The Dissertation Committee for Sunil Gupta
certifies that this is the approved version of the following dissertation:

Improving Magneto-Optic Data Storage Densities using Nonlinear Equalization

Committee:

_____
Baxter Womack, Supervisor

_____
Ishaq H. Unwala

_____
Earl Swartzlander

_____
Lizy John

_____
Margarita Jacome

Improving Magneto-Optic Data Storage Densities using Nonlinear Equalization

by

Sunil Gupta, B.S., M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2004

Acknowledgments

I wish to thank the multitudes of people who helped me. Time would fail me to list all of them here. In general, I would like to thank the Electrical and Computer Engineering Department at the University of Texas at Austin for giving me the opportunity to pursue my doctoral studies.

Improving Magneto-Optic Data Storage Densities using Nonlinear Equalization

Publication No. _____

Sunil Gupta, Ph.D.
The University of Texas at Austin, 2004

Supervisor: Baxter Womack

In a magneto-optic data storage systems, nonlinear intersymbol interference (ISI) is one of the limiting factors for high storage density. The ISI characteristics gives insight into its nonlinear behavior. To reduce its effect equalization is used. Volterra and Hybrid (combination of Volterra and Linear) nonlinear equalizers are proposed and developed. Performance of these equalizers is compared to that of linear equalizers.

The Hybrid equalizer provided the best trade-off of performance. The gain achieved in signal quality by using 3-tap 3rd order Volterra and N-tap linear was about 1.3-1.5 dB over N-tap linear equalizer. It was observed that going to higher order Volterra equalizers (e.g., 5-tap 5th order) does not yield any extra benefit as one reaches a point of diminishing return immediately after 3-tap 3rd order.

Convergence issues related to adaptive algorithms were also investigated. It was found that the eigenvalue spread of the covariance matrix was

large for the Volterra/Hybrid equalizers compared to linear equalizers which resulted in somewhat slower convergence. This can be mitigated by employing other algorithms such as multiple-step LMS or Kalman algorithms.

The extra gain in signal quality obtained can be utilized in various ways such as - achieving higher storage density, less expensive read/write optical assembly, lower BER (bit error rate) and less stringent FEC (Forward Error Correction) in the overall system design. The Volterra/Hybrid equalizers proposed and developed here can be applied to other communication systems (e.g., wireless, optical) exhibiting nonlinear ISI channel characteristics.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Optical data storage has become important for its high storage density, random access, rewritability, portability, and long archival life. An areal storage density of 0.5×10^6 bits/mm^2 has been reported [1]. In addition to the high density, the ability to read and write small bits with a large mechanical clearance leads to many favorable attributes in optical storage. It enables second-surface operation, which greatly improves immunity to surface dirt and scratches, and allows good isolation of the active film from the atmosphere. This in turn permits the removal of media, consequently enabling the drive to be used with any medium (erasable, non-erasable and read only). This disk portability requires no extra servo requirement because of the stricter requirement for high density tracking. Furthermore, because of this mechanical clearance, we can have longer archival life and there is no possibility of head "crashes" that destroy stored data. Therefore, optical drives can have lesser shock and vibration control as compared to magnetic storage.

There are different ways of increasing storage densities, such as 1) Using smaller wavelength laser such as blue lasers which have wavelength of around 0.4 micron (400 nm), which results in smaller spot size (minimum mark size).

1

2) Using better magnetic-optic recording material i.e., material which has less heat diffusion, which may reduce the ISI. 3) Using better encoding/decoding schemes. 4) Using smaller and lighter read/write head assembly. 5) Improving signal quality by using techniques from communication, signal processing, digital filtering and architecture.

In this dissertation the main effort is devoted to improving storage density using option 5, in which nonlinear equalization techniques are examined.

1.1 Issues in Optical Data Storage and Dissertation Objective

Although optical data storage has large storage densities, it has longer access time and lower transfer rates than magnetic storage. Longer access time is due to heavier laser heads, and lower transfer rates are due to limited laser power. Furthermore, not every optical storage system is rewritable. M-O and phase change [2] systems have been developed in the past for rewritability. However, in these systems, writing data takes two cycles: one cycle to erase and one cycle to write. Therefore, this is more time consuming than magnetic storage.

In this dissertation, various techniques are explored to increase the storage density for Magneto-Optic Systems.

Pulse broadening (can be actually shrinking) is the change of the pulse width when it is written and read back again from the original. When a sequence of marks is written to a disk, the real size of each mark is not deter-

2

mined by itself but also by adjacent marks. When marks are read, each pulse width is also subject to itself and neighboring marks. This change in pulse due to neighboring marks is called ISI. When the neighboring marks have large widths, the ISI effect is not important because the central mark is more like an isolated mark. However, when mark widths become smaller as we use shorter wavelengths, this ISI effect becomes important and a careful ISI equalization is necessary.

Thus, the objective of this dissertation is to determine the characteristics of intersymbol interference (ISI) and pulse broadening and to evaluate how much signal to noise ratio can be improved. The ISI characteristics gives us insight into its nonlinear nature. After that to compensate it (ISI), equalization is performed. Various configurations are evaluated with focus on the nonlinear filters such as Volterra and Hybrid. The performance is compared using a quantitative measure.

1.2   Organization of Dissertation

In the remainder of the dissertation, Chapter 2 reviews the principles of magnetic and optical storage. It also describes the pulse width modulation and encoding/decoding mechanism used along with an illustration of ISI. Chapter 3 covers the ISI characteristics in optical data storage system. Chapter 4 states the problem definition, defines the Volterra equalization and shows the convergence analysis. Furthermore, previous work in this area is summarized. Chapter 5 shows the performance comparisons of Linear, Volterra and Hybrid

3

equalization schemes. Chapter 6 summarizes the main conclusions of this dissertation and lists some items for further exploration in this field.

Chapter 2

Principles of Optical Data Storage and Pulse Width Modulation

To serve as the background for the following chapters on ISI characterizations and signal processing, this chapter reviews important principles of optical storage systems. Specifically, we explain the read/write principles of magneto-optical storage systems and pulse width modulation (PWM).

2.1 Configuration of Optical Recording System

The basic configuration of a typical magneto-optical (M-O) recording system is shown in Fig. 2.1 [2]. The system consists of a laser diode, two photodetectors, and optical components including polarizing beam splitters and lenses. The M-O disk used has a pre-grooved substrate and is coated with a ferro-magnetic thin film layer consisting of Tb-Fe-Co alloy. This magnetic layer is polarized vertically (up or down) and used for recording 0's and 1's. To control the polarization or record data, an electromagnet is placed on the disk to supply magnetic field at the recording point.

The first polarizing beam splitter polarizes the light from the laser and focuses it on the disk where it is used for writing and sensing the polarization
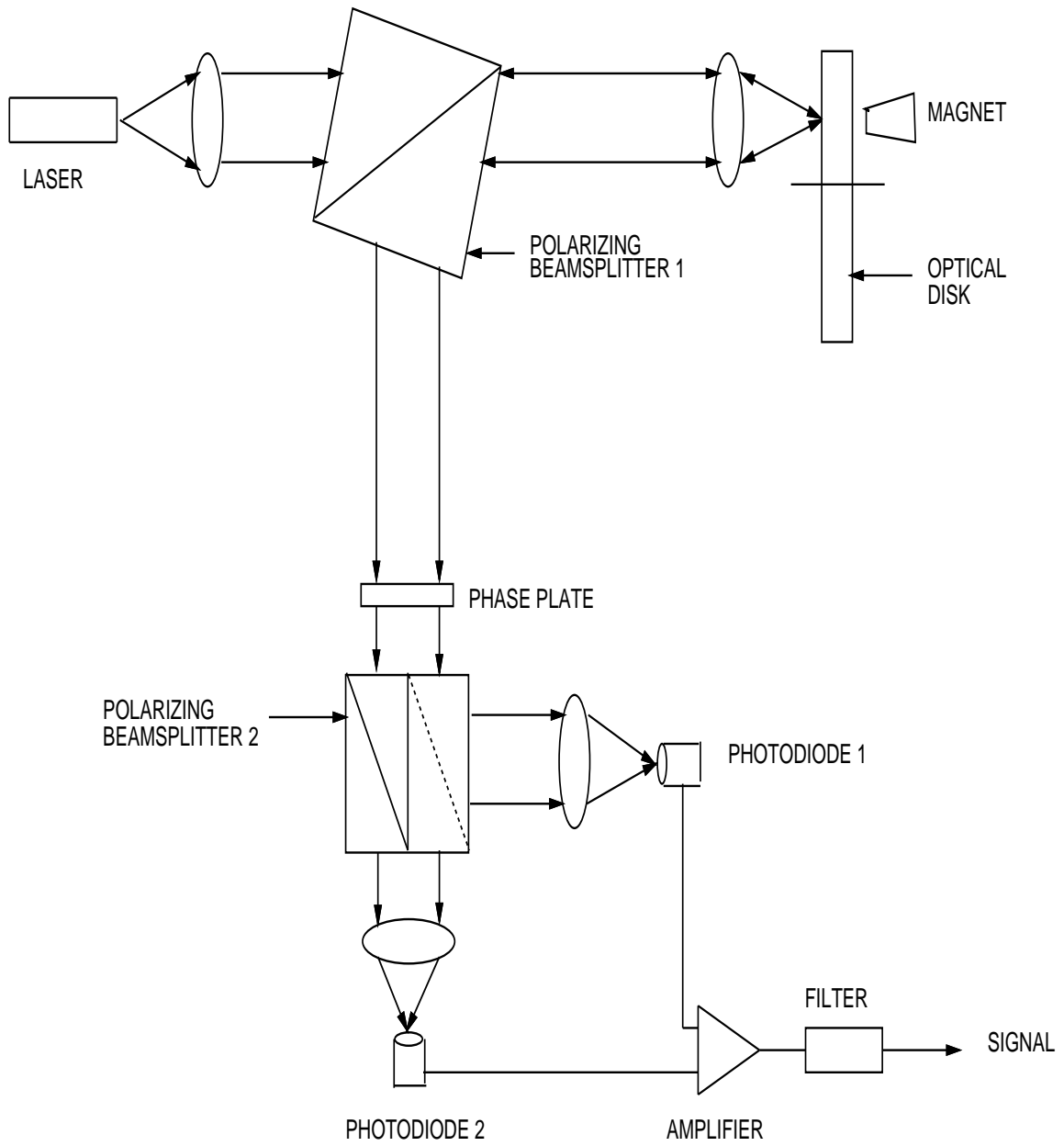
Figure 2.1: Schematic layout of an Optical Data Storage System.

6

when reading. When the polarized light that goes to the disk is reflected, it goes back to the first beam splitter and is reflected to the second beam splitter. This second polarizing beam splitter is rotated 45 degrees about the optical axis so that the incident light is split between the two photodiodes. The detected signals from the photodiodes are used to sense the magnetic polarizations and to provide servo control for laser beam focusing and tracking to maintain the laser spot in precise position relative to the disk surface and data tracks.

2.2   Principles of optical data storage

This section explains in more detail how data can be recorded and read back. Fig. 2.2 shows the basic principle of magneto-optic recording. The data is recorded by changing the direction of polarization of the magnetic film on the platter. On one side of the platter is the laser which heats the material and on the other side is the electro-magnet which causes the change in direction depending upon its polarity.

The write and read processes are further illustrated in Fig. 2.3. To record data, the recording M-O layer is first polarized uniformly in the same direction before recording. To change the polarization say from up to down, the electromagnet head is supplied a different direction current for opposite magnetization direction. However, if there is no laser beam on the same spot, the polarization cannot be changed because the Curie temperature of the medium is higher than the room temperature. When the focused laser spot irradiates

7
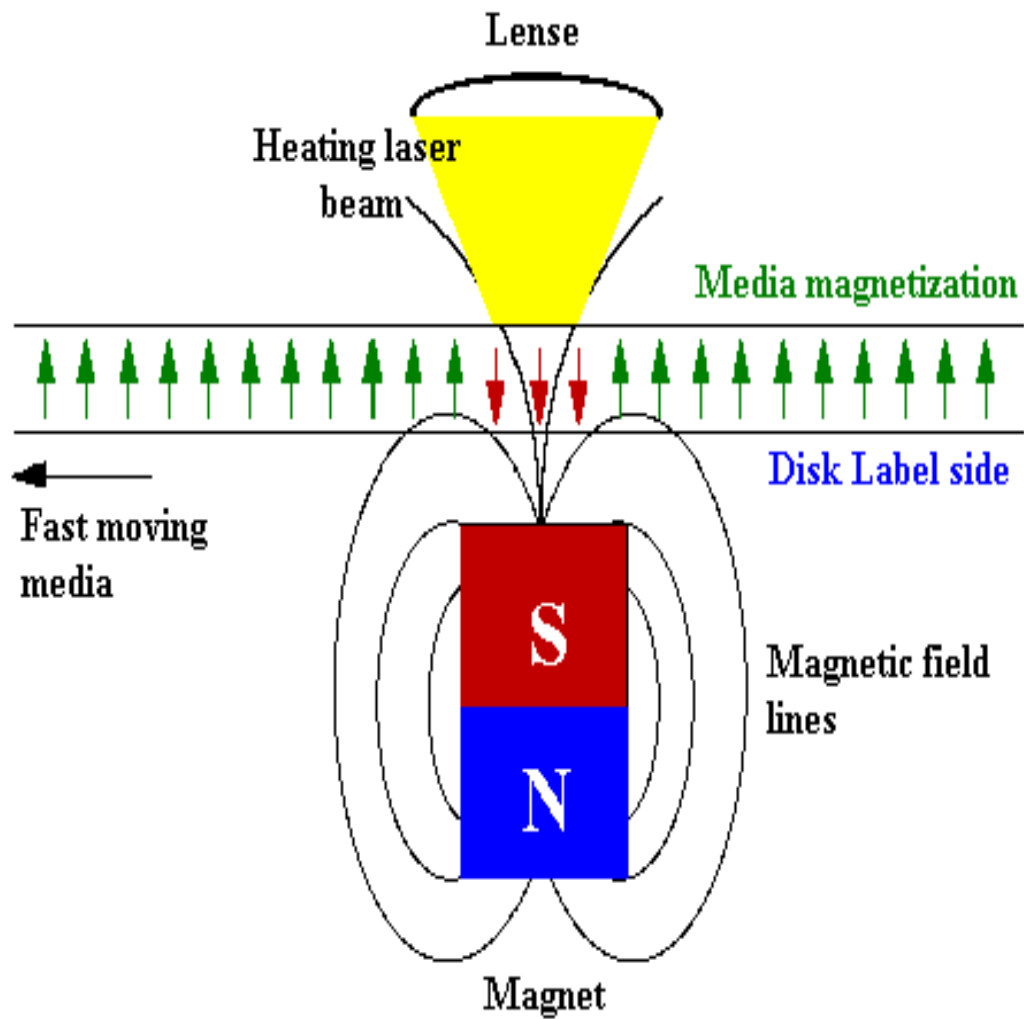
Figure 2.2: Magneto-Optic Data Storage Principle. Figure Courtesy of US-Byte - http://www.usbyte.com/common/MOsystems.htm

8

WRITE PROCESS   READ PROCESS

ELECTROMAGNET

RECORDING LAYER

KERR ROTATION

LASER
BEAM

INCIDENT   REFLECTED
BEAM   BEAM
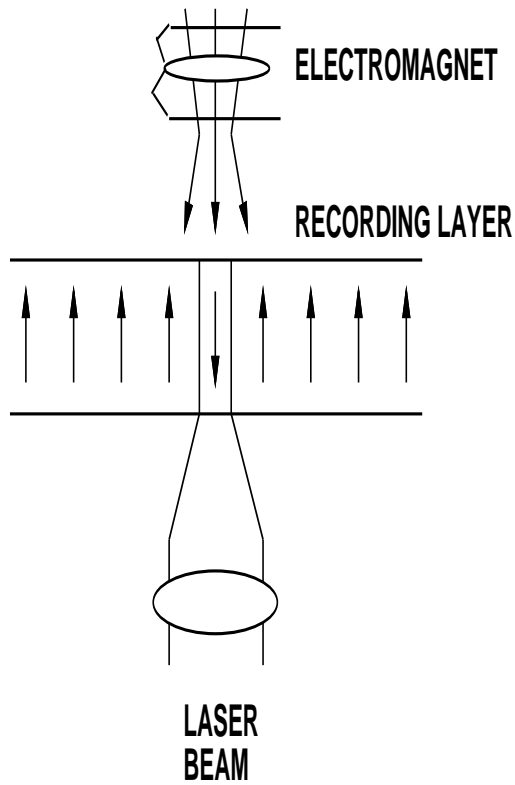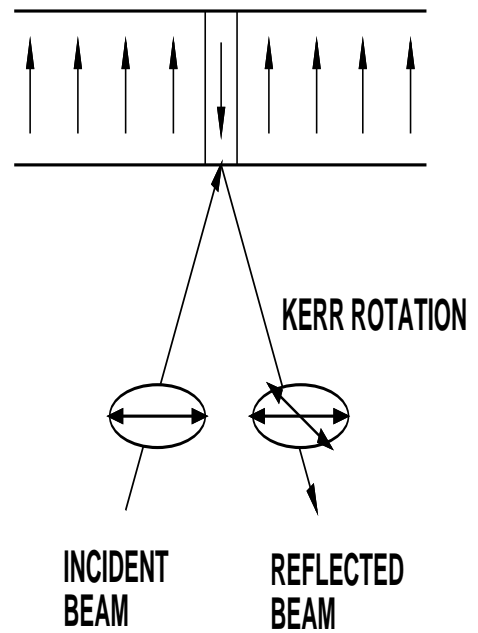
Figure 2.3: Read/Write Mechanism in Optical Data Storage System.

9

and heats the recording layer up to the Curie temperature, the initial magnetization of the recording layer is changed from up to down at the recorded domain. Erasure of the recorded domain can be done simply by repeating the same process with up magnetic polarization.

To read back the recorded data in the read process, we must be able to detect the polarizations in the medium. To do this, the Kerr effect is used [2], which can be understood as follows. When a linearly polarized laser beam irradiates the recording layer (small enough power to avoid changing the polarizations), the polarization of the reflection beam will be determined by the polarization of the irradiated spot. The change in reflected beam polarization angle with respect to incident beam is termed as Kerr rotation. See Fig. 2.3. This rotation is detected by the photodetectors.

From the explanation of the read/write principles, we see that optical recording is done through modulating the magnetic polarization of the recording layer. That is, information bits are stored through modulating marks alternating in polarization or state. In the next section, we will explain how to translate bits into alternating marks and vice versa.

To serve as the background for the following sections on Linear and Nonlinear Equalization, this chapter reviews the important principles of Magnetic and Optical storage systems. Specifically, we explain the read/write principles of magnetic and optical systems and pulse width modulation (PWM).

Optical storage is of two types - Magneto-Optical Technology and Phase

10

change Technology.

2.2.1 Advantages of Optical Storage

Our need for storage is explosive; fueled by multimedia requirement for text, images, video and audio, storage requirements are growing at an exponential rate and are expected to exceed 10^{20} bits (12 exabytes) in 2002. With an increasing amount of information being generated or captured electronically, a large fraction, perhaps as much as 40%, will be stored digitally. To meet this need, the hierarchy of on-line, near-line and off-line storage systems will be composed of many diverse technologies: magnetic disk drives, magnetic tape drives and tape libraries, optical disk drives and optical libraries. The mix of these sub-systems will be very application-specific so as to optimize the performance and cost of the overall system.

An optical storage system is a particularly attractive component of this hierarchy because it provides data access times that are an intermediate solution between a hard disk drive and a tape drive. Access time is the time, including latency, required to start retrieving a random block of data and typically ranges from less than 10 msec for a hard disk drive, to 30-50 msec for an optical disk drive, and several seconds for a tape drive. It becomes an important link in the chain as data are staged up and down between cpu, memory, and storage.

Perhaps the most enabling feature of optical storage is the removability of the storage medium. With separations of a few millimeters between the

11

recording surface and the optical 'head', and with active servos for focusing and tracking, the medium can be removed and replaced with relatively loose tolerances. The infamous head crashes often experienced in hard disk drives do not occur in optical drives. Data reliability and removability are further enhanced by using the 1.2 mm transparent disk substrate as a protective cover to keep contamination away from the recording surface. (The 2003 announcement of digital versatile disk or DVD will have a substrate thickness of only 0.6 mm.)

Removability has created a whole new industry in CD-audio. CD-ROMs have enhanced the efficiency of distribution and use of software, games, and video. These read-only disks containing 680 MB of information can be mass replicated by injection molding in a few seconds for less than 10 cent each, and they are virtually indestructible. One-day express mail of 125 CDs is an effective data rate of 1.0 MB/s.

At the other end of the spectrum, phase change and magneto-optical disks are used in WORM (write-once-read-many) and read/write/erase systems where a single disk can contain almost 5 GB. Using robotics, storage libraries can be assembled with capacities of a terabyte with access to any disk in under 5 seconds.

This review discusses the fundamentals of optical storage, the components that comprise a system, and the emerging technologies that will allow increased performance, higher storage capacities, and lower cost.

12

2.2.2 Recording and Readout of Information on Optical Disks

An optical disk is a plastic (or glass) substrate with one or more thin-film layers coated upon its surface(s). The information may be prerecorded on the surface of the substrate by the disk manufacturer, or it may be recorded on one or more of the thin film layers by the user. Typical diameters of presently available optical disks are as follows:

12 cm, as in compact disk audio (CD), compact disk read-only memory (CD-ROM), compact disc recordable (CD-R), and digital versatile disk (DVD) 2.5", as in digital audio mini disk (MD) 3.5", as in magneto-optical (MO) removable storage devices 5.25", as in both magneto-optical and phase-change (PC) rewritable disks 12" and 14", such as the write-once-read-many (WORM) media manufactured by Kodak and other companies for high volume storage applications In the case of read-only media the information is pressed onto the substrate by injection molding of plastics, or by embossing of a layer of photopolymer coated on a glass substrate (Bouwhuis, *et al.* 1985; Marchant 1990). The substrate is then coated with a thin metal layer (e.g., aluminum) to enhance its reflectivity. In other types of optical disks, some information, such as format marks and grooves, may be stamped onto the substrate itself, but then the substrate is coated with a storage layer that can be modified later by the user during recording of information. Typical storage layers are dye-polymer films for write-once applications, tellurium alloys for ablative recording (also write-once), GeSbTe for phase-change rewritable media, and TbFeCo magnetic films for magneto-optical disks (also rewritable)

13

CD technology was developed jointly by Philips of the Netherlands and Sony of Japan in the late 1970s and early 1980s, then introduced to the market in 1983 as the hugely successful compact audio disk. Today most of the research and advanced development work in DVD (the successor to CD) and future generations of read-only media takes place in Japan. Sony, Matsushita, Pioneer, Toshiba, Sharp, and NEC are some of the leading Japanese companies in this area. The MiniDisc (MD) was pioneered by Sony, and despite a sluggish early start, has become a very successful product in Japan. The major players in the field of magneto-optical data storage are Sony, Fujitsu, Hitachi Maxell, Mitsubishi, Nikon, and Sanyo. In phase-change optical recording, the leaders are Matsushita, Toshiba, and NEC, recently joined by Sony. Many other companies are involved in the development of components (lasers, miniature optics, actuators, detectors, substrates, spindle motors, etc.) as well as mastering machines and test equipment for optical recording products. These include some of the major multinational corporations as well as smaller firms.

