Modeling, Detection and Signal Design for Multichannel Fiber Optic Communications

A Dissertation

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

In Partial Fulfillment

of the requirements for the Degree

Doctor of Philosophy

Electrical and Computer Engineering

by

Kirtan Narendrakumar Modi

May 2009
Abstract

Starting with the TAT-8 transatlantic telephone cable, long-haul fiber optic cables have played a significant role in carrying world-wide voice and data traffic. Research efforts in optical device technology have led to the emergence of all-optical fiber networks that promises to reduce the maintenance cost and the electro-optical conversion bottleneck for long-haul fiber communications. Overcoming the effects of cross-talk is one of the challenges being addressed by researchers to realize the low-cost high-speed promise of all-optical communications. This dissertation contributes to the design of signal processing algorithms being investigated for cross-talk cancelation in multispans multichannel all-optical fiber communications.

Cross-talk effects considered in this dissertation arise primarily from fiber nonlinearity or inter-channel leakages that occur inside optical devices. Discrete-time models characterizing nonlinear fiber effects have previously been proposed but not fully validated for multispans systems. This dissertation begins by characterizing the effect of fiber nonlinearity on standard system performance. The derivation and validation of a discrete-time model for the channel is considered. The tradeoff between the model complexity and accuracy for different system parameters is studied. Results presented can be used to determine the model to be employed as system parameters change and the number of repeaters that are required.

Appropriate receiver based, transmitter based, or joint transmitter-receiver based
signal processing approaches to mitigate cross-talk are applied to the models obtained. For receiver based signal processing, a new low-complexity multiuser detection algorithm that uses a regression based polynomial approximation of the optimal detector is proposed and applied to designing affine and quadratic detectors. Transmitter based signal design problems that use cross-talk information available at the transmitter to jointly choose the transmitted signal set are researched. For cases where transmitter based signal design alone is insufficient for zero-forcing the cross-talk, a joint transmitter-receiver signal processing approach employing signal design and multiuser detection is employed. Results show that depending on channel conditions, different low-complexity signal processing schemes can be implemented to successfully reduce the effects of cross-talk.
Approvals

This dissertation is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy Electrical and Computer Engineering

Kirtan Narendrakumar Modi

Approved:

Maïté Brandt-Pearce (Advisor)  Stephen G. Wilson (Chair)

Zongli Lin  Stephen Patek

Toby Berger

Accepted by the School of Engineering and Applied Science:

James H. Aylor (Dean)

May 2009
Acknowledgments

I believe that an atmosphere supporting the free spirit of investigation is required for researchers to develop a feeling of thorough understanding about a particular field, and make subsequent contributions that help us understand more. I was fortunate to have Dr. Maïté Brandt-Pearce as my advisor. She granted me the freedom to explore, while ensuring that I progressed in the right direction, encouraged abstract thought, was patient and motivated me with her insight as I struggled to understand fiber optic systems and the field of research itself, and worked relentlessly to improve the quality of my research. I sincerely thank her for all her efforts in furthering my academic career, and am indebted to her for the tireless efforts to improve this dissertation.

I would like to express my gratitude to Dr. Stephen Wilson, Dr. Toby Berger, Dr. Stephen Patek and Dr. Zongli Lin for their insightful comments on this research. I thank the National Science Foundation for funding the project related to this research. My graduate experience was made enjoyable by the numerous discussions, often illuminating and profound, sometimes confusing and trite, but always fun, that I had with my friends. I thank Dr. Satya Ponnaluri for the technical, Souvik Das for the nontechnical, Vivek Balasubramanian and Nilav Bose for the motivational, and Nishant Sinha for everything else.

Finally, I would like to thank my parents Uravashi Modi and Narendra Modi,
my brother Mitul Modi. Their undying faith in me, the values they have instilled in me, the lessons of humility and hard work that they have taught kept me grounded and motivated to finish this dissertation. This dissertation may be far from perfect and may not measure to their sacrifices that enabled me to pursue my make believe dreams; nevertheless, I humbly dedicate this work to them.
## Contents

1 Multichannel Fiber Optic Communications 1  
  1.1 Introduction ..................................................... 1  
  1.2 Multispan Fiber Optic Communication Systems ................. 4  
  1.3 Research Problems ............................................... 7  
  1.4 Summary .......................................................... 13  

2 Effects of Nonlinear Cross-talk 14  
  2.1 Introduction ..................................................... 14  
  2.2 System Model ................................................... 16  
  2.3 Nonlinear Schrödinger Wave Equation and Channel Behavior .. 19  
  2.4 Split-step Fourier Simulation .................................. 21  
  2.5 Simulation Results ............................................. 23  
  2.6 Conclusion ....................................................... 31  

3 Modeling Fiber Cross-talk Effects in Discrete-Time 33  
  3.1 Introduction ..................................................... 33  
  3.2 Applying the Volterra Series Transfer Function Approximation to Multispan Systems ........................................... 36  
  3.3 Discrete-Time Models and System Setup ........................ 39  
  3.4 Model Accuracy via Simulation .................................. 49
## Contents

3.5 Summary, Conclusion and Future Research ........................................ 60

4 Low-Complexity Multiuser Detection for WDM Systems ............................ 62
   4.1 Introduction ............................................................................. 62
   4.2 Noisy Channel Models for Multichannel Fiber Optic Communications .. 65
   4.3 Detector Design Problem ...................................................... 68
   4.4 High Complexity Multiuser Detection ...................................... 73
   4.5 Design of Sub-Optimal Detectors ........................................... 74
   4.6 Performance Analysis and Simulation Results ............................ 83
   4.7 Conclusion ............................................................................. 90

5 Precoding Based Signal Design for WDM with Linear Interference .......... 92
   5.1 Introduction ............................................................................. 92
   5.2 System Model ......................................................................... 95
   5.3 Precoder Design .................................................................... 96
   5.4 Performance Analysis .......................................................... 97
   5.5 Estimating the Correlation Coefficients .................................. 100
   5.6 Simulation Results ............................................................... 103
   5.7 Conclusion ............................................................................. 107

6 Signal Design for Nonlinear Interference Limited WDM Systems .......... 109
   6.1 System Model ....................................................................... 111
   6.2 Signal Design for Cross-Phase Modulation Effects ..................... 112
   6.3 Signal Design for Four-Wave Mixing Effects ............................. 118
   6.4 Summary and Conclusion ..................................................... 125

7 Conclusion and Future Research Directions ........................................ 127
   7.1 Effects of Fiber Nonlinearity .................................................. 127
List of Figures

1.1 A typical all-optical multispan fiber optic system with demodulation via optical demultiplexing and electrical filtering at the receiver. 4

2.1 Multispan fiber optic system with demodulation via optical filters that are matched to single channel systems. 17

2.2 Multispan fiber optic system with demodulation via optical demultiplexing and electrical filtering at the receiver. 17

2.3 Ordered scaled received samples for the center channel in 25 and 50 GHz systems with one and 16 spans, and 0.25 mW/channel. The samples are scaled to set the largest value to one. 26

2.4 Normalized EOs for systems with up to 16 spans at multiple transmit power levels. The plots correspond to 50 GHz (above), and 25 GHz (below), systems with optical matched-filtering. The case $\gamma = 0$ corresponds to the linear channel and captures the ACI effects. 28

2.5 Normalized EOs for systems with up to 16 spans at multiple transmit power levels. The plots correspond to 50 GHz (a), and 25 GHz (b), systems with optical bandpass filtering and lowpass electrical filtering. The case $\gamma = 0$ corresponds to the linear channel and captures the ACI effects. 29
List of Figures

2.6 Number of spans for which the EOs close for a given power level for systems with channel spacings of 50 GHz and 25 GHz. Systems with matched optical filter receiver as shown in Fig. 2.1 and with optical bandpass filter and electrical filter for demodulation as shown in Fig. 2.2 are considered. .......................................................................................................................... 30

3.1 Receiver for multispans fiber optic system with demodulation via optical demultiplexing and photodetection. Optical filters that are matched to single channel systems or practical optical bandpass filters are employed in the demultiplexing process. .......................................................................................................................... 34

3.2 Normalized MSE averaged across all power levels in the range of 0.25–1.25 mW/channel is given for the different models for multispans 25 GHz and 50 GHz systems. The system in Fig. 3.1 with single channel matched optical filters is considered in (a) while that with practical bandpass filters is considered in (b). .......................................................................................................................... 51

3.3 The number of spans for which the modeling MSE is within the specified tolerance of 1%, 2%, 5% and 10% for a 25 GHz system. The different power levels in the range of 0.25 mW/channel to 1.25 mW/channel are considered. Single channel optical matched-filter receiver is assumed. Identical results are found when practical bandpass filters are employed instead. .......................................................................................................................... 52
3.4 The number of spans for which the modeling MSE is within the specified tolerance of 1%, 2%, 5% and 10% for a 50 GHz system. The different power levels in the range of 0.25 mW/channel to 1.25 mW/channel are considered. Single channel optical matched-filter receiver is assumed. Identical results are found when practical bandpass filters are employed instead. ............................................................... 53

3.5 Extra error in the average normalized MSE caused when a single set of coefficients is obtained as compared to the average normalized MSE when coefficients are obtained separately for each power level. Only the system with optical demultiplexing via single channel matched filters given in Fig. 3.1 is shown. Identical results are found for the system with practical bandpass filters. The averaging for case with individually obtained coefficients is done over the relevant power range. 56

3.6 Scaled correlation coefficients for ACI-dominant and XPM-dominant systems for the fourth channel in a eight channel system. The coefficients for XPM-dominant systems are multiplied with the transmit power to normalize and compare with \( k = 4 \) in (3.21). Systems with optical demultiplexing via single channel matched filters is considered in (a) while that with practical bandpass filtering is considered in (b). A power level of 0.5 mW/channel is considered. The ACI terms and the XPM terms have no unit due to appropriate scaling. .................. 58

4.1 Decision region for user 1 in a two-user ACI-dominant system with a SNR of 25dB, \( A_1 = A_2 = 1, \rho_{i,j} = \rho = -0.3 \ \forall i \neq j \). (Regr.: Regression). .................................................... 81
4.2 (a) AME and (b) BER of different detectors in a two-user ACI-dominant system with $A_1 = A_2 = 1$. For the AME plot, $\rho_{1,2} = \rho_{2,1} = \rho$ is used and for the BER plot, $\rho = -0.3$ is used. The decision regions corresponding to the BER plot are shown in Fig. 4.1. (Regr.: Regression) 85

4.3 A two-user ACI-dominant system with $A_1 = A_2 = 1$, $\rho_{1,2} = \rho_{2,1} = \rho$ is considered. The figures show relative AME loss w.r.t. the optimal MUD when SUD, the proposed regression-based detectors (Regr.), the locally optimal gradient search based detectors (Grad.) and the genetic search based detectors (GA) are employed. 87

4.4 The AME of different detectors designed for a three-user ACI-dominant system with $A_i = 1$, $\forall i$, $\rho_{i,j} = \rho$ $\forall i \neq j$. (Regr.: Regression) 88

4.5 The AME achieved when different detectors are employed for two-user and three-user XPM-dominant systems with $A_i = 1$, $\forall i$, $\rho_{i,j} = \rho_{j,i} = \rho$, $\forall i \neq j$. (Regr.: Regression) 89

5.1 AME of optimal MUD and SUD without Precoding compared with the proposed joint precoding strategies. 101

5.2 Mean square error (MSE) for blind and training based ML estimation of correlation in a two-user system. 104

5.3 SUD and regression based Quadratic $y$ MUD without precoding are compared with different types of joint precoding strategies, when the correlation is assumed to be 0.4 or $-0.4$ for a two-user system. Performance of single user systems is plotted for comparison. 105
5.4 Achievable BER of ZF precoding and $\rho_F$-precoding with Quadratic–$y$ MUD compared with SUD without precoding for different cross-correlations. The effect of perfect and imperfect correlation estimates is shown. The estimates are generated using 128 samples for both Blind and ML approaches.  

6.1 Multiuser efficiency of precoding and optimal MUD for 2 and 3 users for XPM-dominant systems. Real positive precoding solutions that do not require any phase control at the transmitter are considered.  

6.2 Multiuser efficiency $\eta$ of precoding and optimal MUD for 2 and 3 users for XPM-dominant systems. Complex solutions for precoding requiring phase control of $\{0, \pi\}$ at the transmitter are allowed.  

6.3 Multiuser efficiency $\eta$ for AO MUD with signal design and AO MUD and SUD (without signal design) for a four-user FWM-dominant system.  

A.1 AME, $\eta_\theta$, of a convex combination detectors of two specific Affine–$x$ detectors, $g_0 = \{0.7133, 0.0421\}$ and $g_1 = \{0.8940, 0.0168\}$, and convex combination of AME of specific detectors, vs combination coefficient $\theta \in (0, 1)$ for a two-user system with linear interference, $A = 1$, $\rho_{i,j} = \rho = -0.4 \forall i \neq j$.  

List of Tables

2.1 System Parameters for Simulations ........................................ 24

3.1 Range of magnitudes of correlation coefficients for the system with 16
spans and parameters given in Table 2.1 ................................ 44

3.2 Number of correlation coefficients to be estimated per channel for the
four models given in (3.19), (3.20), (3.21) and (3.22). .............. 49
List of Acronyms

WDM  Wavelength Division Multiplexing
MUX  Multiplexer
DEMUX  Demultiplexer
BER  Bit-error rate
OOK  On-off Keying
EO  Eye Opening
ISI  Intersymbol Interference
SNR  Signal-to-Noise Ratio
DC  Dispersion Compensator
NLSE  Nonlinear Schrödinger Wave Equation
SSF  Split-step Fourier
ACI  Adjacent Channel Interference
SPM  Self-phase Modulation
XPM  Cross-Phase Modulation
FWM  Four-Wave Mixing
CW  Continuous Wave
MUD  Multiuser Detection
SUD  Single User Detection
AO  Asymptotically Optimal
AME  Asymptotic Multiuser Efficiency
ML  Maximum-Likelihood
ZF  Zero-Forcing
List of Symbols

\( K \)  Number of WDM channels

\( b \)  \( K \)-dimensional vector of transmitted bits

\( A_k \)  Transmitted amplitude for channel \( k \)

\( A \)  Transmitted amplitude for all channels when \( A_k \) are assumed equal

\( S^{(n)}(\omega, z) \)  Frequency domain optical field of signal propagating in span \( n \) at distance \( z \) from origin of span

\( s(t, z) \)  Optical field of signal propagating along a fiber at time \( t \) and distance \( z \) from the origin

\( h \)  Simulation step-size for distance in SSF simulation

\( \mathcal{F}\{\cdot\} \)  Time-domain Fourier transform

\( \mathcal{F}^{-1}\{\cdot\} \)  Inverse Fourier transform corresponding to \( \mathcal{F} \)

\( \rho_{k,m} \)  Correlation between channel \( k \) and \( m \) for ACI-dominant and XPM-dominant systems

\( \rho_{k,l,m,n} \)  Correlation between channel \( k, l, m, n \) for FWM-dominant systems

\( \rho \)  Correlation among users when \( \rho_{k,m} \) and/or \( \rho_{k,l,m,n} \) are assumed equal

\( y \)  Vector of received statistics at detector input

\( x \)  Square-root of received statistics in vector \( y \)

\( |\cdot|_o \)  Schür product of a vector
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathbf{\mu}(\cdot) )</td>
<td>Vector of noise-free received statistics in ( \mathbf{x} )-domain</td>
</tr>
<tr>
<td>( A )</td>
<td>Transmit constellation</td>
</tr>
<tr>
<td>( \mathbf{\mu}_A(\cdot) )</td>
<td>Vector of noise-free received statistics in ( \mathbf{x} )-domain when transmit constellation ( A ) is used</td>
</tr>
<tr>
<td>( \mathbb{E}(\cdot) )</td>
<td>Energy of transmit constellation</td>
</tr>
</tbody>
</table>
Chapter 1

Multichannel Fiber Optic Communications

1.1 Introduction

High-speed communications has been the primary application of optical communications. The Chappe telegraph system built in 1793 and credited as being the first practical industrial communication system was an optical system based on transmitting information visually [1]. It replaced post-riders and represented the high-speed long distance communication solution suitable for the era. Wired communication demonstrated by the Gauss-Weber telegraph in 1833 provided a much more reliable communication channel compared to the atmospheric medium used by the optical systems. Subsequent developments through the first half of the twentieth century were centered around wire-line electrical communications. Inventions like the photophone in 1880 failed to commercialize due to relative uncertainty of the atmospheric channel [2]. Unclad optical fibers emerged around 1930 as a way to guide light waves, but high energy loss made them unsuitable for practical applications. Optical fibers with cladding and losses on the order of 1000 dB/km appeared around 1966, [3], and
Chapter 1. Multichannel Fiber Optic Communications

were unsuitable for communication applications. Subsequent developments in man-
ufacturing optical fibers using silica glass with reduced impurities led to fibers with
losses as low as 0.47 dB/km in 1977 [1]. Applications and developments in the field of
fiber optic communications have since been quick, and optical fibers now carry most
of the telephonic and data traffic over short as well as large distances. Mitigating
cross-talk effects resulting from fiber nonlinearity and device leakages presents the
next set of challenges that are being addressed by current researchers. Designing bet-
ter devices and/or employing signal processing algorithms are the directions being
investigated. This dissertation contributes to the signal processing approaches for
treating fiber nonlinearity.

Present day fibers have losses no more than 0.2 – 0.3 dB/km in the large wave-
length region surrounding the 1.55 micrometer band. Engineers and physicists are
now involved in designing devices, communication protocols and transmission schemes
with the goal of optimizing the resources made available by these fibers. Wavelength
division multiplexing (WDM) is one such technology which involves transmitting data
on multiple optical carriers over a single fiber so as to make the terahertz bandwidth
accessible. Optical switches, demultiplexers and amplifiers have evolved to enable
cost-effective all-optical WDM communications [4]. One direction for further optimi-
zing long haul optical systems is to reduce the frequency spacings between adjacent
channel carriers for WDM systems. Current ITU specifications are for channel spac-
ings of 50, 100 and 200 GHz [5]. Systems with spacings of 25 GHz are now being
investigated heavily [6], [7], [8] and references therein. Reducing the channel spacings
has multiple benefits. Some of the key benefits are listed below:

1. Given a single fiber link, tight packing of channels brings in more capacity and
reduces the cost of transmission per bit.
2. Optical amplifiers have limited bandwidth. Increasing capacity by using more channels without decreasing the channel spacings would require advances in optical amplification technology to span the increased bandwidth [9].

3. When multiple optical amplifiers are used in cascade, the channels in the middle of the optical amplifier band get more amplified as compared to those on the edge. As the number of spans increases, this phenomenon reduces the effective flatness of the amplifiers [10].

Cross-talk effects act as a significant efficiency bottleneck when reducing the channel spacings. Cross-talk arises from two sources,

1. Leakages in nonideal optical devices, [11] and references therein, and

2. Fiber nonlinearities as described next.

All dielectric media have a nonlinear response to intense electromagnetic fields. For silica based optical fiber systems, the nonlinearity becomes more pronounced with increasing distances and higher power required to support more channels. This effect can deter the deployment of long-haul systems that have fewer and more widely spaced repeaters. Nonlinearity induces time and frequency domain cross-talk among copropagating signals. For all-optical long-haul WDM systems, the frequency domain inter-channel cross-talk increases significantly with reduced channel spacing, thereby increasing the error rate and reducing the channel capacity. Additionally, higher transmit power levels also experience higher levels of cross-talk due to nonlinearity. This prohibits transmitting more power to increase capacity and also reduces the energy efficiency over long distances.

Various approaches are being investigated to counter the effects of cross-talk. The primary ones include designing optical fibers with low nonlinearity, switches and other
devices with lower leakages, [12], [13] and references therein, advanced network layer solutions based on intelligent routing when applicable [14], and employing signal processing and communication theoretic approaches [15], [16], [17], [18], [19] and references therein. This dissertation builds upon the approach of employing low complexity signal processing, at both the transmitter and the receiver, to compensate for cross-talk effects.

The rest of the chapter is organized as follows: The fiber optic system considered in this dissertation is introduced in the following section. Relevant research problems and the organization of the following chapters contributing to these problems are given in Section 1.3. A brief summary of the contributions is provided in Section 1.4.

1.2 Multispan Fiber Optic Communication Systems

![Figure 1.1: A typical all-optical multispan fiber optic system with demodulation via optical demultiplexing and electrical filtering at the receiver.](image-url)
A typical present-day long haul, all-optical fiber optic communication system employing WDM is shown in Fig. 1.1. The transmitter consists of a bank of lasers and an optical WDM multiplexer (MUX) switch, used to simultaneously launch multiple channels over a single mode fiber. Each laser operates at a wavelength corresponding to a single WDM channel. To enable transmission over large distances, multiple fiber spans are employed as indicated by the figure. Successive spans are separated by a dispersion compensator and an optical amplifier, in order to compensate for material dispersion and attenuation. The receiver consists of an optical demultiplexer (DEMUX), a photodetector bank and an electrical filter bank that together act as demodulators. The demodulated signals are sampled and fed to the threshold detector which detects the transmitted information. The different components of the fiber system have different effects. These are briefly discussed next. Note that the system model here corresponds to that for direct detection systems.

1.2.1 Fiber nonlinearity induced cross-talk

Fiber nonlinearity has two effects, cross-talk among signals and cross-talk among signals and noise:

1. Cross-talk effects: Due to fiber nonlinearity, signals transmitted for a particular channel influence the signals propagating in other channels. This results in inter-channel (frequency domain) and intra-channel (time domain) cross-talk. While inter-channel effects are dominant when channels are packed closely, intra-channel effects are dominant when short pulses are employed. As higher data rate channels are now being packed closely in fibers in dense WDM systems, intra-channel and inter-channel interference are bound to occur significantly in future fiber optic systems.
2. Signal dependent and colored noise: Signal dependent noise due to channel non-linearity arises in multispan systems if amplifiers are placed at the end of each span. Optical amplifiers, employed to avoid electro-optical conversion bottlenecks, introduce amplified spontaneous emission (ASE) noise in the system. In multispan systems, copropagation of this ASE noise with the transmitted signals results in nonlinear amplification of noise. This effect, known as parametric gain [20], results in colored, signal dependent noise at the receiver front-end. The noise coloring also occurs from copropagation of noise over intermediate spans.

1.2.2 Optical amplifier saturation

Optical amplifiers are now widely used in optical systems. Typically a single amplifier operating in the saturation region is employed for all channels. The saturation region operation of the amplifier can act as an additional source of nonlinearity. Amplifiers are also the primary source of noise in the system.

1.2.3 Device leakage and channel overlap

Starting from the transmitter, the MUX switch, the dispersion compensators and amplifiers between spans to the DEMUX switch at the receiver, signal leakage takes places between copropagating channels as they pass through the various devices. The resulting cross-talk which is linear in nature, can significantly degrade performance even in the absence of nonlinear effects.
1.2.4 Square-law channel from photodetection

Direct detection optical communication systems do not use coherent modulation-demodulation schemes since phase detection is not yet possible at such high frequencies for long lengths of nonpolarization preserving fiber based systems. Typically, the transmitted signal in optical systems is constrained to be nonnegative. The photodetector acts as an envelope detector that only detects amplitudes. Since the photodetectors are square-law devices, the channel is referred to as a square-law channel. This introduces an additional nonlinearity into the system. The use of photodetectors in optically amplified systems causes signal dependent noise arising from $signal \times noise$ beat terms.

Treating the channel induced cross-talk as well as the signal dependent noise when designing transceivers can significantly enhance the energy and bandwidth efficiency of fiber based systems. The following section reviews some of the relevant research problems involved.

1.3 Research Problems

A wide range of problems exist for mitigating cross-talk effects via signal processing. Here the focus is on understanding and modeling nonlinear effects, and compensating for cross-talk effects via signal processing at the receiver and the transmitter. The nonlinear effects present a number of fundamental problems in the design of transmission and reception schemes. Efforts to understand and mitigate the effects of intra-channel cross-talk from a communication theoretic view have been well studied in [21], [22], [23], [24], [25]. Some effort has been made recently to understand and mitigate inter-channel cross-talk [16], [26], [27]. While signal processing based techniques used to treat intra-channel and inter-channel effects are in some cases in-
terchangeable, the noncausality and finiteness of the number of channels need to be considered when treating inter-channel effects. The eventual goal of applying communication theoretic principles to fiber optic systems is to design and analyze techniques for countering two-dimensional cross-talk. However, this research focuses on systems that are limited by single dimensional cross-talk, with emphasis on inter-channel cross-talk limited systems.

Before proceeding to applying signal processing to compensate for the cross-talk, it is vital to obtain a more concrete understanding of the nonlinear effects on the system under consideration. This requires two things: characterizing the effect of nonlinearity induced cross-talk on a standard system, and modeling the nonlinear channel behavior in discrete-time. Both of these problems are studied here.

