FP9 - 5:00

Proceedings of 25th Conference
on Decision and Control
Athens, Greece • December 1986

ASYMPTOTIC EFFICIENCY OF LINEAR MULTIUSER DETECTORS

Ruxandra Lupas-Golaszewski and Sergio Verdú

Department of Electrical Engineering
Princeton University
Princeton, NJ 08544

Abstract: **Demodulation of data streams transmitted synchronously by several users over a Gaussian multiple access channel is considered. Each user modulates a different signal from a linearly independent signal set. The asymptotic efficiency criterion is used to evaluate the perfomance of different detection rules. The two currently most important detectors are presented and compared: the optimum multiuser detector whose complexity is exponential in the number of users and the single user detector whose performance degrades very rapidly as the multiuser interference energy increases. In contrast, a memoryless linear transformation on the matched filter outputs of the receiver is proposed, which is shown to guarantee a high lower bound on the asymptotic efficiency, which is independent of signal energies, while the time-complexity per bit is linear in the number of users. An algorithm which finds the best linear transformation is given, along with sufficient conditions on the signal energies and crosscorrelations to ensure optimum asymptotic efficiency of the proposed detector.**

1. Introduction

K users are transmitting simultaneously over a multiple-access channel perturbed by additive white Gaussian noise. To this end each user modulates a different signal waveform. In this paper we will assume that the transmissions occur symbol-synchronously, in which case the waveform received in the j^{th} symbol interval has the form:

$$r(t)=\sum_{i=1}^{K} x_i(j)s_i(t-jT)+n(t), \quad t \in [jT,jT+T] \tag{1.1}$$

where x_i, $s_i(t)$ are the transmitted symbol and the signal waveform corresponding to user i. If the modulating signals are mutually orthogonal, the problem reduces to K single-user problems. It is well-known that the optimal detector is linear in this case: a bank of matched filters followed by thresholds. Unfortunately, bandwidth limitations prevent the usage of orthogonal signal waveforms for a large number of users. In this case, even for low signal crosscorrelations, it is far from optimal to continue using single-user detectors, neglecting the lack of orthogonality of the signal set ([1]). As shown in the sequel, the performance of the single-user detector degrades monotonically with increasing relative energy of the interfering users. The maximum likelihood multiuser detector is nonlinear and has been shown ([2]) to be NP-complete, that is, there is no corresponding decision algorithm which is polynomial in the number of users, unless NP=P. However, the performance of the optimum detector is considerably superior to that of the conventional single-user detector.

The purpose of this paper is to derive and analyze *linear* multiuser detection schemes, in order to improve the error rate in comparison to the bank of single-user detectors used in practice, while maintaining approximately the same ease of computation. We insert a linear memoryless transformation between the matched filters and the thresholds and investigate which linear transformation of the matched-filter outputs has the best performance with respect to bit error rate, and how close this performance gets to that of the optimal receiver. The performance criterion used is the ***asymptotic efficiency***, which is a good measure of the performance loss due to other active users on the channel, in the high SNR region. The k^{th} user asymptotic efficiency η_k of a detector whose k_{th} user bit error rate is equal to P_k has been defined in [1] as

$$\eta_k = \sup\left\{0 \le r \le 1; \qquad P_k(1/\sigma) = O\left(\, Q\left((rw_k)^{1/2}/\sigma\right)\right)\right\} \tag{1.2}$$

where w_k is the energy of user k and Q(x) is the complement of the Gaussian cumulative distribution function. Since in the absence of other users the probability of bit error for user k is equal to $Q(\sqrt{w_k}/\sigma)$, η_k is equal to the limit as $\sigma \rightarrow 0$ of the ratio between the effective energy (that required by a single user to achieve the same error probability) and w_k, the actual energy of the k^{th} user.

To illustrate the concept of effective energy, consider the maximum likelihood, i.e. optimum detector for the binary antipodal synchronous K-user problem, with independent noise in each component. This is known to be the detector which, from the 2^K hypotheses, picks the one which has the minimum Euclidean distance with respect to L_2 norm distance to the received signal. Asymptotically, as the signal-to-noise ratio increases, the bit error probability for the k^{th} user is determined by the distance between the closest pair of possible hypotheses differing in the k^{th} bit. Thus the effective k^{th} user energy is that of a single user decision between the pair of closest K-user hypotheses differing in the k^{th} bit.

This paper is structured as follows. In Section 2, and 4 the two most important detectors are presented and their asymptotic performance is derived, as measured by the asymptotic efficiency : the single-user detector, which is currently used in practice for multiuser communication systems, and the optimum detector, the maximum likelihood detector. The most pertinent results available are presented. It is shown that k^{th} user efficiency of the single-user detector degrades to zero with growing energy of other users, while on the other hand the computational overhead of the optimal detector is exponential in the number of users, which precludes its use in practical situations where, say, $K > 15$.

Motivated by the wide gap between the performances of the optimum and conventional systems, we present in Section 3 another detector, the ***decorrelating*** detector, which decouples the matched filter outputs by multiplication with the inverse crosscorrelation matrix. It is shown to have an asymptotic efficiency which is independent of the energy of other users, while exhibiting a linear time complexity per bit.

In Section 4 the best linear detector is found. The 2-user case is analyzed in detail, then an algorithm is given which finds the best linear detector for the K-user case. Interestingly enough, there is a region of signal energies and crosscorrelations for which the performance of the optimal detector is achieved. Sufficient conditions are given for this case, which are necessary for the 2-user case. Outside this region, the decorrelating detector gives a lower bound on achievable performance, which is tight for another region of signal energies and crosscorrelations. We give sufficient conditions for this case. These conditions are also necessary in the 2-user case. While the maximization procedure given in part 4.2 is exponential, it has to be performed only once, when initially finding the best linear detector. Afterwards, the overall real-time computational effort is quadratic in the number of users.

Notation: K users transmit over a synchronous Gaussian multiple access channel, modulating antipodally a set of signal waveforms $s_k(t), t \in [0,T], k=1,...,K$. It is assumed that the signal set is linearly independent. The signal transmitted in the $(j+1)^{th}$ symbol interval is $s(t) = \sum_{i=1}^{K} x_i s_i(t-jT)$, $t \in [jT,jT+T]$ where x_i is the symbol sent by user i and T is the duration of one symbol interval. On the channel the signal is perturbed by additive white Gaussian noise with power spectral density σ^2. The k^{th} matched filter output is

$$y_k = \int_{jT}^{jT+T} r(t)s_k(t-jT)dt = \sum_{i=1}^{K} x_i R_{ik} + n_k = (\mathbf{R}\mathbf{x})_k + \mathbf{n}_k \tag{1.3}$$

---

[1] This work was supported by the National Science Foundation under grant ECS-8504752

where

$$R_{ik} = \int_0^T s_i(t)s_k(t)dt \tag{1.4}$$

is the crosscorrelation between signals i and k and the received signal is $r(t) = s(t) + n(t)$. In vector notation the matched filter output vector $\mathbf{y}$ is $\mathbf{y}=\mathbf{Rx}+\mathbf{n}$. The assumption of linear independence of the signal set guarantees positive definiteness of the symmetric crosscorrelation matrix $\mathbf{R}$ (i.e. for all $\mathbf{x}\neq 0$ $\mathbf{x}^T\mathbf{Rx} > 0$). Let $\mathbf{r}_i$ be the columns of $\mathbf{R}$, and let $\mathbf{B}$ denote its inverse (symmetric and positive definite), with columns $\mathbf{b}_i$, $i=1,\ldots,K$. If w_i denotes the energy of the signal s_i, $w_i=\int_0^T s_i^2(t)dt$, the elements of the crosscorrelation matrix $\mathbf{R}$ can be expressed as $R_{ik}=\sqrt{w_i}\sqrt{w_k}\tilde{R}_{ik}$ where $\tilde{R}_{ik}$ is the correlation coefficient of signals i and k.

