ASYMPTOTIC EFFICIENCY OF LINEAR MULTIUSER DETECTORS

Ruxandra Lapsa-Golenewski 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 performance of different detection rules. The two most important detection rules are compared and computed, namely, the minimum mean square error detector where complexity is exponential in the number of users and the single user detector whose performance degrades very slowly as the maximum interference energy increases. In contrast, a new linear transformation on the matched filter outputs of an receiver is proposed, which is shown to guarantee a higher lower bound on the asymptotic efficiency, which is independent of signal energies, while the time-complexity per bit in 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 correlation matrices to ensure optimum asymptotic efficiency of the proposed detector.

1. Introduction

K users are transmitting synchronously over a multiple-access channel and permitted with additive white Gaussian noise. To this end each user modulates a different signal waveform, in this paper we will assume that the transmission occurs asynchronously, in which case the waveform received in the $j$-th symbol interval has the form:

$$r(t) = \sum_{n=0}^{N-1} a_n x(t-nT) + n(t), \quad t \in [jT, (j+1)T),$$

where $a_n$ are the transmitted symbols and the signal waveform corresponding to one is $x(t)$. 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 thresholding. Unfortunately, bandwidth limitations prevent the usage of orthogonal signal waveforms for a large number of users. In this case, even for low signal-to-noise ratios, it is far from optimal to construct using single-user detectors, neglecting the lack of orthogonality of the signal set $\{x(t)\}$, for performance of the linear detector and the single-user detector degrades monotonically with increasing relative size of the interfering users. The minimum likelihood user detector is nonlinear and has been shown to be NP-hard, that is, there is no corresponding decision algorithm which is polynomial in the number of users, unless NP=P [10]. 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 receivers, in order to improve the error rate in comparison to the bank of single-user detectors used in practice, while maintaining asymptotic optimum performance as the number of users goes to infinity. The linear combination transformation between the matched filters and the thresholds and the vector which linear combination of the matched filter outputs has the best performance with respect to bit error rate, and how close the performance gets to that of the optimal receiver. The performance receivers used is the asymptotic efficiency, which is a good measure of the performance in the limit to other error metrics on the channel, in the high SNR regime. The $K$-user asymptotic efficiency of a detector whose bit error rate is equal to $P_e$ has been defined as

$$P_e = \exp \left( -\frac{\rho}{2} \right) \left( 1 - \frac{\rho}{2} \right)^{K/2}, \quad \rho \geq 1.$$
where
\[ R_y = \frac{1}{\sigma^2} \left[ \text{cov}(y, \epsilon) \right] \]

is the correlation between the observed signals \( x_i \) and \( y \) and the residual signal in \( \epsilon = x - x_i \). In vector notation, the matched filter output vector \( y \) is \( \mathbf{y} = R_y x \). The assumption of linear independence of the signal set guarantees positive definiteness of the correlation matrix \( R_y \) for all \( x \) and \( y \). Let \( \mathbf{x} \) be the vector of \( n \) and \( \mathbf{y} \) be \( \mathbf{B} \) elements in inverse (symmetric and positive definite) with columns \( b_i \). \( e \) denotes the energy of the signal \( s \), \( w = \sqrt{\sigma^2} \mathbf{y} \); the elements of the correlation matrix \( R_y \) can be expressed as \( R_y = \frac{1}{\sigma^2} \mathbf{B} \mathbf{B}^T \mathbf{R}_y \mathbf{B} \), where \( \mathbf{R}_y \) is the correlation coefficient matrix of signals \( x_1 \) and \( x_2 \).

2. The Conventional Single-User Detector

In practice, single-user detection is applied to a multistatic environment. A threshold decision is performed on each matched filter output \( y_i \) as
\[ \delta_{i} = \text{sign} \left( y_i \right) \alpha_i \]

for \( y_i > \alpha_i \) and \( y_i < -\alpha_i \). This approach is not optimal due to the interference term \( \sum_{j \neq i} R_{yj} x_j \).

