AFRL-RY-WP-TR-2009-1172

WAFER SCALE DISTRIBUTED RADIO

A.M. Niknejad, Elad Alon, Borivoje Nikolic, and Jan Rabaey

University of California, Berkeley

JULY 2009
Final Report

Approved for public release; distribution unlimited.

See additional restrictions described on inside pages

STINFO COPY

AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7320
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the USAF 88th Air Base Wing (88 ABW) Public Affairs Office (PAO) and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RY-WP-TR-2009-1172 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

*//Signature//

POMPEI ORLANDO, Project Engineer
Advanced Sensor Components Branch
Aerospace Components Division

//Signature//

BRADLEY J. PAUL, Chief
Advanced Sensor Components Branch
Aerospace Components Division

//Signature//

TODD A. KASTLE, Chief
Aerospace Components Division
Sensors Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

*Disseminated copies will show "//Signature//" stamped or typed above the signature blocks.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
July 2009	Final	10 July 2008 – 16 July 2009

4. TITLE AND SUBTITLE
WAFER SCALE DISTRIBUTED RADIO

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA8650-08-1-7855

5c. PROGRAM ELEMENT NUMBER
63739E

6. AUTHOR(S)
A.M. Niknejad, Elad Alon, Borivoje Nikolic, and Jan Rabaey

5d. PROJECT NUMBER
ARPR

5e. TASK NUMBER
YD

5f. WORK UNIT NUMBER
ARPRYD0K

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California, Berkeley
Berkeley Wireless Research Center
2108 Allston Way, Suite 200
Berkeley, CA 94704-1302

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Sensors Directorate
Wright-Patterson Air Force Base, OH 45433-7320
Air Force Materiel Command
United States Air Force

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
AFRL/RYDI

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RY-WP-TR-2009-1172

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
PAO Case Number: 88ABW-2009-3632; Clearance Date: 10 Aug 2009. This report contains color.

14. ABSTRACT
Modem silicon technology offers ultrafast transistors, with fT > 200 GHz in today's 45nm CMOS and fT > 300 GHz in SiGe. While extremely fast, these transistors suffer from several limitations which affect the performance of high dynamic range analog and RF circuits. Principally, the low supply voltage hampers the dynamic range, and, combined with the low intrinsic gain, high variability of nanoscale transistors, and increasing 1 = f noise due to high-K materials, it becomes clear that analog operations are increasingly difficult to realize in silicon. In RF applications, the modest noise performance in the microwave and mm-wave band (operating close to fT) has limited the application of silicon technology to short range wireless systems. In this project we explore the potential for a new kind of wafer scale distributed radio that can overcome the limitations of silicon by exploiting its high levels of integration and a more intimate relationship between the electronics and electromagnetic structures.

15. SUBJECT TERMS
radio frequency, phased array, high bandwidth, high frequency, amplifiers, components, trade study

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
a. REPORT	b. ABSTRACT	c. THIS PAGE			Pompei Orlando
Unclassified	Unclassified	Unclassified	SAR	100	19b. TELEPHONE NUMBER *(Include Area Code)* N/A

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

Contents

Chapter 1

Introduction to the Distributed Wafer Scale Radio

The goal of this research project was to determine the feasibility and likely performance for a wafer scale distributed radio. The vision for such a radio includes a monolithic wafer consisting of a repetitive pattern of building blocks, as shown in Fig. 1.1, where each block consists of a wideband transceiver and several antenna elements covering a wide frequency range. By integrating the entire functionality of the radio onto a single wafer, several advantages ensue.

Furthermore, the aim was to quantify the technical challenges and ultimate performance of a wafer scale distributed microwave/mm-wave radio implemented in deeply scaled silicon technology (CMOS or SiGe BiCMOS). The envisioned radio operates over several broad frequency bands spanning 10-90 GHz with multiple spatial beams, in essence exploiting frequency, time, and spatial degrees of freedom. Even though the underlying electronics can be implemented in nanoscale CMOS or SiGe – which is known to suffer from lower power handling, higher noise, and lower reliability than other semiconductor technologies – the distributed radios performance could far exceed todays systems in terms of size, robustness, cost, and power. This improvement in performance (despite the inferior devices) is obtained by utilizing many hundreds to thousands of radiation elements for signal processing, beam forming, redundancy, and a combination of near and far-field spatial power combining.

