Noncoherent Multiuser Detection for CDMA Systems with Nonlinear Modulation: A Non-Bayesian Approach

Eugene Visotsky, *Member, IEEE,* and Upamanyu Madhow, *Senior Member, IEEE*

***Abstract*—This paper considers the problem of multiuser detection for a system in which each user employs nonlinear modulation, with an emphasis on noncoherent detection techniques which do not require knowledge of the users' channel parameters at the receiver. Our goals are to gain fundamental insight into the capabilities of multiuser detection in such a setting, and to provide practical algorithms that perform better than conventional matched-filter reception. We begin by providing fundamental performance benchmarks by considering coherent maximum-likelihood (ML) detection, which requires knowledge of the users' channel parameters, as well as noncoherent detection, formulated in a non-Bayesian generalized likelihood ratio test (GLRT) framework. The asymptotic performance of each detector, as the noise level vanishes, is characterized, yielding simple geometric criteria for near–far resistance. In general, both the ML and GLRT detectors have complexity which is exponential in the number of users. We, therefore, propose the more practical sequential decision projection (SDP) detector, which has complexity which is quadratic in the number of users. It is shown that the SDP detector has nonzero asymptotic efficiency if the users' powers are suitably disparate.**

***Index Terms*—Code division multiple access (CDMA), direct sequence, multiuser detection, nonlinear modulation, spread spectrum**

.

I. Introduction

A CANONICAL problem in multiuser communications is the demodulation of one or more signals based on a received signal that is the sum of the signals generated by several independent (typically asynchronous) users. Multiuser detection attempts to exploit the structure of the multiple-access interference (MAI) due to other users when demodulating one or more desired users, thereby providing substantial performance gains over conventional methods that ignore this structure. The specific structure assumed in much of the past literature [1] on multiuser detection is that of *linearly modulated* signals: for each user, the (lowpass equivalent) signal sent in a given symbol interval consists of a symbol waveform multiplied by a complex-valued symbol. For linear modulation, the *user subspace*, defined to be the subspace spanned by the set of possible signals sent by a given user, has complex dimension one, since it is spanned by the user's symbol waveform. For the purpose of this paper, *nonlinear modulation* refers to any modulation scheme which is "not linear"; that is, with user subspace of complex dimension greater than one. A classical example of nonlinear modulation is M-ary orthogonal modulation, for which the user subspace is of dimension M. More generally, coded modulation obtained by the cascade of an encoder and a modulator (the latter is often linear) is invariably nonlinear; that is, the transmitted signal corresponding to a codeword may be viewed as a nonlinearly modulated "symbol," typically selected from a large alphabet. Note that linearly modulated *subsymbols* often form the building blocks for a nonlinearly modulated symbol.

Multiuser detection methods for linearly modulated systems typically exploit the fact that different users have different symbol waveforms (these may possibly change from symbol to symbol). For example, in direct sequence (DS) code division multiple access (CDMA), each user is assigned a different *spreading waveform* which occupies a bandwidth significantly larger than the symbol rate. In narrow-band applications (i.e., when the bandwidth of the symbol waveform is comparable to the symbol rate), each user may use the same *transmitted* symbol waveform, but the *received* symbol waveform may differ due to each user seeing different channels. Two key notions, applicable to any modulation scheme and any detector, that have emerged in the literature are *asymptotic efficiency* [2] and *near–far resistance* [4]. Asymptotic efficiency for a detector is the rate of decay, as the noise level goes to zero, of its error probability relative to that of the error probability without MAI. Near–far resistance is the value of asymptotic efficiency for the worst possible values of the users' amplitudes. For linearly modulated systems, the optimal maximum-likelihood (ML) detector has nonzero near–far resistance if the following *linear independence condition* (LIC) holds: the symbol waveforms for different users must be linearly independent. While the ML detector has complexity exponential in the number of users, it is possible to obtain the same near–far resistance (but, in general, smaller asymptotic efficiency) using the linear complexity decorrelating detector, which projects the received signal orthogonal to the space spanned by the symbol waveforms for the interfering users or with the linear minimum

Manuscript received December 1, 1998; revised May 1, 2000. This work was supported in part by the Office of Naval Research under Grant N00014-95-1-0647, and in part by Motorola under the University Partnerships in Research Programs. The material in this paper was presented in part at the 1997 IEEE International Symposium for Information Theory, Ulm, Germany, June 1997.

E. Visotsky is with the Communication Systems and Technologies Labs, Motorola Labs, Schaumburg, IL 60196 USA (e-mail: visotsky@labs.mot.com).

U. Madhow is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106 USA (e-mail: madhow@pablo.ece.ucsb.edu).

Communicated by M. L. Honig, Associate Editor for Communications.

Publisher Item Identifier S 0018-9448(01)02842-5.

mean-square error (MMSE) detector [8]. Another suboptimal strategy that has been explored is that of interference subtraction, either sequentially [6], [7], [19] or in parallel [5], in which estimates of the interferers' data are used to reconstruct the MAI and to subtract it out from the received signal.

While we briefly consider coherent ML detection at the beginning as a performance benchmark, the emphasis of this paper is on noncoherent detection techniques, which do not require knowledge of the users' channel parameters at the receiver. In order to focus on the geometry of the signal space rather than on the model for channel uncertainty, we adopt a non-Bayesian approach to the unknown data and channels. Thus, the performance benchmark for noncoherent detection is taken to be the generalized likelihood ratio test (GLRT), which performs joint ML estimation of the data and the channel parameters. The asymptotic efficiency and near–far resistance for both the ML and GLRT detectors are obtained, the results for the latter yielding new geometric insights into noncoherent detection.

An alternative, Bayesian, approach to noncoherent multiuser detection, which models the channel parameters as random variables with a known distribution, is treated in [11].

For linearly modulated systems, treating interferers' amplitudes as unknown parameters and applying the GLRT detection rule, the (linear complexity) decorrelator [1] is recovered. For systems with nonlinear modulation, however, both ML and GLRT detection are exponentially complex, and are therefore impractical. This paper proposes a suboptimal sequential decision projection (SDP) detector, which has complexity quadratic in the number of users. The idea here is to estimate the symbols sent by the users one by one, with the decision in the kth stage being based on conventional reception using the projection of the received signal orthogonal to the signals corresponding to the previous $k-1$ decisions. The order of demodulation is not fixed; rather, it is based on which user (among the ones that remain to be demodulated at a given stage) has the largest decision statistic. Thus, at stage k, the user being demodulated sees users that have not yet been demodulated as noise. For reliable demodulation, this "noise" should be small compared to the power of the user being demodulated. Using such reasoning, we are able to show that the SDP detector has nonzero asymptotic efficiency if the user powers are sufficiently disparate. The SDP detector is not near–far resistant (i.e., its worst case asymptotic efficiency is typically zero), with the worst case occurring when the users are at roughly equal powers. As shown in our numerical results in Section V, for fast fading, the power disparities required for SDP detection to work well would arise naturally even when the users are controlled to arrive at the same average received power. On the other hand, as discussed in our conclusions in Section VI, for channels with either slow or no fading, it might well be advantageous to use power control that enforces disparities in user powers.

Another low-complexity detector which has been recently proposed [14], [17], [9] is the following extension of the decorrelating detector: to demodulate a user of interest, project the received signal orthogonal to the space spanned by all possible signals that any of the interfering users could send, a space that we term the *total interference subspace*. While this decorrelation strategy is attractive because the performance of a desired user does not depend on the symbols sent by the interfering users, such a receiver would exist only if the dimension of the signal space is larger than that of the total interference subspace. Thus, decorrelation is a useful concept mainly when the user subspaces are of small dimension. The concept of decorrelation is used in [10] to develop a decision feedback-type detector which is similar to the SDP detector proposed in this paper in that the users are detected sequentially. The SDP detector, however, is not based on the concept of decorrelation with respect to the total interference subspace and hence is not limited by the dimensionality requirements on the user subspaces. Detailed comparison of the performance of the SDP detector with that of the decision feedback detector proposed in [10] is beyond the scope of this paper. The GLRT detector is briefly discussed in [15, Sec. 5], although no performance analysis appears in that reference.

Yet another suboptimal approach is successive interference cancellation [7], which estimates the users' symbols one by one, subtracting out past decisions from the received signal by using amplitude estimates obtained from the decision statistics. Errors in amplitude estimation lead to residual interference even when the decisions made for the interfering users are correct, so that the detector in [7] exhibits an interference floor in performance even under the best of circumstances. In contrast, SDP detection is immune to errors in amplitude estimation because we use orthogonal projection rather than subtraction, and this is the key to its nonzero asymptotic efficiency (which translates to the error probability tending to zero as the background noise vanishes) under disparate powers. A method similar to SDP detection has been proposed for linearly modulated systems in [22]: the difference is that, while the projection operation does not require knowledge of user amplitudes, the order of demodulation in [22] is fixed *a priori* based on the user powers, unlike the empirical ordering employed by our SDP detector. Prior to [7], most other attempts at successive interference cancellation (e.g., [6], [19]) assumed that the amplitudes of the interfering users were known to the receiver.

For ease of exposition, we restrict attention to a synchronous CDMA system throughout. Extension to asynchronous CDMA is conceptually straightforward, since asynchronous CDMA may be viewed as synchronous by interpreting each symbol of each user as a different virtual user. The number of relevant virtual users for each symbol decision can be restricted by considering a finite observation interval for each symbol decision (see, for example, [20] for an "equivalent synchronous model" corresponding to a linearly modulated asynchronous CDMA system).

The remainder of this paper is organized as follows. In Section II, we present the system model, the performance measures considered, and the coherent ML detector. In Section III, we consider noncoherent detection using the GLRT, and investigate its performance in terms of asymptotic efficiency and near–far resistance. The geometric insights into noncoherent detection developed in Section III are used to provide bounds on asymptotic efficiency and near–far resistance, and to establish criteria for the GLRT detector to have nonzero near–far resistance. Section IV contains the description and analysis of the SDP de-

tector. Numerical results are presented in Section V, and our conclusions are given in Section VI.

II. Preliminaries

We describe in Section II-A the system model and notation to be used in the remainder of this paper. Section II-B introduces the performance measures used to characterize multiuser detectors. In Section II-C, we examine the performance of the (coherent) ML detector. The structure and performance of the ML detector for nonlinear systems is similar to the ML detector considered in the context of linear modulation [2].

A. System Model and Notation

We consider the following synchronous discrete time[1] CDMA system, with received vector given by

$$\boldsymbol{r} = A_1\boldsymbol{u}_1^{j_1} + \sum_{k=2}^{K} A_k\boldsymbol{u}_k^{j_k} + \boldsymbol{w} \tag{1}$$

where $\boldsymbol{u}_k^{j_k}$ denotes the vector sent by user k when symbol j_k is transmitted, A_k is the complex amplitude of the kth user, and $\boldsymbol{w}$ is complex additive white Gaussian noise (AWGN) with independent real and imaginary parts, each with covariance $\sigma^2\boldsymbol{I}$. User k employs an alphabet of size M_k. Without loss of generality, assume that user 1 is the desired user, and that all user waveforms are normalized to have unit energy, i.e., $\|\boldsymbol{u}_k^{j_k}\|^2 = 1$. Denote by $\boldsymbol{A}$ the column vector of complex user amplitudes $\{A_k, 1 \le k \le K\}$.

It is convenient to denote by $\boldsymbol{j}$ the vector of symbols $(j_1, j_2, \ldots, j_K)$, and by $H(\boldsymbol{j})$ the hypothesis that $\boldsymbol{j}$ is sent. The set of all possible values of $\boldsymbol{j}$ is denoted by $\mathcal{J}$, and has size $M_1 \times \cdots \times M_K$. For each user i, the set of all possible symbols j_i is denoted by $\mathcal{J}_i$. With this notation, the received vector under hypothesis $H(\boldsymbol{j})$ can be expressed as follows:

$$H(\boldsymbol{j})\colon \boldsymbol{r} = \boldsymbol{U}_{\boldsymbol{j}}\boldsymbol{A} + \boldsymbol{w} \tag{2}$$

where $\boldsymbol{U}_{\boldsymbol{j}}$ is a matrix whose columns are user-vectors sent under hypothesis $H(\boldsymbol{j})$.

Notation: The following notation is used throughout this paper.

- $\boldsymbol{x}^H, \boldsymbol{U}^H$: Conjugate transpose of vector $\boldsymbol{x}$, matrix $\boldsymbol{U}$, respectively.
- $P[\boldsymbol{j}'|\boldsymbol{j}]$: Probability that the receiver decides that hypothesis $H(\boldsymbol{j}')$ is true, given that hypothesis $H(\boldsymbol{j})$ is true.
- $\langle \boldsymbol{x}, \boldsymbol{y}\rangle$: $\boldsymbol{x}^H\boldsymbol{y}$.
- $\|\boldsymbol{x}\|$: $\sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$.
- $Q(x)$: Complementary cumulative distribution function (cdf) of a standard Gaussian random variable
$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}}\, du.$$
- $P_{\mathcal{S}}\boldsymbol{x}$: Projection of vector $\boldsymbol{x}$ onto subspace $\mathcal{S}$.
- $P_{\mathcal{S}}^{\perp}\boldsymbol{x}$: Projection of vector $\boldsymbol{x}$ orthogonal to the subspace $\mathcal{S}$ ($P_{\mathcal{S}}^{\perp}\boldsymbol{x} = \boldsymbol{x} - P_{\mathcal{S}}\boldsymbol{x}$).
- $\mathcal{S}_{\boldsymbol{j}}$: Subspace spanned by signal vectors $\boldsymbol{u}_1^{j_1}, \boldsymbol{u}_1^{j_1}, \ldots, \boldsymbol{u}_K^{j_K}$ sent under $H(\boldsymbol{j})$.
- $\boldsymbol{s}(\boldsymbol{j})$: Signal sent under $H(\boldsymbol{j})$, $\boldsymbol{s}(\boldsymbol{j}) = \boldsymbol{U}_{\boldsymbol{j}}\boldsymbol{A}$.
- K: Number of users in the system, numbering users 1 through K.
- N: Processing gain.

[1]The restriction to discrete time is for the sake of simplicity, since all results extend to signals in any Hilbert space.

B. Performance Measures

Throughout this paper, we adopt *asymptotic efficiency* [2] and *near–far resistance* [3] as the standard measures characterizing the asymptotic behavior of multiuser detectors. For AWGN channels, the asymptotic efficiency (γ) is the ratio of the rates of exponential decay of error probability for the multiuser detector $P_{e,MU}$, relative to the single-user error probability $P_{e,SU}$ as the noise level tends to zero [2]. Thus, it is defined as the limit

$$\gamma = \lim_{\sigma^2 \to 0} \frac{\log P_{e,MU}}{\log P_{e,SU}}. \tag{3}$$

The near–far resistance (η) of a multiuser detector is the minimum value that its asymptotic efficiency can take if the amplitudes of the interferers can be chosen arbitrarily. Following [12], the near–far resistance for complex amplitudes is formally defined as follows:

$$\eta = \min_{A_2 \cdots A_K \in \mathcal{C}} \gamma \tag{4}$$

where $\mathcal{C}$ is the set of all complex numbers. Nonzero near–far resistance implies that the error probability of the detector always decays exponentially fast as the noise level tends to zero, regardless of the interferers' amplitudes.

