Design and Analysis of Digital Receivers

PER ÖDLING
**Doctoral Thesis 1995:165 D**

**Design and Analysis of Digital Receivers**

Per Ödling

Luleå University of Technology  
Division of Signal Processing  
S-971 87 Luleå, Sweden

Telephone: Int.+46 920 91135  
Telefax: Int.+46 920 72043  
E-mail: per@sm.luth.se

---

**Dissertation**

for the degree of Doctor of Philosophy (Ph.D.) in the subject area of *Signal Processing*, which, with due permission of the Faculty Board at Luleå University of Technology, will be defended in public, in Swedish, in LKAB-salen, room α117 in the Alpha-building at Luleå University of Technology, on Thursday the 1\textsuperscript{th} of June 1995, at 13\textsuperscript{15} am.

---

**Akademisk avhandling**

för avläggande av teknisk doktorsexamen inom ämnesområdet *Signalbehandling*, som med vederbörligt tillstånd av Tekniska Fakultetsnämnden vid Tekniska högskolan i Luleå kommer att offentligen förvaras, i LKAB-salen, α117 i Alfa-huset vid Tekniska Högskolan i Luleå, torsdagen den 1 Juni 1995, kl. 13\textsuperscript{15}.

---

**Supervisor / Handledare**  
Professor Per Ola Börjesson, Luleå University of Technology

**Faculty opponent / Fakultetsopponent**  
Professor Tor Aulin, Chalmers University of Technology

**Examination committee / Betygsnämnd**  
Professor Göran Salomonsson, Lund Institute of Technology  
Professor Åke Wernersson, Luleå University of Technology  
Professor Mats Viberg, Chalmers University of Technology
Design and Analysis
of
Digital Receivers

Per Ödling

Luleå University of Technology
Division of Signal Processing
Luleå, Sweden

May 1995

Supervisor
Professor Per Ola Börjesson, Luleå University of Technology

Faculty opponent
Professor Tor Aulin, Chalmers University of Technology
Abstract

This thesis consists of a summary and nine included papers, grouped into three parts. There is one journal paper, three reports written in the style of articles, four conference papers and one paper submitted to a conference.

The thesis proposes and investigates a number of digital receivers, especially receivers based on the maximum a posteriori and maximum likelihood criteria. Signal processing methods and models are developed and applied to a number of estimation and detection problems in systems with time dispersion and additive Gaussian noise. Digital receivers in two application areas are investigated: telecommunications and ultrasonic distance measurements.

Within telecommunications, particular attention is given to block transmission systems, where digital data is transmitted in independent blocks. With a geometric approach and reflecting on properties of the binary hypercube, it is shown that the minimum bit-error probability receiver (OBER) becomes the maximum likelihood sequence detector (MLSD) when the expected SNR used for designing the OBER goes to infinity. Likewise, the OBER reduces to the whitened matched filter with hard decisions in the limit when the expected SNR decreases. Furthermore, a novel detector is developed that makes MLSD-decisions on scattered bits in a block. This low-complexity detector can, if combined with a sub-optimal receiver such as a linear or decision-feedback equalizer, substantially reduce the system bit-error rate. Finally, using the geometric approach, the genie-aided detector, a device proposed by Forney for deriving performance bounds, is reconsidered and augmented with an explicit statistical description of the side information. This renders a more flexible tool, new performance bounds, and gives an instructive view on earlier work.

Reduced complexity Viterbi detection is addressed by means of combined linear-Viterbi equalizers. These equalizers reduce the complexity of the Viterbi detector, a structure for implementing the MLSD, by linear pre-equalization of received data and by giving the Viterbi detector a truncated channel model. Three receivers in this class are introduced, one of which is intended for multiple-antenna reception in block transmission systems.

In the field of ultrasonics, the problem of estimating the time-of-flight of an ultrasonic pulse is addressed under the assumption that the pulse has been distorted by an unknown, linear and time-dispersive system and by additive Gaussian noise. Two approaches for taking the linear distortion into account are presented. Both assume that the transmitted pulse is known and narrowband. Although the intended application is distance estimation using the ultrasonic pulse-echo method, the assumed basis of the time-of-flight estimation problem is more general: a known narrowband waveform is transmitted through a dispersive, linear system with additive Gaussian noise.
Contents

Preface vi

Acknowledgements vii

Thesis Summary ix

Corrections xxiii

Bibliography xxiv

Part I: Optimal Receivers

Part I.1: Representations for the Minimum Bit-Error Probability Receiver for Block Transmission Systems with Intersymbol Interference Channels
Per Ödling, Håkan Eriksson, Timo Koski and Per Ola Börjesson
A shorter version has been accepted for presentation at the 1995 IEEE International Symposium on Information Theory, Whistler, Canada, September 1995.

Part I.2: A Fast, Iterative Detector Making MLSD Decisions on Scattered Bits
Per Ödling and Håkan Eriksson
Research Report TULEA 1995:12, Division of Signal Processing, Luleå University of Technology.

Part I.3: A Genie-Aided Detector Based on a Probabilistic Description of the Side Information
Håkan Eriksson, Per Ödling, Timo Koski and Per Ola Börjesson
A shorter version has been accepted for presentation at the 1995 IEEE International Symposium on Information Theory, Whistler, Canada, September 1995.
Part II:  Low Complexity Receivers

Part II.1: A Reduced Complexity Viterbi Equalizer Used in Conjunction with a Pulse Shaping Method
Per Ödling, Timo Koski and Per Ola Börjesson

Part II.2: Combined Linear-Viterbi Equalizers – A Comparative Study and A Minimax Design
Nils Sundström, Ove Edfors, Per Ödling, Håkan Eriksson, Timo Koski and Per Ola Börjesson

Part II.3: Multiple-Antenna Reception and a Reduced-State Viterbi Detector for Block Transmission Systems
Håkan Eriksson and Per Ödling
Submitted to the 4th International Conference on Universal Personal Communications (ICUPC'95), Tokyo, Japan, November 1995.

Part II.4: Implementation of a System for Validation of Algorithms Used in Digital Radio Communication Schemes
Mikael Isaksson, Roger Larsson and Per Ödling

Part III: Robust Receivers for Ultrasound

Part III.1: A Robust Correlation Receiver for Distance Estimation
Håkan Eriksson, Per Ola Börjesson, Per Ödling and Nils-Gunnar Holmer

Part III.2: Simultaneous Time of Flight and Channel Estimation Using a Stochastic Channel Model
Håkan Eriksson, Per Ödling and Per Ola Börjesson
Preface

Writing a Ph.D. thesis is like digging a hole in the ground. My analogy is this: research is picking and hacking to loosen the soil. Scientific documentation of results is shovelling and throwing material out of the hole. How much soil one piles up is how much knowledge one contributes, with each paper written being one more swing of the shovel.

I became a Ph.D. student at the Division of Signal Processing at Luleå University of Technology in November 1989, after finishing my Master's degree and a subsequent short period of travelling. The thesis work presented here has been performed at the Division of Signal Processing, with many weeks at the Department of Signal Processing, the University College of Karlskrona/Ronneby and two one-quarter stays at the Information Systems Laboratory (ISL), Stanford University.

I spent the first three years researching in a group that included essentially me and my two advisors, Professor Per Ola Börjesson and Docent Timo Koski, in a field where no research had previously been performed at this university. Being a part of building up a field of research from scratch is a demanding assignment for one single Ph.D. student, a fact I didn't realize at the time. After three and a half years, in March 1993, I was awarded the Licentiate degree. This was to be a turning point in my way of working. One year earlier I had begun hoping for extending the group and in early 1993 my friend and fellow graduate student Håkan Eriksson joined. For six months Håkan and I split our time, working together on estimation theory in ultrasonics, and working individually writing our Licentiate theses. Then, for one and a half years, we worked intensely on one common project in statistical telecommunications, creating what I consider to be the core of this thesis.

There are two elements of style of this thesis that I'd like to comment upon. The first is that some of the papers are written in the style of engineering sciences and that some are written in a more mathematical style (in particular Part I.1 and Part I.3). Having an engineering background, we saw writing in a mathematical style as a part of our education and seized the excellent opportunity given by our cooperation with Timo Koski, Docent in applied mathematics.

The second element of style is our liberal choice of applications of signal processing and the way each report is individually optimized. Our philosophy has been to investigate and write reports about the best of our ideas, with the prospect of submitting these to journals. We have been shovelling and digging the most accessible soil. If, on the other hand, we had sought to optimize the thesis and not the individual reports, viewing the thesis as a single entity, we would have stressed completeness rather than novelty. We would have created, perhaps, a more nicely shaped hole, but shovelled out less soil. The thesis would have treated fewer topics, but more elaborately. This thesis is directed towards developing and documenting our best ideas.
Acknowledgements

This is a section that could easily take ten pages; there is so much to say about so many. Even with the effort I made in my Licentiate thesis, I was only able to thank a fraction of those that should have been mentioned. I cannot mention you all here, but I acknowledge your efforts and importance to me. You deserve all the praise you can be given and you are not forgotten. A few persons will reappear here, however.

I'd like to begin with expressing my most sincere gratitude towards my advisor Professor Per Ola Börjesson. A large portion of the good things I've become during the past five years, I attribute to him. (The things that were left out and all the bad stuff are my own fault. He did his best.) In addition to our professional relation, I also consider him one of my most precious and closest friends.

Docent Timo Koski, my friend and also my second advisor, has been very important for my education, results and daily work. I highly respect his skills as a researcher, but I esteem his personality even more.

My close friend and coauthor, Lic. Tech. Håkan Eriksson, is one of the most influential persons on my work. After working for three years alone in a small room with a large computer (as I'm given to overstate it), I approached my friend Håkan, the person I've worked the best with, and suggested a close cooperation. Fortunately, he had similar ideas. Our cooperation has been so productive and rich in ideas that I fear I will never find its equal.

The above three mentioned are the three definitively most important persons for my research and they are also coauthors of mine. I'd like to thank my other coauthors, too: Mr. Nils Sundström, Lic. Tech. Ove Edfors, Mr. Mikael Isaksson, Mr. Roger Larsson, and Professor Nils-Gunnar Holmer. It has been a pleasure to work with them all, and they have all been important for this thesis and to me.

Some others, although they were not coauthors, have had a great influence on both my development as a researcher and the creation of some of the papers: Dr. Anders Grennberg, Dr. Klas Ericsson, Dr. Tomas Nordström, Professor Sarah Kate Wilson, Mr. Paul A. Petersen, and Mr. Jan-Jaap van de Beek. A number of others have, without being directly involved in any of the papers, acted as role models and made important contributions to the development of my values, ethics and thinking: Dr. Harry Hurd, Professor Tor Aulin, Professor Mats Gyllenberg, Professor Lennart Elfgrén, Professor Lennart Karlsson, Dr. Gerard J. Hoffmann, Professor Göran Salomonsson, and Professor Lars-Erik Persson.

There are five large groups of persons that I'd like to thank: all of my friends (of which some are mentioned above and below); the staff and students at the Division of Signal Processing here in Luleå; the staff at Communication Systems (Ksu), Telia Research AB, Luleå (Hans Lundberg and his group); the staff and students at the Department of Signal Processing, the University College at Karlskrona/Ronneby; and the staff and students at the Information Systems Laboratory (ISL), Stanford University. Some of these individuals are mentioned by name in my Licentiate thesis, and I'd like to thank all of them again, too. If I were to add one single person here, it would be Professor Paulraj of ISL, Stanford, a person worthy of elaborate praise.
I'd also like to acknowledge the financial support from the Swedish National Board for Industrial and Technical Development (NUTEK); Telia Mobitel, Luleå (headed by Assar Lindkvist, Vice President, Upper Northern Region); Telia Research AB, Luleå; and the ISS’90 foundation, Stockholm.

Finally, there's one person that, although only occasionally participating in the academic discussions, has nevertheless been of great importance for all of this: my carefully chosen, Petra Deutgen. As many, including Petra, know, being a Ph.D. student doesn't promote family life, quite the opposite. I have learned many things from this. For instance: if you’re attending a party with a lot of academics, just plain ignore them. Talk to their spouses instead! Here you’ll find a group of very pleasant persons, with patience, tolerance, and an ability to entertain and to have a good time even in boring company. (Besides, they won’t recognize another academic’s awkwardness; they’ve already demonstrated that once.)
Thesis Summary

This thesis consists of this summary and nine included papers. There is one journal paper, three reports written in the style of articles, four conference papers and one paper submitted to a conference. Shortened versions of two of the reports have been accepted for presentation at conferences. The three reports either have been, or will be, submitted to journals. The included papers are grouped into three parts according to their intended application.

This thesis is mostly concerned with maximum likelihood (ML) and maximum a posteriori (MAP) estimation of parameters in signals that have been distorted by additive Gaussian noise and linear time dispersion [41]. The thesis is to some extent characterized by the application of ad hoc system models with optimal methods for these models. We have attempted to use models that are general to the behaviour of our subject; our focus is on clarifying the general principles of a problem. This is in contrast to directly applying ad hoc methods to a problem or searching for “optimal” models.

All the papers in Part I and Part II have a relation to the maximum likelihood sequence detector (MLSD)\(^1\) [9, 14, 19]. The MLSD is the detector that minimizes the probability of choosing the wrong sequence in a communication system transmitting equally probable sequences of data symbols. When the MLSD is employed in practical communication systems under the assumption of a linear channel with memory, it is often in the form of the Viterbi detector [18, 19], the complexity of which increases only linearly with the length of the sequence. Its complexity, however, increases exponentially with the memory-length of the channel. One approach for controlling this is discussed in Part II.1 – Part II.3.

\(^1\)The MLSD is defined in equation (2) below.
Part I: Optimal Receivers

The three reports in this part of the thesis discuss receivers that are optimal in the sense that their output, directly or indirectly, results from Bayesian decisions [41].

The reports are all concerned with block transmission systems [9, 15, 28], systems where digital data are transmitted in independent blocks. We assume that these are blocks of \( N \) equally probable, independent, identically distributed and antipodally modulated bits \( b \in \{+1, -1\}^N \), with \( \beta \) as the outcome of the stochastic variable \( b \). This facilitates geometrical interpretations using the binary hypercube \( C \triangleq \{+1, -1\}^N \) [29], defined as the set of all possible transmitted binary sequences of length \( N \). Our particular interest in the geometry of the binary hypercube is fundamental for the three reports Part I.1 – Part I.3, as well as for Part II.3.

The statistical block transmission system model, shown in Figure 1, is the model used in Part I.1 – Part I.3. This model, which we will refer to as the discrete-time, Gaussian channel (DTGC) with intersymbol interference (ISI) [9, 15, 28], is:

\[ y = Hb + n, \]  

where \( H \) is a deterministic and known \((N + L - 1) \times N\) matrix representing the ISI, \( n \sim N(0, \sigma_n^2 I) \) is a white, jointly Gaussian, zero-mean random noise vector with variance \( \sigma_n^2 \), and \( y \) is the \((N + L - 1) \times 1\) dimensional stochastic variable observed by the receiver. A variation on this model is to allow coloured noise with \( n \sim N(0, R_n) \) and \( R_n \) as the noise covariance matrix. The model (1) can, depending on the structure of \( H \) and \( R_n \), represent an arbitrary symbol-sampled linear system, including time-variant and time-invariant systems where matrix multiplication with \( H \) represents convolution with system impulse response. A definition of the MLSD for this model, which will be referred to often in the following text, is

\[ \hat{b}_{MLSD}(y) \triangleq \arg \min_{\beta \in C} \| y - H\beta \|^2, \]  

where \( \|x\|^2 = x^T x \) is the squared Euclidean norm of \( x \).

The signal-to-noise ratio (SNR) in this block transmission system with white noise, we define as

\[ \text{SNR} \triangleq \frac{\text{trace}\{H^TH\}}{N\sigma_n^2}. \]  

This is a block-SNR definition, which if \( H \) represents a time-invariant filter, will be equal to a definition of SNR often used in continuous transmission systems; the received energy for one transmitted bit over the noise variance. Note that our block-SNR definition does not specifically indicate whether or not the received signal contains information about all bits in the block. The block-SNR could be high despite some bits having completely faded, contributing no received energy.

The DTGC with ISI is a well-established model in communications, cf. [9, 15, 28] and the references therein. Although it is a discrete-time model, it accurately describes some communication systems. One example of such systems is the class of systems presented by Forney [19], where receivers contain a whitened matched filter.
with a symbol-rate sampler [3]. A whitened matched filter consists of three parts: an analog matched filter, a symbol-rate sampler and a discrete-time whitening filter [3, 19]. The output from the whitened matched filter is a sufficient statistic, whose properties are described by the model (1), for the transmitted sequence \( b \), assuming correct synchronization.

In Part II.3 we note that the DTGC with ISI in (1) can also model the output from a system with multiple receiving-antennas used in combination with a maximal-ratio combiner.

However, we are primarily interested in the DTGC with ISI not because it is a precise description of any particular group of receivers, but because it is a general description of many communication systems. Based on the basic model of a DTGC with ISI, we derive results that are relevant to a number of application-specific situations. Our choice of binary modulation is probably a greater restriction than our choice of channel model. All three of the reports in Part I are based on binary pulse amplitude modulation (PAM), i.e., \( b \in \{+1, -1\}^N \), using the geometry of a binary hypercube.

**Part I.1: Representations for the Minimum Bit-Error Probability Receiver for Block Transmission Systems with Intersymbol Interference Channels**

This report reconsiders the receiver that has optimal, i.e., minimal, bit-error probability [1, 2, 11, 23, 26], henceforth referred to as the OBER. It investigates the asymptotic properties of the OBER in the extreme cases where the noise variance used for designing the OBER approaches \(-\infty\) or 0.

Assuming the model (1) with binary antipodal modulation, the OBER for the detection of bit \( k \) is based on the following binary hypothesis:

\[
H_0 : \quad b \in C_k^- \\
H_1 : \quad b \in C_k^+,
\]

where \( C_k^- \) and \( C_k^+ \) are the halfcubes with \( (b)_{\text{bit}_k} = +1 \) and \( (b)_{\text{bit}_k} = -1 \), respectively. The resulting Bayesian test [41], which describes the operation of the OBER, is

\[
\frac{f_{y|H_1}(\eta|H_1)}{f_{y|H_0}(\eta|H_0)} \overset{H_1}{\gtrless} \frac{\Pr\{H_0\}}{\Pr\{H_1\}},
\]

where \( \eta \) is the outcome of the stochastic vector \( y \), \( f_{y|H_0}(\eta|H_0) \) is the probability density of \( y \) given the hypothesis \( H_0 \) and \( \Pr\{H_0\} \) is the a priori probability of
The OBER has to be designed for a specific SNR, as opposed to the MLSD. The reason for this is found in the way that the shape of the densities $f_{y|H_0}(\eta|H_0)$ and $f_{y|H_1}(\eta|H_1)$ changes with the (expected) noise variance $\sigma_n^2$. The shape of the densities $f_{y|b}(\eta|\beta)$ (associated with the MLSD) is more regular with respect to $\sigma_n^2$.

The output of the OBER is a block of bits $\hat{b}_{\text{OBER}} = [\hat{b}_1, \ldots, \hat{b}_N]^T$. As shown in the report, it can be calculated by

$$\hat{b}_{\text{OBER}}(y) \triangleq \text{sign} \left( \sum_{\beta \in \mathbb{C}_1} w(y, \beta, \sigma_n) \beta \right),$$

where $\text{sign}(\cdot)$ is a vector operation taking the sign of its argument componentwise, and

$$w(y, \beta, \sigma_n) = \psi(y|\beta, \sigma_n) \Pr\{b = \beta\} - \psi(y| - \beta, \sigma_n) \Pr\{b = -\beta\},$$

where $\Pr\{b = \beta\}$ is the a priori probability for the sequence $\beta$ being transmitted, cf. [23]. The density $\psi(y|\beta, \sigma_n)$ is the conditional density of $y$ given that the sequence $\beta$ was transmitted. The novel representation (6) of the OBER, a sum of Gaussian kernels, describes a parallel computing structure.

We visualize the behaviour of the OBER by plotting decision surfaces of the OBER, the MLSD and the matched filter for the decoding of two- and three-bit blocks. The decision surfaces are plotted in the source space [9] containing the binary hypercube [35]. The map from the space of received data to the source space is the block zero-forcing equalizer as described by Barbosa in [9], turning all ISI into coloured noise. Figure 2 shows an example of a decision surface of the MLSD for the decoding of a three-bit block. In Part I.1 we demonstrate that this is the same decision surface as for the OBER designed for an infinite SNR. The three-dimensional hypercube is plotted in the centre of the Figure.

With our geometric approach, we show that the OBER becomes the MLSD when the SNR used for designing the OBER goes to infinity; their decision regions become identical. Similarly, the OBER becomes the matched filter with hard decisions in the limit when the design-SNR decreases. It is well-established that the performances of the OBER and the MLSD become equal for high SNR [17, 19, 24, 34, 35]. (I conjecture that every receiver, for which the asymptotically dominating error events are confusing sequences at the minimum distance, will have a performance decaying with the same exponent, cf. [24] and the references therein.) The performances of, e.g., the OBER and the MLSD, have upper and lower bounds proportional to $Q(d_{\text{min}}/2)^2$ [17, 19, 34]. Our findings show that the OBER becomes the MLSD; it is as a consequence of this that their performances coincide for high SNR.

---

\(^2\)The $Q(\cdot)$-function is defined as $Q(x) \triangleq (1/\sqrt{2\pi}) \int_x^{+\infty} e^{-t^2/2} dt$. The minimum distance between any two sequences, $d_{\text{min}}$, is defined in Part I.2.
Part I.2: A Fast, Iterative Detector Making MLSD Decisions on Scattered Bits

This report proposes a detector that is based on the block transmission system model (1) with binary antipodal modulation. The detector is simple in structure and of low complexity; it consists of a matched filter and a threshold device with two variable thresholds per bit in the block, see Figure 3. It is characteristic of this detector that it is not generally able to give an estimate of all bits in a block; its quality is that the decisions it does make on detected bits are the same decisions an MLSD would have made. The output of the proposed detector is a mix of the output that an MLSD would have given and undetermined bits; the output is a block of $N$ ternary symbols, $+1$, $-1$ or "no decision".

\[
\begin{align*}
y & \xrightarrow{H^T R_n^{-1}} \text{Iterative threshold device} \quad \rightarrow \hat{b}
\end{align*}
\]

Figure 3: A block description of the proposed detector.

The detector's final output is calculated iteratively, with the thresholds recalculated in each iteration. The decisions of the first iteration of the detector are demonstrated in Figure 4, showing the decision regions of the MLSD and of the pro-
posed detector for an example with two-bit blocks. The decision regions are plotted in the source space [9] containing the binary hypercube. The decision region $O_1^0$ corresponds to “no decision” for this bit, in the first iteration. Figure 5 shows an example of the iterative calculation of the thresholds.

Figure 4: Examples of 2-dimensional decision regions in the source space containing the binary hypercube. The left part shows the decision regions of the MLSD and the right part the decision regions of the proposed detector, with the $[x_1, x_2]$ as coordinates in $\mathbb{R}^2$. The detection of the first bit in a two-bit block is considered.

Our detector’s most important property, perhaps, is its decision rate, the average percentage of bits on which a decision is made. The decision rate depends on the channel and the SNR, and is in many situations between 20% and 80%. The report contains simulations of the receiver for a two-tap channel, giving three views on the examples: the decision rate versus the SNR, the decision rate versus bit index, and the decision rate versus relative strength of the second tap in the two-tap channel. The decision rate is shown to decrease with stronger ISI and higher SNR. The bit-error rate and the distribution of the number of bits decoded are also investigated.

There are several ways of combining our detector with other receivers. Using the concept we advocate, we simulate the receiver in combination with both linear equalizers and decision-feedback equalizers (DFE). In our simulations, again with a two-tap channel, the performance of the combined receivers clearly surpasses the performance of the unsupported receivers.

For future work in this area, it would be desirable to determine the practical potential of our proposed receiver, preferably including work on other modulation techniques and more application-specific channel models.

Part I.3: A Genie-Aided Detector Based on a Probabilistic Description of the Side Information

It is often difficult to calculate the exact bit-error probability for receivers, e.g., the MLSD, operating in the presence of ISI. In this report, Bayesian estimation theory is applied to deriving lower bounds on performance for receivers operating in a block transmission system. Our work is based on the structure for deriving
lower performance bounds presented by Forney [17, 19], where he employs a good genie and a genie-aided detector (GAD). The idea with using a genie is to create a fictional receiver that is better than any real receiver and whose performance is easy to calculate. This fictional receiver, the GAD, has access to all the information that a real receiver has and, by the intervention of the genie, to additional side information about the transmitted sequence, see Figure 6.

One important result in Forney’s work [17] was a performance bound decaying with the same exponent as the performance of the MLSD. An almost identical bound was later derived by Mazo in [34] without employing the concept of the genie. We use the same basic hypothesis formulation as in Part I.1 and the block transmission system model (1). The focus of our work is on the concept of the GAD as a tool, but we also give some novel performance bounds and re-derive the bound presented by Mazo in [34].

We have enhanced the concept of the GAD with an explicit statistical description
of the side information. That is, we introduce conditional probabilities for any particular side information to appear given the identity of the transmitted sequence, describing the operation of the lower path in Figure 6. This makes it possible to derive the GAD with Bayesian detection theory, which seems to be difficult without otherwise augmenting Forney's work with some extra a priori information. It is sometimes argued that if the totality of information is used optimally, the GAD must have a better performance than any actual receiver. This reasoning must, however, be used with care: the criterion of optimality and the system model must be precisely specified for it to be meaningful. One of the points of this report is to make a derivation of the GAD based on a complete model, using the criterion of minimum bit-error probability.

The side information \( z \) supplied to the GAD is given for every bit to be decoded. The side information for the decoding of bit \( k \) is a pair of sequences \( (\beta^+, \beta^-) \in \mathcal{C}_k^+ \times \mathcal{C}_k^- \) (at least differing in the bit to be decoded), the transmitted sequence being either \( \beta^+ \) or \( \beta^- \). Our transition probabilities \( \Pr \{ z_{i,j} | \beta^+_k \} \) and \( \Pr \{ z_{i,j} | \beta^-_k \} \) are user-selected design parameters. These parameters control the properties of the bound resulting from analysing the GAD's performance. The structure given in Part 1.3 is flexible in that many different performance bounds can be derived, among them Mazo’s bound [34].

When the side information is given with the transition probabilities described by Lemma 4.2 in Part 1.3, the performance of the GAD for bit \( k \) will be determined by a \( Q(\cdot) \)-function scaled with a constant independent of the SNR, as

\[
P_{\text{DE},k} \geq P_{\text{DE},k}^{\text{GA}} = \frac{2|\mathcal{B}_k|}{2^N} Q\left(\frac{d_{\text{min}}}{2}\right),
\]

with \( |\mathcal{B}_k| \) as the number of elements in a set of exclusive pairs of sequences at minimum distance. Equation (8) is Mazo’s bound [35] and of the same form as the bound presented by Forney [17, 19]. We also introduce other bounds on performance, for instance the tightest bounds possible with the structure presented in Part 1.3.

In conclusion, the results of our derivation differ from Forney’s. However, the significance of our work is not found in the bounds we derive; it is found in the application of our formal structure when producing such bounds. This formal structure is an analytic tool, based on the novel probabilistic description of the side information, for deriving performance bounds. It is accessible for future analysis, and, we feel, offers an instructive view on earlier work.
Part II: Low Complexity Receivers

This part of the thesis consists of three conference contributions and one paper submitted to a conference.

Combined linear-Viterbi equalizer (CLVE) [10, 16, 20, 21, 39, Part II.1 – Part II.3] is a term often used for a class of receivers reducing the complexity of the Viterbi detector [19] by assuming the noise to be white and giving the Viterbi detector an approximate, truncated channel model. Linear pre-equalization of the received data is employed to adapt the system impulse response, as seen by the Viterbi detector, to the truncated channel model, see Figure 7. Other approaches to reducing the complexity of the Viterbi detector can be found in [4, 5, 6, 25, 31, 43].

In our matrix formalism, the operation of a CLVE in a block transmission system is given by

\[ \hat{b}_{\text{CLVE}} = \arg\min_{\beta \in \mathbb{C}} \|P\beta - Q\|^2, \]  

where \( P \) is a matrix describing the linear pre-equalizer and \( Q \) describes the truncated channel model defining the metric used by the Viterbi detector. The matrix \( Q \) is a band matrix, where the width of the band determines the complexity of the Viterbi detector, i.e., its number of states. Note that if \( P = I \) and \( Q = H \), the CLVE becomes the full-complexity Viterbi detector, cf. the definition of the MLSD in (2).

CLVE receivers are “mismatched receivers,” a class of receivers working with an incorrect metric. In the case of CLVE, this is a consequence of the complexity restrictions on \( Q \). Estimating the performance of mismatched receivers has been addressed by many authors [4, 7, 9, 30, 36], and there are means of determining tight upper bounds for them [4].

Part II.1: A Reduced Complexity Viterbi Equalizer Used in Conjunction with a Pulse Shaping Method

This conference paper introduces the use of the weighted least-squares (WLS) filter [32] as the pre-equalizer part of a CLVE. This design is intended for transmission systems where the channel \( H \) represents a linear, time-invariant filter.

The WLS method was originally developed as a pulse shaping filter for ultrasonic imaging [13, 32, 33]. It can be described by three requirements:

1. The system impulse response must have a non-zero amplitude at least at one point, say for index \( k = k_0 \).
2. The energy of the system impulse response should be small outside some desired interval, \( |k - k_0| > \frac{m}{2} \).

3. The noise amplification should be low.

A solution for 1 and 3 is the matched filter, and a solution for 1 and 2, with the interval \( m = 1 \), is the inverse filter. The WLS filter combines 1, 2, and 3 with the desired memory length \( m \) and the expected noise power \( \frac{\sigma^2}{2} \) into a cost function [32]. We assume here that the channel matrix \( H \) is Toeplitz describing convolution with a time-invariant filter. Let \( \hat{H} \) be a larger version of \( H \) containing \( H \) as a submatrix (see Part II.1), \( p \) be a vector containing the impulse response of the pre-equalizer \( \{p_k\}, \delta_{k_0} \) be a vector of zeros except for a one (') at position \( k_0 \), and, finally, let \( V \) be a diagonal matrix with \( m - 1 \) zeros in the center of the diagonal and ones at position \( k_0 \) and on the edges. The WLS cost function can then be expressed as

\[
J(p) = \|V(\hat{H}p - \delta_{k_0})\|^2 + \alpha \frac{\sigma^2}{2} \|p\|^2. \tag{10}
\]

Let \( h \ast p \) denote convolution between the sequences \( h \) and \( p \). Equation (10) can be interpreted as a function of the result of sending a unit pulse through the model network in Figure 8, summing the energy of the system impulse response \( h \ast p \) outside the given interval of length \( m \) (often referred to as the residual ISI, cf. [36]) and adding a term proportional to the noise amplification of the WLS pre-equalizer. With this criterion, the optimum pre-equalizer \( p \) is given by

\[
p_{\text{opt}} = \left[ \hat{H}^T V^T V \hat{H} + \alpha \frac{\sigma^2}{2} I \right]^{-1} \hat{H}^T \delta_{k_0}. \tag{11}
\]

The truncated channel model \( q \) consists of the \( m \) elements of the overall system pulse response \( h \ast p \) corresponding to the \( m \)-wide "window" of \( V \), i.e., \( h \ast p \) is calculated and the tails are truncated.

The complexity-reducing procedure is to filter the incoming data stream \( y \) with \( p_{\text{opt}} \), truncating the result to the proper length, \( N + m - 1 \), and feeding this new data stream and the truncated channel model \( q \) to the Viterbi detector.

We compare the bit-error rate of the WLS-CLVE with the bit-error rate of the CLVE of Falconer and Magee [16]. Simulations indicate that the WLS-CLVE has equal or superior performance compared to other CLVE-designs, see Part II.1 and Part II.2.

Part II.2: Combined Linear-Viterbi Equalizers – A Comparative Study and A Minimax Design

This conference paper presents a design criterion for CLVE where the distribution of the residual ISI, the ISI ignored by the CLVE, is taken into account and shaped. Our basic idea is to avoid strong, dominating components in the residual ISI.

We reconsider the weighted least-squares (WLS) design technique for CLVEs (established in Part II.1) by introducing a minimax criterion for suppressing the strongest component of the residual intersymbol interference. We investigate the performance of the minimax design and of some CLVE designs found in the literature.
All these design methods, including the minimax-CLVE and the WLS-CLVE, assume that the channel matrix $H$ describes a time-invariant filter, as opposed to the block-CLVE design method of Part II.3.

We also present a comparison of these CLVE designs based on a common quadratic optimization criterion for the selection of the pre-equalizer and the truncated channel model, see Figure 8. The error $e_i$ in Figure 8, can be expressed as

$$e_i = [b * (h * p - q) + n * p]_i,$$

where $q$ is properly aligned. The quadratic optimization criterion is a function of the variance of $e_i$. By using $\|x\|_M^2 \triangleq x^T M x$, the variance of $e_i$ is given by

$$E\{e_i^2\} = \|H p - q\|_b^2 + \|p\|_n^2,$$

where $p$ and $q$ are vectors containing the impulse response of the pre-equalizer and the truncated channel model, respectively. The matrices $R_b$ and $R_n$ are the covariance matrices for $b$ and $n$. In this paper $H$ is a Toeplitz band matrix such that the multiplication $H p$ describes convolution. This quadratic criterion reflects the basic ideas found in most design methods [10, 16, 21, Part II.1]; the difference between the methods being how the truncated channel model $q$ is chosen.

In our simulations the WLS-CLVE shows the best performance among the CLVEs. It is close in performance to the full-complexity Viterbi detector, despite having a much lower complexity. The performance of the minimax-CLVE is almost up to par with the WLS-CLVE, which indicates a potential for the concept of shaping the distribution of the residual ISI.

**Part II.3: Multiple-Antenna Reception and a Reduced-State Viterbi Detector for Block Transmission Systems**

This paper introduces the combination of a CLVE and a symbol-sampled, multiple receiving-antenna system model with maximal-ratio combining [8, 12, 22] in the receiver. It discusses two topics and uses the same matrix formalism as found in Part I, but in a multiple-antenna setting. The first topic concerns the noise and ISI properties of the output from the maximal-ratio combiner. We observe that, in
Figure 9: The receiver structure for the case that the noise at each particular antenna is uncorrelated with the noise at all other antennas.

general, as the number of available antennas increases, the severity of distortion (ISI and noise) decreases.

We note that a system with a receiver as depicted in Figure 9, with multiple receiving-antennas, a maximal-ratio combiner, and an additional data transformation matrix, can be modelled with our standard block transmission system model (1). The equivalent model has a quadratic, $N \times N$, channel matrix $H$, which cannot be interpreted as a convolution with a dispersive, time-invariant filter. This makes standard methods of CLVE-design inappropriate and calls for new design methods.

The second topic we discuss is the design of CLVEs for block transmission systems, allowing the channel matrix $H$ to have arbitrary structure, see Part II.3 and Part II.2. CLVEs designed for continuous transmission systems presume a time-invariant system impulse response. We give an example of a block-CLVE design method, a block version of the WLS-method from Part II.1, to illustrate possible performance gains. This method is, however, not practical due to its high design complexity; its significance is that it may give insight into block-CLVE design.

