Advanced Signal Processing Seminar,

Institute of Signal Processing and Speech Communication

Technical University Graz

Multi User Detection II

Robin Hofe, 98 30 579

1. Introduction/Overview

The second part of 'Multi User Detection' deals with optimum detection theory. The derived optimum detectors are used as a theoretical bound and as a quality criterion for any real implementation of a multi user detector.

First, the two different theoretical approaches towards an optimal detector will be introduced. The mathematical solutions the different approaches yield will be examined assuming at first the synchronous channel model, then the asynchronous channel model.

To obtain a useful quality criterion for real implementations, we will have to deal with the performances of the derived optimal detectors at the end of this documentation.

1. 1 Goals and Main Principles of Optimal User Detection

The aim is to derive the minimum probability of error (or Bit Error Rate = *BER*), a realization of a multi user detector can ever achieve. This server as a baseline for quality assessment.

In the beginning of multi user detection, the matched filter introduced in 'Multi User Detection I' by Thomas Blocher was thought to be the optimal detector. This is true for a detector that only knows the signal sent by the one sender it wants to detect and its code.
The theoretically best detector could have some additional information: it could know all codes of all senders and the correlations between them, the amplitudes and timings of all senders as well as the noise level and so have additional sources of information that lead to a better BER.

The two main approaches to an optimum detector are the 'individually optimum detector' and the 'jointly optimum detector'. Usually, the latter is referred to as the 'optimum detector'. This neglects the fact that for a single user, the individually optimum detector yields better results. The reason is that the jointly optimum approach is more elegant and leads to suitable results. For an arbitrary number of users, the individually optimum detector does not lead to a closed-form solution.

It might feel strange that there should be any differences between the optimal solution for a single user (individually) and for all users (jointly) if we just look at the actual solution for one user. The following example shows that there is in fact a difference:

Consider the a posteriory probabilities (APP) for two users, each transferring one bit of information:

$$P[(+1,+1) \mid \{y(t), 0 \leq t \leq T\}] = 0.26$$
$$P[(-1,+1) \mid \{y(t), 0 \leq t \leq T\}] = 0.26$$
$$P[(+1,-1) \mid \{y(t), 0 \leq t \leq T\}] = 0.27$$
$$P[(-1,-1) \mid \{y(t), 0 \leq t \leq T\}] = 0.21$$

The jointly optimum decision would yield (+1, -1) as the probability is highest: 0,27.
The individually optimum decision would lead to (+1, +1) with the probabilities 0,53 and 0,52.

2. 1 Jointly Optimum Detector in the Synchronous Channel

The received signal at the detector assuming two senders can be expressed as the following equation:

$$y(t) = A_1 b_1 s_1(t) + A_2 b_2 s_2(t) + \sigma n(t), t \in [0, T]$$

y(t) contains the amplitudes, information bits and signature waveforms of both senders as well as the noise of the channel described by its standard deviation σ.

For two users, each transferring one bit, there are four equiprobable hypotheses:

$$x_1 = A_1 s_1 + A_2 S_2$$
$$x_2 = A_1 s_1 - A_2 S_2$$
$$x_3 = -A_1 s_1 + A_2 S_2$$
$$x_4 = -A_1 s_1 - A_2 S_2$$

This is a standard hypothesis testing problem. As shown in the 'Fundamentals of Detection Theory' papers of this seminar, the pair (b_1, b_2) is needed, which maximizes:

$$f[y(t), 0 \leq t \leq T | (b_1 b_2)] =$$
$$\exp(-\frac{1}{2\sigma^2} \int_0^T [y(t) - A_1 b_1 s_1(t) - A_2 b_2 s_2(t)]^2 dt)$$

This expression can be transformed:

$$f[y(t), 0 \leq t \leq T | (b_1 b_2)] =$$
$$\exp(-\frac{1}{\sigma^2} \Omega_2(b_1, b_2)) exp(-\frac{A_1^2 + A_2^2}{2\sigma^2}) exp(-\frac{1}{2\sigma^2} \int_0^T y^2(t) dt)$$

$$\Omega_2(b_1, b_2) = b_1 A_1 y_1 + b_2 A_2 y_2 - b_1 b_2 A_1 A_2 \rho$$

Maximizing this expression means maximizing Ω_2, as it is the only term dependant on the pair (b_1, b_2).