C. ML Detection

Let

$$\boldsymbol{s}(\boldsymbol{j}) = A_1\boldsymbol{u}_1^{j_1} + \sum_{k=2}^{K} A_k\boldsymbol{u}_k^{j_k}$$

denote the signal received under hypothesis $H(\boldsymbol{j})$. The ML detector implements a minimum-distance rule to arrive at the following vector symbol estimate:

$$\hat{\boldsymbol{j}} = \arg\min_{\boldsymbol{j} \in \mathcal{J}} \|\boldsymbol{r} - \boldsymbol{s}(\boldsymbol{j})\|. \tag{5}$$

Clearly, this rule requires knowledge of the vectors $\boldsymbol{u}_k^{j_k}$ (which involves knowing the signaling waveform for each possible transmitted symbol for each user and the timing for each user), as well as of the complex amplitudes A_k for all users. The complexity of this detector is M^K, the number of possible values for the vector symbol $\boldsymbol{j}$.

Similarly to the case of the ML detection of linearly modulated CDMA systems [1], we obtain from (3) that the asymptotic efficiency of the ML detector for nonlinearly modulated CDMA systems is given by $\gamma_{\mathrm{ML}} = \frac{d_{\min}^2}{(d_{\min}^{su})^2}$, where $d_{\min}$ denotes the minimum distance between signals in the set $\{\boldsymbol{s}(\boldsymbol{j})\colon \boldsymbol{j} \in \mathcal{J}\}$, and $d_{\min}^{su}$ is the minimum distance on the single-user constellation of the first user. Minimizing the asymptotic efficiency with respect to the interferers' amplitudes we obtain Proposition 2.1.

Proposition 2.1: The near–far resistance of the ML detector for the first user is given by

$$\eta_{\mathrm{ML}} = \frac{\min\limits_{\boldsymbol{j},\boldsymbol{j}' \in \mathcal{J}, j_1 \neq j_1'} \left\| P^{\perp}_{\mathcal{S}_{\boldsymbol{V}(\boldsymbol{j},\boldsymbol{j}')}} \boldsymbol{v}_1(j_1, j_1') \right\|^2}{\min\limits_{j_1 \neq j_1'} \|\boldsymbol{v}_1(j_1, j_1')\|^2} \tag{6}$$

where

$$\{\boldsymbol{v}_k(j_k, j_k') = \boldsymbol{u}_k^{j_k} - \boldsymbol{u}_k^{j_k'},\ 1 \leq k \leq K\}$$

and $\mathcal{S}_{\boldsymbol{V}(\boldsymbol{j},\boldsymbol{j}')}$ as the subspace spanned by matrix $\boldsymbol{V}(\boldsymbol{j}, \boldsymbol{j}')$ with columns $\{\boldsymbol{v}_k(j_k, j_k'),\ 2 \leq k \leq K\}$.

The following corollary is immediate.

Corollary 2.1: The near–far resistance of the ML detector is nonzero for the first user if, and only if, $\boldsymbol{u}_1^{j_1} - \boldsymbol{u}_1^{j_1'}$ is linearly independent of $\{\boldsymbol{u}_k^{j_k} - \boldsymbol{u}_k^{j_k'},\ 2 \leq k \leq K\}$ for every possible pair of vector symbols $\boldsymbol{j}, \boldsymbol{j}'$ such that $j_1 \neq j_1'$.

Remark 2.1: The linear independence condition for linearly modulated CDMA systems [4] follows directly from the corollary. For linearly modulated systems, $\boldsymbol{u}_k^{j_k} = b^{j_k}\boldsymbol{s}_k$, where $\boldsymbol{s}_k$ is the symbol vector for user k, and b^{j_k} is a complex symbol. Hence, $\boldsymbol{v}_k(j_k, j_k') = (b^{j_k} - b^{j_k'})\boldsymbol{s}_k$, which lies in the span of the symbol vector of user k. Thus, nonzero near–far resistance for the first user is obtained, if and only if, its symbol vector is linearly independent of the symbol vectors of the interferers.

III. GLRT Detector

The GLRT detector is obtained by maximizing the likelihood of each hypothesis with respect to the unknown parameters, and then choosing the most likely hypothesis. It may be interpreted as the joint ML estimate of the unknown parameters and the transmitted data. For an AWGN multiuser channel with unknown complex amplitudes, the GLRT reduces to the following minimum distance rule:

$$\hat{\boldsymbol{j}} = \arg\min_{\boldsymbol{j} \in \mathcal{J}} \min_{\boldsymbol{A} \in \mathcal{C}^K} \|\boldsymbol{r} - \boldsymbol{U}_{\boldsymbol{j}}\boldsymbol{A}\| = \arg\max_{\boldsymbol{j} \in \mathcal{J}} \|P_{\mathcal{S}_{\boldsymbol{j}}}\boldsymbol{r}\|^2 \tag{7}$$

where the second equality follows because $\boldsymbol{U}_{\boldsymbol{j}}$ is a (possibly nonorthonormal) basis for $\mathcal{S}_{\boldsymbol{j}}$. Minimization over $\boldsymbol{j} \in \mathcal{J}$ does not require knowledge of user amplitudes, but it does require knowledge of all user vectors. As with the ML detector, the complexity is M^K, the number of possible values for $\boldsymbol{j}$.

From (7), we conclude that, since the user amplitudes are allowed to vary over the set of complex numbers, the detector minimizes the distance from the received vector $\boldsymbol{r}$ to the subspace $\mathcal{S}_{\boldsymbol{j}}$. Alternatively, the GLRT detector projects the received vector onto the subspace spanned under each hypothesis, and chooses the hypothesis corresponding to the largest projection. Fig. 1 illustrates the geometric interpretation of the decision rule in a binary hypothesis testing setting. Application of the GLRT to a single-user setting yields the familiar noncoherent detector.

Example 1: Consider binary equal energy modulation for a single user

$$H(i)\colon \boldsymbol{r} = A\boldsymbol{u}^i + \boldsymbol{w}, \qquad i = 1, 2 \tag{8}$$

where $\|\boldsymbol{u}^i\|^2 = E_b$, $i = 1, 2$. By applying (7) and recognizing that in this case $\tilde{\boldsymbol{U}}_{\boldsymbol{j}} = \boldsymbol{u}^1$ and $\tilde{\boldsymbol{U}}_{\boldsymbol{j}'} = \boldsymbol{u}^2$, we obtain the standard binary noncoherent detector

$$\hat{i} = \arg\max_{i=1,2} |\boldsymbol{u}_i^H\boldsymbol{r}|^2. \tag{9}$$

A. Asymptotic Performance of the GLRT Detector

Throughout this section, we fix two hypotheses $H(\boldsymbol{j})$ and $H(\boldsymbol{j}')$ such that $j_1' \neq j_1$. Conditioning throughout on $H(\boldsymbol{j})$, set $\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{j})$ as the received signal in the absence of noise. The net received vector $\boldsymbol{r} = \boldsymbol{s} + \boldsymbol{w}$, where $\boldsymbol{w}$ denotes the noise vector. The incorrect decoding region for pairwise testing using (7) is given by

$$\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j}) = \left\{\boldsymbol{y} \in \mathcal{C}^N\colon \left\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}}}\boldsymbol{y}\right\| \geq \left\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{y}\right\|\right\}. \tag{10}$$

The computation of the pairwise error probabilities of the GLRT detector involves sums of dependent magnitude-squared complex Gaussian random variables, and can be performed directly using the method of characteristic functions [13]. However, the resulting expressions are complicated and are not easily amenable to asymptotic analysis. In order to obtain insight into the performance of the GLRT detector, therefore, we obtain an asymptotic characterization of the pairwise error probabilities as the noise variance tends to zero using signal space concepts.

Consider the pairwise error exponent $e(\boldsymbol{j}'|\boldsymbol{j})$ of the decay, defined as

$$e(\boldsymbol{j}'|\boldsymbol{j}) = -\lim_{\sigma^2 \to 0} 2\sigma^2 \log P[\boldsymbol{j}'|\boldsymbol{j}]. \tag{11}$$

Similarly to the case of the ML detection, the asymptotic efficiency of the GLRT detector is given by

$$\gamma_{\mathrm{GLRT}} = \frac{e_{\mathrm{GLRT}}}{e^{su}_{\mathrm{GLRT}}} \tag{12}$$

where

$$e_{\mathrm{GLRT}} = \min_{\boldsymbol{j},\boldsymbol{j}', j_1' \neq j_1} e(\boldsymbol{j}'|\boldsymbol{j})$$

and e^{su}_{GLRT} denotes the corresponding exponent for GLRT detection for a single user.

We complete the asymptotic analysis by specifying the exponents $\{e(\boldsymbol{j}'|\boldsymbol{j})\}$. Let $d(\boldsymbol{j}'|\boldsymbol{j})$ denote the distance of $\boldsymbol{s}$ from $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$. As stated in Theorem 3.1 below, the pairwise error exponent is governed by this distance, analogous to the distance notions relevant for ML detection.

Theorem 3.1: The pairwise error exponent is given by

$$e(\boldsymbol{j}'|\boldsymbol{j}) = d(\boldsymbol{j}'|\boldsymbol{j})^2 \tag{13}$$

so that the pairwise error probability has the asymptotic form

$$P[\boldsymbol{j}'|\boldsymbol{j}] \sim \exp\left(-\frac{d(\boldsymbol{j}'|\boldsymbol{j})^2}{2\sigma^2}\right).$$

Furthermore, $d(\boldsymbol{j}'|\boldsymbol{j}) > 0$ if and only if $\boldsymbol{s} \notin \mathcal{S}_{\boldsymbol{j}'}$.

Proof: To establish the equality, we derive upper and lower bounds on the pairwise error probability that coincide asymptotically as $\sigma^2 \to 0$. The upper bound on error probability, which corresponds to a lower bound on the exponent (11), is based on the observation that

$$P[\boldsymbol{j}'|\boldsymbol{j}] \leq P(\|\boldsymbol{w}\| > d(\boldsymbol{j}'|\boldsymbol{j})) \tag{14}$$

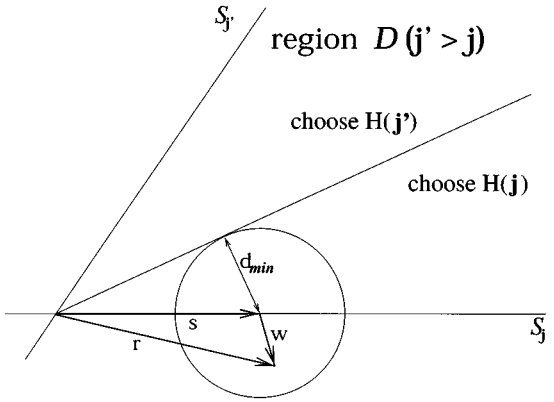

Fig. 1. Illustration of the GLRT detector for binary hypothesis testing.

since, otherwise, the received vector could not fall in the wrong decoding region. This is illustrated in Fig. 1, where we show the idea in a simplified two-dimensional setting. The entries of the noise vector $\boldsymbol{w}$ are zero mean independent complex Gaussian random variables. Hence, the energy

$$\|\boldsymbol{w}\|^2 = \sum_{i=1}^{N} [\mathrm{Re}(w_i)]^2 + [\mathrm{Im}(w_i)]^2$$

of the noise vector has a chi-square distribution with $2N$ degrees of freedom, so that (see [16, p. 43])

$$P(\|\boldsymbol{w}\| > d(\boldsymbol{j}'|\boldsymbol{j})) = e^{-\frac{d(\boldsymbol{j}'|\boldsymbol{j})^2}{2\sigma^2}} \sum_{k=0}^{N-1} \frac{1}{k!} \left(\frac{d(\boldsymbol{j}'|\boldsymbol{j})^2}{2\sigma^2} \right)^k. \quad (15)$$

Applying the definition (11) yields, using (14) and (15), that

$$e(\boldsymbol{j}'|\boldsymbol{j}) \geq d(\boldsymbol{j}'|\boldsymbol{j})^2. \quad (16)$$

We now derive a lower bound on $P[\boldsymbol{j}'|\boldsymbol{j}]$ (corresponding to an upper bound on $e(\boldsymbol{j}'|\boldsymbol{j})$). The idea is to consider a dominant error event in which the noise vector "points" in the direction corresponding to the minimum distance from $\boldsymbol{s}$ to the incorrect region $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$, and the norm of the noise vector exceeds the minimum distance $d(\boldsymbol{j}'|\boldsymbol{j})$. This event is a lower bound because the dominant error event is not the only way of producing an error. Since its probability is asymptotically governed by $d(\boldsymbol{j}'|\boldsymbol{j})$, it coincides asymptotically with the upper bound (15) on $P[\boldsymbol{j}'|\boldsymbol{j}]$.

Consider now the lower bound in detail. Define $\boldsymbol{y}_*$ as the point in $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$ closest to $\boldsymbol{s}$, and define $\boldsymbol{x}_* = \boldsymbol{s} - \boldsymbol{y}_*$ as the vector that points in the direction of the shortest path from $\boldsymbol{s}$ to $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$, so that $d(\boldsymbol{j}'|\boldsymbol{j}) = \|\boldsymbol{x}_*\| = \|\boldsymbol{s} - \boldsymbol{y}_*\|$. This is illustrated in Fig. 2.

We now decompose the noise vector $\boldsymbol{w}$ into two orthogonal components as follows. The first, of length w_*, points along $\boldsymbol{x}_*$, and is given by

$$w_* \frac{\boldsymbol{x}_*}{\|\boldsymbol{x}_*\|} = \langle \boldsymbol{w}, \boldsymbol{x}_* \rangle \frac{\boldsymbol{x}_*}{\|\boldsymbol{x}_*\|^2}.$$

The second is

$$\boldsymbol{w}_*^{\perp} = \boldsymbol{w} - w_* \frac{\boldsymbol{x}_*}{\|\boldsymbol{x}_*\|}$$

the component of $\boldsymbol{w}$ orthogonal to $\boldsymbol{x}_*$. The following lemma, proved in the Appendix, gives the construction for a dominant error event.