**Part II.4: Implementation of a System for Validation of Algorithms Used in Digital Radio Communication Schemes**

A hardware system for testing and evaluating algorithms and methods used in the physical layer of communication systems is presented. The reusable hardware platform facilitates rapid implementation of different digital receivers and is, e.g., suitable for CLVEs. This proposed hardware system partially merges system simulation with real-time validation; its purpose, in the development of communication concepts, is to make the real-time validation phase as flexible as the simulation phase.

The proposed flexible system uses standard signal processors, array signal processors and field programmable gate arrays. It also contains a software controlled waveform generator to be used as a transmitter and a high performance channel simulator. The point of the paper is that the use of this platform would lower the cost in time and equipment involved in developing prototypes and validating communication systems.

To illustrate the use of the proposed system, a GSM-type system with high data rate is used as a design example.
Part III: Robust Receivers for Ultrasound

The following two papers address the problem of time-of-flight estimation when the shape of the received pulse is uncertain, see *e.g.*, [40, 42]. The primary intended application is distance estimation using the ultrasonic pulse-echo method [38]. With the ultrasonic pulse-echo method, a short pulse with most of its energy concentrated in a narrow frequency band, *cf.* Figure 10, is transmitted with an ultrasound transducer. Reflections will appear where the acoustic impedance changes, and the distance to a reflector is assumed to be proportional to the time-of-flight of the pulse. In the two papers we address the problem of estimating the distance to a single reflector by estimating the arrival-time and the changes of shape of a returned pulse whose shape has been changed by additive noise and time dispersion. (For example, the geometry of the reflector affects the shape of the reflected pulse [27].) We present two approaches for mitigating the effect of the pulse-shape distortion. Although the intended application is primarily distance measurements using ultrasound, the two methods for time-of-flight estimation presume only that a known, narrowband waveform is transmitted through a dispersive linear system with additive Gaussian noise.

Our use of the word "robust" to describe our receivers in Part III.1 and Part III.2 deserves a comment. The receivers are robust against time dispersion by modelling and estimating the same. The word "robust" is not used in the sense that the receivers may simply disregard the shape distortion because they are insensitive to it.

![Figure 10: An example of a received ultrasound pulse.](image-url)
Part III.1: A Robust Correlation Receiver for Distance Estimation

In this journal paper we present a method for distance estimation using pulsed ultrasound based on a continuous-time system model. The developed signal model, of variable order, accounts for the time-of-flight as well as for an unknown, linear time dispersion and additive noise. Our method can be described in the following way. We first expand the unknown distorting system with a Taylor series in the frequency domain about the transmitted signal's centre frequency. We then simultaneously estimate the time-of-flight and the coefficients of the Taylor expansion.

The time-of-flight estimator is derived based on the criterion of maximum likelihood. The resulting receiver can be seen as a generalization of the well-known cross-correlation, or "matched filter," estimator described, e.g., by Nilsson in [37]. The proposed receiver has been validated, and compared with the cross-correlation receiver, by means of measurements with an experimental ultrasonic surface scanning system. The receiver is found, by both simulations and analysis, to be more robust against unknown pulse shape distortion than the cross-correlation estimator, giving time-of-flight estimates that are less biased. Furthermore, bias versus noise sensitivity can be controlled by signal-model order selection.

Part III.2: Simultaneous Time of Flight and Channel Estimation Using a Stochastic Channel Model

In this conference paper we address the problem of estimating the time-of-flight of an ultrasonic pulse by modelling the shape of the received waveform as stochastic. We assume the pulse to be distorted by both additive noise and a time-dispersive transition system with a stochastic impulse response. The stochastic impulse response is modelled as a discrete-time impulse response, the taps of which are a sequence of Gaussian random variables. The covariance matrix and mean of the taps are given by the user, based on some assumptions or a priori knowledge about the distorting system. The joint estimation of waveform and time-of-flight is couched in terms of maximum a posteriori (MAP) and maximum likelihood estimation. When deriving the MAP estimator we assume a priori knowledge of the probability density of the stochastic transmission system impulse response, and that the transmitted waveform is known to the receiver.

The ordinary cross-correlation time-of-flight estimator [37] assumes complete knowledge of the received noiseless waveform. In the viewpoint of this paper, it has a one-dimensional transmission system model. This investigation indicates that a more complex model structure is worthwhile when unknown, or partially unknown, time dispersion and additive noise is present.
Corrections

We make a correction to the conference paper Part III.2. In that paper it is assumed that the transmitted pulse is both bandlimited in frequency and identically zero over a semi-infinite time-interval. A signal cannot completely fulfill both these criteria. Instead of being bandlimited in frequency, the signal should have most of its energy concentrated to a narrow frequency interval. (One alternative would have been the formulation of Part III.1.)
Bibliography


Part I.1

Representations for the Minimum Bit-Error Probability Receiver for Block Transmission Systems with Intersymbol Interference Channels
Part I.1:
**Representations for the Minimum Bit-Error Probability Receiver for Block Transmission Systems with Intersymbol Interference Channels**

Per Ödling\(^1\), Håkan Eriksson\(^1\), Timo Koski\(^2\) and Per Ola Börjesson\(^1\)


A shorter version has been accepted for presentation at the *1995 IEEE International Symposium on Information Theory*, Whistler, Canada, September 1995.

\(^1\)Luleå University of Technology, Division of Signal Processing, S-971 87 Luleå, Sweden

\(^2\)Luleå University of Technology, Department of Mathematics, S-971 87 Luleå, Sweden
Representations for the Minimum Bit-Error Probability Receiver for Block Transmission Systems with Intersymbol Interference Channels

Per Ödling\(^1\) Håkan B. Eriksson\(^1\) Timo Koski\(^2\)
Per Ola Börjesson\(^1\)

Luleå University of Technology
\(^1\) Division of Signal Processing
\(^2\) Department of Mathematics
S-971 87 Luleå, Sweden

May 1994

Abstract

We reconsider the minimum bit-error probability receiver for intersymbol interference channels with Gaussian noise using a geometric theory of data transmission of finite blocks of bits. Representations of the receiver are given in terms of an explicit closed form expression consisting of sums of certain Gaussian kernels. Using these representations and assuming that all data sequences are equally probable we prove two results about the behaviour of the minimum bit-error probability receiver, both results based on that the receiver is dependent on the SNR. We show that the minimum bit-error probability receiver when designed for asymptotically high SNR becomes the maximum likelihood sequence detector and that it collapses to a matched filter followed by a hard-limiting device when designed for low SNR. This also implies that these respective receivers attain the minimum bit-error probability in the mentioned extremes of SNR, a well-established result in the case of the MLSD. Simulations are presented to illustrate our results.

Key words: digital block transmission, intersymbol interference, minimum bit-error probability, maximum likelihood sequence detection, matched filter, sufficient statistic, asymptotic receiver properties.
1 Introduction

The optimal, or minimum, bit-error probability receiver (OBER) for intersymbol interference channels was first proposed in the sixties [1, 2, 5, 9], later further investigated in [7, 10, 19] and recently discussed with respect to neural networks in [11, 12]. Much of the attention given to the OBER was transferred to the Viterbi detector [16] when the latter was introduced as a Maximum Likelihood Sequence Detector (MLSD). When compared to the Viterbi detector, the OBER shows only a moderate gain in bit-error probability [6, 14, 15, 16, 19], but it achieves this at a cost of considerably higher computational complexity.

In this work we reconsider the OBER for block transmission systems in order to determine the OBER's properties and its relation to the MLSD. In section 2 we present a binary block transmission system model. In section 3 we use standard detection theory to derive a representation of the OBER. We also point out a parallel block processor structure for the OBER, which has simultaneous detection of all the individual bits in a block. In section 4 we investigate the asymptotic behaviour of the OBER when designed for large and small signal-to-noise ratios (SNR). For the case when all sequences are equally probable, we show that when the SNR used for designing the OBER increases without bound, the OBER and the MLSD coincide in the limit: their respective decision regions become identical. One important consequence of this is that the MLSD will have the minimum attainable bit-error probability when used in a system with a high SNR, as shown in, e.g., [16, 18]. In section 4.3 we find a similar correspondence between the OBER and a matched filter with hard decisions when the OBER is designed for a low SNR. To illustrate the similarities and differences between the three mentioned receivers simulations are presented in section 5.

2 A block transmission system model

Consider the transmission of blocks of binary data, typically interspersed with blocks of bits known to the receiver, through a channel with known intersymbol interference (ISI) and additive Gaussian noise at the receiver. The set of possible blocks or sequences of \( N \) bits are the vertices of the centered binary hypercube \( C \triangleq \{-1, +1\}^N \). Let \( b \in C \) denote the stochastic vector containing the sequence of \( N \) independent bits to be transmitted. We represent the transmission system in matrix notation as

\[
y = Hb + n, \tag{2.1}
\]

where \( H \) is a deterministic and known \( (N + L - 1) \times N \) matrix representing the ISI, \( n \) is a white, jointly Gaussian, zero mean random noise vector with variance \( \sigma_n^2 \), or equivalently \( n \in N(0, \sigma_n^2 I) \), and, finally, \( y \in \mathbb{R}^{N+L-1} \) is the \( (N + L - 1) \times 1 \) dimensional stochastic variable to be observed by the receiver*. Further, let \( \beta \) and \( \eta \) denote the outcomes of the stochastic variables \( b \) and \( y \), respectively. With a few exceptions, we ignore the issue of coloured noise.

We continue with a number of additional definitions that we will find useful in

*How one can derive this symbol-rate sampled representation from a continuous-time channel model is described in [3, 4, 16].
the sequel. Let us first define the SNR in the above block transmission system as

\[ \text{SNR} \triangleq \frac{\text{trace}\{H^T H\}}{N\sigma_n^2}. \]  

(2.2)

This is the block-SNR, the average received energy per transmitted bit over the variance of the noise in a block.

Following Barbosa in [4], a sufficient statistic [23] for \( b \) given the model in (2.1) is provided by

\[ x_{\text{ZF}} = x_{\text{ZF}}(y) \triangleq M^{-1}H^T y, \]  

(2.3)

where \( M = H^T H \). The matrix operation \( M^{-1}H^T \) is an instance of the Moore-Penrose pseudo-inverse, and can be seen as the zero-forcing equalizer for block transmission systems (ZFEB) [4, 20]. This is illustrated by inserting (2.1) in (2.3) giving \( x_{\text{ZF}} = b + M^{-1}H^T n \). Given that \( \beta \) is transmitted, the output of the ZFEB, \( x_{\text{ZF}} \in \mathbb{R}^N \), is a joint Gaussian noise vector with mean \( \beta \) and covariance matrix equal to \( M^{-1} \). Let \( \xi_{\text{ZF}} \triangleq x_{\text{ZF}}(\eta) \) denote the outcome of the stochastic variable \( x_{\text{ZF}} \).

Assuming that all sequences are equally probable to be transmitted, it follows, by the factorization theorem for characterizing sufficiency [23, p. 80] and from the definition of \( x_{\text{ZF}} \), that an MLSD is obtained by solving

\[ \hat{b}_{\text{MLSD}}(x_{\text{ZF}}) \triangleq \arg\min_{\beta \in \mathcal{C}} \|x_{\text{ZF}} - \beta\|_M^2, \]  

(2.4)

where \( \|x\|_M^2 \triangleq x^T M x \).

One is often interested in the bit-error probability, or probability of error per data bit, instead of the sequence error probability. With the output of a detector denoted by \( \hat{b} \), we define the pertinent bit-error probability, \( P_{\text{BER}} \), as

\[ P_{\text{BER}} \triangleq \frac{1}{N} \sum_{k=1}^{N} P_{\text{DE},k}, \]  

(2.5)

where \( P_{\text{DE},k} \) is the probability of detection error in the kth bit. More strictly, we define \( P_{\text{DE},k} \) as

\[ P_{\text{DE},k} \triangleq \sum_{\beta \in \mathcal{C}} \Pr\{b = \beta\} \Pr\{\hat{b}_{\text{bit},k} \neq (\beta)_{\text{bit},k} | b = \beta\}, \]  

(2.6)

where \( \Pr\{b = \beta\} \) is the a priori probability that the sequence \( \beta \) is transmitted, \((\beta)_{\text{bit},k}\) is the kth bit in \( \beta \) and \( \Pr\{\hat{b}_{\text{bit},k} \neq (\beta)_{\text{bit},k} | b = \beta\} \) is the probability that the kth bit of the detector output differs from the corresponding transmitted bit.

We refer to systems, described by the model in (2.1), where optimal detection of the transmitted blocks of data can be performed independently on each block, as block transmission systems, see, e.g., [4, 8, 20]. Many real-life communication systems, like the cellular phone systems GSM (Europe) and D-AMPS/IS-54 (USA), use block transmission, often with MLSD, without, however, strictly falling into our definition of block transmission systems. The coding and interleaving they employ introduce dependency between the blocks (and the individual bits in a block), wherefore optimal decoding of a particular block cannot be done independently of all other blocks, see, e.g., [24]. Here we focus on the properties of the OBER used in block transmission systems with independent blocks.
3 The Optimal Bit-Error Probability Receiver

Given the model of a system with digital block transmission, ISI and additive Gaussian noise, we will next derive the optimal bit-error probability receiver (OBER) [1, 2, 5, 9]. Let us consider the detection of \((b)_{\text{bit }, k}\), which is a binary hypothesis test. A geometric interpretation of this is that of choosing the correct halfcube:

\[
H_0 : \ b \in C_k^-
\]
\[
H_1 : \ b \in C_k^+ ,
\]

where \(C_k^+\) and \(C_k^-\) are the halfcubes with \((b)_{\text{bit }, k} = +1\) and \((b)_{\text{bit }, k} = -1\), respectively.

The Bayes decision rule [25] minimizing the probability of detection error is in this case given by

\[
\Lambda_k(x_{ZFE}) \triangleq \frac{f_{X_{ZFE}|H_0}(x_{ZFE}|H_1)}{f_{X_{ZFE}|H_0}(x_{ZFE}|H_0)} \geq \frac{H_1}{H_0} \Pr\{H_0\} \Pr\{H_1\},
\]

where \(\Pr\{H_1\} = 1 - \Pr\{H_0\}\) is the a priori probability that \(H_1\) is true, and \(f_{X_{ZFE}|H_0}(\cdot)\) and \(f_{X_{ZFE}|H_1}(\cdot)\) designate the probability density functions for \(x_{ZFE}\) given \(H_0\) and \(H_1\), respectively. We use \(\geq\) to represent that the decision is \(H_1\) if the indicated inequality is true. In the proof of the next proposition we give a formula for these densities. Using this formula, the proposition yields a representation for the OBER.

**Proposition 3.1** Denote the sufficient statistic for \(b\), as in (2.3), with \(x_{ZFE}\). Let

\[
(b_{\text{OBER}})_\text{bit }, k = \text{sign} \left( \sum_{\beta \in C_k^+} w(x_{ZFE}, \beta) \right),
\]

where

\[
w(x_{ZFE}, \beta) \triangleq \psi(x_{ZFE}|\beta) \Pr\{b = \beta\} - \psi(x_{ZFE}|\beta) \Pr\{b = -\beta\},
\]

and where \(\Pr\{b = \beta\}\) is the a priori probability for the sequence \(\beta\) being transmitted. The density \(\psi(x_{ZFE}|\beta)\) is the conditional density of \(x_{ZFE}\) given that the sequence \(\beta\) was transmitted. Then \((b_{\text{OBER}})_\text{bit }, k\) is the detector, based on the channel output \(y\), that minimizes the probability of error in bit \(k\), \(P_{DE,k}\).

A formal proof of the proposition can be found in appendix A. The proposition is related to the presentation in [17].

The detector given by equation (3.9) is valid for all channels for which \(\psi(\cdot)\), the conditional density of \(x_{ZFE}\) given a transmitted sequence \(\beta\), is well-defined. In particular, for the block transmission model of section 2, \(\psi(\cdot)\) is a Gaussian density given by

\[
\psi(x_{ZFE}|\beta) = \mathcal{K}(\sigma_n) \exp \left( \frac{-1}{2\sigma_n^2} \|x_{ZFE} - \beta\|_M^2 \right),
\]

where

\[
\mathcal{K}(\sigma_n) = \frac{\sqrt{\det \mathbf{H}^T \mathbf{H}}}{(\sqrt{2\pi})^N \sigma_n^N}.\]
We wish to point out that the function $w(x_{ZFE}, \beta)$ is independent of $k$. If we use (3.9) for the detection of every bit in $b$ given $y$, we can use the same $2^N$ values of $w(x_{ZFE}, \beta)$ for all $N$ bits of the transmitted sequence. Based on this, we present a closed form expression for the OBER that detects all bits in $b$ simultaneously and displays a structure amenable for parallel signal processing. It should be noted that the concept of binary antipodal modulation is crucial for this structure.

**Proposition 3.2** Let

$$\hat{b}_{OBER}(x_{ZFE}) \triangleq \text{sign} \left( \sum_{\beta \in C_+^k} w(x_{ZFE}, \beta) \beta \right). \quad (3.13)$$

Then $\hat{b}_{OBER}(x_{ZFE})$ is the detector, based on the channel output $y$, that minimizes the total bit-error probability, $P_{BER}$, in the sense of (2.5).

**Proof:** Recall that $C_+^k$ is the set of all sequences in the hypercube $C$ with $(\beta)_{bit\ k} = +1$. In (3.9) we need to sum over $C_+^k$ for each $k$, i.e., for each bit to be decoded. We use the sequences in $C_+^1$ to order the sequences in the other sets $C_+^k$ as $C_+^k = \{\beta_1^k, \beta_2^k, \ldots, \beta_{2^{N-1}}^k\}$, where

$$\beta_i^k = \text{sign} \left( (\beta_i^1)_{bit\ k} \right) \beta_i^1 = (\beta_i^1)_{bit\ k} \beta_i. \quad (3.14)$$

Thus, the way of ordering the sequences in $C_+^1$ determines the order of the sequences in the sets $C_+^k$ for $k \in \{2 \ldots N\}$. Using (3.14) in (3.9),

$$\begin{align*}
(\hat{b}_{OBER})_{bit\ k} &= \text{sign} \left( \sum_{i=1}^{2^{N-1}} w(x_{ZFE}, \beta_i^1) \right) = \text{sign} \left( \sum_{i=1}^{2^{N-1}} w(x_{ZFE}, (\beta_i^1)_{bit\ k} \beta_i^1) \right) \\
&= \text{sign} \left( \sum_{i=1}^{2^{N-1}} w(x_{ZFE}, \beta_i^1) \beta_i \right)_{bit\ k},
\end{align*} \quad (3.15)$$

where we have used the property that the function $w(x_{ZFE}, \beta)$ is, for a fixed value of $x_{ZFE}$, odd in the variable $\beta$, i.e., $w(x_{ZFE}, -\beta) = -w(x_{ZFE}, \beta)$. Finally, using (3.15) for the detection of every bit we have

$$\hat{b}_{OBER}(x_{ZFE}) = \begin{bmatrix} (\hat{b}_{OBER})_{bit\ 1} \\
\vdots \\
(\hat{b}_{OBER})_{bit\ N} \end{bmatrix} = \begin{bmatrix} \text{sign} \left( \sum_{i=1}^{2^{N-1}} w(x_{ZFE}, \beta_i^1) (\beta_i^1)_{bit\ 1} \right) \\
\vdots \\
\text{sign} \left( \sum_{i=1}^{2^{N-1}} w(x_{ZFE}, \beta_i^1) (\beta_i^1)_{bit\ N} \right) \end{bmatrix} \quad (3.16)$$

Every bit $(b)_{bit\ k}$ is hereby detected with minimum probability of error, $P_{BER}$, and, consequently, the total probability of bit-error, $P_{BER}$, as defined in (2.5), is minimized.
In minimizing the function in (2.4) the MLSD is processing optimally the information \( x_{ZFE} \), sufficient for the transmitted sequence \( b \). The findings above demonstrate a way the OBER can be constructed to extract optimally from \( x_{ZFE} \) the information for detecting the individual bits \( (b)_\text{bit} \). In (3.9) and (3.10) this is implemented by making pairwise computations and summations of the *posteriori* probabilities for the vertices in the halfcube \( C_+^k \) and for their opposite vertices or poles (having the maximal possible Hamming distance).

For a given \( \alpha \), the function \( \Gamma(x, \alpha) \) in (3.13) is a mapping from \( \mathbb{R}^N \) to \( \{-1, +1\}^N \), the hypercube \( C \) of dimension \( N \). The operations of equation (3.13) can be seen as either adding \( 2^{N-1} \) different vectors along the diagonals of the hypercube, or as weighing together the \( 2^N \) vertices, where the weights are dependent on the received data. This procedure bears some resemblance to computing the projection of \( x_{ZFE} \) onto the vertices in the halfcube \( C_+^k \). We shall show in the next section that for low signal-to-noise ratios (SNR) the argument inside the \( \text{sign}(\cdot) \) function is dominated by a term, which is a standard scalar product.

The argument of the \( \text{sign}(\cdot) \) function in equation (3.13) can also be viewed as the mapping from \( x_{ZFE} \in \mathbb{R}^N \) to a vector of the same dimension that causes the optimal decision boundaries to coincide with the unit axes in \( \mathbb{R}^N \).

In [11, 10, p. 307] it is argued that the OBER can be well approximated by a radial basis function network. Here (3.9) and (3.10) show that the OBER can in the case of a Gaussian channel be represented with the exact structure of such a network [21, pp. 119–120].

## 4 The behaviour of the OBER

In this section we will discuss some asymptotic properties of the OBER for high and low SNR, given the model in (2.1). Let

\[
\Gamma(x_{ZFE}(z), \alpha) \triangleq \text{sign} \left( \sum_{\beta \in C_+^N} w(x_{ZFE}, \beta|\alpha) \beta \right),
\]

where

\[
w(x_{ZFE}, \beta|\alpha) \triangleq \psi(x_{ZFE}|\beta, \alpha) \Pr\{b = \beta\} - \psi(x_{ZFE}|-\beta, \alpha) \Pr\{b = -\beta\},
\]

and

\[
\psi(x_{ZFE}|\beta, \alpha) \triangleq \mathcal{K}(\alpha) \exp \left( -\frac{1}{2\alpha^2} \|x_{ZFE} - \beta\|_M^2 \right),
\]

and where \( \mathcal{K}(\cdot) \) is given in (3.12). Then, given the model in (2.1), \( \Gamma(x, \sigma_n) = \tilde{b}_{\text{MLSD}}(x) \), where \( \sigma_n \) is the variance of the Gaussian noise vector \( n \).

Assuming that all sequences are equally probable to be transmitted we show that

\[
\lim_{\alpha \to 0} \Gamma(x, \alpha) = \tilde{b}_{\text{MLSD}}(x), \quad \text{for all } x \in \mathbb{R}^N,
\]

and that

\[
\lim_{\alpha \to \infty} \Gamma(x_{ZFE}(z), \alpha) = \text{sign}(H^Tz), \quad \text{for all } z \in \mathbb{R}^{N+L-1}.
\]

Equation (4.20) means that the OBER designed for a high SNR becomes the MLSD. As a consequence of this, the MLSD will achieve the minimum bit-error probability
when used in systems with a high SNR. In equation (4.21) we find a similar comparison for low SNR between the OBER and the matched filter with hard decisions. If the true SNR is low, the best possible receiver is actually the matched filter receiver as comes to the BER.

In subsection 4.1 we illustrate (4.20) and (4.21) by plotting decision regions for the statistics $x_{ZFE} \in \mathbb{R}^2$ for a two-dimensional numerical example. We also display an example of a decision surface when transmitting blocks of length $N = 3$. In subsections 4.2 and 4.3 we prove (4.20) and (4.21), respectively.

### 4.1 A numerical example

Let us study the OBER, the MLSD and the matched filter receiver by observing their decision surfaces for the statistics $x_{ZFE}$. Let $D_k^+$ designate the region in $\mathbb{R}^N$ where, when $x_{ZFE} \in D_k^+$, the detector in question decides that bit $k$ is positive and $D_k^-$ the region where it decides that bit $k$ is negative. The decision regions $D_k^+$ and $D_k^-$ are both continuous and disjoint and the surface that separates the two we denote the decision surface for bit $k$, cf. [22, 26].

In the numerical experiments below we will confine our attention to the transmission of blocks of two bits (giving a flat two-dimensional hypercube with four vertices) over a channel with ISI and white, stationary additive Gaussian noise. Furthermore, we assume that all sequences have equal a priori probabilities, i.e., $\Pr\{b = \beta\} = 2^{-N} = 1/4$ for all $\beta \in C$. The matrix $M$ for this system becomes a standard $2 \times 2$ dimensional correlation matrix, i.e.,

$$M \sim \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

where $-1 < \rho < 1$.

In Figure 1 we display decision surfaces for $x_{ZFE} \in \mathbb{R}^2$. On the left hand side are the decision surfaces of the MLSD for bit one with $\rho = 0.1, 0.3, 0.5, 0.7$ and $\rho = 0.9$, respectively. The corresponding decision surfaces for the matched filter receiver are given on the right hand side of Figure 1.

Using the system as defined above, Figure 2 shows some samples of decision surfaces of the OBER for the statistics $x_{ZFE}$. Each one of the three graphs of Figure 2 contains decision surfaces for the first bit in the block, for various SNRs. The situations described by the three graphs differ in the strength of the noise correlation, with $\rho$ equal to 0.1, 0.5 and 0.9, respectively. A correlation $\rho$ of 0.9 corresponds to severe ISI, for instance, in the case of white additive noise $n$, to a time invariant channel with an impulse response consisting of ten consecutive, equally large ISI coefficients, e.g., $[1, 1, 1, 1, 1, 1, 1, 1, 1, 1]/\sqrt{10}$. Sign reversing $\rho$ gives mirrored curves. The SNRs in Figure 2 range from $-12$dB to 12dB in steps of 3dB, where the straight lines correspond to $-12$dB SNR and the “edgy” curves to $+12$dB SNR. The curve corresponding to 0dB is dashed.

In Figure 3 we give an example of a decision surface for the first bit of the MLSD when transmitting blocks of three bits.

As the SNR decreases the decision surface between the decision regions of the two possible values of any given bit $(b)_k$ approaches a hyperplane. This corresponds to a receiver, for optimal detection of the specific bit, consisting of a linear filter followed by a hard limiting device. On the other hand, as the SNR increases the
Figure 1: Decision regions of the MLSD (left) and the matched filter receiver (right) for the statistics $x_{ZFE} = [x_1, x_2]^T \in \mathbb{R}^2$. The decision regions are for the first bit in a two bit block with $p = [0.1, 0.3, 0.5, 0.7, 0.9]$.

decision surfaces of the OBER tend to become more and more "edgy". A comparison of Figure 1 and Figure 2 gives an indication of the relation between the three receivers subject to our interest, and also suggests the method of proof employed in subsections 4.2 and 4.3.

Figure 2: Decision regions of the OBER for the statistics $x_{ZFE} = [x_1, x_2]^T \in \mathbb{R}^2$ when deciding on the first bit in a two bit block. The SNR varies from $-12$dB to 12dB, in steps of 3dB. The dashed curves designate the decision boundaries corresponding to a SNR of 0dB. Left: $\rho = 0.1$, Middle: $\rho = 0.5$, Right: $\rho = 0.9$.

4.2 The MLSD is optimal for high SNR

In this subsection we show that the OBER and the MLSD have identical decision regions when the OBER is designed for asymptotically high SNR. We assume that the a priori probabilities $\Pr \{b = \beta\} = 2^{-N}$ for all $\beta \in \mathcal{C}$. This assumption makes the MLSD independent of $\sigma_n$. Let us order the vertices of $\mathcal{C}$ so that $\mathcal{C} = \{\beta_1, \beta_2, \ldots, \beta_{2N}, \beta_{2N+1}\}$ and let, for $i = 1, 2, \ldots, 2^N$,

$$D(\beta_i) \triangleq \{x \in \mathbb{R}^N \mid \hat{b}_{\text{MLSD}}(x) = \beta_i\}$$
Figure 3: An example of a decision surface of the MLSD for the statistics $x_{ZFE} = [x_1, x_2, x_3]^T \in \mathbb{R}^3$ when detecting the first bit in a block of three bits. In the centre of the Figure the top part of the hypercube is visible.

$$\{x \in \mathbb{R}^N \mid \psi(x|\beta_i) > \psi(x|\beta_j), \ j \neq i\}$$

denote the region in $\mathbb{R}^N$ where $\beta_i$ is the decision of the MLSD ($\hat{b}_{MLSD}$), see (2.4). By the boundary of a decision region we mean the set of $x \in \mathbb{R}^N$ for which the two (or more) largest of the densities $\psi(x|\beta_i)$ are equal. We ignore values on the boundaries because, once the OBER or the MLSD are applied in a communication system, we will be dealing with channel observables having densities, and outcomes on the boundaries will then normally have probability zero conditioned on any transmitted sequence.

**Proposition 4.1** Assume the model (2.1) and $\Pr\{b = \beta\} = 2^{-N}$ for all $\beta \in C$. For all $x \in \mathbb{R}^N$ not lying on the decision boundaries it holds that

$$\hat{b}_{MLSD}(x) = \Gamma(x, \alpha)$$

(4.23)

for all $\alpha$ smaller than some positive number dependent on $x$.

**Proof:** Let us take a fixed arbitrary $x \in \mathbb{R}^N$. Because the decision regions $\{D(\beta_i)\}_{i=1}^{2^N}$ form a partition of $\mathbb{R}^N$, the chosen $x$ lies in one and only one of the $D(\beta_i)$, disregarding $x$ on any of the boundaries. In addition we assume without restriction of generality that this $\beta_i \in C_i^+$. From (4.17) we write

$$\Gamma(x, \alpha) = \text{sign} \left( \psi(x|\beta_i, \alpha) \beta_i - \psi(x|\beta_i, \alpha) \beta_i + \sum_{\beta \in C_i^+, \beta \neq \beta_i} w(x, \beta|\alpha) \beta \right).$$

(4.24)
Because the value of the function \( \text{sign}(\cdot) \) is not altered if its argument is multiplied by a positive number, we obtain
\[
\Gamma(x, \alpha) = \text{sign}(\beta_t + r(x|\beta_t, \alpha)),
\] (4.25)
where
\[
r(x|\beta_t, \alpha) \triangleq - \frac{\psi(x|\beta_t, \alpha)}{\psi(x|\beta_t, \alpha)} \beta_t + \sum_{\beta \in C_t \setminus \beta_t} \frac{w(x, \beta|\alpha)}{\psi(x|\beta_t, \alpha)} \beta.
\] (4.26)

Because \( x \) is in \( D(\beta_t) \), we can choose the parameter \( \alpha \) to be sufficiently small, such that all the ratios of the type \( \psi(x|\beta, \alpha) / \psi(x|\beta_t, \alpha) \) which \( r(x|\beta_t, \alpha) \) consists of, will become as small as desired. We, therefore, select \( \alpha \) so small that the signs of none of the bits \( (\beta_i)_{\text{bit } k} \) is overturned by \( r(x|\beta_t, \alpha) \). In particular, it then holds that
\[
\text{sign} \left( \beta_t - \frac{\psi(x|\beta_t, \alpha)}{\psi(x|\beta_t, \alpha)} \beta_t + \sum_{\beta \in C_t \setminus \beta_t} \frac{w(x, \beta|\alpha)}{\psi(x|\beta_t, \alpha)} \beta \right) = \text{sign}(\beta_t) = \beta_t,
\] (4.27)
the desired conclusion.

Since we are dealing with Gaussian densities \( \psi(\cdot) \) the convergence established above is monotonic. However, the convergence is not uniform, i.e., the value of \( \alpha \) achieving the desired equality must in general be altered with \( x \).

In addition to showing that the MLSD and the OBER are identical in the limit, the above proof also indicates that the parts of the signal space where the receivers make different decisions are quite small for high SNRs.

### 4.3 The matched filter receiver is optimal for low SNR

We will show that the OBER and the matched filter receiver have identical decision regions when the OBER is designed for an asymptotically low SNR. We will disregard input signals lying on the decision boundaries of the matched filter receiver, that is, the values of \( z \in \mathbb{R}^{N+L-1} \) for which the vector \( H^T z \) contains at least one zero.

**Proposition 4.2** Assume the model (2.1) and \( \Pr\{b = \beta\} = 2^{-N} \) for all \( \beta \in C \). For all \( z \in \mathbb{R}^{N+L-1} \) it holds for sufficiently large \( \alpha \) that
\[
\Gamma(x_{\text{MFE}}(z), \alpha) = \text{sign}(H^T z).
\] (4.28)

**Proof:** Consider the receiver with the representation (4.17). We write the generic kernel corresponding to \( \Gamma(x, \alpha) \) as
\[
w(x, \beta|\alpha) = \frac{1}{2^N} A(x, \alpha) \exp \left( -\frac{\beta^T M \beta}{2\alpha^2} \right) \sinh \left( \frac{\beta^T M x}{\alpha^2} \right),
\] (4.29)
where \( A(x, \alpha) > 0 \) is independent of \( \beta \) and is given in appendix A. Let us recall the standard Taylor expansion
\[
\sinh \left( \frac{\beta^T M x}{\alpha^2} \right) = \frac{\beta^T M x}{\alpha^2} + \frac{(\beta^T M x)^3}{6\alpha^6} + \ldots
\]
From (4.17), we write
\[
\Gamma(x, \alpha) = \text{sign} \left( \sum_{\beta \in \mathcal{C}_1^+} w(x, \beta | \alpha) \beta \right)
\]
\[
= \text{sign} \left( \frac{1}{2^N A(x, \alpha)} \sum_{\beta \in \mathcal{C}_1^+} \exp \left( -\frac{\beta^T M \beta}{2 \alpha^2} \right) \beta^T M x + \frac{1}{2^N A(x, \alpha)} t(x, \alpha) \right),
\]
where we are collecting the terms containing powers of $\beta^T M x / \alpha^2$ of order higher than one inside the notation
\[
t(x, \alpha) \triangleq \sum_{\beta \in \mathcal{C}_1^+} \exp \left( -\frac{\beta^T M \beta}{2 \alpha^2} \right) \frac{\beta^T M x \beta}{6 \alpha^6} + \ldots \beta.
\]
(4.30)

Since the value of the sign(⋅)-function does not change if its argument is multiplied by a positive number, we scale the argument in the right hand side above by $2^N A(x, \alpha)^{-1} \cdot \alpha^2$ to obtain
\[
\Gamma(x, \alpha) = \text{sign} \left( \sum_{\beta \in \mathcal{C}_1^+} \exp \left( -\frac{\beta^T M \beta}{2 \alpha^2} \right) \beta^T M x \beta + \alpha^2 \cdot t(x, \alpha) \right). \quad (4.31)
\]

Let us next note that by selecting $\alpha$ sufficiently large we can make the term $\alpha^2 \cdot t(x, \alpha)$ arbitrarily small, in particular so small that the value of $\Gamma(x, \alpha)$ is determined by the first term of the argument in the right hand side of (4.31). Expanding $\exp \left( -\frac{\beta^T M \beta}{2 \alpha^2} \right)$ in its standard Taylor expansion we find that
\[
\Gamma(x, \alpha) = \text{sign} \left( \sum_{\beta \in \mathcal{C}_1^+} \beta^T M x \beta + u(x, \alpha) + \alpha^2 \cdot t(x, \alpha) \right), \quad (4.32)
\]
where all terms but the first from the Taylor expansion are collected into the function $u(x, \alpha)$. As in the case with $\alpha^2 \cdot t(x, \alpha)$, we can make $u(x, \alpha)$ arbitrarily small by choosing $\alpha$ large enough, in particular large enough so that the sign of the right hand side of (4.32) is determined by its first term. By inserting $x = x_{\text{zfe}}(z)$ in (4.32), recalling that $x_{\text{zfe}}(z) = M^{-1} H^T z$ and using the identity
\[
\sum_{\beta \in \mathcal{C}_1^+} (\beta^T H^T z) \beta = \sum_{\beta \in \mathcal{C}_1^+} \beta (\beta^T H^T z) = \sum_{\beta \in \mathcal{C}_1^+} (\beta \beta^T) H^T z = 2^{N-1} I H^T z, \quad (4.33)
\]
we have
\[
\text{sign} \left( \sum_{\beta \in \mathcal{C}_1^+} \beta^T H^T z \beta \right) = \text{sign}(H^T z). \quad (4.34)
\]
Thus, it holds that there is for each $z \in \mathbb{R}^{N+L-1}$ an $\alpha$ large enough so that
\[
\Gamma(x_{\text{zfe}}(z), \alpha) = \text{sign}(H^T z). \quad (4.35)
\]

12
Hence, in the limit when the SNR used for designing the OBER decreases, the OBER reduces to the matched filter for block transmission systems, $H^T y$, followed by a hard limiting device, sign($\cdot$), as was to be proved.