Mathematically, the multichannel WDM system can be considered analogous to some multiuser wireless communication systems. Signal processing techniques for interference mitigation in multiuser/multichannel wireless systems have been heavily investigated. These techniques can be classified as

1. Interference post-cancelation at receiver using multiuser detection (MUD) [28], [29]

2. Interference pre-cancelation at transmitter using signal design/precoding, [30], [31], [32], and

3. Joint transmitter-receiver optimization, [33], [34] and references therein.

The square-law nature of the fiber system, combined with some unconventional interference structures like those resulting from nonlinearity induced cross-talk, differentiates the system investigated here from those for wireless systems. This dissertation investigates the application of the above techniques as suitable to WDM type optical
systems. With this in mind, the following research problems are addressed here.

1.3.1 Characterizing the impact of nonlinear effects

The overall effect of nonlinearity depends on the type of processing that takes place at the receiver. The system in Fig. 1.1 shows a receiver structure with an optical demultiplexer and an electrical filter. Understanding the effects of nonlinearity for different filters employed for the demultiplexer as well as the electrical filter is difficult, if not impossible. Two types of receiver structures for systems with 50 GHz and 25 GHz spacings are considered in Chapter 2.

1.3.2 Discrete-Time models for nonlinear channel effects

Various models have been proposed and studied for single span systems, [15], [35], [36] and references therein. The accuracy of the models when applied to multispan systems for different power levels and channel spacings is not clear. Discrete-time models that make multispan optical systems accessible to discrete-time statistical signal processing are studied in Chapter 3. The continuous-time model for fiber behavior is used to obtain a nonlinear discrete-time model for multispan systems with unmodulated continuous wave (CW) inputs. Low complexity approximations to this discrete-time model are then studied for accuracy when applied to realistic systems with appropriate pulse shaping and modulation.

The author would like to note that the terms interference and user, prevalent in wireless communications literature, are used interchangeably with the terms cross-talk and channel, respectively, throughout this dissertation.
1.3.3 Multiuser detection for interference mitigation at the receiver

Traditional approaches to designing receivers for fiber optic systems have treated cross-talk as noise. Cross-talk or interference results in correlated statistics at the multichannel receiver output. Recent work has focused on using joint channel processing via MUD schemes to improve system performance. High complexity equalization to overcome intra-channel interference effects has been considered in [37] and references therein. Optimal and sub-optimal detectors for suppressing inter-channel interference in systems with signal independent pre-detection noise are obtained in [26] and [27], both having high complexity.

This dissertation contributes to the body of research in [26] and [27] with the following:

1. Evaluate existence conditions for sensible detectors: For linear channels experiencing linear interference like in various wireless CDMA systems, the interference is modeled using a cross-correlation matrix. For such systems, a full rank cross-correlation matrix ensures the existence of sensible detectors, i.e., detectors that achieve bit-error-rates that reduce monotonically with increasing signal to noise ratio (SNR). It turns out that for the fiber systems studied here, a similar cross-correlation matrix can be used to characterize the channel behavior. Existence conditions, related to the sufficiency of the full rank condition to guarantee sensible detectors for such channels, are investigated.

2. Design of polynomial complexity detectors: Unstable laser phases can cause the fiber channel to appear highly time-varying. This along with the high data rates of optical systems motivates the search for low complexity detectors that can be quickly adapted to the channel. Due to signal dependent nature of
the post photodetection noise, conventional sub-optimal low-complexity designs may not necessarily help mitigate interference. Obtaining optimal polynomial complexity MUD for any channel is still an open problem. A new regression based technique for designing polynomial MUDs is proposed and applied to design affine and quadratic. It is found that the regression based approach can result in detectors that perform close to locally optimal detectors obtained using gradient based and genetic algorithm based searches.

1.3.4 Signal design for interference mitigation at transmitter

If quantitative information of the exact channel behavior is available at the transmitter, the transmitted signal can be designed to compensate for cross-talk effects. While there are some proposals for phase modulated transmissions [22], [38], and [39], the receiver is still assumed to employ a square-law detector. The development of direct-detection square-law channels is still not as complete as that for linear Gaussian channels and efforts are underway to develop signal designs, modulation and coding schemes for such systems [40], [41], and references therein. Signal designs based on channel information at the transmitter can significantly improve the overall system efficiency. With the advent of advanced optical device technologies, amplitude and phase controlled transmissions are becoming feasible [39], further motivating such signal designs. Signal design problems for square-law channels, [40] and references therein, and for noncoherent channels, [42] and references therein, that have been investigated recently do not incorporate cross-talk effects. Differential transmission schemes with special demodulation devices are proposed in [19]. The focus here is on understanding and compensating for cross-talk effects seen in a single transmission slot. Differential schemes do provide an alternative for compensating cross-talk by
processing adjacent time domain transmissions. This approach notwithstanding, the choice of appropriate amplitude and phase for transmitting signals in each bit slot, using a simple electronic detector whenever possible, is investigated.

Cross-talk effects in fibers can be modeled as linear or nonlinear. As mentioned earlier, cross-channel linear effects arise mainly from nonideal devices other than the fiber. The fiber channel itself causes nonlinear cross-talk which is dominant over long distances. In the case of linear interference, interference from a particular channel is present when the transmitted bit for that channel is nonzero. As such, signal design schemes for such interference are independent of the transmitted bits and form the focus of Chapter 5. Signal design approaches are sensitive to the quality of channel estimates available at the transmitter. Channel estimation and the effect of estimation errors on signal designs for the linear interference case are also studied in Chapter 5. For nonlinear interference, the interference from a particular channel depends on the bits transmitted over other channels. As such, a different signal design approach is required for such systems. An optimization problem is devised for the general signal design problem and an iterative algorithm is proposed to solve it in Chapter 6.

The author would like to note that in this research, a single shot approach is employed to characterize the frequency domain effects. Differential detection schemes have been heavily investigated recently for optical communications. This motivates the research for characterizing both time-domain and frequency-domain cross-talk effects jointly and developing the signal processing schemes. The detector design mentioned and the procedure for obtaining signal designs for nonlinear cross-talk effects can be easily extended when such characterizations are available.
1.4 Summary

Recent developments in fiber-optic systems have focused on overcoming cross-talk effects in long haul all-optical systems employing closely packed WDM systems. Cross-talk may appear primarily from the nonlinear response of the fiber channel and from nonideal optical devices. This dissertation contributes to signal processing approaches to understand and mitigate cross-talk as follows:

1. In Chapter 2, the cross-talk effects arising from nonlinear fiber behavior change with distance, power level and channel spacing. These effects are studied for multispan systems with different receiver structures, one practical and other idealistic.

2. Modeling the nonlinear effects for multispan systems is studied in Chapter 3, where the tradeoff between model complexity and accuracy with changing distances, power levels and channel spacings is provided.

3. Low-complexity MUD schemes to mitigate the harmful effects of cross-talk via signal processing at the receiver are investigated in Chapter 4. The conditions for existence of reasonable detectors are investigated. A novel regression-based design approach is proposed and shown to be highly effective for the systems considered.

4. In Chapter 5, a signal design scheme that exploits channel information at the transmitter is considered for channels where cross-talk can be modeled via first order linear interference models. Channel estimation, and the effect of estimation errors on the signal design schemes are also investigated.

5. In Chapter 6, signal design algorithms for channels where cross-talk is modeled as being dominated by nonlinear interference are proposed and analyzed.
Chapter 2

Effects of Nonlinear Cross-talk

2.1 Introduction

The effect of nonlinearity increases with distances, transmit power, and decreased channel spacings. For multichannel systems employing intensity modulation and direct detection (IM-DD), when the frequency spacing is reduced from 50 GHz to 25 GHz, nonlinearity induced spectral distortion effects can significantly reduce eye-opening (EO) seen at the receiver. The EO is directly proportional to the effective $Q$-factor seen at the receiver, which in turn implies the bit-error-rate (BER). Reduction in EOs implies a reduction in the $Q$-factor, an increase in the BER, and hence a reduction the communication rate. Increasing transmit power levels is one solution that can help increase the rate for a given noise level by increasing the EO. However, the EOs may not scale linearly due to nonlinearity and this can reduce the energy efficiency of the transmission. The effects of nonlinearity on the EOs realized in multispans operating at varying power levels for 25 GHz and 50 GHz are studied via fiber simulations in this chapter.

Channel linearity preserves signal-space structures and enables the use of single channel/user matched filters to obtain sufficient statistics in the electrical demodu-
Chapter 2. Effects of Nonlinear Cross-talk

Fiber nonlinearity causes signal space expansion making it difficult to obtain sufficient statistics, especially for multichannel systems with square-law receivers. As such, the optimal demultiplexing and demodulation processes at the receiver are not clear. The choice of filters for this process can also impact the EO. Two filtering approaches are employed here, one motivated by the ideal matched-filter and the other by the practical approach of using a standard optical demultiplexer and a much narrower electrical filter.

The effects of different filters in optical receivers have been studied [43], [44] and references therein. These studies however, are focussed on the effects of filter dispersion and completely ignore fiber nonlinear effects. An experimental study of the simultaneous effects of fiber nonlinearity and different receiver filters is carried out in [45]. This research is limited by the fact that while filters of width 25 GHz are employed, the interfering channels considered are spaced 50 GHz apart. Additionally, only three channel systems are considered with a total of 5 spans. Continuous-time analytic models that provide insight into the nature of the nonlinearity are studied in [15], [36], [46], and [47], and compared with simulation results. These investigations do not incorporate the demodulation filters and their effects on EOs. Additionally, channel spacings in excess of 100 GHz are considered in [36], [48], and 50 GHz are considered in [47]. EOs for channel spacings in excess of 100 GHz are studied for a fixed power level of 0.6 mW/channel in [49] via simulations. Our research focuses on spacings of 25 GHz and 50 GHz, for a wider range of input power levels. A simulation is employed in this chapter to solve the nonlinear partial differential equation describing the channel behavior.

The rest of the chapter is organized as follows: The two fiber optic system set-ups are given in Section 2.2. The nonlinear behavior is described in terms of the nonlinear Schrödinger wave equation (NLSE) in Section 2.3 where the different effects
Chapter 2. Effects of Nonlinear Cross-talk

are described. The simulation algorithm employed to solve the NLSE is presented in Section 2.4. The simulation setup and results are presented in Section 2.5 and a conclusion is provided in Section 2.6.

2.2 System Model

The two systems considered in this chapter are shown in Figs. 2.1 and 2.2. The difference between the two systems is in the receiver structures; the rest of the system is identical. On-off keying (OOK) is employed at the transmitter where the transmitted signal is either a one or a zero for each channel. A channel is said to be on when a one bit is transmitted and off otherwise. For a WDM system with $K$ channels, the complex envelope of the optical field of the transmitted signal can be written in the frequency domain as

$$S^{(1)}(\omega, 0) = \sum_{k=1}^{K} A_k b_k P(\omega - k\Delta\omega),$$

(2.1)

where $|A_k|^2$ and $b_k \in \{0, 1\}$ denote the transmit power and the transmitted bit for the $k^{th}$ channel, respectively. $P(\omega)$, denotes the Fourier transform of the unit energy base band pulse employed for all channels. The spacing between channel center frequencies is given by $\Delta\omega$. The notation $S^{(n)}(\omega, z)$ is employed to denote the optical field of the signal propagating in the $n^{th}$ span at a distance of $z$ from the origin of the span. The signal $S^{(1)}(\omega, 0)$ is the signal launched into the first fiber span by the WDM MUX shown in Fig. 2.1 and Fig. 2.2.

The signal passes through the different spans of single mode fiber, the dispersion compensator and optical amplifier at the output of each span, and is received as the overall channel output at the WDM DEMUX shown in the two figures. The signal
Figure 2.1: Multispan fiber optic system with demodulation via optical filters that are matched to single channel systems.

Figure 2.2: Multispan fiber optic system with demodulation via optical demultiplexing and electrical filtering at the receiver.

at the output of the channel is demodulated via filtering and sampled by the receiver to generate decision statistics, \( \{y_1, \ldots, y_k\} \), which are used in the detection process.

As mentioned in the introduction, two demodulation approaches are considered here. The first approach uses an optical filter followed by a photodetector and a
sampler as shown in Fig. 2.1. The demodulation takes place by optically filtering the channel output signal with filters that are matched to each signal received as if in a single channel system. For single channel systems with ideal (noiseless) photodetection, if the optical filter is matched to the signal received when the channel is on, sufficient statistics are obtained. While obtaining the filter response analytically is not feasible, for simulation purposes, keeping a single channel on in a noiseless simulation gives the required filter response for that channel. When multiple channels are transmitted simultaneously, nonlinear channel interactions make obtaining sufficient statistics both analytically and practically difficult. Nevertheless, when nonlinear effects are less dominant, the approach of using a filter matched to the single channel system can result in near optimal performance. For the rest of this chapter, the system is referred to as having matched filters.

The above approach is impractical given the difficulty in implementing precise optical filters. However, it closely approximates the ideal theoretical optimal demodulation process. A more practical approach is shown in Fig. 2.2. Here, simple optical bandpass filters are implemented in the WDM DEMUX block, followed by a bank of photodetectors, and, unlike the first approach, a bank of narrow-band electrical filters, and then the sampling devices to generate decision statistics. The statistics thus generated may not be sufficient, but nevertheless carries reliable information for the detection process.

The sampled statistic is passed through a threshold detector, one for each channel to estimate the transmitted bit. Since a simple threshold detector is typically employed, the accuracy of the detector depends on the EO seen by the detector. To obtain the EO, we need to characterize the input-output relationship for the overall system.

\footnote{In practice, an electrical filter would always be necessary to limit thermal and shot noise introduced by the receiver. In this dissertation, receiver noise is ignored because it is assumed to be orders of magnitudes smaller than amplified spontaneous emission noise.}
system, which is discussed next.

## 2.3 Nonlinear Schrödinger Wave Equation and Channel Behavior

Optical pulses of widths greater than 1 ps are found to follow the simplified NLSE given below in (2.2) [50]. Ignoring the span index, the complex envelope of a signal, $s(t, z)$, propagating through a fiber at time $t$ and position $z$ can be described by

$$\frac{\partial s}{\partial z} = -\frac{\alpha}{2} s + j \frac{\beta_2}{2} \frac{\partial^2 s}{\partial t^2} - j \gamma |s|^2 s,$$  

(2.2)

where $\alpha$ represents the attenuation coefficient characterizing the fiber loss, $\beta_2$ the second order group velocity dispersion parameter that determines the optical pulse broadening, while $\gamma$ denotes the nonlinearity coefficient. In this equation, higher order terms starting with $\beta_3$ for approximating the linear propagation constant are assumed to be zero. In understanding the behavior of the channel it is useful to view the evolution equation in the frequency domain as

$$\frac{\partial S(\omega, z)}{\partial z} = -\frac{\alpha}{2} S(\omega, z) - \frac{j \beta_2 \omega^2}{2} S(\omega, z)$$

$$- j \frac{\gamma}{4\pi^2} \int \int S(\omega_1, z) s^*(\omega_2, z) S(\omega - \omega_1 + \omega_2, z) d\omega_1 d\omega_2,$$  

(2.3)

where $S(\omega, z) = F\{s(t, z)\}$, $F\{\cdot\}$ being the Fourier transform operator applied to $t$.

For shorter distances up to a few kilometers, it can be safely assumed that $\gamma = 0$, i.e., the channel response is perfectly linear. In this scenario, (2.3) implies that the signal for each channel undergoes attenuation and phase distortion as indicated by the linear terms. The use of perfect dispersion compensation and optical amplifiers
can help correct for these effects [15]. The end-to-end channel in this scenario behaves like conventional linear channels that are well studied [51]. In such systems, cross-talk results from leakages at the different devices. In this chapter we assume that the optical repeater devices (the dispersion compensator and the optical amplifier) are ideal. Cross-talk occurs primarily from leakages in the MUX and DEMUX and from frequency domain overlap among channels at the transmission point. This form of cross-talk is identical to adjacent channel interference (ACI) that is experienced by linear wired and wireless channels.

When the receiver is located after few tens of kilometers, or few hundred kilometers as is the case in multispans systems, the nonlinearity can no-longer be ignored and even the noiseless channel can no longer be considered trivial. As indicated by the third term on the RHS of (2.3), the signals transmitted across different channels interact with each other. This interaction is in the form of a frequency domain convolution resulting in signal spreading, which causes inter-channel cross-talk. Depending on the transmitted bits, three forms of inter-channel cross-talk are observed [46]. First, every nonzero transmitted bit convolves with itself and spread in the frequency domain causing linear ACI. Additionally, this causes higher order distortion of the signal itself referred to as self-phase modulation (SPM). Secondly, any two channels with nonzero transmitted bits interact with each other and cause cross-phase modulation (XPM). Finally, two or three channels with nonzero transmitted bit interact and generate a ghost pulse in another channel causing four-wave mixing (FWM). Note that each of these forms of cross-talk builds over the existing ACI arising from device leakages as discussed earlier.

The different forms of cross-talk can be constructive or destructive causing energy expansions in certain channels and energy reductions in others. This can reduce the EOs seen by the different channels. Understanding the effects on EOs requires solving
(2.2) or (2.3) for the input signal given in (2.1). The simulation approach for doing this is described next.

### 2.4 Split-step Fourier Simulation

There are three alternatives to study the effects of nonlinearity: via analysis, experimentation, or simulation. A closed form analytic solution for the NLSE exists only in special cases inapplicable here. The Volterra series transfer function (VSTF) method to solve the NLSE analytically is employed in [46], [36], [47] and [48]. The solution obtained consists of an infinite sum of multiple integrals of increasing orders. Combining this infinite sum effect across multiple spans and with the demodulation process is analytically intractable. Additionally, since the sum is infinite, it needs to be truncated, which leads to inaccuracies. The experimental approach, while accurate, is expensive, inflexible, and requires resources that are not easily accessible. A simulation approach based on the split-step Fourier (SSF) algorithm is found to be the best alternative to actual experimentation [52], [53] and is used widely to study various fiber effects. A short review of the algorithm is provided here.

The symmetrized split-step Fourier method, [52], is described here and used in the simulation results presented. Defining a linear operator \( \hat{d} = -\frac{\alpha}{2} + j \frac{\beta_2}{2} \frac{\partial^2}{\partial t^2} \), and a nonlinear operator \( \hat{n} = -j \gamma |s|^2 \), (2.2) can be re-written as

\[
\frac{\partial s}{\partial z} = \left( \hat{d} + \hat{n} \right) s, \tag{2.4}
\]

The solution to this equation at position \( z + h \) along the fiber is related to that at
Chapter 2. Effects of Nonlinear Cross-talk

position \( z \) by \cite{53}

\[ s(z + h, t) \approx \exp(h(\hat{d} + \hat{n})) s(z, t), \tag{2.5} \]

where \( \hat{d} \) and \( \hat{n} \) are still time domain operators. The solution is computed numerically by first applying the nonlinear step in the time domain, and then applying the linear step in the frequency domain to give,

\[ s(z + h, t) \approx \mathcal{F}^{-1}\left\{ \exp(h\hat{D}(\omega))\mathcal{F}\{\exp(h\hat{n})s(z, t)\}\right\}, \tag{2.6} \]

where \( \hat{D} \) denotes the Fourier domain operation for \( \hat{d} \). During a simulation, the step-size \( h \) is chosen and the above equation is applied successively to compute the signal at various points along the fiber.

In the symmetric SSF method the step-size is split into two halves. The linear operator is applied in the frequency domain for the first half, which is followed by a time-domain nonlinear step and then by a frequency-domain linear step for the remaining half, to give

\[ s(z + h, t) = \mathcal{F}^{-1}\left\{ \exp(h/2\hat{D}(\omega))\mathcal{F}\{\exp(h\hat{n})\mathcal{F}^{-1}(\exp(h/2\hat{D}(\omega))S(z, \omega))\}\right\}. \tag{2.7} \]

This approach, which is superior to the approach indicated in (2.6), can be made more accurate by using an iterative approach to compute the nonlinear operation.

The error from this computation arises primarily from the second order approximation of the operators \( \hat{d} \) and \( \hat{n} \) and is proportional to the second order in the step size. An adaptive algorithm to choose \( h \) helps achieve a tradeoff between simulation time and accuracy. Choosing the step-size such that the corresponding phase
change is limited can help achieve higher accuracy. The nonlinear phase rotation method chooses the step-size adaptively for a given maximum phase change $\phi_{\text{max}}$ as $h = \frac{\phi_{\text{max}}}{\gamma_{\text{max}} |s(z,t)|^2}$. For a few hundred kilometers, using $\phi_{\text{max}} = 0.005$ radians is found to be sufficient [47]. Simulation results described below were obtained with $\phi_{\text{max}}$ set to 0.001 radians for higher accuracy. A number of studies have been conducted on the accuracy vs. computation tradeoff for the SSF algorithm for varying fiber parameters. For the system parameters considered here, and the nonlinear phase rotation method used, the total number of FFTs computed for the simulation is on the order of $10^3$ to $10^4$. This corresponds to a global relative error on the order of $10^{-5}$ to $10^{-4}$ for similar fibers [54]. The relative error in [54] is computed with respect to the results obtained by using fixed but extremely small step-sizes and the symmetric SSF method [52].

### 2.5 Simulation Results

A system with $K = 8$ channels is simulated for frequency spacings, $\Delta \omega = 2\pi \times (25 \text{ GHz})$ and $2\pi \times (50 \text{ GHz})$. The fiber attenuation parameter is chosen identical to that of one of the latest Corning ultra-low loss fibers [55], while the second order dispersion coefficient and the nonlinearity parameters were chosen identical to those seen in other simulation results available in literature. The simulation parameters are given in Table 2.1. Gaussian pulse shapes are employed for the transmitted signals. A noiseless system is simulated since the goal is to study the effect of nonlinearity alone on the EOs. The focus here is to analyze the impact of inter-channel, frequency domain cross-talk during a single transmission. To avoid ISI type effects, single shot bit-vector transmissions are simulated. With $K$ users, $2^K$ separate simulations are performed for a given set of parameters, distance traversed, power level and band-
width spacing. The transmit power level used, although high, is in the range found in [15], [49] and references therein. Importantly, it facilitates the study of nonlinearity on increasing power levels. After channel simulations, the signal obtained at the fiber output is passed through the two different receiver demodulators considered here. The demodulated signals are sampled at time instant $T$ corresponding to the bit-period, which maximizes the performance of the threshold detector.

Table 2.1: System Parameters for Simulations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channels</td>
<td>8</td>
</tr>
<tr>
<td>Channel Spacing</td>
<td>50 GHz &amp; 25 GHz</td>
</tr>
<tr>
<td>Data Rate</td>
<td>10 Gbps</td>
</tr>
<tr>
<td>Attenuation Coefficient</td>
<td>$\alpha = 0.18 \text{ dBkm}^{-1}$</td>
</tr>
<tr>
<td>$2^{nd}$ Order Dispersion Coefficient</td>
<td>$\beta_2 = -2.6 \text{ ps}^2 \text{ km}^{-1}$</td>
</tr>
<tr>
<td>Nonlinearity Coefficient</td>
<td>$\gamma = 2.2 \text{ W}^{-1} \text{ km}^{-1}$</td>
</tr>
<tr>
<td>Span Length</td>
<td>$L = 125 \text{ km}$</td>
</tr>
<tr>
<td>Maximum No. of Spans</td>
<td>16</td>
</tr>
<tr>
<td>Average Transmit Powers</td>
<td>0.25 to 1.25 mW/channel or 3 dBm to 10 dBm total</td>
</tr>
<tr>
<td>Maximum Simulation Phase Change</td>
<td>$\phi_{max} = 0.001 \text{ radians}$</td>
</tr>
</tbody>
</table>

Fig. 2.3 shows the sample values that are obtained for channel four out of eight (near the center of the band) when 0.25 mW/channel is employed with both 25 and 50 GHz systems. The first 128 samples correspond to the received signal when the channel is set to OFF and the next 128 correspond to the received signal when
the channel is set to be ON. With a single span, nonlinear effects are minimal and the only deterioration is caused by channel overlap among signals at the transmitter. With 16 spans, the nonlinear effects dominate and the samples corresponding to a ON transmission reduce as energy leaks into other channels; at the same time samples corresponding to the OFF transmission increase as energy is pumped from other channels that are ON. The nonlinear effects are severe for the case of 25 GHz channel spacing as compared to the 50 GHz case.

The important effect to be captured is the distance between the minimum signal value for the ON channel and the maximum value for the OFF channel condition. This measure is the EO. The EO is set to zero when the value received for a one bit transmission is below that for some zero bit transmission over the channel. The EO value thus obtained is normalized with the value that would be obtained for single channel transmission with $\gamma = 0$ to compute the normalized EO. This normalization ensures that the gains or losses resulting from the presence of nonlinearity induced channel cross-talk are captured and multiple power levels can be compared on the same plot. An EO of zero implies that error-free communication is not possible even in the absence of noise, at least when the traditional threshold detector is employed.