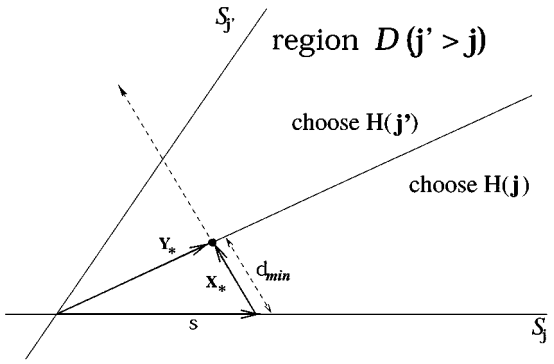

Fig. 2. Illustration of the lower bound.

Lemma 3.1: Define the event E_n

$$E_n = E_{1,n} \cap E_{2,n} \quad (17)$$

where

$$E_{1,n} = \left\{ d(\boldsymbol{j}'|\boldsymbol{j}) \left(1 + \frac{1}{n}\right) \leq w_* \leq \left(A - \frac{1}{n}\right) d(\boldsymbol{j}'|\boldsymbol{j}) \right\}$$

and

$$E_{2,n} = \{ \|\boldsymbol{w}_*^{\perp}\| \leq \epsilon_n \}.$$

Then, there exists $A > 1$ and, for n large enough, there exists $\epsilon_n > 0$ such that E_n lies in the error region $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$.

Proof: See Appendix A.

The lemma implies that

$$P[\boldsymbol{j}'|\boldsymbol{j}] \geq P[E_n]. \quad (18)$$

Since $\boldsymbol{w}$ is white and Gaussian, w_* and $\boldsymbol{w}_*^{\perp}$ are uncorrelated and jointly Gaussian, and hence independent. Therefore, we have

$$P[E_n] = P[E_{1,n}] P[E_{2,n}]. \quad (19)$$

Fixing n, as $\sigma^2 \to 0$, we have

$$P[E_{2,n}] \to 1 \quad (20)$$

since $\epsilon_n > 0$. Furthermore, since w_* is mean zero Gaussian with variance σ^2, it follows that

$$P[E_{1,n}] = Q\left(\frac{d(\boldsymbol{j}'|\boldsymbol{j}) \left(1 + \frac{1}{n}\right)}{\sigma} \right) - Q\left(\frac{\left(A - \frac{1}{n}\right) d(\boldsymbol{j}'|\boldsymbol{j})}{\sigma} \right).$$

As $\sigma^2 \to 0$, we obtain that

$$-2\sigma^2 \log P[E_{1,n}] \to d(\boldsymbol{j}'|\boldsymbol{j})^2 \left(1 + \frac{1}{n}\right)^2. \quad (21)$$

It follows from (18)–(21) that

$$e(\boldsymbol{j}'|\boldsymbol{j}) = \lim_{\sigma^2 \to 0} -2\sigma^2 \log P[\boldsymbol{j}'|\boldsymbol{j}] \leq d(\boldsymbol{j}'|\boldsymbol{j})^2 \left(1 + \frac{1}{n}\right)^2.$$

Letting $n \to \infty$ completes the proof of the upper bound on $e(\boldsymbol{j}'|\boldsymbol{j})$.

It remains to prove the necessary and sufficient condition for a nontrivial error exponent. If $\boldsymbol{s} \in \mathcal{S}_{\boldsymbol{j}'}$, clearly, $d(\boldsymbol{j}'|\boldsymbol{j}) = 0$, since $\mathcal{S}_{\boldsymbol{j}'}$ is contained in $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$. Conversely, suppose that $\boldsymbol{s} \notin \mathcal{S}_{\boldsymbol{j}'}$. Define the functional $F(\boldsymbol{w})$:$\mathcal{C}^N \to \mathcal{R}$ as

$$F(\boldsymbol{w}) = \left\| P_{\mathcal{S}_{\boldsymbol{j}'}}^{\perp}(\boldsymbol{s} + \boldsymbol{w}) \right\|^2 - \left\| P_{\mathcal{S}_{\boldsymbol{j}}}^{\perp}(\boldsymbol{s} + \boldsymbol{w}) \right\|^2.$$

Since $\boldsymbol{s} \in \mathcal{S}_{\boldsymbol{j}}$, we have $F(\boldsymbol{0}) > 0$ if $\boldsymbol{s} \notin \mathcal{S}_{\boldsymbol{j}'}$ (and hence $\boldsymbol{s} \neq 0$). Hence, by the continuity of F, there exists $\delta > 0$ such that $F(\boldsymbol{w}) > 0$ for all $||\boldsymbol{w}|| \leq \delta$. Thus, the translated ball $\{\boldsymbol{s} + \boldsymbol{w}: ||\boldsymbol{w}|| \leq \delta\}$ lies in the region of correct decoding, which implies that $d(\boldsymbol{j}'|\boldsymbol{j}) > \delta$. □

Example 2: The single-user example with binary orthogonal modulation provides a "sanity check" for the results given in this section, because in this case, there is an explicit formula for the probability of error [16] $P_e^{su} = \frac{1}{2} e^{-\frac{E_b}{4\sigma^2}}$. From Fig. 1, we observe that the minimum distance and the energy per bit are related as $2(d_{\min}^{su})^2 = E_b$, giving us $e(\boldsymbol{j}'|\boldsymbol{j}) = (d_{\min}^{su})^2$ as predicted by Theorem 3.1.

Remark 3.1: Since the GLRT decision rule compares $||P_{\tilde{\mathcal{S}}_{\boldsymbol{j}}}^{\perp}\boldsymbol{r}||$ with $||P_{\tilde{\mathcal{S}}_{\boldsymbol{j}'}}^{\perp}\boldsymbol{r}||$, any component of $\boldsymbol{r} \in \mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}$ is irrelevant to the pairwise detection problem being considered. We may therefore project all vectors involved orthogonal to $\mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}$ prior to any performance computations. To this end, define $\tilde{\boldsymbol{x}} = P_{\mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}}^{\perp}\boldsymbol{x}$ for any vector $\boldsymbol{x}$, and $\tilde{\mathcal{S}} = P_{\mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}}^{\perp}\mathcal{S}$ for any subspace $\mathcal{S}$. The effective noiseless signal is, therefore, $\tilde{\boldsymbol{s}}$, the effective subspaces spanned by the vectors involved under the two hypotheses are $\tilde{\mathcal{S}}_{\boldsymbol{j}}$ and $\tilde{\mathcal{S}}_{\boldsymbol{j}'}$, respectively, and the effective incorrect region is $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$. Note that the distance of $\tilde{\boldsymbol{s}}$ from $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$ is $d(\boldsymbol{j}'|\boldsymbol{j})$, as before. Furthermore, the noise vector $\tilde{\boldsymbol{w}}$, restricted to the subspace orthogonal to $\mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}$ remains white.

Corollary 3.1: The minimum distance $d(\boldsymbol{j}'|\boldsymbol{j})$ is strictly greater than zero if, and only if, $\tilde{\boldsymbol{s}} \neq \boldsymbol{0}$.

Proof: From the condition of Theorem 3.1, $d(\boldsymbol{j}'|\boldsymbol{j}) > 0$ if, and only if, $\boldsymbol{s} \notin \mathcal{S}_{\boldsymbol{j}'}$, which implies that $\boldsymbol{s} \notin \mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}$. This is equivalent to the condition that the projection of $\boldsymbol{s}$ onto $(\mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'})^{\perp}$ is nonzero, i.e., that $\tilde{\boldsymbol{s}} \neq \boldsymbol{0}$. □

Proposition 3.2 (Condition for Nonzero Near–Far Resistance): The GLRT detector has nonzero near–far resistance for the first user, if and only if, the following condition is satisfied for every possible pair of vector symbols $\boldsymbol{j}, \boldsymbol{j}'$ such that $j_1 \neq j_1'$.

The vector $P_{\mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}}^{\perp}\boldsymbol{u}_1^{j_1}$ is linearly independent of

$$\{P_{\mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}}^{\perp}\boldsymbol{u}_k^{j_k}, 2 \leq k \leq K\}.$$

Proof: For nonzero near–far resistance, we must have $e(\boldsymbol{j}'|\boldsymbol{j}) = d(\boldsymbol{j}'|\boldsymbol{j}) > 0$ for each hypothesis pair $\boldsymbol{j}, \boldsymbol{j}'$, regardless of the interference amplitudes. For each such pair, therefore, we must have $\tilde{\boldsymbol{s}} \neq \boldsymbol{0}$. Expanding $\tilde{\boldsymbol{s}}$ as follows:

$$\tilde{\boldsymbol{s}} = A_1 P_{\mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}}^{\perp}\boldsymbol{u}_1^{j_1} + \sum_{k=2}^{K} A_k P_{\mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}}^{\perp}\boldsymbol{u}_k^{j_k}$$

where $A_1 \neq 0$, it is evident that $\tilde{\boldsymbol{s}} \neq \boldsymbol{0}$ regardless of $\{A_k\}_{k=2}^{K}$ if, and only if, the linear independence condition in the proposition holds. □

Remark 3.2: The user independence assumption $\boldsymbol{u}_1^{j_1}$ is linearly independent of $\{\boldsymbol{u}_k^{j_k}, 2 \leq k \leq K, \forall\, \boldsymbol{j} \in \mathcal{J}\}$ is necessary, but not sufficient, for the above condition to hold.

Proposition 3.3: Nonzero near–far resistance of the GLRT detector implies nonzero near–far resistance of the ML detector.

Proof: We prove this by contradiction. Assume that the GLRT detector is near–far resistant, while the ML detector is not near–far resistant. From Corollary 2.1 to Proposition 2.1, for some $\boldsymbol{j}, \boldsymbol{j}'$, such that $j_1 \neq j_1'$, there exists a column vector $\boldsymbol{a} = [a_1 \cdots a_K]^T$ such that $a_1 \neq 0$ and $(\boldsymbol{U_j} - \boldsymbol{U_{j'}})\boldsymbol{a} = \boldsymbol{0}$. This implies that $\boldsymbol{U_j a} \in \mathcal{S_j} \cap \mathcal{S_{j'}}$. But this implies that $P_{\mathcal{S_j} \cap \mathcal{S_{j'}}}^{\perp}\boldsymbol{U_j a} = \boldsymbol{0}$, which in turn implies that $P_{\mathcal{S_j} \cap \mathcal{S_{j'}}}^{\perp}\boldsymbol{u}_1^{j_1}$ is linearly dependent on $\{P_{\mathcal{S_j} \cap \mathcal{S_{j'}}}^{\perp}\boldsymbol{u}_k^{j_k}, 2 \leq k \leq K\}$, so that the condition for nonzero near–far resistance of the GLRT is not satisfied. □

B. Pairwise Error Exponents: Geometric Bounds

The exact computation of the pairwise error exponents $e(\boldsymbol{j}'|\boldsymbol{j}) = d(\boldsymbol{j}'|\boldsymbol{j})$ amounts to solving a constrained nonconvex optimization problem for which there appears to be no analytical solution. Our approach in this paper is to obtain bounds on $d(\boldsymbol{j}'|\boldsymbol{j})$ using simple geometric notions. These bounds lead to corresponding bounds on the asymptotic efficiency and the near–far resistance. To facilitate development of the bounds, we make the following geometric definitions.

Definitions:

1) The minimum angle $\theta_{\mathcal{S}}(\boldsymbol{a})$ between a vector $\boldsymbol{a}$ and a subspace $\mathcal{S}$ is defined as

$$\cos\theta_{\mathcal{S}}(\boldsymbol{a}) = \frac{||P_{\mathcal{S}}\boldsymbol{a}||}{||\boldsymbol{a}||}. \tag{22}$$

2) The minimum angle between the subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$ is defined as follows:

$$\begin{aligned}\cos\theta_{\min}(\mathcal{S}_1, \mathcal{S}_2) &= \max_{\boldsymbol{a} \in \mathcal{S}_1} \frac{||P_{\mathcal{S}_2}\boldsymbol{a}||}{||\boldsymbol{a}||} \\ &= \max_{\boldsymbol{a} \in \mathcal{S}_1, \boldsymbol{b} \in \mathcal{S}_2} \frac{\mathrm{Re}\langle \boldsymbol{a}, \boldsymbol{b}\rangle}{||\boldsymbol{a}||\,||\boldsymbol{b}||} \\ &= \max_{\boldsymbol{b} \in \mathcal{S}_2} \frac{||P_{\mathcal{S}_1}\boldsymbol{b}||}{||\boldsymbol{b}||}.\end{aligned}$$

We note that the definition of $\cos\theta_{\min}(\mathcal{S}_1, \mathcal{S}_2)$ is symmetric with respect to the subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$, and that we may restrict $\theta_{\min}$ to $[0, \pi/2]$ without loss of generality. Furthermore, $\theta_{\min}(\mathcal{S}_1, \mathcal{S}_2)$ is trivially zero if $\mathcal{S}_1 \cap \mathcal{S}_2$ contains a nonzero element.

The following theorem provides bounds on $d(\boldsymbol{j}'|\boldsymbol{j})$ (and hence on the pairwise error exponents).

Theorem 3.2: For each hypothesis pair $(\boldsymbol{j}, \boldsymbol{j}')$, the minimum distance for the GLRT detector is bounded as

$$||\tilde{\boldsymbol{s}}|| \sin\frac{\theta_{\min}}{2} \leq d(\boldsymbol{j}'|\boldsymbol{j}) \leq ||\tilde{\boldsymbol{s}}|| \sqrt{f(\tilde{\boldsymbol{s}}, \boldsymbol{v})}$$

where $\theta_{\min} = \theta_{\min}(\tilde{\mathcal{S}}_{\boldsymbol{j}}, \tilde{\mathcal{S}}_{\boldsymbol{j}'})$, $\boldsymbol{v}$ is any vector in $\tilde{\mathcal{S}}_{\boldsymbol{j}'}$ such that $||\boldsymbol{v}|| = ||\tilde{\boldsymbol{s}}||$, and

$$f(\tilde{\boldsymbol{s}}, \boldsymbol{v}) = \frac{1 - \left|\frac{\langle \tilde{\boldsymbol{s}}, \boldsymbol{v}\rangle}{||\tilde{\boldsymbol{s}}||^2}\right|^2}{1 + \frac{\sin^2\theta_{\boldsymbol{v}}}{\sin^2\theta_{\tilde{\boldsymbol{s}}}} + 2\frac{\mathrm{Re}\langle \tilde{\boldsymbol{s}}, \boldsymbol{v}\rangle}{||\tilde{\boldsymbol{s}}||^2}\frac{\sin\theta_{\boldsymbol{v}}}{\sin\theta_{\tilde{\boldsymbol{s}}}}}$$

where $||\boldsymbol{v}|| = ||\tilde{\boldsymbol{s}}||$, $\theta_{\boldsymbol{v}} = \theta_{\tilde{\mathcal{S}}_{\boldsymbol{j}}}(\boldsymbol{v})$, $\theta_{\tilde{\boldsymbol{s}}} = \theta_{\tilde{\mathcal{S}}_{\boldsymbol{j}'}}(\tilde{\boldsymbol{s}})$.