1.1 Motivation

If such a wafer scale distributed radio can be realized, several existing applications (such as radar and communication links) would immediately benefit, while many new applications would emerge. Given the raw signal processing power of an entire silicon wafer, we envision the wafer scale radio modulating/de-modulating tens to hundreds of different channels simultaneously. The additional spatial resolution brought by the use of higher mm-wave frequencies also opens the possibility of realizing a high resolution mm-wave imaging system for applications such as imag-

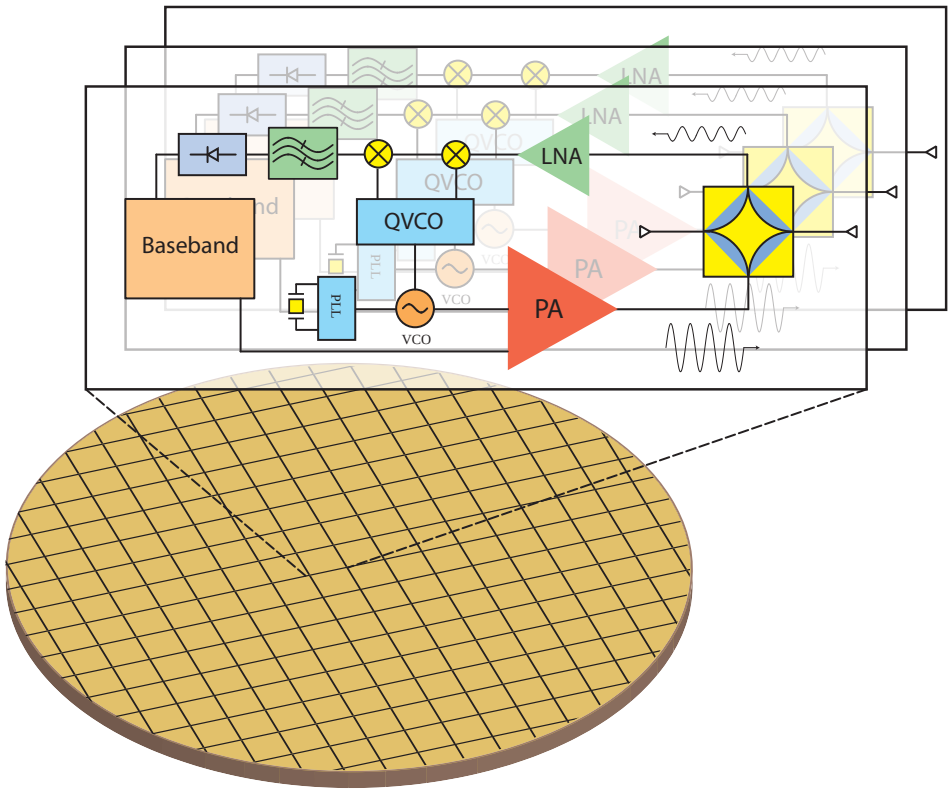


Figure 1.1: A wafer scale distributed radio.

ing in low visibility conditions (fog), hidden weapon detection, and medical applications such as tumor detection and stroke-type identification. Today these systems are not in widespread use due to several limitations, including cost, but more importantly the difficulty in manufacturing and assembling the various components reliably, and the resulting bulky size due to the modular approach of traditional MMIC packaging of sub-components.