If $\min\{A_1|y_1|, A_2|y_2|\} \geq A_1 A_2 |\varrho|$, which means that each of the signals sent by the two users is stronger than the cross correlation between them, the maximizing pair (b_1, b_2) is the same as in the matched filter case (with y_1 and y_2 being the outputs of the matched filters):

$$\hat{b}_1 = sgn(y_1)$$
$$\hat{b}_2 = sgn(y_2)$$

If one of the signal amplitudes is far bigger than the other, then (depending of the amount of cross correlation) the upper unequation is no longer fulfilled.

In that case the maximizing pair will be:

$$\hat{b}_1 = sgn(A_1 y_1 - sgn(\rho) A_2 y_2)$$
$$\hat{b}_2 = sgn(A_2 y_2 - sgn(\rho) A_1 y_1)$$

Figure 2.1 shows a possible implementation for such a detector:

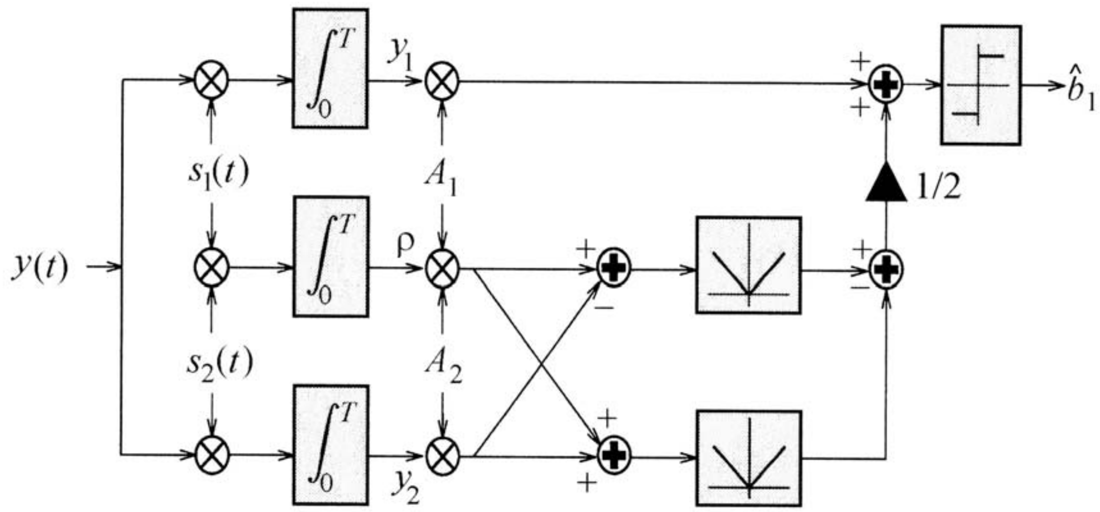


Figure 2.1: Jointly Optimum Detector for Two Users in a Synchronous Channel [1]

The left side of Figure 2.1 shows the matched filters on top and bottom and the calculation of the cross correlation in the middle. On the right side, the influence of both cross correlation and signal amplitudes is visible.

It is easier to understand this behaviour of the jointly optimum detector, when looking at the decision regions that derive from the above solution.

Figure 2.2 shows the decision regions for a two user synchronous channel using a matched filter detector. On the axes, there are the matched filter outputs y_1 and y_2. The signals s_1 and s_2 themselves show a cross correlation and are therefore not exactly orthogonal to each other. The greater the cross correlation would be, the more the angle between the signal vectors would differ from 90°.

The constellation in Figure 2.2 leads to a correct decision as the amplitudes of both signals are equal.
In Figure 2.3 the matched filter detector leads to a wrong decision as the amplitude A_1 dominates. This leads to a peculiar situation: only the noise present in the received signal could lead to a correct detection of the signal sent by the second user.