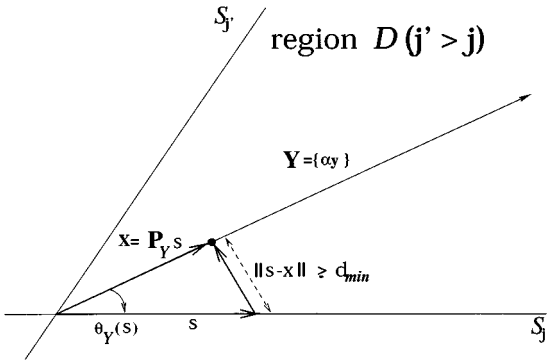

Fig. 3. Illustration of the upper bound construction.

Notation: All vectors and subspaces involved in the proof are projected orthogonal to the irrelevant subspace $\mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}$. However, for notational simplicity, we drop the tilde notation introduced in Remark 3.1 within the proof.

Proof: The proof of the lower bound involves somewhat lengthy algebraic manipulations and is relegated to Appendix B for the sake of continuity and clarity of the presentation.

To prove the upper bound, we note that, for any $\boldsymbol{x} \in \mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$, $||\boldsymbol{s} - \boldsymbol{x}||$ is an upper bound on $d(\boldsymbol{j}'|\boldsymbol{j})$. For constructing such an $\boldsymbol{x}$, let $\boldsymbol{v}$ be any vector in $\mathcal{S}_{\boldsymbol{j}'}$ such that $||\boldsymbol{v}|| = ||\boldsymbol{s}||$. Set

$$\boldsymbol{y} = \frac{\boldsymbol{s}}{\sin\theta_{\boldsymbol{s}}} + \frac{\boldsymbol{v}}{\sin\theta_{\boldsymbol{v}}} \tag{23}$$

where the angles $\theta_{\boldsymbol{s}}$ and $\theta_{\boldsymbol{v}}$ are defined in the statement of the theorem. It is easy to check that $||P_{\mathcal{S}_{\boldsymbol{j}}}^{\perp}\boldsymbol{y}|| = ||P_{\mathcal{S}_{\boldsymbol{j}'}}^{\perp}\boldsymbol{y}||$, and hence $\boldsymbol{y}$ is in the closure of $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$. Furthermore, any vector $\alpha\boldsymbol{y}$ in $\mathcal{Y}$, the one-dimensional complex subspace generated by $\boldsymbol{y}$ is also in $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$. This is illustrated in Fig. 3. In particular, $\boldsymbol{x} = P_{\mathcal{Y}}\boldsymbol{s} \in \mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$ and

$$||\boldsymbol{s} - \boldsymbol{x}||^2 = ||\boldsymbol{s} - P_{\mathcal{Y}}\boldsymbol{s}||^2 = ||\boldsymbol{s}||^2 \sin^2\theta_{\mathcal{Y}}(\boldsymbol{s}) \geq d(\boldsymbol{j}'|\boldsymbol{j}). \tag{24}$$

From (22), we have

$$\cos^2\theta_{\mathcal{Y}}(\boldsymbol{s}) = \frac{||P_{\mathcal{Y}}\boldsymbol{s}||^2}{||\boldsymbol{s}||^2} = \frac{|\langle P_{\mathcal{Y}}\boldsymbol{s}, \boldsymbol{s}\rangle|}{||\boldsymbol{s}||\,||P_{\mathcal{Y}}\boldsymbol{s}||} = \frac{|\langle \boldsymbol{y}, \boldsymbol{s}\rangle|}{||\boldsymbol{s}||\,||\boldsymbol{y}||}.$$

Thus,

$$\sin^2\theta_{\mathcal{Y}}(\boldsymbol{s}) = \frac{||\boldsymbol{s}||^2||\boldsymbol{y}||^2 - |\langle \boldsymbol{y}, \boldsymbol{s}\rangle|^2}{||\boldsymbol{s}||^2||\boldsymbol{y}||^2}.$$

Substituting this expression into (24), and using the expression (23) yields the desired upper bound after straightforward algebraic manipulations. □

Remark 3.3: The upper bound in the Theorem 3.2 depends on the choice of $\boldsymbol{v}$. Unfortunately, explicit minimization of the upper bound as a function of $\boldsymbol{v}$ does not appear to be tractable. Two convenient choices of $\boldsymbol{v}$ which are used in our later computations are as follows. The first is $\boldsymbol{v} = P_{\tilde{\mathcal{S}}_{\boldsymbol{j}'}}\tilde{\boldsymbol{s}}||\tilde{\boldsymbol{s}}||/||P_{\tilde{\mathcal{S}}_{\boldsymbol{j}'}}\tilde{\boldsymbol{s}}||$, with the convention that $\theta_{\boldsymbol{v}} = 0, \boldsymbol{v} = \boldsymbol{0}$ if $\tilde{\boldsymbol{s}} \in \tilde{\mathcal{S}}_{\boldsymbol{j}'}^{\perp}$. (In this degenerate case, the upper bound becomes $||\tilde{\boldsymbol{s}}||^2$.) The second choice is

$$\boldsymbol{v}^* = \arg \min_{\boldsymbol{v} \in \tilde{\mathcal{S}}_{\boldsymbol{j}'},\, ||\boldsymbol{v}|| = ||\tilde{\boldsymbol{s}}||} \theta_{\tilde{\mathcal{S}}_{\boldsymbol{j}}}(\boldsymbol{v}) = \arg \max_{\boldsymbol{v} \in \tilde{\mathcal{S}}_{\boldsymbol{j}'},\, ||\boldsymbol{v}|| = ||\tilde{\boldsymbol{s}}||} \frac{\left|\left|P_{\tilde{\mathcal{S}}_{\boldsymbol{j}}}\boldsymbol{v}\right|\right|}{||\boldsymbol{v}||}. \tag{25}$$

That is, the angle between $\boldsymbol{v}^*$ and $P_{\tilde{\mathcal{S}}_{\boldsymbol{j}}}\boldsymbol{v}^*$ equals the minimum angle between the subspaces $\tilde{\mathcal{S}}_{\boldsymbol{j}}$ and $\tilde{\mathcal{S}}_{\boldsymbol{j}'}$. Computation of such a $\boldsymbol{v}^*$ can be done explicitly using the method of Lagrange multipliers.

Theorems 3.1 and 3.2 yield the following bounds on the pairwise error exponents:

$$g_L(\boldsymbol{j}'|\boldsymbol{j}) \leq e(\boldsymbol{j}'|\boldsymbol{j}) \leq g_U(\boldsymbol{j}'|\boldsymbol{j})$$

where

$$g_L(\boldsymbol{j}'|\boldsymbol{j}) = ||\tilde{\boldsymbol{s}}||^2 \sin^2\frac{\theta_{\min}}{2}$$

$$g_U(\boldsymbol{j}'|\boldsymbol{j}) = ||\tilde{\boldsymbol{s}}||^2 f(\tilde{\boldsymbol{s}}, \boldsymbol{v})$$

$\theta_{\min} = \theta_{\min}(\tilde{\mathcal{S}}_{\boldsymbol{j}}, \tilde{\mathcal{S}}_{\boldsymbol{j}'})$. Minimizing $g_L(\boldsymbol{j}'|\boldsymbol{j})$ and $g_U(\boldsymbol{j}'|\boldsymbol{j})$ over the choice of $\boldsymbol{j}'$ and $\boldsymbol{j}$ gives the bounds on the asymptotic efficiency of the GLRT detector.

Consider now minimization of $g_L(\boldsymbol{j}'|\boldsymbol{j})$ and $g_U(\boldsymbol{j}'|\boldsymbol{j})$ over the intereferers' amplitudes which leads to the bounds on the near–far resistance of the GLRT detector. We have

$$h_L(\boldsymbol{j}'|\boldsymbol{j}) = \min_{A_2 \cdots A_k \in \mathcal{C}} g_L(\boldsymbol{j}'|\boldsymbol{j}) = \left|\left|P_{\tilde{\mathcal{S}}_{\boldsymbol{j}}^{(1)}}^{\perp}\tilde{\boldsymbol{u}}_1^{j_1}\right|\right|^2 \sin^2\frac{\theta_{\min}}{2} \tag{26}$$

where $\tilde{\mathcal{S}}_{\boldsymbol{j}}^{(1)}$ is the subspace spanned by the $\{\tilde{\boldsymbol{u}}_k^{j_k}, 2 \leq k \leq K\}$. The second equality in (26) holds since only the norm in $g_L(\boldsymbol{j}'|\boldsymbol{j})$ exhibits dependence on interferers' amplitudes. Consider now the minimization of $g_U(\boldsymbol{j}'|\boldsymbol{j})$. Although only $\tilde{\boldsymbol{s}}$ depends on the interferers' amplitudes in the $g_U(\boldsymbol{j}'|\boldsymbol{j})$, the minimization of the $g_U(\boldsymbol{j}'|\boldsymbol{j})$ over the interferers' amplitudes does not appear to be tractable. We therefore obtain an upper bound on this minimum by minimizing the $||\tilde{\boldsymbol{s}}||^2$ norm in the expression $g_U(\boldsymbol{j}'|\boldsymbol{j})$, and then evaluating the overall expression with the minimizing set of amplitudes. This amounts to evaluating the overall expression with $\tilde{\boldsymbol{s}} = P_{\tilde{\mathcal{S}}_{\boldsymbol{j}}^{(1)}}^{\perp}\tilde{\boldsymbol{u}}_1^{j_1}$. That is,

$$\min_{A_2 \cdots A_k \in \mathcal{C}} g_U(\boldsymbol{j}'|\boldsymbol{j}) \leq g_U(\boldsymbol{j}'|\boldsymbol{j})|_{\tilde{\boldsymbol{s}} = P_{\tilde{\mathcal{S}}_{\boldsymbol{j}}^{(1)}}^{\perp}\tilde{\boldsymbol{u}}_1^{j_1}} = h_U(\boldsymbol{j}'|\boldsymbol{j}). \tag{27}$$

Minimizing $h_U(\boldsymbol{j}'|\boldsymbol{j})$ and $h_L(\boldsymbol{j}'|\boldsymbol{j})$ over the choice of $\boldsymbol{j}'$ and $\boldsymbol{j}$ gives the bounds on the near–far resustance of the GLRT detector.

Remark 3.4 (Computation of the Minimum Angle): To complete the discussion of the asymptotic performance of the GLRT detector, we give explicit expressions for $\theta_{\min}(\tilde{\mathcal{S}}_{\boldsymbol{j}}, \tilde{\mathcal{S}}_{\boldsymbol{j}'})$. Recall that

$$\cos\theta_{\min}(\tilde{\mathcal{S}}_{\boldsymbol{j}}, \tilde{\mathcal{S}}_{\boldsymbol{j}'}) = \max_{\boldsymbol{a} \in \tilde{\mathcal{S}}_{\boldsymbol{j}}} \frac{\left|\left|P_{\tilde{\mathcal{S}}_{\boldsymbol{j}'}}\boldsymbol{a}\right|\right|}{||\boldsymbol{a}||}.$$

This optimization problem can be explicitly solved to obtain

$$\cos\,\theta_{\min}(\tilde{\mathcal{S}}_{\boldsymbol{j}}, \tilde{\mathcal{S}}_{\boldsymbol{j}'}) = \sqrt{\lambda_{\max}}$$

where $\lambda_{\max}$ is the largest singular value of the matrix $\boldsymbol{W} = \tilde{\boldsymbol{U}}_{\boldsymbol{j}}{}^H \tilde{\boldsymbol{U}}_{\boldsymbol{j}'} \tilde{\boldsymbol{U}}_{\boldsymbol{j}'}^H \tilde{\boldsymbol{U}}_{\boldsymbol{j}}$, where $\tilde{\boldsymbol{U}}_{\boldsymbol{j}}$ and $\tilde{\boldsymbol{U}}_{\boldsymbol{j}'}$ contain orthormal bases of the subspaces $\tilde{\mathcal{S}}_{\boldsymbol{j}}$ and $\tilde{\mathcal{S}}_{\boldsymbol{j}'}$, respectively.

Example 3: Here we present a simple example where the GLRT detector is near–far resistant, while the decorrelator does

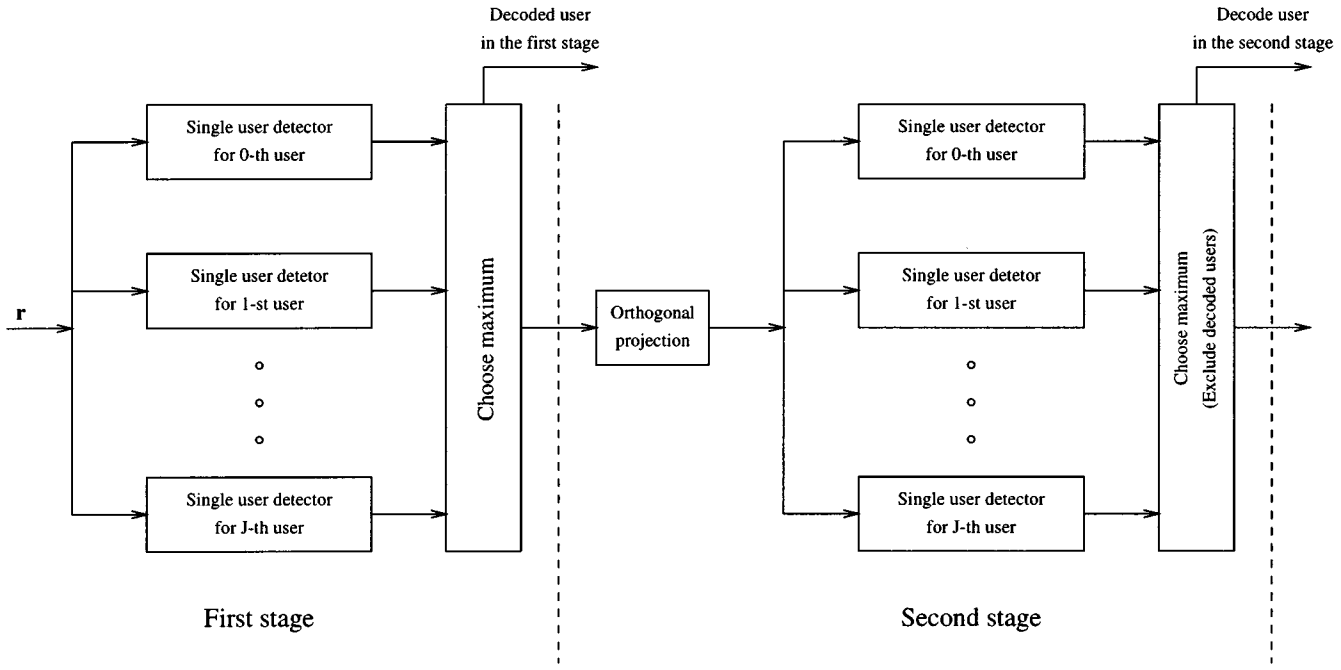

Fig. 4. SDP detector.

not exist. Consider a system with $K = 2$ users, each sending one of $M = 4$ orthogonal signals, in a signal space of dimension $N = 4$. The signaling matrices for users 1 and 2, respectively, are given as follows:

$$\mathcal{S}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathcal{S}_2 = 1/\sqrt{3} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ 1 & -1 & 1 & 0 \end{bmatrix}$$

where the columns of $\mathcal{S}_1$ (respectively, $\mathcal{S}_2$) specify the four possible orthogonal signals transmitted by user 1 (respectively, user 2). In this case, the interfering user spans the entire signal space, so that decorrelation with respect to the total interference subspace cannot be applied. The near–far resistance of the GLRT detector can be explicitly lower- and upper-bounded using (26), (27). To determine the upper bound, we use the two choices for $\boldsymbol{v}$ suggested in Remark 3.3. Performing the numerical computations, we obtain $0.0746 \leq \eta_{\mathrm{GLRT}} \leq 0.0893$.

