbf{W}(l+1)\;\mathbf{b}(l+1) + \mathbf{R}(0)\mathbf{W}(l)\;\mathbf{b}(l) + \mathbf{R}(1)\mathbf{W}(l-1)\;\mathbf{b}(l-1) + \mathbf{n}(l)\;, \qquad (4.72)$$

where $\mathbf{b}(-M-1) = \mathbf{b}(M+1) = 0$. Taking z-transforms and letting N go to infinity we have:

$$Y(z) \;=\; \mathbf{S}(z)\;[WB](z) \;+\; N(z)\;, \qquad (4.73)$$

where $[WB](z)$ is the z-transform of the sequence $\mathbf{wb} = \{[\sqrt{w_1(i)}b_1(i), ..., \sqrt{w_K(i)}b_K(i)]\}$, the matrix $\mathbf{S}(z)$ is

$$\mathbf{S}(z) \;=\; \mathbf{R}^T(1)z \;+\; \mathbf{R}(0) \;+\; \mathbf{R}(1)z^{-1}\;, \qquad (4.74)$$

and $Y(z)$, $B(z)$ and $N(z)$ are, respectively, the vector-valued z-transforms of the matched filter output sequence, the transmitted sequence, and the noise sequence at the output of the matched filters. $\mathbf{S}(z)$ can be interpreted as the equivalent transfer function of the multiuser communication system between transmitter and decision algorithm, as illustrated in Figure 18.

In this setting the optimal receiver problem is to find the transfer function matrix $\mathbf{G}(z)$ of a K-input K-output linear time-invariant filter, at the output of which a sign-decision yields estimates of the transmitted sequence which are optimal in a certain sense. In our case the optimality criterion is the near-far resistance, and we have demonstrated that the optimal filter is the decorrelating filter, which is the filter which eliminates the multiuser interference, i.e. is the K-input K-output

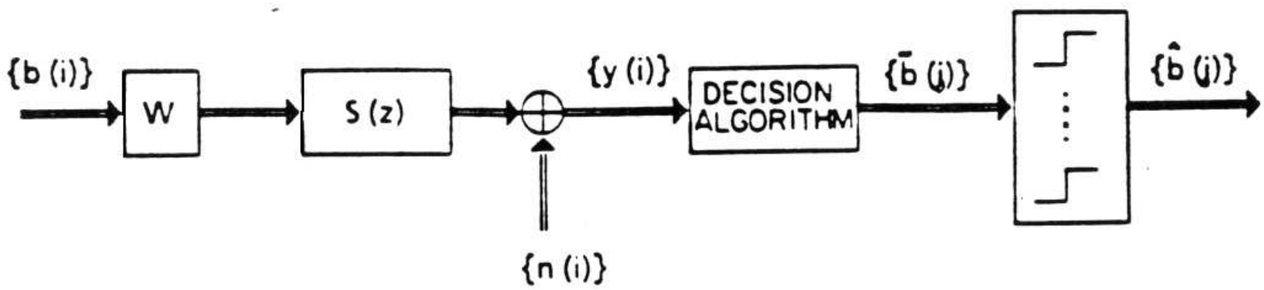


Fig. 18. Equivalent Communication System.

time invariant linear filter which recovers the transmitted data in the absence of noise. Its transfer function is therefore the inverse of the equivalent transfer function $\mathbf{S}(z)$:

$$\mathbf{G}(z) \;=\; [\; \mathbf{S}(z)\;]^{-1}\;. \tag{4.75}$$

◇

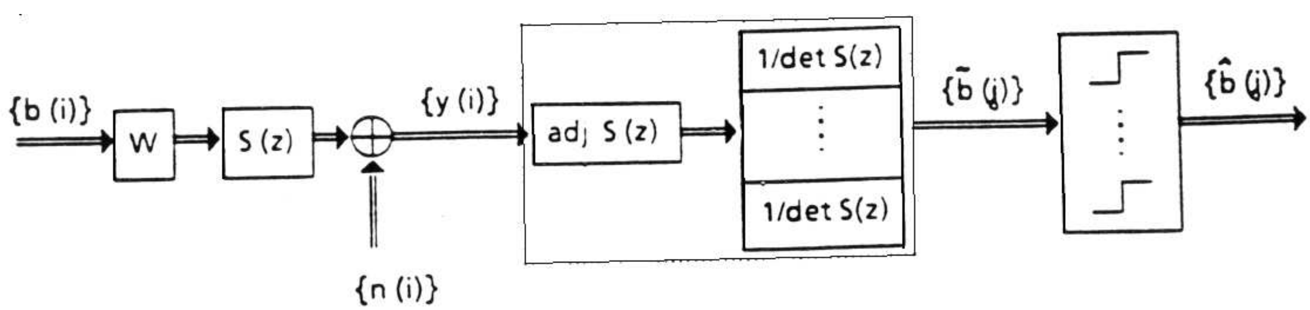


Fig. 19. Interpretation of the Decorrelating Detector.

The effect of the inverse filter $[\; \mathbf{S}(z)\;]^{-1}$ can be interpreted as illustrated in Figure 19. The decorrelating filter can be viewed as the cascade of a finite impulse response filter with transfer function *adjoint* $\mathbf{S}(z)$, which decorrelates the users, but introduces intersymbol interference among the previously noninterfering symbols of the same user, and of a second filter, consisting of a bank of

K identical filters with transfer function $[det\ \mathbf{S}(z)]^{-1}$, which removes this intersymbol interference. Whereas the region of convergence of the z-transform can always be chosen so as to make $\mathbf{S}(z)$ invertible, attention has to be paid to the issue of stability.

Proposition 4.3 : There is a stable, noncausal realization of the decorrelating detector, if and only if the signal cross-correlations are such that

$$\det \mathbf{S}(e^{j\omega}) \;=\; \det\,[\,\mathbf{R}^T(1)e^{j\omega} \;+\; \mathbf{R}(0) \;+\; \mathbf{R}(1)e^{-j\omega}\,] \;\neq\; 0\,, \quad \forall\,\omega \;\in\; [0, 2\pi]\,. \tag{4.76}$$

$\diamondsuit$

Proof : As long as *det* $\mathbf{S}(z)$ has no zeroes on the unit circle, a nonempty convergence region of $\mathbf{S}^{-1}(z)$ can be chosen which includes the unit circle. Thus stability can be achieved. But, since $\mathbf{R}(0)$ is symmetric,

$$\det \mathbf{S}(z) \;=\; \det \mathbf{S}^T(z) \;=\; \det \mathbf{S}(z^{-1})\,.$$

Hence the stable version of the decorrelating detector will be noncausal. ■

In the two-user case condition (4.76) is easily shown to be

$$|\rho_{12}| \;+\; |\rho_{21}| \;<\; 1\,. \tag{4.77}$$

Since $|\rho_{12}| \;+\; |\rho_{21}| \;\leq\; 1$ is always satisfied, condition (4.77) is violated only if the normalized waveforms coincide modulo circular shifts and sign changes.

Condition (4.76) is equivalent to the limit of the LIA as $N \to \infty$. Both are necessary and sufficient conditions for system invertibility. The LIA requires that the output of a system (the system between the user bit-streams and the matched filter outputs) be not identically zero if the input is nonzero. Hence different inputs generate different outputs, i.e. the system is invertible. For a linear system the requirement that nonzero input produce nonzero output is equivalent to requiring that the transfer matrix be nonsingular on the unit circle. Assume the transfer matrix is singular at the angular frequency ω_0. Necessity follows since otherwise the input sequence consisting of a complex exponential at ω_0 times a vector in the nullspace of the transfer matrix evaluated at ω_0 yields zero output, since the transfer function on the unit circle gives the magnitude and phase of the system response to complex exponentials. On the other hand sufficiency can be established by using Parseval's relation extended to multivariable systems:

$$\sum_{n=-\infty}^{\infty} \|\mathbf{y}_n\|^2 = \frac{1}{2\pi}\int_0^{2\pi} \|\mathbf{Y}(e^{j\omega})\|^2 d\omega = \frac{1}{2\pi}\int_0^{2\pi} \|\mathbf{H}(e^{j\omega})\ \mathbf{X}(e^{j\omega})\|^2 d\omega\,. \tag{4.78}$$

Hence for a zero output sequence $\mathbf{y}_n$ the vector $\mathbf{H}(e^{j\omega})\mathbf{X}(e^{j\omega})$ has to vanish for all ω, which implies that $\mathbf{H}(e^{j\omega})$ is singular whenever $\mathbf{X}(e^{j\omega})$ is nonzero. This establishes the claimed equivalence.

The following results quantify the asymptotic efficiency achieved by the limiting decorrelating detector.

Proposition 4.4 : Let

$$[\mathbf{S}\ (z)]^{-1} = \sum_{m=-\infty}^{\infty} \mathbf{D}(m)\ z^{-m} \tag{4.79}$$

Then the asymptotic efficiency of the limiting decorrelating detector for the k^{th} user is given by

$$\eta_k^d = \frac{1}{D_{kk}(0)} \tag{4.80}$$

$$= \left[\frac{1}{2\pi}\int_0^{2\pi} [\ \mathbf{R}^T(1)e^{j\omega} + \mathbf{R}(0) + \mathbf{R}(1)e^{-j\omega}\ ]_{kk}^{-1}\ d\omega \right]^{-1} . \tag{4.81}$$

Proof : With Proposition 4.2 the decision statistic at the output of the limiting decorrelating detector has a z-transform given by

$$\mathbf{G}(z)Y(z) = [WB](z) + [\mathbf{S}(z)]^{-1}N(z) = [WB](z) + N'(z)$$

where $N'(z)$ is the z-transform of the (stationary) filtered Gaussian background noise vector sequence. Its covariance matrix sequence $E\ [\mathbf{n}'(.)\mathbf{n}'^T(. + i)]$ has a z-transform of $\sigma^2[\mathbf{S}(z)]^{-1}$ $(\sigma^2\mathbf{S}^{-1}(z)\mathbf{S}(z)\mathbf{S}^{-1^T}(z^{-1}))$, hence with (4.79) n_k' is a zero-mean Gaussian random variable with variance $\sigma^2 D_{kk}(0)$. Therefore the probability of error for the k^{th} user equals

$$P_k = P(n_k' > \sqrt{w_k}) = Q(\frac{\sqrt{w_k}}{\sigma\sqrt{D_{kk}(0)}})\,. \tag{4.82}$$

From here, using the definition of asymptotic efficiency, the first equality follows. For the second, applying the inverse z-transform and (4.79),

$$D_{kk}(0) = \frac{1}{2\pi}\int_0^{2\pi} [\mathbf{S}\ (e^{j\omega})]_{kk}^{-1}\ d\omega\ ,$$

and the result follows by inserting (4.74) into the above.

◇

Proposition 4.5 : The asymptotic efficiency of the limiting decorrelating detector for the k^{th} user is strictly positive, and lower bounded by

$$\eta_k^d \geq \left[\max_{\omega \in [0,2\pi]} \left| \left[\mathbf{R}^T(1)e^{j\omega} + \mathbf{R}(0) + \mathbf{R}(1)e^{-j\omega} \right]_{kk}^{-1} \right| \right]^{-1} > 0 .$$

Proof : From (4.81)

$$D_{kk}(0) \leq \max_{\omega \in [0,2\pi]} \left| \left[\mathbf{R}^T(1)e^{j\omega} + \mathbf{R}(0) + \mathbf{R}(1)e^{-j\omega} \right]_{kk}^{-1} \right| . \tag{4.83}$$

Hence

$$\eta_k^d = \frac{1}{D_{kk}(0)} \geq \left[\max_{\omega \in [0,2\pi]} \left| \left[\mathbf{R}^T(1)e^{j\omega} + \mathbf{R}(0) + \mathbf{R}(1)e^{-j\omega} \right]_{kk}^{-1} \right| \right]^{-1}$$

$$\geq \frac{\min_\omega \left| \det \left[\mathbf{R}^T(1)e^{j\omega} + \mathbf{R}(0) + \mathbf{R}(1)e^{-j\omega} \right] \right|}{\max_\omega \left| \mathrm{adj}_k \left[\mathbf{R}^T(1)e^{j\omega} + \mathbf{R}(0) + \mathbf{R}(1)e^{-j\omega} \right] \right|} , \tag{4.84}$$

which is positive by Proposition 4.3.

◇

Next we summarize some properties of the matrix encountered in Propositions 4.3, 4.4 and 4.5.

Proposition 4.6 : Let $\mathbf{M}(\varphi) = \mathbf{R}^T(1)\, e^{j\varphi} + \mathbf{R}(0) + \mathbf{R}(1)\, e^{-j\varphi}$. Then

i) $M_{kk}^{-1}(\varphi)$ is real.

ii) $M_{kk}^{-1}(2\pi - \varphi) = M_{kk}^{-1}(\varphi)$

iii) $\mathbf{M}(\varphi)$ is nonnegative definite for all φ. ◇

Corollary :

i) $D^{kk}(0)$ is real.

ii) $D_{kk}(0) = \frac{1}{\pi} \int_0^\pi [\, \mathbf{R}^T(1)e^{j\varphi} + \mathbf{R}(0) + \mathbf{R}(1)e^{-j\varphi} \,]_{kk}^{-1} \; d\varphi$. (4.85)

iii) If one user is added to the system, the asymptotic efficiency of the other users is nonincreasing, and changes according to

$$\frac{1}{\eta_{k|K+1}} = \frac{1}{\eta_{k|K}} + \frac{1}{\pi} \int_0^\pi \frac{\|\mathbf{m}^T(\varphi)\, \mathbf{M}^{-1}(\varphi)\, \mathbf{u}_k\|^2}{1 - \mathbf{m}^T(\varphi)\, \mathbf{M}^{-1}(\varphi)\, \mathbf{m}(\varphi)} \, d\varphi , \tag{4.86}$$

where $[1,\ \mathbf{m}^T(\varphi),\ \mathbf{m}(\varphi),\ \mathbf{M}(\varphi)]$ are the elements of the matrix $\mathbf{M}(\varphi)$ after the additional user has been added to the system. $\diamondsuit$

The proofs are given in Appendix 4.3.

Proposition 4.7: Condition (4.76) of Proposition 4.3 is equivalent to

$$\min_{\substack{\mathbf{x}^*\mathbf{x}=1 \\ \mathbf{x}\in\mathcal{C}}} \left(\mathbf{x}^*\ \mathbf{R}(0)\ \mathbf{x}\ -\ \sqrt{(\mathbf{x}^*\ \mathbf{R}_+\ \mathbf{x})^2\ +\ (\mathbf{x}^*\ \mathbf{R}_-\ \mathbf{x})^2} \right)\ >\ 0\,, \tag{4.87}$$

where $\mathbf{R}_+ = \mathbf{R}^T(1) + \mathbf{R}(1)$ and $\mathbf{R}_- = j(\mathbf{R}^T(1) - \mathbf{R}(1))$. The * denotes the complex conjugate. $\diamondsuit$

Note that both $\mathbf{R}_+$ and $\mathbf{R}_-$ are Hermitian.

Proof: Since from Proposition 4.6 $\mathbf{M}(\varphi) \geq 0$, condition (4.76) is equivalently expressed as

$$\inf_{\varphi}\ \lambda_{\min}\,(\mathbf{M}(\varphi))\ >\ 0\,. \tag{4.88}$$

Since $\mathbf{M}(\varphi)$ is Hermitian, its smallest eigenvalue is [Horn, Thm.4.2.2]

$$\lambda_{\min}\,(\mathbf{M}(\varphi))\ =\ \min_{\substack{\mathbf{x}^*\mathbf{x}=1 \\ \mathbf{x}\in\mathcal{C}}}\ \mathbf{x}^*\ \mathbf{M}(\varphi)\ \mathbf{x}\,. \tag{4.89}$$

Therefore (4.88) becomes

$$\begin{aligned} &\inf_{\varphi}\ \min_{\mathbf{x}^*\mathbf{x}=1}\ \mathbf{x}^*\mathbf{R}(0)\mathbf{x}\ +\ \mathbf{x}^*[\mathbf{R}^T(1)+\mathbf{R}(1)]\mathbf{x}\cos\varphi\ +\ \mathrm{j}\ \mathbf{x}^*[\mathbf{R}^T(1)-\mathbf{R}(1)]\mathbf{x}\sin\varphi \\ &=\ \min_{\mathbf{x}^*\mathbf{x}=1}\ \mathbf{x}^*\ \mathbf{R}(0)\ \mathbf{x}\ -\ \sqrt{[\mathbf{x}^*\ (\mathbf{R}^T(1)+\mathbf{R}(1))\ \mathbf{x}]^2\ +\ [j\ \mathbf{x}^*\ (\mathbf{R}^T(1)-\mathbf{R}(1))\ \mathbf{x}]^2} \end{aligned} \tag{4.90}$$

where we have exchanged the *inf* and *min* operations and used the fact that

$$\inf_{\varphi}\ (a\cos\varphi\ +\ b\sin\varphi)\ =\ -\sqrt{a^2\ +\ b^2}\,. \tag{4.91}$$

$\diamondsuit$

Proposition 4.8 : A necessary condition for Proposition 4.3 is that the matrices $\mathbf{R}(0) + \mathbf{R}(1) + \mathbf{R}^T(1)$ and $\mathbf{R}(0) - \mathbf{R}(1) - \mathbf{R}^T(1)$ be nonsingular. $\diamondsuit$