In a typical disk the storage layer is sandwiched between two dielectric layers, and the stack is capped with a reflector layer (e.g., aluminum or gold), and protected with a lacquer layer. The dielectric layers and the reflective layer perform several tasks: protecting the storage layer, creating an optically tuned structure that has optimized reflectivity and/or absorptivity, allowing the tailoring of the thermal properties of the disk for rapid cooling and reduction of thermal cross-talk during writing, and so on (Mansuripur 1995;

14

McDonald and Victora 1997).

The laser is usually a semiconductor laser diode, whose beam is collimated by a well-corrected lens and directed towards an objective lens through a beam-splitter. The objective lens focuses the beam onto the disk and collects the reflected light. This reflected light is directed (at the beam-splitter) towards the detectors, which produce a data readout signal as well as servo signals for automatic focusing and tracking. The functions and properties of the various elements of this system are described in some detail in the following sections.

2.2.3 The Light Source

All optical recording technologies rely on lasers as their source of light. The lasers used in optical disk and tape data storage are semiconductor laser diodes of the shortest possible wavelength that can provide sufficient optical power for read/write/erase operations over a period of several thousand hours. Presently the shortest wavelength available in moderate power lasers (around 50 mW) is in the neighborhood of 680 nm; these lasers are being used in CD-R, MO, and PC products. For CD and DVD applications where writing is not a concern, low power lasers (e.g., 5 mW) which emit at 630 nm and 650 nm are being considered. The emphasis on short wavelength lasers for optical recording applications is due to the fact that shorter wavelengths can be focused to smaller spots at the diffraction limit. All things being equal, the diameter of a focused spot scales with its wavelength: a reduction of the

15

wavelength by a factor of 2, for example, will result in a reduction of the focused spot diameter by the same factor and, consequently, a fourfold increase in the data storage density can be realized.

The wavelengths of optical data storage have continuously shrunk during the past 15 years; starting at 830 nm, they are now down to 630 nm, and there is every indication that they will continue to shrink in the foreseeable future. Nichia Chemicals Corporation of Japan recently announced that 50 mW blue lasers operating in the wavelength range of 370 nm to 420 nm would be available before the end of 2003. Several other Japanese companies (e.g., Sony, Matsushita, Pioneer) have demonstrated the feasibility of small, inexpensive, second harmonic generation (SHG) green and blue lasers for use in optical disk drives.

What is needed for optical data storage is a compact and inexpensive laser diode that can be incorporated into small, low-cost drives. The power requirement for such lasers is several milliwatts for read-only media and several tens of milliwatts for recordable media. The lasers should be capable of direct modulation (e.g., by modulating the electrical current input to the laser), otherwise the cost and the size of external modulators may be prohibitive. Spatial coherence and single transverse-mode operation is a requirement, because the beam must be focused to the diffraction limit. (In this context vertical cavity surface emitting laser diodes (VCSELs) as well as arrays of such lasers need to be improved, since, at high powers, these lasers tend to operate in high-order modes.) Low-noise operation is very important, especially in applications such

16

as magneto-optical readout, where signal-to-noise ratio is at a premium.

Although operation of laser diodes in several longitudinal modes is presently acceptable, for future devices it may be important to add operation in a single, stable, longitudinal mode to all their other desirable characteristics. Mode hopping, wavelength shifts of several nanometers with temperature fluctuations and with operating current, manufacturing variability of the wavelength from batch to batch, etc., are so severe that, at the present time, it is not possible to consider the use of diffractive lenses either for the collimator or for the objective lens. These high numerical aperture (NA), diffraction-limited lenses are still produced by molding of glass elements. In the future, when microminiaturization becomes a necessity and lenslet arrays begin to appear in commercial products, operation of the laser in a single, stable, longitudinal mode may be required.

2.2.4   The Objective Lens

Presently disk and tape optical recording systems use molded glass lenses for focusing the laser beam to a diffraction-limited spot. These lenses consist of two aspheric surfaces on a single piece of glass, have fairly large numerical apertures (in the range of 0.4 to 0.6), and are essentially free from aberrations. The numerical aperture of a lens is defined as $NA = sin\theta$, where θ is the half-angle subtended by the focused cone of light at its apex. A 0.5 NA lens, for example, will have a focused cone whose full angle is 60°. The diameter of the focused spot is of the order of λ_0/NA, where λ_0 is the vacuum

17

wavelength of the laser beam. It is thus clear that higher numerical apertures are desirable if smaller spots (and therefore higher recording densities) are to be attained. Unfortunately, the depth of focus of an objective lens is proportional λ_0/NA^2, which means that the higher the NA, the smaller will be the depth of focus. It thus becomes difficult to work with high NA lenses and maintain focus with the desired accuracy in an optical disk drive.

But a small depth of focus is not the main reason why the present optical drives operate at moderate numerical apertures. The more important reason has to do with the fact that the laser beam is almost invariably focused onto the storage medium through the disk substrate. The disk substrate, being a slab of plastic or glass, has a thickness of 1.2 mm (DVD substrates are only half as thick, or 0.6 mm). When a beam of light is focused through such a substrate it will develop an aberration, known as coma, as soon as the substrate becomes tilted relative to the optical axis of the objective lens. Even a 1° tilt produces unacceptably large values of coma in practice. The magnitude of coma is proportional to NA^3, and therefore, higher NA lenses exhibit more sensitivity to disk tilt. Another aberration, caused by the variability of the substrate's thickness from disk to disk, is spherical aberration. This aberration, which scales with the fourth power of NA, is another limiting factor for the numerical aperture. In the future, manufacturers will move toward higher numerical apertures by adopting one or more of the following strategies:

1. Use thinner substrates or avoid focusing through the substrate altogether.

18

2. Make the substrates as flat and as uniform as possible.

3. Develop servo mechanisms whereby the tilt and thickness variations of the disk are automatically sensed and corrected for in the optical path.

As was mentioned earlier, the use of smaller lenses is always desirable in optical storage technology, particularly if lens arrays are being considered for parallel accessing of multiple tracks in a system. In this respect, gradient-index (GRIN) lenses, holographic optical elements (HOEs), and binary diffractive optical (BDO) lenses are all being considered for future generations of optical storage devices.

2.2.5   Automatic Focusing

A typical optical disk has a plastic substrate that is not perfectly flat, but is slightly warped. Also, when mounted in a drive, small tilts of the axis could cause vertical motions of the disk surface during operation. It is not unusual to find vertical movements as much as $\pm 100 \mu m$ during the operation of an optical disk. A typical objective lens has a numerical aperture of 0.5 or higher, and therefore, the focused beam has a depth of focus of a fraction of λ_0/NA^2, which is only a fraction of a micron. The focused spot must remain within the depth of focus while the disk rotates at speeds of several thousand rpm and wobbles in and out of focus by as much as $\pm 100 \mu m$ in each revolution. Needless to say, without an autofocus mechanism to maintain the disk continually in focus, the operation of an optical disk drive is unworkable.

19

In practice the objective lens is mounted in a voice coil actuator (bandwidth = several kHz), and a feedback mechanism is used to drive the lens towards and away from the disk in such a way as to maintain focus at all times. The signal needed for this feedback mechanism is derived from the light that is reflected from the disk itself. The light reflected from the disk and collected by the objective lens is either convergent or divergent, depending on whether the disk is further away from best focus or closer to the lens than the plane of best focus. This returned beam goes through an astigmatic lens, which normally focuses the incident beam to a symmetric spot halfway between its focal planes. A quad detector placed at this plane (also called the plane of least confusion) then receives equal amounts of light on its four quadrants. When the disk is out of focus, however, the astigmatic lens creates an elongated spot on the detector. Depending on the sign of defocus, this elongated spot may preferentially illuminate quadrants 1 and 3 or quadrants 2 and 4 of the detector. Therefore, the combination signal (S1+S3) - (S2+S4) provides a bipolar focus-error signal, which is feed back to the voice coil for maintaining focus automatically

2.2.6 Automatic Track-Following

The information on an optical disk is recorded either around a series of concentric circular tracks or on a continuous spiral. Manufacturing errors and disk eccentricities caused by mounting errors, thermal expansion of the substrate, etc., will cause a given track to wobble in and out of position as

20

the disk spins. Typically, a given track might be as much as $\pm 100 \mu m$ away from its intended position at any given time. The focused spot, of course, is only about $1 \mu m$ across and cannot be at the right place at all times. An automatic tracking scheme is therefore desired. Given the mechanical rotation rates of the disks, the frequency response of the actuator needed for track following does not have to cover more then a few kHz, and a voice coil is usually sufficient for the purpose. The feedback signal, for controlling the position of the objective lens within the tracking coil, is again provided by the return beam itself. Several mechanisms have been proposed and have been put to use in commercial devices. Two of these schemes are mentioned here.

The push-pull tracking mechanism relies on the presence of either grooves or a trackful of data on the media. In the case of CD and CD-ROM, the data are prestamped along a spiral on the substrate, and the sequence of marks along the spiral comprises a sort of discontinuous groove structure. The discontinuity is irrelevant to the operation of the tracking servo, however, because it is at a much higher frequency than the tracking servo is designed to follow. Writable media such as CD-R, MO and PC require a tracking mechanism distinct from the data pattern, because prior to the recording of data, the write head must be able to follow the track before it can record anything. Once the data are recorded, the system will have a choice to follow the original tracking mechanism or to follow the recorded data pattern. Continuous grooves are a popular form of preexisting tracks on optical media. A typical groove is a fraction of a micron wide (say, $0.4 \mu m$) and one-eighth of a wavelength ($\lambda/8$)

21

deep. As long as the focused beam is centered on a track, diffraction of light from the adjacent grooves will be symmetric. The symmetry of the reflected beam, as sensed by a split detector in the return path, would produce a zero error signal. However, when the focused spot moves away from the center of the track, an asymmetry develops in the intensity pattern at the split detector. The bipolar signal thus generated from the difference signal is sufficient to return the focused spot to the center of the track.

In read-only media, the three-beam method of tracking has been extremely popular. The laser beam is divided into three beams, one of which follows the track under consideration, while the other two are focused on adjacent tracks, immediately before and after the desired track. Any movement of the central track away from its desired position will cause an increase in the signal from one of the outriggers and, simultaneously, a decrease in the signal from the other outrigger. A comparison of the two outrigger signals provides sufficient information for the track-following servo.

2.2.7 Disk Substrates

Polycarbonate is a fairly strong, inexpensive, moldable plastic that is currently the material of choice in optical disk fabrication. The pattern of pits and grooves is readily impressed onto the surface of this substrate during injection molding. The transparency of polycarbonate at red and near-infrared wavelengths allows the focused beam of light to reach the storage layer through the substrate, an important factor for removable disks, since focusing through

22

the substrate will keep dust, fingerprints, and scratches on the front facet of the disk well out of focus. Thermal and mechanical properties of polycarbonate are acceptable for present day needs, and its multitude of applications outside the field of optical storage has brought its price down to almost negligible values.

On the negative side, polycarbonate is a birefringent plastic and affects the polarization of the beam as it travels through the substrate. Substrate birefringence is particularly troublesome in MO data storage, where the read-out signal is embedded in the polarization state of the return beam. Careful balancing techniques applied during substrate manufacture have made it possible to reduce the in-plane birefringence to negligible values, but the remnant vertical birefringence and the fact that in-plane birefringence may return at elevated temperatures have kept alive the search for better materials. For example, amorphous polyolephin, which is essentially free from birefringence, has been shown to be an excellent substrate material. Its high price (because of its low volume of applications in other areas), however, has so far been an impediment to its use in optical data storage.

Transparency of plastics will become an issue in the future when very short wavelength lasers become available. Also, if disk flatness happens to be an issue, and recent trends toward flying optical heads and near-field recording indicate that it might be, then glass and aluminum substrates may be more suitable alternatives for these applications. Characteristics that are deemed desirable for glass to become an acceptable substrate material are low cost,

23

flatness, hardness, low stress birefringence, and the ability to be patterned with grooves and preformat marks.

2.2.8   Stacked Optical Disks and Double-Layer DVD

Magnetic drive manufacturers generally use a stack of several magnetic disks to achieve the desired storage capacity within the small volume of a hard drive. The small size and the low cost of magnetic heads allows them to use a separate head for each disk surface without compromising the overall size and price of the drive. In contrast, optical heads are very expensive and rather bulky. To achieve volumetric density with optical disks, designers have sought methods that rely on a single head for accessing multiple platters. In 1994 IBM researchers demonstrated a system that could read through a stack of six CD surfaces using a single optical head (Rubin, *et al.* 1994). Their system was very similar to a conventional CD player, except that special care was taken to correct the spherical aberrations that result when the beam of light is focused through a substrate at different depths. The IBM stack consisted of three thin glass disks (thickness $300\mu m$) on both surfaces of which the data pits had been embossed. Unlike standard CDs, these disks were not metallized because the focused laser beam had to pass through several such layers before reaching the desired surface. A bare glass surface reflects about 4% of the incident light which was apparently sufficient to enable the detection system to retrieve the data.

Technically, the method described above is straightforward and requires

www.manaraa.com

only that the objective lens be corrected for focusing through different thicknesses of the substrate. The separation between adjacent surfaces must be large enough to reduce cross-talk from the data marks recorded on neighboring surfaces. With a 0.5 NA objective lens, a separation of 40 or 50 microns is typically enough to assure acceptable levels of cross-talk from these other surfaces.

In the case of writable media, recording and readout of information on multiple platters is more difficult, primarily because storage layers absorb a significant amount of the laser light. (In their 1994 demonstration, IBM researchers also described a four-layer WORM disk and a two layer MO disk.) Recently double-layer MO and PC disks have been proposed and demonstrated in several industrial laboratories around the world. Again the separation between recording layers has been kept at about 40 microns to reduce cross-talk, and the laser beam has been strong enough to write on the second layer even after half of its power has been absorbed by the first layer. (The power density at the first layer is reduced by a factor of almost 2000; hence, no writing occurs at that layer.) Since focusing through an additional 40 microns of plastic is well within the range of tolerance of typical objective lenses, correction for spherical aberration has not been necessary in these double-layer systems.

Stacked optical disks have obvious advantages in terms of volumetric capacity, but the technological barriers for rewritable media are substantial.

2.2.9 Magneto-Optical Recording and Readout

Presently all commercially available MO disks are based on an amorphous terbium-iron-cobalt [Tbx(FeyCo1-y)1-x] magnetic alloy. Typical compositions have values of x=0.2 and y=0.9. This material belongs to a class of materials known as the rare earth-transition metal alloys. (Terbium is a rare earth element, while iron and cobalt both are transition metals.)

2.2.10 Land-Groove Recording

It has been found that by making the land and groove of equal width, and by recording the information on both lands and grooves, it is possible to eliminate (or at least substantially attenuate) the cross-talk arising from the pattern of marks recorded on adjacent tracks (Fukumoto, Masuhara and Aratani 1994). It turns out that for a particular groove depth (typically around $\lambda/6$) the cross-talk from adjacent tracks reaches a minimum. These results are in excellent agreement with the experimental data, indicating that cross-talk cancellation is a direct consequence of diffraction from the grooved surface and interference among the various diffracted orders.

Land-groove recording works well in phase-change media, and, in fact, this is where it was originally discovered. For MO systems, the birefringence of the substrate creates certain problems; specifically, the presence of in-plane birefringence will make the optimum groove depth for land reading different from that for groove reading. To overcome this problem either better substrates (with smaller in-plane birefringence) need to be developed, or a servo

26

system must be deployed to automatically correct the effects of birefringence. So far, the laboratory results of land-groove recording on MO disks have been very encouraging, and there is little doubt that this technique will play a major role in future generations of both PC and MO devices.

2.2.10.1   Magneto-Optical Technology:

It is a rewritable storage technology that uses a combination of magnetic and optical methods. Data is written on an MO disk by both a laser and a magnet. The laser heats the bit to the Curie point, which is the temperature at which molecules can be realigned when subjected to a magnetic field. A magnet then changes the bit's polarity. Writing takes two passes. MO disks do not have to be reformatted when full.

2.2.10.2   Phase Change Technology:

It is an optical storage technology in which the disk drive writes data with a laser that changes dots on the disk between amorphous and crystalline states. An optical head reads data by detecting the difference in reflected light from amorphous and crystalline dots. When full a phase change disk can be erased (or reformatted) using a medium-intensity pulse to restore the original crystalline structure. CD-RW uses phase change technology.

Optical recording is considered a future replacement for magnetic recording. Optical recording systems potentially have much greater reliability than magnetic recording since there is a much larger distance between the read/write

27

element and the moving media. Therefore, there is no wear associated with repeated use of the optical systems. However, there are other possible sources of trouble: the life and stability of the laser, mechanical damage to the relatively soft and exposed-to-the-environment media, mechanical damage due to shock and vibration, etc. Another advantage of the optical recording systems over the best performing magnetic recording systems - hard drives - is their removability.

The main disadvantage of optical storage when compared to magnetic is slower random data access. This partially comes from the design of the relatively large and heavy optical head. Plus, unlike the hard disk, an optical disk is usually removable, which limits rotational velocities and, correspondingly, limits access time. Increasing rpm causes the relatively loosely fixed CD-ROM disk to vibrate significantly compared to a stiff, fixed, and balanced hard disk. Figs. 2.1, 2.2 and 2.3 illustrate the read/write mechanism of optical storage as discussed earlier in this chapter.

2.3   Magnetic Storage

The basic principles of magnetic recording is as follows: Step 1. The drive channel electronics receive data in binary form from the computer and converts them into a current in the head coil. The current in the coil reverses at each 1 and remains the same at each 0.

Step 2. This current interaction with the media results in magnetization of the media, which direction depends on the current direction in the coil.

The reading process includes excitation of the current in the head coil when the head senses changes in the magnetic flux. The read voltage pulses at the flux transitions are then translated into sequences of bits equal to 0 and 1. The spacing loss factor postulates that the loss of magnetic signal power will be proportional to the media - head separation. This requires modern heads to fly at a few nanometers only. Fig. 2.4 illustrates the magnetic read/write mechanism. The original bit stream is feed to the write head where the direction of the current changes depending upon whether it is a one or a zero. This changes the direction of magnetization of the electromagnet which in turns changes the direction of magnetization of the recording media on the platter.

The partitioning of disks is as follows:

The concentric set of magnetic or optical disk bits is called a track. Each track is divided into sectors of 512 bytes (usually). Sectors have a sector header and an error correction code (ECC). In modern drives, sectors are numbered sequentially. A group of tracks with the same radius is called a cylinder. BIOS (Basic Input Output System) contains the number of cylinders, heads, and sectors for each drive. To improve performance and increase data rate, HDD (Hard disk drive) utilize a small amount of fast solid-state memory to store the most frequently used data. This memory is called a cache or a buffer.

The current State of the art is as follows:

The highest density of magnetic recording is 56 GB/sq.inch, belongs to
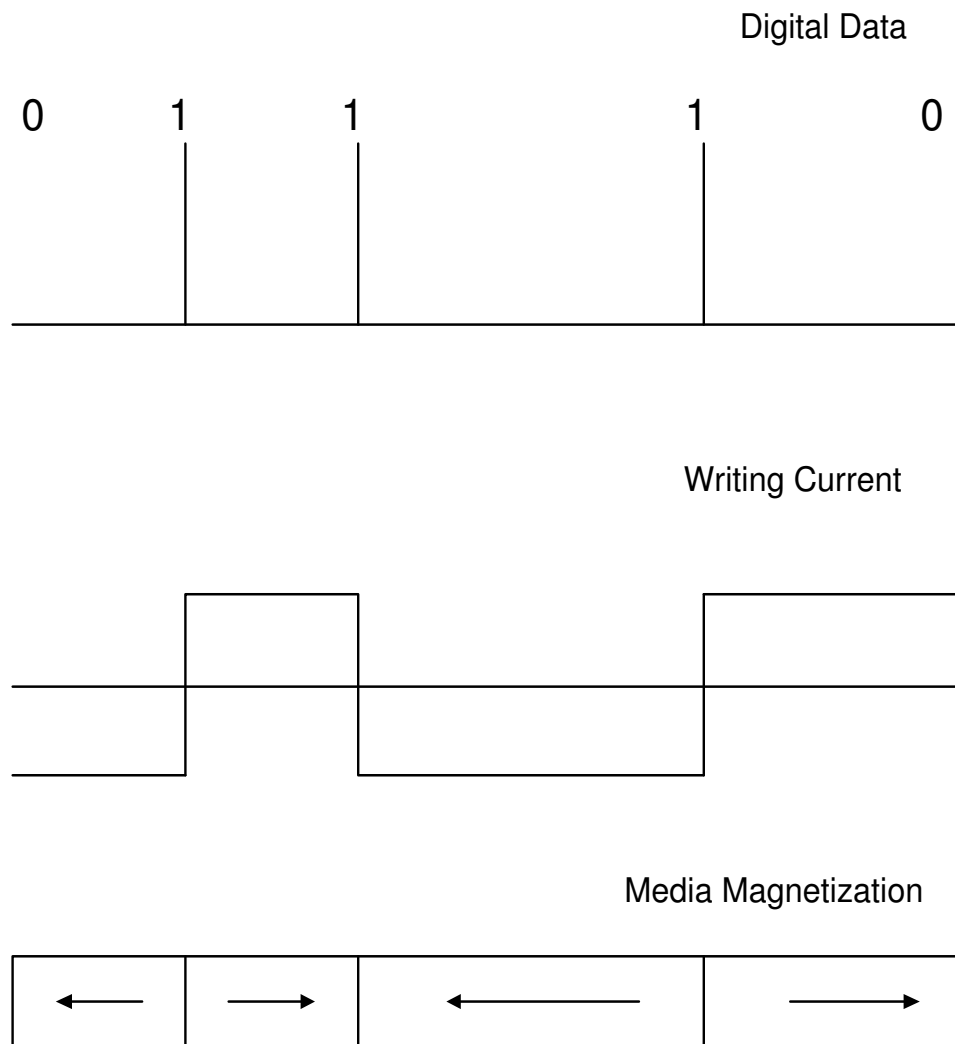
Figure 2.4: Read/Write Mechanism in Magnetic Data Storage System.

Fujitsu (April 7, 2000).

The current optical-storage technology is working close to the diffraction limit of 5 GB/sq.inch. This density can be substantially improved by using shorter wavelength lasers, e.g., blue lasers.

The fastest hard disk drive on the market (smallest average seek time) is the Seagate Cheetah X15 (18GB@15,000rpm) with average seek time of 3.9 ms.

Fastest CD-ROM on the market is the Kenwood 72x, which uses Newest 7-beam pickup technology and is as fast as 90X CD-ROM ($\sim 110Mbps$).

