CROSSTALK-RESILIANT CODING FOR HIGH DENSITY DIGITAL RECORDING

M. Z. Ahmed

Ph.D. June, 2003

CROSSTALK-RESILIANT CODING FOR HIGH DENSITY DIGITAL RECORDING

by

MOHAMMED ZAKI AHMED

A thesis submitted to the University of Plymouth
in partial fulfillment for the degree of

DOCTOR OF PHILOSOPHY

Department of Communication and Electronic Engineering
Faculty of Technology
June, 2003

Crosstalk-Resiliant Coding for High Density Digital Recording

by

Mohammed Zaki Ahmed

Abstract

Increasing the track density in magnetic systems is very difficult due to inter-track interference (ITI) caused by the magnetic field of adjacent tracks. This work presents a two-track partial response class 4 magnetic channel with linear and symmetrical ITI; and explores modulation codes, signal processing methods and error correction codes in order to mitigate the effects of ITI.

Recording codes were investigated, and a new class of two-dimensional run-length limited recording codes is described. The new class of codes controls the type of ITI and has been found to be about 10% more resilient to ITI compared to conventional run-length limited codes. A new adaptive trellis has also been described that adaptively solves for the effect of ITI. This has been found to give gains up to 5dB in signal to noise ratio (SNR) at 40% ITI. It was also found that the new class of codes were about 10% more resilient to ITI compared to conventional recording codes when decoded with the new trellis.

Error correction coding methods were applied, and the use of Low Density Parity Check (LDPC) codes was investigated. It was found that at high SNR, conventional codes could perform as well as the new modulation codes in a combined modulation and error correction coding scheme. Results suggest that high rate LDPC codes can mitigate the effect of ITI, however the decoders have convergence problems beyond 30% ITI.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I would like to express my gratitude to everyone who has helped, and supported me in my research and funding, in particular;

My supervisors, Dr Terry Donnelly, Dr. Paul Davey and Professor Warwick Clegg for their constant support, encouragement, guidance and friendship.

Dr. Yoshitake Kurihara, Dr. Marcel Ambroze and Dr. Hazel Shute for the valuable discussions which have significantly broadened my knowledge.

And, not least of all, my parents and sisters without whose unlimited support and patience I would not have reached this stage.

Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award.

This study was financed partly with the aid of the Overseas Research Studentship Award and the Dean of Technology Award.

A programme of advanced study was undertaken, which included the extensive reading of literature relevant to the research project and attendance of international conferences on coding and magnetic recording.

The author has presented papers at The 2001 IEEE Joint MMM Intermag Conference at San Antonio, USA, The 2002 IEEE Intermag Conference at Amsterdam, The Netherlands, and The 2003 IEEE Intermag Conference at Boston, USA.

In addition the author attended The Second Symposium on Turbo Codes in September 2000 in Brest, France and the IEEE International Symposium on Information Theory in June 2002 in Lausanne, Switzerland.

Mohammed Zaki Ahmed : Mohammed Zaki Ahmed

Date : 04th July 2003

This thesis is dedicated to Uncle Justin and Terry who taught me how to learn.

Glossary

List of Acronyms

AEQ Adaptive Equaliser

APR A Priori Probabilities

APP A Posteriori Probabilities

AWGN Additive White Gaussian Noise

BCJR Bahl Cocke Jelenik and Raviv (Algorithm)

BER Bit Error Rate

BN Belief Network

CITI Constructive Inter-Track Interference

DSP Digital Signal Processing

ECC Error Correction Code

FIR Finite Impulse Response

ISI Inter-Symbol Interference

ITI Inter-Track Interference

KTS Kirkman Triple System

LDPC Low-Density Parity Check

LMS Least-Mean Squared

LUT Look Up Table

MAP Maximum A Posteriori

ML Maximum Likelihood

MLSD Maximum Likelihood Sequence Detection

NITI No Inter-Track Interference

NRZI Non-Return to Zero Inverse

PDF Probability Density Function

PR Partial Response

PR4 Partial Response Class 4

PRML Partial Response Channel with Maximum Likelihood Decoding

RC Recording Code

RLL Run Length Limited

SIHO Soft In Hard Out

SISO Soft In Soft Out

SNR Signal to Noise Ratio

SPA Sum-Product (Algorithm)

STM State Transition Matrix

STD State Transition Diagram

TC Turbo Code

List of Symbols

α Inter-Track Interference

$\alpha(a, b)$ BCJR Forward State Probability, a specifies the State and b specifies the Sample

$\beta(a, b)$ BCJR Backward State Probability, a specifies the State and b specifies the Sample

D Unit Delay operator

$\mathbf{H}$ Parity Check Matrix

PW_{50} Pulse Width at half maximum

1 Introduction

The current trend in magnetic recording can be summarised as "smaller, faster, denser, cheaper". The major stages in the signal processing of magnetic storage technology and multi-track recording can be summarised (from IBM[IBM02a] and other sources) as follows*.

1.1 Magnetic Recording & Two Dimensional Recording

1952 Non-Return to Zero (NRZI) coding was created for tape drives. This coding technique simplified the way data was recorded and read, and it permitted bit densities to exceed the density of magnetic transitions.(IBM)

1956 IBM RAMAC – The first computer disk storage system. The IBM 305 RAMAC (Random Access Method of Accounting and Control) could store five million characters (five megabytes) of data on 50 disks, each 24 inches in diameter.

1967 Trellis decoding was introduced by Forney[Jr.67] as a conceptual means to explain the Viterbi algorithm.

1970 Kobayashi and Tang consider application of PRML techniques to magnetic recording channels[KT70]

1973 First disk drive to use low-mass heads, lubricated disks, and sealed assembly, which came to be known as Winchester drive technology. The IBM 3340 Winchester

*Important events in the history of magnetic recording are also presented

drive had two spindles with a storage capacity of 30 million characters each, hence 30-30 or the "Winchester"[†].(IBM) Maximum Likelihood Sequence Detection was introduced[Jr.73].

1974 The BCJR algorithm was introduced[BCJR74].

1978 Trellises were defined (more rigorously) by Massey[Mal].

1979 First disk drive to feature thin-film inductive heads and a run-length-limited (RLL) coding scheme. The "2-7" RLL code permitted higher performance. (IBM 3370)

1984 Brian Marcus and Paul H. Siegel consider constrained codes for PRML[MS84].

1985 Paul Siegel considers recording codes for digital magnetic storage[Sie85].

1986 First disk drive to use the more efficient "1-7" run-length-limited (RLL) code, then widely used in the industry. (IBM 9332)

1988 Modeling offtrack/inter-track interference was proposed by Abbott et al[ACT88].

1989 Partial Response Maximum Likelihood (PRML) channel used in the IBM 0681 disk drive, a One-gigabit-per-square-inch, 3.5 inch device.

1990 L.C. Barbosa considers simultaneous detection of readback signals from interfering magnetic recording tracks using array heads[Bar90].

1991 Vea et al[VM91a][VM91b] publish work on detection and modeling in the presence of crosstalk.

1992 M.W. Marcellin and H.J. Weber introduce Two-Dimensional Recording codes[MW92], constructing run-length constraints in two dimensions. Robert E. Swanson and Jack K. Wolf also introduce a new class of Two-Dimensional recording codes[SW92] that had an additional constraint across tracks.

[†]The first American designed smokeless powder sporting cartridge apparantly very popular with deer hunters[30-02]

1994 Soljanin[SG94] proposes sliding block codes for two-track two head magnetic recording, and Lee[LM94] proposes a code construction that does not permit certain adjacent transitions as methods of combating Inter-Track Interference.

1997 Bane[BO97] propose cyclic two dimensional codes that do not permit certain types of adjacent transitions similar to Lee's constraints.

1998 Davey et al[DDMD98] publish a multi-track code with immunity to ITI.

2001 Ahmed et al[ADDC01] propose a new class of two-dimensional codes that permit only certain types of transitions. These new codes improve on the capacity (and the achievable code rate) of all previous crosstalk resilient codes for PR4 two-track recording systems.

It is amazing that Kobayashi and Tang[KT70] proposed the use of Partial Response Maximum Likelihood (PRML) channels in 1970, and the technique was implemented 19 years later in Hard Disk Drives.

From 1956 until 1990, peak detection was the method of data detection on a magnetic drive. The major shift in the signal processing technology was the adaption of PRML techniques[CGK+91]. This has led to an increase in linear densities, as bits could be packed closer together and the interaction between adjacent bits was no longer unwanted. Advances in the recording head has led to an increase in track densities, as narrower tracks could be written.

From 1994 to 2002, track and linear densities have increased. Track densities have increased from about 4×10^3 tracks per inch (1600 tracks per cm) to about 9×10^4 tracks per inch or 35400 tracks per cm (an increase in $22\frac{1}{2}$ times). Linear density have increased from 70kbits per inch (approximately 28000 bits per cm) to 600kbits per inch or 236000 bits per cm (an increase of $\approx 8\frac{1}{2}$ times)[IBM02b].

Various authors, notably Wood[Woo00] from IBM, have suggested that storage at 1 Terabit per square inch is feasible. It has been highlighted by Wood[Woo00] that

side-reading of information from adjacent tracks and interaction between adjacent tracks will be a major source of noise in future recording systems. These concerns were also previously highlighted by Barbosa[Bar90].

As track densities are increased, the interaction between adjacent tracks results in a performance degradation. One solution is to treat adjacent tracks as a single unit rather than separate units. This assumes that data can be written and read back simultaneously from multiple tracks. Previous work done at the University of Plymouth investigated the use of multi-track channel coding methods to increase recording densities[Dav94]. The work presented in this thesis extends this; investigating channel codes, signal processing and error correction codes. A two-track model is assumed and new methods and techniques have been found to mitigate the effect of inter-track interference (ITI).

1.2 Thesis Outline

A background to recording codes, signal processing and error correction codes is given in chapter 2. The chapter discusses magnetic recording channels and the model used by the author. The chapter proceeds with basic magnetic recording, PRML Channels and PR equalisation. A brief background to modulation codes and modern error correction codes is also provided.

Very high rate error correction codes are used in chapter 3 in order to improve the performance of conventional codes to match the performance of the new class of codes introduced in chapter 4. Low Density Parity Check (LDPC) codes were investigated at very high rates (rates greater than 0.9) and some insight into the decoding issues of LDPC's has been gained. BCJR decoding and The Sum-Product Algorithm are described and the scheme implemented by the author is discussed.

Two-Dimensional Modulation codes are discussed in chapter 4. ITI is modelled and a new class of codes is introduced and their performance compared to conventional methods. A special case for these new codes that is free from the detrimental effects of ITI is also described, and bounds for the new class of codes are discussed. The new codes are compared with conventional codes.

New signal processing methods for detection of two tracks simultaneously are presented in chapter 5. The new detector operates in the presence of an unknown amount of ITI. The design of the new decoding technique is presented, and a new detection scheme is discussed. The performance of the new class of codes is measured and compared with the performance of conventional codes.

The authors conclusions are presented in chapter 6, with areas of future work.

2 Background to Investigation

2.1 Introduction to Digital Magnetic Recording

Digital Magnetic Recording deals with the storage and retrieval of binary data on magnetic media. It can be modeled as a communications system transmitting data in time from "now" to "later". In digital recording, the data are modulated into a write current, which flows through a write head. The write head induces an almost saturated* magnetic pattern on the media, that moves below the head. The magnetic medium is sufficiently saturated so that non-linearities from the magnetisation curve are not introduced†.

The medium can be saturated in three ways, defined by the direction of magnetisation relative to the direction of motion of the medium.

1. Longitudinal – direction of magnetisation is in the plane of the media surface and parallel to the motion of the media.

2. Perpendicular – magnetisation is normal to the media surface.

3. Transverse – magnetisation is in the plane of the media surface but normal to the direction of motion of the media.

In this thesis, a Longitudinal recording channel is studied as it is currently the most popular way data is stored[CGK+91].

The retrieval of data from the media is done using the read head. The read signal reflects changes of magnetisation on the media. Signal processing techniques are used to

*saturation causes fringing fields around the head and crosstalk in adjacent tracks

†When the magnetic flux density B of a ferromagnetic medium varies, the permeability of the medium also varies; and the relationship between the magnetic flux density B and the (applied) magnetic field strength H is not a straight line but a curve[Map01].

retrieve a clock and the recorded data.

2.1.1 Data Writing

As the magnetic medium has only two saturated states, the record waveform is essentially bipolar. To provide the best Signal to Noise Ratio (SNR), the write current is a bit less than that required to saturate the media[Wat88]. This is because saturation causes fringing fields around the head and crosstalk in adjacent tracks. A preamplifier is sometimes used to correct for losses in the record head material at high frequencies[Map01].

Digital data $\{1, 0\}$ are recorded such that the symbol 1 prescribes a positive amplitude of write current, and the symbol 0 a negative amplitude. This is known as non-return-to zero (NRZ). If the input NRZ sequence is $x_n, n = 0, 1, \ldots$ then the output sequence from the media $y_n, n = 1, 2, \ldots$ is given by equation (2.1)

$$y_n = x_n - x_{n-1} \tag{2.1}$$

Reconstruction of x_n using (2.1) suggests an error in y_n will propagate, and to circumvent that the user data (u_n) are precoded using a $\frac{1}{(1-D)_{mod2}}$ precoder[KT70, MS84]. The effect of precoding the input data sequence leads to an NRZ-Immediate (NRZI) sequence. Using NRZI and this channel model (a $1 - D$ Channel) the input data symbol 1 is recorded as a *transition*[‡][Sie85]. The symbol 0 is recorded as *no transition*, as shown in table 2.1.

Input Data u_n	0	0	1	0	1	1	$0\cdots$
Precoded Sequence $x_n = u_n \oplus x_{n-1}$	0	0	1	1	0	1	$1\cdots$
y_n		0	+1	0	−1	+1	$0\cdots$

Table 2.1: Basic Writing

[‡]changes in polarity of magnetisation on the media, can be positive or negative

2.1.2 Data Readback

The readback waveform is used to derive the data and the data clock. The location of transitions on the medium is the location of the 1's in the data sequence. Sufficient 1's are required in the data sequence so that the data clock can stay locked. The recording system suffers from many sources of noise. In this thesis the following noise and channel effects are considered

Additive White Gaussian noise (AWGN): This emanates from the readback head and electronics.

Inter-Symbol Interference (ISI): This is caused by neighbouring transitions along the track. It leads to peak shifting and waveform drooping in the readback waveform.

Inter-Track Interference (ITI): Caused by the magnetic pattern on adjacent tracks. It introduces (adjacent) data-dependent drop and amplification of the readback waveform.

Other sources of noise which are not addressed in this work include

Dropouts: Dropouts are caused by heads becoming clogged resulting in the loss of the channel. They are a source of error in magnetic tape systems, and are corrected using an error correction code designed for burst errors.

Jitter: This is caused by timing instability in the readback process.

Media noise: This is caused by granularity in the media, leading to a zigzag transition. As the track width is reduced, media noise increases. It is data dependent and is highlighted as an area of future work by the author[§].

[§]For very high-density recording, Wood[Woo00] suggests that patterned media could reduce the effect of random granularity (and hence the effect of media noise).

Non-Linear Transition Shifts (NLTS): This is caused by the mutual attraction between transitions at high-density recording. It occurs during the writing of two transitions, and NLTS is the shift of the latter transition towards the former[Tar95].

These sources are discussed in Bertram's book[Ber94] and Watkinson's book[Wat88]. ITI is considered as the dominant source of noise of future very high-density recording[Woo00].

If readback transitions are too close to one another, undesirable ISI occurs, and if transitions are too far apart, clock recovery may lose synchronisation (which may be catastrophic). Control of the distance between transitions can be achieved using modulation codes, discussed later in this chapter (page 25).

2.2 Partial Response Maximum Likelihood (PRML) Recording

Partial Response Maximum Likelihood (PRML) systems have been used in commercial computer hard disk drives (HDD) since the early 90's[CGK+91]. The use of Partial Response in magnetic recording systems was first proposed by Kobayashi and Tang in 1970[KT70], whereas a method for Maximum Likelihood Sequence Detection was proposed by Kobayashi in 1971[Kob71] and Forney in 1973[Jr.73].

Partial Response (PR) signalling (also known as correlative level coding) is essentially a technique of allowing a controlled amount of Inter-Symbol Interference (ISI) to occur. The classes of PR are defined by Kretzmer[Kre66], and Class 4 has been used extensively in magnetic systems. Class 1 is a system that has all unity coefficients, while Class 2 has only linearly tapered distributions of coefficient values. It is unclear how the remaining class number relate to the coefficient distributions, however a Class described by $1 - D^2$ is labelled as Class 4 in[Kre66]. Class 4 is studied in this thesis, as in the region of interest it provides very good performance[TP87, DUH86].

Applying a PR system to a binary sequence results in the output from the system no longer being binary, but would have more levels depending on the PR polynomial[TP87]. Precoding of the data is necessary to prevent error propagation due to the memory the PR polynomial introduces[KT70]. The precoder for a PR polynomial is given by

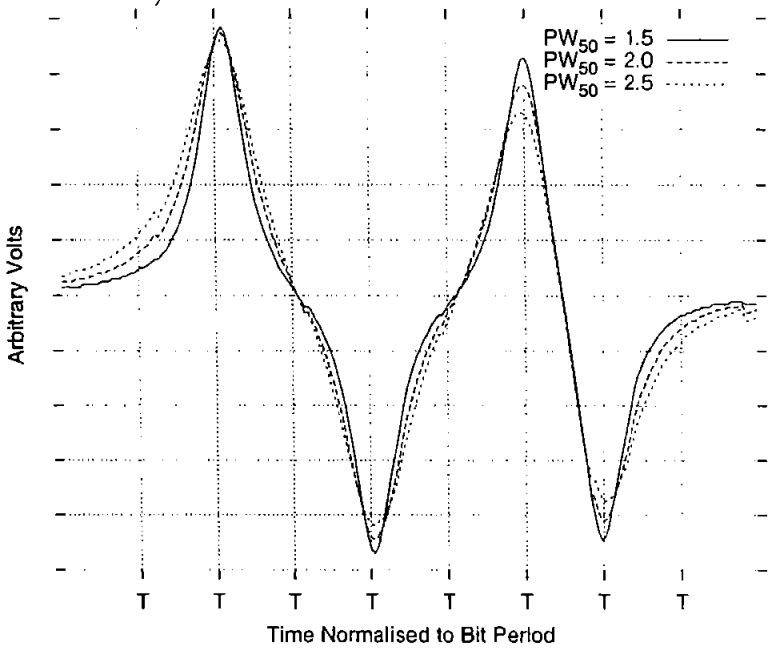
$$\left[\frac{1}{\text{PR Polynomial}}\right]_{\text{mod } 2} = \frac{1}{1 \oplus D^2} \quad \text{for PR4} \tag{2.2}$$

Table 2.2 shows different stages in a typical PR4 magnetic recording system in a top-down manner.

User Data	0	1	1	1	1	1	0	1	0
PR4 Precoded Data (using $\frac{1}{1 \oplus D^2}$)	0	1	1	0	0	1	0	0	0
Write Current	−1	+1	+1	−1	−1	+1	−1	−1	−1

Magnetisation Pattern: → | ← | → | ← | →

Readback Samples (Ideal & Noiseless): 0 +1 0 −1 0 +1 −1 0 0



$(1 + D)$ Equalisation	0	+1	+1	−1	−1	+1	0	−1	0
Detected Data (same as User Data)	0	1	1	1	1	1	0	1	0

The arrowhead corresponds to Magnetic North. The vertical line between two arrows is referred to as the ***transition****. Transitions occur where the write current changes direction.*

Table 2.2: PR4 Longitudinal Recording Data at various stages

A positive transition occurs when the arrowheads point toward each other at a transition, and a negative transition occurs when the arrowheads point away from each other. User data are initially precoded using a $\frac{1}{1 \oplus D^2}$ transfer function. The precoded data are written using an NRZ write current resulting in the magnetisation pattern shown above. The ideal readback samples are shown, however the readback waveform modelled using

a Lorentzian wave is what is read from the channel. The Lorentzian is equalised using a $1 + D$ equaliser (described later) resulting in samples whose magnitude reconstructs the user data.

Additional PR polynomials can be used so that the overall PR response is as close as it can be to the magnetic recording channel response. This is to reduce the amount of equalisation required to convert the pulses from the real channel to the Partial Response channel. Excessive equalisation can lead to undesirable noise amplification.

For the pattern shown in table 2.2, assuming no transitions exist before and beyond the transitions shown, the waveform shown in table 2.2 would be expected. The overall Channel Response is

$$(1 - D)(1 + D) = 1 - D^2 \tag{2.3}$$

The read-back voltage from an isolated transition can be approximated by a Lorentzian function[Ber94] (equation 2.4).

$$v(t) = \frac{1}{1 + \left(\frac{2t}{PW_{50}/T}\right)^2} \tag{2.4}$$

PW_{50}/T is the pulse width at half maximum, normalised to the bit period. It is a measure of how close the transitions are in relation to the minimum length between transitions.

Table 2.2 shows the effect of bit crowding, the reduction in voltage when there are adjacent transitions in the read-back sequence. In addition to the voltage drop, there is a shift in the local maxima away from the sample instance. These effects get worse as the PW_{50} increases.

In a PR system, a controlled amount of ISI is added leading to the waveforms shown in figure 2.1. This is the partial response equalisation. The design method used by the author is described in more detail on page 17.

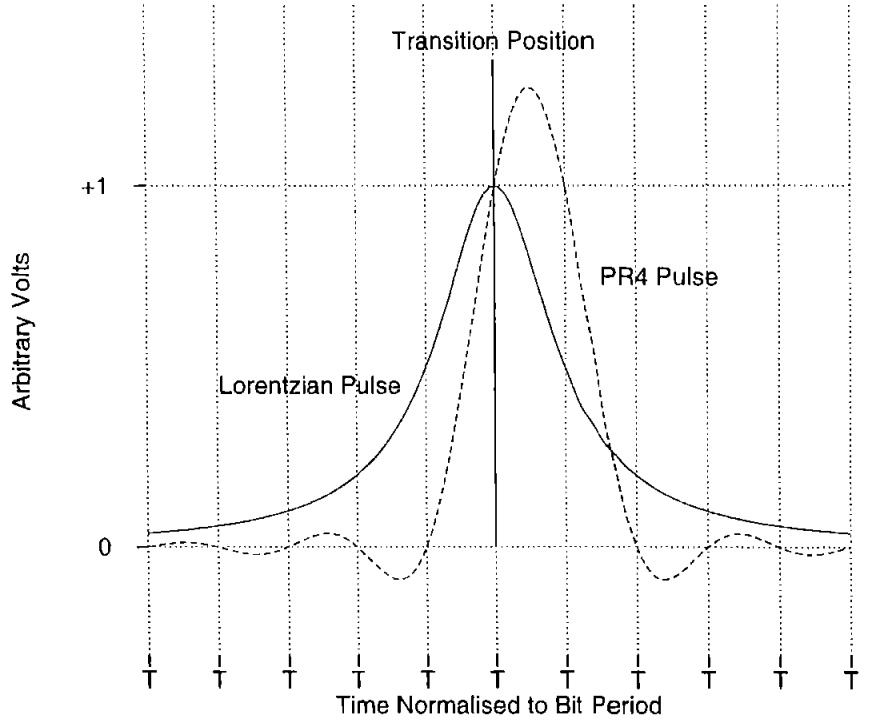


Figure 2.1: Ideal PR4 Read-back Waveform compared with a Lorentzian with $PW_{50} = 2.0$

Figure 2.1 shows, the position where an isolated transition occurs, and the resulting Lorentzian and PR4 waveforms.

PR4 equalisation leads to zero read-back voltage at the all sampling instances except at the location of the transition and the next sample instance from the transition. At the latter two sampling instances, the PR4 Pulse has amplitude +1. The PR4 Pulse does not suffer from ISI at the sampling points (labeled 'T') as shown in figure 2.1, and can be described by equation 2.5

$$s(t) = \frac{\sin\left(\frac{\pi t}{T}\right)}{\frac{\pi t}{T}} + \frac{\sin\left(\frac{\pi(t-T)}{T}\right)}{\frac{\pi(t-T)}{T}} \quad \text{where } T \text{ is the sampling interval} \tag{2.5}$$

When there are only 2 adjacent transitions the resulting waveform is referred to as a "dibit". A Lorentzian dibit can be expressed as

$$d(t) = \frac{1}{1+\left(\frac{2t}{PW_{50}}\right)^2} - \frac{1}{1+\left(\frac{2(t-1)}{PW_{50}}\right)^2} \tag{2.6}$$

The frequency response of equation 2.6 is

$$D(\omega) = (1-e^{-j\omega T})\frac{PW_{50}}{2}\pi e^{-(\omega\frac{PW_{50}}{2})} \tag{2.7}$$

where ω is the frequency term and T is the sampling period.

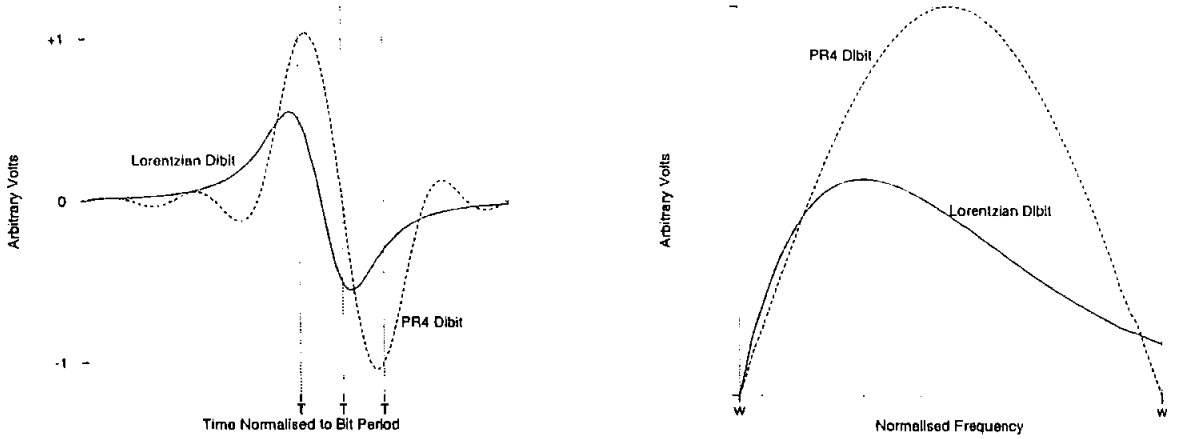
Replacing the 'ones' in the precoded sequence of table 2.2 with the dibit waveform ($d(t)$) and the 'zeros' with $-d(t)$ leads to the read-back waveform (table 2.2); and is also a common way of simulating part of the PRML channel.

The dibit resulting from using equation 2.5 (the PR4 dibit) has a frequency response

$$S(\omega) = 2\sin(\omega T) \tag{2.8}$$

where ω is the frequency term and T is the sampling period. The spectra and waveforms of both these dibits are shown in figure 2.2, and the derivation of the equations is in the

appendix.



Dibits in Time Domain Dibits in Frequency Domain

Figure 2.2: Lorentzian (with $PW_{50} = 2.0$) and PR4 dibit response

2.2.1 Partial Response (PR) Equalisation

A matched filter is a filter that maximises the Signal to Noise Ratio (SNR) at the output of the filter. A signal $s_{mf}(t)$ confined to an interval $0 \leq t \leq T$ has a matched filter impulse response $h(t) = s_{mf}(T - t)$[Pro95]. For a Lorentzian dibit input waveform, the output of a matched filter would be the autocorrelation of the Lorentzian dibit. Figure 2.3 shows the matched filter input and output sampled waveforms for a dibit with $PW_{50} = 2.0$

Designing a filter that achieves this is possible, however an optimum detector (a Viterbi detector) that uses the output of the matched filter has exponential complexity as the PW_{50}/T increases[Moi00]. A partial response filter/equaliser is a sub-optimal scheme normally used as the structure of the MLSD detector for a partial response system is not as complicated as the matched filter system.

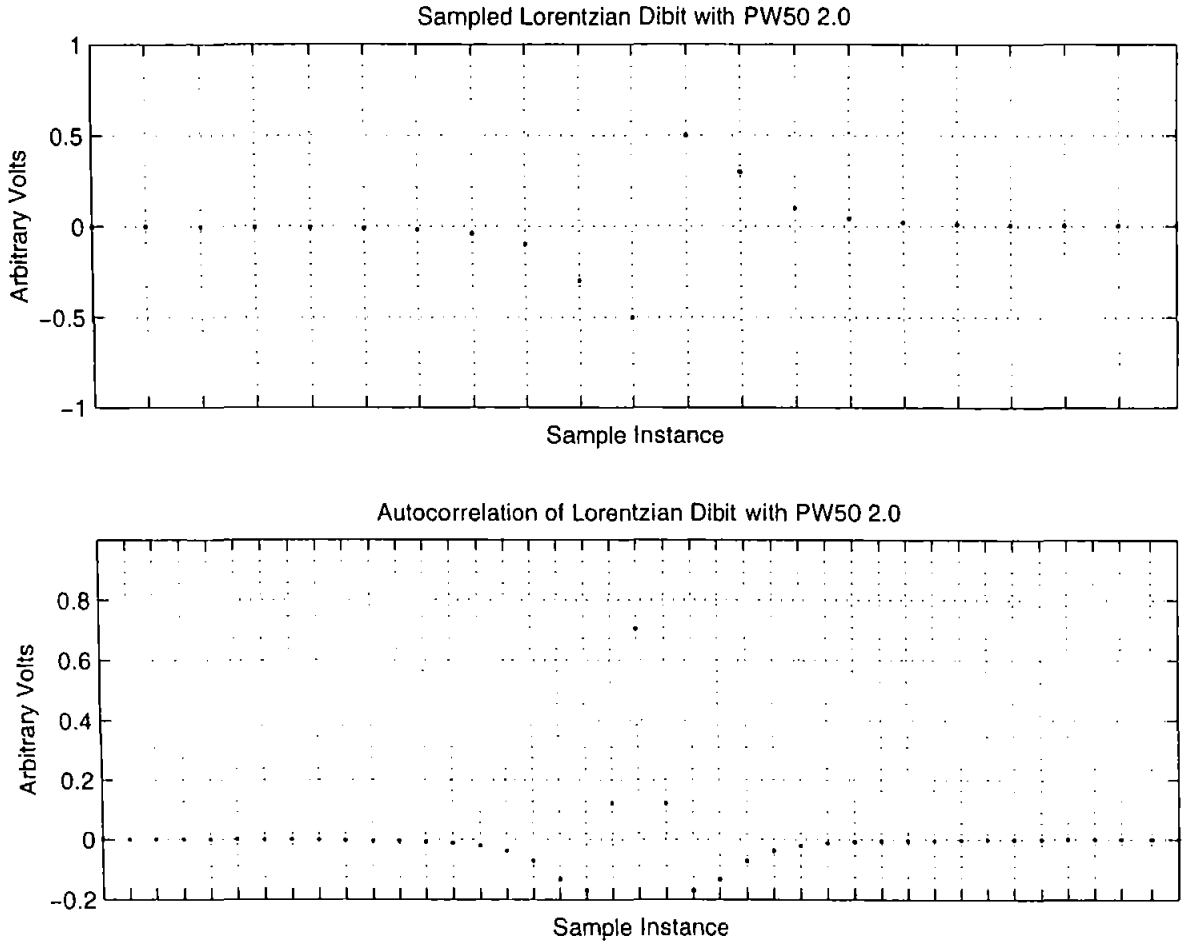
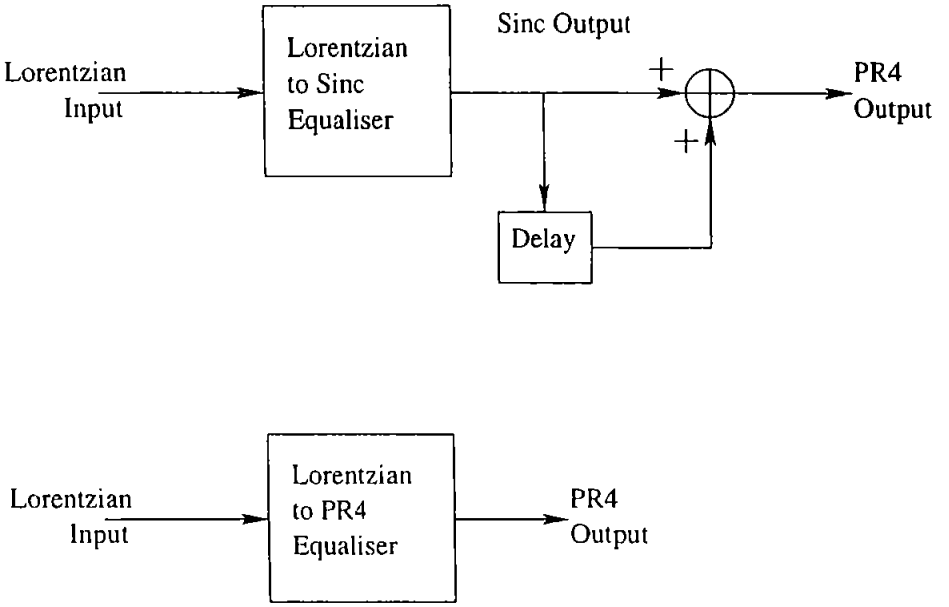


Figure 2.3: Matched Filter input (the upper figure) and output (the lower figure) for $PW_{50} = 2.0$

PR4 equalisation can be viewed as a $1 + D$ operation on an ideal sinc $\left(\frac{\sin(x)}{x}\right)$ pulse, which was the model used in the software simulations described in figure 2.4. The effect of equalising the read-back Lorentzian wave to a PR response is to eliminate the ISI introduced by the Lorentzian function. The equaliser can be implemented using various methods, a combination of analogue and digital schemes being used in commercial products[CGK+91]. A more theoretical description (used in the software simulation, and shown in figure 2.4) can be found in Watkinsons' book[Wat88].



Both schemes were implemented by the author

Figure 2.4: PR4 Equalisation Implementation, adapted from Watkinson

Two methods of equalisation are implemented in the software written by the author. The first equalises the Lorentzian pulses to sinc pulses, then performs a $1 + D$ operation on the resulting samples as shown in figure 2.4. The second technique transforms the Lorentzian directly to a PR4 waveform. Both these methods are based on solving a system of simultaneous equations to obtain the filter tap coefficients.

For a 7-Tap equaliser with tap coefficients $\mathbf{h}$

$$\mathbf{h} = \begin{bmatrix} h_0 & h_1 & h_2 & h_3 & h_4 & h_5 & h_6 \end{bmatrix} \tag{2.9}$$

Using equation (2.4) to evaluate $v(t)$ for a given PW_{50}

$$\begin{bmatrix} v(0) & v(1) & v(2) & v(3) & v(4) & v(5) & v(6) \\ v(-1) & v(0) & v(1) & v(2) & v(3) & v(4) & v(5) \\ v(-2) & v(-1) & v(0) & v(1) & v(2) & v(3) & v(4) \\ v(-3) & v(-2) & v(-1) & v(0) & v(1) & v(2) & v(3) \\ v(-4) & v(-3) & v(-2) & v(-1) & v(0) & v(1) & v(2) \\ v(-5) & v(-4) & v(-3) & v(-2) & v(-1) & v(0) & v(1) \\ v(-6) & v(-5) & v(-4) & v(-3) & v(-2) & v(-1) & v(0) \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \end{bmatrix} = \mathbf{i}' \qquad (2.10)$$

$\mathbf{i}$ is defined as

$$\mathbf{i} = \begin{cases} \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} & \text{Lorentzian to Sinc Equaliser} \\ \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} & \text{Lorentzian to PR4 Equaliser} \end{cases} \qquad (2.11)$$

and $\mathbf{i}'$ is the transpose of matrix $\mathbf{i}$.

Solving equation (2.10) for $\mathbf{h}$ for different PW_{50} results in different filter coefficients. The variation of the filter tap coefficients are shown in figure 2.5 and figure 2.6. The equaliser coefficients are symmetrical about Tap 3 for the Lorentzian to Sinc equaliser. This is due to the that fact the equaliser was designed for linear phase, and the output of the equaliser has a group delay of $\frac{N-1}{2}$ samples, which for a 7 Tap filter is 3 Taps.

The Lorentzian to PR4 equaliser has coefficients symmetrical about 'Tap 3.5'. The group delay of this filter is greater than the previous filter due to the $(1 + D)$ additional PR4 response. This has the effect of increasing the overall filter length to $N + 1$, resulting in a group delay of $\frac{N}{2}$.

