Title

Phase-Based Algorithms for Communication and Signal Processing

Author

Li, Yilei

Publication Date

2016

University of California
Los Angeles

Phase-Based Algorithms for Communication and Signal Processing

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Electrical Engineering

by

Yilei Li

2016

ABSTRACT OF THE DISSERTATION

Phase-Based Algorithms for Communication and Signal Processing

by

Yilei Li

Doctor of Philosophy in Electrical Engineering
University of California, Los Angeles, 2016
Professor Mau-Chung Frank Chang, Chair

In this work, phase-based methods for emerging applications implemented with advanced CMOS technology are proposed. We use phase-based methods to solve problems in high-speed chip-chip/chip-memory data link interconnect, and in terahertz system.

With increasing demand for bandwidth between chip-chip and chip-memory communication, high-speed data link interconnect becomes a key block in modern computer systems. In order to serve portable devices, energy efficiency is a critical metric for data link interconnect. Multiband RF Interconnect, which transmits data by upconverts data streams into multiple frequency bands, can fully exploit the bandwidth of interconnect wire with excellent energy efficiency. In order to demodulate data stream, a correct carrier phase in receiver end is necessary. The conventional way for carrier phase synchronization requires high-speed, high-resolution analog-to-digital converter, which consumes large amount of power and silicon area. In the proposed phase-based method, a pilot signal is sent by transmit side, and in receiver end a simple comparator is sufficient to identify correct phase. This greatly enhances energy efficiency of data link and reduces cost. In addition, optimal demodulation phase for dispersive channel is derived.

The other proposed application for phase-based method is terahertz systems. CMOS terahertz source suffer from low device cut-off frequency, and fundamental frequency is limited.

Conventional time-domain harmonic extraction can selectively enhance certain harmonic of fundamental tone at output radiation and suppress other harmonics, and therefore increase effective radiation frequency. In terahertz band, time-domain harmonic extraction cannot achieve sufficient low-order harmonic suppression. A spatial-domain harmonic extraction method is proposed, which samples signal in specific spatial positions and combines sampled signals together to selectively enhance and eliminate tones. With proposed method, 1.4-THz CMOS imaging is achieved. In addition, a fast characterization method for terahertz source is proposed. This method uses random sampling (compressive sensing), and reconstruct sparse spectral profile by two-step zoom-in algorithm. With proposed method, a 8X time reduction can be achieved for terahertz source characterization, which enables regular self-test of terahertz systems.

The dissertation of Yilei Li is approved.

Wentai Liu
Sudhakar Pamarti
Jingsheng Jason Cong
Mau-Chung Frank Chang, Committee Chair

University of California, Los Angeles
2016

To my family

TABLE OF CONTENTS

List of Figures

ACKNOWLEDGEMENTS

I first wish to express my deepest gratitude to my advisor, Prof. Frank Chang. Prof. Chang is not only a great advisor in research, but also a wise mentor in how to be a better person. This is the most precious lesson that I learned in HSEL.

I also would like to thank Prof. Jason Cong, Prof. Sudhakar Pamarti and Prof. Wentai Liu for serving in my PhD committee. Your comments for my project greatly enhanced the quality of my research.

Further, I wish to acknowledge my lab mates for help and collaboration in my projects. Dr. S. J. Lee is the main contributor to RF Interconnect project, his vision and experience is the key to success of RFI projects. Dr. R. Al Hadi and Dr. Yan Zhao laid the foundation of THz imaging project. Wei-Han Cho is my main collaborator in RFI project, and his hardworking finally made our testchip work. Rod Kim, Yuan Du, Ken Wong, Jieqiong Du and Bug Huang also helped me a lot in RFI project. Besides, Joseph Chen, Hugh Wu, Boyu Hu, Xiao Wang, Chun-Chen Liu, Yen-Cheng Kuan, Yen-Hsiang Wang and Li Du provided valuable discussions that benefited me a lot. I also would like to thank Janet Lin for taking care of my funding.

Finally, I would like to express my most sincere gratitude to my family and my wife. Without your support, it would be impossible for me to finish this journal.

VITA

2009	B.S., Fudan University, Shanghai
2012	M.S., Fudan University, Shanghai
2012-2016	Research Assistant, High-Speed Electronics Lab, Electrical Engineering Department, UCLA

Publications

Y. Li et al., "Carrier synchronisation for multiband RF interconnect (MRFI) to facilitate chip-to-chip wireline communication." Electronics Letters, vol. 52, no. 7, 0216, pp.535-537

Wei-Han Cho, Y. Li, et al. "A 5.4-mW 4-Gb/s 5-band QPSK transceiver for frequency-division multiplexing memory interface." In 2015 IEEE Custom Integrated Circuits Conference (CICC)

Wei-Han Cho, Y. Li, et al., "A 38mW 40Gb/s 4-lane tri-band PAM-4/16-QAM transceiver in 28nm CMOS for high-speed Memory interface." In 2016 IEEE International Solid-State Circuits Conference

Z. Chen, Y. Li, Y. Kuan et al., ″Digital PLL for Phase Noise Cancellation in Ring Oscillator-Based I/Q Receivers”, In IEEE Symp. VLSI 2016

R. Al Hadi, Y. Zhao, Y. Li et al., “Retroactive terahertz detection for imaging and remote sensing applications in a standard CMOS technology”, In SPIE Photonic West 2016

R. Al Hadi, Y. Zhao, Y. Li et al., "A Spectral Profiling Method of mm-Wave and Terahertz Radiation Sources", In 2016 International Microwave Symposium

M. -C. F. Chang, S. –J. Lee, Y. Li and Wei-Han Cho, "Phase synchronization of modulation or demodulation for qam-based multiband tsv-link", Patent Pending, WO2014189651 (A1)

S. –J. Lee, Y. Li, G. Virbila, and M.-C. F. Chang, "LOW JITTER TUNABLE VOLTAGE CONTROL OSCILLATOR WITH SELF CALIBRATION CIRCUITS TO REDUCE CHIP FABRICATION PROCESS VARIATION", U.S. Patent No. 9,356,582

CHAPTER 1
INTRODUCTION

In electronic circuits, the carrier of information is essentially electromagnetic (EM) waves. Any signal S(t) can be approximated by the sum of N EM wave harmonics:

$$S(t) = \mathrm{Re}\left\{\sum_{i=1}^{N} A_i e^{j(\omega_i t+\theta_i)}\right\}$$

where A_i, ω_i and θ_i are the amplitude, frequency and phase for i-th harmonic. When S(t) is a deterministic signal, A_i and θ_i are deterministic and such decomposition is Fourier expansion. When S(t) is a random signal, A_i and θ_i are random variable and such decomposition is Karhunen-Loeve expansion. EM wave can be fully described by its three basic elements: amplitude, frequency and phase. Therefore information of signal can be carried by those three elements. Naturally, phase can be manipulated to carry or retrieve information.

Nowadays, phase, as a degree of freedom, is becoming more and more important as information carrier. First, with the advent of big data era, a huge amount of information needs to be communicated between producer and consumer of information within high speed. If all information is carried by amplitude, media with large bandwidth is required to send/receive error-free data according to Shannon's information theory [1.1]. This can greatly increase the cost of information exchange. Therefore we need to explore other degree of freedom as information carrier. Frequency modulation of EM wave cannot carry information with very high data rate as state-of-art implementation of frequency modulator, phase-locked loop (PLL), has very limited bandwidth (~1 MHz) and therefore frequency change rate cannot be too high [1.2]. On the other hand, modulation of phase can be easily achieved by Cartesian sum [1.3], and high data rate (>1 Gb/s) can be achieved by phase modulation. In current implementations of state-of-art data links [1.4-1.7], both amplitude and phase of EM wave are used to achieve high data rate.

In addition, phase manipulation becomes important today due to the evolution of low-cost CMOS circuits. With the continuous scaling of feature size, CMOS devices are becoming more and more powerful for high speed circuits. Today, from PC processor to mobile phone RF frontend, most high circuits are fabricated with CMOS technology. Meanwhile, the supply voltage of CMOS circuits drops as feature size scales. This means that the amplitude dynamic range of signal keeps decreasing since the largest output amplitude of CMOS circuits can hardly exceed supply voltage. With lower dynamic range, the signal has lower signal-to-noise ratio (SNR) and bit-error rate (BER) is degraded if all information is modulated by amplitude. On the other hand, the supply voltage scaling does not affect phase of signal. Therefore, the BER degradation can be alleviated if we modulate signal in phase domain instead of in amplitude domain.

Due to the two reasons above, phase-domain signal processing is becoming more and more important and attractive in circuits fabricated with deeply scaled CMOS technology for big data applications. In this work, we propose phase-domain signal processing algorithms to solve various problems in high speed circuit. This work can be divided into two parts. In part I, we use phase domain signal processing for carrier synchronization in multiband RF interconnect (MRFI) high speed data link. In Chapter 2, we briefly introduce background of high speed data-link, the concept of MRFI and its benefits. In Chapter 3, we discuss the necessity of carrier synchronization and proposed our phase-based algorithm for carrier synchronization. The experiment results are present in Chapter 4. In part II of this work, we applied phase-domain algorithms for CMOS terahertz imaging systems. We briefly review the background of terahertz imaging systems in Chapter 5. In Chapter 6, we introduce the concept of interferometric imaging over terahertz band. Then we propose Spatial Harmonic Extraction (SHE) method for resolution enhancement in Chapter 7. In Chapter 8 we further proposed random phase sub-sampling method for terahertz system characterization. After that, conclusion is drawn in Chapter 9 and possible future works are described in Chapter 10.

CHAPTER 2

CONCEPT OF MRFI

2.1 BACKGROUND OF HIGH SPEED DATA-LINK INTERCONNECT IN MORE THAN MOORE GENERATION

In 1965, Moore's Law was proposed by Gordon Moore in his famous article entitled "Cramming more components onto integrated circuits" [2.1]. According to Moore's law, the feature size of integrated circuits (IC) scales to half every two years. This is an amazing result. With feature size scaling, the IC chips become more powerful, but their price keeps reducing. As time goes by, the performance of IC chips grows exponentially but still affordable. Thanks to Moore's Law, we are able to enjoy a prosperous market of electronics products.

Today, we have to slow down the pace of feature size scaling due to following reasons. First, the downscaling of feature size is too expensive. Semiconductor manufacturers need huge amount of money to research and development next-generation process technology. In addition, fabrication with nanometer technology is very expensive. This is due to the fact that mask making and designing are all expensive. For example, in 16 nm technology, double patterning [2.2] is adopted with greatly increases the cost of mask. Also the design of mask is expensive as design rule is becoming more and more complicated for nanometer technology. When the upgrade expense is prohibitively high, the technology scaling pace is forced to slow down. Second, the resolution of lithography is limiting the feature size. The diffraction limit can be expressed as [2.3]:

$$d = \frac{\lambda}{2NA}$$

where d is diffraction (resolution) limit, λ is the wavelength of illuminating beam and NA is numeric aperture which is fixed for a given system. With the help of multiple patterning, the estimated achievable minimum feature size is around 5 nm [2.4].

The “More Than Moore” (MTM) plan has been proposed as a roadmap for future semiconductor industry [2.5]. In MTM, the semiconductor evolution is not just downscaling feature size with brute force, but with more efficient ways. One promising solution in MTM is advanced packaging [2.6]. Advanced packaging allows designer to put multiple chips within the same package, and large amount of in-package interconnect wires are possible. The advanced packaging has two main benefits. First, it makes heterogeneous integration possible. Currently, though the minimum wire width and spacing can be as narrow as ~15 nm for on-chip interconnects, the minimum interconnect wire width and spacing in printed circuit board (PCB) is still around 3 mil (~76 μm). Therefore, interconnect resources in PCB is rather limited. Without advanced package technologies, we have to integrate as many blocks on the same chip as possible. This makes system-on-chip (SoC) chips consume large silicon area and greatly raises the cost of such SoC. Moreover, all block in SoC must be designed with the same technology. However, nanometer technologies are not the best match for analog and RF blocks. For one thing, analog and RF design is extremely challenging under low supply voltage [2.7]. For the other, the size of inductor that are commonly used by analog and RF blocks do not scale with feature size, and analog and RF blocks usually consume large area. Therefore the analog and RF blocks in nanometer technology SoC are expensive and hard to design. On the other hand, with advanced packaging, since in-package interconnect resource is rich, one huge SoC can be split into multiple sub-blocks are connect with each other inside package. Each block can be implemented with the technology that matches its purpose best. For example, processor can be implemented with most advanced technology to achieve best transistor density and performance. Analog and RF blocks can be implemented with mature technology which has high supply voltage and low cost. With advanced packaging, SoC can be transformed into low-cost high-performance SiP (system-in-package).

The second important benefit that advanced packaging can bring is memory bandwidth for processors. The performance of processors today is greatly limited by memory bandwidth [2.8]. For modern computers with von Neumann architecture, processor needs to frequently fetch data from memory. When there is not enough memory bandwidth for processor, the processor has to stay idle until data is successfully fetched from memory. Conventional memory modules have limited pin count due to form factor, and the only way to increase memory bandwidth is to enhance the data rate of each pin. However, since the PCB interconnect quality is far from ideal, the increase of data rate per pin is followed by high power consumption, which once again brings in battery life and heat issues. With advanced packaging technology, high-density in-package interconnect with low loss is available. Therefore memory bottleneck can be solved by in-package energy-efficient interconnects.

It is clear the key enabling block for advanced packaging is interconnect data-link block. Conventional data-link circuits include inverter and current-mode logic (CML) (Fig. 2.1).

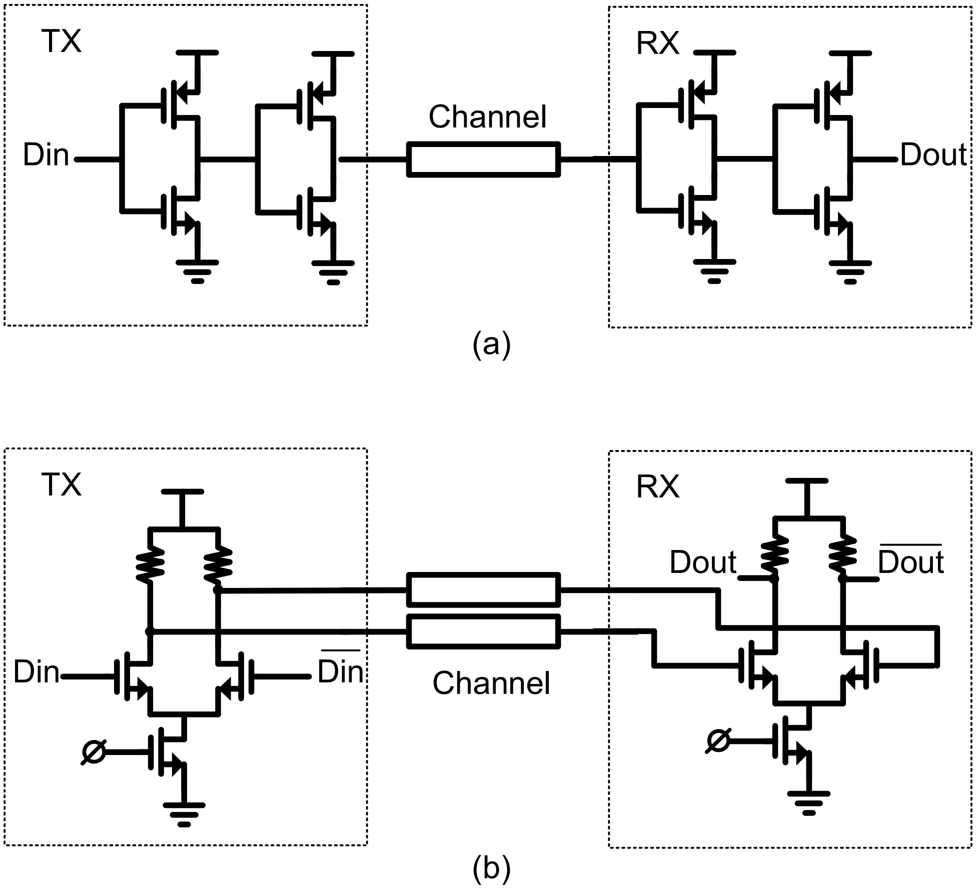


Fig. 2. 1 Schematic of (a) Inverter and (b) CML data link

Inverter data-link is easy to implement and does not consume DC current. However, inverter-based data-link can cause simultaneous switch noise (SSN) when multiple inverters turn on and turn off at the same time. CML data-link has constant total current regardless of switch status, and SSN issue is greatly alleviated. However, CML has DC current and is not as energy-efficient as inverter. In addition, there is a common issue with inverter and CML, which is inability to handle channel with frequency notches. Frequency notches can be caused by impedance discontinuity of channel near open stubs and vias. With notches present, the output signal quality is degraded, which in turn increases BER. Conventional way to overcome channel frequency notch is equalizer. However, equalizer can consume considerable amount of power and is not easy to design [2.9]. It is desirable to keep signal quality while eliminating equalizers as much as possible.

In sum, the interconnect block needs to satisfy following requirements:

(1) High data bandwidth

(2) Low power consumption

(3) Capacity to handle channel

2.2 MRFI

MRFI has been proposed [2.10, 2.11] to meet the above requirements without equalizer. The diagram of MRFI is shown in Fig. 2.2.

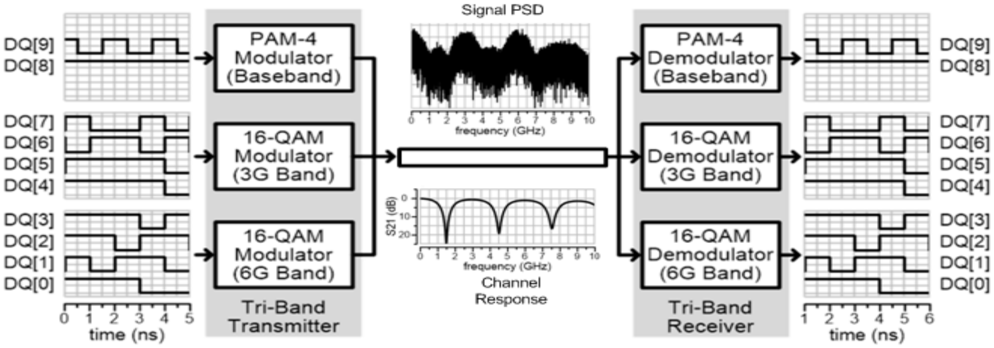
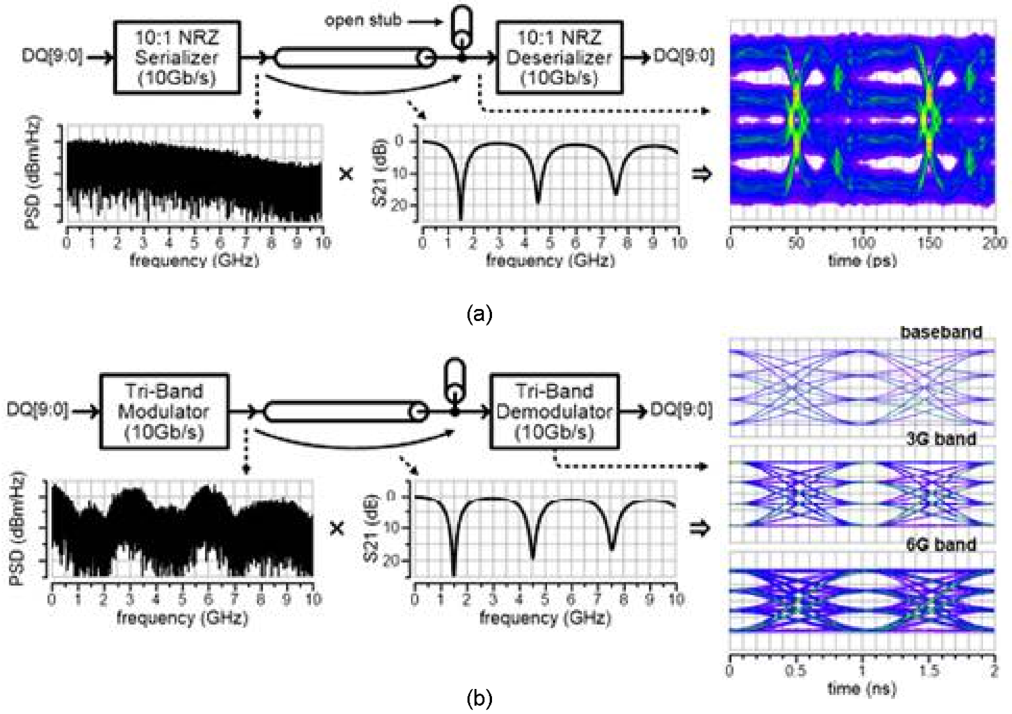


Fig. 2. 2 MRFI Interconnect

MRFI splits high-speed data stream into multiple low-speed data streams and sends/receives low-speed data streams in parallel. In order to overcome frequency notch issue, each low-speed data stream is up-converted to specific carrier frequency which has relative flat channel response around. Since each data stream is transmitted over flat channel, no equalization is required, which is a great advantage over conventional return-to-zero (NRZ) signaling (Fig. 2.3). Therefore, the power and design burden can be saved if MRFI is used. Moreover, since multiple bands are used to transmit data streams in parallel in one physical channel, one physical channel is equivalent to multiple logic channels. This virtual logic channel equivalent is proved to be useful for flexible data transmission among processors [2.12].



(a)

(b)

Fig. 2. 3 Comparison of (a) NRZ and (b) MRFI through same channel with frequency notches without equalizer

In MRFI proposed in [2.11], three bands are used, including baseband (no upconversion), 3 GHz band and 6 GHz band. The total data rate is 10 Gbps. In TX side, each modulator includes digital-to-analog converter (DAC), up-mixer and output buffer (Fig. 2.4). Each DAC has four-level output to represent two bits. For baseband modulator, the carrier for mixer is stuck at DC as no upconversion is required. The output of baseband is equivalently PAM-4 signal. For 3 GHz and 6 GHz band, 3 GHz and 6 GHz I/Q carrier is fed into mixer to generate QAM-16 signal. The signal at buffer output is in current domain, and the output from different buffer outputs are easily combined together.

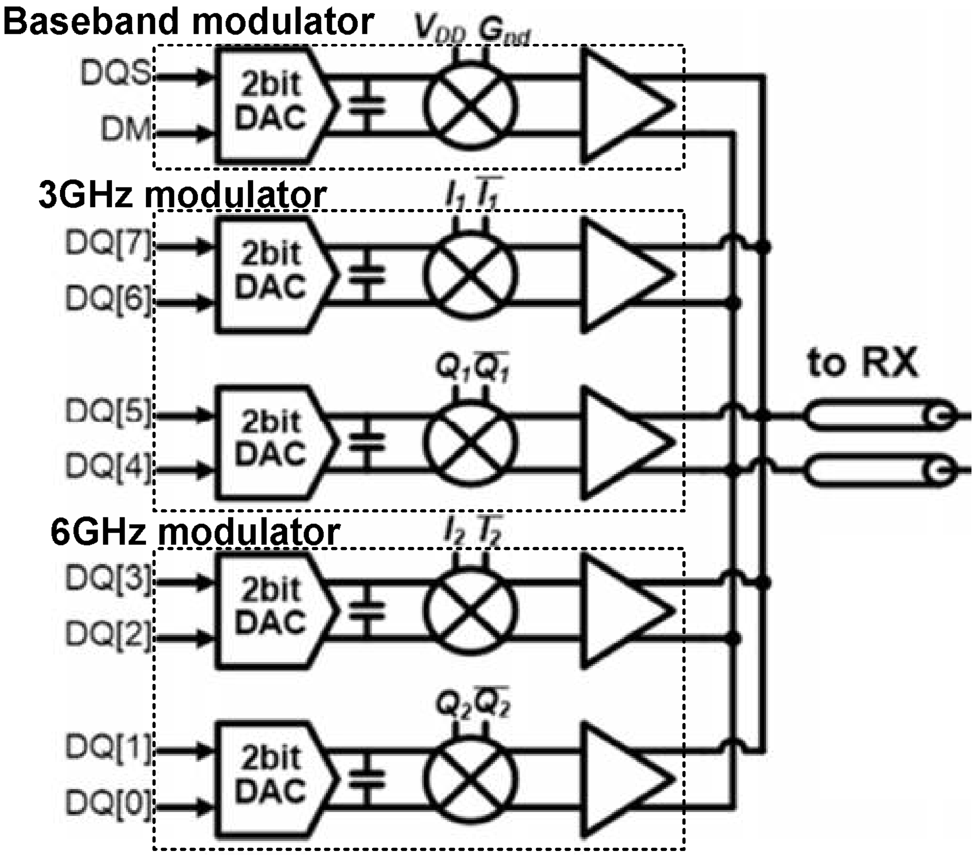


Fig. 2.4. TX diagram of MRFI

In RX side (Fig. 2.5), the signal is first coupled into RF coupler. The coupler mirrors input signal into 5 branches. Each branch down-converts signal with specific carrier, filters down-converted signal with low-pass filter (LPF) and converts filtered analog signal to digital signal by analog-to-digital converter (ADC). The ADC can be implemented with hysteresis to eliminate the effect of ripple due to inter-channel interference.

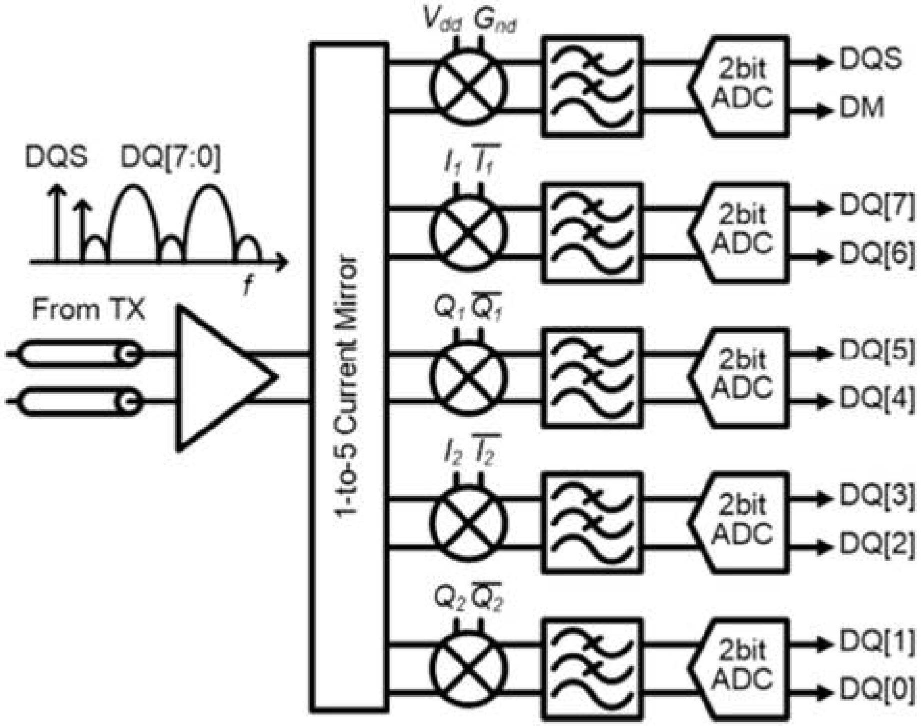


Fig. 2.5. RX side of MRFI

The proposed MRFI can achieve 10 Gbs data rate with power consumption of only 9.5 mW, and an excellent energy efficiency of 0.95 pJ/bit is achieved. In addition, since no equalizer is used, the MRFI transceiver design is very compact, which only occupies 0.01 mm^2 per transceiver pair. The main benefits of MRFI include:

(1) No need to use equalizer. Since MRFI places data streams into frequency bands where channel responses are flat, received data eye-diagram can be near ideal even without equalizer. Power consumption, area occupation and design burden can be greatly reduced for equalizer-free MRFI data-link transceivers.

(2) The bandwidth of wire is fully exploited. In conventional data-link interconnect transceivers, even with advanced equalizer design, the number of frequency notches that can be handled is still limited. Therefore data bandwidth is limited by number of channel frequency notches, and a

great portion of available wire bandwidth is wasted. With MRFI, the data stream can fully exploit flat frequency bands between notches, and the bandwidth of interconnect wires can be used without waste. Therefore high data rate is available with MRFI transceivers.

(3) MRFI has better tolerance for sampling jitter. This is due to the fact that MRFI uses frequency domain multiplexing instead of time domain multiplexing for data transceiving. For example, suppose two 200-Mbps data streams are to be sent through one wire. For conventional time-domain serializer, the transmitted and received data will have data rate of 400 Mbps, and eye window is less than 2.5 ns. On the other hand, if we use MRFI, we modulate two 200-Mbps data streams to two different bands. The eye window for each demodulated 200-Mbps data stream is around 5 ns. Therefore MRFI can tolerate more sampling jitter.

Chapter Conclusion

This chapter introduces concept of MRFI. The concept of MRFI was proposed and verified by teamwork of Wei-Han Cho, Yilei Li, Yanghyo Kim, Yuan Du, Jieqiong Du, Chien-Heng Wong and Sheau-Jiung Lee.

CHAPTER 3

PHASE-BASED ALGORITHM FOR MRFI CARRIER SYNCHRONIZATION

3.1 BACKGROUND

3.1.1 Modulation

Modulation is a transform of signal. By modulate signal S(t) into Su(t), we can add features that original signal S(t) does not have. Original signal can also be extracted from modulated signal Su(t) by demodulation.

Originally modulation is used to solve problems in radio communication [3.1]. Slow-changing signals in our daily life (usually have their center frequency at 0 Hz, called baseband signals) are hard to propagate in air due to physical characteristics of air and electromagnetic wave. In order to propagate properly, baseband signal should have its center frequency upconverted by multiplying baseband signal S(t) with carrier cosωt (frequency modulation):

$$\mathrm{Su}(\mathrm{t}) = \mathrm{S}(\mathrm{t}) \cdot \cos\omega\mathrm{t} \tag{3.1}$$

Demodulation of above modulation is by multiplying Su(t) with carrier again and take away high frequency components by low-pass filter (coherent modulation/demodulation):

$$\mathrm{Su}(\mathrm{t}) \cdot \cos\omega\mathrm{t} = \mathrm{S}(\mathrm{t}) \cdot \cos\omega\mathrm{t} \cdot \cos\omega\mathrm{t} = \mathrm{S}(\mathrm{t})\frac{1+\cos2\omega\mathrm{t}}{2} \xrightarrow{\text{LPF with gain of 2}} \mathrm{S}(\mathrm{t}) \tag{3.2}$$

Frequency modulation has other useful features. One carrier $C_1(t)$ is orthogonal to another carrier $C_2(t)$ if their correlation is (approximately) zero:

$$\mathrm{C}_1(\mathrm{t}), \mathrm{C}_2(\mathrm{t})\text{orthogonal} \Leftrightarrow \frac{1}{\mathrm{T}}\int_0^{\mathrm{T}} \mathrm{C}_1(\mathrm{t})\mathrm{C}_2(\mathrm{t})\mathrm{dt} \cong 0 \tag{3.3}$$

Many baseband signals can be transmitted simultaneously with the same medium if carriers orthogonal to each other are used to modulate them. In receiving end, one carrier multiplies sum of modulated signals to demodulate desired signal. The signal modulated with in-phase carrier is recovered, while signals modulated with orthogonal carriers are eliminated. This is called frequency-domain multiplexing. Carriers with different frequencies are orthogonal since:

$$\frac{1}{T}\int_0^T \cos\omega_1 t\cos\omega_2 t dt = \left[\frac{\sin(\omega_1+\omega_2)t}{2(\omega_1+\omega_2)T} + \frac{\sin(\omega_1-\omega_2)t}{2(\omega_1-\omega_2)T}\right] \Big|_T^0 \xrightarrow{(\omega_1-\omega_2)T\gg1,(\omega_1+\omega_2)T\gg1} 0 \qquad (3.4)$$

In addition, two carriers with the same frequency but with π/2 phase difference are orthogonal since:

$$\frac{1}{T}\int_0^T \cos\omega_1 t\sin\omega_1 t dt = \left[\frac{\sin2\omega_1 t}{4\omega_1 T}\right] \Big|_T^0 \xrightarrow{\omega_1 T\gg1} 0 \qquad (3.5)$$

Thus, with n carriers, it will be possible to simultaneously send 2n modulated signals with same media.