Proof:

$$\min_{\substack{\mathbf{x}^*\mathbf{x}=1 \\ \mathbf{x}\in\mathbb{C}}} \ \mathbf{x}^* \ \mathbf{R}(0) \ \mathbf{x} \ - \ \sqrt{(\mathbf{x}^* \ \mathbf{R}_+ \ \mathbf{x})^2 \ + \ (\mathbf{x}^* \ \mathbf{R}_- \ \mathbf{x})^2}$$

$$\leq \ \min_{\substack{\mathbf{x}^T\mathbf{x}=1 \\ \mathbf{x}\in R}} \ \mathbf{x}^T \ \mathbf{R}(0) \ \mathbf{x} \ - \ \sqrt{(\mathbf{x}^T \ \mathbf{R}_+ \ \mathbf{x})^2 \ + \ (\mathbf{x}^T \ \mathbf{R}_- \ \mathbf{x})^2} \tag{4.92}$$

$$= \ \min_{\mathbf{x}^T\mathbf{x}=1} \ \mathbf{x}^T \ \mathbf{R}(0) \ \mathbf{x} \ - \ \sqrt{(\mathbf{x}^T \ \mathbf{R}_+ \ \mathbf{x})^2} \tag{4.93}$$

$$= \ \min_{\mathbf{x}^T\mathbf{x}=1} \ \mathbf{x}^T \ \mathbf{R}(0) \ \mathbf{x} \ - \ |(\mathbf{x}^T \ \mathbf{R}_+ \ \mathbf{x})^2| \tag{4.94}$$

$$= \ \min_{\mathbf{x}^T\mathbf{x}=1} \ \min \{\mathbf{x}^T \ (\mathbf{R}(0) \ + \ \mathbf{R}_+) \ \mathbf{x}, \ \mathbf{x}^T \ (\mathbf{R}(0) \ - \ \mathbf{R}_+) \ \mathbf{x}\} \tag{4.95}$$

$$= \ \min \ \{ \ \lambda_{\min} \ (\mathbf{R}(0) \ + \ \mathbf{R}_+), \ \lambda_{\min} \ (\mathbf{R}(0) \ - \ \mathbf{R}_+) \ \} \, . \tag{4.96}$$

The result follows from Proposition 4.7, since from Proposition 4.6 iii) both $\mathbf{R}(0) \ + \ \mathbf{R}_+$ and $\mathbf{R}(0) \ - \ \mathbf{R}_+$ are nonnegative definite ($\phi = 0$, respectively $\phi = \pi$). ■

Proposition 4.9: A sufficient condition for Proposition 4.3 is that

$$\lambda^2_{\min}(\mathbf{R}(0)) \ > \ \max \ \{\lambda^2_{\max}(\mathbf{R}_+) \ , \ \lambda^2_{\min}(\mathbf{R}_+)\} \ + \ \lambda^2_{\max}(\mathbf{R}_-) \ , \tag{4.97}$$

or equivalently

$$\lambda_{\min}(\mathbf{R}^2(0)) \ > \ \lambda_{\max}(\mathbf{R}^2_+) \ + \ \lambda_{\max}(\mathbf{R}^2_-) \ . \tag{4.98}$$

♢

Proof:

$$\min_{\substack{\mathbf{x}^*\mathbf{x}=1 \\ \mathbf{x}\in\mathbb{C}}} \ \mathbf{x}^* \ \mathbf{R}(0) \ \mathbf{x} \ - \ \sqrt{(\mathbf{x}^* \ \mathbf{R}_+ \ \mathbf{x})^2 \ + \ (\mathbf{x}^* \ \mathbf{R}_- \ \mathbf{x})^2}$$

$$\geq \ \min_{\mathbf{x}^*\mathbf{x}=1} \ \mathbf{x}^* \ \mathbf{R}(0) \ \mathbf{x} \ - \ \sqrt{\max_{\mathbf{x}^*\mathbf{x}=1} \ [\mathbf{x}^* \ \mathbf{R}_+ \ \mathbf{x}]^2 \ + \ \max_{\mathbf{x}^*\mathbf{x}=1} \ [\mathbf{x}^* \ \mathbf{R}_- \ \mathbf{x}]^2} \tag{4.99}$$

$$= \ \lambda_{\min}(\mathbf{R}(0)) \ - \ \sqrt{\max \ \{\lambda^2_{\max}(\mathbf{R}_+) \ , \ \lambda^2_{\min}(\mathbf{R}_+)\} \ + \ \max \ \{\lambda^2_{\max}(\mathbf{R}_-) \ , \ \lambda^2_{\min}(\mathbf{R}_-)\}}$$

$$= \ \lambda_{\min}(\mathbf{R}(0)) \ - \ \sqrt{\max \ \{\lambda^2_{\max}(\mathbf{R}_+) \ , \ \lambda^2_{\min}(\mathbf{R}_+)\} \ + \ \lambda^2_{\max}(\mathbf{R}_-)} \, . \tag{4.101}$$

Equation (4.101) follows after noticing that if λ is an eigenvalue of $\mathbf{R}_-$, $-\lambda$ is is an eigenvalue of $\mathbf{R}^T_-$, hence of $\mathbf{R}_-$, which allows us to collapse the second set. Condition (4.98), which requires more computational effort to verify than (4.97), but is of simpler structure, follows from the fact that $\mathbf{R}(0)$ is nonnegative definite (it can be easily reasoned) and both $\mathbf{R}_+$ and $\mathbf{R}_-$ are Hermitian, hence diagonalizable. ∎

We now turn our attention to the two-user case, in which the asymptotic efficiency has a closed form expression.

Proposition 4.10 : In the two-user case let $R_{12}(0) = \rho_{12}$ and $R_{12}(1) = \rho_{21}$. Then the asymptotic efficiency of the decorrelating detector for infinite sequence length is given by:

$$\eta_1^d = \eta_2^d = \sqrt{(1-\rho_{12}^2-\rho_{21}^2)^2 - 4\rho_{12}^2\rho_{21}^2}$$
$$= \sqrt{[1-(\rho_{12}+\rho_{21})^2][1-(\rho_{12}-\rho_{21})^2]}\,. \tag{4.102}$$

◇

Proof : This formula can be obtained by particularizing Proposition 4.4 or by minimizing the asymptotic efficiency of optimal multiuser detection in the two-user case with respect to energies. Alternatively, (4.102) can be proved by taking the limit as $N \to \infty$ of the asymptotic efficiency of the decorrelating filter for the central bits in a length N sequence, which is done in the sequel. (The reason for considering bits near the center of the transmitted sequence is that the demodulation process in this region is least affected by the marginal effects due to the finiteness of the sequence.) Thereby it is proved that in the two-user case the limit of the asymptotic efficiency of the finite-length decorrelating detector as $N \to \infty$ is indeed the asymptotic efficiency of the limiting decorrelating detector.

Recall that the asymptotic efficiency of the decorrelating detector is given by the reciprocal of the corresponding diagonal element of $\mathcal{R}^{-1}$. We need to find explicit expressions for the central diagonal elements of the inverse of the matrix $\mathcal{R}$ as a function of N. We have

$$\mathcal{R} = \begin{pmatrix} 1 & \rho_{12} & 0 & 0 & \\ \rho_{12} & 1 & \rho_{21} & 0 & \\ 0 & \rho_{21} & 1 & \rho_{12} & \ddots \\ 0 & 0 & \rho_{12} & 1 & \ddots \\ & \ddots & \ddots & \ddots & \ddots \end{pmatrix}. \tag{4.103}$$

Denote by Δ_n the determinant of the above $n * n$ matrix. It is easy to see from the structure of $\mathcal{R}$ that Δ_n satisfies the recursion:

$$\Delta_n = \Delta_{n-1} - \begin{cases} \rho_{12}^2\ \Delta_{n-2}\,, & \text{n even} \\ \rho_{21}^2\ \Delta_{n-2}\,, & \text{n odd}\,. \end{cases} \tag{4.104}$$

Hence we can write

$$\begin{bmatrix} \Delta_{2n} \\ \Delta_{2n-1} \end{bmatrix} = \begin{bmatrix} 1-\rho_{12}^2 & -\rho_{21}^2 \\ 1 & -\rho_{21}^2 \end{bmatrix} \begin{bmatrix} \Delta_{2n-2} \\ \Delta_{2n-3} \end{bmatrix} . \tag{4.105}$$

If we consider the sequence of $4n * 4n$ matrices for simplicity, the central diagonal element of the inverse of $\mathcal{R}$ is $\Delta_{4n}/(\Delta_{2n-1}\,\Delta_{2n})$. Hence after introducing the state vector

$$\mathbf{x}_n = \begin{bmatrix} \Delta_{2n} \\ \Delta_{2n-1} \end{bmatrix} , \tag{4.106}$$

we see that finding Δ_{2n}, Δ_{2n-1} requires finding the trajectory of the unforced linear dynamic system

$$\mathbf{x}_n = \begin{bmatrix} 1-\rho_{12}^2 & -\rho_{21}^2 \\ 1 & -\rho_{21}^2 \end{bmatrix} \mathbf{x}_{n-1} , \quad \mathbf{x}_1 = \begin{bmatrix} 1-\rho_{12}^2 \\ 1 \end{bmatrix}$$

i.e.,

$$\mathbf{x}_n = \begin{bmatrix} 1-\rho_{12}^2 & -\rho_{21}^2 \\ 1 & -\rho_{21}^2 \end{bmatrix}^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} . \tag{4.107}$$

The eigenvalues of this system are found to be

$$\lambda_{1,2} \; = \; \frac{1-\rho_{12}^2-\rho_{21}^2 \pm \sqrt{(1-\rho_{12}^2-\rho_{21}^2)^2-4\rho_{12}^2\rho_{21}^2}}{2}\,.$$

We see $0 < \lambda_1 < \lambda_2 < 1$. After finding the corresponding eigenvectors it follows that:

$$\begin{aligned}
\mathbf{x}_n \; &= \; \begin{bmatrix} \lambda_1+\rho_{12}^2 & \lambda_2+\rho_{21}^2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} \begin{bmatrix} 1 & -(\lambda_2+\rho_{21}^2) \\ -1 & \lambda_1+\rho_{21}^2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{1}{\lambda_1-\lambda_2} \\
&= \; \begin{bmatrix} \lambda_1+\rho_{12}^2 & \lambda_2+\rho_{21}^2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^n \\ -\lambda_2^n \end{bmatrix} \frac{1}{\lambda_1-\lambda_2}\,.
\end{aligned} \tag{4.108}$$

Hence the central diagonal element of the inverse of $\mathcal{R}$ is

$$\begin{aligned}
\frac{\Delta_{4n}}{\Delta_{2n-1}\,\Delta_{2n}} \; &= \; \frac{[1 \quad 0]\;\mathbf{x}_{2n}}{[0 \quad 1]\;\mathbf{x}_n\;[1 \quad 0]\;\mathbf{x}_n} \\
&= \; \frac{(\lambda_1/\lambda_2)^{2n}\,(\lambda_1+\rho_{21}^2) \;-\; (\lambda_2+\rho_{21}^2)}{[(\lambda_1/\lambda_2)^n \;-\; 1]\,[(\lambda_1/\lambda_2)^n\,(\lambda_1+\rho_{21}^2) \;-\; (\lambda_2+\rho_{21}^2)]}\,(\lambda_1-\lambda_2)\,.
\end{aligned} \tag{4.109}$$

So finally

$$\eta^d \; = \; \lim_{n\to\infty} \frac{\Delta_{4n}}{\Delta_{2n-1}\,\Delta_{2n}} \; = \; \lambda_2-\lambda_1 \; = \; \sqrt{(1-\rho_{12}^2-\rho_{21}^2)^2-4\rho_{12}^2\rho_{21}^2}\,.$$

◇

Figure 20 shows the asymptotic efficiency of the decorrelating detector for infinite transmitted sequence length in the two-user case. Note its invariance with respect to energies. The discrepancy between η^d and η_{min}, defined in (4.25), is due to the fact that η_{min} is higher than the near-far resistance of optimum multiuser detection, since for η_{min} the energies are constrained to be constant over time.

The fact that the stable version of the decorrelating filter turns out to be noncausal is not surprising. Due to the lack of synchronism among the users any decision based on less than the entire received waveform is suboptimal. In practice, since the filter is stable, the more remote symbols will count less heavily, and truncation of the noncausal part will be performed after a suitable delay without affecting performance appreciably. For illustration consider the two-user case. There from (4.74)

$$\mathbf{S}(z) \; = \; \begin{pmatrix} 1 & \rho_{12}+\rho_{21}z^{-1} \\ \rho_{12}+\rho_{21}z & 1 \end{pmatrix}$$

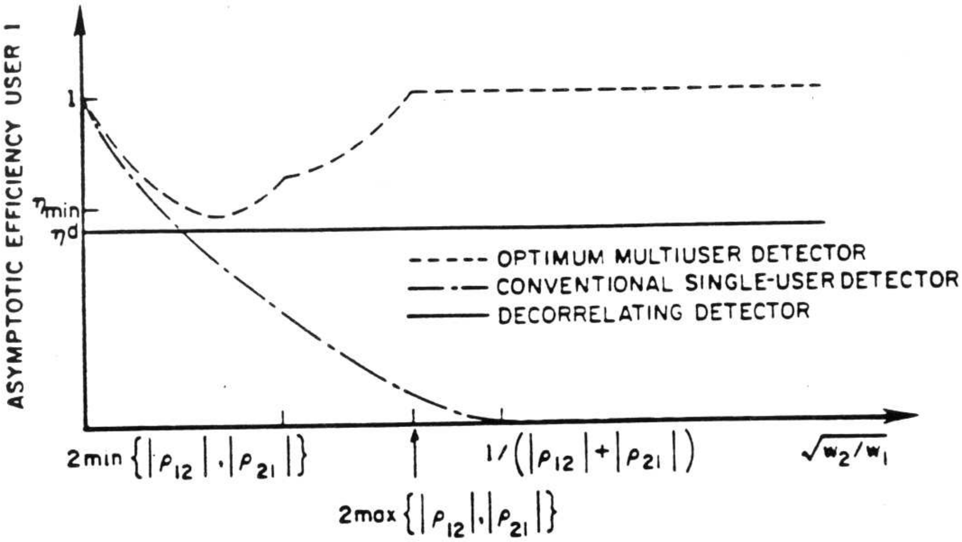


Fig. 20. Asymptotic efficiencies in the 2-user case for infinite transmitted sequence length, when the user energies are constant over time (here we chose $|\rho_{12}|, |\rho_{21}|$=0.3,0.5 which yields $\eta_{min} = 0.68$, η^d =0.59).

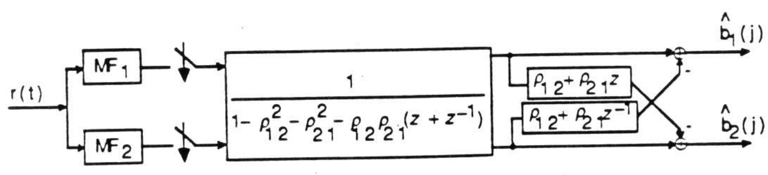


Fig. 21. Limiting decorrelating detector in the 2-user case.

and the transfer function of the decorrelating detector as given by (4.71) is:

$$\mathbf{G}(z) = \frac{1}{1-\rho_{12}^2-\rho_{21}^2-\rho_{12}\rho_{21}z-\rho_{12}\rho_{21}z^{-1}} \begin{pmatrix} 1 & -(\rho_{12}+\rho_{21}z^{-1}) \\ -(\rho_{12}+\rho_{21}z) & 1 \end{pmatrix} . \quad (4.110)$$

The resulting detector is shown in Figure 21. We are interested in the impulse response $f(n)$ of the IIR part of the above filter. Taking the inverse z-transform it is found to be

$$f(n) \;=\; Z^{-1}\left[\frac{1}{1-\rho_{12}^2-\rho_{21}^2-\rho_{12}\rho_{21}z-\rho_{12}\rho_{21}z^{-1}}\right] \;=\; \frac{\xi^{|n|}}{\eta}\,, \tag{4.111}$$

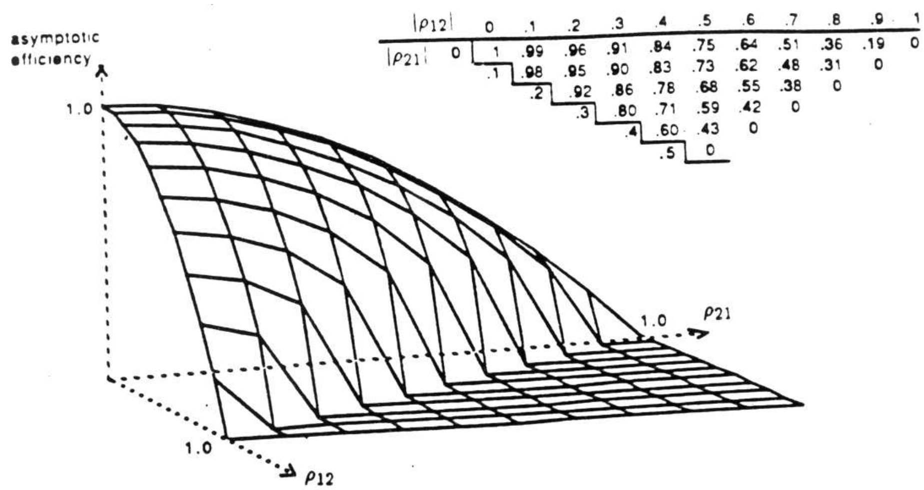
where $\xi \;=\; (1-\rho_{12}^2-\rho_{21}^2-\eta)/(2\rho_{12}\rho_{21})$ and η is the asymptotic efficiency which is given in Proposition 4.10. It can be checked that $|\xi| \leq 1$, with equality if $|\rho_{12}|+|\rho_{21}| = 1$, which can be shown to coincide with the condition imposed by Proposition 4.3 for the two-user case. In the latter case the asymptotic efficiency is zero, which follows from Proposition 4.10. Otherwise since $|\xi| < 1$ the limiting filter is stable, with symmetric coefficients which decay with rate ξ. In practical applications the filter will be approximated up to any desired precision by truncation of the noncausal part to a finite number of filter coefficients. For illustration the decay rate ξ of the filter coefficients and the achievable asymptotic efficiency η are plotted in Figure 22 as functions of ρ_{12} and ρ_{21}.

Poor cross-correlation properties among the signature waveforms could imply that the limiting filter $\mathbf{G}(z)$ does not exist, although the decorrelating detector exists for finite-length transmitted sequences. We give an example to illustrate this fact. As mentioned earlier, for $K = 2$ it is straightforward to show that the condition of Proposition 4.3 is satisfied for all signal constellations for which $|\rho_{12}| + |\rho_{21}| \neq 1$, which is the case unless the normalized waveforms coincide modulo circular shifts and sign changes.