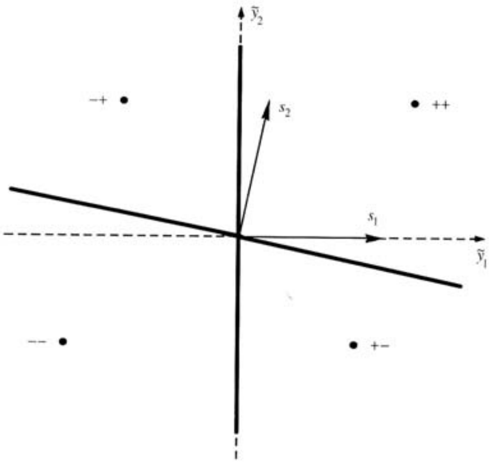


Figure 2.2: Decision Regions for a Two-User Matched Filter Detector ($A_1 = A_2$) [1]

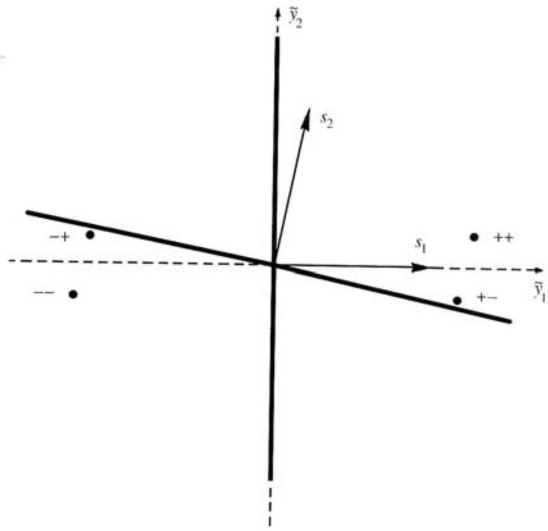


Figure 2.3: Decision Regions for a Two-User Matched Filter Detector ($A_1 = 6A_2$) [1]

Figures 2.4 and 2.5 show how the jointly optimum detector deals with a dominating amplitude:

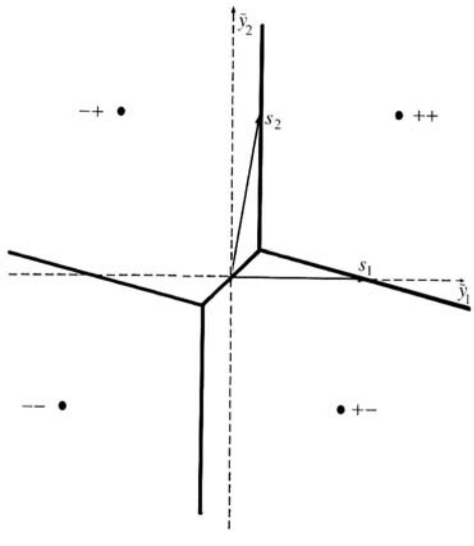


Figure 2.4: Figure 2.3: Decision Regions for a Two-User Jointly Optimum Detector ($A_1 = A_2$) [1]

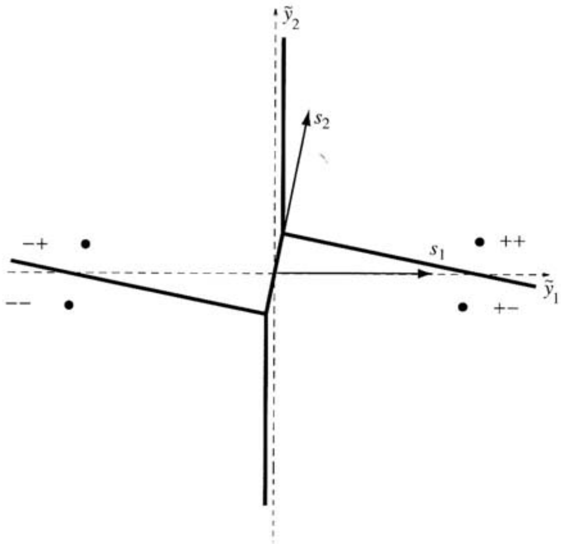
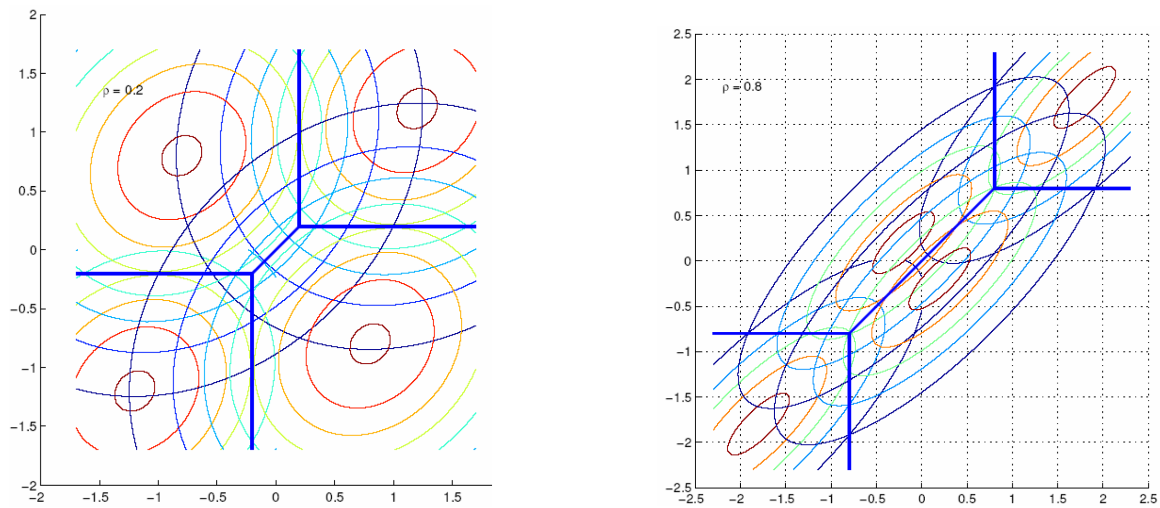


