ERROR PERFORMANCE OF A CHANNEL OF KNOWN IMPULSE RESPONSE

Paul D. Teal — Industrial Research Limited, PO Box 31-310 Lower Hutt, New Zealand, p.teal@irl.cri.nz

Robert C. Williamson — Department of Engineering, Australian National University, Canberra, ACT, 0200, Australia

Rodney A. Kennedy — Telecommunications Engineering Group, RSISE, Australian National University, Canberra, ACT, 0200, Australia

ABSTRACT

It has long been known [1] that the performance of a channel with intersymbol interference (ISI) not only depends on the signal to noise ratio, but on the actual impulse response coefficients as well. In this paper, a technique for approximating the performance of a communications channel in the presence of ISI is presented. It is assumed that the finite length impulse response of the channel is known, and that a maximum likelihood sequence estimation (MLSE) technique is used for equalisation. The technique consists of identifying for the modulation and impulse response duration, those *few* error events which are most likely to occur. This selection process is computationally intensive, but once it has been performed for a given modulation, the results can be used to enable very efficient approximation. Results are presented for some common modulation schemes.

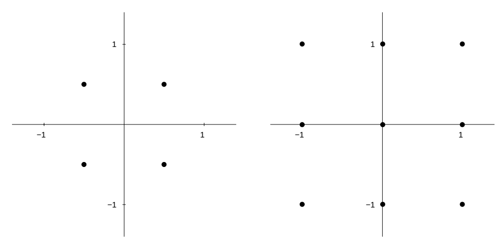
1. INTRODUCTION

THERE are situations in which it would be advantageous to be able to estimate the error performance of a communications system operating on a particular channel, using only knowledge of the channel. The performance of a system using a MLSE equalisation technique is, however, not a simple function of the impulse response; there are many possible error events which can contribute to the overall probability of error. Thus the performance estimate has not been easy to obtain. This paper presents a "reduced complexity" method for very quickly finding an accurate approximation to the exact error performance based on the few most likely error events.

The worst case performance has been investigated in [1–5]. These papers have focussed on finding the performance for the worst possible channel. This gives an approximation to the channel error rate where the channel is unknown. A much better approximation can be obtained in the case where the channel is known. To give one example only, the worst-case length 7 QPSK channel needs transmission power 11.4 dB greater than the best-case channel of the same length to achieve the same error rate. The gap is larger for longer impulse responses and for more spectrally efficient modulations. Some detail of this gap (and thus an improved performance approximation) can be obtained using the method treated here (although we make no claim that this is the *best* method).

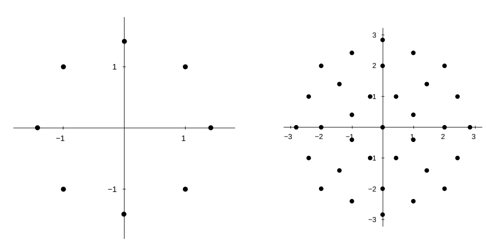
The particular situation requiring a performance estimate which motivates this paper is where the channel impulse response is being predicted in advance, so that there is sufficient time for corrective action of poor performance to be negotiated between the transmitter and receiver [6, 7]. This corrective action may for example take the form of change of time-slot (for a TDMA system), change of frequency (for a FDMA) system, change of power level, or change of coding scheme.

2. SYMBOL ALPHABET



(a) Channel Symbols (b) Difference Alphabet

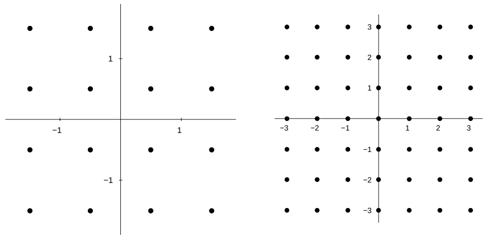
Figure 1: QPSK Symbols and Differences



(a) Channel Symbols (b) Difference Alphabet

Figure 2: 8-PSK Symbols and Differences

Several possible symbol alphabets used in this problem are shown in figures 1 to 3. These correspond to various linear modulation schemes in use in real systems. For each symbol alphabet there is a set of possible differences, arising where one member of the symbol alphabet is mistaken for another. This latter set is referred to here as the *difference alphabet*. Note that each element of the difference alphabet has a different *multiplicity*. For instance, the zero symbol will have multiplicity equal to the number of elements



(a) Channel Symbols (b) Difference Alphabet

Figure 3: 16-QAM Symbols and Differences

in the symbol alphabet. (This multiplicity is not required for calculating the distance of an error event, but is required for bit error rate calculations from the distance, particularly in the low signal to noise ratio case).