Proposition 1: The \( k \)-th user asymptotic efficiency of the conventional detector is
\[ \eta_k^U = \frac{1}{\sigma^2} \left[ 1 - \sum \frac{\left| R_{kj} \right|}{\sqrt{\sum \left| R_{kj} \right|^2}} \right] \]

Proof: The \( k \)-th user probability of error is given by:
\[ P_k = \frac{1}{2} \left[ 1 - \text{erfc} \left( \frac{\sqrt{\sum \left| R_{kj} \right|^2} - \left| R_{kj} \right|}{\sigma} \right) \right] \approx \frac{1}{2} \left( \frac{\sigma}{\sqrt{\sum \left| R_{kj} \right|^2}} \right) \]

Since the matrix \( \mathbf{R}_y \) is Cholesky with zero mean and variance equal to \( \sigma^2 \), the sum in (2.2) is dominated by the term
\[ \sum \left| R_{kj} \right| \approx \sigma \sqrt{\sum \left| R_{kj} \right|^2} \]

Here the \( k \)-th user asymptotic efficiency of the single-user receiver is equal to zero if \( R_{kk} = \sum_{j \neq k} R_{kj} \). Otherwise it is equal to the square of the ratio of the argument of the lower-Q function and the argument corresponding to the single-pole asymptote of \( Q(\sqrt{\sum \left| R_{kj} \right|^2}) \).

\[ \eta_k^U = \frac{1}{\sigma^2} \left[ 1 - \sum \frac{\left| R_{kj} \right|}{\sqrt{\sum \left| R_{kj} \right|^2}} \right] \]

Note that the \( k \)-th user asymptotic efficiency of the conventional detector is equal to zero for sufficiently high signal-to-noise ratios of any user not orthogonal to the \( k \)-th user. In particular, in the case of two active users \( \sum \left| R_{kj} \right| \approx 1 / \sigma^2 \), results in \( \eta_k^U = 0 \).

3. The Decoupling Detector

In the absence of noise, the matched filter output is \( y = Rx \). So, the optimal strategy is to follow the principle \( y \) by the inverse correlation matrix \( R_y = \mathbf{B}^{-1} \mathbf{B} \). The solution \( y = \mathbf{B}^{-1} \mathbf{B} \) will be a decoupled detector, and Proposition 2 captures the performance in the presence of noise. It is extremely common in (1.9) that this detector is optimal in terms of hit-rate error.

Proposition 2: The \( n \)-th user asymptotic efficiency of the decoupled detector is given by
\[ \eta_n^D = \frac{1}{\sigma^2} \left[ \mathbf{R}_n \right]^T \mathbf{R}_n \]

Then the \( k \)-th user asymptotic efficiency of the decoupled detector is independent of the energy of other users.

Proof: We have
\[ \delta_k = \text{sign} \left( y_k \right) \alpha_k \delta_k (k > 1, \text{all terms}) \]

The \( k \)-th user error probability is, if symmetry (for equally likely unambiguously symbolized symbols)
\[ P_k = P_0 \left( 1 - e^{-1} \right) = P_0 (R_{kk} > 1) \]

Also, \( (R_{kk}) \) is Gaussian, with variance equal to the \( k \)-th diagonal element of
\[ \sum R_{kj} \sum_{j \neq k} R_{kj} = \frac{1}{\sigma^2} \left( \frac{1}{\sqrt{n}} \right) \]

By definition of \( \alpha_k \). Therefore, by monotonicity of \( Q(\cdot) \),
\[ \eta_k^D = 1 - \sum_{j \neq k} \frac{1}{\sigma^2} \frac{1}{\sqrt{\sum \left| R_{kj} \right|^2}} \]

Next, we show that \( \eta_k^D > 1 \). Since \( \eta_k^D = 1 \), we prove that \( \eta_k^D > 1 \) for all pairs of symmetric positive definite matrices \( R \) and \( E \) and 

Thus
\[ P_k = P_0 (R_{kk} > 1) = Q(1/\sqrt{\sum \left| R_{kj} \right|^2}) = Q(1/\sqrt{\sum \left| R_{kj} \right|^2}) \]

4. The Optimum Multistatic Detector

We now turn attention to the optimum multistatic detector under the maximum likelihood criterion. That is the detector which decides for the transmitted vector \( x \), which is most likely to have produced the received signal \( y = Rx + W \).