The next section explains how to translate bits into alternating marks and vice versa.

2.4 Pulse Width Modulation, Coding and Detection

In PWM, information bits are stored by modulating mark widths recorded on disks. Modulation codes are used to translate bits into mark widths. One important modulation code is called Run-length-limited (RLL) [2] [3]. Basically, RLL codes introduce additional constraints in the bit sequence before bits are differentially encoded. There are two constraints in RLL codes, denoted by (d,k). The d and k constraints specify that there are at least d and no more than k consecutive zeros between two ones at the output of the RLL encoder. Therefore, alternative marks using differential coding will have width at least (d+1) and not greater than (k+1). The k constraint guaran-

31

Table 2.1: (2,7) RLL Encoding Lookup Table

Data	Code Bits
10	0100
11	1000
000	100100
010	001000
011	000100
0010	00001000
0011	00100100

tees enough transitions for bit timing recovery in reading. The d constraint determines the minimum mark size at a given bit length. Therefore, at a given diffraction limit or minimum beam spot that can be achieved, the larger the d, the more encoded RLL bits can be recorded within the minimum mark. This will improve the recording density as long as the RLL encoding efficiency is not too small to offset the gain from the d factor. The RLL encoding inefficiency is from the fact that when m data bits are encoded into n bits, we have $n > m$. The coding efficiency is the expected value of m/n.

As an example, one very popular RLL code is the (2,7) code. Therefore the minimum mark consists of 3 encoded bits. The coding efficiency of the (2,7) code is found to be 0.5 [2]. Therefore, the storage density can be improved by 50% (3×0.5).

The RLL encoding scheme can be generated by using a finite state machine (FSM). Table 2.1 shows the lookup table for the (2,7) RLL encoding.

The (2,7) code data sequence and PWM is illustrated in Fig. 2.5. The

32

input data bits, encoded RLL bits, alternative polarization marks, and the readback signal are shown as a function of time.

To recover the recorded bits, we need to detect the width of each mark from the readback signal. A popular method used (although perhaps not the best) is known as the peak detection technique. In this method, the readback signal is differentiated to generate peaks at the mark boundaries. By locating the peaks or positions of the mark boundaries, mark widths can be determined. Specifically, a bit timing is regenerated from a phase lock loop (PLL) in the detection circuitry. From the bit timing recovered, time windows can be defined as the time intervals between clock transitions. Therefore, we can locate mark boundaries by determining which window peaks fall into. More sophisticated detection can by done using the maximum likelihood (ML) and Viterbi algorithm.

The readback signal is subjected to a number of degradations which cause detection errors. In addition to noise and material defects, as mentioned earlier in Chapter 1, important sources of degradation include intersymbol interference (ISI) and pulse broadening. The remainder of this dissertation will quantify these degradation sources and use equalization to improve the detection performance.

2.5   Illustration of ISI

Pulse broadening (can be shrinking) is the change of the pulse width when it is written and read back again from the original. When a sequence
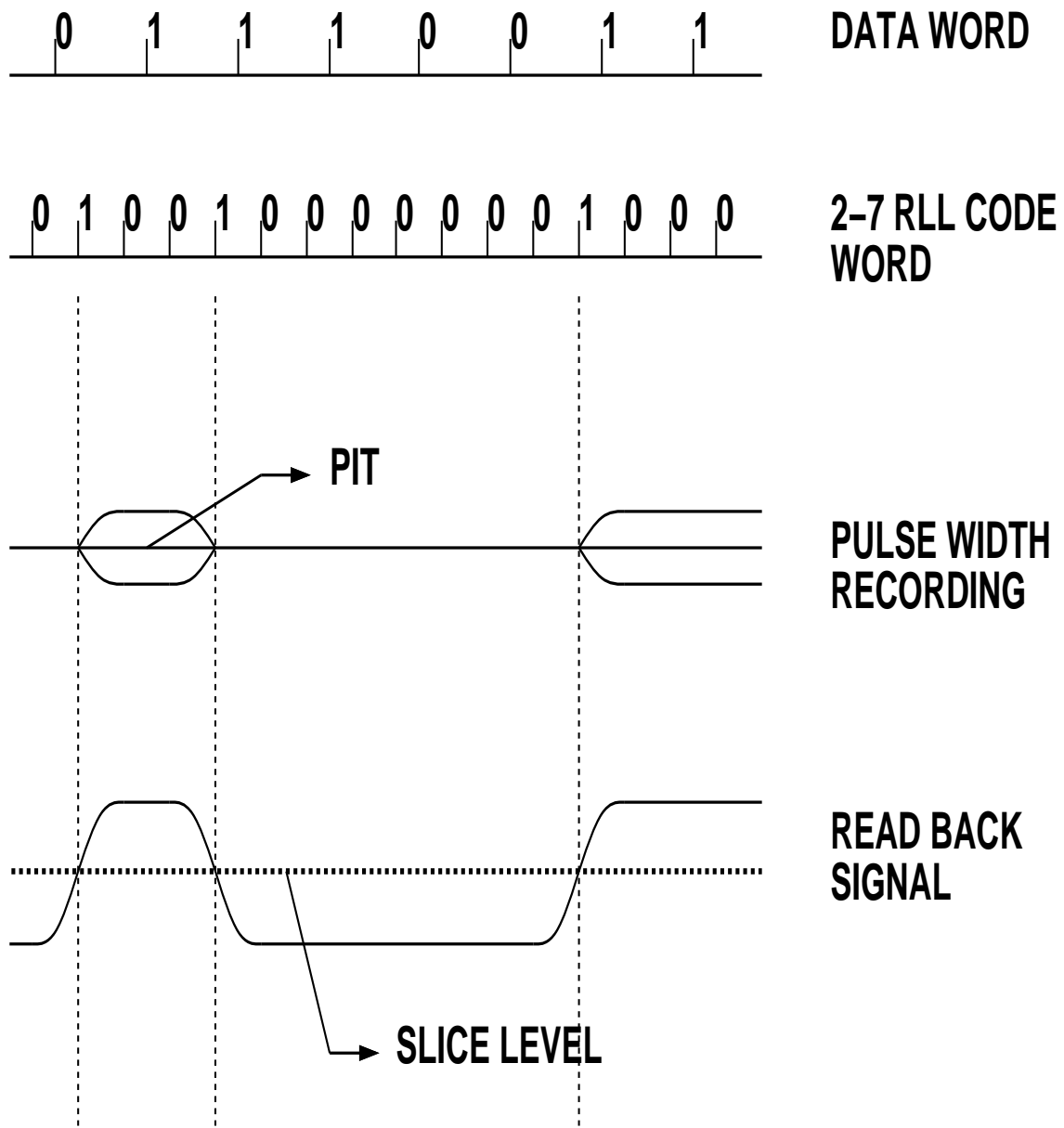
Figure 2.5: Principle of Pulse Width Recording.

34

of marks is written to a disk, the real size of each mark is not determined by itself but also by adjacent marks. When marks are read, each pulse width is also subject to itself and neighboring marks. This change in pulse due to neighboring marks is called ISI. Fig. 2.6 illustrates the ISI in the readback signal. Let W_i be the mark widths originally recorded. Let $\hat{W}_i$ be the readback marks. The ISI is the difference between the originally recorded and readback marks and is given by the following equation.

$$ISI = W_i - \hat{W}_i = f(\cdots, W_{i-1}, W_i, W_{i+1}, \cdots) \qquad (2.1)$$

When the neighboring marks have large widths, the ISI effect is not important because the central mark is more like an isolated mark. When the marks become smaller, this ISI effect becomes important and careful equalization is necessary.
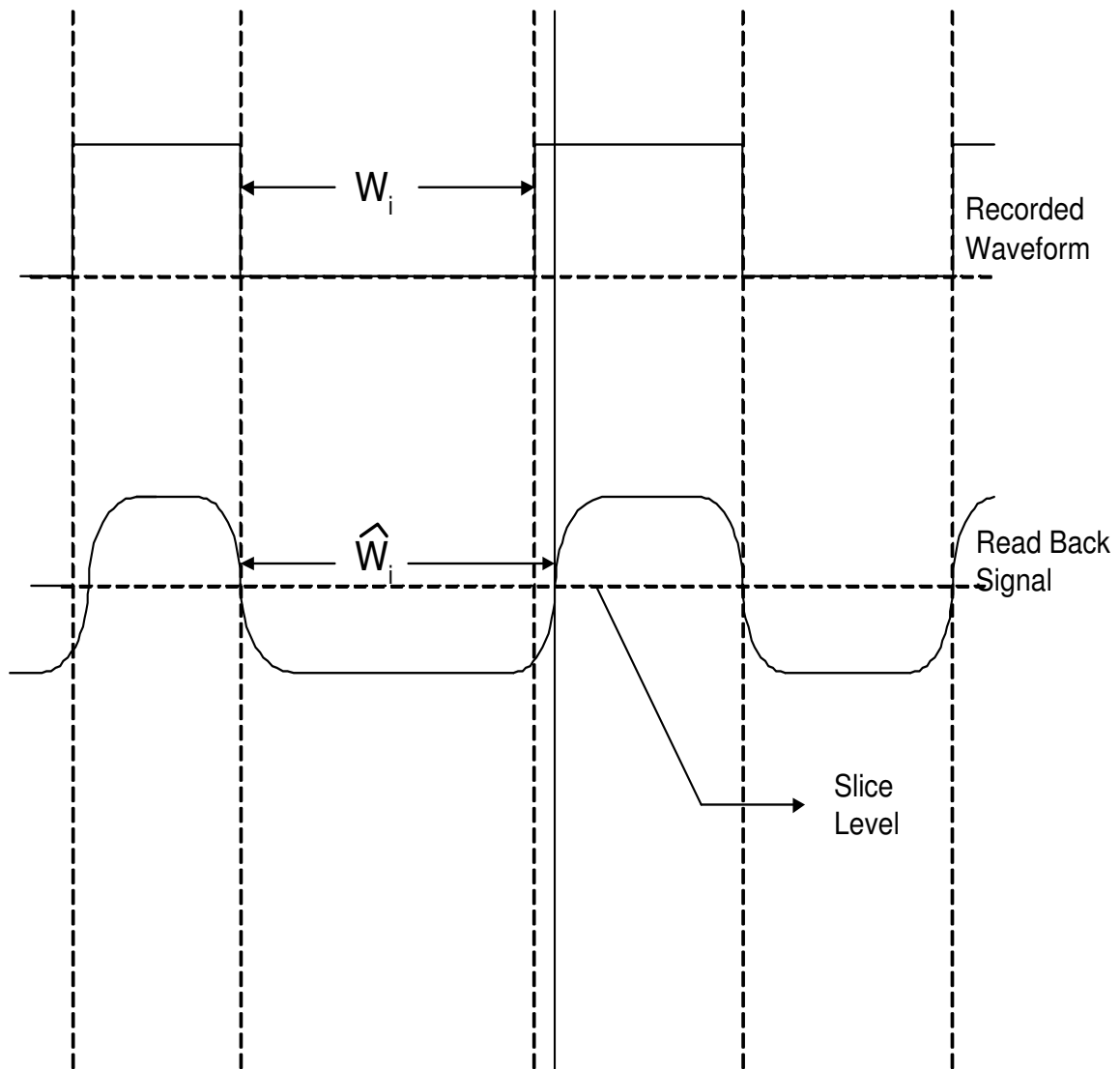
Figure 2.6: Illustration of ISI

$$ISI = W_i - \widehat{W}_i = f(..., W_{i-1}, W_i, W_{i+1}, ...)$$

36

Chapter 3

ISI Characterization

The sampled data pattern obtained is analyzed to compute the width of each individual mark. Since the sampled waveform for each data pattern consists of not only ISI but also random noise, it is necessary to take average of the repetitive data patterns to suppress the noise. In this chapter, these steps are explained and the ISI is quantified with an empirical interpolation function.

3.1   Mark Size Detection

After the read back signal is sampled, mark sizes are detected with the help of a computer program, which computes each mark size by detecting its two boundaries. The basic idea used in the program for boundary detection is to take first and second derivatives of the readback signal. As illustrated in Fig. 3.1, peaks of the first derivative and zero-crossings of the second derivatives correspond to the mark boundaries. By matching the peaks and zero-crossings, mark boundaries and consequently the mark widths are easily obtained.

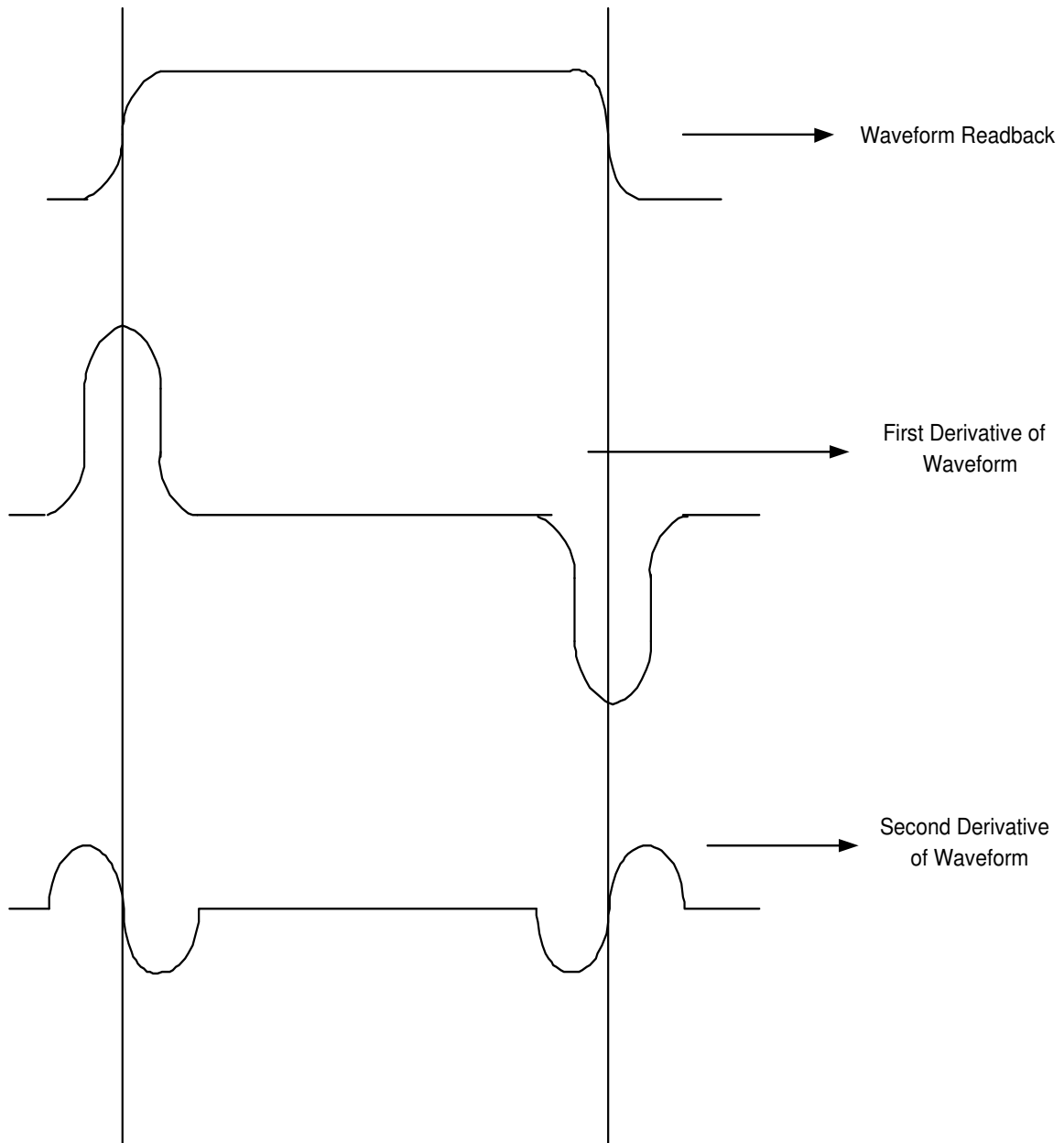All the samples were recorded and readback from a Nakamichi Magneto-

Figure 3.1: Mark Size Detection Method.

38

Optic drive. The drive had the following characteristics. The Write Power of laser = 5.0 mW. The Read Power of laser = 1.8 mW. The diffraction limit (spot size) = $1\mu m$ (996 nm). The wavelength λ=830 nm. Thus spot size = $0.6\lambda/NA$. NA is the numerical aperture (0.5).

Thus spot size = $0.6\lambda/NA$ = 996 nm = $\sim 1\mu m$

The drive was running at constant linear velocity = 4.0 m/s. The sampling rate of readback waveform = 20 Msamples/s.

Since the readback signal is discrete after sampling, regenerating a smooth curve for the readback signal and finding the first and second derivatives are explained below.

First, let y_i be the amplitude of the ith data point of the original waveform, and let x_i be the time point corresponding to the data point y_i. For computational convenience, x_i is actually set to zero so the *absolute position* of the sample data does not have to be known. The basic idea of curve smoothing or curve fitting is to find the coefficients a_2, a_1, and a_0 for a given point x_i such that

$$\hat{y}_{i+k} = a_2 x_{i+k}^2 + a_1 x_{i+k} + a_0 \tag{3.1}$$

$$= \sum_{j=0}^{2} a_j x_{i+k}^j. \tag{3.2}$$

is as close to y_{i+k} as possible for $k = 0, \pm1, \dots \pm5$. More quantitatively, the coefficients a_j's are determined for each data point by minimizing the square

39

sum:

$$\sum_{k=-5}^{5} (y_{i+k} - \hat{y}_{i+k})^2.$$

Detailed derivation of this least squares interpolation is given in Appendix A.

Once the coefficients a_j's are obtained for each data point y_i, its first and second derivatives can be found to detect the mark boundaries. From Eq. (3.1),

$$\dot{y}_i = 2a_2 x_i + a_1 \tag{3.3}$$

and

$$\ddot{y}_i = 2a_2 \tag{3.4}$$

where $\dot{y}_i$ is the value of the first derivative corresponding to x_i, and $\ddot{y}_i$ is the value of second derivative corresponding to x_i. Figs. 3.2, 3.3, 3.4, 3.5 and 3.6 show the smoothed waveform, and the first and second derivatives of various data waveforms.

A large number of mark widths of the same pattern are computed and averaged to suppress random noise. The ISI will be simply the deviation of the averaged mark width from its original recorded width. Results of these computations are summarized in the next section.

3.2 ISI Characterization

From the computed mark widths of different patterns, we can characterize the ISI for the given M-O system. Since ISI is a function of the mark

40

Data Pattern Recorded

| High | nT | T | mT | High |

T=1um n=1 m=1

Figure 3.2: Smoothed Waveform, First and Second Derivatives for the case when n=1 and m=1

41

Figure 3.3: Smoothed Waveform, First and Second Derivatives for the case when n=2 and m=4