For the case of correlated additive noise $n$, the receiver of (4.35) generalizes to the whitened matched filter receiver (WMFR), $H^T R_n^{-1} y$, where $R_n$ is the correlation matrix for the noise $n$.

## 5 Simulations

Consider a communication system that is transmitting blocks of data over a time dispersive channel with additive uncorrelated Gaussian noise, $n \in N(0, \sigma_n^2 I)$, see the model (2.1). We present simulation results for the OBER, the MLSD (using the Viterbi algorithm) and the matched filter receiver in terms of the measured Bit-Error Rate (BER) and the measured Sequence-Error Rate (SER), for various SNRs. We have used the time-invariant channel impulse response $h_k$ given in Table 1 in this example. (This is actually the normalized GSM “Rural Area (non-hilly)” test channel power profile [13, p. 137], here used as static ISI coefficients.) For each simulated SNR value enough blocks of 10 data bits ($N = 10$) were transmitted for each receiver to make a minimum of 4500 errors.

### TABLE 1: THE CHANNEL IMPULSE RESPONSE

<table>
<thead>
<tr>
<th>$h[1]$</th>
<th>$h[6]$</th>
<th>$h[0.1]$</th>
<th>$h[0.2]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6043</td>
<td>0.3813</td>
<td>0.2406</td>
<td>0.1518</td>
</tr>
<tr>
<td>0.0958</td>
<td>0.0604</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of the simulations are presented in Figure 4 and Figure 5. To the left of Figure 4 you find the BER for the three receivers, and to the right the corresponding SER. Numerical values of the BER and the SER are given in tables 2 and 3, respectively, where WMFR refers to the whitened matched filter receiver. In Figure 5 the bit decoding difference rate between the OBER and the MLSD respectively the WMFR is plotted versus SNR. With “bit decoding difference rate” we mean the ratio of bits on which the decision by the OBER differs from the detector of comparison.

### TABLE 2: BIT-ERROR RATE - BER IN % (see Figure 4)

<table>
<thead>
<tr>
<th>SNR</th>
<th>-20dB</th>
<th>-15dB</th>
<th>-10dB</th>
<th>-5dB</th>
<th>0dB</th>
<th>2dB</th>
<th>4dB</th>
<th>6dB</th>
<th>8dB</th>
<th>10dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBER</td>
<td>46.7</td>
<td>41.3</td>
<td>39.5</td>
<td>30.9</td>
<td>20.9</td>
<td>16.1</td>
<td>10.6</td>
<td>5.6</td>
<td>1.9</td>
<td>0.4</td>
</tr>
<tr>
<td>MLSD</td>
<td>46.7</td>
<td>41.5</td>
<td>39.8</td>
<td>32.1</td>
<td>21.5</td>
<td>17.2</td>
<td>11.2</td>
<td>5.7</td>
<td>1.9</td>
<td>0.4</td>
</tr>
<tr>
<td>WMFR</td>
<td>46.7</td>
<td>41.4</td>
<td>39.1</td>
<td>32.0</td>
<td>25.4</td>
<td>22.1</td>
<td>20.8</td>
<td>20.0</td>
<td>19.1</td>
<td>18.6</td>
</tr>
</tbody>
</table>

### TABLE 3: SEQUENCE-ERROR RATE - SER IN % (see Figure 4)

<table>
<thead>
<tr>
<th>SNR</th>
<th>-20dB</th>
<th>-15dB</th>
<th>-10dB</th>
<th>-5dB</th>
<th>0dB</th>
<th>2dB</th>
<th>4dB</th>
<th>6dB</th>
<th>8dB</th>
<th>10dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBER</td>
<td>99.7</td>
<td>99.7</td>
<td>99.0</td>
<td>97.3</td>
<td>90.7</td>
<td>77.7</td>
<td>58.6</td>
<td>33.1</td>
<td>11.6</td>
<td>2.3</td>
</tr>
<tr>
<td>MLSD</td>
<td>99.7</td>
<td>99.5</td>
<td>99.2</td>
<td>96.9</td>
<td>88.5</td>
<td>75.3</td>
<td>55.6</td>
<td>30.8</td>
<td>11.3</td>
<td>2.3</td>
</tr>
<tr>
<td>WMFR</td>
<td>99.7</td>
<td>99.7</td>
<td>99.7</td>
<td>98.6</td>
<td>96.7</td>
<td>95.3</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>93.5</td>
</tr>
</tbody>
</table>

As can be predicted from section 4.2 the OBER and the MLSD are close in performance. The gain in BER by using the OBER instead of the MLSD is here very small, even when the decision regions differ substantially. The receivers will make different decisions when the received signal lies in those areas where the decision regions of the two receivers do not overlap, which are small areas in the centre part of the hypercube. The probability of the receivers making a different decision
on any individual bit is calculated as the integral of the part of the noise probability mass covering the areas where the decision regions do not overlap, with the noise probability density centered at the transmitted sequence. When the SNR is large these areas are small, since the receivers’ decision surfaces are close together. When the SNR is very low, the size of these areas is significant, but the covering probability mass becomes thin, because the high variance of the noise flattens the Gaussian distribution $\psi(x_{zfe}|b)$. The maximum difference in bit-error probability lies somewhere in between these extremes. (For many channels the largest difference occurs around 0dB SNR.)

For low signal to noise ratios the matched filter receiver has a lower BER than the MLSD, although the difference is small. For SNRs below, say 0dB in this example, all three receivers have roughly equally high BERs. It should be noted, though, that the complexity of the whitened matched filter receiver is typically much lower than the complexity of any of the other receivers. Observe that when the ISI is severe (corresponding to a closed so called “signal eye”), as in the above example, the probability of bit-error for the matched filter receiver will not go to zero, even when no noise is present. In these cases, the whitened matched filter receiver is not a consistent estimator of the transmitted sequence.

6 Summary

Assuming that all sequences are equally probable we have shown that the MLSD is identical to the OBER designed for an infinite SNR. Consequently, the MLSD will asymptotically possess the property of minimum bit-error probability, as stated in [16, 18]. It is to be remembered that the MLSD minimizes the sequence-error probability, and is thus not by construction concerned with the bit-error probability. We have also shown that the matched filter receiver and the OBER coincide in the
The OBER and the MLSD have similar performances and often make the same decisions on individual bits. Simulation experiments indicate that the difference in performance for many channels and most SNRs is small. This agrees with the fact that the bit-error rate of the MLSD is close to upper and lower bounds on the bit-error probability, as is discussed in, e.g., [6, 14, 16, 18].

All facts considered, there is little motivation in terms of performance for using the OBER instead of the MLSD as a practical receiver. Whichever could be implemented at the lowest cost or complexity is likely to be preferred. This presently favours the MLSD in the form of the Viterbi detector. However, the OBER, with its displayed structure, constitutes a powerful tool for studying sufficient statistics and operations of receivers such as the MLSD and the matched filter receiver as demonstrated in sections 3 and 4.

Acknowledgements

We would like to acknowledge the improvements provided by Dr. Sarah Kate Wilson, Purdue University, Paul Petersen, West Lafayette, USA and Lic. Tech. Tomas Nordström, Luleå University of Technology.

APPENDICES

A The Minimum Bit-Error Probability Receiver

Let $\psi(x_{ZFE}|\beta, H)$ denote the probability density function for $x_{ZFE}$ given that the sequence $\beta \in C$ is transmitted and the hypothesis $H$, and $\psi(x_{ZFE}|\beta)$ denote the
probability density function for \( x_{zFE} \) given that the sequence \( \beta \in C \) is transmitted. Also, let \( \Pr \{ b = \beta | H \} \) be the probability for the sequence \( \beta \) being transmitted, given the hypothesis \( H \). Note that the geometric formulation (3.7) states in a sense a pair of composite hypotheses [25, pp. 86-87]. Using the above notations, we obtain

\[
\Gamma (x_{zFE}|H_1) = \sum_{\beta \in C} \psi(x_{zFE}|\beta, H_1) \Pr \{ b = \beta | H_1 \} = \sum_{\beta \in C^+_k} \psi(x_{zFE}|\beta) \frac{\Pr \{ b = \beta \}}{\Pr \{ H_1 \}} \tag{A.1}
\]

and

\[
\Gamma (x_{zFE}|H_0) = \sum_{\beta \in C} \psi(x_{zFE}|\beta, H_0) \Pr \{ b = \beta | H_0 \} = \sum_{\beta \in C^-_k} \psi(x_{zFE}|\beta) \frac{\Pr \{ b = \beta \}}{\Pr \{ H_0 \}} \tag{A.2}
\]

Using (A.1) and (A.2) the Bayes test in (3.8) can be written as

\[
\sum_{\beta \in C^+_k} \psi(x_{zFE}|\beta) \Pr \{ b = \beta \} \overset{H_1}{\geq} \sum_{\beta \in C^-_k} \psi(x_{zFE}|\beta) \Pr \{ b = \beta \} \tag{A.3}
\]

or, equivalently,

\[
\sum_{\beta \in C^+_k} w(x_{zFE}, \beta) \overset{H_1}{\geq} 0, \tag{A.4}
\]

where

\[
w(x_{zFE}, \beta) = \psi(x_{zFE}|y|\beta) \Pr \{ b = \beta \} - \psi(x_{zFE}|y| - \beta) \Pr \{ b = -\beta \}. \tag{A.5}
\]

From the preceding we observe that \( w(x_{zFE}, \beta) \) is independent of \( k \). The result of the hypothesis test is given by the sign of the left hand side of (A.4) so that

\[
(b_{\text{ORDER}})_{k} = \text{sign} \left( \sum_{\beta \in C^+_k} w(x_{zFE}, \beta) \right). \tag{A.6}
\]

Note that in the process of rewriting (A.3) to (A.4) we have grouped the sequences \( \beta \) into pairs \((\beta, -\beta)\), where \( \beta \in C^+_k \) and consequently \(-\beta \in C^-_k \). There could be other useful ways of grouping the sequences into pairs, giving other receiver structures. The density \( \psi(x_{zFE}|\beta) \) as given in (3.11) is immediately obtained from (2.3) and the first properties of multivariate Gaussian distributions.

In the important special case of equal \textit{a priori} probabilities, \( \Pr \{ b = \beta \} = 2^{-N} \) for all \( \beta \in C \) we obtain using (3.11) in (A.5) that

\[
w(x, \beta) = \frac{1}{2^N} A(x, \sigma_n) \exp \left(-\frac{\beta^T M \beta}{2\sigma_n^2}\right) \sinh \left(\beta^T M x / \sigma_n^2\right), \tag{A.7}
\]

where

\[
A(x, \sigma_n) = 2 \mathcal{K} (\sigma_n) \exp \left(-\frac{x^T M x}{2\sigma_n^2}\right) > 0 \tag{A.8}
\]

and \( \mathcal{K} (\sigma_n) \) is readily obtained as (3.12).
References


Part I.2

A Fast, Iterative Detector
Making MLSD Decisions on Scattered Bits
Part 1.2:
A Fast, Iterative Detector Making MLSD Decisions on Scattered Bits
Per Ödling\textsuperscript{1} and Håkan Eriksson\textsuperscript{1}
Research Report TULEA 1995:12, Division of Signal Processing, Luleå University of Technology.

\textsuperscript{1}Luleå University of Technology, Division of Signal Processing, S-971 87 Luleå, Sweden
A Fast, Iterative Detector Making MLSD Decisions on Scattered Bits

Per Ödling  
E-mail: per.odling@sm.luth.se

Håkan B. Eriksson  
E-mail: hakan.eriksson@sm.luth.se

Luleå University of Technology  
Division of Signal Processing  
S-971 87 Luleå, Sweden

May 1995

Abstract

This paper presents a fast detector of binary, antipodally modulated data that has been corrupted by intersymbol interference and additive Gaussian noise. This detector will make the same decisions as a maximum likelihood sequence detector on scattered bits in a transmitted sequence. It is simple in structure, consisting of a whitened matched filter and two variable thresholds for each bit to be detected. The thresholds are dependent on the received signal and are calculated using an iterative method. Both the number of bits to be decoded and their positions in the sequence are stochastic, varying from sequence to sequence. Because the proposed detector leaves bits undetermined and has a low bit error rate on the decoded bits, we believe its primary use would be in combination with low-complexity, sub-optimal receivers. Simulation results demonstrate that the probability of bit error of a decision-feedback equalizer is decreased if this detector is used as a pre-processor. They also demonstrate that the performance of linear equalizers can be improved substantially.

Key words: digital block transmission, intersymbol interference, maximum likelihood sequence detection, fast decoding.
1 Introduction

We investigate the detection of blocks of independent antipodally modulated, binary data transmitted over a discrete-time, additive Gaussian channel with intersymbol interference (ISI). We present a fast detector that makes the same decisions as the maximum likelihood sequence detector (MLSD) [6] on scattered bits in the transmitted block. Our use of the phrase “decisions on scattered bits” signifies that only some of the bits in each block are decoded, the remaining bits left undetermined, where the number of decoded bits and their positions are stochastic and vary from one block to another. The most important property of the detector, its decision rate, which we define as the average probability that a decision is made on any particular bit, depends on the ISI as well as the SNR. Because the detector makes MLSD decisions on the decoded bits, these bits have a low probability of error.

Even when implemented with the Viterbi algorithm [6], the MLSD will, in many applications be too computational complexity to be attractive. Our proposed detector offers an efficient way to obtain MLSD decisions on some random bits. The complexity of the proposed detector is stochastic with an average that depends on the ISI. The detector consists of a whitened matched filter and two variable thresholds for each bit. The thresholds are dependent on the received signal and are calculated using an iterative method, where the number of iterations is stochastic but never exceeds the number of bits in the block.

Perhaps our detector’s most important potential use is as an aid to sub-optimal receivers. But besides this possible practical value, it also possesses the conceptual value of helping us to see that some bits are particularly reliable, reliable in the sense that they are MLSD decisions. This could, for instance, be used as soft information when decoding an error correcting code. However, in this paper we will restrict ourselves to discussing two principal methods of combining the detector with other receivers, and to giving simulation examples which use one of these two principal methods, the one we advocate.

Examples of sub-optimal receivers for block transmission systems that could be combined with the proposed detector can be found in literature on digital communication. In [7] Kaleh describes four sub-optimal receivers for block transmission systems: the zero-forcing block linear equalizer, the minimum mean-square error block linear equalizer, the zero-forcing block decision-feedback equalizer and the minimum mean-square error block decision-feedback equalizer. The zero-forcing block linear equalizer is also derived by Barbosa in [1] where also the MLSD is discussed in the context of combined linear-Viterbi detectors [10] for block transmission systems, see also [4, 12].

The presentation proceeds as follows. In Section 2 a model for block transmission systems is described and some auxiliary definitions are given. The proposed detector is derived in Section 3 and described with an explicit algorithm. An attempt at geometrical visualization of its operation is also made in terms of decision regions, and an example of how the thresholds are calculated is given. Section 4 examines the behaviour of the detector, especially its decision rate. Two methods of how to implement the proposed detector in combination with sub-optimal receivers are discussed in Section 5. Also, in this section, simulation results are presented which demonstrate that the probability of bit error of a decision-feedback equalizer would be decreased if it were combined with our proposed detector. It is also shown that the
performance of linear equalizers can be improved substantially. Finally, in Section 6 we summarize our main results.

2 A block transmission system model

Consider the transmission of blocks of binary data, typically interspersed with blocks of bits known to the receiver, through a channel with additive Gaussian noise and known intersymbol interference (ISI). We represent the resulting transmission system with the matrix notation

\[ y = Hb + n, \]

where the transmitted message is coded in \( b \in \{-1, +1\}^N \), \( y \in \mathbb{R}^{N+L-1} \) is a vector of channel observables, \( H \) is a real-valued, deterministic and known \((N+L-1) \times N\) matrix representing the ISI, and \( n \) is a jointly Gaussian, zero-mean random vector with a \( N(0, R_n) \) distribution [1, 7]. We assume the bits in \( b \) to be independent, identically distributed with an equal probability of +1 and -1 occurring.

If the matrix \( H \) represents a linear, time-invariant and causal system, the structure of \( H \) becomes

\[
H = \begin{pmatrix}
h_0 & 0 & \ldots & \ldots & \ldots & 0 \\
h_1 & h_0 & \ldots & \ldots & \ldots & 0 \\
\vdots & h_1 & \ddots & \ldots & \ldots & 0 \\
h_{L-1} & \ldots & \ldots & \ldots & h_0 & \ldots \\
0 & \ldots & h_{L-1} & \ldots & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & h_{L-1} & \\
\end{pmatrix},
\]

(2.2)

where \([h_0, h_1, \ldots, h_{L-1}]\) is the impulse response of the system. If \( h_0 \neq 0 \) and \( h_{L-1} \neq 0 \) then \( L \) is the length of the channel memory.

We continue with a number of additional definitions that we will find useful in what follows. The signal-to-noise ratio (SNR) in this block transmission system we define as

\[
\text{SNR} \triangleq \frac{\text{tr}\{H^TR_n^{-1}H\}}{N},
\]

(2.3)

where \( \text{tr}\{X\} \) denotes the trace of the matrix \( X \), i.e., the sum of the elements on the main diagonal of the matrix \( X \).

The *likelihood function* for \( b \) given \( y \) is the probability density function for \( y \) given \( b \). The MLSD is any processor that, given the received signal \( y \), finds the sequence \( b \) for which the *likelihood function* for \( b \) given \( y \) reaches its maximum. Since all sequences are equally probable to be transmitted, the MLSD* is optimal as measured with the probability of sequence error and asymptotically achieves minimum probability of bit error for high SNR [13]. Defining the matrix \( M = H^TR_n^{-1}H \) and using the model (2.1) with the multivariate Gaussian distribution for \( y \) given \( b \) as \( N(Hb, R_n) \), the MLSD is minimizing

\[
L(b, y) \triangleq \|y - Hb\|_{R_n^{-1}}^2 = \|b - M^{-1}H^TR_n^{-1}y\|_M^2 + \|y\|_{R_n}^2 - \|H^TR_n^{-1}y\|_{M^{-1}}^2.
\]

(2.4)

*The MLSD is often realized as the Viterbi detector [6].
with respect to $b$, where the norm $\| \cdot \|_M$ is defined as $\|x\|^2_M \triangleq x^T M x$. Let $\hat{b}^{\text{MLSD}}$ denote the output of the MLSD, the sequence in $\{-1, +1\}^N$ found to minimize (2.4). The last two terms on the right hand side of (2.4) are independent of $b$. Thus, finding the sequence $\hat{b}^{\text{MLSD}}$ is equivalent to finding the sequence that minimizes the first term on the right hand side of (2.4):

$$\hat{b}^{\text{MLSD}} = \arg \min_{b \in \{-1, +1\}^N} L'(b, z),$$

(2.5)

where

$$L'(b, z) \triangleq \| b - M^{-1} z \|^2_M,$$

(2.6)

and where

$$z \triangleq H^T R_n^{-1} y = M \hat{b}_u = Mb + H^T R_n^{-1} n$$

(2.7)

is the output from the whitened matched filter for block transmission systems (WMF) [1, 7, 13].

3 The proposed detector

From equation (2.6) we will draw conclusions that are sufficient to devise the mentioned detector. Let $b_i, \hat{b}_i^{\text{MLSD}}, z_i$ and $m_{i,j}$ denote the elements in $b, \hat{b}^{\text{MLSD}}, z$ and $M$, respectively, where $i$ is the row index and $j$ is the column index. Focusing on the detection of the $k$th bit, we rewrite (2.6) in Appendix A and obtain

$$L'(b, z) = 2b_k (\Delta(b, k) - z_k) + g(z, b, k),$$

(3.1)

where

$$\Delta(b, k) \triangleq \sum_{i \neq k} b_i m_{i,k}$$

(3.2)

and

$$g(z, b, k) \triangleq \sum_{i \neq k, j \neq k} b_i m_{i,j} b_j - 2 \sum_{i \neq k} b_i z_i + m_{k,k} + z^T M z.$$  

(3.3)

Note that both (3.3) and (3.2) are independent of $b_k$, and that

$$|\Delta(b, k)| \leq \sum_{i \neq k} |b_i m_{i,k}| = \sum_{i \neq k} |m_{i,k}|.$$  

(3.4)

If the output of the WMF at instant $k$ is above the possible maximum of $\Delta(b, k)$, i.e.,

$$z_k > \sum_{i \neq k} |m_{i,k}|,$$

the function $L'(b, z)$ in (3.1) is minimized by $b_k = +1$, independently of the status of the other bits. Hence, if the value in position $k$ of the output of the whitened matched filter is above the positive of the threshold, the corresponding output of the MLSD is positive, $\hat{b}_k^{\text{MLSD}} = +1$. Similarly, if $z_k < -\sum_{i \neq k} |m_{i,k}|$, then $\hat{b}_k^{\text{MLSD}} = -1$. This is true for all positions $k$, and the tests can be done independently for all bits. These observations are sufficient for attempting to decode some bits in a block using only a matched filter and a threshold device. Such a decoding would constitute the first iteration of the proposed detector.

Assume now that some of the bits in $b$ have been decoded as described above. The decisions on decoded bits, once made, does not change due to any subsequent decisions on still-undecoded bits. Using decoded bits as constants, we proceed by
minimizing (3.1) with respect to the remaining (undetermined) bits. With the same technique as described above, the function $\Delta(b, k)$ in (3.2) can be bounded from above by

$$\Delta^+_k = \sum_{i | i \neq k, b_i \text{ decoded}} m_{i,k} \hat{b}^{\text{MLSD}}_i + \sum_{i | i \neq k, b_i \text{ not decoded}} |m_{i,k}|$$

and from below by

$$\Delta^-_k = \sum_{i | i \neq k, b_i \text{ decoded}} m_{i,k} \hat{b}^{\text{MLSD}}_i - \sum_{i | i \neq k, b_i \text{ not decoded}} |m_{i,k}|,$$

where the difference between (3.5) and (3.6) is the sign of the second sum. Comparing the output of the matched filter, $z_k$, with these new and tighter thresholds, additional bits could possibly be decoded. We can again check whether $z_k > \Delta^+_k$, and if it is, then $\hat{b}^{\text{MLSD}}_k = +1$. Analogously, if $z_k < \Delta^-_k$ then $\hat{b}^{\text{MLSD}}_k = -1$. This procedure can be repeated until no more bits are decoded. We conclude that the decisions on the bits that actually are decoded are identical to the corresponding decisions of an MLSD.

The output of the detector is ternary, "$-1$", "$+1$" or "$\text{no decision}$", with the value of all decoded bits equal to the corresponding output of an MLSD. The actual identity and the total number of bits that can be determined depend on the received signal, $y$, and on the channel characteristics through the matrix $M$. Put together into an algorithm, the above procedure is a description of the proposed detector.

### 3.1 An algorithm for the implementation of the detector

We present an iterative algorithm that formally describes the decoding procedure. Let $\hat{b}^{(l)} = [\hat{b}_1^{(l)}, \hat{b}_2^{(l)}, \ldots, \hat{b}_N^{(l)}]^T$ denote the output of the proposed detector after the $l$th iteration and let $\hat{b}_\Delta$ denote its final output. Furthermore, let $\Delta^+_k (\hat{b}^{(l-1)})$ and $\Delta^-_k (\hat{b}^{(l-1)})$ denote the positive and the negative thresholds at iteration $l$, respectively. One algorithm describing the decoding operation is then:

1. $l = 0$
2. $\hat{b}_k^{(0)} = 0$, $\forall k$  

REPEAT

3. $l = l + 1$
4. Calculate the positive threshold:  

$$\Delta^+_k (\hat{b}^{(l-1)}) = \sum_{i \neq k} m_{i,k} \hat{b}_i^{(l-1)} + \sum_{i \neq k} |m_{i,k}| \left(1 - |\hat{b}_i^{(l-1)}|ight), \forall k$$
5. Calculate the negative threshold:  

$$\Delta^-_k (\hat{b}^{(l-1)}) = \sum_{i \neq k} m_{i,k} \hat{b}_i^{(l-1)} - \sum_{i \neq k} |m_{i,k}| \left(1 - |\hat{b}_i^{(l-1)}|ight), \forall k$$
6. Compare and decode:

$$\hat{b}_k^{(l)} = \begin{cases} 1 & z_k > \Delta^+_k (\hat{b}^{(l-1)}) \\ -1 & z_k < \Delta^-_k (\hat{b}^{(l-1)}) \\ 0 & \text{otherwise} \end{cases}, \forall k$$
UNTIL \( \| \hat{b}^{(l)} \|_1^2 = N \) or \( \hat{b}^{(l)} = \hat{b}^{(l-1)} \)

7. \( \hat{b}_\Delta = \hat{b}^{(l)} \)

Note that the output of the detector \( \hat{b}_\Delta \), when implemented by the algorithm described above, belongs to \( \{-1, 0, +1\}^N \), where a zero signifies that the detector did not make a decision on the corresponding bit. The algorithm will continue to iterate until either none of the remaining bits are decoded or there are no more bits to be decoded. One other way of terminating the algorithm’s iterative process would be to iterate exactly \( N \) times every time.

### 3.2 An example of the operation of the detector

To illustrate how the proposed detector calculates the thresholds and makes decisions, we present an example where blocks of binary data of length \( N = 20 \) are transmitted over a time-invariant channel described by (2.1). The impulse response of the system, as in (2.2), is \( [h_0, h_1, h_2] = [1, 0, 1] \), the noise \( n \) is white, stationary and Gaussian with an SNR of 10 dB. In Figure 2 an example of the iterations of the detector is given. The solid lines are the thresholds, \( \Delta_k^\pm(\hat{b}^{(l-1)}) \) and \( \Delta_k^\mp(\hat{b}^{(l-1)}) \), the stars and the circles are the output of the matched filter \( z \), with the circles denoting symbols decoded at iteration \( l \) and the stars undetermined symbols. The upper and lower thresholds will merge when a sufficient number of bits in a sequel have been decoded. In the example, the bits at \( k \) equal to 15, 17 and 19 could not be decoded. Note that bit 12 was decoded as positive although the output from the whitened matched filter was negative. In this example the detector made correct decisions on all bits that were decoded. As we will see later, this example is typical for this channel in the sense that the proposed detector frequently iterates four times.

The detector can be seen as a whitened matched filter followed by two thresholds, which may be different for every bit \( k \) and vary with every iteration, cf. Figure 1. The thresholds are dependent on \( z \) and \( M \) and have to be re-calculated for every block. If the noise \( n \) is white and no ISI is present, the thresholds are all equal to zero and the detector will be equivalent to the optimal threshold detector.

![Figure 1: A block description of the proposed detector.](image)

If the proposed detector is viewed as a matched filter followed by a decision device, the algorithm presented above can be seen as an iterative procedure to calculate the final thresholds defining this decision device. In the example displayed in Figure 2, the thresholds obtained after the third iteration, that is, the thresholds shown in the diagram belonging to the fourth iteration, are the final thresholds to be used.
Figure 2: The output of the whitened matched filter plotted together with the thresholds of the proposed detector in an example where four iterations are needed to obtain $\hat{b}_\Delta$. Bits that are decoded are marked “o” and undecoded bits are marked “*”.

### 3.3 A Geometrical Interpretation

If we compare the decision regions of our proposed detector to the decision regions of the MLSD, we can illustrate not only how these two detectors relate to each other, but also how the proposed detector performs its first iteration in calculating a threshold. Let us study the detectors’s decision regions for the sufficient statistics

$$\hat{b}_u(y) \triangleq M^{-1}H^TR_n^{-1}y,$$

(3.7)

in $\mathbb{R}^N$, where $M = H^TR_n^{-1}H$. The matrix operation $M^{-1}H^TR_n^{-1}$ is an instance of the Moore-Penrose pseudo-inverse, and can be seen as the zero-forcing equalizer for block transmission systems (ZFEB) [1, 7]. This can be illustrated by inserting (2.1) in (3.7) giving $\hat{b}_u = b + M^{-1}H^TR_n^{-1}n$. It is shown by Barbosa in [1] that $\hat{b}_u(y)$ is a sufficient statistic for $b$ given the model in (2.1). Because both the MLSD and the proposed detector can be derived as operating on $\hat{b}_u(y)$, they are completely described by their respective decision regions for $\hat{b}_u(y)$ in $\mathbb{R}^N$.

We plot decision regions for $\hat{b}_u(y)$ when transmitting blocks of two bits over a
time-invariant channel with channel impulse response \([h_0, h_1] = [1, 1]\), see (2.1) and (2.2), and with white, stationary and Gaussian noise \(n\). The left part of Figure 3 contains the decision surface of the MLSD for the detection of the first bit in the block. The four possible transmitted sequences are marked with circles. The set of possible sequences is often referred to as the binary hypercube\(^1\), in this case the 2-dimensional binary hypercube. If \(\hat{b}_u \in \mathcal{D}_1^+\) then bit one of \(\hat{b}_{MLSD}\) is equal to +1, otherwise it is −1. In the right part of Figure 3 the corresponding decision regions of the proposed detector are shown. If \(\hat{b}_u \in \mathcal{O}_1^+\) or \(\hat{b}_u \in \mathcal{O}_1^-\) the output for bit one is equal to +1 and −1, respectively. If \(\hat{b}_u \in \mathcal{O}_1^0\) then the detector is unable to make a decision. In the example, it is clear from Figure 3 that the bits decoded by the proposed detector will be the identical to the corresponding decisions made by the MLSD. The decisions made in the first iteration would also be the same as the decisions made by the minimum bit error probability receiver, which can be shown by using the technique of Appendix A on the results in Appendix A of [13].

![Diagram showing decision regions for the MLSD and the proposed detector.](image)

Figure 3: Examples of 2-dimensional decision regions for the statistics \(\hat{b}_u\) and for the detection of the first bit in a two-bit block. The left part shows the decision regions of the MLSD and the right part the decision regions of the proposed detector, with the \([x_1, x_2]\) as coordinates in IR\(^2\). (Note that the angle of the decision boundaries depends on the matrix \(M\). With this particular impulse response the boundaries happen to cross \([+1, -1]\) and \([-1, +1]\).)

In Figure 4 we show an example of a 3-dimensional decision surface of the MLSD for the detection of the first bit in a block of three bits. As can be noted in both the 2D and 3D pictures there are \(2^{N-1}\) parallel hyperplanes extending from the interior of the hypercube. If these planes had passed through origo, they would have represented the operation of the whitened matched filter, i.e., the sign of the output of the filter at instant \(k = 1\) would have told whether \(\hat{b}_u\) is above or below this hyperplane [13]. It can be observed that, for every bit, the signal space can be divided into three parts by extending the two most offset of the parallel hyperplanes, the uppermost and the lowermost. If the output of the matched filter at some instant \(k\) is larger than this offset, we know that the MLSD would decide on the corresponding bit being positive. Analogously, if the output is below the negative of the same offset, the corresponding output of the MLSD is −1.

\(^1\)Discussions on the geometry of the binary hypercube can be found in, e.g., [8, 9, 13].
Figure 4: An example of a 3-dimensional decision surface of the MLSD for the statistics $\hat{b}_u$ and for the detection of the first bit in a block of three bits. The $[x_1, x_2, x_3]$ are coordinates in $\mathbb{R}^3$.

4 Performance analysis

In section 3 we concluded that the decisions made by the proposed detector are the same as the decisions made by the MLSD. This, we feel, is our detector's most important property. Following this in importance would be its decision rate, its average probability of making a decision on any individual bit. We will first give a definition of the "decision rate" and then, making use of this definition, proceed to use simulations to investigate this and other aspects of the proposed detector's behaviour. The simulations will illustrate that the detector's decision rate is high enough for the detector to be of practical value.

We define the probability that a decision has been made on bit $k$ after iteration $l$ as

$$P_{D,k}^{(l)} \triangleq \Pr \{ \tilde{b}_k^{(l)} \neq 0 \} = 1 - \Pr \{ \tilde{b}_k^{(l)} = 0 \},$$

(4.1)

where $\tilde{b}_k^{(l)}$ denotes the decision on the $k$th bit after the $l$th iteration, see the algorithm described in Section 3.1. The probability that the proposed detector makes a decision on bit $k$ is $P_{D,k} \triangleq P_{D,k}^{(N)}$, where $N$ is the length of the block and the maximal number of iterations needed to process the block. Finally, the decision rate is defined as

$$P_D \triangleq \frac{1}{N} \sum_{k=1}^{N} P_{D,k}.$$  

(4.2)
The decision rate is, of course, dependent on the ISI, as well as on the covariance matrix of the noise. For instance, if no ISI is present and the noise \( n \) is white, then the detector reduces to the optimal threshold detector [13], and decodes all bits.

In appendix B an expression for the probability that the detector has decided on bit \( k \) after iteration \( l \), \( P_{D,k}^{(l)} \), is derived. Evaluating this expression is a computationally demanding task. Besides integrating densities over \( N \)-dimensional boxes, it requires the evaluation and summation of a large number of such integrals. Because an analytic evaluation of the decision rate, \( P_0 \), seems to be a complicated task, it is only addressed by simulations in this paper.

4.1 Studying the behaviour of the detector by simulations

As noted above, the potential of the proposed detector is very dependent on its ability to make decisions. This section presents simulation results where the proposed detector is used to detect bits in a block transmission system. Consider a time-invariant two-tap channel described by (2.1) and (2.2) with

\[
    h_k = \delta(k) + \alpha \delta(k-d),
\]

where \( \delta(\cdot) \) is the Kronecker delta function, \( \alpha \) is the real-valued amplitude of the second tap and \( d \), a positive integer, is the time-spread of the channel. Throughout these simulations the noise \( n \) is white, stationary and Gaussian; the block length \( N = 20 \).