A single monolithic wafer obviates the assembly process and presents a compact and lightweight solution. The wideband capability of such a device could also be used to realize an ultrawideband (UWB) ground-penetrating radar or portable "X-ray" (based on mm-wave radiation) for on-the-field detection of broken bones and other injuries. Higher frequencies offer higher spatial resolution but suffer higher attenuation when traveling through tissue, walls, or the ground. A wafer scale system employing true time delay elements can be used for spatial power combining of a wideband signal, offering the potential for very high effective radiation power in the desired spatial direction. The powerful digital signal processing offered by nanoscale silicon enables extremely intelligent arrays to be realized on-wafer, naturally exploiting the distributed computing offered by silicon without the requirement to route several high data rate digital signals off-chip

as with traditional systems. Given that the power capability of semiconductor technology drops dramatically with frequency, it is fortunate that more radiators can be realized in an array due to the decreasing wavelength. For instance, each reticle on the wafer scale radio may contain only 4-6 elements at 10 GHz (spaced at a fraction of a wavelength). However, using the same spacing, hundreds of elements could be integrated into a reticle at 90 GHz. In this way, the output power and sensitivity of the wafer as a whole would be relatively constant with frequency.

1.2 Technical Rationale

Modern silicon technology offers ultrafast transistors, with $f_T > 200$ GHz in todays 45nm CMOS and $f_T > 300$ GHz in SiGe. While extremely fast, these transistors suffer from several limitations which affect the performance of high dynamic range analog and RF circuits. Principally, the low supply voltage hampers the dynamic range, and, combined with the low intrinsic gain, high variability of nanoscale transistors, and increasing $1/f$ noise due to high-K materials, it becomes clear that analog operations are increasingly difficult to realize in silicon. In RF applications, the modest noise performance in the microwave and mm-wave band (operating close to f_T) has limited the application of silicon technology to short range wireless systems. Long range communication and in particular military communication systems require extremely wide dynamic range front-ends and radios, robust performance in the presence of strong interfering and jamming signals, high output power, and wideband operation. Most of these specifications cannot be met with traditional silicon technology above 10 GHz. In this project we explore the potential for a new kind of wafer scale distributed radio that can overcome the limitations of silicon by exploiting its high levels of integration and a more intimate relationship between the electronics and electromagnetic structures. We plan to build on the success of the TEAM project, where researchers demonstrated highly integrated CMOS and SiGe radio transceivers operating in the 60 GHz band, and circuit building blocks operating beyond 100 GHz, close to the limits of activity for silicon where realized. The limitations of conventional radios and the solutions offered by the wafer scale distributed radio are summarized below.

1.2.1 Wafer Scale Distributed Radio

The proposed system is comprised of approximately 100 25mm$\times$25mm reticles on a 12" wafer. Each reticle forms a distributed radio over several broad frequency ranges, e.g. from 10-30 GHz, 30-60 GHz, and 60-90 GHz. In turn, each reticle contains several sub-element transceivers operating over these same frequency ranges. Each sub-element transceiver is comprised of distributed active near-field and far-field antennas and couplers, a broadband front-end radio with a synthesized true time-delay element, and a bank of high performance baseband signal processing elements.

An important question is concerning the practical realizability of a very large phased array. The performance of an extremely large phased-array is limited by mismatches in the gain and phase of each sub-element. In Chapter 2 we explore these limitations in detail, and in particular