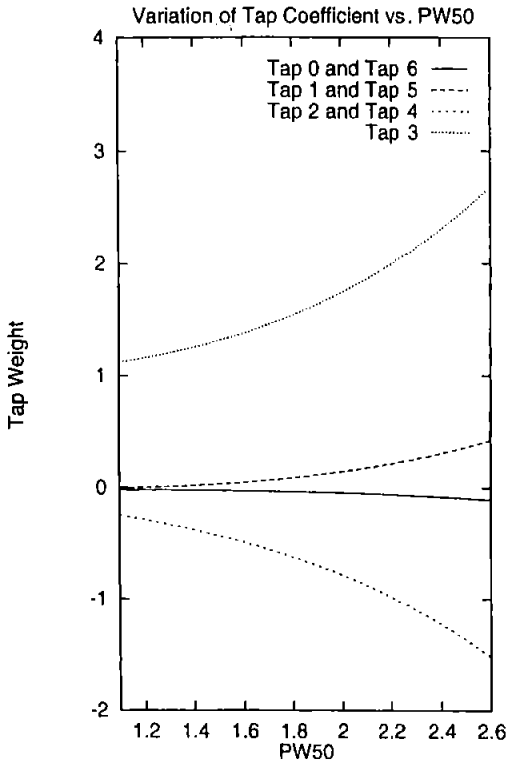


Figure 2.5: Variation of Filter Tap Coefficients with respect to PW_{50} for the Lorentzian to Sinc Equaliser

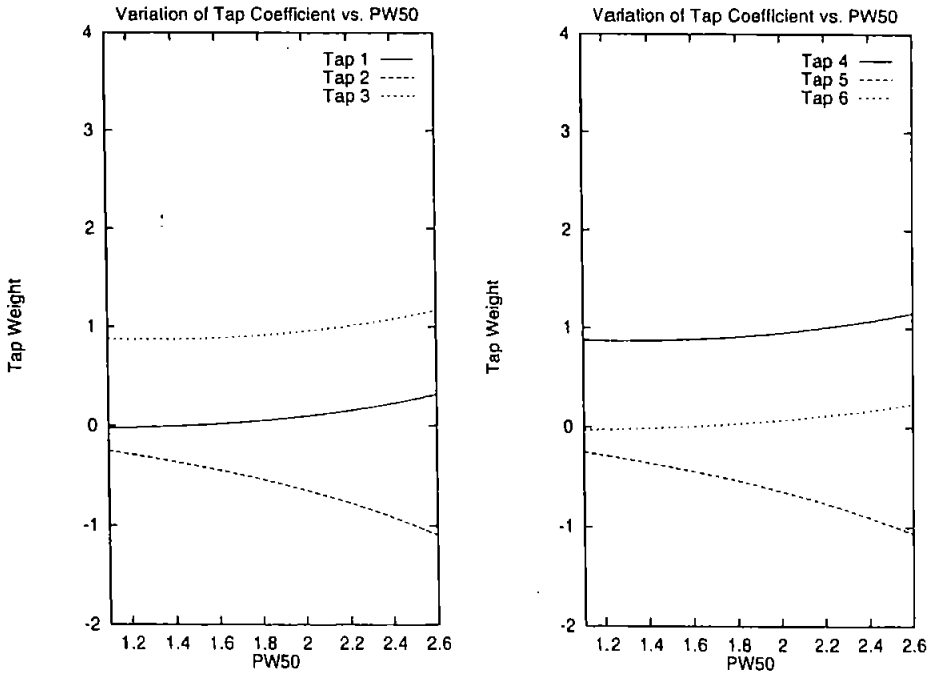


Figure 2.6: Variation of Filter Tap Coefficients with respect to PW_{50} for the Lorentzian to PR4 Equaliser

The frequency response for both these techniques is shown in figure 2.7 and 2.8.

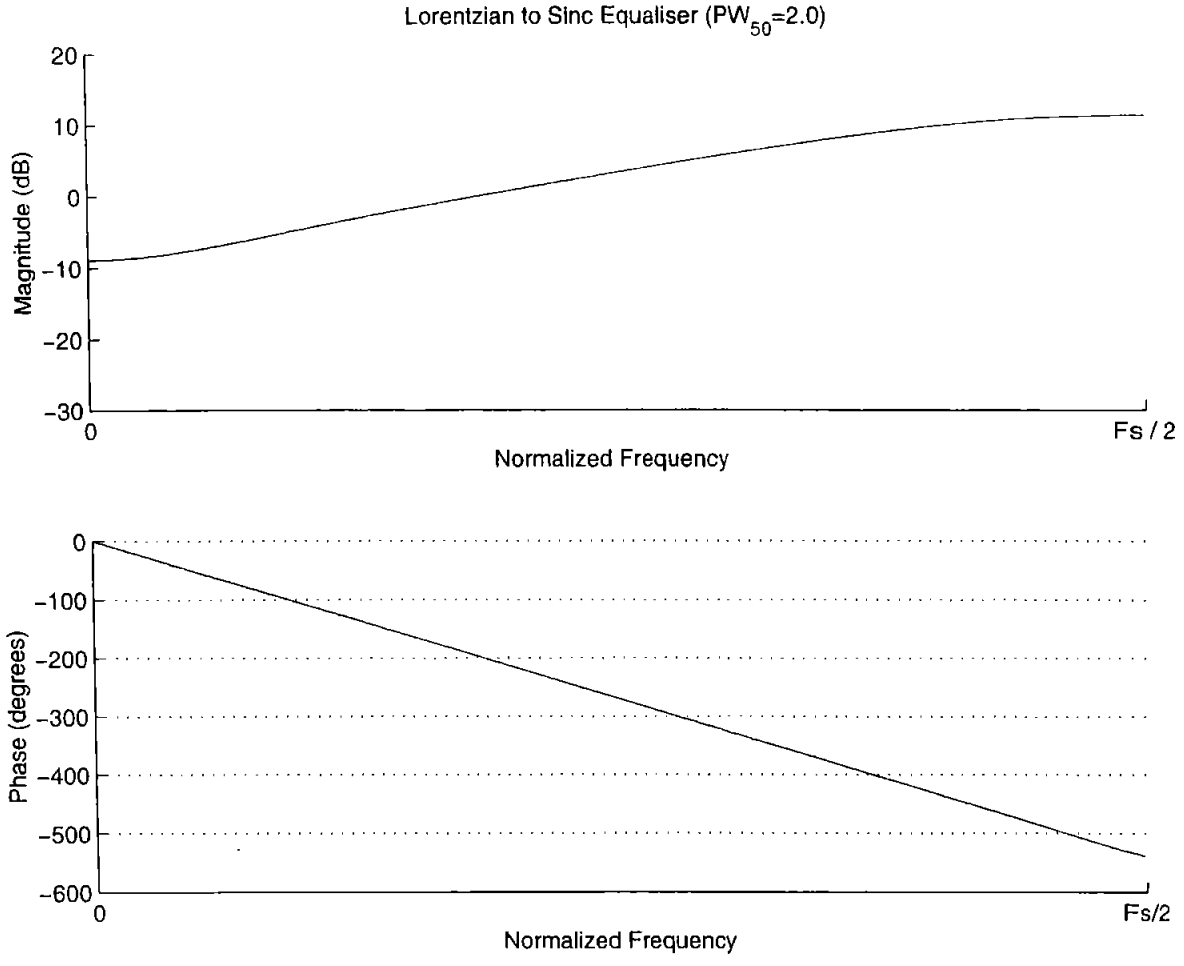


Figure 2.7: Frequency Response of Lorentzian to Sinc Equaliser

It has been found that both equalisers performed relatively equally (figure 2.9).

Beyond a PW_{50}/T of 2.5, the noise amplification introduced by both techniques was similar. This is similar to the results quoted in[TP87].

The overall PR system described here is $(1-D)(1+D)$, commonly called a Class IV PR, or PR4 System. The overall system response is $(1-D^2)$. The $(1-D)$, or dicode PR response is inherent to the magnetic recording system, while the $(1+D)$ modifies the channel response so that it is very close to the actual magnetic channel response for recording densities up to a PW_{50} of approximately 2.5. The closer the target PR response is to the actual response, the easier it is to equalise to the target response.

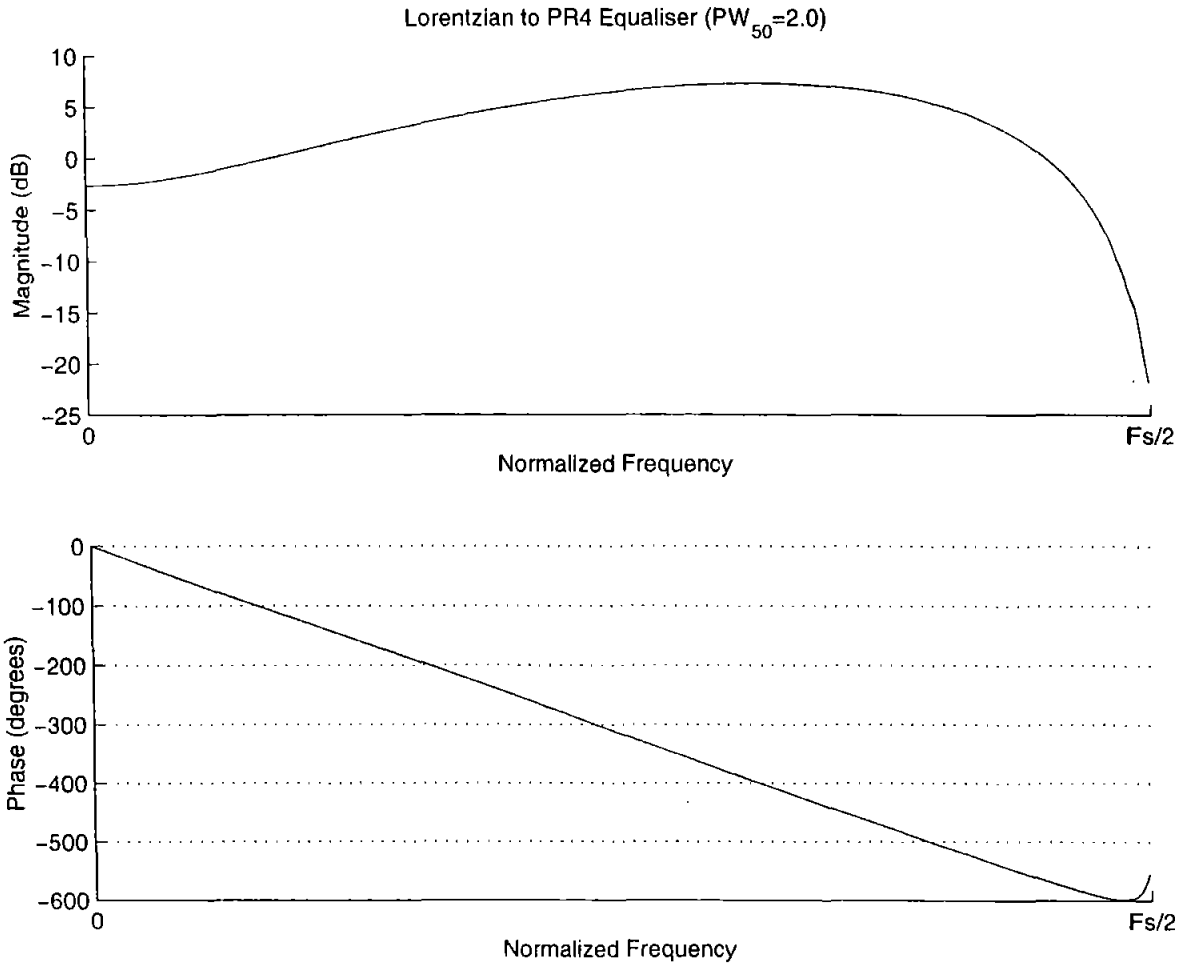


Figure 2.8: Frequency Response of Lorentzian to PR4 Equaliser

The overall idea of converting the magnetic recording channel to a target partial response channel is to introduced a controlled amount of ISI such that the ISI at the sampling points is zero. As PW_{50} increases, the Magnetic Channel response changes, and more equalisation is required to force the read-back wave to a PR4 target. At densities greater than PW_{50} of about 2.5 the Extended PR4 target response is used (with target PR equation $(1-D)(1+D)^2$), and more recently, a combination of PR and Noise Predictive Maximum Likelihood (NPML) targets are used[CEGH00]. Other channel responses that match magnetic channels at high densities include the $(1-D^2)(h_0+h_1D+h_2D^2)$ model, where $h_0 = 2, h_1 = 2, h_2 = 1$ has been found to be a good match[Con98].

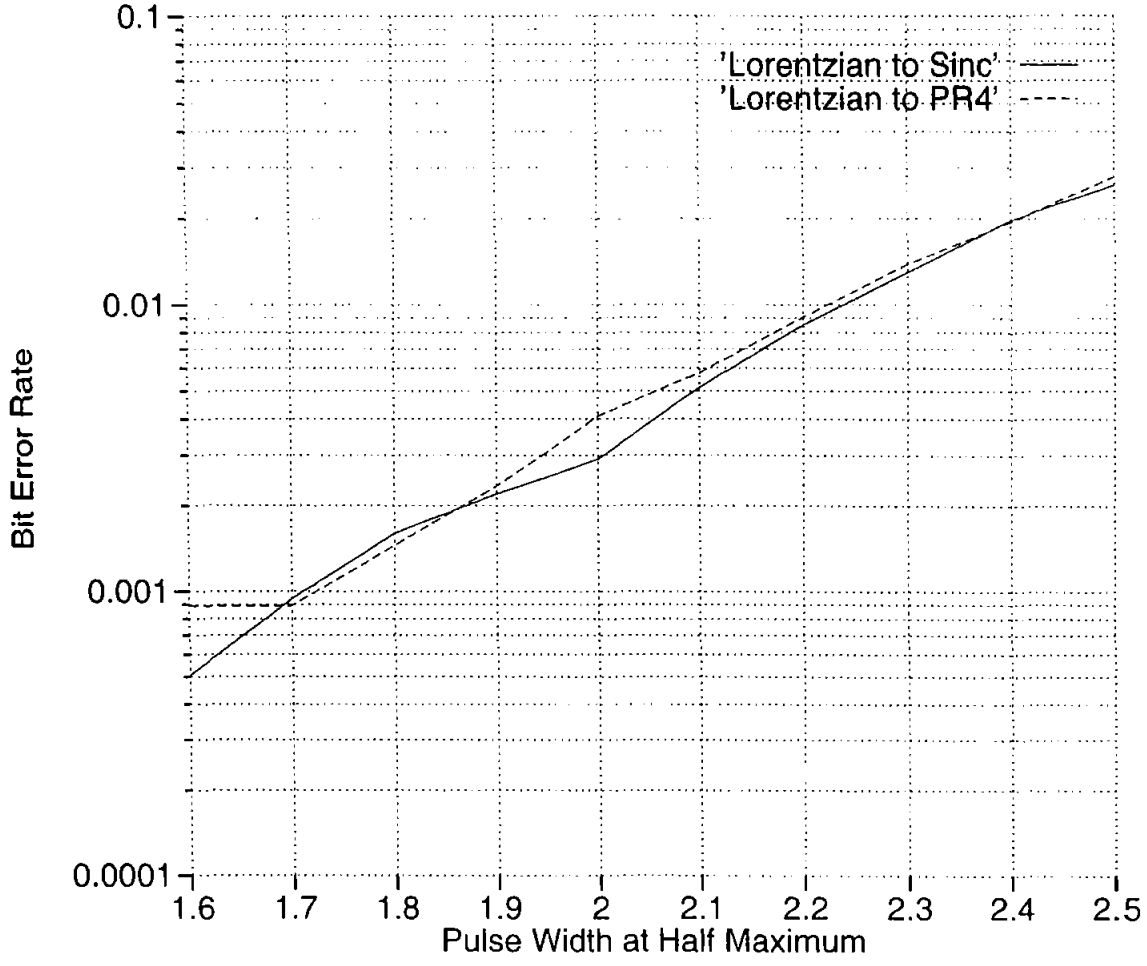


Figure 2.9: 7 Tap FIR Filter Equaliser

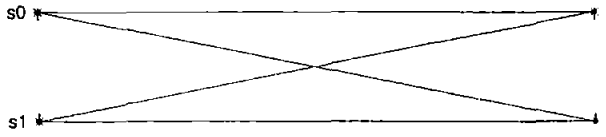
2.2.2 Maximum Likelihood (ML) Detection

The Maximum Likelihood Sequence Detection (MLSD) of the PR samples is achieved using the Viterbi Algorithm described by Forney[Jr.73], and by Kobayashi[Kob71]. The detection scheme used in the software simulations was that described by Dolivio[DUH86], which simplifies the PR4 channel to two interleaved $(1 - D)$ Channels.

Maximum Likelihood Sequence Detection (MLSD) seeks to minimise the probability of error in an Additive White Gaussian Noise (AWGN) channel. The detector selects the sequence that minimises the euclidean distance metric.

The output from the MLSD implementation consists of "hard" decisions (1's and 0's)

and conveys no information on the reliability of each decision. The trellis described by Dolivio[DUH86] is shown in figure 2.10.



Start State	End State	Path Metric
s0	s0	0
s0	s1	−1
s1	s0	1
s1	s1	0

Figure 2.10: Conventional Interleaved PR4 Single-Track Trellis Section.

In the simulated PR4 scheme, two detectors are used, one operating on the even samples, and the other on the odd samples. The detectors used take in soft values (real numbers) and output hard decisions. This detector can be described as a Soft Input Hard Output (SIHO) detector. Soft Input Soft Output (SISO) methods will be discussed later in the chapter on Error Correction.

From the equations that calculate the new metric values, a transition on the recording media occurs when one of the diagonal paths in the trellis is more likely.

Error events that this detector (figure 2.10) is weak to are of the type "+−" or noise which mimics opposite polarity in succession (a positive going noise spike followed by a negative going noise spike in the odd or even samples)[Kob71][Tar95]. This leads to bursts of errors in the detected sequence. This has been verified by the author using computer simulations.

The overall system simulated in software is shown in figure 2.11. The data are generated using a Pseudo Random Binary Sequence (PRBS) generator published in Lin and Costello[LC82]. A variety of two-track run length encoders can be used, resulting in two sequences. The two sequences are precoded and passed through a $1 - D$ filter. The output from this filter is then convolved with a Lorentzian pulse of specified PW_{50}. At this

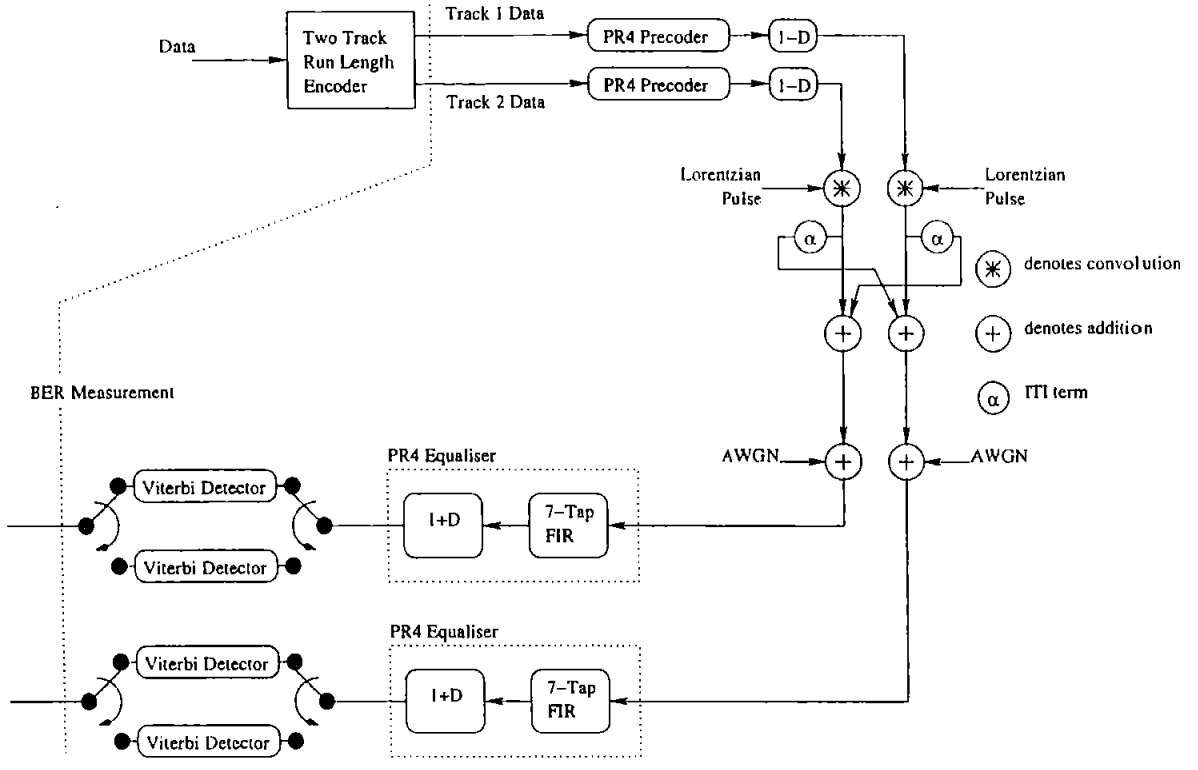


Figure 2.11: System Simulated in Software.

stage a predetermined amount of the adjacent track information is added (simulating the ITI). Additive White Gaussian Noise (AWGN) is added based on either the algorithm described by Devroye[Dev86] or Press[PTVF92]. Unless otherwise stated, the signal to noise ratio (SNR) is assumed to be defined by equation 2.12

$$SNR = 10 \log_{10} \left(\frac{1}{2\sigma^2} \right) \tag{2.12}$$

where σ is the standard deviation of the Gaussian noise distribution. The sampled data are then equalised to a PR4 response, either using

1. A Lorentzian to PR4 filter, or

2. using a Lorentzian to Sinc filter and a $1 + D$ filter

The sampled data are then separated into odd and even samples, and the trellis (figure

2.10) is used to detect the data using a Viterbi detector. The data are recombined and Bit Error Rate measurements are made comparing the detected data with the encoded data as shown in figure 2.11

2.3 Coding for Magnetic Recording

The magnetic recording channel uses two types of codes.

1. An error correction code (ECC).

2. A recording or modulation code.

The object is to obtain reliable retrieval from the channel. Normally the modulation code is applied after the ECC. Various authors have used combined modulation and error correction codes[WU86, FHN86, LW87]. Modulation codes that preserve the Hamming distance of the input sequence are described in[FHN86], and single error correcting RLL codes are described in[LW87].

Codes can also be used to combat synchronisation errors, where due to timing instability there is an insertion or deletion of a symbol[Lev66, Bou94].

2.3.1 Modulation Codes

Modulation codes, also called constrained codes are used to improve the performance of the system by ensuring sufficient transitions in the readback waveform for clock recovery. These codes take the data from the ECC, and generate a constrained sequence which determines the signal that is recorded on the media. The types of constraints can be classified into two categories as suggested below.

Run Length Constraints

Run Length Limited (RLL) Codes were discussed by Franaszek[Fra68]. These codes specify the minimum (d_{rll}) and maximum (k_{rll}) runs of consecutive symbol 0. Thus a (1,7) RLL would have a minimum of 1 zero and a maximum of 7 zeros between consecutive ones. RLL constraints can be extended to the following classes

- RLL constraints introduced by Funk[Fun82], specified as (d_{rll}, k_{rll}, s), specify a parameter s that indicates the runs of zeros are in the form of $d_{rll} + is$ where i is a non negative integer.
- In PRML systems, the maximum runlength constraints are satisfied globally and in the odd and even interleaves. These codes are specified as $(d_{rll}, G/I)$ PRML codes, where G is the global k_{rll} constraint and I is the interleaved k_{rll} constraint.
- A limit on the maximum number of ones in the encoded sequence results in Maximum Transition Run (MTR) codes[MB96]. MTR codes are specified by the k_{rll} constraint, as the d_{rll} is normally zero.

Spectral Null Constraints

Spectral Null constraints specify that the recorded waveforms have no spectra content at a particular frequency f. The code power spectrum (equation (2.13)) is discussed in more depth in Pless and Huffman[PH98] and Immink[Imm91]. A necessary and sufficient condition that an encoder has a spectral null at a frequency $f = \frac{m}{nT}$, where m and n are real numbers, and T is the sampling period, is that there is a constant B such that for all sequences $\mathbf{w} = w_0, w_1, \ldots, w_{N-1}$ and $0 \leq i \leq i' < N$

$$\left| \sum_{h=i}^{i'} w_h e^{-j2\pi hm/n} \right| \leq B \quad \text{where } j = \sqrt{-1} \tag{2.13}$$

A spectral null at $f = 0$ results in a dc-free sequence (very popular in optical recording[Imm91]) and the parameter B is the digital sum variance or the running digital sum of the encoder. Further information can be obtained in Karabed[KS91]. Codes that have both runlength and are dc-free are specified as $(d_{rll}, k_{rll}; c)$ codes with $B \leq 2c$ see Pless[PH98]

Many codes can achieve identical RLL constraints. The capacity due to these constraints is used to provide a measure of code efficiency.

Capacity

Shannon defined capacity C as

$$C = \lim_{T \to \infty} \frac{\log(N(T))}{T} \tag{2.14}$$

where $N(T)$ is the number of allowed signals of duration T[Sha48]. The ratio of the code rate, R, divided by the capacity (due to the constraints), C, provides a measure of code efficiency η as shown in equation (2.15).

$$\eta = \frac{R}{C} \tag{2.15}$$

A RLL code can be described by a State Transition Diagram (STD) and a State Transition Matrix (STM). It can be shown[MS84] that the capacity due to constraints can be computed as

$$C = \log_2(\lambda_{max}) \tag{2.16}$$

where λ_{max} is the largest real eigenvalue (or the spectral radius) of the state transition matrix of the constraints.

This equation has been extended to n_{tracks} track situations as

$$C = \frac{\log_2(\lambda_{max})}{n_{tracks}} \tag{2.17}$$

2.3.2 Error Correction Codes

ECC differ from modulation codes in that they add redundancy in order to correct errors. Typically Reed Solomon codes are used[RS60]. ECC is currently a region of tremendous research with the invention and use of Turbo Codes[BGT93, RMM98, MMH98], and the "rediscovery" of Low Density Parity Check Codes (LDPC) introduced by Gallager[Gal63]. These coding systems use Soft In Soft Out (SISO) decoders and use the *a posteriori* probability from separate component codes to achieve BER performance close (< 1dB) to the theoretical limit (the Shannon Capacity) for Binary Phase Shift Keying (BPSK) systems. There is considerable work being done currently in porting these techniques to magnetic recording[SC00, RMM98, MMH98, STC01, FFKM99]. Reed Solomon codes are also being decoded using SISO techniques[KV00a, KV00b, Sud97]. A few ECC parameters are discussed in this chapter.

- **Minimum distance** : The number of places where two codewords differ is the Hamming distance (or just the distance). The smallest distance Hamming distance between any two codewords is the minimum distance d_{min}.

- **Code word multiplicity** : The number of neighbouring codewords at a distance d from a codeword is defined as the number of neighbours n_n. This is an important factor that determines the performance of a coded system. The multiplicity and the minimum distance of an encoder are the two most important parameters of an error control system.

- **Error floor** : An error floor appears in the FER vs SNR characteristics of coded systems due to very good performance almost upto the Shannon limit. Performance of coded systems are bound by the Shannon limit and the code properties listed in equation (2.18) and (2.19) as shown in figure 2.12.

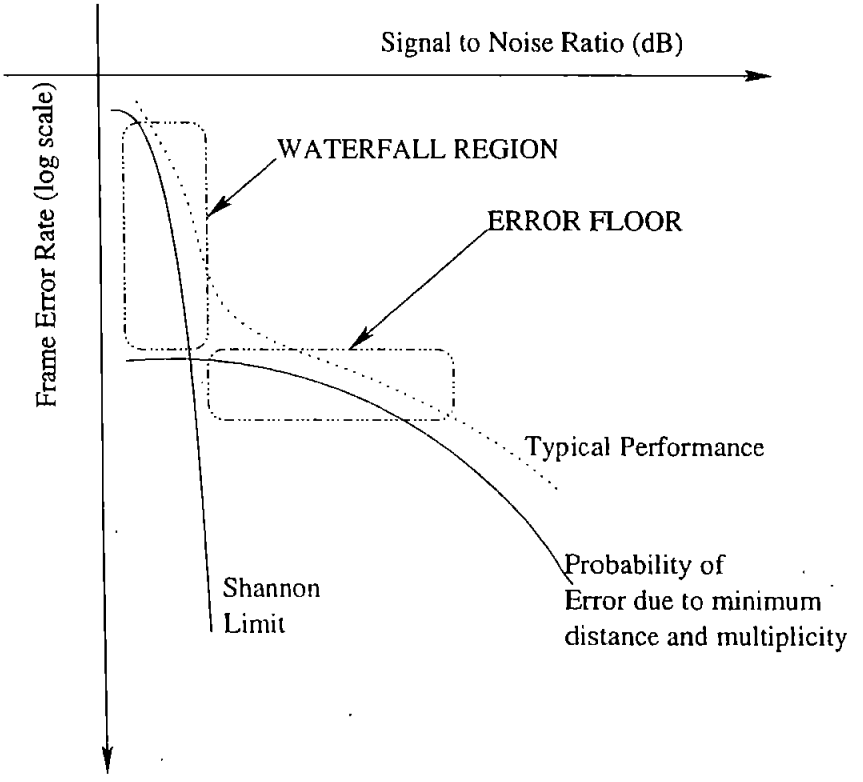


Figure 2.12: Typical Turbo Code Performance Characteristics

The Waterfall Region is due to Shannons limit for the channel, and the Error floor is due to the Probability of Error of the code. The probability of error for an AWGN noise source with standard deviation σ can be upper bound (using the $\mathbf{Q}$ function[¶]) by

$$P_e \leq \sum n_n \mathbf{Q}\left(\frac{d_{min}}{2\sigma}\right) \tag{2.18}$$

when all neighbours are at distance d_{min} and,

$$P_e \leq \sum_j n_{nj} \mathbf{Q}\left(\frac{d_j}{2\sigma}\right) \tag{2.19}$$

when neighbouring codewords are at different distances from each other. See Burr[Bur01] for more details.

- **Apriori Probability (AP)** : Apriori probabilities are the probabilities of the received bits being 1 or 0 (for a binary sequence) before decoding. These are calculated using the probability density function (pdf) of the received samples. For a PR4 channel with ideal samples $\{+1, 0, -1\}$ in the presence of AWGN with standard deviation σ and given a received sample r, the AP's can be calculated using equation (2.20)

$$\begin{aligned} \text{for } S &= 1, 0, -1 \\ \mathbf{P}(r = S) &= k\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(r-S)^2}{2\sigma^2}} \\ &\quad \text{where } k \text{ is a constant such that} \\ &\quad \mathbf{P}(r = 1) + \mathbf{P}(r = 0) + \mathbf{P}(r = -1) = 1 \end{aligned} \tag{2.20}$$

¶

$$\begin{aligned} \mathbf{Q}(z) &= \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-\frac{x^2}{2}} dx = \frac{1}{2}\mathbf{erfc}\left(\frac{z}{\sqrt{2}}\right) \\ \mathbf{erfc}(z) &= \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-x^2} dx \end{aligned}$$

- **A Posteriori Probability (APP)** : This is the probability of bits taking into account all constraints imposed by the ECC. The AP's and the encoder parity check constraints are used to obtain the APP's.

- **Extrinsic Information** : This is the ratio of APP/AP for each bit. The logarithm of the extrinsic information is the definition of Mutual Information[Gal68].

The term turbo-decoding is used to describe decoders that evaluate the APP's using extrinsic information from different decoders.

Detection

The detection of an encoded sequence can be done using trellis based or "block based" algorithms. If reliability information is obtained from the decoder, the algorithm is termed **soft output**; and if no reliability information is generated, the algorithm is termed a **hard output**.

Detection algorithms implemented include

1. Forney's[Jr.73] Viterbi algorithm (a Hard Output algorithm resulting in the most likely codeword also referred to as a Maximum Likelihood Sequence Detector (MLSD)).

2. Bahl[BCJR74] (BCJR) algorithm, a Soft Output algorithm resulting in the most probable bits (which will not necessarily result in the most likely codeword). It outputs the exact APP's.

3. Gallagers Second Method[Gal63] also referred to as the Sum-Product Algorithm or Belief Propagation[MMC98]. These methods refer to the same technique. The parity check constraints are individually brute force decoded to obtain APP's, and all appropriate APP's are combined to given a single APP. The output is normally an approximation of the true APP's. Only under certain circumstances (in cycle-free decoding graphs) does this method result in the exact APP's.

These methods are discussed in Burr[Bur01], Pless[PH98] and several authors notably[RU01a, RU01b].

2.4 Outline of the Investigation

This thesis investigates a two-track PR4 magnetic recording channel. The aim of the investigation is to increase the recording density by bringing tracks closer together while maintaining high linear density.

The effects of ITI are investigated and mitigated, using

1. Error Correction Coding (Chapter 3).
2. Two-Dimensional Modulation code (Chapter 4).
3. Advanced Signal Processing methods (Chapter 5).

The thesis describes the use of advanced error correction codes (Low Density Parity Check Codes, LDPC Codes) to improve resilience to ITI. Very high rate LDPC codes are investigated for a two-track PR4 channel. There is currently very little research activity in using LDPC Codes in multi-track magnetic channels. Modulation codes are investigated and work done by Lee[LM94], Bane[BO97] and Davey[DDMD98] is extended. This looks at the recording channel, and methods of mitigating ITI on the medium using modulation codes. Signal processing methods are then investigated in order to improve the performance of the detectors. The new detectors are compared with previous work done by Soljanin[Sol94] using sliding block decoders. The readback of data from the magnetic channel in the presence of ITI is optimised.

3 Error Control Coding

This chapter explores the use of Error Correction Codes (ECC) to improve the BER performance of a two-track magnetic system in the presence of ITI. Very high rate codes are employed (Code Rate $R > 0.9$). Low Density Parity Check (LDPC) codes are used as they are "very good"* generic codes[Mac99]. Very powerful codes (like Turbo Codes) can be described as LDPC's[Mac00].

LDPC codes were introduced by Gallager[Gal63] and largely abandoned due to computational complexities. The class of codes recently regained attention due to MacKay[MN97] and the fact that decoding was no longer unrealistic. LDPC were applied to magnetic recording by Fan[FFKM99] and Song[STC01, SC00]. The use of soft serially concatenated high rate codes for a PR4 channel was discussed by Siala[SPHK95] in 1995. Turbo codes were applied to magnetic channels by Ryan[RMM98] and McPheters[MMH98]. The magnetic channel is regarded as a very weak inner code in all these schemes, serially concatenated with a more powerful outer ECC.

This section describes very high rate LDPC codes, their encoding and decoding and presents some results. Recording codes normally have rates in the region of 0.9. A combined error correction code (ECC) and modulation code with an overall rate of about 0.8 suggests the ECC would have a rate of about 0.9. The combined rate of 0.8 was chosen arbitrarily by the author. The value of 0.9 was chosen by the author as a lower limit of the ECC code rate.

*A "very good" code family is defined as one that achieves arbitrary small probability of error at any communication rate up to the Shannon limit.