Figure 2.5: Figure 2.3: Decision Regions for a Two-User Jointly Optimum Detector ($A_1 = 6A_2$) [1]

Figures 2.6 and 2.7 show the dependency of the decision regions of a jointly optimum two-user detector on the cross correlation of the two signals:



Figures 2.6 (left, low cross correlation ϱ=0.2) [2] and 2.7 (right, high cross correlation ϱ=0.8) [2]

Advancing from the two-user case to a system with an arbitrary number of users, we have to expand our previous mathematical model in the following way:

The signal in the k-user channel becomes

$$y(t) = \sum_{k=1}^{K} A_k b_k s_k(t) + \sigma n(t), t \in [0, T]$$

$$\mathbf{b} = [b_1, ..., b_k]^T$$

The whole vector **b**, not only the previous pair, has to be jointly optimally demodulated. In the same way as before, **b** has to maximize

$$\exp(-\frac{1}{2\sigma^2}\int_0^T [y(t) - \sum_{k=1}^{K} b_k A_k s_k(t)]^2 dt)$$

As in the two-user case, this leads to the maximization of a parameter Ω:

$$\Omega(\mathbf{b}) = 2\mathbf{b}^T\mathbf{A}\mathbf{y} - \mathbf{b}^T\mathbf{A}\mathbf{R}\mathbf{A}\mathbf{b}$$

R is the cross correlation matrix, **A** is a matrix which has the amplitudes as the diagonal elements. The **y**-vector contains the matched filter outputs.

The time complexity for an algorithm solving the above equation is $O(2^K)$. That means that the number of steps, the algorithm needs to calculate the solution grows exponentially with the number of users K.

It is though possible to design algorithms with a better time complexity behaviour, if certain restrictions are imposed on the **R**-matrix.

For example, the minimum capacity cut algorithm described in [3] reduces the exponential time complexity to a polynomial one: $O(K^3)$.

2. 2 Individually Optimum Detector in the Synchronous Channel

This detector has to maximize the symbol-wise APP of b_k, $P(b_k | y(t))$:

$$P(b_k|y(t)) = \sum_{\mathbf{b}_{\bar{k}}} P(b_k, \mathbf{b}_{\bar{k}}|y(t))$$

A dash on a k indicates a bit transmitted by a different (not the k^{th}) user. As shown earlier in the example in 1.1, every possible state **b** can take has to be considered.

Again, we face a maximization problem. The solution to the two-user case is, without further proof:

$$\hat{b}_1 = sgn[y_1 - \frac{\sigma^2}{2A_1} ln\frac{cosh(\frac{1}{\sigma^2}(A_2y_2 + A_1A_2\rho))}{cosh(\frac{1}{\sigma^2}(A_2y_2 - A_1A_2\rho))}]$$

Figure 2.8 shows a possible implementation for this detector. It looks similar to the jointly optimum detector shown in Figure 2.1 except that the parts forming the absolute values in Figure 2.1 are replaced by nonlinearities in Figure 2.8.