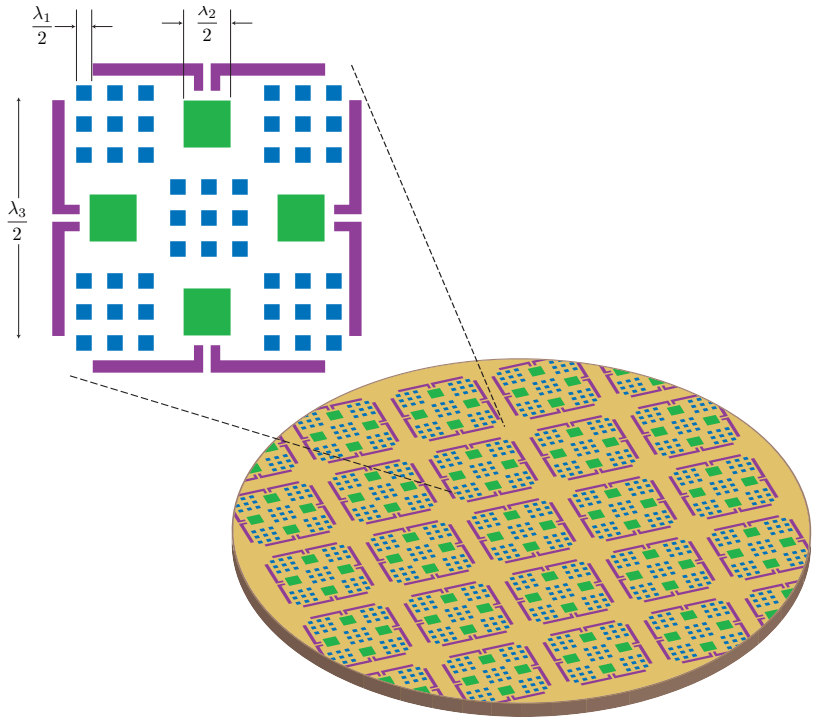


Figure 1.2: Each reticle consists of a pattern of broadband antennas covering lower to higher frequencies with increasing density.

use statistical analysis to show that for a very large array, the effect of inaccuracies can be tolerated if the errors can be kept as random as possible. The antenna pattern on the silicon has a layout density commensurate with the frequency of operation as shown in Fig. 1.2. Since Si on-chip antennas are inefficient and consume large area at the lower edge of the band, a major challenge is the efficient realization of antenna elements. The key issue is to avoid lossy radiation modes in the silicon substrate. One option which was briefly explored is the use of a highly resistive substrate in the silicon processing, such as SOI technology. Alternatively, the antenna can be grown on top of the wafer or using a daughter wafer (Fig. 1.3) with additional low temperature masking steps and electrically shielded from the substrate with a thick dielectric layer. The post-processing option is low cost and mass producible due to the relaxed lithography (10μm) and offers much lower losses due to the spatial isolation from the substrate. This fabrication can be performed with conventional technology, e.g. "redistribution layers" used in a flip-chip process technology. Flip-chip bumps can be used for power and ground whereas all other signals travel off the chip using the integrated antennas. However interconnects from one wafer to another are lossy, especially in the higher frequencies such as 94 GHz, and it is preferable to realize as much functionality in a single wafer as possible.

In this project the on-chip antenna structures were carefully designed and simulated using full-