3.1 Low Density Parity Check (LDPC) Codes

LDPC codes are specified using a sparse parity check or **H** matrix[†]. An (n, j, k) LDPC code has a block length n, with an **H** matrix with j 1's in each column and k 1's in each row.

$$\mathbf{H} = \overbrace{\left.\begin{bmatrix} 1 & 1 & 1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & & & \vdots & \vdots \\ 0 & 1 & 0 & \cdots & 1 \end{bmatrix}\right\} nj/k \text{ rows}}^{n \text{ columns}} \tag{3.1}$$

The code rate R can be lower bound using

$$R \geq \frac{n - \frac{nj}{k}}{n} = 1 - \frac{j}{k} \tag{3.2}$$

The **H** matrix (equation (3.1)) can be subdivided into j matrices, each containing a single 1 in each column. All rows in the sub-matrix are independent of each other[‡]. All submatrices are a random column permutation of the first submatrix. This matrix generation method was proposed by Gallager and LDPC codes were constructed by ensuring the sub-matrices were exhaustive of all bits and no bit in each sub-matrix participated in more than 1 parity check[Gal63]. Equation (3.1) can be written as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_j \end{bmatrix} \tag{3.3}$$

[†]an **H** matrix containing many 0's and very few 1's for a binary system
[‡]bits participate in only one check equation or row in the sub-matrix

where

$$\mathbf{H}_1 = \overbrace{\left.\begin{bmatrix} 1 & 1 & \dots \\ 0 & 0 & \dots \\ \vdots & \vdots & \vdots \\ 0 & 0 & \dots \end{bmatrix}\right\} n/k \text{ rows}^{\P}}^{n \text{ columns}^{\S}} \tag{3.4}$$

and $\mathbf{H}_2$, $\mathbf{H}_3 \dots \mathbf{H}_j$ are all column permutations of $\mathbf{H}_1$. For example an $n = 20$, $j = 5$ and $k = 10$ will have $\mathbf{H}_1 = \begin{smallmatrix} 1&1&1&1&1&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&1&1&1&1&1 \end{smallmatrix}$, and $\mathbf{H}_2$, $\mathbf{H}_3$, $\mathbf{H}_4$ and $\mathbf{H}_5$ being some random column permutation of $\mathbf{H}_1$. A possible $\mathbf{H}$ matrix is shown below

$$\mathbf{H} = \begin{smallmatrix} 1&1&1&1&1&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&1&1&1&1&1 \\ 1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0 \\ 0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1 \\ 1&1&0&0&1&1&0&0&1&1&0&0&1&1&0&0&1&1&0&0 \\ 0&0&1&1&0&0&1&1&0&0&1&1&0&0&1&1&0&0&1&1 \\ 0&0&0&1&1&1&0&0&0&1&1&1&0&0&0&0&1&1&1&1 \\ 1&1&1&0&0&0&1&1&1&0&0&0&1&1&1&1&0&0&0&0 \\ 0&0&0&0&1&1&1&1&0&0&0&0&1&1&1&1&0&1&0&1 \\ 1&1&1&1&0&0&0&0&1&1&1&1&0&0&0&0&1&0&1&0 \end{smallmatrix}$$

If an $\mathbf{H}$ matrix adheres strictly to the j and k constraints, the resulting code is referred to as a regular LDPC, and if the $\mathbf{H}$ matrix does not the code is described as an irregular LDPC. It has been shown by Luby that irregular LDPC's generally have better performance[LMSS98].

[§] with only a single one in each column

[¶] with k ones in each row

In order to "design" an **H** matrix for a given code rate $\hat{R}$, j and user bits b, algorithm 3.1 below can be used.

Algorithm 3.1. Starting with a design code rate $\hat{R}$, b user bits and a Gallager j constraint, the final Gallager n and k constraints can be found using the following algorithm. This is iterated until valid n and k are obtained ‖.

define: $\hat{p} = \left\lfloor \frac{b}{\hat{R}} \right\rfloor$

begin:

if: $\left\lfloor \frac{b+\hat{p}}{\hat{p}/j} \right\rfloor = \frac{b+\hat{p}}{\hat{p}/j}$

then:

$n = \hat{p} + b$

$k = \frac{b+\hat{p}}{\hat{p}/j}$

end:

else: $\hat{p} = \hat{p} - 1$ and goto **begin:**

The new code rate R is then obtained as

$$R = \frac{b}{\hat{p} + b} \tag{3.5}$$

In this work, LDPC codes of rate > 0.9 are investigated. Unless otherwise stated, 4096 user bits are RLL encoded using a rate 8/9 (0,4/4) PRML encoder. The resulting 4608 bits are then encoded at a rate of R resulting in $\frac{4608}{R}$ bits. These $\frac{4608}{R}$ bits are then detected using an iterative combined soft-in soft-out trellis/LDPC detector.

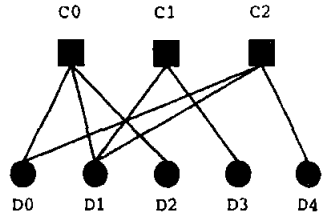
An **H** matrix can be described using a *bipartite* graph[Tan81]. A bipartite or factor graph is a graph with two types of nodes, **check** nodes and **data** nodes and are an

‖the algorithm is also made to terminate for negative $\hat{p}$

alternative method of describing the parity check equations in the **H** matrix[KF98]. An example is shown in Figure 3.1.

	D0	D1	D2	D3	D4
C0	1	1	1	0	0
C1	0	1	0	1	0
C2	1	1	0	0	1

(a)**H** Matrix



(b)Bipartite Graph

The top nodes on the bipartite graph represent check nodes and the bottom nodes represent data nodes. The number of columns of the **H** *Matrix is equal to the number of data nodes and the number of rows of the* **H** *Matrix is equal to the number of check nodes. A path or edge in the graph occurs if a particular data bit participates in a check.*

Figure 3.1: **H** Matrix and its associated bipartite graph

Some definitions associated with graphs are as follows

Path A path or an edge is a connection between a data node (a bit) and a check node (a parity check equation). This exists only if the bit participates in the parity check equation.

Cycles A cycle is a set of different paths that start and end at a particular node. For example in Figure 3.1, **D0 – C0 – D1 – C2 – D0** is a cycle of length 4. Cycles of length 4 are considered as being "bad" when decoding LDPC's using the sum-product algorithm (page 42). This is because short cycles result in information being fed back very quickly (without being influenced by many parity checks) during decoding. A cycle of length 4 occurs for the following combination of ones in the **H** Matrix.

$$\mathbf{H} = \begin{bmatrix} & \vdots & & \vdots & \\ \cdots & 1 & \cdots & 1 & \cdots \\ & \vdots & & \vdots & \\ \cdots & 1 & \cdots & 1 & \cdots \\ & \vdots & & \vdots & \end{bmatrix}$$

Degree This is the number of paths connected to nodes. Regular LDPC's have the same degree on all data nodes and the same degree on check nodes. The degree of the data nodes is equal to the j constraint, and the degree of the check nodes is equal to the k constraint. Irregular LDPC's can have very different data and check node degrees.

3.1.1 H Matrix Generation and Encoding

Gallager describes a methods for generating the **H** Matrix of an LDPC[Gal63]. A slightly different method is used for very high rate LDPC codes.

- The first nj/k columns are made an $nj/k \times nj/k$ identity matrix. This results in systematic codes. The first nj/k columns need to be invertible in order to encode; and at very high rates, it was found by the author that the number of independent rows reduces rapidly, reducing the number of parity checks. In order to maintain the number of parity checks, an identity matrix is placed in the first part of the **H** Matrix. This observation has also been made by Song[STC01].

- For a given n, j and k constraint a randomly generated, regular **H** Matrix is generated in the rightmost $n - nj/k$ columns.

- Cycles of length 4 can be avoided by simply eradicating one of the paths in the cycle. This affects the j and k constraints, affecting the j more than the k. It is ensured that the minimum j is greater than or equal to one.

The **H** Matrix constructed in such a way can be described as

$$\mathbf{H} = [\ \mathbf{I}\ |\ \mathbf{LDPC}\ ] \tag{3.6}$$

where **I** is an identity matrix with nj/k rows and **LDPC** is an LDPC created from a regular Gallager LDPC **H** Matrix.

H matrices have been generated with the following j and k.

j	k	Code Rate
2	18	0.900
3	32	0.914
4	36	0.900
2	24	0.923
3	36	0.923
4	48	0.923
2	32	0.941
3	48	0.941
4	64	0.941
2	48	0.960
3	72	0.960
4	96	0.960
2	128	0.985
3	192	0.985
4	256	0.985

$[\mathbf{I} \mid (4608, j, k)]$ LDPC Code Parameters

Table 3.1: Valid j and k for 4608 data bits using Algorithm 3.1

H matrices that are high rate and free of cycles of length 4 can also be derived from combinatronic designs, in particular Kirkman Triple Systems[JW01]. An introduction to combinatronics and the design of the Kirkman Triple System used by the author can be found in the appendix (page 119).

A comparison between a KTS construction and the scheme due to equation (3.6) is shown in figure 3.2. The Kirkman Triple System compared is a $\nu = 183$, $b = 5551$ system, and scheme due to equation (3.6) has identical n and j parameters. It is clear from figure 3.2 that at high rates (0.97), codes of comparable performance can be obtained from a Kirkman Triple System and LDPC codes based on equation (3.6). An advantage of the latter LDPC is that the resulting **H** matrix is systematic. The matrices obtained from the KTS construction are not systematic. Encoding data using this is more computationally intensive compared to (3.6).

At rates of 0.9, KTS based LDPC codes have better performance at higher E_b/N_o. Magnetic recording systems (hard disk drives) normally operate at a raw channel error

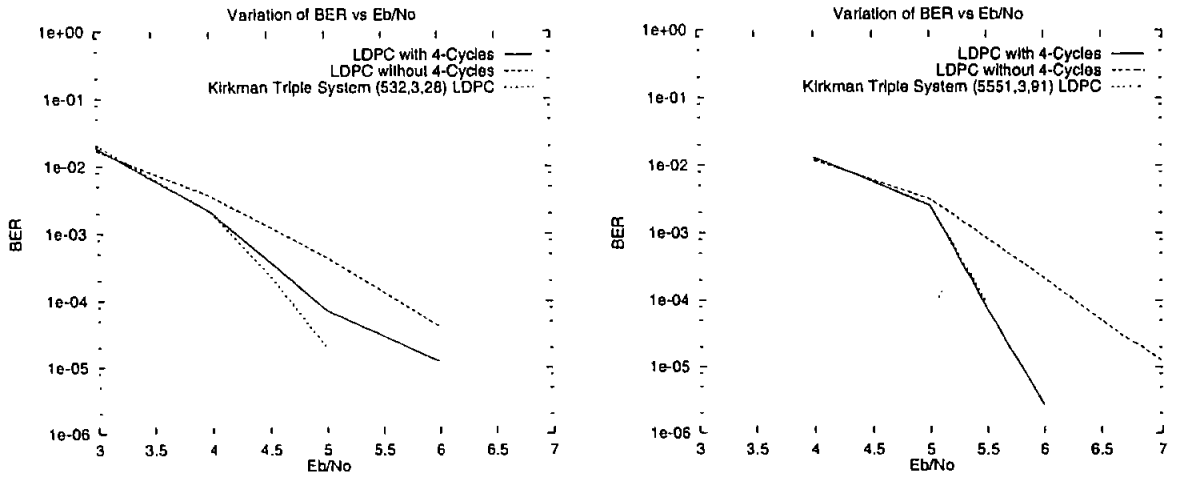
m	q	ν	b	Code Rate
1	7	21	70	0.70
2	13	39	247	0.84
3	19	57	532	0.89
5	31	93	1426	0.93
6	37	111	2035	0.95
7	43	129	2752	0.95
10	61	183	5551	0.97
11	67	201	6700	0.97
12	73	219	7957	0.97
13	79	237	9322	0.97
16	97	291	14065	0.98
17	103	309	15862	0.98
18	109	327	17767	0.98

m is an integer, q is a prime number,
ν is the number of parity check equations and
b is the number of codeword bits
All the $\mathbf{H}$ matrices based on the incidence matrix
of these Kirkman Triple systems have $j = 3$

Table 3.2: Valid Parameters for a Kirkman Triple System

rate of about 10^{-3} to 10^{-4}[Mal]. The gain offered by KTS constructions is about 0.2dB in this error rate region. The presence of cycles of length 4 improves the performance of the code due to equation (3.6).

The authors LDPC coding scheme seems to have an error floor at about 10^{-4} at a rate of 0.9. This is due to a very low minimum distance of the authors LDPC. The decoder used was an SPA decoder with a maximum of 200 iterations, that terminated if the correct codeword was found. Simulations stopped at a maximum of 100 frames in error.



(a) Rate 0.89 (b) Rate 0.97

In the authors scheme, cycles of length 4 were removed by randomly removing a 1 from the pattern in the $\mathbf{H}$ *matrix. It was ensured that the number of 1 in columns of the* $\mathbf{H}$ *matrix was* ≥ 1. *The simulations were done assuming a BPSK channel, and the* E_b/N_o *was normalised for the code rate. This was done to enable the new LDPC codes to be compared with existing codes as most performance graphs are made assuming a BPSK channel.*

Figure 3.2: $\frac{E_b}{N_o}$ performance comparison of KTS LDPC and the scheme described in equation (3.6)

3.2 Decoding Algorithms

This section describes maximum aposteriori probability (MAP) decoding methods used by the author. MAP decoding is used to obtain *aposteriori* probabilities based on the current *apriori* probability and the constraints the code imposes. The ratio of aposteriori probability (the MAP output, or APP) divided by the apriori probability (the MAP input, or AP) is the extrinsic information. Extrinsic information is the most essential ratio information that is passed between decoders in iterated schemes.

3.2.1 The Sum-Product Algorithm

LDPC's are decoded using the Sum-Product Algorithm (SPA). Extrinsic information is obtained from the SPA and in a concatenated scheme and is passed to other decoders.

To understand extrinsic information, Apriori Probabilities and Aposteriori Probabilities consider the following example. Consider the following **H** matrix $H = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$, resulting in a codeword $b_0 b_1 b_2$. Suppose the received probabilities of the three bits are

$$\Pr(b_0 = 0) = 0.5 \quad \Pr(b_1 = 0) = 0.6 \quad \Pr(b_2 = 0) = 0.4$$

The APP can be evaluated using a brute force method by

- enumerating all valid codewords,
- calculating (and normalising) the probability of each codeword and from the codeword probabilities
- re-evaluating the bit probabilities (the APP's).

If all possible combinations (not only valid codewords) are enumerated, the APP would equal the AP. For the example this is shown in Table 3.3

Apriori Probability

$$\Pr(b_0 = 0) = 0.5; \Pr(b_1 = 0) = 0.6; \Pr(b_2 = 0) = 0.4$$

Valid Sequences			Combined	Normalised Combined
b_0	b_1	b_2	Probability	Probability
0	0	0	0.12	0.24
0	1	1	0.12	0.24
1	0	1	0.18	0.36
1	1	0	0.08	0.16

Aposteriori Probability

$$\Pr(b_0 = 0) = 0.48; \Pr(b_1 = 0) = 0.6; \Pr(b_2 = 0) = 0.4$$

The APP is obtained by summing the normalised probability of codewords. For $\Pr(b_0 = 0)$ *this is the sum of first and second codeword.*

Table 3.3: Brute Force APP Decoding

The extrinsic information e is calculated as follows. Taking b_0 for example, the ratio of $e = \frac{\text{APP}}{\text{AP}}$ is

$$e_0 = \frac{0.48}{0.5} = 0.96 \tag{3.7}$$

$$e_1 = \frac{1-0.48}{1-0.5} = 1.04 \tag{3.8}$$

where e_0 is the extrinsic information of 0 and e_1 is the extrinsic information for 1. e_0 and e_1 are normalised to give $e_0 = 0.48$ and $e_1 = 0.52$. These values of extrinsic information suggest that given the *input probabilities* and the *constraints imposed* by the code, we are 52% more sure that b_0 is a 1 and 48% more sure that b_0 is a 0. Comparing this to the extrinsic information for b_1 (for b_1 the APP and AP are equal so $e_0 = e_1 = 0.5$), the extrinsic information for b_1 suggests that given the input probabilities and code constraints, we do not get any more information than what we previously had.

The brute force evaluation of APP's is unrealistic for a large number of bits, and using trellises and graphs to describe the constraints, the APP's can be evaluated using the

BCJR algorithm[BCJR74] and the Sum-Product algorithm[Gal63] respectively.

The sum product algorithm is a graphical method of visualising probability propagation on a bipartite graph. It was proposed by Gallager[Gal63] as a method of obtaining the APP's. The algorithm results in the exact APP's if the check equations are completely independent of each other. It uses the product of difference** to obtain the sum of probabilities that have an even number of 1's. Consider the example on page 42,

$$\begin{aligned} \Pr_0^{b_x} &= \text{Probability } b_x \text{ is a 0} \\ \Pr_1^{b_x} &= \text{Probability } b_x \text{ is a 1} \end{aligned}$$

Then

$$\begin{aligned} \prod_{x=0}^{2}\left(\Pr_0^{b_x}-\Pr_1^{b_x}\right) &= (\Pr_0^{b_0}-\Pr_1^{b_0})(\Pr_0^{b_1}-\Pr_1^{b_1})(\Pr_0^{b_2}-\Pr_1^{b_2}) \\ &= (\Pr_0^{b_0}\Pr_0^{b_1}+\Pr_1^{b_0}\Pr_1^{b_1}-\Pr_1^{b_0}\Pr_0^{b_1}-\Pr_0^{b_0}\Pr_1^{b_1})(\Pr_0^{b_2}-\Pr_1^{b_2}) \\ &= \Pr_0^{b_0}\Pr_0^{b_1}\Pr_0^{b_2}+\Pr_1^{b_0}\Pr_1^{b_1}\Pr_0^{b_2}+\Pr_1^{b_0}\Pr_0^{b_1}\Pr_1^{b_2}+\Pr_0^{b_0}\Pr_1^{b_1}\Pr_1^{b_2} \\ &\quad -\Pr_0^{b_0}\Pr_0^{b_1}\Pr_1^{b_2}-\Pr_1^{b_0}\Pr_1^{b_1}\Pr_1^{b_2}-\Pr_1^{b_0}\Pr_0^{b_1}\Pr_0^{b_2}-\Pr_0^{b_0}\Pr_1^{b_1}\Pr_0^{b_2} \end{aligned} \tag{3.9}$$

All sequences with an even number of 1's have a positive sign, and all sequences with an odd number of 1's have a negative sign. This is due to the negative sign in the difference, all combinations of the product with even 1's have a positive sign and this separates the sum of sequences with odd and even number of 1's.

**(Pr(0) – Pr(1))

Equation (3.9) can be seen as

$$\prod_{x=0}^{n}\left(\Pr_{0}^{b_x}-\Pr_{1}^{b_x}\right) = \Pr(\text{all sequences with an even number of ones in sequence}) - \Pr(\text{all sequences with an odd number of ones in sequence})$$

Using the sum of probabilities of all sequences being unity, for a sequence of n bits in a parity check,

$$\Pr(\text{Even Number of 1's in Sequence}) = \left[1+\prod_{x=0}^{n}\left(\Pr_{0}^{b_x}-\Pr_{1}^{b_x}\right)\right]/2 \quad (3.10)$$

$$\Pr(\text{Odd Number of 1's in Sequence}) = \left[1-\prod_{x=0}^{n}\left(\Pr_{0}^{b_x}-\Pr_{1}^{b_x}\right)\right]/2 \quad (3.11)$$

These equations greatly reduce the computation time required to obtain the APP's of each parity check equation in the **H** Matrix.

Equations (3.10) and (3.11) also show that

1. If a single bit has a probability difference of zero, the extrinsic information for all other bits participating in that parity check will be $e_0 = e_1 = 0.5$.

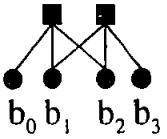
2. If two different bits have a zero probability difference, all parity checks both bits participate in result in no extrinsic information.

Both condition could change from iteration to iteration. These two properties result in little or negligible extrinsic information being generated from an LDPC that has many bits with little or no difference in probability.

Cycle of length 4 can result in non-uniform scaling (double counting[Wei00]) and the introduction of an offset in the computed APP when compared with the exact APP. This is best illustrated using the following example. Consider a simple code with following **H** matrix and bipartite graph. Consider bit b_1, brute force decoding results in the following

	b_0	b_1	b_2	b_3
C0	1	1	1	0
C1	0	1	1	1

(a)**H** Matrix



(b)Bipartite Graph

Figure 3.3: **H** Matrix and its associated bipartite graph for a "toy" code.

extrinsic information

$$e_1 = \overset{b_0}{\underset{1}{\Pr}}\overset{b_2}{\underset{0}{\Pr}}\overset{b_3}{\underset{1}{\Pr}} + \overset{b_0}{\underset{0}{\Pr}}\overset{b_2}{\underset{1}{\Pr}}\overset{b_3}{\underset{0}{\Pr}} \tag{3.12}$$

The belief network will have two extrinsic information for b_1 which will be

$$\hat{e_1} = \overset{b_0}{\underset{1}{\Pr}}\overset{b_2}{\underset{0}{\Pr}} + \overset{b_0}{\underset{0}{\Pr}}\overset{b_2}{\underset{1}{\Pr}} \quad \text{from the first parity check equation}$$

$$\hat{e_1} = \overset{b_2}{\underset{1}{\Pr}}\overset{b_3}{\underset{0}{\Pr}} + \overset{b_2}{\underset{0}{\Pr}}\overset{b_3}{\underset{1}{\Pr}} \quad \text{from the second parity check equation}$$

Multiplying these two extrinsic information results in extrinsic information that differs from equation (3.12) in the two ways described earlier.

$$\hat{e_1} = \underbrace{\overset{b_0}{\underset{1}{\Pr}}\left(\overset{b_2}{\underset{0}{\Pr}}\right)^2\overset{b_3}{\underset{1}{\Pr}} + \overset{b_0}{\underset{0}{\Pr}}\left(\overset{b_2}{\underset{1}{\Pr}}\right)^2\overset{b_3}{\underset{0}{\Pr}}}_{\text{non-uniform scaling}} + \underbrace{\left(\overset{b_0}{\underset{1}{\Pr}}\overset{b_2}{\underset{1}{\Pr}}\overset{b_2}{\underset{0}{\Pr}}\overset{b_3}{\underset{0}{\Pr}} + \overset{b_0}{\underset{0}{\Pr}}\overset{b_2}{\underset{1}{\Pr}}\overset{b_2}{\underset{0}{\Pr}}\overset{b_3}{\underset{1}{\Pr}}\right)}_{\text{offset}}$$

The belief network will result in the exact APP's for b_0 and b_3. The Belief Network can be seen as a very large parallel concatenation of very weak single parity check codes that are turbo decoded. Turbo decoding refers to the iterative exchange of extrinsic information between SISO decoders. Incidentally, if we add the two rows of the parity check matrix, the resulting (equivalent) parity check matrix will be decoded exactly using the sum-product algorithm.

3.2.2 The BCJR Algorithm

Very good mathematical descriptions of the BCJR algorithm exist in almost every modern coding textbook. The algorithm was introduced by Bahl[BCJR74]. The algorithm, as applied to the new trellis derived by the author for PR4 channels is discussed in this subsection. For linear block code minimal trellis construction methods see[MH97].

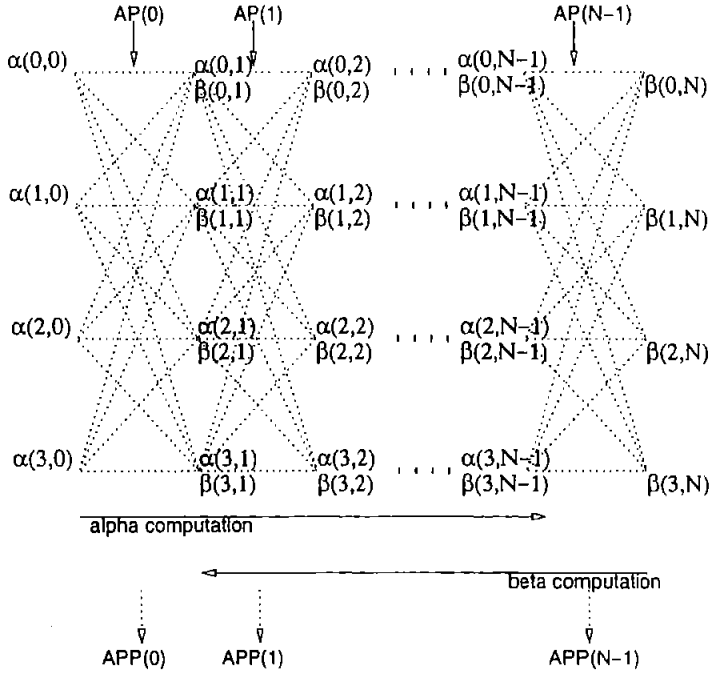


Figure 3.4: BCJR Trellis for New PR4 Trellis

Branch Probabilities The Apriori probabilities are used to compute the branch probabilities. The path probabilities are computed using the paths specified in Figure 5.6. The trellis has 9 different path metrics and 16 different paths with some path metrics appearing more than others. There are two methods of normalising the input probability. The input probability could be normalised so that the sum of probabilities of the 9 different states is unity or the sum of probabilities of the 16 different paths is unity. It was found by the author that normalising the probabilities of the 9 different states resulted in better BER performance. This reinforces the idea that although considering the branch (and state) computations as probabilities

is convenient, the branch (and state) computations are part of a trellis solution to obtain the APP's. The input and output from a trellis are probabilities, however the numbers in between are not.

Forward State Probabilities $\alpha(a,b)$ These are computed by taking the product of the previous α and the probability of the path that links the two states. a specifies the state number and b specifies the sample number. For example,

$$\begin{aligned}\alpha(0,1) &= \alpha(0,0)\Pr(0,0) + \alpha(1,0)\Pr(0,1) \\ &\quad + \alpha(2,0)\Pr(1,0) + \alpha(3,0)\Pr(1,1)\end{aligned}$$

$\Pr(x,y)$ is the probability that track 1 is in state x and track 2 is in state y. $\alpha(a,b)$ is the state probability given the current path probabilities and previous state probabilities. For each section in the trellis, $\alpha(a,b)$ is normalised to unity for numerical stability. At the root of the trellis,

$$\alpha(0,0) = \alpha(1,0) = \alpha(2,0) = \alpha(3,0) = 0.25$$

Backward State Probabilities $\beta(a,b)$ $\beta(a,b)$ are the state probabilities computed from the toor of the trellis to the root. The toor is the end of the trellis (derived by spelling root backwards). The $\beta(a,b)$ are identical to $\alpha(a,b)$ in definition except that the computations for $\alpha(a,b)$ go from the trellis root to the trellis toor. At the toor,

$$\beta(0,\mathrm{N}) = \beta(1,\mathrm{N}) = \beta(2,\mathrm{N}) = \beta(3,\mathrm{N}) = 0.25$$

APP Computation

The APP is computed in the following manner.

1. Channel data are converted to normalised probabilities.

2. The apriori probabilities are passed to the trellis.

3. $\alpha(a, b)$ are computed going from the trellis root to toor.

4. $\beta(a, b)$ are computed from toor to root.

5. The APP of a bit at a trellis section is evaluated using

   - the $\alpha(a, b)$ on the left of a trellis section,
   - the $\beta(a, b)$ on the right of the trellis section and
   - the individual path probabilities of the trellis section

The BCJR decoder and the Sum-Product Algorithm/Belief Network are both used in the scheme implemented by the author.

3.3 Implemented Two-Track Scheme

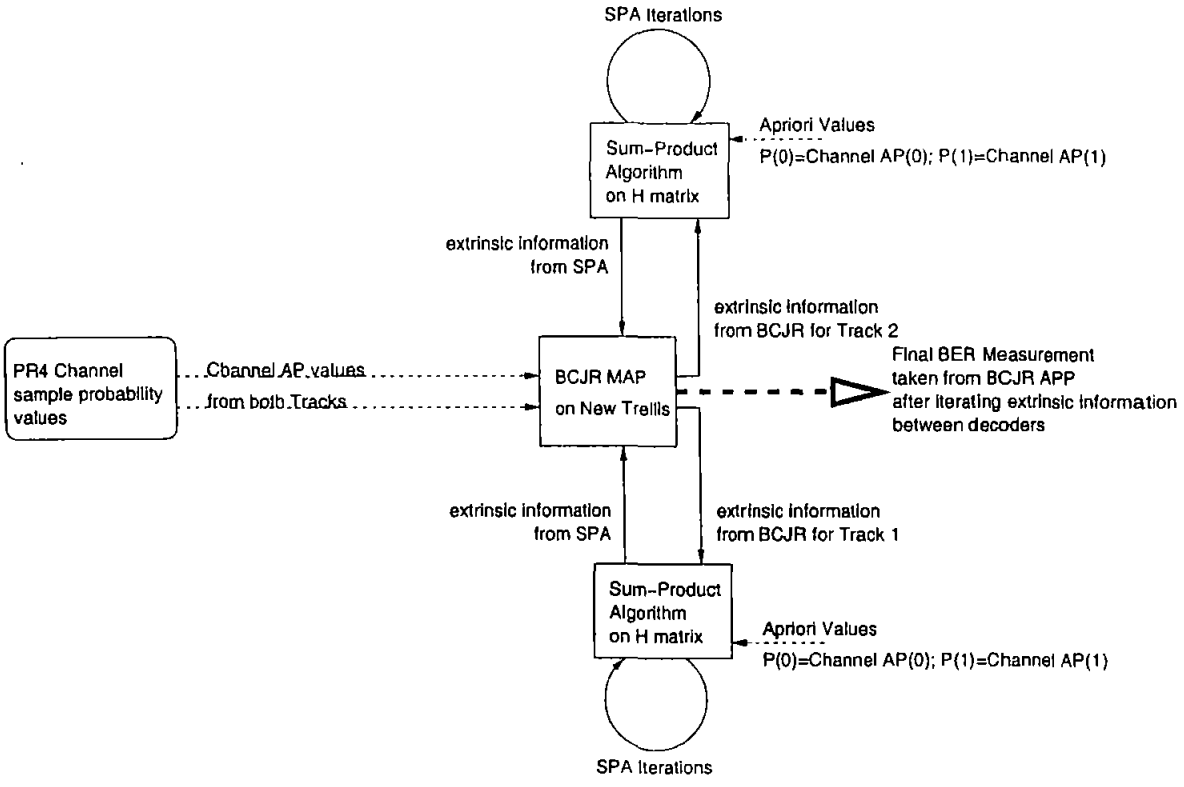


Figure 3.5: Two Track SISO PR4 Detector

Figure 3.5 shows the scheme implemented by the author. The RLL encoded individual track data are encoded using the **H** matrix. The **H** matrices are created using the method described in[JW01] or equation (3.6). The resulting binary sequence is simulated using a PR4 channel experiencing ITI. Sample values are converted to probability values using equation (2.20). These apriori probability are multiplied with the input extrinsic information to the BCJR MAP decoder. The initial extrinsic information supplied to the decoder is $e_0 = e_1 = 0.5$. The product of the input apriori probability and extrinsic information are used to obtain the $\alpha(a, b)$ and $\beta(a, b)$. The APP are then computed using the AP, the $\alpha(a, b)$ and the $\beta(a, b)$. Extrinsic information is then obtained and passed to the two Sum-Product decoders that operate on the individual tracks.

The input AP to the Sum-Product decoders are the channel Apriori probabilities.

These are multiplied by the extrinsic information from the BCJR MAP decoders, and the resulting probability is applied to a single track belief network. The APP are computed, and the extrinsic information obtained. This new extrinsic information is used with the extrinsic information from the BCJR and the belief network is iterated a number of times. At the end of the iterations, the final extrinsic information is passed to the BCJR MAP decoder. The AP of the BCJR MAP and the AP of the Sum-Product decoders are not changed. Only the extrinsic information supplied to each decoder is modified.

An alternative scheme in which the input AP to the Sum-Product decoders was 0.5 was also implemented. The alternative scheme had better performance at lower ITI, however there was no improvement as ITI increased. The alternative scheme performed slightly better than the MLSD decoding (Figure 3.9, page 55). This was abandoned as the performance at higher ITI was not as good as the scheme shown in Figure 3.5.

The **H** Matrix used for encoding and decoding based on equation (3.6) are chosen from an ensemble of matrices with the same n, j and k parameters. A computer search was done on different **H** matrices, and the matrices were graded in terms of the number of inherent cycles of length 4. Simulations were done using matrices with the largest and smallest number of cycles of length 4. The simulations suggested the presence of many cycles of length 4 generally degraded performance.

3.3.1 Decoding Configurations

Two decoding configurations were implemented as shown in Figure 3.6.

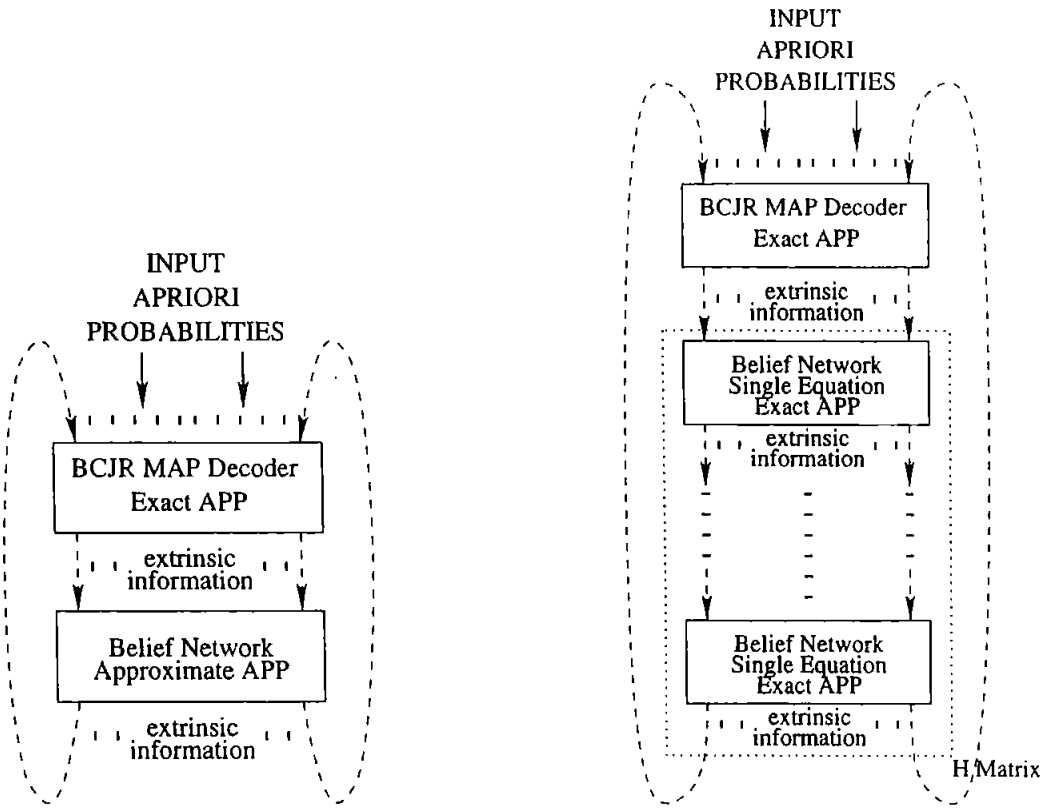


Figure 3.6: Decoding Configurations

"Classical" Decoding of H Matrix: In this scheme the extrinsic information from the entire **H** matrix is updated after evaluating all parity checks. The extrinsic information passed from the belief network is approximate. This is the scheme described by Gallager[Gal63].

Turbo Decoding of H Matrix: Each equation of the **H** matrix is decoded individually in this scheme. The extrinsic information from each equation is passed on to the next parity check. Decoding is much slower compared to the "classical" method as the extrinsic information needs to be updated after each parity check. All extrinsic information is exact, however, treating the **H** matrix in this fashion is an

approximation. This setup treats the **H** matrix as a parallel concatenation of very weak single parity check codes.

It was found by the author that, both in the presence and absence of ITI, the schemes performed relatively identically using codes of rates 0.9.

3.4 Results and Discussion

The effect of cycles and the parameters of the LDPC code are discussed in this section. The decoding setup was a Two Track SISO PR4 Detector(Figure 3.5) using the configuration in Figure 3.6a. The SPA would stop at 200 iterations if a solution was not found, and the MAP decoder and SPA iteratively exchanged extrinsic information.