3.1.2 Constellation

As mentioned above, two carriers with the same frequency but phase difference of π/2 can modulate two independent signals. Let's assume the two carriers to be cosωt and sinωt for simplicity, and baseband signals to be modulated to be a(t) and b(t). Since a(t) and b(t) are independent, number pair (a(t),b(t)) can be mapped to a two-dimension Cartesian coordinator system. This kind of modulation is called *quadrature modulation*, where a(t) (modulated by in-phase carrier cosωt) is called signal of in-phase path (I-path), and b(t) (modulated by quadrature carrier sinωt) is called signal of quadrature path (Q-path).

For digital communication system, we can define all possible (a(t),b(t)) pairs (in the context of digital communication, usually one pair (a(t), b(t)) is called a symbol). Then we can map multiple bits to one symbol. For n bits, (according to information theory)we need the position of symbol (a(t),b(t)) to have 2^n possibilities. For example, suppose a(t), b(t) can be +1 and -1, then the position of symbol (a(t),b(t)) can have 4 possibilities. Then we can map two bits to such one symbol: 00 is mapped to (1,1), 01 is mapped to (-1,1), 10 is mapped to (-1,-1), 11 is mapped to (1,-1). All possible positions of symbols are called *constellation of symbols*.

QAM is a commonly used modulation scheme in digital communication. For QAM-n^2, a(t) and b(t) will have n possible positions, with equal distance between two neighbor positions (such as -3, -1, 1, 3). Thus, one symbol can carry information of $\log_2 n^2$ bits. Usually n is an even number so that all constellation of QAM-n^2 is symmetric (so we can see QAM-4, or QPSK, with n=2, QAM-16 with n=4, QAM-64 with n=8, etc.).

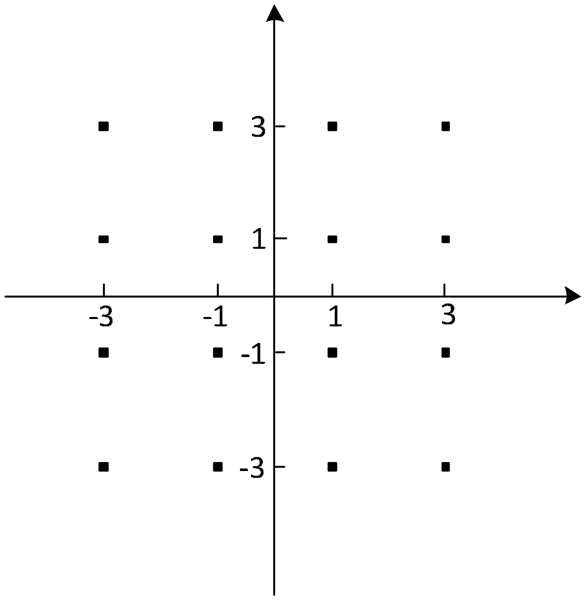


Fig.3.1 Constellations of QAM-16

In RX end, received signals symbols will be corrupted by noise and interference. Therefore, positions of received symbols will be away from ideal constellation. Instead of staying at one definite point, symbols will be moving in a certain area. Based on minimum error probability criterion, we can define decision regions for each symbol. Ideal position of each symbol will stay in the center of its decision region to minimize probability of error. For example, in QAM-16, the possible mapped bits for one symbol can be 0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111. For each possible mapped bits, we have a decision region. Symbol is translated back into bits according to the decision region that it falls into.

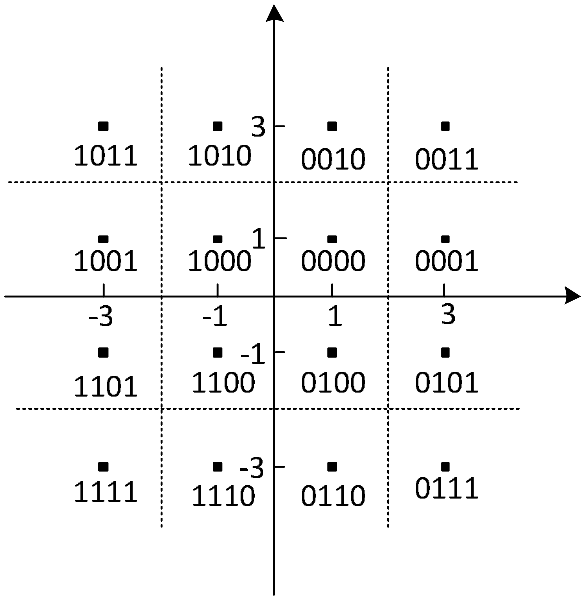


Fig. 3.2 Decision Regions of QAM-16

3.2 CARRIER PHASE ERROR IN QUADRATURE MODULATION

Phase recovery greatly effects the demodulated signal quality in coherent modulation/demodulation. Assume we have modulated signal of $a(t)\cos\omega t+b(t)\sin\omega t$ at receiving end, where $a(t)$ is baseband signal for I-path and $b(t)$ is baseband signal for Q-path. If we use carrier with phase error of θ cos(ωt+θ)t and sin(ωt+θ) to demodulate, then output for I-path and Q-path will become:

$$\mathrm{I-path{:}}\ [\mathrm{a(t)cos\omega t} + \mathrm{b(t)sin\omega t}] \cdot \cos(\omega t + \theta) = 0.5 \cdot [\mathrm{a(t)cos\theta} - \mathrm{b(t)sin\theta}] +$$

$$0.5 \cdot [\mathrm{a(t)} \cos(2\omega t + \theta) + \mathrm{b(t)} \sin(2\omega t + \theta)] \xrightarrow{\text{LPF with gain of 2}} \mathrm{a(t)cos\theta} - \mathrm{b(t)sin\theta}$$

(3.6)

$$\mathrm{Q-path{:}}\ [\mathrm{a(t)cos\omega t} + \mathrm{b(t)sin\omega t}] \cdot \sin(\omega t + \theta) = 0.5 \cdot [\mathrm{a(t)sin\theta} + \mathrm{b(t)cos\theta}] +$$

$$0.5 \cdot [\mathrm{a(t)} \sin(2\omega t + \theta) - \mathrm{b(t)} \sin(2\omega t + \theta)] \xrightarrow{\text{LPF with gain of 2}} \mathrm{a(t)sin\theta} + \mathrm{b(t)cos\theta}$$

(3.7)

If we use carrier with correct phase, e.g., θ=0 to demodulate, then the output of I-path is a(t) and Q-path is b(t). On the other hand, if we use incorrect phase θ to demodulate, we will have part of Q-path signal b(t) in I-path, and we will have part of I-path signal a(t) in Q-path. This is *IQ interference* due to phase error.

We can also use constellation to analyze effect of IQ interference. By using carrier with phase error of θ to demodulate, it is equivalent to rotate the constellation by phase of θ while keeping the original decision region without rotation. It can be seen from Fig. that the symbols in constellation with phase rotation is closer to boundary of decision region than original constellation. This means that with the same noise power, symbol with phase rotation is more likely to fall into wrong decision region since it is closer to the boundary of decision region. This means higher bit error rate for symbols with phase error even with the same noise power. In order to optimize bit error rate, phase recovery is necessary for quadrature modulation.

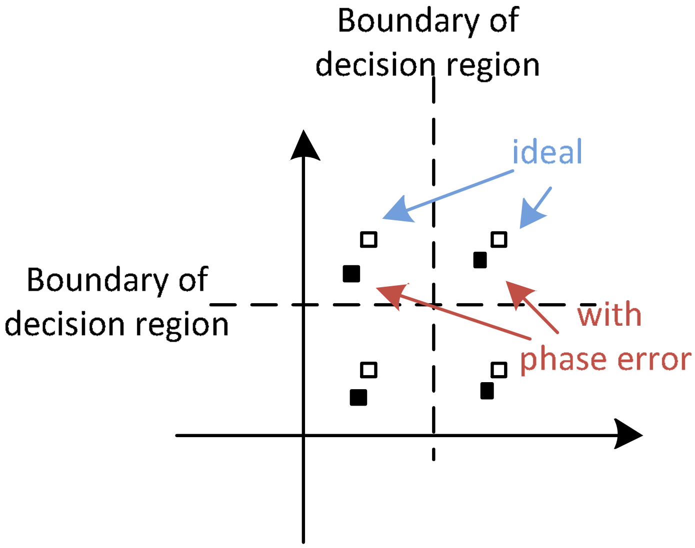


Fig. 3.3 Constellation with and without phase rotation

For wireless communication, phase recovery is usually accomplished by baseband. Basically signal is demodulated and converted into digital domain and constellation is rotated back to correct position by digital baseband. This is due to the fact that channel characteristics of wireless communication is changing with time. When you are making phone call in a car in the highway, the signal path from your cellular phone to base station changes fast with time. Consequently, a baseband algorithm must be implemented to track the channel and find correct phase rotation in real time [3.2].

On the other hand, for communication on channel with almost time-invariant characteristics (which is the case of wireline RF interconnect), phase recovery will be needed only once and real-time phase track is not necessary. Real-time phase tracking will also induces a lot of latency in signal processing, which is not desired in RF interconnect. Thus, for wireline RF interconnect, we will use phase calibration instead of real-time phase recovery.

3.3. PHASE CALIBRATION

3.3.1 Problems of phase calibration in digital domain

In digital domain, we only have state of '0' or '1'. We can only send symbols on pre-defined constellation (which translates into '-1' and '1' for digital state of '0' and '1'). This induces complicity into phase recovery. First, the resolution of phase recovery is strongly affected by the resolution of analog-to-digital conversion (ADC). If the analog-to-digital conversion only has coarse resolution, then it will be impossible for digital domain to tell the difference of symbols with small carrier phase change (as the digital outputs of ADC are the same for two cases). For example, if we have QPSK modulated signal of 00 which is translated into $-1*\cos\omega t-1*\sin\omega t$, and we use carrier phase with a significant phase error of 20° to demodulate. According to (6) and (7), the demodulated I-path signal is -0.59 and Q-path signal is -1.27. This means it will needs ADC with at least 3-bit resolution to tell the difference of 20-degree phase error (as 2-bit ADC with full-swing of ±1 has threshold of {-0.5, 0, 0.5} and signal of -0.59 and -1 fall into the same conversion step). On the other hand, this is QPSK modulation and we only want 1-bit ADC for data conversion. Thus, it will bring much burden to ADC design if we want to do phase calibration in digital domain.

Second, in addition to phase error, IQ mismatch (signal strength difference between I-path and Q-path) is also present. This brings another degree of freedom into constellation distortion. Suppose we have gain error of Δ for I and Q (as shown in Fig. 3.4), which means 00 will be modulated as $-(1+\Delta)\cos\omega t-\sin\omega t$. Since it is impossible to decouple phase error and IQ mismatch in digital domain, the only way to find correct phase and IQ mismatch is to solve a system of transcendental equations (since there is trigonometric equations related to phase rotation), or to build a look-up table with fine resolution. In anyway, it will need ADC with very fine resolution and highly complicated digital baseband, which is not desired in RF interconnect.

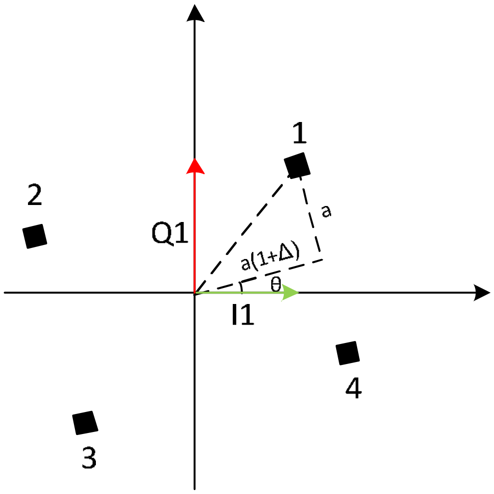


Fig. 3.4 Constellation of QPSK with both phase error and IQ mismatch

3.3.2 Phase calibration with mixed-mode method [3.3]

Phase calibration with small phase error (<90°)

The problems of digital-domain phase calibration are largely due to the fact that we can only send symbols on pre-defined constellation positions when calibrating. To solve this problem, we will need to use mixed-mode calibration method.

The essence of mixed-mode calibration method is to introduce state of 'off' in addition to digital states '0' and '1'. In calibration mode, suppose we turn off a(t), and send '0' in b(t), which is translated into -sinωt, and we use carrier phase with a phase error of θ to demodulate. According to (3.6) and (3.7), the demodulated signal at I-path and Q-path are:

$$\mathrm{I-path:}\,[-\sin\omega t]\cdot\cos(\omega t+\theta)=0.5\cdot[\sin\theta]+0.5\cdot[-\sin(2\omega t+\theta)]\xrightarrow{\text{LPF with gain of 2}}\sin\theta \tag{3.8}$$

$$\mathrm{Q-path:}\,[-\sin\omega t]\cdot\sin(\omega t+\theta)=0.5\cdot[-\cos\theta]+0.5\cdot[\sin(2\omega t+\theta)]\xrightarrow{\text{LPF with gain of 2}}-\cos\theta \tag{3.9}$$

According to (3.8), I-path output will become analog 0 when RX carrier phase is correct (since we assume phase error<90°, there is only one solution for $\sin\theta=0$). Thus, to find correct phase, we only need to find when output of I-path is analog 0. ADC with 1-bit resolution is enough. Suppose comparator in ADC has sufficiently large gain, when θ goes across zero (e.g., -1° to 1°), output of ADC will jump from '-1' to '1'. Thus, the phase corresponding to the zero-crossing point of ADC is the phase that is closest to correct phase. In this way, the resolution of phase calibration is only limited by resolution of phase sweep and will be decoupled from resolution of ADC. One-bit ADC is sufficient for phase calibration for all kinds of QAM.

Another advantage of mixed-mode calibration is that it decouples phase error from IQ mismatch. Suppose we have IQ mismatch of Δ, where the gain of I-path is 1 and gain of Q-path is (1+Δ). Under the presence of phase error of θ, when we turn off I-path and send '0' in Q-path, the received signal will become:

$$\mathrm{I-path:}\ [-(1+\Delta)\sin\omega t]\cdot\cos(\omega t+\theta)=0.5(1+\Delta)\cdot[\sin\theta]+0.5(1+\Delta)\cdot[-\sin(2\omega t+\theta)]$$

$$\xrightarrow{\text{LPF with gain of 2}}\ (1+\Delta)\sin\theta \tag{3.10}$$

$$\mathrm{Q-path:}\ [-(1+\Delta)\sin\omega t]\cdot\sin(\omega t+\theta)=0.5(1+\Delta)\cdot[-\cos\theta]+0.5(1+\Delta)\cdot[\sin(2\omega t+\theta)]$$

$$\xrightarrow{\text{LPF with gain of 2}}\ -(1+\Delta)\cos\theta \tag{3.11}$$

According to (3.8), the criterion of correct phase is still the zero-crossing point in I-path, which is not affected by IQ mismatch (as long as Δ is not -1 to totally cancel the signal in I-path).

In summary, mixed-mode calibration introduces one more state (off) in addition to digital states. Consequently, effectiveness of phase calibration is decoupled from resolution of ADC (1-bit

ADC is enough for phase calibration of all kinds of QAM modulations) and IQ mismatch. With mixed-mode phase calibration, we can first calibrate phase and then calibrate gain. No iteration of phase and gain calibration is required.

An exemplary mixed-mode phase calibration flow chart is shown in Fig. 3.5.

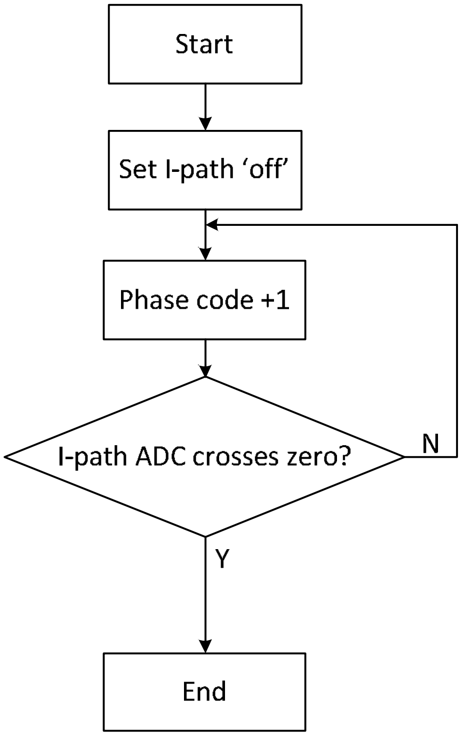


Fig. 3.5 Exemplary flow chart of mixed-mode phase calibration

Phase calibration with arbitrary phase error

When phase error can exceed 90°, expression (3.8) will have more than one solution for θ (0°,180°,360°, etc.). We will need more constraints to find the solution.

Suppose we have phase error of 190°. When we sweep phase to 180°, (3.8) is zero. In this case, I- and Q-paths will have opposite polarity as original signal. The polarity can be checked by turning on I- and Q-path and send predefined training sequence at the end of calibration. By checking the polarity of sequence at receiving end, we can know the phase error is 0° or 180°.

At this time we can either sweep phase back by 180° more, or just change data polarity in digital domain. In most cases, changing polarity by digital means is much easier than by shifting carrier phase by 180° more. Moreover, by changing polarity in digital domain, we can reduce required phase sweep range from 360° to 180°, which relieves much design burden from phase shifter (delay line).

In addition, if we allow IQ swap in digital domain, we can further limit phase sweep range to 90°. Suppose we have initial phase error of 100°, and turn off I-path in calibration. As we sweep the phase from 100° towards 0°, according to (9), Q-path will first become zero when phase error is 90°. At this time, I/Q path is swapped. If we keep phase error of 90° and swap I/Q path data output at digital side, then effectively we have no IQ interference. Thus, phase error of exact 100° can be handled by phase delay of 10° combined with IQ swap.

In sum, by allowing polarity swap and IQ swap in digital domain, we can limit the phase sweep range to 90°. From constellation aspect, the algorithm does follows: first sweep phase to rotate symbol from original position (Fig. 3.6(a)) to nearest correct position in any quadrate (Fig. 3.6(b)), and then map the symbol to correct quadrate by polarity/IQ swap (Fig. 3.6(c)).

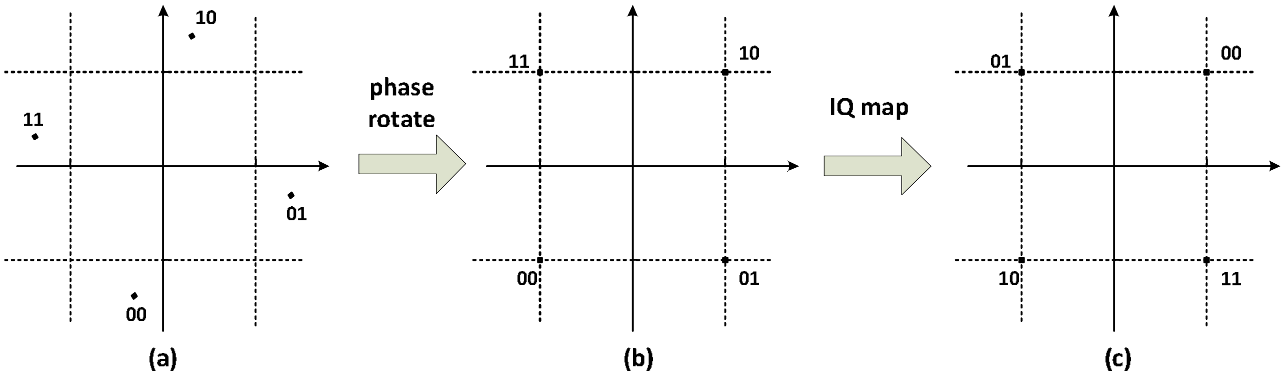


Fig. 3.6 Constellation during calibration

3.3.3 Circuit Implementation of Phase Calibration Algorithm

The proposed algorithm is implemented with a finite-state machine (FSM). The diagram is shown in Fig. 3.7. The simplified flow is described below.

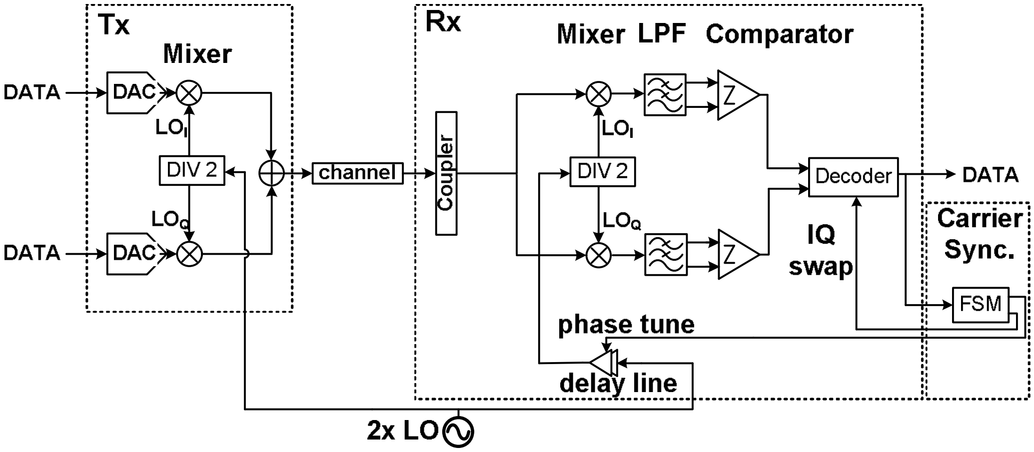


Fig. 3.7 System diagram with carrier synchronization FSM

First, after phase adjustment enable signal is asserted (both in TX and RX), the DAC of I path in TX will turn off, while DAC in Q path will output ‘1’. The FSM in RX side adjust the delay of delay line until output of comparator crosses zero (either from 1 to 0 or from 0 to 1; here assume is from 1 to 0 in the following discussion). Save the phase code at zero-crossing point as P1.

However, P1 might not be the true phase code corresponding to the zero-crossing point of analog output (LPF). This may due to the hysteresis in comparator. In order to avoid the interference of noise, the flipping threshold of comparator from 0 to 1 (T_{0-1}) is greater than 0 mV. Similarly, the threshold from 1 to 0 (T_{1-0}) is lower than 0 mV. Suppose the first flip is from 1 to 0, then phase code P1 is corresponding to hysteresis threshold T_{1-0}. In order to find the true zero-crossing point, phase code is swept backwards. Then the phase code P2 (flip from 0 to 1) corresponds to hysteresis threshold T_{0-1}. Since hysteresis is symmetric, i.e., $T_{0-1}+ T_{1-0}=0$, we can use P=(P1+P2)/2 as the estimate of the phase code corresponding to true analog zero-crossing point

(Fig. 3.8). This finished phase sweep in phase adjustment algorithm. Meanwhile, algorithm will detect the zero-crossing point is in I-path or Q-path as first step of IQ map. If the zero-crossing point is in I-path, then no IQ swap is required (only polarity swap is likely). On the other hand, if the zero-crossing point is in Q-path (while TX turns off I-path not Q-path), then IQ swap is necessary.

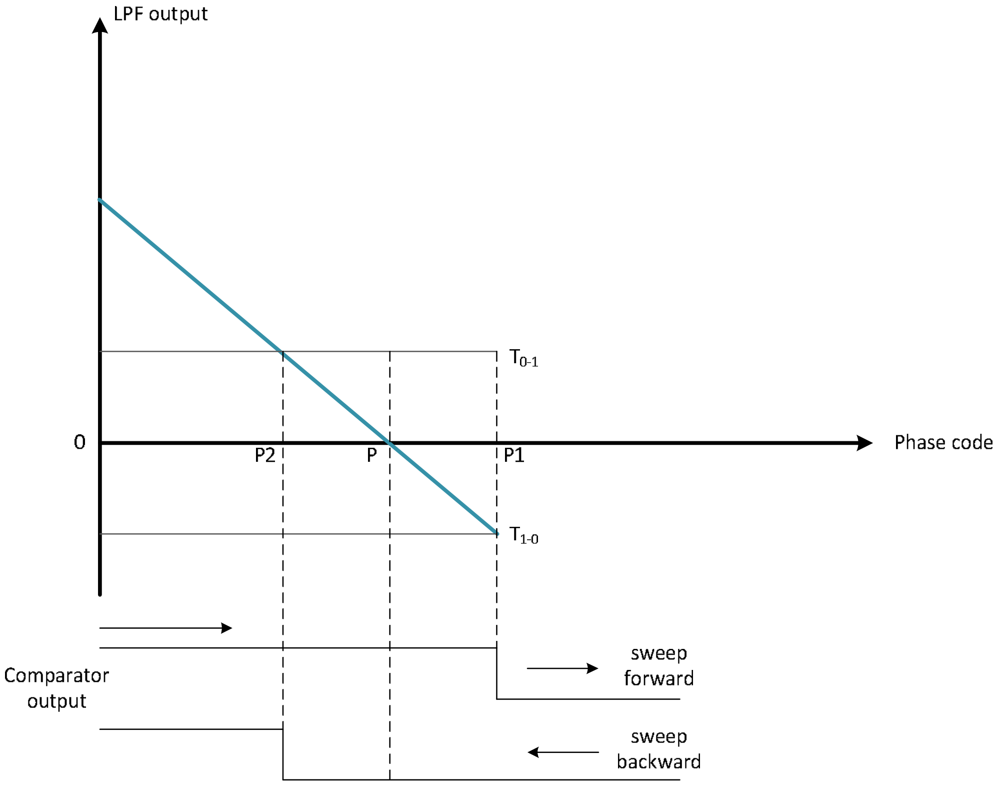


Fig. 3.8 Dual direction sweep for ADC with hysteresis

After phase sweep, polarity check enable signal is asserted in both TX and RX. I-path is turned on again in TX. A pre-defined fix data symbol is sent from TX. After receiving in RX, the FSM detects whether the received symbol has correct polarity compared to pre-defined symbol. Polarity swap will be performed when received symbol has opposite polarity with pre-defined symbol.

The whole flow chart is shown in Fig. 3.9.

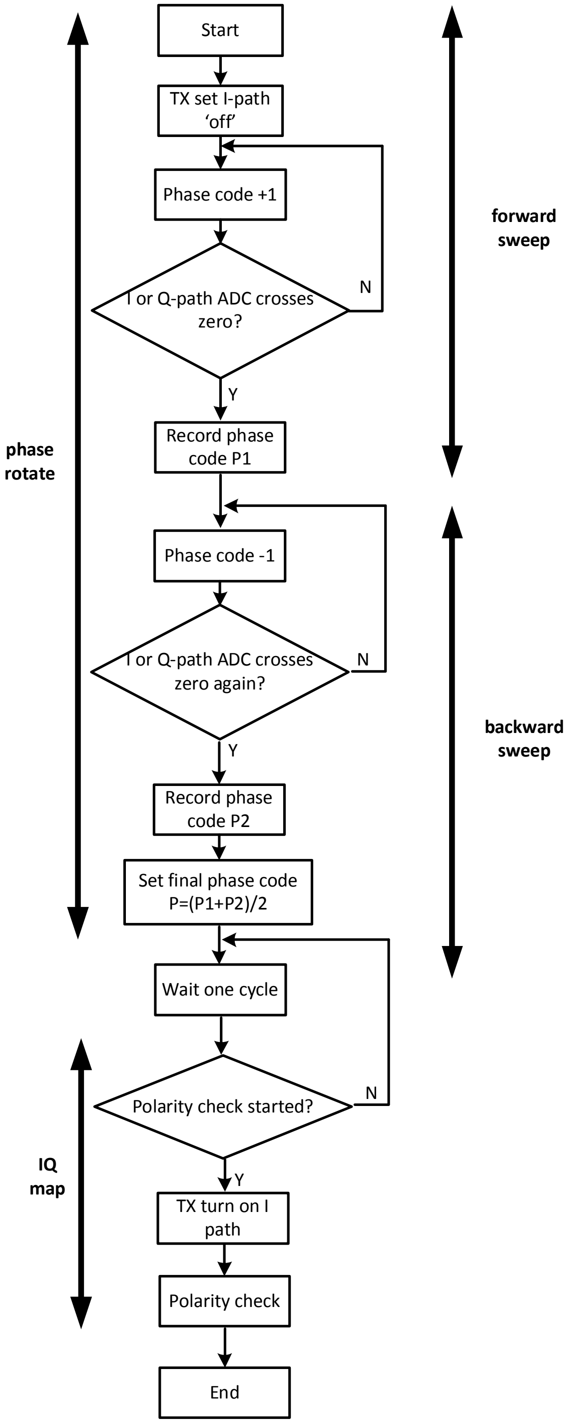


Fig. 3.9 Complete flow chart of algorithm

3.3.4. Phase Calibration with DC Offset

Suppose there is DC offset ε, so that the output of I-path is (with I-path gain of (1+Δ)):

$$(1+\Delta)\sin(\theta-\theta c)+\varepsilon$$

Where θ is current phase and θc is actual phase delay for compensation. And the stop criterion of phase fine tune becomes:

$$(1+\Delta)\sin(\theta-\theta c)+\varepsilon=0 \Rightarrow \theta=\operatorname{asin}\left[-\frac{\varepsilon}{1+\Delta}\right]+\theta c$$

This means DC offset brings in error for phase calibration.

In order to eliminate this error, we can do phase fine tune twice. First, we scan with I off and Q=1 in TX; then we scan with I off and Q=0 in TX. In second case, the output of I-path becomes:

$$-(1+\Delta)\sin(\theta-\theta c)+\varepsilon$$

And the output phase by algorithm is:

$$\operatorname{asin}\left[\frac{\varepsilon}{1+\Delta}\right]+\theta c$$

Due to the fact that asin is an odd function, we have $\operatorname{asin}\left[\frac{-\varepsilon}{1+\Delta}\right]=-\operatorname{asin}\left[\frac{\varepsilon}{1+\Delta}\right]$. Therefore, by averaging the two phase codes obtained in two scans, we can know the correct phase code with DC offset.

This method can also be used to calibrate DC offset (when phase code is the same for Q=1 and Q=0, then DC offset is 0 for I path). The advantage of this method over statistical method (transmit 0 and 1 in TX, and count received 0 and 1 in RX) is that it can have better sensitivity. When offset is small, statistical method may not be able to detect offset (received data are correct), but this offset may be important when the gain of RX drops for some reason. Suppose the signal amplitude is A, then sensitivity of traditional method is A (i.e., can detect offset as small as A). On the other hand, with this method, the sensitivity to offset is $A\sin(2\pi\Delta t\cdot fc)$, where fc is carrier frequency. This sensitivity can be much better than traditional method.

3.4. OPTIMAL PHASE WITH NON-LINEAR-PHASE-DELAY MEDIUM

3.4.1 Basics of Transmission Line

When shall we consider transmission line effect

When we talk about "signal goes through interconnect", we are actually talking about the phenomenon that an electromagnetic (EM) wave traveling through a waveguide. When the waveguide is short compared with wavelength, phase difference of points in the waveguide can be ignored, and the waveguide can be modeled as lumped R-L-C. On the other hand, when waveguide is long compared with wavelength, it should be modeled as a transmission line since points in waveguide can have significant phase difference. As a rule of thumb, when wire is longer than 1/4 of wavelength ($\lambda/4$), then the wire must be modeled as a transmission line. For our application with highest carrier frequency of 2.4 GHz, when trace is longer than 10mm, then transmission line effect must be considered [3.4].

Impedance discontinuity

EM wave will travel smoothly through the medium if the characteristic of medium is constant. However, when the characteristic of medium changes at some point, part of the energy in incident wave will reflect back. For transmission line, we use characteristic impedance (Z0) to describe its characteristic. When one transmission line connects to another transmission line with different characteristic impedance or terminated by an impedance different from its characteristic impedance (impedance discontinuity), then reflection appears at the boundary of two transmission lines. Incident wave and reflected wave add together and become stand wave in transmission line [3.5].

For any practical RF system, transmission line interconnect between modules will not be infinitely long. The end of the transmission line will be the input point of one module (Fig. 3.10(a)). If the termination impedance is not perfectly matched to characteristic impedance of transmission line, then reflection exists in boundary (Fig. 3.10(b)).

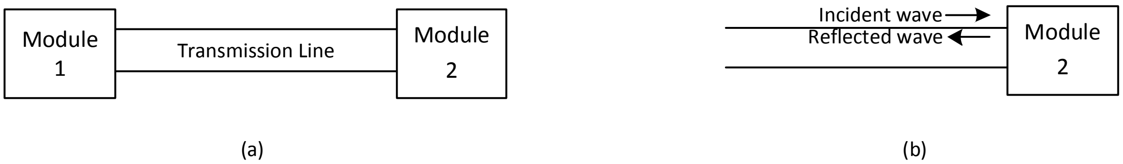


Fig. 3.10 Transmission line effect

The reflected wave will have the same frequency as the incident wave but not necessarily the same amplitude and phase. The phase and amplitude differences between incident wave and reflected wave are decided by reflection coefficient Γ. Γ can be expressed with respect to input impedance Zin and characteristic impedance Z0 as:

$$\Gamma = \frac{Z_{in} - Z_0}{Z_{in} + Z_0} = |\Gamma| e^{-j\theta}$$

At boundary of transmission line and module, suppose we have incident wave of V(t) and reflection coefficient of Γ, then the stand wave at boundary can be expressed as (1+Γ)V(t) .

3.4.2 Cause of Non-linear Phase Delay (Dispersion)

Non-linear phase due to reflection

Consider the incident wave with phase velocity of c and frequency of f. Its initial phase at the start of transmission line is 0. When it goes through transmission line with length of l, the phase of incident wave at the boundary between transmission line and RF module 2 is $\theta_0 = l*c/f$. This phase is linear with frequency f. With reflection coefficient of Γ, the effective phase θ_1 at input of module 2:

$$\theta_1 = \arg\left[e^{-j\theta_0}(1+\Gamma)\right] = \arg\left[e^{-j\theta_0}\left(1+|\Gamma|e^{-j\theta}\right)\right]$$

It can also be shown with vector plot (Fig. 3.12):

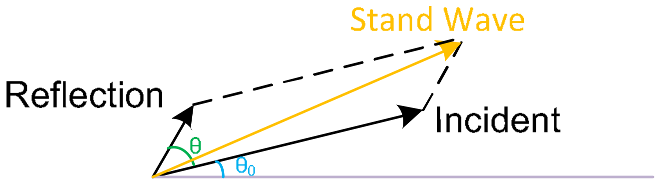


Fig. 3.12 Standing wave formation

Thus, the actual phase at boundary depends on not only electrical length of trace, but also reflection coefficient.

To make things even more complicated, reflection coefficient Γ is not constant with frequency. This can be understood by the expression of Γ. The input impedance of RF module 2 is not constant with frequency, as there are capacitive and inductive components in Z_{in} which changes with frequency. This is the main factor responsible for the change of Γ. On the other hand, practical low-cost waveguides are far from ideal (especially PCB trace), and the Z_0 can also vary with frequency. With those two factors, Γ will change with frequency. With frequency-variant Γ, phase change at the input of module 2 is no longer linear with frequency. Now consider the time delay of stand wave:

$$t_d = \frac{\theta_1}{f}$$

where f is the frequency of traveling wave. Since θ_1 is not linear with frequency f, t_d is no longer constant. Thus, traveling waves of different frequencies have different time delays going through the same waveguide. In other words, traveling waves of different frequencies have different velocity. This means there is dispersion in medium due to reflection.

How significant the dispersion is depends on reflection. When reflection is weak, $|\Gamma| << 1$, stand wave phase θ_1 is very close to phase of incident wave θ_0. On the other hand, when reflection is strong, the magnitude of reflection wave can be significant and phase of stand wave can deviate from incident wave, inducing strong non-linear phase delay (with respect to frequency).

Effect of non-linear phase delay on modulated tone

In the above, we discussed about the effect of non-linear phase delay on single tone. For modulated tone, now consider a baseband sinusoidal wave $s(t)=\cos\omega_b t$ modulated by carrier $c(t)=\cos\omega_c t$. The modulated signal can be expressed as:

$$A(t)= \cos\omega_b t\cdot \cos\omega_c t=0.5*(\cos(\omega_c+\omega_b)t+\cos(\omega_c-\omega_b)t)$$

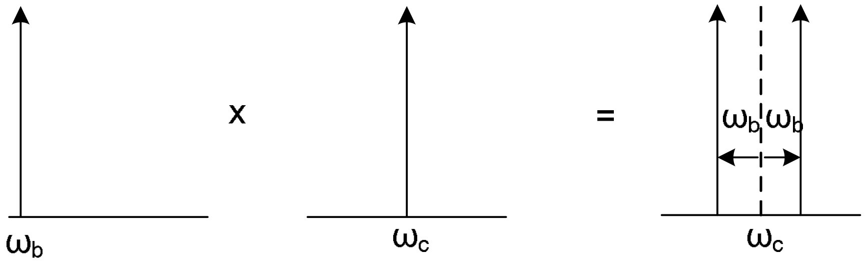


Fig. 3.13 Frequency moving by mixing

Thus, by modulation we have two tones around carrier frequency. Then the modulated signal goes through waveguide and appears at RX side (properly with reflection). Suppose at RX side the effective phase delay for frequency ($\omega_c+\omega_b$) is θ_h, while for frequency ($\omega_c-\omega_b$) is θ_l. On the other hand, the amplitude for frequency ($\omega_c+\omega_b$) is A_h, while for frequency ($\omega_c-\omega_b$) is A_l.

$$R(t)= A_h \cdot\cos[(\omega_c+\omega_b)t+\theta_h]+ A_l \cdot\cos[(\omega_c-\omega_b)t+\theta_l]$$

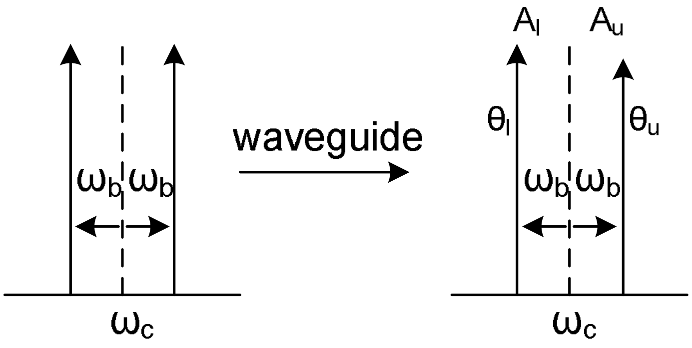


Fig. 3.14 Upconverted signal after waveguide

In RX, we use in-phase carrier with phase θ_{rx} to demodulate the signal. The demodulated signal can be expressed as:

$$R\_mixer\_I(t)= \{A_h \cdot\cos[(\omega_c+\omega_b)t+\theta_h]+ A_l \cdot\cos[(\omega_c-\omega_b)t+\theta_l]\}\cdot\cos(\omega_c t+\theta_{rx})$$

$$=0.5A_h*\{\cos(\omega_b t+\theta_h-\theta_{rx})+ \cos[(\omega_b+2\omega_c)t+\theta_h+\theta_{rx}]\} + 0.5A_l*\{\cos(\omega_b t-\theta_l+\theta_{rx})+ \cos[(-\omega_b+2\omega_c)t+\theta_l+\theta_{rx}]\}$$

$$R\_LPF\_I(t)= A_h\cdot\cos(\omega_b t+\theta_h-\theta_{rx})+ A_l\cdot\cos(\omega_b t-\theta_l+\theta_{rx})$$

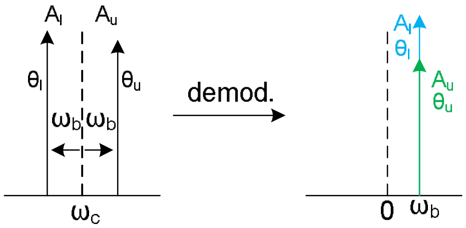


Fig. 3.15 Downconversion

Similarly, for quadrature carrier with phase $\pi/2+\theta_{rx}$, the interference due to in-phase component can be expressed as:

$$R\_LPF\_I(t)= A_h\cdot\sin(\omega_b t+\theta_h-\theta_{rx})- A_l\cdot\sin(\omega_b t+\theta_{rx}-\theta_l)$$

We want to find a phase so that signal strength after in-phase demodulation can be maximal while interference to quadrature demodulation can be minimal. This optimal phase θ_{rx} is found by vector calculation:

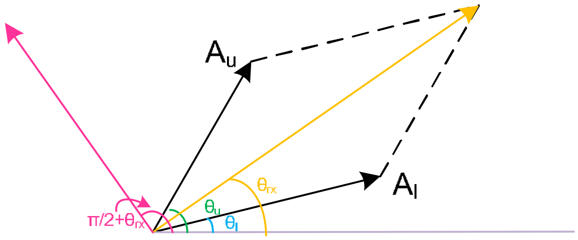


Fig. 3.15 Vector calculation of optimal RX phase

For waveguide with same A_h and A_l (amplitude of frequency response is flat within bandwidth of modulated signal), $\theta_{rx}=(\theta_h+\theta_l)/2$. If the waveguide is further without dispersion (linear phase delay), then phase delay for carrier frequency ω_c is $(\theta_h+\theta_l)/2$ (since $\omega_c=[(\omega_c+\omega_b)+(\omega_c-\omega_b)]/2$, with linear phase delay with frequency $\omega_c=(\theta_h+\theta_l)/2$) (Fig. 3.16). This means that for a waveguide with flat frequency response within modulated signal bandwidth and without dispersion, the optimal carrier phase for in-phase demodulation equals to the phase delay of carrier frequency.

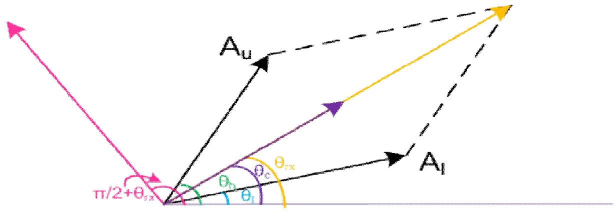


Fig. 3.16 Vector calculation for dispersion-free case

On the other hand, if the waveguide has non-ideal phase-frequency response, the optimal carrier phase to completely cancel IQ interference will deviate from carrier phase, i.e., θ_c no longer equals to optimal RX carrier phase θ_{rx}. This can be seen from Fig. 3.17.

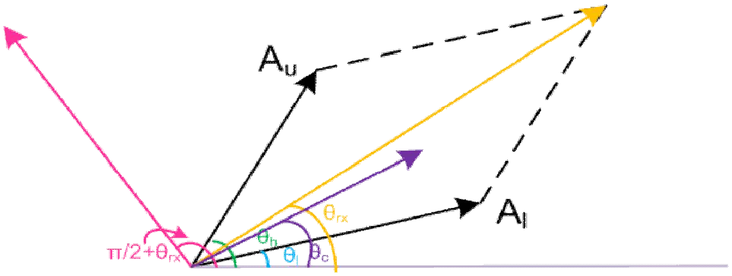


Fig. 3.17 Vector calculation for dispersive case

Generalization for complex single-tone baseband signal

For complex single-tone signal, the result is similar to that of real single-tone signal. The calculation is as follows. Suppose we have complex baseband signal of $e^{j\omega_B t}$, then the signal after mixing with in-phase carrier $(e^{j\omega_C t} + e^{-j\omega_C t})$ will become $\left(e^{j(\omega_C+\omega_B)t} + e^{-j(\omega_C-\omega_B)t}\right)$. After channel, suppose the phase shift for upper sideband is θ_h while for lower sideband is θ_l (and with flat magnitude response), then the RF signal after channel becomes $\big(e^{j[(\omega_C+\omega_B)t+\theta_h]} +$

$e^{-j[(\omega_C-\omega_B)t+\theta_l]})$. When it is down-converted by quadrature carrier $(e^{j(\omega_C t+\theta)} - e^{-j(\omega_C t+\theta)})$ and low-pass filter, the residue signal is $\left(e^{j[\omega_B t-\theta_l+\theta]} - e^{j[\omega_B t+\theta_h-\theta]}\right)$. According to Euler's formula, the IQ interference can be expressed as:

$$\cos(\omega_B t - \theta_l + \theta) + j\sin(\omega_B t - \theta_l + \theta) - \cos(\omega_B t + \theta_h - \theta) - j\sin(\omega_B t + \theta_h - \theta)$$

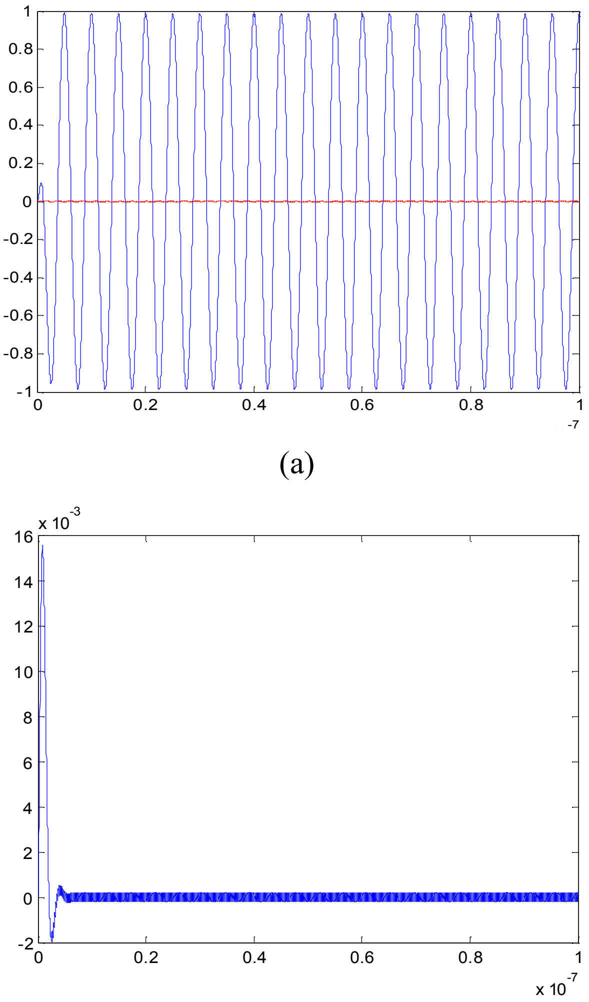
$$= -2\sin\left(\theta - \frac{\theta_h + \theta_l}{2}\right)\sin\left(\omega_B t + \frac{\theta_h - \theta_l}{2}\right) + 2j\cos\left(\omega_B t + \frac{\theta_h - \theta_l}{2}\right)\sin\left(\theta - \frac{\theta_h + \theta_l}{2}\right)$$

Its power can be expressed as:

$$0.5 \cdot \left[4\sin^2\left(\theta - \frac{\theta_h + \theta_l}{2}\right) + 4\sin^2\left(\theta - \frac{\theta_h + \theta_l}{2}\right)\right] = 4\sin^2\left(\theta - \frac{\theta_h + \theta_l}{2}\right)$$

The optimal phase of quadrature carrier to eliminate IQ interference is $(\theta_h+\theta_l)/2$. This is the same as the real signal case.

The above discussions have been verified by Matlab simulation. First, we simulate an dispersion-free channel, where $\theta_h=\pi/6$, $\theta_c=0$ and $\theta_l=-\pi/6$. We simulate two cases: (a) TX turns on I path and sends a single tone baseband signal, and turn off Q path and (b) TX turns on I path, sends DC baseband signal and turns off Q. In RX side, we use carrier phase=θ_c=0 for demodulation. The simulation result is in Fig. 3.18(a) and Fig. 3.18(b). It can be seen that in RX side, there is no I/Q interference in this case. This means the same RX demodulation phase works for both TX sends single tone and sends DC.



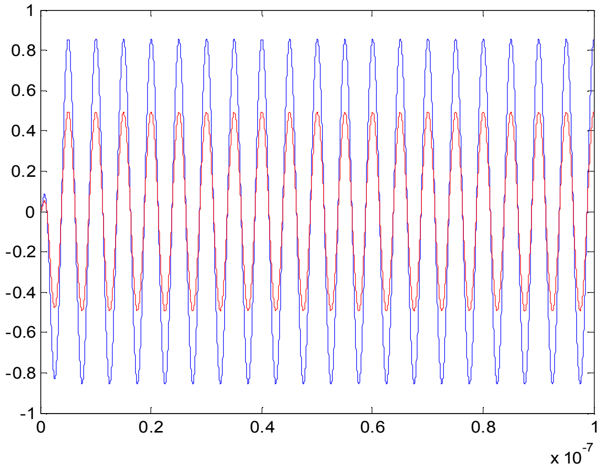
(a)

(b)

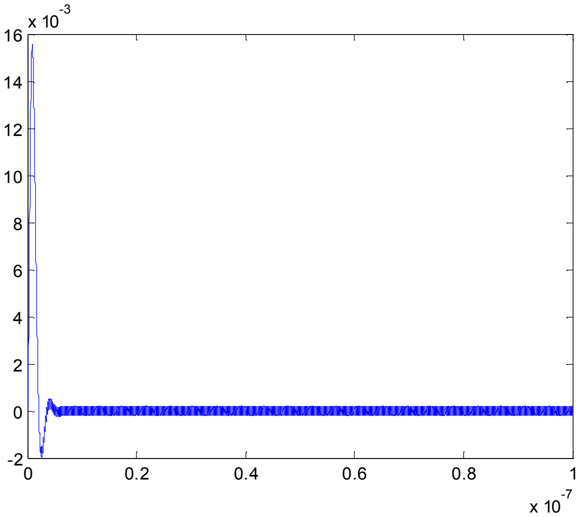
Fig. 3.18 RX demodulation result with TX (a) I sends single tone and Q off (b) I sends DC and Q off in a dispersion-free channel

Then, we simulate an dispersive channel, where $\theta_h=\pi/6$, $\theta_c=0$ and $\theta_l=-\pi/2$. We simulate two cases: (a) TX turns on I path and sends a single tone baseband signal, and turn off Q path and (b) TX turns on I path, sends DC baseband signal and turns off Q. In RX side, we use carrier phase=$\theta_c=0$ for demodulation. The simulation result is in Fig. 3.19(a) and Fig. 3.19(b). It can be seen that in RX side, there is no I/Q interference when TX sends DC baseband signal, but there is

I/Q interference when TX sends single tone baseband signal. Therefore demodulation phase of θ_c only works for the case when TX sends out DC baseband signal.



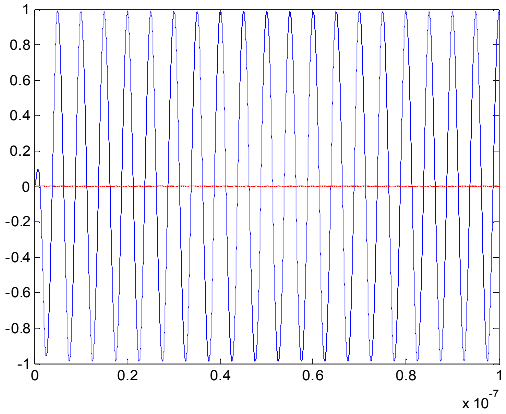
(a)



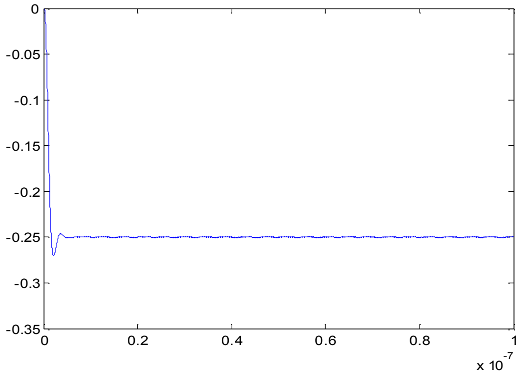
(b)

Fig. 3.19 RX demodulation result with TX (a) I sends single tone and Q off (b) I sends DC and Q off in a dispersion channel. $\theta_{rx}=\theta_c$

After that, we simulate an dispersive channel, where $\theta_h=\pi/6$, $\theta_c=0$ and $\theta_l=-\pi/2$. We simulate two cases: (a) TX turns on I path and sends a single tone baseband signal, and turn off Q path and (b) TX turns on I path, sends DC baseband signal and turns off Q. In RX side, we use carrier phase=$(\theta_h+\theta_l)/2$ for demodulation. The simulation result is in Fig. 3.20(a) and Fig. 3.20(b). It can be seen that in RX side, there is no I/Q interference when TX sends single tone baseband signal, but there is I/Q interference when TX sends DC baseband signal. Therefore demodulation phase of $(\theta_h+\theta_l)/2$ only works for the case when TX sends out single tone baseband signal.



(a)



(b)

Fig. 3.20 RX demodulation result with TX (a) I sends single tone and Q off (b) I sends DC and Q off in a dispersion channel. $\theta_{rx}=(\theta_h+\theta_l)/2$

We can also check the phase relationship of carrier in transmitted signal and carrier in RX. First we have a dispersion-free channel where $\theta_h=\pi/6$, $\theta_c=0$ and $\theta_l=-\pi/6$. TX turns on I-path with baseband single tone signal and turns off Q-path. The time domain waveform is shown in Fig.

3.21. It can be seen that optimal RX demodulation phase $\theta_{rx}=\theta_c$ and they all align well with modulated TX signal after channel.

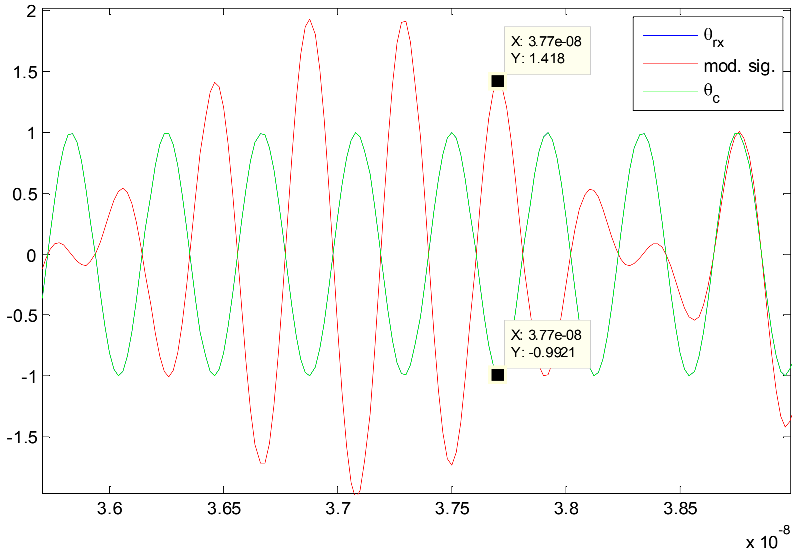


Fig. 3.21 Time-domain waveform with dispersion-free channel

On the other hand, when channel is dispersive with $\theta_h=\pi/6$, $\theta_c=0$ and $\theta_l=-\pi/2$, the optimal demodulation phase no longer equals to θ_c (Fig. 3.22). It can be seen from Fig. 3.22 that in RX side carrier with optimal demodulation phase $\theta_{rx}=(\theta_h+\theta_l)/2$ aligns well with TX modulated signal, but carrier with carrier phase θ_c does not align with modulated signal well.

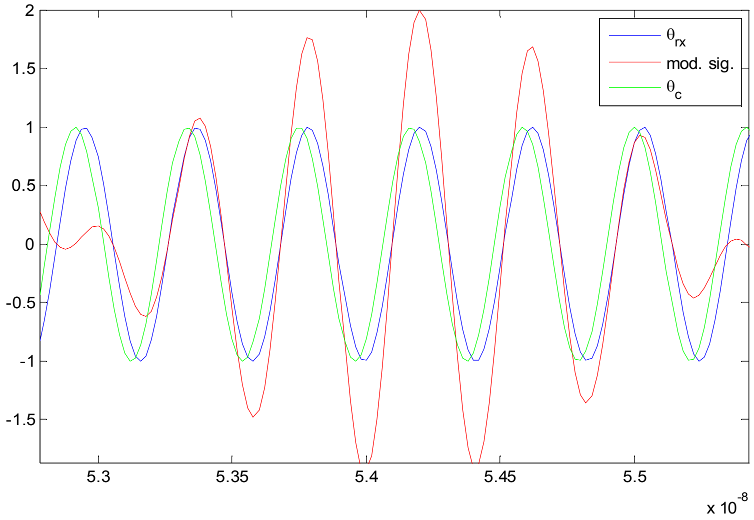


Fig. 3.22 Time-domain waveform with dispersion channel

From discussion above, we can see that the optimal phase for demodulation in dispersive channel depends on the signal power spectral density (PSD). Next we will discuss the optimal phase for a random signal, where channel condition and signal statistics in frequency domain are known *in priori*.

3.4.3 Optimal Phase for Broadband Coherent Demodulation

For single-tone signal, phase that completely cancels IQ interference at RX can be found (though it can be different from carrier phase at RX). On the other hand, if the signal has more frequency components, it is possible that the optimal phase could not completely cancel IQ interference, but reduce IQ interference power to minimum level.

Optimal phase for two-tone signal

Let's first consider simple two-tone case. Suppose we have signal with two tones at ω_{b1} and ω_{b2}. Further assume the magnitude response is flat (unity) for interested band. The received RF signal can be expressed as:

$$R(t)= \cos[(\omega_c+\omega_{b1})t+\theta_{h1}]+ \cos[(\omega_c-\omega_{b1})t+\theta_{l1}] + \cos[(\omega_c+\omega_{b2})t+\theta_{h2}]+ \cos[(\omega_c-\omega_{b2})t+\theta_{l2}]$$

After downconversion and LPF, the residue IQ interference is:

$$R_{LPF-I(t)} = \sin(\omega_{b1}t + \theta_{h1} - \theta_{rx}) - \sin(\omega_{b1}t + \theta_{rx} - \theta_{l1}) + \sin(\omega_{b2}t + \theta_{h2} - \theta_{rx}) - \sin(\omega_{b2}t + \theta_{rx} - \theta_{l2})$$

$$= 2\sin\left(\frac{\theta_{h1} + \theta_{l1}}{2} - \theta_{rx}\right)\cos\left(\omega_{b1}t + \frac{\theta_{h1} - \theta_{l1}}{2}\right) + 2\sin\left(\frac{\theta_{h2} + \theta_{l2}}{2} - \theta_{rx}\right)\cos\left(\omega_{b2}t + \frac{\theta_{h2} - \theta_{l2}}{2}\right)$$

The power of IQ interference is:

$$P_{LPF-I(t)} = 2\sin^2\left(\frac{\theta_{h1} + \theta_{l1}}{2} - \theta_{rx}\right) + 2\sin^2\left(\frac{\theta_{h2} + \theta_{l2}}{2} - \theta_{rx}\right)$$

In order to minimize interference power, its derivative with regard to θ_{rx} should be zero:

$$\frac{dP_{LPF-I(t)}}{d\theta_{rx}} = 4\sin\left(\frac{\theta_{h1} + \theta_{l1}}{2} - \theta_{rx}\right)\cos\left(\frac{\theta_{h1} + \theta_{l1}}{2} - \theta_{rx}\right) + 4\sin\left(\frac{\theta_{h2} + \theta_{l2}}{2} - \theta_{rx}\right)\cos\left(\frac{\theta_{h2} + \theta_{l2}}{2} - \theta_{rx}\right)$$

$$= 2\sin(\theta_{h1} + \theta_{l1} - 2\theta_{rx}) + 2\sin(\theta_{h2} + \theta_{l2} - 2\theta_{rx}) = 0 \rightarrow \theta_{rx} = \frac{\theta_{h1} + \theta_{l1} + \theta_{h2} + \theta_{l2}}{4}$$

Under those assumptions, the phase at RX for minimizing interference power is just the average of θ_h and θ_l of two tones.

Optimal phase for signal with arbitrary power spectral density

For arbitrary wide-sense-stationary random process x(t), we can use Fourier transform for random process:

$$X(\omega) = \int_{-\infty}^{+\infty} x(t)e^{-j\omega t}dt$$

where X(ω) is a random variable with its variance proportional to power spectral density (PSD) of random process x(t). Thus x(t) can be expressed as Riemann integral of X(ω) [3.6]:

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(\omega)e^{j\omega t}d\omega$$

For every differential element $X(\omega)e^{j\omega t}d\omega$, it is up-converted by in-phase carrier, then pass the channel (with phase shift), and down-converted by quadrature carrier. The interference signal can be expressed as:

$$x_i(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(\omega)\cdot\left(e^{j[-\theta_l+\theta]} - e^{j[\theta_h-\theta]}\right)e^{j\omega t}d\omega$$

According to Parseval's theorem, the average interference power P_i will be:

$$P_i = \lim_{T\to\infty}\frac{\int_{-T}^{T}|x_i(t)|^2dt}{2T} = \lim_{T\to\infty}\frac{1}{2\pi\cdot 2T}\int_{-\infty}^{+\infty}\left|X(\omega)\cdot\left(e^{j[\theta_l(\omega)+\theta]} - e^{j[\theta_h(\omega)-\theta]}\right)\right|^2 d\omega$$

$$= \lim_{T\to\infty}\frac{1}{2\pi\cdot 2T}\int_{-\infty}^{+\infty} 4|X(\omega)|^2\sin^2(\theta - \frac{\theta_h(\omega)+\theta_l(\omega)}{2})d\omega$$

Let $\theta_{hl}(\omega) = [\theta_h(\omega) + \theta_l(\omega)]/2$, then the mean average power then is equal to:

$$E[P_i] = \lim_{T\to\infty}E\left[\frac{\int_{-T}^{T}|x_i(t)|^2dt}{2T}\right] = \lim_{T\to\infty}\frac{1}{2\pi\cdot 2T}\int_{-\infty}^{+\infty} E[|X(\omega)|^2]\sin^2(\theta - \theta_{hl}(\omega))d\omega$$

Using Wiener-Khinchin theorem, the power spectral density S(ω) just equals to $\lim_{T\to\infty}\frac{E[|X(\omega)|^2]}{2T}$[3.7]. Thus, the mean average interference power can be expressed with respect to the power spectral density S(ω) of baseband signal:

$$E[P_i] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} S(\omega)\sin^2(\theta - \theta_{hl}(\omega))d\omega$$

Since $\theta_{hl}(\omega)$ (which is constant channel characteristic) and S(ω) (which is statistical characteristic of baseband channel) can be known *a priori*, it is possible to obtain a optimal demodulation carrier phase θ to minimize mean average power. The optimal phase θ can be calculated by set derivative of $E[P_i]$ w.r.t. θ to zero:

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} 2S(\omega)\sin 2[\theta - \theta_{hl}(\omega)]d\omega = 0$$

This is a transcendental equation. For linear phase delay (channel without dispersion), $\theta_{hl}(\omega)= \theta_{hl}$ is constant, then the solution of θ is $\theta_{hl}= (\theta_h+\theta_l)/2$ which agrees with the result in 4.2. Under the assumption of weak dispersion, i.e., $\theta_{hl}(\omega)$ only changes weakly with ω, the optimal θ will be very close to $\theta_{hl}(\omega)$. Consequently we can assume $\sin 2[\theta-\theta_{hl}(\omega)]\approx 2[\theta-\theta_{hl}(\omega)]$. So the optimal θ is:

$$\theta \approx \frac{\int_{-\infty}^{+\infty} S(\omega)\theta_{hl}(\omega)d\omega}{\int_{-\infty}^{+\infty} S(\omega)d\omega}$$

Therefore it is possible to analytically obtain optimal RX demodulation carrier phase in a weakly dispersive channel, assuming signal statistics and channel frequency response is known. In order to measure $\theta_{hl}(\omega)$, we can send single tone baseband signal with frequency ω in TX I path and turn off TX Q path, and sweep optimal demodulation carrier phase in RX for each ω.

Chapter Conclusion

This chapter introduces carrier synchronization of MRFI. The phase-based algorithm was proposed and simulation-verified by Yilei Li, the chip was designed and tested by teamwork of Weihan Cho, Yilei Li, Yanghyo Kim, Yuan Du, Jieqiong Du, Chien-Heng Wong and Sheau-Jiung Lee. The analysis of optimal demodulation phase in dispersive channel was performed by Yilei Li.

CHAPTER 4

EXPERIMENT RESULTS OF MRFI WITH CARRIER SYNCHRONIZATION

The RTL code of FSM was simulated with SPICE netlist of the MRFI system by using mixed-signal simulation tools. The carrier frequency in simulation is 3 GHz, and the delay-line resolution is 1 ps. The FSM runs at 100 MHz clock, and the bottleneck of calibration time is LPF settling time. The simulation results of both LPF and comparator outputs throughout the carrier synchronization process are shown in Fig. 4.1. During the phase tuning process, I-path is first turned off in TX. When the FSM is at the phase tuning state, the I-path filter output decreases from +220 mV and crosses 0 mV roughly at 950 ns, and renders the comparator output jumping from 1 to 0 at the same time. When FSM detects the comparator output change, it stops the fine-tuning of phase and records the resultant phase code. Since it is the I-path LPF output instead of the Q-path that crosses zero, no I/Q swap is required. On the other hand, Q-path LPF output reaches its maxima at this specific phase code, which validates our analysis in the previous section. After finishing phase tuning, the FSM begins the process for polarity calibration by transmitting 1 via both I- and Q-path. The output of LPF becomes maxima for both paths, and comparator output becomes 1 for both paths as well. Correspondingly, polarity swap bits are set to 0. The simulated LPF and comparator output eye diagrams before and after phase synchronization are shown in Fig. 4.2. Initially, the phase error is near 45° and I/Q signals would cancel each other if they have opposite polarities. After synchronization, simulated LPF output/comparator outputs are nearly ideal.

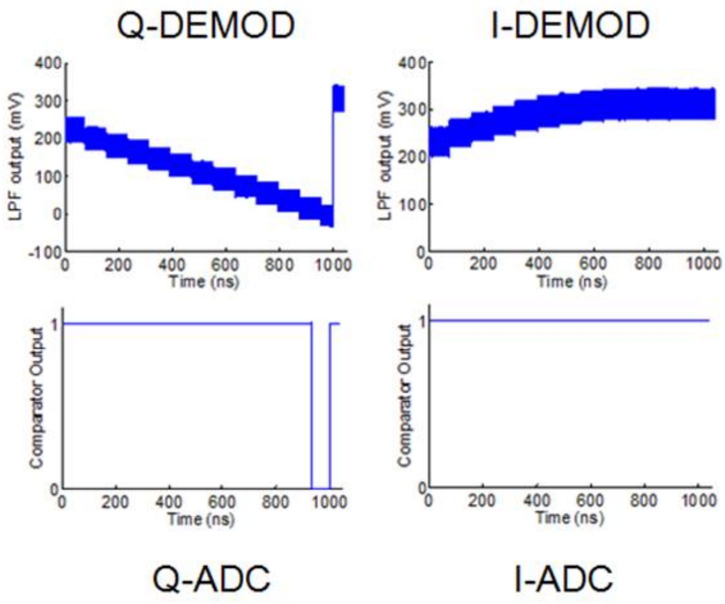


Fig. 4.1 Simulated LPF and ADC output during carrier synchronization

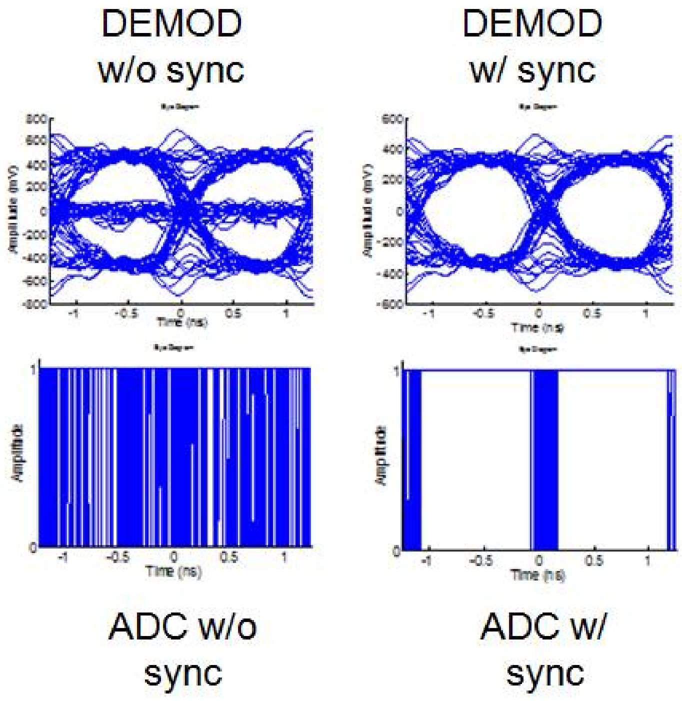


Fig. 4.2 Simulated eye diagram before and after carrier syncrhonization

The MRFI chip [2.9] has been implemented with TSMC 28nm HPC technology. The whole test chip occupies 2 mm x 1.5 mm. The test chip is pad limited, and each 10Gbps data link only occupies 80 µm x 100 µm (Fig. 4.3). The total power consumption of 4-lane MRFI transceivers (including carrier generation) is 38 mW. The chip is configured by UART serial communication protocol.

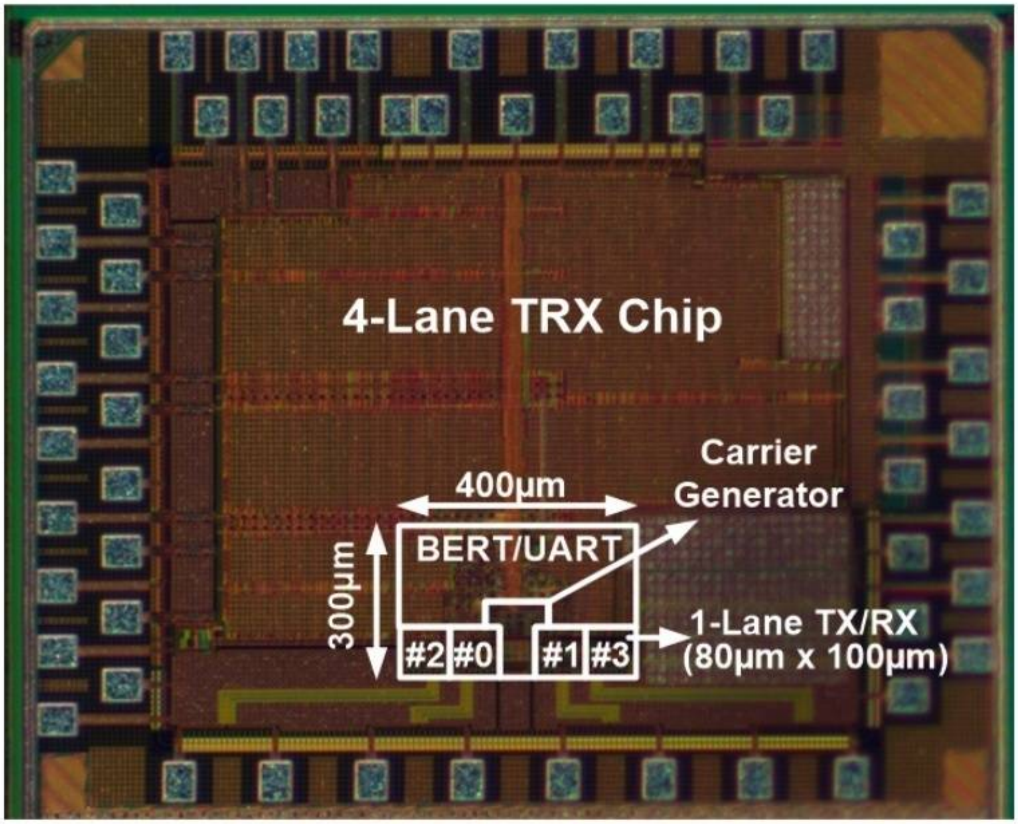


Fig. 4.3 Chip micrograph of MRFI chip

The test board is shown in Fig. 4.4. There is a 2-inch FR4 wire interconnect between one TX chip and one RX chip.

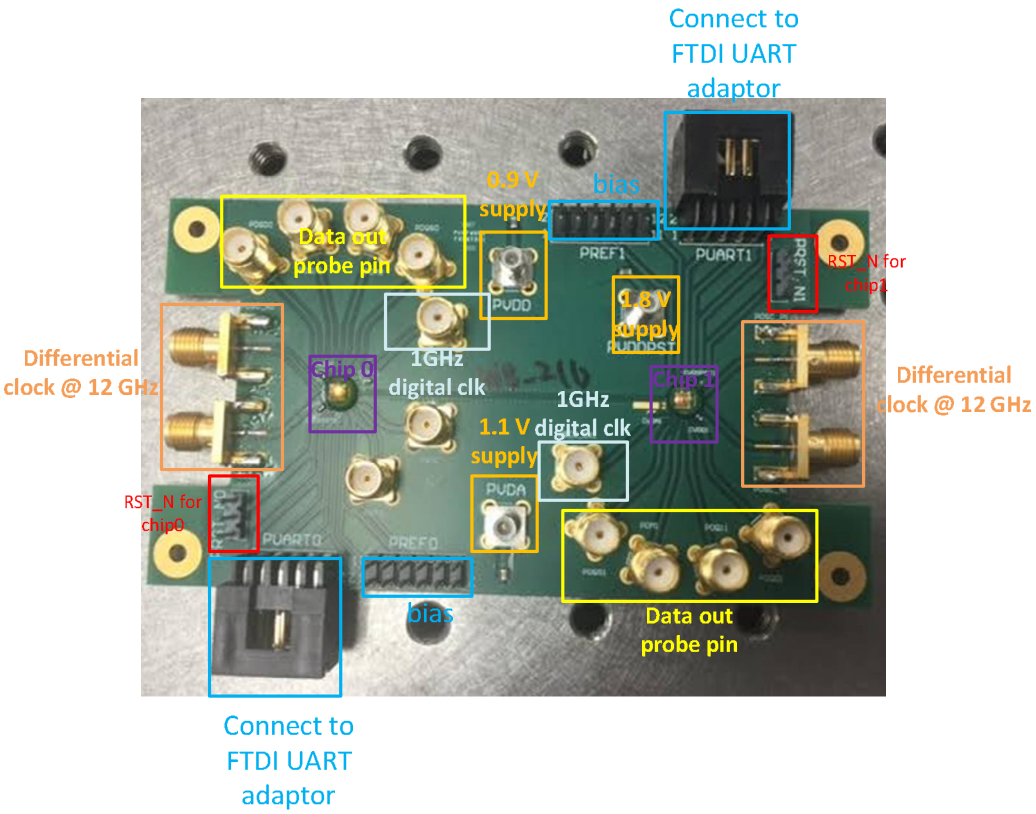


Fig. 4.4 Test board of MRFI interconnect

The RX carrier phase is controlled by FSM. We measure output eye diagram before and after carrier synchronization. The RX output is the output of ADC (comparator). Before carrier synchronization, the eye cannot be observed in the eye diagram of RX output (Fig. 4.5(a)). This is due to I/Q interference. Without carrier synchronization, bit-error rate is very high, and this MRFI transceiver without carrier synchronization cannot be used for data communication. After carrier synchronization, we can observe nearly ideal eye diagram at RX output (Fig. 4.5(b)).

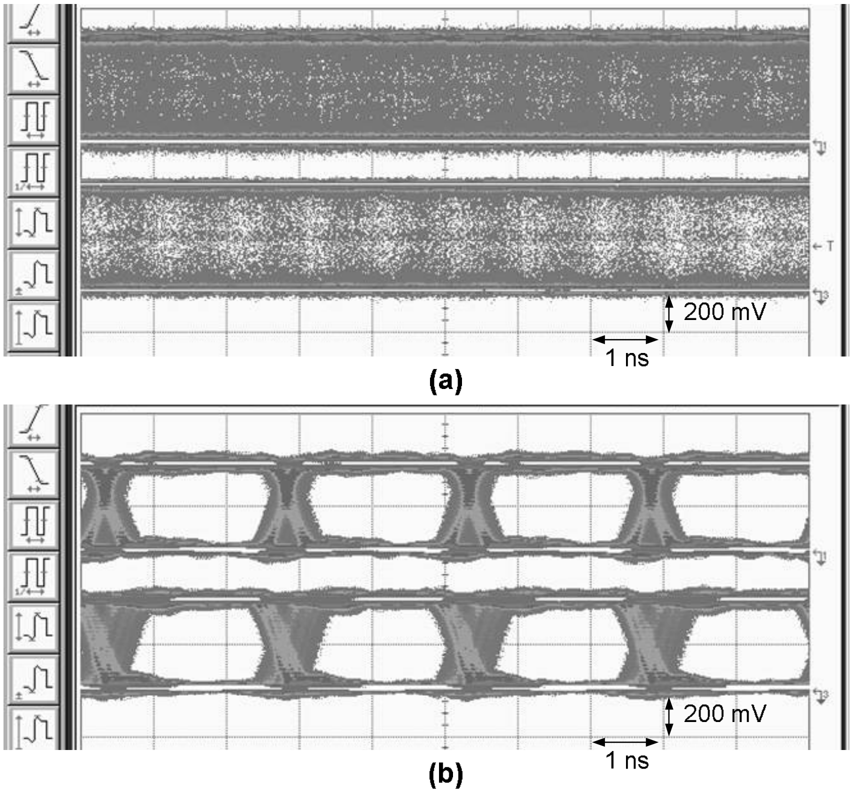


Fig. 4.5 RX output eye diagram (a) without synchronization and (b) with synchronization

Chapter Conclusion

This chapter introduces test results of MRFI. The chip was designed and tested by teamwork of Weihan Cho, Yilei Li, Yanghyo Kim, Yuan Du, Jieqiong Du, Chien-Heng Wong and Sheau-Jiung Lee.

Chapter 5

BACKGROUND OF TERAHERTZ IMAGING SYSTEM

5.1 APPLICATIONS OF TERAHERTZ (THZ) BAND EM WAVES

EM wave was first described by J. C. Maxwell in 1873, and then verified by H. Hertz with experiment. In 1895, W. Roentgen discovered X-ray, which is an EM wave with frequency between 10^{17}-10^{20} Hz. Today, EM wave from 10^{2}-10^{24} Hz has been well studied and understood (Fig. 5.1) [5.1].

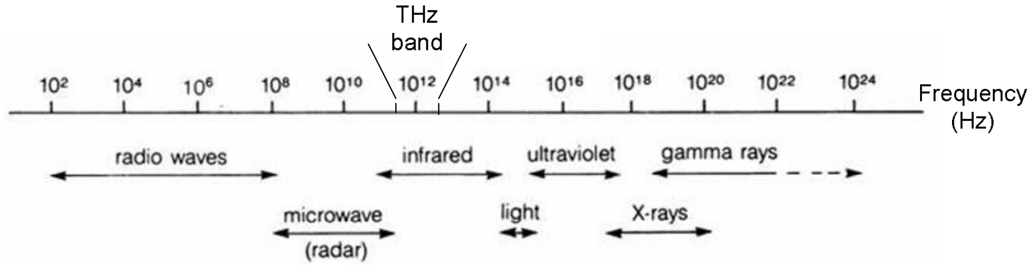


Fig. 5.1 Frequency bands of EM wave

EM wave has been extensively used for various applications. Radio waves (10^{2}-10^{8} Hz) are usually used for radio communication. Microwave and millimeter wave (10^{8}-10^{11} Hz) can be used for object detection (radar) and high speed communication. Also water resonating frequency lies in microwave band, and microwave oven takes advantage of heat produced in resonating for cooking. Infrared and visible light bands (10^{11}-10^{15} Hz) are mostly used for imaging. For instance, human eye uses EM wave in visible light band for imaging. For frequency band above ultraviolet ($>10^{15}$ Hz), EM wave has significant interaction with organic tissues. Ultraviolet EM wave can eliminate certain bacteria. X-ray and gamma ray are used for medical imaging and treatment. However, since X-ray and gamma ray are ionizing radiation, the total exposure amount of human body to EM wave in such high frequency band needs to be strictly controlled.

Terahertz is frequency band between 300 GHz and 3 THz, which lies in infrared band. Terahertz band has unique features that make it highly useful in imaging applications. First, it can penetrate common materials, such as fabric, water and skin. In addition, EM wave in THz is non-ionizing radiation. Therefore it is safe to use THz on human body.

THz wave has been used for security screening [5.2], medical imaging [5.3] and archeology [5.4]. Since THz wave can penetrate clothes and other fabrics, hidden weapons under clothes can be detected by THz wave (Fig. 5.2). In addition, THz wave security screening is safer than X-ray as THz wave does no harm to human body.

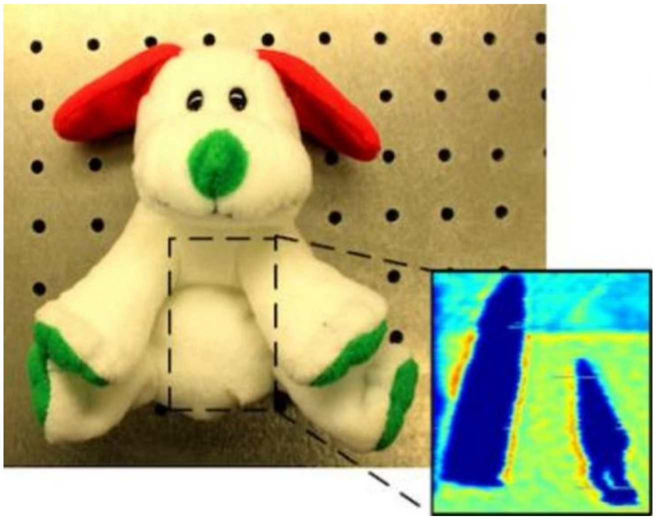


Fig. 5.2 THz security screening

THz wave also finds application in medical imaging. Unlike X-ray, THz wave can penetrate skin without any damage to human tissue. Therefore it is useful for imaging tissues under skin. It is also useful for dentistry, as a potential substitute for X-ray (Fig. 5.3).

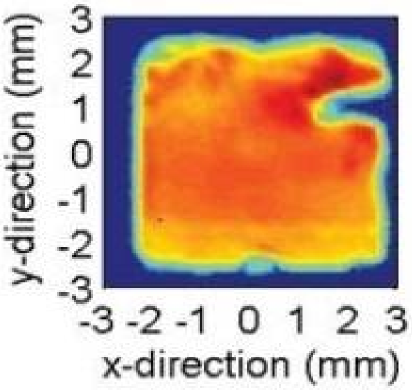


Fig. 5.3 THz imaging for dentistry

In addition, THz wave is useful for archeology. Instead of painting new work on a new canvas, artist in medieval or renaissance era commonly painted their new work on top of old work on the same canvas. Therefore the old work is lost as they have been overridden by new works. This action is especially common for frescos. THz EM wave can help us see the old works underneath the new painting (Fig. 5.4), which has been proved to be extremely useful for history studies.

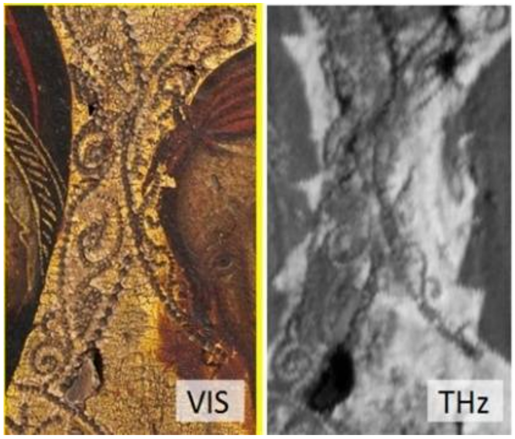


Fig. 5.4 THz application in history studies

5.2 CONVENTIONAL THZ SOURCE AND DETECTORS

Conventional, titanium-sapphire (Ti:Sapphire) based laser is used as THz sources. In order to tune the frequency of source, photo-mixing technique was proposed in [5.5]. The output frequency of laser source depends on the bandgap of material. Photo-mixing takes two continuous wave (CW) lasers with frequency of f1 and f2, and uses interference to generate frequency component at |f1-f2|.

In order to implement laser in semiconductor, quantum cascade laser (QCL) has been proposed [5.6]. Quantum well structures are formed in semiconductor, which determines the spacing of multiple energy sub-bands. Unlike Ti:Sapphire laser, the output frequency of QCL can be designed by tuning the quantum well structures, which brings flexibility to QCL THz source.

An alternative source for low-end of THz band operation is Gunn diode [5.7]. Compared with lasers, Gunn diode is more available in commercial market and is more compact.

For THz detector, conventional solution uses FTIR bolometer or III-V Schottky diode. When EM radiation reaches bolometer, the power of radiation is absorbed by absorptive material inside bolometer and transformed into heat. The heat increases the temperature of absorptive material in bolometer, and therefore the radiation power can be measured from the change of temperature. In order to have good sensitivity, bolometer needs to work at extremely low temperature. This makes bolometer expensive and bulky. Compared with bolometer, Schottky diode has much smaller form factor, but the cut-off frequency is much lower than bolometer.

In short, the conventional THz sources and detectors are bulky and expensive. In order to make THz technique popular, a THz source/detector system with compact size and cost-effective design is necessary.

5.3 CMOS THZ TECHNOLOGY

Nowadays, the most widely used device technology for semiconductor is CMOS technology. CMOS technology uses silicon as substrate, and is cost effective. In addition, the feature size of

CMOS device keeps scaling according to Moore's Law. With feature size scaling, the CMOS devices can run faster, and more CMOS devices can be integrated into the same chip. Therefore if we can implement THz source and detector with CMOS technology, the cost of THz source and detector will be quite low. In addition, analog and digital signal processing blocks can be integrated together with CMOS THz source/detector to form a powerful THz system. The question is, is it possible to implement CMOS THz system? If it is possible, how to design a CMOS THz source/detector system?

The answer for first question is yes. With feature size scaling, the cut-off frequency of CMOS devices grows rapidly. The cut-off frequencies of various CMOS technology nodes are shown in Fig. 5.5 [5.8]. It can be seen that after 45-nm technology node, the cut-off frequency of CMOS devices is already in THz band.

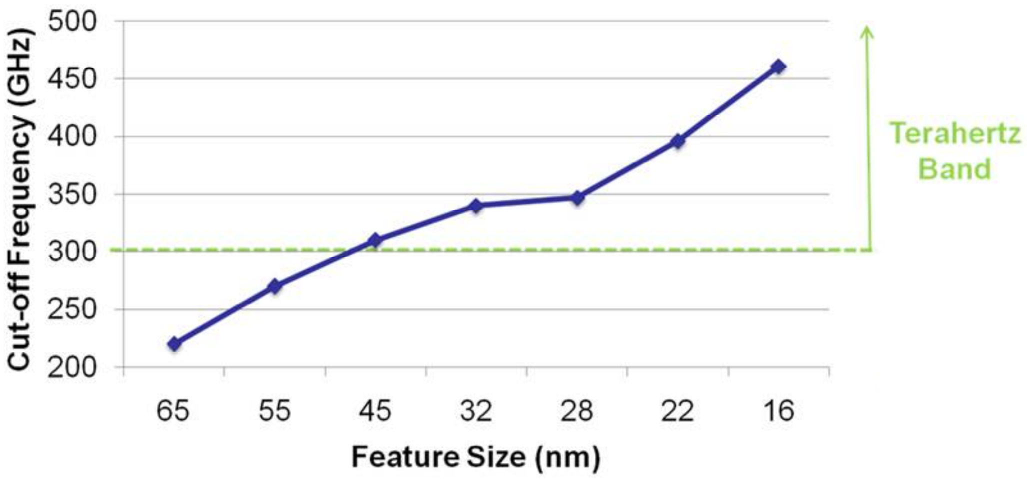
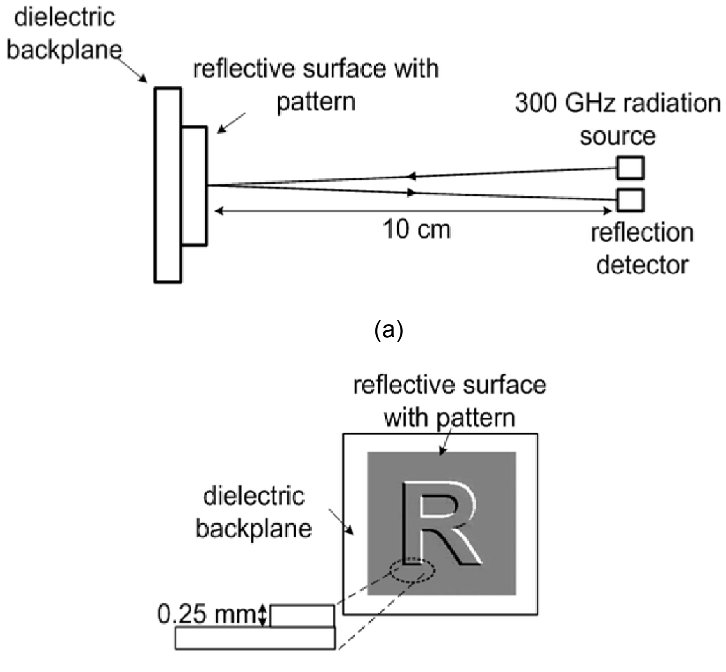


Fig. 5.5 Cut-off frequencies of various technology nodes

The second question is harder to answer. This is due to the fact that CMOS THz systems are not easy to design. CMOS THz system has several challenges.

The first challenge of CMOS THz imaging system is how to obtain surface imaging. The state-of-art CMOS detectors use power imaging [5.9]. Power imaging works for surface imaging if the wavelength of EM wave for imaging is small enough. For example, our eye uses power imaging, and the outline of surface can be detected successfully with EM wave in visible light frequency band. However, with EM wave in THz band, it is almost impossible to identify detailed surface information from power imaging. This is mainly due to the fact that the path loss change due to surface roughness is too small.

As an example, one setup for power imaging is shown in Fig. 5.6.



(a)

(b)

Fig. 5.6 (a) Setup of surface power imaging (b) Surface pattern

In setup, a reflective surface is mounted on a dielectric (non-reflective) backplane. The distance between reflector and THz source/detector is 10 cm, while the surface roughness height of surface pattern is 0.25 mm. The THz source/detector works at 300 GHz. The simulation result of power imaging is shown in Fig. 5.7. It is clear that only the outline between reflector can be seen, but the surface pattern cannot be detected. This is due to the fact that the path loss due to surface roughness height is less than 0.1%, and therefore cannot show in imaging result.

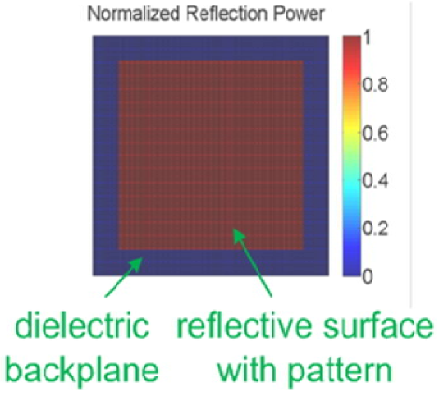


Fig. 5.7 Simulated imaging result of power imaging. The path loss is assumed to be square-law.

The second challenge is low cut-off frequency of CMOS devices. Though the cut-off frequencies of CMOS devices in advanced technology node is already in THz band, it is still not high enough. On the other hand, the resolution of THz imaging system is determined by radiation frequency. The resolution limit defined by diffraction of THz imaging system can be expressed as:

$$d = \frac{\lambda}{2NA} = \frac{c}{2NA \cdot f} \tag{5.1}$$

where d is resolution limit, λ is the wavelength of radiation, NA is numerical aperture, c is speed of light, and f is the frequency of radiation. With higher radiation frequency, resolution limit decreases, which means that better resolution can be achieved.

The third challenge of CMOS THz system is calibration problem. For a reliable THz system, routine self calibration of THz source is necessary, as the work condition and environment may change over time. For example, when supply voltage and/or environment temperature change, the oscillation frequency of THz source may change. Those variation need to be tracked by routine calibration.

However, current standard calibration setup for THz source is very bulky and expensive. Standard test equipment for THz source is FTIR system with bolometer. Since bolometer needs cooling for sensitivity, liquid helium is necessary. This makes FTIR bolometer very bulky and expensive (Fig. 5.8).

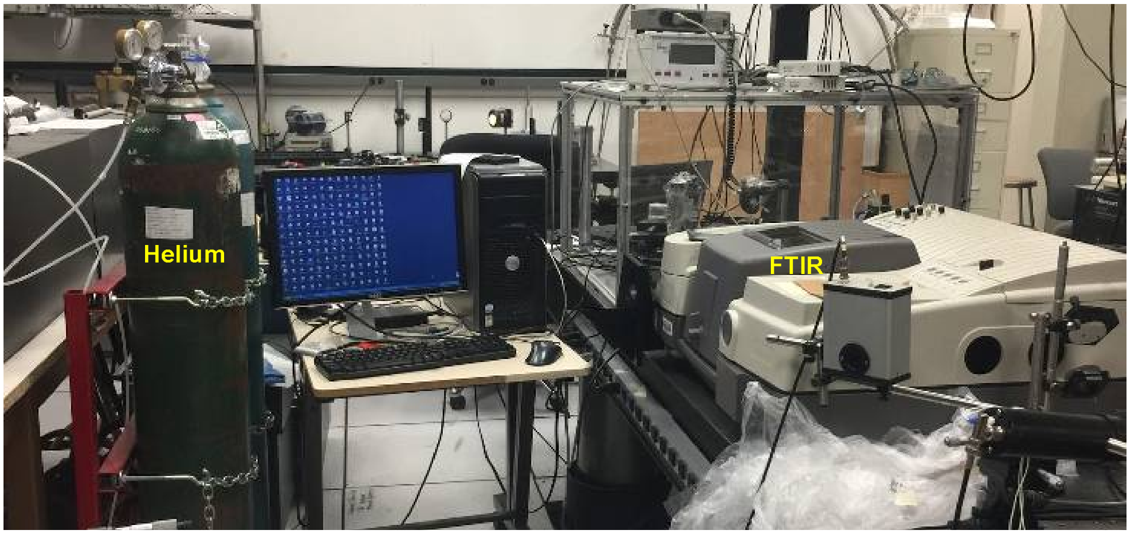


Fig. 5.8 FTIR bolometer with liquid helium

An alternative way for source characterization is harmonic mixer with spectrum analyzer. Harmonic mixer uses super harmonic of carrier to mix down radiation, and spectrum analyzer is used to observe power spectral density (PSD). The drawback of this method is that its characterization range is defined by the working range of spectrum analyzer, and it is hard to capture the wideband (e.g., 300 GHz-2 THz) PSD of radiation.

Chapter Conclusion

This chapter introduces the state-of-art THz applications and CMOS THz technology. For CMOS THz technology, we have many challenges to address.

Chapter 6

THZ INTERFEROMETRIC IMAGING

6.1 PHASE-BASED IMAGING

In Chapter 5, power imaging has been proved to be improper for surface imaging. The relative reflected power change can be expressed as:

$$\frac{\Delta P}{P} = \left(\frac{\Delta h}{D}\right)^2 \tag{6.1}$$

where Δh is the change of surface roughness height, D is the average distance between reflector and source/detector. Δh is usually smaller than 1 mm, while D is greater than 10 cm. Therefore ΔP is too weak to be identified. On the other hand, phase of EM wave is more sensitive to surface roughness. The phase difference of reflected wave due to surface roughness height change can be expressed as:

$$\Delta\theta = \frac{4\pi\Delta h}{\lambda} \tag{6.2}$$

where λ is the wavelength of radiation EM wave. As λ of THz EM wave is within the range of 0.1 mm – 1mm, the phase change is very sensitive to surface roughness height change.

The same setup in Fig. 5.6 was simulated with phase and power imaging, as shown in Fig. 6.1. For 0.25-mm surface roughness height change, phase imaging can clearly identify the surface pattern while power imaging cannot.

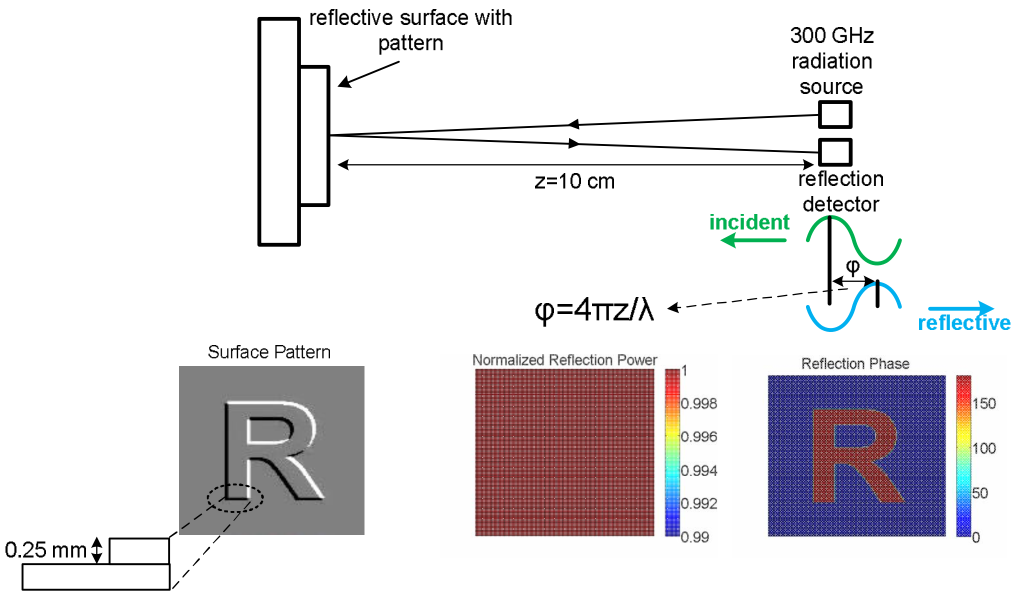


Fig. 6.1 Comparison of power and phase imaging

Conventional way for phase imaging is interferometry with beam splitter. The setup of interferometric imaging is shown in Fig. 6.2 [6.1].

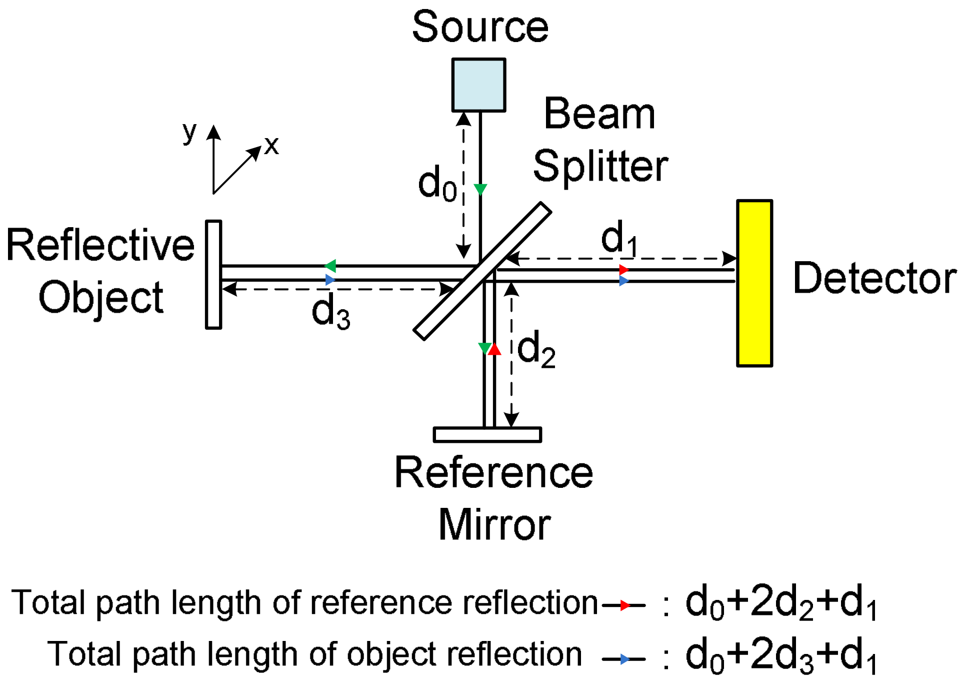


Fig. 6.2 Interferometric imaging

The beam from source is split into two identical beams by beam splitter. One beam (object reflection beam) travels to reflective object, and is reflected to detector. The other beam (reference reflection beam) travels to reference mirror, and is reflected back to detector. The total path length of reference reflection beam is $d_0+2d_2+d_1$, while the total path length of object reflection beam is $d_0+2d_3+d_1$. Suppose we use square-law detector, the output of detector is:

$$\begin{aligned} Y(d_2-d_3) &= \left\{ A_i \cos\left[\omega \mathrm{t} + \frac{2\pi(d_0+d_1+2d_2)}{\lambda} \right] + A_r \cos\left[\omega \mathrm{t} + \frac{2\pi(d_0+d_1+2d_3)}{\lambda} \right] \right\}^2 \\ &= \frac{A_i^2}{2} + \frac{A_r^2}{2} + A_i A_r \cos\frac{4\pi(d_2-d_3)}{\lambda} + A_i A_r \cos\left[2\omega \mathrm{t} + \frac{4\pi(d_0+d_1+d_2+d_3)}{\lambda} \right] \end{aligned} \tag{6.3}$$

where A_i is the amplitude of reference beam, A_r is the amplitude of object reflection beam. The output after low-pass filter (LPF) is:

$$Y_{LPF}(d_2-d_3) = \frac{A_i^2}{2} + \frac{A_r^2}{2} + A_i A_r \cos\frac{4\pi(d_2-d_3)}{\lambda} \tag{6.4}$$

It can be seen that the output of square law detector after LPF is a function of path difference of two beams. The wavelength of frequency component in radiation also plays a part in output. When distance d_3 changes due to surface roughness height change, the output of detector changes due to phase relation between object reflection beam and reference beam changes.

Equation (6.4) is the case for one-tone radiation. If multiple tones exist in radiation, the output of square-law detector after LPF can be expressed as:

$$Y_{LPF}(d_2-d_3) = \sum_{k=1}^{m} \left[\frac{A_{ik}^2}{2} + \frac{A_{rk}^2}{2} + A_{ik} A_{rk} \cos\frac{4\pi(d_2-d_3)}{\lambda_k} \right] \tag{6.5}$$

where A_{ik}, A_{rk} and λ_k are the reference beam amplitude, object reflection beam amplitude and wavelength of k-th harmonic in radiation, respectively.

If we look into (6.5) more carefully, we can see that if we keep d_2 and move d_3, measure output of Y at every d_3 and perform Fourier Transform upon Y with regards to variable d_3, multiple peaks at spatial frequencies of $2/\lambda_1$, $2/\lambda_2$, $2/\lambda_3$,… $2/\lambda_m$ will appear in Fourier Transform. The interferometer is essentially doing time-space transform, which transforms time-domain frequency component f_i in radiation into spatial-domain frequency component $2/\lambda_i$, where $\lambda_i = c/f_i$.

6.2 CMOS INTERFEROMETRIC IMAGING

Interferometric imaging can capture the surface roughness information by detecting phase change of reflective wave, which solve the challenge of surface imaging. However, it would be bulky to put a beam splitter along with CMOS THz imaging system. We need a compact design.

A compact CMOS interferometric imaging has been proposed in [6.2]. We count on two key observations (made by Dr. Al Hadi in [6.2]) for THz interferometric imaging to achieve compact design.

The first key observation is that antenna is THz band is reciprocal. With reciprocity, antenna can transmit and receive THz radiation simultaneously. In conventional interferometric imaging design, optical design is used and coupler is not reciprocal. Therefore we need to split incident beam into two beams and use coupler to receive both in detector. For THz design, we can use the same antenna for receive and transmit. Therefore beam splitter can be saved, which makes design compact. The setup diagram with reciprocal antenna for THz interferometric imaging is shown in Fig. 6.3.

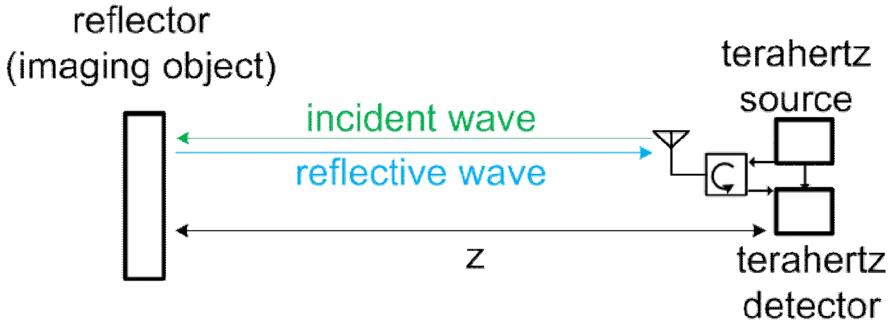


Fig. 6.3 THz interferometric imaging with reciprocal antenna

In Fig. 6.3, an antenna is used for both transmit and receive. A circulator is used to couple incident wave from source to antenna, and to couple reflective wave from antenna to detector. Reference beam is coupled from source directly to detector.

The second key observation is that CMOS THz source and detector for inteferometric imaging can be one. The second-order non-linearity of MOSFET I-V curve makes MOSFET a good square law detector.

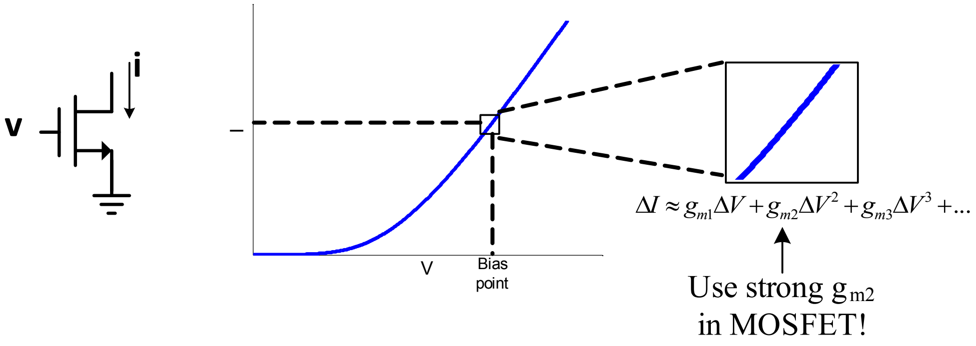


Fig. 6.4 I-V curve of MOSFET

The I-V curve of MOSFET is shown in Fig. 6.4. Around certain bias point, the incremental current ΔI with regards to incremental voltage ΔV can be approximated by Taylor expansion:

$$\Delta I \approx g_{m1}\Delta V + g_{m2}\Delta V^2 + g_{m3}\Delta V^3 + \ldots \tag{6.6}$$

where g_{mi} is i-th order nonlinearity coefficient. For MOSFET, g_{m2} is strong, and therefore it can work as a detector.

The principles of CMOS source/detector is explained as follows. First, we explain how CMOS THz source generates THz radiation. A CMOS THz triple-push oscillator is shown in Fig. 6.5 as an example in explanation.

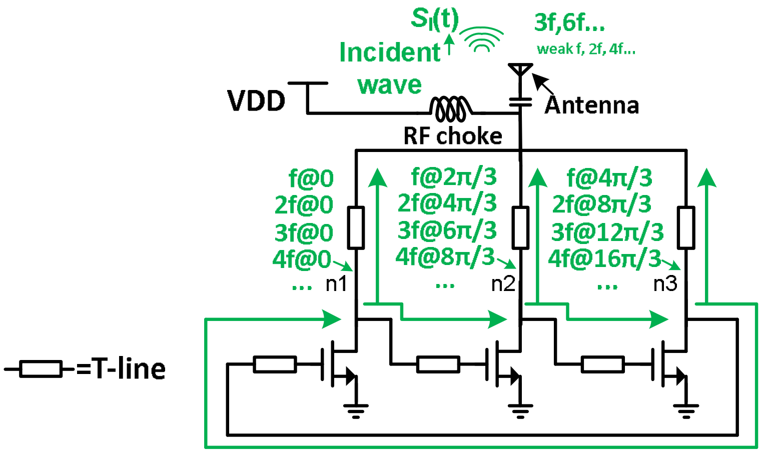


Fig. 6.5 CMOS THz source

For a triple-push oscillator, suppose the fundamental frequency of oscillator is f. In order to oscillate at f, the phase of component at f at node n1 is 0, at node n2 is 2pi/3 and at node n3 is 4pi/3. Due to the non-linearity of MOSFETs, there also exist harmonics of fundamental tone in oscillator. The second harmonic at frequency 2f has 0, 4pi/3, 8pi/3 at three nodes, third harmonic at frequency 3f has 0, 6pi/3, 12pi/3 at three nodes, and etc. The incident waves are combined at antenna. Due to phase relationships, harmonics at 3f, 6f, 9f,…will be enhanced while other harmonics are cancelled. This can be explained as follows. For the general case of n-push oscillator, the output at antenna can be expressed as:

$$\sum_{j=1}^{n}\sum_{k=1}^{m}\mathrm{A}_k \sin\left(k\omega t+\frac{2k\left(j-1\right)\pi}{n}\right)=\sum_{j=1}^{n}\sum_{k=1}^{m}\mathrm{Im}\left[A_k e^{i\left(k\omega t+\frac{2k(j-1)\pi}{n}\right)}\right]$$
$$=\sum_{k=1}^{m}\mathrm{Im}\left[A_k\sum_{j=1}^{n}e^{i\left(k\omega t+\frac{2k(j-1)\pi}{n}\right)}\right] \tag{6.6}$$

where A_k is the amplitude of k-th harmonic.

Using the formula of sum of geometric series

$$\sum_{j=1}^{n}\mathrm{e}^{i\left(k\omega t+\frac{2k(j-1)\pi}{n}\right)}=\begin{cases} ne^{ik\omega t}, \text{if } k=pn\ \ p\in Z^{0+} \\ 0, otherwise \end{cases} \tag{6.7}$$

equation (6.6) can be reduced to

$$\sum_{j=1}^{n}\sum_{k=1}^{m}\mathrm{A}_k \sin\left(k\omega t+\frac{2k\left(j-1\right)\pi}{n}\right)=\sum_{p=1}^{\lfloor m/k \rfloor} nA_k \sin\left(pn\omega t\right) \tag{6.8}$$

Therefore for triple-push oscillator, ideally only 3f, 6f, 9f remain, while other harmonics are eliminated. In real case, mismatch induces remaining frequency components at f, 2f, 4f, … at output.

At the same time the oscillator receives reflected THz radiation, as shown in Fig. 6.6.

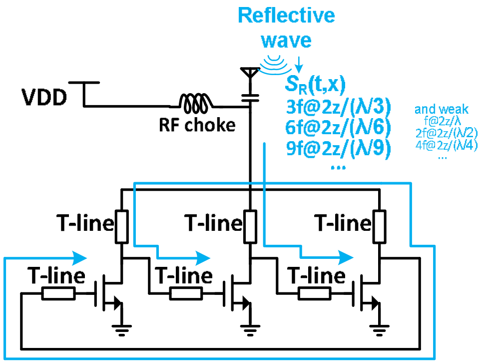


Fig. 6.6 CMOS oscillator as a THz detector

The reflective wave is received by antenna, which includes frequency components at 3f, 6f, 9f… with phase shift of 4πz/(λ/3), 4πz/(λ/6), 4πz/(λ/9)… The phase shift depends on the distance z between source and reflector (imaging object). The reflective wave is coupled to gate of MOSFETs.

Therefore at the gate of MOSFET, the total incremental voltage is the superposition of incident frequency components and reflective frequency components. The phase difference between incident wave and reflective wave is identified by second-order non-linearity of MOSFET in oscillator. The total incremental current flowing through power supply is the sum of three incremental currents of three MOSFETs (Fig. 6.7).

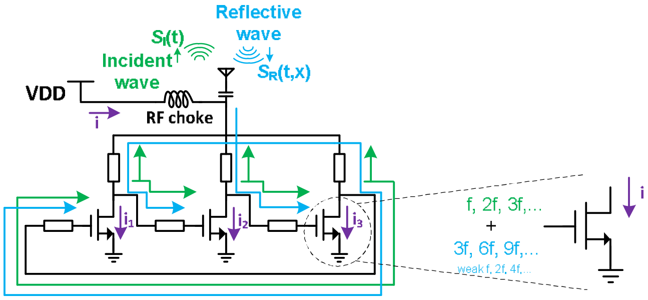


Fig. 6.7 Incremental current in CMOS source/detector

We can analyze the incremental current of one single MOSFET. The input incremental voltage at its gate is the superposition of incident wave and reflective wave. The incident wave part can be expressed as:

$$\Delta V_i = \sum_{k=1}^{3m} A_{ik} \sin\left(k\omega \mathrm{t} + \theta_{ik}\right) \tag{6.9}$$

For reflective wave part:

$$\Delta V_r = \sum_{k=1}^{m} A_{r3k} \sin\left(3k\omega \mathrm{t} + \theta_{r3k} + \frac{4\pi z}{\lambda / 3k}\right) \tag{6.10}$$

where z is the distance between reflector (imaging object) and THz source/detector. The output incremental current after second-order nonlinearity and LPF (by RF choke inductor) is:

$$i_p = \sum_{k=1}^{m}\left[\frac{A_{i3k,p}^2}{2} + \frac{A_{r3k,p}^2}{2} + A_{ik,p}A_{rk,p}\cos\left(\frac{4\pi z}{\lambda / 3k} + \theta_{r3k,p} - \theta_{i3k,p}\right)\right] \tag{6.11}$$

The incremental current is the difference of phases of harmonics between incident wave and reflective wave. The total incremental current can be expressed as:

$$i(z) = \sum_{p=1}^{3} i_p = \sum_{k=1}^{m}\left[\frac{A_{i3k}^2}{2} + \frac{A_{r3k}^2}{2} + A_{ik}A_{rk}\cos\left(\frac{4\pi z}{\lambda / 3k} + \varphi_{3k}\right)\right] \tag{6.12}$$

where φ_k is the phase difference between incident wave and reflective wave of k-th harmonic. Again, this is a time-space transform, and time-domain harmonics are transformed into spatial-domain harmonics. Those spatial-domain harmonics will appear in Fourier Transform plot of i(z) if we measure i(z) with regards to various z. Time-space transform can also be used to verify whether the CMOS interferometric imager works by Fourier Transform output of current.

A prototype CMOS interferometric imager, designed and tested by Dr. Yan Zhao and Dr. Al Hadi in [6.2], is implemented in TSMC 65nm technology. The interferometric imager is a triple-push oscillator with fundamental frequency of 117 GHz. After triple-push, the strongest tone in incident wave is 117 GHz x 3= 351 GHz. The chip micrograph is shown in Fig. 6.8.

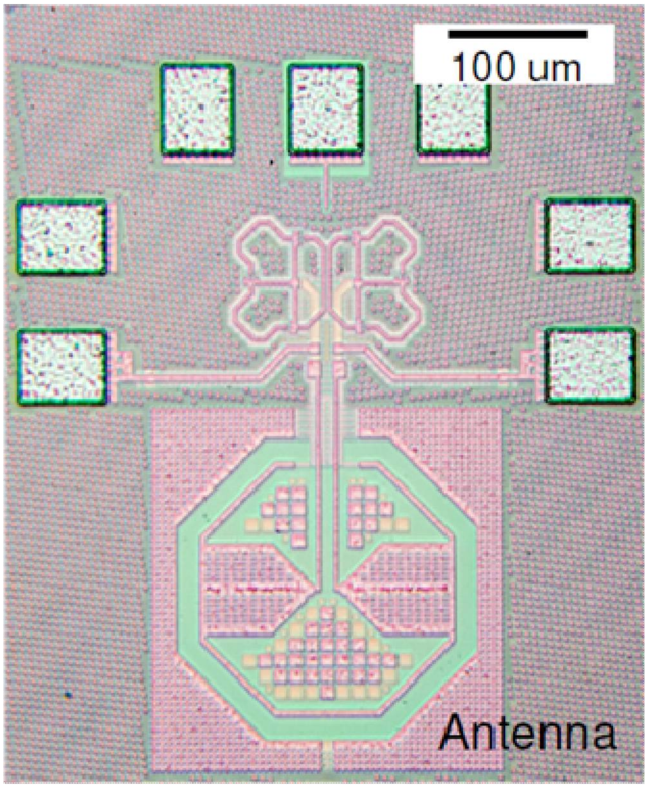


Fig. 6.8 Chip micrograph of CMOS interferometric imager

The frequency components of radiation are characterized by FTIR system, as shown in Fig. 6.9. We can see strongest tone at 351 GHz, and also harmonics at 700 GHz, 1.05 THz and 1.4 THz.

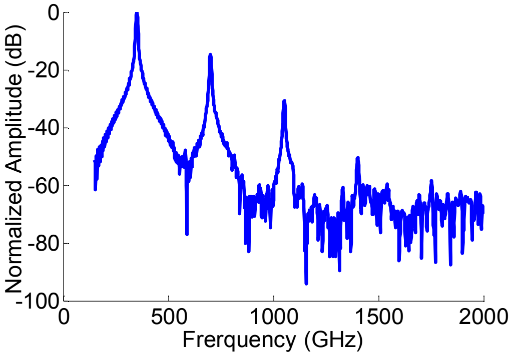


Fig. 6.9 Radiation frequency components detected by FTIR

Then we use time-space transform to verify simultaneous transmit/receive concept. The experiment was carried out by Dr. Al Hadi. The experiment setup is shown in Fig. 6.10. We put CMOS chip and board on one side, and reflector on the other side. We control x,y,z coordinator of reflector by linear stage. In order to enhance signal-to-noise ratio, and optical chopper along with lock-in amplifier is used. We keep x, y of reflector but change z. We measure $Y_{lock\_in}(z)$ with various z. If CMOS oscillator can transmit/receive at the same time and identify the phase difference between incident and reflective wave, the Fourier Transform of $Y_{lock\_in}(z)$ will have peaks at the positions corresponding to the wavelength of components in radiation.

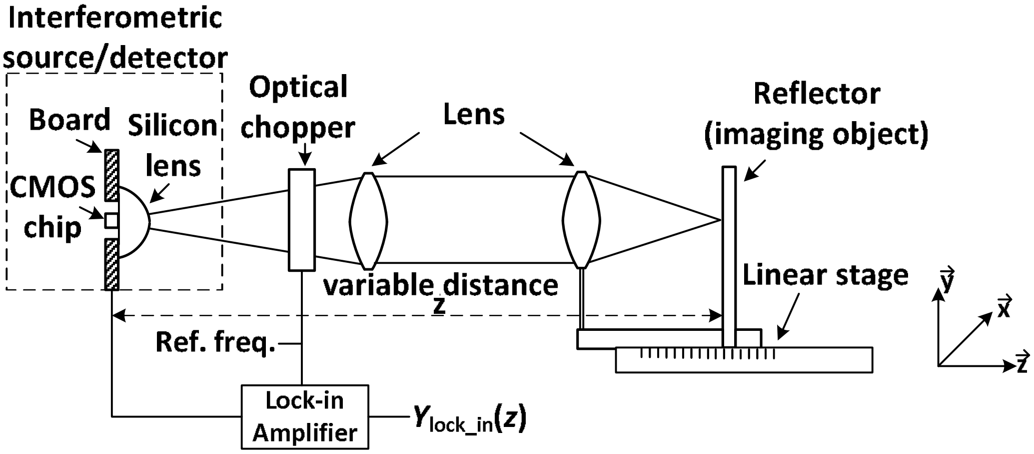


Fig. 6.10 Setup for CMOS interferometric chip experiments

The output amplitude of lock-in amplifier with various z is shown in Fig. 6.11. We can see periodic waveform which corresponding to spatial-domain harmonics.

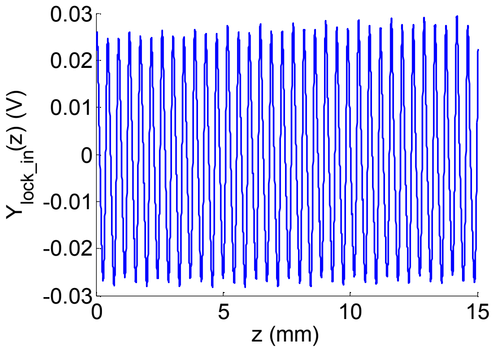


Fig. 6.11 Lock-in amplifier output

The Fourier Transform of $Y_{lock\_in}(z)$ is shown in Fig. 6.12. It is clear that multiple spatial harmonics are present in plot, which verifies that there exists time-space transform by oscillator.

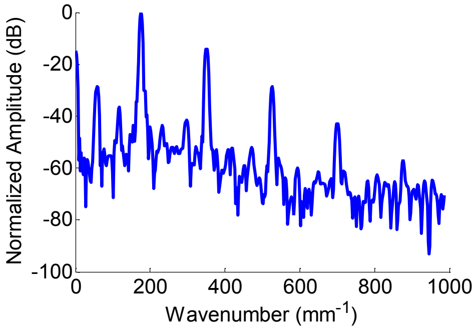


Fig. 6.12 Fourier transform plot of $Y_{lock\_in}(z)$

We can further transform back spatial-domain harmonics back to time-domain. This can be achieved by following mapping relations from wavelength λ to time-domain frequency:

$$\lambda = \frac{c}{f} \tag{6.13}$$

where c is speed of light. The mapped-back time-domain frequency PSD is shown in Fig. 6.13. The frequencies components include tones at 351 GHz, 750 GHz, 1.05 THz, 1.4 THz, which have exactly the same frequencies as FTIR result. This verifies that our CMOS oscillator can indeed work as both a THz source and detector. In addition, some weak tones at 117 GHz, 234 GHz and 1.75 THz are present. The 117 GHz and 234 GHz tones are fundamental tone and its second harmonic of triple-push oscillator. Those two tones are not detected by FTIR since the FTIR system has high-pass characteristics and those tones fall below FTIR's cut-off frequency. In addition, 1.75 THz is the 15th harmonic of fundamental tone. This tone is not detected by FTIR system since SNR of FTIR system is not as good as our interferometric system.

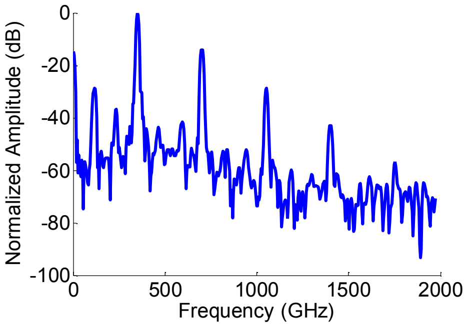


Fig. 6.13 Mapped-back time-domain frequency components plot

Then we use CMOS interferometric system for imaging. When doing imaging, we keep z coordinator but sweep x and y coordinator of object. We record output amplitude of lock-in amplifier at each (x,y) position and plot the density map of lock-in amplifier output as imaging. The imaging result is shown in Fig. 6.14. We can clearly see the outline of inscription of the object, a 50 cent coin. Since the strongest tone in radiation is 351 GHz tone, the resolution of imaging is determined by this frequency.

optical imaging

350 GHz imaging

Fig. 6.14 Interferometric imaging result of a 50 cent coin

Chapter conclusion

This chapter introduces and analyzes CMOS THz interferometric imaging. The concept of CMOS interferometric imaging was proposed by Dr. Richard Al Hadi, the chip was designed and tested by Dr. Richard Al Hadi and Dr. Yan Zhao, and the analysis was performed by Yilei Li.

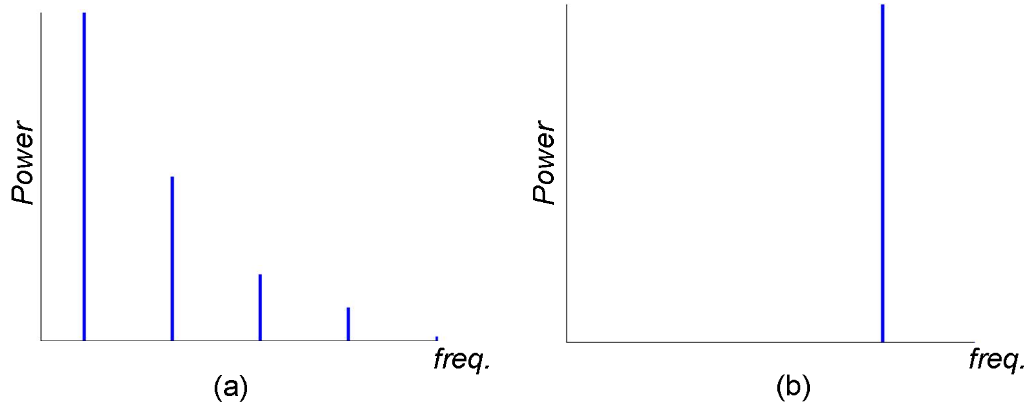
CHAPTER 7

SPATIAL HARMONIC EXTRACTION (SHE) TECHNIQUE

7.1 BACKGROUND

The low cut-off frequency of CMOS is a bottleneck for CMOS THz systems. Beyond cut-off frequency, the gain of MOSFET is not high enough to maintain oscillation. Therefore it is very hard to design a CMOS THz oscillator source with fundamental frequency higher than cut-off frequency. With low cut-off frequency, the resolution of CMOS THz imaging is limited by diffraction limit.

One promising way to generate frequency component higher than device cut-off frequency is Harmonic Extraction (HE). Due to non-linearity of devices, the MOSFETs in oscillator generate the fundamental tone of oscillation as well as harmonics. The idea behind HE is, enhance desired harmonic while cancel all other harmonics, which is shown in Fig. 7.1



(a)

(b)

Fig. 7.1 Output power of frequency components (a) before HE and (b) after HE

The target of HE is to enhance desired harmonic while cancel all other harmonics. There are two key metrics of HE. The first is low-order tone leakage. In reality, the cancellation of low-order harmonics is not perfect. Therefore there still exists residue low-order harmonic in output spectrum. On the other hand, before HE, the low-order harmonics usually have much higher power than desired harmonic. As a result, with insufficient low-order harmonic suppression, the low-order harmonic leakage may have even higher power in output spectrum than desired harmonic after HE. If this case happens in HE for imaging, then effectively the resolution is still determined by low-order harmonic leakage (with low frequency) rather than desired harmonic (with high frequency). Therefore the low-order harmonic leakage should be reduced as much as possible in HE. The second metric is processing loss. HE needs certain operation to eliminate low-order harmonic and enhance high-order harmonic. This operation may induce power loss. With power loss, SNR is degraded. Therefore we need to avoid processing loss as much as possible.

A naive method for HE is by filtering. We can build a band-pass filter to enhance desired harmonic at THz while eliminate all other harmonics. This is very hard to achieve for CMOS THz system. First it is very hard to manipulate poles and zeros in THz band. Due to high frequency nature, the value of inductance and capacitance in filter is extremely small, and can be easily offset by process, voltage and temperature variation. In addition, the insertion loss of filter at THz can be quite large. Therefore filtering is not a good way for THz HE.

A common way for CMOS THz HE is time-domain HE [7.1]. A diagram for time-domain HE is shown in Fig. 7.2. Time-domain HE has n oscillations with the same frequency components but different delays. For k-th delayed branch, the delay is k/n. Then those n delayed versions of oscillation are combined at output. Based on equations (6.6)-(6.8), in output ideally only frequency at nf, 2nf, 3nf… frequency components remain, while all other harmonics are cancelled. Since frequency component at nf usually has much higher power than frequency components at 2nf, 3nf,…, the dominant frequency component at output is nf. An exemplary phasor plot for time-domain HE of 4th order harmonic enhancement is shown in Fig. 7.3. Geometrically it can be seen how the harmonics cancel or enhance after combination.

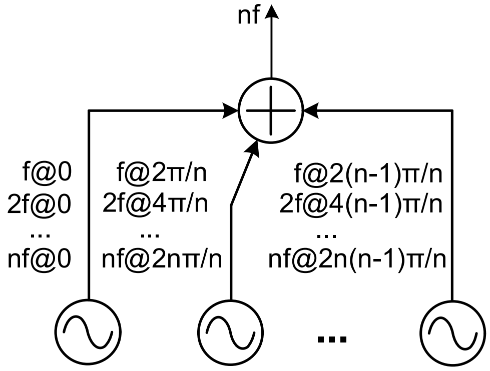


Fig. 7.2 Diagram of time-domain HE

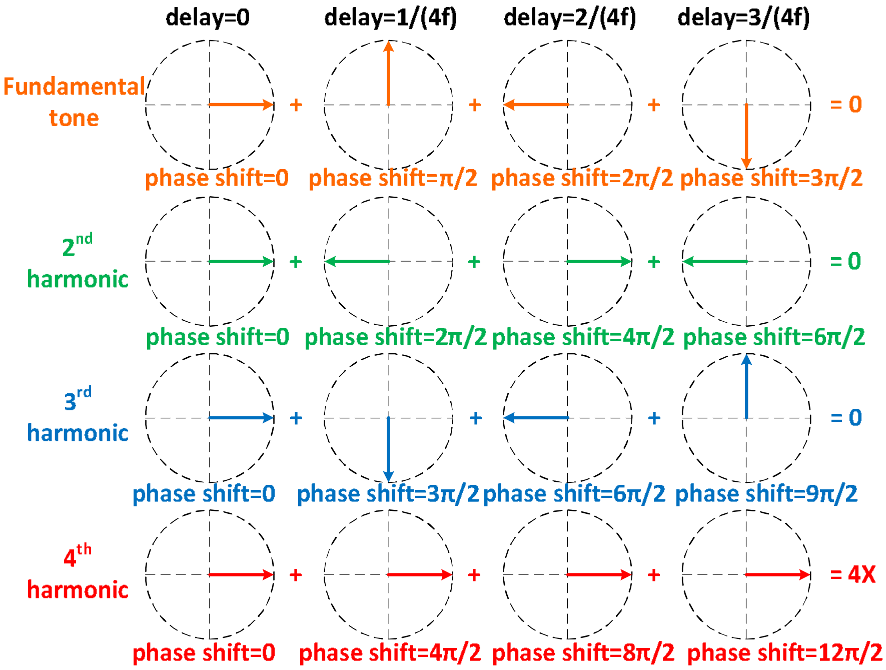


Fig. 7.3 Phasor plot of time-domain HE for 4th order harmonic

However there is issue with time-domain HE. The most important issue is low-order harmonic leakage. From Fig. 6.13, it can be seen that fundamental harmonic has 30-40 dB higher power than high-order harmonics. In order to enhance high-order harmonic while eliminate the effect of fundamental tone, low-order harmonic suppression need to be at least 30-40 dB. On the other hand, device delay mismatch exists in CMOS chips. In order to achieve such a high low-order harmonic suppression, the phase mismatch must be lower than 15 deg is required [7.2], as shown in Fig. 7.4. This phase mismatch translates to delay mismatch of less than 40 fs, which is extremely hard to achieve for CMOS devices. As an example of low-order harmonic leakage, the output spectral profile of state-of-art CMOS THz source with time-domain HE [7.3] is shown in Fig. 7.5. In [7.3], the desired harmonic for enhancement is 4th harmonic at 873 GHz. But we can see due to device mismatch, the power of 1st to 3rd harmonic is even higher than desired tone. If we use it directly for imaging, the resolution is still limited by low-order harmonics.

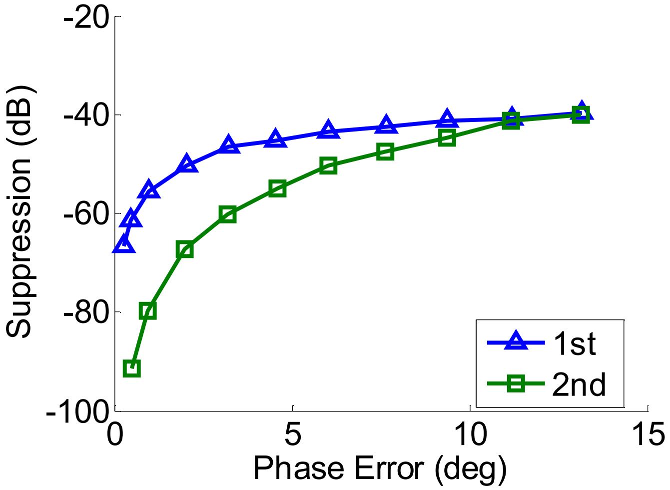


Fig. 7.4 Phase mismatch vs. low-order harmonic suppression

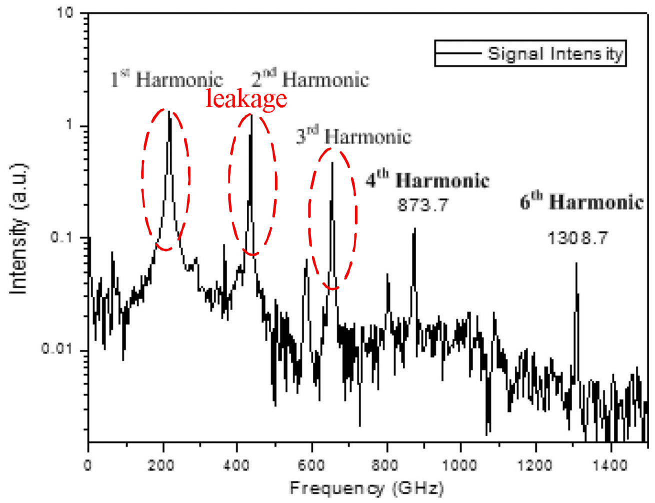


Fig. 7.5 PSD of output in [7.3]

Another issue of time-domain HE is processing loss. Coupler is used to couple THz oscillations with various delays together. The coupler can be lossy at THz band, and therefore SNR is degraded.

7.2 PRINCIPLES OF SHE

As is explained before, time-domain HE has stringent delay mismatch requirement, and it is very hard for state-of-art CMOS technology to achieve required mismatch. For radiation at 1 THz period, the period is only 1 ps. To achieve acceptable low-order harmonic suppression, a delay mismatch of less than 4% of period which translates to <40 fs is required. In short, it's very hard for current technology to achieve the required time resolution for THz time-domain HE.

On the other hand, our technology can already achieve precise displacement with precise linear stage. The resolution that a low-cost linear stage can achieve is less than 1 µm (Fig. 7.6). For radiation at 1 THz, the wavelength in air is 300 µm, so a resolution of 1 µm translates to error of less than 0.3%. This is equal to phase mismatch of less than 1 deg. If we can do phase shift in spatial domain instead of time domain, an excellent low-order harmonic suppression can be achieved.

Low-Cost Linear Translation Stage: VT-80 Series

- Motorized Positioning Range: 25 to 300 millimeters
- Maximum Load: 5 kg
- Maximum Velocity: 20 mm/second
- Maximum Positioning Resolution: 0.2 micrometers
- Position Reproducibility: 0.4 micrometers unidirectional +/ 10 micrometers bidirectional

Fig. 7.6 Specifications of a low-cost linear stage

According to our discussion in Chapter 6, interferometric system performs time-space transform, which transforms time-domain harmonics into spatial-domain harmonics. Therefore we should be able to implement the dual system in spatial domain to achieve spatial-domain phase shift.

In order to do phase shift in spatial domain, what we can do is to very distance z between THz source/detector and reflector. For CMOS interferometric system with distance z_0 between source/detector and reflector, the phase shift of reflective wave traveling back to detector is $4\pi z_0/\lambda$, where λ is the wavelength of radiation frequency component. If we vary distance from z_0 to $z_0+\Delta z$, the phase shift of reflective wave becomes $4\pi(z_0+\Delta z)/\lambda$. Therefore, by varying distance by Δz, an extra phase shift of is $4\pi\Delta z/\lambda$ induced. By sampling data at specific positions and combing data together, the effect of desired harmonic is enhanced while effect of other harmonics is cancelled.

Concretely, suppose we want to enhance n-th harmonic and cancel all other harmonics. According to discussion in previous section, we need data with fundamental tone phase shift of 0, $2\pi/n$, $4\pi/n$,…, $2(n-1)\pi/n$, and add them together. If we want to implement phase shift in spatial domain, we start by sampling interferometric data at position z_0 (Fig. 7.7).

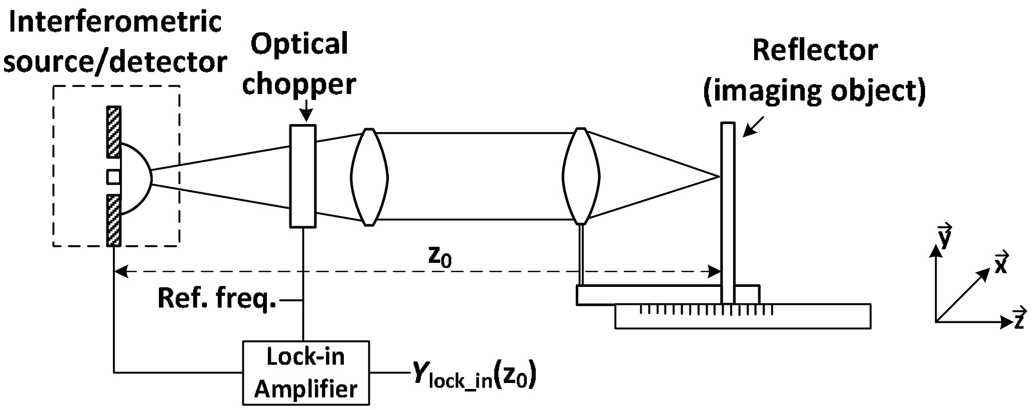


Fig. 7.7 First step of SHE: sample interferometric data at position z_0

At position z_0, the incident wave $S_I(t)$ can be expressed as:

$$S_I(t) = \sum_{k=1}^{m} A_{ik} \cos\left(2k\pi ft\right) \tag{7.1}$$

And the reflective wave $S_R(t)$ can be expressed as:

$$S_R(t,z) = \sum_{k=1}^{m} A_{rk} \cos\left(2k\pi ft + \frac{4\pi z_0}{\lambda / k}\right) \tag{7.2}$$

The output of lock-in amplifier $Y_{lock\text{-}in}$ can be expressed as:

$$Y_{lock-in}\left(z_0\right) = \sum_{k=1}^{m}\left[\frac{A_{ik}^2}{2} + \frac{A_{rk}^2}{2} + A_{ik}A_{rk}\cos\left(\frac{4\pi z_0}{\lambda / k} + \varphi_k\right)\right] \tag{7.3}$$

where φ_k is a constant phase shift (does not depend on distance z) between incident wave and reflective wave. Then we need to have phase shift of $2\pi/n$ for fundamental tone in $Y_{lock\text{-}in}$. This translates to position displacement $\Delta z=\lambda/2n$ (Fig. 7.8).

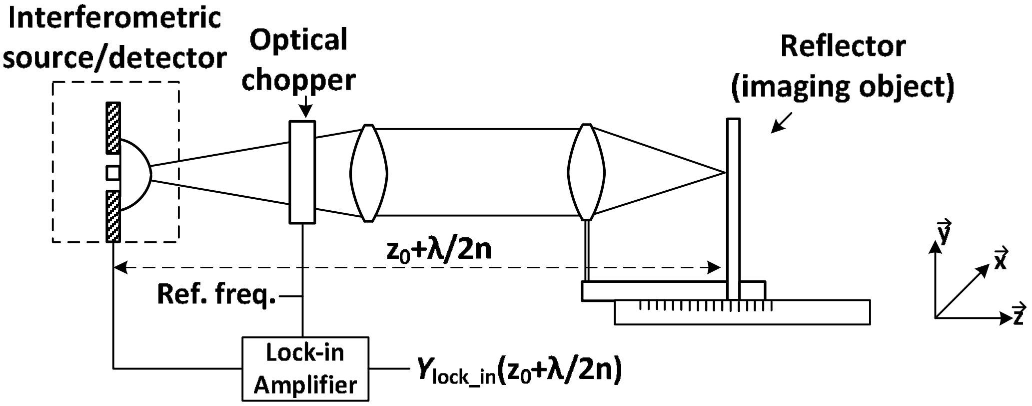


Fig. 7.8 Phase shift of 2π/n by spatial displacement of λ/2n

At position $z_0+\lambda/2n$, the incident wave $S_I(t)$ does not change. The reflective wave $S_R(t)$ can be expressed as:

$$S_R(t,z)=\sum_{k=1}^{m} A_{rk}\cos\left(2k\pi ft+\frac{4\pi z_0}{\lambda/k}+\frac{2k\pi}{n}\right) \tag{7.4}$$

The output of lock-in amplifier $Y_{lock\text{-}in}$ can be expressed as:

$$Y_{lock-in}\left(z_0+\frac{\lambda}{2n}\right)=\sum_{k=1}^{m}\left[\frac{A_{ik}^2}{2}+\frac{A_{rk}^2}{2}+A_{ik}A_{rk}\cos\left(\frac{4\pi z_0}{\lambda/k}+\frac{2k\pi}{n}+\varphi_k\right)\right] \tag{7.5}$$

We keep moving reflector by step of $\lambda/2n$, until $z=z_0+(n-1)\lambda/2n$. At $z=z_0+(n-1)\lambda/2n$, the output of lock-in amplifier can be expressed as:

$$Y_{lock-in}\left(z_0+\frac{(n-1)\lambda}{2n}\right)=\sum_{k=1}^{m}\left[\frac{A_{ik}^2}{2}+\frac{A_{rk}^2}{2}+A_{ik}A_{rk}\cos\left(\frac{4\pi z_0}{\lambda/k}+\frac{2k(n-1)\pi}{n}+\varphi_k\right)\right] \tag{7.6}$$

Then we add those n $Y_{lock\text{-}in}$ together. The sum can be expressed as:

$$
\begin{aligned}
Y_{sum}(z_0) &= \sum_{j=1}^{n} Y_{lock-in}\left(z_0 + \frac{(j-1)\lambda}{2n}\right) \\
&= \sum_{j=1}^{n}\sum_{k=1}^{m}\left[\frac{A_{ik}^2}{2}+\frac{A_{rk}^2}{2}+A_{ik}A_{rk}\cos\left(\frac{4\pi z_0}{\lambda/k}+\frac{2k(j-1)\pi}{n}+\varphi_k\right)\right] \\
&\approx \frac{nA_{in}^2}{2}+\frac{nA_{rn}^2}{2}+nA_{in}A_{rn}\cos\left(\frac{4\pi z_0}{\lambda/n}+\varphi_n\right)
\end{aligned}
\tag{7.7}
$$

The last step of (7.7) is based on (6.8) and the fact that n-th harmonic is much stronger than 2n-th, 3n-th… harmonics.

From equation (7.7), it is clear that after spatial domain phase shift and combination, the n-th harmonic is enhanced while all other low-order harmonics are cancelled. Phasor plot of an example of Spatial Harmonic Extraction (SHE) for 4th-order harmonic extraction is shown in Fig. 7.9.

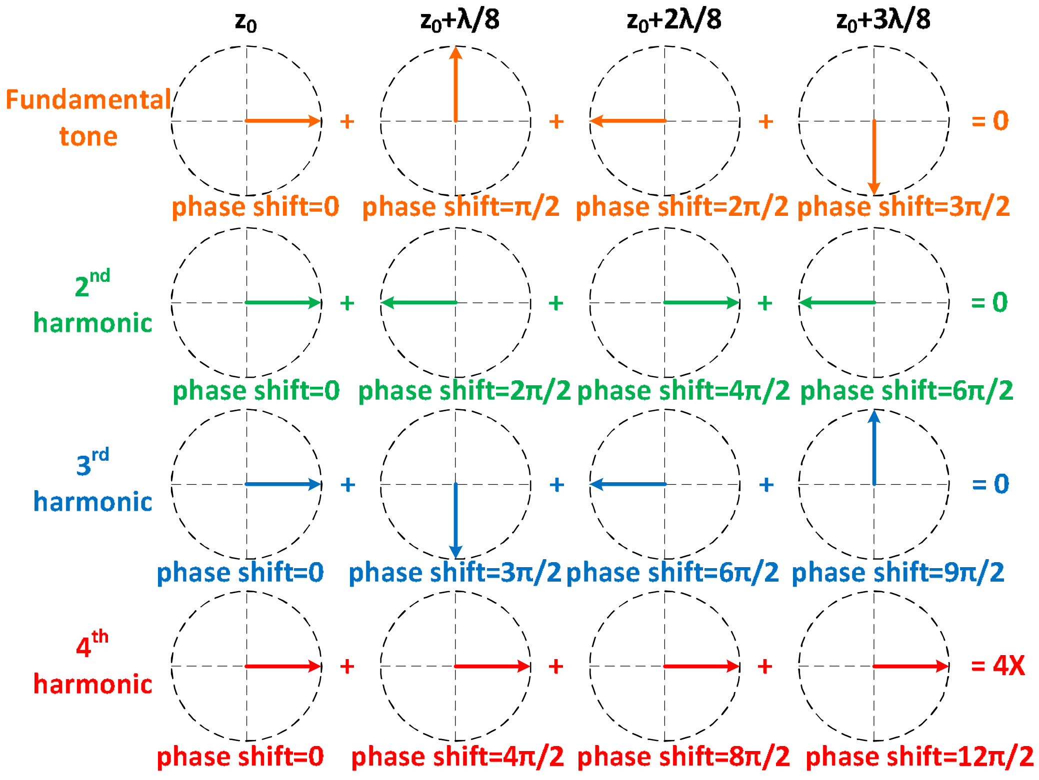


Fig. 7.9 Phasor plot of 4th-order harmonic extraction by SHE

SHE is dual of time-domain HE in spatial domain. For imaging purpose, time-domain HE and SHE both work in ideal situation. Time-domain HE enhances and eliminates harmonics in radiation. After time-domain HE, ideally the radiation only contains desired tone. It samples only once for one image, and ideally the resolution of image is determined by desired tone (Fig. 7.10).

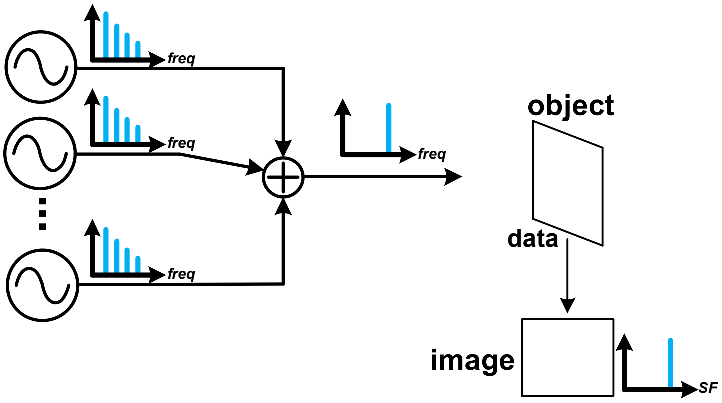


Fig. 7.10 Imaging using time-domain HE

On the other hand, SHE sample multiple times for one image. Each sample contains the effects of multiple tones. After summation, only the effect of desired tone is enhanced while the effect of all other tones is eliminated (Fig. 7.11).

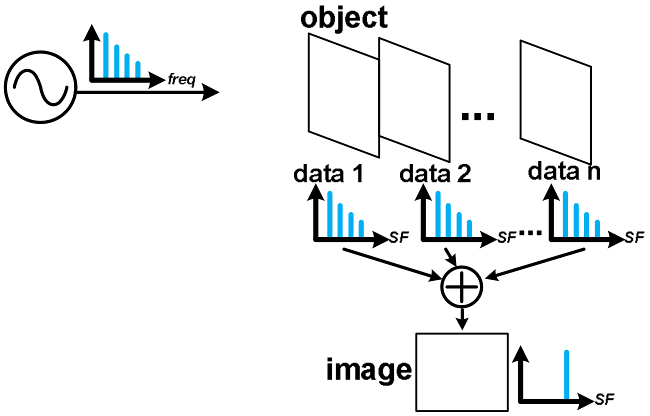


Fig. 7.11 Imaging using SHE

The advantage of SHE is also clear. First, high-accuracy displacement is possible in spatial domain, and therefore phase error of phase shift can be kept low. Therefore good low-order harmonic suppression can be achieved for THz harmonic extraction. In addition, the processing is not lossless for THE. This is due to the fact that data at each position is sampled and digitized. Therefore combination of data is just sum in digital domain. By contrast, combination step of time-domain HE involves power combination at THz band, which can be lossy and degrade SNR of imaging.

7.3 QUADRATURE SAMPLING

The output of lock-in amplifier output of interferometric imaging can be written as:

$$S_R(t,z) = \sum_{k=1}^{m} A_{rk} \cos\left(2k\pi ft + \frac{4\pi z_0}{\lambda / k}\right) \tag{7.8}$$

Our interferometric imaging for surface profiling is based on one important assumption: the amplitude of reflective wave does not change, i.e., the reflection index of surface does not change. Therefore the output change of lock-in amplifier is assumed to be solely caused by the phase change of reflective wave due to surface depth change.

However, if the reflection index of surface also changes (i.e., the amplitude of reflective changes), then we cannot distinguish the output change of lock-in amplifier due to surface depth (which changes phase $4\pi z/\lambda$) from the change due to reflection index (i.e., A_{rk}). The change of both surface roughness height and reflection index is quite common. One example is the mosaic plate in Fig. 7.12.

Fig. 7.12 A common plate with mosaic which has both surface and reflection index change

In order to distinguish amplitude change of reflective wave from phase change of reflective wave, quadrature sampling is proposed. We first assume that in the reflective wave, only one harmonic is dominating. The output of lock-in amplifier can be expressed by:

$$Y_{lock-in}\left(z_0\right)=\frac{A_i^2}{2}+\frac{A_r^2}{2}+A_iA_r\cos\left(\frac{4\pi z_0}{\lambda}+\varphi\right) \tag{7.9}$$

If we move z by λ/8, 2λ/8 and 3λ/8, then the outputs of lock-in amplifier can be expressed as:

$$\begin{aligned}Y_{lock-in}\left(z_0+\frac{\lambda}{8}\right)&=\frac{A_i^2}{2}+\frac{A_r^2}{2}+A_iA_r\cos\left[\frac{4\pi(z_0+\lambda/8)}{\lambda}+\varphi\right]\\&=\frac{A_i^2}{2}+\frac{A_r^2}{2}+A_iA_r\cos\left(\frac{4\pi z_0}{\lambda}+\frac{\pi}{2}+\varphi\right)\end{aligned} \tag{7.10}$$

$$\begin{aligned}Y_{lock-in}\left(z_0+\frac{2\lambda}{8}\right)&=\frac{A_i^2}{2}+\frac{A_r^2}{2}+A_iA_r\cos\left[\frac{4\pi(z_0+2\lambda/8)}{\lambda}+\varphi\right]\\&=\frac{A_i^2}{2}+\frac{A_r^2}{2}+A_iA_r\cos\left(\frac{4\pi z_0}{\lambda}+\pi+\varphi\right)\end{aligned} \tag{7.11}$$

$$Y_{lock-in}\left(z_0+\frac{3\lambda}{8}\right)=\frac{A_i^2}{2}+\frac{A_r^2}{2}+A_iA_r\cos\left[\frac{4\pi(z_0+3\lambda/8)}{\lambda}+\varphi\right]$$
$$=\frac{A_i^2}{2}+\frac{A_r^2}{2}+A_iA_r\cos\left(\frac{4\pi z_0}{\lambda}+\frac{3\pi}{2}+\varphi\right) \tag{7.12}$$

We can manipulate (7.9) - (7.12) to calculate A$_i$A$_r$:

$$Y_{lock-in}(z_0)-Y_{lock-in}\left(z_0+\frac{2\lambda}{8}\right)=A_iA_r\cos\left(\frac{4\pi z_0}{\lambda}+\varphi\right)-A_iA_r\cos\left(\frac{4\pi z_0}{\lambda}+\pi+\varphi\right)$$
$$=2A_iA_r\cos\left(\frac{4\pi z_0}{\lambda}+\varphi\right) \tag{7.13}$$

$$Y_{lock-in}\left(z_0+\frac{\lambda}{8}\right)-Y_{lock-in}\left(z_0+\frac{3\lambda}{8}\right)=A_iA_r\cos\left(\frac{4\pi z_0}{\lambda}+\frac{\pi}{2}+\varphi\right)-A_iA_r\cos\left(\frac{4\pi z_0}{\lambda}+\frac{3\pi}{2}+\varphi\right)$$
$$=2A_iA_r\cos\left(\frac{4\pi z_0}{\lambda}+\frac{\pi}{2}+\varphi\right)=-2A_iA_r\sin\left(\frac{4\pi z_0}{\lambda}+\frac{\pi}{2}+\varphi\right) \tag{7.14}$$

$$\sqrt{\left[Y_{lock-in}(z_0)-Y_{lock-in}\left(z_0+\frac{2\lambda}{8}\right)\right]^2+\left[Y_{lock-in}\left(z_0+\frac{\lambda}{8}\right)-Y_{lock-in}\left(z_0+\frac{3\lambda}{8}\right)\right]^2}$$
$$=\sqrt{\left[2A_iA_r\cos\left(\frac{4\pi z_0}{\lambda}+\varphi\right)\right]^2+\left[-2A_iA_r\sin\left(\frac{4\pi z_0}{\lambda}+\frac{\pi}{2}+\varphi\right)\right]^2}$$
$$=2\sqrt{2}A_iA_r \tag{7.15}$$

With quadrature sampling, the amplitude change of reflective wave can be tracked. This can be very useful to obtain both phase and amplitude information of reflection. With both phase and amplitude information of reflective wave, do can combine power and phase for imaging. This technique can also be used together with SHE for high-harmonic imaging for enhanced resolution. The cost of quadrature sampling is more data required for imaging. This is similar to Heisenberg's uncertainty principle: the product of Δx (position error, which translates to resolution in this case) and Δt (measurement time) must be larger than a certain value.

7.4 EXPERIMENT RESULTS

The experiment of SHE is similar to setup in Fig. 7.7, and was carried out by Dr. Al Hadi. To verify that SHE works, we checked the output DFT before and after SHE. Before SHE, we swept z from z_0 to z_0+6 mm, and sampled output of lock-in amplifier (Fig. 7.13). In DFT result, we can see fundamental tone leakage of the triple-push oscillator at 117 GHz, while the strongest tone is at 350 GHz (Fig. 7.14).

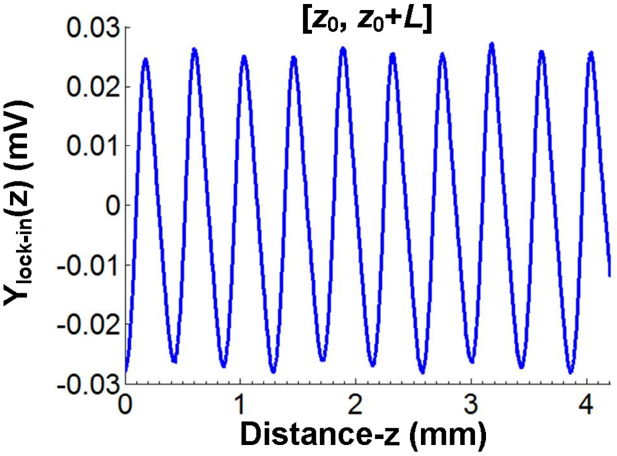


Fig. 7.13 Lock-in amplifier output of z sweep

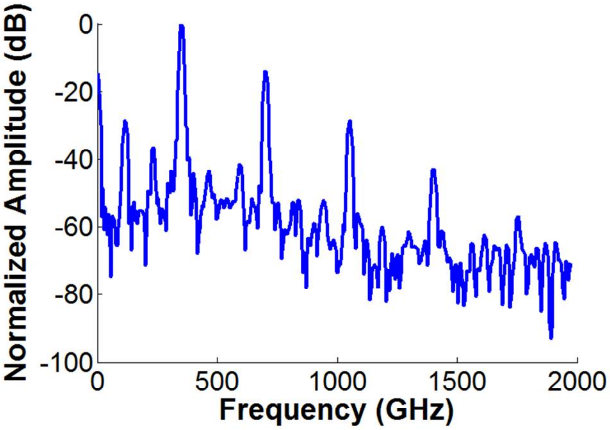


Fig. 7.14 DFT result of lock-in amplifier output

We tried to enhance 1.4-THz tone while cancel other tones. The 1.4-THz tone is the 12th harmonic of 117 GHz. So we need to combine 12 data with spatial displacement step of λ/24. We then swept z, sampled output of lock-in amplifier, and clipped data with 12 different data windows. Each data window has z-axis shift of λ /24. Then we add 12 data clips together (Fig. 7.15).

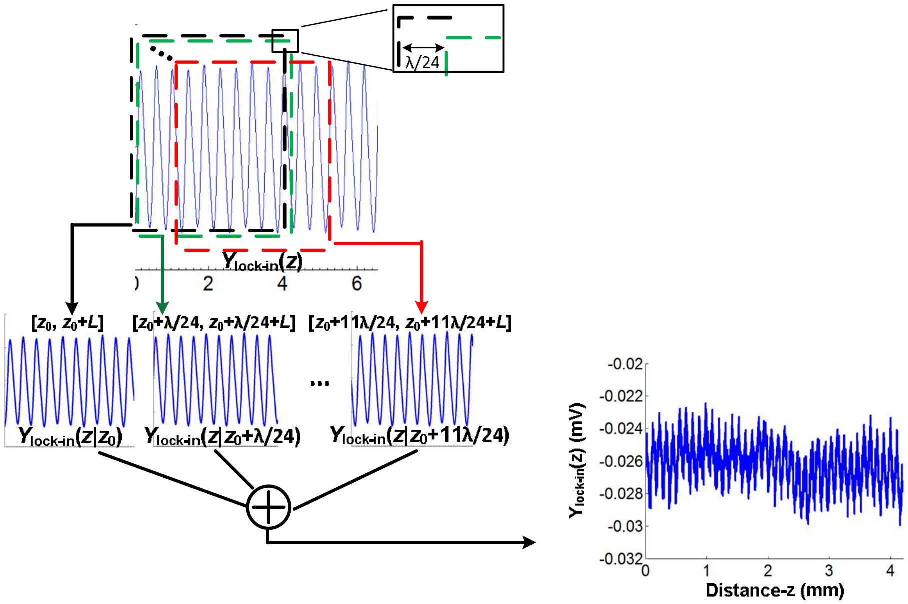


Fig. 7.15 Combination of lock-in amplifier with various phase delay

Then we do DFT on the sum of 12 clips of data. In DFT, we can see that the 12th harmonic enhanced while other harmonics are suppressed (Fig. 7.16). This verifies SHE method.

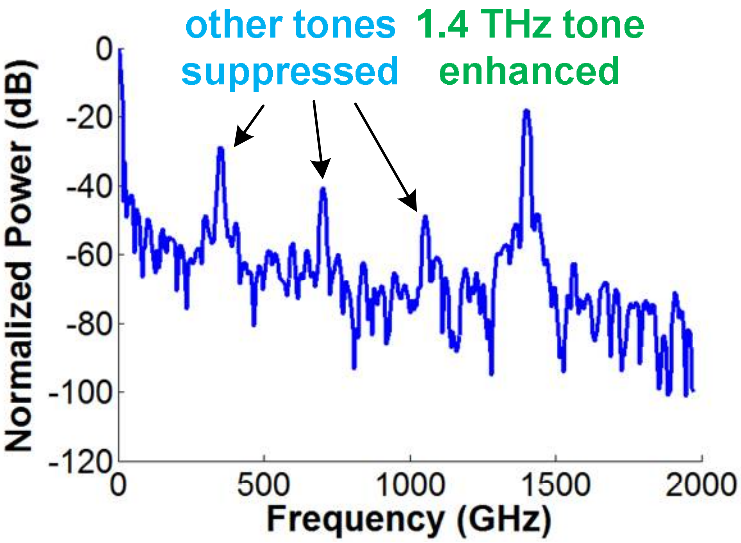


Fig. 7.16 DFT results of waveform after SHE

Another concern in interferometric imaging is the effect of non-linearity. We try to enhance 12th harmonic, which has very small power. The non-linearity terms of MOSFET can alias other harmonics with 12th harmonic. Since 12th harmonic has small power, we need to make sure that output tone at 12th harmonic is actually the effect of 12th harmonic instead of the effect of nonlinear aliasing.

Besides the second non-linearity that can generate a desired DC term corresponding to phase difference, third-order and fourth-order and other higher order harmonics can also generate a DC term with similar form (see below).

Desired	$(A_{r4}\cos(4\omega t+8kz)+A_{i4}\cos 4\omega t)^2 -> A_{r4}A_{i4}\cos 8kz$	linear
Nonlinear alias	$(A_{r2}\cos(2\omega t+4kz)+A_{i4}\cos 4\omega t)^3 -> \frac{3}{4}A_{r2}^2A_{i4}\cos 8kz$	quadratic
Nonlinear alias	$(A_{r2}\cos(2\omega t+4kz)+A_{i2}\cos 2\omega t)^4 -> \frac{3}{4}A_{r2}^2A_{i2}^2\cos 8kz$	quadratic

The third-order and fourth-order nonlinearities both alias 2nd harmonic to 4th harmonic. The output amplitude corresponding to desired second-order nonlinearity has linear relation to reflective wave amplitude. On the other hand, the output amplitude corresponding to third- and fourth-order nonlinearity aliasing has quadratic relation to reflective wave amplitude. Therefore we can identify whether output is dominated by square-law detection or high-order nonlinearity aliasing by changing the amplitude of reflective wave and measuring the output amplitude.

We then insert attenuator with various attenuation between source/detector and reflector, and measure output amplitude of harmonics in Fourier Transform. The results are shown in Fig. 7.17. It can be seen that the change of output amplitude has linear relation with attenuation. Therefore we can conclude that the dominant effect in output is desired second-order nonlinearity instead of other higher order nonlinear terms.

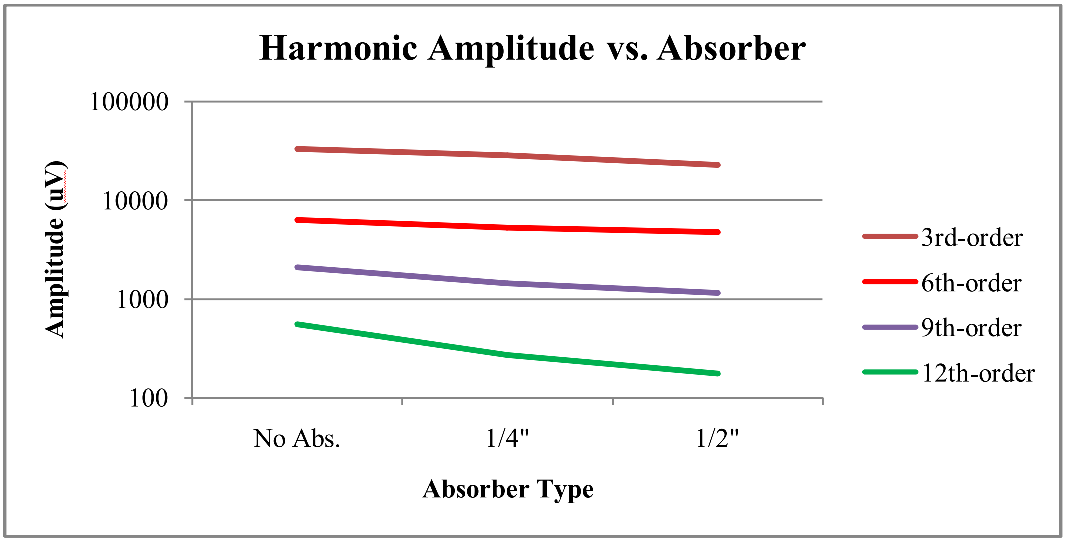
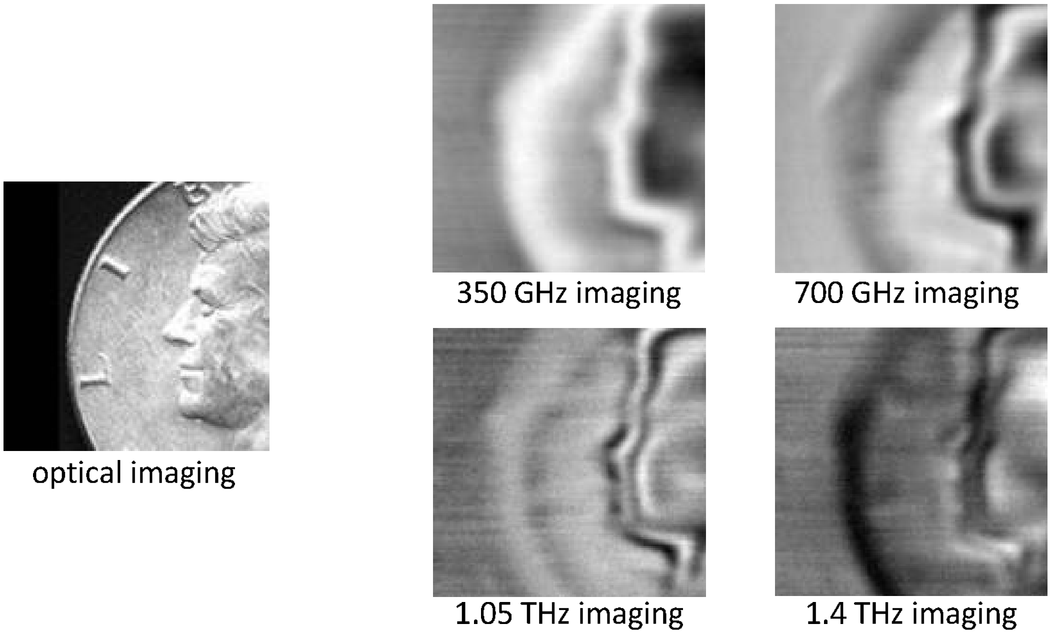


Fig. 7.17 Harmonic amplitudes vs. attenuation

We then use SHE for CMOS THz imaging. The experiment setup is similar to Fig. 7.7. We first sample imaging data at z-position z_0 by sweeping x-y coordinator of object. This data at z_0 is named as $I(z_0)$. $I(z_0)$ is the imaging result with 350 GHz tone. If we want to enhance n-th harmonic by SHE, we sample imaging data at $z_0+\lambda/2n$ (i.e.,$I(z_0+\lambda/2n)$), imaging data at

$z_0+2\lambda/2n$ (i.e.,$I(z_0+2\lambda/2n)$),…, imaging data at $z_0+(n-1)\lambda/2n$ (i.e.,$I(z_0+(n-1)\lambda/2n)$), and add $I(z_0)$, $I(z_0+\lambda/2n)$, $I(z_0+2\lambda/2n)$,…, $I(z_0+(n-1)\lambda/2n)$ together. We measured imaging result with 6th harmonic at 700 GHz, 9th harmonic at 1.05 THz and 12th harmonic at 1.4 THz, as shown in Fig. 7.18.


optical imaging

350 GHz imaging

700 GHz imaging

1.05 THz imaging

1.4 THz imaging

Fig. 7.18 Imaging result with various harmonics in radiation

It can be seen from Fig. 7.18 that with higher harmonic frequency, the imaging resolution becomes better. For 350 GHz imaging, only a brief outline of President Kennedy can be seen. With higher harmonic frequency, more and more details are available. When using 1.4 THz imaging, even the inscription of “L” and “I” on the coin can be seen. By using SHE, we successfully extract high order harmonic with good low-order harmonic leakage. Therefore the resolution is determined by desired harmonic instead of low-order harmonic leakage.

The main limit of SHE is SNR. Since high-order harmonic has very low power, the SNR after SHE is still low. In order to enhance SNR, we can either increase radiation power, or increase the integration time of lock-in amplifier. The SNR of 1.4 THz imaging in Fig. 7.16 is around 6-7 dB, which can be calculated from DFT result in Fig. 7.14. We have 8192-point DFT, which translates to 36 dB noise floor reduction. In one-shot sampling, the SNR is around 32~33 dB. When 12 samples are combined, SNR is enhanced by 11 dB. Therefore SNR is estimated to be around 6-7 dB. SNR can also be estimated by referring to the image quality-SNR comparison table in Fig. 7.19 [7.4, 7.5].

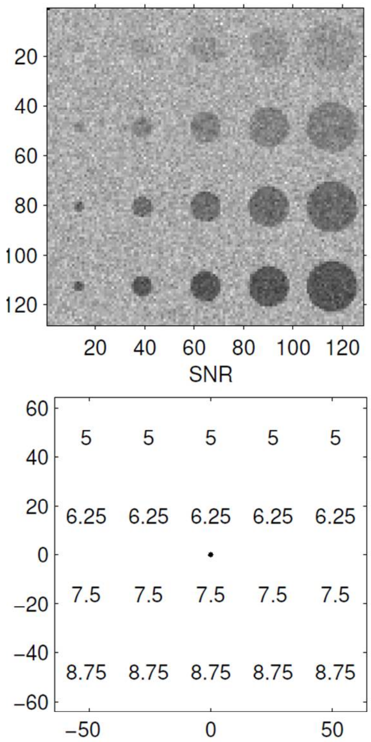


Fig. 7.19 Image quality-SNR comparison table

Chapter Conclusion

This chapter introduces SHE based on CMOS THz interferometric imaging. The concept of SHE and analysis were by Yilei Li, the experiment was carried out by Dr. Richard Al Hadi.

CHAPTER 8

FAST CMOS THZ SOURCE CHARACTERIZATION BY COMPRESSIVE SENSING (CS)

8.1 BACKGROUND

As has been discussed before, regular self-test is a key requirement for reliable THz systems. Due to the possible change of voltage and temperature, the THz source output frequency may drift in THz system. Without regular self-test, this drift cannot be tracked or compensated.

A THz imaging system with SHE technique for resolution enhancement need exact wavelength λ of radiation. As discussed in Chapter 7, the displacement step of SHE is determined by λ. However, λ may drift with time. When supply voltage of oscillator changes, the oscillation frequency changes and the components in radiation can change frequency accordingly. The oscillation frequency can also change with environment temperature. Even if we use phase-locked loop to lock oscillation frequency, λ can still change due to change of dielectric constant of environment (e.g., dielectric constant change due to humidity change). If change of λ is not tracked, then SHE cannot achieve enough low-order harmonic suppression, which in turn degrades resolution of THz imaging.

The requirements of self-test for low-cost compact THz system include:

(1) Self-test must be low-cost. Self-test is an integral part of THz system, and therefore any extra cost of self-test adds to the total cost of THz system.

(2) Self-test must be implemented with low hardware overhead. If instruments for self-test is too bulky, the whole THz system becomes bulky and cannot meet the requirement of portable applications.

(3) Self-test must be fast. If self-test takes too much time, the change of environment cannot be tracked in real time.

Conventional way for THz source characterization is FTIR system or harmonic mixer with spectrum analyzer. Those solutions are bulky and expensive, and cannot meet the requirement of low-cost compact THz system.

An alternative way for THz source characterization for interferometric system is by time-space transform of oscillator with reciprocal antenna, which was first proposed by Dr. Al Hadi in [8.1]. The setup is shown in Fig. 8.1.

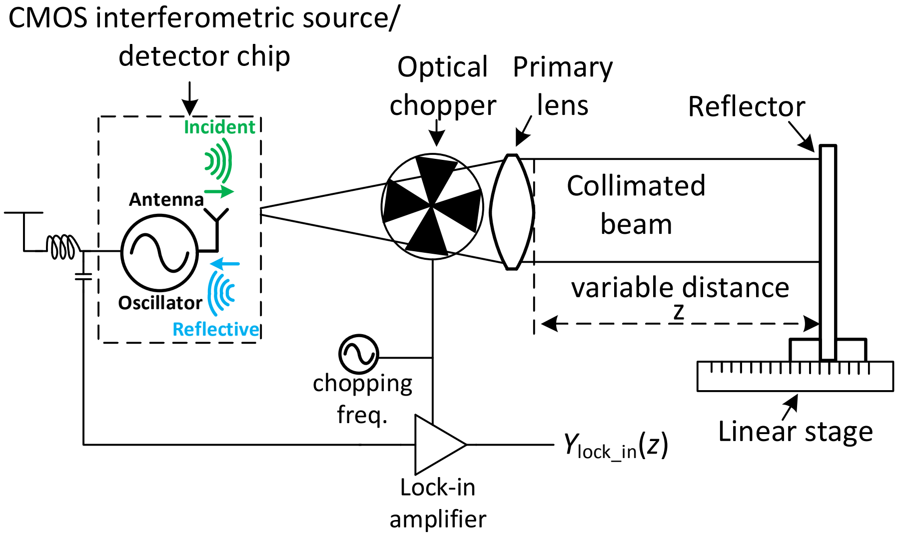


Fig. 8.1 Setup for interferometric THz system with reciprocal antenna

With interferometric spectral profiling (ISP), we just to vary z, measure $Y_{\text{lock\_in}}(z)$ at every z location and calculate $Y_{\text{lock\_in}}(f)$, which is the Fourier Transform of $Y_{\text{lock\_in}}(z)$. Due to time-space transform, every time-domain tone in radiation is transformed into spatial-domain tone in $Y_{\text{lock\_in}}(f)$. From Fig. 8.1, it can be seen that the ISP method does not add any hardware overhead. Therefore the ISP method can meet the requirement of compactness and low-cost.

However, there is issue with ISP, which is measurement time. For each step of z, it may take several hundred milliseconds as integration time of lock-in amplifier. Therefore it takes several milliseconds to take one $Y_{lock\_in}(z)$. On the other hand, the quality of ISP greatly depends on the number of measurement points. An example of 8192-point Discrete Fourier transform (DFT) result of ISP is shown in Fig. 8.2. A total of 7 tones are clearly seen in DFT.

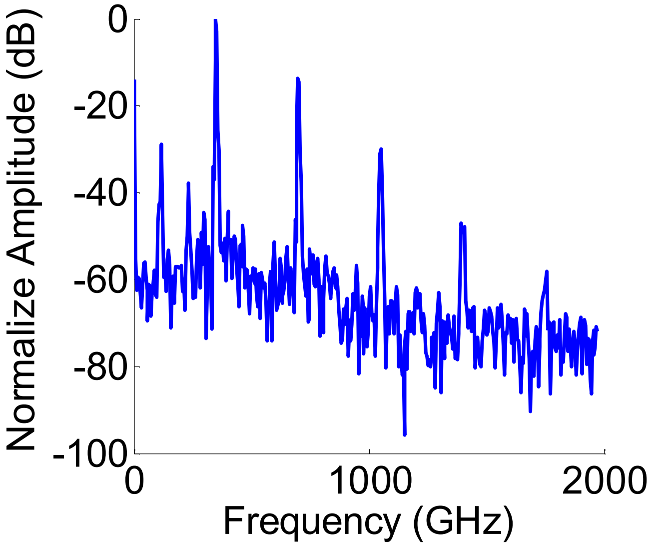


Fig. 8.2 8192-point DFT result of ISP

However, taking 8192 points may take 1 hour to finish. This self-test time is too long to do regularly. On the other hand, if we only have 1024 points for DFT, the ISP quality is degraded. Due to poor quality, we can see only 4 tones in DFT (Fig. 8.3).

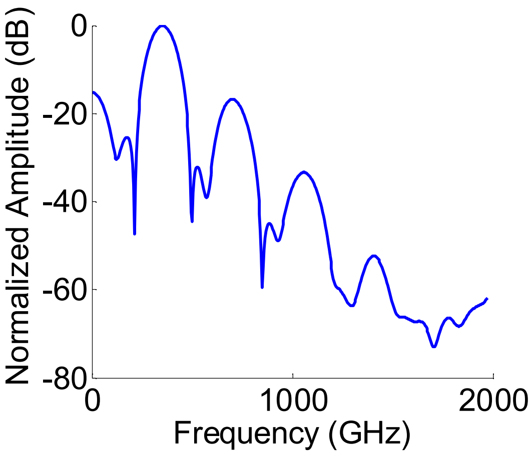


Fig. 8.3 1024-point DFT result of ISP

The DFT is based on uniform sampling method. According to time-frequency duality of Fourier transform [8.2], large time-domain step brings about poor resolution.

If we take a closer look at Fig. 8.2, it can be seen that only a few frequency bins have significant power, while all other bins have nearly zero power. Therefore the PSD is sparse. In other words, the information can be compressed by setting a threshold power, and set power of all frequency bins that have less power than threshold to zero, while maintaining the original power of frequency bins that have larger power than threshold. However, this kind of compression needs sampling first and then compression [8.3]. In our case, the bottleneck is measurement time, and we need a compressing-while-sampling scheme. In this scheme, we can directly reduce the number of samples required by PSD reconstruction without significantly degrading the quality of reconstruction. This scheme is compressive sensing [8.4-8.7].

8.2 INTRODUCTION TO COMPRESSIVE SENSING

Compressed sensing has been proposed by Candes, Romberg, Donoho and Tao in [8.4-8.7]. Instead of uniform time-step sampling in common measurements, compressive sensing uses random sampling (Fig. 8.4).

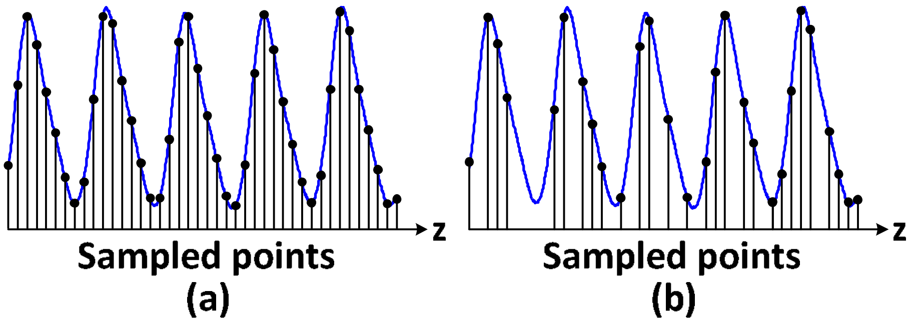


Fig. 8.4 (a) Uniform sampling and (b) random sampling

For a signal x with length of N, it can be expressed by certain basis Ψ as:

$$x = \Psi v \tag{8.1}$$

where v is the projection of x in basis Ψ. Signal x is said to be K-sparse with basis Ψ if v has only K non-zero elements. In real applications, signal x is approximately K-sparse with basis Ψ if v only has K large elements while all other elements are near zero. On example of sparse signal is x=sin(2πft), which is dense in time domain, but is sparse in frequency domain as it only has two non-zero elements (at –f and +f). If x is K-sparse with basis of Ψ, then we can use random sampling matrix Φ to sample x with only M measurements to measure v, where M<<N.

Ideally, v can be reconstructed by L1-norm optimization algorithm, as long as:

$$M > K \log_2 N \tag{8.2}$$

where K is a constant. Unlike DFT transform where frequency domain is calculated by linear transform, compressive sensing reconstructs v in Ψ domain by solving following optimization problem:

$$\min \|v\|_{l1} \text{ subject to } \|\Phi\Psi v - y\|_{l2} < \varepsilon \tag{8.3}$$

where y is the measured signal by random matrix Φ, and ε is error tolerance.

The diagram of CS is shown in Fig. 8.5 [8.3].

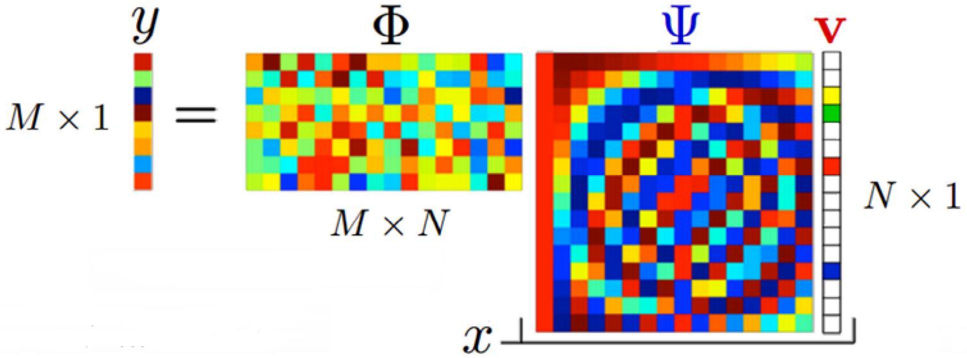


Fig. 8.5 Diagram of CS by random sampling

In real case, due to the presence of noise, the achievable compression ratio is among 8-10. This is due to the fact that random noise present in real world is not sparse with any basis. With compressive sensing, noise is folded and SNR is degraded [8.8]. In order to keep reasonable SNR, compression ratio should be kept around 8-10.

8.3 SOURCE OUTPUT PSD CHARACTERIZATION BY CS

According to the discussion above, we can apply CS to THz isp. This is due to the fact that output of isp is sparse in Fourier domain representation Y_{lock_in}(f). Therefore Ψ in our case in inverse DFT matrix, and we can randomly subsample Y_{lock_in}(z) and reconstruct Y_{lock_in}(f) with L1-norm optimization algorithm:

$$\min \left\|Y(f)\right\|_{l1} \text{ subject to } \left\|\Phi\Psi Y(f) - Y_{lock-in}(z)\right\|_{l2} < \varepsilon \tag{8.4}$$

The diagram of CS for ISP is shown in Fig. 8.6.

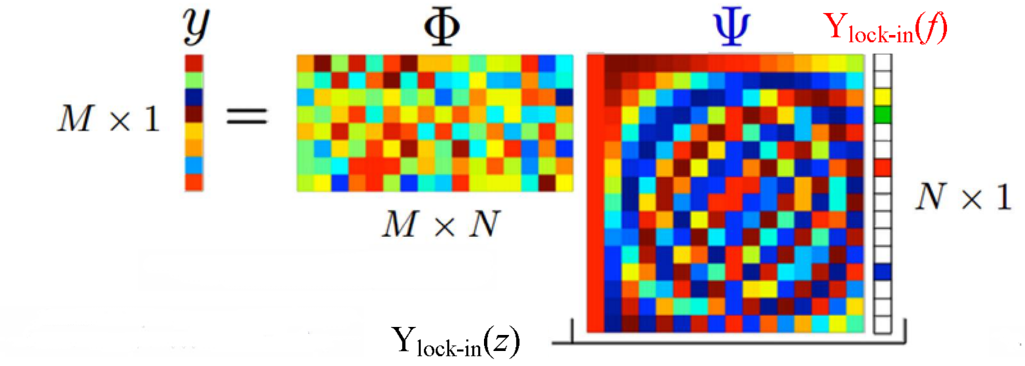


Fig. 8.6 Diagram of CS for ISP

With CS, we can randomly sample Y_{lock_in}(z) and reconstruct Y_{lock_in}(f) by L1-norm algorithm. The random sampling can be controlled by PC. If Y_{lock_in}(z) at certain step is sampled in random matrix, the linear stage will stop at that position and wait for lock-in amplifier to integrate. If Y_{lock_in}(z) is not sampled at certain step in random sampling matrix, the PC will control linear stage to simply skip that step, and therefore measurement does not wait at that position. Also, reconstruction of Y_{lock_in}(f) can be achieved by reconstruction algorithm implemented in PC, such as C program or MATLAB. Note that in reconstruction, random sampling matrix is known to algorithm. This means that we need to record random sampling sequence and then input it to reconstruction program. The setup of CS for ISP shown in Fig. 8.7.

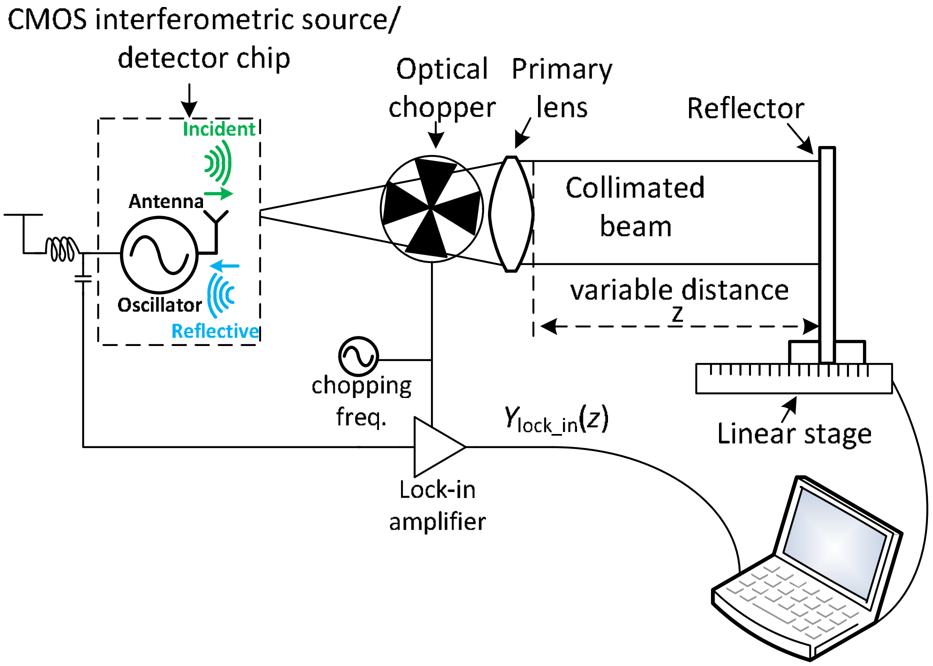


Fig. 8.7 Setup of CS for ISP

In experiment, we used $Y_{lock\_in}(z)$ data set in Chapter 7.3 and use MATLAB program to do random sampling and reconstruction.

We use compression ratio of 8 in our experiment, i.e., 1024-point data in CS. Then we use L1-norm reconstruction algorithm in [8.9]. The result is shown in Fig. 8.8(a). The 8192-point DFT result is also shown in Fig. 8.8(b) for comparison. We can see that the overall quality of reconstruction is good, but one weak tone at 1.75 THz is missing in reconstructed PSD of CS. The reason for missing tone is incoherence noise [8.10], which will be explained in detail in next section.

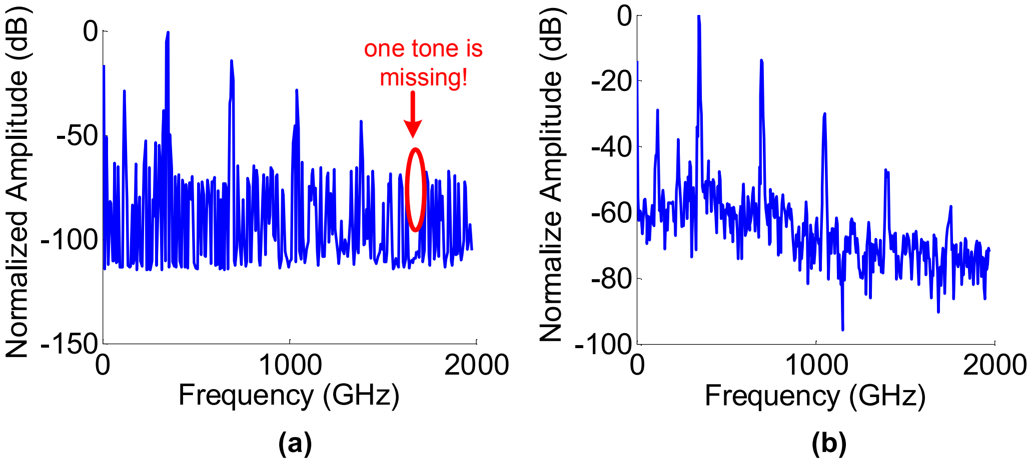


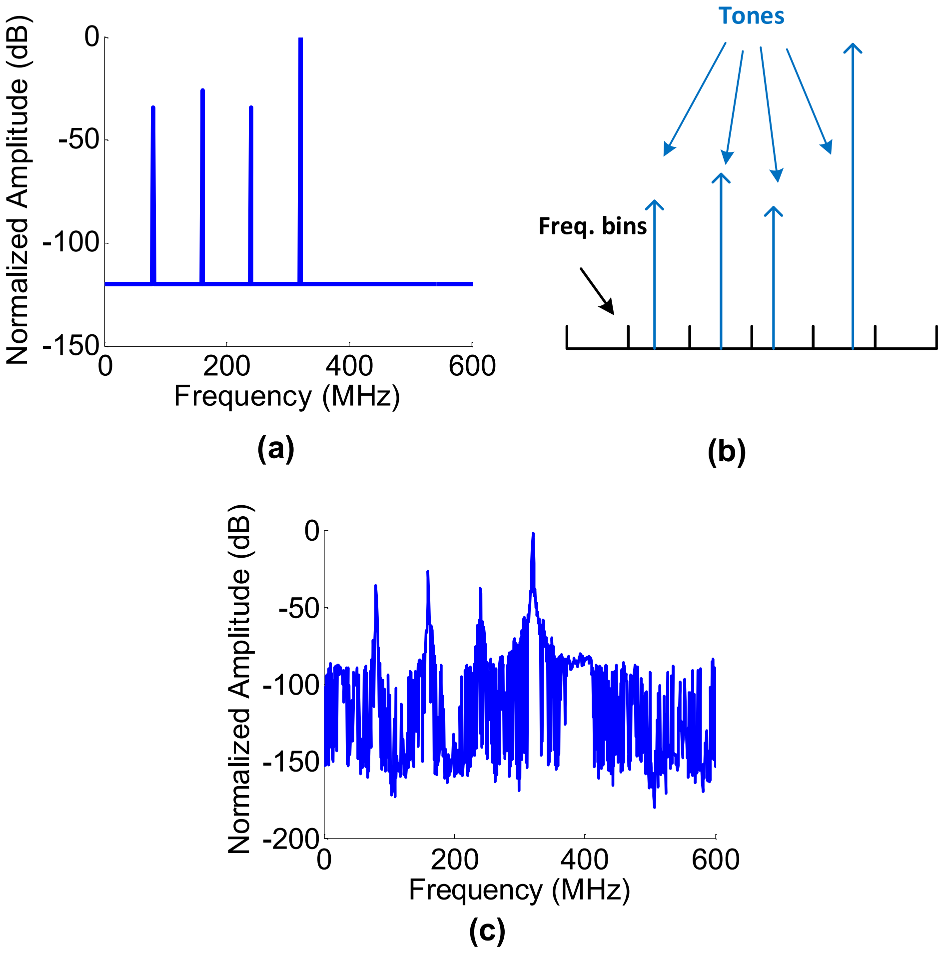
Fig. 8.8 (a) Reconstruction of CS with 8X compression (1024 point) and (b) result of 8192-point DFT

8.4 ZOOM-IN RECONSTRUCTION ALGORITHM

When tones of $Y_{\text{lock-in}}(f)$ do not fall exactly on frequency bins in CS, incoherent noise is generated. This can be understood as follows.

Reconstruction of CS is essentially optimization. It tries to find optimal linear combination of given basis which has minimal error with respect to sampled signal. The ideal case is when sampled signal is exactly linear combination of basis. For example, suppose basis is [cosωt, sinωt, cos2ωt, sin2ωt], and sampled signal is 2cos2ωt at some t samples, then solution is simply [0,0,2,0]. Optimization algorithm can find perfect solution in this case. However, in actual case, sampled signal cannot be exactly represented by linear combination of basis. For example, basis is [cosωt, sinωt, cos2ωt, sin2ωt], but sampled signal is 2cos(1.5ωt). In this case, energy of 2cos(1.5ωt) leaks to many frequency bins and induces incoherent noise. This is not similar to DFT: DFT is linear transform and if one frequency component is off-grid, it only leaks to neighboring bins. However, reconstruction of CS is not linear, and the off-grid frequency components can leak to many bins which increases noise floor. In addition, if one weak tone is off-grid, the tone can be missing at reconstruction result, as in Fig. 8.8(a), as the energy of the weak tone leaks to other bins.

As an example of incoherence noise induced by off-grid frequency components, a noise-free signal (Fig. 8.9(a)) is randomly sampled, and reconstructed by off-grid frequency bins (Fig. 8.9(b)). The reconstruction result is shown in Fig. (8.9(c)). It can be seen that significant incoherence noise is present in reconstruction, and signal quality is greatly degraded.



(a)

(b)

(c)

Fig. 8.9 (a) Noise-free signal, (b) sampling frequency grids and (c) reconstruction results

In order to reduce incoherence noise, a zoom-in CS reconstruction scheme is proposed. First, we reconstruct with original Ψ with uniform frequency bin size. According to reconstruction result, we identify candidate tones, including detected tones and their harmonics (Fig. 8.10(a)).

Then, we update Ψ, and put dense frequency bins near candidate tones. After that, we rerun reconstruction algorithm (Fig. 8.10 (b)).

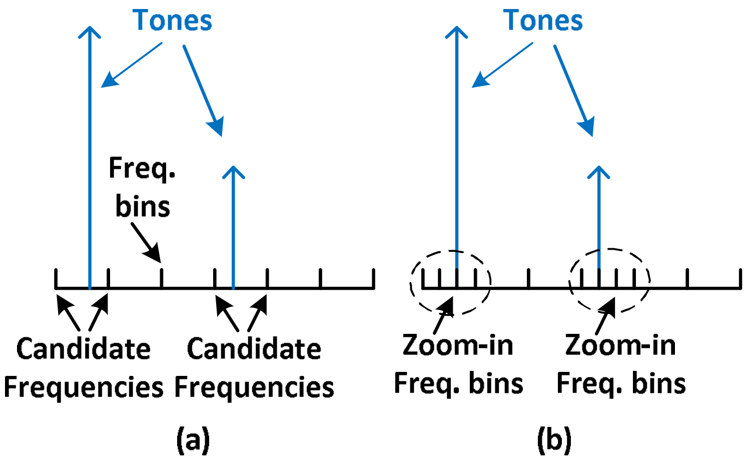


Fig. 8.10 Two-step zoom-in CS reconstruction

We can see before zoom-in, the distance from candidate tones to frequency bin can be large. After zoom-in, tones will have smaller distance from frequency bins. This will decrease incoherent noise. Note that we only need one data set for reconstruction, so measurement time will not increase.

The experiment result with noise-free signal of two-step zoom-in reconstruction is shown in Fig. 8.11. We used 8X compression ratio. The reconstruction result of first step is shown in Fig. 8.11(a). The incoherence noise is significant, but candidate frequencies can still be identified. Then we zoom-in by 3X frequency bin density near candidate frequencies and update Ψ. The reconstruction result with the updated Ψ is shown in Fig. 8.11(b). It can be seen that the incoherence noise is reduced to very low level.

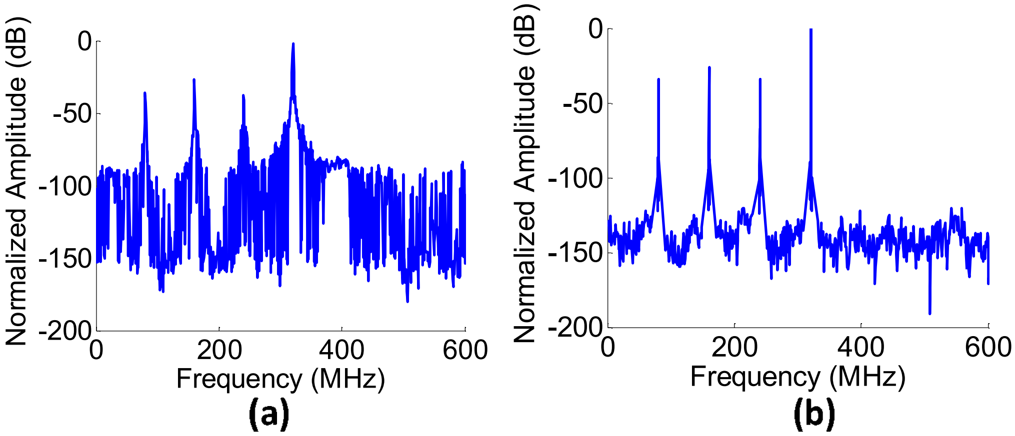


Fig. 8.11 (a) Reconstruction with original Ψ and (b) reconstruction with updated Ψ

We then apply two-step zoom-in algorithm on compressed sense data. The comparison of DFT and CS results are shown in Fig. 8.12. In Fig. 8.12(a), 8192 data points are used for DFT, and we can see seven tones at 117 GHz, 234 GHz, 351 GHz, 700 GHz, 1.05 THz, 1.4 THz and 1.75 THz. However gathering 8192 points may take as long as one hour, and cannot meet the requirement of fast characterization and calibration. If we only sample 1024 data points for DFT with uniform sampling, then the resolution becomes very poor and only 4 tones can be seen. On the other hand, if we use random sampling (compressive sensing) to sample 1024 points, the reconstruction quality can be much better than 1024-point DFT. With naive L1-reconstruction algorithm, we can see six tones at 117 GHz, 234 GHz, 351 GHz, 700 GHz, 1.05 THz and 1.4 THz, but 1.75 THz tone is missing (Fig. 8.11(c)). This is due to incoherence noise as discussed in earlier this chapter. With two-step zoom-in algorithm with 3X zoom-in, the incoherence noise is reduced and we can see all seven tones (Fig. 8.11(d)). At this point the SNR is limited by noise folding instead of incoherence noise, so zoom in with more frequency bin density does not help much in reconstruction quality enhancement. With the zoom-in reconstruction algorithm, we achieved 8X compression ratio without significantly degrading quality. Therefore the time for characterization is reduced to 7 minutes instead of nearly one hour.

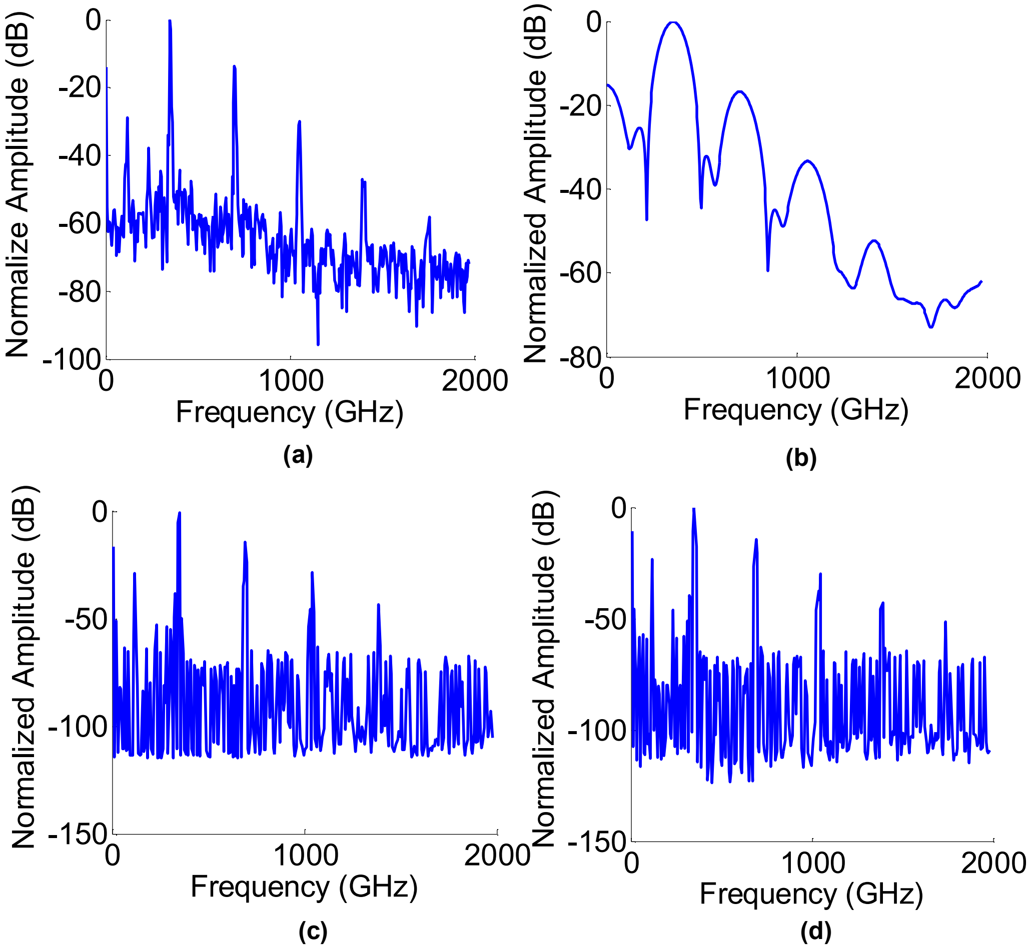


Fig. 8.12 (a) DFT with 8192 data points (b) DFT with 1024 data points (c) CS with 1024 data points without zoom-in (d) CS with 1024 data points with 3X zoom-in

Chapter Conclusion

This chapter introduces CS-based fast THz source characterization. The concept of interferometric spectral profiling was by Dr. Richard Al Hadi. The concept of CS-based fast THz source characterization and analysis were by Yilei Li, the experiment was carried out by Dr. Richard Al Hadi.

CHAPTER 9

CONCLUSION

In this work, applications of phase-based algorithms are proposed. Phase, as a key degree of freedom in EM wave, can be manipulated for various applications.

For data link interconnect, phase can be modulated to carry information. A hardware-efficient algorithm is proposed in this work for carrier phase synchronization. By sending specific information known in priori, the phase difference status of carrier in TX and RX can be identified by simple 1-bit ADC. This greatly reduces hardware overhead for MRFI compared with conventional carrier phase synchronization scheme, where high-speed high-resolution ADCs are required. Furthermore, the dispersion effect of channel for phase synchronization is investigated and optimal demodulation phase for random signal with weak dispersion is derived.

For THz applications, combing signals with specific phase relation can enhance or cancel the effect of certain harmonic component. Conventionally this harmonic extraction is done in time domain, which has mismatch issue in THz band and can induce large low-order harmonic leakage. In this work, we propose to perform harmonic extraction in spatial domain for interferometric imaging, and a 4X resolution enhancement with 1.4 THz imaging is achieved.

In addition, random sampling of phase can help us reduce the data amount needed to reconstruct a sparse signal. This compressive sensing theorem is used in fast characterization of THz source with inteferometric spectral profiling. The spectral profiling of THz source is sparse in frequency domain. Further, a zoom-in two-step reconstruction algorithm is proposed, which takes full advantage of information known in priori about spectral profile of THz source. By applying random sampling and zoom-in reconstruction algorithm, we are able to reduce the amount of data required by spectral profiling by eight times with reasonable characterization quality. This reduces the time needed for measurement of inteferometric spectral profiling to acceptable range.

CHAPTER 10

FUTURE WORKS

10.1 CHARACTERIZATION OF INTERCONNECT WIRE BY COMPACT MRFI TRANSCEIVERS

Currently interconnect wire is characterized by S-parameter analyzer. S-parameter analyzer is bulky, expensive and needs calibration. On the other hand, the S-parameter analyzer is overkill for wireline transceivers. The information that wireline transceivers need includes:

(1) Amplitude-frequency response.

Notches in frequency response can induce unwanted signal distortion and inter-symbol interference (ISI). With MRFI, we are able to place signal frequency bands to flat region of frequency response. Therefore we need to identify the notch positions of frequency bands. The absolute loss is not necessary, what we need is the frequencies where received signal is extremely weak in amplitude. Therefore we can use MRFI to detect notch position. MRFI TX transmits signal with amplitude known in priori, and RX detects the amplitude of received signal. By sweeping carrier frequency within range of interest (this can be done by various structures such as dual mixer, fractional PLL, or simply fine tune reference frequency of PLL), the relative amplitude response is measured by RX.

(2) Phase-frequency response

Dispersive channel displays non-linear phase delay with regards to frequency. The phase response can be detected by MRFI transceivers as follows. For each frequency, carrier phase synchronization routine is performed to find optimal phase code. Then we sweep carrier frequency in both TX and RX. Based on the optimal phase code in various frequencies, we can decide whether the phase delay through channel is linear with regards to frequency or not.

Dispersion can be put into consideration for carrier frequency selection, as in a strongly dispersive frequency band the IQ interference is strong and theoretically cannot be eliminated for random signal.

10.2 FREQUENCY-MODULATED CONTINUOUS WAVE (FMCW) RADAR WITH SHE

Interferometric imaging can obtain outline of a smooth surface whose height does not change much. However, there is aliasing issue if object surface has large change of surface height. This is due to the fact that the output of lock-in amplifier is periodical with change of z. The output of lock-in amplifier is:

$$Y_{lock-in}(z) = \frac{A_i^2}{2} + \frac{A_r^2}{2} + A_i A_r \cos\left(\frac{4\pi z}{\lambda}\right)$$

$Y_{lock\text{-}in}(z)$ is periodical with period of λ/2 (period ambiguity, Fig. 10.1). In addition, $Y_{lock\text{-}in}(z) = Y_{lock\text{-}in}(k\lambda/2-z)$ (non-monotonic). With quadrature sampling, the alias due to non-monotonic $Y_{lock\text{-}in}(z)$ can be eliminated [10.1], but period ambiguity still exists.

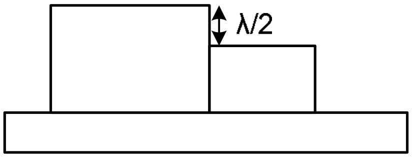


Fig. 10.1 Surface with the period ambiguity

In order to avoid alias, we can use the solution similar to FMCW radar, which sweeps wavelength of radiation to solve ambiguity (Fig. 10.2) [10.2]. For any position z, we have the following equations:

$$S_{\lambda_1}(z) = \frac{A_{i1}^2}{2} + \frac{A_{r1}^2}{2} + A_i A_r \cos\left(\frac{4\pi z}{\lambda_1} + \varphi_1 + 2k_1\pi\right) \quad (10.1)$$

$$S_{\lambda_1}(z-\frac{\lambda_1}{4})=\frac{A_{i1}^2}{2}+\frac{A_{r1}^2}{2}+A_iA_r\cos\left(\frac{4\pi z}{\lambda_1}-\frac{\pi}{2}+\varphi_1+2k_1\pi\right)$$
$$=\frac{A_{i1}^2}{2}+\frac{A_{r1}^2}{2}+A_iA_r\sin\left(\frac{4\pi z}{\lambda_1}+\varphi_1+2k_1\pi\right) \qquad (10.2)$$

$$S_{\lambda_2}(z)=\frac{A_{i2}^2}{2}+\frac{A_{r2}^2}{2}+A_{i2}A_{r2}\cos\left(\frac{4\pi z}{\lambda_2}+\varphi_2+2k_2\pi\right) \qquad (10.3)$$

$$S_{\lambda_2}(z-\frac{\lambda_2}{4})=\frac{A_{i2}^2}{2}+\frac{A_{r2}^2}{2}+A_iA_r\cos\left(\frac{4\pi z}{\lambda_2}-\frac{\pi}{2}+\varphi_2+2k_2\pi\right)$$
$$=\frac{A_{i2}^2}{2}+\frac{A_{r2}^2}{2}+A_iA_r\sin\left(\frac{4\pi z}{\lambda_2}+\varphi_2+2k_2\pi\right) \qquad (10.4)$$

where A_{i1}, A_{r1}, A_{i2}, A_{r2}, φ_1 and φ_2 are constants and can be determined by calibration with reflector at position known a priori. We can solve k_1, k_2 and z from (10.1) – (10.4) with the constraint that k_1 and k_2 are integers. With presence of noise, exact solution may not be possible, but we can still estimate z with highest probability.

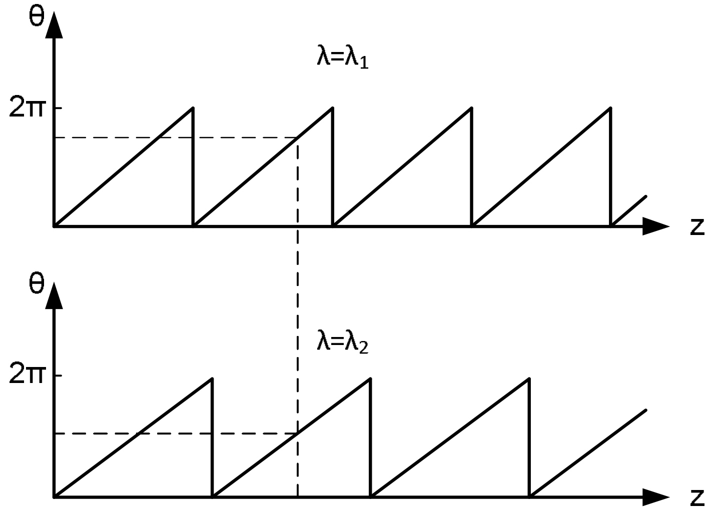


Fig. 10.2 FMCW to solve z

With z determined, the imager is actually radar. The surface plot is the depth distribution of surface.

BIBLIOGRAPHY

[1.1] C. E. Shannon and Warren Weaver. The mathematical theory of communication. University of Illinois press, 2015.

[1.2] B. Razavi. RF Microelectronics, 2nd Ed. Prentice Hall, 2012

[1.3] Xin He and Jan Van Sinderen, "A 45nm low-power SAW-less WCDMA transmit modulator using direct quadrature voltage modulation." In 2009 IEEE International Solid-State Circuits Conference.

[1.4] Thomas, Stewart J., and Matthew S. Reynolds. "A 96 Mbit/sec, 15.5 pJ/bit 16-QAM modulator for UHF backscatter communication." In 2012 IEEE International Conference on RFID.

[1.5] Eberle, Wolfgang, et al. "A digital 72 Mb/s 64-QAM OFDM transceiver for 5 GHz wireless LAN in 0.18um CMOS." In 2001 IEEE International Solid-State Circuits Conference.

[1.6] Okada, Kenichi, et al. "A 64-QAM 60GHz CMOS transceiver with 4-channel bonding." In 2014 IEEE International Solid-State Circuits Conference.

[1.7] Zhao, Dixian, and Patrick Reynaert. "A 3 Gb/s 64-qam E-band direct-conversion transmitter in 40-nm cmos." In 2014 IEEE Asian Solid-State Circuits Conference.

[2.1] Chien, Andrew A., and Vijay Karamcheti. "Moore's Law: The first ending and a new beginning." Computer 12 (2013): 48-53.

[2.2] Natarajan, S., et al. "A 14nm logic technology featuring 2 nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 µm 2 SRAM cell size." In 2014 IEEE International Electron Devices Meeting (IEDM).

[2.3] Max Born and Emil Wolf. Optical Physics. Cambridge Univ. Press, 1997.

[2.4] Kim, Kinam. "Silicon technologies and solutions for the data-driven world." In 2015 IEEE International Solid-State Circuits Conference.

[2.5] W. M. Holt, "Moore’s Law: A Path Forward." In 2016 IEEE International Solid-State Circuits Conference.

[2.6] S. Sutardja, "The Future of IC Design Innovation." In 2015 IEEE International Solid-State Circuits Conference.

[2.7] W. Sansen, "Analog CMOS from 5 Micrometer to 5 Nanometer." In 2015 IEEE International Solid-State Circuits Conference.

[2.8] J. L. Hennsssy and D. A. Patterson. Computer Architecture: A Quantitative Approach, 5th Ed. Elsevier, 2011.

[2.9] Chen, E., Ramy Yousry, and Chih-Kong Ken Yang. "Power optimized ADC-based serial link receiver." IEEE Journal of Solid-State Circuits, vol. 47, no.4, 2012.

[2.10] Wei-Han Cho, et al. "A 5.4-mW 4-Gb/s 5-band QPSK transceiver for frequency-division multiplexing memory interface." In 2015 IEEE Custom Integrated Circuits Conference (CICC)

[2.11] W.-H. Cho et al., "A 38mW 40Gb/s 4-lane tri-band PAM-4/16-QAM transceiver in 28nm CMOS for high-speed Memory interface." In 2016 IEEE International Solid-State Circuits Conference.

[2.12] Xiao, Chunhua, et al. "Stream arbitration: Towards efficient bandwidth utilization for emerging on-chip interconnects." ACM Transactions on Architecture and Code Optimization (TACO), vol.9, no.4,2013.

[3.1] A. Goldsmith. Wireless Communications.Cambridge Univ. Press, 2005.

[3.2] Zhao, Qing, and Gordon L. Stüber. "Convergence behavior of iterative phase synchronization and detection." IEEE Journal on Selected Areas in Communications, vol.23, no.5, 2005, pp: 932-943.

[3.3] Y. Li et al., "Carrier synchronisation for multiband RF interconnect (MRFI) to facilitate chip-to-chip wireline communication." Electronics Letters, vol. 52, no. 7, 0216, pp.535-537

[3.4] Thomas H. Lee. The Design of CMOS Radio-Frequency Integrated Circuits, 2nd Ed. Cambridge Univ. Press, 2003

[3.5] David M. Pozar. Microwave Engineering, 4th Ed. Wiley, 2011.

[3.6] William A. Gardener. Introduction to Random Processes, With Applications to Signals & Systems, 2nd Ed. McGraw-Hill, 1990.

[3.7] A. D. Whalen. Detection of Signals in Noise. Academic Press, 1971.

[5.1] Fawwaz T. Ulaby. Fundamentals of Applied Electromagnetics, 6th Ed. Prentice Hall, 2010.

[5.2] Kaushik Sengupta, and Ali Hajimiri. "A 0.28 THz Power-Generation and Beam-Steering Array in CMOS Based on Distributed Active Radiators." IEEE Journal on Solid-State Circuits, vol.47, no.12, 2012, pp: 3013-3031.

[5.3] Churchley, David, et al. "Terahertz pulsed imaging study to assess remineralization of artificial caries lesions." Journal of biomedical optics, vol. 16 ,no.2 ,2011: 026001-026001.

[5.4] Cosentino, Antonino. "Terahertz and Cultural Heritage Science: Examination of Art and Archaeology." MDPI Technologies, vol.4, no.1, 2016.

[5.5] Siebert, Karsten J., et al. "All-optoelectronic continuous wave THz imaging for biomedical applications." Physics in medicine and biology vol.47, no.21, 2002.

[5.6] Williams, Benjamin S. "Terahertz quantum-cascade lasers." Nature photonics vol.1, no.9, 2007, pp. 517-525.

[5.7] Khalid, Ata, et al. "In0. 53Ga0. 47As planar Gunn diodes operating at a fundamental frequency of 164 GHz." IEEE Electron Device Lett vol.34, no.1, 2013, pp. 39-41.

[5.8] ITRS Reports, available online: http://www.itrs2.net/itrs-reports.html

[5.9] Han, Ruonan, et al. "Active terahertz imaging using Schottky diodes in CMOS: Array and 860-GHz pixel." IEEE Journal on Solid-State Circuits, vol.48, no.10 , 2013, pp.2296-2308.

[6.1] P. Hariharan. Basics of Interferometry, 2nd Edition. Elsevier, 2007.

[6.2] R. Al Hadi et al. "Retroactive terahertz detection for imaging and remote sensing applications in a standard CMOS technology." 2016 SPIE Photonics West.

[7.1] Y. Zhao et al., "An Integrated 0.56THz Frequency Synthesizer with 21GHz Locking Range and -74dBc/Hz Phase Noise at 1MHz Offset in 65nm CMOS", In 2016 IEEE International Solid-State Circuits Conference.

[7.2] Daquan Huanget al., "Terahertz CMOS frequency generator using linear superposition technique", IEEE Journal of Solid-State Circuits, vol.43, no.12, 2008.

[7.3] Qun Guet al., "CMOS THz Generator with Frequency Selective Negative Resistance Tank", IEEE Transaction on Terahertz Science and Technology, vol. 2, no. 2, 2012, pp. 193-202.

[7.4] A. Macovski. Medical imaging systems. Prentice-Hall, New Jersey, 1983.

[7.5] J. L. Prince and J. M. Links. Medical imaging signals and systems. Prentice-Hall, 2005.

[8.1] R. Al Hadi, Y. Zhao, Y. Li et al., "A Spectral Profiling Method of mm-Wave and Terahertz Radiation Sources", In 2016 International Microwave Symposium

[8.2] A. V. Oppenheim and R. W. Schafer. Digital signal processing. Englewood Cliffs, 1975.

[8.3] Richard Baraniuk, "Compressive sensing", IEEE Signal Processing Magazine, 24(4), pp. 118-121, July 2007

[8.4] Emmanuel Candès, Justin Romberg, and Terence Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information", IEEE Trans. on Information Theory, 52(2) pp. 489 - 509, February 2006

[8.5] Emmanuel Candès and Justin Romberg, "Quantitative robust uncertainty principles and optimally sparse decompositions", Foundations of Comput. Math., 6(2), pp. 227 - 254, April 2006

[8.6] David Donoho, "Compressed sensing", IEEE Trans. on Information Theory, 52(4), pp. 1289 - 1306, April 2006

[8.7] Emmanuel Candès, Justin Romberg, and Terence Tao, Stable signal recovery from incomplete and inaccurate measurements", Communications on Pure and Applied Mathematics, 59(8), pp. 1207-1223, August 2006

[8.8] E. Arias-Castro, and Y. C. Eldar, "Noise folding in compressed sensing", IEEE Signal Processing Letters, vol. 18, no. 8, 2011, pp. 478-481.

[8.9] Emmanuel Candes and Justin Romberg. "l1-magic: Recovery of sparse signals via convex programming." http://www. acm.caltech.edu/l1magic/, 2005.

[8.10] Michael Lustig, David Donoho, and John M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging." Magnetic resonance in medicine, vol.58, no.6, pp. 1182-1195, Oct. 2007.

[10.1] Bamji, Cyrus S., et al. "A 0.13 μm CMOS system-on-chip for a 512× 424 time-of-flight image sensor with multi-frequency photo-demodulation up to 130 MHz and 2 GS/s ADC." IEEE Journal of Solid-State Circuits, vol.50, no.1, 2015, pp. 303-319.

[10.2] Wu, Wanghua, Robert Bogdan Staszewski, and John R. Long. "A 56.4-to-63.4 GHz multi-rate all-digital fractional-N PLL for FMCW radar applications in 65 nm CMOS." IEEE Journal of Solid-State Circuits, vol. 49, no.5, 2014, pp. 1081-1096.