**4.1.** Taking into account the Gaussian noise statistics, a few equivalent optimum standard deviation of the above yield
\[ \eta_k = 1 - \frac{1}{\sigma^2} \left( \frac{1}{\sqrt{n}} \right) \]

It has been shown recently [2] that the cumulant minimization problem (4.2) is NP-complete. This is a key result which provides the main result which allows us to seek suboptimal detection schemes which offer both acceptable performance and computations. 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 trrett analysis, [2], has shown that the computation of the optimum kth user asymptotic efficiency is also an NP-complete computational problem. Because of this reason we will state this basic result obtained in [2]. The optimal kth user asymptotic efficiency for isotropic modulation can be expressed as follow:

\[ \gamma_k = \frac{1}{k(N-k)} \min_{w_k} \left( \frac{1}{2} \sum_{i=1}^{k} \left| \frac{2}{\sqrt{v_i^2 + \sum_{j=1}^{k-1} w_j^2}} \right| - 1 \right) = \frac{1}{R_{\text{out}}} \min_{w_k} \frac{1}{2} \sum_{i=1}^{k} w_i^2 \] (4.3)

Proposition 3. The following problem is NP-hard.

Given \( K \leq 2 \), \( z \in [0, \infty) \), and a positive definite matrix \( R \in \mathbb{R}^{m \times m} \), find the kth user asymptotic efficiency \( \gamma_k \) that satisfies

\[ \frac{1}{k(N-k)} \min_{w_k} \left( \frac{1}{2} \sum_{i=1}^{k} \left| \frac{2}{\sqrt{v_i^2 + \sum_{j=1}^{k-1} w_j^2}} \right| - 1 \right) = \frac{1}{R_{\text{out}}} \min_{w_k} \frac{1}{2} \sum_{i=1}^{k} w_i^2 \] (4.8)

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]. In particularization to the synchronous case the following expression holds.

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

\[ \gamma_2 = \frac{1}{2} \left( \frac{\sqrt{v_1} + \sqrt{v_2}}{\sqrt{v_1} \sqrt{v_2}} \right) - \frac{\sqrt{v_1} - \sqrt{v_2}}{\sqrt{v_1} \sqrt{v_2}} \] (4.9)

While the asymptotic efficiency of the decoding 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 energy varies dynamically over a wide range (e.g., near-flat fading), it is important to assess the worst-case asymptotic efficiency with respect to the energy of the interfering users.

Proposition 5. For any user asymptotic efficiency of the maximum likelihood detector, taken over all possible energies of the interfering users, equals the asymptotic efficiency of the decoding detector:

\[ \gamma_{\text{max}} = \gamma_2 \] (4.10)

Proof. Using expression (4.3) for the asymptotic efficiency of the optimum multi-user detector, we have

\[ \gamma_{\text{max}} = \min_{w_2} \left( \frac{1}{k(N-k)} \sum_{i=1}^{k} \left| \frac{2}{\sqrt{v_i^2 + \sum_{j=1}^{k-1} w_j^2}} \right| - 1 \right) \]

\[ = \min_{w_2} \left( \frac{1}{k(N-k)} \frac{1}{\sqrt{v_1^2 + \sum_{j=1}^{k-1} w_j^2}} \right) \]

\[ = \frac{1}{k(N-k)} \] (4.7)

with the obvious substitution \( \alpha = \sqrt{v_1^2 + \sum_{j=1}^{k-1} w_j^2} \). As previously defined, the normalized cochannel matrix \( R \) by removing the kth row and column, a, is the vector a without the kth element, equation (4.7) can be expanded as:

\[ \gamma_{\text{max}} = \min_{w_2} \left( \frac{1}{k(N-k)} \frac{1}{\sqrt{v_1^2 + \sum_{j=1}^{k-1} w_j^2}} \right) \]

\[ = \frac{1}{k(N-k)} \frac{1}{\sqrt{v_1^2 + \sum_{j=1}^{k-1} w_j^2}} \]

\[ = \frac{1}{k(N-k)} \] (4.11)

This is a straightforward minimum quantization problem with solution:

\[ w_2 = \frac{1}{2} \sqrt{M^2 - a} \] (4.12)

for

\[ a = \frac{1}{2} \sqrt{M^2 - a} \] (4.13)