The Figures with the simulations results discussed in this section are collected in Appendix C. In Figure 6 the estimated probability of no decision, \( 1 - P_0 \), is plotted versus the strength of the second tap, \( \alpha \), for the delay \( d \in [1, 2, 5, 10, 15] \) and an SNR of 10 dB. It can be observed that the probability of no decision, given a certain \( d \), has global maxima at around \( \alpha = 1 \) and \( \alpha = -1 \); it is low if \( |\alpha| \) is small or large enough, and also non-increasing with \( d \). If \( d \geq N \) the two received echoes are separable, effectively a no-ISI case, and consequently the detector is able to detect all bits.

Figures 7 to 10 show various data from a simulation varying the SNR in a two-tap channel model with impulse response \( [h_0, h_1, h_2] = [1, 0, 1] \), as described by the model (4.3) with \( d = 2 \) and \( \alpha = 1 \). Among the two-tap channels, this is the worst case as given by Figure 6.

Figure 7 shows the estimated probability that bit number \( k \) is decoded, \( P_{D,k} \), for varying SNR. The plot shows the probability of decision for each individual bit \( k \in [1, 2, \ldots, 20] \) separately. The four bits at the edges of the block, i.e., the bits corresponding to \( k \in [1, 2, 19, 20] \), have the highest probability of being decoded. This is because these bits are subjected to less ISI than the bits in the centre part of the block. As these bits are decoded more frequently, their neighbours, the bits at \( k \in [3, 4, 17, 18] \) are also decoded more frequently. It can be observed that the probability of decision in general decreases with increasing SNR.

In Figure 8 the bit error rate (BER) of the MLSD is plotted versus SNR, with the BER divided into two parts; the BER of the bits that could have been decoded by the proposed detector and the BER of the bits that would have been left undetermined. Note that the bits that could have been decoded have a BER that is a fraction of the BER of the other bits. In this sense the detector apparently finds bits that are "easy" to decode correctly.
Figure 9 presents the number of iterations needed to calculate the output of the proposed detector. The percentage of blocks that needed a certain number of iterations is plotted versus SNR and the number of iterations. With this channel the detector most frequently used three iterations and rarely more than five.

In Figure 10 the received blocks are sorted after the number of decoded bits. The percentages of blocks with certain numbers of bits decoded are plotted versus SNR. For this two-tap channel the complete sequence is quite often decoded when the SNR is low. Observe that if nineteen bits are decoded, the twentieth is always decoded, too, since no ISI remains then.

5 The proposed detector as a pre-processor

We will discuss two principal methods of using the proposed detector in combination with other receivers. The first, and more obvious, is to run the detector in parallel with some other receiver, using the bits decoded by the proposed detector when such are available, otherwise using bits from the complementary receiver. A further enhancement for a decision-feedback equalizer (DFE) would be to let the “known” bits replace the corresponding bits at the output of the threshold detector, i.e., the output of the DFE that is also fed back in the feedback loop. Correspondingly, sub-optimal receivers that are based on the Viterbi algorithm [6], e.g., combined linear-Viterbi equalizers [12, 4, 2, 5, 11, 10], can be used by restraining the trellis in the positions corresponding to the “known” bits.

However, simulation experiments indicate that a principally different technique for combining receivers is preferable. Say that the proposed detector is used to decode bits in a system modelled with (2.1). Assuming that the estimated bits are correct, we subtract their influence from the received signal. In order to detect the remaining bits, consider the random vector

\[ x = y - H \hat{b}_A, \]

where \( y \) is the channel output from (2.1) and \( \hat{b}_A \in \{-1, 0, +1\}^N \) denotes the ternary output of the proposed detector, see Figure 5. Let \( G \) be a matrix containing the columns of \( H \) corresponding to those bits that were not decoded. It follows from the structure of \( G \) and (5.1) that the random vector \( x \) given \( b \) (and given correct decisions) is Gaussian with \( x \sim N(Gc, R_n) \), where \( c \) is a column vector containing the undetermined bits. Equivalently, \( x \) can be modelled as

\[ x = Gc + n, \]

where \( n \) is the same noise vector as in (2.1), thus \( n \sim N(0, R_n) \). Note that if no bits were decoded then all elements in \( \hat{b}_A \) are equal to zero, \( G = H \) and \( x = y \). To decode the undetermined bits we can use (5.1) and (5.2) by giving the vector \( x \) and the characteristics of the channel model (5.2), i.e., \( G \) and \( R_n \) to any detector intended for block transmission systems. Note that \( G \) in general will represent a time-variant system even when the original channel model \( H \) represented a time-invariant channel.

There are many different types of block transmission receivers that could be applied to this reduced system (5.2). One type would be linear equalizers [1, 7],
which could, for instance, be based on the minimum mean-square error [7] or zero-forcing criteria [1, 7]. Another type would be the DFEs described in [7]. Combined linear-Viterbi equalizers for block transmission systems capable of handling time-variant systems [4], could also operate directly on the reduced system (5.2).

When a decision made by the proposed detector is erroneous, i.e., when a decoded bit differs from the corresponding transmitted bit, the reduced system of equation (5.2) is an incorrect model of the signal \( x \). As a consequence, the complementary detector will be inclined to make additional erroneous decisions on the remaining bits. However, this error propagation is likely to be less severe than, for instance, the error propagation in DFEs, because the decisions made by the proposed detector have a lower BER.

In the following section we combine the proposed detector with four receivers described by Kaleh in [7]: two linear equalizers and two decision-feedback equalizers, one of each based on the minimum mean-square error and zero-forcing criteria, respectively.

### 5.1 Simulations of receiver combinations

To illustrate the potential of the proposed detector as a pre-processor, we present results from simulations of block transmission systems with four different sub-optimal, complementary receivers, see Figure 5. The four sub-optimal receivers are: the zero-forcing linear equalizer (ZFE), the minimum-mean-square error equalizer (MMSE), the zero-forcing decision-feedback equalizer (ZF-DFE) and the minimum-mean square error decision-feedback equalizer (MMSE-DFE), all described by Kaleh in [7].

In the simulations, blocks of data of length \( N = 20 \) are transmitted over a time-invariant dispersive channel with additive, stationary, white, Gaussian noise, as described by (2.1). The impulse response of the system is again \([h_0, h_1, h_2] = [1, 0, 1]\). Figures containing the simulation results are presented in Appendix D.

Simulation results in terms of the BER versus SNR are presented in Figures 11-12, with the performance of the MLSD included as a reference. Note that all the complementary receivers improve their performance when our proposed detector is used as a pre-processor. As an example, the performance of the ZFE is improved by approximately 4 dB at a BER of \(10^{-3}\). In contrast to the unaided case, the ZFE combined with the proposed detector has a slightly better performance than

---

\[ z' = G^T R_n^{-1} x. \]

---
the corresponding combination with the MMSE. At first glance, this may seem to contradict other known results, see e.g., [7]. It should be noted, however, that because the proposed detector sometimes makes errors, the model (5.2) used for the two equalizers might be incorrect. In this example this would seem to penalize the MMSE more than the ZFE, but this need not always be the case, especially when the decision rate is low.

In Figure 12 the performances of the ZF-DFE and the MMSE-DFE with and without the proposed detector are compared. The performances of both receivers improve with the pre-processor, although the improvements are much smaller than the improvements for the linear equalizers. This agrees with the DFEs being much closer to the MLSD in performance, making the potential for improvements smaller.

Figures 13–16 contain BER curves where the BER has been divided into two parts; the BER of the bits that were decoded by the proposed detector and the BER of the remaining bits. As can be observed, the bits decoded by the pre-processor have a lower BER than the other bits. Note also that the BER of the bits decoded by the complementary receivers operating on the reduced system (5.2) is improved compared to the corresponding bits from the unaided receivers.

6 Conclusions

We have presented a detector of low computational complexity that makes MLSD decisions on a portion of the transmitted bits. The key components of the detector are a whitened matched filter and a threshold device with two variable thresholds per bit.

We believe that the primary use of the proposed detector is as a complement to sub-optimal receivers inferior in performance to the MLSD. It can aid these receivers either by providing MLSD decisions on scattered bits, or, as in our examples, by letting the complementary receivers operate on a reduced system model. Simulation experiments using the latter technique show an improvement of the BER in both the bits decoded by the pre-processor and the bits decoded by the complementary receivers.

An important property of the proposed detector, determining its potential, is its decision rate: the average percentage of bits on which a decision is made. The percentage of bits decoded in a block is stochastic and varies between 0% and 100%, with the average dependent on the SNR and the channel characteristics. Simulations using a block length $N = 20$ rendered a decision rate better than 65% for all two-tap channel models, while channel models with strong ISI and a high SNR could render decision rates in the order of a few ten percents.

One other potentially useful property of the proposed detector is the existence of unequal error probabilities between the pre-processor and the adjoined receiver. It might, for example, be useful in some applications (e.g., those involving soft decoding information) to know that some bits have been decoded with the MLSD criterion. This property, not developed in this paper, is a subject for future research.
APPENDICES

A The log-likelihood function

Let \( b_i, z_i \) and \( m_{i,j} \) denote the elements in \( b, z \) and \( M \), respectively, where \( i \) is the row index and \( j \) is the column index. Then

\[
L'(b, z) = \left\| b - M^{-1}z \right\|_M^2
\]

\[
= b^TMb - 2z^Tb + z^TM^{-1}z
\]

\[
= \sum_{i,j} b_i m_{i,j} b_j - 2 \sum_i b_i z_i + z^TM^{-1}z
\]

\[
= \sum_{i \neq k, j \neq k} b_i m_{i,j} b_j + \sum_{j \neq k} b_i m_{i,j} b_j + \sum_{i \neq k} b_i m_{i,k} b_k + b_k m_{k,k} b_k - 2 \sum_{i \neq k} b_i z_i - 2 b_k z_k
\]

\[
+ z^TM^{-1}z
\]

\[
= b_k \left( \sum_{i \neq k} b_i m_{i,k} + \sum_{j \neq k} m_{k,j} b_j - 2 z_k \right) + \sum_{i \neq k} b_i m_{i,j} b_j - 2 \sum_{i \neq k} b_i z_i + |b_k|^2 m_{k,k}
\]

\[
+ z^TM^{-1}z
\]

\[
= 2b_k \left( \sum_{i \neq k} b_i m_{i,k} - z_k \right) + m_{k,k} + \sum_{i \neq k} b_i m_{i,j} b_j - 2 \sum_{i \neq k} b_i z_i + z^TM^{-1}z. \tag{A.1}
\]

We will use this in the form

\[
L'(b, z) = 2b_k (\Delta(b, k) - z_k) + g(z, b, k), \tag{A.2}
\]

where

\[
\Delta(b, k) \triangleq \sum_{i \neq k} b_i m_{i,k} \tag{A.3}
\]

and

\[
g(z, b, k) \triangleq \sum_{i \neq k, j \neq k} b_i m_{i,j} b_j - 2 \sum_{i \neq k} b_i z_i + m_{k,k} + z^TMz, \tag{A.4}
\]

see equation (3.1).

B The probability of decision

Let \( \Pr\{\tilde{b}_k^{(l)} = 0 | b = c\} \) denote the probability that \( \tilde{b}_k^{(l)} = 0 \) given that the sequence \( c \) was transmitted. The probability of decision for bit \( k \) at step \( l \) can then be written as

\[
P_{D,k}^{(l)} = 1 - \sum_{c \in \{-1, +1\}^N} \Pr\{b = c\} \Pr\{\tilde{b}_k^{(l)} = 0 | b = c\}, \tag{B.1}
\]
where \( \Pr \{ b = c \} \) is the probability that the sequence \( c \) is transmitted. The probability
\[
\Pr \left\{ \tilde{b}_k^{(l)} = 0 \mid b = c \right\} = 
\sum_{a \in \{-1,0,+1\}^N} \Pr \left\{ \tilde{b}^{(l-1)} = a \mid b = c \right\} \Pr \left\{ \tilde{b}_k^{(l)} = 0 \mid \tilde{b}^{(l-1)} = a, b = c \right\},
\]
\( (B.2) \)
where \( \Pr \left\{ \tilde{b}_k^{(l)} = 0 \mid b^{(l-1)} = d, b = c \right\} \) is the probability that \( \tilde{b}_k^{(l)} = 0 \), given \( b^{(l-1)} = d \) and given that \( c \) was transmitted, and can be calculated as we will show later.

The probability \( \Pr \left\{ \tilde{b}^{(l-1)} = a \mid b = c \right\} \) can be computed recursively as
\[
\Pr \left\{ \tilde{b}^{(l-1)} = a \mid b = c \right\} = 
\sum_{d \in \{-1,0,+1\}^N} \Pr \left\{ \tilde{b}^{(l-2)} = d \mid b = c \right\} \Pr \left\{ \tilde{b}^{(l-1)} = a \mid \tilde{b}^{(l-2)} = d, b = c \right\},
\]
\( (B.3) \)
starting with
\[
\Pr \left\{ \tilde{b}^{(0)} = d \mid b = c \right\} = \begin{cases} 1 & \text{if } d = 0 \\ 0 & \text{otherwise} \end{cases}.
\]
\( (B.4) \)
Calculating the probability \( \Pr \left\{ \tilde{b}_k^{(l)} = a \mid \tilde{b}^{(l-1)} = d, b = c \right\} \) involves the integration of an \( N \)-dimensional density function over an \( N \)-dimensional box. Note that this density is not jointly Gaussian.

It remains to calculate the probability \( \Pr \left\{ \tilde{b}_k^{(l)} = 0 \mid b^{(l-1)} = a, b = c \right\} \). Let \( \tilde{z}_k \) denote the \( k \)th element in \( \tilde{z} \) and \( m_{k,k} \) the \( k \)th diagonal element in \( M \). Then, if \( \tilde{b}_k^{(l-1)} = 0 \),
\[
\Pr \left\{ \tilde{b}_k^{(l)} = 0 \mid \tilde{b}^{(l-1)} = a, b = c \right\} = 
\Pr \left\{ \Delta_k^- \left( \tilde{b}^{(l-1)} \right) < z_k < \Delta_k^+ \left( \tilde{b}^{(l-1)} \right) \mid \tilde{b}^{(l-1)} = a, b = c \right\} = 
\int_{\Delta_k^-}^{\Delta_k^+} \frac{1}{\sqrt{2\pi m_{k,k}}} \exp \left( -\frac{(x - \tilde{z}_k)^2}{2m_{k,k}^2} \right) dx,
\]
\( (B.5) \)
which can be evaluated using the \( Q(\cdot) \)-function [3]. Otherwise, if \( \tilde{b}_k^{(l-1)} \neq 0 \) then \( \Pr \{ \tilde{b}_k^{(l)} = 0 \mid \tilde{b}^{(l-1)} = a, b = c \} = 0 \).
C Simulations: The behaviour of the detector

The following pictures are described in section 4.1.

Figure 6: The probability of no decision, $1 - P_D$, for the proposed detector plotted versus the strength of the second tap, $\alpha$, for a delay $d \in [1, 2, 5, 10, 15]$.

Figure 7: The probability of decision for each individual bit $k$, $P_{D,k}$ versus SNR.
Figure 8: The bit error rate of the MLSD for the bits decoded by the proposed detector (dashed line) and for the undetermined bits (solid line) plotted versus SNR.

Figure 9: The percentage of blocks that needed a certain number of iterations is plotted versus SNR and the number of iterations. The right part of the figure shows the same data as the left part but with a logarithmic scale.
Figure 10: The percentage of blocks where a certain number of bits were decoded is plotted versus SNR and the number bits. The right part of the figure shows the same data as the left part but with a logarithmic scale.
D Simulations: Performances of receiver combinations

The following Figures are described in section 5.1.

Figure 11: The bit error rate of the aided and unaided linear receivers versus SNR.

Figure 12: The bit error rate of the DFE based receivers versus SNR.
Figure 13: The BER of the ZFE (left) and the ZFE combined with the proposed detector (right) plotted versus SNR. The two curves in the left part of the Figure are the BER of the bits that could have been decoded by the proposed detector (dashed line) and the BER of the bits that would have been left undetermined (solid line). The bits in the right part are identically grouped. However, here the proposed detector was used and the ZFE was operating on the reduced system model (5.2).

Figure 14: The BER of the MMSE (left) and the MMSE combined with the proposed detector (right) plotted versus SNR. The two curves in the left part of the Figure are the BER of the bits that could have been decoded by the proposed detector (dashed line) and the BER of the bits that would have been left undetermined (solid line). The bits in the right part are identically grouped. However, here the proposed detector was used and the MMSE was operating on the reduced system model (5.2).
Figure 15: The BER of the ZF-DFE (left) and the ZF-DFE combined with the proposed detector (right) plotted versus SNR. The two curves in the left part of the Figure are the BER of the bits that could have been decoded by the proposed detector (dashed line) and the BER of the bits that would have been left undetermined (solid line). The bits in the right part are identically grouped. However, here the proposed detector was used and the ZF-DFE was operating on the reduced system model (5.2).

Figure 16: The BER of the MMSE-DFE (left) and the MMSE-DFE combined with the proposed detector (right) plotted versus SNR. The two curves in the left part of the Figure are the BER of the bits that could have been decoded by the proposed detector (dashed line) and the BER of the bits that would have been left undetermined (solid line). The bits in the right part are identically grouped. However, here the proposed detector was used and the MMSE-DFE was operating on the reduced system model (5.2).
References


Part I.3

A Genie-Aided Detector Based on a Probabilistic Description of the Side Information
Part I.3:
A Genie-Aided Detector Based on a Probabilistic Description of the Side Information
Håkan Eriksson¹, Per Ödling¹, Timo Koski² and Per Ola Börjesson¹
A shorter version has been accepted for presentation at the 1995 IEEE International Symposium on Information Theory, Whistler, Canada, September 1995.

¹Luleå University of Technology, Division of Signal Processing, S-971 87 Luleå, Sweden
²Luleå University of Technology, Department of Mathematics, S-971 87 Luleå, Sweden
A Genie-Aided Detector Based on a Probabilistic Description of the Side Information

Håkan B. Eriksson¹ Per Ödling¹ Timo Koski² Per Ola Börjesson¹

Luleå University of Technology
¹ Division of Signal Processing
² Department of Mathematics
S-971 87 Luleå, Sweden
April 1995

Abstract

We consider lower performance bounds for the detection of independent blocks of bits transmitted over discrete-time Gaussian channels with intersymbol interference. Building on Forney’s concept of the genie [9, 10], and introducing the idea of an explicit statistical description of the side information provided to the genie-aided detector, we develop a mathematically strict derivation of a lower bound for the bit error rate of any actual receiver. This statistical description makes the lower bounding a simple, transparent application of basic Bayesian detection theory. With this approach, the side information statistics become a design parameter, which may be chosen to give the bound a desired structure. To illustrate this, we choose statistics in order to obtain a special case: the lower bound derived by Mazo [20].

Our derivation’s results differ from Forney’s. However, the significance of our work is not found in the bounds it produces; it is found in the application of a formal structure when producing such bounds. This formal structure is more general, more accessible for future analysis, and, we feel, offers an instructive view on Forney’s work.

Key words: digital block transmission, intersymbol interference, genie aided detection, lower bounds, bit error rate.
1 Introduction

Over the past decades much effort has been put into developing digital receivers for channels with intersymbol interference and additive noise. Because an exact analysis of such a receiver’s performance is generally quite difficult, frequent use is made of the concept of upper and lower performance bounds or of a variety of simulation techniques, see, e.g., [4, 9, 10, 12, 13, 20, 21] and the references therein.

The idea of a good genie with a corresponding genie-aided detector (GAD) has, in particular, often been used to determine a lower performance bound for the probability of bit error [1, 4, 9, 10]. The GAD has access to more information than any actual detector: it has access to the side information supplied by the genie and is expected to handle all information optimally. Because of this, it is argued, it cannot have a worse performance than any detector working without the side information. However, in order that optimal processing of the side information be well-defined in the sense of Bayesian detection theory, an explicit (statistical) description of the side information is required. This paper introduces such a representation of the genie, augmenting the foundational ideas of Forney’s work in [9, 10]. In [20] Mazo derives a lower bound by an analytic calculation without the intervention of a genie, which under some conditions coincides with Forney’s bound. As a special case, we show that when the statistics of the side information are properly chosen, the performance of the optimal GAD is identical to Mazo’s bound.

This presentation is given in a matrix terminology describing a block transmission system environment. Our essential aim is to introduce the side information supplied by the genie as the output of a “side information channel” parallel to the original channel and governed by a probabilistic rule with free parameters. We will demonstrate that the GAD using the side information optimally cannot be outperformed. The analysis of the performance of this optimal detector provides a lower bound on the bit error probability of any actual receiver that lacks access to the output of the side information channel. Furthermore, a lower bound to the performance of the GAD itself, can be analytically derived by using the Bhattacharyya coefficient [5, 16, 14]. Thus the notion of the GAD can be used for a lower bound to the performance of an actual receiver, and the Bhattacharyya coefficient can in turn be used for a lower bound to the performance of the GAD.

The presentation proceeds as follows. In Section 2 we present a model of a transmission system incorporating a stochastic description of a side information channel. Section 3 presents and analyses the GAD that has minimum probability of making an error when detecting a specific bit in the transmitted sequence, with special attention given to the case that the side information consists of pairs of sequences. Section 4 describes strategies for choosing the statistics, or transition probabilities, of the side information. We will, in particular, discuss those statistics for which the performance of the optimal GAD matches Mazo’s lower bound [20]. In Section 5 we relate our work in this paper to the work by Forney [9, 10]. A numerical example is given in Section 6. And in Section 7 we conclude with a summary of our main results.
2 The side information channel

Consider a transmission system where binary data is sent through a discrete-time, additive Gaussian channel (DTGC) with intersymbol interference (ISI), and where additional side information is carried to the detector through a parallel channel (representing the genie) as depicted in Figure 1. This channel, the "genie’s channel", we refer to as the "side information channel".

Let \( b \in \{-1, +1\}^N \) denote the \( N \times 1 \) dimensional stochastic variable containing the \( N \) independent bits to be transmitted in each block. Then the DTGC with ISI and binary block transmission [2, 8, 15, 23] can be described in terms of the representation

\[
y = Hb + n,
\]

(2.1)

where \( H \) is an \((N + L - 1) \times N\) real valued, deterministic matrix given in terms of the ISI coefficients of the channel with \( L \) as the length of the channel memory, \( n \) is an \((N + L - 1) \times 1\) dimensional, jointly Gaussian stochastic variable with a \( N(0, R_n) \) distribution, and \( y \in \mathcal{Y} \) is the \((N + L - 1) \times 1\) dimensional, real valued stochastic variable that is observed by the receiver.

Let the output of the side information channel, \( z \), be a stochastic variable and denote the outcome of \( z \) as \( \zeta \). Furthermore, let \( \zeta \) take on the values in a finite set \( \Omega \) of \( M \) possible outcomes, i.e., \( \zeta \in \Omega = \{\zeta_1, \zeta_2, \ldots, \zeta_M\} \). Also let \( \beta, \eta \) and \( \rho \) denote the outcomes of the stochastic variables \( b, y \) and \( r = (y, z) \), respectively. The joint channel output \( r \) takes its values in the space \( \mathcal{Y} \times \Omega \). The transmission system described above is completely characterized by \( f_{r|b}(\cdot) \) and \( \Pr\{b = \beta\} \), where \( f_{r|b}(\cdot) \) is the conditional distribution for \( r \) given \( b \), and \( \Pr\{b = \beta\} \) is the probability of \( \beta \) being transmitted.

Note that the above model contains a novel concept: the probability distribution \( f_{z|b,y}(\cdot) \) is a user chosen distribution. Additional properties of \( f_{z|b,y}(\cdot) \) and examples of strategies for choosing the distribution are given later.

3 The Genie-Aided Detector

We will formulate our communication problem as a hypothesis test, deciding between hypotheses on \( b \). A detector or receiver for the DTGC is, for the purposes of this discussion, any processor that both divides the observation space \( \mathcal{Y} \) into decision regions and decides in favour of a hypothesis if the outcome of \( y \) lies in the corresponding decision region. Great effort has been put into developing receivers
for the DTGC with ISI, and closely related channel models. We will refer to these actual receivers as receivers acting without side information.\(^1\) Detectors acting with aid of the side information \(z\) as well as the channel observable \(y\) for deciding between the hypotheses we refer to as genie-aided detectors, or, more frequently, GADs.

**Proposition 3.1** A GAD (a detector acting with the aid of side information) deciding between hypotheses and making optimal use of all available information has equal or superior performance compared to every detector acting without side information.

We give a formal proof of proposition 3.1 for binary hypotheses testing in Appendix A.

This statement, known from, e.g., [3, p. 350] or [9, 10], is intuitively plausible. A commonly given rationale for proposition 3.1 is that any detector optimally exploiting both the channel output and the side information must perform as well as, or better than, any similarly placed detector lacking access to the side information. (Optimal handling of the augmented information is defined explicitly in terms of the probabilistic description of the side information channel.) Consequently one interpretation of proposition 3.1 is that the performance of the optimal GAD determines a lower bound on the probability of error for all detectors lacking access to the side information \(z\).

### 3.1 The detection of bit number \(k\)

Let us now turn our attention to analysing the bit error probability of receivers. Consider the detection of \((b)_{bit_k}\), the \(k\)th bit in a sequence \(b\). We define the two hypotheses, \(H_0\) and \(H_1\), formally as the binary hypotheses testing problem:

\[
H_0 : (b)_{bit_k} = -1 \\
H_1 : (b)_{bit_k} = +1.
\]  

(3.1)

The blocks of bits, \(b\), take on the values in the binary hypercube \(C \triangleq \{-1, +1\}^N\), cf. [2, 17]. By partitioning the hypercube \(C\) into two subsets of sequences of symbols according to

\[
C_+^k \triangleq \{\beta_1^+, \beta_2^+, \ldots, \beta_{2^N-1}^+ \mid \beta_i^+ \in C, \ (\beta_i^+\)_{bit_k} = +1\}
\]

\[
C_-^k \triangleq \{\beta_1^-, \beta_2^-, \ldots, \beta_{2^N-1}^- \mid \beta_i^- \in C, \ (\beta_i^-\)_{bit_k} = -1\},
\]  

(3.2)

the hypotheses testing problem may be restated as \(H_0 : b \in C_-, \ H_1 : b \in C_+\). We can now see that our transmission model with a side information channel is realized by the two probability distributions \(\{f_{r|H_i}(\rho|H_i)\}_{i=0}^1\), with

\[
f_{r|H_0}(\rho|H_0) = f_{y,z|H_0}(\eta, \zeta|H_0) = \sum_{\beta \in C_-^k} f_{y,z|b}(\eta, \zeta|\beta) Pr\{b = \beta\} Pr\{H_0\},
\]

(3.3)

\(^1\)Receivers acting without side information can be defined as receivers operating with decision regions that are independent of the side information \(z\), cf. Appendix A.
where \( f_{y,z|b}(\eta, \zeta | \beta) \) is a joint probability expression for \( y \) and \( z \) given \( b = \beta \). Note that this is a mixed probability density, as it is a density in \( y \) and a probability for the random events \( z \). The probability \( \Pr \{H_0\} = 1 - \Pr \{H_1\} \) is the probability of the symbol hypothesis \( H_0 \) occurring. In the examples to follow an additional constraint is introduced: the channel output \( y \) and the side information \( z \) are conditionally independent given \( b \). Therefore:

\[
f_{y,z|H_0}(\eta, \zeta | H_0) = \sum_{\beta \in \mathbb{C}} \frac{\Pr \{z = \zeta | b = \beta\} f_{y|b}(\eta | \beta) \Pr \{b = \beta\}}{\Pr \{H_0\}},
\]

where \( \Pr \{z = \zeta | b = \beta\} \) is the probability for a specific side information \( \zeta \) given that \( \beta \) was transmitted. The analogous rules hold for \( f_{y|H_1}(\rho | H_1) \).

We wish to point out that the performance of an optimal GAD will be dependent on the probability distribution \( \Pr \{z = \zeta | b = \beta\} \). This probability distribution can, thus, be seen as a design parameter. That is, \( \Pr \{z = \zeta | b = \beta\} \) may, for example, be chosen to give a tight bound, or to give the bound some explicit, desirable structure. In Section 4 we will give examples of this.

### 3.2 The probability of bit error

The procedure of detecting bit number \( k \) in \( b \) was in (3.1) defined as a hypothesis testing problem; thus, the probability of the GAD making an error when detecting bit number \( k \) in \( b \) is given by the probability of deciding on the wrong hypothesis. Let us denote the probability of detection error of an actual detector \( P_{DE,k} \). The total probability of bit error in a block is defined as the average of the probability of detection error in a specific bit, \( P_{DE,k} \), over all bits in a block \( b \):

\[
P_{BER} = \frac{1}{N} \sum_{k=1}^{N} P_{DE,k}.
\]

To calculate \( P_{DE,k} \) we can use the Bayesian rule, i.e., averaging over all possible outcomes of the transmitted block \( b, \beta \in \mathbb{C} \). This gives us

\[
P_{DE,k} = \sum_{\beta \in \mathbb{C}} \Pr \{\beta\} \Pr \{(\bar{b})_{bit,k} \neq (\beta)_{bit,k} | b = \beta\},
\]

where \( \Pr \{(\bar{b})_{bit,k} \neq (\beta)_{bit,k} | b = \beta\} \) is the probability that the detector in question (with the output \( \bar{b} \)) makes an erroneous decision on bit \( k \) given that \( \beta \) was transmitted.

We proceed with an inequality on \( P_{DE,k} \) and therefore, through the relation (3.5), on \( P_{BER} \). It follows from proposition 3.1 that

\[
P_{DE,k} \geq P_{DE,k}^{GA},
\]

where \( P_{DE,k}^{GA} \) denotes the bit detection error probability of the GAD that uses the side information in the *optimal* fashion, i.e., for a Bayes decision. The inequality (3.7) is the precise statement that a detector taking its decisions by exploiting optimally the side information cannot be outperformed by any other detector [3, p. 350]. The calculation of \( P_{DE,k}^{GA} \) and \( P_{BER}^{GA} \) through (3.5), then gives a lower bound on the bit error probability of any receiver acting without side information.

Using the relation (3.5) here prescribes the use of \( N \) parallel GADs, each with a different, independent side information channel. The \( N \) GADs would then decode one bit each and their joint output would be the estimate of the entire sequence.
3.3 The side information as a pair of sequences

In this section we will discuss the important special case when the side information consists of a pair of sequences and one of the sequences is equal to the transmitted sequence, cf. [9, 10]. The rationale for this discussion is that the probability of bit error of the corresponding GAD, using all information to minimize $P_{DE,k}^{GA}$, can be derived using sums of the standard error function, $Q(\cdot)$.

For a specific position $k$, let the outcomes of the side information $z$ consist of pairs in $C_k^+ \times C_k^-$ of the form $\zeta_{i,j} \triangleq (\beta_i^+, \beta_j^-)$, for $1 \leq i, j \leq 2^{N-1}$. With the transmitted sequence being $\beta$, let the additional sequence be chosen at random, independently of $n$, among the sequences differing from $\beta$ in bit $k$, according to the known, probabilistic transition rule:

$$\Pr\{z = \zeta_{i,j}|b = \beta\} = \begin{cases} p(j|i) & \text{if } \beta = \beta_i^+ \in C_k^+ \\ q(i|j) & \text{if } \beta = \beta_j^- \in C_k^- \\ 0 & \text{otherwise} \end{cases}. \quad (3.8)$$

Hence, the properties of the genie, or equivalently, the properties of the output $z$ of the side information channel, are defined by the statistics (or transition probabilities) $p(j|i)$ and $q(i|j)$. The GAD and the associated lower bound can now be designed by choosing the transition probabilities. Note, however, that the following conditions must apply if (3.8) is to be a true probabilistic transition rule:

$$0 \leq p(j|i), q(i|j) \leq 1, \quad \forall (i, j),$$

$$\sum_j p(j|i) = 1, \quad \forall i,$$

$$\sum_i q(i|j) = 1, \quad \forall j. \quad (3.9)$$

If the transition probabilities are chosen such that the side information $\zeta_{i,j}$ appears only when one of the two sequences in $\zeta_{i,j}$ is transmitted, then that particular side information $\zeta_{i,j}$ uniquely determines what sequence was transmitted. Such an outcome of the side information forces the optimal detector to always decide on the correct value of the bit to be decoded. We refer to this as singular detection.

Proposition 3.2 Assume conditional independence of $y$ and $z$ given $b$ and the transition rule (3.8). The GAD achieving the minimal probability of error when detecting bit number $k$ after observing $p = (\eta_i, \zeta_{i,j})$ is then given by the Bayesian test

$$f_{y|b}(\eta|\beta_i^+) \leq \gamma_{i,j}, \quad (3.10)$$

where the threshold

$$\gamma_{i,j} = \frac{q(i|j) \Pr\{b = \beta_j^-\}}{p(j|i) \Pr\{b = \beta_i^+\}}. \quad (3.11)$$

and $f_{y|b}(\cdot)$ is the multivariate Gaussian density function for $y$ given $b$.

A proof of proposition 3.2 is found in Appendix B.

As the GAD described by proposition 3.2 achieves the minimal probability of error when detecting bit number $k$ we henceforth refer to it as the optimal GAD.
Proposition 3.3 The probability of error in bit \( k \) of the GAD in proposition 3.2 is given by

\[
P^{GA}_{DE,k} = \sum_{i,j} Q \left( \frac{d_{i,j}}{2} + \frac{\ln \gamma_{i,j}}{d_{i,j}} \right) q(i|j) \Pr \{ b = \beta_{j}^{-} \} + \sum_{i,j} Q \left( \frac{d_{i,j}}{2} - \frac{\ln \gamma_{i,j}}{d_{i,j}} \right) p(j|i) \Pr \{ b = \beta_{i}^{+} \},
\]

where

\[
d_{i,j} \triangleq \sqrt{\left( \beta_{i}^{+} - \beta_{j}^{-} \right)^{T} H^{T} R_{n}^{-1} H \left( \beta_{i}^{+} - \beta_{j}^{-} \right)}
\]

and \( Q(x) \triangleq \left( \frac{1}{\sqrt{2\pi}} \right) \int_{x}^{\infty} e^{-t^{2}/2} dt \).

Proposition 3.3 is derived in Appendix C.

As a consequence of proposition 3.1, the performance of the GAD of proposition 3.2, \( P^{GA}_{DE,k} \) as given in (3.12), serves as a lower bound for the performance of any receiver acting without the aid of side information. This holds for all valid choices of the statistics \( p(j|i) \) and \( q(i|j) \). However, because the calculation of (3.12) requires summing over \( 2^{2N-2} \) terms, using (3.12) as a lower bound is often impractical.

4 Design of Lower Bounds

In this section we will discuss strategies for how to choose the transition probabilities of the side information channel for the purpose of using the performance of the optimal GAD as a lower bound. The transition probabilities of the side information channel, \( p(j|i) \) and \( q(i|j) \), can be seen as design parameters for the lower bound given by the performance expression of proposition 3.3. The problem is how to choose these transition probabilities, given the conditions in (3.9), so that the lower bound becomes tight and easy to compute. In subsections 4.1 to 4.3 three examples of methods for choosing the transition probabilities are described and in Subsection 4.4 a lower bound based on the Bhattacharyya coefficient is given.