2. The Conventional Single-User Detector

In practice single-user detection is applied to a multiuser environment. A threshold decision is performed on each matched filter output y_k, i.e.

$$\hat{x}_k = sgn(\,(\mathbf{Rx})_k + n_k\,) \tag{2.1}$$

$y_k>0$ and for $x_k= -1$ else. This approach is not optimal due to the interference term $\sum_{i\neq k} x_i R_{ik}$.

Proposition 1: The k^{th} user asymptotic efficiency of the conventional detector is

$$\eta_k^c = \max{}^2\left\{0,\, 1 - \sum_{j\neq k}\frac{\sqrt{w_j}}{\sqrt{w_k}}\,\left|\tilde{R}_{kj}\right|\right\} \tag{2.2}$$

Proof: The k^{th} user probability of error is given by:

$$\begin{aligned}P_k &= P(\hat{x}_k= 1\,|\,x_k= -1) = P(\,(Rx)_k + n_k > 0\,|\,x_k= -1)\\ &= P(n_k > R_{kk} - \sum_{j\neq k} R_{kj}\,x_j)\\ &= 2^{1-K}\sum_{\substack{\mathbf{x}\in\{-1,1\}^K\\ x_k=-1}} P(n_k > R_{kk} - \sum_{j\neq k}R_{kj}\,x_j)\end{aligned} \tag{2.3}$$

Since the random variable η_k is Gaussian with zero mean and variance equal to $w_k\sigma^2$, the sum in (2.2) is dominated as $\sigma \to 0$ by the term

$$2^{1-K}Q(\,(R_{kk} - \sum_{j\neq k}|R_{kj}\,|)/\sigma\sqrt{w_k}\,)$$

Hence the k^{th} user asymptotic efficiency of the single-user receiver is equal to zero if $\mathbf{R}_{kk} \leq \sum_{j\neq k}|\mathbf{R}_{kj}\,|$. Otherwise it is equal to the square of the ratio of the argument of the foregoing Q-function and the argument corresponding to the single-user probability of error $Q(\sqrt{w_k}/\sigma)$, i.e.

$$\begin{aligned}\eta_k^c &= (R_{kk} - \sum_{j\neq k}|R_{ij}\,|)^2/w_k^2\\ &= (\,1 - \sum_{j\neq k}\sqrt{w_j/w_k}\,|\tilde{R}_{kj}\,|)^2\end{aligned} \tag{2.4}$$

●

Note that the k^{th} user asymptotic efficiency of the conventional detector is equal to zero for sufficiently high signal-to-noise ratio of any user not orthogonal to the k^{th} user. In particular in the case of two active users $\sqrt{w_j/w_k} \geq 1/\,|\,\tilde{R}_{12}\,|$ results in $\eta_k^c = 0$.

3. The Decorrelating Detector

In the absence of noise the matched filter output is $\mathbf{y} = \mathbf{Rx}$, so the optimal strategy to follow is to premultiply $\mathbf{y}$ by the inverse crosscorrelation matrix $\mathbf{B}$. Then $\hat{\mathbf{x}} = \mathbf{By} = \mathbf{x}$. The detector $\hat{\mathbf{x}} = sgn$ ($\mathbf{By}$) will be called a *decorrelator*, and Proposition 2 quantifies its performance in the presence of noise. In [6] is was erroneously shown (cf.[3]) that this detector is optimum in terms of bit-error rate.

Proposition 2 : The k^{th} user asymptotic efficiency of the decorrelating detector is given by

$$\eta_k^d = 1/(\tilde{\mathbf{R}}^{-1})_{kk} \quad , \tag{3.1}$$

Thus the k^{th} user asymptotic efficiency of the decorrelating detector *is independent of the energy of other users* .

Proof : We have

$$\hat{x}_k = sgn\ (\,\mathbf{By}\,)_k = sgn\ (\,x_k + (\mathbf{Bn})_k) \tag{3.2}$$

The k^{th} user error probability is, by symmetry (for equally likely transmitted symbols)

$$P_k = P\ (\,\hat{x}_k = 1\,|\,x_k= -1) = P(\,(\mathbf{Bn})_k > 1\,)$$

Again, $(\mathbf{Bn})_k$ is Gaussian, with variance equal to the k^{th} diagonal element of

$$E\ [\,(\mathbf{Bn}\,)(\mathbf{Bn}\,)^T\,] = \mathbf{B}\,E(\mathbf{nn}^T)\,\mathbf{B}^T = \mathbf{B}\sigma^2\mathbf{RB}^T = \sigma^2\mathbf{B}$$

Thus

$$P_k = P(\,(\mathbf{Bn})_k > 1) = Q(1/\sigma\sqrt{B_{kk}}\,) = Q(\,\sqrt{\eta_k^d w_k}/\sigma\,)$$

by definition of η_k^d . Therefore, by monotonicity of Q(.),

$$\eta_k^d = \min\ \{\,1,\ 1/w_k B_{kk}\ \}$$

Next we show that $B_{kk}w_k \geq 1$. Since $w_k = R_{kk}$ we prove that $R_{kk}B_{kk} \geq 1$ for all pairs of symmetric positive definite matrices $\mathbf{R}$ and $\mathbf{B}$ inverse to each other. Since $\mathbf{R}$ is symmetric it can be represented as $\mathbf{R} = \mathbf{T}\Lambda\mathbf{T}^T$ with Λ the diagonal matrix of eigenvalues of $\mathbf{R}$ and $\mathbf{TT}^T = \mathbf{I}$ $=>$ $\sum_{i=1}^{K} t_{ki}\,t_{ij} = \delta_{jk}$. Let λ_i be the i^{th} eigenvalue of $\mathbf{R}$. Since $\mathbf{R}$ is positive definite its eigenvalues are strictly positive. Then, $\mathbf{B} = \mathbf{T}\Lambda^{-1}\mathbf{T}^T$ and, with $\mathbf{t}_k^T$ the k^{th} row of $\mathbf{T}$,

$$\begin{aligned}R_{kk}B_{kk} &= (\mathbf{t}_k^T\Lambda\mathbf{t}_k)(\mathbf{t}_k^T\Lambda^{-1}\mathbf{t}_k) = (\sum_{i=1}^{K}t_{ik}^2\lambda_i\,)(\sum_{i=1}^{K}t_{ik}^2\frac{1}{\lambda_i}\,)\\ &= \sum_{i=1}^{K}\sum_{j=1}^{K}t_{ik}^2t_{jk}^2\lambda_i\frac{1}{\lambda_j} = \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K}t_{ik}^2t_{jk}^2\left(\frac{\lambda_i}{\lambda_j}+\frac{\lambda_j}{\lambda_i}\right)\\ &\geq \frac{1}{2}\,2\sum_{i=1}^{K}\sum_{j=1}^{K}t_{ik}^2t_{jk}^2 = (\sum_{i=1}^{K}t_{ik}^2)^2 = 1\end{aligned}$$

We used the fact that x + 1/x $\geq$ 2 for x > 0 and the property of $\mathbf{TT}^T = \mathbf{I}$ given above. We now have that $\eta_k^d = 1/w_kB_{kk}$. Let $C_{ij}(\mathbf{M})$ be the cofactor of element ij of the matrix $\mathbf{M}$. Then

$$B_{kk} = \frac{C_{kk}(\mathbf{R})}{\det\mathbf{R}} = \frac{w_1\cdots w_{k-1}w_{k+1}\cdots w_K\,C_{kk}(\tilde{\mathbf{R}})}{\det\tilde{\mathbf{R}}\,\prod_1^K w_i} = \frac{1}{w_k}\,(\tilde{\mathbf{R}}^{-1})_{kk}$$

●

4. The Optimum Multiuser Detector

We now turn our attention to the optimal multiuser detector under the maximum likelihood criterion. This is the detector which decides for the transmitted vector $\mathbf{x}$ which is most likely to have produced the received signal $\mathbf{y} = \mathbf{Rx}+\mathbf{n}$. We seek

$$\hat{\mathbf{x}} = arg\max_{\mathbf{x}\in\{-1,1\}^K}P(\mathbf{y} = \mathbf{Rx} + \mathbf{n}) \tag{4.1}$$