Remark 3.5 (GLRT Detection for DPSK Modulation): The GLRT detector is not guaranteed to outperform other noncoherent detectors in general, since it does not minimize the error probability. For instance, consider the performance of the noncoherent detector for diferentially encoded phase-shift keying (DSPK) modulation in [12], based on the decorrelation strategy. It can be shown that the near–far resistance of the GLRT detector, when applied to DPSK, is no larger than that of the noncoherent detector in [12].

IV. Suboptimum Detection

In this section, we propose a suboptimum detector that we term sequential decision projection (SDP) (see Fig. 4). The complexity of the detector is quadratic in the number of users, and it yields nonzero asymptotic efficiency under a suitable power control strategy. The idea is to discover and exploit power disparities among users based on demodulation statistics, and to demodulate the users sequentially, from strongest to weakest, projecting the received signal orthogonal to the subspace spanned by the signal vectors corresponding to previous decisions. The detector performs decoding in K stages, with one user being decoded at each stage. This is a natural strategy because a) decisions for stronger users are more reliable, and b) performance gains for weaker users are obtained from suppression of stronger users. Because the projection operation does not depend upon amplitudes, the SDP detector does not require explicit knowledge of user amplitudes. The power ranking necessary for the SDP detector is performed by considering single-user statistics, which can be either coherent or noncoherent matched-filter outputs. Clearly, the preceding scheme functions best in the presence of power disparities among users. Indeed, it is shown that it is possible to attain nonzero asymptotic efficiency for an AWGN channel by enforcing such disparities via power control. Moreover, we demonstrate via simulations that such disparities arise naturally, and can be exploited by the SDP detector, on a Rayleigh fading channel, even when the users have equal average power.

Notation: Let $\mathcal{D}_i$ denote the set of $i-1$ users decoded prior to the ith stage, and let $\overline{\mathcal{D}}_i$ denote the set of $K-i+1$ users yet to be decoded. Let $\hat{\boldsymbol{u}}^{(i)}$ denote the decoded signal vector corresponding to the decision in the ith stage, i.e., $\hat{\boldsymbol{u}}^{(i)} = \boldsymbol{u}_k^{j_k}$ for some $k \in \overline{\mathcal{D}}_i$ and some $j_k \in \mathcal{J}_k$. Finally, let $\hat{\boldsymbol{U}}_i$ denote the subspace spanned by $\hat{\boldsymbol{u}}^{(1)}, \ldots, \hat{\boldsymbol{u}}^{(i-1)}$, i.e., by the decoded vectors for the first $i-1$ stages.

To demodulate in the ith stage, the received vector is first projected orthogonal to $\hat{\boldsymbol{U}}_i$. The receiver then computes the single-user decision statistics for every user in $\overline{\mathcal{D}}_i$, i.e., every

user not decoded thus far, and makes a decision for the strongest of these users. The corresponding operations are

$$\tilde{\boldsymbol{r}}^{(i)} = P^{\perp}_{\hat{\boldsymbol{U}}_i}\boldsymbol{r} \tag{28}$$

$$\left(k(i), \hat{\jmath}_{k(i)}\right) = \arg\max_{k\in\overline{\mathcal{D}}_i}\max_{j_k\in\mathcal{J}_k}\left|\left(\tilde{\boldsymbol{r}}^{(i)}\right)^H \boldsymbol{u}_k^{j_k}\right|^2 \tag{29}$$

where $\hat{\boldsymbol{u}}^{(i)} = \boldsymbol{u}_{k(i)}^{\hat{\jmath}_{k(i)}}$. The inner maximization in (29) decodes every user on the basis of single-user decision statistics. The outer maximization chooses a decoded vector corresponding to the strongest user. Note that the user with the maximum matched-filter output is assumed to be the strongest. The subspace $\hat{\boldsymbol{U}}_i$ and the set of decoded users $\mathcal{D}_i$ are updated to include the vector $\hat{\boldsymbol{u}}^{(i)}$ and the user $k(i)$, respectively. The algorithm terminates at stage K, when all users are decoded.

Remark 4.1 (Notes on SDP Detector Implementation):

1) An orthonormal basis for the subspace $\hat{\boldsymbol{U}}_i$ can be maintained using recursive updates by performing a Gramm–Schmidt procedure on the decoded vectors $\hat{\boldsymbol{u}}^{(1)}, \ldots, \hat{\boldsymbol{u}}^{(i-1)}$. Letting $\hat{\boldsymbol{V}}_i$ denote the matrix containing the orthonormal basis vectors for the $\hat{\boldsymbol{U}}_i$ as columns, we can express the transformation on the received vector as $P^{\perp}_{\hat{\boldsymbol{U}}_i}\boldsymbol{r} = \boldsymbol{r} - \hat{\boldsymbol{V}}_i\hat{\boldsymbol{V}}_i^H\boldsymbol{r}$.
2) It is possible to obtain power ranking by considering matched-filter outputs averaged over several symbols. Although such a strategy would result in a better power estimate for a static channel with no power control, averaging is expected to yield only marginal improvement in performance for a realistic wireless channel due to the dynamic nature of fast fading, shadowing, fast power control, and other effects present in the channel. In fact, in such an environment, averaging over long runs of symbols would make it difficult to track fast changes in user strength, thereby limiting the effectiveness of the SDP detector.

Remark 4.2 (Complexity of Implemenation): The SDP detector computes $M(K-k+1)$ correlations in the kth stage. Since the total number of stages is K, the total number of complex multiplications necessary to perform all correlations is $\frac{M}{2}(K)(K+1)N$, where K is the total number of users, N is the processing gain, M is the size of the signaling alphabet. The detector also performs Gramm–Schmidt orthogonalization procedure on K vectors each of length N at a cost of $\frac{N}{2}(K)(K+1)$ complex multiplications. The total complexity of the proposed SDP detector is $\frac{N}{2}(K)(K+1)(M+1)$ complex multiplications. The complexity is thus quadratic in the number of users. Note that the complexity of this detector can be significantly reduced if the order of demodulation is completely determined on the basis of measurements obtained only in the first stage. While this detector has a lower complexity, $KN(2M+\frac{K+1}{2})$, it is found to give poor performance, and is therefore not considered any further.

A. Performance Analysis

The performance analysis is based on the observation that, if there exists a significant power imbalance among users, then with high probability, the detector decodes users in the order strongest to weakest, and the decisions of the detector are correct. This motivates us to consider a *chain-based* analysis of the SDP detector, where a chain refers to a particular permutation of the set $\{1, \ldots, K\}$. Each permutation corresponds to a particular order of user decoding followed by the detector. As before, we consider asymptotic performance measures, and derive a lower bound on the asymptotic efficiency of the SDP detector. Throughout the analysis, we condition on the vector symbol $\boldsymbol{j}$ being sent and assume that the first user is the desired user.

Notation: Let Π denote the set of all possible permutations of $\{1, \ldots, K\}$ and π denote a particular member of this set. Let $\pi(i)$ denote the stage of demodulation for the ith user and $\pi^{-1}(i)$ denote the user demodulated at the ith stage. Let us now define the following auxiliary events.

- $C_{i,\pi}$: SDP detector follows order π in stages 1 through i, and arrives at the correct decisions in stages 1 through i.
- $E_{i,\pi}$: Complement of event $C_{i,\pi}$.
- E: SDP detector arrived at an incorrect decision for the first user.
- C: SDP detector arrived at the correct decision for the first user.

Consider now, for a given demodulation order π, the operation of a genie-based SDP detector, in which, at the ith stage, the received vector is projected orthogonal to the actual (rather than the estimated) signal vectors of users to be decoded in stages 1 through $i-1$ as specified by π, i.e., at the ith stage, the decision statistic is obtained by projecting the received vector $\boldsymbol{r}$ orthogonal to the subspace spanned by the

$$\{\boldsymbol{u}_k^{j_k} : k = \pi^{-1}(1), \ldots, \pi^{-1}(i-1)\}.$$

Let

- $\tilde{E}_{i,\pi}$: Genie-based detector for order π arrives at an incorrect decision in the ith stage.

Remark 4.3: The event $\tilde{E}_{i,\pi}$ is defined with respect to the given order π, i.e., in the ith stage the detector can make two types of incorrect decisions: 1) decoding the correct user (as specified by the order of demodulation π in the ith stage) but making an erroneous decision on that user, or 2) picking and decoding a user not specified by the order π. Note that the incorrect decision of type two occurs whether or not the user not specified by the order π is decoded correctly or incorrectly, and it accounts for the possibility of the detector not following order π.

Finally, let us define $f_{\boldsymbol{j}}(i, \pi)$, the asymptotic efficiency of the ith stage of the genie-based detector as follows:

$$f_{\boldsymbol{j}}(i, \pi) = \lim_{\sigma^2\to 0}\frac{\log P[\tilde{E}_{i,\pi}]}{\log P_{e,SU}}. \tag{30}$$

Remark 4.4: As shown below, the asymptotic efficiency of the genie-based SDP detector can be explicitly computed by using a special case of the theory developed in the previous section for the GLRT detector. To lower-bound the asymptotic efficiency of the SDP detector, we note that the probability of

the SDP detector providing a correct decision for the desired user is lower-bounded by the probability that all users demodulated prior to and including the desired user are demodulated correctly. The latter probability can be expressed as an intersection of the probabilities of correct decisions for the genie-based SDP detectors (there is one such detector for each decoding order π and hypothesis $\boldsymbol{j}$). This, in turn, leads to an upper bound on the probability of error for the desired user in terms of the error events of the genie-based detectors, which yields the required lower bound in terms of the asymptotic efficiencies of these detectors. Such arguments can be used to show, as stated in the following theorem, that the SDP detector is asymptotically efficient if, for each $\boldsymbol{j}$, at least one order of demodulation exists for which the stages of a genie-based detector prior to and including the stage in which the desired user is demodulated are asymptotically efficient. Under conditions of sufficient power disparity, demodulating users in the order of strongest to weakest would result in asymptotically efficient stages. However, for equal powers, the early stages for any order of demodulation would typically suffer from an interference floor. We therefore do not expect the SDP detector to be near–far resistant (i.e., to be asymptotically efficient regardless of the choice of user powers). This remark is confirmed by the numerical results presented in the next section.

Theorem 4.1: The asymptotic efficiency (γ) of the SDP detector for the first user can be lower-bounded as follows:

$$\gamma_{\mathrm{SDP}} \geq \min_{\boldsymbol{j}} \max_{\pi \in \Pi} \min_{1 \leq i \leq \pi(1)} f_{\boldsymbol{j}}(i, \pi).$$

Proof: Consider an arbitrary order of demodulation π and event $C_{\pi(1),\pi}$. Clearly, if this event occurs then the SDP detector arrives at the correct decision on the first user. Hence, we have

$$P[C] \geq P[C_{\pi(1),\pi}], \qquad \forall \pi \in \Pi$$

or

$$P[E] \leq P[E_{\pi(1),\pi}], \qquad \forall \pi \in \Pi.$$

For a given $\pi \in \Pi$, let us now consider an event $E_{\pi(1),\pi}$, i.e., detector does not follow order π in some stage i, $1 \leq i \leq \pi(1)$, or follows it but arrives at an incorrect decision in some stage i, $1 \leq i \leq \pi(1)$. This can be decomposed into a disjoint union of events as follows:

$$E_{\pi(1),\pi} = \bigcup_{i=1}^{\pi(1)} C_{i-1,\pi} \cap \tilde{E}_{i,\pi} \tag{31}$$

where $C_{i-1,\pi} \cap \tilde{E}_{i,\pi}$ is an event in which the detector follows order π in stages 1 through $i-1$ and arrives at the correct decisions in those stages, but arrives at an erroneous decision in stage i. Note that this event is expressed in terms of the error probability of the genie-based SDP detector since, given that decisions in stages 1 through $i-1$ are correct, the operation of the actual and genie-based detectors in the ith stage are identical. The decomposition in (31) holds since, for the event $E_{\pi(1),\pi}$ to occur, the detector must eventually arrive at an incorrect decision in some stage prior to or at stage $\pi(1)$, while arriving at the correct decisions in all previous stages. Restating this using the notation above, event $E_{\pi(1),\pi}$ occurs if, and only if, event $C_{i-1,\pi} \cap \tilde{E}_{i,\pi}$ occurs for some $1 \leq i \leq \pi(1)$. Hence, (31) follows. Then

$$P[E] \leq \sum_{i=1}^{\pi(1)} P[C_{i-1,\pi} \cap \tilde{E}_{i,\pi}] \leq \sum_{i=1}^{\pi(1)} P[\tilde{E}_{i,\pi}].$$

Since $\pi \in \Pi$ is arbitrary, we have

$$P[E] \leq \min_{\pi \in \Pi} \sum_{i=1}^{\pi(1)} P[\tilde{E}_{i,\pi}].$$

Removing the conditioning on $\boldsymbol{j}$ and applying the definition of the asymptotic efficiency to the upper bound, we obtain the lower bound specified by the proposition. □