We have used the inequality \( a \leq v \) and the fact that the matrix M being diagonal in the matrix R for users 1, 2, 3, ..., K, is positive definite, hence invertible. Remains to show that the expression obtained in (4.9) equals

\[ \gamma_2 = \frac{1}{2} \left( \frac{\sqrt{v_1} + \sqrt{v_2}}{\sqrt{v_1} \sqrt{v_2}} \right) - \frac{\sqrt{v_1} - \sqrt{v_2}}{\sqrt{v_1} \sqrt{v_2}} \] (4.14)

where the last equality follows from the expression for the determinant of a block matrix (see [2], p. 567).

There is a nice geometric explanation for the equality of \( \gamma_{\text{max}} \) and \( \gamma_2 \). For example, by computing the appropriate minors in the geometric setting given below (may 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 y satisfies \( y = Ra + n \) and that the noise autovalue matrix \( R \) is. It is convenient to work in the \( \mathbb{C}^2 \)-domain, call it domain 2, where the hypens are the points \( \mathbb{R}^2 \), with \( \mathbb{R}^2 \rightarrow \mathbb{C}^2 \), and where the noise is hyperbolic symmetry. Therefore the decision region of the maximum likelihood detector, intersected by the minimum Euclidean distance, are given by the perpendicular bisectors of the segments between the different hypothesis. Recall that the kth user asymptotic efficiency corresponds to the sum of the minimum distances between different hypotheses differing in the kth bit \( \mathbb{D} = \mathbb{D} \) (where \( \mathbb{D} \) is a hyperplane). The decision regions of the decodating detector are coher to the one at the origin, such that application of \( \mathbb{R} \) maps them to the domain 1.

Thus in domain 1, the decision region passes through the points \( \mathbb{R}^2 \), with unit vectors in \( \mathbb{R}^2 \). It is interesting to observe that these points are at the center of the sides of the parallelogram formed by the hypens, because the unit vectors can be represented as half the sum of the attributed hypens. Therefore, the decoding detector decision boundaries are parallel to the parallelogram sides and intersect at the center of its sides.

As a consequence of the property a geometric interpretation of the kth user asymptotic efficiency of the decoding detector can be as follows: recall that in order to find the kth user asymptotic efficiency of a detector, we have to approximate its probability of kth bit error by a Gaussian, in the low noise regime. The kth bit error probability - by symmetry - we assume that the transmitted bit is 1- n the sum of two independent 0.5 for each possibility for the remaining bit, of the noise density function over the region in which the kth bit is decided as 1. In our case the kth bit error probability can be easily computed by taking advantage of the properties presented above. To this end we will rotate the mendes system to let the y- axis coincide with the kth bit decision boundary and we set the equal distance property of the domain boundary to the hypothesis, to observe that the two integrals are equal, and thus use the symmetrical symmetry of the same to identify each integral as Gaussian of the distance to the hypotheses to the decision boundary. Hence the kth user asymptotic efficiency of the decoding detector is equal to the square of the distance any hypothesis to the kth bit decision boundary.

For the sake of figure \( X \), \( X \) is the length of the distance of the segments \( A \), \( \mathbb{D} \), and \( L \), in length of \( \mathbb{D} \). After \( \mathbb{R} \) is the row and column \( \mathbb{C} \), the matrices can be computed as:

\[ \det R = \det \frac{1}{2} \left[ \frac{1}{2} \sum \left( \frac{1}{\sqrt{v_i^2 + \sum_{j=1}^{k-1} w_j^2}} \right) - 1 \right] = \det M \left( \frac{1}{M} \right) \]

Fig. 1. Hypotheses and decision regions in domain S

![Fig. 1. Hypotheses and decision regions in domain S](image-url)
The results of Proposition 4 are of special interest in a source for continuous fields in which the received signals have different frequencies and where the energy contents are nonstationary over a broad band of the position of the source evolves dynamically. In this environment, if waveforms perform is considered, the decorrelating detector, with its large time-complexity per bit, offers the same performance as the NB-peak optimum linear detector.