This nonlinearity is expressed by the following function:

$$f_\sigma(x) = \sigma^2 log(cosh(\frac{x}{\sigma^2}))$$

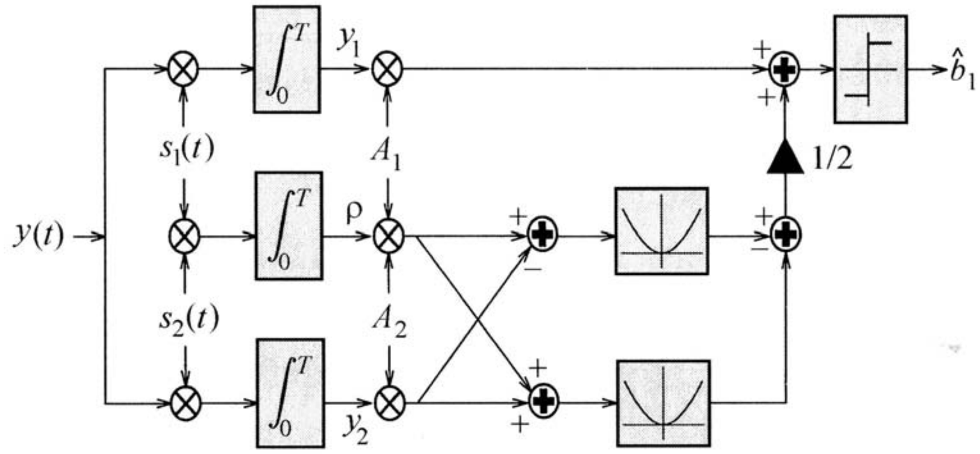


Figure 2.8: Individually Optimum Detector for Two Users in a Synchronous Channel [1]

Since

$$\lim_{\sigma \to 0} f_\sigma(x) = |x|$$

it is clear that for increasing Signal-to-Noise Ratio (SNR), the individually optimum detector converges towards the jointly optimum detector.

On the other hand, the nonlinearity in the individually optimum detector prevents a closed-form solution for an arbitrary number of users. That is why the jointly optimum detector usually is referred to as the 'optimum detector'.

Figure 2.9 shows the decision regions for an individually optimum detector. As we are only interested in one of the two users, there is just one borderline in the picture.

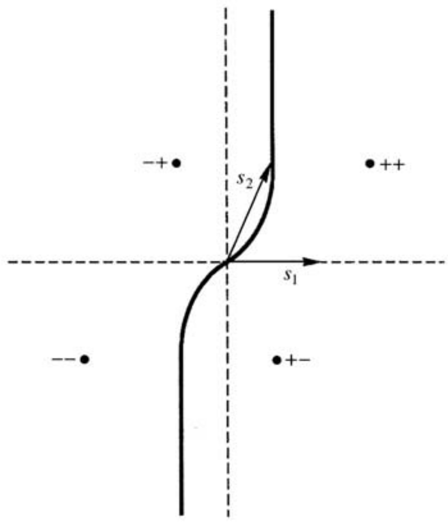


Figure 2.9: Decision Regions for the Individually Optimum Detector, Two-User Case [1]

In figure 2.10 we can see the convergence of the individually optimum detector towards the jointly optimum with increasing SNR.