4.1 Optimizing the side information channel

Corollary 4.1 The GAD, in the class of detectors described by proposition 3.2, whose performance gives the tightest bound on the bit error probability is found by maximizing (3.12) with respect to \( p(j|i) \) and \( q(i|j) \) as

\[
P^{GA}_{DE,k} \triangleq \max_{\{p(j|i)q(i|j)\}_{i,j}} P^{GA}_{DE,k}
\]

given the conditions in (3.9) and where \( P^{GA}_{DE,k} \) is given in (3.12).

Thus, (4.1) optimizes the transition probabilities of the side information channel so that the performance of the optimal GAD becomes as bad as possible. The bound corresponding to \( P^{GA}_{DE,k} \) in corollary 4.1 has, in general, high computational complexity. The number of parameters in the optimization grows exponentially with the length of the transmitted sequence. Also, the location of the optimum of \( P^{GA}_{DE,k} \)
and it’s value, $P_{DE,k}^{GA}$, are dependent on the specific impulse response of the channel and on the noise covariance. Therefore, the maximization has to be repeated for every signal-to-noise ratio (SNR) and for every $k$.

### 4.2 A bound using a single $Q(\cdot)$-function - Mazo’s bound

In this section we use the present framework to re-derive Mazo’s bound which only involves evaluating a single $Q(\cdot)$-function. This will lead us to the customary sequences of minimum distance [1, 3, 9, 10, 20], that is, the pairs of sequences achieving the minimum distance in equation (3.13). For high SNRs, the receiver confusing sequences of minimum distance will often be the dominating error events due to the exponential decrease of the Gaussian probability density. The minimum distance is thus defined as

$$d_{min} \triangleq \min_{i,j} d_{i,j}, \quad (4.2)$$

with $d_{i,j}$ given in (3.13). Note that this $d_{min}$ involves the noise covariance $R_n$ and, thus, increases with the SNR\(^1\). The set of pairs of sequences that are at minimum distance is

$$D_k \triangleq \{(i,j) | d_{i,j} = d_{min}\}. \quad (4.3)$$

**Lemma 4.2**

I: Choose $B_k$ as a non-empty subset of $D_k$, such that no sequence (with another sequence at minimum distance) is included in more than one pair $(\beta_i^+, \beta_j^-)$ for all $(i,j) \in B_k$. Then $p(j|i) = q(i|j) = 1$ is a valid choice of transition probabilities for all $(i,j) \in B_k$.

II: With the transition probabilities $\{p(j|i), q(i|j) | (i,j) \in B_k\}$ chosen as above, it is possible to choose all other transition probabilities, $\{p(j|i), q(i|j) | (i,j) \not\in B_k\}$, so that the optimal detector will make singular detection whenever $z = \zeta_{i,j}$ for $(i,j) \not\in B_k$.

The first part of lemma 4.2 just groups sequences into disjoint pairs; a choice of transition probabilities that does not violate the conditions of (3.9). When a sequence associated with the set $B_k$ is transmitted the corresponding unique pair in $B_k$ (of which the transmitted sequence is a member) is always given as side information. The singular detection of part two can be realized in many ways. As the set of sequences $\{\beta_i^+, \beta_j^- | (i,j) \in B_k\}$ is a non-empty set; one can always pick an arbitrary sequence from this set belonging to, e.g., $C_k^+$ and use it in pair with all the sequences outside $B_k$ belonging to $C_k^-$, when these latter are transmitted. As a result, singular detection occur in the optimal GAD for all sequences outside $B_k$, belonging to $C_k^-$. The transition probabilities for the sequences outside $B_k$ in $C_k^+$ can be chosen analogously.

Let us assume that all sequences are transmitted with equal probability, i.e., $\Pr\{b = \beta\} = 1/2^N$. If we apply the choice of transition probabilities described by lemma 4.2 to proposition 3.3 and include (3.7), we find that

$$P_{DE,k} \geq P_{DE,k}^{GA} = \frac{2|B_k|}{2^N} Q\left(\frac{d_{min}}{2}\right), \quad (4.4)$$

\(^1\)The minimum distance, as defined here by (4.2), is actually similar to the effective SNR of e.g., [11].
where $|B_k|$ denotes the number of elements in the set $B_k$. Note that this side information channel will act deterministically: the transmission of any given sequence always reveal the same side information $z$.

In [20] Mazo presents a bound by analysing the performance of the receiver that gives minimum bit error probability [23]. The bound is derived by a series of analytic simplifications, without the intervention of a genie, resulting in the right hand side of equation (4.4), with $|B_k|$ referred to as "the number of pairings". Although Mazo does not explicitly prescribe the use of the largest possible set $B_k$ in (4.4) we henceforth refer to (4.4) maximized over $B_k$ as Mazo’s bound. This bound is, in comparison with the bounds of sections 4.1 and 4.3, easy to evaluate and closely related to Forney’s work [9, 10], as will be discussed in Section 5.

### 4.3 Optimizing the side information channel for high SNR

As a comparison to Mazo’s bound we will use a side information channel with fixed transition probabilities, where these have been chosen to give a tight bound for high SNR. This is done by maximizing an approximation of (3.12) for an infinite SNR, giving the performance of the optimal GAD for asymptotically high SNR. The work in Subsection 4.2, [20] and [9, 10] also focuses on the high SNR performance and obtains, as we will see, similar results. The set of transition probabilities obtained in the following maximization will be inserted in equation (3.12), rendering a valid bound for any SNR.

We begin with a general limit to be applied to (3.12). If $p \neq 0$ and $q \neq 0$ then

$$\lim_{d \to \infty} \frac{1}{Q(d)} \left[ aQ \left( \frac{d}{2} + \frac{\ln(q/p)}{d} \right) q + bQ \left( \frac{d}{2} - \frac{\ln(q/p)}{d} \right) p \right] = (a + b)\sqrt{pq}. \quad (4.5)$$

To avoid $p = 0$ and/or $q = 0$ note that the pairs $\zeta_{i,j}$ for which singular detection occur do not contribute to the error probability and may be omitted in the sums of equation (3.12). Omitting these terms, let us consider the behaviour of (3.12) for large SNR by applying the result of (4.5) to each term in (3.12). This gives

$$P_{DE,k}^{GA} \approx 2 \sum_{(i,j) \in A_k} Q \left( \frac{d_{ij}}{2} \right) \sqrt{p(j|i) \Pr \{ b = \beta_i^+ \} q(i|j) \Pr \{ b = \beta_j^+ \}}, \quad (4.6)$$

for large SNR, where $A_k$ denotes the set of pairs $\zeta_{i,j}$ that occur with a probability greater than zero and are not subject to singular detection, i.e.,

$$A_k \triangleq \{(i, j) \mid p(j|i) \neq 0, q(i|j) \neq 0\}.$$

Note that because of the Gaussian kernel’s exponential decrease, the terms with the minimum distance $d_{\text{min}}$ will dominate the sum. On this basis we reduce (4.6) to

$$P_{DE,k}^{GA} \approx Q \left( \frac{d_{\text{min}}}{2} \right) 2 \sum_{(i,j) \in D_k} \sqrt{p(j|i) \Pr \{ b = \beta_i^+ \} q(i|j) \Pr \{ b = \beta_j^+ \}}, \quad (4.7)$$

where $D_k$ is given as the set of pairs of sequences at minimum distance in (4.3). The characteristics of the channel model are reflected by the contents of the set $D_k$ and the value of $d_{\text{min}}$.

\(^1\)The largest size of the set $B_k$ is in general difficult to determine. Estimating the largest $|B_k|$ in block transmission systems is discussed in e.g., [6].
Our intended set of transition probabilities are obtained by maximizing equation (4.7) with respect to \( p(j|i) \) and \( q(i|j) \) under the conditions given by equation (3.9). This maximization is easier to perform than maximizing (3.12) and gives parameters that can be inserted in (3.12) to give a lower bound for \( P_{\text{DE},k} \) for all values of SNRs. Also, this bound will, under certain conditions described in corollary 4.3, coincide with Mazo’s bound given in Subsection 4.2.

**Corollary 4.3** If \( B_k \) in lemma 4.2 can be chosen so that \( B_k = \mathcal{D}_k \) and all sequences have equal a priori probability, then the performance of the optimal GAD with a side information channel optimized for high SNR will coincide with the Mazo bound for bit \( k \).

**Proof:** Assume that all sequences have equal a priori probability to be transmitted. By Cauchy’s inequality for sequences of real numbers and because \( 0 \leq p(j|i) \leq 1 \) and \( 0 \leq q(i|j) \leq 1 \), we have

\[
\sum_{(i,j) \in \mathcal{D}_k} \sqrt{p(j|i)q(i|j)} \leq \sqrt{\sum_{(i,j) \in \mathcal{D}_k} p(j|i) \sum_{(i,j) \in \mathcal{D}_k} q(i|j)} \leq |\mathcal{D}_k| \tag{4.8}
\]

with equalities if \( q(i|j) = p(j|i) = 1 \) for all \( (i,j) \in \mathcal{D}_k \). In this way the right hand side of (4.7) is maximized if \( q(i|j) = p(j|i) = 1 \) for all \( (i,j) \in \mathcal{D}_k \). If \( B_k = \mathcal{D}_k \), then this is an admissible choice of transition probabilities. This corresponds to the choice of transitions probabilities producing Mazo’s bound, i.e., the transitions probabilities given in lemma 4.2.

### 4.4 Lower bounds using the Bhattacharyya Coefficient

There are several known ways to calculate a lower bound for the Bayesian probability of error when testing any pair of hypotheses [5, 16, 14]. Applying such a bound to the Bayesian hypotheses test of proposition 3.2 give a lower bound on the performance of the GAD. In a sense, this is a lower bound to a lower bound on the performance of actual receivers. The bound \( P_{\text{DE},k}^{\text{GA}} \geq \frac{1}{2} \left[ 1 - \sqrt{1 - 4h^2} \right] \), from [5, 14], is an example of this. This bound makes sense only if \( h \leq \frac{1}{2} \), where \( h \) is obtained as the Bhattacharyya coefficient

\[
h = \int_{\mathcal{Y} \times \Omega} \sqrt{f_{r|H}(\rho|H_0)\Pr\{H_0\}f_{r|H}(\rho|H_1)\Pr\{H_1\}}d\rho. \tag{4.9}
\]

In view of [18] we can interpret \( h \) as the conditional standard deviation of the mean square error estimate of \( (\beta)_{\text{bit},k} \). The coefficient \( h \) is evaluated for the DTCG with side information in Appendix D.

**Proposition 4.4** Assume the transition rule (3.8) and that the channel output \( y \) and the side information \( z \) are conditionally independent given \( b \). Assume also that

\[
h \leq \frac{1}{2}. \tag{4.10}
\]

Then the probability of error for the GAD of proposition 3.2 is bounded below as [5, 14]

\[
P_{\text{DE},k}^{\text{GA}} \geq \frac{1}{2} \left[ 1 - \sqrt{1 - 4h^2} \right]. \tag{4.11}
\]
where
\[
\bar{h} = \sum_{i,j} \sqrt{p(j|i) \Pr \{ b = \beta_i^+ \}} q(i|j) \Pr \{ b = \beta_j^+ \} e^{-\frac{d_{ij}^2}{8}}
\]  
(4.12)
and \(d_{ij}\) is given in (3.13).

The condition \(\bar{h} \leq \frac{1}{2}\) is obtained, e.g., in the white noise case, \(R_N = \sigma_n^2 I\), if \(\sigma_n\) is low enough. The lower bound of the proposition above then holds for sufficiently high SNR.

Under the current assumptions about the side information channel and, in addition, assuming a sufficiently high SNR, we observe the following further lower bound to the right hand side of the inequality (3.7).

**Corollary 4.5** Under the assumptions of proposition 4.4 we have
\[
P^\text{GA}_{\text{DE},k} \geq e^{-d_{\text{min}}^2 q} \left( \sum_{(i,j) \in \mathcal{D}_k} p(j|i) \Pr \{ b = \beta_i^+ \} q(i|j) \Pr \{ b = \beta_j^+ \} \right)^2.
\]  
(4.13)

The derivation of corollary 4.5 is outlined in Appendix D.

Corollary 4.5 is similar to the expression in (4.7). Consequently, the bound in (4.13) is maximized by the same transition probabilities that maximize (3.12) for asymptotically high SNR.

## 5 The Forney lower bound

We will discuss how the bounds of sections 4.2 and 4.3 relate to the Forney lower bound [9, 10]. For this comparison we need to assume that all sequences are transmitted with equal probability, i.e., \(\Pr \{ b = \beta \} = 2^{-N}\) for all \(\beta \) in \(C\).

Forney investigates the probability of the first error event in a continuous transmission system, and concludes that this is also a lower bound for the bit error probability. The lower bound is presented as
\[
P_{\text{DE},k} \geq \pi^c Q \left( \frac{d_{\text{min}}}{2} \right),
\]  
(5.1)
where \(\pi^c\) denotes the probability of compatibility for \(k = 1\), i.e., the probability that a transmitted sequence has a neighbour at minimum distance differing in the first bit, see, e.g., [6, 9, 10] and [3, pp. 152–153]. With our notations for finite sequences of length \(N\) this would be the probability that the transmitted sequence is represented in at least one of the pairs in \(\mathcal{D}_1\). Noting the similarity to equation 4.4, we give the following corollary.

**Corollary 5.1** If \(B_1\) in lemma 4.2 can be chosen so that \(B_1 = \mathcal{D}_1\) then \(\pi^c = 2 * |B_1|/2^N\) and Forney's bound will be equal to Mazo's bound for bit 1.

For the purpose of illustration we will give an example for which Mazo's and Forney's bounds differ.

**Example:**
Consider a transmission system with a sequence length of (at least) $N = 3$ and two different error patterns achieving minimum distance, say $\pm[2, 2, 0]$ and $\pm[2, 0, 2]$ as illustrated in Figure 2. The sequences of minimum distance differing in bit one, i.e., the pairs of sequences corresponding to $\mathcal{D}_1$, are indicated by a line connecting them with their minimum distance partner(s). The dotted lines correspond to the error pattern $\pm[2, 0, 2]$ and the solid to $\pm[2, 2, 0]$.

In the case illustrated in Figure 2 the probability of compatibility as given by Forney [9, 10] would be $\pi_c = 6/8$ and the constant given by Mazo [20], see Section 4.2, would be $4/8$.

![Figure 2: Sequences differing with the error patterns $\pm[2, 0, 2]$ (dashed lines) and $\pm[2, 2, 0]$ (solid).](image)

Within the framework presented in Section 3 the constant given by Mazo can be reached, cf. lemma 4.2, while the probability of compatibility $\pi_c$ as given in [9, 10] seems impossible to reach in some cases. Either the class of GADs presented in this paper is an insufficient tool, or Forney, perhaps, intended the constant $\pi_c$ to be interpreted as in equation (4.4). The latter, indeed, seems to be a common interpretation, see, e.g., [20] and [4, p. 153].

### 6 Examples

In this section we give a numerical example and display the corresponding lower performance bounds. We only consider the probability of bit error for the first bit in the block, i.e., $P_{DE,1}$. Let $N = 2$ and denote the possible outcomes of $b$, the sequences in $C$ as in equation (3.2), with

$$
\beta_1^+ = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \beta_2^+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \beta_1^- = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \beta_2^- = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.
$$

(6.1)

We further assume that all sequences are transmitted with equal a priori probability, i.e., $\Pr\{b = \beta\} = 2^{-N}$ for all $\beta$.

If the matrix $H$ represents a time-invariant system and the noise $n$ is stationary, white Gaussian noise with zero mean and variance $\sigma_n^2$, i.e., $n \in N(0, \sigma_n^2I)$, then the distance matrix, whose entries $\{d_{i,j}\}$ are the distances defined in (3.13), is

$$
D = 2\sqrt{\text{SNR}} \left[ \frac{1}{\sqrt{2(1+\delta)}} \begin{bmatrix} \sqrt{2(1-\delta)} \\ 1 \end{bmatrix} \right],
$$

(6.2)

where $\text{SNR} = 1/\sigma_n^2$ and $-1 < \delta < 1$. Note that the parameter $\delta$ holds the description of the dispersion in the channel $H$. From the definition (4.3) and equation (6.2) we
see that
\[ D_1 = \{(1,1), (2,2)\} \quad \text{if} \quad |\delta| < 0.5, \]
\[ D_1 = \{(1,2)\} \quad \text{if} \quad \delta > 0.5, \]
\[ D_1 = \{(2,1)\} \quad \text{if} \quad \delta < -0.5, \]
\[ D_1 = \{(1,2),(1,1),(2,2)\} \quad \text{if} \quad \delta = 0.5, \]
\[ D_1 = \{(2,1),(1,1),(2,2)\} \quad \text{if} \quad \delta = -0.5. \] (6.3)

The largest possible subset in accordance with lemma 4.2 of \( T > \) is given by
\[ B_1 = \{(1,1), (2,2)\} \quad \text{if} \quad |\delta| < 0.5, \]
\[ B_1 = \{(1,2)\} \quad \text{if} \quad \delta > 0.5, \]
\[ B_1 = \{(2,1)\} \quad \text{if} \quad \delta < -0.5, \]
\[ B_1 = \{(1,2),(1,1)\} \quad \text{if} \quad \delta = 0.5, \]
\[ B_1 = \{(1,2),(2,2)\} \quad \text{if} \quad \delta = -0.5. \] (6.4)

Applying to this example Forney's bound and the three bounds found in section 6, the performance of the optimal GAD with a side information channel optimized for all SNRs, the performance of the optimal GAD with a side information channel optimized for high SNR, and Mazo's bound, we note the following. If \( |\delta| \neq 0.5 \) then the largest possible \( B_1 = D_1 \) and the bounds presented by Mazo [20] and Forney [9, 10] will coincide, in accordance with corollary 5.1. Also the performance of the optimal GAD using a side information channel optimized for high SNR according to Subsection 4.3, will coincide with these bounds as given by corollary 4.3. From (6.4) and equation (4.4) we have
\[ P_{de,k}^{GA} = \begin{cases} Q \left( \frac{d_{2,1}}{2} \right) & |\delta| < 0.5, \\ \frac{1}{2} Q \left( \frac{d_{2,2}}{2} \right) & \delta > 0.5, \\ \frac{1}{2} Q \left( \frac{d_{2,2}}{2} \right) & \delta < -0.5. \end{cases} \] (6.5)

Figure 3 describes the behaviour of the side information channel corresponding to Mazo's bound through lemma 4.2, which in this same example also corresponds to the GAD optimized for high SNR. (Note that the behaviour of this side information channel is independent of the SNR.) All of the transition probabilities \( p(j|i) \) and \( q(i|j) \) are either one or zero. As a consequence this side information channel will act deterministically in the sense that, given a certain transmitted sequence, the genie will always give one particular pair as side information. Figure 3 indicates with arrows those branches with probability one; branches with probability zero are omitted.

Figures 4-6 present three graphs plotting performance bounds versus SNR. Each graph tracks three bounds: the performance of the optimal GAD with a side information channel optimized for all SNRs, Mazo's bound and the bound based on the Bhattacharyya coefficient given in proposition 4.4. For the bound based on the Bhattacharyya coefficient, the transmission probabilities has been chosen so that (4.13) is maximized. In addition we give the matched filter bound [19, p. 267], i.e., the performance of the optimal receiver operating in an environment without ISI, as a reference. We have used the expression in [7] to evaluate the \( Q(\cdot) \)-function, with the pertinent constants chosen so that the relative deviation was less than 1.2%.

The graphs 4-6 differ according to the assigned values of \( \delta \).

We make several observations. The bound based on the Bhattacharyya coefficient is not tight for high SNRs. For weak ISI (small \(|\delta|\)) Mazo's bound will be close to
the performance of the optimal GAD with a side information channel optimized for all SNRs, but less close for moderate ISI and low SNR.

In Figure 7 are given the transition probabilities for the side information channel of section 4.1, optimized for all SNRs to make the performance of the optimal GAD a tight bound. This side information channel will generally act in a random manner meaning that a transmitted sequence could render one out of several pairs $z$ as side information. However, for high and low SNRs it will act deterministically. Note that the transition probabilities developed in sections 4.1 and 4.3, respectively, as expected coincide for high SNRs.

7 Summary

We have studied the genie-aided detector (GAD) as an analytic device and introduced an explicit statistical description of the genie’s action as a probabilistic rule governing a side information channel. This description is given in terms of transition probabilities, the probabilities that for any given transmitted sequence there will be a certain side information supplied to the detector. The optimality of the processing of the augmented information is defined explicitly in terms of the probabilistic description of the side information channel, representing the genie.

With the complete statistical description of the transmission system, including a set of transition probabilities, the GAD with minimum bit error probability, see proposition 3.2, is derived in terms of a binary Bayesian hypothesis test. By evaluating the performance of this GAD, a lower bound on the performance of any detector, with or without access to the side information, is obtained.

The transition probabilities are free parameters which can be chosen to optimize the performance of the GAD. They might for example be chosen to make the corresponding bound tight, or to give the bound a simple structure. We have chosen several different sets of transition probabilities as examples in order to discuss the properties of their respective performance bounds. In this, we also discuss the relation of the attainable performance bounds to the works by Forney [9, 10] and Mazo [20].
Figure 4: The performance of the optimal GAD using a side information channel optimized for all SNRs, Mazo’s bound, the corresponding bound based on the Bhattacharyya coefficient and the matched filter bound are plotted versus SNR for $\delta = 0.3$.

Acknowledgements

We would like to express our gratitude to Prof. Tor Aulin, Chalmers University of Technology, Sweden, Amos Lapidoth, Stanford University, USA, and Prof. Shlomo Shamai (Shitz), Dept. of Electrical Engineering, Technion, Haifa Israel, for helpful discussions on the topic of genies. We would also like to acknowledge the improvements provided by Dr. Sarah Kate Wilson, Purdue University, and Paul Petersen, West Lafayette, USA.

This research has been supported in part by the Swedish National Board for Industrial and Technical Development (NUTEK), Telia Mobitel AB and the ISS’90 foundation.
Figure 5: The performance of the optimal GAD using a side information channel optimized for all SNRs, Mazo’s bound, the corresponding bound based on the Bhattacharyya coefficient and the matched filter bound are plotted versus SNR for $\delta = 0.6$.

Figure 6: The performance of the optimal GAD using a side information channel optimized for all SNRs, Mazo’s bound, the corresponding bound based on the Bhattacharyya coefficient and the matched filter bound are plotted versus SNR for $\delta = 0.9$. 
Figure 7: The transition probabilities of the globally optimized side information channel of corollary 4.1. Plotted versus SNR for three different values of $\delta$; upper: $\delta = 0.3$, middle: $\delta = 0.6$, lower: $\delta = 0.9$. 
APPENDICES

A Bounding the Error Probability Using Side Information

The significance of the side information can be explained without explicitly referring to the channel model in Figure 1. This can be done in terms of hypothesis testing between two hypotheses, $H_0$ and $H_1$, specified by the respective probability densities $f_{r|H_0}(\rho|H_0)$ and $f_{r|H_1}(\rho|H_1)$ on a product space, $\rho = (\eta, \zeta) \in \mathcal{Y} \times \Omega$. The variables in $r = (y, z)$ are grouped such that $y$ is assumed to have a continuous marginal distribution and $z$ is supposed to be finite, i.e., $z \in \Omega = \{\zeta_1, \ldots, \zeta_M\}$. Hence the two densities $f_{r|H_i}(\rho|H_i)$ are understood as densities with respect to a suitable product measure on $\mathcal{Y} \times \Omega$. Otherwise the distributions may, for the sake of the initial part of the argument in this appendix, be quite arbitrary.

Every binary decision rule (excluding rules with rejection options) partitions $\mathcal{Y} \times \Omega$ into two disjoint subsets, say $V_0$ and $V_1$, and favours the hypothesis $H_0$ if $\rho \in V_0$ and $H_1$ otherwise. For each $\zeta \in \Omega$ we introduce the sets (or $\zeta$-sections) $D_i(\zeta) = \{\eta \in \mathcal{Y} \mid (\eta, \zeta) \in V_i\}$ for $i = 0, 1$. Hence $V_i$ is a disjoint union, $V_i = \bigcup_{\zeta \in \Omega} D_i(\zeta, \zeta)$. The probability of deciding for the wrong hypothesis, denoted by $P_{\text{Error}}(V_0, V_1)$, is expressed as

$$P_{\text{Error}}(V_0, V_1) = \Pr\{H_0\} \int_{V_1} f_{r|H_0}(\rho|H_0) d\rho + \Pr\{H_1\} \int_{V_0} f_{r|H_1}(\rho|H_1) d\rho$$

$$= \Pr\{H_0\} \sum_{\zeta \in \Omega} \int_{D_1(\zeta)} f_{y,z|H_0}(\eta, \zeta|H_0) d\eta$$

$$+ \Pr\{H_1\} \sum_{\zeta \in \Omega} \int_{D_0(\zeta)} f_{y,z|H_1}(\eta, \zeta|H_1) d\eta,$$

where, e.g., $f_{V_i} f_{r|H_0}(\rho|H_0) d\rho$ designates the probability of $V_1$ with respect to the distribution determined by $f_{r|H_0}(\rho|H_0)$ (in mathematical terms, the Lebesgue-Stieltjes integral of the density $f_{r|H_0}(\rho|H_0)$ over $V_1$).

Any special decision rule determined by first partitioning $\mathcal{Y}$ into $D_0$ and $D_1$ independently of $\zeta$, thus disregarding the information in $z$, and then setting $V_i = \bigcup_{\zeta \in \Omega} (D_i, \zeta)$ has by an interchange of the order of integration and summation in (A.1) the probability of error

$$P_{\text{Error}}(\{D_0\}, \{D_1\}) = \Pr\{H_0\} \int_{D_1} f_y(\eta|H_0) d\eta + \Pr\{H_1\} \int_{D_0} f_y(\eta|H_1) d\eta,$$

where $f_y(\eta|H_i) = \sum_{\zeta \in \Omega} f_{y,z|H_i}(\eta, \zeta|H_i)$.

Letting $\{V_i^*\}_{i=0}^1$ designate the corresponding decision regions of the Bayesian decision rule using $f_{r|H_0}(\rho|H_0)$ and $f_{r|H_1}(\rho|H_1)$ to minimize the function $P_{\text{Error}}(\cdot)$ in equation A.1, it holds that $P_{\text{Error}}(V_0, V_1) \geq P_{\text{Error}}(V_0^*, V_1^*)$ for any other decision
B The GAD achieving minimal bit error probability

Let us derive the receiver that achieves the minimal probability of error when detecting bit number $k$ after observing $\rho = (\eta, \zeta_{i,j})$. Assume that $y$ and $z$ are conditionally independent given $\theta$ and that the transition kernel for the side information channel is of the form (3.8). The detection of bit number $k$ is equivalent to the following binary hypotheses testing problem (equation (3.1)):

$$H_0: \quad (b)_{bit,k} = -1$$
$$H_1: \quad (b)_{bit,k} = +1. \quad (B.1)$$

As is known [22], the Bayesian test

$$f_r|H_0(\rho|H_0) = f_y,z|H_0(\eta, \zeta_{i,j}|H_0)$$

It analogously follows that

$$f_r|H_1(\rho|H_1) = f_y|b(\eta|\beta^+_j) p(j|i) \Pr \{ b = \beta^+_j \}. \quad (B.4)$$

Using (B.3) and (B.4) in (B.2) the Bayesian test becomes

$$f_r|H_0(\rho|H_0) \geq \frac{f_y|b(\eta|\beta_j^+) p(j|i) \Pr \{ b = \beta_j^+ \}}{f_y|b(\eta|\beta_j^-)} \geq \Pr \{ H_0 \} \frac{H_0}{H_1}, \quad (B.5)$$

where $\Pr \{ H_1 \} = 1 - \Pr \{ H_0 \}$ is the a priori probability that $H_1$ is true, minimizes the probability of choosing the wrong hypothesis. Then, given the observation $\rho = (\eta, \zeta_{i,j})$ and using the stated assumptions

$$f_r|H_0(\rho|H_0) = f_y,z|H_0(\eta, \zeta_{i,j}|H_0)$$

$$= \frac{1}{\Pr \{ H_0 \}} \sum_{\beta \in C_k} f_{y,z|b}(\eta, \zeta_{i,j}|\beta) \Pr \{ b = \beta \}$$

$$= \frac{1}{\Pr \{ H_0 \}} \sum_{\beta \in C_k} f_{y|b}(\eta|\beta) \Pr \{ z = \zeta_{i,j}|b = \beta \} \Pr \{ b = \beta \} \quad (B.3)$$

$$= \frac{1}{\Pr \{ H_0 \}} f_{y|b}(\eta|\beta_j^-) \Pr \{ z = \zeta_{i,j}|b = \beta_j^- \} \Pr \{ b = \beta_j^- \}$$

$$= \frac{1}{\Pr \{ H_0 \}} f_{y|b}(\eta|\beta_j^-) q(i|j) \Pr \{ b = \beta_j^- \}.$$
where
\[ \gamma_{i,j} = \frac{q(i|j) \Pr \{ b = \beta_j^- \}}{p(j|i) \Pr \{ b = \beta_j^+ \}}. \]  

**C  The bit error probability of the optimal GAD**

Let \( P_F|z = \zeta_{i,j} \) denote the probability of a false alarm and \( P_M|z = \zeta_{i,j} \) denote the probability of a miss given \( z = \zeta_{i,j} \). The probability of error of the test in proposition 3.2 can now be expressed by
\[ P_{DE,k}|z = \zeta_{i,j} = P_F|z = \zeta_{i,j} P_F|z = \zeta_{i,j} + P_F|z = \zeta_{i,j} P_M|z = \zeta_{i,j}, \]
where
\[ \Pr \{ H_0|z = \zeta_{i,j} \} = \frac{1}{\Pr \{ z = \zeta_{i,j} \}} \sum_{\beta \in \mathcal{C}_k} \Pr \{ z = \zeta_{i,j}, \beta \} \]
\[ = \frac{1}{\Pr \{ z = \zeta_{i,j} \}} \sum_{\beta \in \mathcal{C}_k} \Pr \{ b = \beta \} \Pr \{ z = \zeta_{i,j} | b = \beta \} \]
\[ = \frac{1}{\Pr \{ z = \zeta_{i,j} \}} \Pr \{ b = \beta_j^- \} q(i|j). \]

Similarly,
\[ \Pr \{ H_1|z = \zeta_{i,j} \} = \frac{1}{\Pr \{ z = \zeta_{i,j} \}} \Pr \{ b = \beta_i^+ \} p(j|i). \]

Averaging \((C.1)\) over all possible \( \zeta_{i,j} \) we get the probability of bit error for the optimal GAD as
\[ P_{DE,k}^{GA} = \sum_{i,j} \Pr \{ z = \zeta_{i,j} \} P_{DE,k}|z = \zeta_{i,j}. \]

If the density function for \( y \) given \( b \) is multivariate Gaussian as \( N(H\beta, R_n) \), we get
\[ P_F|z = \zeta_{i,j} = Q \left( \frac{d_{i,j}}{2} + \frac{\ln \gamma_{i,j}}{d_{i,j}} \right) \]
and
\[ P_M|z = \zeta_{i,j} = Q \left( \frac{d_{i,j}}{2} - \frac{\ln \gamma_{i,j}}{d_{i,j}} \right), \]
cf. [22].

Using \((C.1),(C.2),(C.3), (C.5)\) and \((C.6)\) in \((C.4)\) we get the probability of bit error as
\[ P_{DE,k}^{GA} = \sum_{i,j} Q \left( \frac{d_{i,j}}{2} + \frac{\ln \gamma_{i,j}}{d_{i,j}} \right) q(i|j) \Pr \{ b = \beta_j^- \} \]
\[ + \sum_{i,j} Q \left( \frac{d_{i,j}}{2} - \frac{\ln \gamma_{i,j}}{d_{i,j}} \right) p(j|i) \Pr \{ b = \beta_i^+ \}. \]
A Bayesian Bound and the Bhattacharyya coefficient

We compute the Bhattacharyya coefficient

\[ \bar{h} = \int_{y \in \mathbb{R}} \sqrt{f_{\|H\|}(\rho|H_0) \Pr \{H_0 \}} \sqrt{f_{\|H\|}(\rho|H_1) \Pr \{H_1 \}} d\rho \]  

(D.1)

for the DTGC with a side information channel. Assuming that \( y \) and \( z \) are conditionally independent given \( b \) as well as that the transition kernel for the side information channel is of the form (3.8), we obtain from (B.3) and (B.4) that

\[ \bar{h} = \sum_{ij} \sqrt{p(j|i) \Pr \{b = \beta^+_i \} q(i|j) \Pr \{b = \beta^-_j \} \int_y f_{y|b}^{1/2}(\eta|\beta^-_j) f_{y|b}^{1/2}(\eta|\beta^+_i) d\eta}. \]  

(D.2)

The channel data \( y \) have, given the transmitted sequence \( b = \beta \), the multivariate Gaussian distribution \( N(H\beta, R_n) \). Hence a straightforward calculation gives that the integrals on the right hand side of (D.1) equal:

\[ \int_y f_{y|b}^{1/2}(\eta|\beta^-_j) f_{y|b}^{1/2}(\eta|\beta^+_i) d\eta = e^{-\frac{d^2_j}{2}}, \]  

(D.3)

as found in [14]. This proves (4.11) in proposition 4.4. Let us note that for \(-1 \leq u \leq 1\) we have \( \sqrt{1 - u} \leq 1 - \frac{1}{2} u \) and this yields that \( \frac{1}{2} \left[ 1 - \sqrt{1 - 4\bar{h}^2} \right] \geq \bar{h}^2 \). Recall that \( \bar{h} \) is a sum of non-negative terms. If we drop the terms corresponding to those pairs of sequences that are not at minimum distance from each other, we obtain (4.13) in corollary 4.5.

References


Part II.1

A Reduced Complexity Viterbi Equalizer Used in Conjunction with a Pulse Shaping Method
Part II.1:
A Reduced Complexity Viterbi Equalizer Used in Conjunction with a Pulse Shaping Method
Per Ödling¹, Timo Koski² and Per Ola Börjesson¹
This material has also been presented at the Nordic Radio Symposium, NRS'92, in Ålborg, Denmark.

¹Luleå University of Technology, Division of Signal Processing, S-971 87 Luleå, Sweden
²Luleå University of Technology, Department of Mathematics, S-971 87 Luleå, Sweden
A Reduced Complexity Viterbi Equalizer Used in Conjunction with a Pulse Shaping Method

P Ödling
Department of Signal Processing
T Koski
Department of Applied Mathematics
P O Börjesson
Department of Signal Processing
Luleå University of Technology, Luleå, SWEDEN

The complexity of the Viterbi algorithm can be reduced by linear pre-equalization. We introduce a linear preprocessor using the weighted least squares filtering (WLS) of [11]. The geometric technique of L. Barbosa [3] is evoked for analytical representation of the complexity reduction by pre-equalization, and for evaluation of error performance. We evaluate the bit error rate of the WLS preprocessor and of the linear preprocessor of Falconer and Magee [6].