Computation of Bound in Theorem 4.1: To compute γ_{SDP}, we need to specify a procedure for computation of $f_{\boldsymbol{j}}(i, \pi)$, as defined in (30). First, consider event $\tilde{E}_{i,\pi}$, which occurs if the maximum of the single-user decision statistics considered in stage i is not the correlator output corresponding to the correct symbol of the user specified by the order π, i.e.,

$$|(\boldsymbol{u}_z^{j_z})^H \tilde{\boldsymbol{r}}^{(i)}|^2 < |(\boldsymbol{u}_k^{j_k'})^H \tilde{\boldsymbol{r}}^{(i)}|^2$$

for some $k \in \overline{\mathcal{D}}_i$ and some $j_k' \in \mathcal{J}_k$, where $z = \pi^{-1}(i)$ and $\tilde{\boldsymbol{r}}^{(i)} = P_{\boldsymbol{U}_{i-1}}^{\perp} \boldsymbol{r}$. Then, event $\tilde{E}_{i,\pi}$ can be expressed as the union of pairwise inequalities in terms of matched filter outputs as follows:

$$\tilde{E}_{i,\pi} = \bigcup_{j_k' \in \mathcal{J}_k,\, k \in \overline{\mathcal{D}}_i,\, j_k' \neq j_z} \left| (\boldsymbol{u}_z^{j_z})^H \tilde{\boldsymbol{r}}^{(i)} \right|^2 < \left| \left(\boldsymbol{u}_k^{j_k'} \right)^H \tilde{\boldsymbol{r}}^{(i)} \right|^2.$$

From the properties of projection operators, $(\boldsymbol{u}_k^{j_k})^H \tilde{\boldsymbol{r}}^{(i)} = (\tilde{\boldsymbol{u}}_k^{j_k})^H \boldsymbol{r}$, where $\tilde{\boldsymbol{u}}_k^{j_k} = P_{\boldsymbol{U}_{i-1}}^{\perp} \boldsymbol{u}_k^{j_k}$. Thus, event $\tilde{E}_{i,\pi}$ can be equivalently described as an error event for a fictitious, single-user AWGN noncoherent detection problem with signaling alphabet given by $\tilde{\boldsymbol{u}}_k^{j_k}$, $k \in \overline{\mathcal{D}}_i$, $j_k \in \mathcal{J}_k$, where $\boldsymbol{r} = \boldsymbol{s}(\boldsymbol{j}) + \boldsymbol{w}$, and $\tilde{\boldsymbol{u}}_z^{j_z}$ is deemed as the correct output vector. The theory developed for the GLRT detector applies directly in this case, since as shown in Section III, single-user noncoherent detection is, in fact, GLRT detection. In particular, the asymptotic behavior of the $P[\tilde{E}_{i,\pi}]$ is governed by the minimum distance described as follows:

$$d_{j_z, j_k'} = \min \|\boldsymbol{s} - \boldsymbol{y}\|$$

where $\boldsymbol{y}$ is subject to the constraint $|(\tilde{\boldsymbol{u}}_z^{j_z})^H \boldsymbol{y}| = |(\tilde{\boldsymbol{u}}_k^{j_k'})^H \boldsymbol{y}|$. Unlike the minimization in the general GLRT case, this minimization can be carried out explicitly by rewriting the constraint in a more convenient form as follows:

$$(\tilde{\boldsymbol{u}}_z^{j_z} - e^{j\theta} \tilde{\boldsymbol{u}}_k^{j_k'})^H \boldsymbol{y} = 0, \qquad \text{for } 0 \leq \theta \leq 2\pi.$$

For a given θ, the minimization problem corresponds to computing the minimum distance from vector $\boldsymbol{s}$ to the null space of $(\tilde{\boldsymbol{u}}_z^{j_z} - e^{j\theta} \tilde{\boldsymbol{u}}_k^{j_k'})^H$. Then

$$d_{j_z, j_k'}(\theta) = \| P_{\tilde{\boldsymbol{u}}_z^{j_z} - e^{j\theta} \tilde{\boldsymbol{u}}_k^{j_k'}} \boldsymbol{s} \|.$$

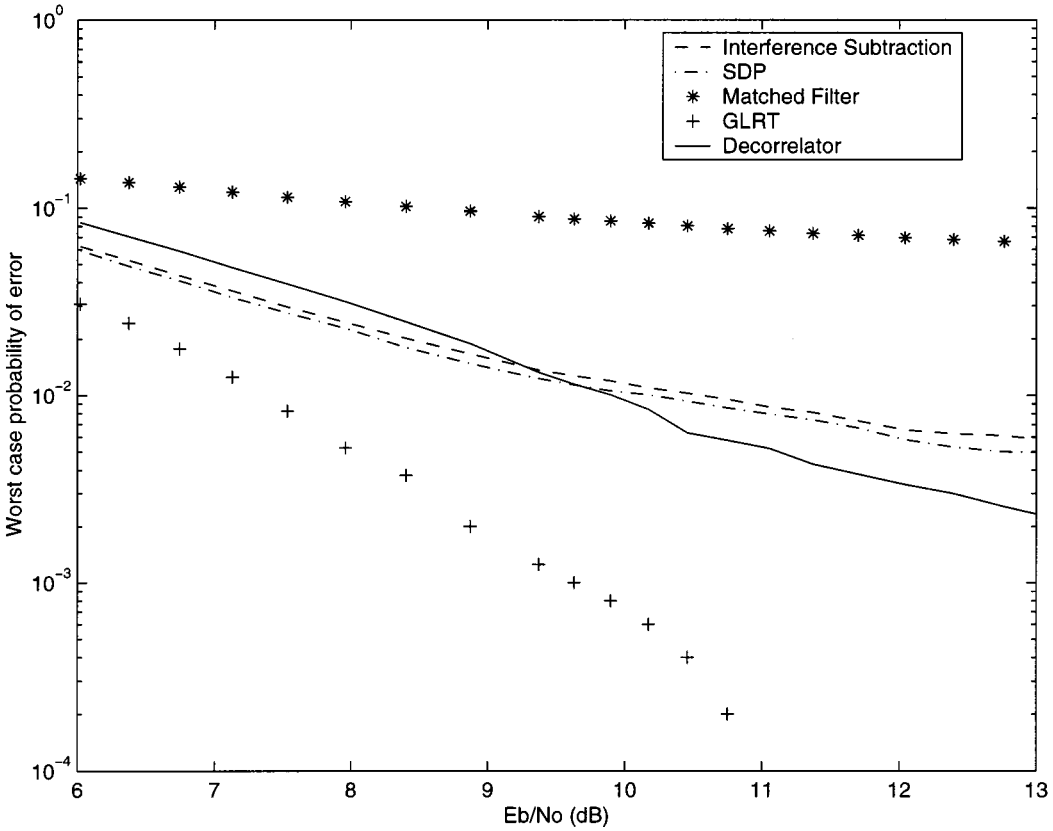

Fig. 5. Numerical comparison of detection schemes for a static AWGN channel. All users are at equal power.

Furthermore, $d_{j_z,j'_k}(\theta)$ can be explicitly minimized over $0 \leq \theta \leq 2\pi$. After some algebra, we obtain

$$d_{j_z,j'_k} = \begin{cases} D & \text{if } \left|(\tilde{\boldsymbol{u}}_z^{j_z})^H \boldsymbol{s}\right|^2 \geq \left|(\tilde{\boldsymbol{u}}_k^{j'_k})^H \boldsymbol{s}\right|^2 \\ 0 & \text{if } \left|(\tilde{\boldsymbol{u}}_z^{j_z})^H \boldsymbol{s}\right|^2 < \left|(\tilde{\boldsymbol{u}}_k^{j'_k})^H \boldsymbol{s}\right|^2 \end{cases}$$

where

$$D = \min\left\{\frac{A_1 + A_2 - 2A_3}{2 - 2A_4}, \quad \frac{A_1 + A_2 + 2A_3}{2 + 2A_4}\right\}$$

$A_1 = |(\tilde{\boldsymbol{u}}_z^{j_z})^H \boldsymbol{s}|^2$, $A_2 = |(\tilde{\boldsymbol{u}}_k^{j'_k})^H \boldsymbol{s}|^2$, $A_3 = \mathrm{Re}[(\tilde{\boldsymbol{u}}_z^{j_z})^H \boldsymbol{s} \cdot ((\tilde{\boldsymbol{u}}_k^{j'_k})^H \boldsymbol{s})^*]$, $A_4 = \mathrm{Re}[(\tilde{\boldsymbol{u}}_z^{j_z})^H \tilde{\boldsymbol{u}}_k^{j'_k}]$.

The ith stage asymptotic efficiency, as a function of π and $\boldsymbol{j}$, can now be expressed as follows:

$$f_{\boldsymbol{j}}(i, \pi) = \frac{\min\limits_{j'_k \in \mathcal{J}_k, k \in \overline{\mathcal{D}}_{i-1}, j'_k \neq j_z} d^2_{j_z,j'_k}}{(d^{su}_{\min})^2}. \tag{32}$$

With this, the asymptotic performance of the SDP detector as in Theorem 4.1 can be computed, for given user amplitudes and signaling alphabets.

V. Numerical Results

In this section, we use computer simulations to compare the performance of the multiuser detectors introduced in previous sections. For the purpose of comparison, the parameters are chosen so that decorrelation with respect to the total interference subspace is feasible. The system considered has a processing gain of 32 and four active users, each employing 4-ary orthogonal modulation. In each symbol period, the spreading sequences are picked at random, with $+1$ and -1 equiprobable in each position. The system is synchronous, and the receiver performs noncoherent reception. Five detection techniques are simulated: matched-filter reception, interference subtraction as in [7], SDP detection, decorrelation with respect to the total interference subspace as in [9], and the GLRT.

Figs. 5 and 6 display simulation results for an AWGN channel. All four users are at equal power for the results in Fig. 5, while the user powers are distributed in 3-dB steps in Fig. 6. The latter corresponds to a significant near–far problem for the weakest user, which has 9 dB less power than the strongest user. The plots display the worst case error probability over all four users versus E_b/N_0 for the weakest user, where E_b denotes the bit energy (which is half the symbol energy in this case). We now make the following observations.

1) The performance of the matched filter under power disparities is poor and not shown. Even for equal powers, which is the regime for which it is designed, its performance is poorer than that of the multiuser detectors considered.
2) The GLRT detector significantly outperforms all other multiuser detectors. With users distributed in 3-dB steps, the GLRT detector exhibits a fast rate of decay of its error probability.
3) Under power disparities, the SDP detector does not exhibit an interference floor (i.e., it has nonzero asymptotic efficiency). Its performance is close to that of the

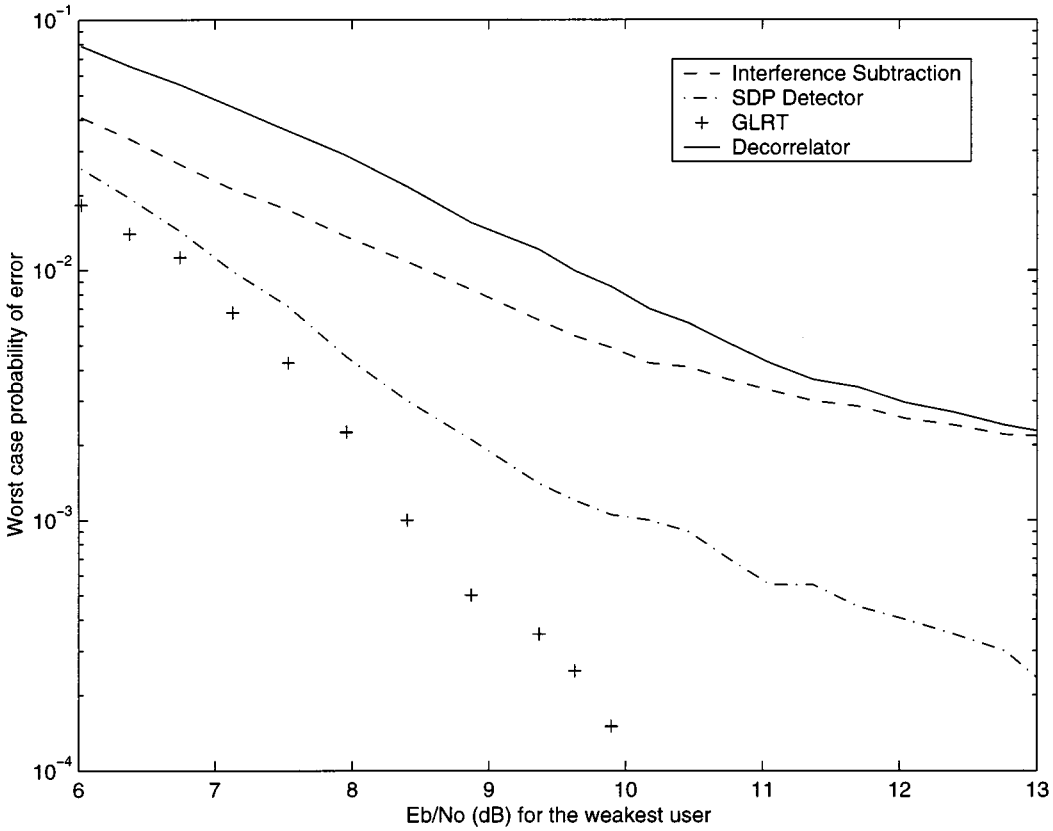

Fig. 6. Numerical comparison of detection schemes for a static AWGN channel. The user powers are distributed in 3-dB steps.

GLRT detector at low signal-to-noise ratio (SNR) to moderate SNR, although its smaller asymptotic efficiency becomes apparent at higher SNR. For equal powers, the SDP detector does have an interference floor indicating zero asymptotic efficiency.

4) Interference subtraction displays an interference floor in its performance under both equal powers and power disparities. This is because of the propagation of residual interference due to errors in amplitude estimation even when correct decisions are made. This problem is avoided by the SDP detector because it projects the received signal orthogonal to the estimated interference vectors rather than attempting to subtract them out.

5) For equal powers, the performance of the decorrelator is comparable to that of the interference subtraction and the SDP detection, even though the latter detector exhibits an interference floor in this regime. On the other hand, under power disparities, the SDP detector performs much better than the decorrelator. This is because the loss in signal energy due to decorrelation, which projects the received signal orthogonal to a 12-dimensional subspace spanned by all possible signals, which are spanned by the three interferers, is much larger than that due to SDP detector. The SDP detector projects the received signal for the user decoded last orthogonal to a three-dimensional space corresponding to the decisions for the other three users. With users distributed in 3-dB steps and at equal powers, the decorrelator does not exhibit an error floor, indicating a nonzero near–far resistance.

Fig. 7 displays lower bounds on the near–far resistance of the GLRT detector and on the asymptotic efficiencies of the GLRT and the SDP detector, with users distributed in 3-dB steps. These quantities were computed for 100 randomly drawn spreading sequences according to the analytical expressions derived in Sections III and IV. Fig. 7 also includes the near–far resistance of the decorrelator which can be computed similarly to the stage asymptotic efficiency of the SDP detector given in (32). From Fig. 7, we observe that the GLRT detector has asymptotic efficiency close to one, which indicates that, with users distributed in 3-dB steps, it incurs almost no loss in performance due to the multiple-access interference present in the channel. Furthermore, the GLRT detector is near–far resistant. Finally, the SDP and the decorrelator detectors exhibit a nonzero value of the asymptotic efficiency. This is consistent with the absence of error floor for these detectors noted in Fig. 6.