Consider the trivial signal case, where both users are assigned the same rectangular waveform, as shown in Figure 23. In this case ρ_{12}, which is the crosscorrelation between bits in the same signaling interval, is $r = (T-\tau)/T \in [0,1)$, where $\tau \in [0,T)$ is the delay between the two users. Then ρ_{21}, which is the crosscorrelation between bits in adjacent intervals, is $1-r$. Then

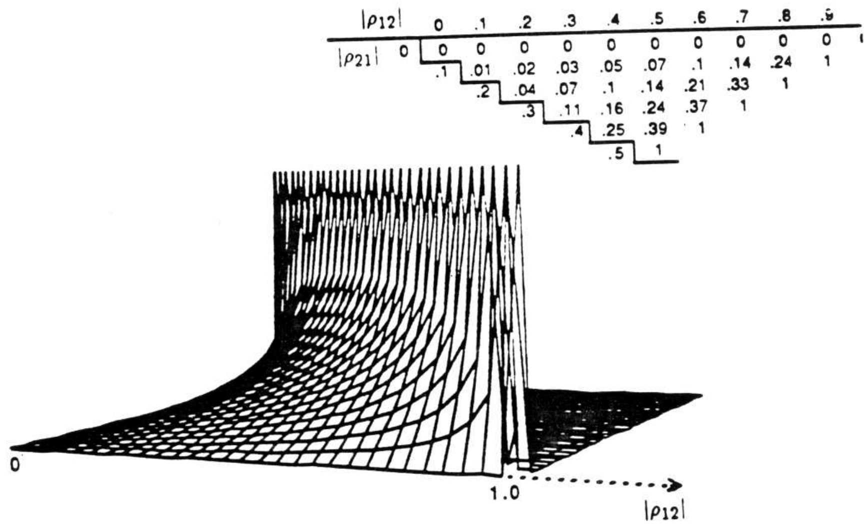
$$\mathbf{S}(z) \;=\; \begin{pmatrix} 1 & r+(1-r)z^{-1} \\ r+(1-r)z & 1 \end{pmatrix} \tag{4.112}$$

becomes singular for $z = 1$, hence there is no stable limiting inverse filter. And if it existed its asymptotic efficiency, as given by (4.102), would be zero. For an infinite sequence of transmitted bits where both users use the same waveform, this is not surprising. However for finite length sequences advantage can be taken of the marginal effects of having bits which are not affected by either past or future bits. For finite N the decorrelating detector exists unless $r = 0$, i.e., when the transmissions are not synchronous. This is in accord with the multiarrival condition given in Appendix 4.1, and with the results obtained in the synchronous case.



$\|\rho_{21}\| \backslash \|\rho_{12}\|$	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1
0	1	.99	.96	.91	.84	.75	.64	.51	.36	.19	0
.1		.98	.95	.90	.83	.73	.62	.48	.31	0	
.2			.92	.86	.78	.68	.55	.38	0		
.3				.80	.71	.59	.42	0			
.4					.60	.43	0				
.5						0					

Fig. 22a. Asymptotic efficiency of the decorrelating detector for 2 users as a function of the partial crosscorrelations of their signature waveforms recall $|\rho_{12}| + |\rho_{21}| \leq 1$, and symmetry in $|\rho_{12}|, |\rho_{21}|$.



$\|\rho_{21}\| \backslash \|\rho_{12}\|$	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	
0	0	0	0	0	0	0	0	0	0	0	1
.1		.01	.02	.03	.05	.07	.1	.14	.24	1	
.2			.04	.07	.1	.14	.21	.33	1		
.3				.11	.16	.24	.37	1			
.4					.25	.39	1				
.5						1					

Fig. 22b. Decay rates of the coefficients of the IIR part of the decorrelating detector for 2 users, symmetric in ρ_{12} and ρ_{21}.

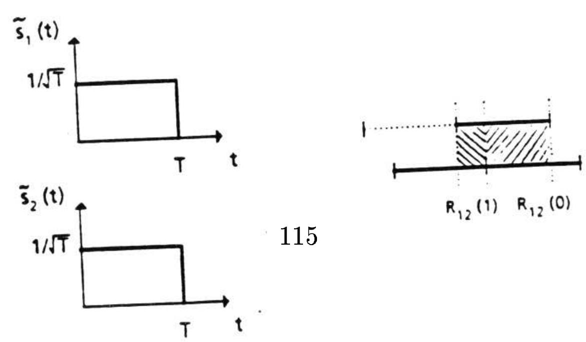


Fig. 23. Signals and crosscorrelations of example (4.112).

Appendix 4.1: Sufficient conditions for linear independence

Suppose that, for a fixed signal set,

i) $\{\tau_1, ..., \tau_K\}$ are continuous random variables,

ii) $\{\tau_1, ..., \tau_K\}$ are independent random variables,

iii) $w_k(i) \neq 0$.

Then almost surely there is no $\mathbf{v} \in L$, $v_k(i) \neq 0$ such that $\tilde{S}(t, \mathbf{v}) = 0$.

Proof : Define the times of effective arrival and departure of the i^{th} signal of the k^{th} user, [1], as

$$\lambda_{i,k}^a = \tau_k + iT + \sup\left\{\tau \in [0, T), \int_0^\tau s_k^2(t)dt = 0\right\} \tag{4.113}$$

and

$$\lambda_{i,k}^d = \tau_k + iT + \inf\left\{\tau \in (0, T], \int_\tau^T s_k^2(t)dt = 0\right\}, \tag{4.114}$$

respectively.

Since $v_k(i) \neq 0$ there is a first and a last symbol that differs from zero. It is readily apparent that in order to have $\tilde{S}(t, \mathbf{v}) = 0$, the effective arrival of the first (and the effective departure of the last) symbol that differs from zero must be a point of effective multiarrival (respectively multideparture). Note that this property does not depend on the particular $\mathbf{v}$ chosen, but only on the set of delays. From (4.113),(4.114), the effective times of arrival and departure inherit from the delays the properties of being continuously valued and mutually independent. Therefore, the result follows, since the set of delays $\{\tau_1, ..., \tau_K\}$ for which multiarrival points result has measure zero.

$\diamond$

Appendix 4.2: If the LIA is not satisfied

In this case $\mathcal{R}$ is nonnegative rather than positive definite. We will concentrate on the i^{th} bit of the k^{th} user. Recall the definition of the decorrelating detector for the i^{th} bit of the k^{th} user, given in (4.58).

Proposition 4.11 : The i^{th} bit k^{th} user decorrelating filter exists if and only if

$$\forall\, \mathbf{x} \in L \text{ with } x_k(i) \neq 0 : \; \|\tilde{S}(t, \mathbf{x})\| \neq 0 \tag{4.115}$$

$\diamondsuit$

It can be seen that this condition is weaker than the LIA of (4.3), which is necessary and sufficient for existence of the decorrelating detector for each bit of each user.

Proof : The proof is given by the equivalences (4.59)-(4.63). ∎

Proposition 4.12 : The decorrelating filter for the i^{th} bit of the k^{th} user does not exist if and only if any decision algorithm for the i^{th} bit of the k^{th} user has a near-far resistance of zero. $\diamondsuit$

Proof : Let's assume that conditions are such that the decorrelating filter for the i^{th} bit of the k^{th} user does not exist. Then by Proposition 4.11 there exists an $\mathbf{x}$ with $x_k(i) = 1$ such that $\tilde{S}(t, \mathbf{x}) = 0$, since $\tilde{S}(t, \mathbf{x})$ is linear in $x_k(i)$. But then (4.28) implies that the near-far resistance of the optimal multiuser detector is zero. Hence unless any decision algorithm has a near-far resistance of zero, existence of the decorrelating filter is ensured. ∎

Proposition 4.13 : Let the existence condition (4.115) for a decorrelating filter for bit i of User k be satisfied. Let $D_{i,k}$ be the set of decorrelating filters for the i^{th} bit of the k^{th} user. Denote the set of generalized inverses of $\mathcal{R}$ by $I(\mathcal{R})$. Denote the set of $(i,k)^{th}$ rows of elements in $I(\mathcal{R})$ by $G_{i,k}$. Then a vector $\mathbf{v} \in L$ is a decorrelating detector if and only if $\mathbf{v} \in G_{i,k}$. $\diamondsuit$

Proof : In other words we need to show $D_{i,k} = G_{i,k}$. The equality is to be interpreted as an isomorphism between otherwise identical sets of row vectors and of column vectors (this is necessary since generalized inverses do not have to be symmetric). We will use the defining property for a decorrelating detector $\mathbf{d}$ for bit i of User k, namely that $\mathcal{R}\mathbf{d} = \mathbf{u}^{i,k}$, and the equivalent existence condition for a decorrelating filter given in (4.62), namely that the $(i,k)^{th}$ column of $\mathcal{R}$ is linearly independent of the others.

a) $G_{i,k} \subseteq D_{i,k}$: Let $\mathcal{B} \in I(\mathcal{R})$ and $\mathcal{S} = \mathcal{B}\mathcal{R} - \mathcal{I}$. By the definition of generalized inverse, it follows that $\mathcal{R}\mathcal{S} = 0$, i.e., every column of $\mathcal{S}$ is in the nullspace of $\mathcal{R}$. But since the $(i,k)^{th}$ column of $\mathcal{R}$ is linearly independent of the other columns of $\mathcal{R}$, it is necessary that the $(i,k)^{th}$ element of each column of $\mathcal{S}$ be zero, i.e. that the $(i,k)^{th}$ row of $\mathcal{S}$ be zero. Hence $\mathcal{B}_{i,k}^T\mathcal{R} - (\mathbf{u}^{i,k})^T = 0$, which implies that $\mathcal{B}_{i,k}$ is decorrelating.

b) $D_{i,k} \subseteq G_{i,k}$: Let $\mathbf{d} \in D_{i,k}, \mathbf{x} \in G_{i,k}$. Then, by the "decorrelating" property of $\mathbf{d}$ and that of $\mathbf{x}$ established in a), $\mathcal{R}(\mathbf{d} - \mathbf{x}) = 0$. Besides implying, by the same reasoning as the one used in a), that the (i,k) element of any k^{th} user decorrelating detector is equal, this equation shows that $\mathbf{x}$ differs from $\mathbf{d}$ by an element in the nullspace of $\mathcal{R}$. But it can be readily checked that $D_{i,k}+\mathcal{N}(\mathcal{R}) = D_{i,k}$.

$\diamondsuit$

It is easiest to think of the equivalent synchronous model when dealing with the case when the LIA does not hold, and use the results obtained in the synchronous case for independent and dependent users.

Appendix 4.3

Proof of Proposition 4.6: i) From a well-known theorem in linear algebra the eigenvalues of a Hermitian matrix are real (e.g. Thm. 2.5.6. in [Horn]). Hence as a corollary the determinant of a Hermitian matrix is real. Write

$$M_{kk}^{-1}(\varphi) \;=\; \det C_{kk}\ (\mathbf{M}(\varphi))\ /\ \det \mathbf{M}(\varphi), \tag{4.116}$$

where C_{kk} is the k^{th} cofactor of a matrix, and note that both $\mathbf{M}(\varphi)$ and (hence) $C_{kk}\ (\mathbf{M}(\varphi))$ are Hermitian.

ii) Making the real and imaginary parts explicit we can write $\mathbf{M}(\varphi) = \mathbf{A} + j\mathbf{B}$, where we omit the dependence on φ for notational convenience. Then $\mathbf{M}^{-1}(\varphi) = \mathbf{C} + j\mathbf{D}$, where $\mathbf{C}$ and $\mathbf{D}$ satisfy

$$\begin{array}{ll} \mathbf{AC} - \mathbf{BD} = \mathbf{I} & \quad \mathbf{CA} - \mathbf{DB} = \mathbf{I} \\ \mathbf{BC} + \mathbf{AD} = \mathbf{0} & \quad \mathbf{CB} + \mathbf{DA} = \mathbf{0} \end{array} \tag{4.117}$$

Hence if $(\mathbf{C},\mathbf{D})$ corresponds to $(\mathbf{A},\mathbf{B})$, then $(\mathbf{C},-\mathbf{D})$ corresponds to $(\mathbf{A},-\mathbf{B})$. On the other hand $\mathbf{A} = \mathbf{R}(0) + [\mathbf{R}^T(1) + \mathbf{R}(1)]\ \cos\varphi$ and $\mathbf{B} = [\mathbf{R}^T(1) - \mathbf{R}(1)]\ \sin\varphi$, so that for $\varphi \to 2\pi - \varphi$, $(\mathbf{A},\mathbf{B}) \to (\mathbf{A},-\mathbf{B})$. Hence $\mathbf{C}$, and therefore, with i), $M_{kk}^{-1}(\varphi)$, is invariant under the transformation $\varphi \to 2\pi - \varphi$.

iii) Given the sets of normalized signature waveforms and delays, define a new set of waveforms as follows:

$$\Phi \;=\; \{\ \phi_i(t),\ i=1,...K \mid \phi_i(t) \;=\; \begin{cases} \tilde{s}_i(t+T-\tau_i)\ e^{j\varphi} & ,\ 0\ \le t\ \le\ \tau_i \\ \tilde{s}_i(t-\tau_i) & ,\ \tau_i < t\ \le\ T\ . \end{cases} \tag{4.118}$$

Define the complex crosscorrelations

$$M_{kj}^{\Phi} \;=\; \int_0^T \phi_k(t)\ \phi_j^*(t)\ dt. \tag{4.119}$$

Then

$$\mathbf{x}^T \mathbf{M}^\Phi \mathbf{x}^* \; = \; \sum_i \sum_j x_i \, x_j^* \, M_{ij}^\Phi \; = \; \int_0^T \| \sum_i x_i \, \phi_i(t) \, \|^2 \, dt \quad \geq 0 \; . \tag{4.120}$$

Therefore $\mathbf{M}^\Phi$ is nonnegative definite. Next we show that $\mathbf{M}^\Phi \equiv \mathbf{M}(\varphi)$, which establishes the desired result. First let $k \leq j$, and recall that the users have been numbered according to increasing delays. Then

$$\begin{aligned} M_{kj}^\Phi \; &= \; \int_0^{\tau_k} \tilde{s}_k(t+T-\tau_k) \, e^{j\varphi} \, \tilde{s}_j(t+T-\tau_j) \, e^{-j\varphi} \, dt \; + \; \int_{\tau_k}^{\tau_j} \tilde{s}_k(t-\tau_k) \, \tilde{s}_j(t+T-\tau_j) \, e^{-j\varphi} \, dt \\ &\qquad + \int_{\tau_j}^{T} \tilde{s}_k(t-\tau_k) \, \tilde{s}_j(t-\tau_j) \, dt \\ &= \; \int_{\tau_j}^{T+\tau_k} \tilde{s}_k(t-\tau_k) \, \tilde{s}_j(t-\tau_j) \, dt \; + \; e^{-j\varphi} \int_{\tau_k}^{\tau_j} \tilde{s}_k(t-\tau_k) \, \tilde{s}_j(t+T-\tau_j) \, dt \\ &= \; R_{kj}(0) \; + \; R_{kj}(1) \, e^{-j\varphi} \; . \end{aligned} \tag{4.121}$$

Similarly, for $k > j$, $M_{kj}^\Phi = R_{kj}(0) \; + \; R_{kj}(-1) \, e^{j\varphi}$. Hence

$$\mathbf{M}^\Phi \; = \; \mathbf{R}^T(1) \, e^{j\varphi} + \mathbf{R}(0) + \mathbf{R}(1) \, e^{-j\varphi} \; = \; \mathbf{M}(\varphi) \; .$$

■

Proof of Corollary: The first two results are immediate consequences of the theorem and of (4.80, 81).

iii) By the LIA $\mathbf{M}(\varphi)$ is invertible, hence, with point iii) of the theorem, positive definite. For a positive definite $K * K$ matrix $\mathbf{M}$ with $\mathbf{M}^T = \mathbf{M}^*$, and for any K-vector $\mathbf{m}$ such that

$$\begin{bmatrix} \mathbf{M} & \mathbf{m}^* \\ \mathbf{m}^T & 1 \end{bmatrix} \; > \; 0 \, , \tag{4.122}$$

for any $1 \leq k \leq K$, we want to express

$$\begin{bmatrix} \mathbf{M} & \mathbf{m}^* \\ \mathbf{m}^T & 1 \end{bmatrix}_{kk}^{-1} \quad \text{in terms of} \quad M_{kk}^{-1} \; . \tag{4.123}$$

But, using formulas for the inverse of a partitioned matrix and then for the inverse of a small-rank adjustment (e.g. [Horn, p. 18]),

$$\begin{bmatrix} \mathbf{M} & \mathbf{m}^* \\ \mathbf{m}^T & 1 \end{bmatrix}_{kk}^{-1} \; = \; M_{kk}^{-1} \; + \; \frac{(\mathbf{u}_k^T \mathbf{M}^{-1} \mathbf{m}^*) \, (\mathbf{m}^T \mathbf{M}^{-1} \mathbf{u}_k)}{1 - \mathbf{m}^T \mathbf{M}^{-1} \mathbf{m}^*} \; = \; M_{kk}^{-1} \; + \; \frac{\| \mathbf{m}^T \mathbf{M}^{-1} \mathbf{u}_k \|}{1 - \mathbf{m}^T \mathbf{M}^{-1} \mathbf{m}^*} \quad . \tag{4.124}$$

From here, using the connection between η_k^d and M_{kk}^{-1} of (4.81), equation (4.86) is immediate. To show mathematically that the asymptotic efficiency is nonincreasing we show that $\mathbf{m}^T \mathbf{M}^{-1} \mathbf{m}^* <$

1 (note that the asymptotic efficiency of a $(K+1)$-user problem cannot decrease if additional information is available, e.g. if the additional user's information sequence is known, which reduces to a K-user problem). For this note that (4.122) is equivalent to

$$\mathbf{x}^T\mathbf{M}\mathbf{x}^* + y\mathbf{m}^T\mathbf{x}^* + y^*\mathbf{x}^T\mathbf{m}^* + yy^* \ > \ 0\,, \quad \forall \mathbf{x} \in \mathbb{C}^K, y \in \mathbb{C}\,. \tag{4.125}$$