1 INTRODUCTION

Consider a discrete time Gaussian channel (DTGC) with intersymbol interference (ISI) described by

\[ y_k = \sum_{l=0}^{L-1} h_l b_{k-l} + n_k, \quad -\infty < k < \infty, \]

where \( \{b_k\}, \ b_k \in \{-1, 1\} \), are i.i.d. channel inputs, \( \{y_k\} \) are channel output observables, \( \{h_k\} \) are real valued ISI coefficients and \( \{n_k\} \) are white Gaussian noise samples with variance \( N_0/2 \). The DTGC has been extensively used in information theory and communications engineering [15]. It models pulse amplitude modulated (PAM) signals transmitted through a time dispersive Gaussian channel, often applicable to e.g. telephone lines, magnetic recording media and radio channels.

The maximum likelihood sequence estimator (MLSE) is a procedure for estimating the input sequence of bits, \( \{b_k\} \), from the observations \( \{y_k\} \). In communications a recursive way of organizing the computation of the MLSE is desired. The Viterbi algorithm (VA) is a fast way of computing the MLSE by a search in a trellis of length \( N \) and width \( 2^L \) [7]. This still becomes impractical when \( L \) is large, since the computational complexity of the VA increases exponentially with \( L \). A number of researchers have addressed this problem suggesting complexity reducing solutions of two basic types, linear [4, 6, 8, 14] and non-linear [2, 9, 16]. Linear preprocessing, treated here, simplifies MLSE by giving the VA a channel model with a memory length shorter than \( L \), and by applying a linear preprocessor, with pulse response \( \{g_k\} \) of length \( J \), to the received data \( \{y_k\} \). The preprocessor shapes the overall system pulse response \( h * g \) to become similar to the channel model \( q \) using some criterion on \( e \), as shown in the figure below.

A convenient way of describing this model (discarding edge effects) is to set \( y = (y_0, \ldots, y_{N+L-2})^T \), \( b = (b_0, \ldots, b_{N-1})^T \) and \( n = (n_0, \ldots, n_{N+L-2})^T \) which gives the matrix formulation

\[ y := Hb + n \]
\[ z := GHb + Gn, \]

where \( H \) and \( G \) are Toeplitz matrices of dimensions \((N+L-1) \times N\) and \((N+L+J-2) \times (N+L-1)\) respectively, representing convolution, and \( \top \) transposes matrices. \( H \) and \( G \) have a band structure
with their corresponding pulse responses \({h_k}\) and \({g_k}\) in their columns, e.g.

\[
H = \begin{bmatrix}
  h_0 & 0 & \ldots & 0 \\
  h_1 & h_0 & \ldots & 0 \\
  \vdots & h_1 & \ldots & 0 \\
  h_{L-1} & \ldots & h_0 & 0 \\
  0 & h_{L-1} & \ldots & h_1 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \ldots & h_{L-1}
\end{bmatrix}
\]  

(1)

2 GEOMETRIC MODEL

In this matrix formalism the operations performed by the MLSE can be seen as operations in two Euclidean spaces, \(Y\), the space of the DTGC outputs \(y\) and the preprocessor outputs \(z\), and \(S\), a space containing the binary, centered hypercube \(C\) of all possible bit strings \(b\) of length \(N\) [3]. The matrix \(GH\) is a map from \(S\) to a subspace of \(Y\), see figure 2.

![Figure 2: A geometry of MLSE.](image)

We define the inner product in \(Y\) as

\[
\langle \alpha, \beta \rangle := \alpha^T R^{-1} \beta,
\]

with \(R := E\{G^n(G^n)^T\} = \frac{N^2}{2} GG^T\), and the norm as \(\|\alpha\|^2 := \langle \alpha, \alpha \rangle\). The conditional density of \(z\) given \(b\) is written as

\[
p(z|b) = Ke^{-\frac{1}{2}||z-GHb||^2}.
\]

The MLSE \(\hat{b}_c\) of \(b\), given \(z\), is thus the \(b\) that minimizes \(||z-GHb||\). We also define the adjoint operator of the linear map \(GH\) from \(S\) to \(Y\) by the equality

\[
\langle GHb, a \rangle = \langle b, H^*a \rangle.
\]

\(H^*\) is then a map from the space \(Y\) to \(S\). Consider \(S_u := (H^*GH)^{-1}H^*z\), the unconstrained estimate of \(b\). The vector \(z-GH\hat{b}_u\) is orthogonal to the range of \(GH\) in \(Y\), and hence

\[
||z-GH\hat{b}_u||^2 = ||z-GH\hat{b}_u||^2 + ||GH(b-b_u)||^2.
\]

Since the first term on the right hand side is independent of \(b\) given \(z\), \(\hat{b}_u\) is a sufficient statistic for \(b\). So, minimizing \(||z-GHb||^2\), a metric in \(Y\), is equivalent to minimizing \(||GH(b-b_u)||^2\), which is the proper metric in \(S\).

In a reduced complexity receiver the VA has a different, simplified metric \(\|\cdot\|_M\) in \(S\). Using \(\|\cdot\|_M\) the event of choosing a sequence \(b\) given \(z\), when \(b_0\) was transmitted, \(b_1 \neq b_0\), corresponds to

\[
||\hat{b}_u - b_0||_M^2 > ||\hat{b}_u - b_1||_M^2 \leftrightarrow (\hat{b}_u - b_0, Me) > \frac{1}{2}||b_1 - b_0||_M^2 = \frac{1}{2}(e, Me),
\]

where \(e := b_1 - b_0\), \(\hat{b}_u - b_0 = M_0^{-1}H^*n = M_0^{-1}H^TGT^{-1}n\) and \(M_0 := H^*GH = H^TGT^{-1}GH\). The scalar random variable \((\hat{b}_u - b_0, Me)\) has a normal distribution with variance \((Me, KMe)\), where \(K\) is the covariance matrix of \(b_u - b_0\) and \(K = E\{(M_0^{-1}H^*n)(M_0^{-1}H^*n)^T\} = M_0^{-1}\). Therefore the probability of the above error event denoted \(E\), i.e. \(||\hat{b}_u - b_0||_M^2 > ||\hat{b}_u - b_1||_M^2\), is

\[
p(E) = \Phi\left(\frac{1}{2}(e, Me)/\sqrt{(e, MMe^{-1}Me)}\right),
\]

where \(\Phi(\cdot)\) is the error function. The probability of a sequence error, i.e. \(\hat{b}_c \neq b_0, b_0\) transmitted,
can be expressed as
\[ p(\text{error}) \propto E(e) \left\{ Q \frac{1}{2} (e, Me) / \sqrt{(e, MM_0^{-1} Me)} \right\}. \]  
(2)

The number of deviation sequences \(e\) is large enough to make (2) computationally not feasible. A lower bound on (2), which corresponds to finding the so called minimum distance \(d_{\text{min}}\), see [10, 17], is [3]
\[ p(\text{error}) \geq \kappa \max_{\{e\}} \left\{ Q \frac{1}{2} (e, Me) / \sqrt{(e, MM_0^{-1} Me)} \right\}, \]
(3)
where \(\kappa\) is a constant with respect to \(e\). Closer lower bounds can be obtained by summing over a subset of \(\{e\}\), c.f. [10, 7, 15, 17]. Upper bounds on (2) can also be derived, see e.g. [10, 13, 7, 15, 17]. Here the bound of equation (3) is used to compare the complexity reducing methods to each other, not to bound their actual performance.

3 COMPLEXITY REDUCTION

We now analyze the Weighted Least Squares method (WLS) and for comparison the method presented in [6], both using linear preprocessing as described in section 1. In both cases the receiver, or equivalently the metric in \(S\), is simplified in two steps. First the noise covariance \(R\) due to the preprocessor is dropped, and secondly the VA is given an approximate channel model with a memory length shorter than \(L\), see figure 1. This yields the kernel of the simplified metric \(\| \cdot \|_{M}\) in \(S\) as
\[ M = QTQ, \]
where the matrix \(Q\) has \(q\) in the columns, the same structure as in (1) and dimension \((N + m - 1) \times N\), where \(m\) is the desired memory length, i.e. the length of \(q\).

An interesting issue is the relative computational complexity of various methods. However, this extra computational complexity involved is very low compared to the complexity reduction accomplished in the VA. See [17] for a more detailed study.

3.1 The WLS-filter

The WLS method was originally developed as a pulse shaping filter for ultrasonic imaging [11, 5, 12]. It can be described by three requirements:

1. The pulse response must have a non zero amplitude at least at one point, say for \(k = k_0\).

2. The energy of the pulse response should be small outside the desired interval, \(|k| > \frac{m}{2}\).

3. The noise amplification has to be low.

An idea related to 1 and 2 is presented in [1]. The solution for 1 and 3 is the matched filter, and the solution for 1 and 2 with \(m = 1\) is the inverse filter. The WLS filter combines 1, 2, and 3 with the desired memory length \(m\) and the expected noise power \(\sigma^2\) into a cost function [11]. Let \(\bar{H}\) be an \((J + L - 1) \times J\) matrix with the same elements as \(H\), \(g\) be a vector containing the pulse response \(\{g_k\}\), \(\delta_{k_0}\) be a vector of zeros except for a one \('1'\) at position \(k_0\), and finally let \(V\) be a diagonal matrix with \(m - 1\) zeros in the center of the diagonal and ones at position \(k_0\) and on the edges. Using \(\|\alpha\|_2^2 = \alpha^T \alpha\) the WLS cost function can be expressed as
\[ J(g) = \|V(\bar{H}g - \delta_{k_0})\|_2^2 + \alpha \sigma^2 \|g\|_2^2. \]

This can be interpreted in figure 1 as sending a unit pulse (deterministic input) through the model network, summing the energy of the response \(h \ast g\) outside the given interval of length \(m\), in [13] referred to as the residual ISI, and adding a term proportional to the noise amplification of the WLS preprocessor. The optimum preprocessing filter \(g\) is given by
\[ g_{\text{opt}} = [\bar{H}^T V^T V \bar{H} + \alpha \sigma^2 I]^{-1} \bar{H}^T \delta_{k_0}. \]

The desired pulse response \(q\) consists of the \(m\) elements of the overall system pulse response \(h \ast g\) corresponding to the \(m\) wide 'window' of \(V\), i.e. \(h \ast g\) is calculated and the tails are truncated.

The complexity reducing procedure is to filter the incoming data stream \(y\) with \(g_{\text{opt}}\), truncating the result to the proper length, \(N + m - 1\), and feeding this new data stream \(z\) and the desired pulse response \(q\) to the VA.
We are investigating the WLS-filter used as a preprocessor by varying the design parameters. In the sequel we have chosen $k_0 = 0$, all non-zero $w_k = 1$ and $a = 1$. These choices are ad hoc, although simulations indicate that the performance of the receiver is not very sensitive to the choice of $a$.

3.2 The Falconer and Magee filter

Figure 1 aids in describing the method of Falconer and Magee [6]. When feeding the network of figure 1 with a white binary sequence $b$ this preprocessor minimizes the expected energy of the output signal $e$, i.e.
\[
\min_{q,S} E\{e^2\} = E\{(q \ast b - g \ast h \ast b - g \ast n)^2\},
\]
with the constraint $q^T q = 1$. Solving this minimization problem in a matrix formulation renders
\[
q_{opt} = \text{the eigenvector of } [I - X^T(H^T H + \frac{2}{b} I)^{-1} X] \text{ corresponding to the minimum eigenvalue},
\]
\[
S_{opt} = (H^T H + \frac{2}{b} I)^{-1} X q_{opt},
\]
where $X$ is a $J \times (m+1)$ channel matrix related to $H$, c.f. [6]. The complexity reducing steps are now by analogy with the WLS method.

4 BIT ERROR RATE ANALYSES

We analyze the bit error performance of these receiver algorithms in two ways. From equation (3) a tractable numerical bound on the bit error probability is derived and plotted. Thereto we simulate the described models and algorithms and 'measure' the bit error rate (BER).

Denote an error sequence containing $k$ errors with $e_k$. The bit error probability bound corresponding to the sequence error probability bound of (3) is then numerically treated by the assessment
\[
\frac{1}{N} \max_{A} \frac{k}{2^k} Q\left\{\frac{1}{2} (e_k, M e_k) / \sqrt{(e_k, MM_0^{-1} M e_k)}\right\},
\]
where $A$ is a subset of $\{e\}$. McLane asserts that the dominating values of the $Q$-function occur for error sequences with only a few errors, i.e. for $k = 1$ or $k = 2$, see [13] and the references therein. In our simulations we have found this to be valid for $p(error)$ analysis independently of the SNR, and valid for high SNR scenarios for BER analysis. For low SNR longer error sequences increase in importance since they contribute more to the bit error probability, although the probabilities of the sequences themselves still are below the probabilities of shorter error sequences. A more complete study of the BER in terms of a Forney type bound is found in [10, 17].

When numerically evaluating (4) we let $A$ contain, i.e. we search through, all error sequences with 1 and 2 errors, a portion of the relevant sequences with 3 and 4 errors and some longer sequences. Following McLane in [13] we set $N = 25$, which is claimed to be sufficient to represent the ISI.

In the first example we use the magnetic recording (Lorentzian) channel model from [16], but with a truncated length of $L = 7$. This truncation is performed to enable us to use an optimal, full complexity Viterbi decoder as a benchmark. The preprocessor length is as above, $J = 31$, and the length of the desired pulse response is $m = 3$. Here the WLS method is compared to the optimal Viterbi decoder and to the lower bound estimate of equation (4).
For the second example we have used the difficult channel 'B' of length $L = 15$ from [6], also used in [13]. As in [6] we choose the filter length as $J = 31$ and the length of the desired pulse response as $m = 5$. We compare the WLS method to the Falconer and Magee method, and to their respective bound estimates according to equation (4).

The BER measurements indicate that the performance of the preprocessors depends heavily on the channel characteristics and that none of the methods is superior for all cases. Work on optimization of the WLS-preprocessor improving its error performance is in progress [17].

References

Part II.2

Combined Linear-Viterbi Equalizers – A Comparative Study and A Minimax Design
Part II.2:
Combined Linear-Viterbi Equalizers – A Comparative Study and A Minimax Design
Nils Sundström, Ove Edfors, Per Ödling, Håkan Eriksson, Timo Koski and Per Ola Börjesson

1Luleå University of Technology, Division of Signal Processing, S-971 87 Luleå, Sweden
2Luleå University of Technology, Department of Mathematics, S-971 87 Luleå, Sweden
Abstract: Combined linear-Viterbi equalizer (CLVE) is a term often used for a class of digital receivers reducing the complexity of the Viterbi detector by assuming an approximate channel model together with linear pre-equalization of the received data.

We reconsider a weighted least squares design technique for CLVEs by introducing a minimax criterion for suppressing the strongest component of the residual intersymbol interference. Previously, in [2], we have studied the performance of some proposed CLVE design methods and evaluated them by simulated bit error rates. Here we investigate the performance of the minimax design and of the CLVE designs found in literature [3, 4, 5, 6] for two GSM test channels.

We also present a comparison of the CLVE designs based on a common quadratic optimization criterion for the selection of the channel prefilter and the desired impulse response.

1 INTRODUCTION

The Maximum Likelihood Sequence Detector (MLSD) is a procedure for estimating a sequence of bits from a sequence of channel output observables, given a model of the communication system. In the presence of intersymbol interference (ISI) the Viterbi algorithm (VA) provides an efficient way of computing the MLSD [7, 8, 9]. However, the VA still becomes impractical when the time spread of the ISI is large because of the exponential relation between ISI time spread and VA complexity. The complexity of a Viterbi detector can be reduced by giving the VA an approximate channel model with a shorter time spread than that of the original channel. The class of receivers employing this technique together with linear pre-equalization of the received data are often referred to as combined linear-Viterbi equalizers (CLVEs), see [1, 3, 4, 5, 6, 2]. Other classes of receivers addressing the same complexity problem can be found in e.g. [10, 11, 12, 13]. In this paper we focus on CLVEs.

When designing CLVEs it is often desirable to minimize the bit error rate of the receiver. The bit error probability depends on the design parameters, such as the channel model and the impulse response of the prefilter, in a complicated and non-linear way. The pre-filtering of the received data perturbs the signal space and colours the channel noise. Ignoring this colouring or giving the VA an approximate channel model result in a displacement of the decision regions from their optimal locations, cf. the residual ISI in [14]. Instead of using the bit error rate as a design criterion, other more feasible criteria are used in CLVE design methods [1, 3, 4, 5, 6, 2].

In this paper we investigate the performance of a minimax CLVE design. Three other principal techniques for designing CLVEs [3, 4, 6] are also overviewed and compared.

2 THE TRANSMISSION SYSTEM MODEL

A continuous time model of a transmission system as described in [7] may, without information loss, be represented by a discrete time model as in Fig. 1. In block transmission systems blocks of information are transmitted so that the decoding can be performed indepen-
dently on each block, see e.g. [15]. For such systems it is convenient to use the matrix formalism [16]
\[ y = Hb + n, \tag{2.1} \]
where the transmitted message is coded antipodally in \( b \in \{-1,+1\}^N \) and the channel observables are \( y \in \mathbb{R}^{(N+L-1)} \). The time invariant channel impulse response, i.e. the ISI coefficients, is represented by a Toeplitz band matrix \( H \). It is a known \((N+L-1) \times N\) matrix with the impulse response \( h \), of length \( L \), in its columns, arranged so that matrix multiplication corresponds to convolution. The noise vector \( n \in \mathbb{R}^{(N+L-1)} \) is a jointly Gaussian, zero mean random vector with a \( N(0, R_n) \) distribution, where \( R_n \) denotes the noise covariance matrix.

### 3 CLVE DESIGN

The performance evaluation of the CLVEs in this paper is performed on block transmission systems, but historically the design methods have focused on continuous transmission systems and VAs with infinite horizon, i.e. \( \{b_n\}_{n=-\infty}^{\infty} \). To relate to existing methods, we have chosen to confine the development of the minimax design to a criterion based on sequences of infinite length. Some remarks on CLVE design for block transmission systems can be found in section 5.

A design model is presented in Fig. 2, cf. [6] where \( q \) is the desired impulse response (DIR), i.e. the channel model given to the VA. If the time delay, modelled by the filter \( d \), is zero, the system is equivalent to the one found in [3]. Since this paper only considers causal channel impulse responses, the introduction of the time delay offers the possibility of placing the energy of the DIR in an arbitrary position. This possibility is accounted for in [3, 4, 5] by allowing the channel impulse response anti-causal components.

![Figure 2: Discrete design model for infinite sequences.](image)

The error \( \varepsilon_i \), in Fig. 2, can be expressed as
\[ \varepsilon_i = [b \circ (h \ast p - d \ast q) + n \ast p]_i, \tag{3.2} \]
where \( \circ \) denotes convolution. By using \( \|x\|_M^2 := x^T M x, \) the variance of \( \varepsilon_i \) is given by
\[ E\{\varepsilon_i^2\} = \|Hp - Dq\|_{R_q}^2 + \|p\|_{R_p}^2, \tag{3.3} \]
where \( p \) and \( q \) are vectors containing the impulse response of the prefilter \( p \) and the DIR \( q \) respectively, and \( R_q \) and \( R_p \) are covariance matrices for the transmitted sequence \( b \) and the noise \( n \). \( H \) and \( D \) are Toeplitz band matrices such that the multiplications \( Hp \) and \( Dq \) describe convolution. Since the length of \( h \ast p \) is greater than the length of \( d \ast q \), the subtraction \( Hp - Dq \) in (3.3) necessitates that the size of \( D \) is chosen such that the dimensions agree. The residual ISI is defined by this difference, as
\[ \text{residual ISI} := Hp - Dq, \tag{3.4} \]
i.e. as the ISI that is not accounted for in the VA, cf. [14].

### 3.1 A SELECTION OF PREFILTER DESIGNS

In this section we compare design methods for \( p \) and \( q \). The methods are due to Falconer and Magee [3], Fredricsson [4] and Odling et al. [6], where the last one is a weighted least squares method inspired by [17].

By introducing a weighting matrix \( W \) in (3.3) selective weighting of the residual ISI becomes possible. This is used to formulate the design criterion
\[ J(p, q) := \|W(Hp - Dq)\|_{R_q}^2 + \|p\|_{R_p}^2. \tag{3.5} \]
In the sequel, both the channel input and the channel noise are assumed white and stationary, hence \( R_q = I \) and \( R_p = \sigma_n^2 I \). Completing the square in (3.5) gives
\[ J(p, q) = \|(p - p_o)\|_A^2 + \|q\|_B^2, \tag{3.6} \]
where
\[ A = H^TW^TWH + \sigma_n^2 I, \]
\[ B = D^TW^T \left( I - WHA^{-1}H^TW^T \right) WD, \]
\[ p_o = -A^{-1}H^TW^TWDq. \]
In the presence of noise, the matrices \( A \) and \( B \) are positive definite, so the minimum of (3.6), with respect to \( p \), is obtained if
\[ p = p_0 = -A^{-1}H^TW^TWDq, \tag{3.7} \]
with a residual error of \( \|q\|_B^2 \).

The most straightforward design approach is to assume \( W = I \), i.e. a uniform weighting of the residual ISI, and to find the global minimum of (3.6) by using
(3.7) in combination with finding a \( q \) that minimizes \( \| q \|_B^2 \) under some constraint e.g. \( \| q \| = 1 \). This is done by Falconer and Magee in [3], where the prefilter obtained is
\[
p_{FM} = A^{-1}H_Dq_{FM},
\]
and the DIR
\[
q_{FM} = \text{the normalized eigenvector corresponding to the smallest eigenvalue of } B.
\]

There are other methods that do not minimize (3.6), but still often render a lower probability of bit error than the method above. One such method was presented by Fredricsson in [4]. For the purpose of preserving the similarity to the expressions in the original reference, we present his result in the Fourier domain:
\[
\text{and } P_{WLS} = \begin{bmatrix} C \end{bmatrix} V_{WLS},
\]
The expression \( D^T H P_{WLS} \) is chosen as in the WLS design method.

A weighted least squares (WLS) design approach for \( p \) and \( q \) was presented by Ödling et al. in [6]. The DIR in this design method is assigned an exact copy of the corresponding positions of \( H_p \), thus giving the VA a correct channel model for those positions, i.e.
\[
q_{WLS} = D^T H P_{WLS}.
\]
The prefilter \( p_{WLS} \) is chosen by modifying (3.7) to
\[
p_{WLS} = A^{-1}H_DW^TW^TD_{q^*},
\]
where the position \( l \) is chosen as a suitable index in \( q_{WLS} \). By choosing the weighting matrix in (3.13) as
\[
W = \text{diag}(1 \cdots 10 \cdots 0.1 \cdots 01 \cdots 1),
\]
where the zeroes coincide with the DIR position and \( k \) is chosen so that it coincides with the \( l \)th position of \( q \). This ensures that the trivial, all zero, solution for \( p_{WLS} \) is avoided and that the energy in \( H_p W_{WLS} \) is concentrated to the DIR interval. Equation (3.13) can now be simplified to the form in which it was presented by Ödling et al.
\[
p_{WLS} = A^{-1}H_D\delta_k
\]
where \( \delta_k \) is a vector with 1 in position \( k \) and zeroes elsewhere.

### 3.2 A MINIMAX DESIGN

The criterion of (3.5) has the total energy of the weighted residual ISI as one part. It can be discussed if also the distribution of the residual ISI is of any importance for the probability of bit error of the resulting receiver. Here we investigate an approach that takes into account the maximum absolute value of the residual ISI, thereby shaping the distribution of the same. Let us use this minimax approach and solve
\[
p_{MM} = \arg \min_p \left\{ \max_i J_i(p) \right\},
\]
where \( J_i(p) \) is the \( i \)th element in the vector
\[
J_i(p) = W(H_p - \delta_k) + \sigma_n \| p \|. \tag{3.17}
\]
The expression \( W(H_p - \delta_k) \) denotes the vector of the absolute values of each element in \( W(H_p - \delta_k) \). The DIR \( q_{MM} = D^T H p_{MM} \), the weighting matrix \( W \) and \( \delta_k \) are chosen as in the WLS design method.

### 4 SIMULATIONS

To evaluate the performances of the CLVEs for the different prefilter and desired impulse response design methods described in section 3, we have simulated the block transmission system of section 2. We have used two GSM test channels [18], the Typical Urban channel (TU) and the Rural Area channel (RA).

![Impulse and frequency responses of the GSM Typical Urban channel (left) and Rural Area channel (right).](image)

When implementing the minimax design, we have used a sequential quadratic programming method provided by the "minimax" routine of the MATLAB™ Optimization Toolbox™ [19] to solve (3.16), starting with an initial prefilter \( p \) given by (3.13).

To determine the prefilter \( p \) and the DIR \( q \) of Fredricsson, we have solved equation (3.11) by means
CONCLUSIONS AND FUTURE WORK

The basic idea in most design methods for CLVEs is reflected by the criterion $J(p, q)$. The main difference between the methods presented in the literature [3, 4, 5, 6] lies in the way the DIR $q$ is chosen. Falconer and Magee find the global minimum of the criterion under the constraint $\|q\| = 1$. The WLS method and the method proposed by Fredricsson has slightly different approaches where the subsequent processing of data by the VA is taken into consideration when applying the criterion. Fredricsson uses a projection on the minimum distance error sequence giving an effective signal to noise ratio [7] tailored for the VA. The WLS method ensures that the DIR given to the VA is a true replica of the total system impulse response in the corresponding time interval.

In our simulations the WLS receiver shows a superior performance and is indeed close to the full complexity Viterbi decoder. The receiver proposed by Fredricsson performs well on the TU channel, which is slightly surprising considering the spectral shape of this channel. The results in the presented simulations agree with applicable observations in [4, 5] and in our earlier investigations [6, 2].

The new method based on the minimax design was introduced as an attempt to shape the distribution of the residual ISI in a fashion favourable to the VA. Being similar to the WLS method it gives a truncated version of the total system impulse response as a DIR to the VA, but instead of minimizing the criterion $J(p, q)$ it suppresses the largest residual ISI coefficient. The performance is almost up to par with the WLS receiver, which indicates a potential for the concept of shaping the residual ISI.

The hitherto discussed methods for CLVE design are derived for continuous transmission systems. However, many contemporary and future communication systems are of block transmission type, e.g. the cellular
telephone systems of Europe (GSM), Japan (JDC) and the USA (ADC). To our knowledge, there are today no CLVE design methods that take advantage of the structure of such systems. An increased understanding of the properties of block transmission systems could result in improved receivers with respect to bit error probability as well as reduced implementation complexity and cost. In CLVEs developed for block transmission systems this could be reflected by the time-invariant linear prefiltering being replaced by, e.g., a general matrix multiplication, in order to utilize the "edge" effects at the block boundaries for performance improvement. The noise correlation due to the prefilter \( p \) is another important issue in connection with CLVEs, cf. [16]. This is recognized by Fredricsson and Beare, but not considered in the other described methods.

We regard the above issues as key components in the development of new CLVEs.

References


Part II.3

Multiple-Antenna Reception and a Reduced-State Viterbi Detector for Block Transmission Systems
Part II.3:
Multiple-Antenna Reception and a Reduced-State Viterbi Detector for Block Transmission Systems
Håkan Eriksson¹ and Per Ödling¹
Submitted to the 4th International Conference on Universal Personal Communications (ICUPC'95), Tokyo, Japan, November 1995.

¹Luleå University of Technology, Division of Signal Processing, S-971 87 Luleå, Sweden
Multiple-Antenna Reception and A Reduced-State Viterbi Detector for Block Transmission Systems

Håkan B. Eriksson  Per Ödling

Luleå University of Technology
Division of Signal Processing
S-971 87 Luleå, Sweden

Abstract—This paper presents a receiver structure employing multiple receiving antennas and reduced-state Viterbi detection. The presentation is couched in terms of matrices representing a discrete-time, symbol-sampled system with frequency-selective, slow fading and additive Gaussian noise. The receiver design involves determining two matrices: one matrix defining the metric of the reduced-state Viterbi detector and one representing a pre-processor operating on the output from a maximal-ratio combiner. The presented receiver is a generalization of those used earlier in single-antenna, combined linear-Viterbi equalizers [3], [5], [7], [11], [19], [16], [17] for continuous transmission systems. Performance comparisons are made with the minimum mean-square error (MMSE) linear equalizer and the MMSE decision-feedback equalizer for block transmission systems [10], indicating that the presented class of receivers offers superior performance.

I. INTRODUCTION

The use of multiple receiving antennas in wireless communication to take advantage of spatial diversity has been widely studied in the literature, see e.g., [1], [4], [8], [9] and the references therein. For flat fading channels it is shown in [8] that the use of multiple receiving antennas can improve the average signal-to-noise ratio (SNR) and lower the probability of a deep fade. In a frequency-selective fading environment the channel will be time-dispersive, which may in turn cause intersymbol interference. Several methods have been put forward to combat this. One such method is the orthogonal frequency division multiplexing (OFDM) modulation technique. There have also been several methods proposed that employ some form of multi-channel equalization. These are based on, for example, the minimum mean-square error criterion [1], decision-feedback techniques [1], or maximum likelihood sequence detection [6].

In this paper we examine one other method of performing multi-channel equalization; we propose a receiver structure that uses multiple receiving antennas and a reduced-state Viterbi Algorithm (VA). A system with a maximal-ratio combiner (MRC) [8], [9] is modelled as a single-antenna, white-noise system by using a data transformation matrix preserving the sufficient statistics. Linear prefiltering and a reduced, or fixed, state VA are then applied, producing a system with an MRC, a linear operation and a fixed complexity VA.

II. MODEL

Consider a system using M receiving antennas and assume that there is one linear, frequency-selective, slow fading channel for each antenna. Each channel is modelled as a discrete-time, symbol-sampled, additive Gaussian channel (DTGC) with intersymbol interference (ISI). Let b be an N × 1 random vector containing the N independent symbols to be transmitted. For each block the received signal at each antenna can then be described in terms of the matrix representation

\[ x_i = H_i b + n_i \]  

(1)

where \( H_i \) is an \((N + L_i - 1) \times N\) stochastic, complex valued matrix representing the ISI, \( n_i \) is a complex, jointly Gaussian, zero mean random vector with a \( N(0, R_i) \) distribution and where \( x_i \) is a vector of observable channel outputs.

In a mobile situation, where a moving transmitter or receiver creates a non-stationary environment, the channels will vary in between the time instants two blocks are transmitted. Throughout this paper we assume that, for each block to be transmitted, the outcomes of all the channel matrices \( H_i \) are known to the receiver. This could be a reasonable assumption under some conditions, such as when each channel impulse response varies slowly.

III. MAXIMAL-RATIO COMBINER

The Maximal-Ratio Combiner (MRC) [8], [9] is an optimal method to combine the received signals \( x_i \) in the sense that it gives a minimal sufficient statistic for the detection of b. In this section we re-derive the MRC by generalizing a result for single-antenna block transmission systems by Barbosa [2] to the multi-antenna case.

With the following block matrices:

\[
\begin{bmatrix}
X \\
N \\
X_M
\end{bmatrix} \triangleq
\begin{bmatrix}
x_1 \\
\vdots \\
x_M
\end{bmatrix},
\begin{bmatrix}
n_1 \\
\vdots \\
n_M
\end{bmatrix} \quad \text{and} \quad
H \triangleq
\begin{bmatrix}
H_1 \\
\vdots \\
H_M
\end{bmatrix}
\]  

(2)

a multi-antenna model can be expressed as

\[ x = Hb + n. \]  

(3)

Define the \( N \times N \) matrix \( M = H^H R^{-1} H \), where \( R = E\{nn^H\} \) and is assumed to be invertible and where
the superscript $H$ denotes Hermitian transpose. Assume that $H$ has a column space of rank $N$, meaning there is information in $x$ about every symbol in $b$. Then $M$ is invertible and

$$\hat{b}_u = M^{-1}H^TR^{-1}x = b + M^{-1}H^TR^{-1}n$$

(4)

is the output of the simultaneous zero-forcing equalizer, cf. the result by Barbosa in [2] for the single-antenna case. From [2] we know that $\hat{b}_u$ is a minimal sufficient statistic [14] for $b$ given $x$. Because $M$ is an invertible $N \times N$ matrix, it follows from (4) that

$$y \triangleq H^TR^{-1}x$$

(5)
is also a minimal sufficient statistic. Hence the vectors $x$ and $y$ both hold the same information for the purpose of detecting $b$, in spite of $y$ being only of dimension $N \times 1$. Equation (5) is the basic appearance of the MRC. In the case that the noise at each particular antenna is uncorrelated with the noise at all other antennas, the outputs of the whitened matched filters are summed to form the output from the MRC, see figure 2 and, e.g., [9].

Before beginning with designing receivers, we would like to transform the model (3) into an "equivalent" single-antenna, white noise model. Given $H$, the filtered additive noise in (4), $M^{-1}H^TR^{-1}n$, is a Gaussian zero mean noise process with a correlation matrix equal to $M^{-1}$. To whiten this noise we use the Cholesky factorization of $M$, producing an upper triangular matrix $V$ such that $V M^{-1}V^H = I$, where $I$ is the identity matrix. Let

$$r \triangleq V \hat{b}_u = V M^{-1}y = Vb + n_v.$$  

(6)

Then $n_v = VM^{-1}H^TR^{-1}n \in \mathcal{N}(0, I)$ given $H$. Equation (6) can be seen as a single-antenna model that is equivalent to (3) in the sense that an optimal receiver will perform equally well operating on either $x$ or $r$, since both are sufficient statistics for $b$, cf. the derivations by Kaleh in [10] for the single-antenna case. The main difference between (6) and (3) lies in the structures of the model matrices $H$ and $V$. The multi-antenna model matrix $H$ is an $(M(N-1) + L_1 + \ldots + L_M) \times N$ matrix while $V$ is an $N \times N$ matrix that can be seen as an anti-causal, time-variant filter.

Note that variations in the signal-to-noise ratio takes the form of varying signal energy in the matrix $V$ while the noise $n_v$ always has unit variance. We define the block signal-to-noise ratio (SNR) as

$$\gamma \triangleq \frac{\text{tr} \{V^H V\}}{N}.$$  

(7)

Note that the block SNR is a stochastic variable (with one outcome for each transmitted block) since the channel matrix $H$, and thereby $V$, is stochastic.

A. The Effects of Using Multiple Receiving Antennas and Maximal-Ratio Combining

Let us look at how the number of antennas and the lengths of the impulse responses at each antenna affect the characteristics of the system model (6). A simple measure of the amount of ISI is the normalized energy in the elements outside the main diagonal of $V$, i.e., with $B = V - \text{diag}\{v_{1,1}, v_{2,2}, \ldots, v_{N,N}\}$,

$$\rho \triangleq \frac{\text{tr} \{B^HB\}}{\text{tr} \{V^HV\}}.$$  

(8)

Consider an example where each channel is frequency-selective and Rayleigh-faded. For each block to be transmitted let the impulse response of each channel be the outcome of a stochastic vector of length $L$ with independent, identically distributed (i.i.d.), complex, zero mean Gaussian random variables. Furthermore let the block length $N = 20$ and the noise correlation $R = \sigma^2 I$. In figure 1 we have plotted the estimated mean value of $\rho$ versus the number of antennas, $M$, and the length of the impulse response of each channel, $L$. (Note that origo is in the lower, right part of the figure.) The mean value of $\rho$ was estimated by averaging (8) over 3000 independent outcomes of $H$.