3. ERROR EVENTS

Once a difference alphabet $\mathbf{D}$ has been defined, a set $\mathbf{D}^R$ of possible *error events* of a particular length R can be defined. This is the set of all sequences

$$\boldsymbol{\epsilon} = \{\epsilon_1, \epsilon_2, \ldots, \epsilon_R\}, \quad \epsilon_k \in \mathbf{D}. \tag{1}$$

A particular channel will have a length L. This is the number of channel co-efficients (usually spaced by one symbol period) required to specify the impulse response. The channel $\mathbf{f}$ is specified by

$$\mathbf{f} = (f_1, f_2, \ldots, f_L). \tag{2}$$

Each of $f_k, k = 1, 2, \ldots, L$ can be complex, although there may be situations in which it is desirable to constrain them to be real [3, p109].
Define $\mathbf{e}$ to be a $(R + L - 1) \times L$ rectangular Toeplitz matrix with first column $(\epsilon_1, \epsilon_2, \ldots, \epsilon_R, 0, \ldots, 0)'$ and first row $(\epsilon_1, 0 \ldots 0)$,

$$\mathbf{e} = \begin{bmatrix} \epsilon_1 & 0 & \ldots & 0 \\ \epsilon_2 & \epsilon_1 & & \vdots \\ \vdots & \vdots & \ddots & \\ \epsilon_R & \vdots & & \vdots \\ 0 & \epsilon_R & & \vdots \\ \vdots & & \ddots & \epsilon_{R-1} \\ 0 & 0 & 0 & \epsilon_R \end{bmatrix} \tag{3}$$

4. DISTANCE AND ERROR PROBABILITY

The *distance* δ of a particular error event $\boldsymbol{\epsilon}$ at a particular channel is given by [5, p619]

$$\delta^2(\boldsymbol{\epsilon}, \mathbf{f}) = |\mathbf{ef}|^2 = \mathbf{f}^H \mathbf{A} \mathbf{f} \tag{4}$$

where the square Hermitian Toeplitz matrix $\mathbf{A}$ is defined as $\mathbf{A} = \mathbf{e}^H \mathbf{e}$. The superscript ^{H} represents Hermitian transpose.
Leaving out derivations that can be found in other references, (e.g.,[3]) the probability of error for a static AWGN channel can be found to have a bound given by a summation over all possible error events:

$$P_e \leqslant \sum_{\boldsymbol{\epsilon} \in \mathbf{D}^R} \overline{m}(\boldsymbol{\epsilon}) \frac{\operatorname{erfc}(\gamma \delta(\boldsymbol{\epsilon}, \mathbf{f}))}{2} \tag{5}$$

where

$\overline{m}(\boldsymbol{\epsilon})$	is	the *average multiplicity* and error weight of $\boldsymbol{\epsilon}$
γ	is	a multiplicative factor including the signal to noise ratio and some other properties of the modulation scheme
erfc	is	the complementary error function.

The very "short" tail of the Gaussian distribution results in the error probability being well approximated (for medium to high signal to noise ratios) by the single term of equation (5) which has the largest likelihood of occurrence. This single term can be found rapidly if it is known to belong to a small set $\mathbf{E}$ of error events.
The set $\mathbf{E}$ is the smallest set of error events which consists of those $\boldsymbol{\epsilon}$ so that for all $\boldsymbol{\epsilon}_N \notin \mathbf{E}$, and for all $\mathbf{f}$, there exists an $\boldsymbol{\epsilon}_E \in \mathbf{E}$ such that

$$\delta^2(\boldsymbol{\epsilon}_E, \mathbf{f}) \leqslant \delta^2(\boldsymbol{\epsilon}_N, \mathbf{f}). \tag{6}$$

Once such a set $\mathbf{E}$ of error events has been found, the error performance of the system can be easily approximated by finding the error rate caused by the most probable element of that set. The method of obtaining the performance estimate for a particular channel $\mathbf{f}$ is thus to evaluate equation 5 for the $\boldsymbol{\epsilon}_1 \in \mathbf{E}$ for which $\delta^2(\boldsymbol{\epsilon}_1, \mathbf{f})$ is smallest.