Taking into account the Gaussian noise statistics, a few equivalence transformations of the above yield:

$$\hat{\mathbf{x}} = arg\min_{\mathbf{x}\in\{-1,1\}^K}(\mathbf{x}^T\mathbf{Rx} - 2\mathbf{x}^T\mathbf{y}) \tag{4.2}$$

It has been shown recently [2] that the combinatorial minimization problem in (4.2) is NP-complete. This is a key result which provides the impetus to seek suboptimal detection schemes which offer both acceptable perfomance and computational ease. As a performance measure for suboptimal detectors we would like to have a closed form expression for the optimum k^{th} user asymptotic efficiency, which is an upper bound on the asymptotic performance. Unfortunately a recent analysis, [2], has shown that the computation of the optimum k^{th} user asymptotic efficiency is also an NP-complete combinatorial problem. Because of its interest we will state this basic result obtained in [2]. The optimal k^{th} user asymptotic efficiency for antipodal modulation can be expressed as follows:

$$\eta_k^* = \frac{1}{R_{kk}} \min_{\substack{b^1, b^2 \\ b_k^1 \neq b_k^2}} \int_0^T \left[\sum_{i=1}^{K} \frac{1}{2} (b_i^1 - b_i^2) s_i(t) \right]^2 dt = \frac{1}{R_{kk}} \min_{\substack{\epsilon \in \{-1,0,1\}^K \\ \epsilon_k \neq 0}} \epsilon^T \mathbf{R} \epsilon \tag{4.3}$$

Proposition 3: The following problem is NP-hard.

Given $K \in \boldsymbol{Z}$, $k \in \{1, \cdots, K\}$, and a positive definite matrix $\mathbf{R} \in Z^{K \times K}$,

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

However, even though general closed-form expressions for the K-user case do not exist, an explicit formula for the 2-user asymptotic efficiency in the general asynchronous case is given in [2]. Its particularization to the synchronous case yields the following expression.

Proposition 4 : The 2-user asymptotic efficiency of the maximum likelihood detector is given by

$$\eta_k^* = 1 - \max \left\{ 0,\ 2 \frac{\sqrt{w_i}}{\sqrt{w_k}} \mid \tilde{R}_{12} \mid - \frac{w_i}{w_k} \right\} \ , \quad (i,k) \in \{\ (1,2),(2,1)\ \} \tag{4.6}$$

While the asymptotic efficiency of the decorrelating detector has been shown to be independent of the energy of the interfering users, this is not the case for the optimum multi-user detector. In environments where the signal energies vary dynamically over a broad range (e.g. near-far problem), it is important to assess the worst case asymptotic efficiency with respect to the energies of the interfering users.

Proposition 5: The worst-case asymptotic efficiency of the maximum likelihood detector, taken over all possible energies of the interfering users, equals the asymptotic efficiency of the decorrelating detector:

$$\min_{\sqrt{w_i} > 0} \eta_k^* = \eta_k^d$$

Proof: Using expression (4.3) for the asymptotic efficiency of the optimum multiuser detector, we have:

$$\min_{w_i > 0} \eta_k^* = \min_{w_i > 0} \min_{\substack{\epsilon_i \in \{-1,0,1\} \\ \epsilon_k \neq 0}} \frac{1}{w_k} \epsilon^T \mathbf{R} \epsilon$$

$$= \min_{w_i > 0} \min_{\substack{\epsilon_i \in \{-1,0,1\} \\ \epsilon_k \neq 0}} \frac{1}{w_k} [\epsilon_1 \sqrt{w_1} \cdots \epsilon_K \sqrt{w_K}]\ \tilde{\mathbf{R}}\ [\epsilon_1 \sqrt{w_1} \cdots \epsilon_K \sqrt{w_K}]^T$$

$$= \min_{\substack{z_i \in \boldsymbol{R} \\ z_k \neq 0}} \frac{1}{w_k} \mathbf{z}^T \tilde{\mathbf{R}} \mathbf{z} \tag{4.7}$$

with the obvious substitution $z_i = \epsilon_i \sqrt{w_i}$. $\tilde{\mathbf{R}}$ is, as previously defined, the normalized crosscorrelation matrix. Letting $\mathbf{M}$ be the matrix obtained from $\tilde{\mathbf{R}}$ by removal of the k^{th} row and column, $\mathbf{a}$ be the k^{th} column of $\tilde{\mathbf{R}}$ without the k^{th} element and $\tilde{\mathbf{z}}$ be the vector $\mathbf{z}$ without the k^{th} element, equation (4.7) can be expanded as:

$$\min_{w_i > 0} \eta_k^* = \min_{\substack{\tilde{\mathbf{z}} \in \boldsymbol{R}^{K-1} \\ z_k \neq 0}} \frac{1}{w_k} (z_k^2 + 2 z_k \tilde{\mathbf{z}}^T \mathbf{a} + \tilde{\mathbf{z}}^T \mathbf{M} \tilde{\mathbf{z}})$$

This is a straightforward continuous minimization problem with solution:

$$\eta_{k\ \min} = 1 - \mathbf{a}^T \mathbf{M}^{-1} \mathbf{a} \tag{4.8}$$

for

$$\tilde{\mathbf{z}} = -z_k\ \mathbf{M}^{-1} \mathbf{a} \tag{4.9}$$

We have used the equality $z_k^2 = w_k$ and the fact that the matrix $\mathbf{M}$, being identical to the matrix $\tilde{\mathbf{R}}$ for users {1,..k-1,k+1,..K} is positive definite, hence invertible. Remains to show that the expression obtained in (4.8) equals η_k^d.

$$\eta_k^d = \frac{1}{\tilde{\mathbf{R}}_{kk}^{-1}} = \frac{\det \tilde{\mathbf{R}}}{(adj\, \tilde{\mathbf{R}})_{kk}} = \frac{\det \tilde{\mathbf{R}}}{cofactor(\tilde{\mathbf{R}}_{kk})} = \frac{\det \tilde{\mathbf{R}}}{\det \mathbf{M}}$$

After $2(K-k-1)$ row and column flips the numerator can be computed as:

$$\det \tilde{\mathbf{R}} = \det \begin{bmatrix} \mathbf{M} & \mathbf{a} \\ \mathbf{a}^T & 1 \end{bmatrix} = \det \mathbf{M}\ (1 - \mathbf{a}^T \mathbf{M}^{-1} \mathbf{a})$$

where the last equality follows from the expression for the determinant of a block matrix (eg.[7],pg 650).

•

There is a nice geometric explanation for the equality of $\eta_{k\ \min}^*$ and η_k^d. In fact, by computing the appropriate metrics in the geometric setting given below (easy in the two-user case), all the previous formulas can be obtained. For the sake of clarity let us consider the two-user case. Recall that the received signal $\mathbf{y}$ satisfies: $\mathbf{y} = \mathbf{R}\mathbf{x} + \mathbf{n}$ and that the noise autocovariance matrix is $\mathbf{R}$. It is convenient to work in the $\mathbf{R}^{-1/2}\mathbf{y}$ domain - call it domain S -, where the hypotheses are at the points $\mathbf{R}^{1/2}\mathbf{x}$, with $\mathbf{x} \in \{-1,1\}^2$, and where the noise is spherically symmetric. Therefore the decision regions of the maximum likelihood detector, determined by the minimum Euclidean distance rule, are given by the perpendicular bisectors of the segments between the different hypotheses. *Recall that the k^{th} user asymptotic efficiency corresponds to the square of half the minimum distance between distinct hypotheses differing in the k^{th} bit* ($Q(d_{\min}/2\sigma) = Q(\sqrt{\eta^*}/\sigma)$).