In particular also $\forall y \in \mathbb{R}$,

$$\mathbf{x}^T\mathbf{M}\mathbf{x}^* + y\,(\mathbf{m}^T\mathbf{x}^* + \mathbf{x}^T\mathbf{m}^*) + y^2 \ > \ 0\,, \quad \forall \mathbf{x} \in \mathbb{C}^K, y \in \mathbb{C}\,. \tag{4.126}$$

The left hand side is a quadratic function in y which is strictly positive, therefore its discriminant is strictly negative, i.e.

$$(\mathbf{m}^T\mathbf{x}^* + \mathbf{x}^T\mathbf{m}^*)^2 \ - \ 4\,\mathbf{x}^T\mathbf{M}\mathbf{x}^* \ < \ 0\ , \quad \forall \mathbf{x} \in \mathbb{C}^K\,. \tag{4.127}$$

Then letting $\mathbf{x} = (\mathbf{M}^{-1})^T\mathbf{m}$, we obtain

$$\mathbf{m}^T\mathbf{M}^{-1}\mathbf{m}^* \ > \ (\mathbf{m}^T\mathbf{M}^{-1}\mathbf{m}^*)^2 \tag{4.128}$$

which implies that

$$0 \ < \ \mathbf{m}^T\mathbf{M}^{-1}\mathbf{m}^* \ < \ 1\,. \tag{4.129}$$

Hence the integrand in (4.86) is nondecreasing when one user is added to the system, which establishes the desired result. ∎

4.3 Numerical examples: Probability of error

In the sequel the performances of the conventional and of the decorrelating detector are compared. Without loss of generality attention is focused on the error probability of User 1 in a channel shared by several active users. The conventional detector decides for the sign of the first component of the matched filter output vector, given by (4.14). Therefore its average error probability over the bit sequences of the interfering users equals

$$\frac{1}{2^{2\,(K-1)}} \sum_{b_j(0), b_j(-1), j \neq 1} Q\left(\frac{\sqrt{w_1} - \sum_{j=2}^{K} [\mathbf{R}(0)_{1j}\, b_j(0) + \mathbf{R}(1)_{1j}\, b_j(-1)]\, \sqrt{w_j}}{\sigma} \right), \tag{4.130}$$

whereas its worst-case error probability over the interfering bit sequences equals

$$Q\left(\frac{\sqrt{w_1} - \sum_{j=2}^{K} [\, |\mathbf{R}(0)_{1j}| + |\mathbf{R}(1)_{1j}|\,]\, \sqrt{w_j}}{\sigma} \right). \tag{4.131}$$

The probability of error of the decorrelating detector equals, from (4.82),

$$Q\left(\frac{\sqrt{w_1\, \eta_1^d}}{\sigma} \right), \tag{4.132}$$

where from (4.85)

$$\frac{1}{\eta_1^d} = \frac{1}{\pi} \int_0^{\pi} [\, \mathbf{R}(1)^T e^{j\omega} + \mathbf{R}(0) + \mathbf{R}(1) e^{-j\omega}\,]_{11}^{-1}\, d\omega\, . \tag{4.133}$$

The delays and phases of the K users enter the above formulas implicitly via the crosscorrelation matrices, which are functions thereof and of the chosen signature waveforms. In the following we consider Direct-Sequence Spread-Spectrum (DS-SS) signaling, where the carrier modulated waveform is given by

$$s_k(t-\tau_k) \;=\; \sqrt{2w_k}\; a_k(t-\tau_k)\, \cos\, (\omega_c t \;+\; \phi_k) \tag{4.134}$$

where $a_k(t)$ is the code waveform, zero outside [0,T], with $\int_0^T a_k^2(t) = 1$, τ_k is the delay of User k due to propagation delay and lack of synchronism between the users, and ϕ_k is the phase angle of the k^{th} carrier. In DS-SS the code waveform $a_k(t)$ is a sequence of N_c nonoverlapping rectangular pulses of amplitude $\pm(1/N_c)^{1/2}$ and duration $T_c = T/N_c$, called chips, so that the k^{th} user waveform is characterized by a sequence of N_c bits, called signature sequence, giving the chip polarities.

4.3.1 On the dependence on delays and phases

In order to compute average error probabilities for a specified set of signature sequences, the average of the error probability expressions (4.130), (4.132) has to be taken with respect to the set of phases and delays, assumed to be independent and uniformly distributed. If $\omega_c T >> 1$, then $R_{kj}(l)$ can be related to the baseband crosscorrelation coefficient, via

$$
\begin{aligned}
R_{kj}(l) &\triangleq \int_{-\infty}^{\infty} \tilde{s}_k(t-\tau_k)\, \tilde{s}_j(t+lT-\tau_j)dt \\
&= \frac{1}{2}\cos(\phi_j-\phi_k) \int_{-\infty}^{\infty} a_k(t-\tau_k)\, a_j(t+lT-\tau_j)dt\,. \qquad (4.135)
\end{aligned}
$$

Since the magnitude of the crosscorrelation coefficients, hence of the multiuser interference, is maximized if the cosine term is unity, in the literature ([Ver 84c], [Var 88]) error probability curves are given for the baseband case. In the case of the conventional receiver it is easy to see from the dependence of the error probability on the crosscorrelation coefficients (4.130) that this corresponds to worst-case conditions. In the case of more complicated detectors the nonlinear dependence of the error probability on the crosscorrelation coefficients precludes a proof of this fact. Nevertheless, intuition tells us that performance should become worse if the absolute values of the crosscorrelations between the users increase. Researchers in the field have worked with this intuitive assumption, e.g. [Ver 84c] for error probability curves of the maximum likelihood detector and [Var 88] for the performance analysis of their M-stage iterative multiuser detector, where the dependence of the error probability on the set of delays and phases is mentioned only for the first stage, which is the conventional detector. This work adheres to this usage, which is supported by the intuition of the problem, because analytical results on phase dependence of the decorrelating detector performance have proved intractable.

Because of the symmetry of the problem assume that the user whose error probability is of interest is User 1. Since only relative delays matter one can set $\tau_1 = 0$ and let $\tau_k \in [0, T)$ denote the delay of User k relative to User 1. Since the chip waveform is rectangular, the correlation coefficients R_{kj} depend linearly on the distance of the respective relative delay between users k and j to the next chip boundary, and are linear combinations of the crosscorrelation values at the adjacent chip boundaries. More precisely, letting R^{τ}_{kj} be the value of R_{kj} when users are delayed by $\tau \cong \tau_{kj} \stackrel{\triangle}{=} |\tau_k - \tau_j|$,

$$R^{\tau}_{kj}(0) = R^{lT_c}_{kj}(0)\,(1 - \frac{\tau - lT_c}{t_c}) + R^{(l+1)T_c}_{kj}(0)\,\frac{\tau - lT_c}{t_c}\,, \qquad l \cong l_{kj} \stackrel{\triangle}{=} \lfloor \frac{\tau_{kj}}{T_c} \rfloor \tag{4.136}$$

and

$$R^{\tau}_{kj}(1) = R^{lT_c}_{kj}(1)\,(1 - \frac{\tau - lT_c}{t_c}) + R^{(l+1)T_c}_{kj}(1)\,\frac{\tau - lT_c}{t_c}\,. \tag{4.137}$$

This property and the convexity of the Q-function imply that the error probability of the conventional detector, averaged over continuous delays, is upper bounded by its average on the discrete grid of time delays corresponding to the chip boundaries. A similar result for the maximum likelihood detector was conjectured in [Ver 84c] on the ground that the argument of the Q-function in the error probability expression is affine (i.e. translated linear) in the vector of delays in every cube $[n_2 T_c, (n_2+1)T_c] \times \cdots \times [n_K T_c, (n_K+1)T_c]$. However this argument is insufficient because $l_{kj} \neq l_k - l_j$, i.e. we encounter nonlinearity, and dependence upon a set of grid boundaries which do not correspond to one delay vector. The same difficulty arises for any detector whose k^{th} user performance depends on crosscorrelations other than those involving the k^{th} waveform, in particular also for the decorrelating detector. Here an additional difficulty in characterizing the delay dependence of the error probability arises because of the nonlinear operation of matrix inversion involved in the expression for the asymptotic efficiency of the limiting decorrelating detector. Probably due to similar difficulties [Var 88] skips the issue, and motivate averaging over the discrete grid of chip delays with the endeavor "to conduct a meaningful comparison" with the conventional and optimum receivers. For the decorrelating detector the following result holds in the two-user case.

Proposition 4.14: For a 2-user DS-SS environment the error probability of the decorrelating detector averaged over a continuous delay between the users is upper bounded by its average over the discrete grid of time delays corresponding to chip boundaries. ◇

Proof: First we show that $\eta^d_k(\tau)$ is concave in τ, i.e.

$$\eta^d_k(\tau) \geq (1 - \alpha_\tau)\,\eta^d_k(lT_c) + \alpha_\tau\,\eta^d_k((l+1)T_c)\,. \tag{4.138}$$

Then, since the function $f(x) = \sqrt{x}$ is concave

$$\sqrt{\eta^d_k(\tau)} \geq (1 - \alpha_\tau)\sqrt{\eta^d_k(lT_c)} + \alpha_\tau \sqrt{\eta^d_k((l+1)T_c)} \tag{4.139}$$

and, denoting for simplicity

$$Q_\tau \triangleq Q\left(\frac{\sqrt{w_k}}{\sigma}\sqrt{\eta_k^d(\tau)}\right), \tag{4.140}$$

since the function $Q(x)$ is convex and decreasing in x

$$Q_\tau \le (1 - \alpha_\tau)\, Q_{lT_c} + \alpha_\tau\, Q_{(l+1)T_c}, \qquad \forall\, \tau \in [lT_c,\ (l+1)T_c). \tag{4.141}$$

Therefore the expected value of the error probability over τ is upper bounded as follows.

$$E\,[Q_\tau] = \sum_{l=0}^{N_c-1} E\,[Q_\tau \mid \tau \in [lT_c,(l+1)T_c)]\,\frac{1}{N_c} \tag{4.142}$$

$$\le \frac{1}{N_c}\sum_{l=0}^{N_c-1} Q_{lT_c} + \frac{E\,[\alpha_\tau]}{N_c}\sum_{l=0}^{N_c-1}[Q_{(l+1)T_c} - Q_{lT_c}] \tag{4.143}$$

$$= \frac{1}{N_c}\sum_{l=0}^{N_c-1} Q_{lT_c} = E\,[Q_\tau \mid \tau \in \{lT_c, l = 0, ..., N_c - 1\}\,]\,. \tag{4.144}$$

It is left to prove that $\eta_k^d(\tau)$ is concave in τ. Fortunately, in the two-user case an explicit expression for the asymptotic efficiency was found in Proposition 4.10, namely

$$\eta_k^d(\tau) = \sqrt{[1-(\rho_{12}(\tau)+\rho_{21}(\tau))^2][1-(\rho_{12}(\tau)-\rho_{21}(\tau))^2]}\,. \tag{4.145}$$

Since both the sum and the difference in this expression are linear combinations of the sum respectively difference values at the chip boundaries, after an obvious change of variable it suffices to show that the function of two variables

$$f(x,y) = \sqrt{1 - x^2}\,\sqrt{1 - y^2}$$

is concave on its domain of definition $\{|x| \le 1\} \times \{|y| \le 1\}$. A necessary and sufficient condition for the latter is that the Hessian matrix of $f(.,.)$ be nonpositive definite ([Horn], p. 392). The Hessian matrix of $f(.,.)$ is

$$\mathbf{H} = \begin{pmatrix} -\dfrac{\sqrt{1 - y^2}}{(\sqrt{1 - x^2})^3} & -\dfrac{xy}{\sqrt{1 - x^2}\,\sqrt{1 - y^2}} \\ -\dfrac{xy}{\sqrt{1 - x^2}\,\sqrt{1 - y^2}} & -\dfrac{\sqrt{1 - x^2}}{(\sqrt{1 - y^2})^3} \end{pmatrix}, \tag{4.146}$$

which is nonpositive definite since the determinant is nonnegative and the trace is nonpositive. Therefore the asymptotic efficiency in the two-user case is a concave function of ρ_{12} and ρ_{21}, as can also be seen from its plot in Figure 22. ∎

The reason why the two-user case is analytically easier to handle than the general case is twofold. First, we have an explicit expression for the asymptotic efficiency, which means that we can avoid handling the complicated expression of (4.133). And second in this case we do not have to worry about the fact that the parameters l_i do not suffice to specify the set $\{l_{ij}\}$, since there is only one parameter l to consider. If rather than using the explicit expression for the asymptotic efficiency, we try to show concavity using the concavity of the determinant in (4.133) in the two-user case, we have not been able to show Proposition 4.14. This is evidence of the toughness of the analytic problem in the general case.

However Proposition 4.14 gives an indication that its extension can be expected to hold in the K-user case. In order to substantiate this assertion we have numerically compared the average error probability of the limiting decorrelating detector for the cases of averaging over chip delays and finer subdivision (of course continuous delays are not feasible numerically), for a three-user DS-SS environment. The results are presented in the next section, and corroborate the expectation that chip delays are worse than continuous delays. We do not know of any counterexample.

Finally, it is shown in [Pur 81] that the crosscorrelation magnitudes are maximized for values of τ which are integer multiples of T_c (even for arbitrary time-limited chip waveforms). Proposition 4.14 in connection with this fact supports our intuition that worst-case conditions are those where the crosscorrelations are maximized, which motivated our choice of baseband analysis.

4.3.2 Numerical Examples

In the following examples we have chosen a set of Spread-Spectrum m-sequences of length 31. First a three-user baseband environment where -for comparison purposes with previous works ([Ger 82], [Ver 86a])- we have used the set of 3 sequences reported in [Gar 80], Table 5 to be optimal with respect to a signal-to-multiple-access interference parameter when the conventional detector is used. To begin with we investigate the dependence on delays of the performance of the decorrelating detector for infinite sequence length. Each chip interval has been subdivided into n subintervals, and the performance has been averaged over the discrete grid of resulting delays, for $n = 1, 5$ and 10. The case $n = 1$ corresponds to the case of delays which are integer multiples of chip intervals, i.e. the case we have discussed in the previous section. Figure 24 shows the error probability for User 1, averaged over the delays of the users, for the different sets of admissible delays. We see that the error probability decreases if a finer subdivision is used, supporting the claim that averaging over chip interval delays leads to an upper bound on error probability. The

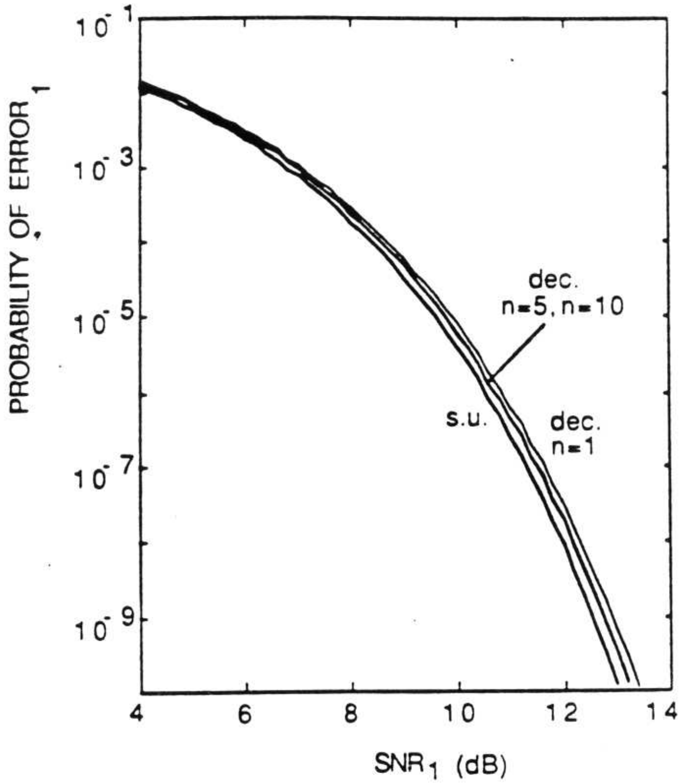
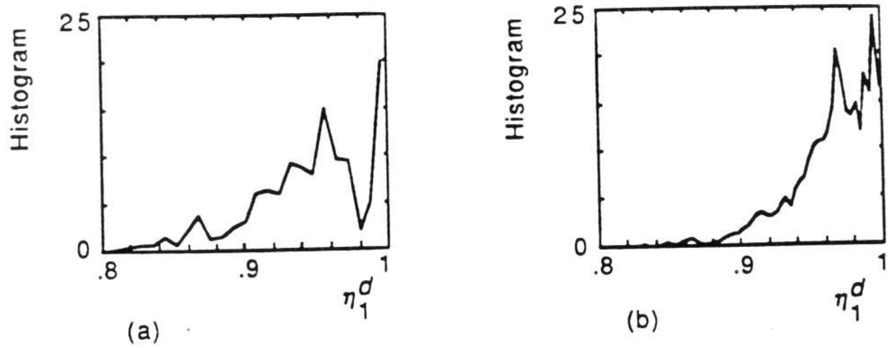
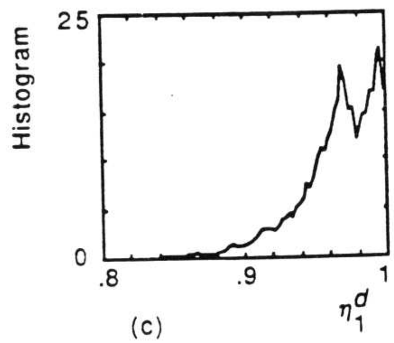


Fig. 24. Error probability of the limiting decorrelating detector for 3 users using m-sequences of length 31, where the average over delays has been made for n subintervals per chip. a) n=1, b) n=5, c) n=10.

single user error probability is also shown for comparison. Observe the good performance of the decorelating detector.