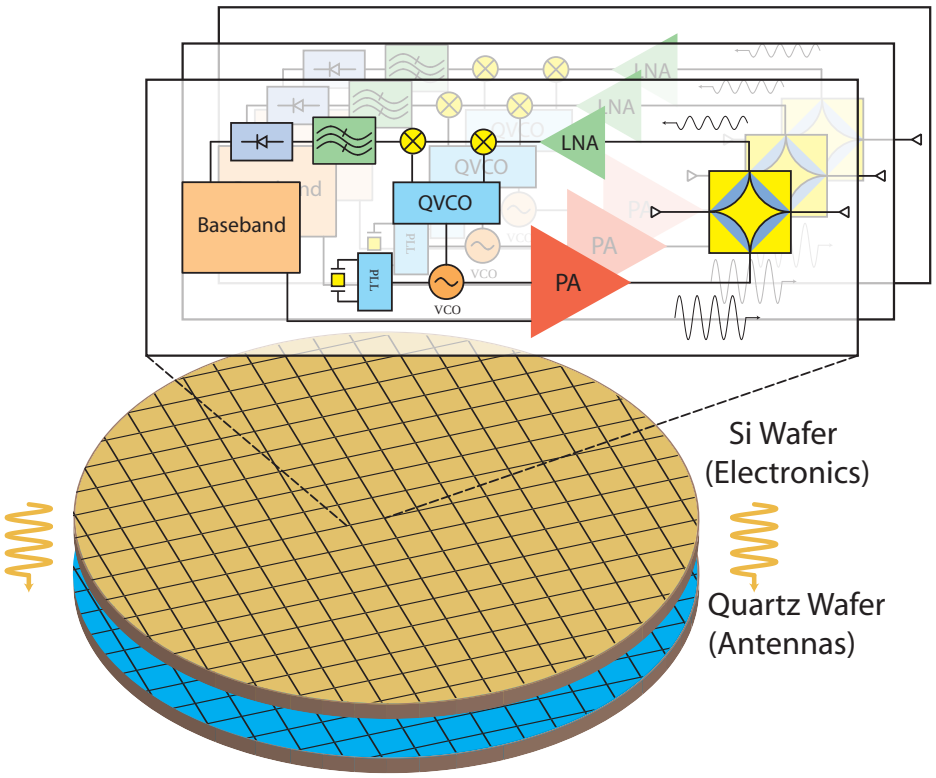


Figure 1.3: A wafer scale distributed radio composed of two wafers. A low-cost antenna wafer is bonded to an advanced silicon wafer containing the electronics.

wave electromagnetic simulation. Simulations of arrays of antennas is key in order to determine practical limits of integration, such as the minimum spacing between antennas, achievable antenna gain and isolation (both isolation in distance and at different frequencies), side lobe radiation levels, the impact of unintentional radiation on other blocks, and spatial nulls achievable to cancel out interference signals. These simulations were carried out and summarized in Chapter 4.

Beam forming is a key ingredient in the function of the proposed system. There are several different techniques possible to introduce the needed phase shift in the signal path, such as RF phase shifters or time-delay elements, LO phase shift, or baseband phase shifting. The baseband phase shifter is the most flexible approach but requires precise timing and synchronization. RF phase shifters are typically narrowband, introduce attenuation unless amplification is used, or require active variable gain elements in an I/Q combiner structure. Active elements such as the VGA generally limit the dynamic range of the system. In Chapter 3, we describe a new phase shifter structure which realizes a synthesized transmission line with variable delay by varying the capacitance of the line through MOS varactors and the inductance of the line through the action of the reflected inductance from the secondary windings of a transformer [3]. The advantage of

the proposed structure is the ability to process wideband signals. An early prototype phase shifter shows a good impedance match and performs up to 11 GHz in 90nm technology. The frequency limit is set by the cut-off frequency of the synthesized transmission line, which improves in direct proportion to the varactor capacitance, which means that technology scaling can lead to a great improvement in the frequency performance of such a structure.

Broadband electronic building blocks are needed for amplification and frequency conversion of the RF energy. Traveling wave amplifiers (TWA) have good output power due to the power combining of multiple devices, good linearity, and reasonable noise performance. We describe new techniques for obtaining higher efficiency from distributed power amplifiers in Chapter 3. A prototype has a simulated bandwidth exceeding 100 GHz in SiGe technology with high output power and efficiency. Due to the extremely broadband operation, these blocks are generic and can be used for an extremely wideband front-end, operational from 10-90 GHz.

The wafer scale radio requires the cooperation of several hundred elements to realize a high power transmitter and a high dynamic range receiver. Self-configuration and testing is enabled by cooperation, eliminating the expensive and time-consuming testing associated with the fabrication of mm-wave systems. Dynamic range is improved by interference cancellation through dynamic beam forming. The noise figure and output power are improved by the phased array spatial power combining and high directivity. For dense urban applications, the sensitivity of the wafer scale is improved due to exploitation of rich spatial diversity in the received signal. None of this is possible unless the system as a whole can dynamically select, adjust, and program the micro radiators and receivers to the appropriate phase, frequency, and power level. In effect, the reticles form a high bandwidth sensor network which must be programmed in a dynamic fashion. Due to variability, the performance of each reticle will be different and some reticles may be non-operational. We investigated the architecture of a high-speed, low-latency "back plane" wired/wireless infrastructure to carry data from the master reticle(s) to the slave reticles. Each packet of data is time stamped with delivery addresses so that a cohort of radiating elements can coordinate with each other to form a beam at the intended frequency. Given that several beams can be formed by such a radiating element carrying different information, the bandwidth of the back-plane (which would effectively need to span the entire wafer) will need to be rather large. We investigated the synchronization limits imposed by technology in the realization of the wafer scale distributed radio to determine the highest data rates and the maximum number of possible beams that can be practically formed. These issues are addressed in Chapter 5.