The EOs normalized with the peak transmit amplitude for 50 GHz and 25 GHz systems employing a matched filter demodulator are shown in Fig. 2.4. For ideal systems (no cross-talk), the normalized EO would be unity. The plot for $\gamma = 0$ corresponds to the linear channel case and the drop in normalized EO from unity is due to ACI effects alone.
Figure 2.3: Ordered scaled received samples for the center channel in 25 and 50 GHz systems with one and 16 spans, and 0.25 mW/channel. The samples are scaled to set the largest value to one.
For the nonlinear channel case where $\gamma \neq 0$, the normalized EO reduces with increasing distances and reaches zero for higher power levels. Thus, nonlinearity has an impact on the maximum distance over which error-free transmission is possible using multiple spans. As expected, the tighter spacing for the 25 GHz systems causes a faster drop in EOs compared to 50 GHz systems for the linear channel case.

For single span systems, the EOs reduce significantly as transmission power increases for both systems. For 25 GHz systems, the reduction in EOs is much higher with increasing power. Therefore, even for single span systems, increasing the power may not necessarily reduce the error-rate, thereby limiting the maximum transmission rate. The figure shows that for power levels above 0.75 mW/channel the EOs for 25 GHz systems go to zero for a number of spans that is less than half the corresponding number of spans for which the EOs for 50 GHz systems go to zero. Thus, due to the nonlinearity, reducing channel spacings by half can reduce the maximum distance over which communication with threshold detection is feasible by much more than half.

The normalized EOs for systems employing a combination of electrical and optical filtering to demodulate the signal are shown in Fig. 2.5. Comparing Fig. 2.4(a) with Fig. 2.5(a) shows that the receiver structure with two filters causes a slight reduction in the EO. Given that the second receiver realization is definitely sub-optimal, this is not unexpected. For transmission powers between 0.75 to 1.25 mW/channel, the EOs close one span earlier. The loss is not insignificant, but, given that this structure is realizable in practice, it indicates that the cost of sub-optimal demodulation may not be high. Comparing Fig. 2.4(b) and Fig. 2.5(b) indicates that the sub-optimal demodulation may in fact be better. This can be explained by the fact that the sharp electrical filtering reduces the ACI as well as other spurious waves created by cross-talk much better than the optical matched filter receiver, (which is optimal
Figure 2.4: Normalized EOs for systems with up to 16 spans at multiple transmit power levels. The plots correspond to 50 GHz (above), and 25 GHz (below), systems with optical matched-filtering. The case $\gamma = 0$ corresponds to the linear channel and captures the ACI effects.

only for single channel systems to begin with), when the channels are tightly packed. Comparing Fig. 2.5(a) and Fig. 2.5(b) again shows that reducing the channel spacing to half reduces the distance over which EOs do not close by more than half.
Chapter 2. Effects of Nonlinear Cross-talk

Figure 2.5: Normalized EOs for systems with up to 16 spans at multiple transmit power levels. The plots correspond to 50 GHz (a), and 25 GHz (b), systems with optical bandpass filtering and lowpass electrical filtering. The case $\gamma = 0$ corresponds to the linear channel and captures the ACI effects.

Fig. 2.6 shows the spans at which the EOs go to zero for the different systems considered at the various power levels. The figure indicates that at power levels below 0.65 mW/channel for 50 GHz systems, no eye closure is seen for up to 16 spans while...
Chapter 2. Effects of Nonlinear Cross-talk

Figure 2.6: Number of spans for which the EOs close for a given power level for systems with channel spacings of 50 GHz and 25 GHz. Systems with matched optical filter receiver as shown in Fig. 2.1 and with optical bandpass filter and electrical filter for demodulation as shown in Fig. 2.2 are considered.

For 25 GHz systems, a maximum for 12 spans can be covered at the lowest power level or 0.25 mW/channel considered. Using the noise power calculation for systems with optical amplification as given in (6.5.18) in [50], the optical SNR corresponding to 0.25 mW/channel is found to be in the range of 8 – 15 dB for various spans. This value is relatively low compared to that specified in typical optical channels. Clearly then, if higher OSNR is needed, operating in this region is not feasible. However, for 25 GHz systems with increasing power, the EO goes to zero more rapidly making it difficult to traverse longer distances.
2.6 Conclusion

The results obtained in this chapter can be summarized as follows:

1. Given a power level, reducing the channel spacing from 50 GHz to 25 GHz drastically reduces the normalized EOs. While the reduction is a function of the distance traversed, it is almost always greater than half. This result implies that electronic regeneration would be required more than twice as often, unless the problem is addressed using signal processing techniques.

2. For a given channel spacing, increasing the transmit power level reduces the relative EO. Note that while the actual EO may be greater, the normalized EO is smaller. Nevertheless, the actual EO goes to zero much quicker with increasing power levels.

3. Given a transmission distance, there exists an upper bound on the power level than can be employed for a given channel spacing. Simply increasing transmission power does not enable long distance communication.

It should be noted here that the EO drops do not involve the ISI effects and real systems may have lower EOs. Clearly, nonlinear effects drastically reduce the error-free communication distance with reduced channel spacing and limit the distance that can be traversed by simply increasing power levels. While a number of receiver parameters can be changed, such as the demodulator structure, the nonlinear effects does not change significantly as shown by the smaller difference in the EO experienced by the two systems considered.

There are three alternatives to overcome the relevant bottlenecks. One is to replace the existing fiber which is beyond the scope of this investigation. The second is to design the input signals to minimize the cross-talk. This is the focus of Chapter 6.
Chapter 2. Effects of Nonlinear Cross-talk

Note that this study was performed assuming threshold detectors. The third approach is to overcome the limitation of this single threshold detector approach by performing joint detection. Different detectors are researched in Chapter 4. The single threshold detector considered here corresponds to the single user detector (SUD) in Chapter 4 and this study demonstrates the limitations of the SUD approach. The following chapter discusses obtaining discrete-time models that enable the signal processing approaches studied in the later chapters.
Chapter 3

Modeling Fiber Cross-talk Effects in
Discrete-Time

3.1 Introduction

In the previous chapter, the harmful effects of nonlinearity-induced inter-channel cross-talk are demonstrated. Designing receiver and transmitter signal processing schemes to compensate for the nonlinear effects can help mitigate some of these effects. This requires a model that characterizes the relation between the discrete input symbols at the transmitter, and the decision statistics at the demodulator outputs at the receiver. Channel nonlinearity makes it difficult to obtain discrete-time models with sufficient statistics as is typical for linear communication channels. This problem is investigated here.

Models that capture both intra- and inter-channel nonlinear effects of not only the fiber medium, but also the nonlinearities introduced in the electro-optical and opto-electrical conversions taking place at various stages in the system are needed. However, the focus here is on modeling inter-channel effects arising from the channel induced nonlinearity. The goal is to model the nonlinear relationship between the
discrete channel input and the discrete sampled output, with a view of characterizing the nonlinear cross-talk. As such, single shot transmissions are modeled. Additionally, a memoryless model is employed. Electrical filters induce memory and hence are ignored here. Only optical demultiplexing is used to obtain the discrete-time received signal. The demodulator at the receiver considered here is given in Fig. 3.1. Similar to Chapter 2, the optical filters employed for optical demultiplexing can be either a single channel matched filters or practical optical bandpass filters. The optical output of these filters is sampled via a photodetector, and the relationship between this sampled signal and the transmitted signal is modeled. It is assumed that the transmitter and receiver components apart from the optical filters at the optical demultiplexer are ideal and linear in operation (except for the ideal square-law photodetector).

Figure 3.1: Receiver for multispam fiber optic system with demodulation via optical demultiplexing and photodetection. Optical filters that are matched to single channel systems or practical optical bandpass filters are employed in the demultiplexing process.

The eventual goal in modeling nonlinear effects in fiber channels is to be able to present viable models, (from a communication theoretic perspective). This requires
that:

1. The model captures the effects of nonlinearity with sufficiently high accuracy.

2. The model is not highly complex as that would increase the eventual channel estimation complexity for implementation.

The continuous-time input output model for fiber channels has been heavily investigated in [15], [36], [46] and [47]. The Volterra series transfer function (VSTF) method can be employed to solve the NLSE [46]. For single mode fibers, it turns out that the Volterra series terms that need to be considered are restricted to terms with odd orders. Based on the continuous-time characterization of the channel, a discrete-time memoryless model with up to third order nonlinearity is proposed for single span systems in [15] and [47]. While multispan systems are briefly addressed in [15], [46] and [47], details for multispan systems are not provided. Additionally, the accuracy and complexity of the different models has not been investigated.

Using the VSTF method it is shown here that for multispan systems, the continuous-time model contains all nonlinear terms that are higher than third order terms. Similar to the approach for single span systems, using a third order approximation of the continuous-time system, the discretization process is investigated. First a system with unmodulated continuous wave (CW) inputs is considered. Using the insight provided by the corresponding model, four models with varying levels of complexity, are proposed for multispan systems with modulated finite duration inputs. Simulations data obtained via the simulation process described in the previous chapter are then employed to validate the accuracy of the models.

The rest of the chapter is organized as follows: The analytic VSTF solution to the NLSE is reviewed in Section 3.2. The section also provides details for applying the VSTF to multispan systems and obtains results identical to that in [15]. Novel
discrete-time models that are employed in the rest of this dissertation are introduced in Section 3.3. The accuracy of models is studied in Section 3.4 and a summary of the chapter is provided in the final section.

3.2 Applying the Volterra Series Transfer Function Approximation to Multispan Systems

A closed-form analytical solution to the general NLS equation given by (2.2) and (2.3) is not known. The Volterra series method approach is employed in [46] to obtain the VSTF solution to the NLSE as an infinite series with no even order terms present. For the baseband equivalent signal $S(\omega, z)$ propagating over the fiber, this solution is given as

$$S(\omega, z) = H_1(\omega, z)S(\omega, 0) + \int \int H_3(\omega, \omega_1, \omega_2, z)S(\omega_1, 0)S^*(\omega_2, 0)S(\omega - \omega_1 + \omega_2, 0)d\omega_1d\omega_2 + ... ,$$

(3.1)

where, $H_1(\omega, z)$ and $H_3(\omega, \omega_1, \omega_2, z)$ represent the first and third order Volterra series kernel coefficients. Ignoring third order dispersion coefficients and Raman effects, the filter kernels are given as

$$H_1(\omega, z) = \exp \left( -\frac{\alpha z^2}{2} - j\frac{\beta \omega^2 z}{2} \right)$$

$$H_3(\omega, \omega_1, \omega_2, z) = -j\frac{\gamma}{4\pi^2} H_1(\omega, z)$$

$$\frac{1 - \exp \left( -\alpha z - j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)z \right)}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)}.$$  

(3.2)

In [47] it is shown that for single span systems, truncating the solution to third order results in a highly accurate continuous-time model. The accuracy measure used there...
is the normalized square deviation which measures the average energy distortion over the entire transmission duration. In this section, the third order approximation for single span systems is applied to multispans systems with dispersion compensation and optical amplification between spans.

Similar to (2.1), the complex envelope of the optical field of transmitted signal at the input of the fiber can be written in the frequency domain as

\[ S^{(1)}(\omega, 0) = \sum_{k=1}^{K} A_k b_k P(\omega - k\Delta\omega). \] (3.3)

Note that in this equation, while \( |A_k|^2 \) denotes the transmit power, the term \( A_k \) represented the transmit amplitude as well as the phase of the the laser for the \( k^{th} \) channel.

This is the field at the output of the WDM MUX block in Fig. 3.1. For the multispans systems considered here, it is assumed that optical amplifiers and dispersion compensators that are located at the end of each span are perfect. That is, the amplifier compensates perfectly for the energy loss and the dispersion compensator perfectly compensates the dispersion due to the \( \beta_2 \) parameter. Let \( S^{(n)}(\omega, z) \) represent the envelope of the signal propagating along the \( n^{th} \) span. Assuming fiber spans of length \( L \), we denote by \( S^{(n)}(\omega, L) \) the output of the \( n^{th} \) span. After the dispersion compensation and amplification, this signal acts as input to the next span, given as \( S^{(n+1)}(\omega, 0) \). Perfect optical amplification and dispersion compensation requires inverting the first order filter operation denoted by \( H_1 \). Assuming this is the case,
the input field for the \((n+1)\)th span can be obtained using (3.1) and (3.2)

\[
S^{(n+1)}(\omega, 0) = H^{-1}_1(\omega, L) S^{(n)}(\omega, L)
\]

\[
= S^{(n)}(\omega, 0) + \int \int -j \frac{\gamma}{4\pi^2} \frac{1 - \exp \left( -\alpha L - j\beta_2(\omega_1 - \omega)(\omega_2 - \omega) \right)}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)} S^{(n)}(\omega_1, 0) S^{(n)*}(\omega_2, 0) S^{(n)}(\omega - \omega_1 + \omega_2, 0) d\omega_1 d\omega_2
\]

(3.4)

with third order approximation. Here \(L\) is the length of a single span.

For a system with \(N\) spans, the channel output that acts as an input to the demodulator is \(S^{(N+1)}(\omega, 0)\). The term can be obtained starting with \(S^{(1)}\) and successively applying (3.4). Doing this implies that even after truncating the VSTF to third order, higher order nonlinearities are present. To see this, consider a two span system. Then the output of the second span after dispersion compensation and optical amplification can be written as the input for a hypothetical third span as

\[
S^{(3)}(\omega, 0) = S^{(2)}(\omega, 0) + \int \int -j \frac{\gamma}{4\pi^2} \frac{1 - \exp \left( -\alpha L - j\beta_2(\omega_1 - \omega)(\omega_2 - \omega) \right)}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)} S^{(2)}(\omega_1, 0) S^{(2)*}(\omega_2, 0) S^{(2)}(\omega - \omega_1 + \omega_2, 0) d\omega_1 d\omega_2
\]

(3.5)

which corresponds to

\[
S^{(3)}(\omega, 0) = S^{(1)}(\omega, 0) + \int \int -j \frac{\gamma}{4\pi^2} \frac{1 - \exp \left( -\alpha L - j\beta_2(\omega_1 - \omega)(\omega_2 - \omega) \right)}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)} S^{(1)}(\omega_1, 0) S^{(1)*}(\omega_2, 0) S^{(1)}(\omega - \omega_1 + \omega_2, 0) d\omega_1 d\omega_2
\]

\[
+ \int \int -j \frac{\gamma}{4\pi^2} \frac{1 - \exp \left( -\alpha L - j\beta_2(\omega_1 - \omega)(\omega_2 - \omega) \right)}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)} S^{(2)}(\omega_1, 0) S^{(2)*}(\omega_2, 0) S^{(2)}(\omega - \omega_1 + \omega_2, 0) d\omega_1 d\omega_2.
\]

(3.6)

The first two terms on the R.H.S. of (3.6) are obtained by using \(S^{(2)}\) in terms of \(S^{(1)}\). A similar substitution for the third term in (3.6) shows that terms higher than the
third order are present in $S^{(3)}$. Fifth, seventh, and ninth order terms are present when only the third order nonlinearity is modeled. However, the nonlinearity coefficient $\gamma$ for these terms also has higher powers which increase with the order of nonlinearity. Since $\gamma \ll 1$, these higher order terms are neglected to obtain an analytically tractable model given as

$$S^{(3)} \approx S^{(1)}(\omega, 0) + 2 \int \int -\frac{j}{4\pi^2} \frac{1 - \exp(-\alpha L - j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)L)}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)} \cdot S^{(1)}(\omega_1, 0)S^{(1)*}(\omega_2, 0)S^{(1)}(\omega - \omega_1 + \omega_2, 0)d\omega_1d\omega_2.$$  \hspace{1cm} (3.7)

Recursively applying the above procedure to a system with $N$ spans implies that the output of the fiber at the end of the last span can be given as

$$S^{(N+1)}(\omega, 0) \approx S^{(1)}(\omega, 0) + N \int \int -\frac{j}{4\pi^2} \frac{1 - \exp(-\alpha L - j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)L)}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)} \cdot S^{(1)}(\omega_1, 0)S^{(1)*}(\omega_2, 0)S^{(1)}(\omega - \omega_1 + \omega_2, 0)d\omega_1d\omega_2, \hspace{1cm} \forall N \geq 2.$$  \hspace{1cm} (3.8)

Apart from the factor $N$, the continuous-time input-output model indicated by (3.8) is identical to that for single span WDM systems and is identical to the transfer function implied by setting $D = 0$ in (2.25) of [15]. Given this form for the channel output, the next step is to obtain a discrete-time model for the sampled statistics $y_k$, $\forall 1 < k \leq K$.

### 3.3 Discrete-Time Models and System Setup

Discrete-time models that relate the sampled statistics $y_k$ to the transmitted bits $b_k$, $\forall k$ are obtained in this section. As shown in Fig. 3.1, the received signal after optical amplification and dispersion compensation is passed through a bank of optical filters or demultiplexers. For a $N$ span system, the effective received signal field
after optical demultiplexing, \( r_k(t) \), received for the \( k^{th} \) channel can be written in the frequency domain as

\[
R_k(\omega) = H_{o_k}(\omega)S^{(N+1)}(\omega, 0). \tag{3.9}
\]

Here, \( H_{o_k}(\omega) \) denotes the response of the optical demultiplexing filter.

For the system in Fig. 3.1, when the matched filter is employed, the filter for the \( k^{th} \) channel can be given as

\[
H_{o_k}(\omega) = P^*(\omega - k\Delta\omega) + N \int \int \left( -j \frac{\gamma}{4\pi^2} 1 - \exp \left( \frac{-\alpha L - j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)L}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)} \right) \right)^* \cdot P^*(\omega_1 - k\Delta\omega)P(\omega_2 - k\Delta\omega)P^*(\omega - \omega_1 + \omega_2 - k\Delta\omega)d\omega_1d\omega_2, \tag{3.10}
\]

denoting the single user channel case for the matched filter process. For the system in Fig. 3.1, if practical demultiplexing is employed, the \( k^{th} \) channel filter corresponds to the response for a generic analog bandpass filter. Nevertheless, using (3.8) and (3.3), the input and output can be expressed in the continuous-time domain as

\[
R_k(\omega) \approx H_{o_k}(\omega) \sum_{i=1}^{K} A_ib_iP(\omega - i\Delta\omega) + H_{o_k}(\omega)N \sum_{l=1}^{K} \sum_{m=1}^{K} \sum_{n=1}^{K} A_lA_m^*A_n b_l b_m b_n \cdot \int \int -j \frac{\gamma}{4\pi^2} 1 - \exp \left( \frac{-\alpha L - j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)L}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)} \right) \cdot P(\omega_1 - l\Delta\omega)P^*(\omega_2 - m\Delta\omega)P(\omega - \omega_1 + \omega_2 - n\Delta\omega)d\omega_1d\omega_2. \tag{3.10}
\]

The optical outputs are then converted to electrical signals via a bank of photodetectors and sampled to obtain the statistics \( y_k \) as shown in Fig. 3.1. Recall, that the signal after photodetection may be filtered via a bank of lowpass narrow
band electrical filters as shown in Fig. 2.2, but this process is ignored here since our goal is to characterize only channel nonlinearity effects. For different receiver demodulator filters and different input spectra, the third order response indicated by (3.10) results in different types of third order terms corresponding to the cross-talk forms of cross-phase modulation (XPM), four-wave mixing (FWM), and self-phase modulation (SPM) discussed in Chapter 2. In the following section, we consider an unmodulated CW input signal to characterize the terms corresponding to these effects.

3.3.1 Channel Model for Single Shot System with Unmodulated Continuous Wave Inputs

Expressing $r_k(t)$ in closed form for different types of optical filters and different input pulse shapes denoted by $P(\omega)$ leads to complexity. To further simplify the model, we assume that $P(\omega) = \delta(\omega)$. This assumption means that a single tone (CW) is transmitted for each nonzero bit for an infinite duration, which is reasonable since our goal is to simply characterize the inter-channel effects for single shot transmissions. Note that each channel in this case occupies infinitesimal bandwidth, and the linear cross-talk effects arising from adjacent channel interference (ACI), and the nonlinear overlap on the resulting cross-talk signal is not visible. (The following section includes
Chapter 3. Modeling Fiber Cross-talk Effects in Discrete-Time

The output is then given as

\[ R_k(\omega) \approx H_{ok}(\omega) \sum_{i=1}^{K} A_i b_i \delta(\omega - i\Delta\omega) \]

\[ + H_{ok}(\omega) N \int \int -j \frac{\gamma}{4\pi^2} \frac{1 - \exp(-\alpha L - j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)L)}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_2 - \omega)} \]

\[ \cdot \sum_{l=1}^{K} \sum_{m=1}^{K} \sum_{n=1}^{K} A_l A_m^* A_n b_l b_m b_n \]

\[ \cdot \delta(\omega_1 - l\Delta\omega) \delta(\omega_2 - m\Delta\omega) \delta(\omega - \omega_1 + \omega_2 - n\Delta\omega) d\omega_1 d\omega_2, \]

\[ = H_{ok}(\omega) \sum_{i=1}^{K} A_i b_i \delta(\omega - i\Delta\omega) - H_{ok}(\omega) N \sum_{l=1}^{K} \sum_{m=1}^{K} \sum_{n=1}^{K} A_l A_m^* A_n b_l b_m b_n \]  (3.11)

\[ j \frac{\gamma}{4\pi^2} \int \delta(\omega_1 - l\Delta\omega) \frac{1 - \exp(-\alpha L - j\beta_2(\omega_1 - \omega)(m\Delta\omega - \omega)L)}{\alpha + j\beta_2(\omega_1 - \omega)(m\Delta\omega - \omega)} \]

\[ \delta(\omega - \omega_1 + m\Delta\omega - n\Delta\omega) d\omega_1 \]

\[ = H_{ok}(\omega) \sum_{i=1}^{K} A_i b_i \delta(\omega - i\Delta\omega) - H_{ok}(\omega) N j \frac{\gamma}{4\pi^2} \sum_{l=1}^{K} \sum_{m=1}^{K} \sum_{n=1}^{K} A_l A_m^* A_n b_l b_m b_n \]

\[ 1 - \exp(-\alpha L - j\beta_2(l\Delta\omega - \omega)(m\Delta\omega - \omega)L)}{\alpha + j\beta_2(l\Delta\omega - \omega)(m\Delta\omega - \omega)} \delta(\omega - l\Delta\omega + m\Delta\omega - n\Delta\omega) \]

To generate the statistics \( y_k \), the signal \( r_k(t) \) is converted to an electrical signal and sampled directly at the appropriate instant as shown in Fig. 3.1. For the CW system in (3.11), the presence of the impulse functions implies that only the filter response value at the center frequencies of \( i\Delta\omega \) affects the sampled signal. The optical filters in Fig. 3.1, whether matched or bandpass, can be scaled to have unit response at these frequencies. The demodulated decision statistic \( y_k \) for the system can then be given as

\[ y_k = |r_k(T)|^2, \]  (3.12)
where

\[ r_k(T) = \mathcal{F}^{-1} R_k(\omega)|_{t=T} \]

\[ \approx A_k b_k \exp(j k \Delta \omega T) - j N \frac{\gamma}{4 \pi^2} \sum_{l=1}^{K} \sum_{m=1}^{K} \sum_{n=1}^{K} A_l A_m^* A_n b_l b_m b_n \exp(j k \Delta \omega T)(3.13) \]

\[ 1 - \exp\left( -\alpha L - j \beta_2 (l \Delta \omega - k \Delta \omega) (m \Delta \omega - k \Delta \omega) L \right) \frac{\delta_{k-l+m-n}}{\alpha + j \beta_2 (l \Delta \omega - k \Delta \omega) (m \Delta \omega - k \Delta \omega)} \]

and \( \delta_i \) represents the Kroneckor delta function [56]. Here \( T \) is the appropriate sampling time that satisfies \( \Delta \omega T = 2\pi n \) for an integer \( n \) (usually very large).

The second term on the R.H.S. above is a result of inter-channel and intra-channel cross-talk experienced by channel \( k \), when the bits transmitted by channels \( l, m \) and \( n \) are nonzero. The summation terms can be decomposed into three classes for which the Kroneckor delta is nonzero. These three types of terms in turn represent three forms of cross-talk as indicated below:

\[
\begin{align*}
\text{XPM} & \quad k = l, m = n, \text{ or } k = n, l = m, \\
\text{FWM} & \quad k = l - m + n, k \neq l, k \neq n, \quad (3.14) \\
\text{SPM} & \quad k = l = m = n.
\end{align*}
\]

Using the above decomposition of the triple summation term, the received signal at the sampling instant can then be rewritten as

\[
r_k(T) \approx b_k A_k + \sum_{m=1}^{K} \rho_{k,m}^{(SPM)} |b_k A_k|^2 + \sum_{m=1}^{K} \sum_{n=1}^{K} \rho_{k,m}^{(XPM)} A_k A_m |b_m A_m|^2 \\
+ \sum_{l=1, l \neq k}^{K} \sum_{n=1, n \neq k}^{K} \rho_{k,l,m=1,n-k,n}^{(FWM)} b_l b_m b_n A_l A_m A_n^* A_m A_n.
\]
where

\[
\rho_{k,m}^{(XPM)} = -\left\{ \frac{1 - \exp(-\alpha L)}{\alpha} + \frac{1 - \exp(-\alpha L - j\beta_2(m-k)^2\Delta\omega^2 L)}{\alpha + j\beta_2(m-k)^2\Delta\omega^2} \right\} \frac{jN\gamma}{4\pi^2}
\]

\[
\rho_{k,l,m,n}^{(FWM)} = -\frac{1 - \exp(-\alpha L - j\beta_2(l-k)(m-k)\Delta\omega^2 L)}{\alpha + j\beta_2(l-k)(m-k)\Delta\omega^2} \frac{jN\gamma}{4\pi^2}
\]

\[
\rho_k^{(SPM)} = -\left\{ \frac{1 - \exp(-\alpha L)}{\alpha} \right\} \frac{jN\gamma}{4\pi^2}.
\]  

(3.16)

Throughout the dissertation we refer to these coefficients as correlation coefficients.

Clearly, these coefficients are complex and depend on the channel spacings and the channel indices. By substituting the parameter values identical to those used for the simulation in Chapter 2, and varying the values for the channel indices, the values for the correlation coefficients are obtained. Table 3.1 shows the range of magnitudes that these values take as the channel indices are varied for an eight channel system with both 25 GHz spacings and 50 GHz spacings. The values indicate that XPM and FWM terms increase with reducing channel spacing as expected. When scaled suitably with the amplitudes, these values have ranges that have magnitude on the order of those obtained in [47] for a single span three user system.

Table 3.1: Range of magnitudes of correlation coefficients for the system with 16 spans and parameters given in Table 2.1

<table>
<thead>
<tr>
<th>Channel Spacing (GHz)</th>
<th>50</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range for $\rho_{k,m}^{XPM}$ (W$^{-1}$)</td>
<td>(20, 35)</td>
<td>(23, 39)</td>
</tr>
<tr>
<td>Range for $\rho_{k,l,m,n}^{FWM}$ (W$^{-1}$)</td>
<td>(2, 14)</td>
<td>(8, 20)</td>
</tr>
<tr>
<td>Value of $\rho_k^{SPM}$ (W$^{-1}$)</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

The values obtained here correspond to a CW system and ignore the effects of modulation bandwidth. The next section builds on the CW model and proposes a
more realistic model for the system.

3.3.2 Channel Models for Systems with Modulated Finite Duration Transmissions

For more realistic systems, the input pulses are of finite duration and have a bandwidth on the order of the transmission rate. Repeating the procedure in the above subsection for different pulse shapes and different optical and electronic filters is analytically intractable. Hence, we resort to a simulation and regression based approach to characterize the input output relations. The model obtained for the CW system above needs modification for practical systems. Effects of modulation bandwidths and resulting overlaps need to be incorporated. These are considered here.

For the simulations employed in Chapter 2 and here, bit rates of 10 Gbps are assumed and Gaussian pulse shapes are used. The spectrum of the baseband Gaussian pulse is given as

\[ P(\omega) = \sqrt{(2\pi)}T \exp \left( -\frac{\omega^2 T^2}{2} \right), \]

for which the spectrum has significant portions beyond the 10 GHz corresponding to the bit rate. As a result, the following changes are seen in the model:

1. Nonnegligible ACI is experienced, especially when the channel spacings are
reduced. As a result, (3.15) has additional linear terms and takes the form,

\[
r_k(T) \approx b_k A_k + \sum_{m=1, m \neq k}^K \rho_{k,m}^{(ACI)} b_m A_m + \rho_k^{(SPM)} b_k |A_k|^2
\]

\[
+ \sum_{m=1, m \neq k}^K \rho_{k,m}^{(XPM)} b_k b_m A_k |A_m|^2
\]

\[
+ \sum_{l=1, l \neq k}^K \sum_{n=1, n \neq k}^K \rho_{k,l,m-n}^{(FWM)} b_l b_m b_n A_l A_m^* A_n
\]

(3.18)

2. Substituting \( P(\omega) \) in (3.10) to obtain the correlation coefficients for SPM, XPM and FWM leads to integrals whose closed form is unknown to the author. Nevertheless, the wide nature of pulse spectrum leads to higher values for the SPM, XPM and FWM terms than indicated in Table 3.1.

3. FWM terms are not restricted to the condition that \( k = l + n - m \), since this constraint that exists due to the Kroneckor delta in (3.13) does not hold anymore. Thus, the complete model is a full third order model given as

\[
r_k(T) \approx b_k A_k + \sum_{m=1, m \neq k}^K \rho_{k,m}^{(ACI)} b_m A_m + \rho_k^{(SPM)} b_k |A_k|^2
\]

\[
+ \sum_{m=1, m \neq k}^K \rho_{k,m}^{(XPM)} b_k b_m A_k |A_m|^2
\]

\[
+ \sum_{l=1, l \neq k}^K \sum_{n=1, n \neq k}^K \sum_{m=1, m \neq k}^K \rho_{k,l,m-n}^{(FWM)} b_l b_m b_n A_l A_m^* A_n
\]

(3.19)
Chapter 3. Modeling Fiber Cross-talk Effects in Discrete-Time

The effects of nonideal filter responses are still not considered. Effects of nonideal optical filters appear as components of the correlation terms and these effects are seen in the simulation results. Nonideal electrical filtering that has been ignored here leads to higher order terms, as well as memory effects. Investigating these is beyond the current scope of this research.

The model given in (3.19) is for a noiseless and time-independent system. In reality the system is both noisy and time-varying. Consequently, the channel parameters are also time varying. Additionally, if the laser phases are unstable as is the case for present day systems, then the channel may appear as a fast fading channel. Any receiver processing to mitigate cross-talk for such a system has to estimate the correlation coefficients. For example, accounting only for ACI and ignoring other effects would require estimating the $K - 1$ terms $\rho_{k,m}^{(ACI)}$ for each channel $k$. However, with a full third order estimate combined with ACI effects would require estimating $K + K^3$ coefficients and then processing accordingly at the receiver or transmitter (which has this estimate information fed back from the receiver for preprocessing). While symmetry arguments can help, the number of coefficients that have to be determined are still on the order of $K^3$.

In [48] it is shown that for single span systems XPM effects are stronger than FWM effects when dispersion compensation is employed. Motivated by this, two low-complexity models are proposed in [26], viz.

ACI-Dominant:

$$ r_k(T) \approx b_k A_k + \sum_{m=1, m \neq k}^{K} \rho_{k,m}^{(ACI)} b_m A_m, $$

(3.20)
and

\[ r_k(T) \approx b_k A_k + \sum_{m=1 \atop m \neq k}^{K} \rho_{k,m}^{(XPM)} b_k b_m A_k |A_m|^2. \]  

Both these channel models have the benefit that only a total of \( K - 1 \) terms need to be estimated per channel. For the 8-user system that is simulated in Chapter 2, both these models require estimates of seven correlation coefficients, whereas the full third order model would require estimating at least 128 coefficients. Clearly these models reduce computational complexity by one order of magnitude in this case. Furthermore, since polynomial basis functions are nonorthogonal, it is likely that some of the effects of higher order interactions are captured in these models. Since estimating the ACI, XPM and SPM coefficients all together would require estimating \( 2K - 1 \) coefficients, the following model is also considered in the results presented here:

\[ r_k(T) \approx b_k A_k + \sum_{m=1 \atop m \neq k}^{K} \rho_{k,m}^{(ACI)} b_m A_m + \rho_{k}^{(SPM)} b_k A_k |A_k|^2 + \sum_{m=1 \atop m \neq k}^{K} \rho_{k,m}^{(XPM)} b_k b_m A_k |A_m|^2. \]  

The number of coefficients to be estimated for each user, for the different models in the above discussion, is summarized in Table 3.2. While the complexity is expressed in terms of the total number of channels, often channels that are hundreds of GHz apart can be ignored and a much smaller value of \( K \) is required below. In the following section, a linear regression approach is employed to approximate the coefficients of the nonlinear models in (3.19)-(3.22) from channel simulation data and used to compare
the accuracy of the models.

Table 3.2: Number of correlation coefficients to be estimated per channel for the four models given in (3.19), (3.20), (3.21) and (3.22).

<table>
<thead>
<tr>
<th>Model</th>
<th>Number of Coefficients to be estimated per channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model with full third order approximation (3.19)</td>
<td>$K^3 + K - 1$</td>
</tr>
<tr>
<td>ACI-dominant model (3.20)</td>
<td>$K - 1$</td>
</tr>
<tr>
<td>XPM-dominant model (3.21)</td>
<td>$K - 1$</td>
</tr>
<tr>
<td>Model with ACI, XPM and SPM (3.22)</td>
<td>$2K - 1$</td>
</tr>
</tbody>
</table>

### 3.4 Model Accuracy via Simulation

The models given in the previous section are compared here. Simulation results obtained in Chapter 2 are employed. A least-squares regression is performed on the sampled data to estimate the coefficients and fit the nonlinear models implied by (3.19), (3.20), (3.21) and (3.22). These pre-detection models for $r_k(T)$ are complex valued, while the actual system models for the post-detection data $y_k = |r_k(T)|^2$, indicate that the practical system output is real and positive. This makes estimating the complex correlation coefficients difficult for practical situations. Since, the pre-detection data corresponding to $r_k(T)$ is available from the actual simulation, it is used here to obtain coefficients as it provides a theoretical validation of the models considered. It should be noted that estimating the coefficients from post-detection data is an open problem that needs to be addressed beyond this thesis.
A linear regression approach is used to estimate the coefficients for each span and power level. The known modulated input sample information about $A_k$ and $b_k$ is used to form the regression matrix. To compute coefficients for a particular span, the output for that span after dispersion compensation and optical amplification is passed through optical filters and sampled. The regression matrix is then used to obtain coefficients that approximate these sampled outputs. When both third and first order terms are present as is the case for (3.19) and (3.22), the first order coefficients are obtained first, and the third order coefficients are then obtained by approximating the residual from the first order term.

The normalized mean square error (MSE) in the modeling is computed as a normalized $MSE = \frac{\sum_k |r_k(T) - \hat{r}_k(T)|^2}{\sum_k |r_k(T)|^2}$, where $\hat{r}_k$ denotes the output computed by applying the coefficients obtained via regression to the models. The average normalized MSE for different spans, averaged across all input power levels, is shown in Fig. 3.2 for the different types of systems considered here. Coefficients are obtained for each simulated power level individually and the averaging operation is performed across the power levels considered. As expected, the ACI-dominant system has the largest error, followed by the XPM-dominant system. The system with ACI, XPM and SPM effects combined as given in (3.22) has minimal benefits over the XPM-dominant model. This implies that modeling the channel as either ACI-dominant or XPM-dominant and ignoring the SPM effects should suffice when the error is tolerable. Clearly, the system with all the forms of cross-talk considered has the lowest error, which comes at the cost of increased estimation and computational complexity. For the system with optical demultiplexing via simple bandpass filtering, the accuracy of modeling 50 GHz systems is always better than that for 25 GHz systems. Since, 50 GHz spacing ensures fewer contributions from higher order terms, (greater than third order that are neglected in the approximation process for obtaining the model), this
Figure 3.2: Normalized MSE averaged across all power levels in the range of 0.25 – 1.25 mW/channel is given for the different models for multispan 25 GHz and 50 GHz systems. The system in Fig. 3.1 with single channel matched optical filters is considered in (a) while that with practical bandpass filters is considered in (b).
behavior is expected. Interestingly, for systems with optical demultiplexers matched to single channel systems, the error for 25 GHz systems is lower than that for 50 GHz systems for some spans. This could be explained by the fact that, while the 25 GHz system clearly experiences more nonlinearity, it is possible that the higher order terms are negligible for a few spans, and the third order models are sufficient to capture the corresponding nonlinearity. Note that for the system with optical bandpass filters, a second order Butterworth filter is used.

Figure 3.3: The number of spans for which the modeling MSE is within the specified tolerance of 1%, 2%, 5% and 10% for a 25 GHz system. The different power levels in the range of 0.25 mW/channel to 1.25 mW/channel are considered. Single channel optical matched-filter receiver is assumed. Identical results are found when practical bandpass filters are employed instead.
Figure 3.4: The number of spans for which the modeling MSE is within the specified tolerance of 1%, 2%, 5% and 10% for a 50 GHz system. The different power levels in the range of 0.25 mW/channel to 1.25 mW/channel are considered. Single channel optical matched-filter receiver is assumed. Identical results are found when practical bandpass filters are employed instead.

As discussed in Chapter 2, nonlinear effects depend on channel spacing, transmit power and distance. The MSE measures shown in Fig. 3.2 are averaged across all power levels. Instead of measuring the average normalized MSE across each power level, measuring normalized MSEs for each power level indicates the validity of the models for each transmit power level, bandwidth and distance. To capture the parameters into a single plot, bar plots are shown in Fig. 3.3 and Fig. 3.4. The number
of spans for which the MSE is within the tolerances of 1%, 2%, 5% or 10% for different power levels are shown. For lower power levels, since the higher order nonlinearities are less dominant, the models are expected to be more accurate. This is confirmed by the two figures from the fact that a given model is more accurate for more spans for low transmit power levels.

For both 25 GHz and 50 GHz systems, the number and power levels across which the ACI-dominant system model is valid is minimal. For higher power levels, the ACI-dominant system model becomes invalid when the tolerance is 1%. For all the power levels, the XPM-dominant model is valid for at least one span for 25 GHz systems and for at least two spans for 50 GHz systems. The number of spans for which modeling ACI, XPM and SPM jointly is more effective as compared to XPM-dominant systems is low for both 25 GHz and 50 GHz systems. As the number of spans increase and the power level increases, nonlinear effects start dominating and considering FWM terms can help, but only up to a certain extent, as indicated by the plot.

The coefficients as expressed in (3.16) for the CW case imply that the actual coefficients are independent of the amplitudes and hence the actual transmit power. This is valid only when amplitude coding is not performed. In the case of amplitude coding, as is done with signal design schemes in Chapter 5 and 6, for each scheme a SSF simulation must be performed on the designed scheme to ensure model conformity. Additionally, the models considered are approximations, and the actual system can have significant higher order terms. The degree of energy contribution from the higher order terms to lower order terms is a function of the actual transmit power. Since polynomial basis functions are not orthogonal, higher order effects may influence the first and third order coefficients estimated here. To test for this effect, two approaches are taken to obtaining the coefficients and the relative error between
the two is noted. In the first, coefficients is obtained individually for each power level and the mean square error is computed. In the second, the simulated transmit powers are divided into two ranges, a low range of $0.25 - 0.75$ mW/channel and a high range of $0.75 - 1.25$ mW/channel. For each range, a single set of coefficients are obtained to approximate the received signals for all the powers in the given range. The excess error resulting in the second approach when compared to the first is shown in Fig. 3.5. The error difference is significantly lower for ACI-dominant systems. This is expected for two reasons:

1. For lower spans, the nonlinear effects are less dominant. As such the channel behavior prior to the photodetection is linear and can be modeled as being transmit power independent.

2. For higher spans, the nonlinear effects dominate. Even though ACI-dominant systems has lower error difference, the actual error is high.

For the three models with XPM terms, the difference in modeling error increases monotonically with increasing power in the low power range. Nevertheless, the error is significantly larger when a single set of coefficients is used. From a practical perspective, optical channels do not suffer from fluctuating power levels, and therefore, estimating and using the individual coefficients for a given transmit power level should be feasible when amplitude coding is not performed.
Figure 3.5: Extra error in the average normalized MSE caused when a single set of coefficients is obtained as compared to the average normalized MSE when coefficients are obtained separately for each power level. Only the system with optical demultiplexing via single channel matched filters given in Fig. 3.1 is shown. Identical results are found for the system with practical bandpass filters. The averaging for case with individually obtained coefficients is done over the relevant power range.
The correlation coefficients are known to be complex. Further, the first order and third order coefficients have different units. For example, the XPM terms are multiplied with the additional $|A_m|A_k$ term when compared to the ACI coefficients. In the XPM-dominant model, the effective interference coefficients can be said to have the form $\rho_{k,m}|A_m|A_k$. Similarly for the SPM and FWM terms there are additional multiplicative amplitude terms. Therefore, for meaningful comparison, the correlation coefficients for third order terms should be appropriately scaled. In wireless and wired systems with interference effects, the models employed are sometimes scaled so that the interference free term has unit amplitude. Fig. 3.6 shows the scaled correlation coefficient magnitudes for the fourth channel in the simulated system that incorporates the effect of modulation. The scaling is performed so that the interference free term, $b_4A_4$ in (3.19), (3.20), (3.21) and (3.22) with $k = 4$, does not have any multiplicative constants from the filtering processes. The XPM terms are additionally scaled to include the additional terms $|A_m|A_k$ into the coefficient. As such, the XPM coefficients, like the ACI coefficients, do not have any units. Clearly, the coefficients for XPM are greater than 0.1, and for ACI-dominant systems the coefficients are found to be in the range of (0, 0.12). In the following chapter we focus on designing detectors for mitigating interference. To analyze the performance we assume that the correlation coefficients have magnitudes less than unity.
Chapter 3. Modeling Fiber Cross-talk Effects in Discrete-Time

Figure 3.6: Scaled correlation coefficients for ACI-dominant and XPM-dominant systems for the fourth channel in an 8-channel system. The coefficients for XPM-dominant systems are multiplied with the transmit power to normalize and compare with $k = 4$ in (3.21). Systems with optical demultiplexing via single channel matched filters is considered in (a) while that with practical bandpass filtering is considered in (b). A power level of 0.5 mW/channel is considered. The ACI terms and the XPM terms have no unit due to appropriate scaling.
3.4.1 Energy Divergence in the Models

In obtaining the models considered, a number of approximation are made. These are listed below:

1. VSTF terms of order greater than three are ignored in modeling a single span.

2. Approximations made in going from (3.5) to (3.7) required ignoring the fifth and higher order terms that arise in the multispans operation even when the VSTF approximation itself is restricted to third order.

3. To reduce complexity, the ACI and XPM dominant models are obtained by ignoring the effects of FWM terms.

4. A nonlinear regression approach was used to obtain the coefficients.

As a result of these approximations, the models may indicate that higher energy is received than transmitted depending on the correlations obtained. These models lead to energy divergence and are often referred to as nondepletion models indicating that higher order effects need to be incorporated into the model. To counter this effect, we assume that only models that indicate the normalized mean square error in modeling is below 2% are valid. As such, the ACI-dominant models are not valid for more than three spans for lower power levels and a single span for higher power levels. XPM-dominant models can be employed for up to five spans for lower power levels and two to three spans for higher power levels. The research in the remaining chapters, which talks about ACI-dominant and XPM-dominant systems, is hence assumed to hold for these systems.
3.5  Summary, Conclusion and Future Research

Discrete-time modeling of nonlinear channel effects is addressed in this chapter. These are summarized below:

1. For multispans systems with dispersion compensation and optical amplification between spans, applying the VSTF method indicates that for the continuous-time model, odd nonlinear terms of order greater than three are present. Limiting the VSTF model to third order terms and ignoring the higher order terms in the multispans operation implies that the multispans model is identical to the single span system up to a constant for the third order coefficients.

2. Modeling a simple system with unmodulated CW inputs in discrete-time by limiting the VSTF continuous-time model to have third order terms enables characterizing the cross-talk terms into XPM, SPM and FWM. Using this, four models with varying levels of complexity are proposed for real systems with modulated inputs.

3. The accuracy vs. modeling complexity tradeoff is studied. Results indicate that for low power levels using a first order ACI-dominant model suffices, and with increasing distances the XPM-dominant model suffices. Modeling SPM and ACI effects along with XPM does not significantly reduce the error from modeling the system using XPM effects alone. The full third order model is more accurate for a large range of power levels and spans but requires higher complexity.

4. The polynomial nature of the model implies that higher order terms influence the accuracy of lower order terms depending on the power level. Obtaining a common set of coefficients for multiple power levels is therefore detrimental.
5. For a simulated realistic system, the appropriately scaled correlations have magnitudes in the range of \((0,0.2)\) for XPM-dominant models and \((0,0.1)\) for ACI-dominant models.

The ACI and XPM-dominant models are valid for low power levels and fewer spans providing a low complexity modeling alternative for such systems. While ACI-dominant models are applicable when device leakages dominate, XPM-dominant models are applied when certain third order nonlinear effects dominate. Due to the nonorthogonality of polynomial basis functions, terms of different orders influence each other. As such, the XPM and ACI-dominant effects capture some higher order nonlinear effects as well. The remaining chapters in this dissertation focus on designing signal processing schemes for systems that are characterized by these models. The following chapter focuses on designing detectors to compensate for these two effects and signal designs are considered in the remaining chapters. Signal design for just the FWM terms in \((3.15)\) is studied from a theoretical perspective in the final chapter. This chapter has focused on modeling the channel ignoring the effects of noise and associated correlation parameter estimation problems. These are considered for a special class of ACI-dominant systems in Chapter 5. A detailed treatment of modeling the effects of noise, electrical filtering, and estimation issues are left for future research.
Chapter 4

Low-Complexity Multiuser Detection for
WDM Systems

4.1 Introduction

The results of Chapter 2 show that with reducing channel spacing, increasing transmission distance and/or transmission power, fiber nonlinearity becomes dominant, causing cross-talk that can severely degrade the overall energy efficiency or reduce the throughput of the system. Tackling the nonlinear cross-talk effects and cross-talk from device leakages requires (i) characterizing the channel behavior in discrete-time, (ii) and designing interference mitigation schemes. Chapter 3 addressed the first need to characterize the channel behavior. Lasers employed in WDM systems have rapidly varying phases as their coherence time can be as low as a few nanoseconds. For example, the IQS-2400 WDM laser source has a frequency linewidth of 20MHz, [57], implying a coherence time of 50 nanoseconds [58]. Consequently, the channel as seen by the receiver is not only nonlinear, but also rapidly changing. As such, the interference mitigation schemes employed should be low-complexity and high speed. Such schemes are investigated in this chapter.
Analysis of WDM systems indicates that joint processing of the different channels can significantly improve the achievable capacity region [16]. Jointly detecting multiple channel signals is identical to multiuser detection (MUD) schemes that were proposed in [59]. This topic has been heavily investigated for linear wire-line and wireless channels [28]. MUD is suggested as a method for increasing spectral efficiency in [60]. Research on MUD based interference cancelation for optical multiple access technologies, specifically optical CDMA, has been carried out extensively in [61], [62] and references therein. These are focussed on the effects of photodetector noise (instead of the ASE noise), which results in Poisson statistics at the receiver, and are not directly applicable to the models considered here. Cross-talk cancelation for dense WDM systems presented in [63] and [64] treats the output and not the input noise of the photodetector. Noncoherent MUD investigated in [65] is based on amplitude measurements at the receiver front-end as opposed to intensity measurements. Due to the symmetric, signal independent nature of the noise resulting from the assumptions made in [63]-[65], the results therein are not directly applicable to the model considered here. Additionally, nonlinear interference induced by the optical channel has not been addressed in [61]-[65] and references therein.

In Chapter 3, models with varying levels of complexity were presented. It is shown that for lower power levels and multispan systems with a few spans, the channel can be modeled as either ACI-dominant or XPM-dominant. In this chapter, the problems of designing low-complexity MUD schemes for WDM systems where these models are valid are addressed. The frequency spacing, power levels and number of spans for which the model is assumed valid can be determined from Fig. 3.3 and Fig. 3.4. After generalizing the structure of the ACI-dominant and XPM-dominant system models to include noise and arbitrary correlations, MUD schemes for these systems are studied.
Optimal and sub-optimal MUD with exponential complexity in the number of channels are investigated for ACI-dominant and XPM-dominant models in [26] and [27]. Fiber optic communication requires high-speed processing since the data rates employed for each channel are on the order of a few Gbps. Additionally, the unstable nature of laser phases causes the channel to appear like a fast fading channel. As such, the detectors need to be low-complexity to enable high-speed processing and fast adaptation. Designing low-complexity affine and quadratic MUDs is the primary focus of this thesis.

Asymptotic multiuser efficiency (AME) [66] indicates the energy efficiency (in terms of reducing BER with vanishing noise) of a given multiuser system relative to that of a single user system. Optical channels operate at extremely low bit-error-rates (BER) or correspondingly at extremely high signal to noise ratios (SNRs). Therefore, the focus here is to design detectors to maximize the AME, which is a direct measure of high SNR performance. A sensible detector is said to exist when it has a nonzero AME. The general problem of existence conditions for detectors with nonzero AME is investigated. Followed by this, low-complexity detector designs are investigated.

It turns out that the AME of the affine and quadratic detectors investigated, and in general all higher order polynomial detectors, is nonconcave in the detector coefficients. This complicates the design of the optimal affine and the optimal quadratic detectors, motivating the search for sub-optimal ones. The asymmetric nature of the noise at the photodetector output can cause well-known sub-optimal designs like the standard affine minimum mean square error (AMMSE) detector intended for wireless and Gaussian channels [28], [29], to perform poorly for optical channels. A novel regression-based approach to designing polynomial detectors is proposed and applied to the design of affine and quadratic detectors. Additionally, search based approaches for obtaining optimal affine and quadratic detectors are investigated. Applying the
regression-based designs to ACI-dominant and XPM-dominant systems indicates that the proposed approach can result in performance that is close to the search based detectors.

The proposed regression-based MUD design technique is robust to the noise statistics and interference structures and hence can be applied to multiuser systems other than the WDM system studied here. As an interesting example, applying this technique to linear Gaussian channels with BPSK modulation results in the decorrelating detector, which is shown to be the asymptotically optimal linear detector for such systems in [67]. The key benefit of this approach is that it enables quick adaptation of detector coefficients and does not necessitate channel estimation.

The chapter is organized as follows: The discrete-time channel models studied in Chapter 3 are modified to consider the effects of noise and vectorized in Section 5.2. Section 4.3 considers the design of MUDs where existence conditions are investigated. High-complexity design are reviewed in Section 4.4. The design of low-complexity MUD is investigated in Section 4.5. Performance of the proposed detection schemes is analyzed in terms of the AME as well as simulated BER in Section 4.6. Conclusions are drawn in Section 5.7.

4.2 Noisy Channel Models for Multichannel Fiber Optic Communications

The models studied in Chapter 3 were noiseless. Here we consider noise introduced by amplifiers. Noise from photodetectors is assumed to be absent. Incorporating noise into the system model, the decision statistics generated at the receiver can be
written as

\[ y_k = |b_k A_k + \rho_k(b) + n_k|^2, \quad k = 1, \ldots, K. \]  

(4.1)

Here \( b = [b_1, \ldots, b_K]^T \), represents the set of transmitted bits, \( \rho_k(b) \) represents the cross-talk introduced by the fiber and the demultiplexing process for channel \( k \), and \( n_k \) represents the complex Gaussian noise term with variance \( \sigma^2/2 \) per complex dimension. In Chapter 3 it was observed the considering ACI and XPM jointly does not significantly affect the accuracy. Hence, the interference term is assumed to have one of the following forms corresponding to (3.20) and (3.21):

\[ \rho_k(b) = \begin{cases} 
\sum_{i \neq k} \rho_{i,k} b_i A_i, & \text{ACI-dominant system} \\
\sum_{i \neq k} \rho_{i,k} b_i b_k |A_i|^2 A_k, & \text{XPM-dominant system} 
\end{cases} \]  

(4.2)

Note that the superscript on the correlation coefficient is ignored for brevity since only of the two forms is sufficient to model the cross-talk present. This convention is followed throughout the rest of the dissertation. The vector of received statistics, \( y = [y_1, \ldots, y_K]^T \), can be written using cross-correlation matrices \( R \) and \( R_b \) in vector form as,

\[ y = \begin{cases} 
|ARb + n|^2_o & \text{ACI-dominant system} \\
|R_bb + n|^2_o & \text{XPM-dominant system} 
\end{cases} \]  

(4.3)

where \( | \cdot |^2_o \) denotes the element-wise Schur product of a vector with itself and \( n = [n_1, \ldots, n_K] \) represents the noise. For ACI-dominant systems, the effective amplitudes of the different users are denoted by the diagonal matrix \( A \) with elements \([A]_{i,i} = A_i\).
and the cross-correlations are represented by the matrix $\mathbf{R}$, defined as

$$[\mathbf{R}]_{i,j} = \rho_{i,j}. \quad (4.4)$$

For XPM-dominant systems, the amplitudes are incorporated along with the cross-correlations in the cross-correlation matrix $\mathbf{R}_b$, which depends on $b$ and takes the form

$$[\mathbf{R}_b]_{i,j} = \begin{cases} 
\rho_{i,j} |A_i| |A_j|^2 b_i, & i \neq j \\
A_i b_i, & i = j.
\end{cases} \quad (4.5)$$

It turns out to be beneficial to analyze and design the detectors after taking the square-root of the received statistics denoted by $x_i = \sqrt{y_i}$ [27]. In vector notation,

$$\mathbf{x} = \begin{cases} 
|\mathbf{A} \mathbf{R} \mathbf{b} + \mathbf{n}| & \text{ACI-dominant Interference} \\
|\mathbf{R}_b \mathbf{b} + \mathbf{n}| & \text{XPM-dominant system}
\end{cases} \quad (4.6)$$

The following notation is used to denote noise-free received signals for the remainder of this chapter:

$$\mu(b) = \begin{cases} 
|\mathbf{A} \mathbf{R} \mathbf{b}| & \text{ACI-dominant system} \\
|\mathbf{R}_b \mathbf{b}| & \text{XPM-dominant system}
\end{cases} \quad (4.7)$$

Some important communication theoretic assumptions made in arriving at the above model are as follows:

1. A single shot approach is used since intersymbol interference effects are assumed absent and perfect timing synchronization is assumed between the transmitter and receiver.
2. The signal space expansion that takes place over the channel due to nonlinearity is ignored. It is shown in [16] that the information loss resulting from neglecting such expansion is minimal.

3. The statistics $\mathbf{y}$ and correspondingly $\mathbf{x}$ may not be sufficient even when the demultiplexing is performed via single channel matched optical filters. Even in the absence of ideal optical matched-filtering the sampled statistics, although insufficient, can still be used in the decision making process.

4. The noise terms $n_k$ are assumed independent. Although, this is not the case when amplifier noise from a middle span travels through the fiber before reaching the destination, or when the optical demultiplexing filters are wide-band, the assumption is made to simplify the design and analysis.

In a real system, estimating the correlation coefficients in the presence of noise is an open problem. Part of this problem is addressed in Chapter 5. In this chapter, we study the design of detectors and apply the proposed designs to ACI- and XPM-dominant systems. For the results presented, real correlation coefficients are assumed for the simplifying the presentation. The extension of the designs to the complex case is straightforward. In the following section, the detection problem is described.

### 4.3 Detector Design Problem

Given the statistics $\mathbf{y}$, the detection problem is to estimate the transmitted bits $b_i$, $\forall i$ with minimal probability of error. Due to cross-talk each element of the vector $\mathbf{y}$ has information about each $b_i$. Optimal processing thus necessitates that each bit be detected using all the elements of $\mathbf{y}$. Without loss of generality user $k = 1$ is

---

1While the correlations may be different, the design scheme for all the users/channels is identical.
chosen as the user of interest. The user index, whenever missing, is assumed to be 1 in the following discussion. The detection problem is to partition the decision space, \( \mathbb{R}^K_+ \), into two complementary sets \( \Omega_1 \) and \( \Omega_0 \), corresponding to estimates \( \hat{b}_1 = 1 \) and \( \hat{b}_1 = 0 \), respectively. The partition is to be chosen to minimize the probability of error. At the same time it is desired that the computational complexity of the decision making process be minimal.

In the following section, the jointly optimal detector that has the exponential complexity in the number of users is described. The AME measure used to compare detectors is introduced next, followed by the study of existence conditions for detectors. High and low complexity detectors are studied in the remaining sections.

### 4.3.1 The Jointly Optimal Detector

The *jointly* optimal detector chooses the transmitted vector \( b \) which maximizes the likelihood of the observed vector, i.e.,

\[
\hat{b} = \arg \max_b L_y(y|b),
\]

where

\[
L_y(y|b) = \prod_k f_{\chi^2_2(\mu_k^2, \sigma^2)}(y_k).
\]

Here each \( f_{\chi^2_2(\mu, \sigma^2)}(x) \) is a conditional central or noncentral chi-squared density function given as

\[
f_{\chi^2_2(\mu, \sigma^2)}(x) = \frac{1}{2\sigma^2} \exp \left( -\frac{x + \mu^2}{2\sigma^2} \right) I_0 \left( \frac{\sqrt{x\mu}}{\sigma^2} \right), \quad x \geq 0 \tag{4.10}
\]
where \( I_0(\cdot) \) represents the zeroth-order modified Bessel function of the first kind.

The optimal detector achieves the lowest error rate among all detectors but has the highest complexity. Apart from the error rate, asymptotic multiuser efficiency is often employed to compare different detectors at low noise levels as is the case for most optical systems. This measure is described below.

### 4.3.2 Asymptotic Multiuser Efficiency

Asymptotic multiuser efficiency (AME) \([66]\) is used to characterize the performance of multiuser schemes, relative to single user systems, in the high signal to noise ratio (SNR) regime. The AME of a particular scheme is the ratio (under vanishing noise) of the average transmit power \( A^2 \) needed in a single user system, to the average transmit power per user, say \( A_{\text{scheme}}^2(\sigma) \), that is required in a multiuser system employing a particular multiuser scheme to achieve the same bit error rate, i.e.,

\[
\eta_{\text{scheme}} \triangleq \lim_{\sigma \to 0} \frac{A^2}{A_{\text{scheme}}^2(\sigma)}. \tag{4.11}
\]

As \( \sigma \to 0 \), the likelihood of the received signal in the \( x \)-domain can be approximated as multidimensional Gaussian, up to a constant with independent components having means \( \mu_k(b) \) and identical variance, as shown in \([27]\), \([68]\), for \( \mu > 0 \). Thus, considering decision regions in the \( x \)-domain is useful for analytic purposes. Let \( P_{e_{\text{scheme}}}(\sigma) \) denote the probability of error resulting from using a particular scheme at a given noise level. Then the AME is given as \([26]\),

\[
\eta_{\text{scheme}} = \sup \left\{ 0 < t : \lim_{\sigma \to 0} \frac{P_{e_{\text{scheme}}}(\sigma)}{Q \left( \frac{\sqrt{t} A}{2\sigma} \right)} < \infty \right\} \tag{4.12}
\]

where \( Q \left( \frac{A}{2\sigma} \right) \) denotes the achieved error probability in a single user channel with a
Chapter 4. Low-Complexity Multiuser Detection for WDM Systems

linear receiver, \( Q(\cdot) \) being the standard Q-function.

For the user of interest, there are a total of \( 2^{K-1} \) possible transmitted vectors corresponding to the null hypotheses, represented by the set \( \mathcal{H}_0 = \{ b : b_1 = 0 \} \) and \( 2^{K-1} \) corresponding to the alternate hypotheses, represented by the set \( \mathcal{H}_1 = \{ b : b_1 = 1 \} \). Let \( \mathcal{R}_0 = \{ \mu(b) : b \in \mathcal{H}_0 \} \) and \( \mathcal{R}_1 = \{ \mu(b) : b \in \mathcal{H}_1 \} \), denote the sets of noise-free received signals, i.e., the effective constellation points as seen at the receiver. The probability of error resulting from a particular scheme is given as

\[
P_e^{(\text{scheme})} = \frac{1}{2^K} \sum_b P_{r}^{(\text{scheme})}(\hat{b} \neq b|b),
\]

(4.13)

where \( P_{r}^{(\text{scheme})}(\hat{b} \neq b|b) \) denotes the conditional probability of error given that \( b \) was transmitted. Under vanishing noise, the conditional error probability conditioned on \( b \) depends on the minimum Euclidean distance of \( \mu(b) \) from the decision boundary \( B_{\text{scheme}} \), implemented at the receiver, i.e.

\[
P_e^{(\text{scheme})} \approx \frac{1}{2^K} \sum_b Q \left( \frac{d_{\text{min}}^{(\text{scheme})}(b)}{\sigma} \right),
\]

(4.14)

where

\[
d_{\text{min}}^{(\text{scheme})}(b) = \begin{cases} 
0 & \mathcal{R}_0 \cap \Omega_1 \neq \emptyset \text{ or } \mathcal{R}_1 \cap \Omega_0 \neq \emptyset \\
\min_{x \in B_{\text{scheme}}} \| x - \mu(b) \| & \text{otherwise}.
\end{cases}
\]

(4.15)

As the minimum conditional error probability dominates the performance in the low noise regime, the AME can correspondingly be shown to be

\[
\eta_{\text{scheme}} = \min_b \frac{d_{\text{min}}^{(\text{scheme})}(b)}{(A/2)^2}.
\]

(4.16)
Among all receiver-based signal processing schemes, optimal MUD directly minimizes the probability of error and clearly has the highest possible AME. It is possible that even optimal detectors may not exist in the sense that they may have zero AME. The nonexistence of an optimal detector implies that no other sensible detector exists and simultaneous communication with negligible error rate at high SNRs is not feasible across the multiuser channel. The following section investigates existence conditions for sensible detectors.

4.3.3 Existence of a Multiuser Detector

Clearly, for the existence of any detector, it is essential that the decision space can be partitioned into $\Omega_1$ and $\Omega_0$ such that $\Omega_1 \cap \mathcal{R}_0 = \emptyset$ and $\Omega_0 \cap \mathcal{R}_1 = \emptyset$. This occurs only when the sets $\mathcal{R}_0$ and $\mathcal{R}_1$ are separable, i.e., $\mathcal{R}_0 \cap \mathcal{R}_1 = \emptyset$. In the absence of interference this condition is trivially satisfied from the fact that $\mathcal{H}_0 \cap \mathcal{H}_1 = \emptyset$. For channels with linear receivers and linear interference, a full rank cross-correlation matrix is sufficient for the existence of the optimal detector [28]. This however, does not hold for square-law channels. This is shown for ACI-dominant systems using the following proposition, the proof of which is available in Appendix A.1 and [69],

Proposition 1: $\mathcal{R}_0 \cap \mathcal{R}_1 \neq \emptyset$ iff one of the following conditions hold

(a) $\mathcal{H}_1 \cap \{\text{Null}\{\mathcal{R}\}\} \neq \emptyset$

(b) $\exists \mathcal{B}, \text{ s.t. } [\mathcal{R}]_{:,1} = \sum_{i=2}^{K} P_i [\mathcal{R}]_{:,i}$

\[ P_i \in \{0, -I, \mathcal{B}, \mathcal{B} - I\}, \forall i \in 2, \ldots, K \]  

(4.17)

where $\mathcal{B}$ is a diagonal matrix with complex exponential entries and $[\mathcal{R}]_{:,i}$ denotes the $i^{th}$ column of the matrix $\mathcal{R}$. A detector exists if $\mathcal{R}_0 \cap \mathcal{R}_1 = \emptyset$, which requires that both conditions (a) and (b) be false. Although a full rank cross-correlation matrix...
ensures that condition (a) above is not satisfied, it is neither necessary nor sufficient to falsify (b). Hence, unlike linear channels, a full rank cross-correlation matrix is not sufficient to ensure the existence of an optimal detector. Furthermore, since the diagonal elements of $B$ can be nonidentical, even if conditions (a) and (b) are both false, $R$ may not necessarily be full rank. Hence, a full rank matrix is not even necessary for the existence of an optimal detector.

A similar fact can be stated for the case of XPM-dominant interference, (Proposition II given in Appendix A.2), albeit in terms of the cross-correlation matrix corresponding to the all ones transmission vector instead of $R$, and with a slightly different set of matrices [69]. Again, using arguments similar to those for ACI-dominant systems, a full rank matrix is neither sufficient nor necessary to guarantee the existence of any detector. Obtaining necessary and sufficient conditions for the existence of the detector beyond the disjointness of $R_0$ and $R_1$ remains an open problem. High complexity detectors that maximize the AME when sensible detectors are known to exist are discussed next.

### 4.4 High Complexity Multiuser Detection

Optimal multiuser detectors partition the decision space to directly minimize the probability of error and clearly have the highest possible AME among all detectors. Employing a high SNR approximation to the likelihood function in the $x$–domain, asymptotically optimal (AO) detectors are obtained in [27]. These take the form

$$
\hat{b} = \arg \min_b ||x - \mu(b)||.
$$

(4.18)
Chapter 4. Low-Complexity Multiuser Detection for WDM Systems

The optimal and AO detectors have complexity that is exponential in the number of channels. A sub-optimal approach studied in [27] consists of using pairwise linear (PL) detectors. These detectors consist of linear spline based approximations to the AO detector used to separate each of the $2^K$ hypothesis corresponding to the transmitted bits $b$. This approach requires $\frac{2^K}{2!(2^K-2)!} \approx 2^{2K-1}$ linear tests which is still exponential in the number of channels.

The AO detector as given in (4.18) is based on a high SNR approximation to the likelihood function in the $x-$domain and has AME identical to that of the optimal detector. Additionally, the PL detector also has AME identical to that of the optimal detector [27]. However, both these sub-optimal detectors have exponential complexity and require complicated implementation and, in case of the PL detector, a complicated design algorithm. The following section considers the design of polynomial detectors with specific focus on affine and quadratic detectors.

4.5 Design of Sub-Optimal Detectors

Given the high data rate nature of the fiber channel and the fact that laser phase changes can make the channel appear to be fast fading, two requirements are essential to designing low-complexity detectors:

1. The detector implementation complexity should be low for high-speed data detection

2. The design time to obtain a detector should be fast as the channel may be fast fading.

With this in mind, the design of two sub-classes of polynomial detectors are investigated here: affine detectors, which have linear complexity in the number of channels,
and quadratic detectors that have quadratic complexity.

The affine detectors operating in the $\mathbf{y}-$domain and $\mathbf{x}-$domain take the forms:

$$y_1 - \sum_{k=2}^{K} c_k y_k - d \begin{bmatrix} \hat{b}_1 = 1 \end{bmatrix} \geq 0, \begin{bmatrix} b_1 = 0 \end{bmatrix}$$

(4.19)

and

$$x_1 - \sum_{k=2}^{K} c_k x_k - d \begin{bmatrix} \hat{b}_1 = 1 \end{bmatrix} \geq 0, \begin{bmatrix} b_1 = 0 \end{bmatrix}$$

(4.20)

and are referred to as Affine-$\mathbf{y}$ and Affine-$\mathbf{x}$, respectively. The quadratic detectors operating on $\mathbf{y}$ and $\mathbf{x}$, denoted by Quadratic-$\mathbf{y}$ and Quadratic-$\mathbf{x}$, take the forms

$$y_1 - \sum_{k=1}^{K} \sum_{j=k}^{K} a_{j,k} y_j y_k - \sum_{k=2}^{K} c_k y_k - d \begin{bmatrix} \hat{b}_1 = 1 \end{bmatrix} \geq 0, \begin{bmatrix} b_1 = 0 \end{bmatrix}$$

(4.21)

and

$$x_1 - \sum_{k=1}^{K} \sum_{j=k}^{K} a_{j,k} x_j x_k - \sum_{k=2}^{K} c_k x_k - d \begin{bmatrix} \hat{b}_1 = 1 \end{bmatrix} \geq 0, \begin{bmatrix} b_1 = 0 \end{bmatrix}$$

(4.22)

Note that computing square-roots is done using iterative algorithms and adds to overall complexity for detectors operating in the $\mathbf{x}-$domain.

The detector design problem is to choose the detector coefficients $\{a_{j,k}, c_k, d\}$. Obtaining the optimal detector coefficients given the order of detector complexity is an open problem for all kinds of channels including wired and wireless. Before delving into the design problem, it is interesting to first establish conditions under which it is possible to obtain a affine/quadratic detector with nonzero AME. According to the well-known separating hyperplane theorem [70], two nonintersecting convex sets are separable by a linear hyperplane. Finite sets are separable if their convex
hulls are nonintersecting. Thus, the Affine–$x$ detector with nonzero AME exists if the convex hulls of $R_0$ and $R_1$ are nonintersecting. Note that $R_0$ and $R_1$ are defined in terms of $x$ and hence their nonintersection guarantees the existence of all polynomial detectors in $x$ that can have linear terms. The Affine–$y$ detector and all other polynomial detectors containing linear terms in $y$ exist if the convex hulls of the sets in the $y$–domain corresponding to $R_0$ and $R_1$ do not intersect. Exploiting the special structure of $R_0$ and $R_1$ to obtain conditions for the existence of the polynomial detectors is beyond the scope of this research. Checking for the intersection of convex hulls is a well known problem in computational geometry and numerical methods proposed in relevant literature can be employed to verify the existence of a polynomial detector before designing one [71]. Throughout this section detectors are designed assuming that they exist.

4.5.1 Single User Detectors and Affine MMSE Detectors

The lowest possible computational complexity is achieved by a single user detector (SUD). Traditionally, a SUD is defined as a detector that treats the presence of other users as increased noise variance. Hence, it corresponds to a typical single threshold detector which can be viewed as the special case of (4.20) with coefficients $\{c_i = 0, \forall i\}$. For a single user optical channel employing on-off keying it is easy to verify that the decision rule at high SNRs corresponds to:

$$y_1 \begin{cases} \hat{b}_1 = 1 & A^2 \\ \hat{b}_1 = 0 & \frac{A^2}{4} \end{cases},$$

(4.23)

or equivalently,

$$x_1 \begin{cases} \hat{b}_1 = 1 & A \\ \hat{b}_1 = 0 & \frac{A}{2} \end{cases}.$$

(4.24)
The AME of the system based on this scheme is given as

\[
\eta_{SUD} = \begin{cases} 
0, & \min_{b \in \mathcal{H}_1} \mu_1(b) \leq \max_{b \in \mathcal{H}_0} \mu_1(b) \\
\min_b \frac{\|\mu_1(b) - \frac{\mathbb{A}_2}{2}\|^2}{\mathbb{A}_2^2}, & \text{otherwise.}
\end{cases}
\] (4.25)

Clearly the SUD scheme completely neglects interference, hence it has an AME that is expected to act as a lower bound to compare with all other interference cancelation schemes. Alternately, it can be viewed as a special case of the affine and quadratic detectors and all designs are focused on improving on the performance of the SUD scheme.

One approach to indirectly minimize the probability of error is to choose the regions \( \Omega_1 \) and \( \Omega_0 \) so as to minimize the mean square error between the detector output \( \hat{b}_1 \), and the transmitted data \( b_1 \). Since the received statistics are not zero mean, an affine MMSE (AMMSE) detector given as

\[
E\{b_1\} + m^T (y - E\{y\}) \underbrace{\hat{b}_1 = 1}_{\hat{b}_1 = 0} \geq \tau
\] (4.26)

where

\[
m^T = E\{b_1(y - E\{y\})^T\} \left[E \{(y - E\{y\})(y - E\{y\})^T\}\right]^{-1},
\] (4.27)

is employed. Closed form expressions for the expectations required to compute \( m \) are easy to obtain but are cumbersome and hence not presented here. The threshold \( \tau \) has to be evaluated numerically. The filter coefficients are obtained analytically or numerically and the threshold are then mapped to the detector coefficients \( \{c_k\}_{k=1}^{K}, d \) of the Affine--\( y \) detector. An AMMSE detector in the \( x \)--domain can be similarly
obtained.

Simulation results indicate that the effective BER and hence the AME of the MMSE detector can be worse than that of the SUD. At high SNRs, the BER depends on the minimum squared error and not the mean squared error. As a result, minimizing the mean squared error does not necessarily minimize the minimum squared error and can yield poor performance as is shown in Section 4.6. Alternative approaches to designing affine and quadratic detectors are proposed in the following section.

### 4.5.2 Designing AME Maximizing Detectors

Choosing the detector coefficients for the affine and quadratic detectors that minimize the error probability at high SNR is equivalent to choosing coefficients that maximize the AME. The resulting detectors are referred to as AME maximizing detectors. Let the set of detector coefficients be represented by \( \alpha = \{ \{ a_{j,k} \}, \{ c_k \}, d \} \). For the set of affine detectors, \( \{ a_{j,k} \} \) are set to 0. The detector coefficients can be obtained by solving the following AME maximization problem:

\[
\hat{\alpha} = \arg \max_{\alpha} \eta(\alpha) = \arg \max_{\alpha} \min_{b} \left( \frac{d_{\min}^{(\alpha)}(b)}{A/2} \right)^2.
\]

The function \( d_{\min}^{(\alpha)}(b) \) is continuous in the coefficients \( \alpha \) and hence the minimum of these distances is also continuous. Therefore, AME can be stated as being a continuous function of the detector coefficients.

Given that AME is a continuous function and obtaining the AME-maximizing detectors would be simplified if the AME was concave with respect to these coefficients. However, except for the single user detector, the AME of all polynomial detectors is not concave with respect to the detector coefficients. This is shown in Appendix
A.3. The search for AME maximizing detectors is consequently not trivial, forcing sub-optimal approaches for obtaining the detector coefficients. Search algorithms employing gradient based approaches or genetic algorithms provide one approach. However, these approaches to obtaining coefficients are again high complexity and time consuming and do not meet the high-speed requirement stated at the beginning of this section. A novel regression-based approach is proposed in the next section.

4.5.3 Regression-Based Polynomial Detectors

The regression-based approach proposed in this section is motivated by the design of the pairwise linear (PL) detector, which can be viewed as a linear spline approximation to the AO detector [27]. The approach differs from the spline approach in that instead of linear splines, a single polynomial, for example, a single affine or quadratic curve, is used to divide the received space into $\Omega_1$ and $\Omega_0$.

The AO detector is implemented by testing all of the $2^K$ possible transmitted vectors against each other. The most likely error (MLE) points of a detector are those points on the boundary of the detector that have the highest likelihood. The MLE point when testing $b^{(l)}$ against $b^{(m)}$ using the AO detector is given as [27],

$$z^{(l,m)} = \arg \max_{x \in \mathbb{R}^K_+} \{ L_x(x|b^{(m)}) : L_x(x|b^{(m)}) = L_x(x|b^{(l)}) \},$$

where $L_x$ represents the likelihood function in the $x$–domain at an arbitrarily high but finite SNR. It is known that detectors with identical MLE points have identical AME [27].

When designing polynomial detectors, it is desirable that the resulting detectors have MLE points identical to those of the AO detector. A simple least-squares regression over the set of the AO detector MLE points can be performed to obtain
Chapter 4. Low-Complexity Multiuser Detection for WDM Systems

detector coefficients. Such a detector closely approximates the decision regions, \( \Omega_1^{(AO)} \) and \( \Omega_0^{(AO)} \) for user 1, resulting from the AO detector. The key step in this process is that of determining the set of points to be used for carrying out the regression and is discussed below.

The set of MLE points, \( \chi_B \), on the boundary separating \( \Omega_1^{(AO)} \) and \( \Omega_0^{(AO)} \) resulting from the AO detector implementation, is a subset of the set \( \chi \) of the \( 2^{2K-2} \) MLE points corresponding to testing all the elements of \( \mathcal{H}_0 \) against all the elements of \( \mathcal{H}_1 \). The remaining points, \( \chi^c_B \), referred as interior points, either lie in \( \Omega_1^{(AO)} \) or \( \Omega_0^{(AO)} \). Clearly the boundary of designed detectors should be away from these interior points. The following fact is used to determine the set \( \chi_B \):

Proposition III: Let \( b^{(m)} \in \mathcal{H}_1 \). Testing this hypothesis against all the hypothesis in \( \mathcal{H}_0 \) results in \( 2^{K-1} \) MLE points that we represent as \( T_m \subseteq \chi \). Let \( z^* \in T_m \), then

\[
  z^* \in \chi_B \iff \nexists \ j \text{ s.t. } L(z^*|b^{(m)}) < L(z^*|b^{(j)}).
\]

(4.29)

A formal proof is provided in Appendix A.4.

Using the fact stated, MLE points can be determined and used in a linear regression to obtain detector coefficients. Fig. 4.1 is an example of the how the MLE points are used for a typical two-user system. The points \( \{A, B, C, D\} \) represent the noise-free received signals \( \mu(b) \), and the set of MLE points \( \chi = \{e, f, g, h\} \). Regression-based detectors shown in the figure are obtained by using the set \( \chi_B = \{e, f, g\} \). \( \chi^c_B = \{h\} \) denotes the interior point. The PL detector is also shown.

If the boundary resulting from implementing a detector has first order derivative identical to that of the AO detector, then the AME of that detector is identical to that of the AO detector [27]. This principle is incorporated in the design of the PL detector which closely matches the slope to the AO decision boundary at the MLE
The minimum number of points required for performing a regression to obtain a quadratic detector is \( K^2 + K \). Observe that \( \forall K > 2 \), the total number of MLE points \( 2^{(2K-2)} \gg \frac{K^2+K}{2} \). This ensures that with a high likelihood, sufficient number of MLE points are available for the regression. In case the number of MLE points available is less than those required for the regression, some of the detector coefficients can be set to 0. The author expects such situations to be degenerate and assumes that sufficient number of points are available for the rest of this chapter. A more detailed
study on this topic, although interesting, is beyond the scope of this research.

The computational complexity of the proposed design algorithm is dominated by the need for computing a pseudoinverse. This requires computing an inverse, thus we expect the complexity of this algorithm to be cubic in the number of coefficients to be estimated. Thus for obtaining the affine detectors, the complexity is third order in the number of channels, while that for the quadratic detector is sixth order. This is lower than that required for an exhaustive search over the space of $K$ coefficients for the linear case and $\frac{K^2+K}{2}$ coefficients for the quadratic case. The proposed design however, may not lead to AME maximizing detectors. The following section discusses the search for quadratic and affine detectors as such.

4.5.4 Search Based Approach for AME Maximizing Detectors

As discussed in Section 4.5.2, obtaining AME maximizing affine and quadratic detectors corresponds to solving an optimization with a nonconcave objective function. The set of coefficients of all polynomial detectors with nonzero AME, if any exist, can be shown to be a cone (and its image) in the space of detector coefficients. Fixing the coefficients corresponding to $x_1$ or $y_1$ in (4.20), (4.21) and (4.22) corresponds to cutting the cone across the corresponding dimension in the space of detectors and results in a convex set of detectors. The space of detector coefficients $\alpha$ that give nonzero AME is thus convex. As such, two search approaches are considered here.

The first approach is to use gradient-based methods implemented using a steepest descent search algorithm. Such methods are sensitive to the initial condition used; if the initial solution is in a region of zero AME, the gradients and Hessian computation may return zero and the search remains at the initial solution pro-
Chapter 4. Low-Complexity Multiuser Detection for WDM Systems

Providing no improvement. The author proposes to use the detectors obtained using regression-based technique as initial solutions to search for locally-optimal detectors using gradient-based methods.

Genetic algorithms provide an alternative to gradient-based search methods and often reach the globally optimum point. For such searches an initial population containing the regression-based detectors and detectors based on the gradient search can be employed. These approaches are employed here primarily to validate the regression-based approach. Results in the following section indicate that for certain cases the regression-based design approach can result in locally optimal detectors. Importantly, the search based approaches are too slow to be suitable for adaptive systems further justifying the use of the regression approach.

4.6 Performance Analysis and Simulation Results

Different detectors are compared in this section via numerical AME computations and BER simulations. Real cross-correlations are considered for simplicity. Additionally, the cross-correlations are considered to be equal across all the users, i.e. $\rho_{i,j} = \rho$, $\forall i \neq j$. Although, this may rarely occur, the condition helps demonstrate the benefits of the different schemes on a two dimensional plot. However, it should be noted that the discussion above is applicable to any cross-correlation structure with real or complex cross-correlation coefficients. Typical two- and three-user systems are considered. Note that for mathematical completeness, all $|\rho| < 1$ are considered.

Obtaining the AME for MUD schemes involves computing Euclidean distances from the effective received constellation points to the effective decision boundary as indicated in (4.15). For linear detectors over $x$ of the form (4.20), this minimum
Chapter 4. Low-Complexity Multiuser Detection for WDM Systems

distance can be computed using

\[
\min_{x \in B(\{c_k\}, d)} ||x - \mu(b)|| = \frac{|\mu_1(b) + \sum_{k=2}^{K} c_k \mu_k(b) + d|}{\sqrt{1 + \sum_{k=2}^{K} c_k^2}}
\]

\[(4.30)\]

in (4.15). Closed form expressions are not available for other detectors and the distances have to be computed numerically, done here using a gradient-based constrained search algorithm. The search space is constrained to a \(K\)-dimensional hypersphere (in the positive quadrant) centered around \(\mu(b)\) and of radius equal to that of the efficiency of the optimal detector. The search begins with \(\mu(b)\) as the initial solution. The regression-based detector designs require likelihood ratio computation as seen in (4.29). An SNR of 25dB was assumed for this purpose.

4.6.1 ACI-dominant system

For ACI-dominant systems, the performance for the equal user transmit power case is independent of the amplitude \(A\). Therefore, without loss of generality, \(A\) is set to unity in this section.

A two user system is considered in Figs. 4.2 and 4.3. With only two users, the quadratic detector based on regression cannot be fully quadratic since sufficient MLE points are not available for the regression. The detectors shown were obtained by setting the quadratic terms for user 1 in (4.21) to zero. Fig. 4.2(a) shows the AME when the AO detector, the AMMSE detector, SUD, and the proposed regression-based detectors are employed. Fig. 4.2(b) shows the achievable BER by the different detectors considered in Fig. 4.2(a) for a two user system with \(\rho = -0.3\). This figure shows that the performance at low SNR mirrors the high SNR AME results.
Figure 4.2: (a) AME and (b) BER of different detectors in a two-user ACI-dominant system with $A_1 = A_2 = 1$. For the AME plot, $\rho_{1,2} = \rho_{2,1} = \rho$ is used and for the BER plot, $\rho = -0.3$ is used. The decision regions corresponding to the BER plot are shown in Fig. 4.1. (Regr.: Regression)
Figs. 4.3(a) and (b) compare the loss in AME relative to the AO detector, defined as

$$\text{Relative AME Loss}_{\text{scheme}} = \frac{\eta_{AO} - \eta_{\text{scheme}}}{\eta_{AO}},$$

(4.31)

when the regression-based detectors and the locally-optimal detectors obtained using gradient-based searches and genetic algorithm searches are employed. Relative performance of the SUD is also plotted for comparison. For a wide range of parameters the regression-based detectors offer significant gains when compared to the SUD. In some cases the loss incurred by the regression-based detectors is identical to that incurred by the locally-optimal detectors obtained using the search algorithms. This implies that the regression-based detectors may in some cases result in locally-optimal detectors. Although the genetic algorithm searches are not guaranteed to lead to the best detectors, they are as close as we can hope to get without an exhaustive search. Results indicate that for most cases, the gradient search based detectors perform close to the detector obtained using genetic algorithm search. Regression based detectors are used as initial conditions for the relevant searchers further motivating their usage. Another important result is that the quadratic detectors achieve nonzero AME for values of $\rho$ where the affine and SUD approaches fail to give a sensible detector.

An AME plot for a three-user system is given in Fig. 4.4. The results for both $K = 2$ and 3 users indicate that the AMMSE detector can perform much worse than the SUD, a fact that motivated this research. The AME of quadratic detectors is almost always better than the affine detectors. For a wide range of parameters the regression-based affine detectors perform better than the AMMSE detector.
Figure 4.3: A two-user ACI-dominant system with $A_1 = A_2 = 1$, $\rho_{1,2} = \rho_{2,1} = \rho$ is considered. The figures show relative AME loss w.r.t. the optimal MUD when SUD, the proposed regression-based detectors (Regr.), the locally optimal gradient search based detectors (Grad.) and the genetic search based detectors (GA) are employed.
Figure 4.4: The AME of different detectors designed for a three-user ACI-dominant system with $A_i = 1, \forall i, \rho_{i,j} = \rho \forall i \neq j$. (Regr.: Regression).

4.6.2 XPM-dominant system

Unlike ACI-dominant systems limited systems, the performance of XPM-dominant systems is sensitive to the effective transmitted amplitude. This stems from the fact that the interference matrix is signal dependent, even for small values of cross-correlations typically encountered. Results in this section assume unit amplitude for the transmitted signal. Note that it can be assumed that the correlation is fixed and the amplitude is scaled.

In the XPM-dominant case, the AME of the AO detector and the SUD are identical and unity for $\rho \geq 0$ [26]. Hence, for such systems detector performance is considered only for $\rho < 0$. MUD schemes are expected to show similar benefits to
Figure 4.5: The AME achieved when different detectors are employed for two-user
and three-user XPM-dominant systems with $A_i = 1$, $\forall i$, $\rho_{i,j} = \rho_{j,i} = \rho$, $\forall i \neq j$. (Regr.: Regression).
those indicated for linear interference limited systems. Figs. 4.5(a) and (b) show the AME when different detectors discussed are applied to nonlinear interference limited systems with two and three users respectively. For two-user systems, the performance of the AO detector improves monotonically with $\rho$ or reducing correlation. This is not the case for the three-user system. The zero AME at $\rho = -0.5$ is caused by the fact that the matrix $R_1$ is no longer full rank and condition (a) in (A.8) is satisfied. For the proposed detectors, most of the results are consistent with those for linear interference. Interestingly, the Affine-$\mathbf{x}$ detector outperforms the Quadratic-$\mathbf{y}$ detector for a wide range of cross-correlation values for systems with nonlinear interference. This anomalous behavior is caused by two phenomena: the MLE points of one detector obtained using regression can be much further from those of the AO detector than those of another, and the designs are not progressive, in the sense that higher order detectors are not obtained as successive refinements of lower order detectors. Importantly, the extra computation spent to take the square-root to obtain $\mathbf{x}$ is rewarded, as the Quadratic-$\mathbf{x}$ detector outperforms all the other detectors for most parameters. In some cases, specifically for small values of $|\rho|$ that are typically realized, the Quadratic-$\mathbf{x}$ detector achieves near-optimal performance.

4.7 Conclusion

Cross-talk can result in error-rates that are nonreducible with increasing SNR. While MUD based techniques can improve performance over simple SUD schemes, the existence of detectors with nonzero AME for any general ACI-dominant and XPM-dominant system is not guaranteed. Schemes with linear and quadratic complexity are considered as alternatives to SUD and previously-proposed optimal techniques. The signal dependency of the noise can render standard symmetric error reduction
based detection approaches like affine MMSE less useful for error reduction. A novel regression-based design methodology is proposed in this chapter to design polynomial complexity detectors. Applying this method to design affine and quadratic detectors for typical two- and three-user systems indicates that significant gains over standard approaches are achievable over a wide range of parameters. Search based algorithms are also applied to obtain detectors and comparing the regression-based detectors to these implies that the proposed approach can in certain cases lead to locally optimal detectors. The proposed algorithm requires a design complexity that is polynomial in the order of the coefficients and hence the number of channels. This complexity is lower than that for an exhaustive search over the space of coefficients.

The following chapters consider the design of low-complexity signal design schemes for fiber-optic systems which can act either as an alternative to MUD based interference cancelation or in combination with MUD schemes as is shown in the final chapter.
Chapter 5

Precoding Based Signal Design for WDM with Linear Interference

5.1 Introduction

Optically amplified regional and long-haul fiber optic networks employing wavelength division multiplexing (WDM) are performance limited by cross-talk among different channels. Receiver based signal processing techniques for minimizing the resulting interference effects have been investigated in Chapter 4. When the optical channel is slowly varying we consider the design of transmitter based signal design schemes to minimize the performance degradation resulting from interference. Our results indicate that depending on the correlation among channels or users the proposed techniques may outperform receiver based multiuser detection techniques. In this chapter the focus is on ACI-dominant systems. Precoding based signal design schemes, estimation of correlation coefficients and the impact of estimation error on the performance of the designed precoding based system are investigated in this chapter. Apart from WDM systems, optical CDMA systems can also be modeled by the ACI-dominant model with correlation coefficients corresponding to correlation among the
code sequences. Therefore, the research results are applicable to those systems as well.

For a system with relatively stable phases, the fiber system can be thought of as being similar to a slowly fading multiuser wireless channel [29]. The correlation can be estimated at the receiver and fed back to the transmitter. For multiuser wireless channels wherein transmission is centralized and interference is known to the transmitter _a priori_, the transmitted signal can be preprocessed to minimize interference effects [30], [31], and [72]. For our optical channel, the transmission and reception are both centralized. Hence, we can combine transmitter based preprocessing or precoding with sophisticated receiver based signal processing to improve the overall system performance. Here we investigate the design and performance of such schemes for the ACI-dominant fiber optic systems.

There are two parts to our design problem. One is to choose a transformation or precoder to operate on the signal to be transmitted while the other is to choose a receiver based signal processing algorithm to operate on the photo-detector output. To limit the search space for the transformation, we constrain the precoding operation at the transmitter to be affine. We propose two types of precoding strategies, viz. zero-forcing (ZF) precoding and \( \rho \text{-forcing (}\rho \text{F) precoding, } \rho \text{ being the desired correlation among users.} \) ZF-precoding strategies are similar to those used for wireless communications wherein the effective interference among users is forced to be zero, thereby resulting in uncorrelated statistics at the receiver. A single user detector (SUD) is employed at the receiver. It has been shown in [27] that for ACI-dominant models, correlated received statistics could improve the energy efficiency of the multichannel systems. Precisely, it is shown that the AME of the AO MUD could be greater than unity for certain cases when the correlation is nonzero. For example, in a 2-user system the energy performance using AO MUD strategies is better than a
single user system provided the correlation $\rho$, among users lies in the range $(0, 0.2)$ as indicated by Fig. 4.2(a). Motivated by this fact, we design $\rho_F$-precoders which force a certain correlation among the users. In this case a MUD needs to be implemented at the receiver. ZF-precoding can be viewed as a special case of $\rho_F$-precoding where the precoder tries to eliminate interference completely by forcing $\rho = 0$.

The design of precoders is a function of the channel correlation coefficients. In Chapter 3, we discussed how the channel can be modeled and employed a least squares approach to obtaining the channel coefficients from noiseless channel simulations. In the presence of noise, the coefficients have to be estimated. The maximum likelihood estimation of coefficients is considered here. The impact of incorrect estimation on precoding is also studied.

In summary, the following are the key contributions of this chapter:

1. ZF-precoders for ACI-dominant systems are obtained and analyzed.

2. $\rho_F$-precoders for ACI-dominant systems are obtained and analyzed.

3. Channel estimation for ACI-dominant systems is considered and the performance of simple estimation schemes is given.

4. The impact of channel estimation errors on the precoding schemes proposed is studied.

We begin by reviewing the system model in Section 5.2. The design of precoders is presented in Section 5.3. Section 5.4 gives insight into the performance of the schemes considered. In Section 5.5 we focus on the estimation of correlation. Simulation results are presented in Section 5.6 and conclusions are drawn in Section 5.7. The author would like to note that in this chapter, the ACI-dominant model derived in Chapter 3 is used for all the estimation and simulation results presented.
5.2 System Model

For low power WDM systems with two or three spans and for optical CDMA systems, the received statistics over different channels are correlated and the system can be modeled via an ACI-dominant model with noise. In the following discussion we assume that all users have identical received amplitude levels, $A_k = A, \forall k$. All schemes can be easily modified to accommodate for nonidentical amplitudes. Similar to the models in Chapter 4, the received signal is

$$y_k = |A b_k + \rho_k(b) + n_k|^2, \quad k = 1, \ldots, K,$$

(5.1)

where the cross-talk term is

$$\rho_k(b) = \sum_{i \neq k} \rho_{i,k} b_i A,$$

(5.2)

and $\rho_{i,k}$ represents the correlation between users $i$ and $k$. In vector notation we can write

$$\mathbf{y} = |A \mathbf{R} \mathbf{b} + \mathbf{n}|^2 \triangleq |A \mathbf{R} \mathbf{b} + \mathbf{n}| \circ |A \mathbf{R} \mathbf{b} + \mathbf{n}|,$$

(5.3)

where $\mathbf{y} = [y_1, \ldots, y_K]^T$, $[\mathbf{R}]_{i,j} = \rho_{i,j}$, $\mathbf{b} = [b_1, \ldots, b_K]$, $\mathbf{n} = [n_1, \ldots, n_K]$ and $\circ$ represents the Schur (element-wise) vector product.

With precoding, the transmitted signal is some transformation of the bit-vector to be transmitted, $\mathbf{b}$. Constraining the set of transformations to be affine, the transmitted signal is then given as $\alpha(M \mathbf{b} + \mathbf{c})$ instead of $\mathbf{b}$. We refer to $M$ as the precoder. The power scaling coefficient $\alpha > 0$ is used to maintain the average transmission energy per bit identical to that without preprocessing while $\mathbf{c}$ is a constant used to
Chapter 5. Precoding Based Signal Design for WDM with Linear Interference

ensure positivity of the transmitted signal (necessary for intensity modulated transmissions), i.e., $Mb + c > 0$. The choice of the constants and the precoder is ideally based on perfect *a priori* knowledge of the correlation among users. This is discussed next.

5.3 Precoder Design

In this section we look into the design of the precoding matrix $M$ and the choice of the constants $\alpha$ and $c$. Since ZF-precoding is a special case of $\rho F$-precoding, we focus on designing for the latter. The goal is to design $\rho F$-precoders such that the received statistics have a specific desired pairwise correlation $\rho_d$. The ideal received statistic before photodetection can be written $A R_d b$, where $[R_d]_{i,j} = \rho_d + (1 - \rho_d) \delta_{i-j}$ where $\delta_k$ represents the Kronecker delta and $\rho_d$ is the desired correlation among users. For ZF-precoding $\rho_d = 0$ and $R_d = I$.

For a given precoding matrix $\tilde{M}$, the positivity constants in $c$ can be chosen as

$$\tilde{c}_i = - \min_b [\tilde{M}b]_i, \quad \forall i \in \{1, \ldots, K\}. \quad (5.4)$$

Given $\tilde{M}$ and $\tilde{c}$, the power scaling coefficient can be obtained using the condition

$$E\{\tilde{\alpha}(\tilde{M}b + \tilde{c})^T \tilde{\alpha}(\tilde{M}b + \tilde{c})\} = A^2 K / 2 \quad (5.5)$$

as

$$\alpha = \sqrt{\frac{A^2 K / 2}{\text{tr}\{\tilde{B} \tilde{M}^T \} + \sum_{i=1}^K \sum_{j=1}^K c_i [\tilde{M}]_{i,j} + ||\tilde{c}||^2}}, \quad (5.6)$$

where $\tilde{B} = E\{bb^T\} = \frac{1}{4} I_{K \times K} + \frac{1}{4} 1_{K \times K}$. 
The received statistics before photodetection resulting from this precoding strategy can be written as $\alpha R(\tilde{M}b + c) + n$. Given the ideal interference-free received signal that is desired at the photodetector input is $AR_d b$, an MMSE precoder design problem can hence be formulated as

$$ (M, \alpha, c) = \arg \min_M E \left\{ \left\| AR_d b - \left( \alpha R(\tilde{M}b + c) + n \right) \right\|^2 \right\}. \quad (5.7) $$

The nonlinear functional dependence of the constants $\alpha$ and $c$ on $\tilde{M}$ makes it difficult to solve this problem analytically. Instead of employing complicated numerical approaches, the following sub-optimal solution is employed:

$$ M = \arg \min_M E \left\{ \left\| AR_d b - (R\tilde{M}b + n) \right\|^2 \right\}. \quad (5.8) $$

Choosing $M = AR^{-1} R_d$ solves this problem [72]. The constants $c$ and $\alpha$ are then chosen so as to satisfy the constraints given in (5.4) and (5.5). Note that although this solution is sub-optimal, it still results in $\rho-$correlated statistics at the receiver.

### 5.4 Performance Analysis

In this section, bounds on the achievable AME with the joint precoding schemes are obtained. To simplify the analysis we assume that all the users have identical correlation $\rho$. Consequently, the precoding matrix is symmetric and the positivity constant can be given as $c = c_1$, where $1$ is a vector of length $K$ and all ones. We first analyze the case with ZF-precoding which is then followed by the $\rho F$-precoding case.
5.4.1 Zero-Forcing Precoding

With ZF-precoding, as the received statistics are uncorrelated, conventional SUD technique can be used to detect the transmitted bit of each user. For each user, the receiver performs a simple hypothesis test on the received statistic corresponding to that user. This hypothesis test corresponds to that between a central and a noncentral chi-square statistic, or to that between two noncentral chi-square statistics depending on the value of $c$. At high SNRs, the threshold corresponding to the required hypothesis test for user $k$ in the $y-$domain is given as

$$\tau_k = \alpha^2 \frac{(A + 2c_k')^2}{4}$$  \hspace{1cm} (5.9)

where $c_k' = c \sum_{j=1}^{K}[R]_{k,j}$. Note that the thresholds and hence the entire precoder transformation has to be computed at the receiver. The AME based on the probability of error is computed next.

The average probability of error for user $k$ can be obtained as

$$P_e = \frac{1}{2} Pr(y_k > \tau_k|b_k = 0) + \frac{1}{2} Pr(y_k \leq \tau_k|b_k = 1)$$  \hspace{1cm} (5.10)

where $y_k$ is the signal received for channel $k$ after photodetection and $\tau_k$ is as given in (5.9). Since the conditional distribution of $y_k$ given $b_k$ is chi-squared with centrality parameters $\alpha(\mathbf{A}b_k + c_k')$, we have [51]

$$P_e = \frac{1}{2} Q_1 \left( \frac{c_k'}{\sigma}, \frac{\sqrt{\tau_k}}{\sigma} \right) + \frac{1}{2} Q_1 \left( \frac{A + c_k'}{\sigma}, \frac{\sqrt{\tau_k}}{\sigma} \right)$$  \hspace{1cm} (5.11)
where $Q_1$ is the first order Marcum-Q function. We use the following inequalities [73]:

$$Q_1(a, b) \leq \exp\left(-\frac{(a-b)^2}{2}\right)$$

\begin{equation}
Q_1(a, b) \leq \exp\left(-\frac{(a-b)^2}{2}\right), \quad b > a \geq 0, \quad (5.12)
\end{equation}

$$1 - Q_1(a, b) \leq \frac{1}{2} \left( \exp\left(-\frac{(a-b)^2}{2}\right) - \exp\left(-\frac{(a+b)^2}{2}\right) \right)$$

\begin{equation}
1 - Q_1(a, b) \leq \frac{1}{2} \left( \exp\left(-\frac{(a-b)^2}{2}\right) - \exp\left(-\frac{(a+b)^2}{2}\right) \right), \quad b > a \geq 0. \quad (5.13)
\end{equation}

and the fact that

$$0 \leq \alpha c_k' \leq \frac{\sqrt{\tau_k}}{\sigma} < \alpha \frac{A + c_k'}{\sigma}$$

\begin{equation}
0 \leq \alpha c_k' \leq \frac{\sqrt{\tau_k}}{\sigma} < \alpha \frac{A + c_k'}{\sigma} \quad (5.14)
\end{equation}

to upper bound the probability of error as

$$P_e \leq \frac{3}{4} \exp\left(-\frac{\alpha^2 A^2}{8\sigma^2}\right) - \frac{1}{4} \exp\left(-\frac{\alpha^2(3A + 4c_k')^2}{8\sigma^2}\right) \leq \exp\left(-\frac{\alpha^2 A^2}{8\sigma^2}\right). \quad (5.15)$$

Using $\lim_{z \to \infty} Q(z) \exp(-z^2/2) = 1$ the AME can be approximated as [27]

$$\eta_{ZF} \approx \sup \left\{ 0 < t : \lim_{\sigma \to 0} \frac{P_e(\sigma)}{\exp\left(-tA^2\right)} < \infty \right\}$$

\begin{equation}
\eta_{ZF} \approx \sup \left\{ 0 < t : \lim_{\sigma \to 0} \frac{P_e(\sigma)}{\exp\left(-tA^2\right)} < \infty \right\} \quad (5.16)
\end{equation}

and then, using (5.15), we can lower bound the AME of linear ZF-precoding as

$$\eta_{ZF} \geq \alpha^2. \quad (5.17)$$

The fact that the bound is independent of the constant $c_k'$ indicates that it is identical for all users. Since in the last inequality of (5.15), all the terms apart from $\alpha$ are independent of the correlation term $\rho$, this fact will hold even for nonequal values of
5.4.2 $\rho$-Forcing Precoding

With $\rho$-forcing precoding the receiver can benefit from employing MUD techniques. Although low-complexity MUD proposed in Chapter 4 can be used, for analytical purposes we assume that an AO MUD is employed at the receiver. In spite of the $\rho$ forcing nature of the precoder, the received constellation being affected by the constant $c$ does not completely match the targeted constellation with correlation $\rho$. The AO detector used has to be obtained as given in (4.18) and the AME is also computed in a manner similar to that for the AO MUD in [27], but with the effective received signals corresponding to that with $\rho_F$-precoding.

Fig. 5.1 shows the AME bound for a 2–user system employing ZF-precoding with SUD and the performance of $\rho_F$-precoding with $\rho_d = 0.2$ and AO MUD. The performance of AO MUD and SUD without precoding is also plotted for comparison. The figure indicates that at least one of the joint precoding strategies outperforms receiver based MUD strategies when the correlation is in the range $(-0.5, 0) \cup (0.5, 1)$. For remaining values of correlation, receiver based MUD strategies should suffice in minimizing interference. These results imply that to optimize the system performance, depending on the estimated correlation, either a suitable joint precoding strategy or simple receiver based MUD should be chosen.

5.5 Estimating the Correlation Coefficients

In this section we focus on estimating the correlation coefficients which are required at both the transmitter and the receiver. We restrict ourselves to the case of real correlations. The correlation can be estimated either blindly or with training. Training
sequences can be used for the estimation process at the expense of added bandwidth. For simplicity we focus only on two-user systems. With two users, employing a training sequence of ones being transmitted for both users, the maximum likelihood (ML) estimate gives the best possible estimate of the coefficients. The likelihood of the coefficient $\rho$ can be obtained as $l(\rho) = L(y|b = 1)$ where $L$ is as given in (4.9). To obtain the ML estimate we approximate the Bessel function $I_0(z)$ in the likelihood function as

$$I_0(z) = \frac{e^z}{\sqrt{2\pi z}}$$  \hspace{1cm} (5.18)
The resulting ML estimate, $\hat{\rho}_{ML}$ is given as

$$\hat{\rho}_{ML} = \frac{\bar{x}}{2} + \frac{\sqrt{\bar{x}^2 - 2\sigma^2}}{2} - 1,$$  \hspace{2cm} (5.19)

where $\bar{x}$ is the sample average of the received statistics in the $x$–domain. In case there are more than two users the correlations can be obtained using appropriate training sequences. For estimating the correlation between any two users, a bit-vector with both the user bits set to one, and all other user bits set to zero is transmitted.

In the previous ML estimation procedure, training sequences are used. Blind estimation using statistics of the received signal is complicated by the presence of quadratic terms when the correlation among users is not identical. However, when the correlation is identical, using first order received statistics, the correlation can be obtained blindly, thus eliminating the need to send a training sequence. All the received statistics for all users then have an identical distribution that can be obtained as a mixture of the distribution in (4.10) over all $b$ vectors. The mean value of each of these statistics is then related to $\rho$. The blind estimate $\hat{\rho}_b$, can be computed by using the statistical average of the received signal for any user, say $\tilde{y}$, as

$$\hat{\rho}_b = \frac{1}{K} \sqrt{1 - 2K \frac{1 + 4\sigma^2 - 2\tilde{y}_1}{K - 1}},$$  \hspace{2cm} (5.20)

where $\sigma^2$ is the variance of the complex noise term $n_k$.

Obtaining a blind estimate when the correlations are not equal is not possible by using a similar procedure. However, without training, the mixture distributions of the received vector signals can be obtained using (4.9) and (4.10). This distribution gives the joint likelihood function for the various correlation coefficients which can then be used to obtain the estimates. This problem is left for future research since our
focus is to understand the impact of estimation errors on the precoding algorithm. Additionally, for systems with nonlinear interference estimation is not addressed here. The topic of estimating coefficients for a nonlinear wireless systems is addressed in [74] and some the results there could help develop further insight into estimation algorithms for optical nonlinear systems. Addressing the estimation problems in more detail is beyond the scope of this thesis and left for future research.

For the ML estimation and blind estimation algorithms considered above, the mean square error (MSE) in the estimation process is a function of the signal to noise ratio (SNR) measured as \( \frac{\sigma^2}{\sigma^2} \), and the length of training sequences or samples used for blind estimation. Fig. 5.2 shows the MSE for a two user system for two different SNRs as the number of samples used (equivalent to the length of training sequence used) varies. As expected, training based ML estimation is superior to the blind technique and achieves MSE which is lower by at least one order of magnitude. Additionally, with blind estimation, the MSE saturates quickly and to a higher value as compared to the training based approach. However, with sufficiently large number of samples used, the blind estimate may have a relative error of less than 0.1%. Simulation results presented in the following section indicate that blind estimation does not degrade the performance of precoded systems severely.

5.6 Simulation Results

Simulation results for a two-user system are presented here. The bit-error-rates (BER) from different schemes are compared assuming that the correlation is perfectly known. With \( \rho \)F-precoding the regression based Quadratic-\( y \) detector proposed in Chapter 4 is employed. The detector coefficients are chosen not only for the value of \( \rho \) being forced but also to incorporate the effect of the power scaling and positivity
constants. Fig. 5.3 shows the simulated BER when the effective correlation due to channel effects is $\rho = -0.4$ and $\rho = 0.4$ respectively. For $\rho F$-precoding the value of the desired correlation is chosen to be $\rho_d = 0.2$. Simulation results indicate that $\rho F$-precoding clearly outperforms strategies which do not employ any kind of precoding. MUD based strategies with and without $\rho F$-precoding outperform ZF-precoding when the correlation is positive; however, with negative correlation ZF-precoding is better. This corroborates with the AME plot in Fig. 5.1. Clearly $\rho F$-precoding with MUD is the best option to guarantee energy efficiency.

The BER plots in Fig. 5.3 are based on the assumption that the correlation is perfectly known when designing precoders. Fig. 5.4 shows the BER performance of ZF-precoding and $\rho F$-precoding (with the Quadratic–$\gamma$ detector), for two SNR levels.
Figure 5.3: SUD and regression based Quadratic−y MUD without precoding are compared with different types of joint precoding strategies, when the correlation is assumed to be 0.4 or −0.4 for a two-user system. Performance of single user systems is plotted for comparison.
Figure 5.4: Achievable BER of ZF precoding and ρF-precoding with Quadratic—y MUD compared with SUD without precoding for different cross-correlations. The effect of perfect and imperfect correlation estimates is shown. The estimates are generated using 128 samples for both Blind and ML approaches.
and varying correlations. The effect of imperfect correlation estimation is shown. The figures indicate that estimation of $\rho$ via blind approaches does not cause significant deviation from the performance achieved with perfect estimation, at least for the two-user system considered. The figures also indicate that at all values of correlation in the range $(-0.4, 0.4)$, ZF-precoding outperforms SUD. $\rho F$-precoding, although combined with MUD, does not always outperform SUD, especially in some of the regions where ZF-precoding outperforms SUD. For $\rho \in (0.15, 0.4)$, $\rho F$-precoding appears to outperform ZF-precoding as well as SUD. These observations are corroborated by the AME plot in Fig. 5.1. In some of the region where the ZF-precoding scheme outperforms the SUD and $\rho F$-precoding schemes, simple optimal MUD schemes outperform ZF-precoding. The results imply that depending on the estimated correlation either a suitable joint precoding strategy or simple receiver based MUD should be chosen.

5.7 Conclusion

For multichannel ACI-dominant fiber optic communication systems, knowledge of correlation at the transmitter can be used to jointly preprocess the signal to be transmitted and minimize the cross-talk/interference. We propose two such preprocessing strategies, namely ZF-precoding and $\rho F$-precoding and compare their performance to that of receiver based interference cancelation techniques without precoding. In spite of sub-optimal approaches taken to the precoding process, these strategies clearly outperform systems employing simple SUD receivers and in some cases even the AO MUD based systems without precoding. Interestingly, $\rho F$-precoding which combines transmitter preprocessing along with a complicated receiver based signal processing approach in the form of an AO MUD, does not always outperform ZF-precoding, which employs a simple SUD at the receiver. Our simulation results indicate that
the performance is not highly sensitive to the accuracy of the correlation estimate which is essential to the strategy. Results indicate that overall system performance could be improved by using an adaptive pre and post processing approach wherein ZF-precoding, ρF-precoding scheme with suitable MUD, or simple MUD is chosen based on the estimated correlations. Future research should focus on estimating complex coefficients from the post-detection signal, applying noisy channel simulations to test estimation algorithms, and finally simulating the signal designs using the SSF method in Chapter 2.
In the previous chapter, signal design via precoding based interference pre-cancelation schemes for ACI-dominant systems is investigated. Precoding is a special case of signal design where the transformation operation on the transmitted bit-vectors is linear. Nonlinear channel models complicate the signal design for XPM-dominant and FWM-dominant systems. In this chapter, heuristic signal design approaches for these systems are investigated.

The channel behavior for both XPM-dominant and FWM-dominant systems is signal dependent, i.e., the correlation matrix is a function of the actual transmitted bit-vector. Therefore, accurate channel estimation plays a crucial role in precoding for such systems. As such, for systems with rapidly varying transmitter phases, signal design schemes for interference pre-cancelation become irrelevant. However, ongoing research on optical transmitters that have a stable frequency of operation and enable phase modulation, as given in [39] and references therein, motivates the research for relevant precoding algorithms.

For ACI-dominant systems, when the interfering channel bit is nonzero cross-talk
is experienced by all channels alike. However, for XPM-dominant systems, the cross-talk is experienced only by channels that are nonzero themselves. For FWM-dominant systems, on the other hand, cross-talk occurs when more than one channel has nonzero transmission and when the so-called phase-matching condition is satisfied [48]. For both XPM and FWM-dominant systems, the cross-talk is thus signal dependent and cannot be characterized in terms of a signal independent matrix. This consequence of nonlinearity prohibits a linear preprocessing approach for such systems.

Similar to signal design for ACI-dominant systems, the ideal transmitted signals can be chosen to result in complete cross-talk mitigation as in zero forcing (ZF) precoding, or to force a particular correlation as in $\rho\text{-F-precoding}$. However, results in Chapter 4 do not provide a specific value of $\rho$ for which nonlinear interference systems can perform better than ZF. Thus, here the focus is on ZF signal design for XPM-dominant systems. It is shown in this chapter that the ZF approach is not feasible for FWM systems. As such, a joint transmit-receive processing approach involving signal design and multiuser detection (MUD) is required. Assuming an asymptotically optimal (AO) MUD at the receiver, a heuristic algorithm to design signals is proposed and investigated. Using asymptotic multiuser efficiency (AME) as a measure, the gains achieved by the signal design schemes over traditional designs with and without MUD based interference rejection are compared. Our results indicate that significant gains are possible with the proposed approaches when compared with OOK based systems with and without MUD based receivers.

The rest of the chapter is organized as follows: Section 6.1 describes the system model where the interference pattern is briefly reviewed. Proposed signal design algorithms for XPM-dominant systems are given and analyzed in Section 6.2. Signal design to combat FWM effects is discussed in Section 6.3. Conclusions and future research directions are presented in Section 6.4.
Chapter 6. Signal Design for Nonlinear Interference Limited WDM Systems

6.1 System Model

The system models considered here are motivated by the continuous wave (CW) system in Chapter 3. When nonlinear effects are significant, the interference terms can be decomposed into XPM and FWM terms among other terms. Here, the FWM terms for the CW system in Chapter 3 are employed. The XPM terms present are found to be identical for CW and practical systems in Chapter 3, and used here.

As in Chapter 4, the sampled signal at the output of the $k$th photodetector at the receiver is given as

\[ y_k = |b_k A_k + \rho_k(b) + n_k|^2, \quad k = 1, \ldots, K, \]

where $b = [b_1, \ldots, b_K]^T$ represents the $K$ transmitted bits, $\rho_k(b)$ represents the channel induced nonlinear interference as experienced by user $k$, and $n_k$ represents the complex Gaussian noise term with variance $\sigma^2/2$ per complex dimension. When XPM dominates, the nonlinear interference term is given as [26],

\[ \rho_k(b) = \sum_{i \neq k} \rho_{i,k} b_i b_k |A_i|^2 A_k, \]

where $\rho_{i,k}$ represents the correlation resulting from cross-talk between users $i$ and $k$.

For the FWM case we focus on the CW system with

\[ \rho_k(b) = \sum_{l-m+n=k} \rho_{k,l,m,n} b_l b_m b_n A_l A_m^* A_n, \]

where $\rho_{k,l,m,n}$ represents the correlation resulting from cross-talk between users $l, m, n$ and user $k$. The correlation coefficients $\rho_{k,l,m,n}$ and $\rho_{k,l}$ are typically complex and can be estimated at the receiver and fed back to the transmitter. In this work it is assumed
that these are perfectly known at the transmitter. With signal design, it is possible that on-off keying (OOK) is no longer employed. In such cases, the amplitude $A_k$ in (6.1) and (6.3) may be a function of $b_k$.

In vector notation the received signal when XPM effects dominate is given as

$$y = |R_bb + n|^2_o$$

where the signal dependent correlation matrix

$$[R_b]_{i,j} = \begin{cases} 
\rho_{i,j}|A_i|^2A_j, & b_i = 1, i \neq j \\
0, & b_i = 0 \\
A_i, & i = j.
\end{cases}$$

and $|\cdot|^2_o$ represents the Schur product of a vector with itself. Note that the interference is referred to as nonlinear since the interference terms do not depend linearly on the bit-vector $b$. For FWM-dominant systems, the received vector cannot be written in such a compact form.

### 6.2 Signal Design for Cross-Phase Modulation Effects

Zero-forcing (ZF) requires choosing transmitted signals such that the received signals appear interference free or uncorrelated at the demodulator output. Such designs for XPM-dominant systems are considered here.

For systems with linear interference studied in Chapter 5, an affine precoding transformation, consisting of a precoding matrix $M$, a power scaling constant $\alpha$ and a
Chapter 6. Signal Design for Nonlinear Interference Limited WDM Systems

positivity constant $c$, is designed to operate on the bit-vectors to be transmitted. The design involves choosing the precoding matrix $M$, which was then employed to choose the power scaling and positivity constants. In the case of XPM, as indicated by (6.5), it turns out that the cross-correlation matrix depends on both the transmitted bit-vector as well as the amplitude levels. If affine precoding is used on XPM-dominant systems, the received signal with precoding becomes

$$|\alpha R_{M,b}(Mb + c) + n|^2,$$

where

$$[R_{M,b}]_{i,j} = \begin{cases} 
\rho_{i,j} \alpha^2 [Mb + c]_i [Mb + c]_j & i \neq j \\
\alpha [Mb + c]_i & i = j 
\end{cases}$$

The subscript $M$ indicates that $R$ depends on the transmitted bit as well as the precoding matrix $M$ (along with $\alpha$ and $c$). Choosing a ZF-precoder then corresponds to solving the problem:

$$M = \arg \min_{\tilde{M},\tilde{c},\tilde{\alpha}} E \left\{ \left\| \left( Ab - \tilde{\alpha} R_{M,b}(\tilde{Mb} + \tilde{c}) + n \right) \right\| \right\}.$$

No solution to the above problem is known. Additionally, due to the dependency of the cross-correlation matrix on the transmitted bits, no single affine transformation that enforces ZF at the receiver exists.

The formulation above indicates that a separate transformation needs to be chosen for each transmit bit-vector. Consequently, the constants $\alpha$ and $c$ are no longer necessary, and the design problem can be reduced to solving the following set of $K$
nonlinear equations

\[ |A_k b_k + \sum_{j \neq k} \rho_{k,j} |A_k|^2 b_k b_j| = |A_k b_k| \quad \forall k \in \{1, \ldots, K\}. \]  

(6.7)

where \( A_k = [M]_{k,k}, \forall k \in \{1, \ldots, K\} \). Considering only real solutions leads to an exactly determined system. For such cases, the inverse function theorem [75] can be invoked to conclude that if a solution exists then it is locally unique. However, the existence of a real as well as positive solution is not guaranteed. Allowing for complex solutions that require phase control at the transmitter results in an under-determined system. Typically, infinite solutions exist for such systems and iterative methods are used to obtain solutions. Such techniques are beyond the scope of this discussion. When multiple feasible solutions exist, the least energy solution should be chosen. It is easy to see that when \( b_k = 0 \), \( A_k = 0 \) is to be chosen. This simplifies the number of equations to be solved for each bit vector. When the correlations \( \rho_{i,j} \) are equal, it turns out that the same amplitude level is to be chosen for all the users that transmit a one. Such scenarios, which are useful from an analytic perspective, result in a single nonlinear equation which can be solved numerically. Choosing a finite \( A \) results in finite average and peak transmit power. The author would like to note that this is not a typical power control where a single power level is chosen for all the transmitted bits. A separate transmit power is chosen for each individual bit-vector and, in general, for each channel in the bit-vector. The following section provides insight into the performance of the proposed signal design method compared with the MUD based methods.
6.2.1 Performance Analysis of Signal Design for Cross-Phase Modulation Effects

The performance of the signal design for XPM-dominant systems is compared to the standard OOK scheme that transmits \((0, A)\) and employs AO MUD or single user detection (SUD) at the receiver. Like in previous chapters, AME is used as a performance measure. For simplicity it is assumed that the correlations are identical and real, i.e., \(\rho_{j,k} = \rho\) and known perfectly at the transmitter. The probability of error for the signal design scheme is identical to that of a single user system employing levels 0 and \(A\), but the transmitted power is clearly different for each bit-vector. The average transmit power with precoding can be computed easily using the \(A_i\) obtained from (6.7) and used to compute the AME as

\[
\eta = \frac{KA^2}{2^{K+1} \sum_b \sum_k |A_k(b)|^2 b_k}.
\]  

(6.8)

Here \(KA^2/2\) is the average power consumed by \(K\) perfectly orthogonal channels to achieve an identical error probability.

Fig. 6.1 compares the achievable AME of the proposed signal design with the AO MUD and the SUD based schemes without precoding for two- and three-user XPM-dominant systems. Typical values of correlation \(\rho\) are small and, as was indicated in Chapter 3, scaled with the transmit power \(A^2\). Fig. 6.1 shown the AME when signal design solutions for transmit amplitudes are constrained to be real and positive. Motivated by recent developments in controlling laser phases, Fig. 6.2 considers designs where real positive as well as real negative solutions (corresponding to a phase shift of \(\pi\)) are allowed. The figures indicate that:

1. For XPM-dominant systems, when the correlation is positive, signal design can
provide significant gains when compared to standard OOK based systems with multiuser processing. Comparing Fig. 6.1 and Fig. 6.2 indicates that allowing for phase modulation results in better performance when the channel induced correlation is negative. This indicates that, in spite of phase insensitive reception, gains are achievable with phase controlled modulation schemes. This is not surprising given the fact that phase modulation to amplitude modulation conversion takes place at the receiver front-end and the cross-talk effects themselves are phase dependent.

2. For XPM-dominant systems, MUD based alternatives can sometimes outperform the signal design based solution. However, it should be noted that the
signal design algorithms considered are not claimed to be optimal in any sense.

3. It should be noted that the range of values of \( \rho A^2 \) that are realized may in practice be smaller than indicated in the plot. Even for such cases gains can be realized when the cross-correlations are positive.

The signal design approach pursued here is not complicated since ZF is feasible. A more complex approach to signal design is required for FWM-dominant systems and is presented in the following section.
6.3 Signal Design for Four-Wave Mixing Effects

Using an approach similar to XPM, it can be shown that linear ZF is not feasible for FWM-dominant systems and a separate set of amplitudes needs to be chosen for each transmitted vector. However, unlike XPM, ZF in any form is not possible for FWM-dominant systems. Let

\[ A = \{ [A_1(b), \ldots, A_K(b)]^T : b \in \{0, 1\}^K \}, \]

(6.9)
denote the transmit constellation, or the set of transmitted amplitudes corresponding to the \(2^K\) bit vectors \(b\). Similar to the XPM case, choosing the amplitudes for a particular bit-vector \(b\) corresponds to solving simultaneous equations of the form

\[ |A_k b_k + \sum_{j+l-i=k} \rho A_j A_i^* A_l b_j b_l b_i|^2 = |A b_k|^2 \quad \forall k. \]

(6.10)

This problem does not have a solution for all the possible sets of \(b\), for example when \(b_k = 0\) but \(b_i = b_j = b_l = 1\). Thus unlike XPM-dominant systems, interference mitigation is not possible by ZF signal design alone and joint-transmit receive processing is required. A more general signal design problem needs to be devised which is discussed in the following section.

6.3.1 Signal Design Problem for Nonlinear Channels

A MUD has to be employed at the receiver. The AME of the system changes with the MUD as well as the signal set chosen. Assuming that an AO MUD is being employed at the receiver, the transmitted signals are designed here so as to maximize the overall energy efficiency. The corresponding problem set-up is not specific to FWM-dominant systems and can be applied to any system with nonlinear channel
effects. Note that the solutions depend on the detector employed at the receiver and the research in this section is relevant only for the AO MUD.

Let

\[ E(A) = \frac{2^{-K}}{K} \sum_b \sum_i |A_i(b)|^2; \]  

(6.11)
denote the average transmit energy for a chosen transmit constellation. Let \( \mu_A(b_i) \) denote the \( x \)-domain received noiseless signal (as in Chapter 4) for the \( i \)th transmit vector \( b_i, i \in \{1, \ldots, 2^K\} \), where the subscript \( A \) denotes the transmit constellation employed. It can be computed using (6.1) and (6.3) letting \( n_k = 0 \). The AME of the system when employing the AO MUD is directly proportional to the Euclidean distance between the noise-free received signals, \( \{\mu_A(b_i), i \in \{1, \ldots, 2^K\}\} \), and the average energy of the signal set employed. The AME for a particular transmit constellation is:

\[ \eta_A = \min_{i \neq j} \frac{||\mu_A(b_i) - \mu_A(b_j)||^2}{2E(A)}, \]  

(6.12)

This expression is obtained from the definition of AME as given in (4.11) and observing that the distance achieved by the signal design can be achieved for the single user linear system case by using a constellation of average energy that is half of the same distance.

To design the constellation \( A \) with minimum energy that satisfies a given effective
received distance, say $\delta_{\text{min}}$ we devise the optimization problem as

$$\hat{A}(\delta_{\text{min}}) = \arg \min_{A} E(A)$$

s.t.

$$||\mu_A(b_i) - \mu_A(b_j)|| \geq \delta_{\text{min}} \quad \forall i, j, \quad i \neq j.$$ (6.13)

The objective function $E(A)$ in (6.13) is a convex function. The function $\mu_A$ is neither concave nor convex except for special cases. Hence the constraint set over the search space for $A$, $\{A : \min_{i \neq j} ||\mu_A(b_i) - \mu_A(b_j)|| \geq \delta_{\text{min}}\}$ in (6.13) cannot be shown to be convex. In the following section an iterative algorithm to solve (6.13) is proposed.

### 6.3.2 Signal Design Algorithm for FWM-dominant Systems

One possible design is obtained by choosing the effective received constellation to be a bounded lattice and then searching and choosing the corresponding transmit constellation, if it exists, so as to minimize the effective transmit energy. This approach is similar to the ZF approaches for XPM-dominant and ACI-dominant, which can be viewed as a design with lattice points on a bounded rectangular lattice in multidimensional space. The AO detectors for these cases is equivalent to the SUD detector. For FWM systems, choosing the received constellation to be a bounded rectangular lattice with points on a scaled unit grid as was done via ZF is not feasible. A computational approach is required to determine both an effective received constellation (that may end up being a bounded lattice) and then obtaining the transmit constellation. Solving the multidimensional optimization problem in (6.13) can be computationally intensive given the number of variables to choose, and the complexity of verifying the constraints. A computational that requires searching a smaller space iterative to solve (6.13) is given here.
For a given $\delta_{\text{min}}$, there are two parts of the problem. First is to identify the feasible region. This is problematic given the fact that the constraint set is not convex. Since we choose $\delta_{\text{min}}$ to correspond to the distance achieved by the OOK method, we already have a starting feasible point. Starting with this point we use a gradient based search in an arbitrary region surrounding it and solve for (6.13). Since the objective function is convex, a local optimal point is achieved. Let $\tilde{A}$ denote this point. To further optimize from this local point we propose the following heuristic iterative algorithm. Given a locally optimal point $\tilde{A}$, we use it as a starting point to obtain another feasible point, say $\tilde{A}$, that lies in the sphere whose boundary passes through $\tilde{A}$. For this, we solve the following optimization problem with $\mathcal{E} = \mathbb{E}(\tilde{A})$, corresponding to the energy of the local optimum $\tilde{A}$,

\[
\tilde{A}(\mathcal{E}) = \arg \max_{\tilde{A}} \min_{\substack{i \neq j \\ i,j \in \{1,\ldots,2^K\}}} \|\mu_{\tilde{A}}(b_i) - \mu_{\tilde{A}}(b_j)\|
\]

\[
s.t. \quad \mathbb{E}(\tilde{A}) \leq \mathcal{E}. \tag{6.14}
\]

This optimization problem starts with a feasible point for (6.13) and the maximization problem is devised such that the resulting solution is also a feasible point for (6.13). Additionally, since the objective function is not concave, the new feasible point, $\tilde{A}$ may lie in the interior of the sphere of radius $\mathbb{E}(\tilde{A})$ and thus has a lower energy since the objective function is not concave. Using the solution $\tilde{A}$ as a starting point, we again solve for (6.13) to obtain a new $\tilde{A}$ which is then used to find another $\tilde{A}$ by solving 6.14 using the new $\tilde{A}$ for the energy constraint and as a starting point. The algorithm below gives the exact steps in details. To further reduce the complexity of the search algorithm, the iterative algorithm searches a $K$-dimensional space iteratively at each
Chapter 6. Signal Design for Nonlinear Interference Limited WDM Systems

...step as opposed to searching in a $K \times 2^K$ dimensional space. The underlying idea is to choose one of the $2^K$ vectors of the constellation at a time.

0. Initialization Step: Obtain an initial feasible point $\tilde{A}$ and $\delta_{\text{min}}$ corresponding to the OOK system with AO MUD.

- Choose the amplitude level, say $A$, and find the corresponding received signal by applying the model in (6.1) and (6.3) and obtain $\delta_{\text{min}}$ as the minimum effective received distance.
- Let $\tilde{A} = \{[\tilde{A}_1(b_i), \ldots, \tilde{A}_K(b_i)]^T = A b_i, \forall i = \{1, \ldots, 2^K\}\}.$

1. Step I: Solve (6.13) starting with $\tilde{A}$ as follows to obtain a locally optimal point.

For $i \in \{2, \ldots, 2^K\}$, let $S_i$ denote a $K-$ dimensional cube or arbitrary width centered around $[\tilde{A}_1(b_i), \ldots, \tilde{A}_K(b_i)]^T$ and solve:

$$[\tilde{A}_1(b_i), \ldots, \tilde{A}_K(b_i)]^T = \arg \min_{a \in S_i} ||a||$$

$$s.t.$$

$$\min_{j \in \{1, \ldots, i-1\}} ||\mu_{\tilde{A}}(b_i) - \mu_{\tilde{A}}(b_j)|| \geq \delta_{\text{min}}$$

(6.15)

to obtain a lowest energy solution about $\tilde{A}$.

2. Step II: Starting with $\tilde{A}$ solve (6.14). For $i \in \{2, \ldots, 2^K\}$ solve:

$$[\tilde{A}_1(b_i), \ldots, \tilde{A}_K(b_i)]^T = \arg \max_{a \in \mathbb{R}^K} \min_{j \in \{1, \ldots, i-1\}} ||\mu_{\tilde{A}}(b_i) - \mu_{\tilde{A}}(b_j)||$$

$$s.t.$$

$$||a||^2 \leq K 2^K \mathbb{E}(\tilde{A})$$

(6.16)

The energy constraint on $a$ ensures that the effective signal design has energy $\mathbb{E}(\tilde{A}) \leq \mathbb{E}(\tilde{A}).$
3. Iterate between Step I and Step II till convergence is attained for Step I.

The convergence of this algorithm is not formally proven here. The fact that the solution obtained in Step I is used for the energy bound for Step II, and the solution to Step II always acts as an upper bound for the constraint in Step I, ensures that the algorithm converges, but not necessarily to the global optimum. In the following section the AME achieved with this signal design algorithm is computed.

**6.3.3 Performance of Signal Design Algorithm for FWM-dominant Systems**

Similar to XPM-dominant systems, for illustration purposes it is assumed that the correlations among users are equal, i.e., $\rho_{j,k,l} = \rho$ for FWM-dominant systems. Again the performance is not insensitive to the amplitude or the total power since the constellation choice does not scale with $A$, and hence the average transmit power. The correlations are scaled with the power as was done for XPM-dominant systems.

Also, comparing the scheme with receiver based schemes requires a suitable choice of $\delta_{\text{min}}$ when solving (6.13). This is ensured in the proposed algorithm by choosing $\delta_{\text{min}}$ to correspond to that obtained in an OOK system. Fig. 6.3 shows the AME when the signal design algorithm as proposed above is employed in conjunction with AO MUD, OOK system is used with AO MUD, and OOK systems is used with SUD, for a four-user FWM-dominant system. The scaled correlations are considered to lie in the interval $(-0.25, 0.25)$. The performance of the scheme obtained after 1, 2, 3 and 4 iterations is shown.

The figure indicates that

1. For the four-user FWM-dominant system, significant gains can be achieved by the signal design algorithm considered when the correlations are negative.
2. The gains from signal design reduce with increasing correlation and for positive correlations when compared to the AO MUD. In fact, when the correlations are positive, it is possible that the OOK method with AO MUD is locally optimal.

3. The proposed algorithm converges, and converges quickly. For positive correlations few gains are available from iteration. However, for negative correlations a few iterations can help maximize the performance. For the system considered only four iterations were needed for most of the values of $\rho A^2$ considered. As expected the AME does not reduce with iterations as expected.

4. It should be noted that the range of values of $\rho A^2$ that are realized in practice may be smaller than indicated in the plot. Even for such cases gains can be
realized when the correlation is negative.

Similar to XPM-dominant systems, the performance depends on the correlations realized. Even when the correlations are small and negative, significant gains are possible, motivating the application of the method. The following section summarizes the results and provides directions for future investigations in this area.

6.4 Summary and Conclusion

Signal designs have been developed for nonlinear interference limited fiber optic channels. ZF is feasible for XPM-dominant systems and a ZF algorithm is proposed. The performance of the algorithm for systems with and without laser phase control is considered by relaxing the positivity constraint on the transmitted signals. Results indicate that constraining the inputs to be real and positive when the correlation is negative causes no benefits from signal design and simple AO MUD should be employed. For positive correlation and when phase control is available, ZF signal design for XPM-dominant systems can provide significant gains. ZF is infeasible for FWM-dominant systems and a novel iterative algorithm is proposed and applied to a four-user FWM dominant system. Results indicate that significant gains are available from designing signals using the proposed algorithm.

The channel induced interference is signal dependent, and any solution for signal design has complexity that is exponential in the number of users. The somewhat heuristic designs presented indicate that significant performance gaps may exist when present day modulation schemes for multichannel fiber optic systems are subjected to channel induced cross-correlation, thereby motivating the search for optimal signal design solutions. The author would also like to note that FWM effects may not occur in isolation. However, the algorithm presented here is generic and applicable to all
forms of cross-talk. Future research should focus on applying the algorithm to a system with the full third order model presented in Chapter 3. Further directions are provided in Chapter 7.
Signal processing approaches for mitigating cross-talk effects, arising from fiber non-linearity and leakages in optical devices, in a multispans, wavelength division multiplexed (WDM) fiber optic communication system are studied in this dissertation. The effects of nonlinear fiber response on the discrete time communication channel are characterized and modeled. Low complexity receiver based, transmitter based, and joint transmitter-receiver based signal processing approaches for mitigating cross-talk are proposed and shown to outperform conventional methods. A summary of the conclusions and future research directions are provided in this chapter.

7.1 Effects of Fiber Nonlinearity

The effects of fiber nonlinearity for multiple transmitting power levels, distances, and channel spacings of 25 GHz and 50 GHz are studied for two different types of receiver structures in Chapter 2 via simulations. The key result is that through reducing channel spacing and increasing power/distances, the eye-openings seen at
Chapter 7. Conclusion and Future Research Directions

the receiver decrease rapidly due to nonlinearity. The performance of the practical receiver structure was found to be almost identical to that of the idealized optical matched filtering based receiver. The results are obtained via simulation. Results with practical experimentation for the said spacings and different receiver structures could form a topic of future research.

7.2 Modeling Fiber Nonlinearity in Discrete-time

Discrete-time models for multispans WDM systems are obtained to model nonlinear fiber behavior in Chapter 3. A least squares regression approach is applied to pre-detection data for verifying the accuracy of the models. Low complexity models that require fewer terms to characterize the input-output relationship are compared with a full third order model for accuracy. The first order low-complexity adjacent-channel interference (ACI) and the third order low-complexity cross-phase modulation (XPM) models for cross-talk are found to be suitable for modeling systems with up to five spans and lower power levels below 0.75 mW/channel for a system with eight channels. For higher power levels as the spans increase, the full third order model with ACI, XPM and four-wave mixing (FWM) cross-talk terms is found insufficient. Higher order models are required for such systems. Channel estimation for the ACI model is considered in Chapter 5 but only for correlations that are constrained to be real. It is found that the blind approach to estimation while not as good as the training based approach to estimation, does not affect the signal designs considered later significantly.

The estimators are restricted to estimating coefficients that are real for the ACI model. Future research should focus on obtaining channel estimators for the nonlinear models proposed in the chapter and address the problem of estimating complex
correlations from post-detection data. An alternate approach employing nonlinear least squares on the post-detection data should also be investigated for model validation. The modeling and estimation problem should itself be considered under the framework of nonlinear system identification problems that have been addressed by control theorists.

7.3 Low-complexity Multiuser Detection

The design of polynomial complexity detectors is investigated and results are applied to affine and quadratic detectors. The high SNR performance of detectors measured using asymptotic multiuser efficiency is found to be nonconcave in the detector coefficient for all affine and higher order polynomial detectors. A new regression-based multiuser detection algorithm is proposed. The key contribution of the regression approach proposed is the choice of points about which to approximate the optimal detector. The affine and quadratic detectors obtained using the proposed approach are found to have performance that is close to the locally optimal detectors and always outperforms the conventional SUD and AMMSE detectors.

Future research could focus on applying the proposed detection algorithm to other communication channels. Analysis of the regression based approach requires some additional investigation. Advanced nonlinear programming methods should also be applied to the detector design problems with a view of further reducing the complexity of obtaining the detectors.
7.4 Signal Design for Mitigating Cross-talk

Signal design approaches for mitigating linear cross-talk arising from ACI, and non-linear cross-talk arising from XPM or FWM is investigated. A linear precoding based approach is found sufficient for zero-forcing (ZF) the cross-talk in ACI systems. The precoding algorithm itself is found to be insensitive to channel estimation errors. ZF to mitigate XPM requires solving a nonlinear system of equations to choose the signal set to be transmitted. The approach may not always give benefits over receiver based MUD schemes unless phase control is possible at the transmitter.

ZF is not feasible for mitigating FWM effects. A novel iterative algorithm is proposed to mitigate FWM effects. The algorithm requires using a MUD tailored to the signal design but the design guarantees that it outperforms simple on-off keying based signal design with receiver based MUD. The algorithm proposed for FWM is generic and can be applied to arbitrary systems.

Future research should focus on applying the FWM algorithm to the system with modulated inputs as modeled in (3.19). A formal proof of the convergence of the FWM algorithm is also required and should be investigated in the future. Duality based approaches to solving the optimization problem in (6.13) should be investigated and could lead to optimal algorithms.

Importantly, the results on detection and signal design indicate that the receiver based MUD schemes, the transmitter based signal design schemes and the joint transmitter-receiver signal processing based schemes need to be applied adaptively for each system depending on the correlations. Additionally, the proposed signal design and MUD algorithms need to be applied to real systems simulations to obtain a more appropriate measure of their performance when applied to actual fiber systems. This requires extensive simulation and appropriate channel estimation beyond the
modeling addressed here and forms an interesting future dissertation topic. Applying some of the results to differential modulation schemes is another topic that merits investigation. Two related topics that are not considered here are employing coding approaches in conjunction with signal design and detection, and joint modeling and processing of time domain and frequency domain interference effects. These topics can lead to exciting theoretical investigations when combined with known results from nonlinear optimization and integer programming.
Appendix A

A.1 ACI-dominant Systems: Proposition I

In Chapter 4, it was shown that a full rank cross-correlation matrix is neither necessary, nor sufficient to ensure the existence of detectors. The conditions in (4.17) that are used to show this are obtained here. Without loss of generality, $A$ is set to unity for linear interference. The following notation is used:

$$|Rb^{(i)}| = D^{(i)}Rb^{(i)}, \quad i \in \{1, \ldots, 2^K\},$$  \hspace{1cm} (A.1)

where $D^{(i)}$ are diagonal matrices (which depends on $b^{(i)}$) such that

$$[D^{(i)}]_{m,m} = \begin{cases} 1, & \text{if } [Rb^{(i)}]_m = 0 \\ -\tan^{-1}\left(\frac{\text{Im}([Rb^{(i)}]_m)}{\text{Re}([Rb^{(i)}]_m)}\right), & \text{if } [Rb^{(i)}]_m \neq 0 \end{cases}$$  \hspace{1cm} (A.2)

**Necessity:** Suppose $\mathcal{R}_0 \cap \mathcal{R}_1 \neq \emptyset$. If the all zeros vector $0 \in \mathcal{R}_0 \cap \mathcal{R}_1$, then clearly (a) must be satisfied. If $0 \not\in \mathcal{R}_0 \cap \mathcal{R}_1$, then there must exist $b^{(i)} \in \mathcal{H}_0$ and $b^{(k)} \in \mathcal{H}_1$. 
Appendix A. Appendix

and corresponding diagonal matrices $D^{(i)}$ and $D^{(k)}$ such that

$$\mathcal{R}_0 \ni D^{(i)} R b^{(i)} = D^{(k)} R b^{(k)} \in \mathcal{R}_1. \quad (A.3)$$

Let $m = \left\{ l \in \{2, \ldots, K : b_l^{(i)} = 1\} \right\}$, and $n = \left\{ l \in \{2, \ldots, K : b_l^{(k)} = 1\} \right\}$ and $B = D^{(k)-1} D^{(i)}$, then (A.3) can be written as

$$B \sum_{p \in m} [R]_{:,p} = [R]_{:,1} + \sum_{q \in n} [R]_{:,q},$$

which then gives

$$[R]_{:,1} = B \sum_{p \in m \setminus n} [R]_{:,p} - I \sum_{q \in n \setminus m} [R]_{:,q} + (B - I) \sum_{l \in m \cap n} [R]_{:,l} + 0 \sum_{r \in (m \cup n) \setminus c} [R]_{:,r}. \quad (A.4)$$

Since $B$ is the product of diagonal matrices with complex exponential entries, it is diagonal with complex exponential entries, and hence necessity of (4.17) holds.

Sufficiency: Suppose either condition (a) or (b) in (4.17) hold. If (a) holds then by definition, $0 \in \mathcal{R}_1$, and hence $\mathcal{R}_0 \cap \mathcal{R}_1 \neq \phi$ since $0 \in \mathcal{R}_0$. Suppose $B$ is any diagonal matrix with complex exponential entries on the diagonal such that (b) holds. Let $\{P_i\}_{i=1}^{2^K}$, be diagonal matrices as defined in (4.17) and let

$$m = \left\{ i \in \{2, \ldots, K : P_i = B \} \right\},$$

$$n = \left\{ i \in \{2, \ldots, K : P_i = B - I \} \right\},$$

$$t = \left\{ i \in \{2, \ldots, K : P_i = -I \} \right\}.$$
Using these we get

\[
[R]_{;1} = B \sum_{i \in m} [R]_{;i} + (B - I) \sum_{k \in n} [R]_{;k} - I \sum_{k \in \bar{t}} [R]_{;k}
\]

\[
\Rightarrow \sum_{k \in n \cup \bar{t} \cup \{1\}} I[R]_{;k} = B \sum_{i \in m \cup n} [R]_{;i}.
\]

(A.5)

Let

\[
b_i^{(p)} = \begin{cases} 
1 & i \in n \cup \{1\} \\
0 & i \not\in n \cup \{1\}
\end{cases} \quad \& \quad b_i^{(q)} = \begin{cases} 
1 & i \in n \cup m \\
0 & i \not\in n \cup m
\end{cases}.
\]

Clearly, \(b^{(p)} \in \mathcal{H}_1\) and \(b^{(q)} \in \mathcal{H}_0\) and (A.5) can be written as

\[
R b^{(p)} = B R b^{(q)}
\]

and since \(B\) is diagonal with complex exponential entries of unit amplitude, we have

\[
\mathcal{R}_1 \ni |R b^{(p)}|_o = |B R b^{(q)}|_o = |R b^{(q)}|_o \in \mathcal{R}_0,
\]

and hence \(\mathcal{R}_0 \cap \mathcal{R}_1 \neq \emptyset\). Thus sufficiency is proved.

A.2 XPM-dominant Systems: Proposition II

With nonlinear interference the following notation is used:

\[
|R_{b^{(i)}} b^{(i)}| = D^{(i)} R_1 b^{(i)} \quad b^{(i)} \in \{1, \ldots, 2^K\},
\]

(A.6)
where $R_1 = R_{\{b=1\}}$ and $\bar{D}(i)$ are again diagonal matrices (which depend on $b^{(i)}$) but with unit phasors or zero elements on the diagonal. Specifically,

$$[\bar{D}^{(i)}]_{m,m} = \begin{cases} 0, & [Rb^{(i)}]_m = 0 \\ -\tan^{-1}\left(\frac{\Re([Rb^{(i)}]_m)}{\Im([Rb^{(i)}]_m)}\right), & [Rb^{(i)}]_m \neq 0 \end{cases}$$

(A.7)

Using this definition and a procedure similar to the one outlined above for linear interference, conditions similar to (4.17) are obtained for the case of nonlinear interference here. This condition is given as

$$\mathcal{R}_0 \cap \mathcal{R}_1 \neq \emptyset \text{ iff at least one of the following conditions hold}$$

(a) $\mathcal{H}_1 \cap \{\text{Null}\{R_1\}\} \neq \emptyset$

(b) $\exists b^{(q)} \in \mathcal{H}_0$ and $b^{(p)} \in \mathcal{H}_1$ s.t. $D^{(p)}[R_1]_{:,1} = \sum_{i=2}^{K} P_i[R_1]_{:,i}$ for some $P_i \in \{0, \bar{D}^{(q)} - \bar{D}^{(p)}\}$, $i \in \{2, \ldots, K\}$

(A.8)

A.3 Nonconcavity of AME of Polynomial Detectors

This Appendix shows that the AME is not concave in detector coefficients of a single threshold polynomial detector, unless the detector is a single user detector.

A single user detector takes the form $y_1 - d \overset{\hat{b}_{1=1}}{\geq} 0$ where nonzero AME is achieved when the threshold $d$ is such that

$$\max\{y_1 : y \in \mathcal{R}_0\} < d < \min\{y_1 : y \in \mathcal{R}_1\}$$

or

$$\max\{y_1 : y \in \mathcal{R}_1\} < d < \min\{y_1 : y \in \mathcal{R}_0\}$$

(A.9)
When using a particular coefficient \(d\) the AME is given as

\[
\eta_{SUD}(d) = \begin{cases} 
\min \{|y_1 - d|^2 : y \in \mathcal{R}_0 \cup \mathcal{R}_1\} & \text{(A.9) holds} \\
0 & \text{otherwise}
\end{cases}
\]  

which is clearly concave in \(d\). The nonconcavity of all other polynomial detectors with linear terms is demonstrated using an example. Consider a two-user system with linear interference, unit amplitude, i.e., \(A = 1\), and identical cross-correlation \(\rho_{1,2} = \rho_{2,1} = \rho = -0.4\). For this system, consider Affine–\(x\) detectors represented as \(\alpha_0 = [c_1, d]^T = [0.7133, 0.0421]^T\) and \(\alpha_1 = [c_1, d]^T = [0.8940, 0.0168]^T\) and having nonzero AME \(\eta_0\) and \(\eta_1\) respectively. Fig. A.1 shows the AME, achieved by all convex combinations of these detectors, \(\theta\alpha_0 + (1 - \theta)\alpha_1, \theta \in (0, 1)\) denoted as \(\eta_\theta\). The convex combination of the the AME of \(\alpha_0\) and \(\alpha_1\), \(\theta\eta_0 + (1 - \theta)\eta_1\), is also plotted. Clearly, \(\eta_\theta < \theta\eta_0 + (1 - \theta)\eta_1\), i.e., the AME is not concave. All higher order polynomials in \(x\) having linear terms thus have nonconcave AME. Similar demonstrations show the AME of Affine–\(y\) detectors is also nonconcave and hence that of all higher order polynomials in \(y\) with linear terms is nonconcave as well.

### A.4 Proof of Proposition III

The fact stated in (4.29) giving a method for choosing the MLE points on the AO decision boundary is proven here via contradiction. Let

\[
F_m = \{z \in \mathcal{R}_+^K : L(z|b^{(m)}) \geq L(z|b^{(k)}), \forall b^{(k)} \in \mathcal{H}_0\}
\]
Figure A.1: AME, $\eta_\theta$, of a convex combination detectors of two specific Affine—$x$ detectors, $g_0 = \{0.7133, 0.0421\}$ and $g_1 = \{0.8940, 0.0168\}$, and convex combination of AME of specific detectors, vs combination coefficient $\theta \in (0, 1)$ for a two-user system with linear interference, $A = 1$, $\rho_{i,j} = \rho = -0.4 \ \forall i \neq j$.

Suppose $z^* \in T_m$, $z^* \in B$ and contrary to (4.29),

$$\exists j \neq m \ \text{s.t.} \ L(z^*|b^{(m)}) < L(z^*|b^{(j)}). \quad (A.11)$$

Now either $b^{(j)} \in H_1$ or $b^{(j)} \in H_0$. If $b^{(j)} \in H_1$, then $z^* \in F_j$ since $\exists q$, s.t. $b^{(q)} \in H_0$ and

$$L(z^*|b^{(q)}) > L(z^*|b^{(m)}),$$

and (A.11) holds. Then either $z^*$ is a boundary point for $F_j$ or it is in the interior of $F_j$. If it is on the boundary, then $\exists q$ s.t. $b^{(q)} \in H_0$ and

$$L(z^*|b^{(q)}) = L(z^*|b^{(j)}) > L(z^*|b^{(m)}),$$
which then implies that $z^* \not\in F_m$ and hence $z^* \not\in T_m$ clearly leading to a contradiction. Hence, $z^*$ must be in the interior of $F_j$, which would imply $z^* \not\in B$. This contradiction indicates that $\not\exists j$ s.t. $b^{(j)} \in \mathcal{H}_1$ and $L(z^*|b^{(m)}) < L(z^*|b^{(j)})$. Now suppose $b^{(j)} \in \mathcal{H}_0$. Then clearly $z^* \not\in T_m$ leading to a contradiction. Hence, the fact holds in the forward direction. The reverse direction can be proved similarly to show that the fact holds.


Bibliography