We have also computed the asymptotic efficiency for each of the $(n * 31)^2$ possible delay configurations. The resulting approximate probability density functions for $n = 1, 5$ and 10, i.e. for 961, 24025 and 96100 samples, are shown in Figure 25. They have been obtained by subdividing the interval [0,1] into respectively 125, 300 and 400 bins, and plotting the probability that the asymptotic efficiency takes values in the corresponding bin, normalized such that the total probability is 1. The number of bins to be used has been chosen by comparing different choices under the criterion that an overly ragged curve probably means that there are too few points per bin to give a reliable result, while an overly smooth curve, which changes a lot when the bin number is increased, is too coarse. Therefore some of the edges in the curves shown may be due to the unavoidable discretization due to the magnitude of the available sample sizes. However the bimodal character of the p.d.f. was present for all bin choices in the cases b) and c).





	sample size	mean	stdev
a)	961	.9425	.0426
b)	24025	.9634	.0296
c)	96100	.9641	.0279

Fig. 25. Histogram of the asymptotic efficiency of the limiting decorrelating detector for the setting of Fig.24, over the ensemble of discretely valued delays, with n subintervals per chip. a) n=1, b) n=5, c) n=10.

Although these curves are only estimates of the true p.d.f. of the asymptotic efficiency of User 1 in this communication environment, several valid observations can be made. The first is that the mass of the p.d.f. shifts towards higher asymptotic efficiencies as the subdivision becomes finer, and at the same time the variance of the values decreases. Aside from the fact that all asymptotic efficiencies have been in the range [.77,.999], their high sample mean and small sample standard deviation is remarkable. The table in Figure 25 shows the respective values for $n = 1, 5$ and 10. For truly continuous delays as found in a real communication environment these curves indicate that the performance will be superior to the one obtained here. Note the high insensitivity of the asymptotic efficiency of the decorrelating detector to the relative delays, as measured by the small standard deviation of the ensemble. Also remember that these values are independent of the transmissions or energies of the interferers, in contrast to the conventional detector.

After having given a basis for averaging over chip interval delays we will adopt this strategy in the following examples, since it results in substantial computation savings. Our aim is to compare the performance of the decorrelating and conventional detector, for selected communication environments with 3 and more users. The computational difficulty using numerical evaluation of the average error probability expression arises from the fact that the performance of the conventional detector also depends on the transmitted bits, two consecutive of each interferer corrupting each bit to be demodulated, therefore when averaging over them in a K user channel $2^{2(K-1)}$ terms have to be averaged for each set of delays. This precludes large values of K. Of course for a real system average error probability is easy to determine using a known transmitted sequence.

We consider a baseband environment with $K-1$ active equal energy interferers, whose delay relative to each other is fixed. Figure 26, for $K = 3$, shows the 1^{st} user error probability of the conventional receiver versus SNR_1, the signal-to-background-noise ratio of User 1, for different values of the energy ratio $SNR_j \ / \ SNR_1$, averaged over the bit sequences of the two interferers and over the delay of User 1. Also shown are the user error probability of the decorrelating detector for User 1 and the error probability of the single user channel.

From Fig. 26 we see the strong dependence of the performance of the conventional receiver on the relative energies of the active users. While the error probability of the decorrelating detector is invariant to the energy of interfering users, the performance of the conventional receiver deteriorates rapidly for increasing interference, till for an energy ratio above $5dB$ the conventional receiver becomes practically multiple-access limited. (For a sufficiently high level of non-orthogonal interference the error probability of the conventional receiver can be seen to become irreducible. E.g. in the two-user synchronous case, for $\sqrt{w_2}/\sqrt{w_1} = (1+\Delta)/\rho$, where ρ is the normalized crosscorrelation coefficient between the two signature signals and $\Delta \geq 0$, the error probability of the conventional receiver tends to $1/4$ if $\Delta = 0$ and to $1/2$ if $\Delta > 0$ for increasing SNR of User 1). Note that if the energies of all the users are equal the decorrelating detector is around two orders of magnitude better than the conventional receiver at 10 dB. Only if the multiple-access interference level plays a subordinate role compared to the background noise does the conventional detector

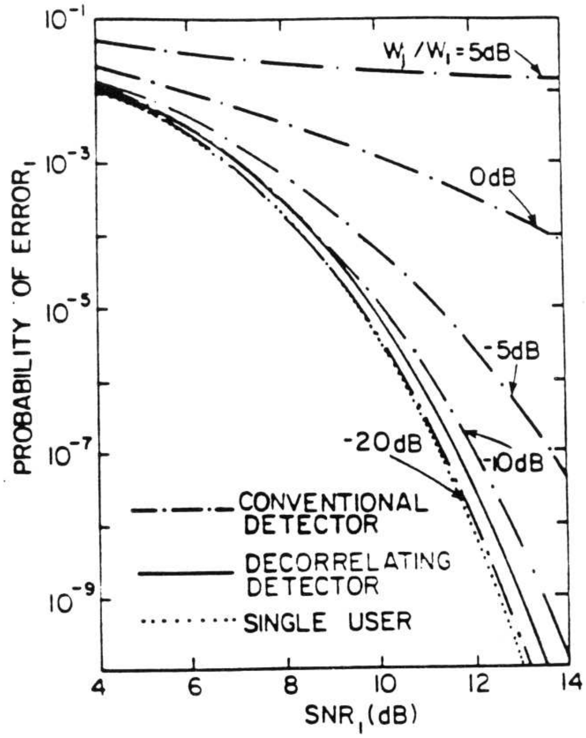


Fig. 26. Error probability of User 1 with 2 active equal energy interferers, each of energy w_j, averaged over the interfering bit sequences and over the delay of User 1, for the decorrelating and conventional receiver versus the SNR of User 1, for m-sequences of length 31 and different interference levels.

outperform the decorrelating detector, which pays a penalty for combatting the interference instead of ignoring it. Similar results were obtained regardless of the actual value of the relative delay between the two interfering users.

Figure 27 shows the same setting as above, in the case $K = 6$. We have used the set of auto-optimal m-sequences of length 31 found in [Pur 79, Fig. A.1], to be optimal with respect to certain peak and mean-square correlation parameters which play an important role in the error probability analysis of the conventional detector.

Comparing Fig. 27 with Fig. 26 we see the same qualitative error probability relation between the two detectors, and again the strong near-far limitation of the conventional receiver. Since there are more active interferers the performance advantage of the decorrelating detector in a near-far

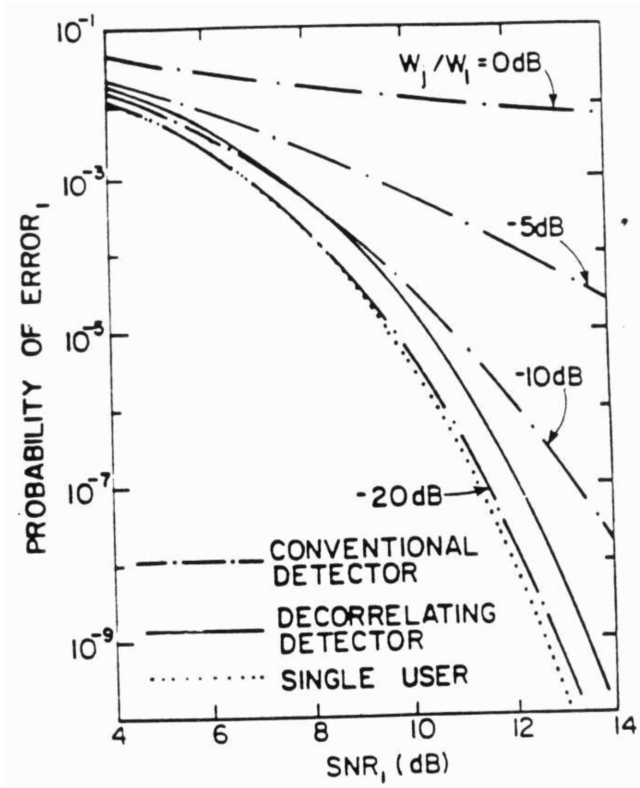


Fig. 27. Same as Fig. 10, with 5 active equal energy interferers.

environment is even more pronounced: if the energies of all the users are equal the decorrelating detector is almost three orders of magnitude better than the conventional receiver at 10 dB.

The same sets of sequences of Figures 26 and 27, were used in Section 3.3, Figures 1 and 2, to compare the error probability of the decorrelating and the conventional detector in the synchronous case. Note that the single error probability curves are lower in the asynchronous case.

Finally, Figure 28 shows the worst-case probability of the conventional detector over the sequences of the interfering users, as given by (4.131), for $K = 10$. The signature sequence set used for $K = 6$ has been expanded - without trying to optimize, as before, with respect to the performance of the conventional detector. The shown error probabilities are typical, varying very little if different sets of delays are used, because of the good crosscorrelation properties of m-sequences.

Overall the generated error probability curves show the pronounced superiority of the decorrelating receiver in a near-far environment, and whenever sufficiently many users are active even if

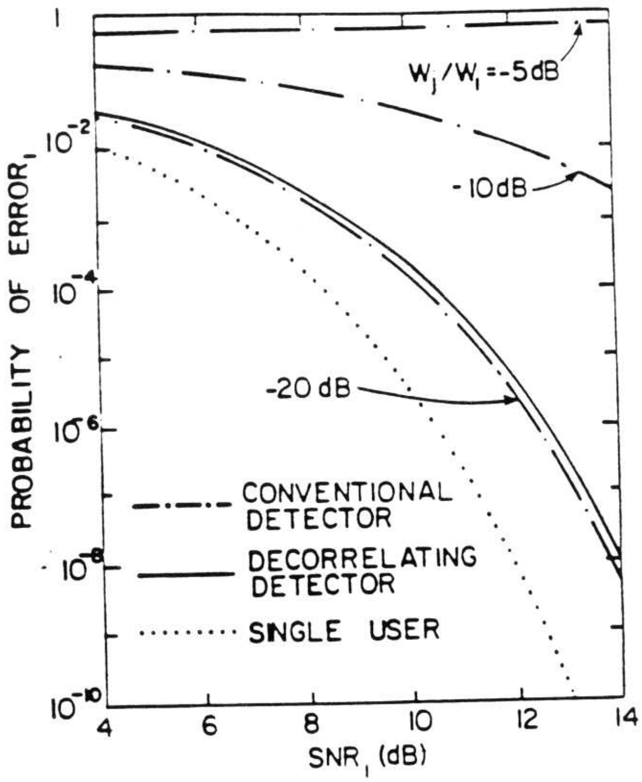


Fig. 28. Worst-case error probability of User 1 w.r.t the bit sequences of the interfering users, with 9 equal energy interferers.

their energies are well below the energy of the desired user. Note, finally, that we have selected signature sequences which have emerged in the literature from attempts to optimize the performance of the conventional receiver. It would be interesting to investigate the possible performance gain of using the decorrelating detector in conjunction with a set of signature sequences optimized for its use (under constraints on bandwidth or structure, e.g. as was done in the conventional case, under the constraint that the sequences be maximal-length shift-register sequences).

4.4 The one-shot decorrelating detector

We now consider a one-shot approach to the decorrelating detector for asynchronous channels, recently suggested in [Ver 88] as an easier-to-compute alternative to the decorrelating detector. For each user the idea is to restrict attention to one bit interval at a time. Let us consider bit 0 of User 1. This bit overlaps with two consecutive bits of each other user, over respectively a subinterval of these bit's durations. Call them the "left sub-bit" and the "right sub-bit". Then if we isolate the waveform received during bit 0 of User 1, the two partial bits corresponding to each interferer can be viewed as two distinct interfering synchronous users, whose waveforms vanish on each other's support set. This situation is equivalent to a $(2K-1)$-user synchronous channel, (see Fig. 15) where each left sub-bit respectively right sub-bit is a distinct user and the waveforms are given by $\{s_1(t), s_i^L(t), s_i^R(t), i = 2, ..., K\}$, where

$$s_i^L(t) = \begin{cases} s_i(t+T-|\tau_i - \tau_1|), & 0 \le t \le |\tau_i - \tau_1| \\ 0, & |\tau_i - \tau_1| \le t \le T \end{cases}$$

$$s_i^R(t) = \begin{cases} 0, & 0 \le t \le |\tau_i - \tau_1| \\ s_i(t-|\tau_i - \tau_1|), & |\tau_i - \tau_1| \le t \le T \ . \end{cases} \tag{4.147}$$

These equivalent waveforms have energies $\{1, e_i, 1-e_i, i = 2, ..., K\}$, where

$$e_i = \int_0^{|\tau_i - \tau_1|} s_i^2\,(t+T-|\tau_i-\tau_1|)\,dt \tag{4.148}$$

is the energy of the left sub-bit. If the delays are continuously valued the probability that $e_i > 0$ is 1. However, in practice, if e_i is too close to 0, users 1 and i would be considered synchronous and the left sub-bit would be discarded. The decorrelating detector for this synchronous $(2K-1)$-user problem is straightforward, and since the resulting matrix $\mathbf{R}$ is nonnegative definite the asymptotic efficiency of User 1 will be strictly positive as long as User 1 is linearly independent of the other users. This requirement is stricter than the LIA, since it requires that the received signal not vanish, regardless of the energies, on each bit-interval, whereas the LIA requires it only for the whole transmission length. However this constraint is still mild. In practice either signature sequences that have this property for all relative delays have to be chosen, or, since the linearly dependent case will occur very infrequently as a function of the relative delays (it will occur with probability 0 for continuously valued delays) in the event of its occurrence the conventional detector decisions can be used, without a measurable effect on error probability. As an example, in the 2-user case, the one-shot matrix $\mathbf{R}$ has the form

$$\mathbf{R} \;=\; \begin{pmatrix} 1 & \rho_{21} & \rho_{12} \\ \rho_{21} & e_2 & 0 \\ \rho_{12} & 0 & 1-e_2 \end{pmatrix} \tag{4.149}$$

The efficiency (and near-far resistance) of the 2-user one-shot decorrelating detector is

$$\eta_1^{os} \;=\; 1 \;-\; \frac{\rho_{21}^2}{e_2} \;-\; \frac{\rho_{12}^2}{(1-e_2)} \tag{4.150}$$

as can be checked by computing R_{11}^{-1}. The first row of the decorrelating detector is, up to a scale factor,

$$\begin{pmatrix} 1 & -\frac{\rho_{21}}{e_2} & -\frac{\rho_{12}}{(1-e_2)} \end{pmatrix} \tag{4.151}$$

which means that for User 1 the received signal is correlated with

$$[\; s_1(t) \;-\; \frac{\rho_{21}}{e_2}\; s_2^L(t) \;-\; \frac{\rho_{12}}{(1-e_2)} s_2^R(t) \;] \,. \tag{4.152}$$

This means that in the absence of noise the output of the detector has a magnitude of

$$\int_0^T [b_1 s_1(t) + b_2^L s_2^L(t) + b_2^R s_2^R(t)]\; [s_1(t) - \frac{\rho_{21}}{e_2}\; s_2^L(t) - \frac{\rho_{12}}{(1-e_2)} s_2^R(t)]\; dt \;=\; b_1 \left(1 - \frac{\rho_{21}^2}{e_2} - \frac{\rho_{12}^2}{(1-e_2)}\right) \tag{4.153}$$

which results in an error probability of $Q(\sqrt{w_1}/\sigma\; \sqrt{1-\rho_{21}^2/e_2 - \rho_{12}^2/(1-e_2)})$. We recognize the efficiency obtained in (4.150).

The one-shot decorrelating detector has a lower complexity than the decorrelating detector $[\mathbf{R}(0) + \mathbf{R}^T(1)z + \mathbf{R}(1)z^{-1}]^{-1}$, at the cost of reduced performance.

Proposition 4.15: The near-far resistance of the one-shot decorrelating detector is upper bounded by that of the limiting decorrelating detector. $\diamondsuit$

Proof: We have established that for a given CDMA environment the decorrelating detector is the linear detector with highest, and moreover optimal, near-far resistance. Therefore for the synchronous one-shot environment the one-shot decorrelating detector has the same attributes. The decision statistic used by the full decorrelating detector is a sufficient statistic, while that used by the one-shot detector is not. That means that the performance achieved by the maximum-likelihood detector for the one-shot model is less or equal than the performance of the maximum-likelihood detector which uses the sufficient statistic. Since this holds for all operating points, in particular the near-far resistance of the maximum-likelihood one-shot detector is less or equal than the near-far resistance of the optimum multiuser detector. But the respective optimum near-far resistances are achieved by the respective decorrelating detectors, thus establishing Proposition 4.15. ∎

Corollary: For each operating point the efficiency of the one-shot decorrelating detector is upper bounded by that of the limiting decorrelating detector, and therefore the error probability of the former is higher or equal to that of the latter. ♢

The first part of the corollary follows from the fact that the efficiency of the decorrelating detector is energy independent, and therefore equal to the near-far resistance. The second is immediate from the definition of efficiency. Note that the error probability of the decorrelating detector is a Q-function. It can be upper bounded using Proposition 3.5.

The traits that distinguish the one-shot decorrelating detector are its memorylessness, its linear time-complexity per demodulated bit, the fact that both its structure and its performance are independent of interfering energies and its near-far resistance. If the relative performance trade-off to the limiting decorrelating detector (which shares all but the first property) is not too severe, the simplicity of the one-shot decorrelating detector makes it an attractive substitute for the limiting decorrelating detector in situations where receiver complexity is a limiting factor.

The following examples illustrate the performance relation between the two detectors. We chose the same set of 3 sequences used when comparing the performance of the decorrelating and conventional detectors in Sections 3.3 and 4.3, Figures 1 and 2. Figures 29 and 30 show the average error probability of the one-shot and limiting decorrelating detector for User 1 for a 2- respectively 3- user baseband environment, averaged over the relative delays, with 10 subdivisions per chip. Note that the average error probability of the one-shot detector is very close to that of the limiting decorrelating detector. While there is always the possibility that the chosen examples might not be representative, they encourage further performance analysis of the one-shot detector.

Figure 31 shows the efficiency of the two detectors as a function of the relative delay of User 1 and 2, for the two-user environment of Figure 29. Ten subintervals per chip were used when discretizing the delay, and there are 31 chips per sequence. Note that the efficiency of the one-shot detector is always less or equal than that of the limiting decorrelating detector, as established by Proposition 4.15. For some delays the performances are equal.