Effect of Cycles : **H** Matrices with and without Cycles of length 4 were decoded using the new decoder. Some results at code rates of 0.96 and a j of 4 (these values were chosen so as to ensure a large number of cycles and a very weak code, in order to assess the effect of cycles) are shown in Figure 3.7 and 3.8. The variations in

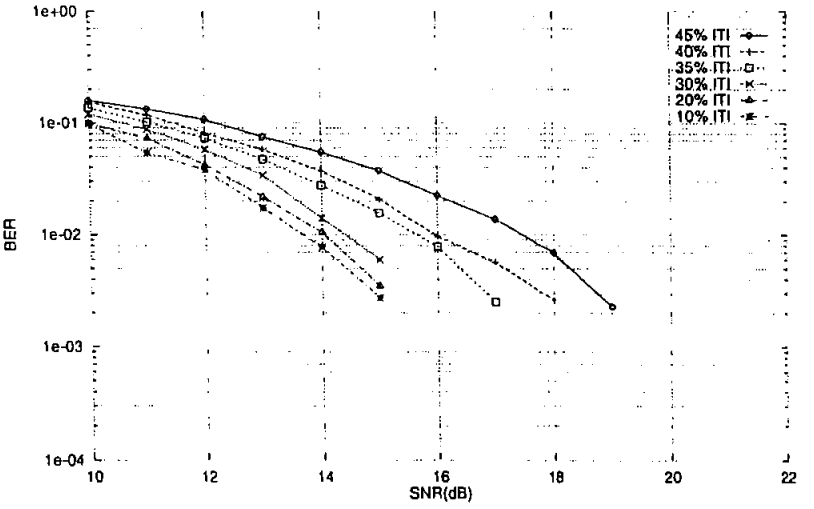


Figure 3.7: Effect of Cycles in decoding using the New Iterative ECC Scheme (**H** Matrix with no cycles of length 4)

Figure 3.7, compared to 3.8, were due to convergence issues in the Belief Network.

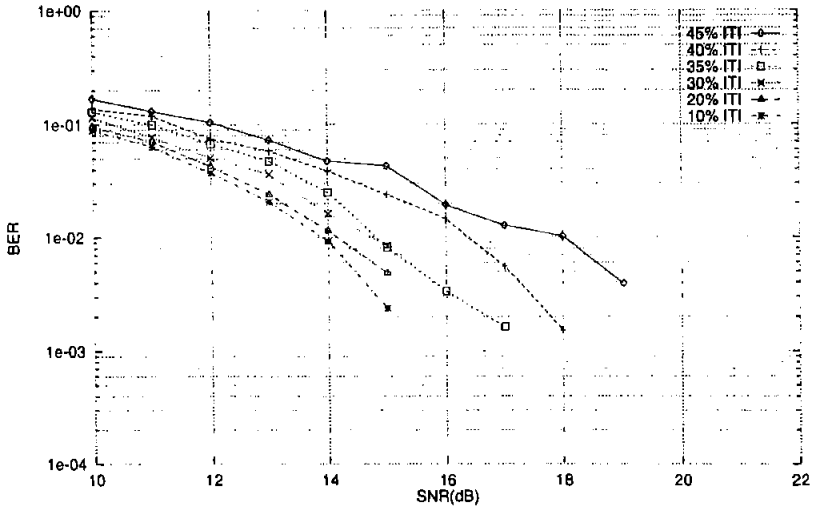
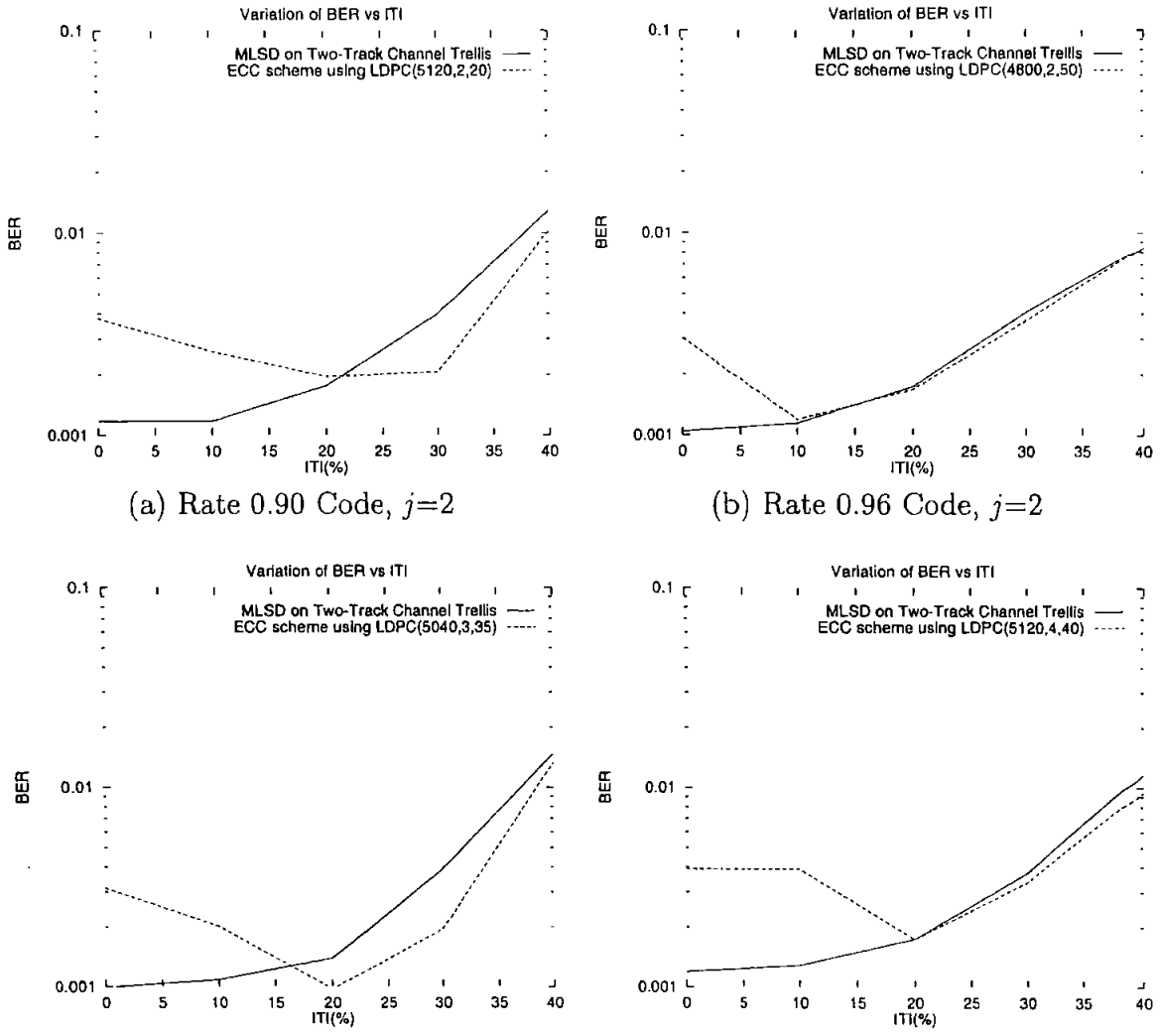


Figure 3.8: Effect of Cycles in decoding using the New Iterative ECC Scheme (**H** Matrix with cycles of length 4)

The instability in the belief network seems to manifest itself more as ITI increases beyond about 30%. **H** matrices with cycles of length 4 seem to be affected more.

Gallager j Constraint : The KTS construction previously described is limited to a j of 3. This can be seen as a limitation of the KTS construction. The authors LDPC codes were constructed with different j constraints and the effect of j was studied. Typical results using various LDPC Codes are shown in Figure 3.9. The use of LDPC Codes is compared to using an MLSD scheme (the Viterbi Algorithm) on the trellis shown in Figure 3.4 on page 47. In all observed cases, at low ITI, the combined ECC and modulation code (referred to as the New Scheme) performs worse than the MLSD scheme. As ITI increases the performance of the New Scheme matches, and sometimes exceeds the MLSD scheme. Comparing Figure 3.9(a) and (b) shows the effect of increasing the code rate and suggest codes of rate 0.96 are very weak and do not result in any improvement in performance using the New Scheme. Figure 3.9(a), (c) and (d) show the effect of increasing the j constraint of the LDPC Codes. The figure suggests that j=3 is the 'best' value. Some resilience to ITI is gained at rates of 0.9, however as the j constraint is increased the number



(a) Rate 0.90 Code, j=2

(b) Rate 0.96 Code, j=2

(c) Rate 0.91 Code, j=3

(d) Rate 0.90 Code, j=4

Figure 3.9: Decoding using the New Iterative ECC Scheme

of cycles of length 4 increases and convergence problems arise again. These results suggest the effect of having cycles can sometimes improve performance when cycles cannot be avoided.

A coding gain is obtained when ITI is greater than about 30%, when the dominant error event changes from the single track error event to the two-track dominant error event. There seems to be a convergence problem in systems with many cycles at high ITI ($\alpha > 30\%$). The theoretical reason for this has not yet been established.

3.5 Contributions

- A new method of iteratively decoding multiple tracks has been presented.
- LDPC codes have been investigated at a very high rate (rates ≥ 0.9) in the presence of ITI with some resilience to ITI observed.
- It has been found from experiments that LDPC codes with $j = 3$ provide the best ensemble of LDPC's for a two track PR4 channel experiencing linear and symmetrical crosstalk.
- The performance of LDPC's based on the incidence matrix of Kirkman Triple System have been compared to the performance of LDPC's suggested by the author.
- A paper entitled "New High Rate Low Density Parity Check Codes for Multi-Track Recording" has been accepted for presentation to the IEEE Intermag 2003 conference in Boston USA.

4 Two Dimensional Codes

4.1 Multi-Track Recording Codes

This section discusses recording codes that use the two dimensional nature of multi-track magnetic recording. Multi-track, or 2 Dimensional Recording codes were introduced by Marcellin and Weber[MW92]. The maximum runlength constraint k_{rll} is satisfied across many tracks in these codes. As the k_{rll} constraint is satisfied across all tracks it relaxes the k_{rll} constraint on individual tracks. As tracks are readback in parallel, the clocking information is derived from all readback signals simultaneously. The result is greater code capacity, permitting the use of higher rate modulation codes.

Swanson and Wolf[SW92] suggest another type of 2 Dimensional Recording code that has an additional k_{rll} constraint along each individual track, as well as the constraint across the group of tracks. This addresses the problem of dropouts, or the loss of information from a single track. The capacity of two dimensional recording codes is discussed by Kato[KZ99].

This section defines types of inter-track interference (ITI), and describes a new class of two-dimensional codes that combat ITI for a PR4 Channel.

4.1.1 Inter-Track Interference (ITI)

Inter-Track Interference (ITI) occurs when tracks get closer together. It is caused by the fringing fields of adjacent tracks. The effects of, and empirical relationships of fringing fields from thin tracks, is discussed by Herk[Her77]. Investigations into ITI for magnetic recording have been done by Soljanin[Sol94], Voois[Voo94] and Vea[Vea93]. ITI is normally modelled as linear and symmetrical across adjacent tracks (see Abbott[ACT88] and Lindholm[Lin78]).

The system simulated considers a two-track two-head model. Data are readback simultaneously, and there is no jitter between the transitions on adjacent tracks. ITI exists between the two adjacent tracks, however it is assumed there is no interaction between groups of tracks. The readback signal as a function of ITI can be described as

$$\begin{aligned} Z_1 &= y_1 + \alpha y_2 + n_1 \\ Z_2 &= y_2 + \alpha y_1 + n_2 \end{aligned} \tag{4.1}$$

where

Z_1 = Readback signal from track 1

Z_2 = Readback signal from track 2

y_1 = Ideal ITI-free signal from track 1

y_2 = Ideal ITI-free signal from track 2

n_1 = AWGN sample

n_2 = AWGN sample

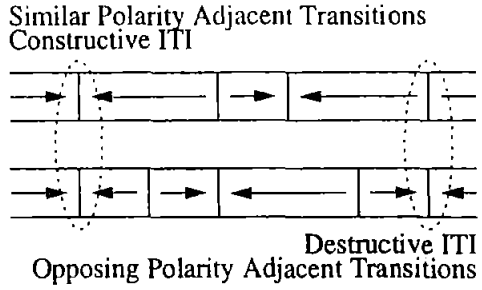
α = Amount of ITI

Assuming ideal equalisation and no noise y_1 and y_2 would belong to $\{+1, 0, -1\}$, and Z_1 and Z_2 would belong to $\{1+\alpha, 1, 1-\alpha, \alpha, 0, -\alpha, -1+\alpha, -1, -1-\alpha\}$.

On the magnetic medium, different types of ITI can be identified depending on the

polarity of the transitions on the adjacent tracks. ITI can be

- **constructive**, when similar polarity transitions occur on adjacent tracks, or
- **destructive** when opposite polarity transitions occur on adjacent tracks.



Arrows pointing towards each other represent a magnetic north pole, and arrows pointing away from each other represent a magnetic south pole. Constructive ITI occurs when there are similar poles on adjacent tracks and Destructive ITI occurs when there are different poles on adjacent tracks.

Figure 4.1: Different types of ITI.

Constructive ITI (CITI) leads to an increased readback voltage whereas Destructive ITI results in a reduction in the readback from both tracks voltage. The amount of ITI is the signal level being read from the adjacent track assuming the readback head stays on track. The maximum amount of ITI possible is 50% according to Herk[Her77], however as read heads go across tracks the relative amount of ITI can be higher than this.

Coding solutions to crosstalk were proposed by Lee[LM94], Bane[BO97] and Davey[DDMD98] Their work impose constraints on a multi-track code to prevent simultaneous adjacent-track transitions, and are referred to as No-ITI (NITI) codes.

A new two-track recording code that allows adjacent track transitions *only if the transition will result in constructive ITI* has been implemented, and published[ADDC01].

4.1.2 Two Dimensional State Constraints

NITI Two dimensional state constraints can be described by the State Transition Diagram (STD) shown in figure 4.2. The STD shown in figure 4.2 does not impose any run length

constraints on the data on each track. State S0 is the state when there are zeros on both adjacent tracks. S1 and S2 occur when there is a 1 recorded on one track and a 0 on the adjacent track.

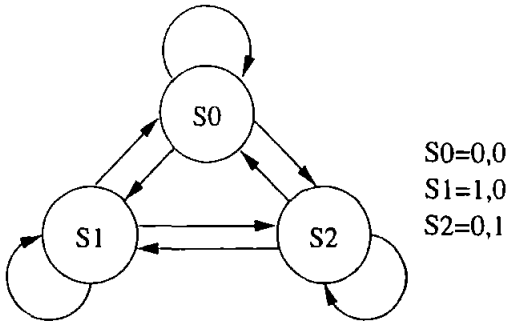


Figure 4.2: STD specifying no simultaneous adjacent track transitions

A State Transition Matrix (STM) describing figure 4.2 is shown in equation 4.2. The STM entries on row i and column j are the number of paths from state i to state j. If there are no path from one state to another, the entry in the STM is 0.

$$STM = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{4.2}$$

The largest real eigenvalue of the STM (equation 4.2) can be evaluated as $\lambda_{max} = 3$. The capacity imposed by these constraints is thus $C = 0.792$ using equation 2.17. This capacity is an upper bound on the code rate that meets these two dimensional constraints. The code described by Davey[DDMD98] is a rate $\frac{4}{6} = 0.667$ code. The efficiency of Davey's code is thus $\frac{0.667}{0.792} = 0.84$ considering the ITI constraints only.

In order to improve the code efficiency, more states are allowed. This results in the Constructive ITI (CITI) Code.

4.2 Constructive ITI Code Design

A new type of code is described in this subsection that does not permit destructive ITI. In a dicode channel (a $(1 - D)$ channel see page 7), the effect of precoding the data sequence results in transitions on the medium occurring where a 1 occurs in the data sequence. This property is exploited in CITI code design. A PR4 Channel $(1 - D^2)$ can be considered as two interleaved $(1 - D)$ channels. The location of transitions on the magnetic media can thus be controlled using this property. The 1's in the encoded data sequence will correspond to transitions on the media as long as data are encoded in an interleaved manner.

On the magnetic media, a positive transition is always followed by a negative transition. When data are written on two adjacent tracks, the tracks can be **in-phase** if the next transition written on either track has the same polarity. The tracks are **out-of-phase** if the next transition written on either track has opposing polarity.

A brute force computer search can be done to obtain codewords and a look-up-table (LUT) can be generated. A block code for a two-track system can be designed based on the following heuristic algorithm.

Algorithm 4.1. 2 Track CITI Codeword Generator Algorithm

1. Starting with all possible two-track sequences.
2. Select valid codewords which
   - do not have all zeros on all tracks and
   - will not lead to destructive ITI whether the tracks are in phase or out of phase*.
3. A two column encoder LUT is produced based on the following.
   - If no adjacent transition exists write codeword to both state I and State O

*for example $\begin{smallmatrix}1&0&1\\1&1&1\end{smallmatrix}$ will always have destructive ITI regardless if the tracks are in phase or out of phase

- If a transition exists, the codeword is placed in state I or state O depending on the ITI. Analysing the transitions in the codeword, it can be determined if the tracks need to be in phase or out of phase for CITI to occur. This determines which column the codeword is written into.
- If at the end of the codeword phase reversal occurs, this is indicated by a slash and the next column the codeword is to be picked from.

The algorithm (4.1) generates codewords, and a LUT is constructed such that a single track does not end up with an all zeros sequence in both the in-phase and out-of-phase tables. The code is specified as a rate $R = k/n$, where k is the number of data bits and n is the total number of codeword bits on both tracks. By inspection, the k_{rll} constraint for two track encoders can be described as shown in equation 4.3 and 4.4

$$k_{rll}^{Track} = \frac{n}{2} + 2\left(\frac{n}{2} - 1\right) \quad \text{The individual track } k_{rll} \text{ constraint} \tag{4.3}$$

$$k_{rll}^{Global} = 2\left(\frac{n}{2} - 1\right) \quad \text{The } k_{rll} \text{ constraint across all tracks} \tag{4.4}$$

The d_{rll} constraint for this is zero.

CITI codes are specified as $\left(d_{rll}, k_{rll}^{Track}\right)$ RLL codes. Thus CITI codes of rate $R = \frac{k}{n}$ have $\left(0, \frac{3n}{2} - 2\right)$ RLL constraints. A table of (d_{rll}, k_{rll}) constraint from Immink[Imm91, page 91] suggests code rates approaching unity (capacity $C > 0.99$) can be achieved with $d_{rll} = 0$ and $k_{rll} > 6$. A graph showing the maximum coderate as the number of bits of the encoded sequence are increased is shown in figure 4.3.

Applying the CITI codeword generator algorithm to an encoder with 5 data bits and 6 codeword bits (3 on each track) results in the LUT shown in table 4.1

The encoding procedure implemented in software takes 10 data bits and encodes 5 bits at a time using the look up table (LUT) shown in table 4.1. The codewords are then bit interleaved to form a larger 12 bit two track codeword (as the simulations are done

Input Data	Codeword/Next State State I	State O	Input Data	Codeword/Next State State I	State O
00000	$\overset{100}{100}/\mathrm{I}$	$\overset{000}{001}/\mathrm{I}$	10000	$\overset{001}{010}/\mathrm{I}$	$\overset{001}{010}/\mathrm{O}$
00001	$\overset{100}{111}/\mathrm{I}$	$\overset{000}{010}/\mathrm{I}$	10001	$\overset{100}{010}/\mathrm{I}$	$\overset{100}{010}/\mathrm{O}$
00010	$\overset{101}{011}/\mathrm{I}$	$\overset{000}{100}/\mathrm{I}$	10010	$\overset{111}{001}/\mathrm{I}$	$\overset{111}{010}/\mathrm{O}$
00011	$\overset{101}{101}/\mathrm{I}$	$\overset{001}{000}/\mathrm{I}$	10011	$\overset{110}{101}/\mathrm{I}$	$\overset{110}{011}/\mathrm{O}$
00100	$\overset{101}{110}/\mathrm{I}$	$\overset{010}{000}/\mathrm{I}$	10100	$\overset{011}{011}/\mathrm{I}$	$\overset{011}{110}/\mathrm{O}$
00101	$\overset{110}{110}/\mathrm{I}$	$\overset{100}{000}/\mathrm{I}$	10101	$\overset{110}{001}/\mathrm{O}$	$\overset{110}{001}/\mathrm{I}$
00110	$\overset{111}{100}/\mathrm{I}$	$\overset{111}{000}/\mathrm{I}$	10110	$\overset{111}{110}/\mathrm{O}$	$\overset{111}{011}/\mathrm{I}$
00111	$\overset{111}{111}/\mathrm{I}$	$\overset{000}{111}/\mathrm{I}$	10111	$\overset{101}{010}/\mathrm{O}$	$\overset{101}{010}/\mathrm{I}$
01000	$\overset{011}{101}/\mathrm{I}$	$\overset{101}{001}/\mathrm{I}$	11000	$\overset{110}{100}/\mathrm{O}$	$\overset{110}{010}/\mathrm{I}$
01001	$\overset{010}{010}/\mathrm{I}$	$\overset{001}{011}/\mathrm{I}$	11001	$\overset{010}{101}/\mathrm{O}$	$\overset{010}{101}/\mathrm{I}$
01010	$\overset{001}{111}/\mathrm{I}$	$\overset{001}{101}/\mathrm{I}$	11010	$\overset{011}{010}/\mathrm{O}$	$\overset{011}{001}/\mathrm{I}$
01011	$\overset{010}{001}/\mathrm{I}$	$\overset{010}{001}/\mathrm{O}$	11011	$\overset{001}{110}/\mathrm{O}$	$\overset{001}{110}/\mathrm{I}$
01100	$\overset{001}{001}/\mathrm{I}$	$\overset{010}{111}/\mathrm{O}$	11100	$\overset{010}{011}/\mathrm{O}$	$\overset{010}{110}/\mathrm{I}$
01101	$\overset{010}{100}/\mathrm{I}$	$\overset{010}{100}/\mathrm{O}$	11101	$\overset{011}{100}/\mathrm{O}$	$\overset{011}{100}/\mathrm{I}$
01110	$\overset{001}{100}/\mathrm{I}$	$\overset{001}{100}/\mathrm{O}$	11110	$\overset{101}{100}/\mathrm{O}$	$\overset{011}{111}/\mathrm{I}$
01111	$\overset{100}{001}/\mathrm{I}$	$\overset{100}{001}/\mathrm{O}$	11111	$\overset{100}{011}/\mathrm{O}$	$\overset{100}{011}/\mathrm{I}$

Table 4.1: Constructive ITI Encoder LUT

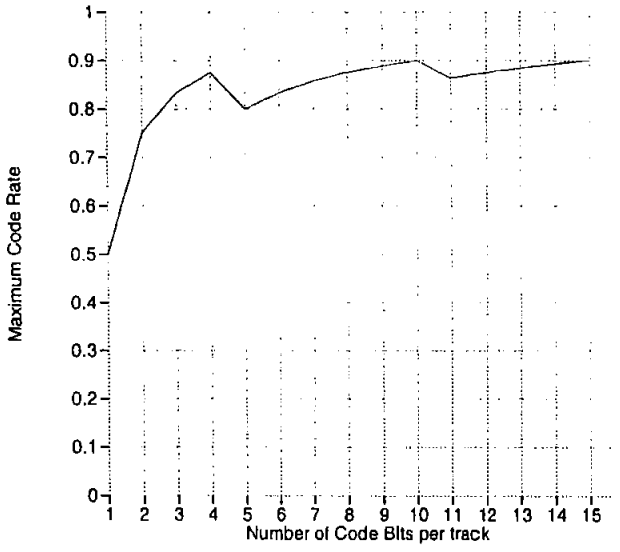


Figure 4.3: CITI Coderate vs. Codelength

for a PR4 Channel). The two separate input data streams have their unique in-phase and out-of-phase identifiers.

A STD describing the CITI constraints is shown in figure 4.4

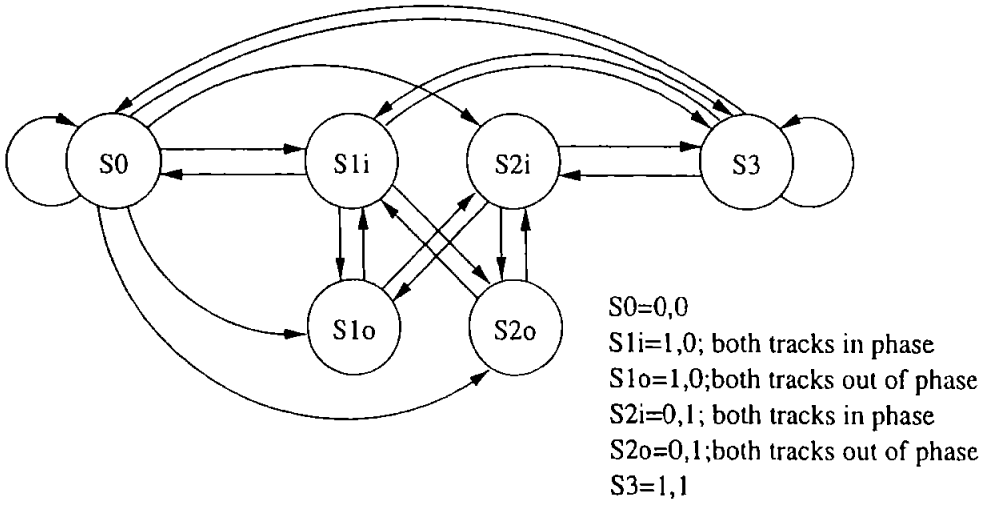


Figure 4.4: STD specifying only constructive adjacent track transitions

An STM that describes figure 4.4 is shown in equation 4.5.

$$STM = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \tag{4.5}$$

The largest real eigenvalue of this matrix is $\lambda_{max} = 3.82$, giving a code capacity of 0.967 (see the method of Faddeev[Jam99]). RLL constraints are not imposed. This CITI code capacity suggests that very high rate codes can be obtained which meet the Constructive ITI constraints. This is a significant improvement over the capacity of NITI constraints (see page 60). A code presented by the author at Intermag 2001[ADDC01] has a rate $\frac{5}{6} = 0.83$, and has RLL constraints of (0,7).

Algorithm 4.1 can be applied to multiple tracks and the maximum code rate achieved is shown in table 4.2.

Number of Tracks	Number of Bits per Track								
	2	3	4	5	6	7	8	9	10
2	0.75	0.83	0.87	0.8	0.83	0.85	0.87	0.88	0.9
3	0.66	0.77	0.83	0.86	0.83				
4	0.75	0.83	0.81						
5	0.80	0.80							
6	0.75								

Designed Multi-Track CITI Codes and their Rates

Table 4.2: CITI Code Rate for various number of tracks and bits per track

Table 4.2 also shows the most efficient CITI code found to date is a 2 track 10 bits per track code with rate 18/20. This code is approximately 93% efficient.

4.2.1 ITI Free Codes

A special sub-class of CITI codes can be constructed that do not "suffer" from ITI. These codes permit only the following states on adjacent tracks $\frac{0}{0}$, $\frac{1}{0}$ and $\frac{1}{1}$. That is one of the tracks does not have any transitions unless there is an adjacent constructive transition. Four states can be defined as

$S_0 = \frac{0}{0}$

$S_1 = \frac{1}{1}$

$S_2 = \frac{1}{0}$ with the next transition being out of phase

$S_3 = \frac{1}{0}$ with the next transition being in phase

The STM for these constraints would be

$$STM_{ITI\ free} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \tag{4.6}$$

This special class will not suffer from ITI, and will have $d^2_{min} = (1+\alpha^2)$ for $0 \leq \alpha \leq 1$ where α is the ITI. There are two reasons why these special codes are not considered any further

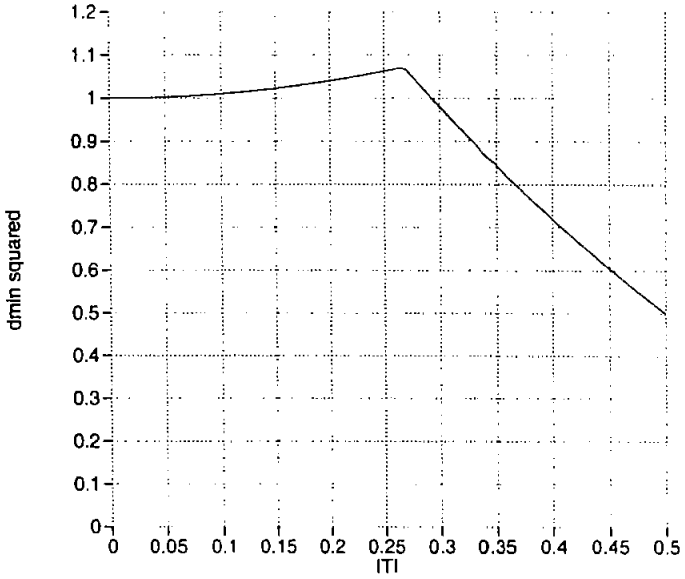
1. These codes would have a maximum rate (based on two track constraints only) of 0.636[†]
2. One track will have many more transitions than the other, resulting in very different RLL constraints on the two adjacent tracks.

For these reasons this special class is not considered further.

[†]the maximum eigenvalue of the STM (equation (4.6)) is 2.414

4.3 Results and Discussion

Soljanin[Sol94] discusses the use of multiple heads and multiple tracks, and shows the effects of ITI on d^2_{min}. For a two-track two-head system, d^2_{min} stays approximately constant upto an ITI of about 27% then degrades rapidly. This is shown in equation 4.7 and figure 4.5 and 4.6.



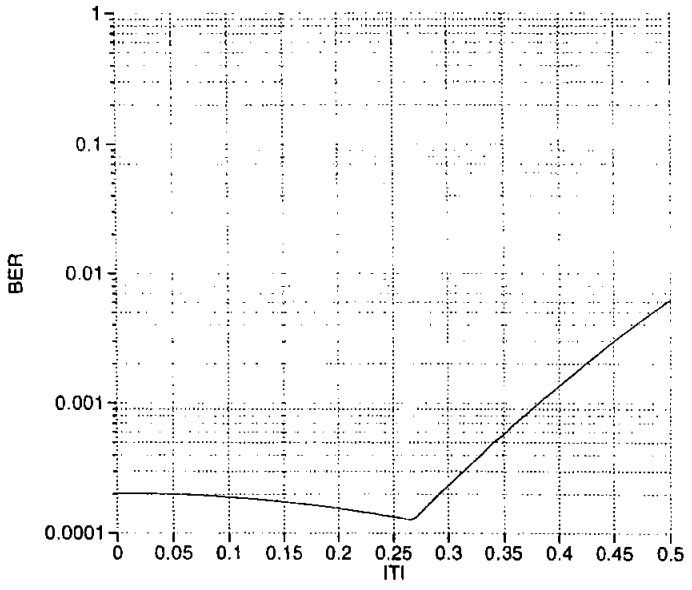
d^2_{min} vs. ITI, taken from [Sol94]

Figure 4.5: Ideal Variation of d^2_{min} with respect to ITI for a two-track two head system

$$d^2_{min} = \begin{cases} (1+\alpha^2) & \text{if} \quad 0 \leq \alpha < 2-\sqrt{3} \\ 2(1-\alpha)^2 & \text{if} \quad 2-\sqrt{3} \leq \alpha \leq 0.5 \end{cases} \tag{4.7}$$

The performance of the CITI code[ADDC01], NITI code[DDMD98] and a code that does not impose any two dimensional constraints (an RLL Code) using a Viterbi Detector is shown in figure 4.7. The results were obtained by computer simulations at 17dB SNR and a $PW_{50} = 2.0$ on a PC. Data were detected using the trellis shown in figure 2.10.

The new distance metric of the trellis (figure 2.10) for an input sample y_n, given



BER vs. ITI for an Arbitrary SNR

Figure 4.6: Ideal Variation of BER with respect to ITI for a two-track two head system

previous distance metrics of $s0_{n-1}$ and $s1_{n-1}$ is computed as

$$\begin{aligned} s0_n &= \min\left\{(y_n - 0)^2 + s0_{n-1}, (y_n + 1)^2 + s1_{n-1}\right\} \\ s1_n &= \min\left\{(y_n - 0)^2 + s1_{n-1}, (y_n - 1)^2 + s0_{n-1}\right\} \end{aligned}$$

Assuming $s0_{n-1}$ and $s1_{n-1}$ are zero, the distance metric is reduced to finding the minimum of

$$\begin{aligned} &\min\{(y_n)^2, \quad (y_n + 1)^2\} \\ &\min\{(y_n)^2, \quad (y_n - 1)^2\} \end{aligned} \tag{4.8}$$

In the presence of ITI $(y'_n\alpha)$[‡], the condition in equation 4.8 become

$$\begin{aligned} &\min\{(y_n \pm y'_n\alpha)^2, \quad (y_n \pm y'_n\alpha + 1)^2\} \\ &\min\{(y_n \pm y'_n\alpha)^2, \quad (y_n \pm y'_n\alpha - 1)^2\} \end{aligned} \tag{4.9}$$

[‡] y'_n is the adjacent track sample, and α is the amount of ITI

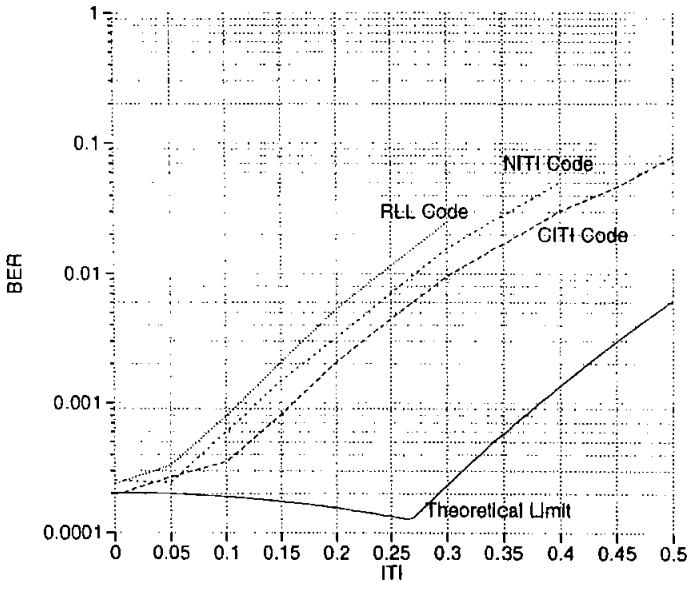


Figure 4.7: BER vs ITI using a normal Viterbi Detector

Analysing and comparing the conditions in equations 4.8 and equation 4.9 results in

$$\begin{matrix} \min\{0, \quad 1 + 2y_n\} \\ \min\{0, \quad 1 - 2y_n\} \end{matrix} \quad \text{from equation 4.8, and} \tag{4.10}$$

$$\begin{matrix} \min\{0, \quad 1 + 2y_n \pm 2y_n'\alpha\} \\ \min\{0, \quad 1 - 2y_n \mp 2y_n'\alpha\} \end{matrix} \quad \text{from equation 4.9} \tag{4.11}$$

Inspecting equation 4.10 and equation 4.11, the distance between each branch metric in a worst case situation is reduced by an amount of 2α. Using the relationship described in equation 2.18, the probability of error P_e can be upper bound using

$$P_e \leq \mathbf{Q}\left(\frac{1-2\alpha}{2\sigma}\right) \tag{4.12}$$

Equation 4.7 provides a lower bound, resulting in figure 4.8.

The upper limit shown describes the degradation in BER, and equations 4.8 to 4.11 describe the nature of the degradation. Table 4.3 lists the conditions of y_n and y_n' that

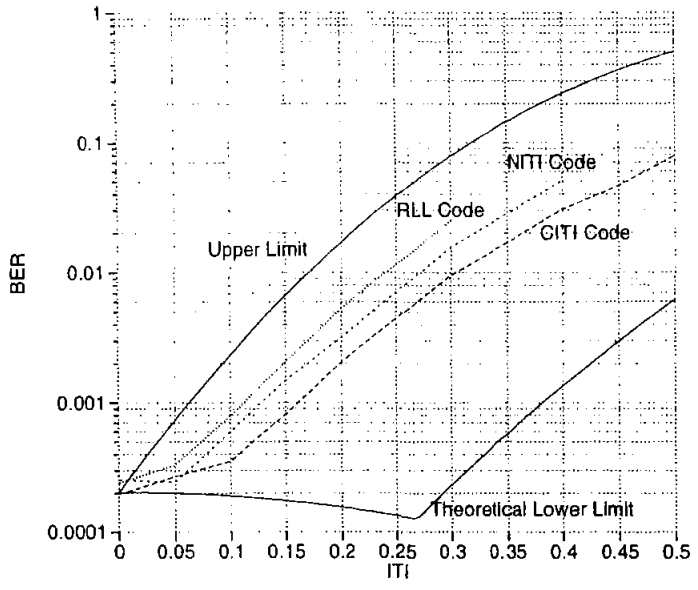


Figure 4.8: BER vs ITI using a normal Viterbi Detector showing Limits

will lead to the worst case situation.

