

Novel Multicarrier Memory Channel Architecture Using Microwave Interconnects:
Alleviating the Memory Wall

by

Brahim Bensalem

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2018 by the
Graduate Supervisory Committee:

James T. Aberle, Chair
Bertan Bakkaloglu
Jennifer Kitchen
Panayiotis A. Tirkas

ARIZONA STATE UNIVERSITY

May 2018

ABSTRACT

The increase in computing power has simultaneously increased the demand for input/output (I/O) bandwidth. Unfortunately, the speed of I/O and memory interconnects have not kept pace. Thus, processor-based systems are I/O and interconnect limited. The memory aggregated bandwidth is not scaling fast enough to keep up with increasing bandwidth demands. The term "memory wall" has been coined to describe this phenomenon[1].

A new memory bus concept that has the potential to push double data rate (DDR) memory speed to 30 Gbit/s is presented. We propose to map the conventional DDR bus to a microwave link using a multicarrier frequency division multiplexing scheme. The memory bus is formed using a microwave signal carried within a waveguide. We call this approach multicarrier memory channel architecture (MCMCA). In MCMCA, each memory signal is modulated onto an RF carrier using 64-QAM format or higher. The carriers are then routed using substrate integrated waveguide (SIW) interconnects. At the receiver, the memory signals are demodulated and then delivered to SDRAM devices. We pioneered the usage of SIW as memory channel interconnects and demonstrated that it alleviates the memory bandwidth bottleneck. We demonstrated SIW performance superiority over conventional transmission line in immunity to cross-talk and electromagnetic interference. We developed a methodology based on design of experiment (DOE) and response surface method techniques that optimizes the design of SIW interconnects and minimizes its performance fluctuations under material and manufacturing variations. Along with using SIW, we implemented a multicarrier architecture which enabled the aggregated DDR bandwidth to reach 30 Gbit/s. We developed an end-to-end system model in SimulinkTM and demonstrated the MCMCA performance for ultra-high throughput memory channel.

Experimental characterization of the new channel shows that by using judicious

frequency division multiplexing, as few as one SIW interconnect is sufficient to transmit the 64 DDR bits. Overall aggregated bus data rate achieves 240 GBytes/s data transfer with EVM not exceeding 2.26% and phase error of 1.07 degree or less.

DEDICATION

To my parents, for instilling a joy for the pursuit of knowledge in me . They inculcated an inquisitive and disciplined erudition in me that made this endeavor possible. To the memory of my father, Aboubaker, who always believed in my ability to be successful in the academic arena. You are gone, but your belief in me has made this journey possible. To the memory of my grand father Omar, you are simply my idol. Although you did not have a formal education, because you did not have a chance to go to the school under the French occupation, I never came across a person who believes in science and education as strongly as you did. You literally funded the education of your son and grand kids by depriving yourself and your family from basic life necessity. I cannot express my indebtedness and gratitude enough.

ACKNOWLEDGMENTS

First and foremost I'd like to acknowledge the tireless work of my advisor Dr. James T. Aberle. I would like to express my sincere appreciation for his guidance, patience, and dedication during my graduate studies at Arizona State University. I am also grateful to Dr. Bertan Bakkaloglu, Dr. Jennifer Kitchen and Dr. Panayiotis A. Tirkas for being a part of my graduate committee and for the interest they took in my success.

Most importantly, I sincerely express my love and gratitude to my parents from whom I have learned endurance, hard work and dedication.

Last but not least, I wish to thank my wife, Sihem, my kids Tarik, Hathamee, Zayneb and Ahmed for their support and patience throughout my PhD endeavor.

TABLE OF CONTENTS

	Page
LIST OF TABLES	ix
LIST OF FIGURES	x
CHAPTER	
1 INTRODUCTION	1
1.1 Challenges	3
1.2 Contribution	5
1.3 Publications	7
1.4 Organization	8
2 Architecture and timing of DDR memory bus	11
2.1 Overview of DDR Memory Interface	11
2.1.1 DDR Timing Overview	15
2.1.2 DDR Performance Saturation	16
3 Transmission Line Theory and Interconnect Technologies	21
3.1 Transmission Line Theory	22
3.1.1 Wave Propagation on a Transmission Line	23
3.1.2 Distributed Versus Lumped Analysis of Electric Network ...	24
3.2 Signal Integrity Impediments in the Design of High-speed Channels	26
3.2.1 Intersymbol Interference	27
3.2.2 Crosstalk	29
3.2.3 Skin Effects	33
3.3 Bandwidth and Frequency Content of Digital Waveform	34
3.4 Rectangular Waveguide	38
3.5 Substrate Integrated Waveguide	41
3.5.1 Connecting SIW to Planar Circuits	43

CHAPTER	Page
3.5.2	Substrate Integrated Waveguide Design Equations 44
3.5.3	Physics of SIW 46
4	Literature Review 50
4.1	Work on Memory Architecture and Throughput 51
4.1.1	Electrical Solution Proposals 51
4.1.2	Optical Interconnect Memory Proposals 59
4.1.3	High-speed Interconnect Using Substrate Integrated Waveguide 63
5	Multicarrier memory channel architecture 68
5.1	Interconnect Proposal 68
5.2	Performance of SIW-based DDR Interconnect 68
5.2.1	Performance of Data and Strobe Interconnect 69
5.3	Multicarrier Memory Channel Architecture Proposal 74
5.3.1	Details of MCMCA Signaling 75
5.3.2	Channel Design for Zero ISI 77
5.3.3	MCMCA Proposal 79
6	Experimental characterization of MCMCA 82
6.1	Characterization of Channel S-Parameters 82
6.2	Study of Channel Distortion Characteristics 84
6.2.1	Error Vector Magnitude and Related Figure of Merit Measurements 86
6.2.2	Distortion Measurements 87
6.3	Group Delay Characterization 90
6.4	Demodulation Set Up 92

CHAPTER	Page
6.5	End-to-end Experimental Validation of Memory Channel Proposal 93
6.6	Performance Comparison Between Classical DDR Bus and MCMCA Proposal 96
6.6.1	System Integration Perspective 97
7	Wideband interconnecting technologies for mutli-GHz MCMCA 99
7.1	Introduction 99
7.2	Haripin Filter Design 100
7.3	Hairpin Filter Performance 102
7.3.1	Full 3-D Model of the Hairpin Filter 104
7.3.2	Roughness Modeling 104
7.3.3	The Hairpin Filter Performance 105
7.4	SIW Characterization 106
7.5	MCMCA End-to-end System Model 107
7.6	System Performance at Different Data Rates 109
7.6.1	System Performance at 100 MHz 110
7.6.2	System Performance at 200 MHz 113
7.7	System Performance at 250 MHz and 400 MHz 114
7.8	256-QAM Modulation of MCMCA 116
7.9	Power Performance of the Channels 118
7.10	Discussion 119
7.11	Conclusion 121
8	Optimization of SIW Interconnect Using Design of Experiment and Re- sponse Surface Method 122
8.1	Introduction 122

CHAPTER	Page
8.2 Substrate Integrated Waveguide Interconnect	122
8.3 Design of Experiment of SIW	122
8.3.1 Design of Experiment Objective and Methodology	123
8.3.2 Response Surface Experiments for the SIW	123
8.3.3 Bandwidth Fit Model	126
8.3.4 Cutoff Frequency Fit Model	128
8.3.5 Cutoff Frequency and Bandwidth Prediction Profiler	129
8.4 Impact of Parameter Variations on Full Channel Performance	130
8.5 Conclusion	131
9 CONCLUSION	133
9.1 Conclusion	133
9.2 Recommendations	134
Bibliography	136

LIST OF TABLES

Table	Page
3.1 Propagation characteristics of common transmission lines	27
5.1 SIW design parameters	69
5.2 RWG and SIW design guide comparison	70
5.3 Usage models of available bandwidth	81
6.1 Impact of channel dispersion on 130 MHz QPSK modulated symbol ...	90
6.2 Performance of the MCMCA channel: single carrier at 500 MSymbols/s	96
6.3 Performance of the MCMCA channel: 4 symbols at 130 MSymbols/s onto 4 different carries	96
6.4 Memory bus comparison	98
7.1 Hairpin filter design parameters	102
7.2 Performance of 64-QAM channels at 100 MHz	112
7.3 Performance of 64-QAM channels at 200 MHz	112
7.4 250 MHz performance	116
7.5 400 MHz performance	116
7.6 Performance of 256-QAM MCMCA channel at 100 MHz	116
7.7 SIW channel performance at 256-QAM 200 MHz and 250 MHz	117
7.8 SIW channel power performance	119
8.1 SIW design parameters	125

LIST OF FIGURES

Figure	Page
1.1 Trends of I/O computing bandwidth demand and I/O data rate	4
2.1 Source synchronous clock architecture.	13
2.2 DDR SDRAM array architecture	14
2.3 Source synchronous timing diagram	16
2.4 DDR3/4 clock and data signals topology	17
2.5 Comparison of CMD signal network (a) in DDR2 standard and (b): in DDR3/4 standard.	18
2.6 Fly-by topology	19
3.1 (a) Transmission line representation. (b) Equivalent circuit for an in- finitesimally short segment of transmission line	22
3.2 Waveform propagation through a network with lumped circuit	25
3.3 Bits smearing due to ISI in the interconnect channel.	28
3.4 Coupled transmission lines (a) Layout, (b) Electrical model	30
3.5 Odd and even fields distribution between two transmission lines	32
3.6 Trapezoidal waveform	34
3.7 Plot of the magnitudes of Fourier coefficients of a square wave	35
3.8 Plot of the envelope bounds of $\frac{\sinh(x)}{x}$ on logarithmic axes	36
3.9 Plot of the magnitude and bounds of Fourier coefficients for both the trapezoidal and rectangular pulse of $\frac{\sinh(x)}{x}$ on logarithmic axes	37
3.10 Geometry of a rectangular waveguide	39
3.11 Substrate integrated waveguide geometry and microstrip transitions captured from Ansoft HFSS.	42
3.12 Microstrip to SIW transition using taper.	43
3.13 Optimization of taper length L.	45

CHAPTER	Page
3.14 Magnitude of the dominant TE_{10} electric field in SIW at 9.0 GHz.	47
3.15 Conduction, dielectric and total loss of an X-band SIW.	49
4.1 Fully buffered DIMM architecture	52
4.2 Advanced memory buffer bloc diagram	54
4.3 Noncoherent ASK memory interconnect	55
4.4 Memory transceiver with crosstalk suppression scheme	57
4.5 DDR transceiver with skew cancellation	58
4.6 Optically connected memory module	60
4.7 experimental setup	60
4.8 4 x 2 Optical SRAM bank architecture	62
4.9 7.6-cm SIW interconnect test structure	64
4.10 S_{21} characteristics of 45° and 90° bend SIW interconnect test structure	65
4.11 Hybrid substrate integrated waveguide	66
4.12 Experimental setup for characterization of multi-mode SIW	67
5.1 Performance of SIW DQ/DQS signal using Ansoft HFSS.	71
5.2 Ohmic loss of SIW with varying guide length	71
5.3 SIW s_{11} (dB) and s_{21} (dB) with optimal taper dimensions (L=3.52 mm, W=4.4mm) using Ansoft HFSS.....	73
5.4 SIW mode coupling simulated using Ansoft HFSS.	73
5.5 Schematic of the multicarrier memory channel architecture proposal ...	74
5.6 QAM modulation and demodulation diagram	77
5.7 Pulses with a raised cosine spectrum	78
5.8 Proposed frequency division multiplexing scheme	79
6.1 SIW VNA test bench.....	83

CHAPTER	Page
6.2 Measured and simulated channel S-parameters	83
6.3 Conceptual setup for measuring distortion	84
6.4 Experimental set-up for channel distortion characterization.....	85
6.5 The error vector magnitude concept (EVM).....	86
6.6 Symbol distortion for 100 MSym/s QPSK symbol when the carrier is in very close to the cutoff frequency (5.6 GHz)	88
6.7 Symbol distortion for 100 MSym/s QPSK when the carrier is farther from the cutoff frequency (6.60 GHz)	89
6.8 Symbol distortion for 100MSym/s QPSK at 9.6 GHz carrier	90
6.9 Measured group delay of the channel.....	91
6.10 SIW channel performance using 64-QAM at 500 MHz symbol rate at 9.6 GHz carrier	92
6.11 SIW channel performance using 64-QAM and sending simultaneously 4 symbols at 130 MSym/s symbol Rate	93
6.12 Signal analysis at the receiver using digital signal analyzer and vector signal analyzer	94
7.1 Structure of half-wavelength (a)Parallel-coupled resonator and (b) U- Shape resonator	101
7.2 Schematic of hairpin bandpass filter.	103
7.3 Performance of the ideal hairpin bandpass filter.	103
7.4 Huray's snowball model of surface roughness	105
7.5 Full wave simulation and 3D model of hairpin filter captured from Ansoft HFSS.....	106
7.6 Fabricated SIW bandpass filter.	107

CHAPTER	Page
7.7 SIW measurement.	108
7.8 Full channel Simulink system model of MCMCA.	109
7.9 Eye diagram and constellation diagram of MCMCA system using SIW at 100 MHz data rate.	111
7.10 Eye diagram of MCMCA using TLBPF at 100 MHz and rolloff=0.2. . . .	111
7.11 Constellation diagram and eye diagram of MCMCA using SIW at 200 MHz and rolloff=0.3.	113
7.12 Eye diagram and constellation diagram of MCMCA using TLBPF at 200 MHz and rolloff=0.8.	114
7.13 Constellation diagram of MCMCA using SIW at 250 MHz and rolloff=0.3.	115
7.14 Constellation and eye diagram of MCMCA using SIW at 400 MHz and rolloff=0.7.	115
7.15 Constellation and eye diagram of MCMCA using SIW at 256-QAM 250 MHz and rolloff=0.8	118
7.16 Symbol power spectrum of 64-QAM SIW channel at 200 Mbps and Rolloff=0.3	120
8.1 Response surface method	124
8.2 DOE flow	124
8.3 SIW RSM experiments	125
8.4 S-parameters of the SIW DOE experiments	126
8.5 Bandwidth RSM model fit and residual distribution	127
8.6 Cutoff frequency FC RSM model fit and residual distribution	128
8.7 Prediction Profiler of the bandwidth and the cutoff frequency RSM models	130

8.8	Eye and constellation diagram of (a) the channel with best case SIW	
	(b): the channel with worst case SIW	132

Chapter 1

INTRODUCTION

Integrated circuit processing speed is increasing exponentially over time. This is mainly due to success in CMOS scaling trend that follows or exceeds Moore's law [2, 3]. It is common nowadays for single chip CPU to exceed one TeraFLOP/s [4] which is equal to 10^{12} FLOPS. The unit FLOPS stands for floating point operation per second and is used as a metric for processing benchmarking [5]. The increase in computing power increased the demand for input/output (I/O) bandwidth. Unfortunately, the speed of I/O and channel interconnects have not kept pace. Thus, processor based systems are I/O and interconnect limited. Recently, many solutions have been proposed[6–8] and implemented for low-pin-count low-density interfaces. Solutions to this class of channels consist of a combination of point-to-point differential architecture, preemphasis, equalization and multilevel coding instead of classical non-return-to-zero (NRZ) signaling. Unfortunately, most of these techniques are not applicable to memory interfaces and the disparity between the speed of memory interfaces and serial interfaces has grown substantially.

The ITRS roadmap predicts that high-performance off chip I/O speed will exceed 40 GHz by 2020 through the use of point-to-point interconnects[9][10]. On the other hand, multidrop memory bus speed will remain at around 4-5 GHz. The memory aggregated bandwidth is not scaling fast enough to keep up with increasing bandwidth demands. This increasing disparity and widening gap between processor performance and memory bandwidth has been termed the "memory wall" by many authors[1].

At one time, many researchers believed that the use of optical interconnects would

be the definitive solution to the memory speed bottleneck. Optical interconnects have many advantages including inherent parallelism, large bandwidth, immunity from crosstalk and electromagnetic interference, and lower signal and clock skew. Unfortunately, the optical solution has not been widely adopted because of the inability to easily integrate optical components such as laser diodes, photo detectors, lenses, and mirrors into high-density, high-scalability applications in a cost-effective manner. We have come to realize that the fundamental advantages of the optical solution can be preserved while overcoming its major disadvantages by shifting the carrier wave frequency from the optical to the microwave frequency range. Microwave transceivers are much easier to realize in standard CMOS processes than optical transceivers and do not require exotic components such as laser diodes. Thus, they can be easily integrated onto the system motherboard or chip package, and perhaps even within a chip itself.

This dissertation focuses on alleviating the memory wall bottleneck by proposing innovative solutions on two fronts: architecture and interconnect. On the architectural front, we propose to map the conventional DDR bus to a microwave link using a multicarrier frequency division multiplexing scheme. We call this approach multicarrier memory channel architecture (MCMCA). In MCMCA, each memory signal is modulated onto an RF carrier using 64-QAM or higher format. The carriers are then routed using substrate integrated waveguide (SIW) interconnects. At the receiver the memory signals are demodulated and then delivered to SDRAM devices. On the interconnect front, we implement the recently introduced substrate integrated waveguide technology as media to transmit DDR signals. We believe that this is the first work that proposes to use substrate integrated waveguide within a memory interconnect. We present the theoretical details of our proposal as well as the results of simulations and experiments that demonstrate the merits of this approach.

1.1 Challenges

Technology scaling continues to follow Moore's law in terms of transistor speed and density, hence, the amount of available computing power at the CPU level becomes enormous. As available computing power keeps increasing with each scaled CMOS process node, computing systems, like servers and personal computers need faster memory interfaces. DDR interface, however, is unable to scale up its bandwidth in response to the increasing bandwidth demand. DDR speed per pin barely reaches 3 Gbps for DDR3/4 and is expected to saturate at speed less than 5 Gbps. DDR buses use both single-ended and differential signaling to transmit memory bits. Therefore, a memory channel is subject to severe signal integrity degradation from impedance mismatch, signal reflection, crosstalk, intersymbol interference (ISI) and jitter.

Serial point-to-point I/O(s) use advanced equalization algorithms, preemphasis, deemphasis and complicated power-hungry noise cancellation techniques in order to mitigate signal degradation problems. Because serial interconnects are not usually a wide bus, they benefit from such techniques while keeping the overall power cost reasonably manageable. These techniques are not applicable for DDR for the obvious reason of excessive power penalty associated with them. As a result, DDR data rate per pin suffers a large gap compared to serial transceivers. The gap continues to grow as transistor critical dimension gets smaller which leads to the so-called memory wall bottleneck. A graphical summary of the memory gap in comparison to other IOs and required bandwidth is shown in Fig. 1.1.

We note that research trends in this field propose either using an optical interconnect as alternative to copper transmission line or adopting a modified version of the solutions implemented in high-speed point-to-point serials. To alleviate memory channel bottleneck, we come to conclude that both the channel architecture and the

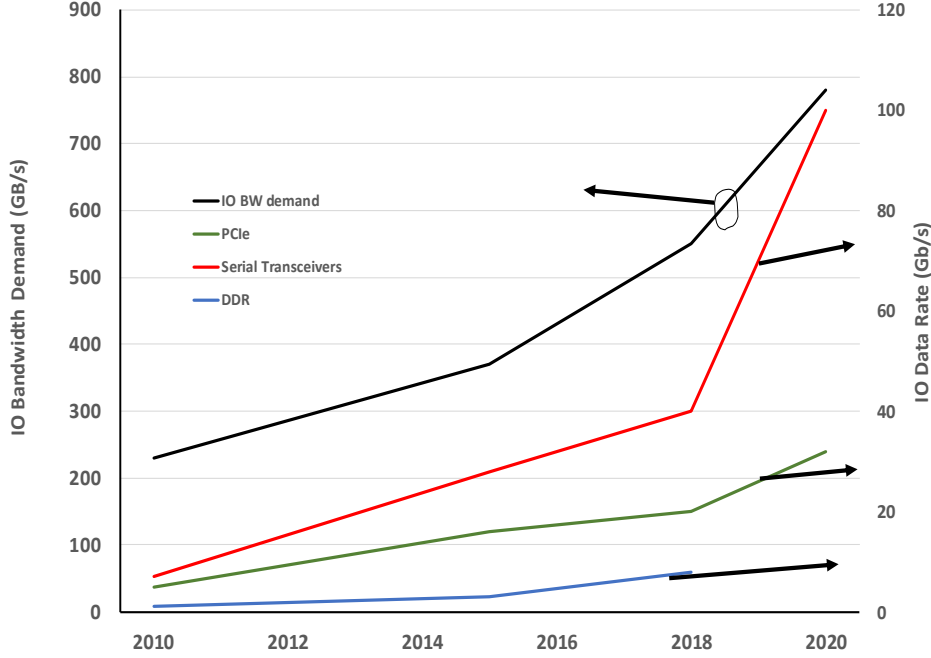


Figure 1.1: Trends of I/O computing bandwidth demand and I/O data rate[9][10]. Note that curves for PCIe and DDR beyond 2018 are an extrapolation based on their respective historical trend[11]

interconnect media need to be re-invented in order to unleash computing power.

This thesis proposes and implements a novel memory channel architecture which we called Multicarrier Memory Channel Architecture (MCMCA) where we introduce the novel idea of transmitting DDR signals onto a multicarrier channel. We also pioneered the use of wide SIW bandpass filters as the interconnect medium for memory signals. The proposed solution achieves 30 Gb/s data rate and considerably alleviates the memory bandwidth bottleneck.

1.2 Contribution

This dissertation presents the reader with a concise overview of the state-of-the-art for DDR bus interconnects. We expose the challenges and bottleneck caused by memory at both sides of the interface. On the CPU side, memory bandwidth cannot keep up with the computational power made available by the CPU. On the peripheral side, the DDR I/O data rate is at disparity with serial I/O transfer rates.

An orthogonal, multicarrier frequency division multiplexing scheme is proposed to serve as a high bandwidth architecture for memory channel. We shift the DDR signals from the baseband to the RF band. The signals are optimally filtered and modulated in a spectrally efficient format using raised cosine filtering and quadrature amplitude modulation format (QAM). The filtering optimally minimizes the ISI while the QAM modulation maximizes the number of bits transferred per cycle. Using the SIW as a novel interconnect for memory augments the solution with an exceptionally wide band channel. From manufacturing perspective, our solution has the advantage of being a low cost solution and highly compatible with planar manufacturing process.

As opposed to optical interconnect whose fabrication process is expensive, bulky and incompatible with planar technologies, MCMCA is an efficient, low cost solution and compatible with planar process. Optical interconnects require a great deal of alignment engineering, and require the addition of many components to convert between electrical and optical domains.

The effectiveness of our proposal is proved in this work by means of design, simulations and experimental characterization. We resolved integration challenges of SIW with planar structures, and demonstrated performance matching of SIW with classical waveguide structures. In brief, we claim the following thesis contributions:

A- **Innovative approach:** By addressing the memory wall problem from a radically

different angle than approaches adopted by researchers in the field, we opened the door for whole new lines of thought, and research opportunities.

By looking at memory as a communication channel and treating the transmitter, receiver and the propagation media as one unit, we were able to leverage higher order modulation, multicarrier concept and orthogonal signaling techniques. Such an approach enabled us to introduce a new concept that can be further developed by our colleagues and the researcher community at large.

B- Architectural innovation: With a new architecture proposal, memory channel is no longer limited by transmission line low pass transfer attenuation. The transmitter and receiver, instead of being a CMOS buffer, are designed as an I-Q modulator, with pulse shaping filter that optimizes spectral efficiency and maximizes the number of bits/symbol. The proposed architecture enabled a breakthrough in memory bandwidth and transfer rate.

C- Interconnect medium: We pioneered the usage for the first time of SIW as medium to carry memory signals and bits. We demonstrate in this thesis the substantial advantage of waveguiding structure compared to transmission lines. The advantages include reduction of attenuation, propagation loss, dispersion, cross-talk and interference. The SIW offers a large bandwidth bandpass channel to carry DDR modulated symbols.

D- Gap closure: In theory, our approach is also applicable to other high-speed digital applications. We specifically targeted DDR signals for the following reason: our proposal is more likely to benefit DDR than any other interconnects, and hence it the optimal candidate to reduce the gap that separates DDR from other high-speed IOs. Serial transceivers have overcome, with remarkable success, the transmission line lowpass attenuation by means of sophisticated equalization and

noise cancellation techniques. Those signaling techniques enabled serial I/O(s) to reach and exceed 25 Gbps. The likelihood that serial I/O(s) benefit from our techniques is slim when we take into account the power and design complexity of adding an up and a down conversion blocks, a filtering and a pulse shaping block. On top of those factors, a serial transceiver is usually used in a single port configuration, or is an element of a narrow bus. Therefore, the cost in power, real estate and design complexity is manageable. DDR however, is intrinsically a very wide bus, where any excess in power or area is replicated on wide bus and the overall cost could quickly exceed the product budget.

E- SIW Interconnect optimization and system perspective: We developed a methodology based on design of experiment and response surface techniques that enables the designer to maximize the MCMCA channel bandwidth while minimizing the effect of material and manufacturing fluctuations on the channel performance. We used a complete end-to-end system and demonstrated that the cost of failing to mitigating manufacturing variabilities could be in the hundred of megahertz and even couple gigahertz in total channel throughput

1.3 Publications

- I-) B. Bensalem and J.T. Aberle. A new high-speed memory interconnect architecture using microwave interconnects and multicarrier signaling. *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, 4(2):322-340, 2014[12]
- II-) J.T. Aberle and B. Bensalem. Ultra-high-speed memory bus using microwave interconnects. In *Electrical Performance of Electronic Packaging and Systems*

(EPEPS), 2012 IEEE 21st Conference on, pages 3-6, 2012[13]

- III-) B. Bensalem and J.T. Aberle. Wideband interconnecting technologies for multi-GHz novel memory architecture. Submitted to *Components, Packaging and Manufacturing Technology, IEEE Transactions on*
- IV-) B. Bensalem and J.T. Aberle. Optimization of substrate integrated waveguide memory interconnect using design of experiment. Submitted to *2018 IEEE Symposia on VLSI Technology and Circuits*.
- V-) B. Bensalem and J.T. Aberle. Effects of manufacturing variation on ultra-high-speed memory interconnects. To be submitted to *Electrical Performance of Electronic Packaging and Systems (EPEPS), 2018 IEEE 21st Conference on, 2018*
- VI-) B. Bensalem and J.T. Aberle. High-bandwidth memory channel architecture using custom OFDM and microwave interconnects. To be submitted to *Microwave Theory and Techniques, IEEE Transactions on*

1.4 Organization

In order to comprehend the need for a radical change in both architecture and signaling to classical DDR channel, an in-depth review of the limitations and the challenges in high-speed transmission line links is necessary. Therefore, Chapter 2 presents an overview of the fundamental architecture and timings of classical DDR bus. In chapter 3, we present a review of the theory of transmission lines and an assessment of high-speed interconnect technology. The limitations of transmission line interconnect like cross-talk, attenuation, skin depth, jitter, and frequency dependent loss are treated in

details and upper frequency limits of transmission line signaling is reviewed. The theory of operation and characteristics of rectangular waveguide and SIW interconnect technology are also reviewed in chapter 3. The waveguiding structure is analyzed in frequency and time domain, and a bandpass channel using SIW technology is analyzed and characterized using full 3D electromagnetic field solver. In chapter 4, we present a concise overview of a selected set of relevant literature representative of the research work aiming to propose solutions to DDR shortcomings and bandwidth bottleneck. Chapter 5 focuses on our novel memory architecture proposal, which consists of dividing the large bandwidth of the SIW into many channels and mapping a DDR signal onto QAM modulated symbols where each symbol occupies a separate channel. The symbols are filtered at the transmitter using a Nyquist filter that minimizes ISI and optimally shaped before being sent over the channel. We detail the trade-off between number of channels, symbol rate and aggregated MCMCA throughput, and we present a set of tables and equations that governs our choice.

Chapter 6 describes the experimental validation of MCMCA proposal. The SIW interconnect is characterized using a vector network analyzer, where we measure the channel bandwidth and channel distortion characteristics. One known issue with SIW in general is the highly dispersive responses near the cutoff frequency. We identify the nonlinear frequency range in order to operate the channel outside that range. We define the figure of merit of the channel, then we measure it end-to-end and quantify the error vector magnitude (EVM) and phase error for different carrier and symbol rate arrangements. We also address the system integration aspects of our proposal and conclude with a summary of MCMCA performance and its potential for further improvement. We do a concise review of the state of the art of multicarrier transceivers in main stream CMOS technology and quantify the chip overhead and the implementation complexity compared to DDR memory transceivers. The

MCMCA throughput and signal quality advantages are viewed in perspective of die area overhead and modulator/demodulator design complexity. Chapter 7 compares the performances of MCMCA when using the SIW interconnect vs. MCMCA using state of the art high-bandwidth transmission line bandpass filter, namely the hairpin filter. We do a thorough comparison and layout the advantages and disadvantages of both competing interconnects. In chapter 8, we address manufacturing deviation and develop a DOE based methodology to maximize the throughput and minimize the impact of material and manufacturing variabilities on MCMCA performance.

Finally in chapter 9, we summarize the dissertation, provide a concise conclusion and present our recommendations for further research and investigation.

Chapter 2

ARCHITECTURE AND TIMING OF DDR MEMORY BUS

A memory bus is essentially a parallel interface that uses printed transmission lines to route memory signals between the memory controller and memory chips. As data rates increase and interconnect nuisances get more difficult to handle in a simplistic single-ended signaling technique, DDR has adopted some of serials and differential signaling schemes to part of its signals (strokes and clocks for DDR3 and DDR4). DDR4 goes one step further in mimicking serials by allowing some sort of equalization and noise cancellation at the controller side. Signal terminations evolved considerably as well. DDR4 termination is not the simple termination adopted in early DDR up to DDR3. DDR4 termination uses a fairly complex logic to implement a dynamically adjustable on-die termination in order to optimize eye opening and minimize impedance mismatch reflections.

We introduce in this chapter the details of DDR signaling protocol, its architecture and the timing constraints that the protocol needs to satisfy. We present the DDR3 and DDR4 fly-by architecture and compare it to previous DDR architecture. We summarize the advantages and limitations of the fly-by topology and demonstrate the need for high memory bandwidth to close the gap with SERDES I/Os and respond to the increasingly large computation bandwidth demand.

2.1 Overview of DDR Memory Interface

Memory is used in computer system to store data. SDRAM, which stands for synchronous dynamic random access memory, is a form of computer data storage. Access

to any piece of SDRAM data can be performed in a constant time regardless of its physical location; hence the origin of the name random access memory. A system clock synchronizes the access mechanism, which makes it a synchronous protocol. Double data rate (DDR), is the dominant operation mode in SDRAM bus. Data is latched on both the rising and falling edge of the strobe, which results in doubling the throughput. The DDR interface is a parallel architecture with a wide bus of data; usually 64 bits wide. It uses Source Synchronous (SS) timing protocol whose architecture is shown in Fig. 2.1 [14].

A DDR SDRAM is composed of numerous arrays of capacitive charge cells used to store data. Each array is organized into banks independent of each other.

The individual cells are accessed using column and row address decoders. A block diagram of a DDR SDRAM array is shown in Fig. 2.2[14]. The memory controller performs data access by sending a command in conjunction with a row/column address using the address bus. An ACTIVATE (**ACT**) command is first sent which sends the entire row of the bank to the sense amplifier. If the operation is a WRITE (READ) operation, data is written (read) into the sense amplifier. The time it takes between an ACTIVATE and WRITE or READ command is the row to column delay timing parameter denoted by $t_{\mathbf{RCD}}$ and is a figure of merit of memory technology that is constrained by JEDEC standard[15].

Following the WRITE(READ) operation to the sense amplifier, the controller issues a PRECHARGE (PRE) command that takes an amount of time $t_{\mathbf{RP}}$ (row precharge time), and resets the sense amplifier and bit lines to prepare for the next row access.

Once the row has been precharged, we need to wait for a time $t_{\mathbf{RC}}$ (row cycle) between subsequent **ACT** commands to the same bank.

The capacitive charge stored in the memory cells is subject to leak over time. The

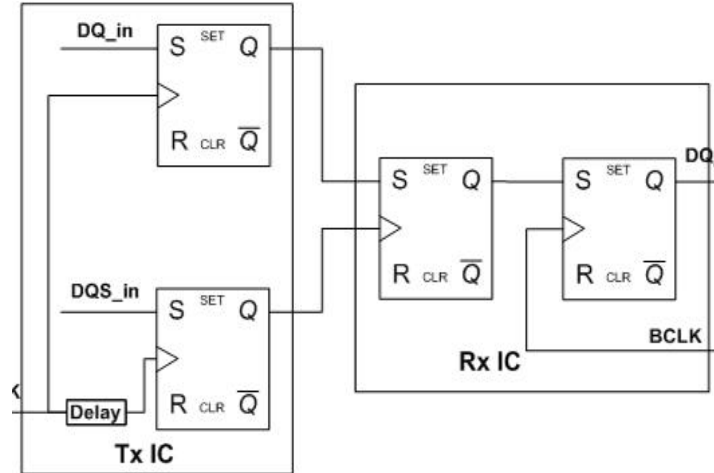


Figure 2.1: Source synchronous clock architecture.

charge needs to be refreshed in order to ensure that data is not lost. The amount of time required to refresh the charges is called t_{RFC} (refresh command) and is the required timing constraint between consecutive REF or ACT command.

DRAM latency The SDRAM latency is quantified in number of clock cycles and refers to the delays incurred in transmitting data between the CPU and the SDRAM. Latency dictates the upper limit on how fast information is transferred between CPU and the SDRAM. The major constituents of latency are:

- t_{CL} Column address strobe tCL: also known as CAS (column address strobe) latency. It is the number of clock cycles from the column address phase to first available input data. It is sometimes referred as tCAS as well.
- t_{RCD} : RAS-to-CAS delay. tRCD stands for row address to column address delay time. It accounts for the number of clock cycles required between an active RAS command and asserting a CAS command during the subsequent

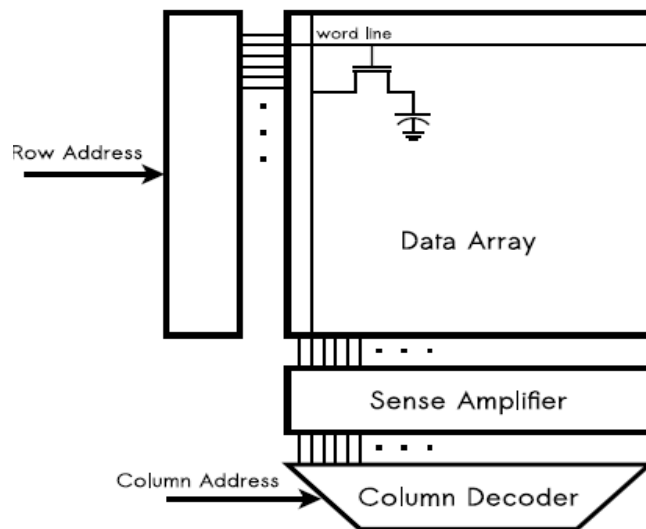


Figure 2.2: DDR SDRAM array architecture[14].

read or write command.

- t_{RP} : row precharge. It is the number of clock cycles needed to terminate access to an open row of memory, and open access to the next row
- t_{RAS} : row address strobe time is the minimum number of clock cycles needed to access a row of data between the data request and the precharge command. It's also known as active to precharge delay.

The memory latency is constrained by the following two equations that need to be satisfied

$$t_{RC} = t_{RAS} + t_{RP} \quad (2.1)$$

$$t_{RAC} = t_{RCD} + t_{CAS} \quad (2.2)$$

2.1.1 DDR Timing Overview

DDR data (DQ), Strobe (DQS), Clock (CK), Command and Address (CMD/ADD) signals use SS bus architecture. A generic SS timing diagram is depicted in Fig. 2.3.