![Fig. 1. The amount of ISI in the matrix $V$ versus the number of antennas and the length of the impulse responses of each channel, $L$, using $\rho$ defined in (8). (Origo is on the right edge.)](image)

From figure 1 we see that if we add more independent diversity paths by means of more antennas, the ISI in the equivalent model (6) tends to decrease in the sense that $\rho$ of definition (8) tends to decrease. Thus, the ISI increases with the length of the channel responses but decreases with the number of antennas. Alternative characterizations of the ISI in block transmission systems, are the colourization of the noise in equation (4) [2] and the eigenvalue spread of the matrix $M$ [11].

In the example the block SNR $\gamma$ in (7) is the sum of the squared absolute value of $2LM$ i.i.d. real valued Gaussian variables giving the block SNR a chi-squared distribution with $2LM$ degrees of freedom. Hence, in this example the number of antennas and the length of the impulse responses have an identical effect on the probability distribution of the block SNR of definition (7): increasing either factor will decrease the probability of a deep fade and increase the mean of the block SNR $\gamma$.

Although we assume the diversity paths to be independent in this example, these characteristics may well apply to a real situation if the antennas are enough separated.
IV. CLVEs for Block Transmission Systems

Consider a receiver where the output of the MRC, $y$, is filtered by a linear system represented by a matrix $G$ and an estimate of the transmitted sequence $b$ is obtained as

$$ b = \arg \min_b ||Gy - Qb||_2^2, \tag{9} $$

where $||x||_2^2 \triangleq x^H A x$. A receiver structure implementing this, consisting of a linear prefilter and a VA, is depicted in figure 2. We refer to this class of receivers as “combined linear-Viterbi equalizers for block transmission systems” (B-CLVE), cf. [5], [7], [13], [16], [17].

For the VA to be a meaningful tool for solving (9), the matrix $Q$ should be a band matrix [10], i.e.,

$$ Q = \begin{bmatrix} 0 & \cdots & 0 \\ 0 & \ddots & \cdots \\ \vdots & \ddots & 0 \\ 0 & \cdots & 0 \end{bmatrix}, \tag{10} $$

where $l$ is the width of the band of non-zero elements. Because wider bands of non-zero elements in $Q$ require more states in the VA trellis, it is the width of these bands that determines the complexity of the VA. By choosing a given value of $l$, the computational complexity of B-CLVEs may be controlled; conversely, for a given VA complexity (determined, perhaps, by hardware), a prescribed value for $l$ may be indicated.

If we apply the model (6) to the structure displayed in figure 2, we get a model suitable for the design of B-CLVEs, see the system described in figure 3. The linear distortion of the received signal is described by the matrix $V$. The matrix $P$ is the prefilter to be designed together with the matrix $Q$, which is given to the VA as a channel model. The function of the prefilter $P$ is to concentrate the energy of the system $PV$ to a band of width $l$, and $Q$ is a description of the maximum like­lihood sequence decoder [6]. The complexity of the VA is then determined by the structure of $V$, or rather, by the width of the band of non-zero elements in $V$. This, in turn, is equal to the length of the largest ISI time spread of the channels, cf. the Cholesky factorization in [15, p. 55].

V. DESIGN METHODS FOR BLOCK-CLVE’S

In this section we will discuss the design of the matrices $P$ and $Q$ with respect to what has been done earlier for continuous transmission systems. We will also, as an example, generalize one of these earlier methods to our block transmission system environment.

Hereinafter we discuss only the case when both $P$ and $Q$ are of dimension $N \times N$. In order to restrict $Q$ to a band-matrix, let $q_i = \hat{q}_i D_i$, for all $i$, where $\hat{q}_i$ is a $1 \times l$ vector with the model of the system’s impulse response given to the VA. The matrix $D_i$ positions the vector $\hat{q}_i$ in the $i$th row of the matrix $Q$, and is given in Appendix A.

In [16] it is recognized by Sundström et al. that in block transmission systems "edge effects" will appear that could be incorporated in B-CLVE design. It is also noted that the CLVEs for continuous transmission systems that are presented in [5], [7], [16], [17] all are related to the same criterion: the variance of a signal formed as the difference between the actual signal given to the VA and the model of the same signal.

Let us here consider the error vector

$$ e = Pr - Qb = PVb + Pn_e - Qb, \tag{12} $$

i.e., the difference between the filtered received signal and the model signal, as displayed in figure 4.

The expected energy in the error vector $e$ is given by

$$ J(P, Q) = E(e^H e) = \sum_{i=1}^{N} j(p_i, q_i), \tag{13} $$

where each term

$$ j(p_i, q_i) = \|p_i V - q_i\|^2_{R_b} + \|p_i^H\|^2_{I} \tag{14} $$

and where $p_i$ and $q_i$ are the $i$th rows of $P$ and $Q$, respectively.

![Fig. 2. The receiver structure for the case when the noise at each particular antenna is uncorrelated with the noise at all other antennas.](image)

![Fig. 3. The structure of CLVEs for block transmission systems.](image)
4. seems, however, to be more difficult

Later investigated by Beare [3], seems, however, to be more difficult to apply directly to B-CLVE.

The following section describes one way of applying the weighted least squares (WLS) method of

where $Wi$ is a weighting matrix and $Pi$ is chosen as the vector $pi$ that minimizes (15), that is

Each row in $Q_{WLS}$, $qi_{WLS}$, is then assigned an exact copy of the positions of the vector $pi_{WLS}$ corresponding to every term in $2$. The absolute value of the entries in the matrices displayed is equal to 10dB. The absolute value of the entries in the matrices displayed is equal to 10dB.

Using simulation, we compare the performance of the B-CLVE presented in section V-A with the performance of the minimum mean-square error linear equalizer (MMSE-LE) and the minimum mean-square error decision-feedback equalizer (MMSE-DFE) for block transmission systems, both described by Kaleh in [10]. The MMSE-DFE and the MMSE-DFE are applied directly to $y$, the output of the MRC. Also using simulation, we illustrate the necessity of prefiltering when the actual impulse response is longer than $l$, the permitted length of the model. We plot the performance of a receiver that assigns a "truncated" version of the channel $V$ as the channel model $Q$ by setting $P = I$ and the rows of $Q$ as in (18). We refer to this receiver as the receiver without pre-filtering (WPF).

In the simulations each block consisted of $N=20$ antipodally modulated, i.i.d. bits, thus $b$ is an $N$-dimensional random vector containing ones, and $D$ is chosen as the vector $pi$ that minimizes (15), that is

Each row in $Q_{WLS}$, $qi_{WLS}$, is then assigned an exact copy of the positions of the vector $pi_{WLS}$ corresponding to every term in $2$. The absolute value of the entries in the matrices displayed is equal to 10dB. The absolute value of the entries in the matrices displayed is equal to 10dB.

Using simulation, we compare the performance of the B-CLVE presented in section V-A with the performance of the minimum mean-square error linear equalizer (MMSE-LE) and the minimum mean-square error decision-feedback equalizer (MMSE-DFE) for block transmission systems, both described by Kaleh in [10]. The MMSE-DFE and the MMSE-DFE are applied directly to $y$, the output of the MRC. Also using simulation, we illustrate the necessity of prefiltering when the actual impulse response is longer than $l$, the permitted length of the model. We plot the performance of a receiver that assigns a "truncated" version of the channel $V$ as the channel model $Q$ by setting $P = I$ and the rows of $Q$ as in (18). We refer to this receiver as the receiver without pre-filtering (WPF).

In the simulations each block consisted of $N=20$ antipodally modulated, i.i.d. bits, thus $b$ is an $N$-dimensional random vector containing ones, and $D$ is chosen as the vector $pi$ that minimizes (15), that is

Each row in $Q_{WLS}$, $qi_{WLS}$, is then assigned an exact copy of the positions of the vector $pi_{WLS}$ corresponding to every term in $2$. The absolute value of the entries in the matrices displayed is equal to 10dB. The absolute value of the entries in the matrices displayed is equal to 10dB.

Using simulation, we compare the performance of the B-CLVE presented in section V-A with the performance of the minimum mean-square error linear equalizer (MMSE-LE) and the minimum mean-square error decision-feedback equalizer (MMSE-DFE) for block transmission systems, both described by Kaleh in [10]. The MMSE-DFE and the MMSE-DFE are applied directly to $y$, the output of the MRC. Also using simulation, we illustrate the necessity of prefiltering when the actual impulse response is longer than $l$, the permitted length of the model. We plot the performance of a receiver that assigns a "truncated" version of the channel $V$ as the channel model $Q$ by setting $P = I$ and the rows of $Q$ as in (18). We refer to this receiver as the receiver without pre-filtering (WPF).

In the simulations each block consisted of $N=20$ antipodally modulated, i.i.d. bits, thus $b$ is an $N$-dimensional random vector containing ones, and $D$ is chosen as the vector $pi$ that minimizes (15), that is

Each row in $Q_{WLS}$, $qi_{WLS}$, is then assigned an exact copy of the positions of the vector $pi_{WLS}$ corresponding to every term in $2$. The absolute value of the entries in the matrices displayed is equal to 10dB. The absolute value of the entries in the matrices displayed is equal to 10dB.
the channel's 1024. Figure 6 shows the simulated Bit Error Rate (BER) for the four receivers versus average SNR.

![Fig. 6. The estimated BER for the B-CLVE, MMSE-DFE and the MMSE-LE against average SNR.](image)

Figure 7 shows the simulated BER for the three receivers versus fixed SNR. Here the energy in $H$ has been normalized to one for each outcome of $H$ and the noise variance $\sigma^2$ has then been varied.

![Fig. 7. The estimated BER for the B-CLVE, MMSE-DFE and the MMSE-LE against fixed SNR.](image)

Looking at the simulations in figures 6 and 7 we see that the performance of the B-CLVE based on the WLS-design is slightly better than the performance of the MMSE-DFE, indicating a potential for the presented type of receivers. We also see that the performance of the VA without pre-filtering (WPF) is the worst among the simulated receivers.

VII. SUMMARY

This paper presents ideas about the use of reduced-state Viterbi detection in block transmission systems employing multiple receiving antennas and maximal-ratio combining. The effect of using a maximal-ratio combiner in terms of sufficient statistics, ISI, and the mean and variance of the SNR over a series of transmitted blocks is briefly discussed. The presented receiver structure, referred to as B-CLVE, is a generalization of the earlier presented combined linear-Viterbi equalizers for single-antenna, continuous transmission systems. Simulations using a version of the WLS-design method in [12], [16], [17] adopted for B-CLVE design indicate a potential for this class of receivers. However, the WLS-design method seems to be too computational complex for most applications to be implemented with today's technology.

APPENDIX

I. APPENDIX: THE POSITIONING MATRIX $D_i$

Let $0_{i \times a}$ denote an $i \times a$ matrix containing zeros and $I_{i \times i}$ denote an $i \times i$ identity matrix. If $0 < i \leq N - l + 1$ let

$$D_i = \begin{bmatrix} 0_{i \times (i-1)} & I_{i \times i} & 0_{i \times (N-i+l-1)} \end{bmatrix},$$  

and if $N - l + 1 < i \leq N$

$$D_i = \begin{bmatrix} 0_{i \times (i-1)} & 0_{i \times (N-i+l-1) \times (N-i+l)} & I_{(N-i+l) \times (N-i+l-1)} \end{bmatrix}.$$  

REFERENCES


Part II.4

Implementation of a System for Validation of Algorithms Used in Digital Radio Communication Schemes
Part II.4:
Implementation of a System for Validation of Algorithms Used in Digital Radio Communication Schemes
Mikael Isaksson¹, Roger Larsson¹ and Per Ödling²
The hardware platform has also been presented at the *Nordic Radio Symposium, NRS'92*, in Ålborg, Denmark.

¹Telia Research AB, Communication Systems, S-977 75 Luleå, Sweden
²Luleå University of Technology, Division of Signal Processing, S-971 87 Luleå, Sweden
IMPLEMENTATION OF A SYSTEM FOR VALIDATION OF ALGORITHMS USED IN DIGITAL RADIO COMMUNICATION SCHEMES

Mikael Isaksson  
Telia Research AB  
S-951 75 Luleå

Roger Larsson  
Telia Research AB  
S-951 75 Luleå

Per Ödling  
Luleå University of Technology  
S-951 87 Luleå

ABSTRACT

A hardware system for testing and evaluating algorithms and methods used in the physical layers of communication schemes is presented. The system aims to make the real time validation phase of developing a communication concept as flexible as the simulation phase, and will in fact partially merge these phases.

The implementation of the system utilizes standard signal processors, array signal processors and field programmable gate arrays.

A realization of a communication scheme of GSM-type, with a throughput rate of an order of magnitude higher than specified for GSM, is given as an example.

1. INTRODUCTION

The development of a modern communication system is usually made up of four phases: design, simulation, real time validation and VLSI-implementation. A system to facilitate the real time validation phase is presented. This system is for the sake of clarity referred to as the RVS (Real time Validation System). The RVS is planned to be ready for use in late 1992.

The development of validation systems is a very expensive process in terms of time and cost. It is often necessary to design specific hardware for each system to be able to achieve the required capacities.

The RVS is a reconfigurable hardware system for the validation of the physical layer of a digital point to point communication system. The main purpose of the RVS is to give the real time validation phase the flexibility of the simulation phase. Adding or changing an algorithm in the RVS should be carried out with the same ease as in a software simulation system. Actually the dividing line between the two phases becomes diffuse. Software modules and algorithms used in the simulations can in general be used when programming the RVS. It could even be convenient to run parts of the simulations on the RVS because of its throughput rate.

2. THE PHILOSOPHY OF THE RVS

A trend towards reusability and standardization has emerged in several technological areas during the 80’s. This has led to high performance standard components, e.g. signal processors and programmable gate arrays (PGA), becoming available on the market.

Developing, expanding and updating a system such as the RVS would not have been possible without strictly standardized interfaces and the ability to achieve high performance through parallel processing.

The RVS is primarily designed to be used for the validation of systems operating in a mobile radio environment, using either time division or spread spectrum as the channel access procedure. However, the structure of the RVS allows validation of a broader class of systems.

It is natural to divide a communication system into three physical parts: transmitter, channel and receiver. In real time validation it is important to have control over, and knowledge of, the environment of the test. Therefore the RVS includes a transmitter and a channel simulator to emulate a transmission system with a given communication scheme and given channel characteristics. Furthermore the RVS consists of a set of hardware modules for implementing a digital receiver.

The real time validation phase often contains tests under radio frequency conditions. The RVS can be used in field measurements by replacing the channel simulator with a radio frequency modulator and demodulator. The transmitter and the receiver are unaffected by this since the interfaces between the three parts of the RVS only consist of the analog baseband signals.

3. HARDWARE DESIGN PHILOSOPHY

The capacity of the RVS determines the class of systems it is possible to implement. In designing the system the major considerations have been: high throughput rate, reconfigurability and potential for incorporating new technology. High throughput rate can be reached by using computationally powerful components and by using parallel processing. The RVS uses standard components with standardized interfaces, which is essential for hardware modularity and the ability to utilize future technology. It is made up of programmable modules which can be reconfigured and replaced in an arbitrary fashion, in order to support the implementation of various types of algorithms.

4. HARDWARE CONFIGURATION

We now describe the implementation of the RVS that is planned to be ready for use in late 1992. All parts of the RVS are supervised by a PC, used for passing environmental parameters and control signals to and from the RVS. The PC is also used for the debugging and downloading of software and data to the various processing units of the RVS. (See figure 1.)
4.1 The receiver

The most powerful part, intended to be used as a receiver, is made up of special purpose signal processing modules, all having VME bus interfaces. The receiver is usually the most complex part of a digital communication system due to the large amount of signal processing needed to estimate data transmitted through a distorting channel.

There are three types of signal processing modules available in the RVS for implementing the receiver. These are standard signal processors, array signal processors and gate array boards. Furthermore the receiver includes a sampling unit for the conversion of the two incoming analog baseband signals (In-phase and Quadrature). Each of the modules uses the VME bus system for control and programming. The data signals are passed through special signal buses.

The sampled signals are stored in a queue memory used as an interface towards the signal processing units. The maximum sampling rate is 30 MHz for two channels, or 60 MHz for one channel, with a resolution of 8 bits. This unit also performs the generation of clock and timing signals for other units.

The parts of the implementation where the algorithm needs to be easily altered are best realized with parallel signal processors. The Texas Instrument TMS320C40 is used in the RVS. This processor model has recently been introduced on the market and is considered to be one of the most powerful [1]. One of the prime features of this processor is the six high-speed communication links, supporting the connection of additional C40 processors. The C40 concept includes software support, standardized interfaces and protocols for a three-dimensional multi-processor structure. The number of processors in one structure can be very large and is easily changed. The C40 signal processor has a computational capacity of 50 Mflops.

For array operations over large amounts of data, e.g. FFT, filtering and correlation, special array signal processors are well adapted. These operations are often difficult to parallelize. The RVS contains the array signal processor DaSP a66210 by Array Microsystems. The DaSP is designed for FFT-transforms and operations in the frequency domain. A complex 1024 points FFT can be computed in 131 microseconds.

For parts where computational power is crucial, specially designed field programmable gate arrays (FPGA) are being used. The FPGAs used are RAM-based LCA-devices* by XILINX. Designing and verifying the FPGAs can be done in a few days, compared to weeks for other gate array technologies.

A high degree of parallelism can easily be obtained, which makes the FPGAs well suited for performing algorithms that can be divided into many small, parallel parts, e.g. the Viterbi algorithm. The FPGAs can be programmed via the VME-buss.

4.2 The channel simulator

The channel simulator is developed at Telia Research [2] and is capable of simulating a time-variant, frequency selective fading channel, e.g. the Rayleigh faded channels specified in COST 207 [3] for GSM. The simulator works on baseband signals. When RF-signals are needed the simulator is complemented with external up and down converters. The channel simulator is based on a VME-bus system and is controlled by a PC. The VME-bus system contains the signal processing units and a CPU-unit for system control. The maximum bandwidth of the analog baseband signal is 15 MHz. The number of reflexes (taps) possible is 10 and each tap is capable of generating a time-variant uncorrelated fading. The doppler frequency of the fading process can be varied in a range of 0-1000 Hz, and the delay between two consecutive taps can be varied in the range of 0.1-12.1 microseconds. The attenuation of each tap can be varied arbitrarily to create a desired power-delay spectrum. In order to simulate slow log-normal fading the tap attenuation can be varied in real time.

A planned extension of the simulator includes the possibility to replay an actual, pre-recorded channel [4].

4.3 The transmitter

As a transmitter the LeCroy Arbitrary Function Generator (AFG) is used. This instrument is capable of transmitting prestored bursts of modulated data in real time, with the possibility of using different types of burst formats and symbol rates. The signal can be generated by a software simulation package, e.g. Cossap™, SPW™ or Mat®. This gives the possibility of distorting the signal before transmission. As an alternative a sampled physical signal can be used. The signal files are downloaded to the RAM-memory of the AFG before starting the test run.

The AFG offers the opportunity of transmitting different parts of data, e.g. bursts, in a pseudo-random way. The transmitters memory capacity of 2 Megabytes is sufficient for the validation of most of todays communication schemes.

5. AN EXAMPLE OF AN IMPLEMENTATION

The features of the RVS are best illuminated by presenting a validation of a well known communication system. As a system specification we choose the physical layer specified for GSM [5][6]. In the GSM specification the receiver algorithm is not strictly specified so we choose the Maximum Likelihood Sequence Estimator (MLSE) described by Ungerboeck [7] for the equalizer part of our receiver. This (optimal) receiver algorithm is well suited as an example because of its high performance and considerable complexity.

* Logical Cell Arrays
GSM is a time division multiple access (TDMA) system where a channel is divided into eight consecutive time slots, one time slot per mobile. A physical GSM channel has a transmission rate of 270 kbit/second. To utilize the capacity of the RVS we design the MLSE-receiver for accessing every time slot, or alternatively accessing eight parallel time-divided GSM channels. This is also equivalent to validating a time divided system with a transmission rate of Mbit/second as long as the channel distortion follows the GSM specification.

The GSM channel is characterized by severe intersymbol interference, additive Gaussian noise and Rayleigh fading. Figure 2 depicts the receiver divided into logical parts.

5.1 The Simulation Phase

The GSM subsystem including transmitter, channel and receiver is simulated in the simulation software package SPW™. Each module in the simulated system corresponds to a logical module in the implementation, see figure 2. A majority of the modules in the simulated receiver is written in the C programming language and can therefore be used almost directly when programming the signal processors of the RVS.

5.2 Algorithm

The algorithm proposed on [7] contains a matched filter and a modified metric computation in the Viterbi algorithm. The metric can be computed recursively as:

$$J_n(l_n) = J_{n-1}(l_{n-1}) + \max(U_n)$$

with

$$U_n = \text{Re} \left[ J_n \left( 2y_n - x_0l_n - 2 \sum_{m=1}^{L} x_ml_{n-m} \right) \right]$$

where $y_n$ is the the sampled output of the matched filter, $x_n$ is the sampled autocorrelation function of the channel and $l_n$ is the symbol sequence. The intersymbol interference spans over a time interval of $L$ symbol durations.

It is convenient to divide the MLSE into three logical and independent parts [9]: the Branch Metric Unit block (BMU), the Add, Compare and Select block (ACS) and the Survivor, Trace and Decode block (STD). These blocks are described below.

5.3 Transmission system

The underlying transmission system model is of TDMA type with data in bursts. The modulation scheme is Gaussian Minimum Shift Keying (GMSK) [5][6].

The bursts contains 116 encrypted data bits consisting of two parts of 58 bits. A 26 bit reference sequence is placed in between the two 58 bit blocks to allow symbol timing, carrier phase recovery and channel estimation. Three additional tail bits complete the burst at both ends. The channel and sequence estimation is performed once per burst, independently of other bursts.

5.4 Synchronization and channel estimation

The 26 bit reference sequence is used for symbol timing and carrier phase recovery [5][6], and for estimation of the channel impulse response. To compute the optimum decision instant and the carrier phase in the received burst a correlation is computed of the entire burst and the 16 central bits of the reference sequence stored in the receiver. The DaSP processor is well suited to perform this operation.

The channel estimate is then computed by an additional correlation of the 16 subsequent symbol samples from the decision instant in the burst and the prestored reference sequence. The channel estimate is computed for a maximum time dispersion of 4 symbol intervals. The channel estimate is essential information for the MLSE receiver and the matched filter.

5.5 Branch Metric Unit (BMU)

Before Viterbi decoding the metric value of each branch in the trellis has to be computed. The trellis contains 16 states according to the time dispersion of 4 symbol intervals. Two branches lead into each state in the trellis which give 32 metric computations for each symbol in the burst. These computations are performed by the BMU unit using the formula [7][8]:

$$U_n = 2 \text{Re}(l_n^* y_n) - 2 \text{Re} \left( l_n^* \sum_{m=1}^{L} x_ml_{n-m} \right) - |l_n| x_0$$

The formula can be further simplified by using the GMSK properties and the fact that terms which do not include $y_n$ can be computed once for each burst.
The simplified formula can be expressed as:

\[ U_n = \text{Re}\{ T_n (y_n - T_s) \} \]

The look-up table \( T_s \) is computed once for each burst and will be used in the metric computations for each state in the trellis by simply fetching the proper value from \( T_s \).

The set of BMU operations is performed by the C40 processors and is one of the parts of the algorithm requiring most computational capacity. The complexity of this operation makes it difficult to use the FPGA technology.

5.6 Add, Compare and Select units (ACS)

For each symbol in the burst the maximum metric values are computed for each of the 16 states in the trellis. Two branches, or paths, lead into each state in the trellis corresponding to the two possible state transitions. The branch metric values, computed by the BMU, are then added to the accumulated metric value in the predecessor state. The path giving the maximum metric value is saved and the state is given the new accumulated metric value.

The FPGA technology is a suitable choice for implementing these ACS operations. The computations are performed by using 16 parallel ACS units. In order to minimize the number of gate array chips and interconnections it is desirable to maximize the number of ACS units in each chip. This can be done by implementing each ACS unit as a one-bit processor using bit-serial design. Bit-serial design requires considerably fewer gate equivalents than a corresponding bit-parallel design [10]. The drawbacks are higher clock rates and the need for parallel/serial conversion.

5.7 Survivor, Trace and Decode unit (STD)

Using the saved paths from the ACS operation the symbol sequence can be decoded. Since only one path leads in to each state, the state transitions uniquely estimate the symbol sequence. These operations are called Survivor, Trace and Decode (STD) and are implemented in a simple hardware design containing dual-port RAMs and counters.

The STD is implemented on a separate board with VME bus interface and is the part of the receiver producing the output of the RVS, i.e. the estimated symbols.

6. COMMENTS AND FUTURE ACTIVITIES

We see several possible ways of using the RVS. Its primary task is to facilitate the real time validation phase of developing a communication system, and this is where the RVS has its economic importance. Another important use of the RVS is as a research tool. As an example of this it will be used for evaluating receiver algorithms as a part of a research project at Luleå University of Technology [11]. Moreover it can be used as a 'number cruncher' for running heavy simulations.

We plan to present the results of the implementation example described in section 5 in more detail during 1993. The results from the cooperation with the research group at Luleå University of Technology, mentioned above, will also be the subject for future presentations.

REFERENCES

[3] Proposal on channel transfer functions to be used in GSM tests late 1986, COST 207 WGI TD(85).
Part III.1

A Robust Correlation Receiver for Distance Estimation
Part III.1:
A Robust Correlation Receiver for Distance Estimation
Håkan Eriksson\textsuperscript{1}, Per Ola Börjesson\textsuperscript{1}, Per Ödling\textsuperscript{1} and Nils-Gunnar Holmer\textsuperscript{2}


\textsuperscript{1}Luleå University of Technology, Division of Signal Processing, S-971 87 Luleå, Sweden
\textsuperscript{2}Lund University, Department of Biomedical Engineering, S-221 00 Lund, Sweden
A Robust Correlation Receiver
for Distance Estimation

Håkan Eriksson
Per Ola Börjesson
Per Ödling
Nils-Gunnar Holmer

Reprinted from
IEEE TRANSACTIONS ON ULTRASONICS, FERROELECTRICS, AND FREQUENCY CONTROL
Vol. 41, No. 5, September 1994
A Robust Correlation Receiver for Distance Estimation

Håkan Eriksson, Student Member, IEEE, Per Ola Börjesson, Member, IEEE,
Per Ödling, Student Member, IEEE, and Nils-Gunnar Holmer, Member, IEEE

Abstract—Many methods for distance estimation, such as the ultrasonic pulse-echo method, involve the estimation of a Time-of-Flight (TOF). In this paper, a signal model is developed that, apart from the TOF, accounts for an unknown, linear frequency dependent distortion as well as for additive noise. We derive a TOF estimator for this model based on the criteria of Maximum Likelihood. The resulting receiver can be seen as an extension or generalization of the well known cross-correlation, or “matched filter”, estimator described, e.g., by Nilsson in [12]. The novel receiver is found to be more robust against unknown pulse shape distortion than the cross-correlation estimator, giving less biased TOF estimates. Also, bias versus noise sensitivity can be controlled by proper model order selection.

Index Terms—Time-of-Flight, delay, estimation, narrow-band, ultrasound, robustness, pulse-echo method.

I. INTRODUCTION

The most widely used method for making ultrasonic measurements is the ultrasonic pulse-echo method, introduced by Pellam and Galt in 1946 [14]. It is used for the investigation and detection of changes in acoustic impedance. Applications can, for instance, be found in medical diagnostics [8] and nondestructive material testing [17]. An electric pulse is applied to an ultrasonic transducer which is in acoustic contact with the sample to be examined. The transducer converts the electric pulse into an acoustic pressure wave that is transmitted into the sample. Reflections will appear where the acoustic impedance changes, e.g., in surfaces between different tissues, and a pressure wave will be reflected to the transducer. The transducer receives and converts the reflected pressure wave into an electric signal. The objective is then to extract the desired information by analyzing this signal. In the case of distance estimation it is often assumed that the distance to be measured is proportional to the corresponding pressure wave flight time. This motivates our current interest in Time-of-Flight (TOF) estimation.

When performing TOF estimation the choice of system model is an intricate matter, also involving defining the TOF. In [12] the assumption that the received signal consists of nonoverlapping delayed and scaled replicas of a known “reference” echo renders the well-known cross-correlation, or “matched filter”, estimator. The received signal is correlated with the reference signal and the position of the maximum of the absolute value of the correlation function yields the estimate of the TOF. This estimator is optimal in the sense that it is the Maximum Likelihood estimator of the TOF given additive white Gaussian noise with zero mean [18]. In general the shapes of the received echoes depend on many parameters, such as the locations and the geometry of the reflectors. Many methods for TOF estimation are based on the assumption that some feature of the received echoes is essentially invariant, for instance the shape [7], [12], the first zero crossing [3], [11], the amplitude peak of the echoes [15], the amplitude peak of the envelope of the echoes [19] or the initial part of the received echoes [5]. Then, in [3], [5], [7], [11], [15], [19] the positions in time of these features gives the TOF estimates. If the assumptions of invariance of the features in question are violated, the mentioned techniques for TOF estimation will be likely to produce biased estimates. In fact, the bias of the TOF estimate could well be greater than the standard deviation of the same, as we will exemplify in the sequel.

Throughout this paper we consider a propagation media containing only one reflector and the estimation of a single TOF. We address the problem of how to derive a receiver for TOF estimation when the shape of the received echo is unknown or uncertain.

The paper proceeds as follows. A method for modeling uncertainties in the shape of the received signal in narrow-band pulse-echo systems is presented in Section II. The modeling method is based on describing the received echo as a known narrow-band reference signal being distorted by an unknown linear and time-invariant system and corrupted with additive noise. This linear system is modeled in the frequency domain with a Taylor series expansion around the center frequency of the reference signal. The coefficients of the Taylor series expansion hold the uncertainty in the system. Inherent in the model is a definition of the TOF. In Section III we derive a TOF estimator, based on the ML criteria, using an $M$th order truncation of the model derived in Section II. The estimation procedure consists of a joint estimation of the TOF and the $2M$ coefficients of the truncated Taylor expansion. In Section IV the estimator is simulated and its performance is compared with the performance of the cross-correlation estimator in terms of bias and standard deviation. Section V describes measurements where the cross-correlation estimator and the estimator of Section III is used to produce ultrasonic images for different model orders $M$. Finally, in Section VI comments on the results can be found.
II. LINEAR DISTORTION OF NARROWBAND SIGNALS

Consider the situation described in Fig. 1. An unknown excitation signal is applied as an input to two unknown linear systems. The output signal from one of the systems, \( s(t) \), is a real valued known narrow-band signal, henceforth called the reference signal. The output from the other system is denoted \( y(t - \theta) \), where \( \theta \) is the unknown delay to be estimated.

Assume that there exists an impulse response \( h(t) \) so that the noiseless waveform

\[
y(t - \theta) = (s * h)(t - \theta),
\]

where * denotes continuous time convolution and \( h(t) \) is an unknown real valued impulse response representing linear distortion. Since the real valued reference signal is narrow-band, we write it as

\[
s(t) = \Re \{ e^{2\pi if_0 t} \alpha(t) e^{j\beta(t)} \},
\]

where \( \Re \{ \cdot \} \) denotes the real part of its argument. The functions \( \alpha(t) \) and \( \beta(t) \) are real valued with the restriction that the Fourier transform of the complex envelop, \( \alpha(t)e^{j\beta(t)} \), is zero outside the frequency interval \([-B, B]\), where \( B \ll f_0 \). We also assume that the complex envelop has finite energy, i.e., that \( \alpha(t)e^{j\beta(t)} \in L_2(\mathbb{R}) = \{ z(t) \mid |z(t)|^2 dt < \infty \} \).

Define \( L_2^F(\mathbb{R}) \) as the space of all signals in \( L_2(\mathbb{R}) \) whose Fourier transforms are zero outside the frequency interval \([-B, B] \). These signals will be referred to as baseband signals. Furthermore, due to the narrow-band assumption \((B \ll f_0)\) on \( s(t) \), it is convenient to define the subspace \( L_2^F(\mathbb{R}) \) of \( L_2(\mathbb{R}) \) as the space of all signals in \( L_2(\mathbb{R}) \) whose Fourier transforms are zero outside the frequency intervals \([-f_0 - B, -f_0 + B]\) and \([f_0 - B, f_0 + B]\). By definition \( \alpha(t)e^{j\beta(t)} \in L_2^F(\mathbb{R}) \) and by construction \( s(t) \in L_2^F(\mathbb{R}) \). Note that \( s(t) \) can be any narrow-band signal having finite energy whose Fourier transform only has support in two intervals, symmetrically around \( f = 0 \) and not containing \( f = 0 \). These signals could typically be bandlimited ultrasonic pulse echo signals decaying fast enough in time to have finite energy.

Models related to (1) can be found in, e.g., [2], [4], [12]. The special case of (1) where \( h(t) \) is modeled as the sum of weighted and delayed Dirac-delta functions has been extensively used in the context of TOF estimation, see, e.g., [12]. Recently in a paper by Boudreau and Kabal [2] a model for the purpose of TOF estimation is presented, allowing an unknown linear frequency dependent distortion in the delay path. The reference signal was there modeled as a broadband stochastic process for the tracking of a time varying TOF with an adaptive algorithm. A discrete time version of (1) is studied in [4], where the distortion is modeled as a linear stochastic system with a known Gaussian statistic.

A. An Mth Order Model of the Distorting System \( h(t) \)

Due to the narrow-band assumption on \( s(t) \) we are only interested in the behaviour of \( h(t) \) in the frequency intervals \([f_0 - B, f_0 + B]\) and \([-f_0 - B, -f_0 + B]\). Let \( H(f) \) denote the Fourier transform of \( h(t) \) and assume that \( H(f) \) is \( M \)-times differentiable in an interval around \( f = f_0 \). Define \( C_M(f) \) as the Taylor series expansion of \( H(f) \) around \( f_0 \) with a remainder term of order \( M \):

\[
C_M(f) = \sum_{k=0}^{M-1} \frac{(-1)^k}{k!} \frac{d^k H(f)}{df^k} \bigg|_{f=f_0}.
\]

Now, let us define a function \( H_M(f) \) such that \( H_M(f) = C_M(f) \) for \( f > 0 \) and the inverse Fourier transform of \( H_M(f) \) is real valued,

\[
H_M(f) = C_M(f)U(f) + C_M^*(f)U(-f),
\]

where \( U(f) \) denotes the unit step function and the superscript * denotes the complex conjugate.

The system model \( H_M(f) \) is then an \( M \)th order model of \( H(f) \), especially adapted for narrow-band signals with their center frequency equal to \( f_0 \). A similar reasoning by Papoulis can be found in [13, page 59] for baseband signals.

B. Evaluation of the Convolution \( (s * h)(t) \)

For the purpose of modeling the convolution \((s * h)(t)\) in (1), we define

\[
y_M(t) \triangleq \mathcal{F}^{-1}\{\mathcal{F}\{s(t)\}H_M(f)\},
\]

where \( \mathcal{F}\{\cdot\} \) denotes the Fourier transform and \( \mathcal{F}^{-1}\{\cdot\} \) its inverse. By using the narrow-band representation of the reference signal in (2), it follows that

\[
y_M(t) = \Re \left\{ e^{2\pi if_0 t} \sum_{k=0}^{M-1} (a_k + j b_k) g^{(k)}(t) \right\},
\]

where \( a_k \) and \( b_k \) are real valued coefficients given by

\[
a_k + j b_k = \frac{1}{(2\pi)^k k!} \frac{\partial^k H(f)}{\partial f^k} \bigg|_{f=f_0},
\]

and

\[
g^{(k)}(t) = \frac{d^k}{dt^k} \left\{ \alpha(t)e^{j\beta(t)} \right\}.
\]

All \( g^{(k)}(t) \) exist and have Fourier transforms that are zero outside the frequency interval \([-B, B]\) since the subspace \( L_2^F(\mathbb{R}) \) is closed under linear operations and \( g^{(0)}(t) \) is analytic having derivatives of all orders in every point on \( \mathbb{R} \).