The Optimum Linear Multiuser Detector

We now apply the improvement obtained by a linear transformation of the matched filter output prior to sign detection and ask the optimal linear detector. This has been used to an important problem since the optimal detector is exponential in the number of users, and the performance of the computationally efficient linear detector depends on how given energy of all users. First we pose 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 compute the achieved asymptotic efficiency in the case reached by the optimum linear and optimum. Thus we ask what mapping $\mathbf{T}$ is the transformation of $\mathbf{Y}$ that will minimize the probability of bit error for the decision scheme $\hat{\mathbf{b}} = \operatorname{sign}(\mathbf{T}\mathbf{y}) = \operatorname{sign}(\mathbf{T}^{*}\mathbf{X}\mathbf{b})$.

(5.4)

In terms of decision region the problem simulates the following: What is the optimal way to partition the multidimensional hypothesis space into $K$ decision regions $K$ sets without creating the surfaces of these sets determine the contours of the constant $T^{*}$ of the rough mapping. Application of $\mathbf{T}$ on the core configuration will map the core on quadratic, which is a top detector is used.

The best error probability of the $\mathbf{b}$ user under this detection scheme, i.e., the detection of the unconditioned bit symbol $(a)$ and $\operatorname{dist}(\mathbf{x}, \mathbf{y})$. The optimal solution of the $\mathbf{T}$ is a function of the $\mathbf{T}^{*}$ and $\mathbf{X}^{*}$, and $\mathbf{T}^{*}$ is defined as $\operatorname{dist}$. The term $T^{*}$ is understood as $\operatorname{dist}$ by the term $\mathbf{Q}(\mathbf{T}) = \{\mathbf{T}^{*}(\mathbf{X})\}^{2} = \mathbb{E}(\mathbf{T}^{*}(\mathbf{X})^{2})$.

(5.5)

The asymptotic efficiency of the best linear detector is bounded away from zero by that of the decorrelating detector which has $\mathbf{T} = 0$. Hence it is equal to the square of the ratio of the asymptotics of the foregoing [7] and the argument corresponding to the null-user probability of error $\mathbf{Q}/\mathbf{Q}(\mathbf{X})$.

(5.6)

In order to minimize $\mathbf{Q}$ we have to maximize the argument of the $\mathbf{T}$ function, and equivalently maximize the asymptotic efficiency $\mathbf{q}^{2}$ with respect to the components of the vector $\mathbf{v}$. Note that $\mathbf{q}^{2}$ is invariant with respect to scaling of $\mathbf{v}$.

First we consider the two-user case for which explicit expression of the minimum asymptotic efficiency is obtainable.

6.1. The Two-User Case

Proposition 5: The $\mathbf{b}$ user optimal linear transformation $\mathbf{T}(\mathbf{y}) = \mathbf{v}^{*} \mathbf{y}$ so that the matched filter outputs prior to threshold decision is $\mathbf{q}^{2} = \left| -\mathbf{e}^{2} \right| \min \left| \mathbf{v}^{*} \mathbf{e} \right|^{2}$, with

(5.5)

Equivalently, $\mathbf{q}^{2} = \left| -\mathbf{e}^{2} \right| \min \left| \mathbf{v}^{*} \mathbf{e} \right|^{2}$ with $\mathbf{v}^{*} \in \mathbb{C}^{L}$, $|v_{i}| = 1$, $i = 1, \ldots, L$. Then $\mathbf{q}^{2}$ is obtained.

(5.6)

Proof: Theorem 3: Without loss of generality, we will consider the two-user case. We have

(5.7)

Then we want to minimize $\mathbf{q}$ with respect to $\mathbf{v}_{0}$, introduce an indicator function for the absolute value term:

(5.8)

Therefore we can take $\mathbf{v}_{0} = 0$ when this is consistent with the definition of $\mathbf{b}$ function $\mathbf{b}_{0}$. Thus,

(5.9)

As can easily be seen, both values correspond to maxima. If one of these conditions is not met, the derivatives do not have a root. The optimal value for $\mathbf{b}$ can be determined from a close look at the behavior of $\mathbf{q}^{2}$ for $\mathbf{b} \neq 0$.

For both $\mathbf{b} = 1$ and $\mathbf{b} = -1$, the derivative of $\mathbf{q}^{2}$ is positive for $\mathbf{b} > 1$ smaller than the absolute of the sum of the derivative $\mathbf{b}$ is negative, and the absolute value of $\mathbf{q}$ for $\mathbf{b} > 0$.