42

Data Pattern Recorded

| High | nT | T | mT | High |

T=1um n=3 m=4
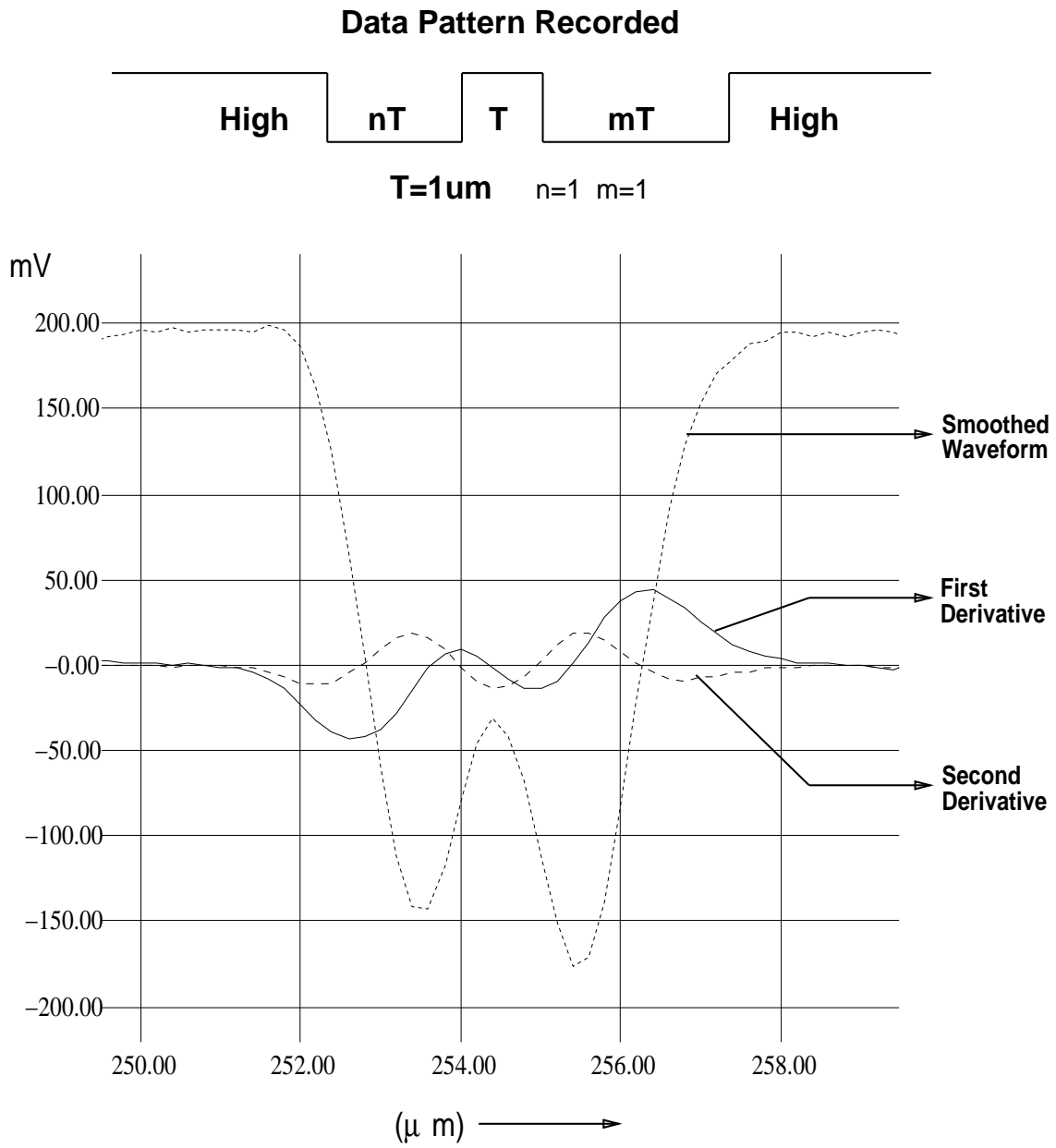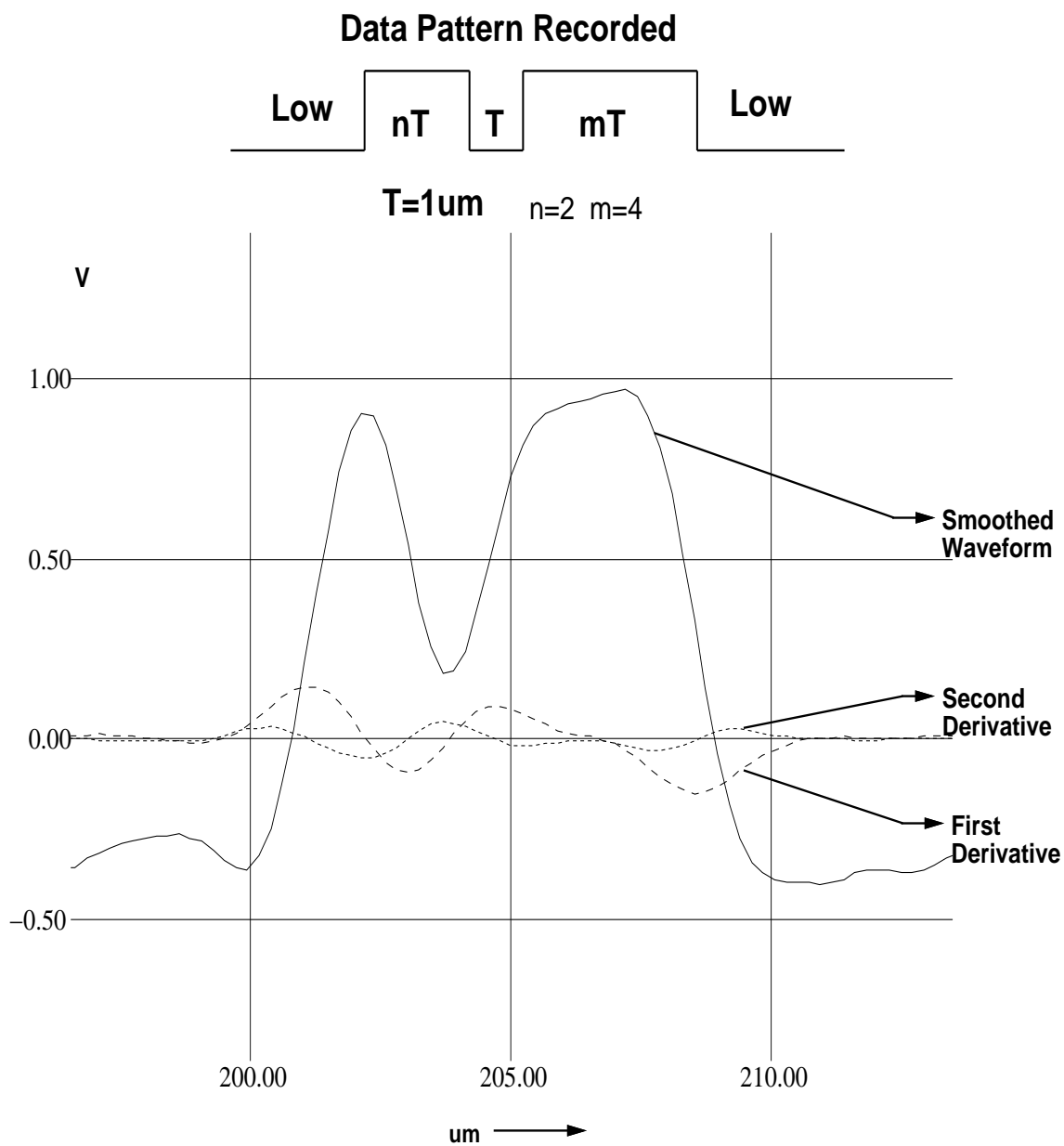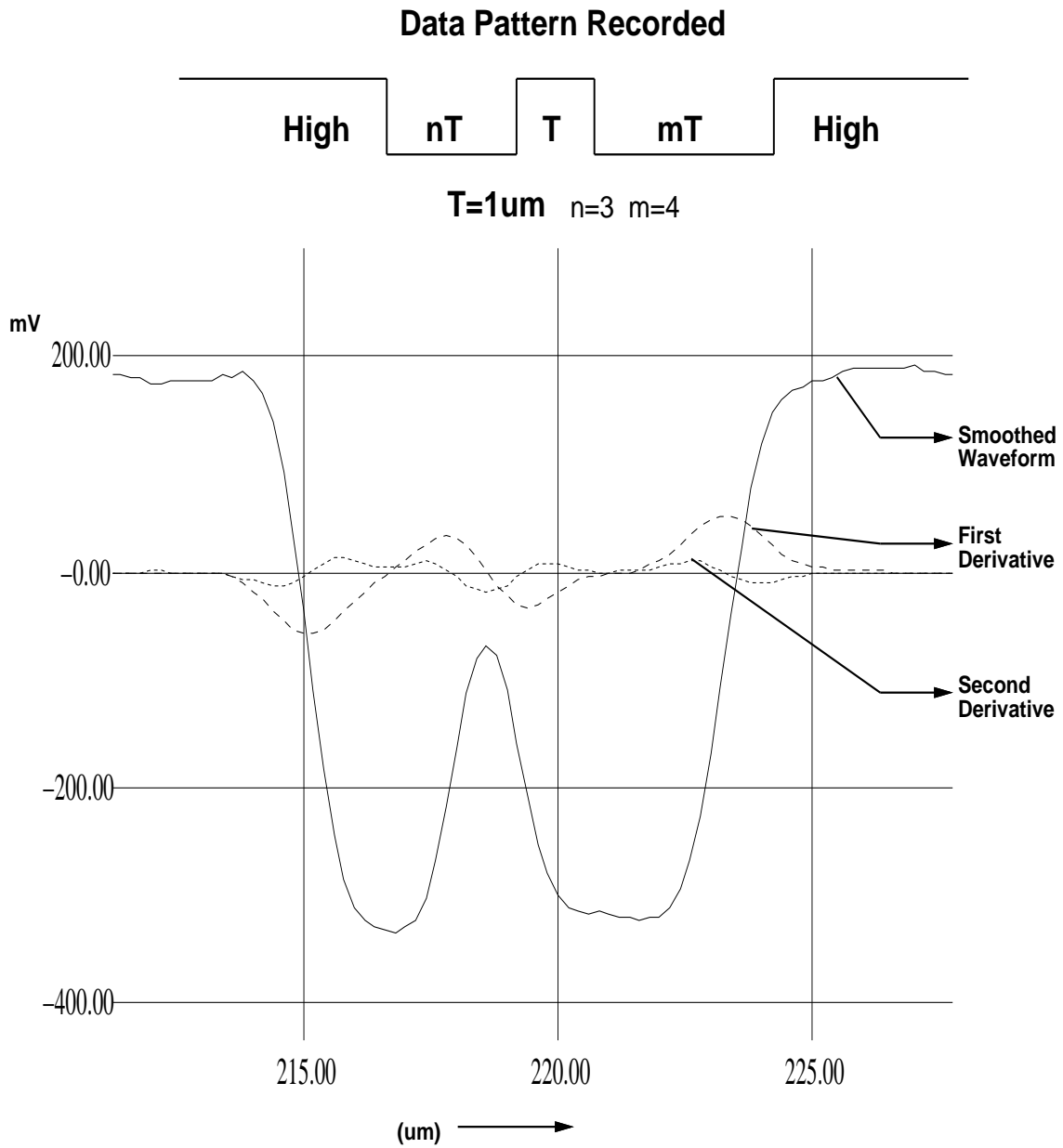
Figure 3.4: Smoothed Waveform, First and Second Derivatives for the case when n=3 and m=4

43
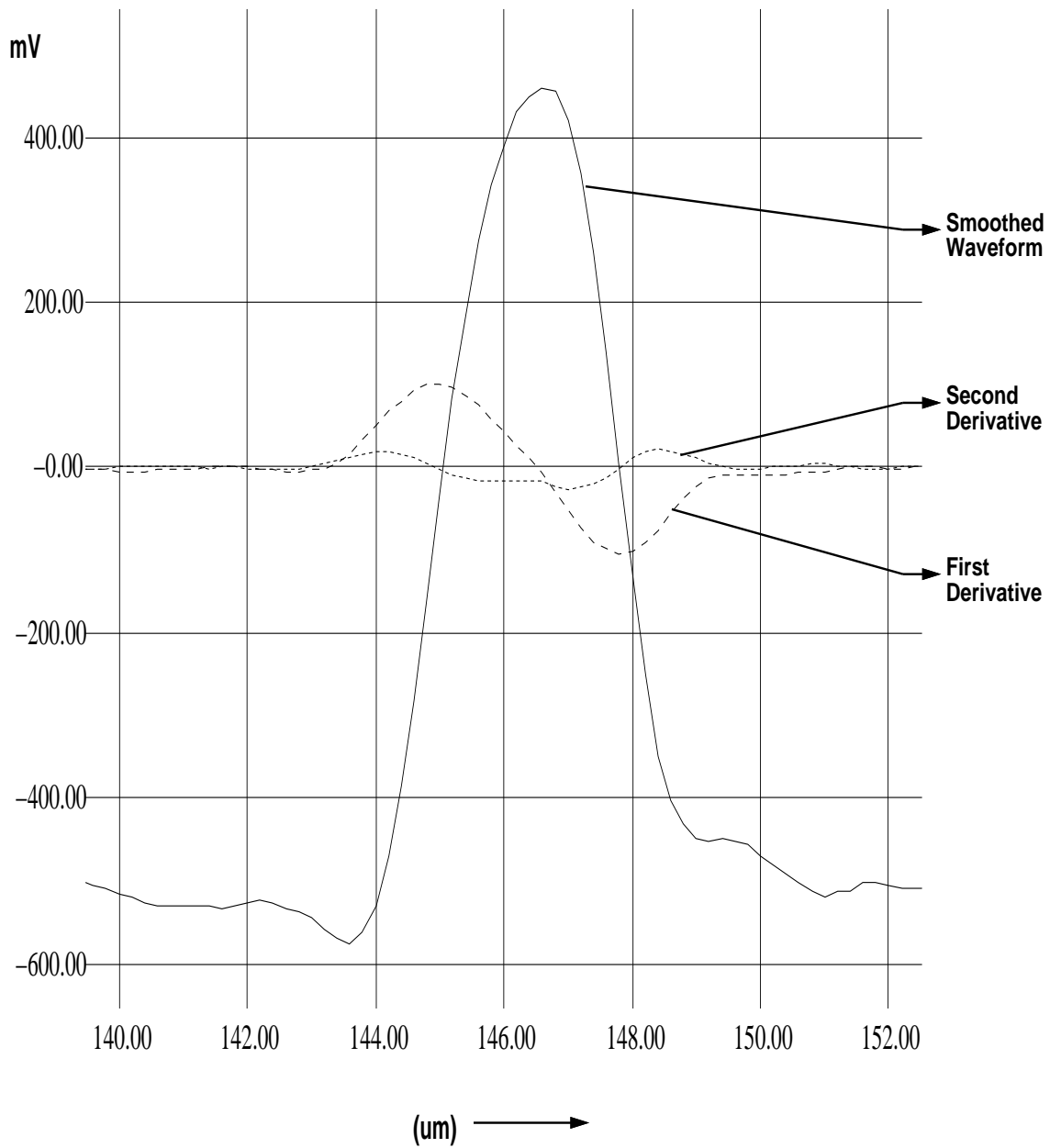
Figure 3.5: Smoothed Waveform, First and Second Derivatives for an isolated high pulse

44
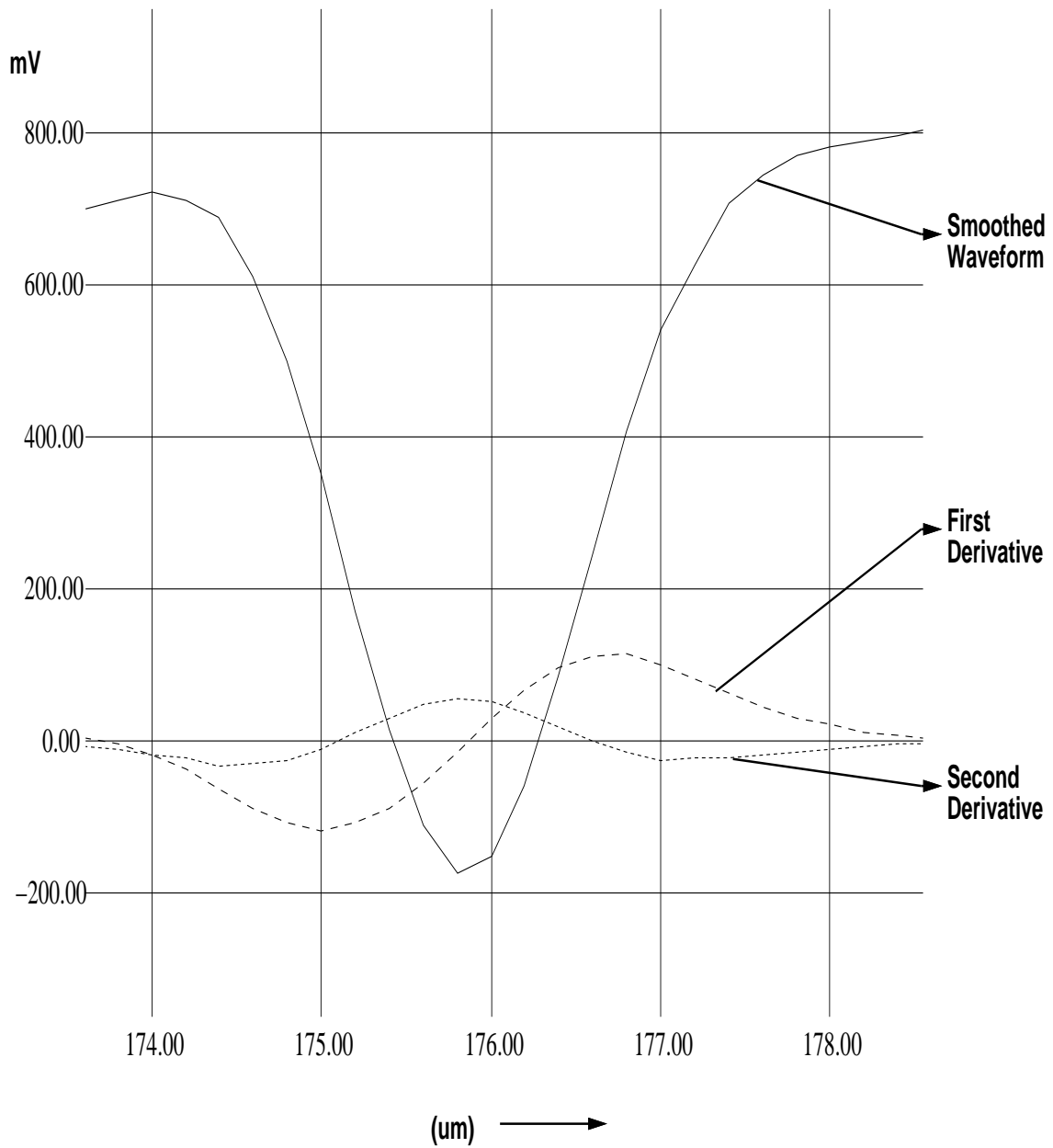
Figure 3.6: Smoothed Waveform, First and Second Derivatives for an isolated low pulse

45

polarization and also a function of the widths of the center mark and its two neighboring marks, we consider four cases as summarized below.

3.2.1 ISI Characteristics for High Center Marks

Fig. 3.7 shows the first type of ISI characteristics where the center mark is a high mark. Each ISI curve shown in the figure is a function of the width of the right neighboring mark, at a fixed width of the left neighboring mark. The width of the center mark is fixed at 1 μm. The ISI behavior is highly non-linear. When the adjacent marks are greater than 2 μm, the ISI approaches a constant value.

3.2.2 ISI Characteristics for Low Center Marks

Fig. 3.8 shows the second type of ISI characteristics where the center mark is a low mark. The ISI curves are plotted very similar to that in the first case. The nature of ISI behavior for this case is very similar to that of the previous case.

46

Figure 3.7: ISI Characteristics for High Center Marks.

47
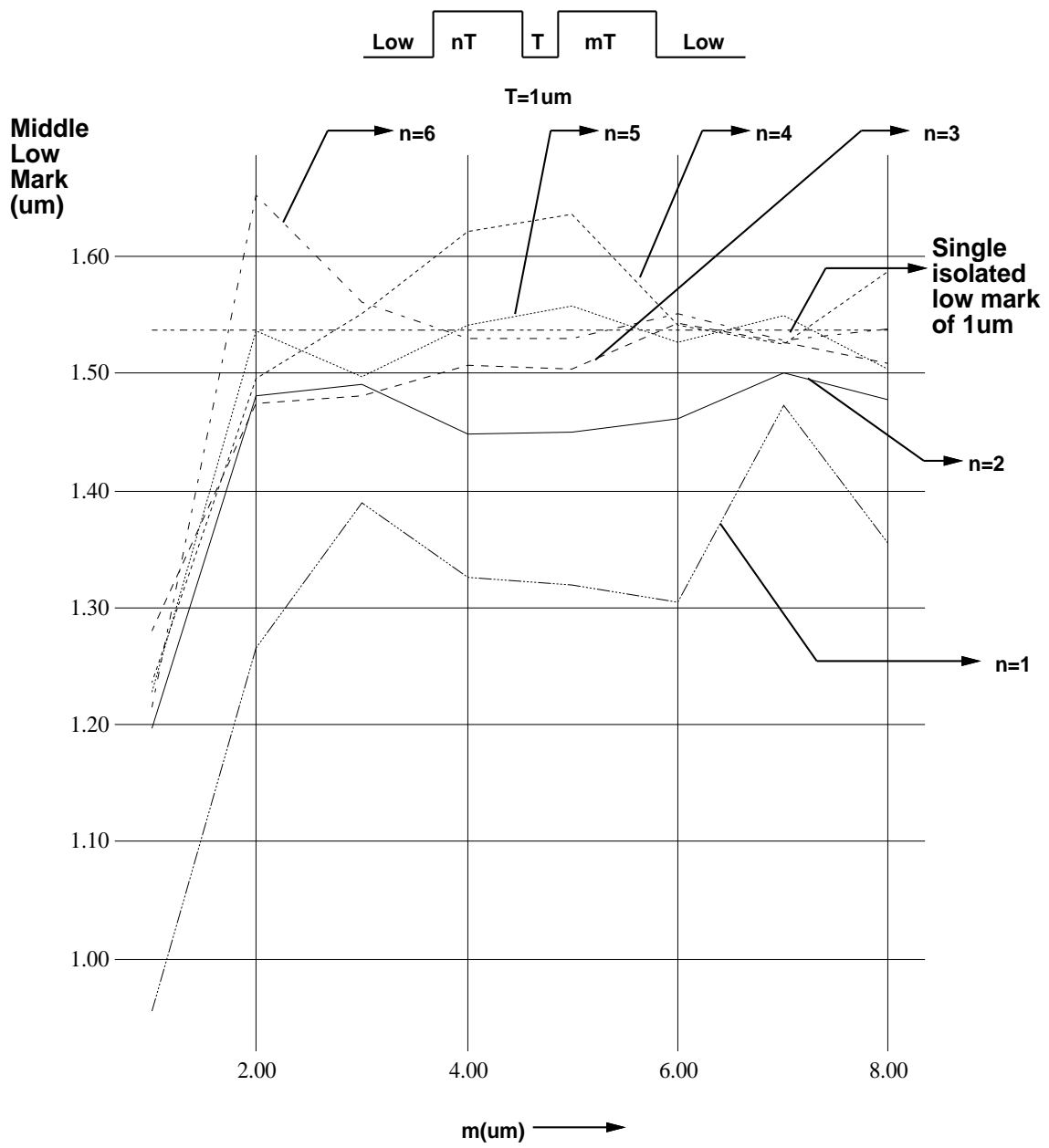
Figure 3.8: ISI Characteristics for Low Center Marks.

48

3.2.3 ISI Characteristics for Isolated High Center Marks

In the first two types of the ISI characteristics summarized, the center mark is fixed at 1 μm. In this third case, the width of the center mark is a variable with large neig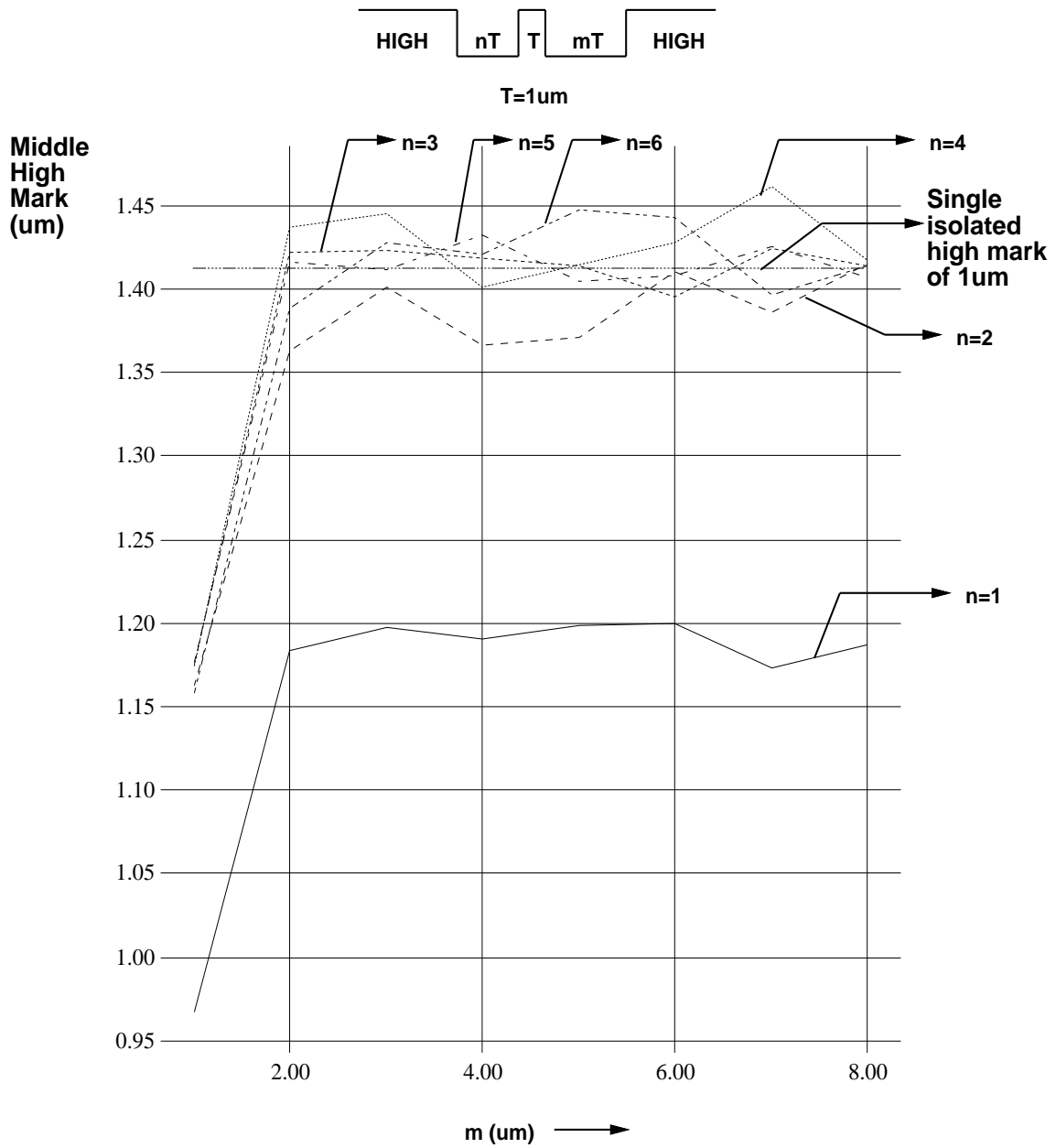hboring marks. Therefore, the center mark is essentially an isolated mark. Fig. 3.9 illustrates the difference between a recovered isolated high mark and recorded mark as a function of width of the recorded high mark. The figure shows that ISI is positive for marks less than 1.5 μm and negative for marks greater than 1.5μm. ISI approaches zero as the mark size increases.