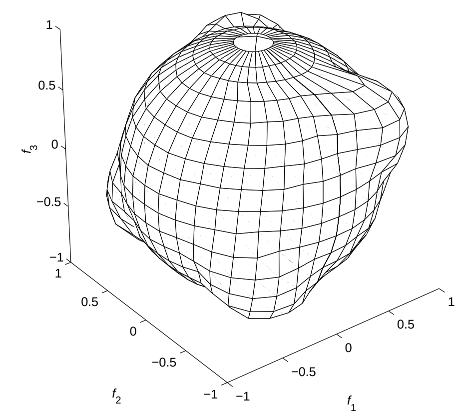


Figure 4: Iso-surface of BER for BPSK, and $L = 3$. $p(E) = 3 \times 10^{-5}$. The SNR of the best channel is 11 dB

An iso-surface of the exact error performance for the modulation Binary Phase Shift Keying (BPSK), and $L = 3$ is shown in figure 4. All points on the surface represent channels with a symbol error probability of 3×10^{-5}. An approximation to this surface based on just three error sequences is shown in figure 5
Note that for a fixed distance δ such as unity, equation 4 defines an ellipsoid in the space of impulse response coefficients $\mathbf{f}$. The

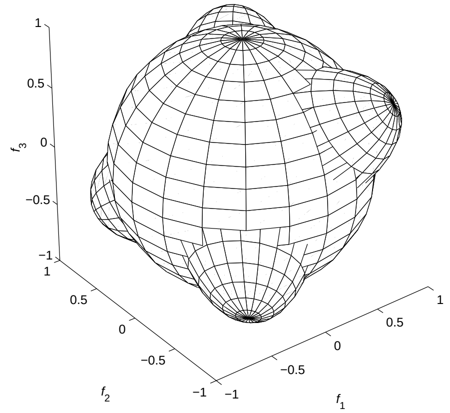


Figure 5: Union of all ellipsoids for BPSK difference alphabet and $L = 3$

axes of this ellipsoid are aligned with the eigenvectors of $\mathbf{A}$, and the lengths of the semi-axes are the inverse of the square roots of the eigenvalues of $\mathbf{A}$ (i.e., the inverse of the singular values of $\mathbf{e}$). Note that since $\mathbf{A}$ is Hermitian, its eigenvalues will be real. An error event consisting of a single symbol error defines a sphere.
The set $\mathbf{E}$ defined in equation 6 can be visualised as the minimum set of ellipsoids required to contain the union of *all* ellipsoids.

5. ALGORITHM FOR FINDING THE SET E

5.1. Stage 1

The first stage is, for a given length of error event R, to search through all possible sequences, and eliminate duplication of the matrix $\mathbf{A}$. There are several error sequences which will produce the same $\mathbf{A}$. For large R this is quite computationally intensive, particularly for large difference alphabets (e.g., 16-PSK has 129 elements in the difference alphabet). The computation time rises exponentially with R.
Sequences containing (but not ending with) a sequence of zeros of length $\geq L - 1$ need not be included. This is because after such a sequence of zero errors (i.e., correct transmissions), a MLSE equaliser will be in the all zero state, and so the behaviour is already covered by another sequence.
Restricting the search of error events to only those of length R or less may mean that a valid member of the set is omitted. Provided however that R is chosen to be sufficiently large, the inaccuracy in the performance approximation will be small [3, pp22,122].

5.2. Stage 2

The second stage of the algorithm is to eliminate all of the matrices $\mathbf{A}$ which *dominate* any others.
We use the following definition of positive semidefinite partial ordering from [8, p469]

Definition 1 *Let* $\mathbf{A}_1, \mathbf{A}_2$ *be* $n \times n$ *Hermitian matrices. We write* $\mathbf{A}_1 \succ \mathbf{A}_2$ *if the matrix* $\mathbf{A}_1 - \mathbf{A}_2$ *is positive definite.*

If $\mathbf{A}_1 \succ \mathbf{A}_2$, then for any $n \times m$ matrix $\mathbf{T}$, $\mathbf{T}^H\mathbf{A}_1\mathbf{T} \succ \mathbf{T}^H\mathbf{A}_2\mathbf{T}$. If $\mathbf{T}$ is a vector, $m = 1$, and the relation $\succ$ is simply $>$. The positive definiteness of $\mathbf{A}_1 - \mathbf{A}_2$ can be established by checking that all its eigenvalues are positive [8, p402].
The analogy with the ellipsoids is that the ellipsoid defined by $\mathbf{A}_1$ is entirely contained within that defined by $\mathbf{A}_2$, and so it is not a member of that minimum set required to define the union.
Since pairs of sequences are being compared in this stage, the computation time required here rises with the square of the number of matrices $\mathbf{A}(\boldsymbol{\epsilon})$ found in stage 1. Hence this stage is also quite computationally intensive. However, many of the matrices can be quickly eliminated by comparison with the largest sphere; if the smallest singular value is larger than the radius of the sphere, further search is not required.