The decision regions of the decorrelating detector are cones with a vertex at the origin, such that application of $\mathbf{R}^{-1}$ maps them to the coordinate axes. Thus in domain S the decision cones pass through the points $\mathbf{R}^{1/2}\mathbf{e}$, with $\mathbf{e}$ the unit vectors in $\boldsymbol{R}^2$. It is interesting to observe that these points are at the center of the sides of the parallelogram formed by the hypotheses, because the unit vectors can be represented as half the sum of adjacent hypotheses. So, the decorrelating detector decision boundaries are parallel to the parallelogram sides and intersect it at the centers of its sides. As a consequence of this property a geometric interpretation of the k^{th} user asymptotic efficiency of the decorrelating detector can be found as follows: recall that in order to find the k^{th} user asymptotic efficiency of a detector, we have to approximate its probability of k^{th} bit error by a Q-function, in the low noise region.The k^{th} bit error probability -by symmetry we can assume that the transmitted bit was 1- is the sum of two integrals -one for each possibility for the remaining bit- of the noise density function over the region in which the k^{th} bit is decoded as -1. In our case the k^{th} bit error probability can be easily computed by taking advantage of the properties presented above. To this end we rotate the coordinate system to let the y-axis coincide with the k^{th}-bit decision boundary and use the equal distance property of the decision boundary to the hypotheses, to observe that the two integrals are equal, and then use the spherical symmetry of the noise to identify each integral as Q-function of the distance of the hypothesis to the decision boundary. *Hence the k^{th} user asymptotic efficiency of the decorrelating detector is equal to the square of the distance of any hypothesis to the k^{th} bit decision boundary.*

Thus, in figure 1, $\sqrt{\eta_1^*}$ is the length of the shortest of the segments AM,AO and DO, and $\sqrt{\eta_1^d}$ is the length of AP. The result of proposition 5 can now be interpreted as follows: since η^* appears as the hypothenuse and η^d as the kathete of a right angled triangle, η^* is lower bounded by the energy independent η^d. However, since the triangle angles vary with increasing energy of the interfering user, there is a particular energy ratio for which the triangle degenerates into a line segment. This is the point when η^* reaches its minimum, which is geometrically identical with η^d. For the parallelogram formed by the hypotheses in domain S this is the case when a diagonal is perpendicular on a side (eg. AP perpendicular on CD).

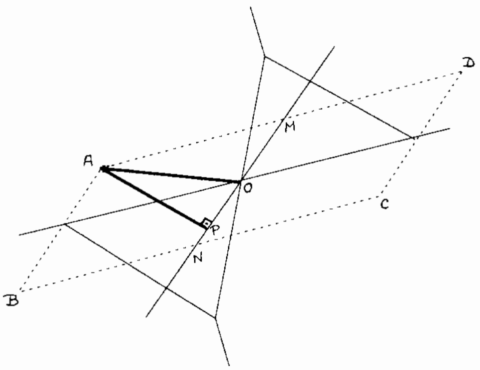

Fig. 1. Hypotheses and decision regions in domain S

The result of Proposition 5 is of special importance in a near-far environment, where the received signals have different energies, and where the energy ratios may vary continuously over a broad scale, if the positions of the users evolve dynamically. In this environment, if worst-case performance is considered, the decorrelating detector, with its linear time-complexity per bit, offers the same performance as the NP-complete optimum multiuser detector.

5. The Optimum Linear Multiuser Detector

We now study the improvement obtained by a linear transformation on the matched filter outputs prior to sign detection and seek the optimal linear detector. This has been seen to be an important problem, since the optimal detector is exponential in the number of users, and the performance of the computationally efficient single-user detector degrades to zero with growing energy of other users. First we give the form of the general linear detector, second we find the linear detector which maximizes the asymptotic efficiency (or equivalently minimizes the probability of bit error) and third we compare the achieved asymptotic efficiency to the ones reached by the conventional and optimal detectors. Thus we ask what mapping $\mathbf{T}$: $\boldsymbol{R}^K \rightarrow \boldsymbol{R}^K$ will minimize the probability of bit error for the decision scheme

$$\hat{\mathbf{x}} = \mathrm{sgn}\ (\ \mathbf{Ty}\ ) = \mathrm{sgn}\ (\mathbf{TRx} + \mathbf{Tn}\ ) \tag{5.1}$$

In terms of decision regions the problem interprets as follows : what is the optimal way to partition the K-dimensional hypotheses space into K decision-cones with vertices at the origin? The surfaces of these cones determine the columns of the inverse $\mathbf{T}^{-1}$ of the sought mapping. Application of $\mathbf{T}$ on the cone configuration will map the cones on quadrants, after which a sign detector is used.
The bit error probability of the k^{th} user under this detection scheme is, (same derivation as for the conventional sign detector) :

$$\begin{aligned} P_k &= P(\hat{x}_k = 1\,|\,x_k = -1) = P(\ (\mathbf{TRx} + \mathbf{Tn})_k\ > 0\,|\,x_k = -1) \\ &= P(\ (\mathbf{Tn})_k\ > (\mathbf{TR})_{kk} - \sum_{j \neq k}(\mathbf{TR})_{kj}\,x_j) \\ &= 2^{1-K} \sum_{\substack{\mathbf{x} \in \{-1,1\}^K \\ x_k = -1}} P(\ (\mathbf{Tn})_k\ > (\mathbf{TR})_{kk} - \sum_{j \neq k}(\mathbf{TR})_{kj}\,x_j) \end{aligned} \tag{5.2}$$

Since the random variable $(\mathbf{Tn})_k$ is Gaussian with zero mean and variance equal to $(\mathbf{TRT}^T)_{kk}\sigma^2$, the sum in (5.2) is dominated as $\sigma \rightarrow 0$ by the term

$$2^{1-K}\, Q(\ ((\mathbf{TR})_{kk} - \sum_{j \neq k} |(\mathbf{TR})_{kj}\ |)/\sigma\sqrt{(\ \mathbf{TRT}^T\ )_{kk}}\ \ ) \tag{5.3}$$

The asymptotic efficiency of the best linear detector is bounded away from zero by that of the decorelating detector (which has $\mathbf{T} = \mathbf{I}$). Hence it is equal to the square of the ratio of the argument of the foregoing Q-function and the argument corresponding to the single-user probability of error $Q(\sqrt{w_k}/\sigma)$, i.e.

$$\eta_k^l = \frac{1}{w_k}\,(\mathbf{r}_k^T\mathbf{v} - \sum_{j \neq k}\ |\ \mathbf{r}_j^T\mathbf{v}\ \ |)^2\,/\mathbf{v}^T\mathbf{R}\mathbf{v} \tag{5.4}$$

In order to minimize P_k, we have to maximize the argument of the Q- function, and equivalently maximize the asymptotic efficiency η_k^l, with respect to the components of the vector $\mathbf{v}$. Note that η_k^l is invariant with respect to scaling of $\mathbf{v}$.

First we consider the two user case for which explicit expressions of the maximum linear asymptotic efficiency are obtainable.

5.1. The Two-User Case

Proposition 6 : The k^{th} user optimal linear transformation $\mathbf{T}_k(y) = \mathbf{v}^T\mathbf{y}$ on the matched filter outputs prior to threshold detection is :

$$\tilde{\mathbf{v}}^T = [\ 1\ ; -\mathrm{sgn}\tilde{R}_{12}\ \ \min\left(1, |\tilde{R}_{12}\,|(w_k/w_i)^{1/2}\ \right)],\ (i,k) \in \{(1,2),(2,1)\} \tag{5.5}$$

Equivalently,

$$\tilde{\mathbf{v}}^T = \begin{cases} [1 \quad -\mathrm{sgn}\tilde{R}_{12}] & ,\ (w_i/w_k)^{1/2} \leq |\tilde{R}_{12}\,| \\ \qquad \mathbf{b}_k^T & ,\ \text{otherwise} \qquad\qquad \text{(decorrelating detector)} \end{cases} \tag{5.6}$$

Proof : Without loss of generality, we will consider user 1. We have

$$\mathbf{R} = \begin{bmatrix} w_1 & \sqrt{w_1 w_2}\tilde{R}_{12} \\ \sqrt{w_1 w_2}\tilde{R}_{12} & w_2 \end{bmatrix} \quad , \qquad \mathbf{v}^T\ =\ [1\ \ ;\ v_2] \tag{5.7}$$

$$\begin{aligned} \sqrt{\eta_1^l} &= \frac{1}{\sqrt{w_1}}\,\frac{\mathbf{r}_1^T\mathbf{v}\ -\ |\ \mathbf{r}_2^T\mathbf{v}\ \ |}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}} \\ &= \frac{1 + \tilde{R}_{12}(w_2/w_1)^{1/2}\ v_2\ -\ |\ \tilde{R}_{12}(w_2/w_1)^{1/2} + (w_2/w_1)v_2\ \ |}{\sqrt{1 + 2\tilde{R}_{12}(w_2/w_1)^{1/2}v_2 + (w_2/w_1)v_2^2}} \end{aligned} \tag{5.8}$$