As an easy way to check, both values correspond to maxima. If one of these conditions is not met, the derivatives do not have a root. The optimal value for $\mathbf{b}$ can be determined from a close look at the behavior of $\mathbf{q}^{2}$ for $\mathbf{b} \neq 0$.

For both $\mathbf{b} = 1$ and $\mathbf{b} = -1$, the derivative of $\mathbf{q}^{2}$ is positive for $\mathbf{b} > 1$ smaller than the absolute of the sum of the derivative $\mathbf{b}$ is negative, and the absolute value of $\mathbf{q}$ for $\mathbf{b} > 0$.

Note that for $\mathbf{b} = 0$ we get $\mathbf{q}^{2} = 0$, i.e., the identity transformation, as expected, since this the users are decoding and a single-user detector is optimal. By taking the inverse of $\mathbf{b}$ we also see that in the "invert" case the optimal transformation vector is exactly the corresponding row of the inverse correlation matrix.

Proposition 7: The $\mathbf{b}$ user asymptotic efficient of the optimal linear two-user detector equals

\[ \mathbf{q}^{2} = \left| -\mathbf{e}^{2} \right| \min \left| \mathbf{v}^{*} \mathbf{e} \right|^{2} \cdot \mathbf{b}_{0}^{*} \quad \text{or} \quad \left( \begin{array}{c} \mathbf{b}_{0}^{*} \mathbf{b}_{0} \end{array} \right) \]

(5.6)

\[ \mathbf{q}^{2} = \left| -\mathbf{e}^{2} \right| \min \left| \mathbf{v}^{*} \mathbf{e} \right|^{2} \cdot \mathbf{b}_{0}^{*} \quad \text{or} \quad \left( \begin{array}{c} \mathbf{b}_{0}^{*} \mathbf{b}_{0} \end{array} \right) \]

(5.6)

\[ \mathbf{q}^{2} = \left| -\mathbf{e}^{2} \right| \min \left| \mathbf{v}^{*} \mathbf{e} \right|^{2} \cdot \mathbf{b}_{0}^{*} \quad \text{or} \quad \left( \begin{array}{c} \mathbf{b}_{0}^{*} \mathbf{b}_{0} \end{array} \right) \]

(5.6)

\[ \mathbf{q}^{2} = \left| -\mathbf{e}^{2} \right| \min \left| \mathbf{v}^{*} \mathbf{e} \right|^{2} \cdot \mathbf{b}_{0}^{*} \quad \text{or} \quad \left( \begin{array}{c} \mathbf{b}_{0}^{*} \mathbf{b}_{0} \end{array} \right) \]

(5.6)

\[ \mathbf{q}^{2} = \left| -\mathbf{e}^{2} \right| \min \left| \mathbf{v}^{*} \mathbf{e} \right|^{2} \cdot \mathbf{b}_{0}^{*} \quad \text{or} \quad \left( \begin{array}{c} \mathbf{b}_{0}^{*} \mathbf{b}_{0} \end{array} \right) \]

(5.6)
The asymptotic efficiency obtained in the range \(\omega_0/\omega < |\delta_0|\) equals the optimum attainable asymptotic efficiency; that of the maximum likelihood detector. Even outside the limits of optimality, the best linear detector shows a far lower performance than the optimum one, with the ratio of the energy of the white noise signal to the signal energy. This is the intuitive interpretation of the destination of the best linear detector and of the bound between \(\omega_0/\omega < |\delta_0|\) and \(\omega_0/\omega > |\delta_0|\). The means to the threshold device corresponding to the zero error, 
\[
\text{Re}_{\omega_0/\omega} = \frac{1}{\delta_0}.
\]

Proof: The transversal of the optimal linear transformation found in Proposition 6 in the asymptotics of the expressions for the asymptotic efficiency in the range \(\omega_0/\omega < |\delta_0|\).

\[
\begin{align*}
\text{Re} \left[ \frac{1}{\delta_0} \right] &= \frac{1}{\delta_0} \\
\text{Re} \left[ \frac{1}{\delta_0} \right] &= \frac{1}{\delta_0} \\
\end{align*}
\]

For \(\gamma = \begin{bmatrix} \begin{array}{c} 1 \\ \omega_0/\omega \end{array} \end{bmatrix} \left[ \begin{array}{c} 1 \\ \omega_0/\omega \end{array} \right]^T \).