Condition	On Track Sample	Adjacent Track Sample
1	$y_n = 0$	$y'_n = \pm 1$
2	$y_n = +1$	$y'_n = -1$
3	$y_n = -1$	$y'_n = +1$

Table 4.3: Conditions leading to the minimum distance for PR4 Two Track

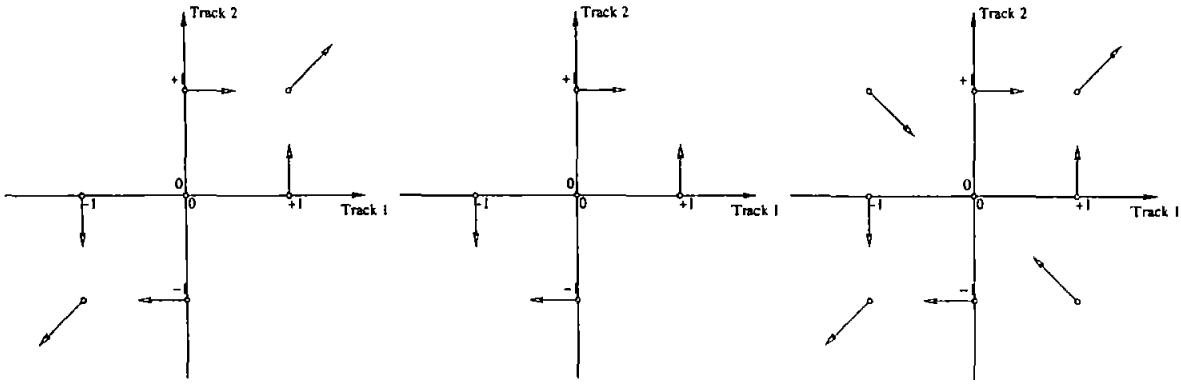
The RLL code suffers from all three conditions shown in table 4.3. The CITI and NITI code only experience the first condition. The RLL code and NITI code measured performance difference is due to the multiplicity of these events. The most likely error sequence is a "$\pm 1, \mp 1$" sequence on adjacent tracks. The CITI code does not suffer from this, however the "$\pm 1, 0$" is d_{min} away from the "$0, \pm 1$" sequence. The most likely error event is this, and the difference in performance of the three codes is entirely due to multiplicity.

The variation of path metrics with track condition {+1,0,-1} is shown in Table 4.4

y_n \ y'_n	CITI Code −1	0	+1	y_n \ y'_n	NITI Code −1	0	+1	y_n \ y'_n	RLL Code −1	0	+1
+1		1	1+2α	+1		1		+1	1−2α	1	1+2α
0	1−2α	1	1−2α	0	1−2α	1	1−2α	0	1−2α	1	1−2α
−1	1+2α	1		−1		1		−1	1+2α	1	1−2α

Table 4.4: Worst-case variation of branch metrics for different on-track and adjacent track conditions

The conditions described in Table 4.4 is shown graphically in Figure 4.9



(a) CITI Code (b) NITI Code (c) RLL Code

The white circles represent the possible combination of two-track data and the arrows show the effect of increasing ITI. As ITI increases the minimum separation between points can be described using equation (4.7)

Track 1 in this Figure is y_n in Table 4.4, and Track 2 is y'_n

Figure 4.9: Effect of ITI on Adjacent Track Samples

The CITI code performs better than the NITI and RLL codes using this detection scheme. It has about 5% more immunity to ITI compared to the NITI code, and about 10% more immunity compared to the RLL code. The measured performance is very different compared to the performance suggested by figure 4.6. This is due to ITI affecting the Viterbi metric computation as shown in figure 4.8.

A modified Viterbi detector with adaptive path metrics significantly improves the performance. This new Viterbi detector is discussed in the chapter on Adaptive Equalisation and Trellis Decoding.

4.4 Contributions

- A new 2 Dimensional RLL Coding scheme has been described. The new code imposes restrictions on the type of ITI that occurs between two adjacent tracks. Run Length Limited constraints for these new codes have also been presented.
- A sub-code of the new 2 Dimensional code has been discussed that will not degrade performance with increasing ITI.
- The capacity of 2 Dimensional Constraints has been evaluated for No ITI constraints, and Constructive ITI constraints.
- Performance of these new codes with respect to ITI has been obtained by computer simulations. These codes have been compared to, and found to be superior to previous published codes by Davey[DDMD98] in systems affected by ITI.
- Comparisons have been made with previous work done by Soljanin[Sol94], and some results have been verified.
- A paper has been published in the IEEE Transactions on Magnetics[ADDC01].

5 Adaptive Equalisation and Trellis Decoding

5.1 Adaptive Equalisation

In this chapter, an adaptive filter is used to estimate the ITI (α). This is developed further and incorporated into a trellis detector, resulting in a trellis that is very robust to ITI. The Least Mean Squared (LMS) algorithm was chosen to estimate α due to the simplicity of the algorithm. A filter with a tap coefficient based on α was also designed.

5.1.1 Least Mean Squared (LMS) Algorithm

The LMS algorithm was first described by Widrow[WJ60]. The algorithm minimises the instantaneous square of the error term (ideal filter output - actual filter output). A complete treatise on the algorithm can be found in Haykin[Hay84]. The main result of the algorithm can be stated as

$$\mathbf{h}_{n+1} = \mathbf{h}_n + \mu e \mathbf{u}_n \tag{5.1}$$

$$0 < \mu < \frac{2}{\text{Power in Filter}} \quad \text{see Haykin[Hay84]} \tag{5.2}$$

where $\mathbf{h}_{n+1}$ is a vector containing the new filter coefficients

$\mathbf{h}_n$ are the existing filter coefficients

$\mathbf{u}_n$ are the existing values in the filter

μ is a constant term

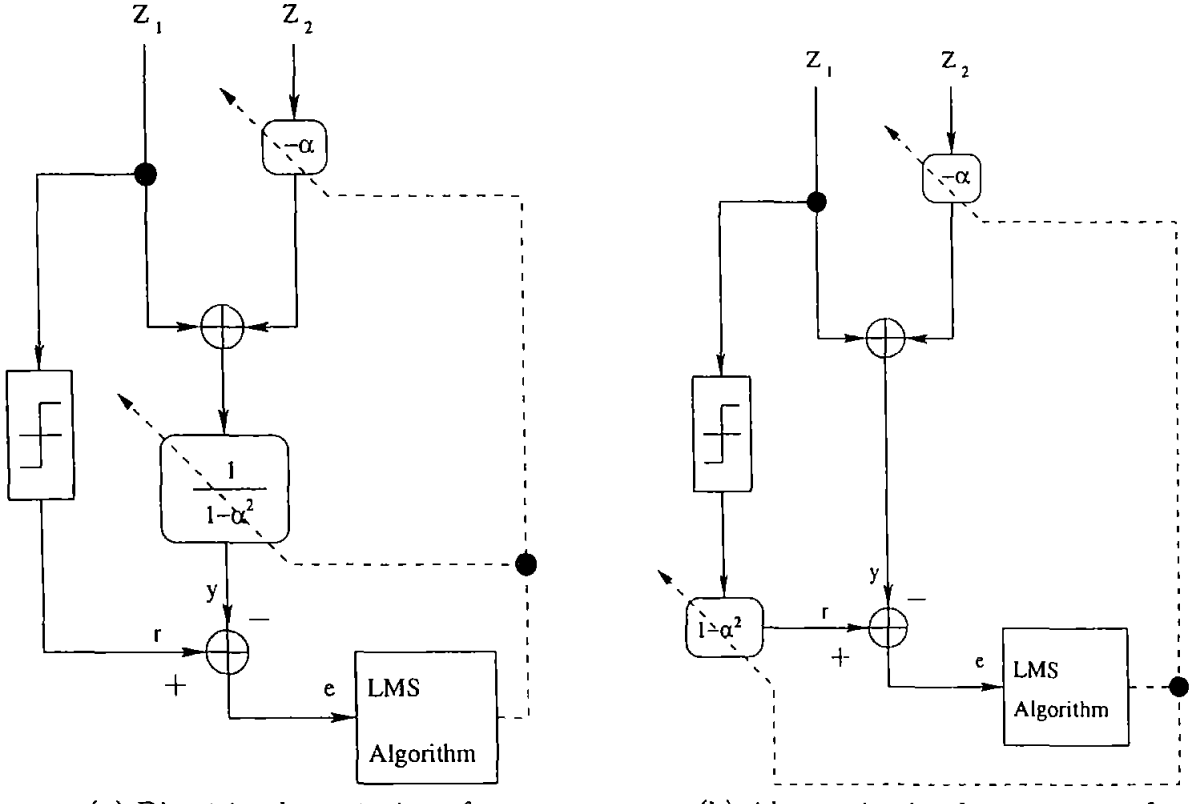
e is the instantaneous error term

ITI Estimation

Solving equation (4.1) for y_1 results in a solution to estimate the ITI-free ideal sample y_1:

$$\begin{aligned} Z_1 - \alpha Z_2 &= y_1 + \alpha y_2 + n1 - \alpha y_2 - \alpha^2 y_1 - \alpha n_2 \\ &= y_1(1-\alpha^2) + [n_1 - \alpha n_2] \qquad (5.3) \\ \therefore \quad y_1 &= \frac{Z_1 - \alpha Z_2 - [n_1 - \alpha n_2]}{1-\alpha^2} \qquad (5.4) \end{aligned}$$

The direct implementation of equation (5.4) (neglecting the noise terms) results in the equaliser shown in figure 5.1a. An ideal value 'r' is derived from thresholding the sample



(a) Direct implementation of equation (5.4)

(b) Alternative implementation of equation (5.4)

Figure 5.1: ITI Adaptive Equaliser

from track 1 (Z_1). For high SNR this thresholding does not lead to significant errors of the ideal response. The filter taps of figure 5.1 are $1(= h_{unity})$ and $-\alpha$. Applying the

LMS algorithm to these taps results in the following new taps

Tap 1

$$h_{unity} = 1 + \mu e Z_1 \tag{5.5}$$

Tap 2

$$\begin{aligned} -\alpha_{new} &= -\alpha_{old} + \mu e Z_2 \\ \Longrightarrow \alpha_{new} &= \alpha_{old} - \mu e Z_2 \end{aligned} \tag{5.6}$$

Normalising so that Tap 1 is unity,

equation (5.6) becomes

$$\alpha_{new} = \frac{\alpha_{old} - \mu e Z_2}{1 + \mu e Z_1} \tag{5.7}$$

Equation (5.7) is used to update the value of ITI (α) in the filter. A similar filter is used on track 2.

Evaluating the error term for this filter results in equation (5.8)

$$\begin{aligned} e_{aeq_{01}} &= y_1 - \frac{Z_1 - \alpha Z_2}{1 - \alpha^2} \\ &= \frac{[\alpha n_2 - n_1]}{1 - \alpha^2} \quad \text{using (5.3)} \end{aligned} \tag{5.8}$$

This error term would be Gaussian as n_1 and n_2 are Gaussian. The mean of the error term would be zero, however the standard deviation $\sigma_{e_{aeq_{01}}}$ would be given by equation (5.9) as long as n_1 and n_2 have zero covariance.

$$\sigma_{e_{aeq_{01}}} = \frac{\sqrt{\alpha^2 + 1}}{1 - \alpha^2} \tag{5.9}$$

The error term for the filter shown in figure 5.1b can be evaluated as

$$\begin{aligned} e_{aeq_{02}} &= y_1(1-\alpha^2) - Z_1 + \alpha Z_2 \\ &= [\alpha n_2 - n_1] \end{aligned} \tag{5.10}$$

Under the same assumptions made for equation (5.9), this error will have a standard deviation given by

$$\sigma_{e_{aeq_{02}}} = \sqrt{\alpha^2 + 1} \tag{5.11}$$

5.1.2 ITI Noise Amplification

An SNR measurement of $e_{aeq_{01}}$ and $e_{aeq_{02}}$ is shown in figure 5.2. In the figure, the bold line is the variation for $e_{aeq_{01}}$, and the dotted line represents the variation for $e_{aeq_{02}}$. The

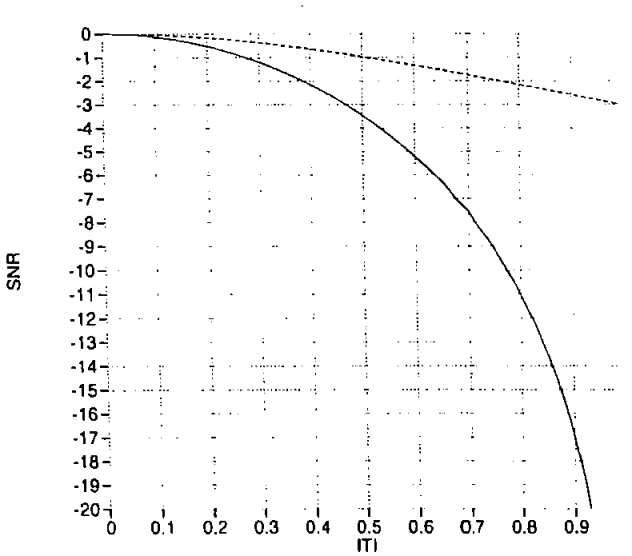


Figure 5.2: Variation in SNR due to ITI, as described in equation (5.12) and (5.13)

SNR is computed using equation (2.12) to give equation (5.12) and (5.13).

$$SNR_{aeq_{01}} = 10\log\left(\frac{(1-\alpha^2)^2}{2(1+\alpha^2)}\right) \tag{5.12}$$

$$SNR_{aeq_{02}} = 10\log\left(\frac{1}{2(1+\alpha^2)}\right) \tag{5.13}$$

The variation shown in figure 5.2 has been verified using computer simulations. It is clear that for high ITI, the ITI Adaptive Equaliser shown in figure 5.1b is better as the error term has smaller variance.

Two algorithms were developed by the author to estimate ITI, as follows.

Algorithm 5.1. ITI Estimation Algorithm 01

Starting with $\alpha = 0$ and a fixed μ, for each ontrack sample Z_1 and adjacent track sample Z_2 the following is performed

1. Obtain filter output
   - $y_1 = \frac{Z_1 - \alpha Z_2}{1-\alpha^2}$
   - $y_2 = \frac{Z_2 - \alpha Z_1}{1-\alpha^2}$
2. Threshold input samples Z_1 and Z_2 to obtain ideal samples r_1 and r_2.
3. Compute the error term
   - $e_1 = r_1 - y_1$
   - $e_2 = r_2 - y_2$
4. Update the value of α using the following
   - $\alpha_{new_1} = \frac{\alpha - \mu e_1 Z_2}{1+\mu e_1 Z_1}$
   - $\alpha_{new_2} = \frac{\alpha - \mu e_2 Z_1}{1+\mu e_2 Z_2}$
   - $\alpha = \frac{\alpha_{new_1} + \alpha_{new_2}}{2}$
5. If $\alpha > 1$ or $\alpha < 0$, $\alpha \leftarrow 0$

Algorithm 5.2. ITI Estimation Algorithm 02

Starting with $\alpha = 0$ and a fixed μ, for each ontrack sample Z_1 and adjacent track sample Z_2 the following is performed

1. Obtain filter output
   - $y_1 = Z_1 - \alpha Z_2$
   - $y_2 = Z_2 - \alpha Z_1$
2. Threshold input samples Z_1 and Z_2 to obtain ideal samples s_1 and s_2.
3. Obtain
   - $r_1 = s_1(1 - \alpha^2)$
   - $r_2 = s_2(1 - \alpha^2)$
4. Compute the error term
   - $e_1 = r_1 - y_1$
   - $e_2 = r_2 - y_2$
5. Update the value of α using the following
   - $\alpha_{new_1} = \frac{\alpha - \mu e_1 Z_2}{1 + \mu e_1 Z_1}$
   - $\alpha_{new_2} = \frac{\alpha - \mu e_2 Z_1}{1 + \mu e_2 Z_2}$
   - $\alpha = \frac{\alpha_{new_1} + \alpha_{new_2}}{2}$
6. If $\alpha > 1$ or $\alpha < 0$, $\alpha \leftarrow 0$

The maximum value of μ is bound by equation (5.2). For this particular case, the

maximum power in the filter can be approximately* bound by

$$\begin{aligned} P &= \left[\underbrace{(1+\alpha+3\sigma)}_{\max\{Z_1\}}\right]^2 + \left[\underbrace{(1+\alpha+3\sigma)}_{\max\{Z_2\}}\right]^2 \\ &= 2(1+\alpha+3\sigma)^2 \qquad (5.14) \\ 0 < \mu &< \frac{1}{(1+\alpha+3\sigma)^2} \qquad (5.15) \end{aligned}$$

The upper μ of equation (5.15) has been tabulated for different SNR and ITI in table 5.1. These values of μ provide a good upper limit, at a penalty of slower convergence.

SNR (dB)	σ	$\alpha = 0.0$	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
10	0.22	0.36	0.29	0.23	0.19	0.16	0.14
12	0.18	0.43	0.33	0.27	0.22	0.18	0.16
14	0.14	0.49	0.38	0.30	0.24	0.20	0.17
16	0.11	0.56	0.42	0.33	0.27	0.22	0.18
18	0.09	0.62	0.46	0.36	0.29	0.23	0.19
20	0.07	0.68	0.50	0.38	0.30	0.25	0.20
22	0.06	0.73	0.53	0.41	0.32	0.26	0.21

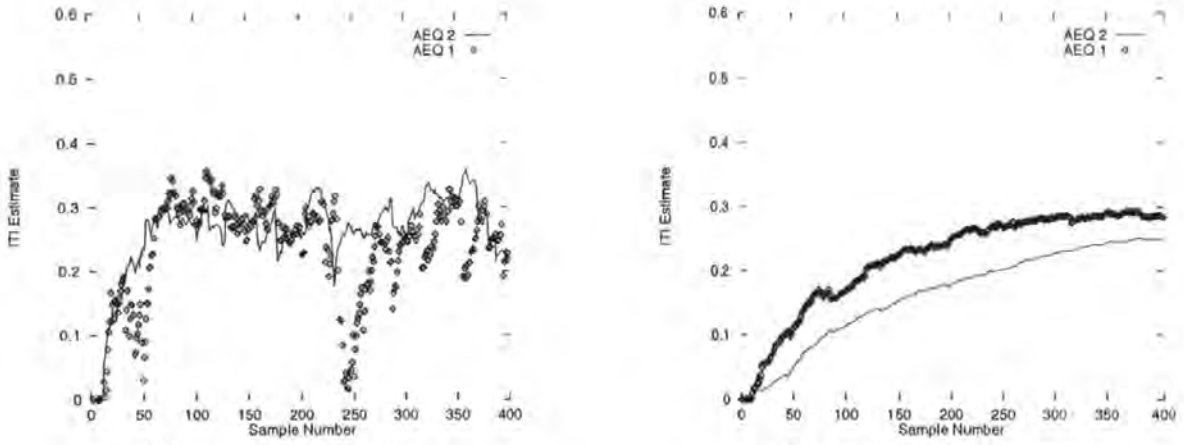
Table 5.1: Maximum μ for the AEQ

Bounding μ using the signal and noise power† results in a higher bound forμ which has been found by the author to be unstable especially at high α. In simulations, half the maximum μ was used as the μ is based on Gaussian probability and not very tight rules.

*for a Gaussian pdf, 99% of the samples will be within $\pm 3\sigma$. Thus $\max\{Z_1\} = \max\{Z_2\} \approx 1+\alpha+3\sigma$

†$\mu < \dfrac{1}{(1+\alpha)^2 + 2\sigma^2}$

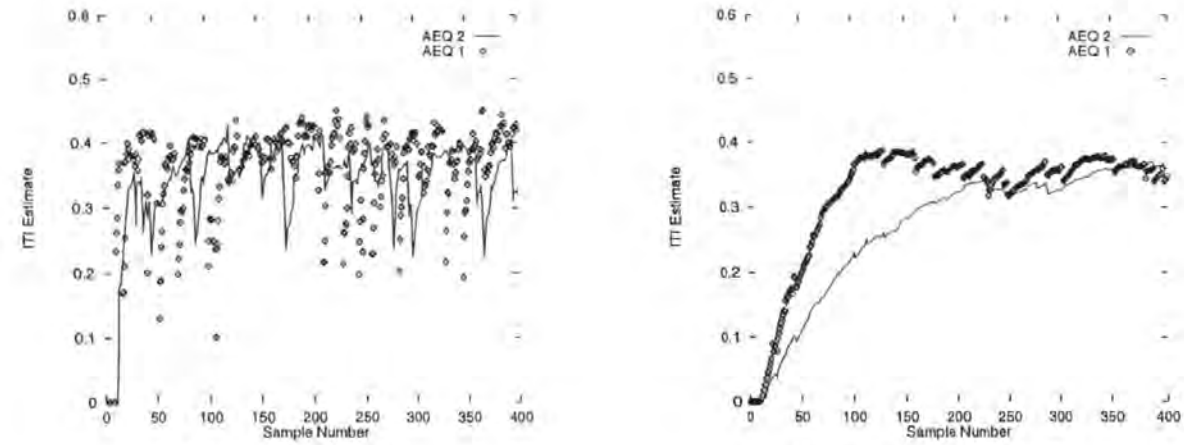
Algorithm 5.1 has been published by the author[ADDC02] as part of an *adaptive ITI trellis detector*. Typical responses for α are shown in figure 5.3 for both algorithms and SNRs of 20dB. The target α was 0.3. A smaller step size μ will result in a smaller variation in the estimated α and slower convergence as is clear from Figure 5.3 (a) and (b) respectively.



(a) using a μ of 0.1 (b) using a μ of 0.01

Figure 5.3: Estimated α for an SNR of 20dB, Ideal $\alpha = 0.3$

Estimating an α of 0.4, at an SNR of 25dB using both algorithms is shown in figure 5.4. μ is chosen to be 0.2 and 0.02, which is about half and a fiftieth of the maximum bound provided in table 5.1.



(a) using a μ of 0.2 (b) using a μ of 0.02

Figure 5.4: Estimated α for an SNR of 25dB, Ideal $\alpha = 0.4$

The performance of both algorithms is shown in figure 5.5. Both algorithms were tested using a SNR of 17dB. The second algorithm is clearly more resilient to changes in μ compared to the first algorithm. The improved performance of algorithm 2 is attributed

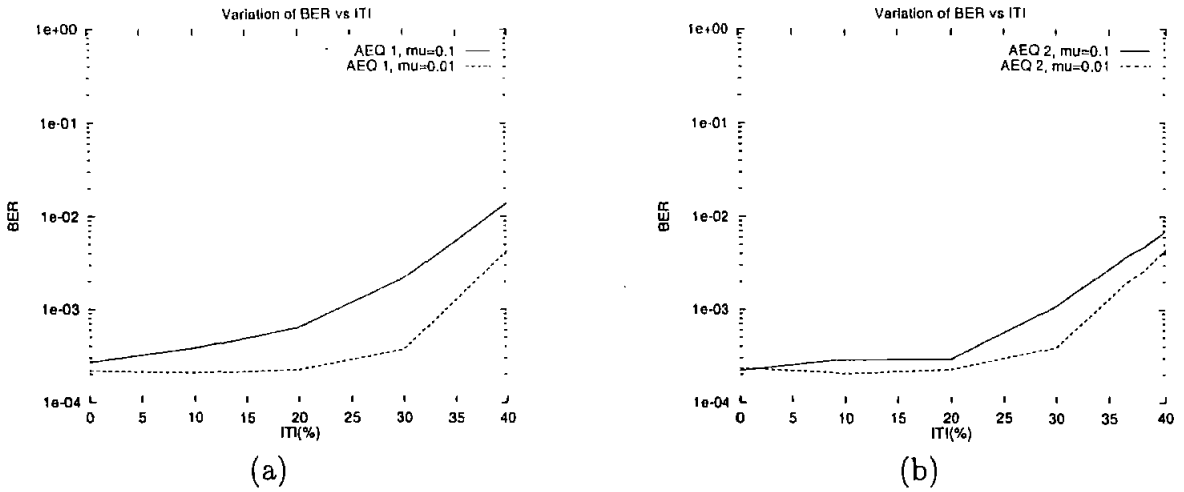


Figure 5.5: Performance of both algorithms with different μ

to a more stable estimate of α compared to the estimate made by algorithm 1. At small μ[‡], both algorithms performed relatively equally. In all results where α was estimated, Algorithm 2 was used by the author, unless stated otherwise.

[‡] $\mu < \frac{1}{10}$ the value shown in table 5.1

5.2 Trellis Decoding

Trellis decoding was originally introduced by Forney[Jr.73] for decoding convolutional codes. A new two track trellis for a PR4 magnetic channel is described in this section, based on the trellis described by Dolivio[DUH86]. The Adaptive ITI estimation is incorporated into this new trellis structure, resulting in a new trellis with immunity to ITI. The performance and limitations of this trellis are discussed, and comparisons made in using it with existing codes and new CITI codes.

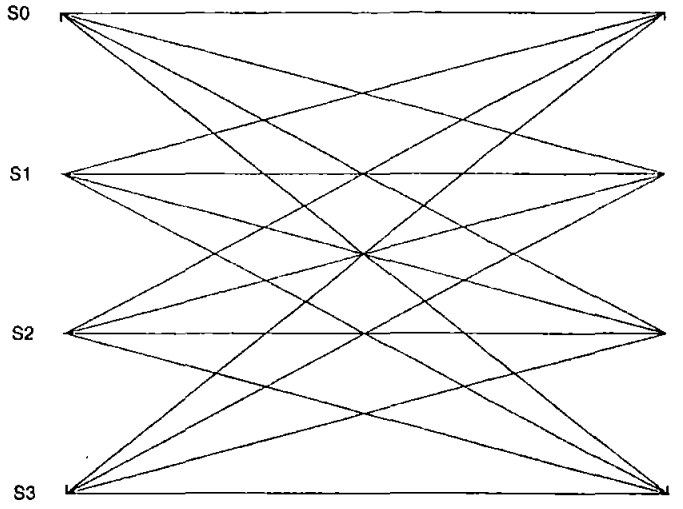
5.2.1 Multi-Track Trellis

A new trellis is constructed using the existing states of the single-track Viterbi Detectors (labelled as s0 and s1 in figure 2.10) and combining them using table 5.2.

New State	Single Track State of Track 1	Single Track State of Track 2
S0	s0	s0
S1	s0	s1
S2	s1	s0
S3	s1	s1

Table 5.2: Combination of 2 trellis (figure 2.10) states resulting in new states (figure 5.6)

For example, going from S0 to S1 means track 1 will stay in state s0 and track 2 will change from s0 to s1. Consequently the ideal path metrics will be 0 and -1 for track 1 and track 2 respectively. Other methods of obtaining trellis products are described in[HM97]. The new trellis is shown in figure 5.6.



Start State	End State	ITI Affected PM (Tr 1,Tr 2)	Ideal PM (Tr 1,Tr 2)
S0	S0	$0,0$	$0,0$
S0	S1	$-\alpha,-1$	$0,-1$
S0	S2	$-1,-\alpha$	$-1,0$
S0	S3	$-1-\alpha,-1-\alpha$	$-1,-1$
S1	S0	$\alpha,1$	$0,1$
S1	S1	$0,0$	$0,0$
S1	S2	$-1+\alpha,1-\alpha$	$-1,1$
S1	S3	$-1,-\alpha$	$-1,0$
S2	S0	$1,\alpha$	$1,0$
S2	S1	$1-\alpha,-1+\alpha$	$1,-1$
S2	S2	$0,0$	$0,0$
S2	S3	$-\alpha,-1$	$0,-1$
S3	S0	$1+\alpha,1+\alpha$	$1,1$
S3	S1	$1,\alpha$	$1,0$
S3	S2	$\alpha,1$	$0,1$
S3	S3	$0,0$	$0,0$

Figure 5.6: New Interleaved PR4 2-Track Trellis Section.

Combining two conventional PR4 detection trellis, yields a trellis with Ideal Path Metrics (PM). The effects of ITI can then be imposed on the ideal PM to obtain the ITI-affected path metrics using equation (4.1). For the previous example this would mean the path metrics would become $-\alpha$ and -1[§].

The equations used to perform MLSD on this trellis are given in equations (5.16) to

[§]using (4.1) and assuming ideal conditions of $n_1 = 0$, and $n_2 = 0$

(5.19) assuming the current state is at n, the next state is at $n+1$ and the samples from the track are Z_1 and Z_2.

$$S0_{n+1} = \min \left\{ \begin{array}{l} S0_n + (Z_1)^2 + (Z_2)^2 \\ S1_n + (Z_1 - \alpha)^2 + (Z_2 - 1)^2 \\ S2_n + (Z_1 - 1)^2 + (Z_2 - \alpha)^2 \\ S3_n + (Z_1 - (1+\alpha))^2 + (Z_2 - (1+\alpha))^2 \end{array} \right\} \tag{5.16}$$

$$S1_{n+1} = \min \left\{ \begin{array}{l} S0_n + (Z_1 - (-\alpha))^2 + (Z_2 - (-1))^2 \\ S1_n + (Z_1)^2 + (Z_2)^2 \\ S2_n + (Z_1 - (1-\alpha))^2 + (Z_2 - (-1+\alpha))^2 \\ S3_n + (Z_1 - 1)^2 + (Z_2 - \alpha)^2 \end{array} \right\} \tag{5.17}$$

$$S2_{n+1} = \min \left\{ \begin{array}{l} S0_n + (Z_1 - (-1))^2 + (Z_2 - (-\alpha))^2 \\ S1_n + (Z_1 - (-1+\alpha))^2 + (Z_2 - (1-\alpha))^2 \\ S2_n + (Z_1)^2 + (Z_2)^2 \\ S3_n + (Z_1 - \alpha)^2 + (Z_2 - 1)^2 \end{array} \right\} \tag{5.18}$$

$$S3_{n+1} = \min \left\{ \begin{array}{l} S0_n + (Z_1 - (-1-\alpha))^2 + (Z_2 - (-1-\alpha))^2 \\ S1_n + (Z_1 - (-1))^2 + (Z_2 - (-\alpha))^2 \\ S2_n + (Z_1 - (-\alpha))^2 + (Z_2 - (-1))^2 \\ S3_n + (Z_1)^2 + (Z_2)^2 \end{array} \right\} \tag{5.19}$$

Analysing this trellis, the minimum square of the euclidean distance is described by equation (4.7).

The minimum distance squared for the new trellis is as shown in figure 5.7 for all combinations of y_1 and y_2 except

1. $y_1, y_2 = +1, +1$

2. $y_1, y_2 = -1, -1$

Thus **only** when Constructive ITI exists, equation (4.7) is not valid. In this case, d^2_{min} is given by **only** the first part of equation (4.7). This suggests that two adjacent identical tracks in the presence of ITI will have a slightly higher d_{min}.

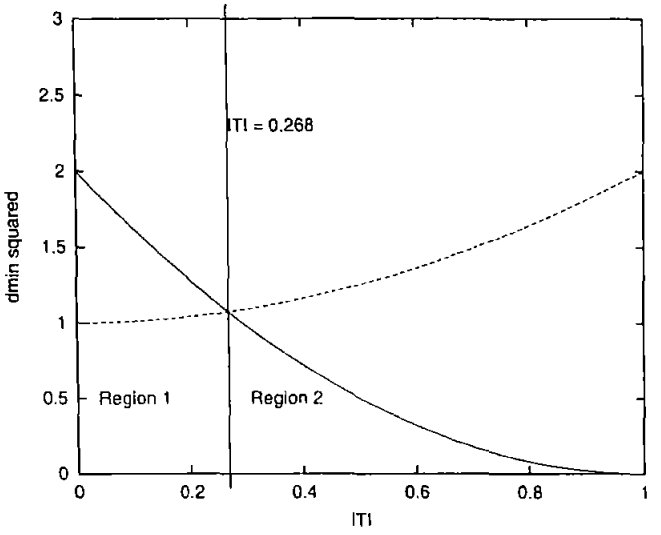


Figure 5.7: Variation of d^2_{min} with ITI for new trellis

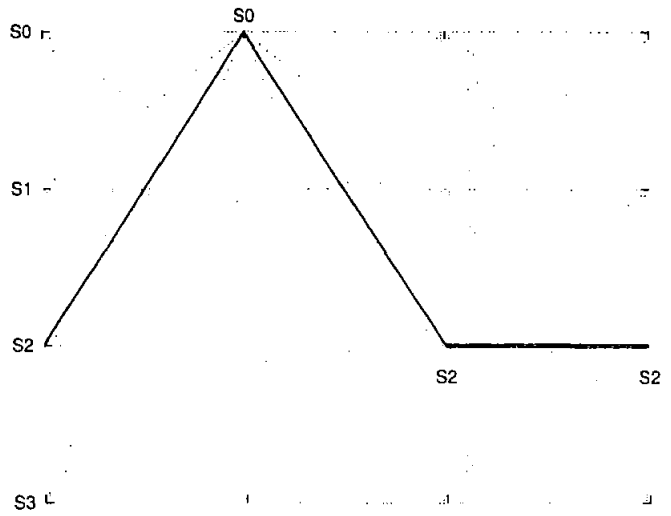
From figure 5.7, there are two separate variations in d^2_{min} that lead to the overall d^2_{min}. The first region results in an improvement in d^2_{min}, whereas the second degrades the d^2_{min}. This explains the flat regions of the measured performance in figure 5.10. The observations and results using this trellis agree with previous work done by Soljanin[Sol94] using sliding block decoders. Sequences leading to d^2_{min} on the trellis are shown in table 5.3. There are different dominant sequences that lead to errors depending on the amount of ITI. In region 1, the sequence $+-$ or $-+$ on either track results in the d^2_{min} event, and in region 2 the recuring theme is a two track destructive ITI pattern $\begin{smallmatrix}+ & - \\ - & +\end{smallmatrix}$. The CITI code

Pattern	Region (figure 5.7)	Error Event Trellis State Sequences
$\begin{smallmatrix}+&-\\0&0\end{smallmatrix}$	1	202, 313
$\begin{smallmatrix}+&0&-\\0&-&+\end{smallmatrix}$	1	2012
$\begin{smallmatrix}+&-&0\\0&-&+\end{smallmatrix}$	1	2032
$\begin{smallmatrix}+&-&0\\-&0&+\end{smallmatrix}$	1	2132
$\begin{smallmatrix}+&-\\-&+\end{smallmatrix}$	2	212
$\begin{smallmatrix}-&+\\+&-\end{smallmatrix}$	2	121

A + sign denotes a positive transition, a − sign denotes a negative transition and 0 denotes no transition. For clarity, leading and trailing zeros have not been tabulated.

Table 5.3: Some events leading to d^2_{min} for the new trellis

eliminates the shortest region 2 error event. The sequence 2022 in table 5.3 is shown in figure 5.8



Sequence 2022 *in table 5.3 is equivalent to the path in the trellis* S2 − S0 − S2 − S2

Figure 5.8: Error Sequence '2022'

5.3 Combined Trellis & Adaptive Equaliser

The new detection scheme uses the adaptive equaliser and trellis described earlier as shown in figure 5.9. Data are equalised and differentiated into odd and even interleaves. The odd and even samples from the adjacent tracks are detected simultaneously. Adaptive equalisers are used to estimate the ITI from adjacent samples, and the ITI estimate is supplied to the trellis. Maximum Likelihood Sequence Detection using the Viterbi algorithm applied to the trellis[Jr.73] is then performed to obtain the user data. The recovered samples are then recombined to obtain the individual track data.