Fig. 8 considers a system with the same parameters as before, except that the users are subject to flat Rayleigh fading. All users are at equal average powers, and the fades for different users are independent. The worst case error probability is plotted against E_b/N_0, where E_b now determines the average bit energy for each user. The performance is averaged over a number of symbol intervals, with the fading assumed to be independent across symbol periods. We may think of such a channel in each symbol period as an AWGN channel with imperfect or randomized power control. Consequently, we would expect the SDP detector to outperform both interference subtraction and decorrelation. Fig. 8 shows that this is in fact the case, with the performance disparity especially pronounced at high SNR. The

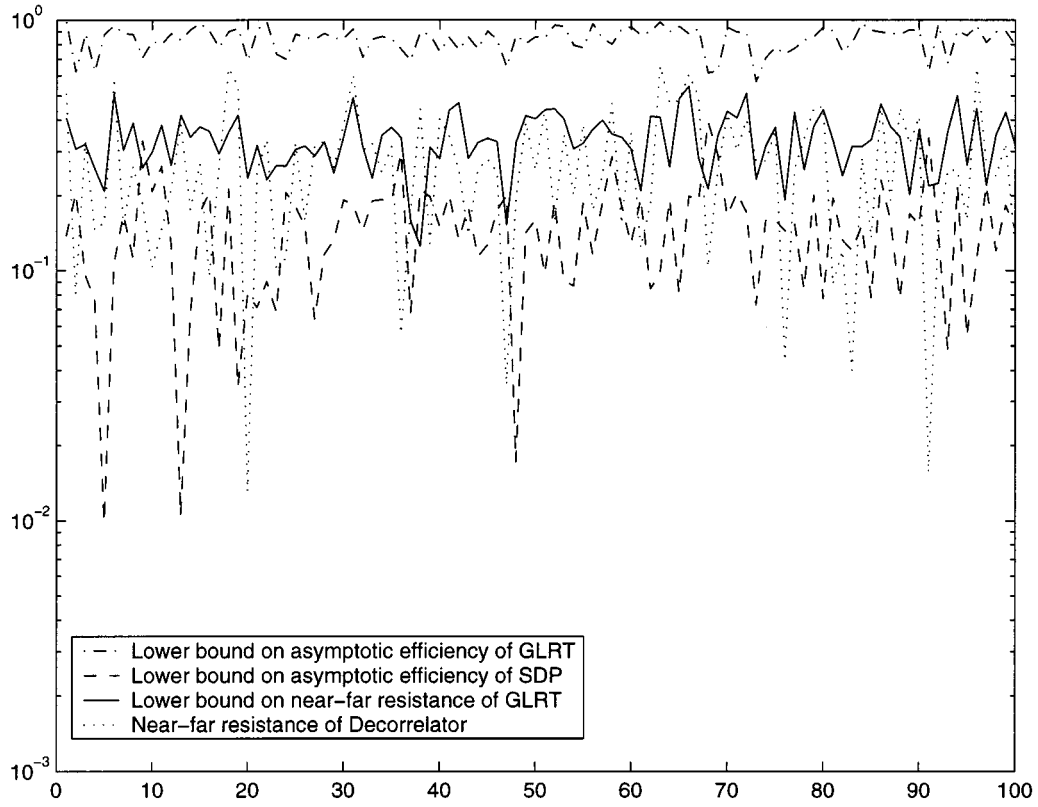

Fig. 7. Comparison of the asymptotic performance measures for the GLRT and the sequential decision projection algorithm.

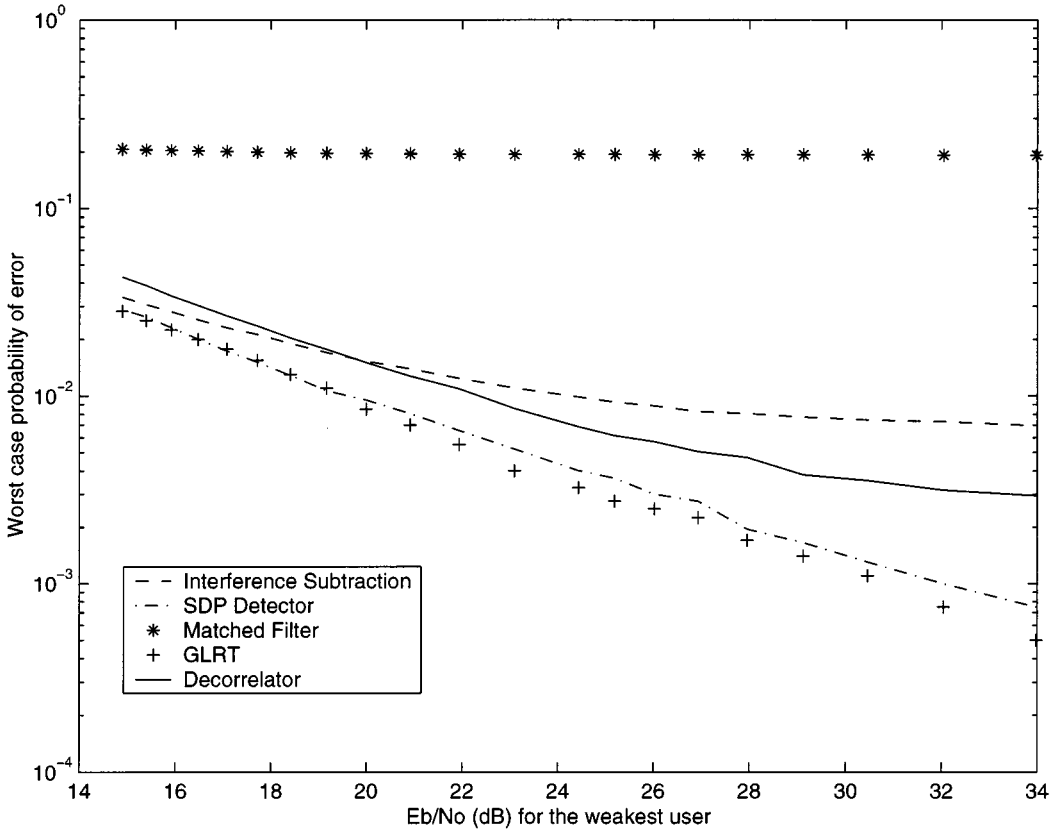

Fig. 8. Numerical comparison of detection schemes in flat-fading environment, with all users at equal average power.

performance of all the multiuser detectors (including the GLRT detector) is very similar for low-to-moderate SNRs, while the GLRT detector outperforms the other detectors at high SNRs. The matched filter is decisively outperformed by the multiuser detectors.

VI. Conclusion

The results in this paper provide insight into the performance of coded, or nonlinearly modulated, multiuser systems for both coherent and noncoherent demodulation. The non-Bayesian formulation for noncoherent multiuser detection presented here provides an attractive alternative to formulations which explicitly model the complex amplitudes of the users (e.g., using a complex Gaussian distribution for Rayleigh or Rician fading). The asymptotic performance analysis of the GLRT (as the noise level tends to zero) requires the development of geometric insights into noncoherent detection that are analogous to classical signal space notions for coherent detection.

The quadratic complexity (in the number of users) SDP detector is a practical approach to successive decoding that does not require knowledge of user amplitudes, and that does not suffer from an interference floor in its performance provided that the user powers are suitably disparate. Since the ordering of user powers is not required to be known *a priori*, SDP detection is ideally suited for systems with fast fading. As shown in our numerical results, even when all users have equal average powers, SDP detection is able to exploit the instantaneous power disparities due to Rayleigh fading. For systems with slow or no fading, SDP detection combines well with power control that enforces disparities, provided that the level of thermal noise and of other types of background interference is low enough so that some minimal post-interference suppression SNR can be guaranteed for the low-power users. Enforcing a controlled amount of power disparity, allowing users with smaller path gains to be received at lower powers, has the combined virtue of reducing the transmitted power of users at cell edges, and of allowing successive decoding to work well. Under idealized assumptions, it has been shown in [21], [18] that such an approach can lead to up to a twofold gain in capacity over systems with conventional power control.

Appendix A
Proof of Lemma 3.1

We first establish the following preliminary result.

Lemma A.1: There exists $A > 1$ such that $\{\boldsymbol{s} + \alpha \boldsymbol{x}_*\colon 1 < \alpha < A\}$ lies in the error region $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$.

Proof: Consider function $f(\alpha)\colon [0, \infty) \to \mathcal{R}$

$$\begin{aligned} f(\alpha) &= \|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}(\boldsymbol{s} + \alpha \boldsymbol{x}_*)\|^2 - \|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}}}(\boldsymbol{s} + \alpha \boldsymbol{x}_*)\|^2 \\ &= \alpha^2(\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{x}_*\|^2 - \|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}}}\boldsymbol{x}_*\|^2) \\ &\quad + 2\alpha \mathrm{Re}\langle P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{s}, \boldsymbol{x}_*\rangle + \|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{s}\|^2. \end{aligned}$$

The function $f(\alpha)$ is quadratic in α, $f(\alpha) > 0$ for $0 \le \alpha < 1$, and $f(1) = 0$. This implies that there are only four possible shapes for $f(\alpha)$ shown in Fig. 9. In cases Fig. 9 (a)–(c), clearly there exists an $A > 1$ such that $f(\alpha) < 0$ for $1 < \alpha < A$, which is equivalent to $\boldsymbol{s} + \alpha \boldsymbol{x}_* \in \mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$ for $1 < \alpha < A$. It remains to show that case of Fig. 9 (d) cannot occur.

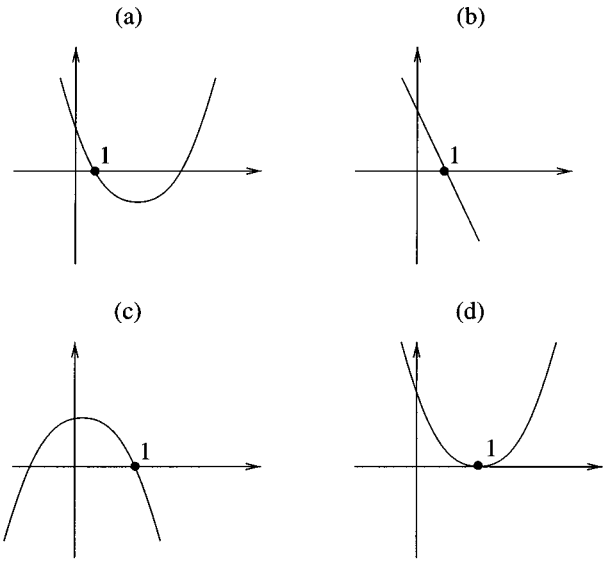

Fig. 9. Possible shapes of $f(\alpha)$.

To this end, we note that case of Fig. 9 (d) corresponds to f having repeated roots at $\alpha = 1$. In order for f to have repeated roots at $\alpha = 1$, we must have $f(\alpha) = c(\alpha^2 - 2\alpha + 1)$ for some constant c, so that

$$\left\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{x}_*\right\|^2 - \left\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}}}\boldsymbol{x}_*\right\|^2 = \left\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{s}\right\|^2 - \mathrm{Re}\left\langle P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{s}, \boldsymbol{x}_*\right\rangle. \quad \text{(A1)}$$

We observe the following.

1) $\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{s}\|^2$ is the distance from $\boldsymbol{s}$ to a point in $\mathcal{S}_{\boldsymbol{j}'}$, which must be at least as large as the minimum distance from $\boldsymbol{s}$ to $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$, since $\mathcal{S}_{\boldsymbol{j}'} \subset \mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$. Thus, we have $\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{s}\|^2 \ge \|\boldsymbol{x}_*\|^2$.
2) $\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{x}_*\|^2 \le \|\boldsymbol{x}_*\|^2$ with equality if and only if $\boldsymbol{x}_* \in \mathcal{S}^{\perp}_{\boldsymbol{j}'}$.
3) $\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}}}\boldsymbol{x}_*\|^2 \ge 0$ with equality if and only if $\boldsymbol{x}_* \in \mathcal{S}_{\boldsymbol{j}}$.

Thus,

$$\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{x}_*\|^2 - \|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}}}\boldsymbol{x}_*\|^2 \le \|\boldsymbol{x}_*\|^2$$

with equality if and only if $\boldsymbol{x}_* \in \mathcal{S}^{\perp}_{\boldsymbol{j}'} \cap \mathcal{S}_{\boldsymbol{j}}$

Combining these observations with (A1) we obtain the following chain of relationships:

$$\left\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{x}_*\right\|^2 - \left\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}}}\boldsymbol{x}_*\right\|^2 \le \|\boldsymbol{x}_*\|^2 \le \left\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{s}\right\|^2 - \mathrm{Re}\left\langle P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{s}, \boldsymbol{x}_*\right\rangle. \quad \text{(A2)}$$

Since we require equalities throughout (A2), we conclude that $\boldsymbol{x}_* \in \mathcal{S}^{\perp}_{\boldsymbol{j}'} \cap \mathcal{S}_{\boldsymbol{j}}$, which implies that $\mathrm{Re}\langle P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{s}, \boldsymbol{x}_*\rangle = 0$, which in turn implies that all terms in (A1) are zero. In particular, we have $\|P^{\perp}_{\mathcal{S}_{\boldsymbol{j}'}}\boldsymbol{s}\|^2 = 0$ which can occur only if $\boldsymbol{s} \in \mathcal{S}_{\boldsymbol{j}'}$. Thus the case of Fig. 9 (d) can occur only if $\boldsymbol{s} \in \mathcal{S}_{\boldsymbol{j}'}$, which we assume is not the case to avoid the trivial case of $d(\boldsymbol{j}'|\boldsymbol{j}) = 0$. □

We are now ready to prove Lemma 3.1, which we restate for convenience.

Lemma 3.1: There exists $A > 1$ and, for n large enough, there exists $\epsilon_n > 0$ such that

$$E_n = \left\{ d(\boldsymbol{j}'|\boldsymbol{j}) \left(1+\frac{1}{n}\right) \le w_* \le \left(A-\frac{1}{n}\right) d(\boldsymbol{j}'|\boldsymbol{j}) \right\} \cap \left\{ \left\| \boldsymbol{w}_*^{\perp} \right\| \le \epsilon_n \right\}$$

lies in the error region $\mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$.