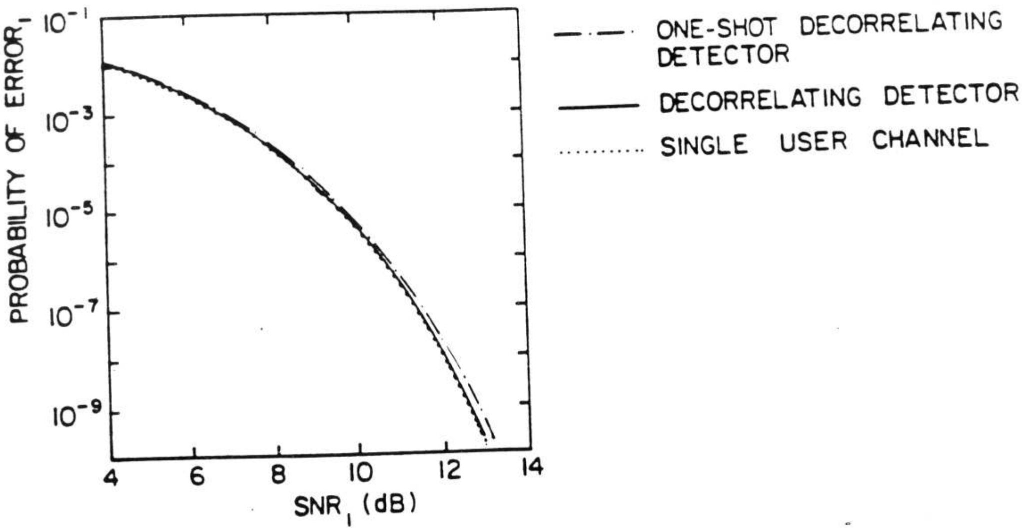


Fig. 29. Error probability of the one-shot and full decorrelating detector for 2 users, averaged over the relative delay, with 10 subdivisions per chip.

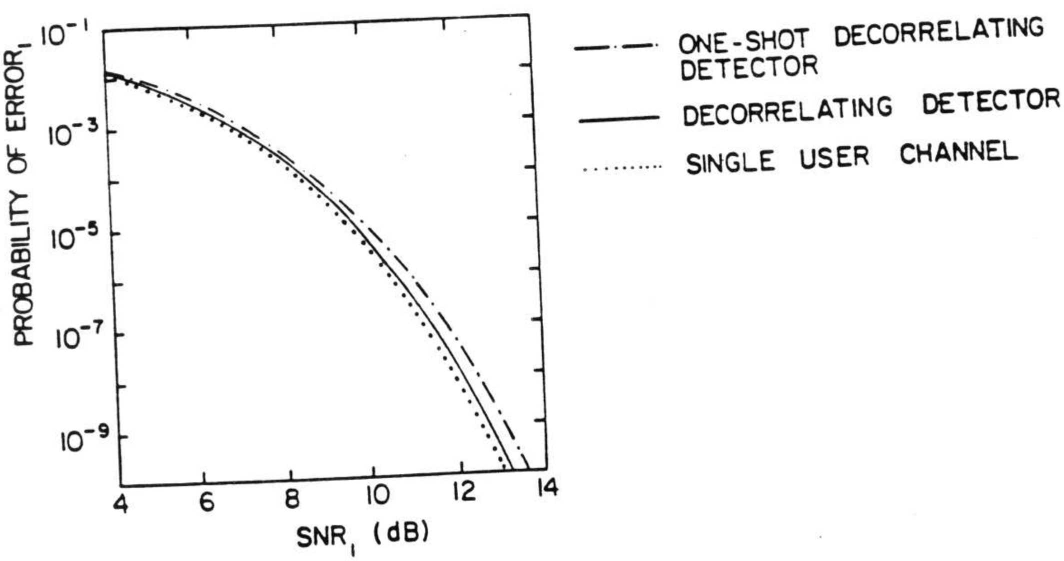


Fig. 30. Same as Fig. 29, for 3 active users.

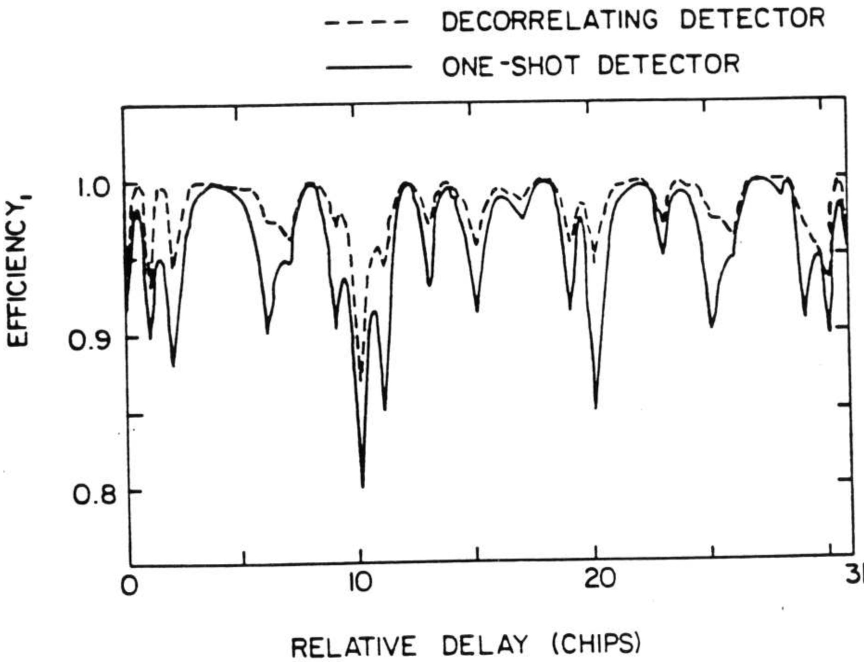


Fig. 31. Asymptotic efficiency of the limiting and of the one-shot decorrelating detector, as a function of the relative delay, for the 2-user environment of Fig. 29.

5. An adaptive algorithm for synchronous channels with unknown signature sequences

An interesting question to ask is what the receiver can do if it does not know all the modulating waveforms, as assumed previously. Such a situation is typical of a decentralized setting, where each receiver is interested only in the information sent by one user, or a proper subset of the user population. Then it is unrealistic to assume that each decentralized receiver knows the waveforms of all the interfering users. Therefore, in the sequel we consider a decentralized DS-SS situation for which proper multiuser demodulation can be achieved despite lack of initial knowledge of the interfering waveforms, due to cooperation between the users in the form of symbol synchronism and usage of a common chip waveform. In this model, also considered in [Poor 88], the K users use the same chip waveform, but the signature sequences of other users are unknown. [Poor 88] gives the maximum likelihood receiver for the general asynchronous case, a receiver which is very complicated for more than two users. Here we consider the case where the transmissions are synchronous, and find an adaptive steepest descent algorithm which converges (in the sense in which stochastic gradient algorithms converge to the solution of the corresponding true gradient algorithm) to a detector which has the desirable property of being asymptotically equivalent to the decorrelating detector as the signal-to-background-noise ratio tends to infinity, while converging to the conventional detector as the interference level tends to zero.

The baseband version of the normalized modulating waveform of each user has the form

$$s_k(t) \;=\; \begin{cases} \sum\limits_{n=1}^{N} c_{kn}\, p\left(t-(n-1)T_c\right), & t\in[0,NT_c) \\ 0, & \text{else} \end{cases} \tag{5.1}$$

where N and T_c are the number of chips per symbol respectively the chip duration, equal for all users, and $p(t)$ is the common unit energy chip waveform, zero outside $[0,T_c)$. Thus each user is characterized by his signature sequence $c_{kn}\in\{\pm 1/\sqrt{N}\}$, $n=1,\ldots,N$. Since the users are synchronous a sufficient statistic for decision on the transmitted information bit of each user is obtained by passing the received waveform through a filter matched to the common chip waveform and sampling at the end of each chip interval. Thus N samples $y_1,\ldots,y_N$ are obtained per symbol interval. Analogously to the representation we had previously, the vector $\mathbf{y}\in\mathbb{R}^N$ of matched filter output samples depends on the transmitted information vector $\mathbf{b}\in\mathbb{R}^K$ via

$$\mathbf{y} \;=\; \mathbf{C}\,\mathbf{W}\,\mathbf{b} + \mathbf{n}\,, \qquad \mathbf{n} \sim N(0,\sigma^2\,\mathbf{I}) \tag{5.2}$$

where

$$\mathbf{C}\in\mathbb{R}^{N\times K} \;\triangleq\; \begin{pmatrix} c_{11} & c_{21} & \cdots & c_{K1} \\ \vdots & \vdots & \vdots & \vdots \\ c_{1N} & c_{2N} & \cdots & c_{KN} \end{pmatrix} \tag{5.3}$$

i.e. the columns of $\mathbf{C}$ are the signature sequences of the users. As before, $\mathbf{W}\in\mathbb{R}^{K\times K}$ is a diagonal matrix containing the square roots of the received energies, $\mathbf{b}\in\mathbb{R}^K$ is the vector containing the K interfering transmitted bits, and $\mathbf{n}\in\mathbb{R}^N$ is the vector of noise components in the matched filter outputs due to the white Gaussian background noise, and has uncorrelated elements because they depend on the noise process on disjoint time intervals.

If the matrix $\mathbf{C}$ were known, the maximum likelihood detector for this problem would select the decisions

$$\hat{\mathbf{b}}^* \in \arg\min_{\mathbf{b}\in\{-1,1\}^K} \; 2\,\mathbf{y}^T\mathbf{C}\mathbf{W}\mathbf{b} \;-\; \mathbf{b}^T\mathbf{W}\mathbf{C}^T\mathbf{C}\mathbf{W}\mathbf{b}\,. \tag{5.4}$$

The case where $\mathbf{C}$ is not known is treated in [Poor 88], and leads to a detector with an extremely complicated structure for more than two users. On the other hand the conventional detector for User k simply decides for

$$b_k^c = \text{sgn} \sum_{n=1}^{N} y_i \, c_{ki} \,, \tag{5.5}$$

which is the particularization of the maximum likelihood detector to the single user case. Note that the c_{ki}, $i = 1, \ldots, N$ are known to the receiver of User k. In the following we derive the structure of the decorrelating detector, i.e. the detector which has the highest near-far resistance among all linear detectors, assuming for the moment that the matrix $\mathbf{C}$ is known.

Since the noise is spherically symmetric the worst-case k^{th} user error probability P_k of the maximum likelihood detector is that of a binary decision between the two closest hypotheses differing in the k^{th} bit, i.e.

$$P_k = Q\,(\frac{1}{\sigma} \min_{\substack{\mathbf{b}_1,\mathbf{b}_2 \\ (\mathbf{b}_1)_k \neq (\mathbf{b}_2)_k}} \frac{1}{2} \parallel \mathbf{CW}(\mathbf{b}_1 - \mathbf{b}_2) \parallel) \,. \tag{5.6}$$

Therefore, the optimum asymptotic efficiency is

$$\eta_k = \frac{1}{w_k} \min_{\substack{\epsilon \in \{-1,0,1\}^K \\ \epsilon_k = 1}} \epsilon^T \mathbf{W} \mathbf{C}^T \mathbf{C} \mathbf{W} \epsilon \tag{5.7}$$

and the optimum near-far resistance is given by

$$\overline{\eta_k} \triangleq \min_{\substack{w_j \geq 0 \\ j \neq k}} \eta_k = \min_{\substack{\mathbf{x} \in R^K \\ x_k = 1}} \mathbf{x}^T \mathbf{C}^T \mathbf{C} \mathbf{x} = \frac{1}{(\mathbf{C}^T\mathbf{C})^{-1}_{kk}} \tag{5.8}$$

where the last equality was proved in the proof of Proposition 3.2. and we have assumed that the matrix $\mathbf{C}^T\mathbf{C}$ is invertible, i.e. that the signature sequences are linearly independent. A linear detector $\mathbf{v} \in \mathbb{R}^K$ decides for

$$\hat{b}_k = \text{sgn } \mathbf{v}^T \mathbf{y} \tag{5.9}$$

and has an error probability equal to

$$P_k = P\,(\hat{b}_k = 1 \,|b_k = -1) = P\,(\mathbf{v}^T \mathbf{n} > -\mathbf{v}^T \mathbf{CWb} \,|b_k = -1) \,. \tag{5.10}$$

Therefore its asymptotic efficiency and near-far resistance are, respectively,

$$\eta_k^v = \max{}^2 \{0, \frac{1}{w_k} \min_{\substack{\mathbf{b} \in \{-1,1\}^K \\ b_k = 1}} \frac{\mathbf{v}^T \mathbf{CWb}}{\sqrt{\mathbf{v}^T \mathbf{v}}}\} \tag{5.11}$$

and

$$\overline{\eta_k^v} = \max{}^2 \{0, \inf_{\substack{\mathbf{y}\in R^K \\ y_k=1}} \frac{\mathbf{v}^T\mathbf{C}\mathbf{y}}{\sqrt{\mathbf{v}^T\mathbf{v}}}\} . \tag{5.12}$$

The optimum near-far resistance achievable by an energy-independent linear detector is

$$\overline{\eta_k^{v*}} = \max{}^2 \{0, \sup_{\substack{\mathbf{v}\in R^N \\ \|\mathbf{v}\|\neq 0}} \inf_{\substack{\mathbf{y}\in R^K \\ y_k=1}} \frac{\mathbf{v}^T\mathbf{C}\mathbf{y}}{\sqrt{\mathbf{v}^T\mathbf{v}}}\} . \tag{5.13}$$

Restricting $\mathbf{v} \in S$, where $S = \{\mathbf{x} \in I\!\!R^N |\ \mathbf{x} = \mathbf{C}\mathbf{z},\ \mathbf{z} \in I\!\!R^K,\ \mathbf{z} \neq 0\}$,

$$\overline{\eta_k^{v*}} \geq \max{}^2 \{0, \sup_{\mathbf{v}\in S} \inf_{\substack{\mathbf{y}\in R^K \\ y_k=1}} \frac{\mathbf{v}^T\mathbf{C}\mathbf{y}}{\sqrt{\mathbf{v}^T\mathbf{v}}}\} \tag{5.14}$$

$$= \max{}^2 \{0, \sup_{\substack{\mathbf{z}\in R^K \\ \|\mathbf{z}\|\neq 0}} \inf_{\substack{\mathbf{y}\in R^K \\ y_k=1}} \frac{\mathbf{z}^T\mathbf{R}\mathbf{y}}{\sqrt{\mathbf{z}^T\mathbf{R}\mathbf{z}}}\} , \qquad \mathbf{R} \triangleq \mathbf{C}^T\mathbf{C} . \tag{5.15}$$

But the last equation is the particularization of (4.39) to the synchronous case, where we obtained as a solution that the maximum is achieved by the decorrelating detector, $\mathbf{z}^* = \mathbf{R}^{-1}\mathbf{u}_k$, and that the value of the maximum equals the near-far resistance $\overline{\eta_k^*}$ of the optimum multiuser detector (Section 4.2, property v)). Therefore from (5.14) $\overline{\eta_k^{v*}} \geq \overline{\eta_k^*}$, i.e. the near-far resistance of the best linear detector is *lower* bounded by that of the optimum detector. This implies that the inequality in (5.14) is an equality and the optimum linear detector $\mathbf{v}^* = \mathbf{C}\mathbf{z}^*$ lies in the K-dimensional subspace spanned by the signature sequences of the K users (this is because the information signal lies in this space and the receiver does not gain anything by correlating with a component outside this space). Therefore the decorrelating detector for the k^{th} user has the form

$$\mathbf{v}^* = \mathbf{C}\ (\mathbf{C}^T\mathbf{C})^{-1}\ \mathbf{u}_k . \tag{5.16}$$

Since the matrix $\mathbf{C}$ is not known to the k^{th} user decentralized decorrelating detector of (5.9), (5.16) can not be implemented. However, assuming an initial training sequence for the k^{th} user (i.e. a sufficiently long sequence of known transmitted bits), we will propose a stochastic gradient algorithm, which has the property that the true gradient algorithm from which it is derived converges to a detector which is asymptotically equivalent to the decorrelating detector as the Gaussian background noise level tends to zero, i.e. under the conditions for which optimality of the decorrelating detector was derived.

First consider the following true gradient algorithm

$$\mathbf{v}^{n+1} = \mathbf{v}^n - \frac{\beta}{2} \frac{\delta}{\delta \mathbf{v}} E\,(\mathbf{v}^T \mathbf{y} - b_k)^2 \,|_{\mathbf{v}=\mathbf{v}^n} \tag{5.17}$$

where

$$E\,(\mathbf{v}^T \mathbf{y} - b_k)^2 = \mathbf{v}^T\,(\mathbf{C}\mathbf{W}^2\mathbf{C}^T + \sigma^2 \mathbf{I})\,\mathbf{v} - 2\,\mathbf{v}^T \mathbf{C}\mathbf{W}\,\mathbf{u}_k + 1\,, \tag{5.18}$$

which is unrealizable because the derivative with respect to $\mathbf{v}$ depends on $\mathbf{C}$, which is unknown. The expectation is taken over the noise and the bits of the interfering users. Assuming that a bypass to the difficulty of not being able to realize the expectation can be found, the behavior of $\mathbf{v}$ as a function of n is as follows. Abbreviate for simplicity the $N \times N$ matrix

$$\Phi \triangleq \mathbf{C}\mathbf{W}^2\mathbf{C}^T + \sigma^2 \mathbf{I} \tag{5.19}$$

which is a positive definite matrix, and

$$\mathbf{p} \triangleq \mathbf{C}\mathbf{W}\,\mathbf{u}_k\,. \tag{5.20}$$

With this notation (5.18) becomes

$$E\,(\mathbf{v}^T \mathbf{y} - b_k)^2 = \mathbf{v}^T\,\Phi\,\mathbf{v} - 2\,\mathbf{v}^T \mathbf{p} + 1 \tag{5.21}$$

$$= (\mathbf{v} - \Phi^{-1}\mathbf{p})^T\,\Phi\,(\mathbf{v} - \Phi^{-1}\mathbf{p}) + 1 - \mathbf{p}^T \Phi^{-1} \mathbf{p} \tag{5.22}$$

which is a quadratic form in the coefficient vector $\mathbf{v}$ and therefore achieves a unique minimum at

$$\mathbf{v}_{\text{opt}} = \Phi^{-1}\mathbf{p}\,. \tag{5.23}$$

From (5.17) the coefficient vector $\mathbf{v}$ of the true gradient algorithm is updated according to

$$\mathbf{v}^{n+1} = \mathbf{v}^n - \frac{\beta}{2} \frac{\delta}{\delta \mathbf{v}} (\mathbf{v}^T\,\Phi\,\mathbf{v} - 2\,\mathbf{v}^T \mathbf{p} + 1)\,|_{\mathbf{v}=\mathbf{v}^n} \tag{5.24}$$

$$= (\mathbf{I} - \beta\,\Phi)\,\mathbf{v}^n + \beta \mathbf{p}\,. \tag{5.25}$$

To see that $\mathbf{v}^n$ converges to the minimizing value (5.23) of the associated cost function, define the error vector

$$\mathbf{q}^n = \mathbf{v}^n - \mathbf{v}_{\text{opt}}\,. \tag{5.26}$$

Then from (5.25), using (5.23)

$$\mathbf{q}^{n+1} = (\mathbf{I} - \beta\Phi)\,\mathbf{q}^n = (\mathbf{I} - \beta\Phi)^n \mathbf{q}^o\,, \tag{5.27}$$

which implies that $\mathbf{q}^n \to 0$ as long as

$$0 \;<\; \beta \;<\; \frac{2}{\lambda_{\max}(\Phi)} \tag{5.28}$$

where $\lambda_{\max} = \max\{\lambda_i | \lambda_i$ is an eigenvalue of $\Phi\}$. With this we have shown that for the range of β of (5.28) the coefficient vector of the true gradient algorithm converges to $\mathbf{v}^\infty = \mathbf{v}_{\mathrm{opt}}$, or, from (5.23),

$$\mathbf{v}^\infty \;=\; (\mathbf{C}\mathbf{W}^2\mathbf{C}^T + \sigma^2\mathbf{I})^{-1}\;\mathbf{C}\mathbf{W}\;\mathbf{u}_k\;. \tag{5.29}$$

It is intuitively apparent that the coefficient vector converges to the value which minimizes the chosen cost function (here the expected value of the mean-square output error), if the step-size β is well chosen, since the coefficient vector is adjusted in a direction opposite to the gradient of the cost-function at each iteration.