Timing equations are derived based on the fact that the sum of timing delays at the receiver must equal the sum of timing delays at the transmitter. This constraint yields

$$T_{va} = T_{HD} + T_{HDMargin} + T_{HDSkew} \quad (2.3)$$

$$T_{vb} = T_{SU} + T_{SUMargin} + T_{SUSkew} \quad (2.4)$$

where T_{SU} is the setup time which specifies the minimum amount of time that the valid data must be present prior to the input clock edge in order to guarantee successful capture of the data. T_{HD} is the hold time which specifies the minimum amount of time that the valid data must remain after the input clock edge. T_{va} is the minimum driver phase input for hold time, T_{vb} is the minimum driver phase offset for the setup. T_{HDSkew} is the hold flight time skew, T_{SUSkew} is the setup flight time skew. $T_{SUMargin}$ is the setup margin, and $T_{HDMargin}$ is the hold margin

Improvement in I/O transceiver technology has led to reasonably small I/O setup (T_{SU}) and hold (T_{HD}) values. However, skew has only got worse. Thus, SS bus performance is skew limited. As platform complexity increases, design of the memory bus is getting more difficult to achieve within an ever shrinking timing budget. It is common now to have 10, 12 or even 30 layers on the platform board where the memory bus has to be routed. A typical network of DDR multidrop signals is shown in Fig.2.4, which shows the topology of clock (CLK) and data signals for raw card A, unbuffered by 8 single rank DDR3/4 DIMM[16]. Routing of DDR signals over many layers causes them to suffer from impedance discontinuities, via stub resonance, jitter, simultaneous switching noise, broken return path reference plane, ISI, and excessive

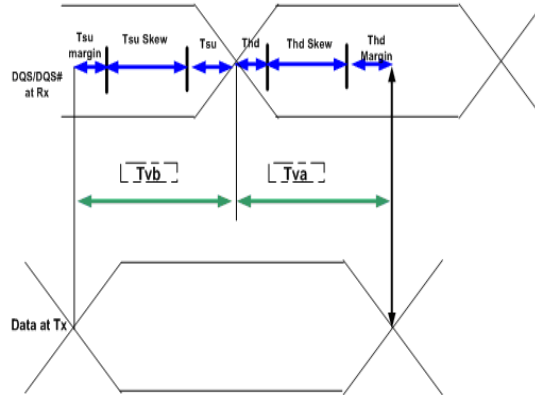


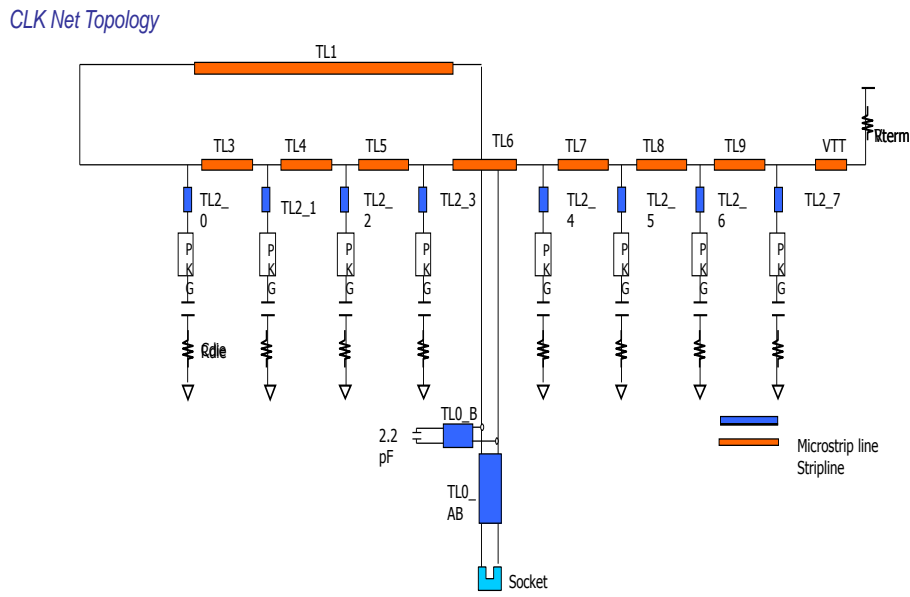
Figure 2.3: Source synchronous timing diagram[14].

crosstalk, in particular in the escape and break out area. The cumulative effects of these factors can easily lead to severe degradation of the memory signal at the receiver.

2.1.2 DDR Performance Saturation

Signal network topologies are standardized by the IEEE JEDEC organization [17] to allow maximum compatibility among different vendors. Since memory signals are digital bits of information, they are intrinsically very wide band signals, and thus are very sensitive to dispersive effects.

DDR3 and DDR4 standards use fly-by topology departing from older DDR2 topology which suffers from multidrop signaling effects[18]. The switch to the fly-by topology is by far the most radical change that DDR has undergone since its inception. The reflection effects of multidrop stubs causes excessive noise that limited DDR2



Data Net Topology

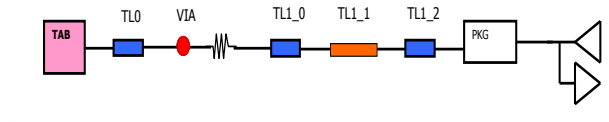
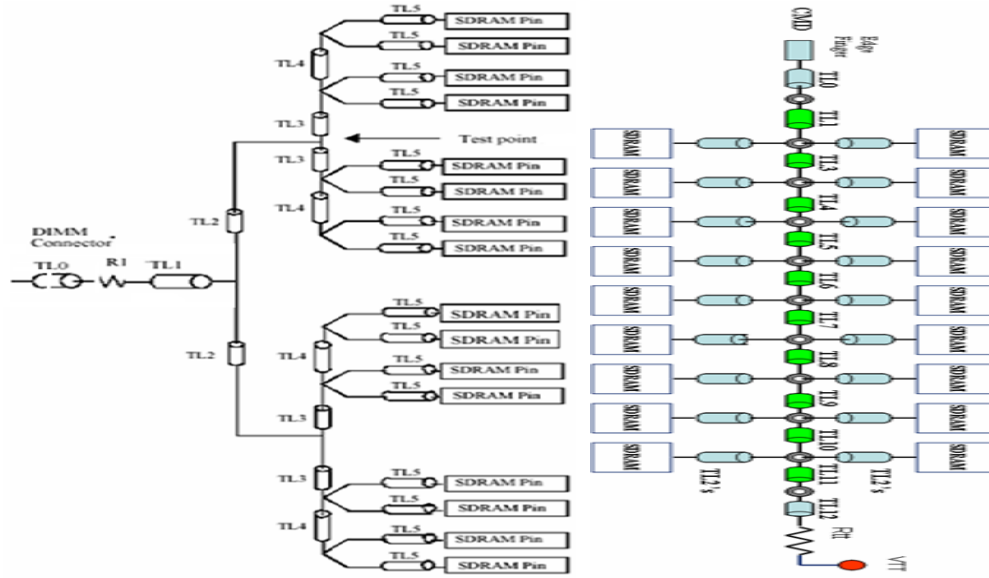


Figure 2.4: DDR3/4 clock and data signals topology[16].

interface to a transfer rate of less than 1 Gb/s. The introduction by JEDEC of fly-by topology enabled DDR to reach much higher data rate. It is expected that DDR4 will reach 4 Gb/s. The Fig.2.5 display side-by-side CMD/ADD topology in DDR2 style and in the fly-by style of DDR3/4. The stubs used in DDR2 are eliminated in DDR3 and the branching out to the SDRAM chip is made very short in fly-by so that it has very negligible effects on the net signal integrity.

In reference to the general fly-by topology shown in Fig. 2.6, DDR3 and DDR4 implement many new features to cope with signals reaching SDRAM at different times. Among those features:



(a) DDR2 multidrop network

(b) DDR3 Fly-by network

Figure 2.5: Comparison of CMD signal network (a) in DDR2 standard and (b): in DDR3/4 standard[17][16]

- Read calibration: Fly-by topology causes CMD/ADD/CTRL/CLK to arrive at different times per memory chip, while DQ/DQS arrive at the same time to all SDRAM chips in the rank. DRAM is augmented with a built-in predefined pattern stored in a special register and is trained at memory initialization. A set of delays is stored in the register memory and is used to adjust signal delays and calibrate offset between signals.
- Write leveling: During Write Leveling the memory controller needs to compensate for the additional flight time skew delay introduced by the fly-by topology with respect to strobe and clock. DRAM contains built-in mode phase detector. The controller uses a programmable delay element on DQS with fine enough granularity so that the proper delay can be inserted to compensate for the additional skew delay.

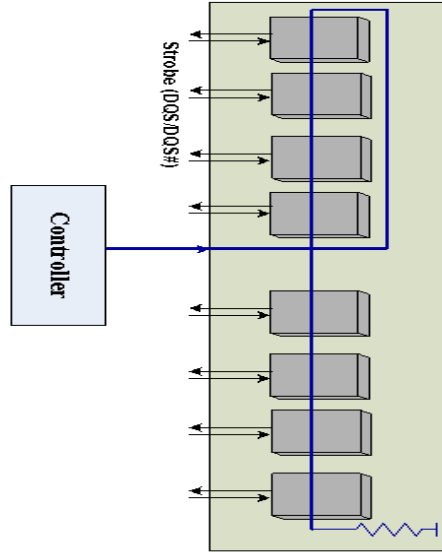


Figure 2.6: Fly-by topology [18].

The calibration features in DDR3/4 controllers ease some of the routing constraints and provide the system designer with an added level of flexibility compared to old DDR systems.

While Fly-By topology resolves most of the problems caused by multidrop stubs, it has its own problems and limitations such as:

- Improvement in overall bus data rate is not enough to close the gap with serial interfaces. It is expected to run at 2.5-to 4 Gbit/s maximum.
- Controller design for Fly-By bus becomes much more complicated than that of DDR2.
- Signal de-skewing becomes very challenging. De-skew needs to be performed on a device basis, rather than globally as in DDR2. Depending on the SDRAM device order in the chain, a different flight time de-skewing number needs to be programmed by the controller.

- The Fly-By architecture achieves only an incremental improvement to signal integrity of DDR signals. The fundamental susceptibility of the signals to impedance mismatch, crosstalk, and attenuation remains considerable and these impediments are not resolved in a radical manner.

DDR3 and DDR4 are both incremental, evolutionary improvements to existing DRAM standard. Both standards improve power consumption and transfer rate relative to DDR2, but they are not a large upswing.

A considerable amount of effort aiming to alleviate the memory bottleneck is being deployed both in commercial and academic world. As a result, many standards and alternative architectures are proposed and/or under development like wide I/O[19], hybrid memory cube (HMC)[20] and fully buffered dual in-line memory module (DIMM)[21] to name few. However, they all remain incremental improvements and are unable to resolve in an economical way the thorny problems of signal integrity.

To appreciate the difficulty of the hurdles that limit the success of these efforts, we first present an in-depth account of signal integrity challenges encountered by memory interface. We then review a set of relevant and representative standards and proposals. We highlight their achievements and outline their shortcomings.

Chapter 3

TRANSMISSION LINE THEORY AND INTERCONNECT TECHNOLOGIES

A printed circuit board (PCB) electrical circuit network uses traces of copper as transmission lines to connect the different parts of the network. As frequency increases, DDR signals become subject to severe signal integrity impediments that consume a considerable percentage of signal margin which can exceed 40% of the available budget as in the case of DDR3 and DDR4.

The signal integrity problems briefly mentioned in section 2.1.2 are attributable to the distributed nature of high-speed interconnections. A deep understanding of transmission line theory and electromagnetic coupling mechanisms are mandatory prerequisites for any serious attempt to comprehend and alleviate high-speed signaling. We present in this section a concise overview of transmission line theory and details of signal integrity impairments in DDR bus. The SI analysis demonstrates the performance limitations of the DDR bus and sets the stage for a radically different approach that enables the required breakthrough in the memory bandwidth. With that perspective, we then present the physics and performance of the SIW which is a viable interconnect for multi-gigahertz data rates as an alternative to the transmission line interconnect used in classical DDR interconnect. In doing so, we lay the foundation for our novel memory channel proposal that is detailed in chapter 5.

3.1 Transmission Line Theory

A transmission line is represented schematically as a two wire-line as depicted in Fig. 3.1 (a). An incremental of transmission line of a length ΔZ can be modeled as lumped-element equivalent circuit as shown in Fig. 3.1 (b).

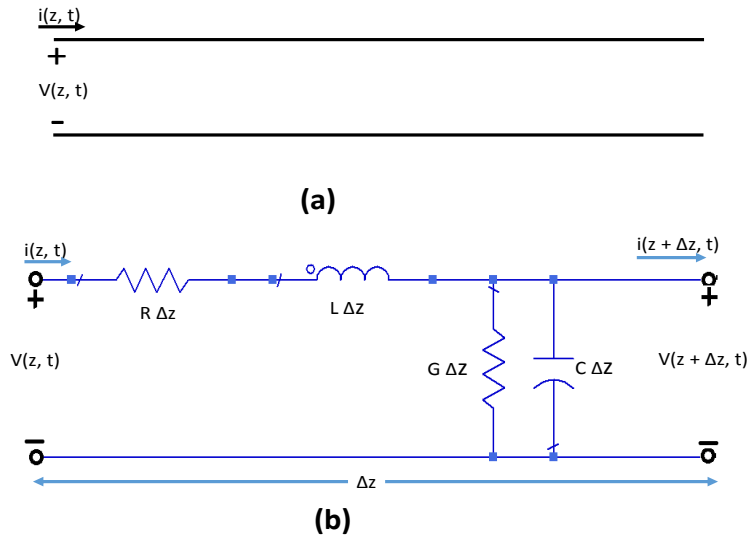


Figure 3.1: (a) Transmission line representation. (b) Equivalent circuit for an infinitesimally short segment of transmission line[22]

$R \Delta Z$ models the incremental Ohmic loss due to finite conductivity of the copper. $L \Delta Z$ accounts for the inductive effect caused by current flow in the transmission line. $G \Delta Z$ and $C \Delta Z$ model respectively the dielectric loss and capacitance between signal and return path of the line.

3.1.1 Wave Propagation on a Transmission Line

It can be shown [22] that wave equations for $I(z)$ and $V(z)$ obey the second order differential equation:

$$\frac{d^2V(z)}{dz^2} - \gamma^2V(z) = 0 \quad (3.1)$$

$$\frac{d^2I(z)}{dz^2} - \gamma^2I(z) = 0 \quad (3.2)$$

where

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)} \quad (3.3)$$

is the propagation constant. Solution to the above equation can be found as:

$$V(z) = V_0^+ e^{-\gamma z} + V_0^- e^{\gamma z} \quad (3.4)$$

$$I(z) = I_0^+ e^{-\gamma z} + I_0^- e^{\gamma z} \quad (3.5)$$

Characteristic impedance Z_0 is defined as:

$$Z_0 = \frac{R + j\omega L}{\gamma} = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \quad (3.6)$$

The ratio of reflected wave to the incident voltage wave is known as the reflection coefficient and is a measure of the impedance mismatch between the transmission line and the load[22].

$$\Gamma = \frac{V_0^-}{V_0^+}; \quad -1 \leq \Gamma \leq 1 \quad (3.7)$$

$$= \frac{Z_L - Z_0}{Z_L + Z_0} \quad (3.8)$$

$$(3.9)$$

where Γ is the reflection coefficient, V_0^+ the incident wave, V_0^- the reflected wave, Z_0 the transmission line characteristic impedance, and Z_L the load impedance.

When the load is not matched, part of the delivered power is reflected. A quantity called return loss (RL) can be defined as the ratio of the incident power to the reflected power expressed in dB.

$$RL = 10 \log_{10} \left[\frac{P_{in}}{P_{ref}} \right] \text{ dB} \quad (3.10)$$

$$= -20 \log |\Gamma| \text{ dB} \quad (3.11)$$

where P_{in} and P_{ref} are the incident and reflected power respectively. In contrast to most "losses", it is usually desired to make the return loss as large as possible.

Transmission coefficient T is defined as $1 + \Gamma$, from which we define the insertion loss (IL) between two points in a circuit expressed in dB.

$$IL = -20 \log_{10} |T| \text{ dB} \quad (3.12)$$

3.1.2 Distributed Versus Lumped Analysis of Electric Network

In high-speed regime, conventional circuit theory based on Kirchhoff's law, might not be accurate in describing signal behavior. Depending on the electrical size of the circuits, a distributed description of signal propagation is required. Circuit analysis theory assumes that the physical dimensions of an electric circuit are negligible, i.e., much smaller than the electrical wavelength of the highest frequency present. As the frequency of operation increases, the wave length decreases. At high enough frequency, the wavelength becomes comparable to the circuit dimension. Under these circumstances, one needs to treat the circuit as a distributed network rather than a lumped network. The waveform at one point of the circuit is different than the waveform at another point of the circuit. We develop here a set of guidelines that

helps one decide when to use the lumped assumption and when to revert to distributed treatment and analysis.

Using similar notations and approach as in [23]; Fig. 3.2 shows a lumped circuit connected to the rest of the electric network with electric wires. If the total network length is \mathcal{L} , the relationship between circuit dimension and wavelength is governed by the following set of equations[23]:

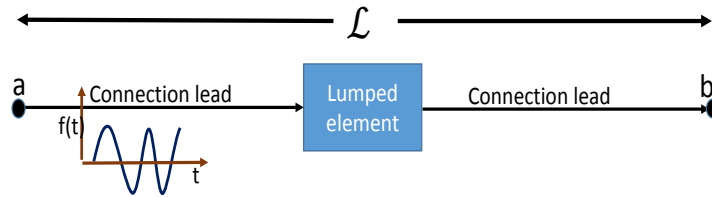


Figure 3.2: Waveform propagation through a network with lumped circuit

$$\nu = \frac{\omega}{\beta} \quad m/s \quad (3.13)$$

$$T_D = \frac{\mathcal{L}}{\nu} \quad s \quad (3.14)$$

$$\lambda = \frac{2\pi}{\beta} \quad m \quad (3.15)$$

$$\lambda = \frac{\nu}{f} \quad m \quad (3.16)$$

$$\begin{aligned}
\Phi &= \beta \mathcal{L} \\
&= 2\pi \frac{\mathcal{L}}{\lambda} \quad \text{rad} \\
&= \frac{\mathcal{L}}{\lambda} \times 360 \quad \text{deg}
\end{aligned} \tag{3.17}$$

Where, ν is the phase velocity, T_D is the time delay the signal takes to travel from point a to point b, ω is the radiant frequency, λ is the signal wavelength, β is the phase constant, and Φ is the phase shift caused by propagation along the leads. It takes the waveform a distance of one full wavelength for the phase to shift by 360 deg. A practical rule of thumb is that for any distance smaller than $\frac{\lambda}{10}$, the traveling wave incurs a negligible phase shift and the distance is said to be electrically short. For any electric circuit whose size is smaller than $\frac{\lambda}{10}$, it can safely be assumed to be a lumped-circuit and its behavior can be analyzed with Kirchhoff's law.

3.2 Signal Integrity Impediments in the Design of High-speed Channels

The propagation characteristic of a transmission line is governed by equation 3.3 as previously stated. However, there are conditions for which the transmission line behaves very similar to ideal lossless case. The case of low-loss transmission line is well approximated by lossless transmission line in terms of impedance, speed of propagation and phase constant. We summarize the special cases of lossless, low-loss and the general case in Table 3.1.

Note that the equations for β , Z_0 , and phase velocity V_p are identical for lossless and low-loss cases.

More importantly, we can see that phase velocity is independent of w (hence f) which means that the propagation along the transmission line does not suffer from

the known dispersion spread that occurs for more lossy transmission line.

For the general case, speed of propagation is frequency dependent as seen in the case of lossy transmission line. This effect is the dispersion effect; and it is getting more severe as frequency increases. The dispersion simply means that low harmonics travel slower than the high harmonics. A digital pulse traveling into a dispersive medium would go through a spread and its pulse width widens at the receiving end.

Table 3.1: Propagation characteristics of common transmission lines

	Z_0	β	α	V_p
The lossy line	$\sqrt{\frac{R+j\omega L}{G+j\omega C}}$	$\text{Im}(\gamma)$	$\text{Re}(\gamma)$	$\frac{\omega}{\beta}$
The lossless line	$\sqrt{\frac{L}{C}}$	$\omega\sqrt{LC}$	0	$\frac{1}{\sqrt{LC}}$
The low-loss line	$\approx \sqrt{\frac{L}{C}}$	$\approx \omega\sqrt{LC}$	$\frac{1}{2}(\frac{R}{Z_0} + GZ_0)$	$\approx \frac{1}{\sqrt{LC}}$

3.2.1 Intersymbol Interference

Digital pulses in theory, have an infinite bandwidth and therefore cannot propagate in physical medium of finite bandwidth unless a certain degree of distortion can be tolerated. The combination of the intrinsically large bandwidth of digital signals and the limited bandwidth of the channels, like transmission line RC effects, results in a bit smearing into the subsequent bit or even into many subsequent bits. The sampling of the bit being transmitted might happen before the precedent bits have completely settled. This phenomena is called channel memory effects in communication. The channel "remembers" the previous bit. A widely used technical term for this phenomena is intersymbol interference (ISI) and is represented pictorially in the

Fig. 3.3.

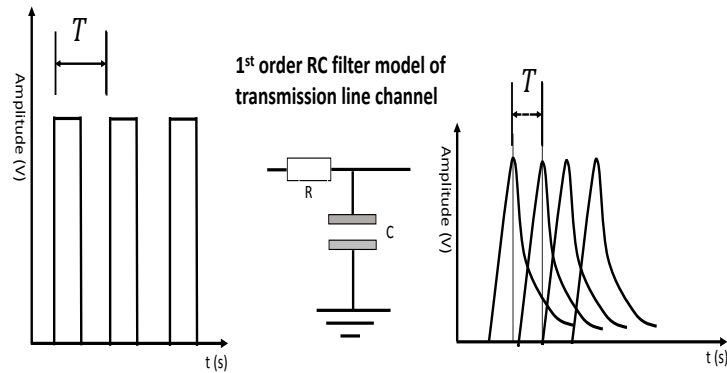


Figure 3.3: Bits smearing due to ISI in the interconnect channel[24].

ISI is aggravated with increasing speed. As unit interval gets smaller with increasing bit transfer rate, the pulse has little time to completely discharge or charge-up the line capacitance, which increases the chance of smearing. It is also known that ISI is very dependent on the pattern sequence. For example, if we compare a clock-like bit pattern '1010101010' to the '1111011111', we can see that in the latter pattern, the signal has enough time to fully charge up to the '1' logic voltage level after five consecutive ones, then quickly goes into a '0' before it goes back to one. The transition to zero would likely not have enough time to fully discharge before it transitions back into a '1'. That '1' will feel the reminiscent charge from the short 'zero'. For the case of the clock pattern, the logic '1' was short and most likely the signal did not rise all the way to the full voltage level, then it transits to '0' with a charge that most likely can be discharged in one UI before the next '1' happens. ISI is clearly a data-dependent class of noise.

3.2.2 Crosstalk

As computing systems are getting denser, designers need to cram more channels into an increasingly crowded real estate. Routing traces close to each other results in electromagnetic coupling between them. The coupling could be capacitive and/or inductive coupling. Crosstalk is the electromagnetic interaction between the nets of electric network.

In modeling the crosstalk, the aggressor net is the net carrying the signal while the victim is the quiet neighboring net feeling the coupling. Capacitive coupling occurs via mutual capacitance that appears between the aggressor and the victim nets. It results into the injection of a current onto the victim line proportional to the rate of change of voltage on the aggressor line. Therefore, the impact of mutual capacitance increases with the speed of the aggressor and gets more significant at higher frequency.

Mutual inductance injects current from the aggressor line into the quiet line by means of the magnetic field. That results in the injection of a voltage noise onto the victim that is proportional to the rate of change of the current on the driver line. Similar to the capacitive coupling, the impact of inductive coupling increases with frequency.

Fig. 3.4 (a) shows nets arranged in an aggressor net, where the signal is launched, and a victim net, where the amount of coupling caused by the aggressor is measured. Fig. 3.4 (b) is a model of the coupling[25]. Current change at the aggressor net causes voltage drop at the victim net via inductive coupling. Similarly, voltage change the voltage change at the aggressor net, generates current in the victim net via capacitive coupling. Crosstalk can be divided into far-end crosstalk (FEXT) and near-end crosstalk (NEXT) in reference to the terminals of the victim net with respect to the aggressor.

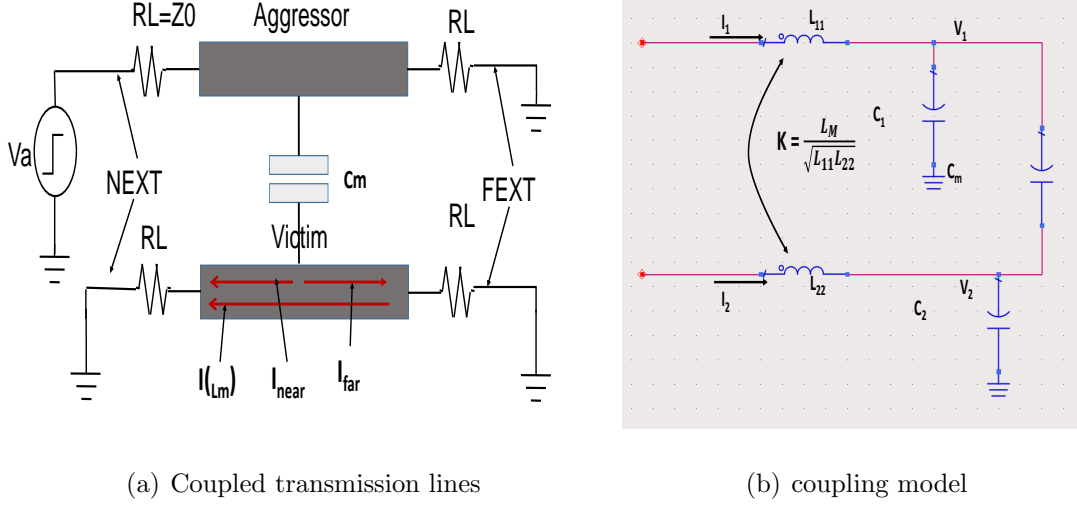


Figure 3.4: Coupled transmission lines (a) Layout, (b) Electrical model [25].

From the equivalent model of the two coupled lines, the current and voltage relation can be described as

$$I_{CM} = C_m \frac{dV_{agg}}{dt} \quad (3.18)$$

$$V_{LM} = L_m \frac{dI_{agg}}{dt} \quad (3.19)$$

$$I_{near-end} = I(LM) + I_{near-end}(C_M) \quad (3.20)$$

$$I_{far-end} = I_{far-end}(C_M) - I(LM) \quad (3.21)$$

where I_{CM} is the current flowing into the mutual capacitance, C_m is the mutual capacitance between the aggressor and the victim transmission lines, V_{LM} is the voltage across the mutual inductance L_m , $I_{near-end}$ is the near-end current flowing into the victim line, and $I_{far-end}$ is the far-end current flowing into the victim line.

It can be shown that the solution to the electromagnetic wave propagation in two coupled lines can be decomposed into two orthogonal modes, namely odd mode and even mode propagation. Each mode is analyzed separately, then we use linear

superposition techniques and recombine them for the final solution.

Fig. 3.5 shows the decomposition of the coupled transmission line pair into even and odd modes and their respective field distribution. In odd mode, the two lines are excited with opposite waveform polarities. The two lines are coupled to each others as well as to the return path. The field lines are distributed in a way dictated by capacitive and magnetic coupling as result of difference in potential between the two lines. A pictorial field line distribution is depicted in the Fig. 3.5 (a). The effective capacitance and inductance in the odd mode are $C_{11} + C_m$ and $L_{11} - L_{12}$ respectively.

In even mode, both lines are at the same potential, therefore no coupling current flows between the lines and field lines around them are at equipotential configuration. The effective inductance is $L_{11} + L_{12}$.

It can be demonstrated that the odd and even parameters and impedances are

$$C_{even} = C_{11} - C_m \quad (3.22)$$

$$L_{even} = L_{11} + L_{12} \quad (3.23)$$

$$C_{odd} = C_{11} + C_m \quad (3.24)$$

$$L_{odd} = L_{11} - L_{12} \quad (3.25)$$

$$Z_{even} = \sqrt{\frac{L_{even}}{C_{even}}} = \sqrt{\frac{L_{11} + L_{12}}{C_{11} - C_{12}}} \quad (3.26)$$

$$Z_{odd} = \sqrt{\frac{L_{odd}}{C_{odd}}} = \sqrt{\frac{L_{11} - L_{12}}{C_{11} + C_{12}}} \quad (3.27)$$

Since the modes have different inductance and capacitance parameters, their propagation delays are different and read

$$TD_{odd} = \sqrt{L_{odd}C_{odd}} = \sqrt{(L_{11} - L_{12})(C_{11} + C_m)} \quad (3.28)$$

$$TD_{even} = \sqrt{L_{even}C_{even}} = \sqrt{(L_{11} + L_{12})(C_{11} - C_m)} \quad (3.29)$$

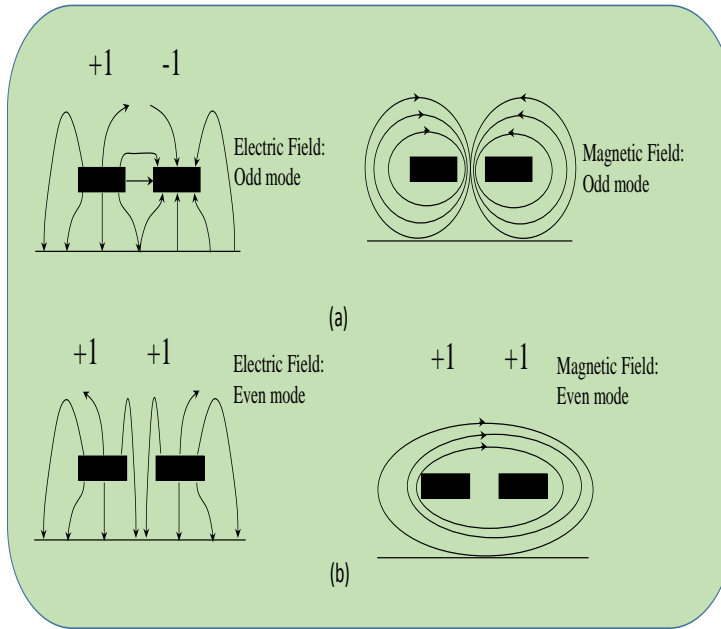


Figure 3.5: Odd and even fields distribution between two transmission lines [25].

where TD is the time delay for the signal to propagate down the transmission line. the speed of the two modes are equal if $\frac{L_{12}}{L_{11}} = \frac{C_m}{C_{11}}$. This condition is true in homogeneous transmission line, such as stripline. For microstrip lines, part of the field propagates in the air and part propagates inside the dielectric which causes odd and even mode waves to propagate at different speeds, and results into a modal induced crosstalk at the far end.

Practical analytic expression for FEXT and NEXT could be derived assuming that the lines are lossless, and properly terminated so that there is no reflection. In

that case, it is possible to develop simple crosstalk formulas to be[25]

$$NEXT = 2K_b td \frac{dV_a}{dt} \quad (3.30)$$

$$FEXT = K_f Length \frac{dV_a}{dt} \quad (3.31)$$

$$K_f = \frac{-1}{2} \left(\frac{L_{12}}{Z_0} - Z_0 |C_{12}| \right) \quad (3.32)$$

$$K_b = \frac{1}{4\sqrt{L_{11}C_{11}}} \left(\frac{L_{12}}{Z_0} + Z_0 |C_{12}| \right) \quad (3.33)$$

$$Z_0 = \sqrt{\frac{L_{11}}{C_{11}}} \quad (3.34)$$

Crosstalk can cause two types of failures: delay induced failure and logic failure. The delay induced failure happens when excessive crosstalk impacts the bottom line timing margin and causes violation of setup or hold time. This is increasingly likely to happen as memory unit interval shrinks with every generation of the protocol.

Logic failure happens when voltage and current induced noise causes an excessive glitch on the line which in the severe situation causes a false "1" or a false "0", and results in higher BER.

3.2.3 Skin Effects

As the frequency increases, the current density tends to get higher toward the surface of the conductor and decreases with greater depth in the conductor. For high-enough frequencies, the current flows mainly in a very shallow layer at the surface of the conductor. The skin depth is defined as the depth from the conductor surface to the point where the current density falls to $\frac{1}{e}$ of its value at the surface and is defined in equation 3.35. The skin effect causes the resistance to increase as function of the frequency according to the equation 3.36.

$$\delta = \sqrt{\frac{2\rho}{\omega\mu}} \quad (3.35)$$

$$R_{ac} \simeq \frac{\rho}{W\delta} = \sqrt{\frac{\rho\pi\mu f}{W}} \quad (3.36)$$

3.3 Bandwidth and Frequency Content of Digital Waveform

Digital signals are intrinsically high bandwidth signals. In fact an ideal digital signal with zero rise and fall time has an infinite bandwidth, and in theory, requires a channel with infinite bandwidth for a distortion-free transmission. In practice, a digital signal has finite rise and fall time and can be approximated by a trapezoidal waveform as shown in figure 3.6.

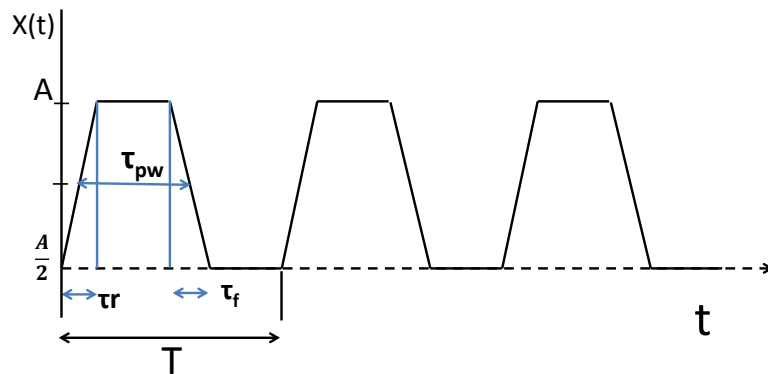


Figure 3.6: Trapezoidal waveform

To estimate the occupied bandwidth of a digital clock waveform $x(t)$, it is instructive to expand it into its Fourier series under the trapezoidal assumption. Let's assume that the rise and fall time are equal, i.e. $\tau_r = \tau_f = \tau$. This assumption greatly simplifies the mathematics while providing the same physical insights into the frequency properties of digital waveform as the general case. The Fourier series

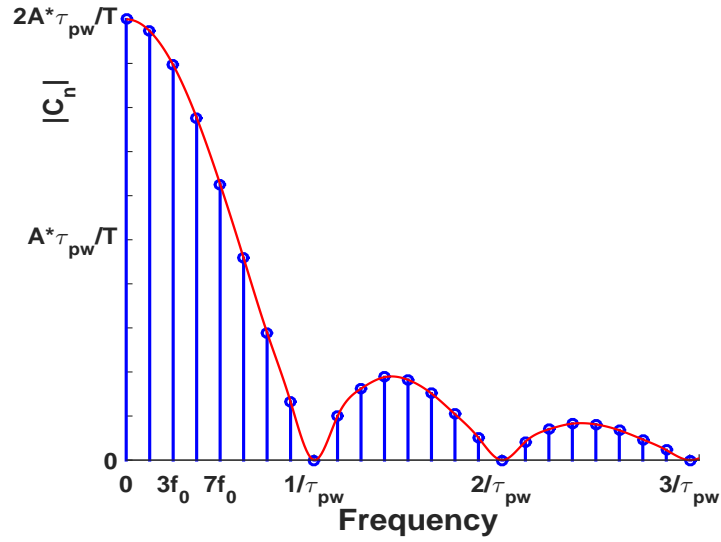


Figure 3.7: Plot of the magnitudes of Fourier coefficients of a square wave

expansion is given by

$$x(t) = c_0 + \sum_{n=1}^{\infty} c_n \cos(n\omega_0 t + \theta_n) \quad (3.37)$$

where $\omega_0 = 2\pi f_0 = \frac{2\pi}{T}$. c_0 is the DC component of the signal and is equal to the average value of the waveform over one period:

$$c_0 = \frac{1}{T} \int_0^T x(t) dt \quad (3.38)$$

In addition,

$$c_n = \frac{2}{T} \int_0^T x(t) e^{-j\omega_0 t} dt \quad (3.39)$$

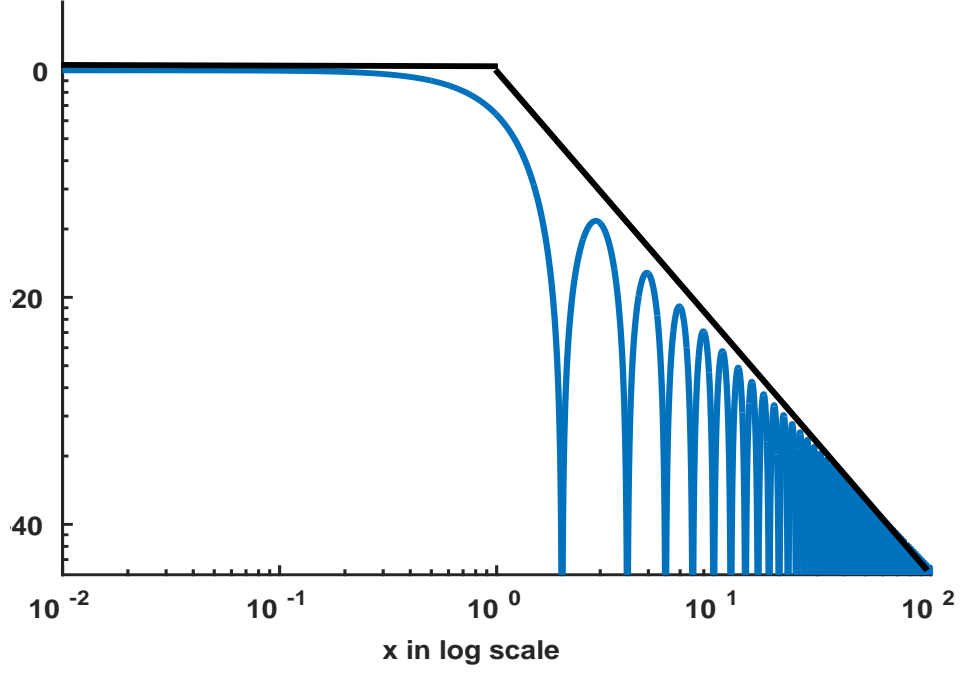


Figure 3.8: Plot of the envelope bounds of $\frac{\sinh(x)}{x}$ on logarithmic axes

For $\tau_r = \tau_f$,

$$c_0 = A \frac{t_{pw}}{T} \quad (3.40)$$

$$c_n = 2A \frac{t_{pw}}{T} \frac{\sin(\frac{n\pi t_{pw}}{T})}{\frac{n\pi t_{pw}}{T}} \frac{\sin(\frac{n\pi\tau}{T})}{\frac{n\pi\tau}{T}} \quad (3.41)$$

$$\theta_n = -n\pi \frac{(t_{pw} + \tau)}{T} \quad (3.42)$$

The Fourier coefficients of the trapezoidal pattern is a product of two sinc functions $\frac{\sin(x_1)}{x_1} \times \frac{\sin(x_2)}{x_2}$, one is a function of the pulse period and one is a function of the rise/fall

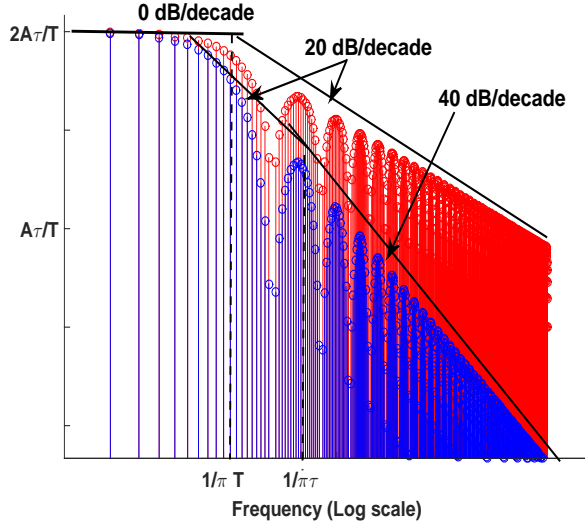


Figure 3.9: Plot of the magnitude and bounds of Fourier coefficients for both the trapezoidal and rectangular pulse of $\frac{\sinh(x)}{x}$ on logarithmic axes

time of the pulse τ .

For the ideal case of rectangular waveform, where rise time and fall time τ are equal to zero, the Fourier coefficients expression can be derived from the trapezoidal equation 3.38 - 3.42

$$c_0 = A \frac{t_{pw}}{T} \quad (3.43)$$

$$c_n = 2A \frac{t_{pw}}{T} \frac{\sin\left(\frac{n\pi t_{pw}}{T}\right)}{\frac{n\pi t_{pw}}{T}} \quad (3.44)$$

$$\angle \theta_n = -n\pi \frac{t_{pw}}{T} \quad (3.45)$$

The amplitude of the Fourier coefficients is captured in the Fig. 3.7. The envelope follows a $\frac{\sin(x)}{x}$ function and goes to zero when $\frac{\pi n}{T}$ becomes a multiple of π at $f =$

$\frac{1}{\tau}, \frac{2}{\tau}, \dots$

More insights can be gain if the $|C_n|$ are plotted in a log-log plot as shown in Fig. 3.8. We plotted the envelop and the bounds. The bounds show that the spectrum decreases at a slope of 20 dB per decade when $x \geq 1$ and has a flat slope at 0 dB per decade for $0 \leq x < 1$.

The trapezoidal and rectangular spectrum are plotted in the same graph shown in figure 3.9 along with asymptotic bounds for both curves. We notice that the bounds on the amplitude of the spectral coefficients start flat at 0 dB/decade from DC to the frequency $\frac{1}{\pi\tau_{pw}}$ for the trapezoidal case . The rectangular bounds then decrease at a rate of 20 dB/decade . On the other hand, the trapezoidal case decreases originally at a rate of 20 dB/decade up to $\frac{1}{\pi\tau}$, and then decreases at a rate of 40 dB/decade beyond that. The shorter the rise/fall time, the more high frequency spectrum content shows up in the pulse. The other finding is that the higher harmonics are attenuated more that lower harmonics. This analysis suggests that bandwidth of digital trapezoidal can be approximated by

$$BW \simeq \frac{1}{\tau} \tag{3.46}$$

3.4 Rectangular Waveguide

Waveguides are structures that are used to propagate electromagnetic waves. The most common waveguides are circular and rectangular waveguides (RWG). In our research, we used a structure widely known as substrate integrated waveguide (SIW). SIW are very similar to RWG as we will see in upcoming sections. The main properties of SIW can be easily derived from similar properties of RWG, hence the theory of operation of RWG is relevant to our work before delving into SIW theory of operation. We present here a summary of rules and equations that governs the operation of RWG,

and subsequently, we present the theory of operation of SIW as well as the difference between RWG and SIW.

A figure showing the geometry of RWG is shown in Fig. 3.10. To develop the equations that describe the behavior of RWG, the following properties are assumed:

- The waves propagating inside the guide are time-harmonic waves with $e^{j\omega t}$ dependence, where $\omega = 2\pi f$ and f is the frequency of propagation of the wave.
- Waves propagate along the z-axis
- The largest edge of the guide is along the x-axis so that $a > b$
- The waveguide region is source free.

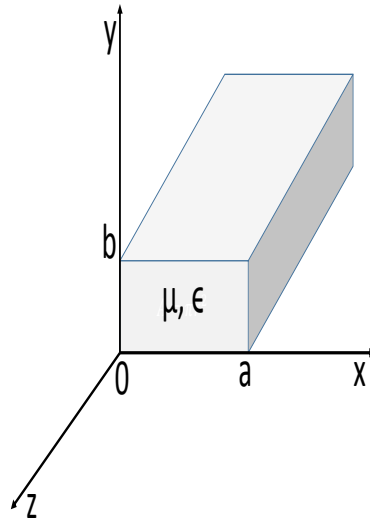


Figure 3.10: Geometry of a rectangular waveguide

The fields can then be written as[22]

$$\vec{E}(x, y, z) = [\vec{e}(x, y) + \hat{z}e_z(x, y)]e^{-j\beta z}. \quad (3.47)$$

$$\vec{H}(x, y, z) = [\vec{h}(x, y) + \hat{z}h_z(x, y)]e^{-j\beta z}. \quad (3.48)$$

where $\vec{e}(x, y)$ and $\vec{h}(x, y)$ are the transverse electric and magnetic field components, whereas $\hat{z}e_z(x, y)$ and $\hat{z}h_z(x, y)$ are the longitudinal electric and magnetic field components.

Since we assumed a source free case in the guide region, Maxwell equations read

$$\nabla_{\times} \vec{E} = -j\omega\mu\vec{H} \quad (3.49)$$

$$\nabla \times \vec{H} = j\omega\epsilon\vec{E}. \quad (3.50)$$

The above equations 3.47-3.49 apply equally to transverse electromagnetic (TEM), transverse electric (TE) and transverse magnetic (TM) waves. Since SIW can support only TE propagation mode as we will explain in the SIW section, we will restrict the RWG treatment to the TE mode. The objective is to develop the common theory that is shared between RWG and SIW to help establish the analogy and subtle difference caused by contiguous via rows on the wall side of the guide in SIW.

In TE modes, $E_z = 0$ and $H_z \neq 0$, the x, y components of the fields read[22]

$$H_x = \frac{-j\beta}{k_c^2} \frac{\partial H_z}{\partial x} \quad (3.51)$$

$$H_y = \frac{-j\beta}{k_c^2} \frac{\partial H_z}{\partial y} \quad (3.52)$$

$$E_x = \frac{-j\omega\mu}{k_c^2} \frac{\partial H_z}{\partial y} \quad (3.53)$$

$$E_y = \frac{j\omega\mu}{k_c^2} \frac{\partial H_z}{\partial x} \quad (3.54)$$

where $H_z(x, y, z) = h_z(x, y)e^{-j\beta z}$ is determined by solving Helmholtz equation which

reduces to the equation:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + k_c^2 \right) h_z(x, y) = 0 \quad (3.55)$$

where $k_c^2 = k^2 - \beta^2$ is the cutoff wavenumber. Solutions to equation 3.55 can be written as

$$E_x = \frac{jw\mu n\pi}{k_c^2 b} A_{mn} \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} e^{-j\beta z} \quad (3.56)$$

$$E_y = \frac{-jw\mu m\pi}{k_c^2 a} A_{mn} \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b} e^{-j\beta z} \quad (3.57)$$

$$H_x = \frac{j\beta m\pi}{k_c^2 a} A_{mn} \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b} e^{-j\beta z} \quad (3.58)$$

$$H_y = \frac{j\beta n\pi}{k_c^2 b} A_{mn} \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} e^{-j\beta z} \quad (3.59)$$

where m, n are mode indices and can take on non-negative integer values with the exception of m=n=0. The propagation constant is

$$\beta = \sqrt{k^2 - k_c^2} = \sqrt{k^2 - \left(\frac{m\pi}{a} \right)^2 - \left(\frac{n\pi}{b} \right)^2} \quad (3.60)$$

For the wave to propagate, β needs to be real, therefore

$$k > k_c = \sqrt{\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2} \quad (3.61)$$

The mode cutoff frequency f_{mn} is given by

$$f_{cmn} = \frac{k_c}{2\pi\sqrt{\mu\epsilon}} = \frac{1}{2\pi\sqrt{\mu\epsilon}} \sqrt{\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2} \quad (3.62)$$

$$f_{c10} = \frac{1}{2a\sqrt{\mu\epsilon}} \quad (3.63)$$

3.5 Substrate Integrated Waveguide

Substrate integrated waveguide (SIW) is a relatively new interconnect technology introduced in the early 1990's[26]. It consists of two parallel metallic conductor plates linked with two rows of metalized posts as shown in Fig. 3.11.

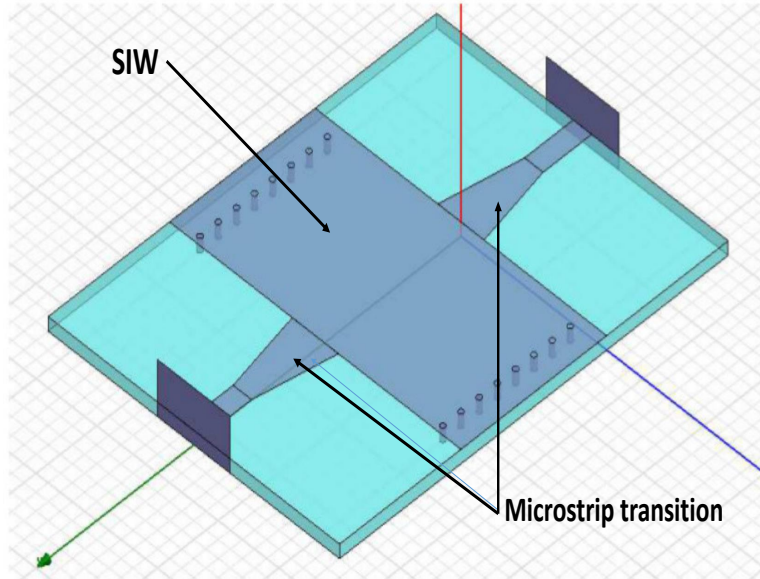


Figure 3.11: Substrate integrated waveguide geometry and microstrip transitions captured from Ansoft HFSS.

It has been shown that with careful optimization of via post diameters and via pitch, leakage of energy through side wall can be made negligible and SIW performance can be made comparable to classic rectangular wave guides (RWG)[27]. In contrast to RWG, SIW has the advantage of being compatible with printed circuit board (PCB) technology.

SIW interconnects have been used in microwave devices such as circulators, filters, power dividers, resonators, antennas, phase shifters and more. A plethora of filters with diverse topologies have been reported [28–37].

A variety of couplers have been successfully designed using SIW [38–43]. SIW have been also deployed as a component in a system such as in six-port-based circuits[44–48]. For digital application, SIW have been proposed in many works that we will cover in details in section 4.1.3.

3.5.1 Connecting SIW to Planar Circuits

SIW has demonstrated relatively smooth integration with traditional planar circuits. Many techniques and schemes connecting the SIW to planar circuits have been described in the literature. A microstrip-to-waveguide transition using an optimized taper was demonstrated in [13, 49–52], and proved to be wideband and practical. Coupling of two SIWs through an aperture was also demonstrated [53]. We chose the taper transition from microstrip-to-SIW for its simplicity and efficiency. The taper transition details are shown in the Fig. 3.12

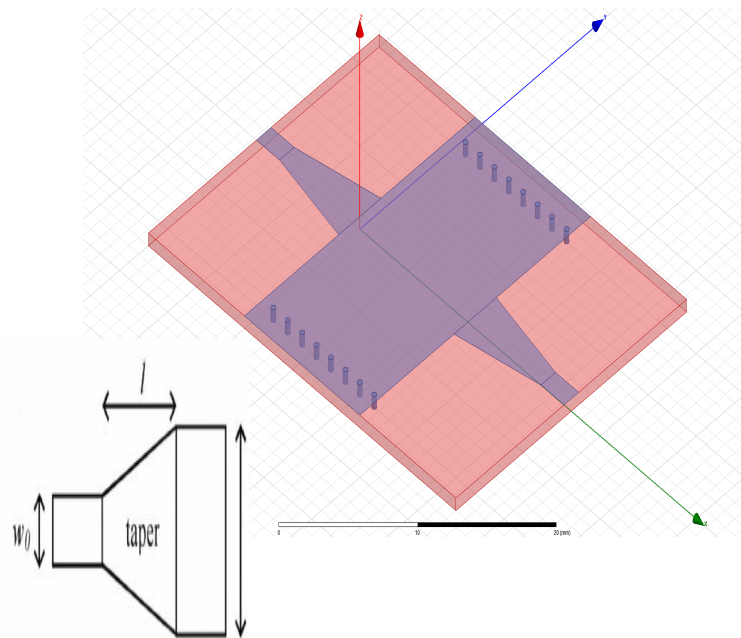


Figure 3.12: Microstrip to SIW transition using taper.

A more formal and rigorous treatment of tapered lines can be found in [54]. The transition between different sections of microwave interconnecting transmission lines is modeled as multisection matching transformers. Using the theory of small reflections,

the reflection coefficient of a taper can be approximated as

$$\Gamma(\theta) = \frac{1}{2} \int_{x=0}^L e^{-2j\beta x} \frac{d}{dx} \ln \left(\frac{Z}{Z_0} \right) dx \quad (3.64)$$

where L is the taper length, Z is the impedance of the taper, and $\theta = 2\beta L$.

By optimizing $Z(x)$, the reflection coefficient can be minimized. The general case of an arbitrary $Z(x)$ taper shape is difficult to model, and detailed analysis can be found in [55]. However, two practical tapers can approximate transition between a transmission line and a waveguide, namely the exponential taper and the triangular taper. For the triangular taper, the reflection coefficient is

$$\Gamma(\theta) = \frac{1}{2} e^{-j\beta L} \ln \left(\frac{Z_L}{Z_0} \right) \left[\frac{\sin(\beta L/2)}{\beta L/2} \right]^2 \quad (3.65)$$

where Z_0 is the impedance at $x=0$, and Z_L is the load impedance.

The reflection coefficient for the exponential taper is

$$\Gamma(\theta) = \frac{\ln \frac{Z_L}{Z_0}}{2} e^{-j\beta L} \frac{\sin(\beta L)}{\beta L} \quad (3.66)$$

In both cases, we see that the reflection coefficient is a function of βL , which led us to perform a geometry optimization for the design of our circuit taper. The simulated reflection coefficient results for taper optimization can be found in Fig. 3.13

3.5.2 Substrate Integrated Waveguide Design Equations

Propagation characteristics of SIW structures are similar to those of rectangular metallic waveguides, provided that the metallic vias dimensions and spacing are judiciously designed such that the sidewall radiation leakage is negligible. The transverse electric modes of SIW coincide with TE modes of rectangular waveguides, and the propagation constants have similar expressions. The gaps between the via posts create conductor discontinuity in the z -direction (wave propagation direction), and the

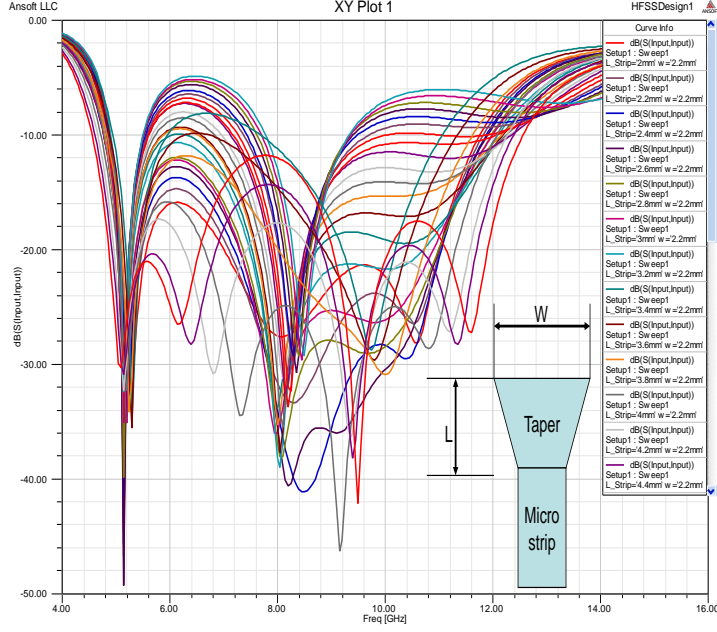


Figure 3.13: Optimization of taper length L.

current cannot flow in the z-direction. As a result, the SIW does not support TM mode propagation.

Due to the similarity between SIW and rectangular waveguide, a set of design rules and constraints have been developed for dimensioning SIW components perturbation from their rectangular waveguide counterpart. Empirical formulas for mapping RWG into SIW had been proposed in [27] [56] and refined in [57] and are reproduced here

$$W_{SIW} = W_{eff} = W_{RWG} - 1.08 \frac{d^2}{5} \quad (3.67)$$

where W_{SIW} is the width of the SIW; W_{RWG} is the equivalent rectangular waveguide used as reference to generate the SIW; and d is the diameter of the SIW wall vias.

The SIW cutoff frequency reads

$$f_c = \frac{c}{2\sqrt{\epsilon_r}} \left(W_{eff} - \frac{d^2}{0.95s} \right)^{-1} \quad (3.68)$$

provided that

$$s < \frac{\lambda_0}{2\sqrt{\epsilon_r}} \quad (3.69)$$

and

$$s < 4d \quad [58] \quad (3.70)$$

where c is the speed of light in the vacuum, ϵ_r is the dielectric permittivity, s is the center-to-center separation between two adjacent vias, and λ_0 is the wavelength in free space.

These empirical design formulas constitute a good starting point for the design of the desired SIW. Then, using a fullwave field solver the design is fine tuned and optimized.

3.5.3 Physics of SIW

As explained earlier, SIW does not support the propagation of TM modes. For TE modes, SIW circuit is described by the same physical equations and constraints as the metallic waveguide. TE_{10} is the dominant mode in SIW. We designed and simulated an X-band SIW. The magnitude of the dominant mode along the guide is shown in the Fig. 3.14.

The loss mechanisms in SIW

Three loss mechanisms occur in SIW: metal conductivity loss, dielectric loss and radiation loss. The radiation loss is a characteristic of the SIW whereas conductor and dielectric loss occur in both the SIW and the RWG.

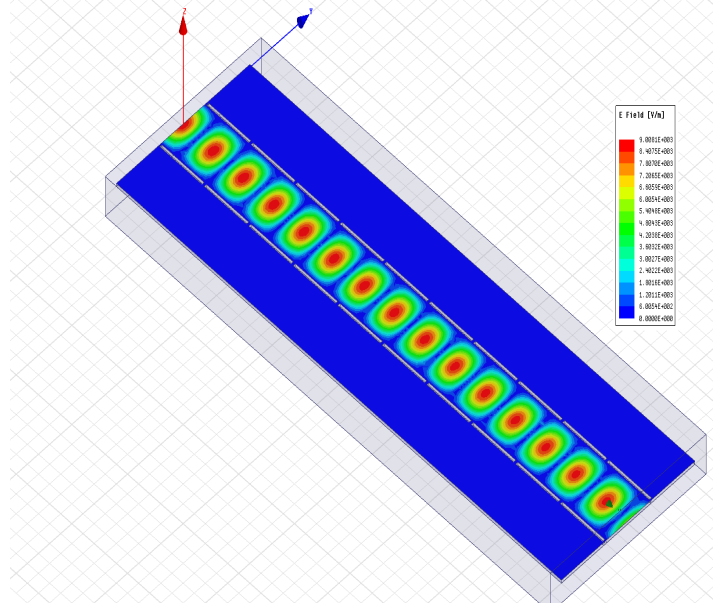


Figure 3.14: Magnitude of the dominant TE_{10} electric field in SIW at 9.0 GHz.

The conductor loss is caused by the finite conductivity of the metal planes and the vias. It is given by:

$$\alpha_c(f) = \frac{\sqrt{\pi f \epsilon_0 \epsilon_r}}{h \sqrt{\sigma_c}} \frac{1 + 2\left(\frac{f_0}{f}\right)^2 \frac{h}{W_{eff}}}{\sqrt{1 - (f_0/f)^2}} \quad (3.71)$$

where h is the substrate thickness, ϵ_0 is the dielectric permittivity of vacuum, ϵ_r is the relative dielectric permittivity of the substrate, f_0 is the cutoff frequency of the SIW, f is frequency of the propagating wave, and σ_c is the metal conductivity.

The dielectric loss is related to the loss tangent $\tan \delta$ of the dielectric substrate. The SIW has the same loss mechanism as the RWG, and the formula is given by:

$$\alpha_d = \frac{\pi f \sqrt{\epsilon_r}}{c \sqrt{1 - (f/f_0)^2}} \tan \delta \quad (3.72)$$

where c denotes the speed of light in the vacuum.

In the SIW, in contrast to the RWG, the side walls have gaps between the vias that comprise the side walls. This gap allows electromagnetic energy to leak in the form

of radiation, a phenomena that does not exist in conventional waveguides. Authors of [59] developed an empirical formula given by:

$$\alpha_R = \frac{\frac{1}{w} \left(\frac{d}{w}\right)^{2.84}}{4.85 \sqrt{\left[\frac{2w}{\lambda}\right]^2 - 1}} \left[\frac{dB}{m} \right] \quad (3.73)$$

where λ is the wavelength of the guided wave.

Understanding the impact of geometrical parameters on SIW performance is important. These parameters can have non negligible effects on the quality of the design.

The dielectric thickness, h is directly related to conductive loss of the SIW. As h increases, the metal surface of the SIW increases and current density decreases which decreases the conductive loss. The current density $|J|$ is proportional to \sqrt{h} while $|J|^2$ is proportional to h . Since conductor loss is inversely proportional to $|J|^2$, it varies proportional to $\frac{1}{h}$.

The diameter d of the via has very little effect on the conductor loss, since most of the current flows on the top and bottom metal layers.

Conductor loss is directly proportional to the separation parameter s between two consecutive via posts. Shrinking the spacing between two vias results in augmenting the overall copper surface used by the current. Hence, the current density is reduced and conductor loss is reduced accordingly.

Radiation loss is also proportional to the separation parameter s . As separation increases, the electromagnetic field confinement to the guide decreases and more fields are radiated from the slots between the posts. The Fig. 3.15 capture the simulation results for conduction, radiation and total loss of the SIW, as a function of frequency.

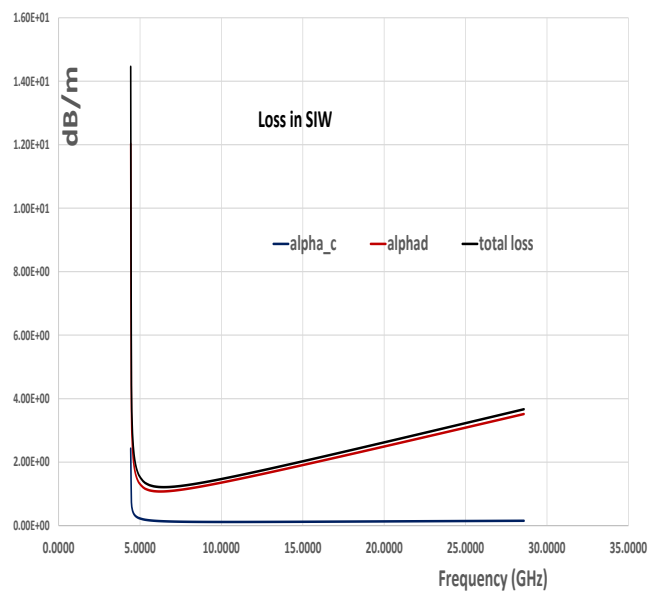


Figure 3.15: Conduction, dielectric and total loss of an X-band SIW.

Chapter 4

LITERATURE REVIEW

The past decade has seen enormous efforts to resolve memory wall bottleneck. A representative sample can be found in [6, 60–71].

Approaches to memory wall issue are predominantly a combination of incremental architecture improvement with circuit techniques that mitigate crosstalk, ISI and SI noise. The work [72] is a representative sample that typifies the proposals attempting to alleviate memory bottleneck. This is done mainly through using advanced noise cancellation techniques, preemphasis and equalization to mitigate SI obstacles that we explained in section 3.2.

A group of researchers adopted the use of differential signaling to replace the conventional single-ended multidrop signaling that dominates DDR interface[62, 63, 73] while other researchers chose to switch to optical interconnect[74–76].

As far as we are aware, we are the only researchers to use SIW for DDR interconnect. However, there are a number of works that uses SIW for high-speed signaling[77–83].

In the following section, we present a concise summary of representative works. We detail their contributions, highlight their achievements, and identify the opportunities to improve upon the art.

4.1 Work on Memory Architecture and Throughput

4.1.1 Electrical Solution Proposals

Among the many proposals introduced in the last decade, fully buffered DIMM (FB-DIMM) is probably the most hyped one. At the initiation of the standard, it was presented as the ultimate solution to the memory shortcomings. What is interesting about FB-DIMM from a research perspective is that it is an eloquent illustration of the challenges, difficulties, and conflicting constraints that the memory standard can pose. The standard was adopted by JEDEC in 2004-2005 time frame. The latest version on JEDEC site is JESD206 2007[21] revision. The standard aims to resolve two major obstacles faced by the standard DDR interface, namely:

- Capacity: The number of DIMM per channel is reduced with every higher speed generation. DDR2, allows up to 2 DIMMs per channel. DDR3 and DDR4 are limited to 1 DIMM per channel. One of the FB-DIMM standard goal is to enable a scalable architecture that allows user to scale-up the number of DIMMs to meet their system needs.
- Transfer rate: The per-pin transfer rate quickly leveled off due to multidrop signaling and the excessive SI noise in transmission line with increasing speed. The fly-by architecture alleviates some of the SI obstacles of multidrop topology, but it is still limited by transmission line attenuation and by the stubs that branch out from the transmission line into SDRAM devices. FB-DIMM objective is to solve the per-pin speed disparity between DDR and serial IO.

In order to resolve these fundamental architecture limitations, FB-DIMM committee adopted an approach where the DDR signals on the DIMM are connected to

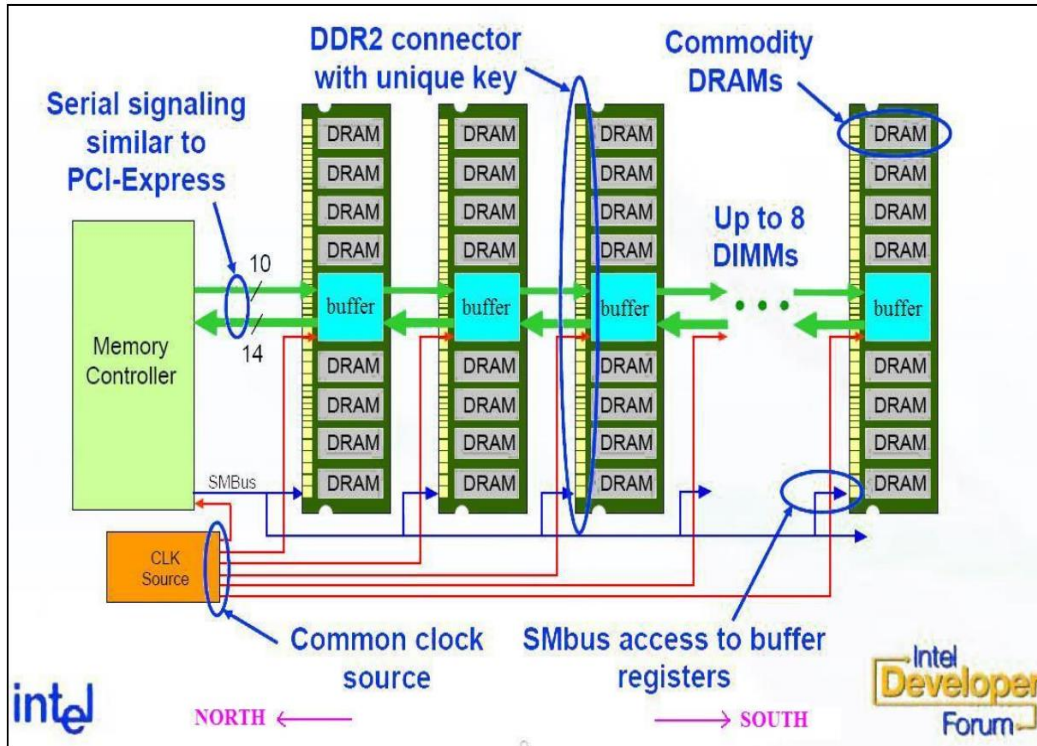


Figure 4.1: Fully buffered DIMM architecture [84].

the controller through a buffer on the DIMM, and the DIMMs are interconnected differentially as illustrated in Fig. 4.1.

The details of FB-DIMM operation is explained below:

- Each DIMM comprises a logic control block called advanced memory buffer (AMB) responsible for interpreting the packetized protocol and controlling the DRAM devices on each DIMM .
- Unlike the parallel bus architecture of traditional DDR, FB-DIMM has a point-to-point serial interface between the memory controller and the AMB. This results in increase in the memory bandwidth. The controller does not write directly to the memory, but through the AMB. The communication is enabled by a narrow daisy chained point-to-point serial protocol. Fig. 4.2 shows a

functional diagram AMB and surrounding logic.

- The AMB performs compensation for signal attenuation by buffering, and re-sending the signal. The advanced memory buffer includes two bidirectional link interfaces using high-speed differential point-to-point electrical signaling.
- The southbound input link is 10 lanes wide and carries commands and write data from the host memory controller or the adjacent DIMM in the host direction. The southbound output link forwards this same data to the next FB-DIMM.
- The northbound input link is 13 to 14 lanes wide and carries read return data or status information from the next FB-DIMM in the chain back towards the host. The northbound output link forwards this information back towards the host and multiplexes-in any read return data or status information that is generated internally.

At first glance, this standard promises to solve both capacity and speed bottleneck of DDR interfaces with high fidelity signals. Indeed, FB-DIMM [84] succeeded in considerably exceeding the speed and density of mainstream DDR3/4. Unfortunately, the complex logic of AMB, the frequent need for the costly operation of regenerating and compensating signals, resulted in the generation of excessive level of heat and unacceptable power dissipation. A 32 GB FB-DIMM channel using fully populated 8 DIMMs consumes up to 90 watts. Work in [65] showed that FB-DIMM consumes over 800% more power than a comparable DDR memory. The standard ultimately failed and was quickly abandoned.

Work reported in [63] and in [71] proposes a technique where the transmission of both the baseband and a non coherent ASK modulated RF signal is done using the same physical channel. An on-chip transformer couples the signals between memory

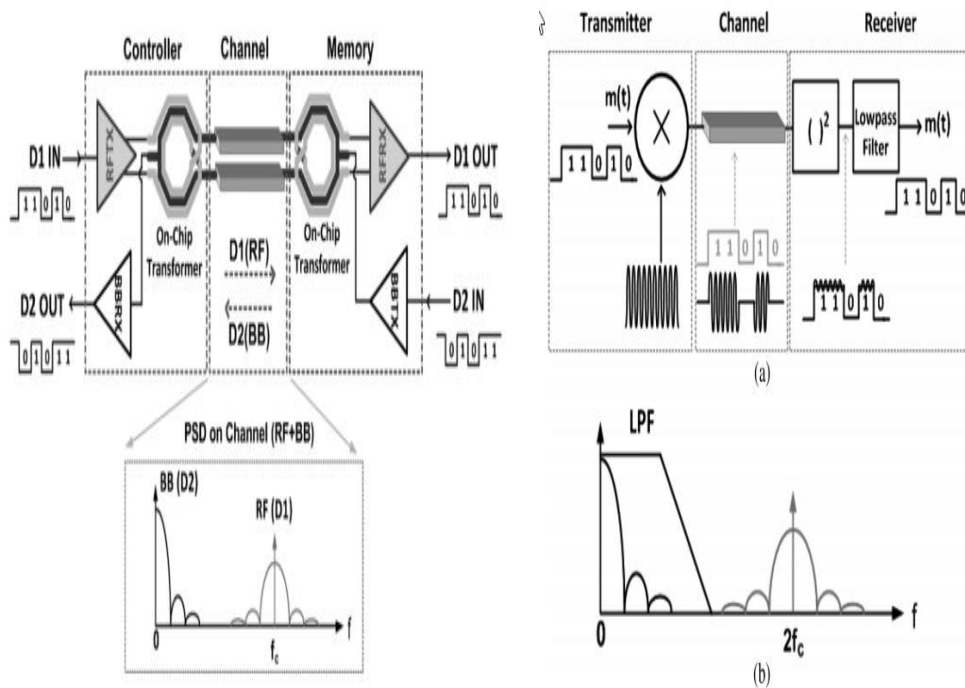


Figure 4.3: Noncoherent ASK memory interconnect [63].

itself to generate the original baseband. A filter at the receiver removes the RF result of the multiplication. Since the BB and the RF are far apart on the spectral scale, they can travel simultaneously on the channel without coupling. A block diagram of the proposal is shown in Fig.4.3.

The authors choose ASK for ease of implementation and to avoid the bit to symbol conversion. This choice negatively impacted the throughput since ASK is only capable of carrying one bit per symbol. The scheme enables carrying half of the data onto normal baseband signaling and the other half onto ASK RF carrier. This arrangement helps reduce the total physical signals to half that of conventional memory bus. This comes with the penalty of considerable complexity of the transmitter and receiver, the

coupling transformers and the additional logic to handle synchronization and signal recovery. An obvious shortcoming of the work is that it uses only one RF band out of many possible bands, resulting in loss of bandwidth and missing opportunity to maximize throughput. The authors acknowledged this weakness in the most recent version of the work[71].

One other issue is that signals are not identical at the receivers. The data that comes from the RF side suffers from common mode noise and RF impairment. The data transmitted on baseband suffers from the usual transmission line attenuation and signal integrity impediments. This incoherence might result into different kind of bit errors that confuses the receiver. The main shortcoming of this approach is that it remains fundamentally identical to conventional memory bus when it comes to maximum data rate possible. The new bus is limited by the transmission line lowpass transfer function and hence is limited to saturate at 4-5 Gbps.

Work in [62] uses low-swing voltage-mode differential signaling with forwarded clock instead of classical PLL/DLL clock recovery. Data link is a differential bidirectional point-to-point channel. The controller is augmented with transmit and receive phase mixers responsible for skew compensation. The skew is inherently caused by high-speed differential signaling. The phase mixers operate from three fully differential clocks spaced at 60° intervals and generate six phase chunks spanning one unit interval. A fourth differential clock is also needed for logic operation. The proposal achieves a data rate of 4.3 GB/s.

The use of differential signaling and complex mixers, phase interpolator and logic to de-skew signals came at a high per-bit power cost. In order to keep the power under control, the authors needed to restrict the bus to 8 bits-wide interface and to reduce the number of control and command signals. In addition, the authors implemented very complex logic with multiple power state levels that are automatically enabled or

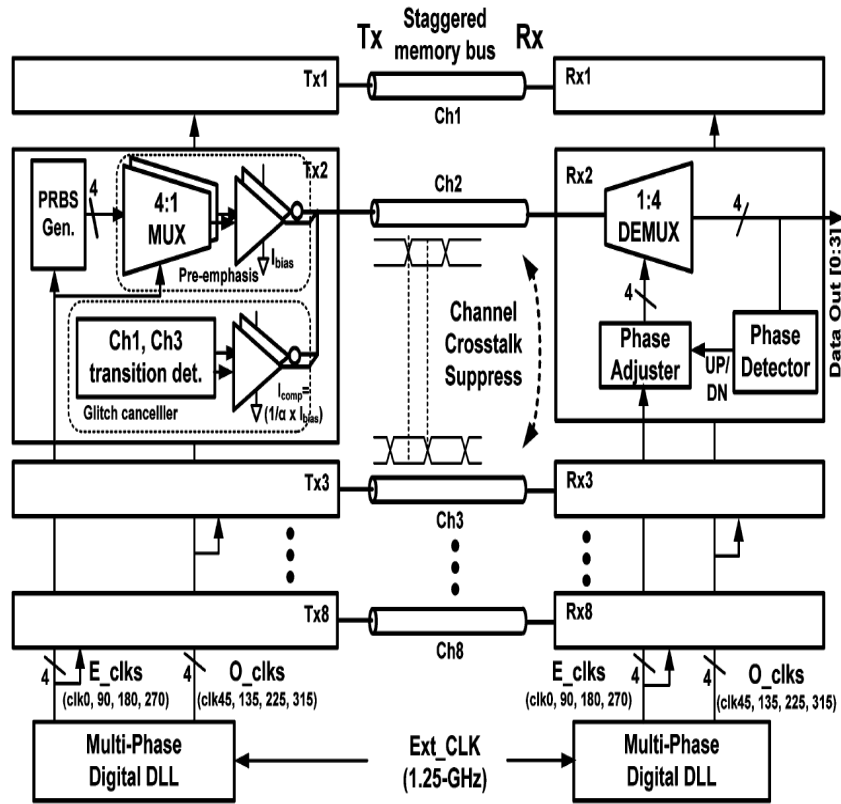


Figure 4.4: Memory transceiver with crosstalk suppression scheme [60].

disabled to reduce power consumption.

Work reported in [60] is illustrated in Fig. 4.4. At the transmitter, data goes through a preemphasis circuit and is then serialized via a 4:1 multiplexing stream. The preemphasis step is used to reduce high speed ISI noise while the serializer boosts the speed. The serial interface runs at 5 Gb/s after muxing 4 parallel inputs at a speed of 1.25 Gb/s each. At the receiver side, the reverse operation happens. A serial-to-parallel operation is performed by a 1:4 DEMUX block.

To reduce the jitter, a multi-phase DLL generates 8 evenly-spaced clocks. The DLL selects the output clock edge which is closest to the reference clock edge, and

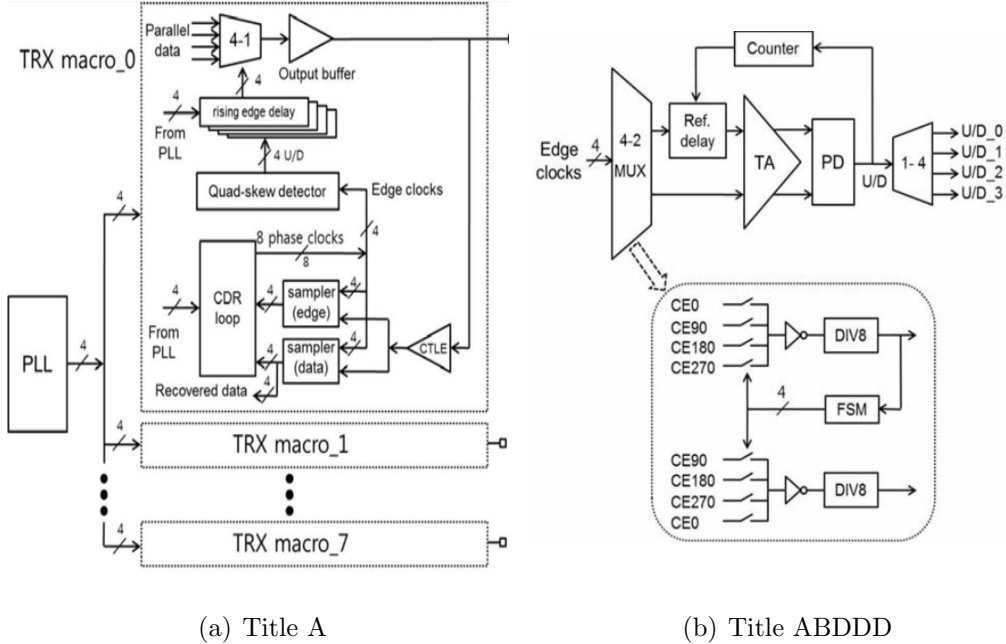


Figure 4.5: DDR transceiver with skew cancellation [86] (a): Transceiver architecture (b): Skew detector circuit

hence reduces the jitter. A phase detector and a phase adjuster at the receiving end align the sampling clock to the middle of the data bit.

The different channels are laid out in a staggered fashion and use clocks derived from independent clock domains. By controlling the relative phase offset between different channel clock domains, the crosstalk is reduced substantially. Since the different channels use clocks originating from different domains, the staggering of traces from two different channels acts as mutual shield and a return path for signals. The configuration behaves in a similar manner as if it was designed to be a ground-signal-ground-signal configuration and effectively double the separation between signals from the same channel.

As explained earlier, the aggressor channels toggle with a controlled phase shift that lands at the middle of victim channel unit interval. Therefore a glitch appears at

the middle of the victim eye pattern. The interface is equipped with a glitch canceling unit which consists of a transition detector and a sub-driver . The measurement of the eye diagram with and without crosstalk canceling mechanism and staggering topology shows an improvement of about 15 mV in eye opening and a reduction of more than 20 ps in total jitter compared to conventional memory bus.

Another work whose block diagram is shown in Fig. 4.5 is reported in[86]. The work in [60] and the work in [86] use very similar approaches. They achieve improvements in memory controller and augment them with ISI and noise cancellation capabilities. However, improvement is marginal due to excessive power penalty inherent with equalization and noise cancellation circuitry.

4.1.2 Optical Interconnect Memory Proposals

In the work reported in [75], the system is composed of a microprocessor implemented on a field programmable gate array (FPGA). The microprocessor communicates with three optically-connected memory modules via a wavelength-stripped 4 x 4 hybrid packet and circuit-switched optical network. Eight 10-Gb/s transceivers and eight wavelength channels are modulated and then combined using wavelength-division multiplexing (WDM) to generate a 240 Gb/s aggregate memory bandwidth (80 Gb/s per channel times 3 channels). Fig. 4.6 shows a diagram of the proposed optically connected memory module.

A Xilinx Virtex-V FPGA controls output connections and process the appropriate packet or circuit-switching routing information.

The circuit-switching protocols uses an electronic control plane while all packet switching is performed optically through the use of dedicated header wavelengths.

The experimental setup is shown in Fig. 4.7. The authors report error free memory traffic of 10 Gb/s on each port with clean eye diagrams. 24% of the traffic

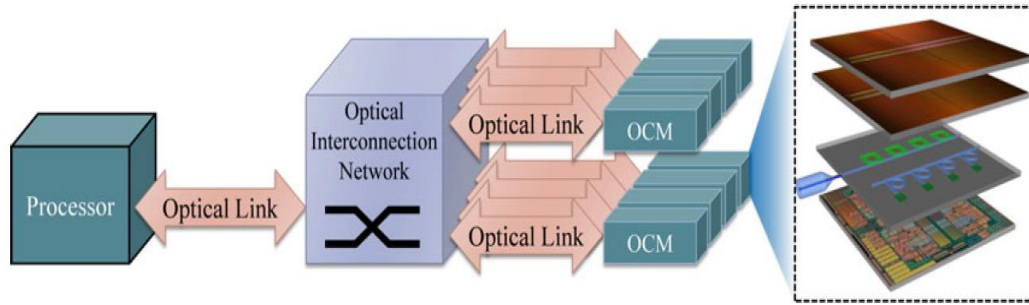


Figure 4.6: Optically connected memory module[75]

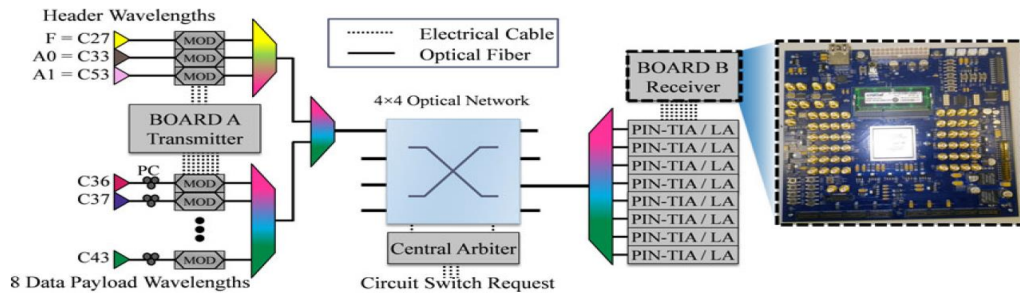


Figure 4.7: experimental setup[75]

uses circuit-switching mode rather than packet switching mode. This proves that the ability to select the mode based on the message length has a positive impact on reducing memory latency.

The proposal requires assembly of a set of bulky and complex optical and electrical circuits in order to operate adequately. Below is a partial list of the circuits required for operation:

- An optical network which consists of 4x4 or 2x2 switching fabric nodes using a set of semiconductor optical amplifier nodes
- Many optical couplers
- Many optical splitters

- A WDM multiplexer
- A set of optical transceivers

The solution exhibits the classical limitations of optical interconnect. It is too bulky, non scalable, and requires a complex aggregation of circuits to work. A big part of the circuits are optical devices known for incompatibility with planar technologies. Hence, it is very difficult to integrate them with electrical circuits in a compact manner. However, it remains a good solution if deployed to a large system such as a commercial networking hub or data center where the system requires a long-haul interconnect. For small system and mainstream computing platform, the solution is not practical.

In the work[74], the authors deploy an optical all passive solution for row address selection (RAS) and columns address selection (CAS) logic in memory matrix. The logic implements the selection using a WDM scheme.

The optical SRAM unit uses a 2-bit-long wavelength division multiplexing (WDM) WL address and four wavelengths. The optical word is broadcasted into the the four access gate modules (AG). Each RAS row output signal is injected into the respective SRAM row AG that controls the access to the incoming WDM optical word.

The full logic truth table is shown at the bottom of the Fig.4.8. The proposal uses a wavelength filtering matrix, an array waveguide grating that forms the CAS unit, a cross-phase modulation semiconductor optical amplifier Mach-Zendher interferometer modulator used as access gate, an optical cache peripheral circuitry using silicon-on-insulator ring resonators. The proposal achieved 10 Gb/s for a 16 x 4 optical SRAM bank. We briefly summarized the architecture of the row and column address selection unit implemented by the authors, however it is worth mentioning that the circuit is only for row and column address. If we add to that the optical-electrical required interface between the CPU and the memory bank, one can gauge the complexity of

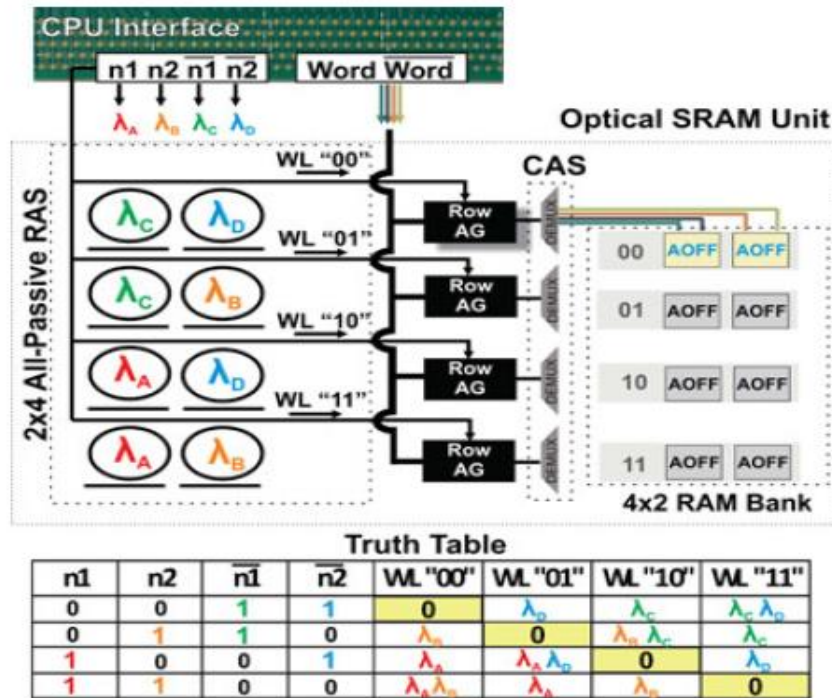


Figure 4.8: 4 x 2 Optical SRAM bank architecture[74]

the overall system.

The authors could not implement part of the circuits using the mainstream CMOS process and they needed to revert to a SOI process instead. This demonstrates again the integration difficulties typically associated with all the optical interconnect initiatives.

The proposal is interesting conceptually, however, it is very complex and bulky as a result of difficulty in making optical and optical-to-electronic conversion circuitry compatible with planar process.

4.1.3 High-speed Interconnect Using Substrate Integrated Waveguide

As mentioned earlier, there is no proposal in the prior art literature that uses SIW for memory interconnect. Therefore, this section is a survey of the most important literature describing the use of SIW as a high-speed digital interconnect.

SIW was originally proposed by [26] in 1994. It is used mainly to replace rectangular waveguides in the objective of miniaturization of RF and microwave circuits.

SIW application as an interconnect had been report as early as 2006[77] where the bandwidth of SIW is characterized. The work also studied the fundamental phenomena of crosstalk and coupling between adjacent SIWs. Effects of 45° and 90° bends were also characterized and reported in that work.

A 7.6 cm-long SIW operating at a center frequency of 50 GHz where the input and output are coupled via two coaxial lines terminated with short vertical probes is described in [77]. The authors also fabricated two extra 12.7 and 25.4 cm-long SIWs used to extract the loss per cm in the SIW. At midband, the measured loss is of 0.31 dB/cm.

Also in [77], a set of two SIWs that share one via row is fabricated and measured to characterize the crosstalk between two adjacent SIWs. The measured crosstalk is 0.2% when the two SIWs are transmitting in the same direction. When the transmitting and receiving ports of the two SIWs are at opposite ends, the crosstalk is 1%.

The authors demonstrated that if the arc in 45° and 90° bends is designed carefully, the S-parameters of the SIW with bends matches the s-parameters of the straight SIW. The straight SIW, the SIW with 45° bends, and the SIW with 90° bends achieve crosstalk performance better than -30 dB over the entire bandwidth.

In the work [79], the authors propose to use SIW as high-speed interconnect. A



Figure 4.9: 7.6-cm SIW interconnect test structure[77]

transmission line is embedded inside the SIW to carry signals from DC all the way to the transmission line bandwidth (about 6.6 GHz for the stripline). The fabricated circuit consists of a 48 mm long SIW that embeds an equally long 50- Ω microstrip. The substrate is a Rogers 4003C material, 0.76 mm thick and with a guide width of 5.8 mm. The SIW transmits the high-frequency signals (from 13.9 GHz to 29 GHz) using an up-converter mixer and an amplifier at the transmit end and a down-converter mixer and filter at the receiving end. Because the interconnect supports both DC and high-frequency, the authors call it hybrid. The SIW achieves 15 GHz bandwidth with insertion loss of 2.65 dB or less. Isolation between the transmission line and the SIW is measured to be about 20 dB. A pulse pattern generator generates 10 Gb/s NRZ PRBS pulse stream for both the SIW and the transmission line. The aggregated throughput is 20 Gb/s.

It is not clear though from the paper how the transmission line achieves a 10 Gb/s transfer rate while the authors stated that the bandwidth is of 6.6 GHz. Our guess is that the microstrip performs well even beyond the 3 dB attenuation point due to

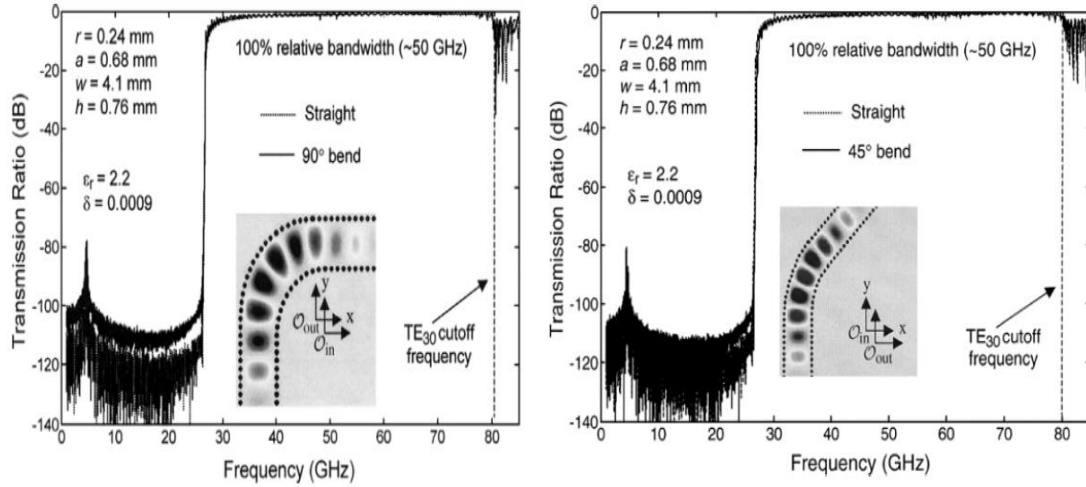


Figure 4.10: S_{21} characteristics of 45° and 90° bend SIW interconnect test structure[77]

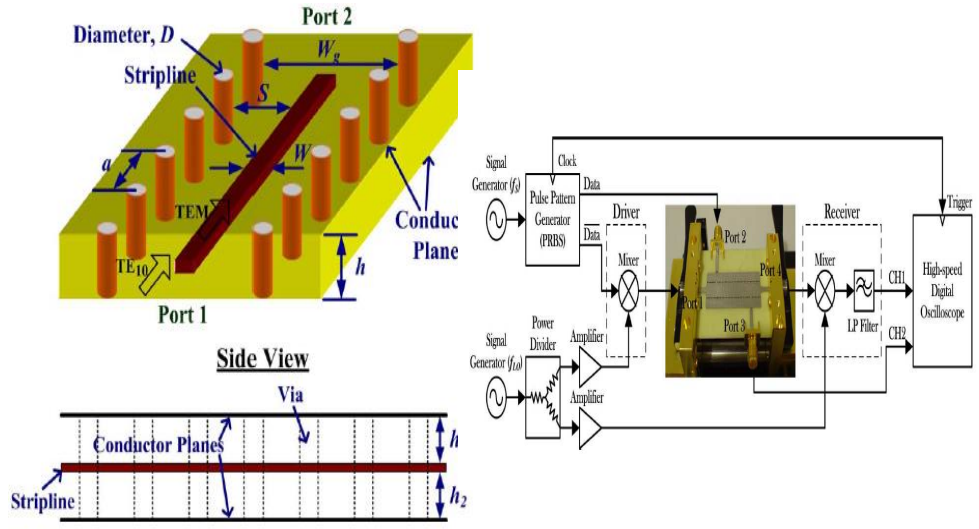
the short length of the link (48 mm).

Fig. 4.11 shows the circuit and the experimental characterization setup.

In a different work [80], the authors propose to leverage multi-mode guiding property of the SIW. Similar to the RWG, many modes can propagate into the guide. TE_{10} and TE_{20} are orthogonal modes; if excited properly they can form the basis for a dual channel interconnect medium with large bandwidth and the advantage of enabling dual-channel communication using a single physical channel.

The excitation of TE_{20} requires a differential feed. This causes the multi-mode guide to become very sensitive to common-mode noise. To avoid the generation of common-mode noise, a balun that connects the single mode guide with the dual-mode guide is designed. The purpose for the balun is to provide a perfect 180° phase between both inputs of the TE_{20} guide. Fig. 4.12 shows the designed circuit and the characterization setup.

The authors demonstrated a simultaneous transmission of PRBS signals of 1 Gb/s



(a) Circuit layout

(b) Characterization setup

Figure 4.11: Hybrid substrate integrated waveguide [79]

each onto TE_{10} and TE_{20} channels with very good isolation.

The excitation of the second mode is very sensitive to unbalance in the balun. The author did not show the coupling noise when manufacturing imperfections causes the balun differential output to have some level of asymmetry. Also, the necessity to provide different up/down conversion circuitry and filtering adds a lot of complexity to the design and requires a sophisticated synchronization mechanism.

Although the concept is elegant, we believe that the reasons described above limited the data rate to 1 Gb/s.

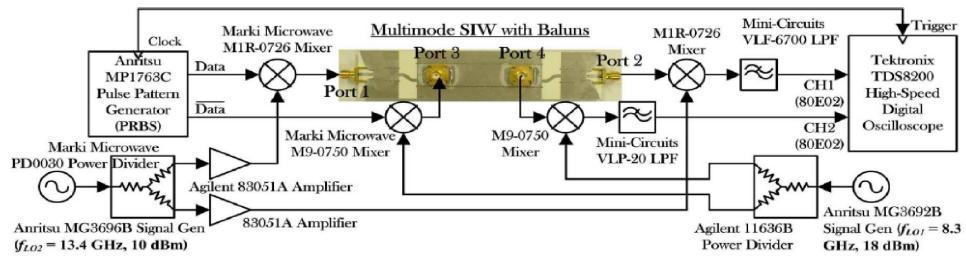
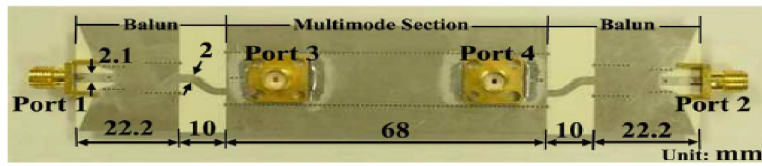


Figure 4.12: Experimental setup for characterization of multi-mode SIW [80].

Chapter 5

MULTICARRIER MEMORY CHANNEL ARCHITECTURE

5.1 Interconnect Proposal

Our interconnect proposal consists of using SIW interconnects to transmit DDR signals. The proposal takes advantage of the large bandwidth of SIW channel, and multiple DDR signals are transmitted in parallel on the same SIW channel by modulating them onto RF carriers and using frequency division multiplexing scheme. Experimental validation, which is detailed in section 6.5, shows that by using 64-QAM modulation, as few as 1 SIW can transmit all 72 data streams (64 data + 8 strobes).

We elected to demonstrate operation of an interconnect at X-band with center frequency at 9.0 GHz. This choice is a balance between compactness and cut-off frequency. Rogers Duroid®5880 dielectric was chosen due to its good quality and small loss tangent at this frequency. Dimensions and design parameters of the SIW structure are summarized in Table 5.1

5.2 Performance of SIW-based DDR Interconnect

In our proposed DDR channel, we plan to use SIW as the interconnect between memory controller and SDRAM devices. To validate that this approach has merit, we performed a signal integrity evaluation of the channel. For that purpose, we used

Table 5.1: SIW design parameters

Parameter	Value
Dielectric	Rogers RT/Duroid 5880
Dielectric Constant ϵ	2.2
Loss Tangent	0.0009/@ 10GHz
Guide Frequency	9.0 GHz
Frequency BW target	≥ 7.0 GHZ /@ -10dB
Via Radius	0.4 mm
Via to Via Spacing (center to center)	1.2 mm
SIW Width	25.62 mm \sim 1.0 in
SIW length	12.0 mm
Substrate height	0.762 mm (30 mil)

the full wave 3D field solver Ansys HFSS to simulate the structures.

5.2.1 Performance of Data and Strobe Interconnect

An SIW interconnect is shown in Fig. 3.11. Design parameters are listed in Table 5.1. In the HFSS simulations, the SIW is excited at the input waveport and its response is measured at the output waveport. The structure is enclosed in a box that extends laterally and vertically for a distance longer than $\frac{\lambda}{4}$ to avoid any non-physical reflections. We used 0.018 mm (1/2 oz) copper layers at the top and bottom plates of the SIW. The vias are made of copper as well. As many as three modes are allowed to be excited at each port. This enables modeling of coupling between higher order modes and the fundamental mode. Fig. 5.1 shows S_{11} (dB), S_{21} (dB) and normalized propagation constant of the SIW channel. The structure has a cut-off frequency around 4.7 GHz.

Despite similarities between RWG and SIW, there is a distinct difference in the

modes that can propagate in both structures. In SIW, the space between via posts is filled with dielectric, and current can not flow in the direction of propagation. Therefore only TE_{m0} modes can be supported in SIW. For WR-90 and its equivalent SIW, modal cut off frequencies are given by:

$$f_{c_{RWG}} = \frac{1}{2a\sqrt{\mu\epsilon}} \quad (5.1)$$

$$f_{c_{SIW}} = \frac{1}{2a_{SIW}\sqrt{\mu\epsilon}} \quad (5.2)$$

Design rules used in this paper for optimal SIW performance are summarized in Table 5.2.

Table 5.2: RWG and SIW design guide comparison

RWG Design Equation	SIW Design Guide
$RWG\ Width = a_{RWG}$	$a_{SIW} = a_{RWG} - 1.08\frac{d^2}{5}$ $+ 0.1\frac{d^2}{a_{RWG}}$
$f_{mc0} = \frac{m}{2a\sqrt{\mu\epsilon}}$	$f_{mc0} = \frac{m}{2a_{eq}\sqrt{\mu\epsilon}}$
$\gamma_{m0} = \sqrt{(\frac{m\pi}{a_{RWG}})^2 - \omega^2\mu\epsilon}$	$\gamma_{m0} = \sqrt{(\frac{m\pi}{a_{SIW}})^2 - \omega^2\mu\epsilon}$
$\lambda_g = \frac{\lambda_{vac}}{\sqrt{\epsilon_r(1-(\frac{f}{f_c})^2)}}$	$\lambda_g = \frac{\lambda_{vac}}{\sqrt{\epsilon_r(1-(\frac{f}{f_c})^2)}}$
NA	Post Diameter $d \leq \frac{\lambda_g}{4}$
NA	Post Pitch (center to center) $P \leq 2d$ or $\frac{\lambda}{2}$. $P = 1.2$ mm (47.2 mils) in our case

The SIW propagation characteristic of the TE_{10} mode is shown in Fig. 5.1. Propagation constant shows highly dispersive behavior in the vicinity of the cut-off frequency, and more or less linear behavior in the rest of the band. This suggests that signal propagation will not show dispersive behavior if the operating frequency

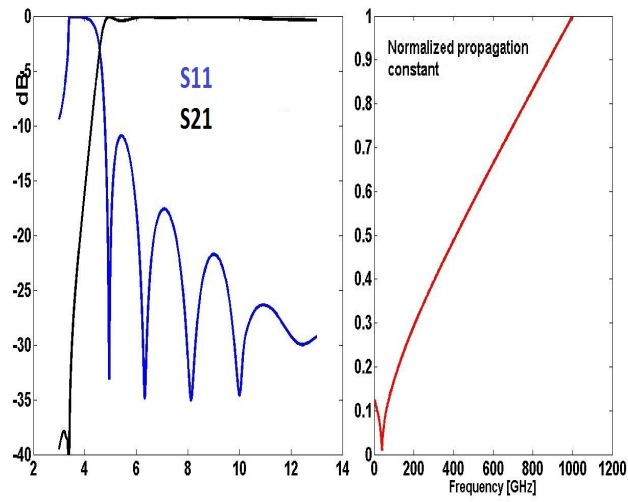


Figure 5.1: Performance of SIW DQ/DQS signal using Ansoft HFSS.

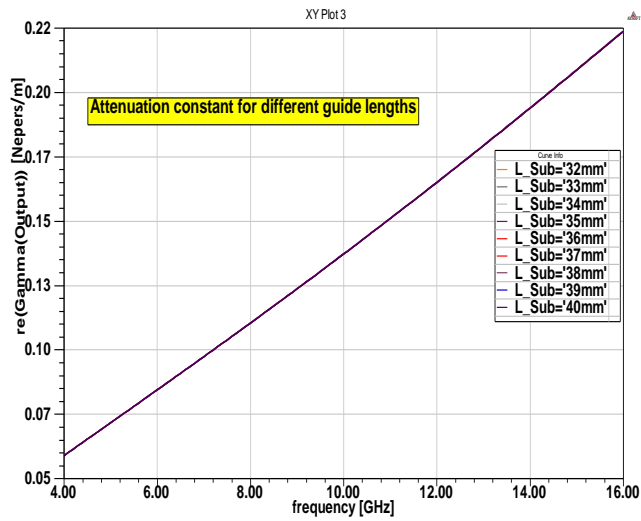


Figure 5.2: Ohmic loss of SIW with varying guide length

is judiciously selected to be far above the highly dispersive frequency range.

Since the TE_{30} field maxima location coincides with the maxima of the fundamental mode, this mode shows higher coupling to the fundamental mode than TE_{20} as shown in Fig. 5.4. Therefore, interconnect performance is affected mainly by TE_{30} mode.

Depending on system topology, total interconnect length varies from one system to another. Fig. 5.2 shows attenuation constant of SIW for guides with lengths from 32 mm to 40 mm (1.26 in to 1.575 in). We can see that in the frequency range of interest, attenuation constant is negligible and does not show dependence on guide length.

To connect the waveguide to the SDRAM and/or to the DDR I/O, we use a tapered microstrip to SIW transition as shown in Fig. 3.11. This is known to cause some signal degradation due to impedance mismatch caused by discontinuity introduced in the H-plane. With careful design of the transition and optimization of taper dimension, the loss is greatly minimized. We performed full 3D optimization of taper dimensions and we achieved an insertion loss better than 10 dB and a return loss of ~ -0.4 dB over the entire channel bandwidth. Fig. 5.3 shows results of SIW with optimal taper dimension. This result is in line with performances reported in literature[49–51].

Coupling to higher modes is negligible. Intensity of TE_{30} is below -18 dB across the entire bandwidth of the channel, and TE_{20} is below -60 dB.

In DDR3 and DDR4, strobe is implemented using differential logic. Simulations of data/strobe performance show that SIW is very immune to skew and noise, therefore, we propose to exclusively use single ended topology for the strobe. This is a significant benefit in comparison to conventional DDR signaling. It enhances power efficiency of the interface and reduces the number of required signals by 8 pins.

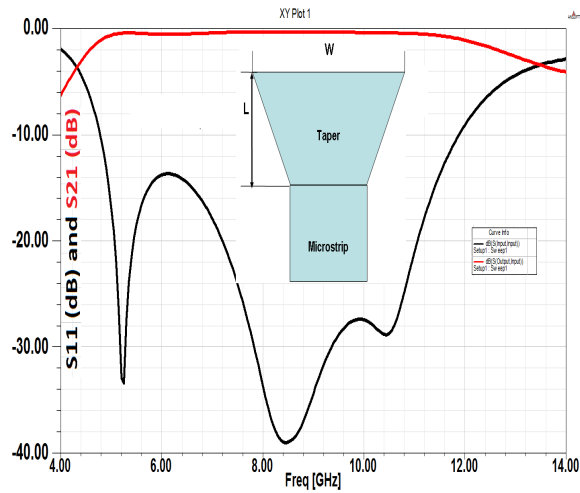


Figure 5.3: SIW s_{11} (dB) and s_{21} (dB) with optimal taper dimensions ($L=3.52$ mm, $W=4.4$ mm) using Ansoft HFSS.

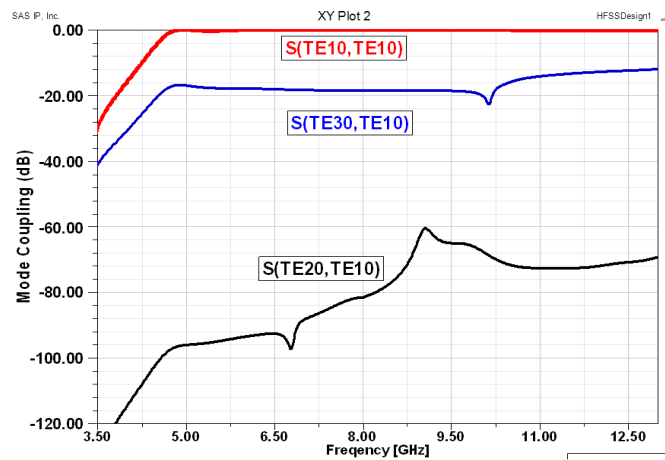


Figure 5.4: SIW mode coupling simulated using Ansoft HFSS.

5.3 Multicarrier Memory Channel Architecture Proposal

Fig. 5.5 shows the high level architecture of MCMCA.

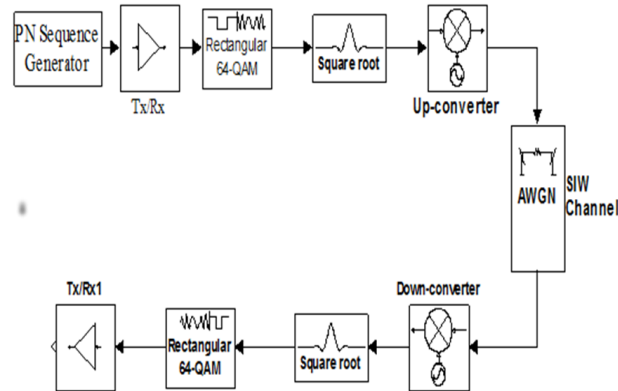


Figure 5.5: Schematic of the multicarrier memory channel architecture proposal

At the transmitter, the memory signals are modulated into a symbol using one of the popular digital modulation formats as appropriate. The symbols are then up-converted onto different carrier frequencies f_{ci} that divide up the SIW bandwidth. At the receiver, the signals are down-converted and then demodulated. We used frequency division multiplexing to maximize the number of signals transmitted per SIW. The channel is divided into a set of frequency bands and each symbol is modulated into a carrier frequency. Spacing between symbols and choice of carrier frequency is determined by the symbol rate and bandwidth occupied by each symbol.

5.3.1 Details of MCMCA Signaling

The MCMCA channel shown in 5.5 can be modeled as digital communication channel where data rate R is limited by signal power, noise power caused by noise interference from the surrounding environment and by distortion related to the physical channel limitations.

Shannon derived a formula that relates the capacity of a channel, i.e. maximum data rate, to the signal to noise ratio under the assumption that the noise can be modeled as additive white Gaussian noise (AWGN)[87]

$$C = B \log_2(1 + SNR) \quad (5.3)$$

where C is the channel capacity in bits/s, B is channel bandwidth in hertz and SNR is the signal to noise ratio.

In digital communication, a normalized version of SNR is rather used as figure of merit. We define E_b as the bit energy which is the signal power S times the bit time T_b . We define N_0 as the noise power N divided by the bandwidth W , often termed noise power spectral density. The bit rate R_b is the inverse of bit duration, T_b . The figure of merit can be written as

$$\frac{E_b}{N_0} = \frac{S}{N} \left(\frac{W}{R_b} \right) \quad (5.4)$$

In communication theory, we assume that the systems are linear time-invariant systems (LTI) for which the system response is fully described by the impulse response by using the convolution theorem.

Knowing the impulse response, and using the superposition principles for linear systems, the overall system response $y(t)$ is given by the convolution relation

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)h(t - \tau)d\tau \quad (5.5)$$

where $x(t)$ is the system input, and $h(t)$ is the system response to the impulse input.

Quadrature amplitude modulation: Quadrature amplitude modulation (QAM) is a digital modulation scheme that modulates both the phase and the amplitude of a carrier wave. QAM belongs to the family of coherent M-ary modulation technique. Instead of using a binary alphabet with 1 bit of information per symbol; it uses an alphabet of M symbols. This technique is known for achieving higher transfer rate and reducing required channel bandwidth. The number k of bits transferred per symbol in an M-ary is

$$k = \log_2 M \quad (5.6)$$

therefore, for a fixed data rate, use of M-ary modulation reduces the required bandwidth by a factor K . QAM modulation consists of two independent bit streams. Let assume k is even, one $(k/2)$ -bit stream modulates the cosine function of a carrier wave and a second $(k/2)$ -bit stream modulates the sine function.

$$\varphi_{QAM} = m_1(t) \cos(w_c t) + m_2(t) \sin(w_c t) \quad (5.7)$$

at the receiver, each of the two signals is independently detected using matched filters.

QAM modulator and demodulator At the demodulator:

$$\begin{aligned} x_1(t) &= \varphi_{QAM}(t) 2 \cos w_c t \\ &= (m_1(t) \cos w_c t + m_2(t) \sin w_c t) 2 \cos w_c t \\ &= m_1(t) + m_1(t) \cos 2w_c t + m_2(t) \sin 2w_c t \end{aligned}$$

The original signal $m_1(t)$ can be recovered by passing $x_1(t)$ through a low pass filter.

The sampling theorem

A bandlimited signal whose Fourier transform is zero outside a finite region of frequencies beyond the frequency f_m can be uniquely reconstructed by samples at a

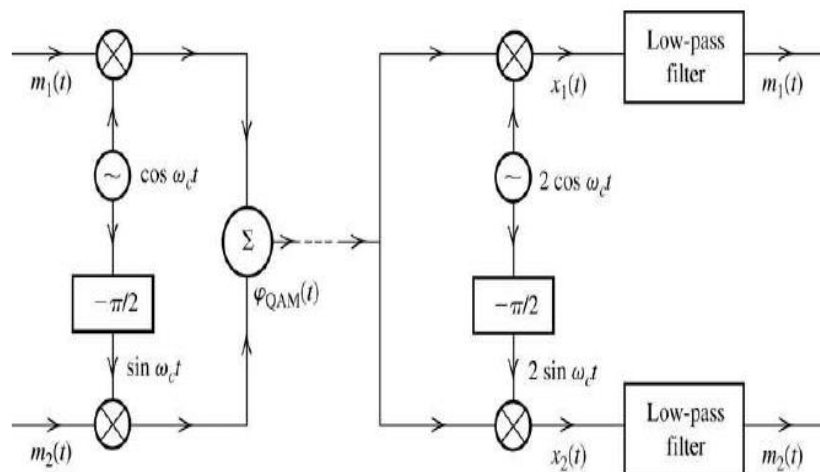


Figure 5.6: QAM modulation and demodulation diagram

sampling rate T_s for which $T_s \leq \frac{1}{2f_m}$. The theorem states that a sufficient condition for reconstructing an analog signal out of discrete samples by establishing the threshold sampling rate.

The sampling rate $f_s = \frac{1}{T_s} = 2f_m$ is called Nyquist rate.

5.3.2 Channel Design for Zero ISI

A necessary and sufficient condition for a bandlimited signal $x(t)$ to have zero ISI is

Let

$$x(nT) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases} \quad (5.8)$$

is that its Fourier transform $X(f)$ satisfies

$$\sum_{m=-\infty}^{\infty} X(f + \frac{m}{T}) = T \quad (5.9)$$

where $\frac{1}{T}$ is the symbol rate.

There are many signals that can satisfy this property. A popular signal that has a raised-cosine frequency response is widely used in signal processing. The frequency response of the raised cosine signal is

$$X_{rc}(f) = \begin{cases} T, & 0 \leq |f| \leq \frac{1-\alpha}{2T} \\ \frac{T}{2} \left[1 + \cos \frac{\pi T}{\alpha} \left(|f| - \frac{1-\alpha}{2T} \right) \right], & \frac{1-\alpha}{2T} < |f| \leq \frac{1+\alpha}{2T} \\ 0, & |f| > \frac{1+\alpha}{2T} \end{cases} \quad (5.10)$$

where α is called the roll-off factor, which takes values in the range $0 \leq \alpha \leq 1$. The bandwidth occupied by the symbol beyond the Nyquist frequency is called the excess bandwidth.

The frequency response $X_{rc}(f)$ corresponds in the time domain to the pulse $x(t)$

$$x(t) = \text{sinc} \left(\frac{\pi t}{T} \right) \frac{\cos \left(\frac{\pi \alpha t}{T} \right)}{1 - 4\alpha^2 \frac{t^2}{T^2}} \quad (5.11)$$

The first part of the equation is the sinc pulse, which ensures that the function transitions at the integer multiples of the symbol rate. The second part, containing the cosine, controls the excursion between the sampling instants beyond Nyquist rate.

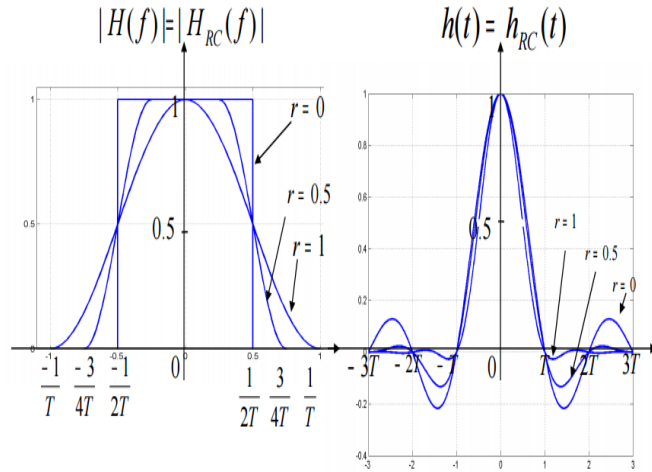


Figure 5.7: Pulses with a raised cosine spectrum [87].

5.3.3 MCMCA Proposal

We propose using a multicarrier frequency division multiplexing scheme as shown in Figs. 5.5 and 5.8. The channel is divided into a set of frequency bands centered around f_{ci} ; [$i = 1 \dots n$]. Each data stream occupies one band. Various modulation schemes could be used. We chose 64-QAM modulation format as it represents an optimal compromise that maximizes throughput while preserving good EVM and phase noise characteristics. The symbols pass through a filter to be filtered and optimally shaped before being sent over the channel. There are many filter types available to shape digital symbols. We used a Nyquist filter because it minimizes ISI.

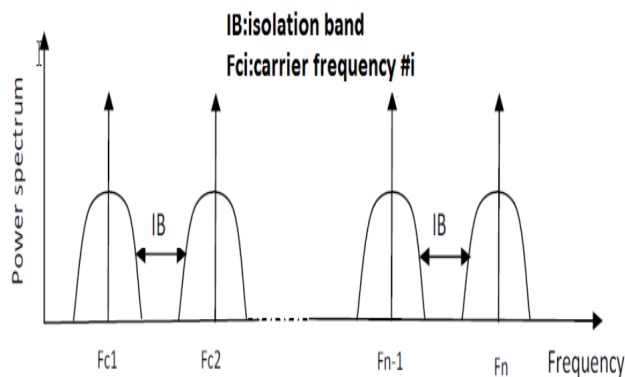


Figure 5.8: Proposed frequency division multiplexing scheme[88].

The time domain impulse response of the Nyquist filter has its peak amplitude at the symbol cursor ($t=0$), and its zeros at the integer multiples of the symbol period. Since other symbols happen at integer multiples of the period, there is zero inter-symbol interference [87]. Roll-off constant α accounts for the sharpness of the filter and is an indicator of excess bandwidth occupied by the symbol beyond the theoretical minimum. Nyquist theory states that the minimum bandwidth required to transmit a signal is equal to one half the signal rate [89]. This requires a perfect brick-wall filter with $\alpha = 0$ and hence it occupies the least bandwidth which equals the symbol rate. In practice, a theoretical minimum filter is not realizable and a

typical α ranging from 0.25 to 0.35 is used. The occupied bandwidth is given by[90]:

$$Occupied\ BW = (1 + \alpha) * Symbol\ rate \quad (5.12)$$

An isolation band IB between channels is introduced to avoid inter-channel crosstalk. We choose a spacing of 40 MHz and experimentally verified that it is indeed adequate for good inter-channel isolation. The choice of symbol rate and the number of channels allowed within the SIW bandpass could be done in different ways. Obviously, as the symbol rate increases, so does bandwidth. One can maximize the symbol rate so as to occupy the whole SIW bandwidth SIW_{BW} . In such a case, the maximum allowable symbol rate SR_{MAX} is:

$$SR_{MAX} = \frac{SIW_{BW}}{1 + \alpha} \quad (5.13)$$

In our case, $SIW_{BW} = 6.5\ GHz$, so for $\alpha = 0.3$, $SR_{MAX} = 5.0\ GSymbol/s$. With 64-QAM, 6 bits are transmitted with each symbols. Which yields a maximum possible rate of 30 Gbit/s per channel. It takes 11 channels to transmit the 64 Bits which yields a bus bandwidth of 240 GBytes/s.

With maximum symbol rate option, one might be restricted to only use a modulation with high spectral efficiency at the expense of data rate throughput. Another disadvantage of this choice is the complexity of modulator/demodulator design. The higher the symbol rate, the more difficult and challenging the design of modulator and demodulator. A more practical usage model of available bandwidth is to divide the bandwidth into channels and choose a modulation scheme that maximizes the number of bits transmitted per symbol and use a lower symbol rate instead. The number N of allowable channels is then given by:

$$N * SR(1 + \alpha) + (N - 1) * IB \leq SIW_{BW} \quad (5.14)$$

where IB is the isolation band, α is the raised cosine filter roll-off coefficient. N is rounded to the smaller integer value. $SIW_{BW} = 6.5 \text{ GHz}$; and $IB = 40 \text{ MHz}$.

The number of available channels is a function of the symbol rate and roll-off constant for a given isolation band. Table 5.3 gives the number of channels and aggregated channel bandwidth for different combinations of roll-off constants and symbol rates.

Table 5.3: Usage models of available bandwidth

Channel Usage scenarios			
	Symbol rate	# of channels	Aggregated BW
$\alpha = 0.25$	130 MHz	29	22.6 Gbit/s
	500 MHz	9	27.0 Gbit/s
	1.0 GHz	4	24.0 Gbit/s
$\alpha = 0.3$	130 MHz	28	21.8 Gbit/s
	500 MHz	8	24.0 Gbit/s
	1.0 GHz	4	24.0 Gbit/s
$\alpha = 0.35$	130 MHz	28	21.8 Gbit/s
	500 MHz	8	24.0 Gbit/s
	1.0 GHz	4	24.0 Gbit/s

Chapter 6

EXPERIMENTAL CHARACTERIZATION OF MCMCA

We devised three experiments to validate the MCMCA proposal. The first experiment aims to characterize the network property of the channel by measuring 2 port S-parameters of the SIW from DC to 20.0 GHz and comparing against HFSS simulation results. The second experiment is used to quantify the effects of dispersion on the propagation of modulated signals and to identify the bandwidth range over which dispersion effects do not compromise the integrity of transmitted signals. The third experiment is intended to characterize dispersion characteristics of the channel by measuring group delay.

6.1 Characterization of Channel S-Parameters

We connected the SIW channel to a VNA through two SMA connectors. The SMA is connected to the SIW via a stripline taper optimized to minimize the transition loss[90]. We used Agilent 8720ES, GigaTest probe station and 20.0 GHz rated coaxial cables. A photograph of the characterization bench is shown in Fig. 6.1.

Fig. 6.2 shows an overlay of simulation and measured channel S-parameter waveforms. Cutoff frequency is 4.7 GHz and the usable bandwidth spans from 4.7 GHz up to about 12.2 GHz, which is about 7.5 GHz.

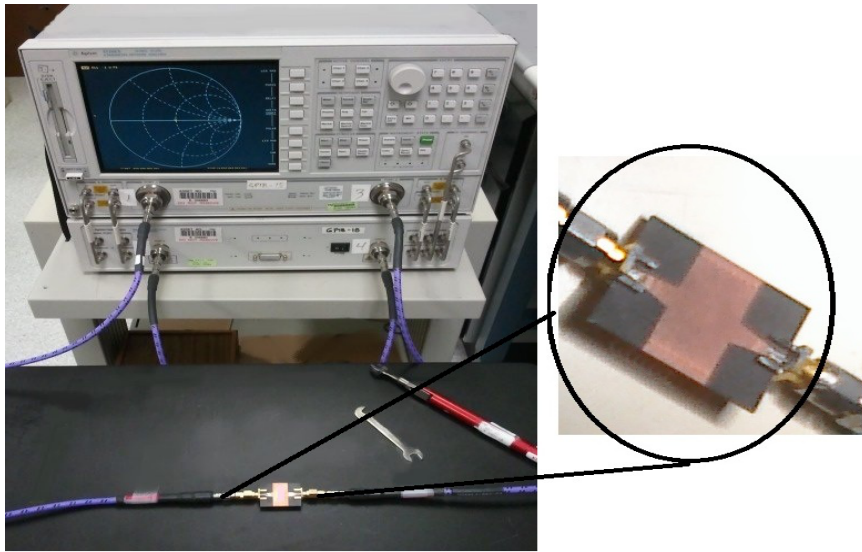


Figure 6.1: SIW VNA test bench

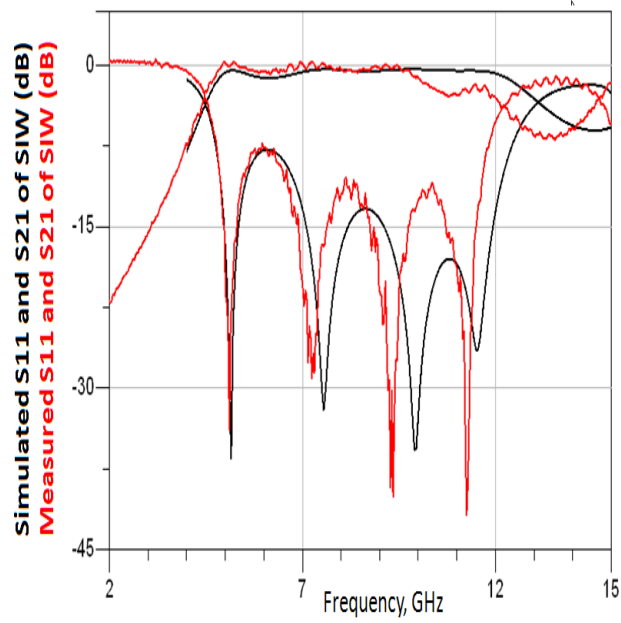


Figure 6.2: Measured and simulated channel S-parameters

6.2 Study of Channel Distortion Characteristics

In section 5.2, we presented results of full-wave 3D simulations of the propagation characteristics of SIW structure. Fig. 5.1 showed that the channel exhibits high dispersion behavior if operated in the vicinity of its cutoff frequency. In our experiment, we characterized the impact of the channel dispersion property on a modulated signal by measuring symbol figures of merit like EVM and phase distortion at different carrier frequencies f_c .

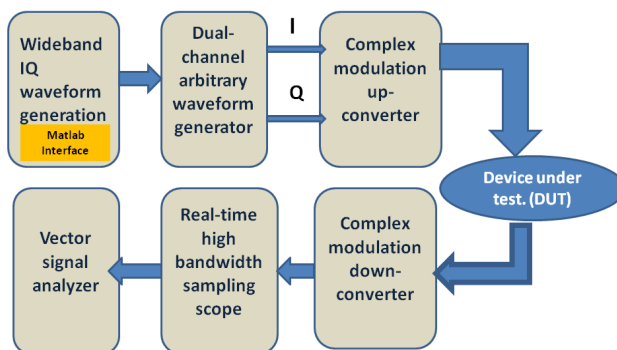


Figure 6.3: Conceptual setup for measuring distortion

Fig. 6.3 shows the conceptual diagram of the test bench used in characterizing modulated signal distortion. An external PC with MatlabTM is used to generate the I and Q wideband signals. It is then downloaded into an Agilent 81180A arbitrary waveform generator with 12 bits DAC resolution up to 4.2 GSa/s and 2 GHz IQ modulation bandwidth. The IQ outputs of the arbitrary waveform generator are fed into an Agilent E8267D RF and microwave signal generator which modulates the IQ onto a carrier in the passband range of the SIW channel. The output of the vector signal generator is applied in closed loop without DUT to Agilent high performance 13.0 GHz digital oscilloscope DSO 91304A with vector signal analyzer (VSA). The

set up is then calibrated to correct for equipment imperfections. In this phase a great deal of equalization and noise cancellation is performed until the auto-calibration algorithm reaches its best possible set up. This sets the performance floor of the test bench, which needs to be de-embedded later from the DUT measurement. The intrinsic EVM and phase error $\Delta\Phi$ of the setup are recorded and called EVM_{ref} and $\Delta\Phi_{ref}$. The DUT input is then connected to the output of the signal generator while the DUT output is connected to the oscilloscope. The wideband signal is then demodulated at the RF carrier frequency. The magnitude and phase response of the channel output and the IQ waveforms are analyzed using the VSA software in the oscilloscope. The complete set up is shown in Fig. 6.4.

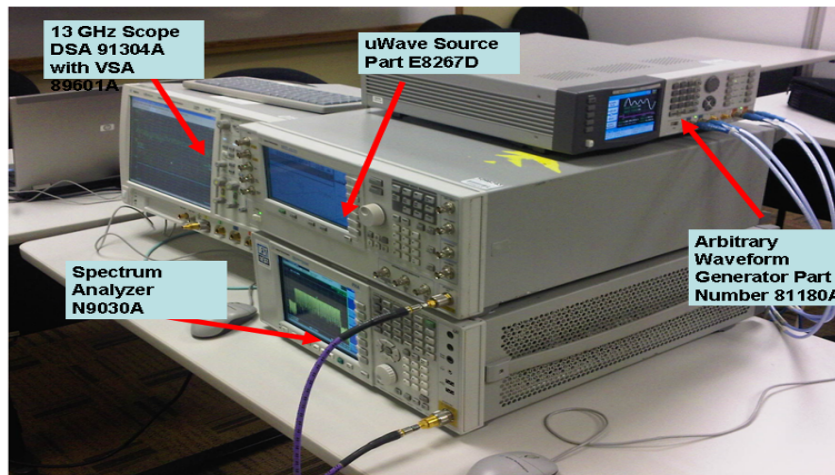


Figure 6.4: Experimental set-up for channel distortion characterization.

6.2.1 Error Vector Magnitude and Related Figure of Merit Measurements

Error vector magnitude (EVM) is a measure of signal quality that provides a quantitative figure-of-merit for a digitally modulated signal. Since EVM processes the signal in vector form, it is capable to provide a measure of both amplitude and phase.

In order to fully understand the measurements and the underlying causes of signal impairments, we start by providing an accurate definition of the measurements an EVM instrument is capable of performing as well as the nomenclature EVM uses.

EVM : Error Vector Magnitude (EVM) is a measurement of modulated or demodulated signal performance in the presence of impairments. EVM is the vector difference at a given time between the ideal signal and the measured signal. The ideal signal is often called the reference signal and the error vector magnitude is measured relatively to the reference vector as shown in the Fig. 6.5[91].

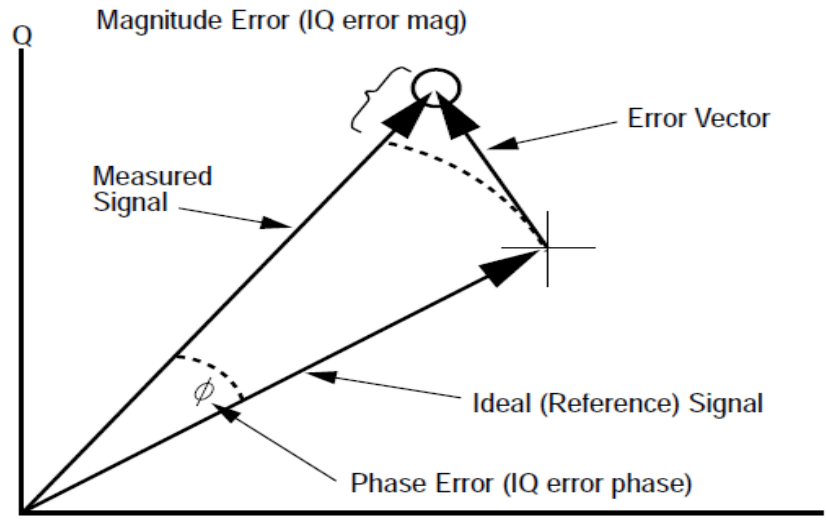


Figure 6.5: The error vector magnitude concept (EVM)[90].

Phase error : error in degree between the ideal expected phase and the actual signal phase.

IQ phase error : The instantaneous angle difference between the measure signal and reference signal. When viewed as a function of time, it shows the modulating waveform of any residual phase modulated signal.

Constellation diagram : A polar graphical representation of the phase and magnitude of the vector-modulated signal relative to the carrier as function of time or symbol. The instrument requires the knowledge of the precise carrier and symbol clock frequencies and phases in order to construct the diagram.

6.2.2 Distortion Measurements

The cutoff frequency of the SIW used in the experiment is 4.7 GHz; We performed this experiment using three carrier frequencies that are at different distances from the cutoff: one carrier at 5.6 GHz, a second carrier at 6.6 GHz and a third carrier at 9.6 GHz (deep in the middle of the channel passband).

The signal is first modulated into a 100 MSym/s QPSK symbol and then up-converted onto one of the three test carriers.

At the receiving end, we capture on the scope display the I and Q eye diagram, the constellation and/or transition diagram, the received symbol spectrum and tabulated performance numbers like EVM, phase error and I-Q imbalance.

Since EVM is an RMS value, de-embedding set up induced EVM error from channel measured EVM is performed using the formula[90]:

$$EVM = \sqrt{(EVM)_{measured}^2 - (EVM)_{ref}^2} \quad (6.1)$$

While phase error is obtained using

$$\text{Tan}^{-1} \frac{Q}{I} = \left[\text{Tan}^{-1} \frac{Q}{I} \right]_{\text{measured}} - \left[\text{Tan}^{-1} \frac{Q}{I} \right]_{\text{ref}} \quad (6.2)$$

Fig. 6.6 shows the received symbol performance when the carrier is 5.6 GHz; not far from the cutoff frequency. The symbols on the constellation diagram (red dots in the figure) resembles more of a cloud than a focused point location. The distance between the orthogonal symbols is so close that one can visually notice the likelihood of bit error and symbol overlapping possibility. The measured SNR is 13.35 dB which is low and correlates to a high BER. The I and Q eye diagrams show poor eye opening and large jitter at zero crossing.

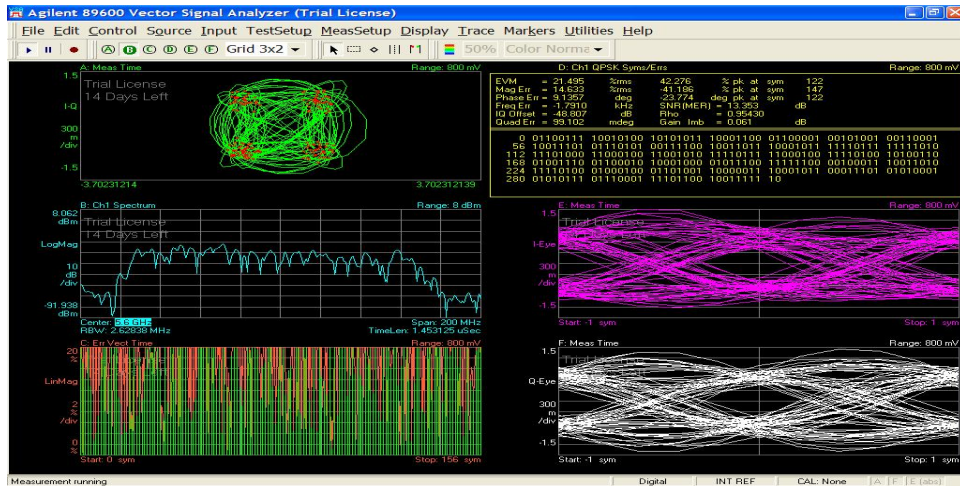


Figure 6.6: Symbol distortion for 100 MSym/s QPSK symbol when the carrier is in very close to the cutoff frequency (5.6 GHz)

The EVM reads 21.5 %; phase error is 9.14°; frequency error is -1.8 KHz, and IQ offset is -48.8 dB.

When the carrier moves farther away from the cutoff frequency and is set at 6.6 GHz, the improvement is dramatic. Fig. 6.7 shows the performances of the same symbol as the previous case with the difference of the carrier, which is in this case at

6.6 GHz. The EVM drops to 7.3 % which is about 200% improvement. The phase error drops to 3.1; frequency error is 1.7 KHz, while the IQ offset becomes -53.3 dB.

The SNR is at 22.6 dB. The I and Q eye diagrams show much larger eye opening while the jitter at zero crossing remains comparable to the previous case.

In Fig. 6.8, the carrier is set to a more optimal value of 9.6 GHz, deep in the middle of the passband. Sure enough, the performance shows an impressive improvement. The EVM reads 2 %, the phase error is less than 1°, the frequency error is -62 Hz (three orders of magnitude smaller than the previous both cases), the IQ offset is -63.9 dB and the SNR is 34 dB.

The I and Q diagrams show very large eye opening and the jitter at zero crossing is reduced substantially.

On the constellation diagram the symbols are compact and show extremely small spread.

Table 6.1 summarizes the performances of the three carriers.

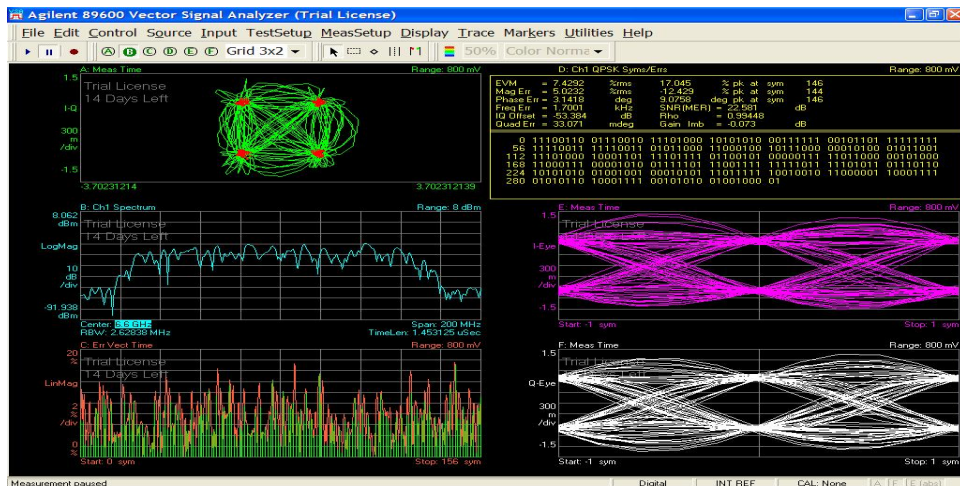


Figure 6.7: Symbol distortion for 100 MSym/s QPSK when the carrier is farther from the cutoff frequency (6.60 GHz)

The channel distortion experiment demonstrates the behavior of the channel in

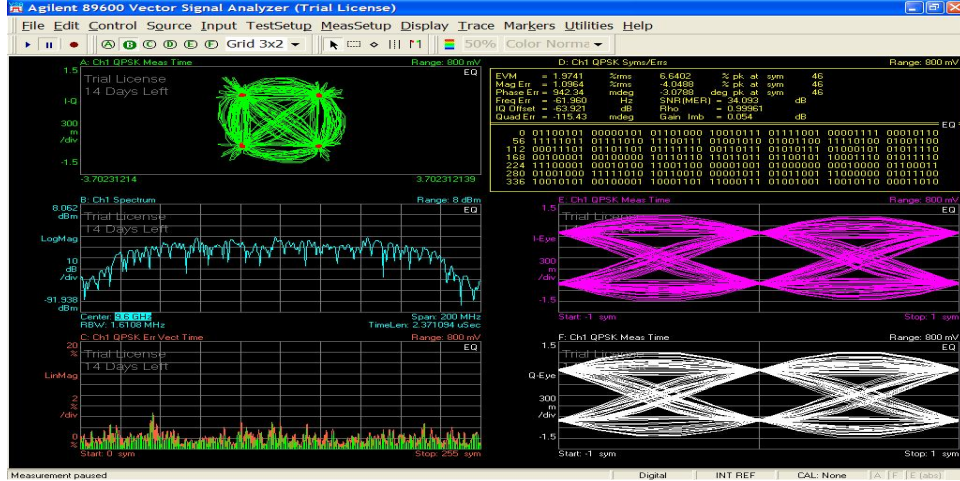


Figure 6.8: Symbol distortion for 100MSym/s QPSK at 9.6 GHz carrier

Table 6.1: Impact of channel dispersion on 130 MHz QPSK modulated symbol

	$f_{c1} = 5.6$ GHz	$f_{c2} = 6.6$ GHz	$f_{c3} = 9.6$ GHz
EVM[%]	21.5	7.4	1.97
Phase Error ^o	9.1	3.1	0.94
SNR[dB]	13.35	22.6	34
Freq. error [Hz]	-1.8K	1.7K	-62
IQ offset	-48.8	-53.3	-63.9

the highly dispersive portion of its passband and shows a degradation in dispersion characteristic of the channel as we move closer to the cutoff edge. This information is of crucial importance for the designer as it determines the usable bandwidth that meets the target distortion budget of the channel. Once the target distortion budget and the BER are specified, the designer can determine the minimum carrier frequency that can be used.

6.3 Group Delay Characterization

To further validate this finding, we characterized the SIW group delay using Agilent N5242A PNA-X. The frequency is swept from 4.0 GHz to 12.0 GHz in increments of

20 MHz. The measured group delay result is shown in Fig. 6.11. At the vicinity of cutoff frequency from 4.0 GHz to 6.2 GHz, the group delay exhibits large variability. This is indicative of high dispersion. In the frequency range from 6.2 GHz to 12.0 GHz, except for a bump around 9.0 GHz, the group delay waveform is flat, which explains the exceptional phase performance of the channel. We ran HFSS simulations of the structure with and without taper transition from SMA to the waveguide, and we noticed that the bump at 9.0 GHz is attributable to this transition. With more detailed optimization of the taper shape and dimension, we believe the SIW group delay flatness from 6.2 GHz to 12.0 GHz would improve.

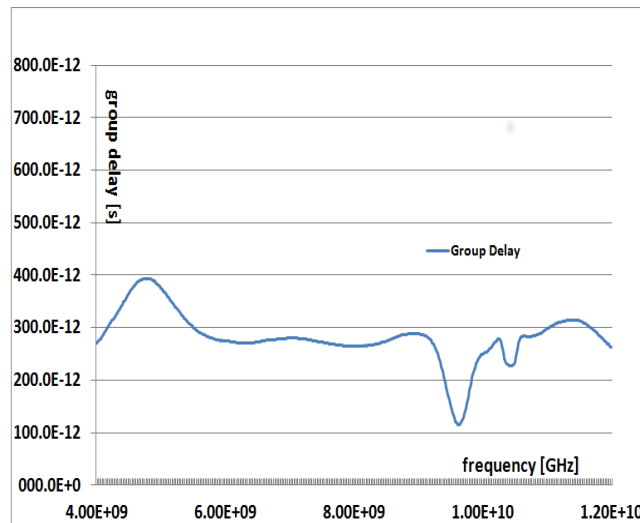


Figure 6.9: Measured group delay of the channel.

To avoid symbol degradation due to channel dispersion, we recommend keeping a 1.0 GHz guard band (GB) away from the cutoff frequency. This GB will reduce the overall channel available bandwidth, but it will substantially reduce distortion effects. Taking GB into consideration, the number of available channels given previously in equation 5.14 becomes

$$N * SR(1 + \alpha) + (N - 1) * IB + GB \leq SIW_{BW} \quad (6.3)$$

where GB is the keep-out guard band needed to avoid the dispersive region of the channel.

6.4 Demodulation Set Up

Fig. 6.12 shows a snapshot of the demodulation set up. The DUT is connected on one end to microwave source output, and on the other end, it is connected through high bandwidth SMA connector to digital signal analyzer (DSA) sampling scope. The VSA application offers versatile features for user to verify quality of the demodulated signal. Most important features are EVM, symbol spectrum, phase noise, equalizer set up, I and Q eye diagrams and constellation diagram. Transition diagram can also be displayed.

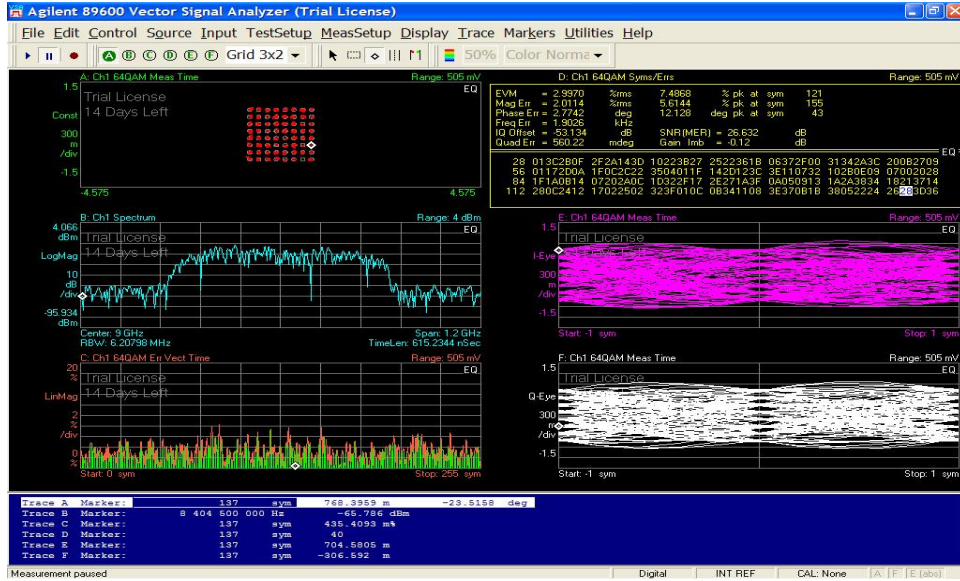


Figure 6.10: SIW channel performance using 64-QAM at 500 MHz symbol rate at 9.6 GHz carrier

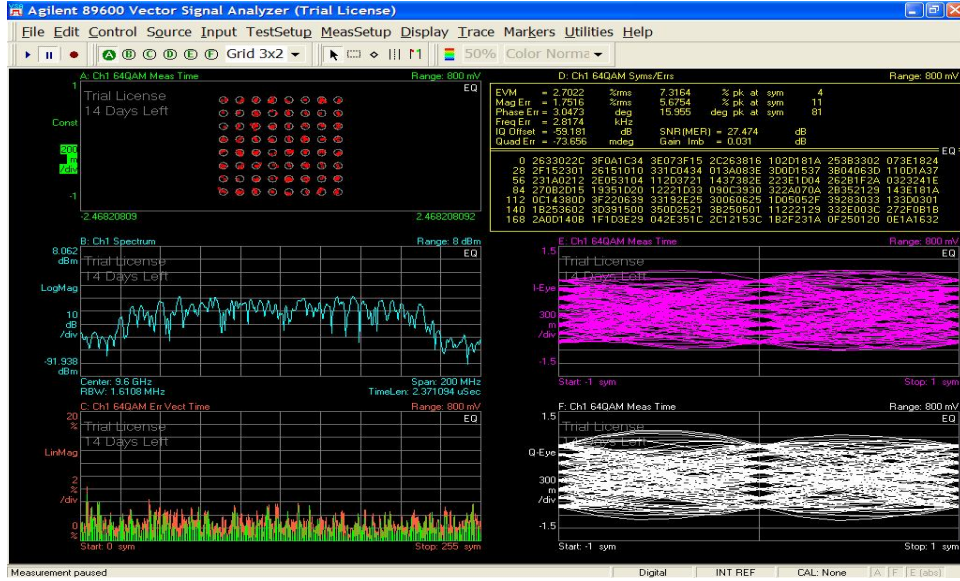


Figure 6.11: SIW channel performance using 64-QAM and sending simultaneously 4 symbols at 130 MSym/s symbol Rate

6.5 End-to-end Experimental Validation of Memory Channel Proposal

The channel characterization experimental setup is identical to the used in characterizing channel distortion and is shown in Fig. 6.4. We used the 64-QAM modulation format to modulate a PN15 pseudo-random $2^{15} - 1$ bit stream representing digital memory signals. MatlabTM is used to generate the 64-QAM wideband IQ symbols. The bit stream is then fed into the Agilent 81180A arbitrary waveform generator which generates the appropriate I-Q format. I-Q outputs are fed into the Agilent E8267D microwave source which up-converts the symbol onto the desired carrier. The signal is then fed into an SMA connector that connects to a tapered microstrip which excites the SIW. At the receiving end, another tapered stripline delivers the SIW output signal to a second SMA which is connected to the instrument demodulator. The received signal is then sampled and demodulated using the high sampling

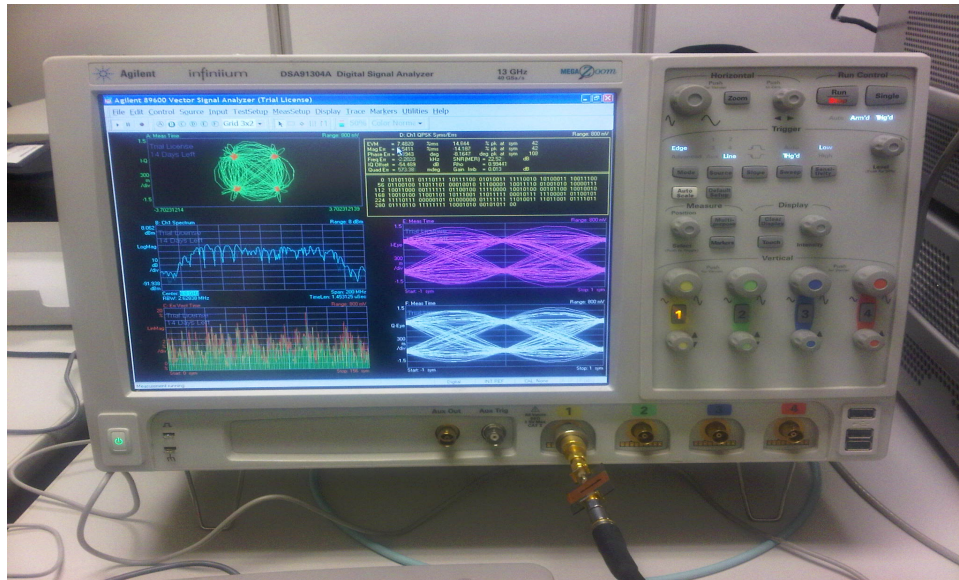


Figure 6.12: Signal analysis at the receiver using digital signal analyzer and vector signal analyzer

scope (13GHz), the Agilent 91304A loaded with vector signal analyzer demodulation application Agilent VSA 89601A.

We performed two sets of experiments. In the first set, we maximized the symbol rate to 500 MSym/s and used single carrier at 9.0 GHz. In the second set, we sent 4 symbols at 130 MSym/s each onto four different carriers at 9.2, 9.4, 9.6 and 9.8 GHz with 40 MHz channel-to-channel spacing.

Although we showed in Table 5.3 that the SIW is capable of hosting a high number of channels, we were in practice limited by the arbitrary waveform generator bandwidth and sampling rate of the demodulator as explained below. For 64-QAM signals at 500 MSymbol per second, taking into account the fact that each symbol requires 8 samples; therefore the necessary bandwidth is $0.5 \cdot \text{GHz} \cdot 8 \text{ samples} = 4.0 \text{ GHz}$ which is about the maximum bandwidth available in the set up.

We measured the performance of the test bench at 500 MSym/s over 9.0 GHz carrier when no DUT is attached. We used 0.35 as the filter roll-off coefficient. We

can see that the symbol spectrum occupies a bandwidth about 670-700 MHz both without and with the DUT. This indicates that Eq. 5.12 is a good approximation for bandwidth budget. The reference test bench performed an auto calibration process to settle into optimal setup with respect to equalization adjustment and synchronization between transmitter and receiver. Auto-lock algorithms enables the test bench to precisely lock at carrier and symbol clock frequencies and phases. Constellation diagram is uniform and shows good symmetry about the origin. No imbalance between I and Q is observed and the overall diagram is well squared which indicates that the auto calibration algorithm succeeded in finding an optimal setup. EVM of the reference set up is 1.97 % and phase error is 1.70°.

Similar to waveforms of the reference setup, visual inspection of the constellation diagram and spectrum response of measured waveforms for 64-QAM at 500 MSym/s with DUT inserted between set up transmitter and receiver, shows a well squared and symmetric one. The SIW channel contributed to overall EVM and phase error and the combined response of set up and DUT reads an EVM of 3.00 and phase error of 2.77°. The intrinsic contribution of the channel after de-embedding the setup contribution yields an EVM of 2.26% and a phase error of 1.07°.

The same experiment is performed for multicarrier usage of the channel where 4 symbols at 130 MSym/s rate are simultaneously sent over the channel onto 4 different carrier frequencies, namely 9.2, 9.4, 9.6, and 9.8 GHz. We used 40 MHz separation between the carriers. Error vector spectrum shows that there are no in-channel spurs. No origin offsets had been observed, which is an indicator that there were no carrier feedthrough nor DC offsets at the I or the Q signals. Performance of MCMCA when using single carrier at 500 MSym/s and when 4 symbols of 130 MSym/s are sent simultaneously onto 4 different carriers is summarized in Table 6.2 and Table 6.3.

The auto-calibration algorithm settled into an EVM and phase noise floor values

Table 6.2: Performance of the MCMCA channel: single carrier at 500 MSymbols/s

	1 symbol at 500 MSym/s		
	Reference set up	Setup with DUT	SIW channel contribution
EVM [% rms]	1.97	3	2.26
Phase error[°]	1.7	2.77	1.07

Table 6.3: Performance of the MCMCA channel: 4 symbols at 130 MSymbols/s onto 4 different carries

	4 symbols at 130 MSym/s		
	Reference set up	Setup with DUT	SIW channel contribution
EVM [% rms]	2.4	2.7	1.24
Phase error[°]	2.5	3.05	0.55

higher than that of the first experiment. EVM reads 2.40 vs. 1.97 for first experiment EVM floor. Phase error reads 2.50° vs. 1.70° for first experiment. By de-embedding the reference floor from the measured channel performance, one can see that channel contribution to total EVM in the case of a single carrier at 500 MSym/s is 2.26 % while phase error caused by the channel is 1.07°. For MCMCA with 4 symbols case, channel contributed EVM is 1.24% and channel contributed phase error is 0.55°.

6.6 Performance Comparison Between Classical DDR Bus and MCMCA Proposal

A summary of the performance comparison between our channel proposal and traditional DDR bus is presented in Table 6.4. Mainstream DDR3/4 channel is limited in speed to about 3 Gbit/s. Although, GDDR5 achieves and exceeds 5.0 Gbps, it is a different architecture that relies on taking advantage of the parallel nature of GPU, and it is not relevant to this paper. MCMCA proposal has the potential to achieve 30

Gbit/s per pin and shows superior signal fidelity and immunity to noise. Therefore there is no need for differential signaling nor for most of the DDR terminations that are designed to cope with impedance mismatch reflections. We showed that with large SIW bandwidth and modulation with high bit per symbol rate, we can transmit all 64 data bits into one SIW. Physical bus width for a typical DDR3 channel with transmission line width between 4 to 5 mils and data to data signal spacing of 10 to 12 mils, will be between 1 to 2 inches. Our SIW width is about 1 inch wide. Therefore, real estate wise, SIW is comparable to standard DDR signaling.

6.6.1 System Integration Perspective

In our design, we used the instrument modulator/demodulator to generate and to recover the DDR signal. For the solution to be practical, one needs to assess the feasibility of integration of the transceiver using mainstream CMOS technology. RF signal generation and reception in CMOS chips is now routine; CMOS modulator/demodulator for Gigabit millimeter-wave applications has been reported in the literature. In [92–94], the design of millimeter-wave IQ modulator using CMOS, RF CMOS and CMOS SOI processes achieved modulation bandwidth of 1 GHz, 4GHz and 14 GHz respectively. Reported chip area is $0.65 \times 0.58 \text{ mm}^2$, $1.54 \times 1.77 \text{ mm}^2$ and $0.6 \times 0.7 \text{ mm}^2$ respectively. In brief, the expected IQ transceiver overhead in MCMCA will be less than $1 \times 1 \text{ mm}^2$.

In this document, we propose and demonstrate a potential solution to the so-called memory-wall problem. We call the solution (MCMCA). Characterization of MCMCA proposal demonstrates that it is a promising solution for alleviating the memory wall problem in digital platforms. With the MCMCA proposal, memory speed can reach as high as 30 Gbit/s for single and multicarrier schemes. The interconnect solution using the SIW structure proved to be compact and compatible

Table 6.4: Memory bus comparison

	Classical DDR	MCMCA
Data Rate [GTS]	0.8-3	16-30
Clock frequency [GHz]	0.4-1.5	8-15
Address/Control [GHz]	0.2-0.75	up to 8
Strobes	8 Diff. Strobes	2 Symbols within single SIW channel
Clock Network	3 Diff. Network	1 Symbol
ADD/CMD Signals	20 signals	1-2 power splitter

with mainstream PCB technology. The MCMCA compactness and ease of integration make it preferable to optical interconnects. Immunity of the channel to skew and cross talk noise enables it to overcome intrinsic stripline and microstrip limitations witnessed in conventional DDR interconnect. The multicarrier 64-QAM modulation scheme maximizes throughput and enables the transmission of many DDR signals using only one physical channel which results in a compact memory bus suitable for high density small form factor platforms. Although the contribution of the channel to overall distortion is higher at higher symbol rates, we proved that the channel is not the limiting factor in choosing the appropriate symbol rate and the number of channels. The trade off between symbol rate and number of channels is determined by the availability and design complexity of the modulator/demodulator. The channel performance is acceptable for low as well as high symbol rates.

Chapter 7

WIDEBAND INTERCONNECTING TECHNOLOGIES FOR MUTLI-GHZ MCMCA

7.1 Introduction

We showed in chapter 3 and in chapter 5 that with careful optimization of via post diameters and via pitch, leakage of energy through side wall can be made negligible and SIW performance can be made comparable to classic rectangular wave guides (RWG) [95–97]. SIW performance matches very closely the performance of the rectangular waveguide and hence is immune to crosstalk, exhibits very large bandwidth and has substantially less conductive loss than microstrip and stripline. Therefore, it seems intuitively obvious that SIW is a superior interconnect medium compared to the traditionally deployed transmission line-based interconnect.

Nonetheless, transmission line has the advantage of operating from the DC to high microwave frequencies without the need for up-converting and down-converting steps in the signal chain. Therefore, a judicious evaluation of both interconnect techniques that takes into consideration the system level end-to-end performance is needed.

In this chapter, we validate the merit of our proposal by comparing it to leading transmission line-based bandpass filter (TLBPF). We use an end-to-end MCMCA platform and perform a thorough comparison between our SIW solution and TLBPF-based solution through using system figures of merit such as BER, EVM, and eye diagram as performance metrics for the appropriate interconnect choice.

We start first by designing the transmission line-based bandpass channel, then we

build the end to end MCMCA channel using Matlab® and simulink ®.

After validating the MCMCA system, we use the SIW measured channel (measured SIW S-Parameters) and perform the system simulation. We perform this operation while varying the symbol rate and collecting the performance data.

Since the objective is a comparison between interconnect solutions, no attempt has been made to implement advanced equalization or signal recovery techniques.

7.2 Hairpin Filter Design

Bandpass filter topologies that operate at microwave frequencies are based on transmission lines and transmission line sections. When the design requires high bandwidth and compact footprint, hairpin filter topology is usually the topology of choice.

Hairpin filter is based on U-shape half-wavelength resonator as the main building block which is called the hairpin resonator [98]. The starting design equations are based on the theory of the parallel-coupled, half-wavelength resonator filter equations [98–100], and are summarized in equations (7.1)-(7.3).

$$\frac{J_{01}}{Y_0} = \sqrt{\frac{\pi FBW}{2 g_0 g_1}}, \quad (7.1)$$

$$\frac{J_{jj1}}{Y_0} = \frac{\pi FBW}{2} \frac{1}{\sqrt{g_j g_{j+1}}} \quad j=1 \text{ to } n-1 \quad (7.2)$$

$$\frac{J_{n,n+1}}{Y_0} = \sqrt{\frac{\pi FBW}{2 g_n g_{n+1}}} \quad (7.3)$$

where $g_0, g_1 \dots g_n$ are the element of a ladder-type lowpass prototype, and FBW is the fractional bandwidth of bandpass filter. $J_{j,j+1}$ are the characteristic admittances of J-inverters and Y_0 is the characteristic admittance of the terminating lines.

The parallel-coupled resonators are mapped onto folded U-shape resonators as shown in Fig. 7.1. The design arranges the adjacent resonators into a parallel scheme

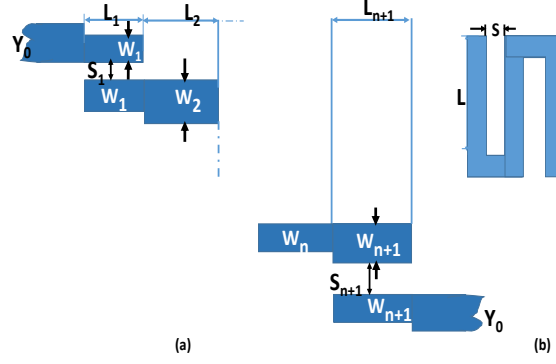


Figure 7.1: Structure of half-wavelength (a) Parallel-coupled resonator and (b) U-Shape resonator .

where coupling is maximized and hence makes it suitable for high-bandwidth applications. The hairpin resonators are U-shape versions of the parallel resonators which resulted into a compact filter structure. However, since the resonators are folded in the hairpin design, the resonator length is reduced. Also, the arms of adjacent resonators behave as coupled lines [101]. Therefore, there is a subtle differences between the electrical behavior of parallel resonators and U resonators that necessitate the usage of full-wave EM simulation and iterative fine tuning in order to meet the target specs [102].

The design equations of hairpin filter are give by [103]

$$Q_{e1} = \frac{g_0 g_1}{FBW}, \quad (7.4)$$

$$Q_{en} = \frac{g_n g_{n+1}}{FBW} \quad (7.5)$$

$$M_{i,i+1} = \frac{FBW}{\sqrt{g_i g_{i+1}}} \quad i=1 \text{ to } n-1 \quad (7.6)$$

Where Q_{e1} and Q_{en} are the external quality factor of the resonator at the input and output. $M_{i,i+1}$ are the coupling coefficients between the adjacent resonators and n is the filter order.

The location of the tap points connecting to input and output ports may be

approximated using [104]

$$\frac{t_{in}}{\lambda_g} = \frac{1}{2\pi} a \sin \left[\sqrt{\frac{\pi \Delta}{2g_0 g_1} \frac{Z_0}{Z_r}} \right] \quad (7.7)$$

$$\frac{t_{out}}{\lambda_g} = \frac{1}{2\pi} a \sin \left[\sqrt{\frac{\pi \Delta}{2g_N g_{N+1}} \frac{Z_0}{Z_r}} \right] \quad (7.8)$$

It is fairly straightforward to achieve hairpin filter with 20-30% fractional bandwidth (FBW). In theory, hairpin filter can achieve FBW of 50% and more. However for a higher FBW, the design becomes very difficult and sensitive to loss, metal etching, and roughness.

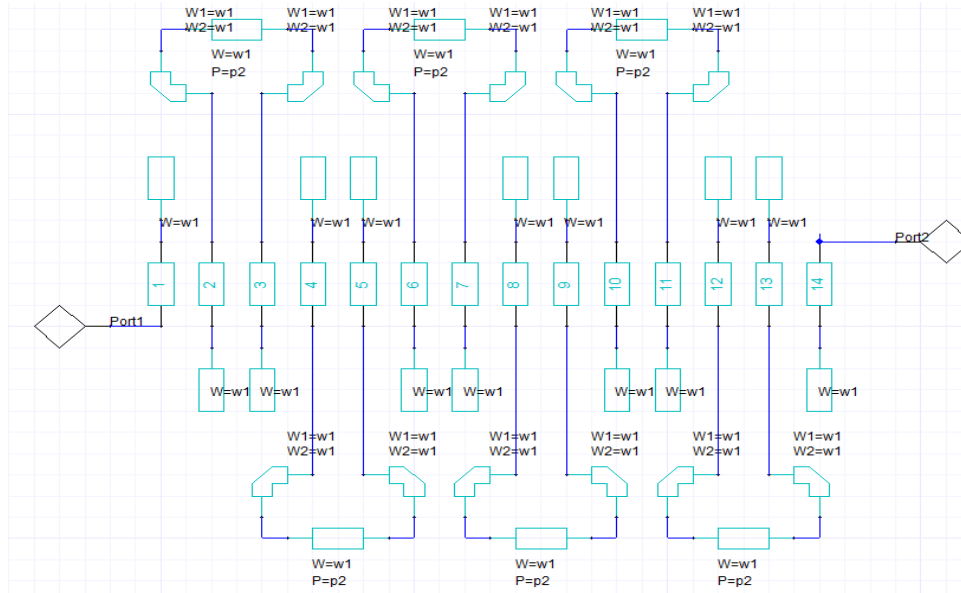
We designed a six-order bandpass hairpin filter that operates at the X-band. The center frequency is at 9.0 GHz with target bandwidth of 4 GHz (45% FBW). The design parameters are summarized in Table. 7.1

Table 7.1: Hairpin filter design parameters

Parameter	Value
Dielectric	Rogers RT/Duroid 5880
Dielectric Constant ϵ	2.2
Loss Tangent	0.0009/@ 10GHz
Substrate height	0.508 mm (20 mil)
Center Frequency	9.0 GHz
Fractional BW target	≥ 4.0 GHZ (45%)
Filter order	6

7.3 Hairpin Filter Performance

The design circuit schematic is shown in Fig. 7.2 and the ideal filter performance is shown in Fig. 7.3.



6-Section hairpin design schematic

Figure 7.2: Schematic of hairpin bandpass filter.

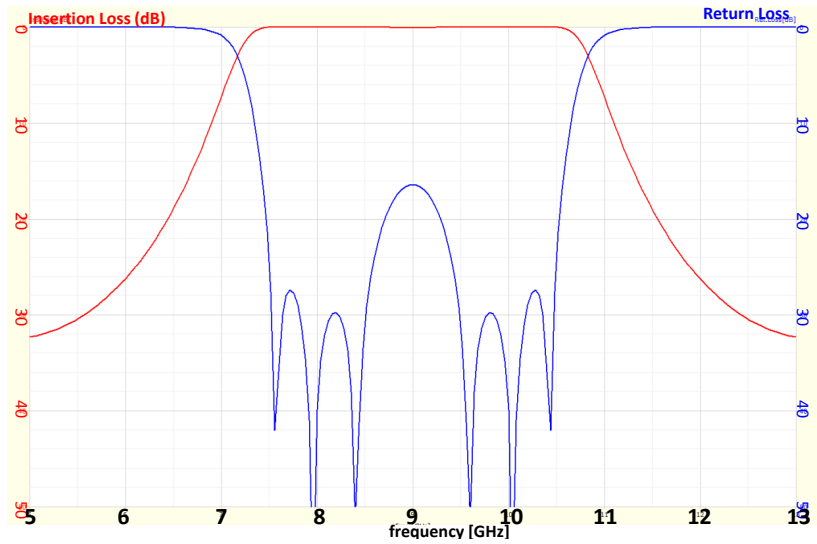


Figure 7.3: Performance of the ideal hairpin bandpass filter.

7.3.1 Full 3-D Model of the Hairpin Filter

Following this first step, we generate the layout, parameterized it and use full-wave EM tool to optimize the circuit. We used Ansys HFSS®3D full-wave electromagnetic simulator. The model is shown in Fig. 7.5.

At those frequencies, we need to model the effect of roughness and manufacturing etching on the channel performance [105–110]. We used Huray’s Model [109–111] which uses a snowball scheme to model copper surface roughness.

7.3.2 Roughness Modeling

Scanning electron microscope (SEM) image of PCB copper specimen exhibits a 3-D snowball profile of copper surface roughness [112]. Huray’s model improves on previous roughness models by taking into account roughness 3-D properties and SEM profiles. The model solves Maxwell equations into copper fine structures where he considers the copper foils as a 3-D pile of snowball-like spheres with different radius and numbers. A pictorial representation of SEM snowball properties of conductor foils is shown in Fig. 7.4. The equation of the loss model is:

$$\alpha_{snowball} = \frac{10dB(\frac{w}{9.4\mu m})}{2(8.14\mu m)} \sum_{i=1}^j \left(\frac{\sigma_{i,abs}(1)}{\sigma_{i,incoming}} \right) \left(\frac{6\pi a_i^2 N_i}{A} \right) \left[\frac{1 + \delta}{a_i + \frac{\delta^2}{2a}} \right]^{-1} \quad (7.9)$$

where $\alpha(1)$ is the first coefficient ($l=1$) of the spherical Hankel function expansion of \vec{E} and \vec{H} electric and magnetic fields solutions of Maxwell equations in the conductor using spherical coordinates. a_i is the average sphere radius, and N_i is the total number of snowballs.

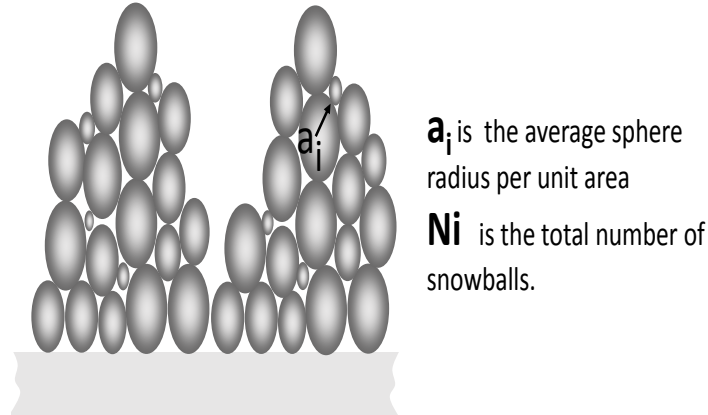


Figure 7.4: Huray's snowball model of surface roughness .

7.3.3 The Hairpin Filter Performance

The performance of the hairpin filter is summarized in Fig. 7.5. We notice the difference between the equation-based model of Fig. 7.3 and the full-wave 3D model shown in Fig. 7.5. Considering the fact that the FBW target is 45%, we consider the return loss to be acceptable within the passband as it is larger than 10 dB from 7.3-10 GHz. On the other hand, the insertion loss shows a substantial attenuation larger than 7dB at the higher end of the passband (from 9.5-10 GHz). It is between 5-6 dB loss from 6.5 GHz through 9.4 GHz.

This attenuation will manifest itself as a collapse of the outer symbols of the constellation diagram and as excessive BER degradation as we will see in the sections 7.6. This will require amplification at the receiving filter to correctly recover the symbols. The eye diagram also shows degradation and overlap between adjacent symbol eyes.

The combined impact of roughness and manufacturing etching (we used a conservative 5% etch factor) account for a loss of about 1 dB across the bandpass.

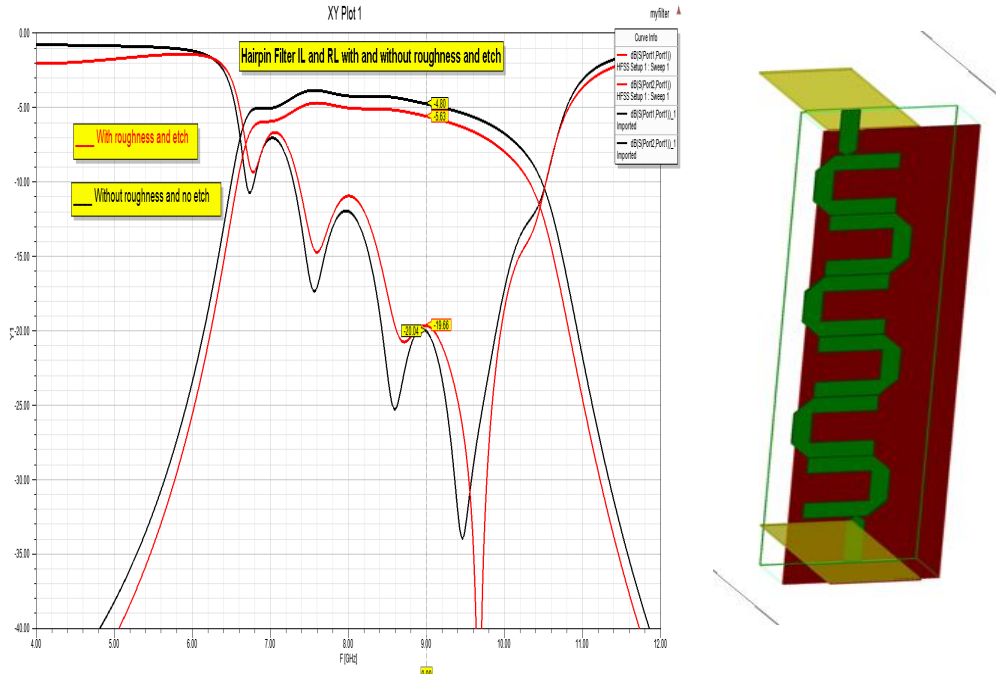


Figure 7.5: Full wave simulation and 3D model of hairpin filter captured from Ansoft HFSS.

7.4 SIW Characterization

We presented the details of the design of the SIW circuit in section 5.2. We fabricated and assembled an SIW circuit. The circuit is shown in Fig. 7.6. The coupling to the input and output ports of the SIW is achieved with an optimized microstrip-to-waveguide taper. The optimization of the taper coupling had been discussed in section 3.5.2.

The characterization setup of the SIW and the measured $S_{11}(\text{dB})$, $S_{21}(\text{dB})$ results are shown in Fig. 7.7. The structure has a cut-off frequency at 4.7 GHz and a bandwidth of about 7.5 GHz.

The measured data and bandwidth of the SIW shows better performance than the hairpin filter. The hairpin BPF achieves a 3 GHz bandwidth. The SIW bandwidth is about 2.5 times higher than the hairpin BPF. The SIW insertion loss achieves

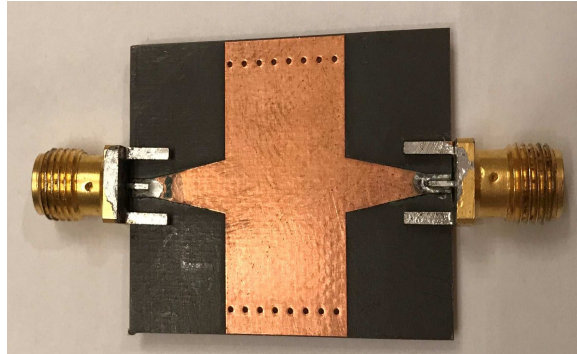


Figure 7.6: Fabricated SIW bandpass filter.

about 0.4 dB over the entire channel bandwidth, while the hairpin BPF insertion loss is between 5 dB to 6.5 dB across the bandwidth. Another advantage of the SIW is that the insertion loss is nearly perfectly flat over the entire bandwidth, which enhances uniformity and predictability of the channel system response. In the case of the hairpin filter, we expect performance to vary depending on where we land in the passband, which impacts uniformity and predictability of the response. We will quantify those effects in the system performance section.

7.5 MCMCA End-to-end System Model

The MCMCA system model in Simulink is shown in Fig. 7.8. The model consists of a transmitter block, a channel, a receiving block, a set of display/measurement outputs, and a carrier synchronizer as needed.

At both ends of the channel, there are up-converter and down-converter blocks that model the mixing and conversion electronic needed to leverage the channel bandpass.

The transmitter consists of a random integer generator, an integer to bit converter, a QAM modulator and a raised cosine transmit filter. The modulator maps the bits



Figure 7.7: SIW measurement.

into symbols. We use gray coding for our model. The 64-QAM modulation results in 6 bits per symbol. The raised cosine filter optimally shapes the symbol using Nyquist filter to cancel intersymbol interference. The filter has a roll-off parameter α which accounts for the excess bandwidth of the symbol beyond the theoretical minimum.

The channel model consists of the S-parameters of the bandpass filter in use. We alternate the model between the measured model of the SIW and the 3D EM model of the hairpin filter.

The receiver performs the reverse operation of the transmitter. It comprises a receive raised cosine filter and a QAM demodulator. The RF section consists of an up-converter and down converter operating at carrier frequency of 9.0 GHz.

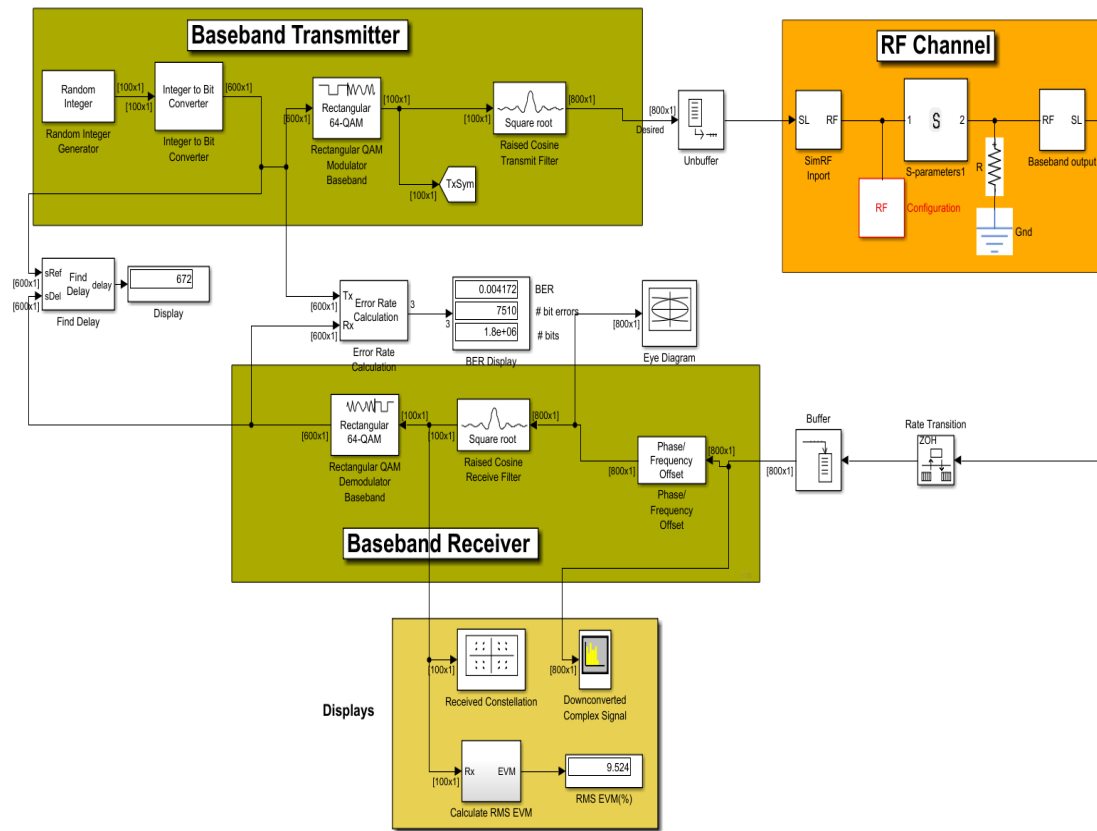


Figure 7.8: Full channel Simulink system model of MCMCA.

7.6 System Performance at Different Data Rates

A random integer stream is generated using a random bit generator. The integers are converted into a bit stream with the appropriate transfer rate. The bits are fed into a QAM modulator that is programmed to use gray coding to minimize bit errors [87]. The bits are converted into symbols where each symbol represents 6 bits of data. The aggregated throughput is the product of the bit rate times the QAM number of bits/symbol times 2 for double data rate memory protocol (DDR transfer data on both rising and falling edge).

The raised cosine filter shapes the symbol optimally to minimize intersymbol interference. The filter uses an excess bandwidth rolloff factor ranging from 0.1 to 0.3 for the 100 MHz case. A higher excess bandwidth factor is used for higher bit rates in order to achieve acceptable BER and eye diagram. The filtered symbol is then up-converted and modulated onto RF carrier of 9.0 GHz, which represents the mid frequency point of our bandpass X-band filter. At the other end of the channel, the symbol is down-converted and sent through a receive raised cosine filter that performs exactly the reverse of the transmit filter. The filtered symbol is then fed into a QAM demodulator which demodulates the symbol and convert it back into its original base-band bit format. The received signal is then compared against the transmitted signal after removing the appropriate path delay and the BER is displayed.

In performing the experiment, we used exclusively a simple static phase offset correction to the channel distortion. This is motivated by the fact that the purpose of the research is to evaluate the performance of both channels with minimal error correction algorithms.

7.6.1 System Performance at 100 MHz

At 100 MHz, the SIW-based channel yields a BER of $4.2e^{-4}$, when the rolloff excess factor α is equal to 0.1 and an EVM of 6.9%. The constellation diagram and the eye diagram are shown in Fig. 7.9. When we swap the channel and use the hairpin S-parameter model, the system performance shows a substantial degradation.

The EVM of the TLHPF-based channel at 100 MHz bit rate is 9.1%. A 32% degradation relative to the SIW-based channel. The BER reads $7.6e^{-3}$. Almost one order of magnitude higher than the SIW case. The constellation diagram and the eye diagram are shown in Fig. 7.10.

When the excess bandwidth factor is changed to 0.2 and 0.3, the EVM for the

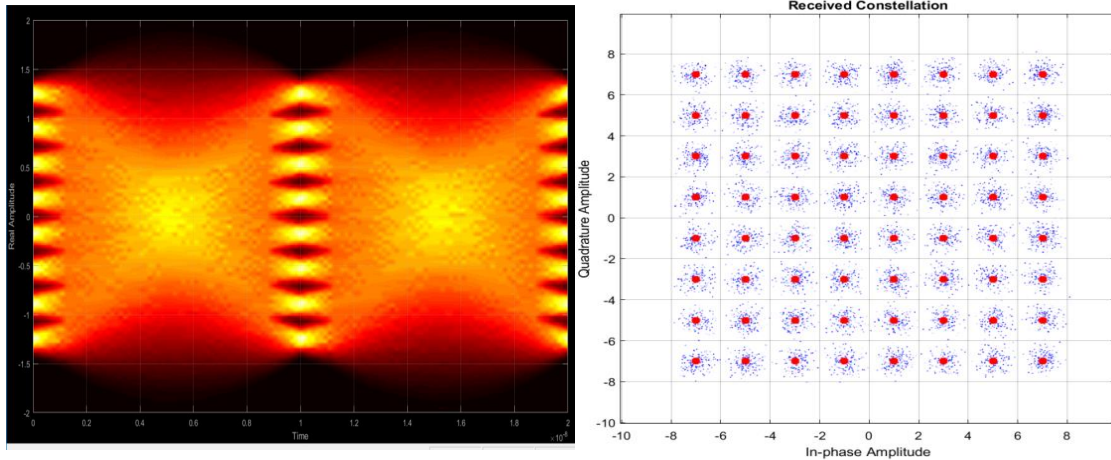


Figure 7.9: Eye diagram and constellation diagram of MCMCA system using SIW at 100 MHz data rate.

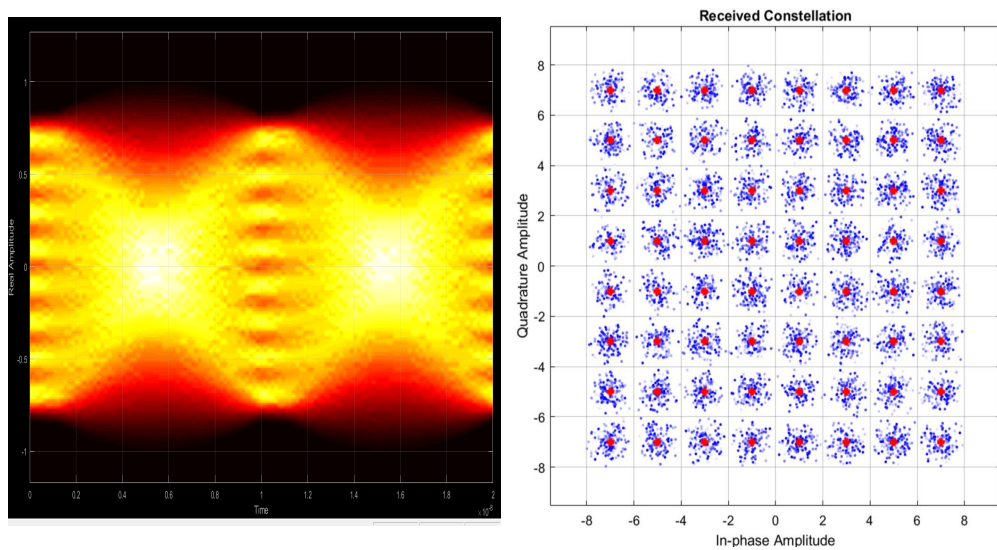


Figure 7.10: Eye diagram of MCMCA using TLBPF at 100 MHz and rolloff=0.2.

TLBPF-based system shows some improvement and yields 7.6% and 7% respectively. The BER shows a noticeable improvement of one order of magnitude improvement for every 0.1 increase of rolloff factor. BER is $3.2e^{-4}$ for $\alpha = 0.2$ and $3.5e^{-6}$ for when $\alpha = 0.3$. This is two orders of magnitude improvement. If compared to SIW-based

channel, things look otherwise different. EVM degrades by 73% and 75% for rolloff of 0.2 and 0.3 respectively. BER for SIW-based channel is 0 for rolloff of 0.2 and 0.3. On the other hand BER for TLBPF-based channel is $3.2e^{-4}$ and $3.5e^{-6}$ respectively.

Table 7.2: Performance of 64-QAM channels at 100 MHz

64-QAM Channel Performances at 100 MHz			
		SIW Channel	TLBPF Channel
EVM[%]	$\alpha = 0.1$	6.9	9.1
	$\alpha = 0.2$	4.4	7.6
	$\alpha = 0.3$	4	7
BER	$\alpha = 0.1$	$4.4e^{-4}$	$7.6e^{-3}$
	$\alpha = 0.2$	0	$3.2e^{-4}$
	$\alpha = 0.3$	0	$3.5e^{-6}$

The table 7.2 summarizes both channels performances at 100 MHz across different settings. The quality of the eye diagram can be verified visually to show a much better eye diagram opening patterns for the SIW-based system.

Table 7.3: Performance of 64-QAM channels at 200 MHz

64-QAM Channel Performance at 200 MHz			
		SIW Channel	TLBPF Channel
EVM[%]	$\alpha = 0.1$	11	See footnote ^a
	$\alpha = 0.2$	8.6	See footnote ^a
	$\alpha = 0.3$	8	See footnote ^a
	$\alpha = 0.8$	-	7.2
BER	$\alpha = 0.1$	0.013	See footnote ^a
	$\alpha = 0.2$	0.02	See footnote ^a
	$\alpha = 0.3$	$3.0e^{-4}$	See footnote ^a
	$\alpha = 0.8$	-	0

^aChannel failed to achieve an acceptable performance

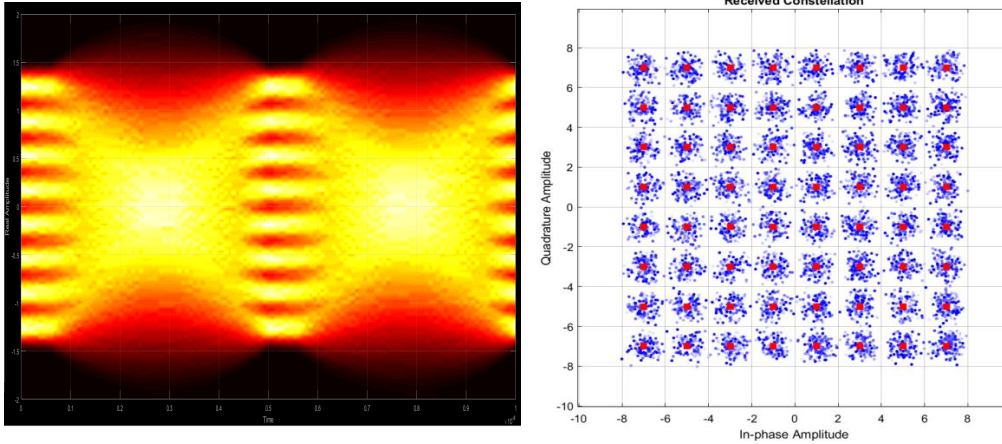


Figure 7.11: Constellation diagram and eye diagram of MCMCA using SIW at 200 MHz and rolloff=0.3.

7.6.2 System Performance at 200 MHz

At 200 MHz, the SIW channel achieves low EVM numbers and good opening of the eye diagram. The BER though is high for rolloff factor of 0.1 and 0.2.

When the rolloff factor is set to 0.3, the BER drops down into the e^{-4} level. The EVM is 8% and the eye diagram shows a high quality eye height and eye width as can be seen in Fig. 7.11

On the other hand, the TLBPF filter fails to achieve an acceptable performance at any of those rolloff factors. The constellation diagram and the eye diagram are of non acceptable quality. It takes a rolloff factor of 0.8 for the TLBPF channel to yield a good constellation and eye diagrams. Eye diagram for TLBPF channel for 200 MHz with $\alpha = 0.8$ is shown in Fig. 7.12.

Therefore, although TLBPF-based channel numbers improve greatly as we increase the excess bandwidth factor, it shows a large disadvantage compared to SIW-based channel.

This is attributable to the difference in the quality of insertion loss and return loss

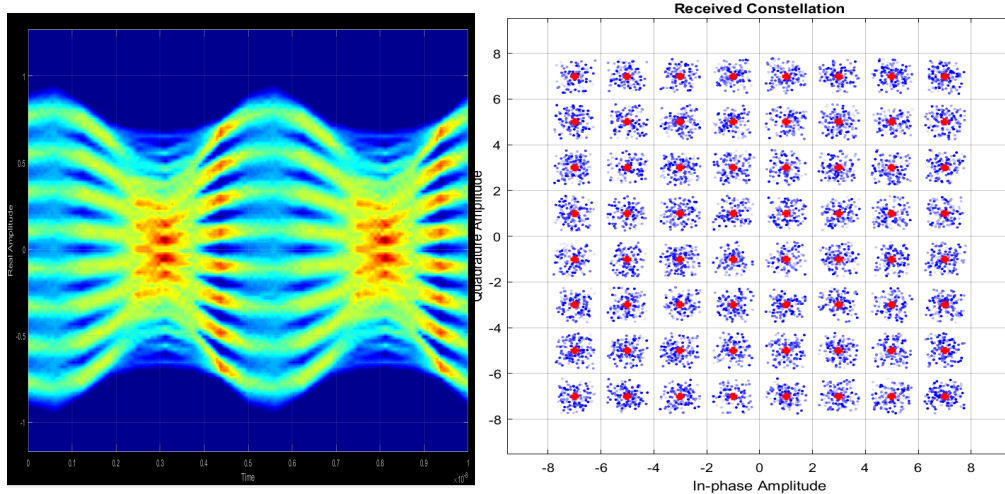


Figure 7.12: Eye diagram and constellation diagram of MCMCA using TLBPF at 200 MHz and rolloff=0.8.

curves observed earlier. The impact of roughness and conductive loss in the TLBPF channel starts to show up as system degradation when we push the data rate into the 100 MHz and above. The TLBPF-based channel can still provide a conductive channel with good BER and EVM performance numbers but at the cost of more complicated transmit and receive raised cosine filters.

7.7 System Performance at 250 MHz and 400 MHz

While the TLBPF channel saturates at 200 MHz, the SIW channel can run at higher bit rates. We run the SIW channel at 250 MHz and 400 MHz. The results are summarized in Table 7.4 and in Table 7.5 respectively

At 250 MHz with $\alpha = 0.3$, the channel achieves EVM of 9.7% and BER of $5.6e^{-3}$. The constellation diagram have good visual quality as can be seen in Fig. 7.13.

At 400 MHz, we need to increase the rolloff factor to 0.7 in order to achieve acceptable performance. At those settings, the channel BER is $8.3e^{-3}$ and the EVM

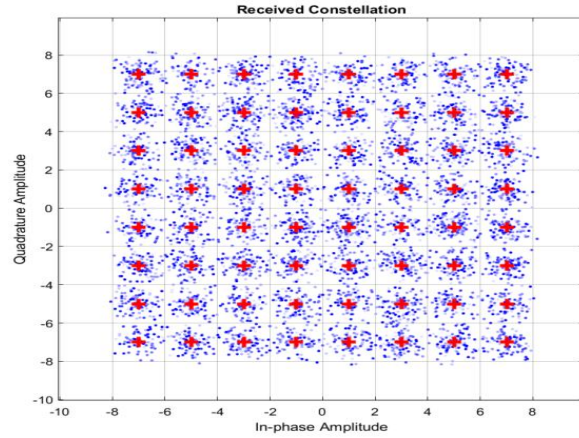


Figure 7.13: Constellation diagram of MCMCA using SIW at 250 MHz and rolloff=0.3.

is 10.7.

The eye diagram and the constellation diagram are shown Fig. 7.14. For a bit rate higher than 400 MHz, the channel would require addition of recovery algorithms and noise cancellation circuits.

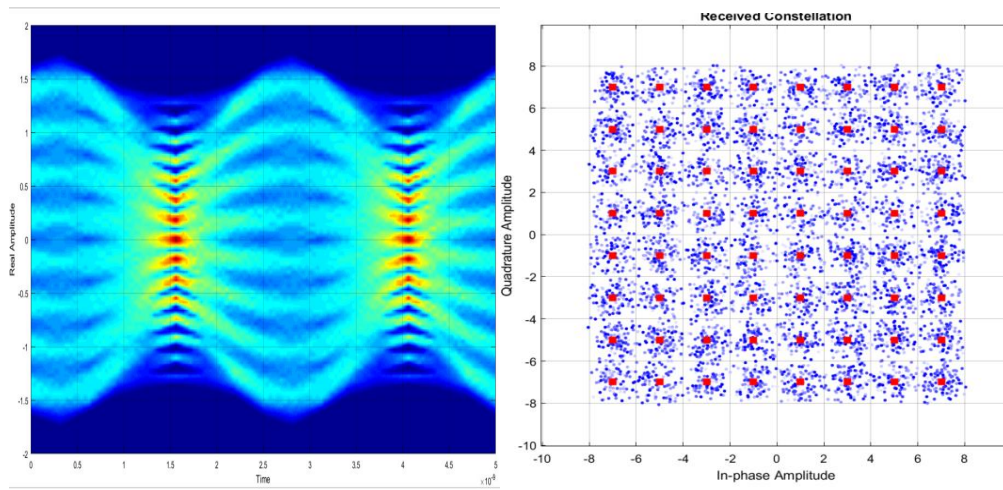


Figure 7.14: Constellation and eye diagram of MCMCA using SIW at 400 MHz and rolloff=0.7.

Table 7.4: 250 MHz performance

64-QAM Channel Performances at 250 MHz		
	SIW Channel	
EVM [%]	$\alpha = 0.3$	9.7
BER	$\alpha = 0.3$	$5.6e^{-3}$

Table 7.5: 400 MHz performance

64-QAM Channel Performances at 400 MHz		
	SIW Channel	
EVM [%]	$\alpha = 0.7$	10.7
BER	$\alpha = 0.7$	$3.3e^{-3}$

7.8 256-QAM Modulation of MCMCA

Since higher order modulation transmits more bits/symbol, it could be beneficial for the channel throughput when the noise level is low.

For higher order modulation though, the spacing between the symbols gets smaller which increases the BER for a given noise level. Therefore, simulation is needed to determine the optimal combination of the modulation order and the transfer rate. In

Table 7.6: Performance of 256-QAM MCMCA channel at 100 MHz

256-QAM Channel Performances at 100 MHz			
		SIW Channel	HPF Channel
EVM [%]	$\alpha = 0.4$	3.6	See footnote ^a
	$\alpha = 0.8$		6
BER	$\alpha = 0.4$	$5.4e^{-5}$	See footnote ^a
	$\alpha = 0.8$		0.02

^aChannel failed to achieve an acceptable performance

this section, we simulate the channel with 256-QAM modulation scheme at different speeds and determine the trade-off between higher modulation, bit rate and channel throughput.

For 256-QAM modulation, each symbol represents 8 bits of data while the distance

between adjacent symbols on the I-Q plane gets shorter.

Table 7.7 summarizes SIW performance at 256-QAM modulation for 200 MHz and 250 MHz. The table does not include TLBPF channel since TLBPF filter fails to deliver symbols beyond 100 MHz bit rate. The level of attenuation incurred in the TLBPF filter caused by roughness and conductive loss causes the channel to become noisy for 256-QAM with data rate bigger than 100 MHz. The symbols overlap each others and the BER becomes high enough to prohibit an acceptable data transmission.

For SIW, with a rolloff factor of 0.8, the BER at 200 MHz is $9.0e^{-5}$ and EVM of 4.2%. For 250 MHz, the BER is $5.0e^{-3}$ and EVM is 5.6.

The eye diagram and constellation diagram for SIW channel at 256-QAM with 250 MHz bit rate are shown in Fig. 7.15.

To compare the different modulations and transfer speeds we explored, let's summarize the results obtained so far. The 64-QAM 400 MHz enables 4800 Mbps ($400\text{MHz} \times 6 \text{ bits/symbol} \times 2 \text{ bits/period}$ for double data rate), while 256-QAM 250 MHz delivers 4000 Mbps ($250\text{MHz} \times 8 \text{ bits/symbol} \times 2 \text{ bits/period}$). In order for the 256-QAM to be advantageous, we need to push the data rate higher than 300 MHz, which will require equalization, and then we need to evaluate the pros and cons accordingly. With no addition of symbol recovery circuits and algorithms, 64-QAM delivers higher data rates than 256-QAM 250 MHz scheme.

Table 7.7: SIW channel performance at 256-QAM 200 MHz and 250 MHz

SIW Channel Performance at 200 MHz and 250 MHz		
200 MHz and $\alpha = 0.8$	EVM [%]	4.2
	BER	$9.0e^{-5}$
250 MHz and $\alpha = 0.8$	EVM [%]	5.6
	BER	$5.0e^{-3}$

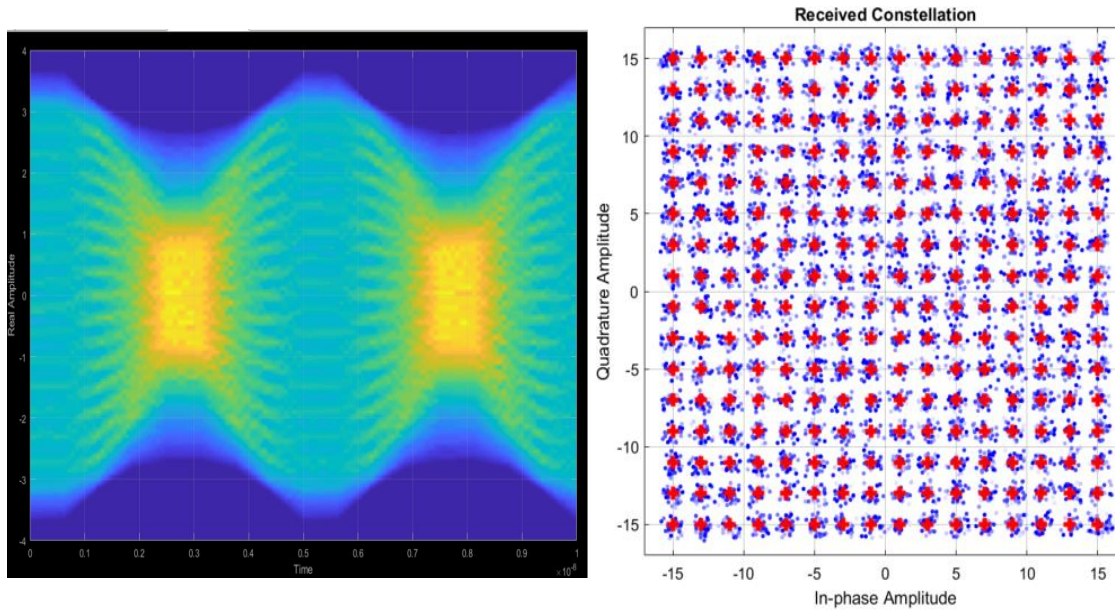


Figure 7.15: Constellation and eye diagram of MCMCA using SIW at 256-QAM 250 MHz and rolloff=0.8

7.9 Power Performance of the Channels

The matching quality of the channel determines its power performance and the power loss of the whole channel. S_{11} of SIW channel is shown in Fig. 7.7, while S_{11} of HPF channel is shown in Fig. 7.3. To measure the full channel power performance, we inserted two power spectrum measurement scopes, one at the RF input block right after the transmitter filter and one at the output of the RF block in front of the receiver filter. The power spectrum of the output of the RF network for 64-QAM at 200 Mbps is shown in 7.16.

The matching network consists of a series resistance at the input and a shunt load resistance at the output. For a perfectly matched network, each matching network will cause a 3 dB power division. If the channel is perfectly matched, the power

Table 7.8: SIW channel power performance

		alpha	OCBW	BB pwr	Rx Pwr	Power Loss
64-QAM	100 SIW					
	100 HPF		107	105	7.3	-11.58
	200 SIW	0.3	226	102	25	-6.11
	200 HPF	0.8	290	106	7.15	-11.71
	250 SIW	0.3				
256-QAM	SIW 400	alp=0.7	560	80	19	-6.24
		alp=0.8	585	76	14	-7.35 dB
	100 SIW	alp=0.4	120	422	101	-6.21 dB
	100 HPF	alp=0.8		428	28.7	-11.73 dB
	SIW 200			412	100	-6.15 dB
	SIW 250	alp=0.8	310	414	100.2	-6.16 dB

delivered to the receiver should be 1/4 of the available power.

The measured power at the receiver for the SIW channel at different bit rates ranges from -6.1 dB to -7.35 dB which is very close to ideally matched network. On the other hand, the power delivered to the HPF receiver ranges from -11 dB to -12 dB. The measured powers and occupied bandwidth are summarized in the Table 7.8.

7.10 Discussion

We demonstrated the merit of SIW bandpass filter as the interconnect of choice for ultra-high speed memory channel compared to transmission line-based bandpass filter as typified by a hairpin bandpass filter of order 6. One advantage of SIW that stands out is its large bandwidth. At the X-band and higher frequencies, the SIW are intrinsically high bandwidth bandpass devices. In contrast, the transmission line-based bandpass filter presents paramount difficulties to achieve FBW higher than 30%.

Other advantage of the SIW is its low conductive loss. Since the currents travel

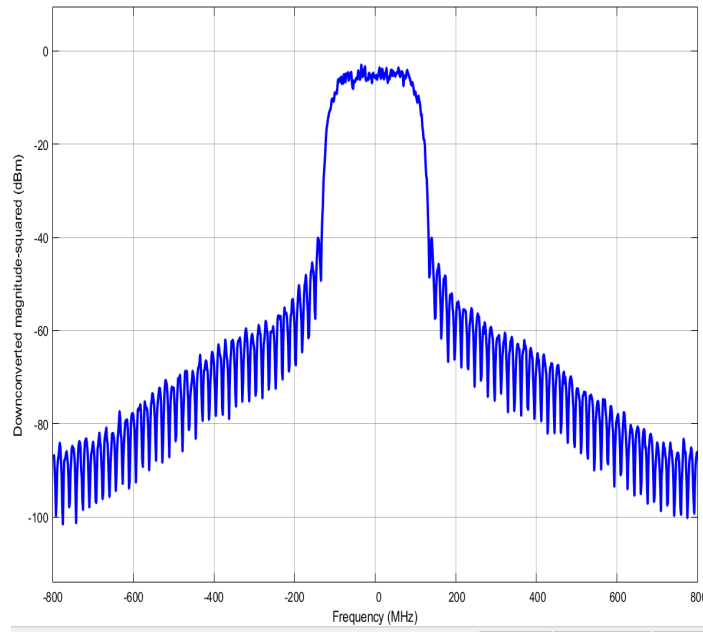


Figure 7.16: Symbol power spectrum of 64-QAM SIW channel at 200 Mbps and Rolloff=0.3

on the SIW surface, SIW is a large flood of metal, hence it has very low resistance and behaves distinctively better than transmission line when it comes to skin depth, roughness and manufacturing etching.

These advantages translate into robust BER, EVM, constellation diagram and eye diagram quality. These features turn out to be critical as the high order modulation schemes pile eye diagrams in a dense vertical arrangement and the low attenuation due to copper loss, low peak-to-peak jitter and large eye opening are of crucial importance.

Although transmission line-based filters look as though having the advantage of transmitting all the frequencies from DC all the way to a high RC attenuation level, a closer look reveals its limitation and inadequacy for ultra-high data rate memory channels. In fact, we strongly believe that memory protocols have to follow the path we pioneered, i.e. compact the symbols, wrap them via nearly Nyquist filters and

up-convert them to take advantage of microwave high bandwidth interconnect. This is simply due to the excessive attenuation, stringent skin effects and severe roughness and etching effects. SIW is the ideal candidate for ultra-high speed interconnect and this will remain till the optical interconnect resolves its compactness and compatibility with planar technology. This work demonstrated the benefits of SIW and laid the foundation for an interconnect testbench platform.

We demonstrated that the SIW interconnect can transmit memory signals at 4800 Mbps for double data rate memory protocol without the addition of signal recovery algorithms.

7.11 Conclusion

We demonstrated that SIW interconnect outperforms transmission line-based interconnect for multicarrier channels. We showed that a transmission line-based bandpass filter quickly runs out of bandwidth. In particular, it shows severe degradation due to conductive loss, skin effects, roughness, and etching. We used our MCMCA platform as a testbed for quantifying the effect of impairments on both SIW and TLBPF affected and concluded that TLBPF saturates at around 200 MHz transfer rate. On the other hand, SIW-based MCMCA reaches 400 MHz with a fixed phase compensation. The SIW-MCMCA is a promising architecture for high-data rate and we continue to optimize it. We believe that we can push the SIW-MCMCA throughput much higher with reasonable carrier and phase synchronization algorithms.

Chapter 8

OPTIMIZATION OF SIW INTERCONNECT USING DESIGN OF EXPERIMENT AND RESPONSE SURFACE METHOD

8.1 Introduction

In this chapter, we address the performance of the SIW as function of process and material fluctuations. We aim to develop a methodology that maximizes the robustness of the SIW with respect to parameter variabilities. We generate design of experiments and use response surface method to build a quadratic model that relates the input parameters and their cross-product to the system responses.

8.2 Substrate Integrated Waveguide Interconnect

In section 5.2, we detailed the 3D modeling and design of SIW circuit. We use the same design parameters for the analysis of impact of manufacturing on guide performance.

8.3 Design of Experiment of SIW

There have been many designs of SIW circuits reported in the literature targeting variety of applications[30, 36, 113]. The impact of process and manufacturing variability on the SIW though have not had similar attention if any. In contrast, there

is an abundant literature on impact of manufacturing variations on transmission line structures like microstrip, stripline and derivative structures.

As SIW getting more and more adoption in academia and commercial products, there is a need to quantify robustness and controlling parameters in minimizing SIW performance dependence on manufacturing fluctuations.

8.3.1 Design of Experiment Objective and Methodology

The design of experiment techniques provide a systematic method for sampling responses and constructing predictive models. The objective of DOE is to build a rigorous model identifying critical parameters and eliminating with high confidence input parameters that do not influence the responses. More importantly, the DOE enables SIW designers to make educated choices and decisions towards maximizing circuit performances by tightening tolerances and deploying technology options that have the highest impact on the performance. DOE method and DOE flow are shown in Fig. 8.1 and 8.2 respectively.

A quadratic response surface model contains the effects of a second degree polynomial fit and additionally the two-way interaction effects of the input variables.

For an experiment with n continuous independent variables, a quadratic response Y model would be:

$$Y = \beta_0 + \underbrace{\sum_{i=1}^n \beta_i X_i}_{\text{linear model}} + \underbrace{\sum_{i=1}^n \beta_{ii} X_i^2}_{\text{quadratic terms}} + \underbrace{\sum_{i,j \neq j}^n \beta_{ij} X_i X_j}_{\text{interaction terms}} \quad (8.1)$$

8.3.2 Response Surface Experiments for the SIW

The main SIW design parameters are listed in Table 5.1 We applied the RSM technique which generates the set of experiments listed in Table 8.3

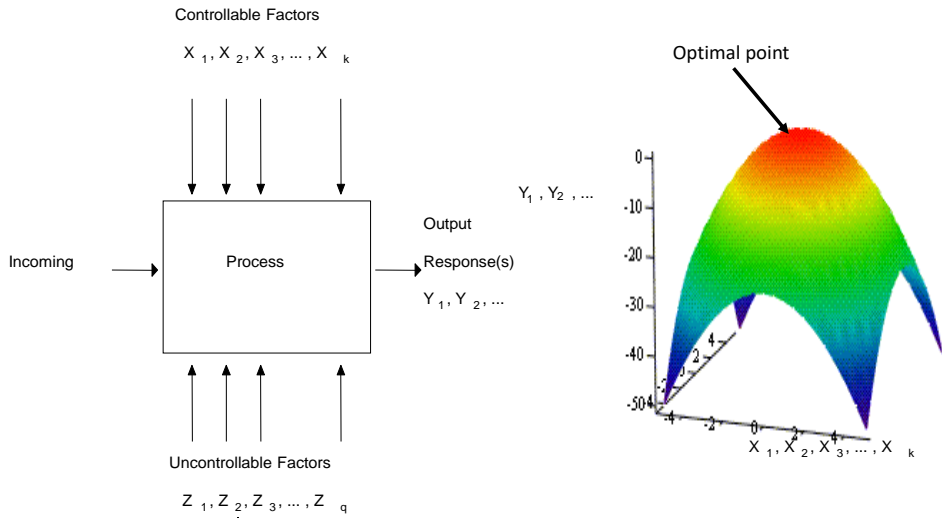


Figure 8.1: Response surface method

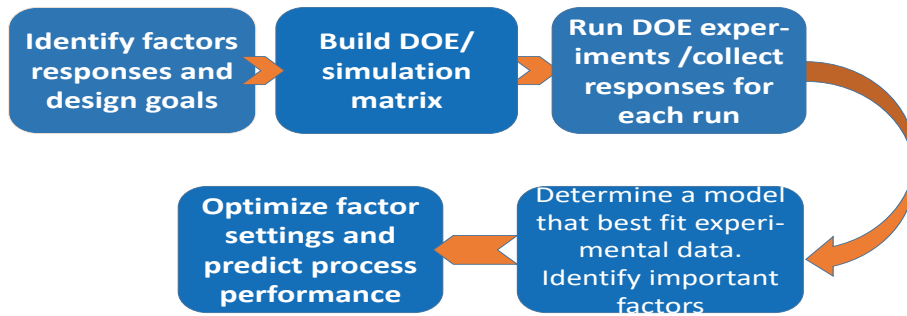


Figure 8.2: DOE flow

DOE factors

Six parameters are identified for the input factors: via radius r , via-to-via pitch d which controls the field leakage, dielectric permittivity ϵ_r , loss tangent (equally

Table 8.1: SIW design parameters

DOE Table	
Parameter	Range
Via radius r [mm]	0.18-0.21
via to via pitch d [mm]	1.3 - 1.7
Dielectric constant ϵ_r	2 - 2.4
Loss Tangent DF	0.0008 - 0.001
Guide Width W [mm]	10 - 12
Dielectric thickness h [mil]	25 - 35

known as dissipation factor) DF, guide width W , and the dielectric thickness h .

	r	d	Er	DF	h	W
1	0.18	1.3	2	0.0008	25	-12
2	0.18	1.3	2	0.0009	30	-10
3	0.18	1.3	2.2	0.001	35	-11
4	0.18	1.3	2.4	0.0008	35	-11
5	0.18	1.3	2.4	0.001	25	-10
6	0.18	1.5	2.2	0.0008	30	-10
7	0.18	1.5	2.2	0.0009	25	-11
8	0.18	1.5	2.4	0.0009	35	-12
9	0.18	1.7	2	0.0008	35	-11
10	0.18	1.7	2	0.001	30	-12
11	0.18	1.7	2.4	0.0009	25	-11
12	0.18	1.7	2.4	0.001	35	-10
13	0.195	1.3	2	0.0008	35	-10
14	0.195	1.3	2.4	0.0008	25	-10
15	0.195	1.3	2.4	0.0009	30	-11
16	0.195	1.3	2.4	0.001	25	-12
17	0.195	1.5	2	0.0009	35	-12
18	0.195	1.5	2	0.001	25	-11
19	0.195	1.5	2.2	0.0009	30	-11
20	0.195	1.5	2.2	0.0009	30	-11
21	0.195	1.5	2.2	0.0009	30	-11
22	0.195	1.7	2	0.0008	25	-10
23	0.195	1.7	2.2	0.0008	30	-12
24	0.21	1.3	2	0.001	30	-12
25	0.21	1.3	2.2	0.0008	35	-12
26	0.21	1.3	2.2	0.0009	25	-10
27	0.21	1.3	2.4	0.001	35	-10
28	0.21	1.5	2	0.0008	30	-11
29	0.21	1.5	2.4	0.0008	25	-12
30	0.21	1.7	2	0.0009	25	-12
31	0.21	1.7	2	0.001	35	-10
32	0.21	1.7	2.4	0.0008	35	-10
33	0.21	1.7	2.4	0.001	25	-10
34	0.21	1.7	2.4	0.001	35	-12

Figure 8.3: SIW RSM experiments

Response functions definition

We run the 34 DOE experiments and plotted the S_{11} and S_{21} family of waveforms. The results are captured in Fig. 8.4. We defined the bandwidth of the guide as the frequency range where $S_{11} \leq -10dB$. F_c is measured from the curve and recorded into the DOE table.

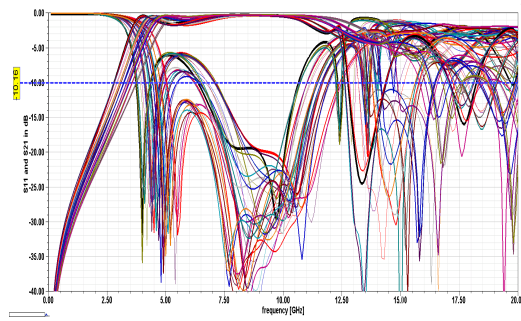


Figure 8.4: S-parameters of the SIW DOE experiments

Cutoff frequency F_c and the bandwidth BW are the responses and the DOE parameters listed in the table 8.3 are the input factors.

8.3.3 Bandwidth Fit Model

The bandwidth response model is shown in Fig.8.5. The fitting curve is a very good normal distribution and the statistics indicate a solid fit where all the important factors are included as detailed in the section below.

Quality of the fit and model interpretation

The quality of the fit is determined by the R^2_{value} , the R^2_{Adj} , and the normality of the residual plot.

1. R^2_{value} is the fraction of the total variability accounted for by the model. A good R^2_{value} value must be ≥ 0.9 . In the BW case, it reads 0.99, which indicates that the model predicts all the important variabilities and it includes all relevant input factors.
2. R^2_{Adj} is reduced when insignificant factors are added to the model. $R^2_{Adj} \geq 0.9$ is an indication that there are no insignificant parameters in the model. In our

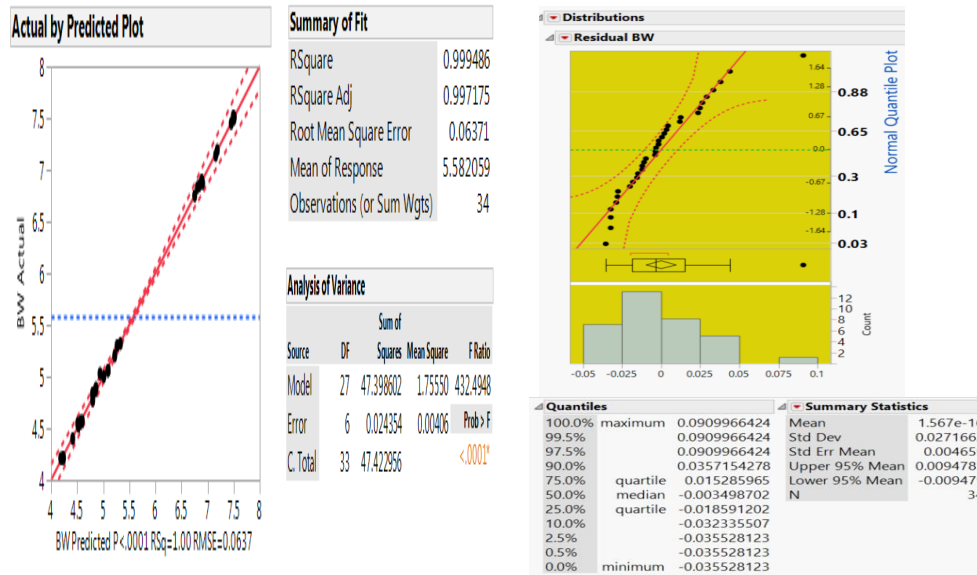


Figure 8.5: Bandwidth RSM model fit and residual distribution

case, the R_{Adj}^2 value is 0.99.

3. Residual plot: A normally distributed residual indicates that the model has accounted for all the significant parameters in the response.

The residual curve of bandwidth is shown in Fig. 8.5. All the points are within the normality boundary markers, and predominantly along the perfect normal line. The model is a very accurate representation of circuit variability that includes all relevant parameters.

Bandwith polynomial fitting model

The RSM builds an analytical model relating the factors to the responses and allows the designer to predict the response for the other input factors that are not included in the original RSM experiments. To illustrate this, we show below the linear portion of the model to be concise. The prediction model inherits the same accuracy as the

fit model and allows the designer to predict the response values without running long simulations.

$$\begin{aligned}
 BW = & 5.1 - 0.008 \times \left(\frac{Rvia - 0.195}{0.015}\right) + 0.01 \times \left(\frac{d - 1.5}{0.2}\right) - 0.23 \times \left(\frac{\epsilon_r - 2.2}{0.2}\right) \\
 & - 0.02 \times \left(\frac{DF - 0.0009}{0.0001}\right) - 0.013 \times \left(\frac{h - 30}{5}\right) - 1.36 \times (W + 11) \\
 & + \text{(square terms and cross terms quantifying the interaction between factors)}
 \end{aligned}
 \tag{8.2}$$

8.3.4 Cutoff Frequency Fit Model

The cutoff frequency response model is shown in Fig. 8.6. The fitting curve is a normal distribution and the statistics indicate a solid fit where all the important factors are accounted for.

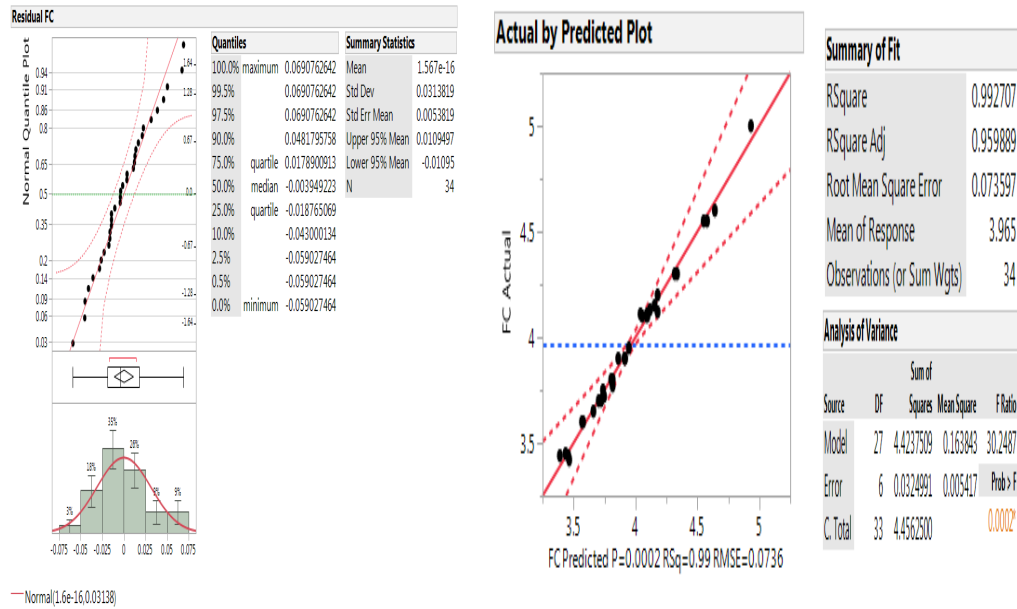


Figure 8.6: Cutoff frequency FC RSM model fit and residual distribution

Quality of the fit and model interpretation

The cutoff frequency model meets all the statistical quality criteria for a robust fit.

1. $R_{value}^2 = 0.992$ which indicates that the model predicts all the important variabilities and that it includes all relevant input factors.
2. $R_{Adj}^2 = 0.95$ which indicates that there are no insignificant parameter in the model.
3. The cutoff frequency residual curve is shown in Fig. 8.6. All the points are within the normality boundary markers, and predominantly along the perfect normal line. The model is a very accurate representation of circuit variability that includes all relevant parameters.

Cutoff frequency polynomial fitting model

The RSM builds the cutoff frequency predictive model. The model is composed of a linear fit portion, a square dependence on the input parameters and interaction terms between the different input factors. The linear portion of the model is shown in the equation below.

$$\begin{aligned}
 f_c = & 3.9 - 0.01 \times \left(\frac{Rvia - 0.195}{0.015} \right) - 0.03 \times \left(\frac{d - 1.5}{0.2} \right) - 0.21 \times \left(\frac{\epsilon_r - 2.2}{0.2} \right) \\
 & + 0.01 \times \left(\frac{DF - 0.009}{0.0001} \right) + 0.02 \times \left(\frac{h - 30}{5} \right) + 0.4 \times (h + 11) \\
 & + (\text{square terms and cross terms quantifying the interaction between factors})
 \end{aligned} \tag{8.3}$$

8.3.5 Cutoff Frequency and Bandwidth Prediction Profiler

The prediction profiler shown in the Fig. 8.7 is generated by the DOE RSM algorithm. It is a powerful graphical tool that allows the designer to interactively explore the solution space and determine the system response to input parameter variations. We can easily notice from the picture that the width fluctuation is the factor that has the most impact on the SIW performance. The second important factor is the

dielectric permittivity. The profiler shows that those two parameters impact the SIW BW and FC into opposite direction. The profiler enables the designer to lock some parameters and maximize the response by varying the desired set of parameters. This capability could turn out to be beneficial if the designer is faced by manufacturing limitation or if the alternative combination offers a lower cost solution for equivalent performance. The red colored numbers on the profiler figures shows the input parameter combination that yields the worst case BW and FC.

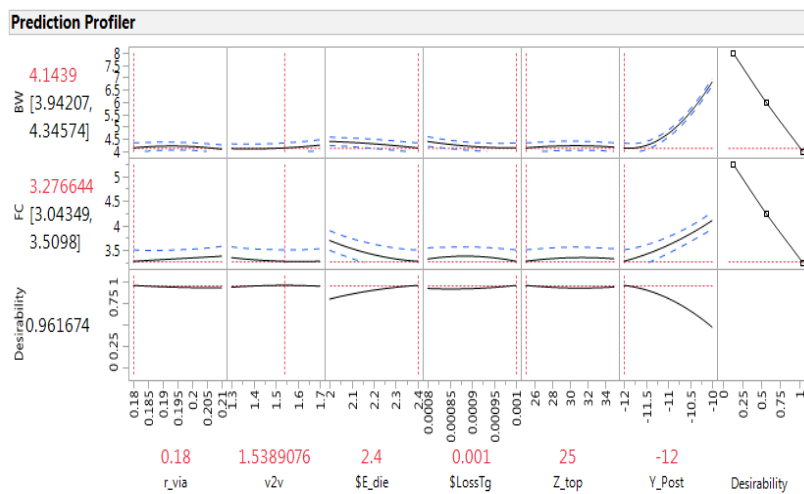


Figure 8.7: Prediction Profiler of the bandwidth and the cutoff frequency RSM models

8.4 Impact of Parameter Variations on Full Channel Performance

We demonstrated in the previous sections the impact of parameters variations on SIW circuit performance. We integrate the SIW into full channel system and evaluate the impact of those fluctuations on the system response.

The system consists of bit generator, a 64-QAM, a transmit root-cosine filter, an

RF up-converter, the SIW channel, a down-converter, a 64-QAM modulator, a root-cosine receive filter, a demodulator and a set of displays and measurement functions to measure the full channel performance. The end-to-end channel is shown in Fig. 7.8. The test consists of simulating the full channel end-to-end system using the worst case SIW channel and comparing the results to the system with the best case SIW channel.

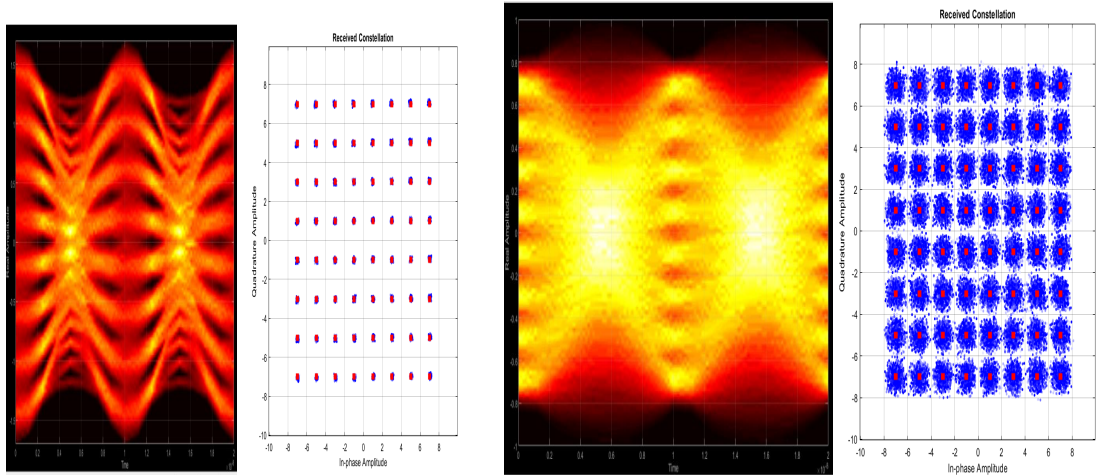
The performance of the channel with the best case SIW (case#1) is shown in the Fig. 8.8(a) while the Fig. 8.8(b) shows the performance of the channel with the worst case SIW (case #2).

At equal data rate (100 MHz) and equal channel setting, case#1 shows an obviously much better eye diagram quality than the case #2 which shows a blurry eye pattern. At this data rate, the best case SIW reaches a $BER = 0$, while the worst case SIW reaches a $BER = 2.7e^{-4}$. The constellation diagram of the case#1 shows the position of the states to land on the reference constellation states with small spread around the ideal position. In the case of the worst case SIW, the states have a large spread around the ideal position and the overlap with adjacent states can be seen on the graph. That explains the BER number of the case #2.

The channel with the best case SIW was able to run at higher data rate (400 MHz) without the need for equalization nor frequency and symbol recovery algorithms, while the case #2 fails to run at a speed higher than 100 MHz.

8.5 Conclusion

We implemented a systematic approach for optimizing the SIW interconnect performances by using a rigorous RSM DOE techniques. The generated models accurately predict the variabilities caused by manufacturing fluctuations and build analytical



(a) Best case SIW

(b) Worst case SIW

Figure 8.8: Eye and constellation diagram of the channel using (a) the best case SIW (b): the worst case SIW

response functions that account for the design parameters and the interactions among the critical inputs. The generated responses with the interactive prediction profiler enable the designer to trade off parameter tolerances for the performance as well as identifying the parameters with higher impact on the responses.

We showed that by controlling the critical parameters, the SIW performance can be improved by up to 40%. We compared the performance of best case SIW and worst case SIW in end-to-end channel and showed that the performance difference is significant. We also showed that the lower end SIW can impede the maximum transfer rate of the full channel, while the high-performance SIW can boost the throughput to higher rates.

Chapter 9

CONCLUSION

9.1 Conclusion

Memory interconnect is a challenging problem. The evolution of the DDR throughput has been, to some extent, monotonic and confined to the classical architecture with incremental improvements at each generation. Today's DDR channel is more or less the same as the one in the early days of computing system. There are many reasons explaining the quasi-static architecture of DDR channel. We discussed some of those reasons and constraints in different sections of this dissertation. One main factor we can enumerate is the pace in process scaling which has been much faster than innovation in architecture and novel interconnects. Another reason is the nature of the DDR bus as an intrinsically very wide bus. Any change to the protocol will have a multi-fold ripple effects on the subsystems and across the ecosystem.

We were among the first to advocate for a radical paradigm shift in the memory channel design. In order to close the gap between memory throughput and processing power, essentially lowering the "memory wall height", a completely different mind set needs to be deployed in rethinking both the architecture and the interconnect. In the last 3 to 4 years, we have begun to see a shift in the mentality of other parts of the industry towards reinventing the memory data transfer paradigm along the lines we have been exploring.

This dissertation is one stone in the complex endeavor of redesigning memory channels. We proposed and validated a novel interconnect based on SIW. The pro-

posal proved to be a very wide-band solution, with superior immunity to crosstalk and electromagnetic interference. We demonstrated that SIW preserves the benefits of waveguide interconnect while remaining compatible with planar PCB manufacturing. A feature that has a tremendous cost and manufacturing advantage.

In addition to the interconnect innovation, our contribution has an architectural component as well. In the architectural arena, we proposed to model the memory as a frequency division multiplexed channel. The drivers and receivers become transceiver based on baseband modulator/demodulator that set the memory to take full advantage of the wide bandwidth of the novel interconnect. The theoretical and experimental details of the proposed architecture have been outlined in this dissertation and the advantages have been quantified.

We addressed the manufacturing aspect of the proposal. Our DOE analysis using the response surface method enabled us to develop a flow that maximizes the channel performance and, nicely mitigates the effects of material and manufacturing disturbances.

9.2 Recommendations

The dissertation addressed the problem of memory throughput from a system and PCB perspective. Our recommendations for a natural continuation along the same line of the dissertation would be to address the problems of packaging.

Packaging accounts for more than 50% of the total loss incurred in a memory system. Any improvement to the packaging solution would have a substantial impact on the whole channel and processing system.

Deploying SIW into the packaging is a promising idea. The main challenges would be the dimensions of the SIW and the dense packaging environment. If SIW is made

smaller, the cutoff and the guiding frequencies will be high. That will result in a more involved design of the up and down converting circuits and filtering.

The density constraint will add another serious challenge for the designer. There is a need to deploy a meticulously designed compact SIW-based splitters and combiners inside the package. The manufacturing fluctuations will become of crucial importance. The designer will have to keep a tight control of the process in addition to the design of the SIW circuits.

Bibliography

- [1] Wm A Wulf and Sally A McKee. Hitting the memory wall: implications of the obvious. *ACM SIGARCH computer architecture news*, 23(1):20–24, 1995.
- [2] Gordon E. Moore. Cramming more components onto integrated circuits, reprinted from electronics, volume 38, number 8, april 19, 1965, pp.114 ff. *Solid-State Circuits Society Newsletter, IEEE*, 11(5):33–35, 2006.
- [3] Gordon E. Moore. Progress in digital integrated electronics. In *Electron Devices Meeting, 1975 International*, volume 21, pages 11–13, 1975.
- [4] <http://www.top500.org/lists/2015/06/>, 2015.
- [5] J. J. Dongarray, P. Luszczeky, and A. Petitetz. The linpack benchmark: Past, present, and future, 2001.
- [6] Yu Kunzhi, Cheng Li, Tsung-Ching Huang, A. Seyedi, Dacheng Zhou, C. Wilson, D.A. Berkram, S. Palermo, J.Q. Smela, M. Fiorentino, and R. Beausoleil. 56 Gb/s PAM-4 optical receiver frontend in an advanced FinFET process. In *Circuits and Systems (MWSCAS), 2015 IEEE 58th International Midwest Symposium on*, pages 1–4, 8 2015.
- [7] Bo Zhang, A. Nazemi, A. Garg, N. Kocaman, M.R. Ahmadi, M. Khanpour, Heng Zhang, Jun Cao, and A. Momtaz. A 195mW / 55mW dual-path receiver AFE for multistandard 8.5-to-11.5 Gb/s serial links in 40nm CMOS. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013 IEEE International*, pages 34–35, 2013.
- [8] E-Hung Chen, R. Yousry, and C.-K.K. Yang. Power optimized adc-based serial link receiver. *Solid-State Circuits, IEEE Journal of*, 47(4):938–951, 4 2012.
- [9] Mozhgan Mansuri, James E Jaussi, Joseph T Kennedy, Tzu-Chien Hsueh, Sudip Shekhar, Ganesh Balamurugan, Frank O’Mahony, Clark Roberts, Randy Mooney, and Bryan Casper. A scalable 0.128–1 tb/s, 0.8–2.6 pj/bit, 64-lane parallel i/o in 32-nm cmos. *IEEE Journal of solid-state circuits*, 48(12):3229–3242, 2013.
- [10] Semiconductor Industry Association (SIA). International technology roadmap for semiconductors 2013 report. Technical report, 2013.

- [11] Scott Kipp. storage networking roadmaps. IEEE 802.3 2.5 Gb/s and 5 Gb/s Backplane and Short Reach Copper Study Group, 2015.
- [12] J.T. Aberle and B. Bensalem. Ultra-high speed memory bus using microwave interconnects. In *Electrical Performance of Electronic Packaging and Systems (EPEPS), 2012 IEEE 21st Conference on*, pages 3–6, 2012.
- [13] B. Bensalem and J.T. Aberle. A new high-speed memory interconnect architecture using microwave interconnects and multicarrier signaling. *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, 4(2):332–340, 2 2014.
- [14] Brent Keeth and R. Jacob Baker. *DRAM Circuit Design*. IEEE Press, New York, 2001.
- [15] JEDEC. Jedec solid state technology association, 2018.
- [16] JEDEC. PC3-14900/PC3L-12800 1 rank x8 planar UDIMM SDRAM Unbuffered DDR3 240-pin DIMM Design Standard. Technical report, December 20, 2010.
- [17] JEDEC. PC2-3200/PC2-4200/PC2-5300/PC2-6400 DDR2 SDRAM unbuffered DIMM design specification, January 5, 2005.
- [18] JEDEC. PC3-12800 SPECIFICATION, 2006.
- [19] JEDEC. Wide I/O 2 (WideIO2), 2014.
- [20] HMC Consortium et al. Hybrid memory cube specification 2.1. Retrieved from *hybridmemorycube.org*, 2014.
- [21] JEDEC. *FBDIMM : Architecture and Protocol, JESD206. ISO*, 2007.
- [22] David M Pozar. *Microwave engineering*. John Wiley & Sons, 2009.
- [23] Clayton R. Paul. *Transmission Lines in Digital and Analog Electronic Systems: Signal Integrity and Crosstalk*. John Wiley & Sons, 2010.
- [24] Ian Glover and Peter M Grant. *Digital communications*. Prentice Hall, 1998.
- [25] S.H. Hall, G.W. Hall, and J.A. McCall. *High speed digital system design: a handbook of interconnect theory and design practices*. A Wiley-Interscience publication. Wiley, 2000.
- [26] Fuyumara Shigeki. Waveguide line; Japan Patent JP-06-053711, 2 1994.
- [27] Feng Xu and Ke Wu. Guided-wave and leakage characteristics of substrate integrated waveguide. *IEEE Transactions on microwave theory and techniques*, 53(1):66–73, 2005.

- [28] David E Senior, Xiaoyu Cheng, Melroy Machado, and Yong-Kyu Yoon. Single and dual band bandpass filters using complementary split ring resonator loaded half mode substrate integrated waveguide. In *Antennas and Propagation Society International Symposium (APSURSI), 2010 IEEE*, pages 1–4. IEEE, 2010.
- [29] Xiao-Ping Chen, Ke Wu, and Zhao-Long Li. Dual-band and triple-band substrate integrated waveguide filters with chebyshev and quasi-elliptic responses. *IEEE Transactions on Microwave Theory and Techniques*, 55(12):2569–2578, 2007.
- [30] M Almalkawi, L Zhu, and V Devabhaktuni. Dual-mode substrate integrated waveguide (siw) bandpass filters with an improved upper stopband performance. In *Infrared, Millimeter and Terahertz Waves (IRMMW-THz), 2011 36th International Conference on*, pages 1–2. IEEE, 2011.
- [31] Jan Schorer, Jens Bornemann, and Uwe Rosenberg. Comparison of surface mounted high quality filters for combination of substrate integrated and waveguide technology. In *Microwave Conference (APMC), 2014 Asia-Pacific*, pages 929–931. IEEE, 2014.
- [32] Mahbubeh Esmaeili, Jens Bornemann, and Peter Krauss. Substrate integrated waveguide bandstop filter using partial-height via-hole resonators in thick substrate. *IET Microwaves, Antennas & Propagation*, 9(12):1307–1312, 2015.
- [33] Yuan Dan Dong, Tao Yang, and Tatsuo Itoh. Substrate integrated waveguide loaded by complementary split-ring resonators and its applications to miniaturized waveguide filters. *IEEE Transactions on Microwave Theory and Techniques*, 57(9):2211–2223, 2009.
- [34] Natalia Leszczynska, Lukasz Szydlowski, and Jakub Podwalski. Design of substrate integrated waveguide filters using implicit space mapping technique. In *Microwave Radar and Wireless Communications (MIKON), 2012 19th International Conference on*, volume 1, pages 315–318. IEEE, 2012.
- [35] Yu Tang, Ke Wu, and Nazih Khaddaj Mallat. Development of substrate integrated waveguide filters for low-cost high-density rf and microwave circuit integration: Pseudo-elliptic dual mode cavity band-pass filters. *AEU-International Journal of Electronics and Communications*, 70(10):1457–1466, 2016.
- [36] Y Dong, C-TM Wu, and T Itoh. Miniaturised multi-band substrate integrated waveguide filters using complementary split-ring resonators. *IET microwaves, antennas & propagation*, 6(6):611–620, 2012.
- [37] Tarek Djerafi, Ke Wu, and Dominic Deslandes. A temperature-compensation technique for substrate integrated waveguide cavities and filters. *IEEE Transactions on Microwave Theory and Techniques*, 60(8):2448–2455, 2012.
- [38] Peng Chen, Guang Hua, De Ting Chen, Yuan Chun Wei, and Wei Hong. A double layer crossed over substrate integrated waveguide wide band directional coupler. In *Microwave Conference, 2008. APMC 2008. Asia-Pacific*, pages 1–4. IEEE, 2008.

- [39] SS Sabri, BH Ahmad, and AR Othman. Design and fabrication of x-band substrate integrated waveguide directional coupler. In *Wireless Technology and Applications (ISWTA), 2013 IEEE Symposium on*, pages 264–268. IEEE, 2013.
- [40] Tarek Djerafi, Jules Gauthier, and Ke Wu. Quasi-optical cruciform substrate integrated waveguide (siw) coupler for millimeter-wave systems. In *Microwave Symposium Digest (MTT), 2010 IEEE MTT-S International*, pages 716–719. IEEE, 2010.
- [41] Ritvik Srivastava, Soumava Mukherjee, and Animesh Biswas. Design of broadband planar substrate integrated waveguide (siw) transvar coupler. In *Antennas and Propagation & USNC/URSI National Radio Science Meeting, 2015 IEEE International Symposium on*, pages 1402–1403. IEEE, 2015.
- [42] Pei-Ling Chi and Tse-Yu Chen. Dual-band ring coupler based on the composite right/left-handed folded substrate-integrated waveguide. *IEEE Microwave and Wireless Components Letters*, 24(5):330–332, 2014.
- [43] R. Tiwari, S. Mukherjee, and A. Biswas. Design and characterization of multi-layer substrate integrated waveguide (siw) slot coupler. In *2015 9th European Conference on Antennas and Propagation (EuCAP)*, pages 1–4, 4 2015.
- [44] T. Djerafi, M. Daigle, H. Boutayeb, Xiupu Zhang, and Ke Wu. Substrate integrated waveguide six-port broadband front-end circuit for millimeter-wave radio and radar systems. In *Microwave Conference, 2009. EuMC 2009. European*, pages 77–80, 9 2009.
- [45] S Mann, S Erhardt, S Lindner, F Lurz, S Linz, F Barbon, R Weigel, and A Koelpin. Diode detector design for 61 ghz substrate integrated waveguide six-port radar systems. In *Wireless Sensors and Sensor Networks (WiSNet), 2015 IEEE Topical Conference on*, pages 44–46. IEEE, 2015.
- [46] Wu Li-nan, Zhang Xu-chun, Tong Chuang-ming, and Zhou Ming. A new substrate integrated waveguide six-port circuit. In *Microwave and Millimeter Wave Technology (ICMMT), 2010 International Conference on*, pages 59–61. IEEE, 2010.
- [47] Martin Dušek and Jiří Šebesta. Design of substrate integrated waveguide six-port for 3.2 ghz modulator. In *Telecommunications and Signal Processing (TSP), 2011 34th International Conference on*, pages 274–278. IEEE, 2011.
- [48] Y.J. Cheng and Y. Fan. Compact substrate-integrated waveguide bandpass rat-race coupler and its microwave applications. *Microwaves, Antennas Propagation, IET*, 6(9):1000–1006, 2012.
- [49] Ching-Kuang C Tzuang, Kuo-Cheng Chen, Cheng-Jung Lee, Chia-Cheng Ho, and Hsien-Shun Wu. H-plane mode conversion and application in printed microwave integrated circuit. In *Microwave Conference, 2000. 30th European*, pages 1–4. IEEE, 2000.

- [50] Ville S Mottonen. Wideband coplanar waveguide-to-rectangular waveguide transition using fin-line taper. *IEEE Microwave and Wireless Components Letters*, 15(2):119–121, 2005.
- [51] D. Deslandes. Design equations for tapered microstrip-to-substrate integrated waveguide transitions. In *Microwave Symposium Digest (MTT), 2010 IEEE MTT-S International*, pages 704–707, 2010.
- [52] D. Deslandes and Ke Wu. Integrated microstrip and rectangular waveguide in planar form. *Microwave and Wireless Components Letters, IEEE*, 11(2):68–70, 2 2001.
- [53] Hiroshi Uchimura, Takeshi Takenoshita, and Mikio Fujii. Development of a” laminated waveguide”. *IEEE Transactions on Microwave Theory and Techniques*, 46(12):2438–2443, 1998.
- [54] R.E. Collin. *Foundations for Microwave Engineering*. 2001.
- [55] R.E. Collin. The optimum tapered transmission line matching section. *Proceedings of the IRE*, 44(4):539–548, 1956.
- [56] D. Deslandes and Ke Wu. Accurate modeling, wave mechanism, and design consideration of a substrate integrated waveguide. *IEEE Trans. Microw. Theory Tech.*, 54(6):2516–2526, 6 2006.
- [57] Songnan Yang and A.E. Fathy. Synthesis of an arbitrary power split ratio divider using substrate integrated waveguides. pages 427–430, June 2007.
- [58] M. Bozzi, Feng Xu, D. Deslandes, and Ke Wu. *Modeling and Design Considerations for Substrate Integrated Waveguide Circuits and Components*. 2007.
- [59] M. Bozzi, M. Pasian, and L. Perregrini. Modeling of losses in substrate integrated waveguide components. In *2014 International Conference on Numerical Electromagnetic Modeling and Optimization for RF, Microwave, and Terahertz Applications (NEMO)*, pages 1–4, May 2014.
- [60] Kwang-Il Oh, Lee-Sup Kim, Kwang-Il Park, Young-Hyun Jun, Joo Sun Choi, and Kinam Kim. A 5-gb/s/pin transceiver for ddr memory interface with a crosstalk suppression scheme. *Solid-State Circuits, IEEE Journal of*, 44(8):2222–2232, Aug 2009.
- [61] Elliott Cooper-Balis. Buffer-on-board memory system, 2012.
- [62] B. Leibowitz, R. Palmer, J. Poulton, Y. Frans, S. Li, J. Wilson, M. Bucher, A.M. Fuller, J. Eyles, M. Aleksic, T. Greer, and N.M. Nguyen. A 4.3 gb/s mobile memory interface with power-efficient bandwidth scaling. *Solid-State Circuits, IEEE Journal of*, 45(4):889–898, 2010.
- [63] Yanghyo Kim, Sai-Wang Tam, Gyung-Su Byun, Hao Wu, Lan Nan, G. Reinman, J. Cong, and M.-C.F. Chang. Analysis of noncoherent ask modulation-based rf-interconnect for memory interface. *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, 2(2):200–209, 2012.

- [64] A.B. Kahng and V. Srinivas. Mobile system considerations for sdram interface trends. In *System Level Interconnect Prediction (SLIP), 2011 13th International Workshop on*, pages 1–8, 2011.
- [65] Micron. FBDIMM - Channel Utilization (Bandwidth and Power)- TN-47-21. Technical Note, Micron Technology, Inc., December, 2009.
- [66] J.A. Kash, F. Doany, D. Kuchta, P. Pepeljugoski, L. Schares, J. Schaub, C. Schow, J. Trehwella, C. Baks, Y. Kwark, C. Schuster, L. Shan, C. Patel, C. Tsang, J. Rosner, F. Libsch, R. Budd, P. Chiniwalla, D. Guckenberger, D. Kucharski, R. Dangel, B. Offrein, M. Tan, G. Trott, D. Lin, A. Tandon, and M. Nystrom. Terabus: a chip-to-chip parallel optical interconnect. pages 363–364, 10 2005.
- [67] Yin-Jung Chang, D. Guidotti, Lixi Wan, and Gee-Kung Chang. Hybrid interconnects using silicon/FR-4 substrates for board-level 10 Gb/s signal broadcasting. pages 161–165, 6 2006.
- [68] J.W. Goodman, F.J. Leonberger, Sun-Yuan Kung, and R.A. Athale. Optical interconnections for vlsi systems. *Proceedings of the IEEE*, 72(7):850–866, July 1984.
- [69] M.H. Nazari and A. Emami-Neyestanak. A 24-gb/s double-sampling receiver for ultra-low-power optical communication. *Solid-State Circuits, IEEE Journal of*, 48(2):344–357, 2 2013.
- [70] D.A.B. Miller, A. Bhatnagar, S. Palermo, A. Emami-Neyestanak, and M.A. Horowitz. Opportunities for optics in integrated circuits applications. In *Solid-State Circuits Conference, 2005. Digest of Technical Papers. ISSCC. 2005 IEEE International*, pages 86–87, 2 2005.
- [71] Gyung-Su Byun, Yanghyo Kim, Jongsun Kim, Sai-Wang Tam, and M.-C.F. Chang. An energy-efficient and high-speed mobile memory i/o interface using simultaneous bi-directional dual (base+rf)-band signaling. *Solid-State Circuits, IEEE Journal of*, 47(1):117–130, 2012.
- [72] K. Gharibdoust, A. Tajalli, and Y. Leblebici. 10.3 a 7.5mw 7.5gb/s mixed nrz/multi-tone serial-data transceiver for multi-drop memory interfaces in 40nm cmos. In *Solid- State Circuits Conference - (ISSCC), 2015 IEEE International*, pages 1–3, Feb 2015.
- [73] JEDEC Solid State Technology Association. Fbdimm advanced memory buffer (amb), 2009.
- [74] T. Alexoudi, S. Papaioannou, G.T. Kanellos, A. Miliou, and N. Pleros. Optical cache memory peripheral circuitry: Row and column address selectors for optical static ram banks. *Lightwave Technology, Journal of*, 31(24):4098–4110, 12 2013.

- [75] D. Brunina, Dawei Liu, and K. Bergman. An energy-efficient optically connected memory module for hybrid packet- and circuit-switched optical networks. *Selected Topics in Quantum Electronics, IEEE Journal of*, 19(2):3700407–3700407, March 2013.
- [76] D. Brunina, C.P. Lai, Dawei Liu, A.S. Garg, and K. Bergman. Optically-connected memory with error correction for increased reliability in large-scale computing systems. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, 3 2012.
- [77] Jamesina J. Simpson, Allen Taflove, Jason A. Mix, and Howard Heck. Substrate integrated waveguides optimized for ultrahigh-speed digital interconnects. *IEEE Transactions on Microwave Theory & Techniques*, 54(5):1983–1990, 05 2006.
- [78] V.P.R. Magri, M.M. Mosso, R.A.A. Lima, and J.F. Mologni. Fr-4 waveguide electronic circuits at 10 gbit/s. In *Microwave Optoelectronics Conference (IMOC), 2011 SBMO/IEEE MTT-S International*, pages 181–184, 10 2011.
- [79] A. Suntives and R. Abhari. Experimental evaluation of a hybrid substrate integrated waveguide. In *Antennas and Propagation Society International Symposium, 2008. AP-S 2008. IEEE*, pages 1–4, 7 2008.
- [80] A. Suntives and R. Abhari. Design and application of multimode substrate integrated waveguides in parallel multichannel signaling systems. *Microwave Theory and Techniques, IEEE Transactions on*, 57(6):1563–1571, 6 2009.
- [81] A. Suntives and R. Abhari. Dual-mode high-speed data transmission using substrate integrated waveguide interconnects. In *Electrical Performance of Electronic Packaging, 2007 IEEE*, pages 215–218, 10 2007.
- [82] A. Suntives, Arash Khajooeizadeh, and R. Abhari. Using via fences for crosstalk reduction in pcb circuits. In *Electromagnetic Compatibility, 2006. EMC 2006. 2006 IEEE International Symposium on*, volume 1, pages 34–37, 8 2006.
- [83] A. Suntives and R. Abhari. Transition structures for 3-d integration of substrate integrated waveguide interconnects. *Microwave and Wireless Components Letters, IEEE*, 17(10):697–699, 10 2007.
- [84] Vogt. P. Fully Buffered DIMM (FB-DIMM) Server Memory Architecture: Capacity, Performance, Reliability, and Longevity , Intel Developer Forum, February, 2004.
- [85] FBDIMM JEDEC. Advanced memory buffer (amb) <http://www.jedec.org/download/search>. *JESD82-20. pdf*.
- [86] Young-Sik Kim, Seon-Kyoo Lee, Seung-Jun Bae, Young-Soo Sohn, Jung-Bae Lee, Joo Sun Choi, Hong-June Park, and Jae-Yoon Sim. An 8GB/s quad-skew-cancelling parallel transceiver in 90nm CMOS for high-speed DRAM interface. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*, pages 136–138, Feb 2012.

- [87] J. G. Proakis. *Digital Communications*. McGraw-Hill Inc., New York, N.Y., 2001.
- [88] Samuel C. Yang. *OFDMA System Analysis and Design*. Artech House, 2010.
- [89] A. V. Oppenheim and A. S. Willsky. *Signals and Systems*. Prentice Hall; 2 edition., New York, N.Y., 1996.
- [90] Agilent Application Note. 1298. digital modulation in communications system-san introduction. *Hewlett-Packard Company*, 1997.
- [91] Agilent Technologies. *Digital Modulation in Communications Systems An Introduction*. Agilent AN 1298.
- [92] Hong-Yeh Chang, Pei-Si Wu, Tian-Wei Huang, Huei Wang, Chung-Long Chang, and J.G.J. Chern. Design and analysis of CMOS broad-band compact high-linearity modulators for gigabit microwave/millimeter-wave applications. *Microwave Theory and Techniques, IEEE Transactions on*, 54(1):20–30, 2006.
- [93] C. Loyez, A. Siligaris, P. Vincent, A. Cathelin, and N. Rolland. A direct conversion IQ modulator in CMOS 65nm SOI for multi-gigabit 60ghz systems. In *Microwave Integrated Circuits Conference (EuMIC), 2012 7th European*, pages 5–7, 2012.
- [94] Shou-Hsien Weng, Che-Hao Shen, and Hong-Yeh Chang. A wide modulation bandwidth bidirectional CMOS IQ modulator/demodulator for microwave and millimeter-wave gigabit applications. In *Microwave Integrated Circuits Conference (EuMIC), 2012 7th European*, pages 8–11, 2012.
- [95] B. Bensalem and J. T. Aberle. A new high-speed memory interconnect architecture using microwave interconnects and multicarrier signaling. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 4(2):332–340, Feb 2014.
- [96] Y. Cassivi, L. Perregrini, P. Arcioni, M. Bressan, K. Wu, and G. Conciauro. Dispersion characteristics of substrate integrated rectangular waveguide. *IEEE Microwave and Wireless Components Letters*, 12(9):333–335, Sept 2002.
- [97] M. Bozzi, L. Perregrini, and K. Wu. Modeling of radiation, conductor, and dielectric losses in SIW components by the bi-rme method. In *2008 European Microwave Integrated Circuit Conference*, pages 230–233, Oct 2008.
- [98] J. T. Bolljahn and G. L. Matthaei. A study of the phase and filter properties of arrays of parallel conductors between ground planes. *Proceedings of the IRE*, 50(3):299–311, March 1962.
- [99] G.L. Matthaei. *Microwave filters, impedance-matching networks, and coupling structures*. Number v. 1. McGraw-Hill, 1964.
- [100] U. H. Gysel. New theory and design for hairpin-line filters. *IEEE Transactions on Microwave Theory and Techniques*, 22(5):523–531, May 1974.

- [101] E. G. Cristal and S. Frankel. Hairpin-line and hybrid hairpin-line/half-wave parallel-coupled-line filters. *IEEE Transactions on Microwave Theory and Techniques*, 20(11):719–728, Nov 1972.
- [102] J. Ye, D. Qu, X. Zhong, and Y. Zhou. Design of X-band bandpass filter using hairpin resonators and tapped feeding line. In *2014 IEEE Symposium on Computer Applications and Communications*, pages 93–95, July 2014.
- [103] HONG J.-S. G. and M. J. LANCASTER. *Microstrip filters for RF microwave applications*. Wiley, New York, 2001.
- [104] J.T. Aberle. EEE 545 Microwave Circuit Design. University Lecture; Arizona State University, 2014.
- [105] E. Hammerstad and O. Jensen. Accurate models for microstrip computer-aided design. In *1980 IEEE MTT-S International Microwave symposium Digest*, pages 407–409, May 1980.
- [106] B. Curran, I. Ndip, and K. D. Lang. A comparison of typical surface finishes on the high frequency performances of transmission lines in PCBs. In *2017 IEEE 21st Workshop on Signal and Power Integrity (SPI)*, pages 1–3, May 2017.
- [107] B. Curran, I. Ndip, S. Guttowski, and H. Reichl. A methodology for combined modeling of skin, proximity, edge, and surface roughness effects. *IEEE Transactions on Microwave Theory and Techniques*, 58(9):2448–2455, Sept 2010.
- [108] M. Koledintseva, T. Vincent, and S. Radu. Full-wave simulation of an imbalanced differential microstrip line with conductor surface roughness. In *2015 IEEE Symposium on Electromagnetic Compatibility and Signal Integrity*, pages 34–39, March 2015.
- [109] P. G. Huray, S. Hall, S. Pytel, F. Oluwafemi, R. Mellitz, D. Hua, and Peng Ye. Fundamentals of a 3-D ”snowball”; model for surface roughness power losses. In *2007 IEEE Workshop on Signal Propagation on Interconnects*, pages 121–124, May 2007.
- [110] S. Hall, S. G. Pytel, P. G. Huray, D. Hua, A. Moonshiram, G. A. Brist, and E. Sijercic. Multigigahertz causal transmission line modeling methodology using a 3-D hemispherical surface roughness approach. *IEEE Transactions on Microwave Theory and Techniques*, 55(12):2614–2624, Dec 2007.
- [111] Paul G. Huray. *The Foundations of Signal Integrity*. Wiley-IEEE Press, 2010.
- [112] M. Schlesinger and M. Paunovic. *Modern Electroplating*. The ECS Series of Texts and Monographs. Wiley, 2011.
- [113] D. Deslandes, M. Bozzi, P. Arcioni, and Ke Wu. Substrate integrated slab waveguide (sisw) for wideband microwave applications, 2003.