5.3. Stage 3

Having eliminated all sequences which define ellipsoids which are entirely contained within others still leaves many others, which, while they are not contained by any other *single* ellipsoid, are still entirely contained by the union of all ellipsoids, and do not add to that union. An illustration of this behaviour is shown in figure 6.

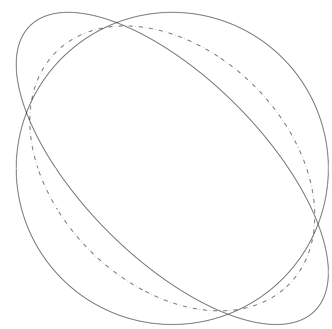


Figure 6: An ellipse (broken line) which is not contained by either of two others, and yet does not contribute to the union

The smallest set which *does* define the union is generally much smaller than the set of matrices which do not dominate any others. Hence the third stage of the algorithm is to perform further elimination so the smallest set required to define the union remains.
One method of performing this elimination is to find a set of points on the unit (hyper-) sphere in the space of $\mathbf{f}$, and perform a search over the surface of the sphere to find the sequence with the smallest distance δ at every point. This will find the smallest set required to define the union provided there isn't too much numerical inaccuracy, and the grid is fine enough. Unfortunately for dimensions greater than $\mathbb{R}_6$ or $\mathbb{C}_4$ the operation becomes too time consuming with the computational power available.
Another method which will eliminate some sequences, but perhaps not necessarily result in the smallest *possible* set, is to consider triples of error sequences. Consider that we have matrices $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{A}_3$ corresponding to three different error events $\boldsymbol{\epsilon}_1$, $\boldsymbol{\epsilon}_2$ and $\boldsymbol{\epsilon}_3$, and $\delta_i^2(\mathbf{f}) = \mathbf{f}^H\mathbf{A}_i\mathbf{f}$ for each of these. The intersections of the pairs $\delta_i^2 = \delta_j^2$ all exist since otherwise one of the matrices would have been eliminated in stage 2. If however
$\delta_1^2 < \delta_3^2$ for all the points of the intersection $\delta_2^2 = \delta_3^2$, *and*
$\delta_2^2 < \delta_3^2$ for all the points of the intersection $\delta_1^2 = \delta_3^2$, *and*
$\delta_3^2 > \delta_1^2$ for all the points of the intersection $\delta_1^2 = \delta_2^2$
then we can safely omit $\boldsymbol{\epsilon}_3$ from the set $\mathbf{E}$.
Each of the three requirements above can be established in the

following manner. For the third requirement, for example, we need to show that the minimum of $\mathbf{f}^H(\mathbf{A}_3 - \mathbf{A}_1)\mathbf{f} > 0$, subject to the constraint that $\mathbf{f}^H(\mathbf{A}_2 - \mathbf{A}_1)\mathbf{f} = 0$, and the constraint that $\mathbf{f}^H\mathbf{f} = 1$. Using the method of Lagrange multipliers, we wish to find points at which the derivative of augmented function

$$\begin{aligned}\frac{\partial \zeta}{\partial \mathbf{f}} &= \mathbf{f}^H(\mathbf{A}_3 - \mathbf{A}_1)\mathbf{f} \\ &\quad -\nu_1\left(\mathbf{f}^H(\mathbf{A}_2 - \mathbf{A}_1)\mathbf{f}\right) - \nu_2\left(\mathbf{f}^H\mathbf{f}\right) \\ &= 0. \end{aligned} \tag{7}$$

The solution to this can be obtained by finding the value of ν_1 for which the eigenvectors of $(\mathbf{A}_3 - \mathbf{A}_1) - \nu_1((\mathbf{A}_2 - \mathbf{A}_1))$ satisfy $\left(\mathbf{f}^H(\mathbf{A}_2 - \mathbf{A}_1)\mathbf{f}\right) = 0$. If the matrix $\mathbf{A}_2 - \mathbf{A}_2$ is first diagonalised, and $\mathbf{A}_3 - \mathbf{A}_1$ correspondingly transformed, a solution can be obtained fairly readily.

6. RESULTS

This section presents a few of the sets $\mathbf{E}$ found using the algorithm described in this paper.

Note that there are many sequences which "touch" these sets at one or more points, and the sets specified below may not even be the smallest *possible* sets. For the complex symbol sets, $j = \sqrt{-1}$.