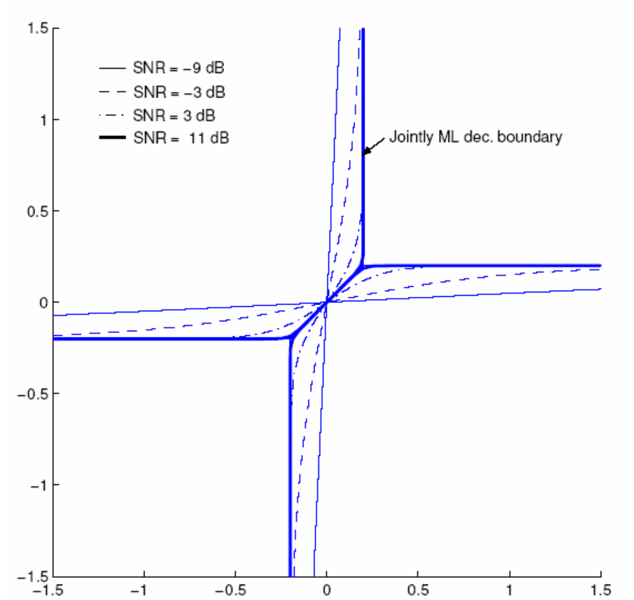


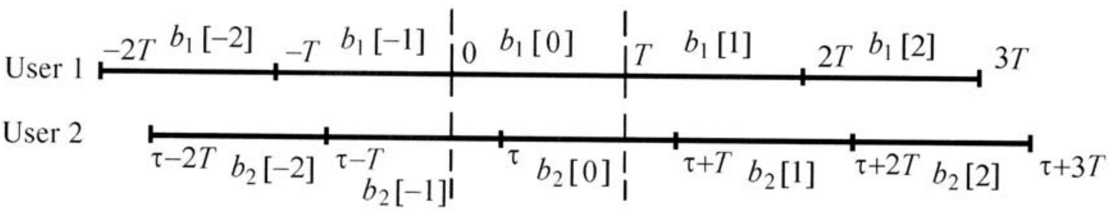
Figure 2.10: Convergence of the Individually Optimum Detector with growing SNR

The decision regions shown here is only one way to process the received data stream. The decision made is a so-called 'hard decision'.

'Soft decision' detectors also store some information about the certainty of the decision. If there is additional information about the signal sent, e.g. when a coding algorithm has been used that assures a certain ratio of zeros and ones transmitted, further assessment of the received bit stream is possible.

2. 3 Optimum Detectors in an Asynchronous Channel

A bit stream transmitted by two users in an asynchronous channel could look like this:



The optimum detector, as defined earlier, knows everything a detector could possibly know in a real system. To reach an optimum decision about $b_1[0]$ in the case shown in the example, the detector would need information about $b_2[0]$ and $b_2[-1]$. But for b2[0], information about $b_1[-1]$ would be needed as well.

That leads to the conclusion that the optimum detector in an asynchronous channel needs to decode all the bits in one frame at once.

2. 4 Jointly Optimum Detector in the Asynchronous Channel

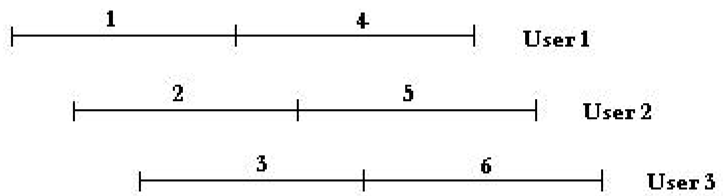
As shown in 'Multi User Detection I', a K-user asynchronous channel in which every user transmits M+1 bits per frame (M-frame) can be treated like a 1-user synchronous channel with inter-symbol interference (ISI) or as a K(M+1)-users synchronous channel in which every user transmits only one bit (without ISI).

As the optimal detector has to deal with the whole frame at once, it does not matter which of those models we use. Important is the influence of both models on the **R**-matrix. In the case with ISI, **R** contains the correlations due to the ISI. In the K(M+1)-user case, **R** holds the cross correlation between the users.

As can be seen in the example in 2.3, some elements of **R** are 0. There is, for instance, no interference between $b_1(-1)$ and $b_2(1)$. Also, none of the b_1-bits interfere with each other.

The **R**-matrix therefore has a banded structure. The following example shows, how this structure can be derived.

Example: 3 Users, each transferring 2 bits:



K=3, M=1 (2 bits transmitted per user)

b = [b1, b2, b3, b4, b5, b6]^T, b1=b1(0), b4=b1(1), etc.

A = diag(A1, A2,…, A6)

The correlation matrix for this example would be:

$$\mathbf{R} = \begin{pmatrix} 1 & \rho_{21} & \rho_{13} & 0 & 0 & 0 \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} & 0 & 0 \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} & \rho_{35} & 0 \\ 0 & \rho_{42} & \rho_{43} & 1 & \rho_{45} & \rho_{46} \\ 0 & 0 & \rho_{53} & \rho_{54} & 1 & \rho_{56} \\ 0 & 0 & 0 & \rho_{64} & \rho_{65} & 1 \end{pmatrix}$$

It can be easily found by just looking at the picture above: clearly, bits number 1 and 4 do not interfere, bit 3 interferes with bits 1, 2, 4 and 5 but not with bit number 6, and so on.