1.3 Executive Summary

This research project has focused on four primary areas related to the wafer scale distributed radio. First and foremost we investigated the possibility of exploiting a very large array for beam forming and beam nulling (spatial filtering). A careful analysis of the antenna element gain and phase mismatch shows that a very large array, when designed correctly, is actually insensitive to the random errors in the components. This is an important result which contradicts some of the well

known relations between element accuracy and beam pattern. This is very encouraging, though, because in a wafer scale radio we can easily exploit a great number of transceivers and antenna elements but it is much harder to control the exact phase and amplitude of signals distributed across the face of an entire wafer.

Second, we analyzed the performance limitations for integrated antenna elements. To date all measurements on integrated silicon antennas exhibit radiation efficiencies below 10%, which is prohibitively low. In our work we have found that it is possible to realize antenna efficiencies as high as 70% through wafer thinning, and entire arrays can be realized with efficiency of 40%. These results are critical, especially for higher frequencies, since routing signals off chip, or even to a daughter high resistivity wafer, at 94 GHz for instance, is extremely lossy. When one includes the losses of a pad, ESD, and bonding and interconnect, the on-chip antenna solution is very compelling.

Third, we demonstrated that it is possible to build extremely wideband building blocks using silicon technology. These building blocks could form the front-end of a broadband transceiver or be utilized in an ultra-wideband system (such as a radar imager). We focused on the most difficult block, the power amplifier. We show that using a new distributed amplifier topology, where the devices and transmission line impedance are tapered, it is possible to realize bandwidths over 100 GHz with good gain and high output power (17 dBm) and high efficiency (25%). We also demonstrate new architectures for phase shifters which are compact and can provide true time delay, which is critical in wideband communication systems. The new phase shifter can be embedded into a amplifier to realize a Variable Delay Amplifier (VDA).

Finally, we studied the power and area requirements in order to properly synchronize an entire wafer. Synchronization is required since we desire collaboration between transceivers operating across a wafer. In order to beam form or to perform space-time coding, the signals applied to the antennas must be time synchronized. Phase synchronization is also required in a phased-array for beam forming/nulling. Even though systematic offsets can be tolerated, any drift or unknown phase error can lead to degraded performance. We find that by carefully designing the clock-tree as a network of standing wave oscillators, and by utilizing injection locking, a clock tree can be realized with small skew (1ps) while dissipating only 170W on the wafer.

Chapter 2

Phased Array System Design Considerations

This chapter analyzes the effect of errors in antenna weights on the performance of adaptive array systems, both in the case when an array is used to maximize the gain in a desired direction and in the case when an array is used to null interfering signals. We begin by deriving an explicit characterization of the loss in array gain due to phase errors in the optimal antenna weights. Then, we examine interference rejection in the presence of amplitude and phase errors in the antenna weights. We prove that the loss in interference rejection is independent of the number of antennas. For both cases, we give numerical simulations that validate our analysis.

2.1 Impact of phase and amplitude errors on array performance

Many modern communication systems employ adaptive antennas in order to improve their capacity, coverage, and reliability. Unlike conventional fixed antenna systems, adaptive antenna arrays dynamically adjust their beam patterns in response to their environment. Adaptive arrays can extend the range by focusing most of the radio frequency (RF) power on a desired target. This is known as beamforming or beam-steering. Adaptive arrays can also reject unwanted interference signals by placing nulls in the direction of the interferers, which is known as beam-nulling or null-steering. Even if the directions of the interferers are unknown, adaptive arrays can still reduce signal propagation in undesired directions by using side lobe suppression.