We want to maximize $\sqrt{\eta_1^l}$ with respect to v_2. Introduce an indicator function for the absolute value term :

$$I = \begin{cases} 1 & ,\ \tilde{R}_{12} + (w_2/w_1)^{1/2}v_2\ \ > 0 \\ -1 & ,\ \tilde{R}_{12} + (w_2/w_1)^{1/2}v_2\ \ < 0 \\ 0 & ,\ \text{else} \end{cases} \tag{5.9}$$

Then

$$\frac{d\sqrt{\eta_1}}{dv_2} = -\ \frac{(\ 1 - \tilde{R}_{12}^2\ \ )\,(w_2/w_1)}{(\ 1 + 2\tilde{R}_{12}\,(w_2/w_1)^{1/2}\ v_2 + (w_2/w_1)\ v_2^2\ \ )^{3/2}}(\ I\ +\ v_2) \tag{5.10}$$

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

$$v_2 = 1 \quad \text{if } I\ = -1\ \ \Longleftrightarrow\ \ (w_2/w_1)^{1/2} < -\ \tilde{R}_{12}\ \ \text{(possible if } \tilde{R}_{12} < 0\ )$$

$$v_2 = -1 \quad \text{if } I\ = 1\ \ \Longleftrightarrow\ \ (w_2/w_1)^{1/2} <\ \tilde{R}_{12}\ \ \ \text{(possible if } \tilde{R}_{12} > 0\ ) \tag{5.11}$$

As can easily be seen, both values correspond to maxima. If neither of these conditions is met, the derivative does not have a zero. The optimal value for v_2 can be determined from a closer look at the behavior of $d\sqrt{\eta_1}/dv_2$, in figure 2.
For both $I = 1$ and $I = -1$, the derivative of $\sqrt{\eta_1}$ is positive for v_2 smaller than the abscissa of the zero of the derivative (-1 respectively 1), and negative afterwards. Due to the nonlinearity of $\sqrt{\eta_1}$ the derivative has the form corresponding to $I = -1$ for $v_2 < -\ \tilde{R}_{12}\,(w_2/w_1)^{1/2}$ and the form corresponding to $I = 1$ afterwards. Since the second branch (for $I = 1$) turns negative before the first one, we have to take the largest value of v_2 yielding a positive derivative on the first branch. It can easily be seen that in the "no-zero" case, $-1 < -\tilde{R}_{12}(w_2/w_1)^{1/2} < 1$, this is the point of discontinuity, i.e. $v_2 = -\tilde{R}_{12}(w_2/w_1)^{1/2}$. •

Note that for $\tilde{R}_{12} = 0$ we get $\tilde{\mathbf{v}}^T = [1 \quad 0]$, the identity transformation, as expected, since then the users are decoupled and a single-user detector is optimal. By taking the inverse of $\mathbf{R}$ we also see that in the "no-zero" case the optimal transformation vector is exactly the corresponding row of the inverse correlation matrix.

Proposition 7 : The k^{th} user asymptotic efficiency of the optimal linear two-user detector equals :

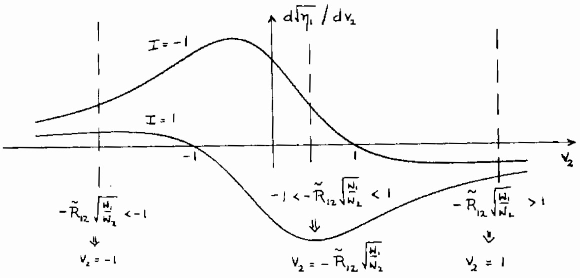

Fig.2

$$\eta_k^l = \begin{cases} 1 - 2\,|\tilde{R}_{12}|(w_i/w_k)^{1/2} + w_i/w_k \ , & \text{if } (w_i/w_k)^{1/2} \le |\tilde{R}_{12}| \\ 1 - \tilde{R}_{12}^2 \ , & \text{otherwise} \end{cases} \tag{5.12}$$

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

Proof : Substitution of the optimal linear transfomation found in Proposition 6 in the expression for asymptotic efficiency. •

The asymptotic efficiency obtained in the range $(w_2/w_1)^{1/2} < |\tilde{R}_{12}|$ equals the optimum attainable asymptotic efficiency, that of the maximum likelihood detector. Even outside the region of optimality, the best linear detector shows a far better performance than the conventional single user detector, since η_k^l is independent of the energy of other users (Proposition 3), whereas η_k^c goes to zero (Corollary to Proposition 1). There is an intuitive interpretation of the discontinuity of the best linear detector and of the boundary point $(w_2/w_1)^{1/2} = |\tilde{R}_{12}|$. The input to the threshold device corresponding to the first user, $z_1 = \mathbf{v}^T\mathbf{R}\mathbf{x}$, has three components:

$$\begin{aligned} z_1 = w_1\ &[(1 - \tilde{R}_{12}^2) + \tilde{R}_{12}(\tilde{R}_{12} + v_2(w_2/w_1)^{1/2})]\ x_1 \\ &+ w_1[(w_2/w_1)^{1/2}(\tilde{R}_{12} + v_2(w_2/w_1)^{1/2})]\ x_2 \\ &+ \tilde{n} \end{aligned} \tag{5.15}$$

where $\tilde{n}$ is Gaussian noise of power spectral density $w_1\sigma^2\ [(1 - \tilde{R}_{12}^2) + (\tilde{R}_{12} + v_2(w_2/w_1)^{1/2})^2]$. For $(w_2/w_1)^{1/2} > |\tilde{R}_{12}|$, the second term outweighs the second part of the first term, so the best one can do is to eliminate it, by choosing $v_2 = -\tilde{R}_{12}(w_2/w_1)^{1/2}$ (the decorrelating detector). Since this minimizes the noise variance at the same time, it is the best strategy in this region. If, however, $(w_2/w_1)^{1/2} < |\tilde{R}_{12}|$, and if additionally the term $\tilde{R}_{12}(\tilde{R}_{12} + v_2(w_2/w_1)^{1/2})$ is positive, it is a better policy to allow interference from user 2, which is compensated by the second part in the first term, and use the residual positive contribution in the first term to increase the SNR compared to the decorrelating case. We have seen that this strategy leads to the same performance as the more complicated maximum likelihood detector.

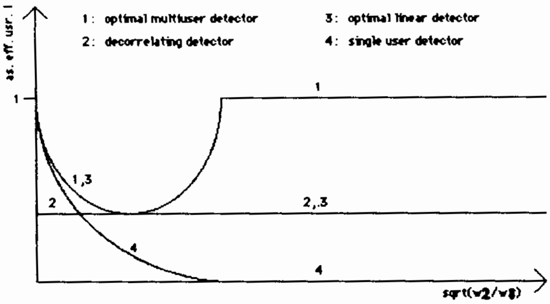

Fig.3 Asymptotic efficiencies in the two-user case

Note that in the two user case the signal energies and crosscorelations can't be picked such as to allow *both* users optimal performance at the same time :

for user 1 need $(w_2/w_1)^{1/2} < |\tilde{R}_{12}| < 1$

for user 2 need $(w_1/w_2)^{1/2} < |\tilde{R}_{12}| \ => \ (w_2/w_1)^{1/2} > \dfrac{1}{|\tilde{R}_{12}|} > 1$

The two cases are mutually exclusive. Nevertheless this is not a substantial shortcoming, since for small values of $\tilde{R}_{12}$ the decorrelating detector has an asymptotic efficiency close to unity.