Proof: We rewrite $f(\alpha)$ of Lemma A.1 in a more general form as $F(\boldsymbol{w})$: $\mathcal{C}^N \to \mathcal{R}$

$$F(\boldsymbol{w}) = \|P_{\mathcal{S}_{\boldsymbol{j}'}}^{\perp}(\boldsymbol{s}+\boldsymbol{w})\|^2 - \|P_{\mathcal{S}_{\boldsymbol{j}}}^{\perp}(\boldsymbol{s}+\boldsymbol{w})\|^2.$$

Also, for simplicity of notation, we define the following sets

1) $\hat{E}_n = \{(1+\frac{1}{n})\, d(\boldsymbol{j}'|\boldsymbol{j}) \le w_* \le (A-\frac{1}{n})\, d(\boldsymbol{j}'|\boldsymbol{j})\} \cap \{\|\boldsymbol{w}_*^{\perp}\| = 0\}$
2) $B_n = \{(1+\frac{1}{n})\, d(\boldsymbol{j}'|\boldsymbol{j}) \le w_* \le (A-\frac{1}{n})\, d(\boldsymbol{j}'|\boldsymbol{j})\} \cap \{\|\boldsymbol{w}_*^{\perp}\| \le M\}$

where M is a large constant. Observe that

1) The functional $F(\boldsymbol{w})$ is continuous and the set B_n is closed and bounded, which implies that $F(\boldsymbol{w})$ is uniformly continuous on the B_n. Furthermore, the functional $F(\boldsymbol{w}) < 0$ whenever $\boldsymbol{s}+\boldsymbol{w} \in \mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$.
2) From Lemma A.1, there exists an N, such that for all $n > N$ we have $\hat{E}_n \in \mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$.

Let $n > N$. Then, $F(\boldsymbol{w}) < 0$ for $\boldsymbol{w} \in \hat{E}_n$. By continuity of $F(\boldsymbol{w})$, there exists $\epsilon(\boldsymbol{w})$, which depends on $\boldsymbol{w}$, such that $F(\boldsymbol{w}) < 0$ for each $\boldsymbol{w} \in \hat{E}_n$ and $\|\boldsymbol{w}_*^{\perp}\| < \epsilon(\boldsymbol{w})$. Furthermore, by uniform continuity of $F(\boldsymbol{w})$ there exists a single ϵ_n, not dependent on $\boldsymbol{w}$, such that $F(\boldsymbol{w}) < 0$ for all $\boldsymbol{w} \in \hat{E}_n$ and $\|\boldsymbol{w}_*^{\perp}\| < \epsilon_n$, or, equivalently, $F(\boldsymbol{w}) < 0$ for $\boldsymbol{w} \in E_n$. Thus, for large enough N, there exists ϵ_n such that $E_n \in \mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$. □

Appendix B
Proof of the Lower Bound in Theorem 3.2

Theorem 3.2: For each hypothesis pair $(\boldsymbol{j}, \boldsymbol{j}')$, the minimum distance for the GLRT detector is lower-bounded and upper-bounded as

$$\|\tilde{\boldsymbol{s}}\| \sin\frac{\theta_{\min}}{2} \le d(\boldsymbol{j}'|\boldsymbol{j}) \le \|\tilde{\boldsymbol{s}}\| \sin\theta_{\tilde{\mathcal{S}}_{\boldsymbol{j}'}}(\tilde{\boldsymbol{s}})$$

where $\theta_{\min} = \theta_{\min}(\tilde{\mathcal{S}}_{\boldsymbol{j}}, \tilde{\mathcal{S}}_{\boldsymbol{j}'})$.

Notation: All vectors and subspaces involved in the proof are projected orthogonal to the irrelevant subspace $\mathcal{S}_{\boldsymbol{j}} \cap \mathcal{S}_{\boldsymbol{j}'}$. However, for simplicity, we drop the tilde notation introduced in Remark 3.1 within the proof.

Proof: The upper bound is established in the main body of the paper. Here we concentrate on the lower bound. We have $d(\boldsymbol{j}'|\boldsymbol{j}) = \min_{\boldsymbol{y} \in \mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})} \|\boldsymbol{s}-\boldsymbol{y}\|$. Any component of vector $\boldsymbol{y}$ which does not lie in the $\mathcal{S}_{\boldsymbol{j}} \cup \mathcal{S}_{\boldsymbol{j}'}$ is irrelevant for the $\boldsymbol{j}$ versus $\boldsymbol{j}'$ binary hypothesis testing problem. Thus, without loss of generality, we may write $\boldsymbol{y} = \boldsymbol{y}_1 + \boldsymbol{y}_2$, where $\boldsymbol{y}_1 \in \mathcal{S}_{\boldsymbol{j}}$ and $\boldsymbol{y}_2 \in \mathcal{S}_{\boldsymbol{j}'}$. Since $\boldsymbol{y} \in \mathcal{D}(\boldsymbol{j}' > \boldsymbol{j})$, $\|P_{\mathcal{S}_{\boldsymbol{j}}}^{\perp}\boldsymbol{y}\| \ge \|P_{\mathcal{S}_{\boldsymbol{j}'}}^{\perp}\boldsymbol{y}\|$.Thus,

$$\left\| P_{\mathcal{S}_{\boldsymbol{j}}}^{\perp}\boldsymbol{y}_2 \right\| \ge \left\| P_{\mathcal{S}_{\boldsymbol{j}'}}^{\perp}\boldsymbol{y}_1 \right\|. \tag{B1}$$

The distance $\|\boldsymbol{s}-\boldsymbol{y}\|$ can be decomposed as

$$\|\boldsymbol{s}-\boldsymbol{y}\|^2 = \left\|\boldsymbol{s}-\boldsymbol{y}_1 - P_{\mathcal{S}_{\boldsymbol{j}}}\boldsymbol{y}_2\right\|^2 + \left\|P_{\mathcal{S}_{\boldsymbol{j}}}^{\perp}\boldsymbol{y}_2\right\|^2.$$

Let $\|\boldsymbol{y}_1\| = \alpha$ and $\|P_{\mathcal{S}_{\boldsymbol{j}}}\boldsymbol{y}_2\| = \beta$. Define θ_1 as the minimum angle between $\boldsymbol{y}_1$ and the subspace $\mathcal{S}_{\boldsymbol{j}'}$, and θ_2 as the minimum angle between $\boldsymbol{y}_2$ and the subspace $\mathcal{S}_{\boldsymbol{j}}$. Then, from (22), we have $\|P_{\mathcal{S}_{\boldsymbol{j}}}^{\perp}\boldsymbol{y}_2\| = \beta\tan\theta_2$ and $\|P_{\mathcal{S}_{\boldsymbol{j}'}}^{\perp}\boldsymbol{y}_1\| = \alpha\sin\theta_1$. By the Cauchy–Schwartz inequality, we have

$$(\|\boldsymbol{s}\| - \alpha - \beta)^2 + \beta^2\tan^2\theta_2 \le \|\boldsymbol{s}-\boldsymbol{y}\|^2 \tag{B2}$$

where, from (B1), $\beta\tan\theta_2 \ge \alpha\sin\theta_1$, and $\theta_{\min}(\tilde{\mathcal{S}}_{\boldsymbol{j}}, \tilde{\mathcal{S}}_{\boldsymbol{j}'}) \le \theta_1$, $\theta_2 \le \pi/2$. Our strategy for obtaining the lower bound is to minimize the left-hand side of (B2). For conciseness, let

$$T = (\|\boldsymbol{s}\| - \alpha - \beta)^2 + \beta^2\tan^2\theta_2$$

and $B = \alpha + \beta$.

Minimization of T will be performed in two steps. First, we minimize over β for fixed B, then we minimize over all $B \ge 0$. From (B1), we have

$$\beta \ge \frac{\alpha\sin\theta_1}{\tan\theta_2}.$$

Then

$$\beta_{\min} = \frac{\alpha_{\min}\sin\theta_1}{\tan\theta_2}$$

and

$$\alpha_{\min}\left(\frac{\sin\theta_1}{\tan\theta_2}+1\right) = B.$$

Substituting, we obtain the following values of $\beta_{\min}$ and T as a function of B:

$$\beta_{\min}(B) = \frac{B\sin\theta_1}{\sin\theta_1 + \tan\theta_2}$$

$$T(B) = (\|\boldsymbol{s}\| - B)^2 + B^2\frac{\sin^2\theta_1\tan^2\theta_2}{(\sin\theta_1+\tan\theta_2)^2} = (\|\boldsymbol{s}\| - B)^2 + B^2\mu$$

where

$$\mu = \frac{1}{\left(\frac{1}{\tan\theta_2}+\frac{1}{\sin\theta_1}\right)^2}.$$

$T(B)$ is a convex quadratic function of B and can be easily minimized by setting the first derivative with respect to B to zero. It is easily verified that the minimum value is achieved at $B^* = \|\boldsymbol{s}\|/(1+\mu)$, and the minimum value is

$$T(B^*) = \frac{\|\boldsymbol{s}\|^2}{1+\frac{1}{\mu}} = \frac{\|\boldsymbol{s}\|^2}{1+\left(\frac{1}{\tan\theta_2}+\frac{1}{\sin\theta_1}\right)^2}.$$

To complete the derivation of the lower bound, we consider the following two cases.

Case 1: Let

$$\cos\theta_{\min} \ge \sin\theta_1/\tan\theta_2, \qquad \text{where } \theta_{\min} = \theta_{\min}(\tilde{\mathcal{S}}_{\boldsymbol{j}}, \tilde{\mathcal{S}}_{\boldsymbol{j}'}).$$

In this case, we have the following chain of inequalities:

$$d(\boldsymbol{j}'|\boldsymbol{j}) \geq T(B^*) = \frac{||\boldsymbol{s}||^2 \sin^2\theta_1}{\sin^2\theta_1 + \left(\frac{\sin\theta_1}{\tan\theta_2} + 1\right)^2}$$
$$\geq \frac{||\boldsymbol{s}||^2 \sin^2\theta_1}{\sin^2\theta_1 + (\cos\theta_{\min} + 1)^2}$$
$$\geq \frac{||\boldsymbol{s}||^2 \sin^2\theta_{\min}}{\sin^2\theta_{\min} + (\cos\theta_{\min} + 1)^2} = ||\boldsymbol{s}||^2 \sin^2\frac{\theta_{\min}}{2}.$$

Case 2: Let

$$\cos\theta_{\min} \leq \sin\theta_1/\tan\theta_2, \qquad \text{where } \theta_{\min} = \theta_{\min}(\tilde{\mathcal{S}}_{\boldsymbol{j}}, \tilde{\mathcal{S}}_{\boldsymbol{j}'}).$$

In this case, we have the following chain of inequalities:

$$d(\boldsymbol{j}'|\boldsymbol{j}) \geq T(B^*) = \frac{||\boldsymbol{s}||^2 \tan^2\theta_2}{\tan^2\theta_2 + \left(\frac{\tan\theta_2}{\sin\theta_1} + 1\right)^2}$$
$$\geq \frac{||\boldsymbol{s}||^2 \tan^2\theta_2}{\tan^2\theta_2 + \left(\frac{1}{\cos\theta_{\min}} + 1\right)^2}$$
$$\geq \frac{||\boldsymbol{s}||^2 \tan^2\theta_{\min}}{\tan^2\theta_{\min} + \left(\frac{1}{\cos\theta_{\min}} + 1\right)^2}$$
$$= ||\boldsymbol{s}||^2 \sin^2\frac{\theta_{\min}}{2}. \qquad \square$$

REFERENCES

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

[2] ——, "Minimum probability of error for asynchronous Gaussian multiple-access channel," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85–96, Jan. 1986.

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

[4] ——, "Near–far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Commun.*, vol. 38, pp. 496–508, Apr. 1990.

[5] M. K. Varanasi and B. Aazhang, "Near optimum demodulation for coherent communications in asynchronous Gaussian CDMA channels," *IEEE Trans. Commun.*, vol. 39, pp. 725–736, May 1991.

[6] A. Duel-Hallen, "A family of multiuser decision-feedback detectors for asynchronous code division multiple access channels," *IEEE Trans. Commun.*, vol. 43, Feb./Mar./Apr. 1995.

[7] P. Patel and J. Holtzman, "Analysis of a simple successive interference cancellation scheme in a DS/CDMA system," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 796–807, June 1994.

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

[9] M. K. Varanasi and A. Russ, "Noncoherent decorrelative detection for nonorthogonal multipulse modulation over the multiuser Gaussian channel," *IEEE Trans. Commun.*, vol. 46, pp. 1675–1684, Dec. 1998.

[10] M. K. Varanasi and D. Das, "Noncoherent decision feedback multiuser detection," *IEEE Trans. Commun.*, vol. 48, pp. 259–269, Feb. 2000.

[11] A. Russ and M. K. Varanasi, "Noncoherent multiuser detection for nonlinear modulation over the Rayleigh fading channel," *IEEE Trans. Inform. Theory*, to be published.

[12] M. K. Varanasi and B. Aazhang, "Optimally near–far resistant multiuser detection in differentially coherent synchronous channels," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1006–1018, July 1991.

[13] M. K. Varanasi and M. Brehler, "A systematic approach to noncoherent detection for DPSK modulation in multiuser correlated diversity Rayleigh fading channels," in *Proc. Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 1998, pp. 236–241.

[14] M. McCloud and L. L. Scharf, "Generalized likelihood detection on multiple access channels," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1997.

[15] M. McCloud, L. L. Scharf, and L. T. Mc Whorter, "Subspace coherence for detection in multiuser additive noise channels," in *Proc. Signal Processing Appl.ications in Wireless Communication (SPAWC'97)*, Paris, France, Apr. 1997.

[16] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.

[17] J. Hegarty and B. R. Vojcic, "Noncoherent multistage multiuser detection of M-ary orthogonal signals," *Wireless Networks*, vol. 4, pp. 319–324, 1998.

[18] D. G. Warrier and U. Madhow, "On the capacity of cellular CDMA with controlled power disparities," in *Proc. IEEE Vehicular Technology Conf.*, 1998, pp. 1873–1878.

[19] L. B. Nelson and H. V. Poor, "Iterative multiuser receivers for CDMA channels: An EM-based approach," *IEEE Trans. Commun.*, vol. 44, pp. 1700–1710, Dec. 1996.

[20] U. Madhow, "Blind adaptive interference suppression for the direct-sequence CDMA," *Proc. IEEE*, vol. 86, pp. 2049–2069, Oct. 1998.

[21] C. C. Chan and S.V. Hanly, "The capacity improvement of an integrated successive decoding and power control scheme," in *1997 IEEE 6th Int. Conf. Universal Personal Communications Rec.*, Oct. 1997, pp. 800–804.

[22] G. E. Bottomley, "Improved successive cancellation of DS-CDMA signals using signal orthogonalization," in *1996 IEEE 5th Int. Conf. Universal Personal Communications Rec.*, Sept. 1996, pp. 141–144.