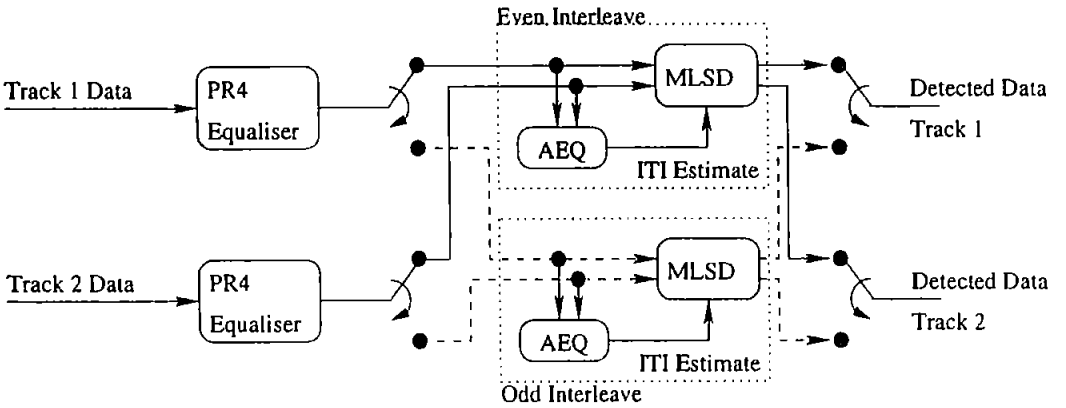
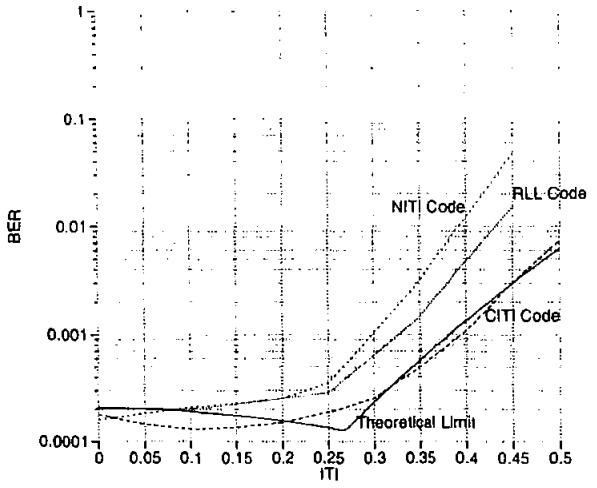


Figure 5.9: New Detection Scheme

5.4 Results and Discussion

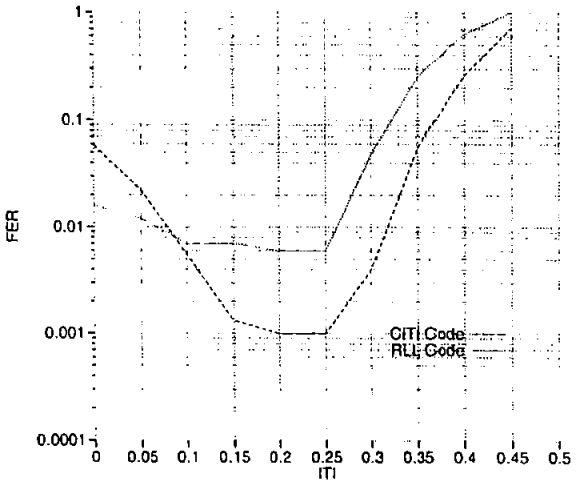
Results using this new trellis are shown in figure 5.10.



The theoretical limit is obtained using equation (4.7) and is for comparison purpose only, to show that the CITI code follows the general trend described by the equations

Figure 5.10: BER vs ITI using new Adaptive Trellis Scheme

This new scheme has been presented by the author at Intermag 2002[ADDC02]. The new CITI code is about 10% more resilient to ITI compared to the NITI and RLL codes, when compared using Bit Error Rates. In terms of Frame Error Rate, the CITI code is about 5% more resilient compared to the RLL code at high ITI. The improvement in FER performance is much more pronounced using the CITI codes. As the ITI equaliser is adaptive, it can cope with varying amounts of ITI. The equaliser trains itself using data samples, and no training sequence is required.



FER performance obtained for an SNR of 22dB; a PW_{50}/T of 2.0 and 4096 user data bits

Figure 5.11: FER vs ITI using new Adaptive Trellis Scheme

Using the number of states and the number of paths in a trellis as a measure of complexity, table 5.4 can be created. Table 5.4 does not take into consideration the

Decoding Using New Method

Number of Tracks	Number of States	Number of Paths	Number of Path Metrics per Path
1	2	4	1
2	4	16	2
3	8	64	3
⋮	⋮	⋮	⋮
k	2^k	2^{2k}	k

Individual Track Decoding

Number of Tracks	Number of States	Number of Paths	Number of Path Metrics per Path
k	$2k$	$4k$	1

Table 5.4: Complexity of k-Track Trellis for a PR4 Channel

additional complexity of the AEQ[¶]. It is clear from table that for many tracks, the relative complexity of the new method can get astronomical. For a few tracks, this new method is realistic.

[¶]the AEQ will be more complicated for more tracks

Results in figure 5.12 show negligible (about 0.5dB) deterioration in SNR for a fixed Bit Error Rate (BER) for ITI up to 30%. The results are for a normal code that does not have any ITI constraints.

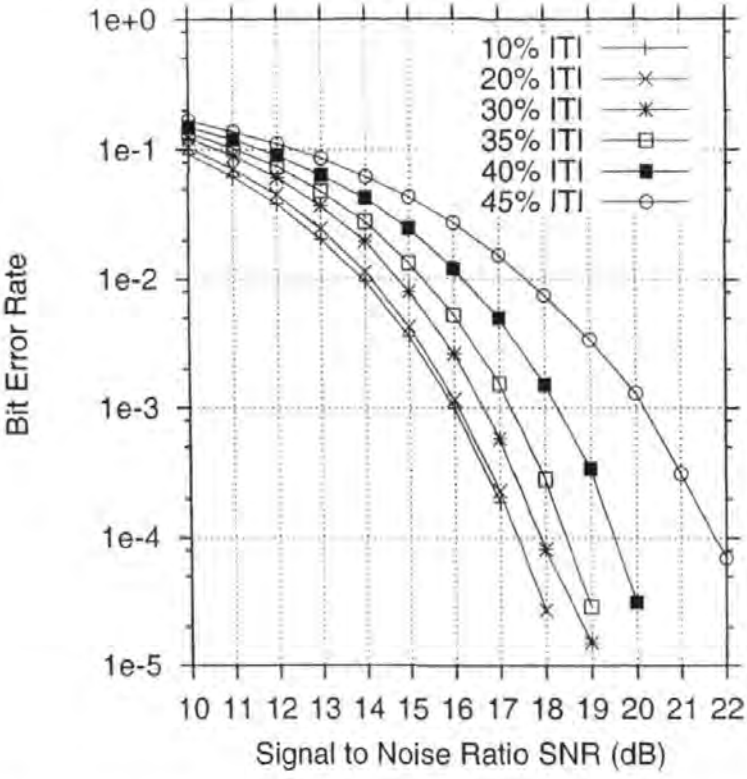


Figure 5.12: SNR vs BER for a PR4 Channel Using New MLSD Decoder

There is a loss of about 2dB SNR for 40% ITI at a PW_{50}/T of 2.0. Previous techniques experience about 7dB loss in SNR at 40% ITI and similar PW_{50}/T[RML99][‖].

The most common error event is when there are opposite transitions on adjacent tracks. Using the CITI code to eliminate this results in a further 1dB SNR gain at 40% ITI.

[‖]This is discussed in the paper by the author[ADDC02]

5.5 Contributions

- A new trellis for simultaneously decoding 2 PR4 tracks is presented.
- An adaptive algorithm has also been developed which estimates the amount of ITI.
- Comparisons have been made, some results published by Soljanin have been verified[Sol94].
- A paper has been published in the IEEE Transactions on Magnetics[ADDC02].

6 Conclusions and Discussions

Magnetic recording densities are expected to continue rising. Existing track separation is expected to reduce, resulting in interference between adjacent tracks. Inter-track interference has been highlighted as one of the dominant sources of noise in very high density recording. The aim of the research was to investigate, and find solutions to inter-track interference (ITI). A target model using a two-track PR4 magnetic channel was investigated using computer simulations; with software written using C++. Digital filters were designed to equalise a Lorentzian pulse to a PR4 pulse. Linear and symmetrical ITI between two adjacent tracks was simulated, and it was assumed that there was no interaction between groups of tracks. Several modulation and error correction codes were simulated, and bit error rates (BER) were measured.

Previous work by Davey[DDMD98, Dav94] was studied, and extended. A new class of two dimensional codes that are more resilient to ITI (named CITI codes) were evaluated and published as a result of the investigation[ADDC01]. The codes are based on controlling the location of transitions on adjacent tracks, and impose constraints that do not allow certain adjacent track conditions from occurring. Previous work on coding solutions to crosstalk by Lee[LM94], Bane[BO97] and Davey[DDMD98] imposed a very strict constraint, basically no (simultaneous) adjacent track transitions were allowed. It was found by the author that some adjacent track transitions resulted in an increase in the readback signal and codes that enforced this were dubbed Constructive ITI codes. The constraints imposed by the recording codes were modified to reflect this, and higher recording code rates could be achieved (upto 0.97, compared to 0.79), leading to higher recording densities. These codes were compared with some existing codes and were found

to outperform existing codes in the presence of ITI. The existing codes compared were a single-track code and a two-dimensional code. Study of CITI codes suggested the euclidean distance variation between different two track conditions with increasing ITI could be exploited to create a subset of CITI codes that would be completely immune to ITI. This subset imposes constraints that result in a maximum code rate of 0.64 and an unequal number of transitions on adjacent tracks. This renders it impractical, as almost all the timing information would be derived from a single track, and this special subset was not pursued further.

The second stage of the investigation focused on the detection scheme. A new multi-track trellis with adaptive path metrics was designed and published[ADDC02]. Data from both tracks is decoded simultaneously, using a trellis that has provision for the effect of ITI in its branches. This requirement led to the development of algorithms to determine ITI from the adjacent track data samples. The new detection scheme resulted in large gains in performance in the presence of ITI, and the performance was very similar to previous work done by Soljanin using sliding block decoders[SG94, Sol94]. The scheme proposed by the author utilised the Viterbi algorithm on the new trellis, and results in the best frame error rate performance. The trellis structure is not very complicated, and a MAP decoding using the BCJR algorithm[BCJR74] would give the best bit error rate performance. As the number of tracks increases, the complexity of the new trellis would render its use prohibitive. It was observed that there were two distinct regions that had different error events. For low* ITI, the dominant error event was found to be the single-track dominant error pattern whereas for higher ITI, the dominant error pattern was different. The dominant error pattern at high ITI was dependent on adjacent track data patterns. The new class of codes were more resilient to ITI (by about 10%) using the new trellis compared to previous codes.

The final stage of the investigation looked at high rate LDPC codes in order to improve

*ITI < 27%

the performance of conventional recording codes. A paper containing some results of this stage has been accepted by the IEEE Intermag 2003 conference[ADDC03]. The new trellis was employed and the BCJR algorithm was implemented on it. LDPC codes were serially concatenated with the magnetic channel, and the overall scheme was iteratively decoded. LDPC codes were decoded using the Sum-Product algorithm. Investigations into different LDPC's showed that cycles of length 4 were generally unavoidable, and an LDPC j constraint of 3 resulted in the best performance. Specific constructions that resulted in cycle 4 free matrices were made based on Kirkman Triple Systems (KTS), however in the error region of interest (10^{-3} to 10^{-4}) and the code rate of interest (≥ 0.9)the gain from using KTS constructions was not much (< 0.2dB). The encoding complexity of KTS matrices was also much greater compared to the systematic matrix of the author. KTS constructions are also much more restrictive in possible code rates compared to the authors LDPC codes. The iterative decoding scheme did have an improvement as ITI increased, however in the case when there was no ITI, the performance was worse than Viterbi decoding. This observation can be understood as iterative decoding is an approximation (to MAP decoding). The use of LDPC codes can improve the performance of conventional codes to that observed for the new CITI codes only in the presence of some ITI. The serially concatenated scheme was also much more complicated than the CITI code. There are convergence issues with the Sum-Product decoder at high ITI that are still open problems.

The work done by the author has led to the following contributions to knowledge in the field of multi-track magnetic recording

Main Contributions to Knowledge

- A new class of Two-Dimensional Modulation codes were introduced and published[ADDC01
- A new Adaptive Multi-Track Trellis was introduced[ADDC02].

Further salient points are

- A new method of iteratively decoding multiple tracks has been presented.
- LDPC codes have been investigated at a very high rate (rates ≥ 0.9) in the presence of ITI, and it has been found from experiments that LDPC codes with $j = 3$ provide the best ensemble of LDPC's for a two track PR4 channel experiencing linear and symmetrical crosstalk.
- The performance of LDPC's based on the incidence matrix of Kirkman Triple System have been compared to the performance of LDPC's suggested by the author.
- A sub-code of the new two dimensional code has been discussed that will not degrade performance with increasing ITI.
- The capacity of two dimensional Constraints has been evaluated for different ITI constraints.
- Comparisons have been made with previous work done by Soljanin[Sol94], and some results have been verified.
- Adaptive algorithms have been developed that estimate the amount of ITI.
- A journal paper on multi-track recording has been reviewed for the IEEE Transactions on Magnetics.

6.1 Future Work

During the course of the work presented in this thesis the following questions/directions still remain unsolved, and are described as potential future work.

- Inclusion of jitter and a micro track model. This is to model the uncertainty in the readback clock information and the provision of granularity in the magnetic media. As linear densities are increased, jitter and zigzag transitions (media noise) are dominant noise sources.

- Extension of the new decoding schemes to more tracks and more dimensions (for example in multi-layer discs). As the number of tracks is increased, the trellis described in this thesis would increasingly become complicated, and other schemes like PR equalisation across tracks/dimensions may offer a solution.

- The convergence issues raised in the ECC chapter.

- Extending the CITI concept to channels with different response like the EPR4 channel. The CITI code is based on the ability to control transitions on the media. A PR4 channel can be viewed as two interleaved dicode channels, and in dicode channels transitions on the media occur where there is a 1 in the data sequence. In other PR responses, this is not the case and further research into this would be valuable.

- Modifying the new adaptive trellis to include the effect of Off-Track Interference (OTI). The adaptive equaliser used to estimate the ITI could also be modified and used to estimate off-track interference.

- Investigation into non-binary LDPC decoding and other SISO algorithms like the Kotter Vardy algebraic decoding algorithm[KV00a][KV00b]. The future of ECC seems to be in designing codes that can be decoded well using SISO and iterative

decoding schemes. Techniques that investigate the flow of extrinsic information (like density evolution) could provide a new way of designing ECC codes.

- New media (like patterned or perpendicular media) result in different wave shapes and different constraint as that imposed on the longitudinal channel. This could be exploited and more powerful recording codes could be designed.

Papers Published

Intermag 2001 Conference Paper

RLL Coded PR(1,2,3,4,3,2,1)ML System for Double Layer Perpendicular Magnetic Recording

This paper was presented by Dr. Kurihara. The author suggested the codes that were used.

RLL Coded PR(1,2,3,4,3,2,1)ML System for Double Layer Perpendicular Magnetic Recording

Y. Kurihara, H. Osawa, Y. Okamoto, P. J. Davey, D. J. Mapps, M. Z. Ahmed, and T. Donnelly

Abstract— **This paper describes a new high order PR(1,2,3,4,3,2,1)ML system together with a 2/3 rate (1,6) run-length limited (RLL) code for double layer perpendicular magnetic recording. The computer simulation employs an arctangent function to represent the isolated reproducing waveform for each recorded transition. Bit error rate performance is compared to conventional perpendicular partial response maximum likelihood (PRML) systems over a range of user bit densities. The result shows that the proposed PR(1,2,3,4,3,2,1)ML system exhibits excellent performance about 13.4dB over the conventional schemes in high density recording.**

Index Terms— **Double layer perpendicular magnetic recording, MR head, PRML system, RLL code.**

I. Introduction

DOUBLE layer perpendicular magnetic recording is regarded as the prime candidate to achieve the recording density of one tera-bit per square inch [1]. For magneto-resistive (MR) readback, unlike longitudinal magnetic recording, the isolated reproducing waveform produces a step-function-like response for each recorded transition [2]. Signal processing technologies such as partial response maximum likelihood (PRML) systems, especially 8/9 or 16/17 coded PR4ML or EPR4ML systems, have played significant roles for increasing linear density for longitudinal magnetic recording. However, for double layer perpendicular magnetic recording, such detection systems prove to be disadvantaged due to characteristic differences in the read-back signal. It has been shown for perpendicular recording that (1,7) run-length limited (RLL) coded systems employing PR(1,2,2,1)ML and PR(1,2,3,2,1)ML yield improved performance over conventional longitudinal PRML schemes [2].

This paper reports on the results of a new high order PR(1,2,3,4,3,2,1)ML scheme together with a 2/3 rate (1,6) RLL code [3] as well as (1,7) RLL code. The computer simulation employs an arctangent function to represent the isolated reproducing waveform for the double layer perpendicular system. Bit error rate (BER) performance is compared to conventional perpendicular PRML systems over a range of user bit densities.

Manuscript received October 10, 2000.

Yoshitake Kurihara is visiting C.R.I.S.T., University of Plymouth, Plymouth, PL4 8AA U.K. (e-mail: Y.Kurihara@plymouth.ac.uk), on leave from Niihama National College of Technology, Niihama, 792–8580 Japan (telephone: +81–897–37–7765, e-mail: kurihara@ect.niihama-nct.ac.jp).

Hisashi Osawa and Yoshihiro Okamoto are with the Faculty of Engineering, Ehime University, Matsuyama, 790-8577 Japan (e-mail: {osawa, okamoto}@rec.ee.ehime-u.ac.jp).

Paul. J. Davey, Des J. Mapps, Mohammed Z. Ahmed, and Terry Donnelly are with the Centre for Research in Information Storage Technology, University of Plymouth, Plymouth, PL4 8AA United Kingdom (telephone: +44–1752–23 {2516, 2570}, e-mail: {pdavey, dmapps, tdonnelly}@plymouth.ac.uk).

II. Simulation Model

The block diagram of a PRML system for double layer perpendicular magnetic recording is shown in Fig. 1, $\{a_{k'}\}$ is the input data sequence at time $k'T_\mathrm{b}$, and $\{b_k\}$ is the channel encoded sequence at time kT_s, where T_b is the user bit interval and T_s the symbol (channel bit) interval. Precoded sequence $\{c_k\}$ is non-return to zero (NRZ) recorded. Additive White Gaussian Noise (AWGN) is added at the read stage to simulate the double layered perpendicular media noise. The equalized signal sequence $\{d_k\}$, at the discrimination point, is the input sequence of Viterbi detector. The output of Viterbi detector $\{\hat{b}_k\}$ is the detected channel code sequence, and $\{\hat{a}_{k'}\}$ is the output data sequence. The bit error rate (BER) of $\hat{a}_{k'}$ for $a_{k'}$ is obtained by computer simulation based on this block diagram.

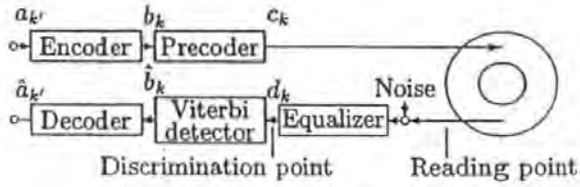


Fig. 1. Block diagram.

III. Channel Coding

The 2/3 rate (1,6) RLL code belongs to the family of the run-length limited code similar to the well known 2/3 rate (1,7) RLL code. The (1,6) RLL code has the same coding rate $T_\mathrm{s}/T_\mathrm{b} = 2/3$, and the same minimum constraint $d = 1$ as the (1,7) RLL code. Employing these codes to the software model, illustrated in Fig. 1, the data sequence $\{a_{k'}\}$ is encoded into RLL coded sequence $\{b_k\}$. After the detection of the equalized sequence $\{d_k\}$, the output sequence of the Viterbi detector $\{\hat{b}_k\}$ is decoded into $\{\hat{a}_{k'}\}$ corresponding to the original $\{a_{k'}\}$.

In Fig. 1, precoded $\{c_k\}$ is related to $\{b_k\}$ by

$$c_k = b_k + c_{k-1} \pmod 2, \tag{1}$$

hence the consecutive run length of 0's or 1's in $\{c_k\}$ is limited to at least 2. This constraint can simplify the Viterbi detector.

IV. Double Layer Perpendicular Magnetic Recording

For MR readback, unlike longitudinal magnetic recording, the isolated reproducing waveform is a step-function-like for each recorded transition. Therefore assume that the isolated reproducing waveform at the reading point is given by

$$h(t) = \frac{2A}{\pi} \tan^{-1} \frac{2t}{T_{50}}, \tag{2}$$

as shown in Fig. 2, where A is the amplitude at infinite time and T_{50} is the width at time domain from minus half amplitude to plus half amplitude.

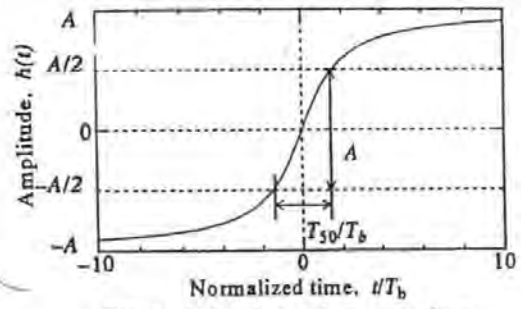


Fig. 2. Isolated reproducing waveform.

Here the normalized linear density (user density) for double layer perpendicular magnetic recording is defined as $K_p = T_{50}/T_b$. At the reading point, we assume that the noise is an AWGN with zero-mean and with two-sided power spectral density $N_0/2$. Hence, the signal to noise ratio (SNR) at the reading point can be defined as the ratio of the rising amplitude of $h(t)$ to the rms value of noise in a bandwidth equal to bit rate $f_b = 1/T_b$:

$$\text{SNR} = 20 \log_{10} \frac{A}{\sqrt{N_0 f_b}} \quad [\text{dB}]. \tag{3}$$

In order to compare the results against other similar systems employing different types of equalization, the SNR is mesured at the reading point and not at the discrimination point.

V. Partial Response Equalization

For $c_k = 1$, the recording waveform is square waveform shown in Fig. 3(a). Then the input waveform of the equalizer becomes

$$g(t) = \frac{1}{2}\{h(t) - h(t - T_s)\}, \tag{4}$$

as shown in Fig. 3(b). The equalized waveform of PR(u_0, u_1, $\cdots$, u_{L-1})ML system for $g(t)$ must be

$$w(t) = \frac{1}{2} \sum_{m=0}^{L-1} u_m r(t - mT_s), \tag{5}$$

where $r(t)$ is the Nyquist waveform with the roll-off factor β given by

$$r(t) = \frac{\sin(\pi t/\eta T_b)}{\pi t/\eta T_b} \cdot \frac{\cos(\pi \beta t/\eta T_b)}{1 - (2\beta t/\eta T_b)^2}. \tag{6}$$

Here η is the parameter to prescribe the amount of intersymbol interference and the bandwidth of the equalizer [2]. For example, employing PR(1,2,3,4,3,2,1), the equalizer output waveform for $c_k = 1$, $\beta = 0.5$ and $\eta = T_s/T_b$ is shown in Fig. 3(c).

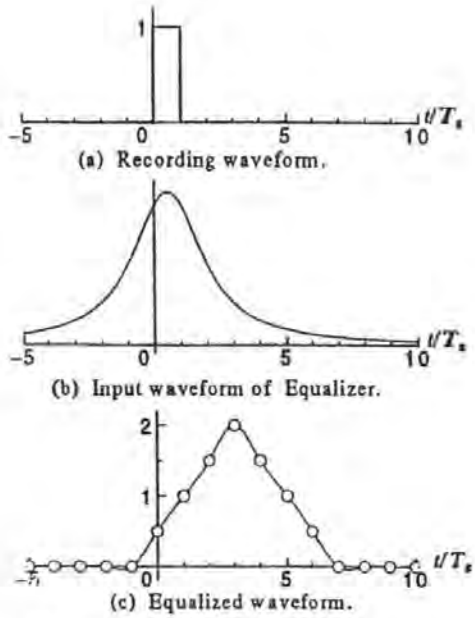


Fig. 3. PR(1,2,3,4,3,2,1) equalization.

Taking the ratio of $W(f)$ to $G(f)$ where $W(f)$ and $G(f)$ are the Fourier transforms of $w(t)$ and $g(t)$ respectively, the transfer function of the equalizer $E(f)$ is obtained. Integrating $N_0|E(f)|^2/2$, the noise power at the discrimination point for Fig. 1 can be obtained [2]. This noise may also become colored noise even if assumed AWGN at the reading point in Fig. 1. Using an impulse response, or the inverse Fourier transform of $E(f)$, the colored noise is calculated [4].

VI. PR(1,2,3,4,3,2,1)ML Detection

When $\eta = T_s/T_b$, the noise-free equalized sequence $\{d_k\}$ of PR(1,2,3,4,3,2,1)ML system is given by

$$\begin{aligned} d_k &= c_k + 2c_{k-1} + 3c_{k-2} + 4c_{k-3} + 3c_{k-4} \\ &\quad + 2c_{k-5} + c_{k-6} - 8. \end{aligned} \tag{7}$$

From (1) and (7), 64 states for Viterbi detection of PR(1,2,3,4,3,2,1)ML can be reduced to 26 by using run length constraint of (1,6) or (1,7) RLL code [5]. These RLL codes can not only eliminate 38 states, but also reduce 10 more add compare select (ACS) operations within the remained 26 states. Therefore this PRML detector can be extremely simplified. Using the constraint of $d = 1$, this PRML system can also increase the minimum Euclidian distance $d_{\min}$ from $\sqrt{8}$ to $\sqrt{26}$ [6].

VII. Bit Error Rate Performance

Fig. 4 shows the bit error rate (BER) performance of PRML systems obtained by the computer simulation

for double layer perpendicular magnetic recording where $K_p = 4$, $\beta = 0.5$, and $\eta = \eta_{opt}$. Here η_{opt} means the optimal value of η to give the minimum BER. As can be seen from Fig. 4, all RLL coded PRML systems show better BER performance than the conventional 16/17 coded EPR4ML system. Especially, the new PR(1,2,3,4,3,2,1)ML systems using both (1,6) and (1,7) RLL codes which demonstrate greatly improved performance. The required SNR's at the reading point to achieve a BER of 10^{-4} are 31.5dB and 32.0dB for (1,7) and (1,6) RLL coded PR(1,2,3,4,3,2,1)ML systems, respectively while 45.4dB for 16/17 coded EPR4ML system. The improvement in SNR using (1,6) RLL code at $K_p = 4$ is 13.4dB for conventional 16/17 coded EPR4ML, and 4.6dB and 2.2dB for PR(1,2,2,1)ML and PR(1,2,3,2,1)ML, respectively.

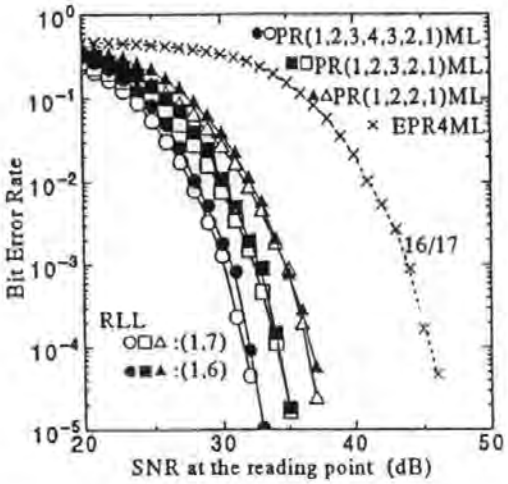


Fig. 4. BER vs. SNR ($K_p = 4$, $\beta = 0.5$, $\eta = \eta_{opt}$).

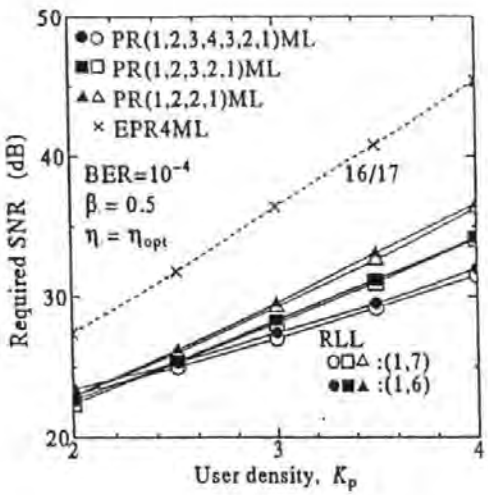


Fig. 5. Required SNR vs. user density.

Fig. 5 shows the relationship between K_p and the required SNR to achieve a BER of 10^{-4}. As can be seen, the new RLL coded PR(1,2,3,4,3,2,1)ML system provides enhanced performance especially at high densities.

VIII. Conclusions

Results from the computer simulations show that the new high order PR(1,2,3,4,3,2,1)ML scheme, together with the 2/3 rate (1,6) RLL code, give a 13.4dB gain in performance over conventional 16/17 coded EPRML systems at high densities.

The results also show that the 2/3 rate (1,7) code exhibits slightly better performance than the (1,6) code. This is possibly due to the differences in their power spectra. The model employed assumes ideal timing recovery, and thus no timing jitters is experienced at the sampling point on the eye opening. However, in a practical system it is expected that the use of (1,6) RLL code with its smaller minimum run length constraint should yield an improvement over (1,7) RLL code.

Acknowledgment

The authors wish to thank Prof. S. Tazaki of Matsuyama University, Prof. Y. Nakamura and Prof. H. Muraoka of Tohoku University, and Dr. H. Saito of Ehime University for their helpful discussions.

References

[1] R. Wood, "The feasibility of magnetic recording at 1 terabit per square inch," *IEEE Trans. Magn.*, vol. 36, pp. 36-42, Jan. 2000.

[2] H. Osawa, Y. Kurihara, Y. Okamoto, H. Saito, H. Muraoka, and Y. Nakamura, "PRML systems for perpendicular magnetic recording," *J. Magn. Soc. Japan*, vol. 21, no. S2. pp. 399-405, Oct. 1997.

[3] R. L. Adler, R. K. Brayton, M. Hassner, and B. P. Kitchens, "A rate 2/3, (1,6)RLL code," *IBM Techn. Discl. Bul.*, vol. 27, pp. 4727-4729, 1985.

[4] H. Osawa, S. Yamashita, and S. Tazaki, "Improvement on error rate performance for FM recording code," *IEEE Trans. Magn.*, vol. 27, pp. 4464-4469, Nov. 1991.

[5] A. M. Patel, "A new digital signal processing channel for data storage products," *IEEE Trans. Magn.*, vol. 27, pp. 4579-4584, Nov. 1991.

[6] G. D. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 18, pp. 363-378, May 1972.

IEEE Magnetics Journal Paper

Increased Areal Density Using a 2-Dimensional Approach

This paper was presented by the author at Intermag 2001 and introduces the new class of Constructive ITI codes.

Increased Areal Density Using a 2-Dimensional Approach

Mohammed Zaki Ahmed, Terry Donnelly, Paul J. Davey, and Warwick W. Clegg.

Abstract—**There is a drive towards achieving much higher track density in magnetic recording. Increase in track density will produce more interaction/crosstalk between adjacent tracks. The simulated variation of bit error rate (BER) as a result of increased inter-track interference (ITI) using a Viterbi Detector is presented in this paper. Results show that for a given BER and signal-to-noise ratio (SNR), different levels of ITI can be tolerated depending on the code. Calculations suggest 2-Dimensional modulation codes designed for a channel experiencing ITI can result in up to 23% gain in areal density.**

Index Terms— **Inter-track Interference, 2-Dimensional Code.**

I. Introduction

THERE is an estimated annual doubling in the areal density of magnetic recording systems [1]. It is predicted that track densities will increase to about 175k tracks/inch within the next year to accommodate this growth [2]. This will lead to greater interaction or cross talk between adjacent tracks. One of the most significant issues raised by Wood [1] is the side reading of information from adjacent tracks.

In our approach 2-Dimensional codes are used to control adjacent-track information. Inter-Track Interference (ITI) can be viewed as either constructive or destructive. Constructive ITI occurs when the polarities of the magnetic transitions on adjacent tracks are the same, leading to an increased read-back voltage. Destructive ITI exists when the fields oppose and result in a reduced read-back voltage.

II. Channel Model

A software channel model is used to simulate a PR4 ($1-D^2$) ideally-equalized channel operating in Additive White Gaussian Noise (AWGN), figure 1. The amount of ITI is assumed linear and symmetrical and no interaction is assumed between groups of tracks[3]. This paper does not concern itself with loss caused by miss equalization, therefore ideal equalization is assumed to offset any additional errors introduced by the equalizer. To obtain this, the read-back pulses are assumed sinc (sin(x)/x) pulses [4]. Negligible jitter is also assumed at the sampling point. The sampled sequence is split into odd and even samples and Viterbi detectors are used to obtain the encoded data. 32k bytes of data generated by a pseudo-random binary sequence were encoded and the BER was measured at the input of the decoder. This measured BER is normalized using code rate to give the user BER. The amount of ITI is modified and the experiment repeated. The Viterbi detector operated on the samples of the PR4 channel and is code independent.

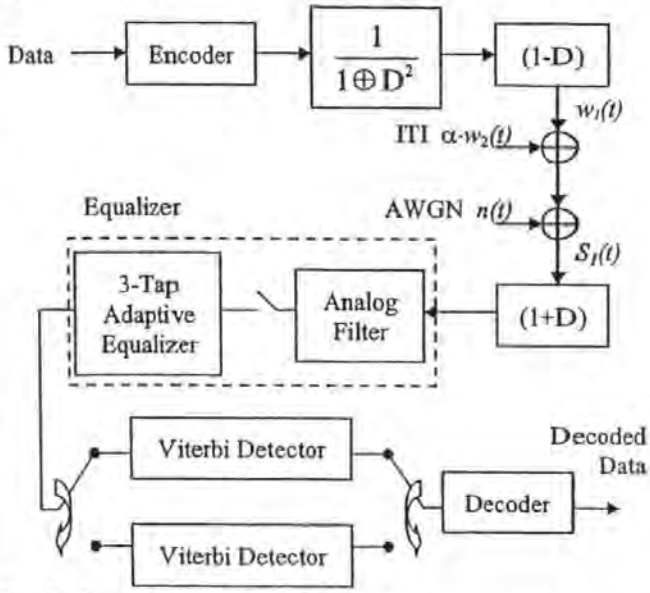


Fig. 1. Block diagram of PR4ML System

The read-back waveforms from the channel can be described by the following equations

$$S_1(t) = w_1(t) + \alpha \cdot w_2(t) + n(t)$$

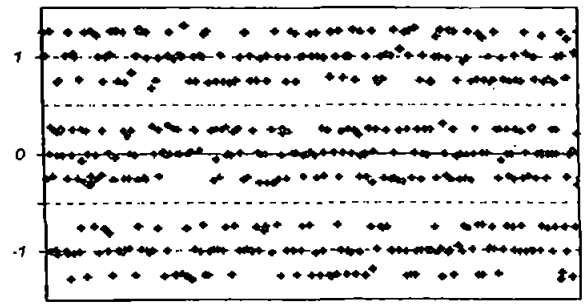
$$S_2(t) = w_2(t) + \alpha \cdot w_1(t) + n(t)$$

where $S_1(t), S_2(t)$ are the read-back waveforms with ITI; $w_1(t), w_2(t)$ are the waveforms assuming no ITI; α is the amount of ITI; and $n(t)$ is AWGN.

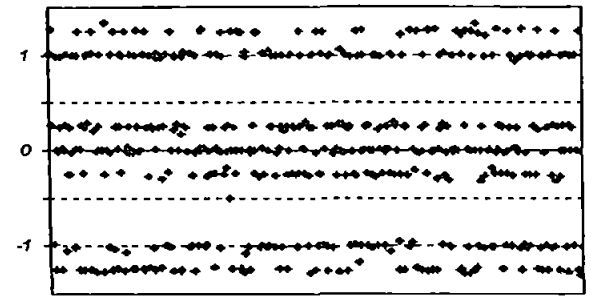
An ideal, sampled, noise and ITI-free read-back sequence comprises three states {+1, 0, -1} for a PR4 channel model. In the presence of ITI, these three states normally split into nine states as shown in figure 2a. The two additional sub-states due to ITI are equally spaced around each of the three basic states. A code designed for an ITI-affected channel experiences only seven states (figure 2b), since the code is designed not to have any transitions of opposite polarity on adjacent tracks. This constraint results in the elimination of two sub-states. The degradation in BER as a function of ITI using a Viterbi detector is due to the sub-states (1-ITI), (-1+ITI), (0+ITI) and (0-ITI).

Manuscript received 03 August 2000; revised 25 January 2001.

M. Z. Ahmed, T. Donnelly, P. J. Davey and W. W. Clegg are with the Center for Research in Information Storage Technology (C.R.I.S.T.), Department of Communication and Electronic Engineering, University of Plymouth, Drake Circus, PL4 8AA, UK. (e-mail: zaki@ieee.org).