We now have obtained a signal representation, (6)–(8), with the model order \( M \) to choose. Observe that the waveforms \( \{g^{(k)}(t)\}_{k=0}^{M-1} \) are independent of the system \( h(t) \) and that the coefficients \( \{a_k, b_k\}_{k=0}^{M-1} \) are independent of the reference signal once the center frequency is given. Thus, the knowledge of the reference signal gives us a signal basis and the
coefficients \( \{ a_k, b_k \}_{k=0}^M \) hold the uncertainty in the shape of the waveform \( y_M(t) \).

For the development of the estimator in Section III we need an orthogonal signal basis. Define the inner product \( \langle \cdot, \cdot \rangle \) in \( L_2(\mathcal{R}) \) as

\[
x, y \triangleq \int_{\mathcal{R}} x(t) y^*(t) \, dt,
\]

and the norm \( ||x|| \triangleq \langle x, x \rangle^{1/2} \). Since \( g^{(0)}(t) \in L_2^B(\mathcal{R}) \) it can be shown that the signal set \( \{ g^{(k)} \}_{k=0}^{M-1} \) spans an \( M \)-dimensional signal space \( \Omega_M \) in \( L_2^B(\mathcal{R}) \). Let \( \{ z_k \}_{k=0}^{M-1} \) denote an orthonormal signal basis that spans \( \Omega_N \), for every \( N \) in \( \{1, \ldots, M\} \), which ensures the signal basis \( \{ z_k \}_{k=0}^{M-1} \) to be in a specific order. This signal basis can, for instance, be obtained by applying Gram-Schmidt orthogonalization to \( \{ g^{(k)} \}_{k=0}^{M-1} \). Let us modulate the signal basis \( \{ z_k \}_{k=0}^{M-1} \) up to the frequency band of interest and define

\[
\psi_k(t) = \sqrt{2 \Re} \left\{ z_k(t) e^{j 2 \pi f_0 t} \right\}, \quad k = 0, 1, \ldots, M - 1,
\]

where \( \Re \{ \cdot \} \) denotes the imaginary part of its argument. Since the Fourier transform of \( z_k \) is narrow-band \( (B \ll f_0) \) and \( (z_k, z_l) = \delta_{k_l} \) the function set \( \{ \psi_k, \bar{\psi}_k \}_{k=0}^{M-1} \) is a real orthonormal basis of dimension \( 2M \) in \( L_2^B(\mathcal{R}) \). The model signal \( y_M(t) \) in (5) can then be expressed as a linear combination of \( \{ \psi_k, \bar{\psi}_k \}_{k=0}^{M-1} \) according to

\[
y_M(t) = \sum_{k=0}^{M-1} c_k \psi_k(t) + d_k \bar{\psi}_k(t). \tag{11}
\]

The unknown coefficients \( \{ c_k, d_k \}_{k=0}^{M-1} \) are the new coordinates for \( \{ a_k, b_k \}_{k=0}^{M-1} \) after the change of signal basis. The coefficients \( \{ c_k, d_k \}_{k=0}^{M-1} \) now hold the uncertainty in the shape of the waveform \( y_M(t) \). Note that \( \psi_k(t) \) is a normalized replica of the reference signal, i.e., \( \psi_0(t) = s(t) ||s|| \), and that \( \psi_0(t) \) and \( \bar{\psi}_k(t) \) by construction form a Hilbert transform pair for all \( k \in \{0, \ldots, M - 1\} \).

C. A Summary of the Signal Basis Design

In this section we have presented a method for modeling linear distortion of narrow-band signals. The analysis results in a signal representation consisting of an orthonormal signal basis. The design process for obtaining this signal basis can be summarized as follows:

- For \( k = 0, 1, \ldots, M - 1 \), take the \( k \)th order derivative of the bandlimited complex envelope of the reference signal, see (8).
- Use for instance the Gram-Schmidt orthogonalization process to obtain a complex orthonormal baseband signal basis, \( \{ z_k \}_{k=0}^{M-1} \).
- Calculate the high frequency signal basis, \( \{ \psi_k, \bar{\psi}_k \}_{k=0}^{M-1} \), as described by (10).

In the case when \( s(t) \) is acquired by a calibration or measurement procedure, the design process will be affected by noise. To reduce this effect, a band-pass filter matched to the bandwidth of \( s(t) \) should be introduced when acquiring \( s(t) \), thereby reducing the noise power. The design process is described in Fig. 2 where LP denotes an ideal low pass filter with a bandwidth equal to \( B \), i.e., the bandwidth of \( a(t) e^{j \beta(t)} \).

III. TIME OF FLIGHT ESTIMATION

Consider the joint estimation of a single TOF and the coefficients describing the linear distortion. We model the received signal as

\[
r(t) \triangleq y_M(t - \theta) + n(t), \tag{12}
\]

where \( y_M(\cdot) \) is given by the \( M \)th order model (11) and \( n(t) \) is an additive white Gaussian noise process, with zero mean and a power spectral density equal to \( N_0/2 \). The parameter \( \theta \) is the TOF. With this definition of the TOF the shape of the waveform \( y_M(\cdot) \) is allowed to vary, cf., [3], [5], [7], [11], [12], [15], [19].

To get a compact notation we rewrite \( y_M(\cdot) \) of (11) as

\[
y_M(t - \theta) = p^T \psi(t - \theta), \tag{13}
\]

where \( \psi = [\psi_0, \ldots, \psi_{M-1}], \bar{\psi}_0, \ldots, \bar{\psi}_{M-1}]^T \) is a vector representation of the orthonormal signal basis and \( p = [c_0, \ldots, c_{M-1}, d_0, \ldots, d_{M-1}]^T \). With this notation we write the Signal to Noise Ratio (SNR) as

\[
\text{SNR} \triangleq \frac{||y_M||^2}{N_0/2} = \frac{p^T \bar{p}}{N_0/2}. \tag{14}
\]

A. The ML Estimator

The estimation problem is to determine the unknown TOF \( \theta \) when observing \( \{ r(t) \} \in \mathcal{I} \), where \( \mathcal{I} = [T_1, T_2] \) is the observation interval \( (T_1 < T_2) \). Define the deterministic parameter vector

\[
q \triangleq [\theta, p^T]^T \tag{15}
\]

and denote the estimate of \( q \) as \( \hat{q} = [\hat{\theta}, \hat{p}^T]^T \). The function

\[
l(q \mid r(t)) = \frac{2}{N_0} \int_{T_1} r(t) y_M(t - \theta) \, dt - \frac{1}{N_0} \int_{T_1} \bar{y}_M^2(t - \theta) \, dt, \tag{16}
\]
is a log likelihood function for $q$ given the observation $(r(t) \mid t \in \mathcal{I})$, c.f. [18, p. 274]. Using (13), (16) can be written as

$$l(q \mid r(t)) = \frac{2}{N_0} p^T r_{\theta} - \frac{1}{N_0} p^T p + \frac{1}{N_0} p^T E(\theta)p. \quad (17)$$

where $r_{\theta}$ is a vector resulting from $2M$ parallel correlations,

$$r_{\theta} = \int_{\mathcal{I}} r(t) \psi(t - \theta) dt, \quad (18)$$

and where the matrix

$$E(\theta) = \int_{\text{outside } \mathcal{I}} \psi(t - \theta) \psi^T(t - \theta) dt. \quad (19)$$

The only term involving the measurements $r(t)$ in (17) is the vector $r_{\theta}$, where $r_{\theta}$ is a sufficient statistic for the parameter vector $q$.

For a given observation $(r(t) \mid t \in \mathcal{I})$, the ML estimate of $q$ is obtained by locating the global maximum of the log likelihood function $l(q \mid r(t))$ with respect to $q$. Let $\hat{\theta} \in \mathcal{I}_{\theta}$, where $\mathcal{I}_{\theta}$ is a set of intervals, on $\mathbb{R}$. If the observation interval $\mathcal{I}$ is chosen sufficiently large for $\psi_k(t - \theta)$ and $\psi_k(t - \theta)$ to have negligible energy outside $\mathcal{I}$ for all $k \in \mathcal{I}_{\theta}$, then, according to Schwarz's inequality, each element in $E(\theta)$ is not greater than this negligible energy. Approximating $E(\theta)$ with zero and completing the square in (17) we have

$$l(q \mid r(t)) = -\frac{1}{N_0} (p - r_{\theta})^T (p - r_{\theta}) + \frac{1}{N_0} r_{\theta}^T r_{\theta}. \quad (20)$$

For any given $\theta$ the maximum of $l(q \mid r(t))$, with respect to $p$, is given by $p = r_{\theta}$. Using $r_{\theta}$ for $p$ in (20), it follows that the location of the global maximum of $l(q \mid r(t))$, with respect to $q$ (resulting in an approximate ML estimate of $q$), is given by

$$\hat{\theta}_{ML}(r) = \arg \max_q \{x_M(\theta)\} \quad (21)$$

\begin{equation}
\text{where} \quad x_M(\theta) = r_{\theta}^T r_{\theta}. \quad (22)
\end{equation}

Fig. 3 depicts a receiver structure for generating $x_M(t)$ from $r(t)$ according to (22). We henceforth refer to this receiver as the $M$th order extended correlation receiver. Note that an ML estimator, in its general form, need not to be unbiased. This also applies to the $M$th order extended correlation receiver.

B. A Comparison With the Cross-Correlation Estimator

The cross-correlation, or “matched filter”, estimator of [12] is derived using the Maximum Likelihood criterion and assumptions of additive white Gaussian noise and perfect knowledge of the linear distortion, except for an unknown amplification factor. The received signal is correlated with a known waveform corresponding to $y_M(t)$ in (16), and the position of the maximum of the absolute value of the correlation function yields the estimate of the TOF. This receiver, when applied to a situation where the assumptions are valid, is unbiased and has a performance that meets the Cramer-Rao lower bound for high SNR’s, see, e.g., [9] and the references therein. For low SNR’s the cross-correlation estimator will suffer from a threshold effect when applied to narrow-band signals in the sense that the performance decreases rapidly with decreasing SNR below a certain SNR [9].

The $M$th order extended correlation receiver can be described as a cross-correlation estimator that first estimates the received waveform $y_M(t - \theta)$ and then correlates the received signal $r(t)$ with this estimate. This can be seen, cf. [18, p. 354], by substituting the right-hand side of (18) for one of the $r_{\theta}$’s in (22), giving

$$x_M(\theta) = \int_{\mathcal{I}} r(t) \tilde{y}_M(t - \theta) dt, \quad (23)$$

where

$$\tilde{y}_M(t - \theta) = r_{\theta}^T \psi(t - \theta). \quad (24)$$

The function $\tilde{y}_M(t - \theta)$ can be interpreted as an estimate of $y_M(t - \theta)$ for a given $\theta$, cf. (13) with $p$ estimated by $r_{\theta}$.

IV. ESTIMATION EXAMPLE

Let us investigate the influence of the model order $M$ on the bias and the standard deviation of the TOF estimate by means of Monte Carlo simulations. We compare the performance of the cross-correlation estimator and the performance of the $M$th order extended correlation receiver for different values of the model order $M$ and various SNR’s.

In these simulations the reference signal $a(t)$ is given by (2) with $\beta(t) = 0$ and $\alpha(t)$ as a Hanning window in the frequency domain. With a Hanning window of bandwidth $B$, the Fourier transform of $\alpha(t)$ becomes

$$F(a(t)) = \begin{cases} 1/2 + 1/2 \cos \left( \frac{\pi f}{B} \right), & |f| < B \\ 0, & \text{otherwise} \end{cases} \quad (25)$$

The center frequency of the reference signal was chosen as $f_0 = 2B$ Hz. In all of these simulations the cross-correlation estimator is matched to the reference signal. For each estimator and every integer value of SNR, the TOF of one thousand echoes has been estimated.

In Figs. 5–6 the absolute value of the bias and the standard deviation of the receivers are shown for a case when the received signal has been subjected to a specific “unknown”
Fig. 4. The reference signal \( s(t) \) and the model \( y_M(t) \) in (11) in the case when \( M = 2 \) and \( p = [1/3, 2/3, 1/3, 1/6]^T \).

Fig. 5. The absolute value of the estimated and normalized bias of the receivers plotted versus SNR when a specific frequency dependent linear distortion is present.

linear distortion. Note that the bias and the standard deviation in the figures have been normalized with \( 1/B \). The distortion is described by \( y_M(t) \) in (13), with \( M = 2 \) and \( p = [1/3, 2/3, 1/3, 1/6]^T \). Fig. 4 shows \( y_M(t) \) together with the reference signal \( s(t) \). In Fig. 7 the standard deviation of the receivers is given for the case when no frequency dependent linear distortion is present, i.e., \( y_M(t) = a s(t) \) where \( a \) is an unknown constant. The bias was, in this case, almost equal to zero for all receivers.

We make the following observations. When linear distortion is present the bias of the cross-correlation estimator can be much greater than the standard deviation of the same. The TOF estimates of the first-order extended correlation receiver are almost as biased as the estimates of the cross-correlation estimator. As expected, since the linear distortion is generated with \( M = 2 \), the second-order extended correlation receiver has very little bias compared with the cross-correlation estimator. It can be noted that the variance of the estimates in general increases with the model order \( M \), especially when no linear distortion is present. An exception can be seen in Fig. 6 for SNR less than 22 dB. Here both the bias and the variance of the TOF estimate are small for the second-order extended correlation receiver whose model order is matched to the linear distortion.

From these and other simulations we draw the following conclusions. The cross-correlation estimator may give biased estimates when linear distortion not accounted for is present and the error due to bias can be more severe than the error caused by the additive white Gaussian noise. By using the \( M \)th order extended correlation receiver the bias caused by linear distortion can be reduced. Robustness against linear distortion of the shape of the received pulse can be traded for noise sensitivity by increasing the model order \( M \).

V. ULTRASONIC MEASUREMENTS

In this section we present results from using an experimental ultrasonic system to produce relief pictures of a surface structure by estimating the TOF of echoes reflecting from different points on the surface. This is a typical example of when it is difficult to choose a specific waveform as the reference signal for the cross-correlation estimator.

A. Measurement Setup

Experiments were carried out by scanning over a Swedish five crown coin in a 100 by 100 point grid, with a distance of 0.3 mm between adjacent points. The coupling medium between the transducer and the surface was air at room temperature. A circular transducer with a radius of 10.5 mm and a nominal frequency of 1 MHz was used. (The wavelength in air at room temperature is approximately 0.3 mm). The transducer was a focused transducer acoustically adapted to air, with a focal distance of 25 mm. It was designed by H. W. Persson for the investigation presented in [16]. The data were acquired with a LeCroy 9430 sampling oscilloscope. The sampling rate was chosen as 5.5 MHz and the sampling was synchronized with the transmitter by the use of an external crystal oscillator. The distance between the transducer and the surface of the coin was equal to the focal distance of the transducer.

B. Signal Processing and Measurements

A reference echo was acquired for the design of the signal basis \( \{\psi_k, \tilde{\psi}_k\}_{k=0}^{M-1} \), i.e., the receiver filter bank in Fig. 3,
Fig. 7. The estimated normalized standard deviation of the receivers plotted versus SNR when no frequency dependent linear distortion is present.

Fig. 8. The basis \( \{\psi_k(t),\tilde{\psi}_k(t)\}_{k=0}^{M-1} \) corresponding to the acquired reference signal.

Fig. 9. The proportions of the features on the surface of the coin in comparison with the length of the reference signal.

Fig. 10. A grey scale picture of a Swedish five crown coin, generated with the cross-correlation estimator matched to the reference signal \( s(t) \).

by letting the acoustic wave be reflected on the flat tip of a cut steel needle with a diameter of 1 mm, placed in the focal point of the transducer, cf., [10]. When acquiring the reference signal an averaging over 1000 B-scans was performed. The design process, described in Section II, was then used to generate the signal set \( \{\psi_k(t),\tilde{\psi}_k(t)\}_{k=0}^{M-1} \) for \( M = 3 \), see Fig. 8. As mentioned before the signal \( \psi_0(t) \) is identical to the normalized reference signal. The reference signal is also illustrated in Fig. 9 and compared with the features on the surface of the coin.

The \( M \)th order extended correlation receiver, (21)-(22), and the cross-correlation estimator were used to estimate the TOF for each point in the 100 by 100 point grid. Fig. 10 shows the result when the cross-correlation estimator matched to the reference signal \( s(t) \) was used. Figs. 11-13 show the results when the \( M \)th order extended correlation receiver was used with \( M \in \{1, 2, 3\} \), respectively. All figures are derived from the same measured data.

In the grey-scale pictures the brightness of each pixel increases with the TOF. No post-processing has been performed to the data, except for a histogram normalization of the grey-scale level thresholds in the grey-scale pictures.

The SNR, when defined as in (14), varies from point to point in the grid, since the reflected signal energy depends on, e.g., the angle of the reflecting surface. Therefore, no SNR is presented in the figures. The SNR was though well above the threshold region [9] for the cross-correlation estimator (SNR > 30 dB).

C. Comments on the Measurements

In the estimation examples presented the image corresponding to \( M = 1 \), Fig. 11, shows a significant improvement in quality compared to the image corresponding to the cross-correlation estimator, Fig. 10. The performance of the cross-correlation estimator is apparently sensitive to shape changes in the received waveform. The errors in the picture are mainly due to bias, since the SNR is high, and can for that reason not be removed by averaging repeated measurements. When
applying the receiver with \( M = 3 \) (Fig. 13) the disturbances due to the additive noise seem to dominate over the effects due to model errors. For this measurement situation the choices \( M = 1 \) and \( M = 2 \) appear to be the most appropriate.

The choice and the acquisition of the reference signal is a delicate matter, see, e.g., [6]. In this investigation we have used the echo from a cut needle located in the focal point of the transducer. We have not found any reference signal that gives the cross-correlation estimator a noticeable better performance in the presented measurement situation. However, other reference echoes could be used, for instance an echo from a flat surface where total reflection occurs, giving similar images.

VI. CONCLUSION

This paper treats the problem of how to design a receiver for TOF estimation when the received signal has been subjected to an unknown linear distortion. Based on a narrow-band assumption on the interrogation pulse, a model of the received signal incorporating unknown linear frequency dependent distortion has been presented together with a TOF estimator based on the criteria of Maximum Likelihood. This ML estimator can be seen as an extension or generalization of the cross-correlation estimator [12], and is referred to as the extended \( M \)th order correlation receiver. The receiver can be implemented as a correlation receiver using a filter bank consisting of \( 2M \) parallel filters.

Simulations have shown that the performance of the cross-correlation receiver is sensitive to the choice of reference signal. The bias of the TOF estimate can be large compared with the standard deviation. By using the \( M \)th order extended correlation receiver the bias caused by linear distortion can be reduced. In the presented ultrasonic measurements we observe that the \( M \)th order extended correlation receiver achieves robustness against pulse shape distortion at the cost of increased noise sensitivity. By increasing the model order the tendency of bias in the TOF estimate can be traded for variance.

In this paper we only examine the case of estimating a single TOF. The receiver could be extended to estimating multiple TOF’s, possibly with the exception of closely spaced echoes. We would like to point out that although this investigation is focused on ultrasound applications, the model presented only presumes a narrow-band pulse being distorted by a linear time-invariant system. Such conditions could also be present in other application areas.

REFERENCES


Håkan Eriksson was born in Stockholm, Sweden, in 1964. He received the M.Sc. degree in computer science from the Department of Computer Science and Electrical Engineering at Luleå University of Technology, Luleå, Sweden. In July 1989, he began working towards the Ph.D. degree at the Division of Signal Processing at Luleå University of Technology. He presented his licentiate thesis on "Modelling of Waveform Deformation and Time-of-Flight Estimation" in June 1993. His current research involves parameter estimation in ultrasound signals and the properties of reduced complexity digital receivers.

Per Öling was born in Örnsköldsvik, Sweden, in 1966. He received the M.Sc. degree in 1989 and the Tech. Lic. degree in 1993, both from Luleå University of Technology, Sweden. Currently, he is a Ph.D. student at the Division of Signal Processing, Luleå University of Technology. His research interests include the signal processing aspects of receivers used in telecommunications and receivers used in ultrasonics.

Nils-Gunnar Holmer was born in 1943 in Malmö, Sweden. He received the M.Sc. degree in electrical engineering 1968 and the Tech. Lic. degree in 1971, both from the Lund Institute of Technology. In 1978 he received the Ph.D. degree with a doctoral thesis entitled "New methods in medical ultrasound". In 1981 he was accepted as an unpaid associate professor in electrical measurements at the Lund Institute of Technology. During 1973–1975 he was employed at the Department of Food Engineering where he designed and constructed measurement instruments for physical and chemical analysis. In 1975 he was with the Department of Biomedical Engineering, Malmö General Hospital as Vice Head of the department. In 1982 he became the Head of the Department of Biomedical Engineering at the Lund University Hospital. In 1986 he was appointed Professor of Biomedical Engineering at the University of Lund and in 1991 became the Supervising Chairman of the Biomedical Engineering Organization at Malmöhus County Council, which consists of five hospitals.
Part III.2

Simultaneous Time of Flight and Channel Estimation Using a Stochastic Channel Model
Correction

We make a correction to this conference paper, Part III.2. In this paper it is assumed that the transmitted pulse is both bandlimited in frequency and identically zero over a semi-infinite time-interval. A signal cannot completely fulfill both these criteria. Instead of being bandlimited in frequency, the signal should have most of its energy concentrated to a narrow frequency interval. (One alternative would have been the formulation of Part III.1.)

Part III.2:

Simultaneous Time of Flight and Channel Estimation Using a Stochastic Channel Model

Håkan Eriksson\(^1\), Per Ödling\(^1\) and Per Ola Börjesson\(^1\)


\(^1\)Luleå University of Technology, Division of Signal Processing, S-971 87 Luleå, Sweden
Simultaneous Time of Flight and Channel Estimation Using a Stochastic Channel Model

Håkan Eriksson\textsuperscript{1,2} Per Ödling\textsuperscript{1} Per Ola Börjesson\textsuperscript{1}
\textsuperscript{1} Luleå University of Technology
Div. of Signal Processing
S-951 87 Luleå
\textsuperscript{2} University of Karlskrona/Ronneby
Dep. of Signal Processing
S-372 25 Ronneby

Abstract
In this paper we address the problem of estimating the Time-of-Flight of a transmitted signal when the shape of the received waveform is stochastic. Specifically, we examine the case when the transmission system model is stochastic, linear and time discrete, with additive Gaussian noise, and where the transmitted waveform is known to the receiver. The joint estimation is couched in terms of Maximum a Posteriori (MAP) and Maximum Likelihood estimation. When deriving the MAP estimator we assume a priori knowledge of the probability density of the transmission system impulse response. The MAP estimator is then compared to estimators derived using less a priori information and lower order system models.

The ordinary correlation based Time-of-Flight estimator assumes knowledge of the received waveform, that is has a one-dimensional transmission system model. This investigation indicates that a more complex model structure is worthwhile when distortion in excess of low additive noise is present.

1 Introduction
A good example of Time-of-Flight estimation is the pulse-echo method when used to measure an unknown distance. Idealized, these applications utilize the assumption that the distance to be measured is proportional to the time it takes for the waveform to travel the path to be measured, the Time-of-Flight (TOF). Various methods have been proposed for estimating the TOF. Most methods assume that the received waveform is known, except for a finite set of undetermined parameters. If the TOF only corresponds to a time delay of the received waveform, classical delay estimation can be performed using e.g. the criteria of Maximum Likelihood (ML) \cite{7}, Minimum Mean-Square Error (MMSE) \cite{3} or Maximum a Posteriori (MAP) \cite{2}.

When no parameter-dependent signal model is known to be appropriate, the choice of estimation method is more intricate. In the literature on ultrasound, several papers \cite{1, 4, 5, 8} have been concerned with this problem.

Our interest is in methods for modelling the uncertainty in the shape of the received waveform, where the amount of a priori information used is a critical issue. In this paper we model the uncertainty in the received signal waveform as an uncertainty in the impulse response of a time discrete linear system, and as an effect of additive noise.

By using a Gaussian assumption on the system statistics, we derive a MAP estimator of the TOF. A class of estimators using less a priori information about the second order statistics, and a system model of lower order, is also derived. These estimators are derived in a matrix formalism using an orthonormal basis for the system model, given by Singular Value Decomposition. If the transmitted waveform is narrow band, the ordered singular values produced by the SVD will rapidly approach zero, thereby encouraging the model order truncation.

The contents of this paper is as follows. Section 2 presents a stochastic model for the transmission system. In section 3 estimation algorithms for the TOF are derived. A numerical example using a narrow band transmitted waveform is given in section 4. Finally, section 5 comments on the result.
2 Model
The received waveform is modelled as a filtered and delayed version of a known waveform, and corrupted by additive noise, see Figure 1.

\[ s\theta[k] = s(kT - \theta) \cdot h[k] \rightarrow r[k] \]

Figure 1: The signal model.

The transmitted waveform \( s(\cdot) \) is assumed to be bandlimited in the frequency domain and sampled with the sample rate \( \frac{1}{T} \), greater than the Nyquist rate. The parameter \( \theta \) denotes the TOF, and is independent of the system impulse response \( h[k] \). The impulse response \( h[k] \) represents a stochastic linear system which is time-invariant during the time the waveform passes through it. The uncertainty in the received waveform is contained in the impulse response \( h[k] \), the TOF \( \theta \), and in the measurement noise \( n[k] \). The range of the parameter \( \theta \) is restricted to \( \Theta = [\theta_1, \theta_2] \), usually a continuous interval.

Assume that the system \( h[k] \) is causal with length \( L \), and that \( s(kT - \theta) = 0 \) when \( k < 0 \) and \( \theta \in \Theta \). The signal model can then be described in a matrix formalism as:

\[ r = S\theta h + n, \]

where \( r = [r[0], \ldots, r[M-1]]^T \), \( h = [h[0], \ldots, h[L-1]]^T \), \( n = [n[0], \ldots, n[M-1]]^T \) and

\[ S\theta = \begin{bmatrix} s\theta[0] & 0 & \cdots & 0 \\ s\theta[1] & s\theta[0] & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ s\theta[M-2] & \cdots & 0 & s\theta[0] \\ s\theta[M-1] & s\theta[M-2] & \cdots & s\theta[M-L-2] \end{bmatrix}. \]

(2)

The covariance matrix of \( h \), \( R = E{(h - h_0)(h - h_0)^T} \), and \( E{h} = h_0 \) are a priori known. The matrix \( R \) is symmetric by definition, and we assume that \( R \) is positive definite such that the factorization \( R = R^\frac{1}{2} R^{-\frac{1}{2}} \) and \( R^\frac{1}{2} R^{-\frac{1}{2}} = I \).

We would like to have a system description in terms of a vector of uncorrelated stochastic variables and an orthonormal basis. This can be achieved with the Singular Value Decomposition (SVD) of \( S\theta R^{-\frac{1}{2}} = U\theta \Sigma\theta V\theta^T \) as

\[ r - n = S\theta h = S\theta R^\frac{1}{2} R^{-\frac{1}{2}} h = U\theta \Sigma\theta V\theta^T R^{-\frac{1}{2}} h = U\theta x, \]

where \( U\theta \) is an \( M \)-by-\( L \) matrix. If the range of \( S\theta \) is a space of dimension \( \gamma \), we know from the theory of SVD, see e.g. [6], that the first \( \gamma \) columns of \( U\theta \) span the same space.

Observe that \( R_x = E{(x - x_0)(x - x_0)^T} = \Sigma\theta \), and \( x_0 = E{X|\theta} = \Sigma\theta V\theta^T R^{-\frac{1}{2}} h_0 \) in general is dependent of \( \theta \).

3 Time-of-Flight estimation
The estimation problem is to estimate the parameter \( \theta \) as well as possible by using the observations \( \{r[k] | k = 0, \ldots, M - 1\} \). Assume that \( n \in \mathcal{N}(0, \sigma_n^2 I) \), and that \( h \in \mathcal{N}(h_0, R) \). If perfect knowledge of the system statistics is available, we can derive the Maximum a Posteriori (MAP) estimator of \( \theta \) as

\[ \hat{\theta}(r) = \arg \max_{\theta \in \Theta} \{ L(\theta, x) \}, \]

(4)

where

\[ L(\theta, x) \triangleq -\frac{||r - U\theta x_0||^2}{2\sigma_n^2} - \frac{(x - x_0)^T R_x^{-1} (x - x_0)}{2} \]

is the logarithm of the a posteriori density function [7]. To find the maximum of \( L(\theta, x) \) with respect to \( x \), \( \frac{\partial L(\theta, x)}{\partial x} = 0 \) is solved for a fixed \( \theta \). This yields

\[ \hat{x}(\theta) = (I + \sigma_n^2 \Sigma\theta^{-2})^{-1} (U\theta^T r + \sigma_n^2 \Sigma\theta^{-2} x_0) \]

(6)
Inserting (5) and (6) into (4), we get

$$
\hat{r}(r) = \arg_{r \in \Omega_{0}} \max \{ L(\theta, \hat{x}(\theta)) \} \tag{7}
$$

and the minimization in equation (4) is reduced to a one-dimensional optimization.

Define \( p(\theta) \triangleq y(\theta) + \sigma_{n}^{2} \Sigma_{\theta}^{-1} V^{T} R^{-\frac{1}{2}} h_{0} \), where \( y(\theta) \triangleq U_{0}^{T} r \), which can be interpreted as the output of a filter bank. Then

$$
L'(\theta) = L(\theta, \hat{x}(\theta)) = p^{T}(\theta) \left( I + \sigma_{n}^{2} \Sigma_{\theta}^{-2} \right)^{-1} p(\theta) \tag{8}
$$

where we have omitted additive constants, cf. [7, page 364, Figure 4.74]. Solutions for other a priori density functions of \( x \) can be found in the literature, e.g. [2]. Several observations can now be made about \( L'(\theta) \). If all diagonal elements of \( \Sigma_{\theta} \rightarrow \infty \), i.e. that the only information contained in the model is a space of possible impulse responses, equation (8) becomes \( L'(\theta) = y^{T}(\theta)y(\theta) \). The estimation problem has then been reduced to determining when the maximum waveform energy in this space arrived. If all diagonal elements of \( \Sigma_{\theta} \rightarrow 0 \), there is no uncertainty in the received waveform and \( L'(\theta) = r^{T}(U_{0} x_{0}) \), resulting in the matched filter receiver.

In an implementation of an estimator according to (7) and (8), the function \( y(\theta) \) has to be calculated for all \( \theta \in \Omega_{0} \). This can be done by \( (u_{i} * r)[k] = \{ y((k - M + 1)T - \theta_{0}) \}_{i} \), where \( * \) denotes time-discrete convolution, and the filter \( u_{i} \) is chosen as the reversed \( i \)th column of \( U_{0} \). Here \( M \) must be sufficiently large for \( s(kT - \theta_{0}) \) to have negligible energy outside \( k = \{ 0, 1, \ldots, M - 1 \} \), compared to the variance of the measurement noise \( n \). Given that \( s(\cdot) \) is bandlimited, \( y(\theta) \) is also bandlimited and can be obtained from \( y((k - M + 1)T - \theta_{0}) \) by interpolation. Furthermore, with the above choice of \( M \), the dependence of \( x_{0} \) and \( \Sigma_{\theta} \), evaluated for \( \theta = (k - M + 1)T - \theta_{0} \in \Omega_{\theta} \), on \( k \) will be weak. This substantially simplifies an implementation, enabling the receiver to be of a correlator - matched filter type.

The ordered singular values \( \lambda_{i} \), i.e. the diagonal elements of the matrix \( \Sigma_{\theta} \), are monotonically decreasing. If \( s(\cdot) \) is narrow band many of the singular values will be close to zero, and the major part of the expected energy of \( h \) will be contained in the first singular values.

In the absence of perfect a priori knowledge of the system statistics an approximation must be made. Here we assume that the choice of signal basis is given, but that we have no knowledge of the singular values. (A discussion of the sensitivity for the choice of signal base is not within the scope of the present paper.) We propose to approximate the singular values as

$$
\lambda_{i} := \begin{cases} C, & i = 1, 2, \ldots, p \\ 0, & \text{otherwise}, \end{cases} \tag{9}
$$

thus, to consider only the \( p \) most important directions of variation in \( h \). This is then essentially a choice of the model order \( p \). This estimator corresponds to the Maximum Likelihood estimator for the case when the true system is of order \( p \), i.e. when \( L = p \).

### 4 Estimation example

To illustrate the estimation method presented in Section 3, and to compare the performance of the MAP-estimator with the approximated structure suggested in (9), a series of simulations was performed. The transmitted signal was \( s(t) = a(t) \cos(2\pi f_{0} t) \). The envelope \( a(t) \) was bandlimited with Fourier transform \( A(f) = 0.5 + 0.5 \cos \frac{\pi f}{W} \), when \( |f| \leq W \), and zero otherwise. The measurement noise \( n \in N(0, \sigma_{n}^{2} I) \) and the Signal to Noise Ratio (SNR) was defined as \( \text{SNR} \triangleq x^{T} x/\sigma_{n}^{2} \), making it a function of a stochastic variable, and thus stochastic itself. For purposes of comparison, the SNR has been normalized to the given values in these simulations.

Now, the estimators were simulated for \( f_{0} = 1/10 \), \( M = 400 \), \( \theta_{0} = \theta = 200 \), and with a sampling rate \( f_{s} = \frac{1}{2} = 1 \). The system statistics was \( h \in N(0, R) \), with the covariance as \( R = \text{diag}(1, 1/2, \ldots, 1/L) \), and \( L \), the length of \( h \), equalled 10. The performances of the estimators were evaluated in terms of the sample mean squared error, calculated for \( N = 5000 \) realizations. The results are presented in Figure 2.
Figure 2: The graphs has the sample mean squared error, $1/N\sum_{n=1}^{N}(\theta - \hat{\theta}(r_n))^2$, on the y-axis, and truncation order, $p$, of the approximated estimator, suggested by (7), (8) and (9), on the x-axis. There are performance plots of the approximated estimator and of the MAP estimator (appearing as horizontal lines) for various SNRs and bandwidths.

5 Discussion and Conclusions

Modelling uncertainties in a received waveform as a Gaussian variation of the transmission system impulse response seems to be a tractable method for TOF estimation, especially when the SNR is high.

Using a low model order $p$, e.g. assuming a known signal shape, can give large systematic errors. Increasing the model order decreases the systematic errors but amplifies the effects of the additive noise, which is illustrated by the estimation example in section 4.

The optimum choice of model order depends on the SNR and on the bandwidth of the transmitted waveform, but the optimum is most often flat.

The performances of these algorithms will be evaluated using an ultrasonic A-scan pulse-echo measurement system.

References