5.2. The K-User Case

Proposition 8 : The k^{th} user asymptotic efficiency of the best linear detector equals:

$$\eta_k^l = \max_{\substack{e_j \in \{-1,1\} \\ j \ne k}} \frac{1}{w_k}\eta^2(\mathbf{e}) \qquad \text{with} \qquad \eta(\mathbf{e}) = \max_{\substack{\mathbf{v} \in \mathbb{R}^K \\ \mathbf{v}^T\mathbf{R}\mathbf{v} = 1 \\ e_j r_j^T \mathbf{v} \ge 0 \\ j \ne k}} \mathbf{v}_o^T\mathbf{R}\mathbf{v} \tag{5.16}$$

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

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

$$(1) \qquad \tilde{\mathbf{v}} = \frac{\mathbf{v}_o + \sum_{j \ne k} \lambda_j e_j \mathbf{u}_j}{(\mathbf{v}_o^T\mathbf{R}\mathbf{v}_o + \mathbf{v}_o^T\mathbf{R} \sum_{j \ne k} \lambda_j e_j \mathbf{u}_j)^{1/2}} \ , \quad (\mathbf{u}_j)_i = \begin{cases} 0 & , i \ne j \\ 1 & , i = j \end{cases} \tag{5.17}$$

$$(2) \qquad e_j r_j^T \tilde{\mathbf{v}} \ge 0 \quad \text{for} \quad j \ne k$$

$$(3) \qquad \mathbf{r}_j^T \tilde{\mathbf{v}} \ne 0 \ => \ \lambda_j = 0$$

$$(4) \qquad \lambda_j \ge 0 \quad j \ne k$$

Proof : Let

$$\begin{aligned} S_j^+ &= \{\mathbf{x} \in \mathbb{R}^K \ : \ \mathbf{r}_j^T\mathbf{x} \ge 0\ \} \\ S_j^- &= \{\mathbf{x} \in \mathbb{R}^K \ : \ \mathbf{r}_j^T\mathbf{x} \le 0\ \} \end{aligned} \tag{5.18}$$

Then, since $\eta_k^l = \dfrac{1}{w_k}\left((\mathbf{r}_k^T\mathbf{v} - \sum_{j \ne k} |\mathbf{r}_j^T\mathbf{v}\ |)/\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}\right)^2$ we seek

$$\begin{aligned} &\max_{\mathbf{v} \in \mathbb{R}^K} \frac{1}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}}(\mathbf{r}_k^T\mathbf{v} - \sum_{j \ne k} |\mathbf{r}_j^T\mathbf{v}\ |) \\ &= \max_{\substack{e_j \in \{-1,1\} \\ j \ne k}} \ \max_{\mathbf{v} \in \cap_j S_j^{e_j}} \frac{1}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}}(\mathbf{r}_k^T\mathbf{v} - \sum_{j \ne k} |\mathbf{r}_j^T\mathbf{v}\ |) \end{aligned} \tag{5.19}$$

$$= \max_{\substack{e_j \in \{-1,1\} \\ j \ne k}} \eta(\mathbf{e}) \ , \quad \eta(\mathbf{e}) = \max_{\mathbf{v} \in \cap_j S_j^{e_j}} \frac{1}{\sqrt{\mathbf{v}^T\mathbf{R}\mathbf{v}}}(\mathbf{r}_k^T\mathbf{v} - \sum_{j \ne k} |\mathbf{r}_j^T\mathbf{v}\ |) \tag{5.20}$$

From the defintion of $\mathbf{v}_o$ we see that the term in parentheses equals $\mathbf{v}_o^T\mathbf{R}\mathbf{v}$. Now $\mathbf{v} \in \cap_j S_j^{e_j} \ <=> \ e_j \mathbf{r}_j^T\mathbf{v} \ge 1$, $j \ne k$ and recall that η_k^l is invariant to scaling of $\mathbf{v}$, so that maximization of the given functional over $\mathbb{R}^K$ is equivalent to minimization over the ellipsoid $\mathbf{v}^T\mathbf{R}\mathbf{v} = 1$.

This proves the first part of Proposition 9. We now have a sequence of two maximizations to perform, where the second one has the explicit form of an exhaustive search. We turn our attention to the inner maximization.

$$\text{Claim: } \eta(\mathbf{e}) = \max_{\substack{\mathbf{v} \in \mathbb{R}^K \\ \mathbf{v}^T\mathbf{R}\mathbf{v} = 1 \\ e_j \mathbf{r}_j^T\mathbf{v} \ge 0 \\ j \ne k}} \mathbf{v}_o^T\mathbf{R}\mathbf{v} = \max_{\substack{\mathbf{v} \in \mathbb{R}^K \\ \mathbf{v}^T\mathbf{R}\mathbf{v} \le 1 \\ e_j \mathbf{r}_j^T\mathbf{v} \ge 0 \\ j \ne k}} \mathbf{v}_o^T\mathbf{R}\mathbf{v} \tag{5.21}$$

Proof : We will obtain the above result by considering an equivalence transformation on $\eta(\mathbf{e})$. Let $\mathbf{y} = \mathbf{R}^{1/2}\mathbf{v}$, $\mathbf{z}_j^T = j^{th}$ row of $\mathbf{R}^{1/2}$

$$\begin{aligned} \mathbf{r}_j^T\mathbf{v} &= \mathbf{r}_j^T\mathbf{B}^{1/2}\mathbf{y} = \mathbf{z}_j^T\mathbf{y} \\ \mathbf{v}_o^T\mathbf{R}^{1/2} &= \mathbf{y}_o^T \\ \mathbf{v}^T\mathbf{R}\mathbf{v} &= \mathbf{y}^T\mathbf{y} = |\mathbf{y}|^2 \end{aligned}$$

$$\eta(\mathbf{e}) = \max_{\substack{\mathbf{y} \in \mathbb{R}^K \\ |\mathbf{y}| = 1 \\ e_j \mathbf{z}_j^T\mathbf{y} \ge 0 \\ j \ne k}} \mathbf{y}_o^T\mathbf{y} = \max_{\substack{\mathbf{y} \in \mathbb{R}^K \\ |\mathbf{y}| = 1 \\ e_j \mathbf{z}_j^T\mathbf{y} \ge 0 \\ j \ne k}} |\mathbf{y}_o|\,|\mathbf{y}|\cos\alpha \tag{5.22}$$

where α is the angle between the vectors $\mathbf{y}_o$ and $\mathbf{y}$. We know that there is at least one feasible $\mathbf{v}$ such that $\mathbf{v}_o^T\mathbf{R}\mathbf{v}$ is strictly positive: With $\mathbf{v} = \mathbf{B}_k/\sqrt{B_{kk}}$ we have already shown (decorrelating detector) that $\eta(\mathbf{e}) = 1/B_{kk} > 0$. This $\mathbf{v}$ is feasible, since $\mathbf{r}_j^T\mathbf{v} = 0$, for all $j \ne k$, and $\mathbf{v}$ is scaled such that $\mathbf{v}^T\mathbf{R}\mathbf{v} = 1$. Therefore the maximum over all feasible vectors is strictly positive. So the optimal $\cos\alpha$ is positive. Since the inequality constraints are linear and partition the space into convex cones with vertex at the origin, the optimal angle α is independent of $|\mathbf{y}|$. Thus the maximum of $\eta(\mathbf{e})$ over the convex set $|\mathbf{y}| \le 1$ is attained for $|\mathbf{y}|$ maximal, i.e. $|\mathbf{y}| = 1$, which completes the proof of the claim.