The reason why the unrealizable true gradient algorithm is interesting for the multiuser demodulation problem at hand will become clear from Proposition 5.1 and the discussion thereafter. Here we first address the realizability issue, in order to make clear that we are not assuming something we want to obtain. The difficulty we are facing, namely that we want to adjust the detector such as to minimize a cost function which depends on unknown parameters, is a standard problem in adaptive filtering. A well-established and much used bypass consists in using a so-called *stochastic gradient* instead of the true gradient. The corresponding (realizable) algorithm is obtained from the true gradient algorithm by dropping the expectation, thus avoiding the need to know $\mathbf{C}$. Denote by $b_k(n)$ the n^{th} bit transmitted by User k, and by $\mathbf{y}^n$ the matched filter output vector corresponding to the n^{th} bits. Let $\mathbf{e}_n$ denote the output error at each iteration, i.e. $\mathbf{e}_n = \mathbf{v}^{nT}\mathbf{y}^n - b_k(n)$, which is a random variable. Then the stochastic gradient algorithm (SGA) derived from the true gradient algorithm of (5.17) is the following.

$$\mathbf{v}^{n+1} \;=\; \mathbf{v}^n \;-\; \frac{\beta}{2}\,\frac{\delta}{\delta\mathbf{v}^n}\,\mathbf{e}_n^2 \;=\; \mathbf{v}^n \;-\; \beta\,\mathbf{e}_n\,\mathbf{y}^n \tag{5.30}$$

$$=\; (\mathbf{I} \;-\; \beta\,\mathbf{y}^n\mathbf{y}^{nT})\,\mathbf{v}^n \;+\; \beta\,b_k(n)\,\mathbf{y}^n\;. \tag{5.31}$$

Since we assume an initial training sequence, this algorithm is realizable. The question that arises is whether the convergence of the expected value of $\mathbf{v}^n$, the latter now being random, to $\mathbf{v}^\infty$ of

(5.29) is preserved. The expected value of the coefficient vector $\mathbf{v}^n$ evolves from (5.31) according to

$$E\left[\mathbf{v}^{n+1}\right] = E\left[\left(\mathbf{I} - \beta\, \mathbf{y}^n \mathbf{y}^{nT}\right) \mathbf{v}^n\right] + \beta\, E\left[b_k(n)\, \mathbf{y}^n\right] \tag{5.32}$$

$$= \left(\mathbf{I} - \beta\, E\left[\mathbf{y}^n \mathbf{y}^{nT}\right]\right) E\left[\mathbf{v}^n\right] + \beta\, E\left[b_k(n)\, \mathbf{y}^n\right] \tag{5.33}$$

$$= \left(\mathbf{I} - \beta\, \Phi\right) E\left[\mathbf{v}^n\right] + \beta\, \mathbf{p} \tag{5.34}$$

which implies that the expected value of the error vector tends to zero since

$$E\left[\mathbf{q}^{n+1}\right] = (\mathbf{I} - \beta\Phi)\, E\left[\mathbf{q}^n\right] = (\mathbf{I} - \beta\Phi)^n E\left[\mathbf{q}^o\right], \tag{5.35}$$

similarly to (5.27). Equation (5.33) makes use of the independence of the vectors $\mathbf{v}^n$ and $\mathbf{y}^n$, which holds since $\mathbf{v}^n$ only depends on the past values of $\mathbf{y}^n$, which are independent of the present due to synchronism. This independence is crucial for the convergence analysis of the SGA, and while it is given in the problem at hand, this is not the case in most applications. Nevertheless the *independence assumption* is made in the above references, though the authors elaborate on both its necessity and its inaccuracy. The present problem is noteworthy in this respect for the fact that the independence assumption is true.

Another convergence measure of interest for an adaptive algorithm is the decrease in time of the output error (this is usually done via a mean-squared error analysis). In the literature the convergence of the output mean-square error of the SGA has been much studied, although a rigorous analysis has apparently not yet been given (cf. [Hon, pg 247]). [Hon, 7.1] or [Ben, 8.2.3] give a convergence analysis of the output mean squared error for a SGA of the form of (5.31), under certain approximations. We showed via the proof of convergence of the mean coefficient value that the problem at hand is better behaved than the problem treated in the references, while allowing for an analogous solution.

In the above references it is obtained that the minimum mean-square error of the SGA is higher than that of the true gradient algorithm, due to the statistical fluctuation of the filter coefficients. More precisely, if E_{min} is the residual mean-square error of the true gradient algorithm, which can be seen to equal $1 - \mathbf{p}^T \Phi^{-1} \mathbf{p}$ by inserting (5.23) into (5.22), then the excess mean-square error E_{ex}^n of the SGA is shown to converge to

$$E_{\text{ex}}^{\infty} = \frac{\beta \sum\limits_{i=1}^{N} \lambda_i}{2 - \beta \sum\limits_{i=1}^{N} \lambda_i} E_{\text{min}} \tag{5.36}$$

as long as β is in the range

$$0 \; < \; \beta \; < \; \frac{2}{N\sigma^2 + \sum_{k=1}^{K} w_k} \; . \tag{5.37}$$

This condition on β is stricter than the one in (5.28), since

$$\begin{aligned} \lambda_{\max}(\Phi) \; &< \; \sum_{i=1}^{N} \lambda_i(\Phi) \; = \; \sum_{i=1}^{N} \Phi_{ii} \; = \; \sum_{i=1}^{N} (\mathbf{C}\mathbf{W}^2\mathbf{C}^T + \sigma^2\mathbf{I})_{ii} \\ &= \; N\sigma^2 \; + \; \sum_{i=1}^{N} \sum_{k=1}^{K} w_i \; c_{ik}^2 \; = \; N\sigma^2 + \sum_{k=1}^{K} w_k \; . \end{aligned} \tag{5.38}$$

Despite the longstanding want of analytical results which do not make use of approximations, the convergence of the SGA is a well-investigated issue. The SGA is widely used in practice and versions with good convergence properties are well established. Therefore, we feel we are reducing the problem we are examining to a known problem, if we reduce it to a solution in terms of the true gradient algorithm of (5.17), and propose use of the stochastic gradient modification in its implementation. (For the sake of completeness, the existence of other modifications should be mentioned, e.g. time-averaging to substitute the expectation in (5.17)). In the following result, we show that the true gradient algorithm has the property that the coefficient vector converges to the decorrelating detector in the limit of $\sigma \to 0$, and converges to the conventional detector in the opposite case, when the power of the multiuser interference goes to zero. While the first property was what we were looking for, the second one is equally desirable, as will be explained in the subsequent discussion.

Proposition 5.1 :

$$\lim_{\sigma \to 0} \; \mathbf{v}^{\infty} \; = \; \frac{1}{\sqrt{w_k}} \; \mathbf{C} \; (\mathbf{C}^T\mathbf{C})^{-1} \; \mathbf{u}_k \tag{5.39}$$

$$\lim_{\substack{w_j \to 0 \\ j \neq k}} \; \mathbf{v}^{\infty} \; = \; \frac{\sqrt{w_k}}{\sigma^2 + w_k/N} \; \mathbf{C} \; \mathbf{u}_k \tag{5.40}$$

i.e., up to a scale factor the true gradient algorithm of (5.17) converges to the decorrelating detector as the Gaussian background noise level goes to zero, and converges to the optimum single-user detector as the power of the multiuser interference goes to zero. $\diamondsuit$

Note that since a sign decision is taken, scale factors do not matter.

Proof : We first show that, from (5.29),

$$\mathbf{v}^{\infty} = (\mathbf{C}\mathbf{W}^2\mathbf{C}^T + \sigma^2\mathbf{I})^{-1}\,\mathbf{C}\mathbf{W}\,\mathbf{u}_k \tag{5.41}$$

$$= \frac{1}{\sqrt{w_k}}\,\mathbf{C}\,(\mathbf{C}^T\mathbf{C} + \sigma^2\mathbf{W}^{-2})^{-1}\,\mathbf{u}_k\,. \tag{5.42}$$

To show (5.42) we use the matrix inversion lemma (e.g. [Hay]), which says that for two positive definite matrices $\mathbf{B}$ and $\mathbf{D}$

$$(\mathbf{B}^{-1} + \mathbf{C}\mathbf{D}^{-1}\mathbf{C}^T)^{-1} = \mathbf{B} - \mathbf{B}\mathbf{C}(\mathbf{D} + \mathbf{C}^T\mathbf{B}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{B}\,. \tag{5.43}$$

Using this lemma

$$\mathbf{v}^{\infty}\,[\frac{1}{\sigma^2}\,\mathbf{I} - \frac{1}{\sigma^4}\,\mathbf{C}\,(\mathbf{W}^{-2} + \frac{1}{\sigma^2}\mathbf{C}^T\mathbf{C})^{-1}\,\mathbf{C}^T]\,\mathbf{C}\mathbf{W}\,\mathbf{u}_k \tag{5.44}$$

$$\frac{1}{\sigma^2}[\mathbf{C}\mathbf{W}\,\mathbf{u}_k - \mathbf{C}\,(\sigma^2\mathbf{W}^{-2} + \mathbf{C}^T\mathbf{C})^{-1}(\mathbf{C}^T\mathbf{C} + \sigma^2\mathbf{W}^{-2} - \sigma^2\mathbf{W}^{-2})\mathbf{W}\mathbf{u}_k] \tag{5.45}$$

$$\frac{1}{\sqrt{w_k}}\,\mathbf{C}\,(\mathbf{C}^T\mathbf{C} + \sigma^2\mathbf{W}^{-2})^{-1}\,\mathbf{u}_k\,. \tag{5.46}$$

From here (5.39) follows directly by taking letting $\sigma \to 0$ in (5.46), and (5.40) follows by letting $w_j \to 0, j \neq k$ in (5.41), and noticing that

$$\lim_{\substack{w_j \to 0 \\ j \neq k}} \mathbf{C}\mathbf{W}^2\mathbf{C}^T = \mathrm{diag}\,(\,x_j\,|\,x_j = 0,\ j \neq k,\ x_k = w_k/N)\,. \tag{5.47}$$

∎

Returning to the true gradient algorithm, the decision statistic after convergence of the tap weights is

$$\mathbf{v}^{\infty^T}\mathbf{y} = b_k - \frac{\sigma^2}{\sqrt{w_k}}\,\mathbf{u}_k^T\,(\mathbf{C}^T\mathbf{C} + \sigma^2\mathbf{W}^{-2})^{-1}\mathbf{W}^{-1}\mathbf{b} + \mathbf{v}^{\infty^T}\mathbf{n} \tag{5.48}$$

and the resulting mean-squared error is given by

$$E\,(\mathbf{v}^{\infty^T}\mathbf{y} - b_k)^2 = \frac{\sigma^2}{w_k}\,(\mathbf{C}^T\mathbf{C} + \sigma^2\mathbf{W}^{-2})^{-1}_{kk}\,. \tag{5.49}$$

Hence unless the inverse is very ill behaved, the mean-squared error is very small in the high SNR region and will matter very little above a certain SNR level, the more so since we are only interested in the sign of $\mathbf{v}^{\infty^T}\mathbf{y}$. To estimate the deviation of $(\mathbf{C}^T\mathbf{C} + \sigma^2\mathbf{W}^{-2})^{-1}$ from $(\mathbf{C}^T\mathbf{C})^{-1}$ we expand the former as

$$(\mathbf{C}^T\mathbf{C} + \sigma^2\mathbf{W}^{-2})^{-1} = [(\mathbf{C}^T\mathbf{C})\,(\mathbf{I} + \sigma^2(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{W}^{-2})^{-1}$$

$$= \sum_{i=0}^{\infty} \sigma^i\,[-(\mathbf{C}^T\mathbf{C})^{-1}\,\mathbf{W}^{-2}]^i\,(\mathbf{C}^T\mathbf{C})^{-1} \tag{5.50}$$

which is possible iff the spectral radius of the matrix $(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{W}^{-2}\sigma^2$ is less than 1 ([Ben, p. 587], [Horn, 5.8]), which is satisfied if the SNR is high enough. The first term in the above series is the desired term. From here the error can be estimated as in [Horn, 5.8], to obtain

$$\frac{\|(\mathbf{C}^T\mathbf{C})^{-1} - (\mathbf{C}^T\mathbf{C}+\sigma^2\mathbf{W}^{-2})^{-1}\|}{\|(\mathbf{C}^T\mathbf{C})^{-1}\|} \leq \frac{\kappa \parallel \sigma^2\mathbf{W}^{-2} \parallel / \parallel \mathbf{C}^T\mathbf{C} \parallel}{1 - \kappa \parallel \sigma^2\mathbf{W}^{-2} \parallel / \parallel \mathbf{C}^T\mathbf{C} \parallel} \tag{5.51}$$

$$= \frac{\sigma^2 \; / \; w_{\min}}{\lambda_{\min} \; (\mathbf{C}^T\mathbf{C}) \; - \; \sigma^2 \; / \; w_{\min}} \tag{5.52}$$

as long as $\sigma^2 < w_{\min} \; \|\mathbf{C}^T\mathbf{C}\|$, when $\| \; . \; \|$ denotes the spectral norm. In this case $\kappa \triangleq \lambda_{\max}(\mathbf{C}^T\mathbf{C})/\lambda_{\min} \; (\mathbf{C}^T\mathbf{C})$, $\|\mathbf{C}^T\mathbf{C}\| = \lambda_{\max}(\mathbf{C}^T\mathbf{C})$ and $\|\sigma^2\mathbf{W}^{-2}\| = \sigma^2/w_{\min}$, and (5.52) results. Hence, as long as $\lambda_{\min} \; (\mathbf{C}^T\mathbf{C})$ is sufficiently large, the relative error is of the order of the inverse of the smallest SNR ratio. A similar analysis can be carried out for the deviation from the conventional detector.

Discussion: The mathematical relation between the unrealizable gradient algorithm of (5.17) and its realizable stochastic modification according to (5.31) are well-known, and there is much evidence that the average behavior of the stochastic algorithm is such that if the step size β is appropriately chosen, the function of the desired unrealizable algorithm can be closely approximated. The unrealizable detector of (5.19) was seen to converge to the decorrelating detector in the limit as the Gaussian background noise level goes to zero, and to deviate little from it in the high SNR region. This is desirable, since in this region the multiuser interference is the main impairment on the common channel, and the decorrelating detector is the detector which eliminates this interference. However we have seen in Figures 1 and 2 for the synchronous case and Figures 26, 27, 28 for the asynchronous case, that the conventional detector performs better than the decorrelating detector if the multiuser interference is very low. This is because the conventional detector is not penalized in this region for ignoring the interference from other users, whereas the decorrelating detector eliminates this small interference at the expense of enhancing the background noise, which in this region is the main distorting factor. Therefore, the property of the detector to converge to the conventional detector in this region is desirable.

For these reasons, inasmuch as the true gradient algorithm can be approximated satisfactorily by the SGA or an equivalent modification, we have obtained a detector which has the capability of *adapting to the dominant cause of channel distortion* in the limit as one single effect (multiuser interference respectively background noise) predominates. We conjecture that the compromise achieved by (5.42) in the region where no single factor significantly outweighs the other is a good demodulation strategy for an operating region where both distortion causes are of comparable magnitude.

It would be interesting to find an equivalent adaptive algorithm which does not require an initial training sequence ("blind adaptation"). Previous research on the subject has not been successful. Since the decorrelating detector effectively inverts the channel transfer function, the following citation from [Ver 84b] is pertinent to this problem: "no such function is known to result in global convergence to the inverse of the channel when the input consists of binary data".

6. Conclusions

The main contribution of this thesis is to have shown that the near-far resistance of optimum multiuser detection of Code-Division multiplexed signals in white Gaussian channels can be achieved by linear detectors, thus providing an effective remedy to the well-known *near-far problem*, which - in contrast to the maximum likelihood detector - is implementable even for a large number of users. This is possible due to the fact that the asymptotic efficiency functional of linear multiuser detection has a saddle point, and the near-far resistance of the optimum multiuser receiver can be written as that of a variable linear receiver which is optimally suited for the respective operating point. Making use of the aforementioned saddle-point property the optimum near-far resistance is equal to the highest near-far resistance achievable by a (fixed) linear receiver. This receiver is then found explicitly. Thus there exists a linear detector which does not exhibit the main limitation of the conventional single-user detector used in practice, which occurs even if signals with very low crosscorrelations are assigned to the users, namely the near-far problem.