3.2.4 ISI Characteristics for Isolated Low Center Marks

This case is similar to the third case except that the center mark is now a low mark. Fig. 3.10 illustrates the difference between a recovered isolated low mark and recorded mark as a function of width of the recorded low mark. This figure shows that the ISI value is always positive and approaches zero as the mark size increases.

3.3   Interpolation of ISI Characteristics

To perform simple ISI equalization in the read process, instead of using an ISI table, an empirical function may be used that gives an estimated amount of ISI from the detected widths of the center mark and its two neighbors.

By observing the ISI characteristics shown in Figs. 3.7 and 3.8, a simple

49
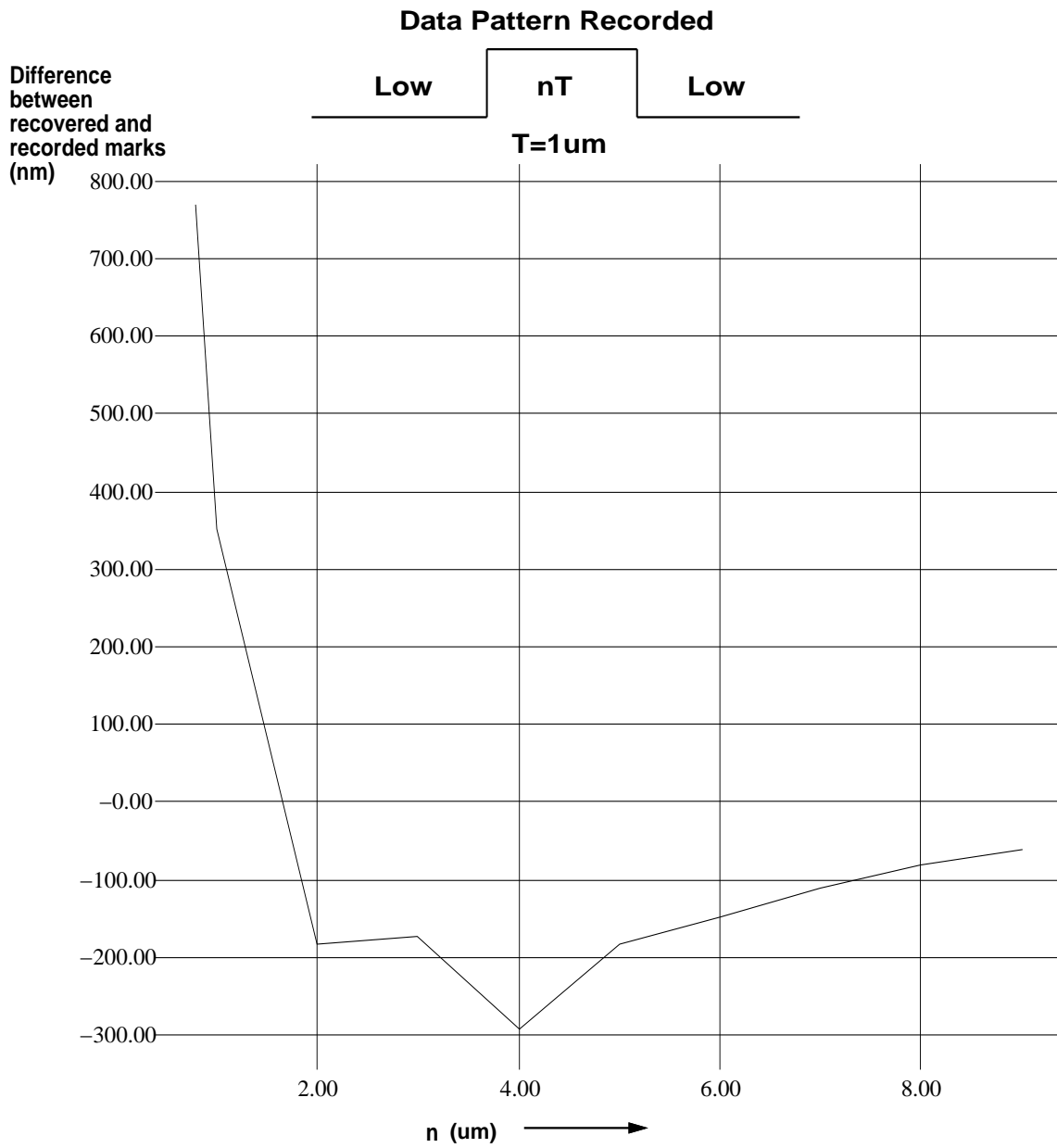
Figure 3.9: ISI Characteristics for Isolated High Center Marks.

50
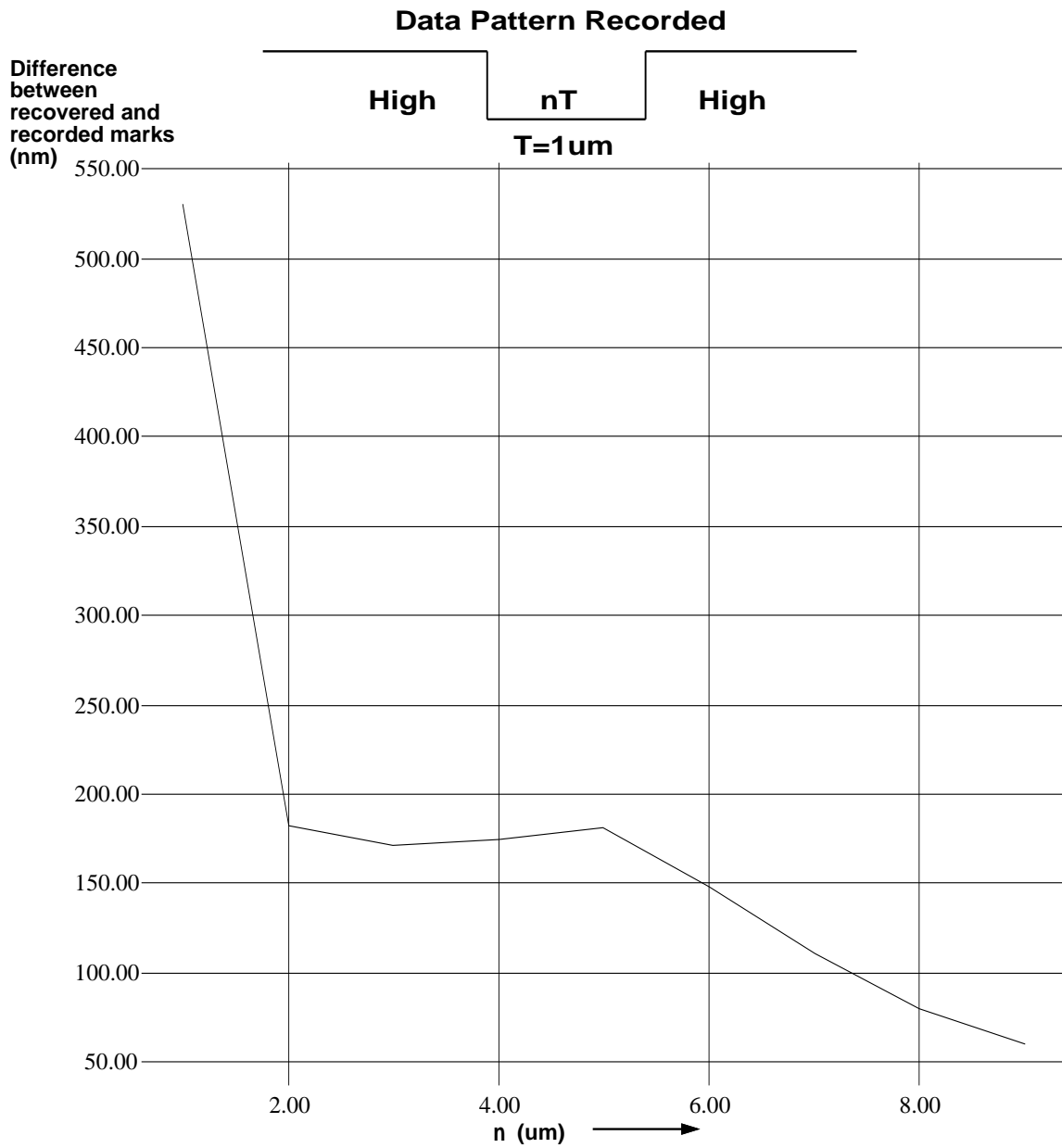
Figure 3.10: ISI Characteristics for Isolated Low Center Marks.

51

empirical function is proposed to be of the following form:

$$W(n_l, n_m, n_r) = I(n_m) \left\{ \frac{2}{\pi} \arctan([\frac{n_l}{k_1}]^\alpha) \right\} \left\{ \frac{2}{\pi} \arctan([\frac{n_r}{k_1}]^\alpha) \right\} \qquad (3.5)$$

where n_l is the recorded mark width on left side, n_m is the recorded mark width in the middle, n_r is the recorded mark width on the right side, $I(n_m)$ is the width of the isolated mark n_m, and $W(n_l, n_m, n_r)$ is the actual mark width of the center mark. k_1 and α are constants. Therefore, the amount of ISI is determined as

$$ISI(n_l, n_m, n_r) = n_m - W(n_l, n_m, n_r). \qquad (3.6)$$

From this empirical function, when both n_r and n_m are large, the actual mark widths approach the isolated mark width $I(n_m)$. Also, at fixed n_r and n_m, the $W(n_l)$ is an arctan function, which is similar to the experimental results. Figs. 3.11 and 3.12, superimpose the empirical curves on the actual curves according to Eq. (3.5).

The constants k_1 and α are determined as follows:

First normalize the actual mark widths $W(n_l, n_m, n_r)$ by dividing them by the width of the isolated mark $I(n_m)$.

We then set particularly two values:

$$\frac{2}{\pi} \arctan([\frac{2.0}{k_1}]^\alpha) = \frac{W(n > 1, n_m, n_r = 2)}{I(n_m)} \qquad (3.7)$$

$$\frac{2}{\pi} \arctan([\frac{1.0}{k_1}]^\alpha) = \frac{W(n = 1, n_m, n_r >> 2)}{I(n_m)}. \qquad (3.8)$$

52
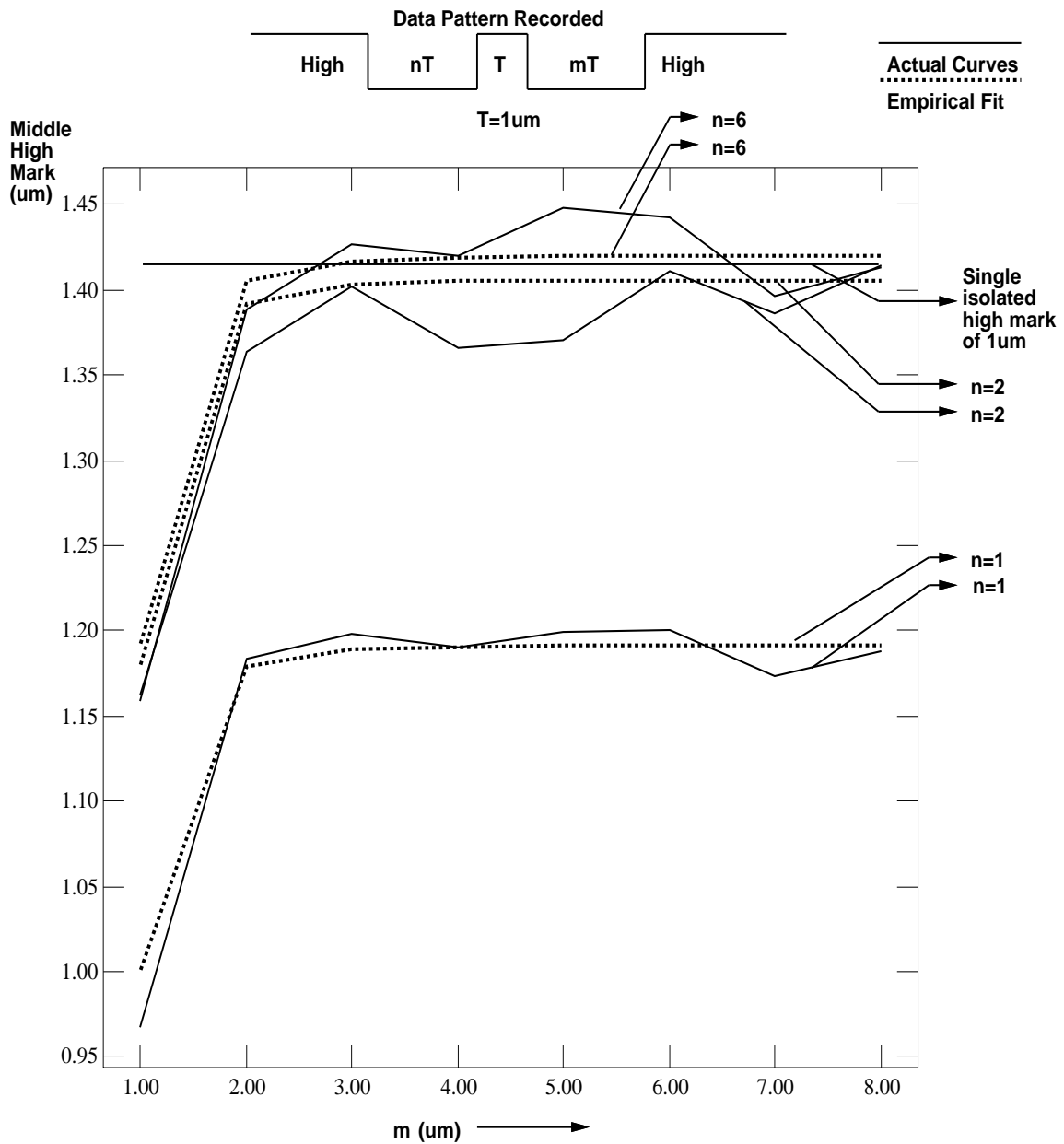
Figure 3.11: Superimposition of Actual and Empirical Curves of ISI for High Marks.

53
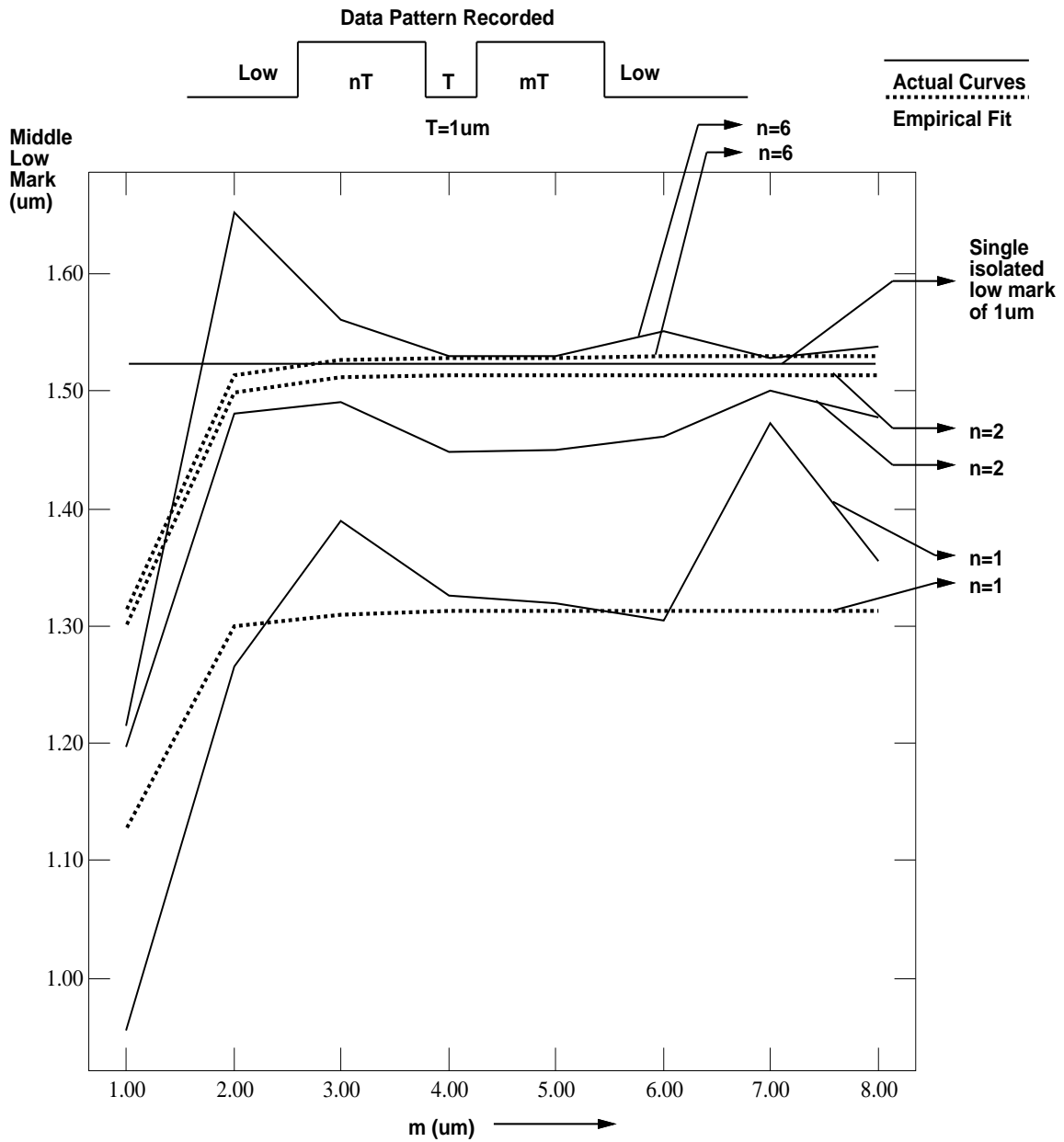
Figure 3.12: Superimposition of Actual and Empirical Curves of ISI for Low Marks.

54

From Eqs. 3.7 and 3.8 solve for k_1 and α whose expressions are given below.

Defining two new parameters a and b as

$a = W(n > 1, n_m, n_r = 2)/I(n_m)$ and

$b = W(n = 1, n_m, n_r >> 2)/I(n_m),$

the equations for α and k_1 become

$$\alpha = \log\left(\frac{\tan(a\frac{\pi}{2})}{\tan(b\frac{\pi}{2})}\right)/\log(2.0) \tag{3.9}$$

and

$$k_1 = 1.0/\{\tan(b\frac{\pi}{2})\}^{\frac{1}{\alpha}}. \tag{3.10}$$

In Fig. 3.11, $\alpha = 3.039685$ and $k_1 = 0.640727$ μm, and in Fig. 3.12, $\alpha = 2.801347$ and $k_1 = 0.581585$ μm.

The work related to ISI characterization was performed at the University of Arizona [4]. The ISI characteristics are included here for completeness sake to give a feel for how the readout waveforms look. In the equalization process readout symbols with ISI are taken and various filtering operations are then performed.

In the next chapter the equalization process is investigated for compensation of ISI. Nonlinear equalization using Volterra filters and Hybrid equalization (combination of Linear and Volterra) are examined. Also, the convergence properties are explored.

55

In this work the biggest difficulty was in obtaining the experimental data. It was not possible to obtain recent experimental data from the industry, but it was necessary to use the older data obtained using the Nakamichi drive.

Although the data is old, it does not diminish its value. The only thing different in the recent drives will be the smaller wavelength of the laser which will result in a smaller spot size.

Chapter 4

Problem Definition, Volterra Filter, Convergence Analysis & Previous Work

Although data storage densities have increased tremendously in the last four decades, there is a pressing need to further improve the densities to meet the growing demands of new applications such as real time video/audio files, 2D/3D image files, digital libraries, etc. This chapter investigates techniques to increase the storage densities of disk drives using nonlinear equalization of the detected marks to compensate for the ISI and pulse broadening.

Thus, the objective is to use Volterra/Hybrid equalization to compensate nonlinear ISI in the read-back signal.