We now have to consider the following problem:

$$\eta(\mathbf{e}) = \max_{\substack{\mathbf{v} \in \mathbb{R}^K \\ \mathbf{v}^T\mathbf{R}\mathbf{v} \le 1 \\ e_j \mathbf{r}_j^T\mathbf{v} \ge 0 \\ j \ne k}} \mathbf{v}_o^T\mathbf{R}\mathbf{v} = \min_{\substack{\mathbf{v} \in \mathbb{R}^K \\ \mathbf{v}^T\mathbf{R}\mathbf{v} - 1 \le 0 \\ -e_j \mathbf{r}_j^T\mathbf{v} \le 0 \\ j \ne k}} - \mathbf{v}_o^T\mathbf{R}\mathbf{v} \tag{5.23}$$

Since this is a minimization problem of a convex functional on a convex set, we know it achieves a minimum on the set, and that a local minimum is a global minimum.Since all the functions are differentiable, we can apply the Kuhn-Tucker conditions [2] e.g. [4], to get : from condition (1) :

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

So $$\tilde{\mathbf{v}} = \frac{1}{2\lambda_o}\left(\mathbf{v}_o - \sum_{j\neq k} \lambda_j e_j \mathbf{u}_j \right) \tag{5.24}$$

with $\mathbf{u}_j$ the j^{th} unit vector, as defined above. Conditions (2),(3) are exactly the corresponding Kuhn-Tucker conditions, condition (4) expresses the nonnegativity requirement for the λ_i. There is one more constraint to satisfy, which is $\tilde{\mathbf{v}}^T\mathbf{R}\tilde{\mathbf{v}} = 1$:

$$1 = \tilde{\mathbf{v}}^T\mathbf{R}\tilde{\mathbf{v}} = \frac{1}{2\,\lambda_o}(\mathbf{v}_o^T\mathbf{R}\tilde{\mathbf{v}} - \sum_{j\neq k}\lambda_j e_j \mathbf{r}_j^T\tilde{\mathbf{v}}) = \frac{\mathbf{v}_o^T\mathbf{R}\tilde{\mathbf{v}}}{2\,\lambda_o} \tag{5.25}$$

We used condition (3) to get the last equality. So

$$2\,\lambda_o = \mathbf{v}_o^T\mathbf{R}\tilde{\mathbf{v}} \tag{5.26}$$

and since

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

we get

$$2\,\lambda_o = (\mathbf{v}_o^T\mathbf{R}\mathbf{v}_o + \mathbf{v}_o^T\mathbf{R}\sum_{j\neq k}\lambda_j e_j \mathbf{u}_j\,)^{1/2} \tag{5.27}$$

Together with equation (5.24) this completes the proof of Proposition 8. •

We would now like to have an explicit procedure to find the maximizing vector $\tilde{\mathbf{v}}$ given implicitly by Proposition 8. Next we give an algorithm which solves this problem. The idea is the following: condition (4) states that if the maximizing vector $\tilde{\mathbf{v}}$ lies in the intersection of n of the delimiting hyperplanes with equations $\mathbf{r}_j^T\tilde{\mathbf{v}} = 0, j \in S_n, n = 0,\ldots,K-1$, with S_n the index set of the specific hyperplanes, only the λ_j , $j \in S_n$ are possibly nonzero and enter into the expression defining $\tilde{\mathbf{v}}$. Thus we have $|S_n|$ equations with $|S_n|$ unknowns, which we can solve to get the λ_i, and then $\tilde{\mathbf{v}}$. First some notation.

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

$$D_{S_n}^n(j) = \det \begin{vmatrix} \mathbf{r}_j^T\mathbf{v}_o & r_{jj_1} & \cdots & r_{jj_n} \\ \mathbf{r}_{j_1}^T\mathbf{v}_o & r_{j_1j_1} & \cdots & r_{j_1j_n} \\ \cdot & \cdot & & \cdot \\ \mathbf{r}_{j_n}^T\mathbf{v}_o & r_{j_nj_1} & \cdots & r_{j_nj_n} \end{vmatrix} \tag{5.28}$$

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

$$\text{If} \quad e_j\, D_{S_n}^n(j) > 0 \quad \text{then} \quad C_{S_n}^n(j) = \text{yes} \;, \; \text{else} \quad C_{S_n}^n(j) = \text{no} \tag{5.29}$$

Definition 3 : An n-tuple S_n of $\{1,2,\ldots,k-1,k+1,\ldots,K\}$ is *matched* if for all i $\in S_n$: $C_{s_n-\{i\}}^{n-1}(i) =$ no.

Proposition 9 : The following algorithm finds a vector $\tilde{\mathbf{v}}$ which achieves the maximum in Proposition 8 :

[A] Search for the index set with least cardinality S_n , $n \in \{1,2,\ldots,K-1\}$, for which λ_i, $i \in S_n$, are possibly nonzero

```
n := 0
all n-tuples := untried $C_o$ := matched
while n $\leq$ K-1
    while there is still an untried matched n-tuple
        pick untried matched n-tuple := $S_n$
        for all $j \notin S_n$ do
            if $e_j\, D_{S_n}^n(j) > 0$ then $C_{S_n}^n(j) =$ yes
            else $C_{S_n}^n(j) =$ no
        if for all $j \notin S_n$ $C_{S_n}^n(j) =$ yes, return $S_n$, stop
        else $S_n$ := tried
    return
n := n+1
return
print "decorrelating detector is optimal", stop
```

[B] Solve the $|S_n|$ equations in the $|S_n|$ unknowns λ_i , $i \in S_n$: $\mathbf{r}_i^T\tilde{\mathbf{v}} = 0$, $i \in S_n$.

[C]

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

Comment : Recall that this procedure has to be repeated for all the different $\{ e_j \}$ in search of the maximal $\eta(\mathbf{e})$ value, until either the eficiency $\eta(\mathbf{e})$ reaches the upper bound given by the optimal detector, or all 2^K possibilities have been exhausted. Prior to running the algorithm, the sufficient conditions given in Propositions 10 and 11 should be checked.

Proof: Conditions (1) and (3) are obviously satisfied by construction of $\tilde{\mathbf{v}}$ in [C], and the requirement $\mathbf{r}_j^T\tilde{\mathbf{v}} = 0$ for the possibly nonzero λ_i in [B]. To prove conditions (2) and (4),consider the system of $|S_n|$ linear equations in $|S_n|$ unknowns, with $S_i = \{j_1, j_2, \ldots, j_n\}$. S_n is the set returned by step [A], hence it is matched, and satisfies $C_{S_n}^n(j) =$ yes for all $j \neq k, j \notin S_n$. We have to show:

a) $\lambda_{j_i} \geq 0$, for all i = 1,2,...,n

b) $C_{S_n}^n(j) =$ yes for all $j \neq k, j \notin S_n$ is equivalent to (or implies) condition (2).

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

$$\mathbf{r}_{j_1}^T\mathbf{v}_o + \lambda_{j_1}\mathbf{e}_{j_1}r_{j_1j_1} + \cdots + \lambda_{j_n}e_{j_n}r_{j_1j_n} = 0 \tag{5.30}$$

$$\mathbf{r}_{j_2}^T\mathbf{v}_o + \lambda_{j_1}\mathbf{e}_{j_1}r_{j_2j_1} + \cdots + \lambda_{j_n}e_{j_n}r_{j_2j_n} = 0$$

$$\cdots$$

$$\mathbf{r}_{j_n}^T\mathbf{v}_o + \lambda_{j_1}e_{j_1}r_{j_nj_1} + \cdots + \lambda_{j_n}e_{j_n}r_{j_nj_n} = 0$$

Denote by D_{S_n} the determinant of the coefficient matrix of the λ_{j_i}. Since this is exactly the crosscorrelation matrix of the subset S_n of the users, it has the same properties as the full matrix, specifically, it is positive definite. Hence its determinant is strictly positive. Then, by Cramer's rule,

$$\lambda_{j_i} = \frac{-\,e_{j_i} D_{S_n-\{j_i\}}(j_i)}{D_{S_n}}. \tag{5.31}$$

The numerator is obtained by i row flips and i column flips in order to get j_i into position (1,1). Since the set S_n is matched, the numerator is nonnegative. As obtained above, the denominator is positive, hence $\lambda_{j_i} \geq 0$ for all i = 1,...,n. All the remaining λ are zero, by definition and construction of the index set S_i. This completes the proof of a).