6.1. BPSK $L = 3, R \leqslant 15$

$$\begin{aligned}\boldsymbol{\epsilon}_1 &= (1) \\ \boldsymbol{\epsilon}_2 &= (1, 1) \\ \boldsymbol{\epsilon}_3 &= (1, -1)\end{aligned}$$

6.2. BPSK $L = 4, R \leqslant 15$

In addition to those for L=3:

$$\begin{aligned}\boldsymbol{\epsilon}_4 &= (1, 1, 1) \\ \boldsymbol{\epsilon}_5 &= (1, -1, 1)\end{aligned}$$

6.3. BPSK $L = 5, R \leqslant 15$

In addition to those for L=4:

$$\begin{aligned}\boldsymbol{\epsilon}_6 &= (1, 0, 1) \\ \boldsymbol{\epsilon}_7 &= (1, 0, -1) \\ \boldsymbol{\epsilon}_8 &= (1, 1, 1, 1) \\ \boldsymbol{\epsilon}_9 &= (1, -1, 1, -1)\end{aligned}$$

6.4. QPSK $L = 3, R \leqslant 8$

$$\begin{aligned}\boldsymbol{\epsilon}_1 &= (1) \\ \boldsymbol{\epsilon}_2 &= (1, 1) \\ \boldsymbol{\epsilon}_3 &= (1, -1) \\ \boldsymbol{\epsilon}_4 &= (1, j) \\ \boldsymbol{\epsilon}_5 &= (1, -j)\end{aligned}$$

6.5. QPSK $L = 4, R \leqslant 8$

In addition to those for L=3:

$$\begin{aligned}\boldsymbol{\epsilon}_{10} &= (1, j, -1) \\ \boldsymbol{\epsilon}_{11} &= (1, 1, 1) \\ \boldsymbol{\epsilon}_{12} &= (1, -j, -1) \\ \boldsymbol{\epsilon}_{13} &= (1, -1, 1)\end{aligned}$$

7. CONCLUSIONS

This paper has demonstrated the feasibility of approximating the error performance of a system operating over an ISI impaired channel and utilising a MLSE equaliser. The efficiency of the approximation technique lies in knowing *a priori* the set of error events which contribute most to the error rate. The approximation is obtained by evaluating equation (5) for this set. An algorithm for finding these has been presented along with some results of its application.

Further work is required to develop a more efficient method of finding the set $\mathbf{E}$. Each of the steps of the algorithm can probably be considerably refined.

For instance, the range of R which can be searched could be considerably extended if the following conjecture is true.

Conjecture 1 *All of the members $\boldsymbol{\epsilon}$ of the set $\mathbf{E}$ defined by equation (6) are symmetric in absolute value. That is, if $\boldsymbol{\epsilon} = \{\epsilon_1, \epsilon_2, \ldots, \epsilon_R\}$, then $|\epsilon_i| = |\epsilon_{R-i+1}|$ for $i \in \{1, 2, \ldots, R\}$.*

The conjecture appears to be true for all the results obtained so far, but a proof has not yet been obtained.

8. REFERENCES

[1] F. Magee and J. Proakis, "An estimate of the upper bound on error probability for maximum-likelihood sequence estimation on channels having a finite-duration impulse response," *IEEE Trans. Information Theory*, vol. IT-19, pp. 699–702, Sep 1973.

[2] R. Anderson and G. Foschini, "The minimum distance for mlse digital data systems of limited complexity," *IEEE Trans. Information Theory*, vol. IT-21, pp. 544–551, Sep 1975.

[3] T. Larsson, *A State-Space Partitioning Approach to Trellis Decoding*. PhD thesis, Chalmers University of Technology, Göteberg, Sweden, 1991. Technical Report No. 222.

[4] W. Ser, K.-C. Tan, and K.-C. Ho, "A new method for determining "unkown" worst-case channels for maximum-likelihood sequence estimation," *IEEE Transactions on Communications*, vol. 46, pp. 164–168, Feb 1998.

[5] J. G. Proakis, *Digital Communications*. New York: McGraw Hill, 2nd ed., 1989.

[6] J. B. Andersen, J. Jensen, S. H. Jensen, and F. Frederiksen, "Prediction of future fading based on past measurements," in *VTC-99/FALL, Delft, The Netherlands*, Sep 1999.

[7] T. Eyceoz, A. Duel-Hallen, and H. Hallen, "Deterministic channel modelling and long range prediction of fast fading mobile radio channels," *IEEE Communications Letters*, vol. 2, pp. 254–256, Sep 1998.

[8] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 1991.