Now the mathematical model has exactly the same form as in the synchronous channel part before. That means that the same results can be used: The optimal solution has to maximize:

$$\Omega(\mathbf{b}) = 2\mathbf{b}^T\mathbf{Ay} - \mathbf{b}^T\mathbf{ARAb}$$

As stated in 2.1, the time complexity of the solution $O(2^K)$ can be improved imposing certain restrictions on **R**. These restrictions are given by the banded structure.

The optimal solution can be found using the Viterbi Algorithm. This algorithm was introduced in 'Maximum Likelihood Sequence Detection' by Klaus Dums during this seminar.

To apply the Viterbi Algorithm, $\Omega(\mathbf{b})$ has to be dependant only on a state vector $\mathbf{x}_j$ and b_j:

$$\Omega(\mathbf{b}) = \Sigma_{j=1}^{K(M+1)}\lambda_j(\mathbf{x}_j, b_j)$$

As $\Omega(\mathbf{b})$ can easily be transformed into

$$\Omega(\mathbf{b}) = 2\Sigma_{j=1}^{K(M+1)}A_j b_j y_j - \mathbf{b}^T\mathbf{ARAb}$$

we have to find an $\alpha_j(\mathbf{x}_j, b_j)$, so that

$$\lambda_j(\mathbf{x}_j, b_j) = A_j b_j y_j - \alpha_j(\mathbf{x}_j, b_j)$$

It is easy to see that

$$\alpha_j(\mathbf{x}_j, b_j) = \mathbf{b}^T\mathbf{ARAb}$$

$$\alpha_j(\mathbf{x}_j, b_j) = A_j b_j(A_j b_j + 2\overrightarrow{\rho}_j^T\Lambda_j\mathbf{x}_j)$$

if we define the following additional variables:

$$\overrightarrow{\rho}_j = [\rho_{j-2,j}, \rho_{j-1,j}]^T$$

$$\Lambda_j = diag(A_{j-2}, A_{j-1})$$

$$\mathbf{x}_j = [b_{j-2}, b_{j-1}]^T$$

Now that $\Omega(\mathbf{b})$ is split up into two parts, one only depending on b_j, the other only on $\mathbf{x}_j$, the Viterbi Algorithm can be used to find the optimal solution for the maximization problem.

Figure 2.11 shows the corresponding Trellis Diagram for a numerical example. The grey lines denote rejected paths, a solid line means a '1' bit, a dashed line a '0' bit.

The column on the left side of the figure (00/01/10/11) represent the different states $\mathbf{x}_j$ can take. $\mathbf{x}_j$ always consists of the two last b_j.