Fig. 4.1 illustrates the block diagram of the read/write and equalization process. As shown on this figure, first the binary data is encoded into a PWM (pulse width modulated) data sequence. The PWM marks are then written/recorded on a magneto-optic disk drive. During the read process the data marks are readout with ISI. Then the linear or nonlinear equalization is performed to compensate for it. After that decoding is done to recover the original data sequence.
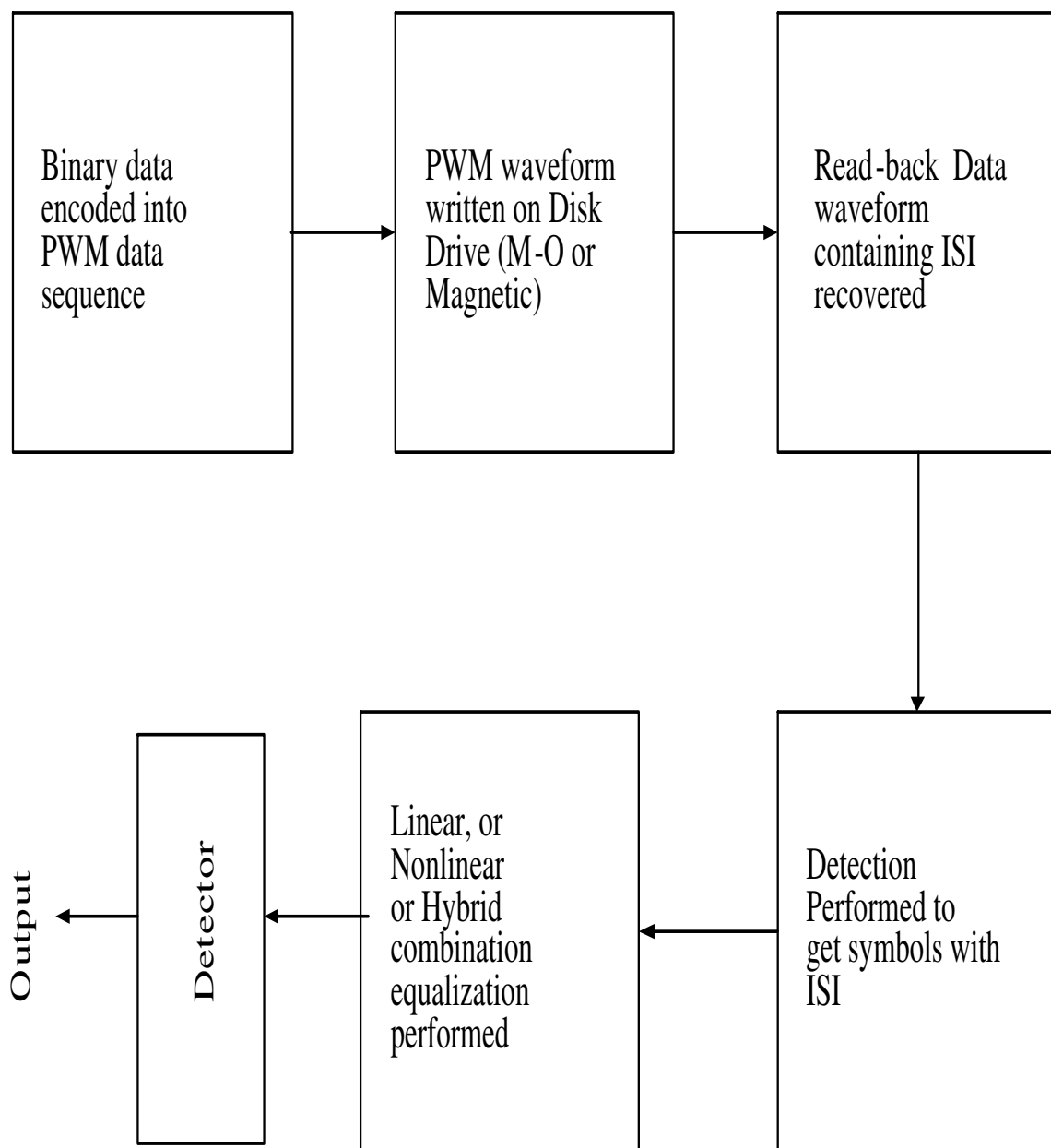
Figure 4.1: Block diagram of Read/Write and Equalization Process

58

4.1 Linear Decision Feedback Equalizer

Fig. 4.2 shows the configuration of a linear decision feedback equalizer (DFE). In this figure there is tap delay line and 5 inputs are considered are compensation of ISI i.e., the mark under consideration and the 2 neighboring marks on either direction (before and after it). The two decided marks are fedback and used in the determination of the compensated mark under consideration. These marks are multiplied with a weighting factor and fed to a summer. The weights are determined using an LMS (Least Mean Square) algorithm. These weights are updated adaptively using the error function and the step-size multiplying factor.

The LMS algorithm is described in more detail in the following sections. In the next section Volterra filters are explored in detail and issues related to number of terms generated examined. Hybrid equalizers are then defined.

4.2 Why Volterra Filters ?

Nonlinearity is encountered in many physical systems and engineering applications. There are two well-accepted approaches to modelling polynomial nonlinearities: the Wiener and the Volterra approaches [29]. In the Wiener approach, the kernels are estimated separately in closed form; however, the theory is applicable only to Gaussian inputs. The Volterra approach is attractive for the reason that it is a straightforward generalization of linear models. The nonlinear combining occurs before the tap weights of the equalizer, the output of the nonlinear combiner may be considered inputs to a linear adap-

59
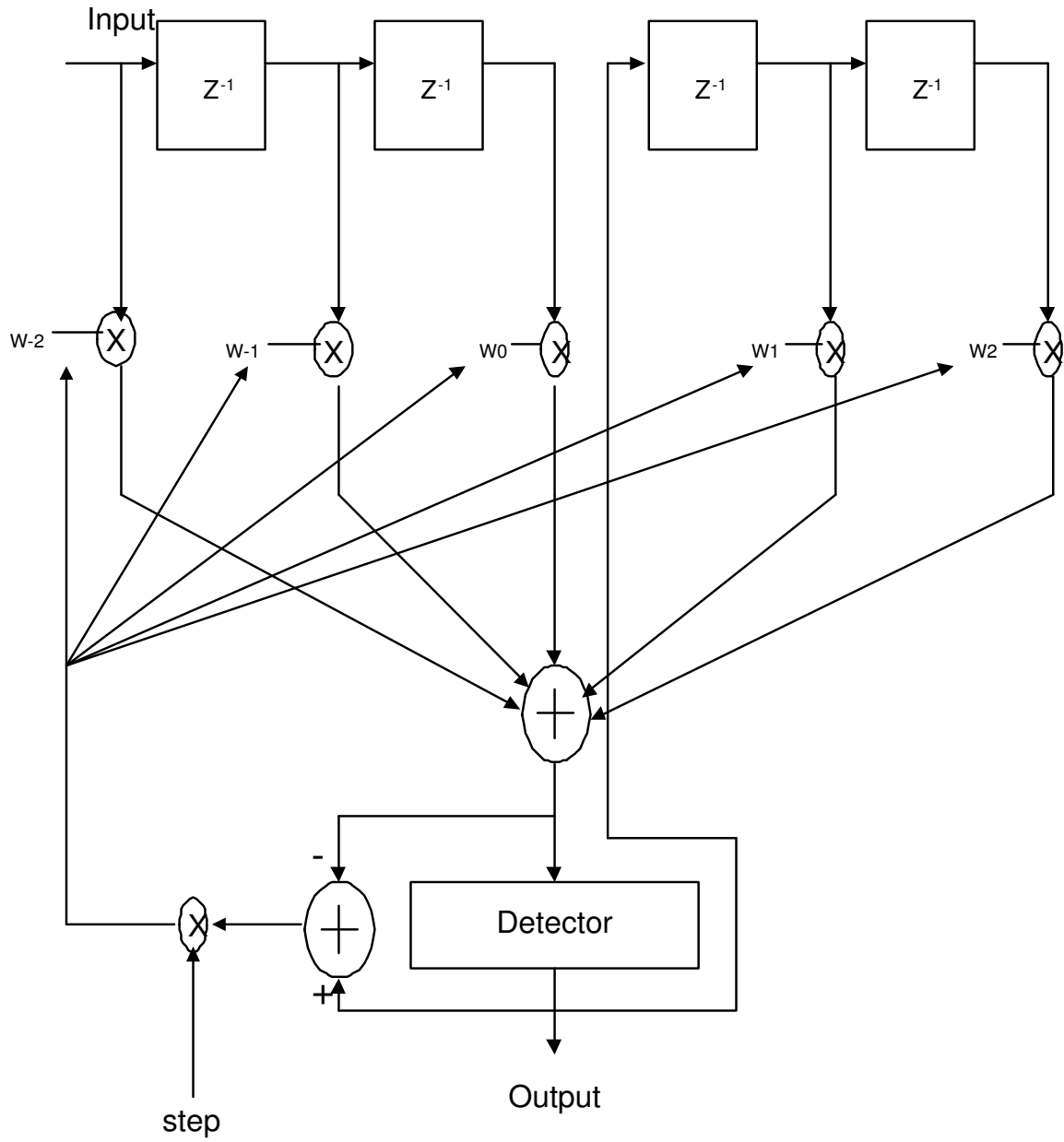
Figure 4.2: Linear Decision Feedback Equalization (DFE)

60

tive filter. Thus, the LMS algorithm may be employed for adaptation and the results of linear filtering apply.

Thus Volterra equalization is used for correcting nonlinear ISI as it has been successfully applied in the other nonlinear physical systems.

The Volterra series characterization of a nonlinear system provides a relationship between the discrete-time input symbols Y_n, and the discrete-time output symbols Z_n as -

$$Z_n = \sum W_{n1} Y_{n-n1_{n1}} + \sum \sum \sum W_{n1n2n3} Y_{n-n1} Y_{n-n2} Y_{n-n3} +$$
$$\sum \sum \sum \sum \sum W_{n1n2n3n4n5} Y_{n-n1} Y_{n-n2} Y_{n-n3} Y_{n-n4} Y_{n-n5} + \cdots \quad (4.1)$$

Where, Y_n are the inputs of the nonlinear Volterra equalizer and Z_n are the output of the equalizer. Y_n consists of linear and nonlinear ISI terms. The Volterra equalizer consists of linear combination of all linear terms and all possible combination of nonlinear terms of Y_n of odd degree. This is because ISI is assumed to be equally influenced by the symbols on either side of each symbol.

Fig. 4.3 illustrates a block diagram of a 3-tap 3rd-order Volterra equalizer, where Y_0 in the figure corresponds to the Y_n. The samples form a tapped delay line are the inputs to a nonlinear combiner. The nonlinear combiner then outputs all single taps and all combinations of three taps. Each output from the nonlinear combiner is then multiplied by a weight to form an input to

61

the summing device. In general there are terms out of the nonlinear combiner, where N is the number of taps in the tapped delay line. Order is the highest degree term at the nonlinear combiner output and is an odd number. Thus, the equalizer in Fig. 4.3 contains total of 30 outputs terms, 3 linear and 27 of degree 3.

$$L = \sum_{i=1}^{(orderr+1)/2} N^{2i-1} \qquad (4.2)$$

The main drawback of going to higher order Volterra series is that the number of terms required for computation tends to explode as shown in Eq. 4.2 and Table 4.1. Thus, there is a compromise between the computational complexity and the gain achieved by going to higher order Volterra series in the equalization.

Table 4.1: # of taps, Volterra order and # of terms

# of taps & Volterra order	# of terms
3-tap 3rd order	30
5-tap 3rd order	130
7-tap 3rd order	350
3-tap 5rd order	273
3-tap 7rd order	2460
5-tap 5rd order	3255

Fig. 4.3 shows the implementation block diagram of a 3-tap $3-^{rd}$ order Volterra equalizer. The implementation is quite similar to that of a linear equalizer. The difference is in that in Volterra equalizer there is a nonlinear

62
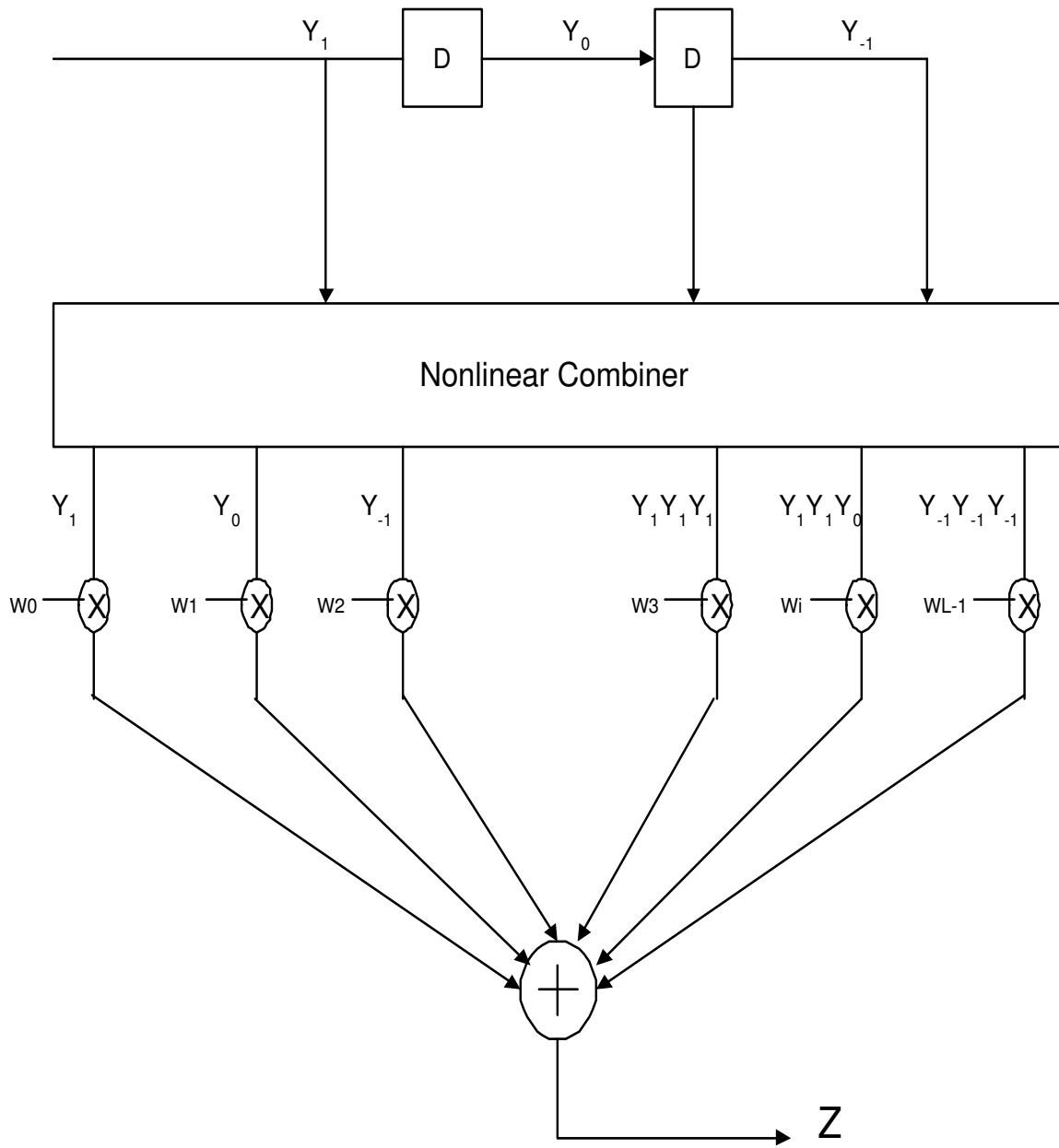
Figure 4.3: Illustration of Volterra Equalization

63

combiner where all the Volterra terms are calculated. After that the weights are determined corresponding to each term using LMS (Least Mean Square) algorithm. One can use other algorithms also such as the Kalman filter for faster convergence or use multiple step sizes in the LMS algorithm.

Fig. 4.4 illustrates the schematic of the Hybrid Equalization. In this figure there is a 5-tap linear filter and a short 3-tap 3rd order Volterra in the middle.

To overcome the issue of large number of terms, Hybrid equalizers are proposed which are a combination of linear and short-tap small-order Volterra filter in the middle of the tap delay line. It is based on the assumption that most of the nonlinearity is because of the most adjacent marks to the mark under consideration. The Hybrid equalizer considered consisted of a 3-tap 3rd order Volterra filter in the middle and a large N-tap (N=5,7,9,11) linear filter. The short Volterra filter in the middle compensates for most of the nonlinear ISI.

A further reduction in the number of nonlinear terms can be obtained by discarding the terms with small relative weights.

Fig. 4.5 shows the block diagram of the adaptive Hybrid equalizer. The implementation approach is very similar to the traditional linear decision feedback equalizer.
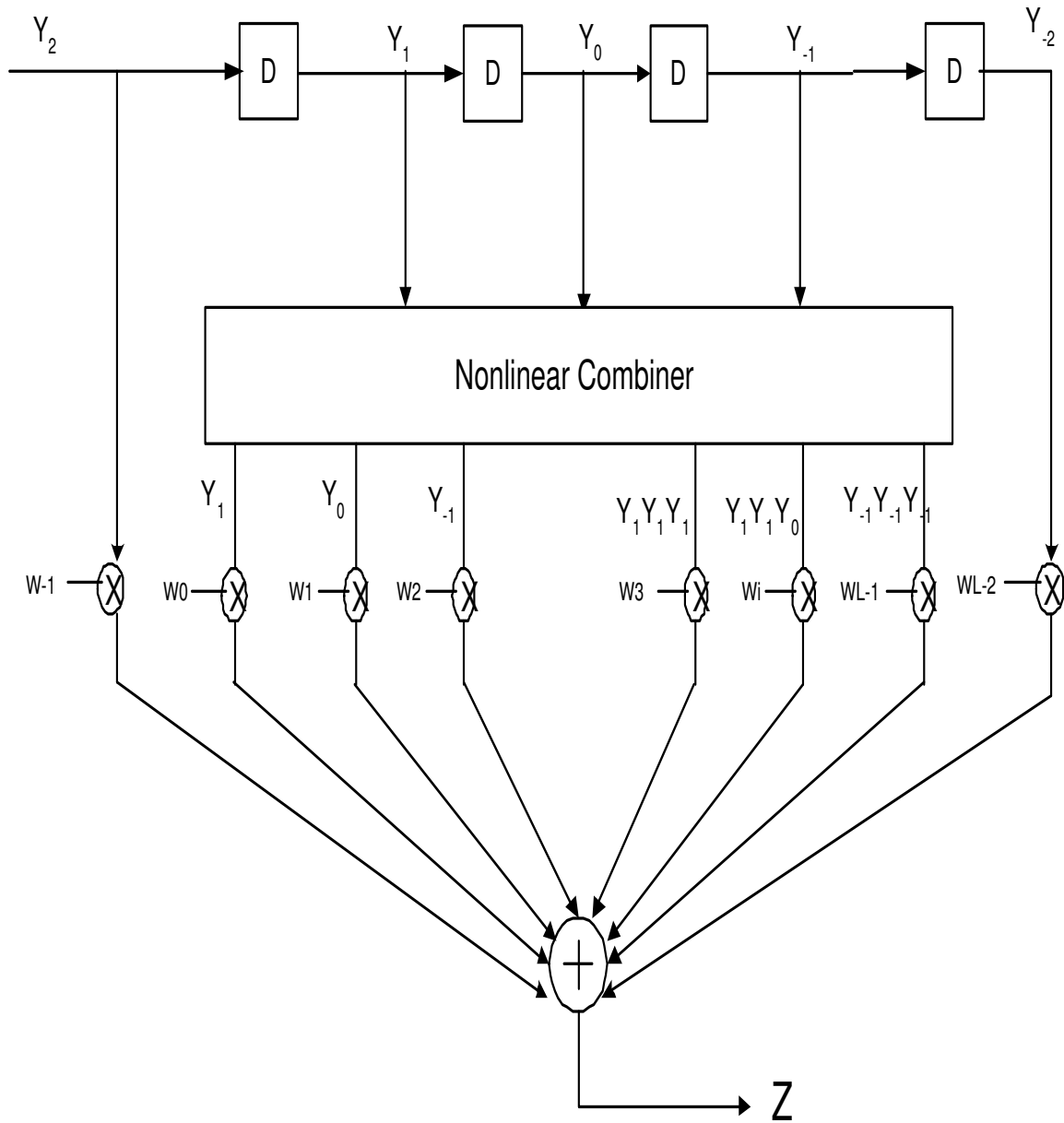
Figure 4.4: Illustration of Hybrid Equalization

65
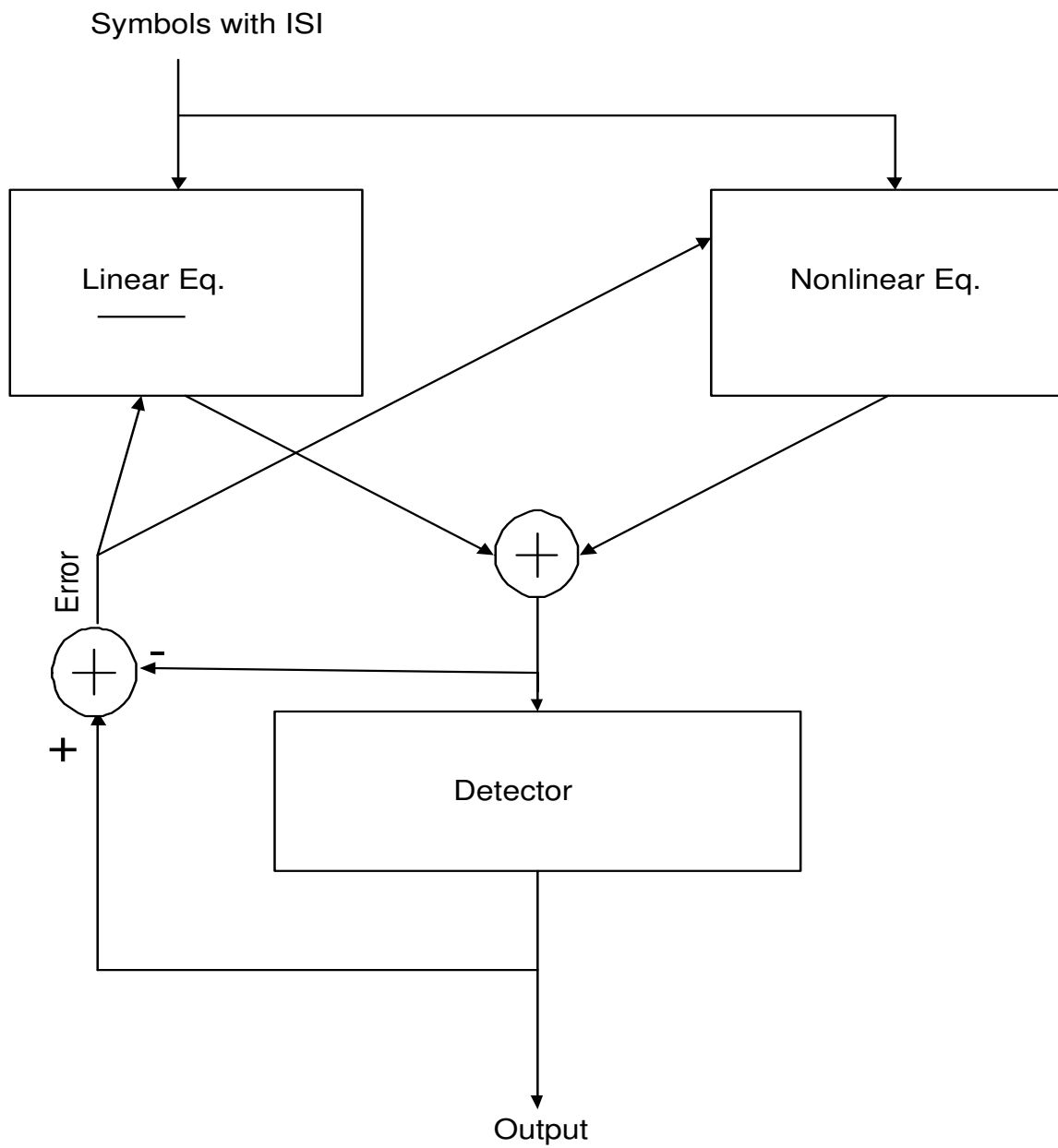
Figure 4.5: Block Diagram of Hybrid DFE Volterra Equalization

66

4.3   Volterra Equalizer Adaptation and Convergence

This section explores the convergence issues related to the nonlinear equalizer. The eigenvalue spread of the covariance matrix is examined.

Since the nonlinear combining occurs before the tap weights of the equalizer, the outputs of the nonlinear combiner may be considered as inputs to a linear adaptive filter. Thus, the LMS algorithm, may be employed for adaptation and the results of linear filtering apply.

Let Y_i be the input to the nonlinear combiner.

Let $U = [U_0,\ U_1,\ U_2,\ \cdots,\ U_{L-1}]^T$, be the output vector of the nonlinear combiner, T denotes transpose.