We considered both the synchronous and the asynchronous CDMA channel shared by simultaneous users: they use the same bit duration T when transmitting and have made their respective signature sequences common knowledge. This setting is the same as that in [Ver 86a], [Ver 86b], where the maximum-likelihood multiuser detector was derived. In a system where this knowledge is restricted, the presented results predict significant performance improvement if the signature sequences of the more powerful interferers are known to each user's receiver.

For the *asynchronous* channel we derived the near-far optimum linear receiver, the decorrelating detector, which does not have the near-far problem of the conventional single-user receiver and, it turns out, of any energy-independent linear receiver except the decorrelating one. Its structure is that of an inverse to the equivalent transfer function between transmitter and receiver, such that the users are decoupled before the sign decision. Obviously then, since each user's decision statistic is now independent of other users transmissions, even very powerful interference can be combatted, evidently as long as the synchronization and carrier phase acquisition mechanisms of the weak user's receiver do not fail. The price paid for eliminating the multiuser interference is an increase in the variance of the Gaussian background noise, which is also the reason why the decorrelating detector is not the optimum multiuser detector. This means that in applications a decision first has to be made as to whether the background noise or the multiuser interference is the dominating factor. In the first case the single-user detector should be used, in the second the decorrelating detector.

Three properties make the decorrelating detector particularly attractive in near-far environments with a large number of users: its linear time-complexity per demodulated bit, the fact that its implementation does not require knowledge of the received energies, and the desirable attributes of its bit-error rate, namely that it is independent of received energies and that it offers the same degree of near-far resistance as the optimum multiuser detector.

We gave conditions for existence of the decorrelating detector and showed that for continuous, independent delays - the usual conditions in a completely asynchronous channel - these conditions are satisfied almost surely. Also, the decorrelating detector does not exist if and only if the optimum multiuser near-far resistance is zero.

In a multiuser environment where K users transmit N-bit sequences, the decorrelating detector was described as the inverse of an $NK * NK$ equivalent synchronous system matrix. In this case the receiver is the near-far optimum linear combination of the front-end matched filter outputs, i.e. is a new matched filter, matched to the multiuser environment. It was shown that as the transmitted sequence length tends to infinity the decorrelating detector tends to a time invariant linear filter which is stable and noncausal. Since the filter is stable the noncausal part of the impulse response can be truncated in practice after a suitable delay, to the desired degree of accuracy. In applications where each receiver is interested in demodulating the information transmitted by only one user, it is easy to decentralize the K-user decorrelating receiver since it can be implemented as K separate (continuous-time) single-input (discrete-time) single-output filters. Each of those filters can be viewed as a modification of the conventional single-user matched filter, where instead of correlating the channel output with the signature waveform of the user of interest, we use its projection on the subspace orthogonal to the space spanned by the interfering signals. Here a comment is appropriate: if the filter is actually an approximation to the decorrelating receiver, due to, for example, finite accuracy in the computation of the crosscorrelations or truncation of the impulse response, it will no longer be orthogonal to the subspace of the interfering signals and therefore it will not be near-far resistant in the worst-case sense adopted in this work. However, for practical purposes we do not need near-far resistance with respect to *all* possible interfering energies - even after the sync and acquisition mechanisms of the weak user have long failed - but rather with respect to a region of interfering energies, e.g., dictated by the signal processing front-end. Since decorrelating corresponds to a projection orthogonal to the multiuser interference, and truncation effects a tilt in the projection plane, by truncating appropriately far the effect on the bit error rate can be made arbitrarily small, preserving near-far resistance *with respect to the desired energy range.*

A general expression as well as a lower bound for the asymptotic efficiency achieved by the limiting decorrelating detector were given. Since the asymptotic efficiency of the decorrelating detector is a nonincreasing function in the number of users, it would be beneficial to investigate a practical implementation where the size of the filter is modified to take into account only the subset of active users, which could be significantly smaller than the total user population.

Computation of the transfer function of the limiting decorrelating detector for K users involves inverting a K by K matrix whose elements are monomials in z respectively z^{-1}. This is due to the memory involved in the observed multiuser process. As proposed in [Ver 88], a simple though clearly suboptimal way to get around this difficulty is to take a one-shot approach, where the k^{th} user receiver considers the received process during each symbol interval of User k independently of all the rest, as if it had resulted from a synchronous process where the two interfering bits of each asynchronous user are viewed as coming from two different synchronous users with smaller energy. We have investigated using this simple model and incorporating the decorrelating philosophy to ensure near-far resistance, and have shown that the performance of the one-shot decorrelating detector is upper bounded by that of the limiting decorrelating detector for each operating point. The performance results obtained for a two- and three-user example yielded only a small performance reduction compared to the decorrelating detector, and motivate consideration in practical situations of the computationally much simpler one-shot approach.

The previous results can be particularized to the *synchronous* channel. For the latter case the best linear detector has been derived, as a function of the received energies, for comparison purposes with the decorrelating detector. One interesting result we obtained is that there is a region of energies and crosscorrelations where the best linear detector achieves the asymptotic efficiency of the optimum multiuser detector, while in another such region it coincides with the decorrelating detector. Other precise analytic results were not feasible. The worst-case complexity of the algorithm obtained in order to find the best linear detector is exponential in the number of users. In a fixed-energy environment this computation needs to be carried out only once, hence the real-time time-complexity per bit is linear, in contrast to the optimum multiuser detector. Nevertheless, this feature and the energy dependence of both its structure and its performance penalize the best linear detector in comparison with the decorrelating detector. The only requisite necessary for the signal of a user to be detected reliably by the decorrelating detector regardless of the level of multiple-access interference, is that it does not belong to the subspace spanned by the other signals - a very mild constraint that should be compared to the condition necessary for reliable detection by the conventional single-user detector, i.e. that the signal is *orthogonal* to all the other signals.

An iterative decision-feedback scheme has been proposed, with the decorrelating detector in the first stage to ensure near-far resistant initial decisions. Previous-stage decisions on the bits transmitted by the other users are used to obtain estimates of the j^{th} noise components in the decision statistic for all $j \neq k$, then the correlation of the noise components is exploited to obtain and then subtract an estimate of the k^{th} noise component. Lower bounds on the second-stage asymptotic efficiency and near-far resistance of this scheme have been obtained and it has been shown that unit asymptotic efficiency can be achieved in a given energy range. Conditions of the received energies have been given to ensure a performance improvement over the decorrelating detector. The scheme is no longer energy independent, and an example illustrated that there is a range of energies where feedback can decrease performance, due to the fact that for that specific parameter range the obtained estimates may not be reliable enough. However near-far resistance was shown to be preserved. The benefits of partial feedback of previous decisions were investigated, where feedback from unreliable users is omitted, and an algorithm was given which finds the best feedback set. A special case thereof is the empty set, which corresponds to just the decorrelating detector. The use of this algorithm guarantees optimum near-far resistance, together with an asymptotic efficiency which is lower bounded by that of the decorrelating detector. The only additional feature required by this algorithm is estimation of the operational energy range. An example showed that performance can be significantly increased in this way.

Finally, the situation has been considered when the receiver for each user has no knowledge of the modulating waveforms of the other users, a situation which is typical for decentralized reception. For the special case of a synchronous channel and DS-SS signaling with common chip waveform but unknown signature sequences, an adaptive algorithm has been found, which uses a training sequence and converges to a detector which adapts to the communication impairment situation on the channel. Namely, in the low background noise region it approaches the decorrelating detector, while in the low multiuser interference region it approaches the conventional detector, each being the detector of choice under the corresponding conditions. An interesting question for further work is whether an equivalent scheme can be found for the asynchronous channel, and/or without knowledge of an initial training sequence (i.e., blind adaptation).

References

[Ben] S. Benedetto, E. Biglieri and V. Castellani, *Digital Transmission Theory,* Prentice-Hall, 1987.

[Bou] T.L. Boullion, P.L. Odell, *Generalized Inverse Matrices,* Wiley-Interscience, 1971.

[Bro] I.N. Bronshtein and K.A.Semendyayev, *Handbook of Mathematics,* Frankfurt: Verlag Harri Deutsch, 1985.

[Due 87] A. Duel-Hallen, *"Detection algorithms for intersymbol interference channels",* Ph.D. dissertation, Cornell University, Aug.1987.

[Ett 75] W. Van Etten, *"An optimum linear receiver for multiple channel digital transmission systems", IEEE Trans. on Communications,* pp. 828-834, Aug.1975.

[Ett 76] W. Van Etten, *"Maximum likelihood receiver for multiple channel transmission systems", IEEE Trans. on Communications,* vol. COM-24, pp. 276-283, Feb.1976.

[For 72] G.D. Forney, *"Maximum likelihood sequence estimation of digital sequences in the presence of intersymbol interference", IEEE Trans. on Information Theory,* vol. IT-18, no.3, pp. 363-378, May 1972.

[Gal 88] P.A. Galko and S. Pasupathy, *"Optimization of linear receivers for data communication signals", IEEE Trans. on Information Theory,* vol. 34, pp. 79-93, Jan.1988.

[Gar 79] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness,* San Francisco: Freeman, 1979.

[Gar 80] F.D. Garber and M.B. Pursley, *"Optimal phases of maximal-length sequences for asynchronous spread-spectrum multiplexing", Electron. Letters,* vol. 16, pp. 756-757, Sept. 1980.

[Ger 82] E.A. Geraniotis and M.B. Pursley, *"Error probability for direct-sequence spread- spectrum multiple-access communications-Part II: Approximations", IEEE Trans. Comm.,* vol. COM-30, pp. 985-995, May 1982.

[Hay] S. Haykin, *"Introduction to Adaptive Filters",* Macmillan, 1984.

[Hel] C.W. Helstrom, *Statistical Theory of Signal Detection,* 2nd edition, London: Pergamon, 1968.

[Hon] M.L. Honig and D.G. Messerschmitt, *Adaptive Filters: Structures, Algorithms and Applications,* Kluwer Academic Publishers, 1985.

[Hor 75] D. Horwood and R. Gagliardi, *"Signal design for digital multiple access communications", IEEE Trans. on Communications,* vol. COM-23, pp. 378-383, March 1975.

[Horn] R.A. Horn and C.R. Johnson, *Matrix Analysis,* Cambridge University Press, 1985.

[Kay 70] A.R. Kaye and D.A. George, *"Transmission of multiplexed PAM signals over multiple channel and diversity systems", IEEE Trans. on Communication Technology,* vol. COM-18, pp. 520-525, Oct.1970.

[Kob 71] H. Kobayashi *"Correlative level coding and maximum-likelihood decoding", IEEE Trans. Information Theory* vol.IT-17, pp.586-594, Jan 1971.

[Lan] P. Lancaster and M. Tismenetsky, *The Theory of Matrices,* Academic Press, 1985.

[Luc 71] R.W. Lucky *"Automatic equalization for digital communication", Bell System Technical Journal* vol. 44, nr. 4, pp. 547-588, Apr 1965.

[Lue] D. Luenberger, *Optimization by Vector Space Methods,* Wiley, 1969.

[Lup 86] R. Lupas and S. Verdú, *"Asymptotic Efficiency of Linear Multiuser Detectors", Proceedings of the 25th IEEE Conference on Decision and Control,* Athens, Greece, December 1986, pp.2094-2100.

[Lup 88] R. Lupas and S. Verdú, *"Optimum near-far resistance of linear detectors for Code-Division Multiple-Access channels", Abs. 1988 IEEE Int. Symp. Inf. Th.,* p.14, Kobe, Japan.

[Lup 89a] R. Lupas and S. Verdú, *"Linear multiuser detectors for synchronous Code-Division Multiple-Access channels", IEEE Trans. on Information Theory,* vol. IT-35, No.1, Jan. 1989.

[Lup 89b] R. Lupas and S. Verdú, *"Near-far resistance of multiuser detectors in asynchronous channels", IEEE Trans. on Communications,* vol. COM-38, No.4, April 1990.

[Pap] C.H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity,* Englewood Cliffs, NJ: Prentice-Hall, 1982.

[Pic 82] R.L. Pickholtz, D.L. Schilling and L.B. Milstein *"Theory of Spread-Spectrum-A Tutorial", IEEE Trans. on Communications,* vol. COM-30, May 1982.

[Poor 80] H.V. Poor, *"Signal detection in multiple access channels",* U.S. Army Research Office Proposal (Contract DAAG29-81-K-0062 to Coordinated Science Laboratory, University of Illinois), 1980.

[Poor] H.V. Poor, *An Introduction to Signal Detection and Estimation,* Springer Verlag, 1988.

[Poor 88] H.V. Poor and S. Verdú, *"Single-user detectors for multiuser channels", IEEE Trans. on Communications,* vol. COM-36, pp. 50-60, Jan. 1988.

[Pur 79] M.B. Pursley and H.F.A. Roefs, *"Numerical evaluation of correlation parameters for optimal phases of binary shift-register sequences", IEEE Trans. Comm.,* vol. COM-27, pp.1597-1604, Oct. 1979.

[Pur 81] M.B. Pursley, *"Spread-Spectrum Multiple-Access Communications", in " Multi-User Communication Systems", ed. G. Longo, CISM Courses and Lectures no. 265,* Springer Verlag,1981.

[Pur 82] M.B. Pursley, D.V. Sarwate and W.E. Stark, *"Error probability for direct-sequence spread-spectrum multiple-access communications-Part I: Upper and lower bounds", IEEE Trans. Comm.,* vol. COM-30, pp. 975-984, May 1982.

[Rud] W. Rudin, *Principles of mathematical analysis,* McGraw Hill, 1976.

[Rus 88] C. Rushforth and Z. Xie, *"Multiuser signal detection using sequential decoding", 1988 IEEE Military Communications Conf.,* San Diego, Oct. 1988.

[Sar 80] D.V. Sarwate and M.B. Pursley, *"Crosscorrelation properties of pseudorandom and related sequences", Proceedings of the IEEE,* vol. 68, no. 5, May 1980.

[Scho 77] R.A. Scholtz, *"The Spread-Spectrum Concept", IEEE Trans. on Communications,* vol. COM-25, no.8, Aug.1977.

[Sch 79] K.S. Schneider, *"Optimum Detection of Code Division Multiplexed Signals", IEEE Trans. Aerosp. Electron. Syst.,* vol. AES-15 (1), pp. 181-185, Jan 1979.

[Sch 80] K.S. Schneider, *"Crosstalk resistant receiver for M-ary multiplexed communications", IEEE Trans. Aerosp. Electron. Syst.,* vol. AES-16, pp. 426-233, Jul.1980.

[Shn 67] D.A. Shnidman, *"A generalized Nyquist criterion and an optimal linear receiver for a pulse modulation system", The Bell System Technical Journal,* pp. 2163-2177, Nov.1967.

[Ska 82] R. Skaug and J.L. Robinson, *"Operation of portable RF CDMA networks with controlled transmitter powers", IEE Proc.,* vol. 129, pt. F, no.6, Dec.1982.

[Tew 69] R.P. Tewarson, *"On some representations of generalized inverses", SIAM Review,* vol. 11, No. 2, Apr.1969.

[Var 88a] M.K. Varanasi and B. Aazhang, *"Near optimum demodulation for coherent communications in asynchronous Gaussian CDMA channels", Proc. of the 22nd Princeton Conf. on Information Sciences and Systems,* pp. 832-839, March 1988.

[Var 88b] M.K. Varanasi and B. Aazhang, *"Probability of error comparison of linear and iterative multiuser detectors"*, *ComCon 88 Transactions & Lecture Notes in Control and Information Science Series, Springer-Verlag Publication,* to appear.

[Var 89] M.K. Varanasi and B. Aazhang, *"Iterative multiuser detection for asynchronous Code-Division Multiple-Access communications"*, *submitted to IEEE Trans. Commun.,* June, 1988.

[Ver 83a] S. Verdú and H.V. Poor, *"Minimax robust discrete-time matched filters"*, *IEEE Trans. on Communications,* vol. COM-31, pp. 208-215, Feb. 1983.

[Ver 83b] S. Verdú, *"Minimum probability of error for asynchronous multiple access communication systems"*, *Proc. 1983 IEEE Military Communications Conf.,* pp. 213-219, Washington DC, Nov. 1983.

[Ver 84a] S. Verdú and H.V. Poor, *"On minimax robustness: A general approach and applications"*, *IEEE Trans. on Information Theory,* vol. IT-30, pp. 328-340, March 1984.

[Ver 84b] S. Verdú *"On the selection of memoryless adaptive laws for blind equalization in binary communications"*, *Sixth Int. Conference on Analysis and Optimization of Systems, Lecture Notes.* NY: Springer Verlag, 1984.

[Ver 84c] S. Verdú, *"Optimum multiuser signal detection"*, Ph.D. dissertation, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Coordinated Science Lab., Urbana, IL, Rep. T-151, Aug.1984.

[Ver 85] S. Verdú, *"Optimum multi-user detection is NP-hard"*, *1985 Int. Symposium on Information Theory,* Brighton, UK, June 1985.

[Ver 86a] S. Verdú, *"Minimum probability of error for asynchronous Gaussian multiple-access channels"*, *IEEE Trans. on Information Theory,* vol. IT-32, pp. 85-96, Jan.1986.

[Ver 86b] S. Verdú, *"Optimum multi-user asymptotic efficiency"*, *IEEE Trans. on Communications,* vol. COM-34, pp. 890-897, Sept.1986.

[Ver 88] S. Verdú, *"Recent progress in multiuser detection"*, *Advances in Communication and Control Systems, ComCon 88, and* Lecture Notes in Control and Information Sciences Series, NY: Springer-Verlag, 1988.

[Ver 89] S. Verdú, *"Computational complexity of optimum multi-user detection"*, *Algorithmica,* to appear.

[Woz] J. Wozencraft and I. Jacobs, *Principles of Communication Engineering,* Wiley, 1965.