(a) RLL Code



(b) Constructive ITI Code

Fig. 2. PR4 Samples at 14.5dB SNR with 25% ITI showing the effect of ITI converting 3-levels to 9-levels in (a) and 7-levels in (b)

III. Coding and Detection

Two modulation codes are compared. A code permitting only the same polarity of transitions on adjacent tracks [5] (referred to as a Constructive ITI, or CITI, code) is compared with a run-length limited (RLL) code. The CITI code is a rate 5/6 code with 2-dimensional run length constraints of (0,7) along the track and (0,4) across the tracks. The (d,k) run length constraints relate to the minimum (d) and maximum (k) number of "zeros" between "ones" in the encoded sequence. This fixes the minimum and maximum run length of transitions on the recording medium. A rate 8/9 (0,4/4) PRML RLL code was used as a benchmark. A (0,G/I) PRML code has (0,G) global run length constraint and (0,I) interleaved run length constraint. As the RLL code generates only single-track data, two independent random data sequences were used to generate codewords for the two tracks. The CITI code generates two track codewords from a data sequence.

The CITI code maps 5 data bits to 6 code bits arranged in a 2x3 matrix (2 tracks, 3 bits per track). The encoding and decoding is implemented using a lookup table (Table I). Two states, marked "I" and "O" are required to monitor the polarity of transitions on both tracks. State "I" refers to the condition when transitions on both tracks are in-phase, such that adjacent transitions will result in constructive ITI. Opposing phase, state "O" occurs when the first transition on either track leads to an opposite polarity transition.

TABLE I
CITI Encoder Lookup Table

Data	Codeword/Next State State I	State O	Data	Codeword/Next State State I	State O
00000	100	000	10000	001	001
	100/I	001/I		010/I	010/O
00001	100	000	10001	100	100
	111/I	010/I		010/I	010/O
00010	101	000	10010	111	111
	011/I	100/I		001/I	010/O
00011	101	001	10011	110	110
	101/I	000/I		101/I	011/O
00100	101	010	10100	011	011
	110/I	000/I		011/I	110/O
00101	110	100	10101	110	110
	110/I	000/I		001/O	001/I
00110	111	111	10110	111	111
	100/I	000/I		110/O	011/I
00111	111	000	10111	101	101
	111/I	111/I		010/O	010/I
01000	011	101	11000	110	110
	101/I	001/I		100/O	010/I
01001	010	001	11001	010	010
	010/I	011/I		101/O	101/I
01010	001	001	11010	011	011
	111/I	101/I		010/O	001/I
01011	010	010	11011	001	001
	001/I	001/O		110/O	110/I
01100	001	010	11100	010	010
	001/I	111/O		011/O	110/I
01101	010	010	11101	011	011
	100/I	100/O		100/O	100/I
01110	001	001	11110	101	011
	100/I	100/O		100/O	111/I
01111	100	100	11111	100	100
	001/I	001/O		011/O	011/I

When encoding, the initial states of both tracks are known, and codewords are taken from the appropriate state column. The next state is determined by the current codeword. To employ the code in a PR4 channel, bit interleaving of codewords is performed at the encoder prior to recording. This maintains the k=7 maximum run-length constraint for the odd and even interleaves of the Viterbi Detector.

The separation of tracks for a given amount of ITI was calculated using the relationship given by Herk [6], simplified to

$$\alpha = -55(d/\lambda) - 6 \text{ dB} \quad (1)$$

where d is the distance from the edge of the track, and
λ is the wavelength of the recorded signal
It was assumed that $\lambda/w = 12$ where w is the track width.

For a given ITI, equation 1 is used to determine the track separation, illustrated in figure 3. The area of interest is that shaded in figure 3. No interaction is assumed between groups of tracks and it is also assumed that the recording density is the same. The number of user bits in this area for a given amount of SNR and BER was kept constant for the RLL and CITI

codes, and the areas compared. The ratio of the penalty due to rate loss is constant and does not affect the result at different PW50/T.

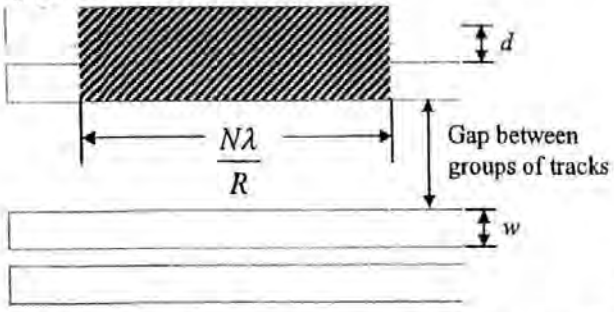


Fig. 3. Schematic of track area occupied by N user bits using a rate R code.

The variation of BER with ITI was measured for the two codes and is presented in figure 4. Calculations show the loss in linear density for the CITI code is more than compensated for by the gain in lateral/track density at the points of interest.

IV. Performance Comparison

The results of variation of BER vs. ITI (figure 4) suggest a sudden degradation in BER beyond an "allowable" ITI. The CITI code has a greater margin of allowable ITI compared to the RLL code, using the same detector. The greater margin is attributed to the fact that there are fewer (7) levels in the detected sequence of the CITI code. This also leads to a lower BER when the ITI is greater than 50%.

Assuming the code bit density is the same for both codes; at an SNR of 14.5dB, with the RLL code suffering 20 % ITI and the CITI code suffering 30% ITI, the expected BER is 10^{-4}. This difference in allowable ITI can be converted to an areal density using equation 1, and an overall gain in areal density of 23% is achieved using the CITI code even though it is a lower rate code.

V. Conclusion

An insight into the variation of BER with ITI has been presented. The fact that the CITI code can suffer greater ITI can be used to provide a gain in areal density. Using a Viterbi Detector, there is a "waterfall" region of degradation of BER with respect to ITI. As the ITI reaches and goes beyond 50%, a saturation level in BER is reached.

Controlling the type of crosstalk that exists is essential as the track density increases. The CITI code gives lower linear density due to the lower code rate, however the improved tolerance to ITI can offer an overall gain in areal density of 23% over conventional RLL codes. Furthermore, the CITI code described is applicable to perpendicular recording and any PR channel that suffers ITI. Current research in advanced detection algorithms may demonstrate further improved performance. Emphasis has been laid on the use of a modulation code designed for ITI-affected channels to increase areal density.

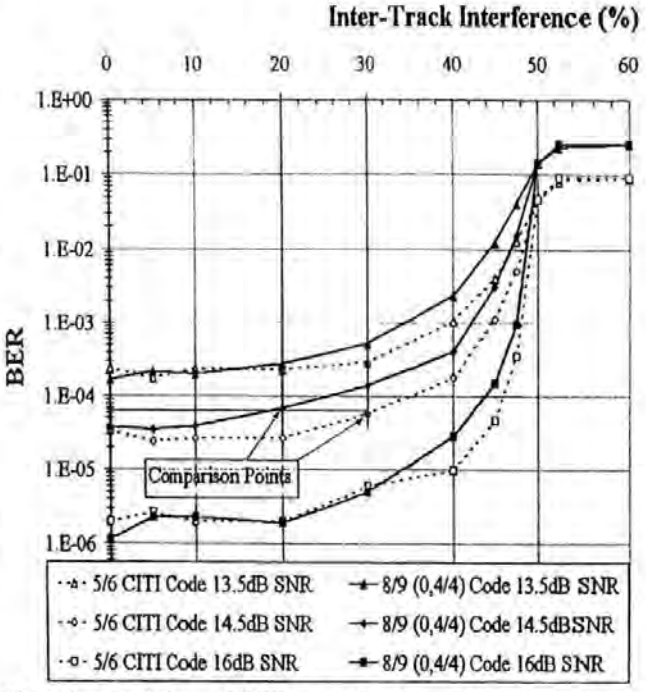


Fig. 4. Variation of BER as a function of ITI showing added immunity provided by the CITI code over the RLL code.

VI. Acknowledgments

The authors would like to thank Neil Darragh (Maxtor) for the use of his PRMLLab software and John Mallinson for guidance on inter-track interference. The authors would also like to thank the referees' for their useful comments.

References

[1] Wood R., "The Feasibility of Magnetic Recording at 1 Terabit per Square Inch", IEEE Transactions on Magnetics, Vol. 36 No. 1 Jan 2000 pages 36-42.

[2] Mittleholzer T., "On Iterative Decoding For Magnetic Recording Channels", Invited Talk at The 2nd International Symposium on Turbo Codes and Related Topics, 7th September 2000, Brest, France.

[3] Kurtas E., Proakis J.G., Salehi M., "Reduced Complexity Maximum Likelihood Sequence Estimation for Multitrack High-Density Magnetic Recording Channels", IEEE Transactions on Magnetics, Vol. 35 No. 4 July 1999 pages 2187-2193.

[4] Dolivio et al "Method and Apparatus for Decoding The Output Signal of a Partial-Response Class-IV Communication or Recording-Device Channel" United States Patent Number 4,571,734 dated Feb 18, 1986

[5] Davey P. J., Donnelly T., Mapps D.J., "Two Dimensional Coding for Controlled Inter-Track Interference", unpublished

[6] Herk A.V., "Side Fringing Fields and Write and read Crosstalk of Narrow Magnetic Recording Heads", IEEE Transactions on Magnetics, Vol. 13 No. 4 July 1977 pages 1021-1028.

IEEE Magnetics Journal Paper

Track Squeeze Using Adaptive Inter-Track Interference Equalisation

This paper was presented by the author at Intermag 2002 and introduces the new two track adaptive trellis decoder.

Track Squeeze Using Adaptive Inter-Track Interference Equalisation

Mohammed Zaki Ahmed, Paul J. Davey, Terry Donnelly and Warwick W. Clegg

Abstract— **A new Maximum Likelihood Sequence Detection (MLSD) technique to reduce the effects of Inter-Track Interference (ITI) is presented. As track densities increase, ITI and Offtrack Interference (OTI) occur, and have detrimental effects on the performance of current MLSD decoders. A solution, described in this paper, is to perform MLSD with an adaptive ITI equaliser constructed within the decoder trellis.**

Keywords— **Inter-Track Interference, Equalisation, Trellis**

I. Introduction

INCREASED recording densities can be achieved by an increase in track densities. As tracks are squeezed closer together, Inter-Track Interference (ITI) occurs. ITI and Off-Track Interference (OTI) have detrimental effects on the performance of current Maximum Likelihood Sequence Detection (MLSD) systems [1],[2]. A new detection trellis for a PR4 Channel is presented in this paper. It incorporates an adaptive equaliser (AEQ) to reduce the effect of ITI. This trellis construction technique can also be applied to other system which experience crosstalk. A block diagram of the new system is shown in Fig 1. This is based on the scheme described in [3].

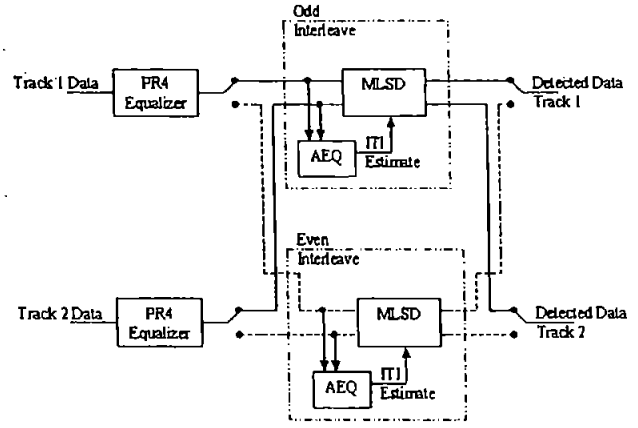


Fig. 1. Block Diagram of New 2-Track Detection System For a PR4 Channel

II. Multi-Track ITI Based Trellis

A. Channel and ITI Model

A two-track PR4 Channel model is used; assuming linear and symmetrical ITI [1]. Linear and Symmetrical ITI can be described using the following.

Assume
Z_1 and Z_2 are the received samples,
n_1 and n_2 are additive white gaussian noise (AWGN)
α is the amount of ITI and
y_1 and y_2 are the ideal ITI-free sample values,
then

$$\begin{aligned} Z_1 &= y_1 + \alpha y_2 + n_1 \\ Z_2 &= y_2 + \alpha y_1 + n_2 \end{aligned} \quad (1)$$

The channel model assumes the Signal to Noise Ratio (SNR) of both tracks is identical. A rate 8/9 $(0, 4/4)$ Run Length Limited (RLL) code is used to encode data on both tracks. The samples Z_1 and Z_2 are equalized to a PR4 Channel response using a 7-Tap FIR filter. This is done prior to the trellis decoding. In all simulations, the read-back signal from the channel is assumed Lorentzian with pulse-width at half-maximum (PW_{50}/T) of 2.0

B. ITI Estimation

An Adaptive Filter is used to estimate the amount of ITI. The coefficients of the filter are updated using a Least-Mean-Squared (LMS) algorithm [4]. An algorithm to estimate α is described below.

2-Track PR4 ITI Estimation Technique

1. Initialize
   $\alpha = \alpha_{old} = \alpha_{new} = 0$
   $\mu = 0.01$
2. Input samples from each track Z_1 and Z_2
3. Threshold Z_1 and Z_2 to obtain $\hat{R}_1$ and $\hat{R}_2$, estimates of the ideal PR4 samples $\{1, 0, -1\}$
4. Obtain an estimate of the ITI-Free sample $\hat{y_1}$
   $\hat{y_1} = (Z_1 - \alpha_{new} Z_2)/(1 - \alpha_{new}^2)$
5. Calculate the error $e_1 = \hat{R}_1 - \hat{y_1}$
6. Update α using
   $\alpha_{old} = \alpha_{new}$
   $\alpha_{new} = \alpha_{old} - \mu e_1 Z_2$
7. Obtain an estimate of the ITI-Free sample $\hat{y_2}$
   $\hat{y_2} = (Z_2 - \alpha_{new} Z_1)/(1 - \alpha_{new}^2)$
8. Calculate the error $e_2 = \hat{R}_2 - \hat{y_2}$
9. Update all α using
   $\alpha_{old} = \alpha_{new}$
   $\alpha_{new} = \alpha_{old} - \mu e_2 Z_1$
   $\alpha = \alpha_{new}$
10. Output α and go to 2

This algorithm seeks to estimate α from the simultaneous equations (1), with Z_1 and Z_2 known. The variable μ in the algorithm is the step size parameter of the LMS algorithm

All authors are with the Centre for Research in Information Storage Technology (C.R.I.S.T.), Department of Communication and Electronic Engineering (D.C.E.E.), University of Plymouth, Drake Circus, PL4 8AA, UK.
email: zaki@ieee.org

and the bounds relating to it are well documented in most good signal processing texts. For the software simulations μ was kept at 0.01

Fig 2 shows the convergence of α for different amounts of ITI and SNR.

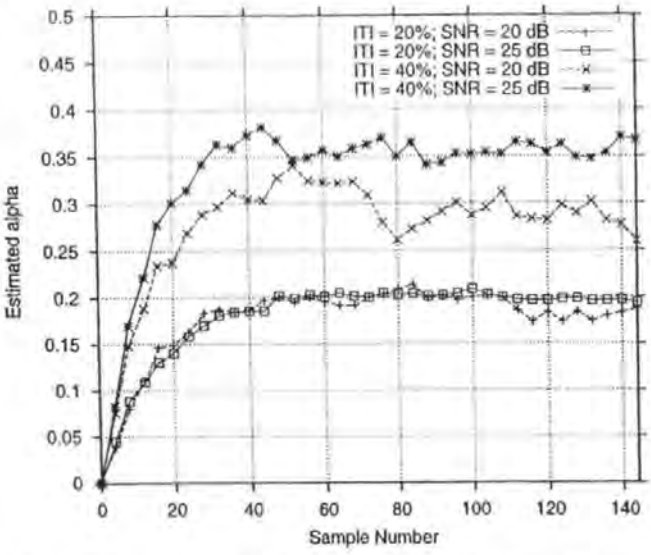
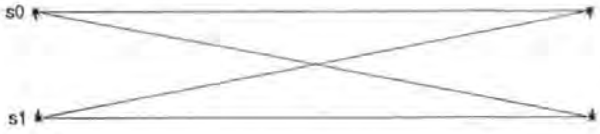


Fig. 2. Convergence of α for Different ITI and SNR for $\mu = 0.01$

As α increases, its variance from the steady-state value increases. This can be mitigated by using a greater SNR or a smaller μ. The exact inter-relationship between α, SNR and μ is still under investigation.

C. Trellis Construction

A trellis that decodes PR4 samples is shown in Fig 3. This trellis is used to decode odd and even interleaves of the received data from the channel [3].



Start State	End State	Ideal Path Metric
s0	s0	0
s0	s1	−1
s1	s0	1
s1	s1	0

Fig. 3. Conventional Interleaved PR4 Single-Track Trellis Section.

A new trellis is constructed using the existing states of the single-track Viterbi Detectors (labelled as s0 and s1 in Fig 3 and combining them using Table I.

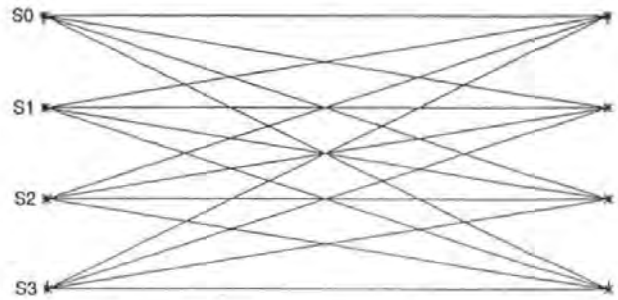
For example, going from S0 to S1 means track 1 will stay in state s0 and track 2 will change from s0 to s1. Consequently the ideal path metrics will be 0 and -1 for track 1 and track 2 respectively.

The new trellis is illustrated in Fig 4. Just combining two conventional PR4 detection trellis (Fig 3), yields a trellis

TABLE I
COMBINATION OF 2 CURRENT TRELLIS (FIG 3) STATES TO GIVE NEW TRELLIS STATES (FIG 4)

New State	Single Track State of Track 1	Single Track State of Track 2
S0	s0	s0
S1	s0	s1
S2	s1	s0
S3	s1	s1

with Ideal Path Metrics (PM). The effects of ITI can then imposed on the ideal PM to obtain the ITI-affected Path Metrics, shown in Fig 4. For the previous example this would mean the path metrics would become -α and -1 (using (1)and assuming ideal conditions of $n_1 = 0$, and $n_2 = 0$)



Start State	End State	ITI Affected PM (Tr 1,Tr 2)	Ideal PM (Tr 1,Tr 2)
S0	S0	$0, 0$	$0, 0$
S0	S1	$-\alpha, -1$	$0, -1$
S0	S2	$-1, -\alpha$	$-1, 0$
S0	S3	$-1-\alpha, -1-\alpha$	$-1, -1$
S1	S0	$\alpha, 1$	$0, 1$
S1	S1	$0, 0$	$0, 0$
S1	S2	$-1+\alpha, 1-\alpha$	$-1, 1$
S1	S3	$-1, -\alpha$	$-1, 0$
S2	S0	$1, \alpha$	$1, 0$
S2	S1	$1-\alpha, -1+\alpha$	$1, -1$
S2	S2	$0, 0$	$0, 0$
S2	S3	$-\alpha, -1$	$0, -1$
S3	S0	$1+\alpha, 1+\alpha$	$1, 1$
S3	S1	$1, \alpha$	$1, 0$
S3	S2	$\alpha, 1$	$0, 1$
S3	S3	$0, 0$	$0, 0$

Fig. 4. New Interleaved PR4 2-Track Trellis Section.

MLSD is performed on this new trellis to obtain the data. The complexity of a trellis is normally measured by the base-2 logarithm of the number of states, and the base-2 logarithm of the number of paths that require computing [5]. A conventional PR4 Detector operating on 2 tracks simultaneously would have 4 states and 8 paths to compute for each interleave. This results in a state complexity of 2

and a path complexity of 3. The new trellis described in this paper has a state complexity of 2 and path complexity of 4. Each path metric for the new trellis also has 2 terms to evaluate, whereas the number of terms in the conventional trellis path is 1. The ITI estimation is an additional computational overhead.

Neglecting the ITI estimation, the path complexity for the new trellis increases by 1, or the number of paths that need computation increases by a factor of 2. There are also twice as many terms to compute for each path, resulting in an overall increase in path computations of 4 times. The state complexity remains unchanged.

In the computer simulations, the ITI-affected PM were used in decoding the received samples (Z_1 and Z_2), using the estimated α from the algorithm described earlier. MLSD was then performed on the trellis using the Viterbi Algorithm.

III. Results

Results show negligible (about 0.5dB) deterioration in SNR for a fixed Bit Error Rate (BER) for ITI up to 30% (Fig 5).

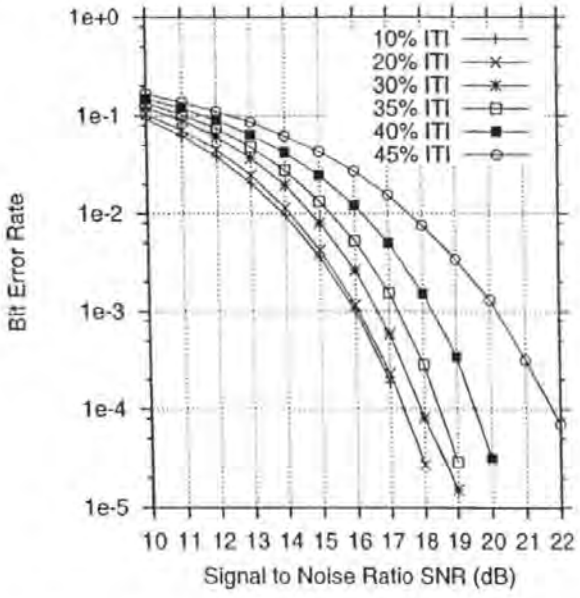


Fig. 5. SNR vs BER for a PR4 Channel Using New MLSD Decoder

There is a loss of about 2dB SNR for 40% ITI at a PW_{50}/T of 2.0. Previous techniques experience about 7dB loss in SNR at 40% ITI and similar PW_{50}/T [2]. The previous technique uses joint PRML to estimate adjacent track information from single-track information and assuming a known amount of OTI. For comparison purposes, the amount of OTI (λ) in [2] and ITI(α) can be equated using $\alpha = \lambda/(1-\lambda)$. As the ITI equaliser is adaptive, it can cope with varying amounts of ITI. The equaliser trains itself using data samples, and no training sequence is required.

The most common error event is when there are opposite transitions on adjacent tracks. Coding to eliminate this was done and a further 1dB SNR was gained at 40% ITI using codes designed for channels experiencing crosstalk [1].

The results obtained were compared with the more general work done in [6]. The variation of the BER performance with respect to ITI observed in this paper follows the theoretical bounds of minimum distance (d_{min}) for a two-track two-head system described in [6].

IV. Conclusions

A new trellis for simultaneously decoding 2 PR4 tracks is presented. The trellis path metrics include the effects of ITI. An algorithm has also been described which estimates the amount of ITI. Computer simulations suggest there is a significant gain in using this hybrid technique over previous methods. For the new trellis, there is an increase in path complexity, however the state complexity remains unchanged.

The new trellis construction scheme can be extended to channels experiencing crosstalk. As the number of adjacent tracks increases, the complexity of the trellis will increase significantly. The ITI estimation algorithm will also be more complex.

A trellis structure lends itself to *maximum a posteriori* (MAP) decoding [7]. MAP decoding can be used as part of a more robust error correction scheme. The output from the MAP decoder can be used to provide the ideal estimates for the algorithm (step 3 in the algorithm). This could be a possible enhancement of the system described in this paper, resulting in a multi-track error correcting system with some immunity to ITI.

Acknowledgment

The authors would like to thank two unknown referees for their comments and suggestions.

References

[1] M.Z. Ahmed, T. Donnelly, P.J. Davey and W.W. Clegg, "Increased Areal Density Using a Two-Dimensional Approach," *IEEE Trans. Magnetics* Vol. 37 No 4 pp 1896 – 1898.

[2] B.G. Roh, J. Moon and S.U. Lee, "Single-Head Detection In The Presence of Offtrack Interference," *GlobeCom* 99 pp 955 – 959.

[3] Dolivio et al, "Method and Apparatus for Decoding The Output Signal of a Partial-Response Class-IV Communication or Recording-Device Channel," U.S. Patent Number 4,571,734 dated Feb 18, 1986.

[4] S. Haykin *Introduction to Adaptive Filters*, Macmillan, New York, 1984.

[5] Pless & Huffman, Editors, Chapter 24 *Handbook of Coding Theory*, Elsevier, Amsterdam, 1998.

[6] E. Soljanin, *Coding for improving noise immunity in multi-track, multi-head recording systems*, PhD Thesis, Texas A & M University, U.S.A., 1994.

[7] L.R. Bahl, J. Cocke, F. Jelinek and J. Raviv, "Optimal Decoding of linear codes for minimising symbol error rate", *IEEE Trans. Inform. Theory* IT-20(1974), 284 – 287.

Intermag 2003 Conference Paper

New High Rate Low Density Parity Check Codes for Multi-Track Recording

This paper was presented by the author at Intermag 2003.

New High Rate Low Density Parity Check Codes for Multi-Track Recording

Mohammed Zaki Ahmed, Paul J. Davey, Terry Donnelly, and Warwick W. Clegg

Abstract—**A new decoding configuration for a serial concatenation of a Low Density Parity Check (LDPC) code and a Partial Response Class 4 (PR4) channel is presented. The input *apriori* probabilities to the LDPC decoder (the outer code) is made equal in the new approach taken. LDPC codes of rate greater than 0.9 are used at block lengths of about 5000 bits. The new scheme is compared to a conventional scheme where the input *apriori* probability to both inner and outer codes is identical. The new scheme performs better than the conventional scheme in the presence of Inter-Track Interference (ITI).**

Index Terms—**Low Density Parity Check (LDPC) Codes, Multi-Track recording, Inter-Track Interference.**

I. Introduction

HIGH rate (Code Rate $>$ 0.9) Low Density Parity Check (LDPC) error correction codes are applied to a multi-track PR4 channel experiencing cross-talk in this paper. Simulations were done for LDPC codes of a block size 4800 bits, and 192 parity-check equations (resulting in a code rate of 0.96), up to a block size of 5120 and rate 0.9. The number of user data bits was constant at 4096 in all simulations. At such high rates and block lengths, the bipartite[1] graphs of the codes have cycles of length 4. These cycles generally result in detrimental performance when using the sum-product algorithm (SPA) in decoding[2], however according to[3] codes with many cycles of length 4 can be decoded with good performance.

Designing 'good' codes (that is codes that do not have length 4 cycles) at these rates usually results in very weak codes. High-rate LDPC codes with good performance are difficult to formulate[4].

It has been found that in the presence of Inter-Track Interference (ITI), performance similar to codes designed for ITI channels[5], [6] can be achieved. This is done by iteratively decoding a high rate LDPC and treating the recording channel as an 'inner' code of a serially concatenated scheme at moderately high Signal to Noise Ratio (SNR).

II. System Model

The recording channel simulated is a two track PR4 channel with linear and symmetric ITI. On readback, the noise is assumed Additive White Gaussian Noise (AWGN) and the readback waveform is a Lorentzian with a PW_{50}/T of 2.0 An overall channel model is shown in Fig. 1.

Manuscript received 02 October, 2002; revised 08 December, 2002.

All authors are with with C.R.I.S.T., D.C.E.E., University of Plymouth, Drake Circus, PL4 8AA, UK. email: zaki@ieee.org

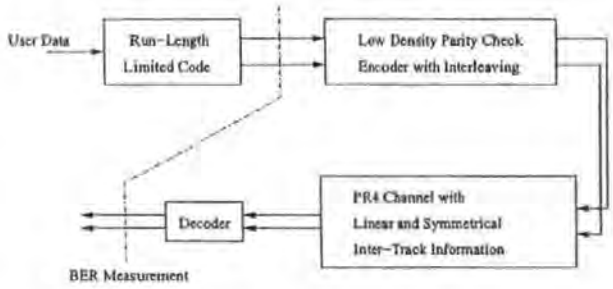


Fig. 1. Overall Channel Model

The Run-Length Limited (RLL) code used is a rate 8/9 (0,4/4) RLL code that has a minimum runlength 0 and maximum (global and interleaved) runlength of 4. Half the user data are encoded to each of two adjacent tracks using the RLL code. RLL data are encoded using the LDPC H matrix and interleaved at regular intervals in the overall sequence written to the channel. This results in very few instances of the maximum run-length constraint k_{rll} changing to $k_{rll}+1$. The LDPC H matrix was made systematic in order to simplify the encoding procedure. LDPC Codes are specified as (n, j, k), where n is the number of codeword bits, j is the number of '1' in each column of the H matrix and k is the number of '1' in each row of the H matrix. The LDPC code rate R is lower bound by

$$R \geq 1 - j/k \tag{1}$$

Regular LDPC's strictly adhere to the j and k constraints, whereas irregular LDPC's do not. As the parameters of the LDPC code do not specify a particular code, an ensemble of codes with the same n, j and k parameters are studied. It has been verified by the authors that $j = 3$ resulted in the best performing ensemble. This observation was also made in[2].

The LDPC H matrix is defined as

$$\mathrm{H} = [I|RLDPC] \tag{2}$$

where I is an identity matrix and $RLDPC$ is a regular LDPC. For example, if I is a 432 $\times$ 432 identity matrix, $RLDPC$ would be a regular $(4608, 3, 32)$ LDPC code. The resulting (irregular) LDPC code has a code rate $4608/(4608+432)$ that is just over 0.91. The inclusion of an identity matrix simplifies the encoding procedure at the expense of reducing the LDPC search set. It has been verified by the author that at rates of about 0.96 this does not affect the code performance significantly. Efficient methods of encoding LDPC are discussed in[7].

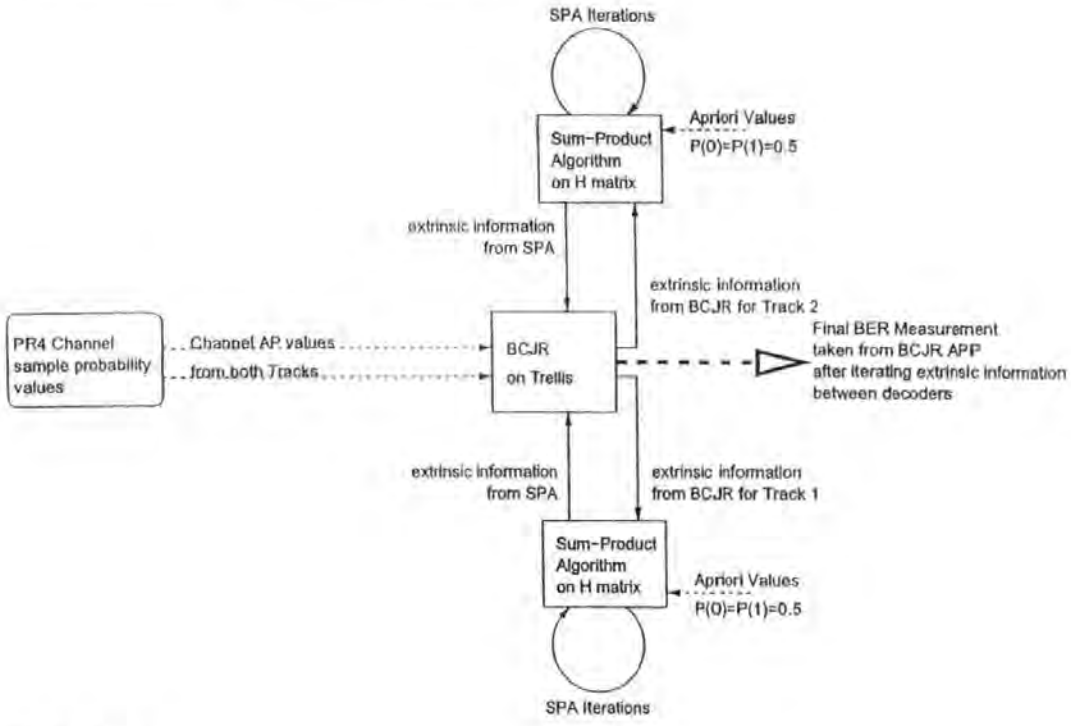


Fig. 2. New Decoding Scheme

Fig. 2 shows the implemented decoding scheme shown as 'Decoder' in Fig. 1. The 'inner' code was the PR4 channel, decoded using the multi-track trellis described by Ahmed[6] and the Bahl algorithm[8]. The *A Posteriori Probability* (APP) was evaluated from the input *Apriori Probabilities* (AP), and the extrinsic information obtained. The 'outer' code was an LDPC code decoded using the SPA. The stopping criteria for the SPA was when a valid codeword was reached, and the iterative procedure started with decoding the inner code. A maximum number of 200 iterations was set for the SPA that terminated the algorithm in the case of a non-convergent block.

The inner and outer decoders exchange only extrinsic information (e) defined as

$$e = \frac{APP}{AP} \tag{3}$$

The rationale in setting the SPA *apriori* values to 0.5 was in order to have a completely different estimate of the channel values as part of the decoding scheme. Iterative decoding relies on the fact that the information from decoders is uncorrelated. The BCJR algorithm results in the exact APP, hence the exact/true extrinsic information from the channel is obtained. The trellis also dynamically takes into account the effect of ITI (something the SPA cannot without increasing the complexity significantly) as described in[6]. SPA decoding results in an approximate APP, and in the new scheme completely uncorrelated *apriori* information is supplied to the SPA decoder. The SPA thus complements the trellis decoder.

A 'classical' decoding scheme would have the channel *apriori* values as the input to the SPA. The performance of these two types of decoding schemes is presented in the results.

III. Results

Fig. 3 shows the performance of the new decoding scheme in the presence of ITI.

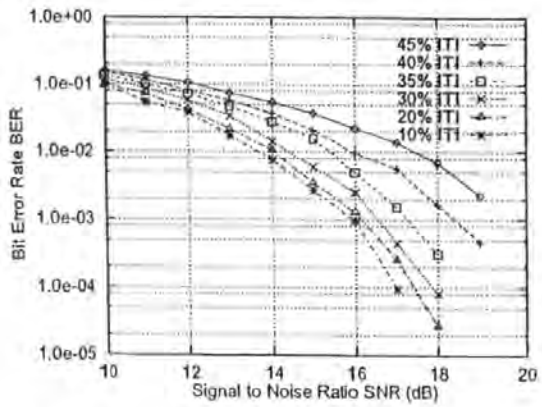


Fig. 3. Performance of New Decoding Scheme

SNR is defined as

$$\mathrm{SNR} = 10\log_{10}\left(\frac{1}{2\sigma^2}\right) \tag{4}$$

where σ^2 is the variance of the AWGN. The SPA on the H matrix starts with AP values of 0.5 as this gave better performance compared to using the channel AP values (see Fig. 4) Better performance was also obtained by starting the decoding using the channel AP values and the BCJR decoder compared to using the SPA decoder at the start. This is because there are no approximations in the APP estimates from the BCJR decoder and the initial estimates of extrinsic information are better than if the SPA is used.

Fig. 4 shows the performance of the new scheme compared to a classical scheme for an SNR of 17dB. The H matrix was a (5040, 3, 35) irregular LDPC constructed as described earlier. For comparison purposes a MAP decoder using the BCJR algorithm on the trellis is also shown.

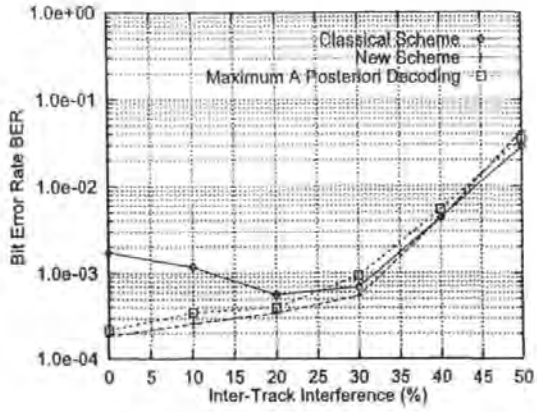


Fig. 4. Comparison of New Decoding Scheme with a Classical scheme

The new scheme out performed the classical scheme for ITI up to 40%. As ITI exceeds 40% the classical scheme shows better performance, however the classical scheme was observed to have convergence issues. The new scheme did not appear to be affected by convergence issues. MAP decoding resulted in similar performance compared to the new scheme, however the new scheme does have slightly better performance. The improvement in performance is more pronounced at an ITI of 30% where the dominant error event changes, and has been estimated to be about 1dB improvement in SNR. At 30% ITI, the classical scheme also out performs the MAP decoder, with an estimated improvement in SNR of about 0.5dB.