With the obtained values for λ compute the " feasibility " expressions :

$$e_j\mathbf{r}_j^T\tilde{\mathbf{v}} = e_j\,(\,\mathbf{r}_j^T\mathbf{v}_o + \sum_{j_i \in S_n} \mathbf{r}_j^T\lambda_{j_i}e_{j_i}\mathbf{u}_{j_i}\,)$$

$$= e_j\,(\,\mathbf{r}_j^T\mathbf{v}_o + \sum_{j_i \in S_n} \frac{-D_{S_n-\{j_i\}}^{n-1}(j_i)}{D_{s_n}}\, r_{jj_i}\,)$$

$$= \frac{e_j}{D_{S_n}}(\mathbf{r}_j^T\mathbf{v}_o D_{S_n} - \sum_{j_i \in S_n} D_{S_n-\{j_i\}}^{n-1}(j_i)\, r_{jj_i}\,) = \frac{e_j}{D_{S_n}} D_{S_n}^n(j) \tag{5.32}$$

The last equality is obtained by expanding along the first row of $D_{S_n}^n(j)$. This completes the proof of b).

By construction the algorithm terminates after at most K steps. •

---

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

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

$$\tilde{\mathbf{v}} = \frac{1}{\eta(\mathbf{e})} (\mathbf{v}_o - \frac{\mathbf{r}_j^T \mathbf{v}_o}{r_{jj}} \mathbf{u}_j) \tag{5.33}$$

and

$$\eta^2(\mathbf{e}) = \mathbf{v}_o^T \mathbf{R} \mathbf{v}_o - \frac{(\mathbf{r}_j^T \mathbf{v}_o)^2}{r_{jj}} \tag{5.34}$$

The asymptotic efficiency achieved in this case is bounded above by the one for n=0, since the second term is nonnegative. If the matrix **R** does not have a lot of structure, which is to be expected in practical applications, this is the most probable case. For increasing n the computational effort grows fast, but in most cases the algorithm will terminate for very small n. We also have explicit solutions for the terminal case n = K, which correspond to the decorrelating detector case. Then $\tilde{\mathbf{v}} = \mathbf{B}_k / \sqrt{B_{kk}}$, and $\eta(\mathbf{e}) = 1/B_{kk}$.

Proposition 10 : The following are sufficient conditions on the signal energies and crosscorrelations for the best linear detector to achieve optimal k^{th} user asymptotic efficiency:

$$\sqrt{w_k} > \max_{j=1,\dots,K} \left(\frac{1}{|\tilde{R}_{1j}|} \sum_{i \neq k} \sqrt{w_i} \, |\tilde{R}_{ij}| \right) \tag{5.35}$$

Proof : In the optimality case the optimality conditions $e_j \mathbf{r}_j^T \mathbf{v}_o > 0$ are satisfied for all $j \neq k$. If we introduce $e_k = 1$ they have to be satisfied also for j = k, else we get negative asymptotic efficiency. Rewrite these conditions as :

$$\mathbf{DRD} \, (-1 \; -1 \; \cdots \; 1 \; \cdots \; -1)^T > 0$$

where **D** is the diagonal matrix with i^{th} diagonal element equal to e_i. By expanding one gets :

$$\begin{bmatrix} r_{11} & e_1 e_2 r_{12} & \cdot\;\cdot & e_1 r_{1k} & \cdot\;\cdot & e_1 e_K r_{1K} \\ e_1 e_2 r_{21} & r_{22} & \cdot\;\cdot & e_2 r_{2k} & \cdot\;\cdot & e_2 e_K r_{22} \\ \cdot & \cdot & \cdot\;\cdot & \cdot & \cdot\;\cdot & \cdot \\ \cdot & \cdot & \cdot\;\cdot & \cdot & \cdot\;\cdot & \cdot \\ e_1 e_K r_{K1} & e_K e_2 r_{K2} & \cdot\;\cdot & e_K r_{Kk} & \cdot\;\cdot & r_{KK} \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ \cdot \\ 1 \\ \cdot \\ -1 \end{bmatrix} > 0 \tag{5.36}$$

We now see that a sufficient condition for the above inequality to hold is

$$|r_{jk}| > \sum_{i \neq k} |r_{ji}| \quad , j = 1,\dots,K.$$

Proposition 10 follows by replacing r_{ij} by $\sqrt{w_i}\sqrt{w_j}\tilde{R}_{ij}$. •
Note that this condition is in fact necessary for K = 2. Also, parallel to the case K = 2, the above condition is satisfiable only for one user, since

$$\sqrt{w_k} > \cdots > \sqrt{w_j} / |\tilde{R}_{kj}| > \sqrt{w_j} \text{ , for all } j \text{ ,since } |\tilde{R}_{kj}| < 1.$$

Proposition 11 : The following condition is sufficient for the decorrelating detector to be the best k^{th} user linear detector for a given set of signal energies and crosscorrelations :

$$|B_{jk}| \leq B_{kk} \quad , \quad \text{for all} \quad j \neq k. \tag{5.37}$$

Proof : We have

$$\tilde{\mathbf{v}} = \frac{(\mathbf{B})_k}{\sqrt{B_{kk}}} = \frac{1}{(\mathbf{v}_o^T \mathbf{R} \mathbf{b}_k / \sqrt{B_{kk}})^{1/2}} \; (\mathbf{v}_o + \sum_{j \neq k} \lambda_j e_j \mathbf{u}_j)$$

or

$$\frac{1}{B_{kk}} \mathbf{b}_k = \left[(\lambda_1 - 1) e_1 \ldots 1 \ldots (\lambda_K - 1) e_K \right]^T$$

so

$$\lambda_j = 1 + e_j B_{jk} / B_{kk} \geq 0 \tag{5.38}$$

It is clear that the condition given in Proposition 11 is suficient to ensure $\lambda_j \geq 0$ regardless of $\{e_i \; , \; i \neq k\}$.

•

For K = 2 the above condition is necessary and reduces to the condition in Proposition 6.

6. Conclusions

The main contribution of this paper is to have found that an appropriately chosen memoryless linear transformation on the matched filter outputs of a multiuser receiver structure will exhibit substantially higher asymptotic efficiencies than the conventional single user detector, while maintaining a comparable ease of computation (a sign decision is performed on each output of the K-dimensional linear transformation).

An algorithm is derived which finds the best linear transformation in the general case. One of the most interesting results of this paper is that there is a region of signal energies and crosscorrelations in which application of this linear transformation ensures optimal asymptotic efficiency. Sufficient conditions to lie in this region are given. Another main result is a lower bound on the asymptotic efficiency achievable by insertion of the best linear transformation. This lower bound is achieved by the decorrelating detector and ensures an asymptotic efficiency which is independent of the energy of the interfering users. Moreover this lower bound equals the minimum of the asymptotic efficiency of the (energy dependent) optimal multiuser detector, taken over the energies of the interfering users. This result makes the decorelating detector particularly attractive in a near-far environment where the energies of the interfering users vary continuously in a range containing the minimum: then as far as worst case performance is concerned the optimum multiuser detector and the decorelating detector are equivalent, while in terms of time complexity per bit the decorelating detector is far superior.

The main interest of this work is that it constitutes a prelude to the study of asynchronous linear decision rules. Due to the fact that in the asynchronous case the channel has memory a K-input K-output linear discrete-time filter will replace the memoryless linear transformation studied in this paper.

References

[1] S.Verdu, "Minimum Probability of Error for Asynchronous Gaussian Multiple-Access Channels", *IEEE Trans. on Information Theory*, Vol.IT-32 (Jan.1986),pp.85-96.

[2] S.Verdu, "Optimum Multi-User Detection is NP-hard", 1985 International Symposium on Information Theory, Brighton, UK, June 1985.

[3] S.Verdu, "Optimum Multi-User Asymptotic Efficiency", *IEEE Trans. on Communications*,Vol.COM-34 (Sept.86),pp.890-897.

[4] D. Luenberger, *Optimization by Vector Space Methods*, Wiley, 1969.

[5] J. Wozencraft and I. Jacobs, *Principles of Communication Engineering*, Wiley, 1965.

[6] K.S.Schneider, "Optimum Detection of Code Division Multiplexed Signals", *IEEE Trans. Aerosp.Electron.Syst.*, AES-15 (1), pp 181-185 (Jan 1979).

[7] T.Kailath, *Linear Systems*, Prentice-Hall, 1980.