The asymptotic efficiency obtained in the range \(\omega_0/\omega < |\delta_0|\) equals the optimum attainable asymptotic efficiency; that of the maximum likelihood detector. Even outside the limits of optimality, the best linear detector shows a far lower performance than the optimum one, with the ratio of the energy of the white noise signal to the signal energy. This is the intuitive interpretation of the destination of the best linear detector and of the bound between \(\omega_0/\omega < |\delta_0|\) and \(\omega_0/\omega > |\delta_0|\). The means to the threshold device corresponding to the zero error, 
\[
\text{Re}_{\omega_0/\omega} = \frac{1}{\delta_0}.
\]

Proof: The transversal of the optimal linear transformation found in Proposition 6 in the asymptotics of the expressions for the asymptotic efficiency in the range \(\omega_0/\omega < |\delta_0|\).

\[
\begin{align*}
\text{Re} \left[ \frac{1}{\delta_0} \right] &= \frac{1}{\delta_0} \\
\text{Re} \left[ \frac{1}{\delta_0} \right] &= \frac{1}{\delta_0} \\
\end{align*}
\]

For \(\gamma = \begin{bmatrix} \begin{array}{c} 1 \\ \omega_0/\omega \end{array} \end{bmatrix} \left[ \begin{array}{c} 1 \\ \omega_0/\omega \end{array} \right]^T \).
Since this is a minimization problem of a convex function 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 condition 1 \( g \leq 0 \) to get from condition (3):
\[
-\mu_x \lambda + \lambda x^T R x = 0
\]
for all \( x \). Since \( x \) is a vector, we define above. Condition (2) is exactly corresponding Kuhn-Tucker condition, condition (1) expresses the non-negativity requirement for the \( \lambda_i \). There is one more constraint to satisfy, which is \( x^T R x = 1 \):
\[
1 = x^T R x = x^T R x = \sum_{i=1}^{n} x_i \lambda_i x_i
\]
We used condition (1) to get the last equality. So
\[
\lambda = \frac{1}{x^T R x} \sum_{i=1}^{n} x_i x_i \lambda_i
\]
and since
\[
\nu = \frac{1}{\lambda} \sum_{i=1}^{n} x_i \lambda_i x_i
\]
we get
\[
\lambda = \frac{1}{\nu^2} \sum_{i=1}^{n} x_i \lambda_i x_i
\]
Together with equation (26) this completes the proof of Proposition 8.

We would now like to have an explicit procedure to find the maximizing vector \( x \) given implicitly by Proposition 8. Next we give an algorithm which solves this problem. The idea is the following conditions by states that if the maximizing vector \( \lambda \) is the sum of a of the dominant \( \rho \) values with equal \( \nu \) and \( \nu_2 \) the index set of the specific hyperplane, only the \( \lambda_i \), \( i \in \rho \) are possibly nonzero and since into the expression defining \( \nu \). Thus we have \( \lambda \) is defined with \( \lambda_2 \) unknown, which we can solve to get \( \nu_2 \lambda_2 \), and then \( \nu_2 \).

First some notation.

Definition 1: Let \( \lambda \) be an index set \( \{i_1, i_2, \ldots, i_\rho\} \) \( 0 \leq \rho \leq K-1 \), with \( i_1, i_2, \ldots, i_\rho \in \{1, 2, \ldots, n\} \), labeled in increasing order. Define
\[
D_{\rho}(\lambda) = \sum_{i=1}^{\rho} x_i \lambda_i x_i
\]
(26)

Definition 2: We introduce an indicator for the second Kuhn-Tucker condition:
\[
\tau = D_{\rho}(\lambda) > 0 \quad \text{then} \quad \nu_2(\lambda) = \nu_2 > 0 \quad \text{also} \quad \nu_2(\lambda) = 0
\]
(29)

Definition 3: An optimal set \( \lambda \) is \( \{1, 2, 3, \ldots, K-1\} \) is maximized \( D_{\rho}(\lambda) > 0 \), for all \( \lambda \in \Omega \), \( \Omega_0(\lambda) = 0 \).