The number of elements L, in the vector is given by Eq. 4.2.

Let $W = [W_0,\ W_1,\ W_2,\ \cdots,\ W_{L-1}]^T$, be the weight vector, where W_i, is the weight multiplying U_i.

Then, the output of the Volterra equalizer is given by

$$Z = W^T U \qquad (4.3)$$

And the weight update equation is given by

$$W(k+1) = W(k) + \mu e(k)U \qquad (4.4)$$

Where, k is the k_{th} update time, $e(k)$ is the error at the output of the equalizer at time k, and μ is the step size.

67

For a linear adaptive filter (LMS algorithm) convergence is guaranteed only if the step size μ is chosen from

$$0 < \mu < 2/\sum_{i=1}^{L} \lambda_i \tag{4.5}$$

where λ_i is the i_{th} eigenvalue of the covariance matrix obtained from the terms of the nonlinear combiner. L is the number of terms of the nonlinear combiner.

When the eigenvalues are widely spread (i.e., the ratio of $\lambda_{max}/\lambda_{min}$ is large) the rate of convergence is limited by the smallest eigenvalues.

The eigenvalue spread for the nonlinear equalizer is greater than that of the corresponding linear equalizer with N taps.

Consider the correlation matrix of the inputs to the N-tap delay line, for example the correlation matrix of the elements Y_i in figure 4.3, and denote this matrix as R and its eigenvalues as α_i. Define the correlation matrix of the nonlinear combiner outputs, U, as $\widetilde{R}$ and its eigenvalues as λ_j. It is clear that $\widetilde{R}$ has dimensions much larger than R and that the elements of R are all contained in $\widetilde{R}$. Reference [5] gives the following two expressions for the maximum and minimum eigenvalues of $\widetilde{R}$

$$\lambda_{max} = max \frac{x^H \widetilde{R} x}{x^H x}, \ x\epsilon c^L, \ x \neq 0 \tag{4.6}$$

$$\lambda_{min} = min \frac{x^H \widetilde{R} x}{x^H x}, \ x\epsilon c^L, \ x \neq 0 \tag{4.7}$$

where L is given by 4.2 and C^L is complex vector space. Equation 4.6 states that the largest eigenvalue of $\widetilde{R}$ is obtained by the largest amount

68

by which any vector is amplified by vector multiplication. But since $R \subset \widetilde{R}$ and $C^N \subset C^L$, then $\lambda_{max} \geq \alpha_{max}$. Similarly, $\lambda_{min} \leq \alpha_{min}$, so that the eigenvalue spread for the nonlinear equalizer is likely to be greater than for the corresponding linear equalizer. This also implies that the maximum step size is likely to be less for the nonlinear equalizer.

The convergence ratio ω is defined as $\omega = (\rho - 1)^2/(\rho + 1)^2$, where ρ is the ratio of the largest to the smallest eigenvalue. It is clear from the above expression that the convergence rate decreases if the eigenvalue spread increases and vice-versa. In fact, if the eigenvalues are all equal, then convergence is obtained in just one step.

Eigenvalue spread gives a measure of the correlation of the data. Larger values of correlation mean larger levels of ISI which translates into greater eigenvalue spread and vice versa.

Table 4.2 shows the eigenvalue spread and convergence ratio for various linear and Volterra and Hybrid equalizer configurations. The convergence ratio increase as one goes to a higher order Volterra filters.

Table 4.2: Equalizer Configuration and Values of Eigenvalue Spread & Convergence Ratio

Equalizer Configuration	Eigenvalue Spread	Convergence Ratio
5-tap Linear EQ	22.36	0.836
3-tap 3rd order Volterra EQ	20.43	0.8373
5-tap 3rd order Volterra EQ	33.45	0.887
7-tap 3rd order Volterra EQ	34.24	0.89
5-tap linear, 3-tap 3rd order Volterra EQ (Hybrid)	31.23	0.879

4.4   Algorithms for Adaptive Equalization

Since an adaptive equalizer compensates for an unknown and time-varying channel, it requires a specific algorithm to update the equalizer coefficients and track the channel variations. A wide range of algorithms exist to adapt the filter coefficients. The development of adaptive algorithms is a complex undertaking. This section describes some practical issues regarding equalizer algorithm design and outlines three of the basic algorithms for adaptive equalization.

The performance of an algorithm is determined by various factors which include:

1. Rate of convergence - This is defined as the number of iterations required for the algorithm, in response to stationary inputs, to converge close enough to the optimum solution. A fast rate of convergence allows the algorithm to adapt rapidly to a stationary environment of unknown

70

statistics. Furthermore, it enables the algorithm to track statistical variations when operating in a non-stationary environment.

2. Computational complexity - This is the number of operations required to make one complete iteration of the algorithm.

3. Misadjustment - For an algorithm of interest, this parameter provides a quantitative measure of the amount by which the final value of the mean square error, averaged over an ensemble of adaptive filters, deviates from the optimal minimum mean square error.

4. Numerical properties - When an algorithm is implemented numerically, inaccuracies are produced due to round-off noise and representation errors in the computer. These kinds of errors influence the stability of the algorithm.

The three equalizer algorithms describes here are - zero forcing (ZF) algorithm, the least mean square (LMS) and the recursive least square (RLS) algorithm.

4.4.1 Zero Forcing Algorithm

In a zero forcing equalizer, the equalizer coefficients are chosen to force the samples of the combined channel and equalizer impulse response to zero at all but one of the NT spaced sample points in the tapped delay line filter. It was developed by Robert Lucky for wireline communication. The zero forcing

71

equalizer has the disadvantage that the inverse filter may excessively amplify noise at frequencies where the folded channel spectrum has high attenuation. The ZF equalizer thus neglects the effect of noise altogether and is not often used.

4.4.2 Least Mean Square Algorithm

A more robust equalizer is the LMS equalizer where the criterion used is the minimization of the mean square error (MSE) between the desired equalizer output and the actual equalizer output.The derivation of it is described in the previous sections of this chapter.

4.4.3 Recursive Least Squares Algorithm

The convergence of the gradient-based LMS algorithm is very slow, especially when the eigenvalues of the input covariance matrix have a very large spread. In order to achieve faster convergence, complex algorithms which involve additional parameters are used. Faster convergence is based on a least squares approach, as opposed to the statistical approach used in the LMS algorithm. That is, rapid convergence relies on error measures expressed in terms of a time average of the actual received signal instead of a statistical average. This leads to a family of powerful, albeit complex, adaptive signal processing techniques known as recursive least squares (RLS) which significantly improves the convergence of the adaptive equalizers. The Kalman RLS algorithm is one algorithm that falls under this category.

72

In this work LMS was used for weight determination.

4.5   Previous Work

To correct the ISI, DFE (Decision Feedback Equalizer) is traditionally used. An adaptive DFE is used to compensate for channel variations.

R. W. Lucky in 1965 [41-42], first proposed and analyzed ADFE (Adaptive DFE), building on earlier work in adaptive filtering by B. Widrow and M. E. Hoff, Jr. In 1960. DFE and ADFE is implemented using feedforward and feedback linear FIR filters as the assumption is that the storage channel response is linear.

K. Fisher, J. Cioffi and C. M. Melas [6] proposed an adaptive DFE for storage channels suffering from nonlinear ISI. They propose a modified version of DFE called the RAM- DFE (Random access memory - DFE) by replacing the transversal filter in the DFE feedback with a random access memory. The RAM is a general nonlinear function in which the most recent decisions form a binary address for the RAM with the addressed value being used as feedback. One flaw of this method is slow adaptation.

The work in this dissertation presents a method based on adaptive Volterra equalizers. Volterra equalizers have been used in satellite communication channels which produce high nonlinear distortions. In this work linear equalizers, nonlinear (Volterra) equalizers and some optimum combination of linear and nonlinear equalizer are analyzed to obtain the best performance for

73

the system.

Another class of equalizers which are considered as nonlinear are MLSE (Maximum Likelihood Sequence Estimator) Equalizers. The MLSE tests all possible data sequences (rather than decoding each received symbol by itself) and chooses the data sequence the data sequence with the maximum probability as the output. An MLSE usually has a large computational requirement. if L be the most recent input samples, then there are M^L states, where M is the size of the symbol alphabet of the modulation. MLSE is optimal in the sense that it minimizes the probability of a sequence error. For more details refer to [7].

Yet another class of nonlinear equalizers that have been studied in the literature are derived from Neural Networks.

In the next chapter the performance evaluation of the various combination of linear and nonlinear equalization methods are presented.

Chapter 5

Performance Evaluation of Linear, Volterra & Hybrid Equalization

This chapter presents quantitative results of various equalization techniques. Firstly, a SNR comparison measure is developed and then performance evaluation of Linear, pure Volterra and Hybrid equalization is examined.

5.1 SNR Computation for Performance Analysis

The computation of SNR can be understood as follows:

Let W_1, W_2, W_3, $\cdots$, W_N be the mark widths recorded on optical disk.

Let $\hat{W}_1$, $\hat{W}_2$, $\hat{W}_3$, $\cdots$, $\hat{W}_N$ be the marks recovered from the readback signal.

Let d_1, d_2, d_3, $\cdots$, d_N be the difference between W_i and $\hat{W}_i$. Then

$$d_i = W_i - \hat{W}_i; \quad i = 1, 2, \cdots N. \tag{5.1}$$

The variance is defined as

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} d_i^2. \tag{5.2}$$

Therefore, SNR is

$$SNR = 10 \log_{10}(\frac{W_{unit}^2}{\sigma^2}) \tag{5.3}$$

75

where W_{unit} is the width of the minimum mark size recorded.

Eq. 5.3 shows that as the W_{unit}, the minimum mark size increases the SNR increases. This is because the amount of ISI as a percentage of the mark size decreases. Also, the variance of the difference between recorded and recovered marks decreases.

To do the performance comparison, following steps were followed: The data symbols with ISI were obtained. Then, weight vectors for different configurations of equalizer were determined. The equalizers considered include - a Linear Decision feedback Equalizer, a pure Volterra Equalizer and a Hybrid equalizer which is a combination of a Linear equalizer and a pure Volterra equalizer.

5.2   SNR with and without Linear Equalization

Fig. 5.1 shows the SNR performance with and without Linear equalization as a function of the minimum mark size. The 3-tap Linear equalizer gives a gain of about 1.5 to 2 dB over the un-equalized system. This shows the importance of doing equalization. The Linear equalization is a very well established traditional method of correcting ISI.

5.3   SNR Comparison with and without Volterra Equalization

Fig. 5.2 shows the SNR as a function of the minimum mark width of the raw marks, marks obtained with a Linear 3-tap equalizer and marks
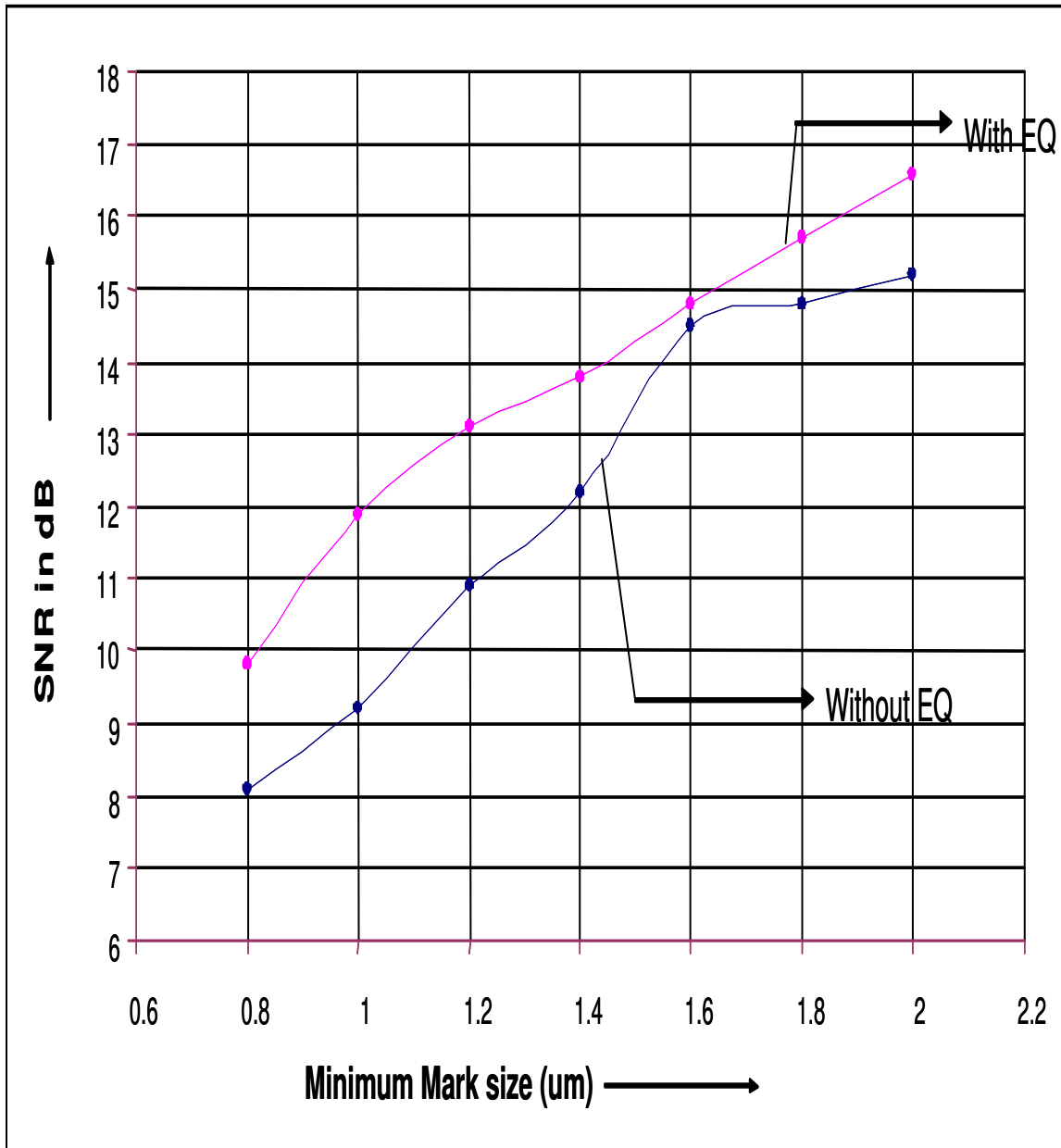
Figure 5.1: SNR with and without linear Equalization.

obtained using a 3-tap 3rd order Volterra equalizer. From the graph it is clear that the nonlinear equalization using a Volterra equalizer yields an additional performance gain of about 1.3-1.5 dB over a Linear equalizer and about 3 dB gain over un-equalized readback marks. The SNR increases as the minimum mark width is increased because the relative amount of ISI decreases.

5.4 SNR Comparison of Linear, Pure Volterra and Hybrid Equalization

Fig. 5.3 shows the SNR comparison of N-tap linear, N-tap 3rd order Volterra and Hybrid (N-tap Linear + 3-tap 3rd Volterra) equalizers versus the number of taps. Using hybrid equalization one can obtain very good performance in signal quality at a very modest increase in computational resources. Using pure Volterra equalization does not payoff much as it reaches a point of diminishing returns after 5 taps. The gain in SNR going from Linear to Hybrid equalization is about 1.3-1.5 dB at 5-taps and beyond.

In the literature DFE (decision feedback equalizers) are also described as nonlinear because its transfer function is a rational function of z^{-1} and is an IIR (infinite impulse response) filter with poles and zeros. The equalizer basically has two linear tap delay lines in the feedforward and feedback directions.

In the case of Volterra equalizers one has higher order terms that is why it is called nonlinear.

It is also worth mentioning that in-spite of the good performance of

78
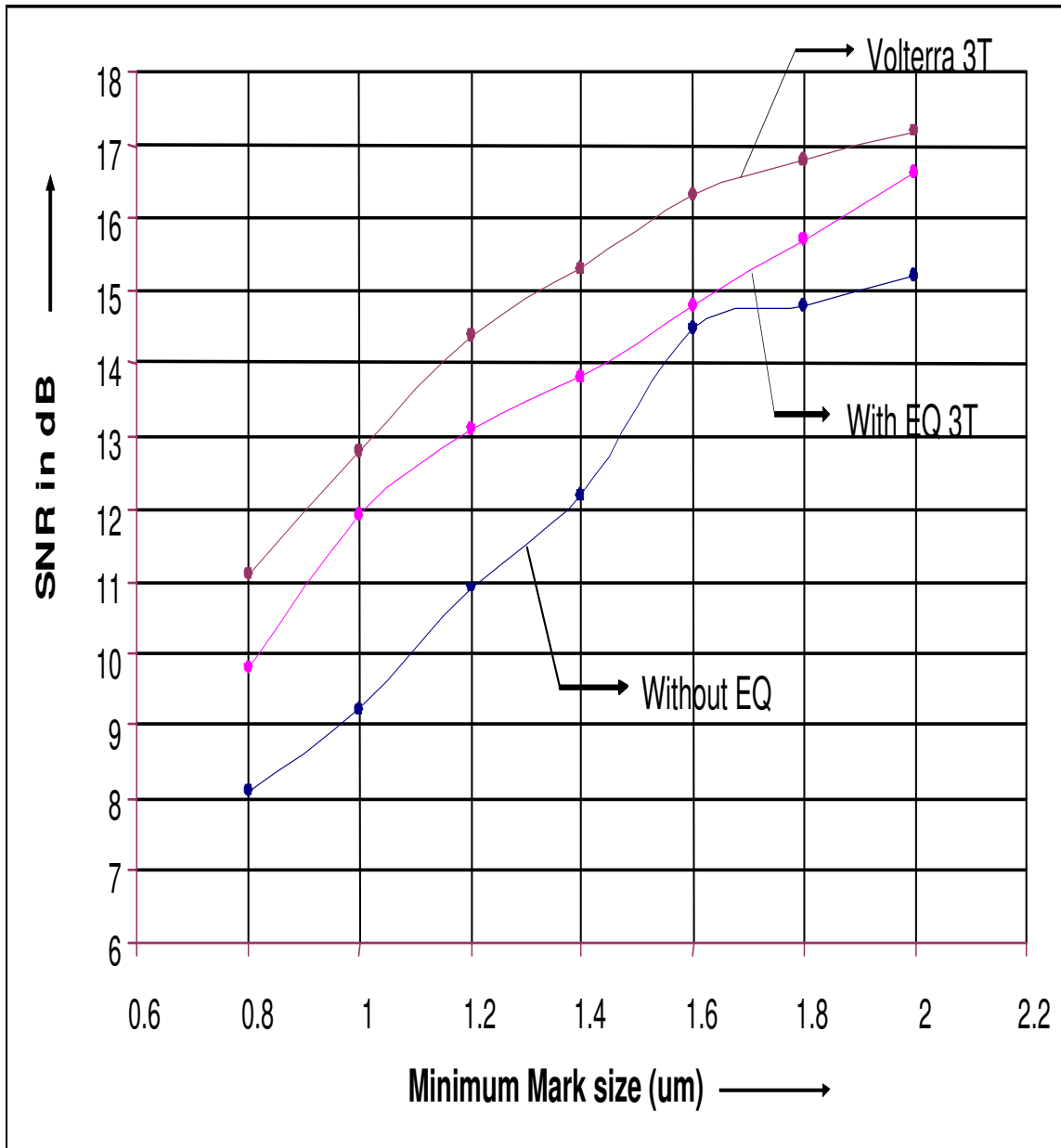
Figure 5.2: SNR with and without Volterra Equalization.
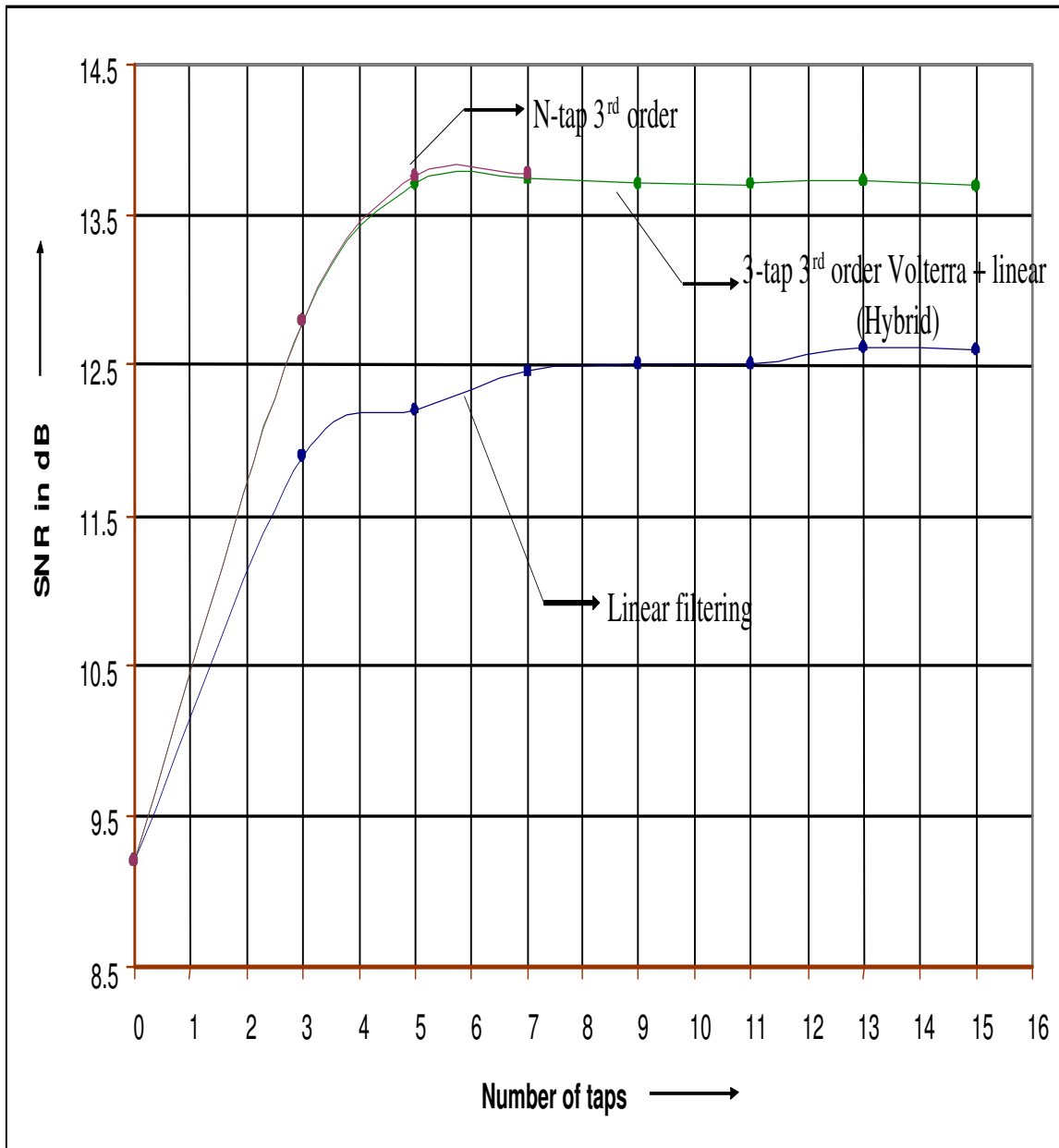
Figure 5.3: SNR with Linear, pure Volterra and Hybrid Volterra Equalization.

the decision feedback equalizers (Linear, Volterra and Hybrid), they suffer from error propagation problem. If a mark is decoded erroneously, it will be feedback into the feedback delay line and will affect the performance of the equalization. It can be kept in control by keeping the error probability low and keeping the feedback filter short enough. Another solution is to use Tomlinson-Harashima precoding (this issue is not investigated here but it is something one can incorporate in the overall system design). The forward error correction incorporated in the overall system also aids in the prevention of error propagation.