At very high ITI (ITI $>$ 40%), both MAP decoding and classical decoding perform (slightly) better than the new scheme. The theoretical reasons for this are currently under investigation. It was observed that at high ITI, the classical scheme experienced convergence problems whereas the new scheme did not. An interesting observation is that the classical scheme has slightly better performance even though it is non-convergent.

IV. Conclusions

A new method of decoding a serial concatenation of a high rate LDPC and a two track PR4 channel is presented. The choice of *apriori* probability to the SPA does affect the performance of the decoder in the presence of ITI. As the dominant error event changes, simulations suggest that the choice of starting *apriori* probability does not matter very much.

The use of a conventional RLL code and interleaving the parity bits ensures that a maximum run-length constraint can be guaranteed. Parity bits need to be interleaved at least $k_{rll}+1$ bits apart, where k_{rll} is the maximum run length constraint of the RLL code.

Recent results[9] suggest that LDPC codes with cycles of length 4 can be decoded with a gain of up to 0.5dB at code rates of 0.8. The application and investigation of this new decoding method to the multi-track channel is currently under investigation. A non-binary equivalent of the SPA that caters for the effect of ITI is also currently under investigation.

References

[1] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Transactions in Information Theory*, vol. 27, no. 5, pp. 533–547, Sep 1981.

[2] R. Gallager, "Low-density parity-check codes," PhD, MIT, 1963.

[3] J. Xu, H. Tang, Y. Kou, S. Lin, and K. Abdel-Ghaffar, "A general class of LDPC finite geometry codes and their performance," *International Symposium on Information Theory*, p. 309, jun 2002.

[4] H. Song, R. M. Todd, and J. Cruz, "Low density parity check codes for magnetic recording channels," *IEEE Journal on Selected Areas in Communication*, vol. 19, pp. 918–923, may 2001.

[5] M. Z. Ahmed, T. Donnelly, P. J. Davey, and W. W. Clegg, "Increased areal density using a 2-dimensional approach," *IEEE Transactions on Magnetics*, vol. 37, no. 4, pp. 1896–1898, jul 2001.

[6] ——, "Track squeeze using adaptive inter-track interference equalisation," *IEEE Transactions on Magnetics*, vol. 38, no. 5, pp. 2331–2333, sep 2002.

[7] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Transactions in Information Theory*, vol. 47, no. 2, pp. 599–618, feb 2001. [Online]. Available: citeseer.nj.nec.com/richardson98capacity.html

[8] L. Bahl, J. Cocke, F. Jelenik, and J. Raviv, "Optimal decoding of linear codes for minimising sybmol error rate," *IEEE Transactions in Information Theory*, vol. 20, pp. 284–287, mar 1974.

[9] M. A. Ambroze, M. Z. Ahmed, M. Tomlinson, and M. Ferrari, "Iterative decoding of binary codes by partitioning the H matrix," *in preparation for publication*, 2002.

Frequency Response

Lorentzian

The frequency response of equation 2.6 can be obtained from the isolated Lorentzian frequency response (equation 2.4) as follows

The Fourier transform of equation 2.4 can be evaluated as

$$V(\omega) = \int_{-\infty}^{\infty} \frac{1}{1+(\frac{2t}{PW_{50}})^2} e^{-j\omega t} dt$$

Substituting $x = \frac{2t}{PW_{50}}$ and $\frac{PW_{50}}{2} dx = dt$ gives

$$V(\omega) = \frac{PW_{50}}{2} \int_{-\infty}^{\infty} \frac{e^{-j\frac{PW_{50}}{2}\omega x}}{1+x^2} dx$$

$$= \frac{PW_{50}}{2} \left(\int_{-\infty}^{\infty} \frac{\cos(\frac{PW_{50}}{2}\omega x)}{1+x^2} dx - j \int_{-\infty}^{\infty} \frac{\sin(\frac{PW_{50}}{2}\omega x)}{1+x^2} dx \right)$$

The imaginary term over the limits of the integral evaluate to zero, as it is anti-symmetrical

$$\therefore V(\omega) = \frac{PW_{50}}{2} \int_{-\infty}^{\infty} \frac{\cos(\frac{PW_{50}}{2}\omega x)}{1+x^2} dx$$

$$= PW_{50} \int_{0}^{\infty} \frac{\cos(\frac{PW_{50}}{2}\omega x)}{1+x^2} dx \qquad (1)$$

Using Gradshteyn pp 418, equation 3.723.2[GR00] to solve (1) (substituting $a = \omega\frac{PW_{50}}{2}$ and $\beta = 1$ in [GR00] equation 3.723.2)

$$V(\omega) = \frac{PW_{50}}{2} \pi e^{-\omega\frac{PW_{50}}{2}}$$

The Lorentzian dibit spectrum $D(\omega)$ can thus be obtained as

$$D(\omega) = (1 - e^{-j\omega T})\frac{PW_{50}}{2}\pi e^{-\omega\frac{PW_{50}}{2}} \tag{2}$$

where ω is the frequency term and T is the sampling period.

Partial Response

The dibit resulting from using equation 2.5 (the PR4 dibit) has a frequency response

$$\begin{aligned}
S(D) &= (1-D)^p(1+D)^q \\
S(\omega) &= (1-e^{-j\omega T})^p(1+e^{j\omega T})^q \\
&= [(e^{+j\frac{\omega T}{2}} - e^{-j\frac{\omega T}{2}})e^{-j\frac{\omega T}{2}}]^p \cdot [(e^{+j\frac{\omega T}{2}} + e^{-j\frac{\omega T}{2}})e^{+j\frac{\omega T}{2}}]^q \\
&= \left[+j2\sin\left(\frac{\omega T}{2}\right)\right]^p \cdot e^{-jp\frac{\omega T}{2}}\left[2\cos\left(\frac{\omega T}{2}\right)\right]^q \cdot e^{+jq\frac{\omega T}{2}} \\
&= \left[+j2\sin\left(\frac{\omega T}{2}\right)\right]^p \cdot \left[2\cos\left(\frac{\omega T}{2}\right)\right]^q \cdot e^{+j(q-p)\frac{\omega T}{2}} \\
|S(\omega)| &= \left[2\sin\left(\frac{\omega T}{2}\right)\right]^p \cdot \left[2\cos\left(\frac{\omega T}{2}\right)\right]^q \\
\arg(S(\omega)) &= (q-p)\frac{\omega T}{2}
\end{aligned}$$

where ω is the frequency term and T is the sampling period. The expression above can be used to evaluate the frequency spectra of any PR system, given p and q. For a PR4 channel $p = 1$ and $q = 1$, and the magnitude of the frequency response is given by $S(\omega) = 4\sin(\frac{\omega T}{2})\cos(\frac{\omega T}{2}) = 2\sin(\omega T)$ *

*$\sin(a)\cos(a) = \frac{\sin(2a)}{2}$

Shannon Capacity

If a matrix $\mathbf{M}$ is the STM of a system, the rowsum of the transpose of matrix $\mathbf{M}$ is the total number of paths or sequences. A matrix can be decomposed into its spectrum using procedure described in Bickley chapter 9[BT64]. An $n \times n$ matrix $\mathbf{M}$ can be decomposed into its spectra as

$$\mathbf{M} = \lambda_0 \mathbf{SC_0} + \lambda_1 \mathbf{SC_1} + \cdots + \lambda_n \mathbf{SC_n}$$

where $\mathbf{SC_0}$ is the spectral component associated with eigenvalue λ_0

If all the components of $\mathbf{M}$ are real numbers, the largest eigenvalue is real

$$\begin{aligned}
\mathbf{M}^k &= \lambda_0^k \mathbf{SC_0} + \lambda_1^k \mathbf{SC_1} + \cdots + \lambda_n^k \mathbf{SC_n} \\
\lim_{k\to\infty} \mathbf{M}^k &= \lim_{k\to\infty} \lambda_0^k \mathbf{SC_0} \quad \text{assuming } \lambda_0 \text{ is the largest eigenvalue} \\
&\quad \text{thus the rowsum of } \lim_{k\to\infty} \mathbf{M}^k \text{ is} \\
&\quad\quad \text{the rowsum of } \lim_{k\to\infty} \lambda_0^k \mathbf{SC_0} \\
\text{rowsum} \lim_{k\to\infty} \lambda_0^k \mathbf{SC_0} &= \text{rowsum} \lim_{k\to\infty} \lambda_0^k \cdot \zeta \\
&\quad\quad \text{where } \zeta \text{ is a constant}
\end{aligned}$$

Using the property above, and equation 2.14

$$\begin{aligned}
\lim_{T\to\infty} \frac{\log(\lambda_0^T \zeta)}{T} &= \log(\lambda_0) + \lim_{T\to\infty} \frac{\log(\zeta)}{T} \quad \text{assuming } \lambda_0 \text{ is the largest eigenvalue} \\
&= \log(\lambda_0)
\end{aligned}$$

This provides the link between equation 2.14 and equation 2.16.

Combinatronics

Combinatorial Mathematics

Consider the magic[†] square that Chinese Emperor Yu (2200BC) observed on the back of a tortoise.

$$\begin{bmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{bmatrix}$$

Combinatronics is the study of arrangements of elements **into** sets. There are two basic types of problems

Existence Problems The existence of the prescribed configuration is in doubt, and the study attempts to settle this issue.

Enumeration Problems The existence of the configuration is known, and the study attempts to determine the number of configurations or the classification of these configurations according to types.

A combinatorial design is an arrangement of a set of ν points into b subsets, called blocks. The arrangements satisfy pre-determined conditions.

Regular A design is regular if the number of points in each block γ and the number of blocks which contain each point ρ are the same for every point and every block.

Covalency This is defined for two points. The number of blocks that contain both points is the covalency λ_{xy}.

[†]all rows, columns and diagonals add up to 15, and all the numbers are different. Pure magic!!!

Balanced If the covalency is constant for all x and y then the design is balanced.

A regular balanced design is denoted as a $(\nu, b, \rho, \gamma, \lambda)$ design and is one type of combinatorial design. Every design can be described by a $b \times \nu$ incidence matrix I, where each row represents a block B_i and each column a point P_j and

$$I_{i,j} = \begin{cases} 1 & \text{if } P_j \in B_i \\ 0 & \text{otherwise} \end{cases}$$

For a regular design, the number of 1's in I is $\nu \times \rho = b \times \gamma$. The incidence matrix (or the transpose of the incidence matrix) of a combinatorial design can be used as the H matrix of an LDPC.

Consider the transpose of the incidence matrix of a design is the H matrix. The H matrix would have ν rows, b columns, ρ row weight and γ column weight and the condition of no 4 cycles imposed by $0 < \lambda \leq 1$, or $\lambda = \{0, 1\}$ The resulting LDPC will be a regular LDPC. With this convention, combinatronic designs of interest (free of 4-cycles) for regular LDPC are thus

$(\nu, b, \rho, \gamma, \lambda)$=(**num. parity equations,num. bits, Gallager** k**, Gallager** j**,** $\lambda = \{1, 0\}$)

Other properties of the $(\nu, b, \rho, \gamma, \lambda)$ designs.

1. The elements ν are called the *varieties* and b are blocks (the nomenclature comes from the design of experiments in statistics).

2. The precise range of values of $(\nu, b, \rho, \gamma, \lambda)$ for which configurations exist is (apparently) an unsolved problem[Rys63].

3. These two equations are not the only conditions for a design to exist

$$\rho(\gamma - 1) = \lambda(\nu - 1) \tag{3}$$

$$b\gamma = \nu\rho \tag{4}$$

4. $b \geq \nu$

Steiner Triple System (STS)

A Steiner triple system is a $(\nu, b, \rho, \gamma, \lambda)$ design with γ = Gallager $j = 3$ and $\lambda = 1$. A Steiner triple system of order ν (= num parity equations) is a set of 3-subsets of blocks b such that each 2-subset of blocks is a subset of exactly one triple.

For example a

$$\begin{aligned}(7, b, \rho, 3, 1) &= \{1,2,4\},\{2,3,5\},\{3,4,6\},\\ &\quad \{4,5,7\},\{5,6,1\},\{6,7,2\},\{7,1,3\}\end{aligned}$$

$$\rho = \text{Gallager } k = \frac{\nu - 1}{2} \tag{5}$$

$$b = \text{num codeword bits} = \frac{\nu(\nu - 1)}{6} \tag{6}$$

If a Steiner Triple System exists of order v_1 and another Steiner Triple System exists of order v_2, then there exists a Steiner Triple system of order $v_1 v_2$. Starting with systems of order 3,7,9,13 and 15 more systems can be described based on these multiples. A Steiner system suggests very high rate codes are possible.

Order of Steiner Triple System ν	Number of Combinations	Number of data bits b	Gallager k ρ	Code Rate
3	1	1	1	
7	7	7	3	
9	12	12	4	0.25
13	2	26	6	0.50
15	80	35	7	0.57
21		70	10	0.70
27		117	13	0.77
39		247	19	0.84
45		330	22	0.86
63		651	31	0.90
91		1365	45	0.93
105		1820	52	0.94
117		2262	58	0.95
135		3015	67	0.955
195		6305	97	0.969

Table 1: Steiner Triple System Table

Kirkman Triple System (KTS)

A Kirkman Triple System is a Steiner Triple System that is resolvable. The blocks b can be arranged into ρ groups such that the $\frac{\nu}{3}$ blocks of each group are disjoint. Each group contains every point only once.

The order ν (= num parity equations) of a Kirkman triple system (n is a non-negative integer) is

$$\nu = 6n + 3$$

The values of ρ which is the Gallager k and b the number of codeword bits can be evaluated using the Steiner Triple System equations (page 121).

If all blocks in a group are removed from H what remains is a parity check matrix H_o with ν parity checks, row weight $\rho - 1$ and $n = b - \frac{\nu}{3}$(see [JW01]).

Construction of a Kirkman Triple System (taken from [JW01])

Let $q = 6m + 1$ be a prime power, m any integer, then a Kirkman triple system with $\nu = 3q$ exists.

Construction : Take θ a primitive element of $GF(q)$, so that $\theta^{6m} = 1$, $\theta^{3m} = -1$ and $\theta^{2m} + 1 = \theta^m$, then use three copies of each element of $GF(q)$ to construct the sets

$$\begin{aligned}
A &= \{0_1, 0_2, 0_3\} \\
B_{ij} &= \{\theta_j^i, \theta_j^{i+2m}, \theta_j^{i+4m}\} \qquad 1 \le i \le m, 1 \le j \le 3 \\
C_{ij} &= \{\theta_j^{i+m}, \theta_{j+1}^{i+3m}, \theta_{j+2}^{i+5m}\} \quad 1 \le i \le m, 1 \le j \le 3 \pmod 3 \\
D_{ij} &= \{\theta_j^i, \theta_{j+1}^{i+2m}, \theta_{j+2}^{i+4m}\} \qquad 1 \le i \le m, 1 \le j \le 3 \pmod 3
\end{aligned}$$

The sets A, B_{ij}, C_{ij} form one resolution class, and the translates of the class give a further $6m$ classes. Next each D_{ij} with its translates gives a further resolution class; so a total of $9m + 1$ resolution classes are obtained.

Example of KTS

Using the algorithm, for $m = 1$ and $GF(7)$ the following is obtained (the subscript denotes the set of $GF(7)$ the particular element is taken from, and three sets are used) $A = \{0_1, 0_2, 0_3\}$, $B_{1,1} = \{3_1, 6_1, 5_1\}$, $B_{1,2} = \{3_2, 6_2, 5_2\}$, $B_{1,3} = \{3_3, 6_3, 5_3\}$, $C_{1,1} = \{2_1, 4_2, 1_3\}$, $C_{1,2} = \{2_2, 4_3, 1_1\}$, $C_{1,3} = \{2_3, 4_1, 1_2\}$, $D_{1,1} = \{3_1, 6_2, 5_3\}$, $D_{1,2} = \{3_2, 6_3, 5_1\}$, $D_{1,3} = \{3_3, 6_1, 5_2\}$. This results in $9m + 1$ "resolution classes"[‡].

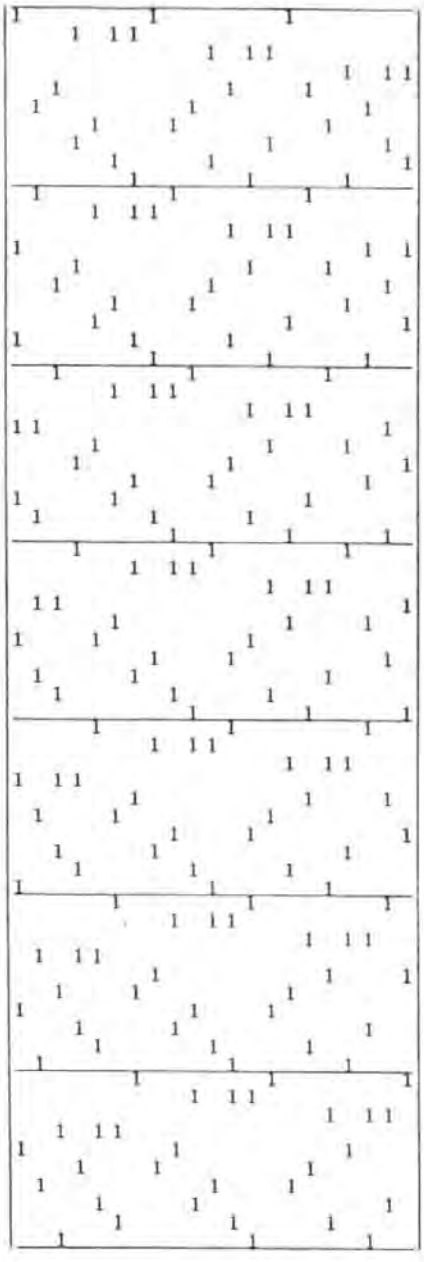
The Incidence Matrix can be evaluated using the rule that the element in the matrix $a_{i,j}$ is 1 if a_j is a member of set S_i. S_i are the sets $A, B_{1,1}, \cdots, D_{1,3}$. The columns of the incidence matrix are the elements, and the rows the sets (zeros have been omitted for clarity)

	0_1	1_1	2_1	3_1	4_1	5_1	6_1	0_2	1_2	2_2	3_2	4_2	5_2	6_2	0_3	1_3	2_3	3_3	4_3	5_3	6_3
A	1							1							1						
$B_{1,1}$				1		1	1														
$B_{1,2}$											1		1	1							
$B_{1,3}$																		1		1	1
$C_{1,1}$			1									1				1					
$C_{1,2}$		1								1									1		
$C_{1,3}$					1				1								1				
$D_{1,1}$				1									1							1	
$D_{1,2}$						1					1										1
$D_{1,3}$							1						1					1			

The translates of this set (shifting the original set to the right by 1 and wrapping

[‡]see kts.cpp to create Different KTS

around) gives the other sets.



The transpose of this matrix gives the H matrix of an LDPC with 21 check equations and 70 bits and no cycles of length 4.

Other valid parameters of a Kirkman Triple System are shown in table 2

m	q	ν	b	Code Rate
1	7	21	70	0.70
2	13	39	247	0.84
3	19	57	532	0.89
5	31	93	1426	0.93
6	37	111	2035	0.95
7	43	129	2752	0.95
10	61	183	5551	0.97
11	67	201	6700	0.97
12	73	219	7957	0.97
13	79	237	9322	0.97
16	97	291	14065	0.98
17	103	309	15862	0.98
18	109	327	17767	0.98

Remember ν are the number of parity checks and b the number of codeword bits.

Table 2: Valid Parameters of a Kirkman Triple System

LDPC Notes

1. The min distance of a code is the minimum number of non-zero columns in the parity check matrix for which the sum equals zero[Wic95].

2. The LDPC codes from a Kirkman system will have low d_{min}[MD00].

3. For short block lengths and high rates, the girth of the belief networks is large and assists in decoding.

4. KTS have better d_{min} compared to STS.

5. STS have $d_{min} \leq 10$[MD00].

6. Simulations in [JW01] done for lengths ≈ 500 suggest randomly constructed LDPC's perform as well as KTS for half rate codes.

7. Simulations on rate $\frac{2}{3}$, codes with block length ≈ 100 suggest KTS outperform randomly generated LDPC's. Euclidean Geometry codes outperform both these methods at a cost of computational complexity.

References

[30-02] 30–30 Cartridge. http://www.gmdr.com/lever/3030wtext.htm. nov 2002.

[ACT88] W.L. Abbott, J.M. Cioffi, and H.K. Thapar. Offtrack interference and equalization in magnetic recording. *IEEE Transactions on Magnetics*, 24(6):2964–2966, nov 1988.

[ADDC01] Mohammed Zaki Ahmed, Terry Donnelly, Paul J. Davey, and Warwick W. Clegg. Increased areal density using a 2-dimensional approach. *IEEE Transactions on Magnetics*, 37(4):1896–1898, jul 2001.

[ADDC02] Mohammed Zaki Ahmed, Terry Donnelly, Paul J. Davey, and Warwick W. Clegg. Track squeeze using adaptive inter-track interference equalisation. *IEEE Transactions on Magnetics*, 38(5):2331–2333, sep 2002.

[ADDC03] Mohammed Zaki Ahmed, Terry Donnelly, Paul J. Davey, and Warwick W. Clegg. High rate low density parity check codes for multi-track recording. *Conference Proceedings of Intermag, Boston*, 2003.

[Bar90] L.C. Barbosa. Simultaneous detection of readback signals from interfering magnetic recording tracks using array heads. *IEEE Transactions on Magnetics*, 26:2163–2165, sep 1990.

[BCJR74] L.R. Bahl, J. Cocke, F. Jelenik, and J. Raviv. Optimal decoding of linear codes for minimising sybmol error rate. *IEEE Transactions in Information Theory*, 20:284–287, mar 1974.

[Ber94] H. Neal Bertram. *Theory of Magnetic Recording*. 1994. ISBN 0 5214 4512 4.

[BGT93] C. Berrou, A. Glavieux, and P. Thitimajshima. Near shannon limit error-correcting coding and decoding: Turbo-codes. *IEEE International Conference on Communication*, pages 1064–1070, 1993.

[BO97] V. Bane and M. Olgica. Cyclic two dimensional IT reducing codes. *IEEE Intl. Symposium on Information Theory*, page 414, 1997.

[Bou94] P.A.H. Bours. *Codes for correcting insertion and deletion errors.* Phd, Dept. of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands, june 1994.

[BT64] W.G. Bickley and R.S.H.G. Thompson. *Matrices Their Meaning and Manipulation.* The English University Press, 1964.

[Bur01] Alister Burr. *Modulation and Coding for Wireless Communications.* Prentice Hall, 2001. ISBN 0 2013 9857 5.

[CEGH00] Jonathan Darrel Coker, Evangelos Stavros Elsftheriou, Richard Leo Galbraith, and Walter Hirth. Method and apparatus for decoding the output signal of a partial-response class iv communication or recording-device channel. *United States Patent*, (6,104,766), sep 2000.

[CGK+91] J.D. Cocker, R.L. Galbraith, G.J. Kerwin, J.W. Rae, and P.A. Ziperovich. Implementation of prml in a rigid disk drive. *IEEE Transactions on Magnetics*, 27(6):4538–4543, nov 1991.

[Con98] T. Conway. A new target response with parity coding for high density magnetic recording channels. *IEEE Transactions on Magnetics*, 34(4):2382–2386, jul 1998.

[Dav94] P.J. Davey. *Channel Coding Techniques for a multiple track digital magnetic*

recording system. Phd, School of Electronic, Communication and Electrical Engineering, University of Plymouth, oct 1994.

[DDMD98] Paul J. Davey, Terry Donnelly, Des J. Mapps, and Neil Darragh. Multitrack RLL codes for the storage channel with immunity to intertrack interference. *IEEE Transactions on Magnetics*, 34(4), sep 1998.

[Dev86] Luc Devroye. *Non-Uniform Random Variate Generation.* Springer Verlag, jul 1986. ISBN 0 0387 9630 5.

[DUH86] Francois B. Dolivo, Gottfried Ungerboeck, and Thomas D. Howell. Method and apparatus for decoding the output signal of a partial-response class iv communication or recording-device channel. *United States Patent*, (4,571,734), feb 1986.

[FFKM99] J. Fan, A. Friedmann, E. Kurtas, and S. McLaughlin. Low density parity check codes for magnetic recording. *Proc. Allerton Conference on Communication, Control and Computing*, 37, 1999.

[FHN86] H.C. Ferreira, J.F. Hope, and A. Nel. Binary rate 4/8 runlength constrained, error correcting magnetic recording modulation code. *IEEE Transactions on Magnetics*, 22:1197–1199, sep 1986.

[Fra68] P.A. Franaszek. Sequence-state coding for digital transmission. *The Bell System Technical Journal*, (47):143–157, jan 1968.

[Fun82] P. Funk. Run-length-limited codes with multiple spacing. *IEEE Transactions on Magnetics*, (18):772–775, 1982.

[Gal63] R.G. Gallager. *Low-Density Parity-Check Codes.* Phd, MIT, 1963.

[Gal68] Robert G. Gallager. *Information Theory and Reliable Communication.* John Wiley & Sons, Inc., 1968. ISBN 0 4712 9048 3.

[GR00] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products.* Academic Press, 6th edition, jul 2000. ISBN 0 1229 4757 6.

[Hay84] Simon Haykin. *Introduction to Adaptive Filters.* Macmillan Publishing Company, 1984. ISBN 0 0294 9460 5.

[Her77] A. Herk. Side fringing fields and write and read crosstalk of narrow magnetic recording heads. *IEEE Transactions on Magnetics*, 13(4):1021–1028, jul 1977.

[HM97] Bahram Honary and Garik Markarian. *Trellis Decoding of Block Codes: A Practical Approach.* 1997.

[IBM02a] IBM History of Storage. http://www.storage.ibm.com/hdd/firsts/index.htm. nov 2002.

[IBM02b] IBM Track, Linear and Areal Density Growth Chart. http://www.storage.ibm.com/hdd/technolo/grochows/g08.htm. nov 2002.

[Imm91] Kees A.S. Immink. *Coding Techniques for Digital Recorders.* Prentice Hall, 1991. ISBN 0 1314 0047 9.

[Jam99] Glyn James. *Advanced Modern Engineering Mathematics.* Prentice Hall, 2nd edition, 1999. ISBN 0 2015 9621 1.

[Jr.67] G.D. Forney Jr. Final report on a coding system designed for advanced solar missions. *NASA Ames Research Center*, dec 1967.

[Jr.73] G.D. Forney Jr. The viterbi algorithm. *Proceedings of IEEE*, 61(3):268–278, mar 1973.

[JW01] Sarah J. Johnson and Steven R. Weller. Construction of low-density parity-check codes from kirkman triple systems. *IEEE Globecom*, 2:970–974, nov 2001.

[KF98] F.R. Kschischang and B. Frey. Iterative decoding of compound codes by probability propagation in graphical models. *IEEE Journal on Selected Areas in Communication*, 16(2):219–230, 1998.

[Kob71] H. Kobayashi. Application of probabilistic decoding to digital magnetic recording systems. *IBM Journal on Research and Development*, (15):64–74, jan 1971.

[Kre66] E. R. Kretzmer. Generalisation of a technique for binary data communication. *IEEE Transactions on Communication Technology*, 14:67–68, feb 1966.

[KS91] R. Karabed and Paul H. Siegel. Matched spectral null codes for partial response channels. *IEEE Transactions in Information Theory*, 37:818–855, 1991.

[KT70] H. Kobayashi and D.T. Tang. Application of partial-response channel coding to magnetic recording systems. *IBM Journal on Research and Development*, (14):368–375, nov 1970.

[KV00a] R. Koetter and A. Vardy. Algebraic soft-decision decoding of Reed-Solomon codes. *IEEE International Symposium on Information Theory*, jun 2000.

[KV00b] R. Koetter and A. Vardy. Algebraic soft-decision decoding of Reed-Solomon codes. *38th Annual Allerton Conference on Communication, Control and Computing, Monticello, IL*, oct 2000.

[KZ99] A. Kato and K. Zeger. On the capacity of two-dimensional RLL recording codes. *IEEE Transactions in Information Theory*, 45(5):1527–1540, jul 1999.

[LC82] Shu Lin and Daniel J. Costello. *Error Control Coding: Fundamentals and Applications*. Prentice Hall, oct 1982. ISBN 0 1328 3796 X.

[Lev66] V.I. Levenstein. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Phys. Dokl.*, 10:707–710, 1966.

[Lin78] D.A. Lindholm. Spacing losses in finite track width reproducing systems. *IEEE Transactions on Magnetics*, 14(2):55–59, mar 1978.

[LM94] Jaejin Lee and Vijay K. Madisetti. Multitrack RLL codes for the storage channel with immunity to intertrack interference. *IEEE Globecom*, pages 1477–1481, nov 1994.

[LMSS98] Michael G. Luby, Michael Mitzenmacher, M. Amin Shokrollahi, and Daniel A. Spielman. Improved low-density parity-check codes using irregular graphs and belief propagation. *IEEE International Symposium on Information Theory*, aug 1998.

[LW87] Patrick Lee and Jack Keil Wolf. Combined error correction/modulation codes. *IEEE Transactions on Magnetics*, 23(5):3681–3683, sep 1987.

[Mac99] David J. C. MacKay. Good error-correcting codes based on very sparse matrices. *IEEE Transactions in Information Theory*, 45(2):399–431, mar 1999.

[Mac00] David J. C. MacKay. Turbo codes are low density parity check codes. *personal communication*, MacKay Draft 0.2, 2000.

[Mal] John Mallinson. personal communication.

[Map01] D.J. Mapps. *Data Storage Lecture Notes.* University of Plymouth, UK, 2001. In House Publication.

[MB96] J. Moon and B. Brickner. Maximum transition run codes for data storage systems. *IEEE Transactions on Magnetics*, 32(5):3992–3994, sept 1996.

[MD00] David J. Mackay and M.C. Davey. Evaluation of gallager codes for short block length and high rate applications. *Proc. of IMA Workshop on Codes, Systems and Graphical Models*, 123:113–130, 2000.

[MH97] H. H. Manoukian and B. Honary. BCJR trellis construction for binary linear block codes. *IEE Proceedings — Communications*, 144(6):367–371, December 1997.

[MMC98] Robert J. McEliece, David J.C. MacKay, and Jung-Fu Cheng. Turbo decoding as an instance of pearl's "belief propagation" algorithm. *IEEE Journal on Selected Areas in Communication*, 16(2):140–152, feb 1998.

[MMH98] Laura L. McPheters, Steve W. McLaughlin, and Eric C. Hirsch. Turbo codes for PR4 and EPR4 magnetic recording. *Asil Omar Conference Pacific Grove, CA On Signals Systems and Computers*, pages 1778–1782, nov 1998.

[MN97] David J.C. MacKay and R.M. Neal. Near shannon limit performance of low density parity check codes. *Electronic Letters*, 33(6):457–458, mar 1997.

[Moi00] B.E.B. Moision. *Constrained coding and detection for magnetic recording channels*. Phd, University of California, San Diego, 2000.

[MS84] Brian Marcus and Paul H. Siegel. Constrained codes for PRML. Technical Report 4371 (47629), IBM Research Division, Almaden Research Center, USA, jul 1984. Declassified 1989.

[MW92] M.W. Marcellin and H.J. Weber. Two dimensional recording codes. *IEEE Journal on Selected Areas in Communication*, 10(1):254–266, jan 1992.

[PH98] V.S. Pless and W.C. Huffman. *Handbook of Coding Theory*. Elsevier, Amsterdam, 1998. ISBN 0 4445 0088 X.

[Pro95] John G. Proakis. *Digital Communications*. McGraw-Hill, third edition, 1995.

[PTVF92] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C : The Art of Scientific Computing*. 2nd edition, 1992. ISBN 0 5214 3108 5.

[RML99] B.G. Roh, J. Moon, and S.U. Lee. Single-head detection in the presence of offtrack interference. *IEEE Globecom*, pages 955–959, sept 1999.

[RMM98] William E. Ryan, Laura L. McPheters, and Steve W. McLaughlin. Combined turbo coding and turbo equalization for PR4-equalized lorentzian channels. *Conference on Information Sciences and Systems*, pages 489–494, mar 1998.

[RS60] I. S. Reed and G. Solomon. Polynomial codes over certain finite fields. *SIAM Journal of Applied Math.*, 8:300–304, 1960.

[RU01a] Thomas J. Richardson and Rüdiger L. Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Transactions in Information Theory*, 47(2):599–618, feb 2001.

[RU01b] Thomas J. Richardson and Rüdiger L. Urbanke. Efficient encoding of low-density parity-check codes. *IEEE Transactions in Information Theory*, 47(2):638–656, feb 2001.

[Rys63] Herbert John Ryser. *Combinatorial Mathematics.* The Mathematical Association of America, 1963. Library of Congress Catalog Number: 63-12288.

[SC00] Hongxin Song and J.R. Cruz. Block turbo codes for magnetic recording channels. *IEEE International Conference on Communication*, pages 85–88, sept 2000.

[SG94] Emina Soljanin and Costas N. Georghiades. Sliding-block codes for two-track magnetic recording channels. *Annual Conference on Information Sciences and Systems*, pages 29–33, mar 1994.

[Sha48] Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, jul 1948.

[Sie85] Paul H. Siegel. Recording codes for digital magnetic storage. Technical Report 4720 (50198), IBM Research Laboratory, Almaden Research Center, USA, may 1985.

[Sol94] Emina Soljanin. *Coding for improving noise immunity in multi-track, multi-head recording systems.* Phd, Texas A&M University, 1994.

[SPHK95] Mohammed Siala, Eckhard Papproth, Kaïs Hajtaieb, and Ghassan Kawas Kaleh. A new iterative soft-output viterbi algorithm and its application to the "dicode" channel. *Ann. Télécommun.*, 50(n 9-10):798–816, 1995.

[STC01] Hongxin Song, Richard M. Todd, and J.R. Cruz. Low density parity check codes for magnetic recording channels. *IEEE Journal on Selected Areas in Communication*, 19:918–923, may 2001.

[Sud97] Madhu Sudan. Decoding of Reed Solomon codes beyond the error-correction bound. *Journal of Complexity*, 13(1):180–193, 1997.

[SW92] Robert E. Swanson and Jack K. Wolf. A new class of two-dimensional RLL recording codes. *IEEE Transactions on Magnetics*, 28(6):3407–3416, nov 1992.

[Tan81] R. Michael Tanner. A recursive approach to low complexity codes. *IEEE Transactions in Information Theory*, 27(5):533–547, Sep 1981.

[Tar95] Alexander Taratorin. *PRML: A Practical Approach Introduction to PRML Concepts and Measurements.* Focal Press, 1995.

[TP87] H. K. Thapar and A.M. Patel. A class of partial response systems for increasing storage density in magnetic recording. *IEEE Transactions on Magnetics*, 23(5):3666–3668, sept 1987.

[Vea93] Mathew P. Vea. *Intertrack Interference in high density recording: Modelling and equalisation.* Phd, Carnegie Mellon University, 1993.

[VM91a] Mathew P. Vea and Jose M.F. Moura. Detection in magnetic recording with cross-track interference. *Proc. Asilomar Conference on Signals, Systems and Computers*, (25):25–29, 1991.

[VM91b] Mathew P. Vea and Jose M.F. Moura. Magnetic recording channel model with intertrack interference. *IEEE Transactions on Magnetics*, 27(6):4834–4836, nov 1991.

[Voo94] Paul Augustine Voois. *Two-dimensional signal processing for magnetic storage systems.* Phd, Stanford University University, 1994.

[Wat88] John Watkinson. *The Art of Digital Audio.* Focal Press, 1988.

[Wei00] Yair Weiss. Correctness of local probability propagation in graphical models with loops. *Neural Computation*, 12(1):1–41, 2000.

[Wic95] S.B. Wicker. *Error Control Systems for Digital Communication and Storage.* Prentice Hall, 1995.

[WJ60] B. Widrow and M.E. Hoff Jr. Adaptive switching circuits. *IRE Wescon Conv. Record*, pages 563–587, 1960.

[Woo00] Roger Wood. The feasibility of magnetic recording at 1 Terabit per square inch. *IEEE Transactions on Magnetics*, 36(1):36–42, jan 2000.

[WU86] J.K. Wolf and G. Ungerboeck. Trellis coding for partial response channels. *IEEE Transactions on Communication*, 34:765–773, aug 1986.

C and C++ Code

Please see attached CD. The subfolder complete has an electronic printed version, and the subfolder code has the code.