Proposition 8: The following algorithm finds a vector \( x \) which achieves the maximum

\[
A \text{ search for the index set with largest } \nu_2 \text{ i.e. } \nu_2 = \max(\Omega_0(\lambda))
\]

for which \( \lambda \), \( i \in \Omega_0 \), is possibly nonzero.

\[\text{Algorithm:}\]

1. if there exists a vector \( \lambda \) such that \( D_{\rho}(\lambda) = 0 \), then stop.
2. if \( \tau = D_{\rho}(\lambda) > 0 \), then \( \nu_2(\lambda) = \nu_2 > 0 \) and go to step 3.
3. if \( \tau = D_{\rho}(\lambda) < 0 \), then \( \nu_2(\lambda) = 0 \).

The last (equality) is obtained by expanding along the first term of \( D_{\rho}(\lambda) \).

By construction the algorithm terminates after at most \( n_0 \) steps.
In [1], it is noted that a = 0 corresponds to a solution in the transim fringe that is realizable, while all λ equal to zero, and γ = γ, the corresponding asymptotic efficiency 
$$e^2 \approx 0 = \sqrt{\text{Re}(\gamma)}$$ 
which equals to the asymptotic efficiency of the maximum likelihood detector. Cell this case the optimality case a. On the other hand, γ ≈ 1 corresponds to a solution on exactly one of the optimal hypotheses with exactly one error for be 1, and

$$e = 1$$

and

$$e(\gamma) = \sqrt{\text{Re}(\gamma)}$$

The asymptotic efficiency achieved in this case is bounded above by the one for a = 0, since the second term is non-negative. If the matrix A does not have a set of eigenvalues, which is to be expected in practical evaluations, this is the most probable case. For increasing λ the computational time grows fast, but in some cases the algorithm will terminate for very small λ. We also have explicit solutions for the special case γ = γ, which correspond to the deterministic detector case. Then

$$e = \sqrt{\text{Re}(\gamma)}$$

and

$$\lambda(\gamma) = 0$$

Proposition 10: The following are sufficient conditions for the signal and constraints for the best linear detectors to achieve optimality in asymptotic sense:

$$\sqrt{e^2} > \frac{1}{\lambda} \sum_{i=1}^{d} \sqrt{\text{Re}(|\lambda_i|)}$$

Proof: In the optimal case the optimality conditions are e/\sqrt{e^2} > 0 also satisfied for all γ. If we introduce \( \gamma \approx 1 \) they have to be satisfied also for \( \lambda = 1 / \lambda \), the non-asymptotic sense. Rewrite these conditions as:

$$\lambda(\gamma) = 1$$

where \( \gamma \) is the diagonal matrix with \( \gamma \) diagonal elements equal to \( \gamma \). By expanding our results:

$$\sum_{i=1}^{d} \gamma_{ii} > \sqrt{\text{Re}(\lambda)}$$

where \( \gamma_{ii} > 1 \)

We see that a sufficient condition for the best linear detector is to have:

$$\lambda(\gamma) = 1$$

Note that this condition is in fact necessary for K = 2, also, parallel to the case K = 2, the above condition is satisfied only for one error, since

$$\sqrt{e^2} > \frac{1}{\lambda} \sum_{i=1}^{d} \sqrt{\text{Re}(|\lambda_i|)} > \frac{1}{\lambda} \sum_{i=1}^{d} \sqrt{\text{Re}(|\lambda_i|)}$$

Proposition 11: The following condition is sufficient for the best Kth order linear detector for a given set of signal source constraints:

$$|\lambda_i| \leq 1 \quad \forall i \in \{1, 2, \ldots, K\}$$

$$\sum_{i=1}^{d} \gamma_{ii} = \lambda(\gamma) = 1$$

where \( \lambda(\gamma) \) is the non-asymptotic sense. It is clear that the condition given in Proposition 11 is sufficient to ensure

$$\lambda(\gamma) \geq 0$$

For γ = 1 the above condition is necessary and reduces the condition in Proposition 6.

References:


