CROSSTALK MITIGATION OF HIGH-SPEED INTERCONNECTS USING MODAL SIGNALING

BY

PAVLE MILOSEVIC

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering in the Graduate College of the University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor José E. Schutt-Ainé, Chair
Professor Jennifer T. Bernhard
Professor Elyse Rosenbaum
Professor Naresh R. Shanbhag
Today’s high-speed I/O signaling links are faced with difficult challenges: due to manufacturing and technology limitations, the resources (pins and interconnects) available for off-chip signaling remain almost constant, while the throughput needed is increasing and the required aggregate bandwidths are moving into the Tb/s range. The solutions need to simultaneously satisfy the requirements for low power (typically 1-2 mW/Gb/s for short chip-to-chip links) and extremely reliable signaling (with target bit error rates as low as $10^{-21}$ for memory links). The routing density is being increased, and faster edge rates of signals are being used, thus causing increased levels of electromagnetic coupling between interconnects. This results in various signal integrity impairments, which in turn limit the system performance and signaling rate. The focus of this research is to explore the application of modal decomposition of coupled transmission lines to crosstalk mitigation of high-speed interconnects, in particular far-end crosstalk (FEXT), which is the dominant noise source for modern single-ended memory links. Special attention is devoted to addressing the issues that arise over realistic tightly coupled cascaded channels with discontinuities in the signal path.

First, we propose the application of generalized modal decomposition theory to the class of tightly coupled, nonhomogeneous and nonuniform channels with discontinuities. The proposed approach offers a robust method of extracting modal properties of the channel starting from the presumed channel geometry and structure, or from actual measured channel data, or a combination of both. Based on the results of generalized modal decomposition, optimal encoder, decoder and termination blocks for the modal signaling system are extracted from
channel geometry or measurements. Due to the nonuniform structure of the communication channels, we propose the use of a frequency-dependent termination network for optimal signaling performance. We demonstrate the performance benefits over the suboptimal resistive grid termination network in terms of decision margin improvement and crosstalk-induced jitter reduction.

In order to facilitate transceiver design, we explore a MIMO system perspective of modal signaling. In this context, some of the important system performance metrics are analyzed. We demonstrate the method of obtaining the modal decoder coefficients for near-optimum SNR for a given channel. We outline a methodology for determining the required number of bits of modal encoder and decoder precision given the target bit-error rate. Finally, we propose two approaches for practical system implementation of modal signaling, using (a) digital cores present in ADC/DAC based transceivers, and (b) analog frontend transceiver structure. For the analog transceiver, the design flow is demonstrated using the low-power digital CMOS process, using a case study of a typical controller-memory microstrip bus.
To my family, for their love and support
ACKNOWLEDGMENTS

First and foremost, I would like to extend my utmost gratitude to my adviser, Dr. José E. Schutt-Ainé, for his guidance during the course of this work. I thank him for having faith in me and recognizing my potential to achieve academic and research success in the exciting field of electromagnetics, and for doing so at the time when those goals were not entirely clear to myself. His gentle guidance, patience and motivation have made my research and dissertation efforts both a fruitful and a joyous experience.

I would like to humbly thank my committee and the SAMS research group members, Drs. Naresh Shanbhag, Elyse Rosenbaum, Jennifer Bernhard and Andrew Singer, for their valuable assistance and suggestions, always bringing a different perspective and real-time calibration of my research efforts.

I acknowledge the Department of Electrical and Computer Engineering of the University of Illinois at Urbana-Champaign for providing me with financial support during my graduate studies.

The summer internship positions held over the course of my graduate studies provided me with the unique opportunity to acquire hands-on experience and the industry point of view on our research problems. In particular I am fortunate to have had the patience and understanding of my managers Neffody Kraskoff at Intel and Ralf Schmitt at Rambus, as well as the privilege to meet and work with Wendem Beyene, Dan Oh, Marko Aleksic, Craig Hampel, Avery Herrick and all the exceptional colleagues I had in Los Altos and Hillsboro.

Of great importance to me was not only the scientific but also the personal support of the unique people I met at Illinois and beyond. I will always reflect fondly on the great
memories I shared with my research group colleagues, Dr. Dmitry Klokotov, Patrick Goh, Dr. Yidnek Mekonnen, and all others with whom I worked over the years. I would like to thank my friends Dr. Rajan Narasimha, Rami Abdallah and Adam Faust for their swift assistance with numerous technical aspects of this dissertation work.

The years spent in graduate school would have passed by much more slowly and would never have provided me with such a deep personal satisfaction if it were not for my wonderful friends, some of whom have cheered me on from back home in Serbia, and others whom I have had the good fortune to meet in the USA: Vanja Lazarevic, Marko Stojkovic, Alen Milosevic, Nasim Suterwala, Dr. Marko and Mara Orescanin, Milos Curcic, Evelyne Marino, Dr. Tijana Prodanovic, Dr. Aleksandra Kijac, Dr. Zarija Lukic, Jackie Schiffer, Dr. Benedict Ofuonye, Dr. Joyram Chakraborty, Dr. Nirali Shah, Bijoya Chakraborty, Dr. Christopher Stewart, Billy Joe Mills, Branko Radan, Rade Djordjevic, Dusa Vukovic, and my best man Milos Stojanovic.

Lastly, I am forever indebted to my parents Zoran and Ljiljana, my lovely wife Mila, uncles Dragoslav and Stevan, and dearest cousins Irena and Ksenija, for their love, understanding and encouragement throughout this long and winding journey. This work is dedicated to them.
# TABLE OF CONTENTS

LIST OF TABLES .................................................................................. ix

LIST OF FIGURES .............................................................................. x

CHAPTER 1 INTRODUCTION ................................................................. 1
  1.1 Background .............................................................................. 1
  1.2 Signaling Channels and Signal Integrity Impairments ................. 3
  1.3 Crosstalk-Induced Jitter and the Signaling Budget .................. 7
  1.4 Objective and Thesis Organization ........................................ 8

CHAPTER 2 CROSSTALK MITIGATION OF HIGH-SPEED INTER-
  CONNECTS ....................................................................................... 10
  2.1 Crosstalk Sources and Impact on System Performance .......... 10
  2.2 Overview of Proposed Crosstalk Mitigation Solutions ........... 12
  2.3 Modal Approach to Crosstalk ................................................ 15
  2.4 Generalized Modal Decomposition Theory ........................... 17
  2.5 Summary ................................................................................. 19

CHAPTER 3 APPLICATION OF MODAL DECOMPOSITION TO CROSSTALK-
  FREE SIGNALING .......................................................................... 21
  3.1 Concept of Modal Signaling: A Time Domain Example .......... 22
  3.2 General Multiline System: A Frequency-Domain Formulation .... 25
  3.3 Modal Signaling System Structure ......................................... 29
  3.4 Modal Extraction and Simulation Framework ........................ 31
  3.5 Summary ................................................................................. 34

CHAPTER 4 MODAL SIGNALING OVER NONUNIFORM COUPLED INTER-
  CONNECTS WITH DISCONTINUITIES ............................................. 36
  4.1 Case Study A: Uniform Interconnects .................................... 36
  4.2 Case Study B: Cascaded Nonuniform Interconnects .............. 38
  4.3 Frequency-Dependent Optimal Termination Network ............ 45
  4.4 Some System-Level Metrics of Modal Signaling ................. 55
  4.5 Summary ................................................................................. 66
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Physical and geometric parameters of different cascaded segments.</td>
<td>37</td>
</tr>
<tr>
<td>4.2</td>
<td>Forward voltage eigenwaves for uniform and cascaded lines.</td>
<td>38</td>
</tr>
<tr>
<td>4.3</td>
<td>Termination resistor values used for modal matching.</td>
<td>39</td>
</tr>
<tr>
<td>4.4</td>
<td>Pole values for the fitted termination model.</td>
<td>48</td>
</tr>
<tr>
<td>4.5</td>
<td>Optimal decoding coefficients for the nonuniform system.</td>
<td>62</td>
</tr>
<tr>
<td>4.6</td>
<td>SNR levels at the slicer input for common and uncorrelated input-referred noise after modal decoding.</td>
<td>62</td>
</tr>
<tr>
<td>5.1</td>
<td>Current changes to generate modal signals and the reference voltage level.</td>
<td>75</td>
</tr>
<tr>
<td>5.2</td>
<td>Peak-to-peak jitter for NRZ signaling over nonuniform bus with Ci=0.5 pF.</td>
<td>82</td>
</tr>
<tr>
<td>5.3</td>
<td>Peak-to-peak jitter for modal signaling over nonuniform bus with Ci=0.5 pF.</td>
<td>83</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1.1 The number and speed of I/Os based on the ITRS roadmap. .......................... 2
1.2 A typical controller-memory interface, forming a cascaded channel with discontinuities. .......................................................... 3
1.3 Insertion loss of the 5-in graphics channel for different device effective capacitances. .......................................................... 4
1.4 Measurement of direct (S21) and crosstalk (S41) signal for coupled microstrip lines with varying microstrip separation. ......................... 5
1.5 A typical DDR timing budget. .......................................................... 7
2.1 Even and odd propagation modes in a two-conductor system. ......................... 10
2.2 Generation of a FEXT noise pulse on a quiet victim line. .......................... 12
2.3 Far-end crosstalk (FEXT) resulting in crosstalk-induced jitter (CIJ). .................. 13
3.1 Block diagram of a two-line signaling link with modal signaling blocks. ......... 22
3.2 The blocks of a generalized MIMO modal transceiver system, including the channel $H$. .......................................................... 26
3.3 Unidirectional modal signaling block diagram for an N-line channel. ............ 29
3.4 Frequency domain modal propagation model in matrix form. ....................... 30
3.5 Raw eigenvalue decomposition of the resonant system resulting in non-physical modal velocity plots. .......................... 33
3.6 Raw eigenvalue decomposition of the resonant system resulting in encoder coefficients misinterpretation. .......................... 34
4.1 Cross-section of a four-line interconnect bundle. ........................................ 37
4.2 Propagation constants of modes for uniform PCB channel. ........................ 38
4.3 Propagation constants (3) of modes for the cascaded channel. .................... 40
4.4 Real part of (a) voltage and (b) current forward eigenvector coefficients for the cascaded channel. ........................................ 41
4.5 S-parameters of direct and strongest crosstalk paths for the uncoded cascaded channel. .................................................. 42
4.6 S-parameters of direct and strongest crosstalk paths for the modal encoded cascaded channel. .................................................. 43
4.7 Real part of (a) voltage and (b) current forward eigenvector coefficients for the cascaded channel with discontinuities. ....................... 44
4.8 A full resistive termination grid for a N=4 line system. .......................... 47
4.9 Real (top) and imaginary (center) part of the Y11 diagonal parameter of matching admittance, extracted (solid) and modeled (dashed). Mean absolute fitting error values are also shown (bottom). 49
4.10 Real (top) and imaginary (center) part of the Y12 off-diagonal parameter of matching admittance, extracted (solid) vs modeled (dashed). 50
4.11 S-parameters of direct and strongest crosstalk paths for the modal encoded cascaded channel with modeled optimal terminations. 51
4.12 S-parameters of direct signal paths for the modal encoded cascaded channel with full-frequency optimal terminations. 52
4.13 S-parameters of direct signal paths for the modal encoded cascaded channel with low-order model of the optimal terminations. 53
4.14 S-parameters of direct signal paths for the modal encoded cascaded channel with low-frequency extracted resistive terminations. 54
4.15 Decoded modal signals corresponding to modes 1-4 for the modal encoded cascaded channel with resistive terminations. 55
4.16 Decoded modal signals for the modal encoded cascaded channel with modeled terminations. 56
4.17 Modal signaling system MIMO block diagram, including noise. 57
4.18 ZF (X) and LMMSE (o) coefficients for the 4-line channel. 60
4.19 BER plot as a function of SNR at the slicer for NRZ modulation. 64
4.20 Two approaches to calculating BER of decoded signal with encoder quantization. 65

5.1 Block diagram of a DSP-based modal transmitter system. 68
5.2 A 3-bit current steering DAC realization. 69
5.3 Typical transition time plot of a 3-bit current steering DAC. 70
5.4 Estimated power of the DSP Tx block as a function of number of lines N and DSP precision M (normalized to 1 Gb/s). 72
5.5 Proposed transmitter architecture for modal signaling. 73
5.6 Proposed receiver architecture for modal signaling. 74
5.7 (a) Open-drain drivers producing the common-voltage levels; (b) current-steering for shared currents, and (c) for non-shared currents. 75
5.8 Layout of the encoder block. 76
5.9 Current subtractor for modal decoder realization. 78
5.10 Layout of the decoder block. 79
5.11 Pulse response on inner and outer lines of the uncoded bundle, and crosstalk pulse responses on other lines. 81
5.12 Unit pulse responses of signals over equivalent modal channels. 83
5.13 Eye diagrams of normalized decoded modal signals. 84
5.14 Eye diagram for XTC equalization for outer line 1, with XTC2-1 FIR and aggressor 2 only. 86
5.15 Pulse response of line 2 and crosstalk on neighboring lines observed at the far-end PCB-package discontinuity. 89

6.1 One-port Foster admittance realization of the transfer function. 95
6.2 Step excitation of ground mode of the cascaded system. 96
1.1 Background

Ever since the revolutionary invention of the transistor paved way for the development of integrated circuits in the 1950s, we have been witnessing the rapid development of all aspects of computer technology, in terms of processing power, communication speed and storage capacity. All of those trends have been especially intensified in the last decade, due to the now-ubiquitous Internet dictating the advancements in both the server platforms and the mobile device performances. In particular, there is an ever-increasing volume of data generated in all aspects of computer-related technologies. In order for this huge volume of data to be processed, it also needs to be reliably and efficiently transported at high speeds: demands are increased for long-haul (e.g. distant Internet servers connected via optical fibers), medium-haul (e.g. computers in a LAN network) and even short-haul communication (e.g. daughter cards sharing a backplane; processor, controller and memory chips on a motherboard; multiple cores of a CPU; and even on-chip buses).

This work focuses on high-speed chip-to-chip communication, an area which has historically been considered less critical than the on-chip signaling performance; however, due to today’s huge demands for data throughput, off-chip interconnect bandwidth is increasingly becoming the system bottleneck, and creative solutions are needed to meet the requirements, now measured even in terabytes per second [1] for graphic processing cores. The challenges in this communication area stem from the disparity between the demand for bandwidth and available resources over which to provide it.
Due to manufacturing and technology limitations, the physical dimension scaling of the passive package and interconnect systems is unable to keep up with the rapid miniaturization benefits enjoyed by active devices (such as silicon chips in CMOS technology nodes) which is dictated by Moore’s law. Therefore, the resources (pins and interconnects) available for off-chip signaling remain almost constant, while the throughput needed is increasing, as shown in Fig. 1.1. This trend is expected to persist in the future [2].

![Figure 1.1: The number and speed of I/Os based on the ITRS roadmap.](image)

At the same time, in order to keep up with the data throughput needed over limited resources, high edge rates are being used for signaling. This causes various signal integrity impairments, which in turn limit the system performance and force the data rates to be well below the Shannon limit of the channel capacity [2], [3]. Coupled with the limited channel resources are the conflicting requirements of low power (the signaling budget of modern systems is typically 10-20 mW/Gb/s for backplane and 1-2 mW/Gb/s for chip-to-chip links) and extremely reliable signaling (with target bit error rates of $10^{-12}$ for backplane or even as
low as $10^{-21}$ for memory links). All of these limitations combined call for creative solutions in order to keep up with the performance demands of next-generation systems.

1.2 Signaling Channels and Signal Integrity Impairments

The structure of a typical high-speed off-chip communication channel is shown in Fig. 1.2 [4]. Note that the channel is heterogeneous, consisting not only of different cascaded trace segments, but also of different discontinuities (vias, connectors). Also, each data bus will consist of multiple channels, placed in close proximity to one another. The active devices used in the transmit and receive circuits will also contribute to deviation of signals from their ideal shapes. All of these effects impair signal integrity and limit both the signaling speed and, as a result, the usable bandwidth of the signaling channel for reliable communication.

![Figure 1.2](image)

Figure 1.2: A typical controller-memory interface, forming a cascaded channel with discontinuities.

1.2.1 Inter-Symbol Interference

In Fig. 1.3, the magnitude of the insertion loss of an isolated 5-inch graphics controller-memory channel is shown for different values of effective device capacitance $C_i$ (average of input capacitances of transmitter and receiver) [4]. Reduced bandwidth at higher signaling frequencies translates into inter-symbol interference (ISI): the signal value at the receiver
depends not only on the current bit sent, but also on the value of previous bits. Observing the simple non-return-to-zero (NRZ) signaling, the vertical eye opening at the receiver will be reduced from its maximum value and can, at high signaling speeds, be completely closed, leaving no margin for reliable detection. This limits the channel performance in terms of achievable signaling speeds, and requires special equalization techniques [5] which mitigate ISI and open the data eye.

Figure 1.3: Insertion loss of the 5-in graphics channel for different device effective capacitances.

Analyzing the single memory signaling channel with various types of equalization, it was shown that the optimum data rate for a channel with equalization applied would be around 7.5 Gb/s [4]. However, depending on the density of lines in the data bus, ISI may not be the dominant signal integrity impairment; in other words, the system performance for high-density interconnects is usually not determined by ISI.
1.2.2 Crosstalk

One of the key signal integrity impairments arises as the result of electromagnetic coupling effects between tightly coupled lines, the effect of which increases rapidly with reduced line separation, as shown in Fig. 1.4 [6]. This translates to increased levels of both near- and far-end crosstalk. In typical high-speed links, there are numerous potential points of crosstalk generation, resulting from signal coupling at discontinuities (vias, connectors, terminations), coupling in package (wirebonds, escape traces), as well as coupling in PCB traces (bus or adjacent layers for wide bus).

Figure 1.4: Measurement of direct (S21) and crosstalk (S41) signal for coupled microstrip lines with varying microstrip separation.

In particular, for interconnects in an inhomogeneous medium, such as microstrip parallel links on a low-cost PCB substrate, far-end crosstalk (FEXT) is the dominant noise source [7], resulting in both amplitude noise and crosstalk-induced jitter (CIJ) [3]. A typical FEXT-
dominated application would be a parallel signaling link (bus) between a memory controller and one of many memory modules in a DDR3 RAM system.

1.2.3 Other Noise and Jitter Sources

Depending on the signaling link structure, off-chip links can suffer from other impairments, such as thermal noise, transmit and receive jitter, simultaneous switching output noise, etc., which can have a significant impact on link performance at the low error rates of interest.

One of the main contributors to bit errors in single-ended signaling systems is the simultaneous switching output (SSO) noise [7], generated by switching of a large number of logic gates switches. This causes the I/O voltage supply to fluctuate, producing SSO noise, which couples between the power and signal distribution systems, causes false switching and degrades signal integrity.

Thermal noise is generated by the resistive terminations at the receiver and the device noise of receiver circuits. However, as estimated by [8], total input-referred random noise for a 5 GHz receiver is on the order of 0.3 mVrms, or 40 dB down from the equalized signal at the receiver. Therefore, thermal noise should not present the dominant cause of bit errors in short haul (chip-to-chip and backplane) links.

More importantly, the transmitter and receiver usually contain a phase locked loop (PLL) or a delay locked loop (DLL). The clock signal from these components is impaired by timing jitter [9]. Supply noise and reference clock phase noise are the dominant noise sources in a PLL and can contribute to a substantial timing jitter at the clock output. This jitter results in a deviation of the sampling point from the optimum sampling instant, where the eye opening is maximum.
1.3 Crosstalk-Induced Jitter and the Signaling Budget

Using a controller-memory subsystem signaling as the example, the total unit interval (UI) timing is broadly composed of three budgets: transmitter (Tx), interconnects and receiver (Rx). Nominally, each of these three budgets is allocated about one third of the total timing budget; both Tx and Rx subsystems will have inherent noise and mismatch sources (e.g. control logic delay, skew, difference in delays of rising and falling edges, skew in clock path, PLL/DLL jitter and DLL phase offset), resulting in timing jitter. This jitter affects the timing budget, which in turn limits the system performance, a standard metric of which is the maximum data rate for a required bit error rate specification.

In modern high-speed systems this split can become even more unfavorable to the interconnect portion of the system, imposing even stricter requirements on the signal integrity associated with interconnects, as shown in Fig. 1.5 [10]. As shown in the figure, the majority of the timing budget is distributed to Rx jitter (orange), routing skew (green), and Tx jitter (purple). The remaining portion needs to cover all the timing uncertainties due to interconnects. Since the data rates are increasing, the unit interval will shrink, and so will the absolute quantity of time allocated to the interconnect budget.

Figure 1.5: A typical DDR timing budget.
1.4 Objective and Thesis Organization

Comparing the two trends described in the preceding discussion (increased dominance of off-chip interconnect bandwidth as the overall system bottleneck, and the shrinking timing budget allocated to the interconnect jitter), it is obvious that a design solution which would allow for mitigation of a dominant CIJ source (in this case, FEXT) would be of great benefit. The target metrics for a particular crosstalk mitigation design should take into account not only the effectiveness of the method itself (FEXT and CIJ reduction), but also the cost (in terms of additional signaling power, area overhead and design complexity) of the solution.

This thesis provides a comprehensive overview of theoretical aspects of a particular type of bus signaling called modal signaling (which in theory promises a complete elimination of FEXT), and proposes its application to a realistic chip-to-chip signaling environment, in contrast to the idealized target environment of previous work in the field.

In order to provide the starting point towards crosstalk elimination, the sources of crosstalk in non-homogeneous signaling environments are reviewed in Chapter 2. Various previously proposed crosstalk mitigation solutions are first reviewed in this chapter, summarizing the ideas, strengths and drawbacks of each method. Then, based on the modal nature of crosstalk sources, the concept of generalized modal decomposition theory is briefly reviewed, in order to introduce the modal approach to crosstalk mitigation.

We first propose the application of generalized modal decomposition theory to nonuniform channels with discontinuities, as detailed in Chapter 3. This is in contrast to the traditional modal decomposition of channel matrices, which presumes a uniform channel cross-section throughout the entire signaling path. The proposed implementation of the decomposition method is robust and applicable to a large class of channels used in modern low-cost high-performance signaling systems. Also discussed is the realization of a modal extraction and simulation framework appropriate for analyzing nonuniform channels, due to their special properties in terms of attenuation and resonances.
Having demonstrated the method of generalized modal decomposition, Chapter 4 explores modal signaling in a realistic chip-to-chip coupled interconnect scenario. Two case studies are presented, in order to contrast the modal properties of uniform and nonuniform channels. Based on the channel properties, a suitable modal signaling system structure is proposed, and the characteristics of each building block are discussed. Further, the potential performance benefits and design tradeoffs are analyzed in terms of relevant system-level metrics.

Based on the results of case studies, in Chapter 5 we propose a particular CMOS circuit-level realization of a transceiver system which employs modal signaling. The proposed realization is designed to operate over a low-cost chip-to-chip signaling bus, in which the off-chip design resources (routing area and PCB quality) are limited, and the uncoded channel performance is therefore heavily crosstalk-dominated. The benefits of the realized modal signaling system are demonstrated using circuit-level simulation. The performance and cost metrics of the proposed design are compared with other crosstalk mitigation techniques proposed for similar applications.

Chapter 6 summarizes the contributions of this thesis. Also discussed are research directions that could be considered as the natural extension of the presented work.
2.1 Crosstalk Sources and Impact on System Performance

In order to gain insight into properties of crosstalk, it is essential to examine the concept of transmission line modes as a source of far-end crosstalk (FEXT). To that end, we will use a simple example of a two-line conductor system. The two fundamental propagation modes are the even mode (with both lines switching in the same direction) and the odd mode (with the lines switching in the opposite direction). In a homogeneous medium (such as stripline), the two mode velocities will theoretically be equal. However, due to lower manufacturing costs, many practical applications perform signaling over a non-homogeneous medium, such as a microstrip (with signal propagating partially through substrate, and partially through air), causing the mode velocities to be different due to the difference of the dielectric constants of two mediums, as shown in Fig. 2.1 [11].

Figure 2.1: Even and odd propagation modes in a two-conductor system.
In a typical case where this two-line system would be used to convey two independent signal sequences using single-ended signaling, the difference in mode propagation velocities will cause the phenomenon called far-end crosstalk (FEXT). To understand its impact, let us observe the signal on the quiet (victim) line when its neighboring (aggressor) line is excited from its value corresponding to logical 0 to logical 1. This excitation will induce both modes of propagation in this two-line system – some of the signal energy will travel over the even mode, and some of it will be contained in the odd mode. The times of flight of the two modes can be calculated from the parameters per-unit-length (PUL) of the system, namely from the self- and mutual-inductance and capacitance of lines, as follows:

\[ T_o = l\sqrt{(L_s - L_m)(C_s + C_m)} \]
\[ T_e = l\sqrt{(L_s + L_m)(C_s - C_m)} \]  \hspace{1cm} (2.1)

where \( T_o \) and \( T_e \) are times of flights of odd and even modes, and \( l \) is the length of each of the two lines. These two equations can be rewritten in the form which leaves only ratios of mutual- and self-inductance, and likewise for capacitances. Assuming a typical microstrip configuration of a microstrip PCB bus, usually the ratio is such that the odd mode will travel faster:

\[ \frac{L_m}{L_s} > \frac{C_m}{C_s} \]  \hspace{1cm} (2.2)

This means that the two signals will arrive at the far end of the victim line at different times; their superposition will produce a noise pulse at a nominally quiet line, as shown in Fig. 2.2 [11].

It is now straightforward to see that, over the course of many billions of bits transmitted by a typical signaling system, these noise pulses will translate into timing jitter (deviation in time-domain of the crosstalk-affected signal away from its reference waveform), as shown in Fig. 2.3 [11]. This jitter is called the crosstalk-induced jitter (CIJ), and belongs to the class of bounded uncorrelated jitter, since its effect is bounded by the difference in time of
flight extremes of two modes, but at the same time uncorrelated to the actual signal that is being transmitted over the line under observation.

2.2 Overview of Proposed Crosstalk Mitigation Solutions

There are numerous design techniques to prevent tight coupling, but not all of them may be practical to apply for a given situation. Crosstalk could be reduced by using differential signaling (as opposed to single-ended), but at the cost of using twice as many signaling lines; also, to achieve the required total aggregate bandwidth, the signaling would have to be performed as twice the original rate, which consumes more power and is not always possible due to CMOS switching speed limitations. The channel could be redesigned with homogeneous environment (striplines) and increased spacing or guard lines, and the discontinuities could be smoothened out by careful design, but again this may not be feasible due to manufacturing limitations, lack of routing resources, or prohibitively increased system cost.
This thesis is concerned with crosstalk mitigation solutions for a tightly coupled non-homogeneous channel which is given as the design input, and it is assumed that the solutions which modify the channel itself are prohibited due to any of the reasons described above. Hence, the solution space would have to encompass the design of the receiver and transmitter to take into account crosstalk. Various active and passive approaches have been suggested in order to mitigate the effects of FEXT, but at the expense of area overhead, power or design complexity. We can divide the crosstalk mitigation approaches in several broad classes, described in the following sections.

2.2.1 Signal Coding

A well-known technique for on-chip buses [12], the coding methods for off-chip signaling include forbidden transition codes [13] and incremental signaling [14], as well as using differential instead of single-ended signaling [15]. While the benefit in utilizing those techniques is tangible, in particular in the case of differential signaling (the widely adopted solution for high-end signaling links of today [16]), the main drawback of those methods is the increased
number of lines needed for the same aggregate bandwidth (or the increase in data rate if the
interconnect real-estate is constant), as well as the increased complexity in particular of the
receiver implementations needed for decoding.

2.2.2 Crosstalk-Induced Jitter (CIJ) Compensation

Recognizing that most of FEXT-coupled energy is introduced at signal transitions, and fur-
thermore that CIJ is independent of signal swing, and is not sensitive to transition slope [17],
these techniques do not directly remove crosstalk, but attempt to correct its effects on tim-
ing. The idea is to detect the induced modes produced by a particular signal combination
present on the bus, and adjust the timing of received signals appropriately. This approach
can be implemented in the receiver [17], as well as in the transmitter [18]. The main issues
connected with this approach are the sensitive variable delay circuits needed to retime the
signals, the difficulty of reliably detecting transmitted modes at the receiver, as well as the
complexity of generalizing the technique to multilane buses.

2.2.3 FEXT Cancelation

The principal idea behind this approach is to estimate the noise caused by crosstalk, and
inject the opposite signal that will cancel FEXT at transitions. The cancelation signal can
be inserted at the transmitter [19], or at the receiver [20]. FEXT estimators can be realized
by discrete-time (DT-FIR) [19] or continuous-time (CTLE) equalizers [20]. The main issues
of these techniques over tightly coupled buses are the difficulty in generating an accurate
replica of the crosstalk signal, the requirement of multiple FIR taps per line, and glitch
introduction at eye center (in case of DT implementations). An improvement would be to
shift coupled FEXT to occur away from data transitions [21], and then remove the voltage
noise using a glitch canceler [22], but this raises the issues of the design complexity and
power.
2.2.4 Passive Equalization

The passive solutions attempt to reduce the even-odd mode velocity mismatch, usually by increasing mutual capacitance $C_m$ to reduce the difference between inductive and capacitive coupling ratio 2.2. Suitable implementations can take the form of stubs [23], parallel capacitors [24], and guard traces [25]. This provides a passive, local solution, but generalization to tightly coupled buses is not straightforward, and channel redesign or additional board real-estate are not always viable options (in case of legacy channels, or low-cost systems).

2.3 Modal Approach to Crosstalk

The main common characteristic of all the previously proposed crosstalk mitigation methods would be that they treat the electromagnetic coupling between transmission channels as an undesired phenomenon, and attempt to mitigate its effect. Therefore, they become harder (more complex or power-demanding) to implement as coupling gets tighter, since there is more crosstalk that needs to be canceled.

An alternative approach named modal signaling [26], which will be explored in depth in the remainder of this work, takes advantage of tight coupling using channel diagonalization, and thus enables increased routing density. The modal approach was explored in several previous works, such as [27] and [28], but only for some special classes of problems: uniform lossless interconnect bundles, where the modal propagation parameters are purely real and frequency-independent; or homogeneous medium (stripline) which does not suffer from FEXT. The aim of this work is to provide a comprehensive treatment of realistic microstrip channels with several cascaded segments and discontinuities in the signal path.
2.3.1 Modal Coordinates

Having introduced the idea of FEXT caused by the transmission-line modes, the notion of modal space can be discussed next. The effect of signals excited in the observable line space can be separated into their odd- and even-mode components. In line space, the signals are coupled, meaning that exciting one line will produce an effect on the other, as shown. However, in modal space, the two modes are linearly independent of one another, meaning that there is no coupling between the two. This can be observed starting from the classical voltage and current representation method in the form of telegrapher’s equations [29]. The line bundle can be described by its resistance $R$, conductance $G$, inductance $L$ and capacitance $C$ matrices per unit length as follows:

$$Z = R + j\omega L, \quad Y = G + j\omega C$$  \hspace{1cm} (2.3)

Telegrapher’s equations in frequency domain reveal coupling of currents and voltages:

$$\frac{d^2V}{dz^2} = (ZY)V, \quad \frac{d^2I}{dz^2} = (YZ)I$$  \hspace{1cm} (2.4)

With regard to (2.4), the decoupling of equations will take place with the substitution of the line space voltage vector $V$ with a modal transformation $\hat{V} = EV$, where the transformation matrix $E$ is chosen such that the matricial product leads to diagonal matrix:

$$E^{-1}ZYE = \gamma^2 = \text{diag}\{\gamma_1^2, \ldots, \gamma_N^2\}$$  \hspace{1cm} (2.5)

In the case of a two-line system $N=2$, and the diagonal matrix $\gamma^2$ collects the eigenvalues of $ZY$. The similar transformation may be written for the vector of currents $I$. From a physical point of view, the matrix elements in each of the column vectors $e_k$ of the transformation matrix $E$ show the distribution of the conductor voltage phasors at a given point of the line, when only mode $k$ is present on the line [30]. It is also important to note that the
values of elements in the matrix $E$ by themselves do not have a special meaning, but rather their ratios do; the multiplication of $e_k$ by an arbitrary constant does not affect (2.5). The physical interpretation of the elements in the diagonal eigenvector matrix $\gamma^2$ is that each element corresponds to a square of the modal propagation constant of one propagation mode.

In the special case of a two-line interconnect system, matrix $E$ will be an arbitrarily column-scaled version of the following matrix:

$$E = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

(2.6)

Revealing the well-known odd- (first column) and even-mode (second column) waveshapes, as discussed before. The inverse $E^{-1}$ of matrix $E$ will in this special two-line case be just an appropriately scaled version of the matrix $E$ as given in (2.6). This will define the transformation from modal space back to the line space.

However, for the case of a multi-conductor system with $N$ lines where $N \geq 3$, the waveshapes will be dependent on the geometry of the system (line spacing, width, dielectric height, and properties of conductors and dielectric) and must be calculated or estimated.

### 2.4 Generalized Modal Decomposition Theory

Modal decomposition of uniform interconnects is well understood and explored [29]. However, most practical interconnect channels consist of multiple cascaded uniform line segments with different cross-sections (i.e. package and PCB traces), as well as other discontinuities (i.e. vias and connectors). The issue with directly applying the traditional modal decomposition to the case of such nonuniform transmission line bundles is that the $RLCG$ matrices per unit length will not remain constant throughout the entire length of the line. Therefore, any modal extraction which diagonalizes $ZY$ and $YZ$ matrices will have a local quality. For the concept of signaling, however, we are mainly interested in the signals as observed at line
terminals (at the transmitter and the receiver), and would require a suitable decomposition technique which will observe the channel as a whole, and in the process take into account (average out) the nonuniformity of the channel structure. Traditional modal decomposition of uniform transmission lines described in the previous chapter has been generalized for the nonuniform case, as presented in [31]. In this section, some of the key ideas are reviewed, in order to examine their application to the multimode encoder/decoder signaling system.

Observing the general nonuniform transmission line system modeled from its terminals, in the absence of nonlinearities, its sending end (near-end) and receiving end (far-end) voltages and currents can be related at each frequency of interest using, for example, the $ABCD$-matrix representation as follows:

$$
\begin{bmatrix}
  v_S \\
  i_S
\end{bmatrix} =
\begin{bmatrix}
  A & B \\
  C & D
\end{bmatrix}
\begin{bmatrix}
  v_R \\
  i_R
\end{bmatrix}
$$

(2.7)

where, in a system with $N+1$ conductors (including the reference conductor or plane),

$$v_S = [v_{S1} v_{S2} \ldots v_{SN}]^T, \quad i_S = [i_{S1} i_{S2} \ldots i_{SN}]^T$$

$$v_R = [v_{R1} v_{R2} \ldots v_{RN}]^T, \quad i_R = [i_{R1} i_{R2} \ldots i_{RN}]^T$$

(2.8)

are the column vectors representing the line voltages $v$ and currents $i$ at sending end (subscript $S$) and receiving end (subscript $R$) of each of the lines $1..N$, referenced to a common reference conductor or plane.

If the $ABCD$-matrix is diagonalizable, it can be rewritten in the following form:

$$
\begin{bmatrix}
  v_S \\
  i_S
\end{bmatrix} =
\begin{bmatrix}
  W_{Fv} & W_{Bv} \\
  W_{Fi} & W_{Bi}
\end{bmatrix}
\begin{bmatrix}
  \Lambda_F \\
  \Lambda_B
\end{bmatrix}
\begin{bmatrix}
  W_{Fv} & W_{Bv} \\
  W_{Fi} & W_{Bi}
\end{bmatrix}^{-1}
\begin{bmatrix}
  v_R \\
  i_R
\end{bmatrix}
$$

(2.9)

The resulting submatrices of eigenvalues $\Lambda$ and eigenvectors $W$ describe the forward- and backward-propagating waves. In case of an $N$-line system, the waveshapes or modeshapes
(set of voltages and currents) of the forward waves are given by the first $N$ columns of the $W$ matrix, and they propagate forward according to the corresponding elements of $\Lambda_F$. The waveshapes of the backward waves are described by the remaining $N$ columns of the $W$ matrix, and they propagate backward according to the corresponding elements of $\Lambda_B$. In contrast to the uniform transmission lines, the forward and backward waveshapes of a nonuniform system are not equal in general [31], and neither are the characteristic admittance matrices of the corresponding waves, as defined by:

$$Y_{C,F} = W_F W_F^{-1}$$
$$Y_{C,B} = W_B W_B^{-1}$$

(2.10)

This important decomposition formulation has so far been used mainly in the area of analysis of electromagnetic transients and harmonics in electric power systems [31]. Applying this technique to high-speed communication channels enables us to analyze the channel behavior with a blackbox approach, where the internal constitution of the channel is not of interest. The starting point is a suitable set of calculated, simulated or measured parameters which relate the input and output ports of the channel (such as ABCD or S-parameters), so the total effect of nonuniformities of the channel structure is taken into account. The technique can be applied to channels with different geometries of constituent coupled sections, and also to channels with discrete discontinuities, rather than simple smooth transitions between sections, as will be shown later.

2.5 Summary

This chapter provided a review of the sources of crosstalk in non-homogeneous signaling environments, in order to gain an understanding of the underlying crosstalk generating mechanism. Various previously proposed crosstalk mitigation solutions are reviewed in this chapter, summarizing the ideas, strengths and drawbacks of each. The review of previously
proposed methods revealed the key common theme of treating the electromagnetic coupling as an undesired phenomenon, and therefore attempting to negate its effect. Based on the modal nature of crosstalk sources, the concept of generalized modal decomposition theory is briefly reviewed in order to introduce the modal approach to crosstalk mitigation. Emphasis was placed on the blackbox approach of generalized modal decomposition to channel description, which facilitates its application to nonuniform channels.
CHAPTER 3

APPLICATION OF MODAL DECOMPOSITION TO CROSSTALK-FREE SIGNALING

As discussed previously, the primary cause of FEXT in coupled signaling links in non-homogeneous media are different velocities of propagation of fundamental transmission-line modes. At the same time, the modes themselves are linearly independent from one another, except in special degenerate cases, which are of little practical interest [30]. Namely, the special cases for which the generalized eigendecomposition results in degenerate eigenmatrices are in practice exceptionally rare, in particular for complex channels with which this work is mainly concerned, and can be additionally handled by the perturbation approach if needed [32]. Therefore, utilizing the linear independence of modes, a natural crosstalk mitigation method for coupled interconnects would be to ensure that the energy of each of the signal streams to be transmitted in parallel is carried via only one propagation mode.

A signaling scheme for coupled interconnects called multimode signaling was suggested by [26], which takes advantage of the multiconductor transmission line theory to encode the parallel signals onto fundamental transmission line modes. Due to the linear independence of modes, the signal streams are decoupled; such signaling is theoretically free of crosstalk, and therefore could allow the data transfer at the channel capacity. It has been demonstrated [33] that this method allows line spacing similar to conventional differential signaling for the same performance, with the density of a single line per signal.
3.1 Concept of Modal Signaling: A Time Domain Example

To illustrate the concept of crosstalk mitigation using modal signaling, we first turn to an intuitive time-domain description using single-ended signaling over a simple two-line structure, as shown in Fig. 3.1.

![Figure 3.1: Block diagram of a two-line signaling link with modal signaling blocks.](image)

If the signaling system does not use the encoder/decoder (uncoded signaling), referring to the system diagram, the two input single-ended signals are \( V_{1n}, V_{2n} \), and the system output is at \( V_{1f}, V_{2f} \). Exciting the near end of the lines produces even- and odd-modal voltages as follows:

\[
V_{en}(t) = \frac{V_{1n}(t) + V_{2n}(t)}{2} \\
V_{on}(t) = V_{1n}(t) - V_{2n}(t)
\]  

(3.1)

Assuming the two lines are ideal lossless coupled lines, and since the two modes are orthogonal, the modal signals at the far end will be simply the time-delayed versions of the modal signals at the near end, with delay times related to the time of flight delays \( T_{De}, T_{Do} \)
as follows:

\[ V_{ef}(t) = V_{en}(t - T_{De}) \]
\[ V_{of}(t) = V_{on}(t - T_{Do}) \]  

(3.2)

Expressed in terms of near-end source signals:

\[ V_{ef}(t) = \frac{1}{2} (V_{1n}(t - T_{De}) + V_{2n}(t - T_{De})) \]
\[ V_{of}(t) = V_{1n}(t - T_{Do}) - V_{2n}(t - T_{Do}) \]  

(3.3)

Finally, the far-end modal voltages result in the following line voltages:

\[ V_{1f}(t) = V_{ef}(t) + \frac{1}{2} V_{of}(t) \]
\[ = \frac{1}{2} (V_{1n}(t - T_{De}) + V_{1n}(t - T_{Do}) + V_{2n}(t - T_{De}) - V_{2n}(t - T_{Do})) \]  

(3.4)

\[ V_{2f}(t) = V_{ef}(t) - \frac{1}{2} V_{of}(t) \]
\[ = \frac{1}{2} (V_{1n}(t - T_{De}) - V_{1n}(t - T_{Do}) + V_{2n}(t - T_{De}) + V_{2n}(t - T_{Do})) \]

In the case of homogeneous channel (e.g. stripline), \( T_{De} = T_{Do} = T_D \) and therefore the system exhibits no far-end crosstalk:

\[ V_{1f}(t) = V_{1n}(t - (T_D)) \]  
\[ V_{2f}(t) = V_{2n}(t - (T_D)) \]  

(3.5)

However, for the nonhomogeneous channel (e.g. microstrip), each far-end line voltage depends on both excitations, as seen in (3.4), therefore exhibiting FEXT, which in turn induces jitter and closes the horizontal eye opening of the received signals.

Now, assume that the information signals \( V_{1s}, V_{2s} \) are linearly combined using the encoder matrix T such that each of the signals is mapped to one of the two orthogonal modes, as
follows:

\[ V_{1n}(t) = V_{1s}(t) + V_{2s}(t) \]  
\[ V_{2n}(t) = V_{1s}(t) - V_{2s}(t) \]  

(3.6)

Note that this corresponds to the encoding matrix \( T \) being equal to the modal decomposition matrix \( E \) from (2.6):

\[ T = E = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \]  

(3.7)

Now the far-end signals are:

\[ V_{1f}(t) = \frac{1}{2}(V_{1s}(t - T_{De}) + V_{2s}(t - T_{De}) + V_{1s}(t - T_{De}) - V_{2s}(t - T_{De})) \]
\[ + V_{1s}(t - T_{Do}) + V_{2s}(t - T_{Do}) - V_{1s}(t - T_{Do}) + V_{2s}(t - T_{Do})) \]  

\[ V_{2f}(t) = \frac{1}{2}(V_{1s}(t - T_{De}) + V_{2s}(t - T_{De}) + V_{1s}(t - T_{De}) - V_{2s}(t - T_{De})) \]
\[ - V_{1s}(t - T_{Do}) - V_{2s}(t - T_{Do}) + V_{1s}(t - T_{Do}) - V_{2s}(t - T_{Do})) \]  

(3.8)

After simplifying, we get:

\[ V_{1f}(t) = V_{1s}(t - T_{De}) + V_{2s}(t - T_{Do}) \]  
\[ V_{2f}(t) = V_{1s}(t - T_{De}) - V_{2s}(t - T_{Do}) \]  

(3.9)

Extracting the signals using the following decoder matrix \( S \):

\[ S = T^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \]  

(3.10)
results in the following decoded voltages:

\[
V_{1d}(t) = \frac{1}{2}(V_{1f}(t) + V_{2f}(t))
\]
\[
V_{2d}(t) = \frac{1}{2}(V_{1f}(t) - V_{2f}(t))
\]

Substituting \(V_{1f}, V_{2f}\), we obtain:

\[
V_{1d}(t) = \frac{1}{2}(V_{1s}(t - T_{De}) + V_{2s}(t - T_{Do}) + V_{1s}(t - T_{De}) - V_{2s}(t - T_{Do}))
\]
\[
V_{2d}(t) = \frac{1}{2}(V_{1s}(t - T_{De}) + V_{2s}(t - T_{Do}) - V_{1s}(t - T_{De}) + V_{2s}(t - T_{Do}))
\]

Finally, we get:

\[
V_{1d}(t) = V_{1s}(t - T_{De})
\]
\[
V_{2d}(t) = V_{2s}(t - T_{Do})
\]

This shows that both signals can be perfectly reconstructed, with no crosstalk between them, since the two orthogonal modes that the signals are mapped on do not couple with each other. Note that the reconstructed signals will exhibit different time delays, due to different times of flight of the modes they were mapped on.

### 3.2 General Multiline System: A Frequency-Domain Formulation

In the previous subsection, a time-domain modal signaling example for a two-line lossless system was given. However, a more realistic scenario which would involve multiple nonuniform lines with discontinuities is better described in frequency domain, as an extension of the modal analysis shown previously.

Referring to the system diagram as shown in Fig. 3.2, \(H\) is the channel (diagonalized using generalized modal decomposition as shown previously); \(G_{MIMO}\) and \(Q_{MIMO}\) represent the transmitter and receiver MIMO structures, with each having an encoder or decoder
block, respectively, along with the optional SISO equalizer blocks, if additional equalization of the equivalent modal channels is needed (due to losses and/or ISI produced by internal reflections, as will be discussed later); and \( n(f) \) represents the random noise at the receiver input (ignored for this initial analysis).

![Figure 3.2: The blocks of a generalized MIMO modal transceiver system, including the channel \( H \).](image)

The voltage vectors associated with the \( N \)-input/\( N \)-output system are as follows:

\[
V_s = [V_{1s} \, V_{2s} \, \ldots \, V_{Ns}]^T, \quad V_n = [V_{1n} \, V_{2n} \, \ldots \, V_{Nn}]^T \\
V_f = [V_{1f} \, V_{2f} \, \ldots \, V_{nf}]^T, \quad V_d = [V_{1d} \, V_{2d} \, \ldots \, V_{Nd}]^T
\]

We can introduce the modal voltage vector associated with the \( N \)-line system as:

\[
V_{mn} = EV_n
\]

where \( V_n \) is the line voltage vector and \( V_{mn} \) is the modal voltage vector at the near end. Matrix \( E \) is the voltage eigenvector matrix associated with the multiconductor system, as discussed in the previous sections. In general, \( E \) will be complex and a function of frequency. The modal voltage vector at the far end \( V_{mf} \) will be given by:

\[
V_{mf} = H_d V_{mn}
\]
where $H_d$ is the complex propagation matrix function given by:

$$H_d = \text{diag}\{e^{-\alpha_1 l - j\beta_1 l}, e^{-\alpha_2 l - j\beta_2 l}, \ldots e^{-\alpha_N l - j\beta_N l}\}$$  \hspace{1cm} (3.17)

where $\alpha_i + j\beta_i$ is the complex propagation constant associated with the $i$-th mode and $l$ is the length of the lines. In terms of near-end signals this will give:

$$V_{mf} = H_d E V_n$$  \hspace{1cm} (3.18)

The far-end line voltage vector $V_f$ can be recovered using:

$$V_f = E^{-1} V_{mf} = E^{-1} H_d E V_n$$  \hspace{1cm} (3.19)

Now, assume that the information signals $V_s$ are encoded before transmitting with the encoder $T$ such that each of signals is mapped to one of the linearly independent modes, and optionally equalized using a linear equalizer $G_{SISO}$, as follows:

$$V_n = TG_{SISO} V_s$$  \hspace{1cm} (3.20)

At the far end the decoded voltage vector would be given by:

$$V_d = Q_{SISO} T^{-1} V_f$$  \hspace{1cm} (3.21)

where $Q_{SISO}$ is an equalization matrix representing any equalizer that might be implemented at the output of the channel. Using (3.19), we get:

$$V_d = Q_{SISO} T^{-1} E H_d E^{-1} T G_{SISO} V_s$$  \hspace{1cm} (3.22)
Now, if we choose $T = E^{-1}$, we obtain:

$$V_d = Q_{SISO}H_dG_{SISO}V_s \quad (3.23)$$

Or stated explicitly, in terms of decoded and source signals:

$$
\begin{bmatrix}
V_{1d} \\
V_{2d} \\
\vdots \\
V_{Nd}
\end{bmatrix} = Q_{SISO}
\begin{bmatrix}
e^{-\alpha_1-j\beta_1l} \\
e^{-\alpha_2-j\beta_2l} \\
\vdots \\
e^{-\alpha_N-j\beta_Nl}
\end{bmatrix}
G_{SISO}
\begin{bmatrix}
V_{1s} \\
V_{2s} \\
\vdots \\
V_{Ns}
\end{bmatrix} \quad (3.24)
$$

In addition, we can implement one or both of the presumed SISO equalizers to remove channel attenuation (the real part of the diagonal coefficients of diagonal matrix $H_d$). The combination of the equalizers can be chosen to give the following response:

$$Q_{SISO}G_{SISO} =
\begin{bmatrix}
e^{+\alpha_1} \\
e^{+\alpha_2} \\
\vdots \\
e^{+\alpha_N}
\end{bmatrix} \quad (3.25)$$

Substituting this expression into (3.24) would remove channel attenuation, and ultimately result in:

$$
\begin{bmatrix}
V_{1d} \\
V_{2d} \\
\vdots \\
V_{Nd}
\end{bmatrix} =
\begin{bmatrix}
e^{-j\beta_1l} \\
e^{-j\beta_2l} \\
\vdots \\
e^{-j\beta_Nl}
\end{bmatrix}
\begin{bmatrix}
V_{1s} \\
V_{2s} \\
\vdots \\
V_{Ns}
\end{bmatrix} =
\begin{bmatrix}
e^{-j\omega_l\frac{v_1}{p_1}} \\
e^{-j\omega_l\frac{v_2}{p_2}} \\
\vdots \\
e^{-j\omega_l\frac{v_N}{p_N}}
\end{bmatrix}
\begin{bmatrix}
V_{1s} \\
V_{2s} \\
\vdots \\
V_{Ns}
\end{bmatrix} \quad (3.26)
$$
where we used the relation $\beta_i = \omega / v_{pi}$, where $v_{pi}$ are the velocities of propagation of individual modes. This shows that if the proper encoder, decoder and equalizer can be implemented, all signals can be perfectly reconstructed, with no crosstalk, no attenuation and no dispersion. In a practical system some level of signal degradation can be tolerated, depending on the receiver sensitivity and detection algorithm used. Note that the source signals will not suffer from signal integrity degradation, but will arrive at the receiver with different overall delays, which would need to be taken into account if a clock recovery or synchronization scheme is implemented.

3.3 Modal Signaling System Structure

With the previous analysis in place, we are now in a position to suggest a physical system which would take advantage of the described modal properties. A straightforward implementation for unidirectional signaling which matches the modal encoder and decoder matrices with the MIMO blocks is given in Fig. 3.3, and can easily be generalized to a bidirectional implementation.

![Unidirectional modal signaling block diagram for an N-line channel.](image)

Figure 3.3: Unidirectional modal signaling block diagram for an N-line channel.

However, an additional block needs to be added to the system, namely, a termination network at the receiving end of the system. This block is crucial to the system performance,
for the following reason: in the modal propagation analysis of previous subsections, the translation from modal space to line space was shown to operate on the total modal voltages present at the receiving end. In case of a properly matched system (no modal reflections at the far end), the signals in modal space at the far end would simply be delayed versions of transmitted modal signals carrying information. In case of any reflections present at the termination, however, the signals would represent the superposition of the incoming waves and the reflected ones, suffering modal redistribution in the process, as shown in Fig. 3.4 [34]. This modal conversion process would translate into crosstalk between modal channels, and therefore into crosstalk between decodes signal as well.

Figure 3.4: Frequency domain modal propagation model in matrix form.

The signaling system configuration proposed in Fig. 3.3 is not the only option for taking advantage of the linear independence of modes. If only the receiver-side circuitry was allowed (effectively using encoder of $T=I$), and ignoring the equalizer blocks, the decoder block $D$ (previously equal to $T^{-1}$) would then have to implement the following:

$$D = E^{-1}H_d^{-1}E = H^{-1}$$  \hspace{1cm} (3.27)
This implementation effectively performs the channel inversion, which can result in amplification of noise at the frequencies where the channel exhibits stopbands, and therefore in an increased signal-to-noise ratio (SNR) after decoding. Also note that in this case timing can become an issue, since the decoder block needs to match all of the modal delays, which can be difficult to accurately know in advance. These issues should be examined in more detail prior to system implementation.

3.4 Modal Extraction and Simulation Framework

The previous sections have introduced the concept of modal signaling for elimination of far-end crosstalk. Next, we will illustrate the proposed design flow and examine the properties of individual system components using a case study of a typical memory signaling system of nonuniform lines with discontinuities in the signal path. In order to explore the performance of the proposed modal signaling configuration, a flexible simulation framework has been set up, implemented as a combination of Agilent ADS [35] and MATLAB [36] software packages.

3.4.1 Eigenvalue Decomposition of Resonant Systems

MATLAB routines have been coded to perform the generalized modal decomposition as described previously, where special care has been devoted to proper identification and extraction of modes due to the resonant nature of the system. The code can operate with RLCG per-unit-length matrix description of cascaded channel subsections (useful for rapid design space exploration), as well as with the (simulated or measured) S-parameters of the overall channel, which is particularly useful when components other than transmission lines (such as vias and connectors) are present, for which the analytical description may be overly complex or unavailable.

It should be noted that, when analyzing resonant systems such as nonuniform interconnects, special attention needs to be given to the eigenvalue decomposition. Namely, even
if the channel matrix can be successfully diagonalized at each frequency of interest, the re-
sulting eigenvector/eigenvalue pairs obtained from standard commercial eigensolvers such
as MATLAB are not guaranteed to be sorted in any particular order. However, to pro-
vide insight into mapping the information bits onto transmission line modes, the order of
eigenvectors needs to be consistent across the entire frequency range. Otherwise, the ex-
tracted modal propagation constant will not exhibit the expected physical behavior. This
is shown in the example of Fig. 3.5, where the propagation velocity of modes extracted by
the MATLAB eigensolver without correction seems to vary in an inconsistent way across the
frequency range.

If not post-processed appropriately, this will lead to pessimistic results and misinterpre-
tation of the encoder/decoder coefficient values, potentially resulting in a non-realizable
system. One such example is shown in Fig. 3.6, where the improperly sorted eigenvec-
tors would seem to indicate that voltage modal encoding coefficients change by an order of
magnitude between system resonances at approximately 8 GHz and 16 GHz.

The solution implemented in this work therefore uses an additional step of per-frequency
sorting of eigenvectors based on the minimum Euclidean distance (in an N-dimensional
space) from the previous frequency point. The main code of this post-processing routine is
provided in Appendix A.

3.4.2 Channel Parameter Extraction and Simulation

The ADS portion of the framework is used for quasi-static matrix $RLCG$ extraction, as
well as analog simulation of the complete system. Statistical eye diagram simulations of
the linear system provide us with a metric of the performance of the proposed signaling
schemes at very low bit error rates needed (in modern chip-to-chip signaling systems, down
to one error in $10^{21}$ bits, or less), which would be impossible to simulate in a reasonable
amount of time using traditional transient simulation techniques. The encoder and decoder
blocks have been implemented as ideal formula-based linear combination elements. This approach is adequate for the channels under investigation and data rates of several Gb/s, since there is not a lot of frequency variation in encoder and decoder coefficients, and their imaginary components are negligible, even with the discontinuities in place [37]. Terminations can be implemented as a resistive grid, as the S-parameter description of the extracted
frequency-dependent impedance matrix, or as a pole/residue description of a passive, physically realizable, frequency dependent network. All three approaches are used in following simulations.

3.5 Summary

This chapter introduced the concept of modal signaling, first using an illustrative two-line time-domain example, and then describing the signaling using the more general frequency-domain formulation for a general multiline system. Based on the results of channel diagonalization obtained using generalized modal decomposition theory, the elements of the system were identified to be the encoder, the decoder, and the termination network, and their relation to eigenvectors of the system was derived. The setup of a flexible framework for modal decomposition and signaling simulation was described in detail, since special attention needs
to be devoted to diagonalizing nonuniform channels, due to the resonant behavior inherently present in the system.
CHAPTER 4

MODAL SIGNALING OVER NONUNIFORM COUPLED INTERCONNECTS WITH DISCONTINUITIES

With the generalized modal theory of the previous section in place to provide us with the frequency-dependent modal waveshapes and the characteristic admittance matrix, we are now ready to explore its applicability to crosstalk-free signaling. We will analyze the feasibility of using fundamental transmission line modes for high-speed signaling by applying modal decomposition to realistic channels with discontinuities. Special attention is devoted to proper matching of the propagation modes, due to the resonant nature of the cascaded system which results in characteristic admittance matrix (and therefore the optimal matching network) exhibiting strong frequency dependence. Also analyzed are some of the relevant system-level characteristics of the building blocks in terms of robustness, immunity to noise and required implementation precision.

4.1 Case Study A: Uniform Interconnects

In order to gain insight into the effects of discontinuities on modal propagation, a typical uniform line interconnect system will be briefly examined first. The uniform system consists of four tightly coupled PCB copper traces on FR-4 substrate, whose cross-section is as shown in Fig. 4.1, with the length of the lines L=4 inch, and other parameters set as given in Table 4.1.

After applying the decomposition procedure, eigenvalues of the four distinct propagation modes are extracted over the frequency range DC–20 GHz. Next, the modal propagation constants are calculated to examine the frequency dependence of the mode attenuation
constant $\alpha$ and the phase constant $\beta$, using the following relation:

$$\lambda_{F,B} = e^{\pm \gamma L}, \text{ where } \gamma = \alpha + j\beta$$  \hspace{1cm} (4.1) 

As observed in Fig. 4.2, where marker families correspond to frequency data points ($\times=10$ GHz, $\circ=20$ GHz), the modal propagation constants exhibit no resonances across the frequency range, as expected. This translates to the values of all the waveshapes and the corresponding matching termination network being nearly constant over the entire frequency range, with the imaginary component several orders of magnitude below the real component. The forward voltage waveshapes and termination resistor values, as needed for realization of a multimode signaling system, are summarized in Table 4.2.A and Table 4.3, respectively.
4.2 Case Study B: Cascaded Nonuniform Interconnects

A typical chip-to-chip interconnect system is shown in Fig. 1.2 [16]. Ignoring the impact of vias and solder balls in the initial analysis, it consists of four traces in each of the three cascaded uniform line sections, representing the controller package, PCB and the memory.
Table 4.3: Termination resistor values used for modal matching.

<table>
<thead>
<tr>
<th>Termination resistors (Ohm)</th>
<th>Uniform (PCB)</th>
<th>Uniform (Package)</th>
<th>Cascaded (100 MHz)</th>
<th>Cascaded (optimized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R11, R44</td>
<td>87.3</td>
<td>125.8</td>
<td>97.5</td>
<td>101</td>
</tr>
<tr>
<td>R22, R33</td>
<td>108.6</td>
<td>187.1</td>
<td>123.9</td>
<td>131</td>
</tr>
<tr>
<td>R12, R34</td>
<td>273.6</td>
<td>181.2</td>
<td>267.1</td>
<td>280</td>
</tr>
<tr>
<td>R23</td>
<td>280.6</td>
<td>187.9</td>
<td>277.2</td>
<td>312</td>
</tr>
<tr>
<td>R13, R24</td>
<td>2955.4</td>
<td>2133.7</td>
<td>2318.2</td>
<td>1450</td>
</tr>
<tr>
<td>R14</td>
<td>5904.6</td>
<td>4451.4</td>
<td>4901.2</td>
<td>2453</td>
</tr>
</tbody>
</table>

package. The dimensions and the cross-section of the PCB segment are the same as for the uniform line case, while the two package segments have lengths of $L_{cont}=18 \text{ mm}$, $L_{mem}=8 \text{ mm}$, and other parameters set as given in Table 4.1.

After applying the decomposition procedure, modal propagation constants of the four distinct propagation modes are shown in Fig. 4.3, where marker families correspond to frequency data points ($x=10 \text{ GHz}$, $o=20 \text{ GHz}$). The plots reveal the presence of frequency bands exhibiting a stopband behavior, with the central frequencies, width and magnitude of the stopbands varying from mode to mode, as dictated by the interaction of modes in different cascaded segments and the overall geometry of the system [32].

It can be observed that the propagation mode 4 (the "ground" mode, whose voltage eigenvector has all the signs positive, as shown in Table 4.2.B) is propagating at the slowest velocity, and that it is attenuated at resonances much more than the other modes. In contrast, mode 2, with to the two inner conductors switching in the direction opposite to the outer two (as shown in Table 4.2.B), is almost unaffected by the introduction of discontinuities.

As shown in Fig. 4.4(a), all four forward voltage eigenvectors are almost constant with frequency over a very large range (approximately up to 8 GHz), with the imaginary component (not shown) two to three orders of magnitude below the real component. This justifies a simple approximation of the complex frequency-dependent encoder and decoder matrices
with their constant real values extracted at low frequencies. Table 4.2.B shows the values of the forward voltage eigenvectors. Note that, compared to the values for the uniform channel, the eigenvector values exhibit only a small amount of change, due to the cross-section of both types of cascaded sections having similar geometric structures.

In contrast, Fig. 4.4(b) reveals a strong frequency dependence of the forward current eigenvectors, due to the channel being dominantly inductively coupled. According to (2.10), this will also hold for the values of the characteristic impedance (or admittance) matrix. In order to eliminate reflections of the incoming modal waves, the lines would need to be terminated with the matching admittance which is equal to the characteristic admittance.
The signaling system from Fig. 3.3, if realized with the optimal voltage encoder and decoder matrix and the termination network extracted as given here, would itself be optimal from the modal signaling perspective, and should theoretically provide complete crosstalk elimination, creating N uncoupled equivalent channels using N signaling lines. To verify the extraction results of eigenmodes and terminations, and also to observe the limit on system performance, a modal signaling simulation framework was created in Agilent ADS.
4.2.1 Simulation Results

To analyze the impact of stopbands and terminations on the signaling performance of the channel, the setup of Fig. 3.3 has been implemented as discussed previously, using the encoder coefficients from Table 4.2.B and resistive terminations optimized for minimal modal reflection from Table 4.3. Assessing the impact of the channel alone (with no capacitive loading), we observe the S-parameters of direct and crosstalk paths for a cascaded channel, shown in Fig. 4.5 for uncoded, and in Fig. 4.6 for the modal-coded channels. Ports 1-4 and 5-8 correspond respectively to near- and far-end terminals of the system.
From Fig. 4.5, we can further confirm that the uncoded system is crosstalk-dominated, with the magnitude of crosstalk approaching that of the direct signal paths already around the frequency of 2 GHz. This is confirmed by the inset in Fig. 4.5 of a statistical eye diagram of a 4 Gb/s NRZ signal: observing the inner line when all four lines are transmitting unsynchronized source signals, we observe a completely closed data eye due to crosstalk.

In contrast, from Fig. 4.6 we note that crosstalk is significantly suppressed using modal encoding, with a guard band of at least 25 dB at frequencies up to 2.4 GHz, which was our primary motivation for exploring modal signaling. We also note that modes are attenuated to varying degrees, in particular the mode labeled as 4 (also called "ground mode," with
all the lines switching in the same direction), exhibiting stopbands. This attenuation is due to the inherent modal propagation properties in a cascaded system, and corresponds to the attenuation shown in the modal propagation constant plot of Fig. 4.3. We will explore the implications of this when discussing the low-order approximation to the extracted optimal termination network.

4.2.2 Impact of Discontinuities

Figure 4.7: Real part of (a) voltage and (b) current forward eigenvector coefficients for the cascaded channel with discontinuities.

The previous analysis concentrated on the impact of cascading different sections of uniform transmission lines. However, a realistic signaling channel also includes a number of non-transmission line artifacts in the signal path, such as vias, solder balls, connectors and
other discontinuities. It is important to verify the impact of those discontinuities on modal decomposition and the assumption of voltage encoding matrix having constant values.

Performing the eigenvalue decomposition on the complete channel model from Fig. 1.2 results in forward voltage- and current-eigenvector coefficients as shown in Fig. 4.7. The impact of discontinuities is reasonably low at lower frequencies, where the system behavior is predominantly determined by coupling over transmission line segments. At higher frequencies, the discontinuities start to introduce significant deviations from the transmission line model of the channel structure, potentially limiting the overall signaling speed using frequency-independent coefficients. Fortunately, in this case the frequency band DC-2.4 GHz (which contains most of the 4 Gb/s NRZ signal energy) is still well-approximated using constant coefficient assumption.

4.3 Frequency-Dependent Optimal Termination Network

As stated before, for the $N$-line bundle there will be $N$ fundamental propagation modes, resulting in the voltage- and current-waveshape matrices with dimensions $N \times N$. Therefore, the characteristic admittance matrix given by (2.10) will also be of dimensions $N \times N$. The admittance matrix can be decomposed into a complete graph where each branch represents a two-port admittance, as follows:

$$ Y_{C,F} = \begin{bmatrix} Y_{C,F11} & Y_{C,F12} & \cdots & Y_{C,F1N} \\ Y_{C,F21} & Y_{C,F22} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ Y_{C,FN1} & \cdots & \cdots & Y_{C,FNN} \end{bmatrix} $$

(4.2)
If the characteristic admittance matrix $Y_{C,F}$ from (2.10) is written as:

$$
Y = \begin{bmatrix}
\sum_{i=1}^{N} Y_{C,Fi1} & -Y_{C,F12} & \ldots & -Y_{C,F1N} \\
-Y_{C,F21} & \sum_{i=1}^{N} Y_{C,Fi2} & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots \\
-Y_{C,FN1} & \ldots & \ldots & \sum_{i=1}^{N} Y_{C,FiN}
\end{bmatrix}
$$

(4.3)

then it is proven by [30] that the termination admittance network is given by:

$$
Y = \begin{bmatrix}
\sum_{i=1}^{N} Y_{C,Fi1} & -Y_{C,F12} & \ldots & -Y_{C,F1N} \\
-Y_{C,F21} & \sum_{i=1}^{N} Y_{C,Fi2} & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots \\
-Y_{C,FN1} & \ldots & \ldots & \sum_{i=1}^{N} Y_{C,FiN}
\end{bmatrix}
$$

(4.4)

In general, for an $N$-line bundle the termination admittance network will consist of a total of \( \binom{N}{2} \) two-port network elements, creating a complete grid between all the lines and the reference terminal. The frequency dependence of those elements is given by individual matrix elements of (4.4), where $Y_{ij} = Y_{ji}$ is the two-port element placed between lines $i$ and $j$ if $i \neq j$, or between line $i$ and the reference terminal if $i = j$. For a lossless uniform interconnect system, the values of the two-port elements are purely real and constant with frequency, representing the conductances of resistors which will allow matching for all the transmission line modes of such a system, as shown in Fig. 4.8.

In the case of an interconnect system with discontinuities, this decomposition can still be performed, but the resulting two-port elements will exhibit frequency dependence. In cases where an approximate match of the system is sufficient for target signaling performance, the low-frequency values of the matrix elements from (4.4) can be used to extract the resistive...
grid which will match the characteristic admittance of the system over a limited bandwidth. Starting from those values, resistor values could also be optimized to minimize overall total reflections [28], as done in the initial case study. However, it will be shown that any purely resistive approach is sub-optimal from the point of view of proper termination of all the modal waves, and can impact the signal integrity of the received modal signals and potentially result in the increase in bit error rate after decoding. Instead, a frequency-dependent approximation of the optimal termination network should be used for complete elimination of modal reflections.

4.3.1 Low-Order Model of the Optimal Termination Network

Examining the frequency dependence of characteristic admittance matrix given by (2.10), we notice that, for the nonuniform channel of the Case Study B, the matrix will exhibit periodic frequency dependence due to its constituent transmission line segments producing a periodic response with respect to frequency. Therefore, the modeling process needs to start with determining an appropriate frequency range BW over which the approximation is performed, with respect to the spectral content of signaling that is to be used in the system. The admittance matrix given in (4.4) is then approximated by a rational function, with the
approximation accuracy (order) over the frequency range BW chosen to achieve the target performance of the signaling system. This approximation can be performed using the vector fitting method [38], which will produce the poles and residues of the low-order termination network.

For the case study, a simple NRZ signaling with the data rate of 4 Gb/s and rise/fall times of 67 ps is assumed. Based on the spectrum of NRZ signals, frequency range DC-2.4 GHz was assumed to contain most of the signal energy. The $S$-parameter representation of the characteristic admittance matrix as extracted from (4.4) was then fitted over the 2.4 GHz bandwidth to a lower-order representation of poles and residues using a commercial macromodeling tool [39]. The order of the approximation was chosen to obtain a eye opening within 1% of that achieved by the extracted optimum terminations, as verified by statistical eye generation. The required model order was 14 poles per port, out of which two were real and six were complex-conjugate, as given in Table 4.4.

Due to the reciprocity of the passive channel, as well as the symmetrical configuration of the lines, only 6 out of 10 $Y$-parameters of the termination network will be distinct; for example, we can choose to observe $Y_{23}$, $Y_{22}$, and $Y_{11}$ through $Y_{14}$. The other elements will then be given by:

$$Y_{33} = Y_{22}, Y_{44} = Y_{11}, Y_{24} = Y_{13}, Y_{34} = Y_{12}$$ (4.5)

<table>
<thead>
<tr>
<th>Pole</th>
<th>Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Real</td>
<td>-5.166E+10</td>
</tr>
<tr>
<td>2</td>
<td>Real</td>
<td>-5.476E+6</td>
</tr>
<tr>
<td>3,4</td>
<td>Complex</td>
<td>-3.442E+8±j*1.426E+10</td>
</tr>
<tr>
<td>5,6</td>
<td>Complex</td>
<td>-4.179E+8±j*1.247E+10</td>
</tr>
<tr>
<td>7,8</td>
<td>Complex</td>
<td>-2.806E+8±j*9.695E+9</td>
</tr>
<tr>
<td>9,10</td>
<td>Complex</td>
<td>-3.844E+8±j*8.259E+9</td>
</tr>
<tr>
<td>11,12</td>
<td>Complex</td>
<td>-1.670E+8±j*4.861E+9</td>
</tr>
<tr>
<td>13,14</td>
<td>Complex</td>
<td>-2.502E+8±j*4.072E+9</td>
</tr>
</tbody>
</table>
Figure 4.9: Real (top) and imaginary (center) part of the Y11 diagonal parameter of matching admittance, extracted (solid) and modeled (dashed). Mean absolute fitting error values are also shown (bottom).

It is also worth noting that, due to the resonant nature of the system, exploring the admittance matrix provides a better graphical insight into the termination frequency behavior than the (mathematically equivalent) impedance representation, which was shown in [37].

Out of six distinct Y parameters of the termination network, for clarity we chose to plot one of the diagonal terms (Y_{11}, shown in Fig. 4.9) and one of the off-diagonal terms (Y_{12}, shown in Fig. 4.10), as representative examples of the two groups. Mean absolute error values of the real and imaginary part of the fitted model are also shown in Fig. 4.9.

From the plots of mean error values as the result of vector fitting, shown in Fig. 4.9, we can observe that the system resonances impact the values of the termination network matrix mainly at lower frequencies, with the effect less pronounced at the high end of the spectrum. It is also worth noting that an adequate model which exhibits a very accurate match of each pole of the optimal termination admittance matrix will alleviate the signal attenuation at the frequencies of system resonance, as were shown previously in Fig. 4.6.
Figure 4.10: Real (top) and imaginary (center) part of the Y12 off-diagonal parameter of matching admittance, extracted (solid) vs modeled (dashed).

4.3.2 System Performance with the Frequency-Dependent Terminations

A simple approach of using the low-frequency resistive values (before the first of resonances, in this case at the frequency chosen as 100 MHz) was explored in detail in [37]. This approach is commonly used to terminate low-loss uniform line bundles, since their optimal terminations will be purely real and frequency independent. It will be contrasted with using the developed low-order model in the following section.

Note that the low-frequency termination network values for the cascaded case will differ significantly from those of both its constituent sections [37], due to ABCD-parameter representation averaging out the nonuniformity of the channel structure [32]. Therefore, a termination approach based only on matching the PCB or package channel impedance matrix would result in suboptimal signaling performance.

To demonstrate the impact of modeled optimal terminations on the signaling performance of the channel, the simulation has been implemented in Agilent ADS, with the system setup
as discussed in previous sections. The S-parameters of the modal-coded channel with modeled optimal terminations are shown in Fig. 4.11. Ports 1-4 and 5-8 correspond respectively to near- and far-end terminals of the overall system. Signals on input terminals were mapped to propagation modes in sequential order: binary signal at port 1 was mapped to excite propagation mode 1, etc.

![Graph showing S-parameters](image)

Figure 4.11: S-parameters of direct and strongest crosstalk paths for the modal encoded cascaded channel with modeled optimal terminations.

From the S-parameter plot given in Fig. 4.11 (where the parameters not shown are below -100 dB), we note that crosstalk of modes 1-3 is significantly suppressed using modal encoding, with a guard band of at least 30 dB at frequencies up to 2.4 GHz, and also that modes are attenuated to varying degrees in particular the ground mode. This attenuation is a
combination of resonances due to the inherent modal propagation properties and the far-end reflections. As will be shown next, the attenuation due to reflections is minimized compared to the one produced by low-frequency resistive termination network.

For comparison purposes, the detail of the transfer characteristic of the optimally terminated modal signaling system is shown in Fig. 4.12. It corresponds to the plot of propagation constant magnitude $\lambda_F$ from (4.1), revealing the presence of stopbands, as expected from the plot of the mode attenuation constant of Fig. 4.3. This gives the limits of performance of modal signaling alone, when using the system approach of Fig. 3.3.

The transfer characteristic of the modal signaling system terminated with low-order model of the optimal termination network is shown in Fig. 4.13. As expected, the response with the synthesized network matches that of the optimal (shown in Fig. 4.12) over the range
Figure 4.13: S-parameters of direct signal paths for the modal encoded cascaded channel with low-order model of the optimal terminations.

of interest, up to 2.4 GHz. Outside of this band, the modal signal attenuation due to far-end reflections is not controlled, and exhibits a behavior pattern similar to the system with low-frequency extracted resistive terminations, as shown in Fig. 4.14.

Based on the transfer characteristic for the two scenarios, we can expect the performance of the system terminated with the modeled optimal terminations to be superior to the one which uses a resistive network, due to the impact of reflections. This is explored via the statistical eye diagrams of decoded signals for the two scenarios, as shown in Figs. 4.15 and 4.16. Single-ended NRZ signals of unit amplitude are used as the encoder inputs, with the bitrate of 4 Gb/s and rise/fall times of 67 ps. All the lines are being excited with random switching phases. In both cases, the observed impact on CIJ is minimal, which confirms the crosstalk mitigation predicted by the S-parameters. However, when using a resistive network,
modal signals are susceptible to the reflections arising from the non-perfect matching, the ground mode in particular, resulting in a significantly reduced vertical eye opening. The improvement for the ground mode obtained by using the low-order model of the optimal frequency dependent network is the vertical eye opening increase of 39% at the optimal sampling point, and the reduction in peak-to-peak jitter of 27%.

This demonstrates that a low-order network synthesis method accurately approximates the optimal network over the bandwidth of interest [37]. The performance of such a system is shown to be superior to that of the system using a full resistor matrix, in particular for the ground mode (exhibiting the most resonances). The model order of the synthesized system can be chosen to trade off performance for design complexity if so desired.
Figure 4.15: Decoded modal signals corresponding to modes 1-4 for the modal encoded cascaded channel with resistive terminations.

4.4 Some System-Level Metrics of Modal Signaling

As was demonstrated in the previous discussion, diagonalizing the channel using modal decomposition could be used to successfully mitigate the crosstalk. However, a practical realization of such a signaling system will consist of realistic circuit elements whose performance will unavoidably deviate from the ideal mathematical modal decomposition operations, in terms of precision, linearity, and quantization noise introduced to the system. Any deviation from these ideal coefficients introduced by circuit non-idealities will translate to redistribution of energy mapped onto modes and therefore to equivalent modal noise after decoding, resulting in voltage noise, crosstalk and timing jitter. Along with these deterministic noise sources, also present in the system is the random noise induced by both active (transistors)
and passive (thermal noise of the resistors) circuit elements. Random noise is not a limiting factor in today’s high-speed crosstalk-dominant links, but may prove to be a major factor due to low-power design trends [40]. In order to meet the signaling performance at low bit-error rates (typically less than $10^{-12}$, and lower than $10^{-15}$ for state-of-the-art short-haul links [41]) required of modern high-speed signaling systems, those issues need to be examined in detail.

4.4.1 MIMO System Perspective of Modal Signaling

As discussed in the previous sections, the eigenvalue decomposition of the interconnect system is performed starting from the channel description or measurements. Channel matrix
\( H(f) \) is decomposable into the following form:

\[
H(f) = W^{-1}(f)H_m(f)W(f)
\]  \hspace{1cm} (4.6)

where \( W(f) \) is the matrix of current- or voltage-waveshapes (eigenvectors) used for signaling, and \( H_m(f) \) is the diagonal matrix describing modal propagation over total length \( l \), in terms of attenuation coefficients \( \alpha \) and propagation delays \( \beta \):

\[
H_m(f) = \text{diag}(e^{-\alpha l - j\beta l})
\]  \hspace{1cm} (4.7)

The block diagram of the complete signaling MIMO system (revised from Fig. 3.2 to include the input-referred noise at the receiver input), showing the transmitter \( GMIMO(f) \), channel \( H(f) \) and the receiver \( YMIMO(f) \) is shown in Fig. 4.17.

![Figure 4.17: Modal signaling system MIMO block diagram, including noise.](image)

This multiple-input-multiple-output (MIMO) system takes as its input a vector of independent input signals \( x(t) \) and produces the decoded versions of the signals \( x'(t) \) as its output. The goal of modal signaling is to create the set of uncoupled equivalent modal channels by appropriately selecting the encoder and decoder matrices. The encoder \( E \) maps the energy of information signals onto fundamental modes of propagation. At the input of the receiver this energy is mapped back into uncoupled channels via the decoder \( D \). Optionally, the uncoupled channels can be equalized on a per-channel basis using equalization at the transmitter \( G_{SISO}(f) \) and/or receiver \( Y_{SISO}(f) \) to account for intersymbol interference (ISI)
due to limited channel bandwidth; in this analysis, a low-loss short channel is assumed, for which equalization is not needed, but it can be added without loss of generalization.

### 4.4.2 Decoder Coefficients for Near-Optimum SNR

In the absence of random noise $n(t)$, the optimal encoder and decoder matrices would be determined solely by channel characteristics. In order to keep the complexity of the decoder realization at a reasonable level, the encoder and decoder matrices can be chosen as $E = W^{-1}$ and $D = W$ to remove coupling:

\[
x' = D(W^{-1}H_mW)Ex = WW^{-1}H_mWW^{-1}x = H_mx
\] (4.8)

while leaving the modal delay and attenuation present in each individual modal channel, to be accounted for by synchronization circuitry and receiver-equalization if needed. However, this approach, which effectively inverts the channel matrix, does not take into account the effect of random noise present in the system. If we refer this noise at the input of the receiver, there is a danger that the decoding operation might also augment the level of noise and therefore decrease the overall signal-to-noise ratio (SNR) at the decision circuitry.

In order to investigate the optimal decoder matrix in presence of noise, we can relate the proposed system to the model given in [42], where modal signaling is mapped onto a flat-fading MIMO system over a quasi-static channel with equal number of transmit and receive antennas, equal to the number of fundamental modes of the channel (for an $N$-line interconnect bundle of coupled cascaded segments and a reference plane, this number is equal to $N$). The decoding approach which does not take random noise into account would then be mapped into a zero-forcing (ZF) interface [43] as follows (labeling $M = EH_m$):

\[
D = (M^*M)^{-1}M^* = M^{-1}(M^*)^{-1}M^* = M^{-1}
\] (4.9)
where $M^*$ denotes the Hermitian conjugate of $M$.

There are different decoding (detection) strategies that take into account the effect of noise, but the optimum procedures (decision feedback, sphere decoding) suffer from increased complexity [44]. A reasonable tradeoff for high-speed communication links is to utilize a linear minimum mean-square error (LMMSE) interface [43], which minimizes the joint effect of off-diagonal elements of modal channel matrix and of the filtered noise:

$$D = (M^*M + \frac{N_0}{E_S}I)^{-1}M^*$$ \hspace{1cm} (4.10)

where $E_S$ is the average energy of one component of source vector $x$, $N_0$ is the noise variance and $I$ is the $N \times N$ identity matrix. In the limiting case of noise being dominant, the LMMSE interface will assume the form of:

$$D = (\frac{N_0}{E_S}I)^{-1}M^* = \frac{E_S}{N_0}M^*$$ \hspace{1cm} (4.11)

Depending on the particular channel in question, the two criteria can result in very different decoder coefficients. In order not to enhance random noise present in the channel, the encoder matrix can therefore be chosen based on the eigenvalue decomposition of the channel, as $E = W^{-1}$, and the decoder matrix will be calculated according to (4.10) for the near-optimum overall SNR at the receiver.

For the special, but very important, case of a channel being a symmetric line bundle, the matrix $M$ will also be symmetric; therefore, the encoder matrix $E$ will be complex orthogonal (as stated by Theorem (4.4.13) from [45]). For the low-loss channel case, the imaginary part of $E$ can be ignored [37], leaving $E$ as real orthogonal and therefore unitary ($E^* = E^{-1}$). Also, the modal propagation matrix $H_m$ will be purely imaginary, in addition to being diagonal. In this case:

$$M^*M = H_m^*E^*EH_m = H_m^*E^{-1}EH_m = I$$ \hspace{1cm} (4.12)
which leads to the LMMSE equation 4.10 rewritten as follows:

\[
D = \left( I + \frac{N_0}{E_S} I \right)^{-1} M^* = kM^* = kH_m^* E^* = kH_m^{-1} E^{-1} = kM^{-1}
\] (4.13)

Therefore, in this case the two interfaces will differ only by a scaling factor \( k \), which can be chosen without affecting SNR, depending on the suitable circuit realization (i.e. whether the decoder coefficients with less than unity gain can be implemented accurately). This also reveals a very important property of the low-loss symmetric channel (which can have cascaded segments, as long as the symmetry of the overall channel matrix is maintained), for which the eigenvalue decomposition is equal to the singular value decomposition.

![Figure 4.18: ZF (X) and LMMSE (o) coefficients for the 4-line channel.](image)

The difference between decoder coefficients chosen using ZF or near-optimum LMMSE criteria will be illustrated by applying them to two PCB channels, one symmetric (as given in 4.1), and the other asymmetric, with widths of traces changed from the uniform 173 to 150, 200, 100 and 80 \( \mu m \), and their spacings changed from uniform 132 to 100, 150 and 130 \( \mu m \) respectively. The results for the 16 decoding coefficients of the symmetric system are
shown in Fig. 4.18a, and for the asymmetric system in Fig. 4.18b. In both cases the noise variance $N_0$ was varied from 0 (no random noise) to 100 times $E_S$ (noise-dominated system). We can observe that, in case of the symmetric bundle, the ZF and LMMSE coefficients converge to the same normalized values, whereas the two interfaces produce significantly different values in the case of the asymmetric bundle.

4.4.3 Impact of Common and Input-Referred Noise

With the near-optimal encoder and decoder matrices chosen as discussed in the previous section, the impact of noise present in the system to overall SNR of the decoded signal can now be analyzed. Referring the noise vector $n$ to receiver input, as shown in Fig. 4.17, we have, after decoding:

$$x' = D(W^{-1}H_mW)Ex + Dn = H_mx + Dn$$  \hspace{1cm} (4.14)

From the above, it can be observed that the signal energy is a function only of the properties of the channel, via the modal propagation matrix $H_m$, while the noise of each decoded signal will depend both on the decoder matrix $D$ (which is again a function of the channel) and the noise distribution $n$. We can distinguish between two types of noise at the receiver input:

- **Correlated (common) noise**

  This type of noise is present in the system due to the close proximity of interconnects (since strong coupling is assumed). Depending on the decoder coefficients as obtained by eigenvalue decomposition, some of the modes will benefit from partial or total common noise cancelation (i.e. common-mode rejection in a two-line system using differential signaling), which is most pronounced for a channel that exhibits physical symmetry (e.g. a microstrip bus consisting of parallel uniform segments).
Table 4.5: Optimal decoding coefficients for the nonuniform system.

<table>
<thead>
<tr>
<th>Mode</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.2837</td>
<td>-0.5684</td>
<td>+0.5684</td>
<td>+1.2837</td>
</tr>
<tr>
<td>2</td>
<td>-1.0780</td>
<td>+0.9005</td>
<td>+0.9005</td>
<td>-1.0780</td>
</tr>
<tr>
<td>3</td>
<td>-0.5992</td>
<td>+1.2698</td>
<td>-1.2698</td>
<td>+0.5992</td>
</tr>
<tr>
<td>4</td>
<td>+0.9423</td>
<td>+1.0460</td>
<td>+1.0460</td>
<td>+0.9423</td>
</tr>
</tbody>
</table>

Table 4.6: SNR levels at the slicer input for common and uncorrelated input-referred noise after modal decoding.

<table>
<thead>
<tr>
<th>Mode</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNRcom [dB]</td>
<td>∞</td>
<td>58.5</td>
<td>∞</td>
<td>37.5</td>
</tr>
<tr>
<td>SNRuncor [dB]</td>
<td>38.2</td>
<td>37.6</td>
<td>38.1</td>
<td>37.5</td>
</tr>
</tbody>
</table>

- Uncorrelated (input-referred) noise

This type of noise is generated by the circuitry which realizes the decoder and other receiver blocks, such as the termination network, referred to the input of the receiver (far-end of the channel). Being uncorrelated, the powers of individual noise components on each line are multiplied with the appropriate decoder coefficients and summed up to obtain the total noise power for SNR calculation.

To demonstrate the impact of two types of noise, the cascaded nonuniform channel of Case Study B will be used, for which the optimal decoding coefficients are given in Table 4.5. We compare the modal impact of common noise of \(2 \, mV_{rms}\) at each decoder input with the uncorrelated noise of the same RMS level, assuming the decoded NRZ signal of \(300 \, mV_{pp}\), all typical values for modern high-speed off-chip signaling links. The SNR values at slicer input after decoding are summarized in Table 4.6.

From Table 4.6, it is observed that common noise is theoretically completely suppressed in modes 1 and 3, due to decoder coefficient symmetry and quasi-differential nature of those modes. Due to decoder nonlinearities, mismatch and variations, this common-mode rejection will in practice be limited to a finite value. At the same time, mode 4 suffers the most, since common noise from all four lines is directly added during the decoding operation. All the
modes will have their SNR reduced due to uncorrelated noise, because the decoder operation in the form of a simple linear combination cannot suppress this type of noise.

4.4.4 Eigenvector Coefficient Quantization

The ideal encoder and decoder coefficients obtained by eigenvalue decomposition would need to be realized with finite precision, due to physical design limitations (e.g. minimum realizable feature size), but also due to mismatch and process variations. Furthermore, coefficient implementation should be tunable in order to adapt the signaling system to different channels or to deviations from the assumed channel description. If the encoder and decoder coefficients are quantized, we can rewrite them as follows:

\[ \hat{E} = E - Q_E \]
\[ \hat{D} = D - Q_D \]  

(4.15)

where \( \hat{E}, \hat{D} \) are the actual encoder and decoders implemented in the system, which deviate from the ideal \( E \) and \( D \) by the encoder and decoder quantization errors \( Q_E \) and \( Q_D \), respectively. Now the decoded signal with encoder quantization will be:

\[ x' = H_d x - H_d E^{-1} Q_E x + D_n \]  

(4.16)

And if the decoder coefficients are quantized, we will have a similar form of the decoded signal, with the main difference being that input-referred noise is quantized as well:

\[ x' = H_d x - Q_D D^{-1} H_d x + \hat{D} n \]  

(4.17)

Propagation delay of modes will not affect the signal-to-quantization noise ratio (SQR). Total SQR can be expressed as the ratio of signal energy, which is the energy of the first terms in (4.16) and (4.17), to quantization noise energy, described by the second terms
in (4.16) and (4.17). To determine the minimum number of quantization bits for a given bit error rate (BER), a particular type of signaling and detection need to be assumed. The most commonly used form of modulation in high-speed chip-to-chip single-ended links is the simple NRZ signaling, for which the probability of error at the slicer can be related to SNR as follows [46]:

\[
P_{\text{err}} \approx \frac{1}{2} \text{erfc} \left( \frac{\sqrt{\text{SNR}}}{2\sqrt{2}} \right)
\]  

(4.18)

Plotting the BER curve from (4.18) results in the trend observed in Fig. 4.19. For a typical target error rate of \(10^{-15}\), we can observe that the overall SNR needs to be at least 24 dB, and even more for the state-of-the-art links with BER less than the \(10^{-21}\) requirement. Therefore, in order for the quantization noise not to limit the overall performance, the number of quantization bits needs to be set to a value such that SQR is significantly higher than the minimum SNR.

![SNR vs. BER](image)

Figure 4.19: BER plot as a function of SNR at the slicer for NRZ modulation.

The traditional approximative approach to treating quantization noise in filter theory is to calculate the power of quantized noise, and then add its variance to the variance of the random noise [46]. However, in quantizing the eigenvector coefficients we need to take...
into account that the quantization noise is strongly signal-dependent, especially for a small number of quantization bits.

The alternative approach proposed is to observe all combinations of signal and quantization noise as the combined signals that the slicer will operate on, and then to calculate BER based on the energy of such equivalent signals versus the random noise energy. This deterministic approach is valid if all the input vectors $x$ appear with equal probability, a fair assumption for a general-purpose memory link.

The two approaches for determining the required number of encoder bits based on a target BER will be contrasted using the system described in Case Study B. In both cases, random noise present at the slicer is set to result in a BER of $10^{-21}$ in absence of quantization noise. Calculated results starting from (4.16) are shown in Fig. 4.20.

From Fig. 4.20a, we can observe that directly summing all the noise variances produces BER estimates that are pessimistic, which would lead to a design with an unnecessarily large number of quantization bits, resulting in more area and power consumption. In contrast,

Figure 4.20: Two approaches to calculating BER of decoded signal with encoder quantization.
Fig. 4.20b reveals the average of one fewer bit needed to reduce quantization noise sufficiently to meet the target BER performance. Similar analysis can be performed for the required number of decoder bits, starting from (4.17). Note that, in both cases, more than the minimum number of bits would be needed, to account for nonlinearities, mismatch and variations in practical design.

4.5 Summary

This chapter expanded on the well-explored concept of modal signaling over uniform interconnects by contrasting it with modal signaling over realistic channels with discontinuities. This was done through case studies of representative channels for each channel class, analyzing the effect of cascaded segments, as well as vias, solderballs and other artifacts. It was shown that the modal encoder and decoder exhibit relatively weak frequency dependence, and can be approximated by constant linear combination blocks. In contrast, it was revealed that the optimal termination network for nonuniform systems possesses a strong frequency dependence. For the representative test channel, we demonstrated that the improvement of 4 Gb/s NRZ signaling over ground mode obtained with the low-order model of the optimal frequency dependent network results in the vertical eye opening increase of 39% at the optimal sampling point, and the reduction in peak-to-peak jitter of 27%. Utilizing MIMO system perspective, relevant system-level characteristics of the building blocks are analyzed in terms of robustness, immunity to noise and required implementation precision.
CHAPTER 5

PROPOSED SIGNALING SYSTEM REALIZATION

With the modal decomposition and extraction procedure of the parameters of modal signaling system blocks as detailed in previous chapters, we are now ready to discuss the issues related to the physical implementation of each individual block. The first proposed approach makes use of the digital cores, due to their straightforward applicability in performing matrix multiplication (by means of calculating linear combinations) to generate modal waveshapes. The second approach relies on custom-designed analog blocks to realize the functions of encoding and decoding, and is more suited to the shorter chip-to-chip interconnects in which the digital core is rarely present.

5.1 Modal Encoding/Decoding Using Digital Cores

The recent trend in high-speed backplane design has seen an increased use of digital cores [47], both in the receiver and the transmitter, to allow for implementation of complex equalizer schemes needed for high-speed signaling over bandwidth-limited channels. As discussed previously, modal encoder and decoder can very accurately be represented by frequency-independent linear combination blocks with coefficients constant for a given channel. If the transceiver is equipped with a digital processing block, this linear combination operation can be implemented naturally and with minimal overhead. A block diagram of one possible encoder implementation is shown in Fig. 5.1.

A digital implementation of the encoder was proposed in [48], but there are practical issues with either of the two suggested methods. If the signals to be sent from the transmitter are in
analog form, the digital processing block would have to perform sampling and processing at a Nyquist rate in order to preserve the signal shape. In case of a high-speed transmitter with a multi-Gb/s data rate, the digital block would have to run at a prohibitively high frequency, from both speed and power standpoints. The system proposed here takes advantage of the fact that with binary NRZ signaling the information to be transmitted is in essence digital; therefore, we can assume that the digital block has at its inputs the digital bits (for example, obtained by a 1-bit slicer). Now the digital processing operations (linear combinations, realized by matrix multiplication) can be performed at the digital rate by the encoder blocks. Practical speeds of operation of modern synthesized digital cores depend on the process used, but are typically below 1 Gb/s. If faster signaling is desired, the encoder blocks can be multiplexed by employing serializing, with a serialization factor of 1:S chosen as appropriate for the target signaling rate.

Another issue that needs to be taken into consideration is the nature of the analog waveforms produced by the encoder DAC blocks, due to the inherently analog nature of modal signaling. Modeling the D/A converter and line driver as being limited only by the slew rate of the driver, it is clear that the transition time on each line will depend on the previous and new voltage values. In the modal space, this causes the decoded signals to exhibit different
rise/fall times, or to change the slope during transition, all resulting in deviations from the orthogonal transmission modes and CIJ which is increased [48]. In order to control the transition time and preserve the modal content of the coded signals, one option is to digitally control the slew rate of output drivers, by adjusting the biasing current depending on the previous and current transmitted bits. The physical realization of this can vary depending on the driver layout; one possible implementation would be to use a combinational network to multiplex the appropriate current for each output symbol to be sent.

On the other hand, if a DAC implementation is chosen such that the output transition time is not signal-dependent, but rather settling with a fixed output RC-time constant, the modal waveforms would be correctly generated, avoiding the need to control the output slew rate. One such DAC implementation for high-speed signaling is the commonly used current steering DAC, a 3-bit implementation of which is shown in Fig. 5.2, along with the plot of all the possible output transitions in Fig. 5.3, which confirms that the settling time is not signal-dependent.

![Figure 5.2: A 3-bit current steering DAC realization.](image)
In the case of digital input signals, there are $2^N$ values possible as the inputs to each D/A converter. However, the exact values needed for modal generation are in general nonuniformly spaced, and furthermore are functions of the physical configuration of the channel. Therefore, it is assumed that the output DACs are M-bit with uniform output levels, where $M \geq N$. Examining the RMS and peak-to-peak CIJ values at the receiver reveals the exponential improvement to both metrics with increasing resolution [48]. However, increasing the DAC resolution adds to encoder area, design complexity and ultimately transmitter power, so a compromise must be made depending on the overall specifications of the system being designed.

To explore the feasibility of the digital transmitter, the power of the DSP system (including the serializer/deserializer, but excluding ADC/DAC) was estimated for various combinations of number of lines N and DSP resolution M. For the assumed data rate of 12.8 Gb/s per line in 0.18 $\mu$m CMOS technology, serialization factor S of 16 was assumed. With no pipelining,
the digital core runs at 800 Mb/s. In each DSP block, there are N additions of M-bit numbers for each of N channels. The main critical path delay is dominated by 1-bit adders. To achieve the required delay of 1.25 ns, the unit 1-bit full adder is first scaled so that its delay does not exceed $1.25 / (N \times M)$ ns, and its power $P_{1FA}$ was extrapolated accordingly. Assuming the multiplier (1-bit AND gate) power $P_{1AND}$ to be approximately $1/4$ of $P_{1FA}$, and with total serializer/deserializer power estimated at $P_{serdes} = 100$ mW for every bit, the total DSP power is then given by:

$$P_{DSP} = S \cdot (P_{mult} + P_{add}) + N \cdot P_{serdes}$$

$$P_{mult} = N^2 M P_{1AND}$$

$$P_{add} = N \log(N) M P_{1FA}$$

(5.1)

The power consumption of the transmitter’s digital core per Gb/s is shown in Fig. 5.4, for various combinations of the number of lines $N$ and DSP precision $M$. As expected, for low values of $N$ the dominant factor is the serializer/deserializer power, while with increasing $N$ the power grows quadratically with the number of lines, eventually limiting the bundle size. Increasing the number of DSP bits also adds to the power, but not as drastically, and its effect could be minimized with custom DAC levels for each line. A DSP-based receiver is estimated to have comparable levels of performance and complexity.

Note that the digital encoder/decoder implementation discussed would mainly be suitable for the transceiver systems already equipped with ADC/DAC and serializer/deserializer blocks, because of the power overhead associated with those blocks. At the present time, this would limit the application of such a type of modal signaling mainly to long (typically 20 inch and above) and lossy ISI-dominated backplane channels. For chip-to-chip signaling, which usually involves shorter parallel channels (4-10 inch) with loss not being the dominant signal integrity issue, a purely analog solution would be better suited.
5.2 Prototype Analog Frontend Transceiver Structure

Referring to the canonical block diagram of the modal signaling system in Fig. 3.2, modal encoder and decoder observed as stand-alone blocks would need to realize matrix multiplication operations, creating linear combinations of multiple inputs. However, some of the modal signaling blocks can be integrated with existing transceiver blocks for better efficiency. One such example is the modal encoder block, which for the case of binary NRZ signaling lends itself naturally to integration with the line drivers. There are different possible realization options for implementing the building blocks, depending on the channel over which the signaling is to take place, the process that is available to the designer, and speed, power and design complexity requirements. This section will present the design flow suitable for short tightly coupled chip-to-chip interconnects, in particular the 4-line system as detailed in Case Study B of this work, consisting of 4-inch PCB traces on FR4 substrate, but also the package traces, vias and connectors. Each line has 0.5 pF of equivalent capacitive loading, due to pad
and receiver input capacitance. The technology node used is 1.2 V IBM 90 nm low-power
digital CMOS, a typical process used for controller-memory links.

In the following subsections, the details of the proposed modal signaling implementation
will be given, referring to Fig. 5.5 for the proposed transmitter architecture blocks and
Fig. 5.6 for the receiver blocks.

![Figure 5.5: Proposed transmitter architecture for modal signaling.](image)

5.3 Encoder/Driver Block

As shown previously in this work, it is preferred to use voltage waves for modal signaling, as
opposed to current waves, due to the frequency-dependent behavior of current eigenvectors.
For each combination of binary data signals at the input of the encoder, a particular linear
combination of modal voltages needs to be excited at the near ends of the transmission line
bundle. Based on the peak power limit for the low-voltage power supply used, the maximum
swing at the output of each transmitter block was set to 250 mVpp, by appropriately scaling
the modal voltage excursions produced by eigenvalue decomposition.

For high-speed data transmission, the preferred output stage configuration for single-ended
signaling is a push-pull topology [49]. However, if such an output stage is treated as a voltage
source, in advanced low-power CMOS processes its output resistance of push-pull transistors
will be comparable to the characteristic impedance of the line [50], thus preventing accurate
generation of linear combinations of voltages.

Rather than attempting to reduce the output impedance by placing the buffers in paral-
lel [50], a converse approach of using current sources with high output impedance is proposed.
To achieve the target 250 mVpp swing, an appropriate set of currents that need to be in-
jected into the line bundle was calculated and is presented in Table 5.1. An NMOS-only
realization is preferred for better high-speed performance (due to higher carrier mobility).

Figure 5.6: Proposed receiver architecture for modal signaling.
Table 5.1: Current changes to generate modal signals and the reference voltage level.

<table>
<thead>
<tr>
<th>Line</th>
<th>Mode 1 [mA]</th>
<th>Mode 2 [mA]</th>
<th>Mode 3 [mA]</th>
<th>Mode 4 [mA]</th>
<th>Common [mA]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+1.0650</td>
<td>+1.1950</td>
<td>-0.7970</td>
<td>+0.5270</td>
<td>+0.6740</td>
</tr>
<tr>
<td>2</td>
<td>+0.4670</td>
<td>-1.0000</td>
<td>+1.6773</td>
<td>+0.6060</td>
<td>+0.0252</td>
</tr>
<tr>
<td>3</td>
<td>-0.4670</td>
<td>-1.0000</td>
<td>-1.6773</td>
<td>+0.6060</td>
<td>+0.0252</td>
</tr>
<tr>
<td>4</td>
<td>-1.0650</td>
<td>+1.1950</td>
<td>+0.7970</td>
<td>+0.5270</td>
<td>+0.6740</td>
</tr>
</tbody>
</table>

The common mode at each line was set to Vcm=1 V using open drain drivers realized as four self-cascode structures, shown in Fig. 5.7(a), with the gate bias set to Vbias1=0.9 V, to ensure operation in saturation at all times. Self-cascode with m=4 is chosen for its large effective channel length and low effective output conductance [51].

![Figure 5.7](image)

Figure 5.7: (a) Open-drain drivers producing the common-voltage levels; (b) current-steering for shared currents, and (c) for non-shared currents.

Examining the changes in currents required to create modal voltages, we can notice from Table 5.1 that, due to channel symmetry, lines 1 and 4 in mode 1 can share the same current sources, as well as lines 2 and 3, and similarly for mode 3. A current-steering differential pair
structure shown in Fig. 5.7(b) is therefore used to fully steer the constant current between lines, based on the bit value dictated by the pre-drivers. For modes 2 and 4, not all the currents can be shared; additional currents need to be injected (e.g. in case of mode 2, outer lines should receive 0.195 mA more). For this purpose, a structure shown in Fig. 5.7(c) is used, also in a differential pair configuration to ensure linearity during transitions. Compared to the implementation presented in [27], only the currents steered into dummy resistors R and the common-level currents are not used in actual information signal generation, thus resulting in a much increased power efficiency.

Each current source is realized as a set of pre-programmable binary-sized NMOS transistors (with those turned on receiving a constant gate bias of $V_{bias2}=0.6$ V to ensure self-cascode saturation), so that the encoder coefficients can be adapted to the particular channel.

![Layout of the encoder block.](image)

The total area of the encoder block is approximately $6500 \, \mu m^2$, as shown in Fig. 5.8. The total average power of encoder/driver blocks is about 11 mW. Both of these figures need to be increased by 30% in case of a non-symmetric bus, since in that case each of the encoder coefficients could take any of the full range of values, and also be either positive or
negative, precluding the use of current steering shared between two coefficients in the linear combination.

The power estimation does not include the pre-drivers, realized in a simple differential pair structure producing a swing of 300 mVpp at each drain node. In this case the number of differential pairs the pre-drivers are driving (compared to the uncoded transmitter) is increased by a factor of 4 (since each source bit takes part in all four linear combinations driving the lines), but the current steered through each differential pair is reduced (since the energy of each bit is spread out to all four lines), so the total pre-driver power should remain roughly unchanged, and is excluded from the power calculation.

5.4 Decoder Block

Once the line voltages have been sensed, they need to be decoded back from the modal space to the line space, in order for the SISO receiver blocks to decide which binary symbol was transmitted. The decoder block is most naturally realized by converting the received voltages to current, and then creating the desired linear combinations by simple KCL summation. A current subtractor similar to the one presented in [52], which allows simultaneous voltage-to-current conversion and reference level removal, is shown in Fig. 5.9. The cell operation is based on its large-signal behavior in response to a differential input signal, as given by (5.2):

\[ I_O = I_1 - I_2 = \frac{1}{2} k_n' W \left( \frac{4 I_{SS}}{k_n' W L} - (V_1 - V_2)^2 \right) \approx \sqrt{k_n' W L} I_{SS} (V_1 - V_2) \]  

(5.2)

The linear approximation in (5.2) is valid for suitably chosen transistor sizes and bias current (which ensures complete switching does not take place); these values are also used to set the decoding coefficient. This versatile block can produce either the positive (when
Figure 5.9: Current subtractor for modal decoder realization.

\[ V_1 = V_{far}, \quad V_2 = V_{cm} \] or negative current (with the opposite wiring) proportional to received signal on the far-end of a line, or the direct difference of two line signals (when \( V_1 = V_{farA}, \quad V_2 = V_{farB} \)). The currents are then summed over a suitably sized resistor to produce the slicer input voltage. Again, each current source is realized as a set of pre-programmable binary-sized NMOS transistors, so that the decoder coefficients can be adapted to the particular channel. The target coefficients should take into account both the channel properties (through eigenvalue decomposition) and the noise present in the system, since simply inverting the channel might enhance noise at particular frequencies, as discussed in Chapter 3.

The total decoder block area in 6-bit coefficient resolution is about 4300 \( \mu m^2 \), as shown in Fig. 5.10. The average power is 14.5 mW. Again, both of these figures need to be increased by 30% in case of a non-symmetric bus, since in that case each of the decoder elements could be in the full coefficient range, and also be either positive or negative, precluding the use of a differential pair structure shared between two coefficients in the linear combination.
5.5 Other System Blocks

5.5.1 Line Termination Block

An important requirement for modal signaling is to provide an appropriate termination for modal waves, in order to eliminate reflections and modal conversion (which directly translates to crosstalk). For a realistic memory bus system, due to the presence of discontinuities and loss, the matching terminations will be frequency dependent, all the ideal termination elements will be complex (due to finite conductivity and dielectric loss), and finally, the termination network will be a full matrix; all of these characteristics could present practical challenges in circuit realization.

One possible solution is the implementation of a low-order frequency dependent matching network. However, as shown in this work, that requires utilizing a number of precision capacitors and inductors, which might prove impractical to implement on chip; pushing these passive elements out into the package would increase the system cost and design complexity. A compromise solution explored in this work is to implement a full grid of on-chip resistors between each line pair, and to utilize mode 4 (the ”ground mode,” most dependent on a
frequency dependent matching network) to transmit a clock signal with a reduced data rate (source-synchronous signaling, a standard technique in DDR memory chips). The resistor grid is terminated to VDD to provide the pull-up termination network to the open-drain transmitter drivers, and their values are given in Table 4.3.

5.5.2 Modal Skew Compensation Block

After decoding the signals transmitted over fundamental modes, there will be additional modal skew introduced, based on the propagation velocities of the modes onto which the signals were encoded. Deskew circuitry on the transmitter side can delay the generation of modal signals, so that they all arrive at the far end of the line bundle synchronized with each other. This pre-delay can be realized using simple double-inverter delay elements [53], or using a current-starved inverter with tunable switching threshold [54]. As discussed in previous sub-sections, in order to minimize the impact of the limited bandwidth of mode 4, a half-rate (2 Gb/s) source-synchronous clock is transmitted over this mode, and full-rate (4 Gb/s) data is transferred over other three modes and sampled on both edges of the clock. For a realistic chip-to-chip communication system consisting of many bundles, some type of master clock realignment circuitry might still be needed [53], such as a simple delay locked loop (DLL) to realign the signals to a particular global reference.

5.5.3 Receiver Block

Assuming that the system under consideration is crosstalk-dominated, the SISO receiver blocks have a simple task of deciding on the signals sent by comparing the received signals with the generated reference voltage; this can be performed by using a standard differential pair configuration operating as a slicer. If the channel had a significant level of intersymbol interference (ISI) present, receiver-side equalization [5] could be used on a per-channel basis.
5.6 Simulation Results and Performance Benefits

To gain insight into crosstalk cancelation performance of the proposed system, we first examine the pulse responses of uncoded NRZ signaling with data rate of 4 Gb/s, as shown in Fig. 5.11. To demonstrate a realistic application, the system is loaded with the typical value of 0.5 pF at each line end, due to equivalent receiver input capacitance. We can see that the uncoded system exhibits strong ISI at the first post-cursor tap, due to line length and capacitive loading. Also, the system exhibits additional ISI (coming from reflections at the package-PCB discontinuities in signal path), arriving to the receiver several unit intervals later. Finally, we can observe strong crosstalk pulses at victim lines (with peak interference summing up almost to signal level), due to very strong coupling at each of the inhomogeneous channel segments.

![Figure 5.11: Pulse response on inner and outer lines of the uncoded bundle, and crosstalk pulse responses on other lines.](image)

Uncoded NRZ signaling results in the peak-to-peak jitter values as given in Table 5.2, with PRBS excitation of $2^8 - 1$ bits still producing a partially open eye, but statistical simulation resulting in completely closed eye for both inner and outer lines if 4 Gb/s signaling is used. If a half-rate clock is transmitted over one of the lines, the results improve somewhat, but
Table 5.2: Peak-to-peak jitter for NRZ signaling over nonuniform bus with Ci=0.5 pF.

<table>
<thead>
<tr>
<th>Bit Rate</th>
<th>Line</th>
<th>$J_{pp-prbs}$ [ps]</th>
<th>$J_{pp-stat}$ [ps]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 Gb/s</td>
<td>2,3</td>
<td>246.3</td>
<td>(eye closed)</td>
</tr>
<tr>
<td></td>
<td>1,4</td>
<td>149.0</td>
<td>(eye closed)</td>
</tr>
<tr>
<td>2 Gb/s</td>
<td>2,3</td>
<td>185.0</td>
<td>296.6</td>
</tr>
<tr>
<td></td>
<td>1,4</td>
<td>100.1</td>
<td>241.1</td>
</tr>
</tbody>
</table>

the jitter value is still at least half of the unit interval, which greatly exceeds the allocated jitter budget.

To demonstrate the crosstalk mitigation performance of the system, Fig. 5.12 shows the circuit-level simulation of 4 Gb/s NRZ pulse responses\(^1\) of signals transmitted over the equivalent modal channels. As expected, each modal signal creates a greatly reduced crosstalk disturbance on the other modes, mainly due to circuit linearity limitations. Also notable is the low amount of ISI present, due to the full matching matrix for this short channel, but the internal reflections which cannot be removed this way are still visible, mainly in mode 4. Comparing Fig. 5.12 to the uncoded pulse responses, we can expect a significant jitter reduction benefit.

To verify the overall system performance, the eye diagrams of decoded modal signals using $2^8 - 1$ PRBS sequences are shown in Fig. 5.13. The modal skew compensation block was not activated, to demonstrate arrival times of source-synchronous clock in relation to the three data lines. As summarized in Table 5.3, maximum peak-to-peak jitter present in the system is now reduced to 15.6% of UI, in case of Mode 1 for which $J_{pp-circuit}=39$ ps (mainly due to ISI), which is only 47% of the target jitter value allocated for interconnects (in this case UI/3=83 ps).

Time-domain circuit simulations were limited to a relatively small number of bits. Table 5.3 also shows the results expected using the ideal (equation-based) encoder and decoder structure, both for the $2^8 - 1$ PRBS sequence and the result from the statistical eye dia-

\(^1\)Note that in this case S-parameters cannot be used for complete system characterization, due to the integrated encoder/driver structure. A large-signal description such as X-parameters [55] could be used instead.
Figure 5.12: Unit pulse responses of signals over equivalent modal channels.

Table 5.3: Peak-to-peak jitter for modal signaling over nonuniform bus with Ci=0.5 pF.

<table>
<thead>
<tr>
<th>Bit Rate</th>
<th>Mode</th>
<th>$J_{pp\text{-prbs}}$ [ps]</th>
<th>$J_{pp\text{-stat}}$ [ps]</th>
<th>$J_{pp\text{-circuit}}$ [ps]</th>
<th>Min. Improvement [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 Gb/s</td>
<td>1</td>
<td>37.2</td>
<td>40.1</td>
<td>39.0</td>
<td>73.8</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>31.0</td>
<td>41.8</td>
<td>34.7</td>
<td>76.7</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>29.5</td>
<td>37.0</td>
<td>32.1</td>
<td>78.4</td>
</tr>
<tr>
<td>2 Gb/s</td>
<td>4</td>
<td>56.4</td>
<td>76.9</td>
<td>58.0</td>
<td>42.1</td>
</tr>
</tbody>
</table>

gram. The PRBS results are in good agreement with the circuit-level simulations, confirming the performance of the modal signaling system. Statistical simulations reveal that uncoded 4Gb/s signaling is not possible for this channel (the eye being completely closed due to FEXT), so only the PRBS results of the uncoded signals from Table 5.2 can be compared to those of the circuit simulation, giving the worst-case minimum jitter reduction improvement for each signaling speed.

This analysis would for completeness need to include Tx jitter and power supply fluctuations as well; however, the uncoded NRZ signaling suffers from the same impairments;
Figure 5.13: Eye diagrams of normalized decoded modal signals.

therefore, it is expected that the degradation in performance will be similar for both systems. The eye opening of 150 mV should present a sufficient voltage margin for the slicer to operate on for reliable detection (BER less than $10^{-12}$), with assumed noise sigma=2 mVrms and offset and sensitivity of 10 mV total. The 12 Gb/s of aggregate bandwidth obtained using source-synchronous clocking with the reduced data rate, represents an increase of 2.5 times compared to the conventional NRZ signaling achieved over a similar channel, where each of the coupled bus lines is limited to a maximum of 1.6 Gb/s [25].
5.7 Comparison with Other Crosstalk Mitigation Techniques

As discussed previously, a number of crosstalk mitigation techniques have been proposed in recent years. None of the techniques are in standard use in the industry today, due to the various issues discussed in Chapter 2. Furthermore, the previously proposed techniques are designed for lines with weak coupling (where the capacitive and inductive coupling coefficients are much less than unity), whereas the modal signaling technique proposed in this work purposefully assumes strong coupling (with the signaling real estate utilized as compactly as possible) and assumes discontinuities in the signal path. Still, it would be useful to gain some insight into the performance of the previously proposed techniques in terms of power and crosstalk cancelation ability.

Note that, for purposes of this comparative study, the lines are terminated with the fully matched resistor grid, in order to focus on FEXT cancelation performance. The traditional receiver design which does not have parallel termination elements would suffer from reflections as well, making it ISI-dominated and exceptionally challenging for signaling using any of the proposed techniques.

5.7.1 FEXT Cancelation Techniques

A typical example of a crosstalk mitigation technique which relies on predistorting the data signal to cancel out FEXT interference [19] attempts to eliminate FEXT, rather than simply compensate for CIJ, so it would be the one most directly comparable to the modal signaling technique. The reported power efficiency of 25 mW/Gb/s, an order of magnitude greater than the proposed modal technique, was given for the 3-lane microstrip channel consisting of three pairs of differential lines for signaling, which is more resistant to crosstalk than single-ended signaling [15].

However, for this class of channels it is actually not the high power utilization which precludes the use of predistorting, but the peak power limit at the transmitter circuit.
Namely, for the simpler case of canceling FEXT from line 2 onto line 1 only, the FIR crosstalk equalizer taps were calculated to be as follows, using equation (4) of [19]:

\[ h_{\text{XTC}2\to1}[n] = [-0.3795, +0.4089, -0.0190, +0.0033] \]  (5.3)

which results in the peak power of the direct signal reduced to a mere 19% of its unit value. When the direct signaling swing is adjusted to match those requirements, by changing \( W_0 \) in equation (4) of [19] to 0.19 instead of its unit value, the peak-to-peak jitter is reduced from 160 ps to 70 ps; however, the eye opening at the receiver is reduced to less than 10 mV, as shown in Fig. 5.14. Since the system is tightly coupled, XTC equalizers from lines 3 and 4 to line 1 would also be needed, thus completely closing the vertical eye. The situation is even more extreme for inner two lines of the bundle, which suffer from much more crosstalk from their neighbors.

![Eye diagram for XTC equalization for outer line 1, with XTC2-1 FIR and aggressor 2 only.](image)

Figure 5.14: Eye diagram for XTC equalization for outer line 1, with XTC2-1 FIR and aggressor 2 only.

A different crosstalk removal approach is proposed in [20], which creates a mimicked version of received crosstalk signals, and subtracts them from the sampled input signals at
the receiver, thus circumventing the peak power limit imposed at the transmitter. However, this method has several drawbacks which limit its usability for the non-uniform system, the main of which is the accurate mimicking of the crosstalk signals. As seen in Fig. 4.5, for tightly coupled lines with discontinuities, the crosstalk signals themselves exhibit resonances; impulse response simulation shown in Fig. 5.11a shows that in time domain, the crosstalk signal induced by exciting line 1 will result in a complex crosstalk shape on line 4 of significant magnitude, which cannot be approximated with a simple RC filter.

Furthermore, the realization proposed in [20] assumes the existence of DFE equalization with dedicated sample/hold and summer circuitry at the receiver to implement subtraction of the mimicked signal. In a non-ISI dominated environment as discussed here, this will typically not be the case. A modified mimicking circuit would have to be realized for continuous-time signal detection, where it is critical to keep the RC time constant at the summation node low to achieve high speed operation. For example, an ideal RC filter with \( R = 9500 \, \text{Ohm} \) and \( C = 9 \, \text{fF} \) would approximate crosstalk from line 1 to line 2 reasonably well. However, even in today’s deep submicron CMOS technologies, the total capacitance already present at the decision node of the high-speed receiver greatly exceeds this value, so the mimicked signal would have to be buffered first.

5.7.2 CIJ Compensation Technique

This technique has been demonstrated [56] on a two-line 4-inch bus on FR-4 substrate, coupled with 400 fF of mutual capacitance (for comparison, Case Study B of this work assumes around 1340 fF capacitive coupling). The presented system operating at 5 Gb/s per channel consumed 80 mW (giving 8 mW/Gb/s), and the implementation area in 130 nm CMOS was 14000 \( \mu m^2 \). Peak-to-peak jitter was reduced from 43% UI to 21% UI. As can be seen from the performance metrics, for a system with lower coupling, the area overhead requirement was comparable to the method proposed in this work, but the power utilization
was greater by a factor of about 4x, mainly due to the use of high-speed current-mode logic (CML) gates to detect the transitions and adjust the variable time delay.

Moreover, it is challenging to achieve the increased complexity of the CIJ compensation technique with number of possible threshold crossing time deviations, with [17] suggesting the approach of compensating only the worst case deviations. As such, this approach seems difficult to extend to the tightly coupled multiline bus channels, where even a non-adjacent aggressor induces a significant amount of FEXT on a bundle line.

5.7.3 Passive Compensation Techniques

Finally, the passive compensation methods, which rely on placing passive elements to create a homogeneous channel (and thus equate the velocities of all the modes in the system), were explored for the nonuniform channel. Some of the channel modifications proposed by passive methods [23], [24] can be too costly (in terms of board real-estate) or impractical, when presented with a legacy channel which cannot be entirely redesigned to improve signal integrity. The most promising techniques were the ones based on addition of lumped capacitor elements [25], [57], but they could not be directly applied to this type of channel (due to their synthesis procedures founded on the assumption of a uniform lossless system, with frequency-independent L and C matrices, and with no discontinuities in the signal path).

An attempt was made to equalize the longest coupling section of the channel, the 4-inch PCB bundle, due to its contribution to modal latencies being the largest of the three. The parallel capacitor values to be placed were extracted according to eqs. (5)-(9) of [57], and the capacitive contribution of the remaining 8 mm package bundle section was subtracted, to arrive at the following parallel capacitance values to be placed at the far-end of the PCB segment:

\[ C_{12} = C_{34} = 1.5742 \, pF; C_{13} = C_{24} = 0.1940 \, pF; C_{14} = 0.1081 \, pF; C_{23} = 1.5093 \, pF \]  \hspace{1cm} (5.4)
Placing those capacitors resulted in an amount of modal velocity equalization over the PCB segment, reducing the peak-to-peak crosstalk noise at this discontinuity from uncompensated 622 mV (Fig. 5.15a) to 393 mV in the compensated case (Fig. 5.15b). However, at the same time it resulted in increasing the internal reflections at this discontinuity by an order of magnitude, resulting in overall worse FEXT at the receiver input. An attempt was made to reduce this reflection by modifying the values of termination resistors at the far end of the package segment. However, this resulted in increased reflection at that final termination point, again increasing overall FEXT and CIJ of the system.

Figure 5.15: Pulse response of line 2 and crosstalk on neighboring lines observed at the far-end PCB-package discontinuity.

In order for the proposed techniques to be applicable to this type of tightly coupled channels, they would need to be adapted to take into account the properties of the overall channel (similar to the generalized modal approach versus the traditional $LC$ matrix diagonalization), since solving the problem locally (on a segment-per-segment basis) does not guarantee an improvement in overall signaling performance.
5.8 Summary

This chapter proposed two approaches for practical system implementation of modal signaling. The first approach relies on the assumption that modal encoding and decoding can be performed using digital cores present in ADC/DAC based transceivers. Although simple and straightforward to implement, at the present time its application is limited to backplane links, due to the high power overhead needed for ADC implementation. The second approach details the system architecture of an analog modal transceiver. The encoder and decoder blocks are designed using a low-power digital CMOS process, for signaling over a 4-lane controller-memory microstrip bus. The proposed design is placed in a context by comparing it with estimated performance of previously proposed crosstalk mitigation techniques over the same tightly coupled channel.
CHAPTER 6
CONCLUSION AND FUTURE WORK

This thesis has illustrated that a crosstalk mitigation solution based on the modal channel decomposition can be successfully applied to realistic high-speed signaling over dense nonuniform channels with discontinuities. This technique provides a very effective reduction of FEXT and associated CIJ. Overall signal integrity is enhanced, thereby providing increased density and/or higher available bandwidth over chip-to-chip channels made with low-cost materials and with limited design resources. In terms of design complexity and crosstalk reduction effectiveness, modal signaling outperforms other previously proposed techniques, which demonstrate significant implementation issues when presented with a realistic, tightly coupled channel. In this chapter, we summarize the contributions of the thesis and discuss potential future research directions that could be considered as the natural extension of the presented work.

6.1 Contributions

In this thesis, we have presented a comprehensive approach used to provide a design flow of a high-speed transceiver employing modal signaling techniques to mitigate far-end crosstalk. The approach was validated on the test case of a typical controller-memory signaling bus.

1. We proposed the application of generalized modal decomposition theory to the class of tightly coupled, nonhomogeneous and nonuniform channels with discontinuities. The proposed approach offers a robust method of extracting modal properties of the channel starting from the presumed channel geometry and structure, or from actual measured
channel data, or a combination of both. The realized framework allows for exploration of different implementations of signaling system elements, and their impact on the overall system performance using statistical simulation techniques.

2. Based on the nonuniform structure of the communication channels, we proposed the use of a frequency-dependent termination network for optimal signaling performance. For the representative test channel, we demonstrated that the improvement of 4 Gb/s NRZ signaling over ground mode obtained with the low-order model of the optimal frequency dependent network results in the vertical eye opening increase of 39% at the optimal sampling point, and the reduction in peak-to-peak jitter of 27%.

3. In order to facilitate transceiver design, we proposed a MIMO system perspective of modal signaling. We have demonstrated the method of obtaining the modal decoder coefficients for near-optimum SNR for a given channel. Also explored was the issue of encoder and decoder coefficient quantization. We demonstrated a methodology for determining the required number of bits of precision given the target bit-error rate.

4. Starting from the properties of the channel as revealed by generalized modal decomposition, we proposed two approaches for practical system implementation of modal signaling, using (a) digital cores present in ADC/DAC based transceivers, and (b) analog frontend transceiver structure. For the analog transceiver, the design flow is demonstrated using the 90 nm low-power digital CMOS process, over a 4-lane controller-memory microstrip bus. The 12 Gb/s of aggregate bandwidth obtained using source-synchronous clocking with the reduced data rate represents an increase of 2.5 times compared to the conventional NRZ signaling achieved over a similar channel. Maximum peak-to-peak jitter on the data lines present in the system is reduced to 15.6% of UI, whereas using NRZ signaling would result in a completely closed data eye. For the symmetric bus, the total area of the encoder/driver block is about 6500 $\mu m^2$, and
the average power is about 11 mW. The total area overhead for the decoder block is about 4300 $\mu m^2$, and the average power is 14.5 mW.

6.2 Practical Synthesis of the Optimal Termination Network

By applying the vector fitting procedure and passivity enforcement during construction of the low-complexity termination network for optimal matching, its physical realizability using a passive termination network can be ensured. However, the general algorithm method of obtaining the particular constituent component values for practical circuit realization does not seem to be readily available.

6.2.1 Overview of Synthesis Techniques

Synthesis is the realization step needed to map the reduced order model into a netlist consisting of electrical circuit components. Modern macromodeling techniques can be used to ensure the positive real property of the model in the process of passivity enforcement [58]. It was shown [59] that passive systems with positive real transfer functions can be synthesized with positive RLC elements and transformers.

Later, it was proven [60] that any positive-real function could be realized as the driving-point immittance of a network consisting of resistors, capacitors and inductors only, thus alleviating the need for transformers in the realization. However, there are issues with the direct application of this approach, since the resulting networks would be in the form of a balanced bridge, which is extremely sensitive to any deviations of constituent component values. Also, the realization is distinctly non-minimal, so the number of components generated during synthesis is too large for practical on-chip implementation.

Today, the main body of research in the area of rational transfer function synthesis is directed towards the synthesis of RLC models for efficient simulation purposes [61]. The drawback of applying most of those techniques is their inability to guarantee physically
realizable (positively valued) RLC constituent elements, or alternatively the use of controlled
sources.

Out of the two potentially realizable approaches which do not involve controlled sources,
the unstamping methods \cite{62} require the original system structure to be available in the
modified nodal analysis (MNA) representation, which the generalized eigenvalue decompo-
sition method for modal analysis does not seem to provide (in particular when starting from
the measured channel description). The other synthesis method which seems most suitable is
Foster synthesis \cite{63}, where the realization is done via the system’s transfer function (readily
available as the result of vector fitting).

6.2.2 Foster Synthesis

This class of synthesis methods relies on observing the \(N\)-port network to be modeled as
a full grid of unit one-port network elements between each pair of \(N+1\) nodes (including the
reference node). Then, rewriting the elementary one-port driving-point admittance functions
in the Foster’s canonical form \cite{64}:

\[
Y(s) = sY_\infty + Y_0 + \sum_{m=1}^{N_R} \frac{a_m}{s-p_m} + \sum_{n=1}^{N_C} \left( \frac{a_n}{s-p_n} + \frac{a_n^*}{s-p_n^*} \right)
\]  

(6.1)

where the rational function is expanded into partial fraction form with \(N_C\) conjugate poles
\(p_n\) and \(N_R\) real poles \(p_r\).

Depending on the assumed topology of the resulting network, an appropriate synthesis
method is chosen. One approach that guarantees the error-free and reversible realization
which does not change the passivity of the realized system is proposed by \cite{65}. The final
realized \(N\)-port network will be a complete graph of RLCG elements, where each branch
represents an admittance given by (6.1) realized by the one-port realization method, as
shown in Fig. 6.1. The relations determining R, L, C, G elements based on the poles and
residues from (6.1) are given in \cite{65}.

94
Figure 6.1: One-port Foster admittance realization of the transfer function.

The one-port elements combined in the full grid are mathematically equivalent to the original N-port termination network, which is passive, due to the passivity of the system it fully matches. However, Foster synthesis realized using the standard relations for element values [65] does not guarantee that the individual R, L, C element values would be physically realizable. Therefore, for practical synthesis purposes this approach would need to be modified to ensure realizability, potentially at the expense of an increased number of components.

In addition to the passive RLC network synthesis, it would be worthwhile to investigate the tradeoffs involved in the realization of the required reflection function using active network synthesis [66], due to its smaller footprint and streamlined practical synthesis methods.
6.3 Integration with Transmitter/Receiver Equalization and Coding

After mitigating of FEXT as the major source of signal integrity impairment, it may also be necessary to mitigate other signaling issues in order to achieve the required level of performance. Therefore, it is important to investigate the compatibility of the crosstalk mitigation technique with standard transmitter- and/or receiver-side equalization techniques [5]. Along with the inherent ISI due to channel bandwidth limitation, ISI due to internal resonances present in the system can affect signaling performance of equivalent modal uncoupled channels. In the case study explored in this work, a 2-tap FIR filter would remove the majority of ISI caused by limited bandwidth for 4 Gb/s NRZ signaling. However, for longer channels (e.g. backplane links), more advanced equalization due to increased channel ISI might be necessary. For channels with multiple cascaded segments and discontinuities, removing the ISI caused by internal resonances might also prove necessary to achieve target BER.

6.3.1 Internal Resonances due to Discontinuities

![Figure 6.2: Step excitation of ground mode of the cascaded system.](image)

(a) Eye diagram  
(b) Step excitation and response (indicated time-of-flight for the first reflection)
As discussed, the optimal matching ensures full crosstalk mitigation by eliminating far-end reflections and modal conversions. However, as seen from Fig. 4.16, the ground mode (mode 4) in particular suffers from reduced eye opening at the receiver.

This is due to the internal modal reflections of the system, which appear in the form of ISI, impairing each of the uncoupled modal channels. The internal resonances can be observed in Fig. 6.2, which shows the decoded response of the step excitation encoded onto the ground mode in the cascaded system, and terminated with the optimal matching admittance network. Since this effect belongs to the class of inter-symbol interference, it can be treated on a per-channel (SISO) basis using any of the standard mitigation techniques, such as transmit- or receive-side equalization.

However, due to the effect being spread over many unit intervals (in this example UI=250 ps), the memory of the channel is long, potentially making equalization challenging. Low-complexity signal coding [3] can also present an interesting approach to alleviating the bandwidth limitation of resonant channels. A system-aware approach [67] should be employed to maximize overall performance of the low-power signaling system.
APPENDIX A

MATLAB CODE FOR EIGENVALUE DECOMPOSITION OF RESONANT SYSTEMS

1
2 % physical constants
3 eps0=8.8542*1e-12; mu0=4*pi*1e-7;
4 % needed for optimization engine
5 global ZCSOLVER
6 % chose freq range for channel diagonalization:
7 freq=0:0.01e9:2.4e9;
8 Nf=numel(freq);
9 % subroutine to read RLCG parameters of channel segments
10 % produces Zch=R+jwL and Ych=G+jwC matrices
11 chsetup
12 % set up parameters of segments:
13 Nseg=numel(Lseg); % number of uniform segments
14 L=sum(Lseg(1:Nseg)); % total length
15 dL=min(Lseg)/100; % 1/2-transition between segments
16 % line up the segments on the X-axis
17 X=[0];
18 for i=1:Nseg-1
19 X=[X sum(Lseg(1:i))−dL sum(Lseg(1:i))+dL];
20 end
21 X=[X L];
22
23 % set up for ODE solving, according to Semlyen 2003, (B.3)
24 Om=zeros(N,N);
25 P0=eye(2*N);
26 Xspan=[0 L];
27 Pfar=zeros(2*N,2*N,Nf);
28 for k=1:numel(freq)
29  % set up ODE P'(x)=F(x)P(x), P(x=0)=I
30  for i=1:Nseg % at each junction position i
31    F(:,:,2*i-1)=[Om squeeze(Zch(i,:,:,:))];
32    F(:,:,2*i )=F(:,:,2*i-1);
33  end
34  [Xdiv P]=ode45(@(x,p) myderiv(x,p,X,F,2*N),Xspan,P0);
35  Pfar(:,:,k)=reshape(P(numel(Xdiv),:),2*N,2*N);
36 end
37 A=Pfar(1:N, 1:N,:); B=Pfar(1:N, N+1:2*N,:);
38 C=Pfar(N+1:2*N,1:N,:); D=Pfar(N+1:2*N,N+1:2*N,:);
39
40 % if reading channel from S-parameters instead:
41 Np=2*N; % number of ports
42 [S,freqsp]=reads8p('sparamfile.s8p');
43 I=eye(Np);
44 Z0p=50*I;
45 Nfsp=numel(freqsp);
46 for i=1:Nfsp
47  % eq. 16 from Joong-Ho Kim's TADVP paper 2010:
48  Z=Z0p*(I+S(:,:,i))*inv(I-S(:,:,i));
49  Z11=Z(1:N, 1:N); Z12=Z(1:N, N+1:2*N);
50  Z21=Z(N+1:2*N,1:N); Z22=Z(N+1:2*N,N+1:2*N);
51  % eq 17:
52  TA(:,:,i)=Z11*inv(Z21); TB(:,:,i)=Z11*inv(Z21)*Z22-Z21;
53  TC(:,:,i)=inv(Z21); TD(:,:,i)=inv(Z21)*Z22;
54 end
55 % replacing the RLCG calculations, if no comparison needed
freq=freqsp; Nf=Nfsp; A=TA; B=TB; C=TC; D=TD;

% eigenvalue decomposition:
llambda=zeros(2*N,Nf);
WW=zeros(2*N,2*N,Nf);

[WWsh,llambdash]=eigenshuffle([A B; C D]); % "smart" eigendecomposition
% note: eigenshuffle is a public domain code from John D'Errico, from:
% http://www.mathworks.com/matlabcentral/fileexchange/22885-eigenshuffle

ltmp=llambdash(:,end); % sort out the fw and the bw modes
[dummy,sortl]=sort(abs(ltmp),1,'ascend');
sortl(1:N)=flipud(sortl(1:N));
llambda=llambdash(sortl,:);
WW=WWsh(:,sortl,:);

% need to check if eigenshuffle did a good job at resonances
% check by observing the angular distance of voltage (more stable) ... eigenvectors
% compared to the low-frequency values
Wfv0=WW(1:N,1:N,2); % at low freq

for f1=3:Nf
    Wfv=WW(1:N,1:N,f1);
    dist = (1-abs(Wfv0'*Wfv)).*sign(real(Wfv0'*Wfv));
    reorder = munkres(dist); % --
    for i=1:N
        sortev(i)=find(reorder(:,i));
    end
    sortmod=[sortev N+sortev];
    llambda(:,f1)=llambda(sortmod,f1);
    WW(:,f1)=WW(:,sortmod,f1);
end
% assuming consistent modes, sort by final unwrapped phase
for i=1:N
    lph(i,:) = squeeze(unwrap(unwrap(angle(llambda(i,:)))));
end
[dummy,sortlph] = sort(lph(:,end),1,'descend');
sortl = [sortlph; N+sortlph];
llambda = llambda(sortl,:);
WW = WW(:,sortl,:);

% calculate alpha and beta coeffs from gamma
gg = -log(llambda)/L;
aa = real(gg);
bbwrap = imag(gg);
for i=1:2*N
    bb(i,:) = unwrap(unwrap(bbwrap(i,:)));
end

% calculate perfect matching terminations
I = eye(N);
Z0 = 50*I;
Y0 = inv(Z0);
for f1 = 1:Nf
    Wfv = WW(1:N, 1:N, f1); Wbv = WW(1:N, N+1:2*N, f1);
    Wfi = WW(N+1:2*N, 1:N, f1); Wbi = WW(N+1:2*N, N+1:2*N, f1);
    Wfi = Wfi.*sign(real(Wfi)).*sign(real(Wfv));
    Wbi = Wbi.*sign(real(Wbi)).*sign(real(Wbv));
    Ycf(:, :, f1) = Wfi*inv(Wfv); Ycb(:, :, f1) = Wbi*inv(Wbv);
    Zcf(:, :, f1) = inv(Ycf(:, :, f1)); % can be resonant for non-uniform systems
end

% calculate S-parameters of the perfect matching box
% go through Y, not the resonant Z
Scf(:,:,f1)=(Y0−Ycf(:,:,f1))*inv(Y0+Ycf(:,:,f1));

% calc the grid of perfect fw termination impedances
% for perfect matching, Ytermf=Ycf
for i=1:N
    for k=1:N
        Ytermf(i,k,f1)=−squeeze(Ycf(i,k,f1));
    end
end

% the diagonal terms are calculated differently
for i=1:N
    Ytermf(i,i,f1)=sum(Ycf(:,i,f1));
end

% calculate S-parameters of each 2-port from admittance values
for i=1:N
    for k=1:i
        Y=Ytermf(i,k,f1); % 2-port admittance, not Y-parameters
        Stermf(i,k,f1,1)=1/(1+2*50*Y); % S11=S22, symmetry
        Stermf(i,k,f1,2)=(2*50*Y)/(1+2*50*Y); % S12=S21, reciprocity
    end
end

% optimization of resistive terminations to achieve minimal reflection ... 
% coefficient Gamma
options = optimset('Algorithm','interior-point');
initcond=50*ones(1,N/2);
for f1=2:Nf
    YL(:,:,f1)=−1./squeeze(Ztermf(:,:,f1));
    % the diagonal terms are calculated differently
    for i=1:N
YL(i,i,f1)=sum(1./Ztermf(i,:,f1));

end

Zcf(:, :, f1)=inv(Ycf(:, :, f1));
Zcb(:, :, f1)=inv(Ycb(:, :, f1));
ZLideal(:, :, f1)=inv(squeeze(YL(:, :, f1)));
ZLideal(:, :, f1)=ZLideal(:, :, f1).*eye(N);
ZLideal(:, :, f1)*eye(N);
ZLideal(:, :, f1)*eye(N);
ZCSOLVER=−squeeze(Zcf(:, :, f1));
[x, fval]=fmincon(@Gcost3sym, initcond,[],[],[],[],[1 ... 1], 10*initcond, [], options);
ZLopti3diag(:, :, f1)=−diag([x fliplr(x)]);
end

ZL=ZLopti3diag;
for i=2:Nf
Gf(:, :, i)=(ZL(:, :, i)-Zcf(:, :, i))*inv(ZL(:, :, i)+Zcf(:, :, i));
Gb(:, :, i)=(ZL(:, :, i)-Zcb(:, :, i))*inv(ZL(:, :, i)+Zcb(:, :, i));
end