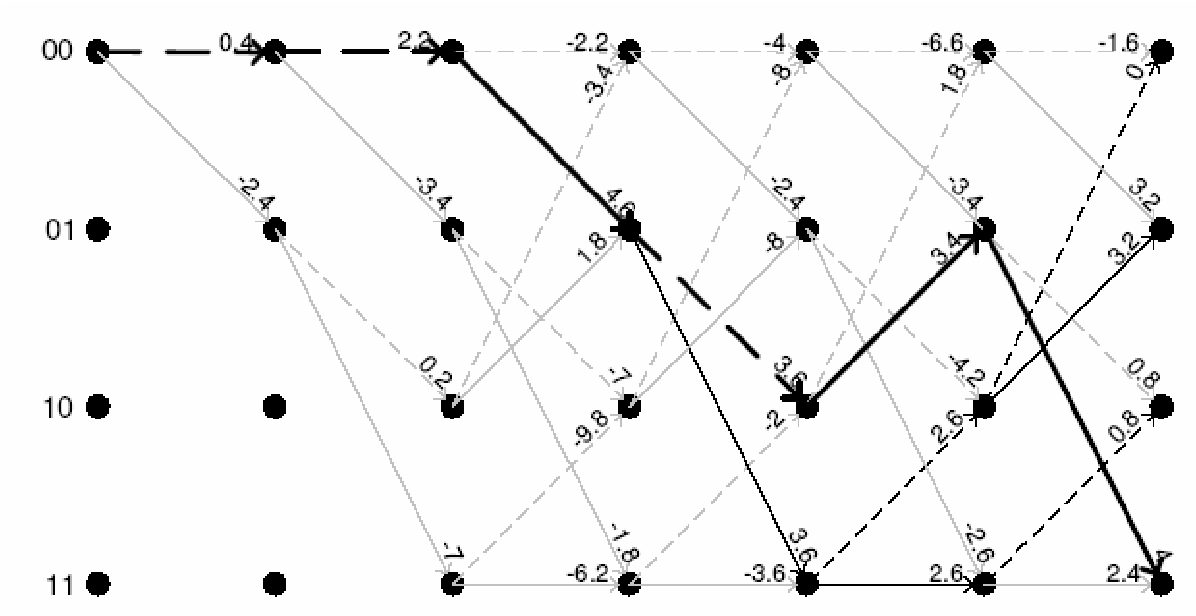


Figure 2. 11: Trellis Diagram for a numerical 3-User 1-Frame example

The optimal solution in this example would be $\mathbf{b}=[0\ 0\ 1\ 0\ 1\ 1]^T$.

So, user 1 has sent [0 0], user 2 [0 1] and user 3 [1 1].

3. 1 Performance of the Optimum Detector

In order to compare any real implementation with the derived optimum detector we have to know about the performance of the optimum detector in the first place. The performance criterion is the BER the detector can reach.

The minimum BER is delivered by the individually optimum detector. But this border is hard to obtain. Therefore, the optimal BER has to be approximated by an upper and by a lower bound.

As we saw earlier, the behaviour of the jointly optimum detector converges towards the individually optimum with growing SNR.
The BER of the jointly optimum detector serves as an upper bound for the BER we are looking for.

The lower bound is trickier: A better detector than the one we stated to be the best one possible has to be found. Luckily, we only need this super-detecoder theoretically. We can assume a 'genie-aided' decoder which has access to information that a real one could never obtain.

The genie should not give the detector too much information though, as we want the lower BER bound to be close to the one really achieved by the individually optimum detector.

A popular genie (for a two-user channel) is one that tells the receiver the value of b_1 whenever $b_1=b_2$.

This is what obtain for the upper and lower bounds:

$$P_1^m(\sigma) \leq Q(\frac{A_1}{\sigma}) + \frac{1}{2}Q(\frac{\sqrt{A_1^2 + A_2^2 - 2A_1A_2|\rho|}}{\sigma})$$

$$P_1^m(\sigma) \geq \max[Q(\frac{A_1}{\sigma}), \frac{1}{2}Q(\frac{\sqrt{A_1^2 + A_2^2 - 2A_1A_2|\rho|}}{\sigma})]$$

The Q-function was introduced in 'Multi User Detection I'. Figure 3.1 shows these boundaries together with the BER of the single-user matched filter.

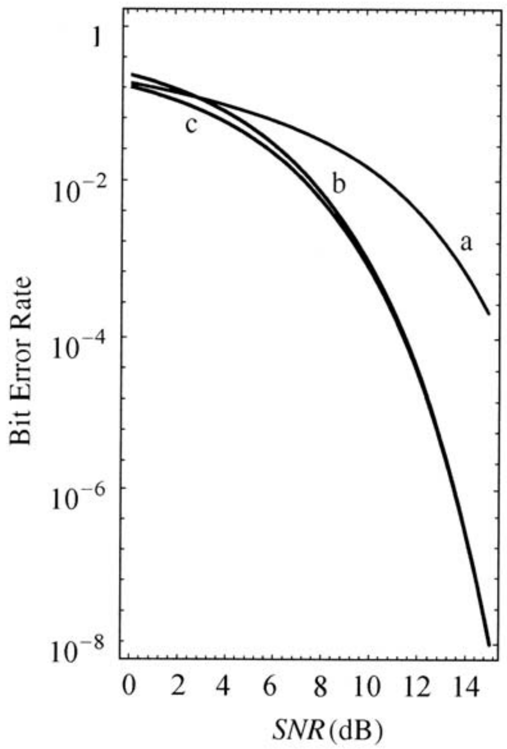


Figure 3.1: BER in a two-user channel, ϱ=0,4, $A_1=A_2$,
(a) Single-user matched filter
(b) Upper bound
(c) Lower bound

References

[1] ‚Multiuser Detection', Sergio Verdu, Cambridge University Press

[2] ‚Optimum Multiuser Detection', T. J. Lim, Notes for ECE1530S 2002/2003, Part III

[3] ‚Optimum Multiuser Detection is Tractable for Synchronous CDMA Systems Using M-Sequences', S. Ulukus, R. D. Yates, IEEE Communications Letter, Vol. 2, No. 4, April 1998