Thus, this chapter focused on the linear and nonlinear equalization techniques.

Chapter 6

Conclusion, Discussion and Future Work

Optical data storage has become important for its high storage density, random access, re-writability, portability, and long archival life. It offers a solution for storing huge amount of data being generated. It has very large shelf life, predicted to be more than 100 years. It is immune to electromagnetic interference. At present storage densities of around 5-10 GB on a 4.75" platter are commonly available. Under development are 20-40 GB and higher capacities.

This dissertation mainly investigated Volterra and Hybrid (combination of Volterra and Linear) equalizers for compensating ISI (Intersymbol Interference) in magneto-optic data storage systems. It also presented the nature of ISI.

In a magneto-optic data storage systems, nonlinear intersymbol interference (ISI) is one of the limiting factors for high storage density. The ISI characteristics were determined by making use of the various data patterns. This gives us insight into its nonlinear behavior. To reduce its effect equalization is used. Volterra and Hybrid nonlinear equalizers are proposed and developed. Performance of these equalizers is compared to that of linear equalizers.

The Hybrid equalizer provided the best performance. The gain achieved in signal quality using a 3-tap 3rd order Volterra and an N-tap linear equalizer was about 1.3-1.5 dB over an N-tap linear equalizer. It was observed that going to higher order Volterra equalizers (e.g., 5-tap 5th order) does not yield any extra benefit as one reaches a point of diminishing returns immediately after 3-tap 3rd order.

Convergence issues related to adaptive algorithms were also investigated. It was found that the eigenvalue spread of the covariance matrix was large for the Volterra/Hybrid equalizers compared to linear equalizers which resulted in somewhat slower convergence. This can be mitigated by employing other algorithms such as multiple-step LMS or Kalman.

The extra signal gain/quality obtained can be used/dissipated in a number of ways in the overall system design, such as -

1. Reduce the amount of redundancy in FEC (forward error correction). FEC consists of block encoding followed by interleaving followed by convolutional encoding. Redundancy reduction will result in increased storage density.

2. Making use of less expensive optical read/write assembly/detection.

3. Lower BER (bit error rate).

4. Higher storage density, as the resulting additional ISI can be compensated.

5. Smaller length (memory) of Viterbi decoding.

The gains of Volterra/Hybrid equalizers are not just limited to the field of data storage. Any communication system exhibiting nonlinear ISI characteristics can utilize this method for improving the signal quality. The communication system may utilize cooper wire where the source of ISI is the limited bandwidth or it can be a fiber optic system where the source of ISI is the chromatic dispersion or it can be a wireless system where a similar effect is produced by multipath dispersion.

A few items of interest for further investigation are -

1. To compare the performance of proposed equalization with a MLSE (Maximum Likelihood Sequence Estimation). It will be most computationally demanding and it will be of interest to see how much gain in signal quality is achieved say for constraint length of 5 or 7. This method will have to analyze M^L states, where M is the size of the alphabet (number of symbols) and L is the constraint length.

2. Hardware implementation issues. The equalization filters have to be implemented in hardware in the overall system design. There are three options for implementation a) Making an ASIC b) Using programmable logic devices such as FPGA/CPLD. c) Using a DSP processor. Out of the three choices best option in terms of the speed of execution will be to make an ASIC but making an ASIC demands a mass market as it is expensive and does not permit design changes.

84

Of the remaining two choices PLDs offers the best combination of hardware execution and the flexibility to make design changes. Benchmarking of common signal processing algorithms such as FFT, convolution, FIR filtering, etc on FPGAs and comparable DSP processors have shown that FPGAs out perform the DSP processors. In the FPGAs execution can be performed in hardware and in parallel unlike a DSP where instructions are fetched, decoded and executed. At the same time FPGAs offer the flexibility to make design changes which can be easily re-synthesized and reprogrammed into them.

The hardware implementation issues relate to efficient HDL (VHDL/Verilog) coding i.e., how to describe the task in an efficient manner so that execution is faster. There are various ways one can accomplish that such as reducing the length of critical path, reducing the number of logic levels in the critical path. One can code such that one gives priority to late arriving signals, avoiding nested if statements and instead using case statements, plus similar other techniques.

Appendix

Appendix 1

In this appendix we describe the algorithm used for smoothing the digitized waveform. Let y_i be the amplitude of the ith data point of the raw waveform and $\hat{y}_i$ be the amplitude of the ith data point of the smoothed waveform.

$\hat{y}_i$ is determined by taking 5 data points before and 5 data points after the ith data point.

$\hat{y}_i$ is estimated as

$$\hat{y}_{i+k} = a_2 x_{i+k}^2 + a_1 x_{i+k} + a_0 \tag{1.1}$$

$$= \sum_{j=0}^{2} a_j x_{i+k}^j \tag{1.2}$$

where $k = 0, \pm 1, \ldots, \pm 5$, and x_i is the time point corresponding to y_i shifted by i positions. Therefore, we simply set $x_i = 0$ The coefficients a_j are determined as follows:

Let z be the square sum of differences between y_{i+k} and $\hat{y}_{i+k}$ for $k = 0, \pm 1, \ldots, \pm 5$. We then have

$$z = \sum_{k=-5}^{5} (y_{i+k} - \hat{y}_{i+k})^2 \tag{1.3}$$

$$= \sum_{k=-5}^{5} (y_{i+k} - \sum_{j=0}^{2} a_j x_{i+k}^j)^2. \tag{1.4}$$

87

To Minimize z with respect to each of a_j, we set

$$\partial z/\partial a_j \;=\; -2\sum_{k=-5}^{5}(y_{i+k} - \sum_{j=0}^{2} a_j x_{i+k}^j)x_{i+k}^j \tag{1.5}$$

$$=\; 0. \tag{1.6}$$

The above expression yields the 3 following equations:

$$\sum_{k=-5}^{5} y_{i+k}x_{i+k}^2 \;=\; \sum_{k=-5}^{5} a_2 x_{i+k}^4 + \sum_{k=-5}^{5} a_1 x_{i+k}^3 + \sum_{k=-5}^{5} a_0 x_{i+k}^2 \tag{1.7}$$

$$\sum_{k=-5}^{5} y_{i+k}x_{i+k} \;=\; \sum_{k=-5}^{5} a_2 x_{i+k}^3 + \sum_{k=-5}^{5} a_1 x_{i+k}^2 + \sum_{k=-5}^{5} a_0 x_{i+k} \tag{1.8}$$

$$\sum_{k=-5}^{5} y_{i+k} \;=\; \sum_{k=-5}^{5} a_2 x_{i+k}^2 + \sum_{k=-5}^{5} a_1 x_{i+k} + \sum_{k=-5}^{5} a_0. \tag{1.9}$$

Eqs. (1.7), (1.8) and (1.9) expressed in matrix form gives

$$\begin{pmatrix} \sum_{k=-5}^{5} x_{i+k}^4 & \sum_{k=-5}^{5} x_{i+k}^3 & \sum_{k=-5}^{5} x_{i+k}^2 \\ \sum_{k=-5}^{5} x_{i+k}^3 & \sum_{k=-5}^{5} x_{i+k}^2 & \sum_{k=-5}^{5} x_{i+k} \\ \sum_{k=-5}^{5} x_{i+k}^2 & \sum_{k=-5}^{5} x_{i+k} & \sum_{k=-5}^{5} \end{pmatrix} \begin{pmatrix} a_2 \\ a_1 \\ a_0 \end{pmatrix} \;=\; \begin{pmatrix} \sum_{k=-5}^{5} y_{i+k}x_{i+k}^2 \\ \sum_{k=-5}^{5} y_{i+k}x_{i+k} \\ \sum_{k=-5}^{5} y_{i+k} \end{pmatrix}.$$

The above set of equations are solved for each data point to get a new set of values which represent the smoothed waveform.

88

Bibliography

[1] S. Esener, "Future of optical data storage technologies," *NIST Panel Report, http://www.wtec.org/loyola/hdmem/01-01.htm*, 2001.

[2] E. D. Daniel and C. D. Mee, *Magnetic Recording Handbook*, New York: McGraw-Hill, 1989.

[3] D. Coppersmith R. L. Adler and M. Hassner, "Algorithms for sliding block codes," *IEEE Trans. Information Th.*, vol. IT-29, pp. 5–22, 1983.

[4] S. Gupta, "Intersymbol interference characterization in a magneto-optic data storage system," *M.S. Thesis*, May 1993.

[5] S. Haykin, *Adaptive Filter Theory, 2nd ed.*, Englewood Cliffs, NJ: Prentice Hall, 1991.

[6] J. Cioffi K. Fisher and C. M. Melas, "An adaptive DFE for storage channels suffering from nonlinear ISI," *IEEE ICC*, vol. 3, pp. 1638–1642, 1989.

[7] T. S. Rappaport, *Wireless Communications*, Upper Saddle River, NJ: Prentice Hall, 2001.

[8] M. W. Marcellin and H. J. Weber, "Two-dimensional modulation codes," *IEEE Journal on Selected Areas in Communications*, to appear.

[9] G. R. Davis and B. Kamali, "Characterization of digital magnetic recording systems," *Proceedings of IEEE, Phoenix Conference on Computers and Communications (PCCC85)*, pp. 134–141, Mar. 1985.

[10] J. W. Goodman, *International Trends in Optics*, Boston: Academic Press Limited, 1991.

[11] H. K. Thapar C. M. Melas J. H. Cioffi, W. L. Abbott and K. D. Fisher, "Adaptive equalization in magnetic-disk storage channels," *IEEE Communication Magazine*, vol. 28, pp. 14–29, Feb. 1990.

[12] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, Boston: Kluwer Academic Publishers, 1988.

[13] M. Mansuripur and G. A. N. Connell, "Signal and noise in magneto-optical readout," *J. Appl. Phys.*, vol. 53, pp. 4485–94, 1982.

[14] Nakamichi U.S.A. Corp., *Optical Memory system - 1000 Nakamichi, Torrance, CA*, U.S.A. Corp., 1989.

[15] G. V. Jacoby, "Signal equalization in digital magnetic recording," *IEEE Trans. on Magnetics*, vol. MAG-4(3), pp. 302–5, 1968.

[16] Hewlett Packard Application Note 358-3, *Time domain characterization of magnetic disk drives*, Palo Alto, CA, circa 1991.

[17] Dennis G. Howe and Hugh M. Hilden, "Shift error propagation in 2,7 modulation code," *IEEE J. on Selected Areas in Comm.*, vol. 10, pp. 223–32, 1992.

90

[18] S. K. Nair and J. Moon, "Nonlinear equalization for data storage channels," *IEEE International Conference on Communications, Cat. No.94CH3403-3*, vol. 1, pp. 250–4, 1994.

[19] A. Gutierrez and W. E. Ryan, "Performance of adaptive Volterra equalizers on nonlinear satellite channels," *IEEE International Conference on Communications, Cat. No.95CH35749*, vol. 1, pp. 488–92, 1995.

[20] C. Eun and E. J. Powers, "A new Volterra pre-distorter based on the indirect learning architecture," *IEEE Trans. on Signal Processing*, vol. 45, pp. 223–7, 1997.

[21] R. J. P. DeFigueiredo, "Beyond Volterra and Wiener: Some new results and open problems in nonlinear circuits and systems," *IEEE Symposium on Circuits and Systems, Cat. No. 98CB36268*, pp. 124–7, 1999.

[22] G. M. Raz and B. D. Van Veen, "Blind linearization, and identification of nonlinear systems - a least squares approach," *IEEE Trans. Signal Process.*, vol. 48, pp. 192–200, 1998.

[23] B. I. Finkelstein P. Asthana and A. A. Fennema, "Rewritable optical disk drive technology," *IBM R&D Journal*, vol. 40, pp. 543–58, 1996.

[24] C. H. Cheng and E. J. Powers, "A reconsideration of the pth-order inverse predistorter," *IEEE 49th Vehicular Technology Conference (Cat. No.99CH36363)*, vol. 2, pp. 1501–4, 1999.

[25] A. Canella L. Agarossi, S. Bellini and P. Migliorati, "A Volterra model for the high density optical disc," *IEEE ICASSP*, vol. 3, pp. 1605–8, 1998.

[26] F. Bregoli L. Agarossi, S. Bellini and P. Migliorati, "Equalization of nonlinear optical channels," *IEEE International Conference on Communications, Cat. No.98CH36220662*, vol. 2, pp. 662–7, 1998.

[27] *Magneto-Optical Storage Systems*, US BYTE, http://www.usbyte.com, 2000.

[28] D. Wright, "Optical data storage," *www.cs.man.ac.uk/eissrg/optstor.htm*, Sept. 2000.

[29] C. H. Cheng and E. J. Powers, "Optimal third-order Volterra kernel estimation algorithms for a nonlinear communication system for PSK and QAM inputs," *IEEE Transactions on Signal Processing*, vol. 49, pp. 147–63, 2001.

[30] A. K. Nandi M. Ghogho and A. Swami, "An overview of the field of optical disk data storage," *http://itri.loyola.edu/hdmem/0502.htm*, June 1999.

[31] Optical Storage Technology, "http://www.osta.org," Sept. 2000.

[32] B. Liners M. Schaenzer and L. Schulz, "Matching burnish head design to next-generation media," *Seagate Tech, Data Storage Magazine, http://ds.pennet.com*, Sept. 2000.

92

[33] S. Im and E. J. Powers, "A block LMS algorithm for third-order frequency-domain Volterra filters," *IEEE Signal Processing Letters*, vol. 4, pp. 75–8, March 1997.

[34] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, New York: Wiley, 1980.

[35] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*, Englewood Cliffs, NJ: Prentice Hall, 1975.

[36] D. Woods, "Improved disk drives through the computational power of DSP," *Data Storage Magazine, http://ds.pennet.com*, Sept. 2000.

[37] M. K. Loze and C. D. Wright, "Strategies for reducing the effects of thermal interference in optical recording," *www.cs.man.ac.uk/eissrg/optstor.htm*, Sept. 2000.

[38] E. C. Gage T. K. Hatwar D. R. Kaiser, Charles F. Brucker and G. O. Simmons, "Large format multifunction 2-terabyte optical disk storage system," *http://esdis-it.gsfc.nasa.gov/MSST/conf1996/C44Kaiser.html*.

[39] Optical Storage Glossary, "PC tech. guide," *http://www.pctechguide.com*.

[40] Volterra V., *Theory of Functions and of Integral and Integro-Differential Equations*, New York: Dover Publications, 1959.

[41] Wiener N., *Nonlinear Problems in Random Theory*, Cambridge, MA: M.I.T. Press, 1959.

[42] R. J. de Figueiredo and T. A. W. Dwyer III, "A best approximation framework and implementation for simulation of large-scale nonlinear systems," *IEEE Trans. on Circuits and Systems*, vol. CAS-27, pp. 1005–1014, 1980.

[43] R. W. Lucky, "Automatic equalization for digital communication," *BSTJ*, vol. 44, pp. 547–588, 1965.

[44] J. P. Harris, W. B. Philips, J.F. Wells, and W. D. Winger, "Innovations in the design of magnetic tape subsystems," *IBM J. Res. Dev.*, vol. 25, pp. 691–699, 1981.

[45] L. D. Stevens, "The evolution of magnetic storage," *IBM J. Res. Dev.*, vol. 25, pp. 663–675, 1981.

[46] D. W. Brede, J. M. Harker, R. E. Pattison, G. R. Santana, and L. G. Taft, "A quarter century of disk file innovation," *IBM J. Res. Dev.*, vol. 25, pp. 677–689, 1981.

[47] P. A. Franaszek, "Sequence-state methods for run-length-limited coding," *IBM J. Res. Dev.*, vol. 14, pp. 376–383, 1970.

[48] H. Kobayashi, "A survey of coding schemes for transmission or recording of digital data," *IEEE Trans. Commun. Tech.*, vol. COM-19, pp. 1087–1100, 1971.

[49] A. M. Patel, "Adaptive cross-parity check (axp) code for a high density magnetic tape subsystem," *IBM J. Res. Dev.*, vol. 29, pp. 546–562, 1985.

94

[50] L. R. Bahl and D. T. Tang, "Blocks codes for a class of constrained noiseless channels," *Inform. Contr.*, vol. 17, pp. 436–461, 1970.

[51] J. K. Wolf and E. Zehavi, "On runlength codes," *IEEE Trans. Inform. Th.*, vol. IT-34, pp. 45–54, 1988.

[52] R. E. Blahut, *Principles and Practice of Information Theory*, Reading, MA: Addison-Wesley, 1987.

[53] A. S. Hoagland, *Digital Magnetic Recording*, New York, NY: Wiley, 1963.

[54] L. E. Franks, *Signal Theory*, Englewood Cliffs, NJ: Prentice Hall, 1969.

[55] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, pp. 379–423 and 623–656, 1948.

[56] H. Kobayashi and D. T. Tang, "Application of partial-response channel coding to magnetic recording systems," *IBM J. Res. Dev.*, pp. 368–375, 1970.

[57] Y. Takasaki, "Timing extraction in baseband pulse transmission," *IEEE Trans. Commun.*, vol. COM-20, pp. 877–884, 1972.

[58] N. D. MacKintosh, "A superposition-based analysis of pulse-slimming techniques for digital recording," *The Radio and Electronic Engineer*, vol. 50, pp. 307–314, 1980.

[59] J. P. Bubrouski and L. E. Franks, "Statistical properties of timing jitter in a PAM timing recovery scheme," *IEEE Trans. Commun.*, vol. COM-22, pp. 913–920, 1974.

[60] W. A. Gardner, "Characterization of cyclostationary random signal process," *IEEE Trans. Inform. Th.*, vol. IT-21, pp. 4–14, 1975.

[61] L. E. Franks, "Synchronization subsystems-analysis and design," in *Digital Communications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[62] P. H. Siegel, "Recording codes for digital magnetic storage," *IEEE Trans. Mag.*, vol. MAG-21, pp. 1344–1349, 1985.

[63] D. S. Bloomberg and K. Norris, "Channel capacity of charge-constrained run-length-limited codes," *IEEE Trans. on Mag.*, vol. MAG-17, pp. 3452–3455, 1981.

[64] L. J. Fredrickson and J. K. Wolf, "Coding using multiple block (d,k) codes," in *IEEE International Conference on Communications*, 1989, vol. 3, pp. 1623–1627.

[65] Y. Lin and J. K. Wolf, "Combined ECC/RLL codes," *IEEE Transactions on Magnetics*, vol. 24, pp. 2527–9, 1988.

[66] G. Ungerboeck and J. K. Wolf, "Trellis coding for partial-response channels," *IEEE Trans. Commun.*, vol. COM-34, pp. 765–773, 1986.

[67] A. R. Calderbank, C. Heegard, and T. A. Lee, "Binary convolutional codes with application to magnetic recording," *IEEE Trans. Inform. Th.*, vol. IT-32, pp. 797–815, 1986.

[68] P. Kabal and S. Pasupathy, "Partial-response signaling," *IEEE Trans. Commun.*, vol. COM-23, pp. 921–934, 1975.

[69] A. Lender, "Correlative digital communication techniques," *IEEE Trans. Commun. Tech.*, pp. 128–135, 1964.

[70] G. S. Dixon, C. A. French, and J. K. Wolf, "Results involving (d,k) constrained M-ary codes," in *IEEE Transactions on Magnetics*, vol. MAG-23.

[71] L. E. Franks, "Carrier and bit synchronization in data communication-a tutorial review," *IEEE Trans. Commun.*, vol. COM-28, pp. 1107–1121, 1980.

[72] P. Ho, *The Power Spectral Density of Digital Continuous Phase Modulation with Correlated Data Symbols*, Ph.D. thesis, Queen's University, Kingston, Ontario, Canada, Aug. 1985.

[73] I. Sasase, S. Mori, and K. Feher, "Bandwidth efficient modulation technique by asymmetrical pulse shaping," in *Communications in the Information Age (Cat. No. 84CH2064-4)*, Atlanta, GA, Nov. 1984, IEEE Global Telecommunications Conference, GLOBECOM '84, pp. 1086–1090.

97

[74] I. Sasase, K. Feher, and S. Mori, "Asymmetrical pulse shaped QPSK modulation systems," *IEEE Trans. Commun.*, vol. COM-33, pp. 1146–1150, 1985.

[75] F. Amoroso, "Pulse and spectrum manipulation in the minimum (frequency) shift keying (MSK) format," *IEEE Trans. Commun.*, vol. COM-24, pp. 381–384, 1976.

[76] M. K. Simon, "A generalization of minimum-shift-keying (MSK)-type signaling based upon input data symbol pulse shaping," *IEEE Trans. Commun.*, vol. COM-24, pp. 845–856, 1976.

[77] E. Biglieri and M. Visintin, "A simple derivation of the power spectrum of full-response CPM and some of its properties," *IEEE Trans. Commun.*, vol. 38, pp. 267–269, 1990.

[78] J. G. Proakis, *Digital Communications, 2nd ed.*, New York: McGraw-Hill Book Company, 1989.

[79] F. Harris, "On the use of windows for harmonic analysis with the DFT (Discrete Fourier Transform)," *Proceedings of the IEEE*, vol. 66, pp. 51–83, 1978.

[80] K. Feher, *Digital Communications: Satellite/Earth Station Engineering*, Englewood Cliffs, NJ: Prentice-Hall Inc., 1983.

98

Vita

Sunil Gupta received his B.Tech (Bachelor of Tecnology) degree in Electrical Engineering from Indian Institute of Technology, New Delhi, India in 1989. He received his M.S. (Master of Science) degree in Electrical and Computer Engineering from The University of Arizona, Tucson, U.S.A. in 1993. Presently, he is pursuing his Ph.D. studies in Electrical and Computer Engr. at the University of Texas at Austin and expect to receive his degree in 2004.

He has been employed at the Intel Corp., Lattice Semiconductor Corp., I.B.M. and the Electrical & Computer Engineering Departments at the University of Texas at Austin and the University of Arizona.

He has worked in the areas of Flash memory, DFT (design for testability), Digital/Analog design, Programmable Logic Devices (FPGA, CPLD) CAD flow development, Communication, Magneto-Optic Data Storage, Software development and Image processing/Computer vision.

Permanent address: A-8-G D.D.A. Flat
                         Munirka
                         New Delhi
                         India - 110067

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.