Adaptive arrays are composed of multiple antenna elements that can be arranged in different geometries (antennas are usually spaced at least half a wavelength apart) [10]. Larger arrays provide more gain and degrees of freedom. The beam pattern (the locations of peaks and nulls, and the heights of the side lobes) is shaped by controlling the amplitudes and phases of the RF signals transmitted and received from each antenna element. For this reason, adaptive arrays are often referred to as phased arrays. Precise control over both amplitudes and phases is required

to achieve good performance. However, various factors such as finite resolution, noise, mismatch in circuit elements, and channel uncertainty limit the precision that can be achieved in practice. Many of these error sources are random, and cannot be compensated for using pre-calibration or adaptive signal processing techniques. The limited precision will degrade the performance of the array (gain and interference rejection). In this chapter, we examine the impact of phase and amplitude errors on the array gain (beamforming) in Section 2.2 and on interference rejection (beam-nulling) in Section 2.3. We provide both mathematical proofs and simulation results that characterize the array performance as a function of phase and amplitude errors.

2.2 Beamforming

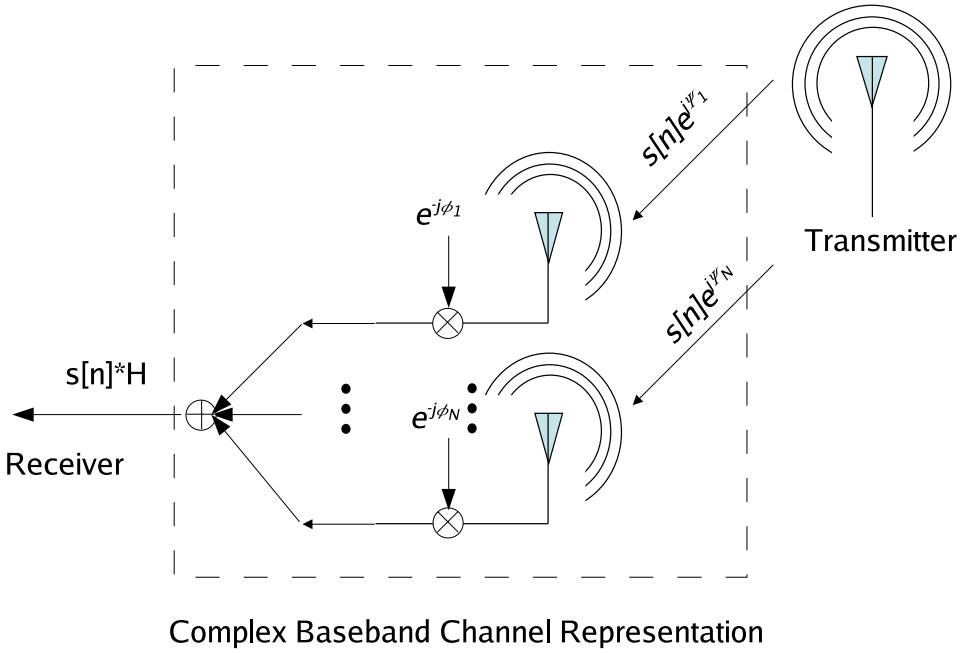


Figure 2.1: A communication system with an adaptive array at the receiver. The narrowband signal $s[n]$ arrives at each antenna shifted in phase by ψ_i. The receiver applies a phase shift of ϕ_i at each antenna and sums the signals.

Consider the array of N elements shown in Figure 2.1. A signal $s[n]$, sent by a remote transmitter, arrives at each antenna i in the array shifted in phase by ψ_i[1]. Antenna i will then apply a

[1]In general, signals arrive at different antenna elements with different delays. However, for narrowband signals, time delays can be approximated with phase shifts [41]. We define signals as narrowband when the fractional bandwidth (the ratio between the signal bandwidth and carrier frequency) is very small (e.g. less than 1%). In this chapter, we shall assume that all signals are narrowband.

phase-shift ϕ_i to the incoming signal. Therefore, the overall complex (baseband) channel response H at the output of the receiver array is given by[2]:

$$H = \sum_{i=1}^{N} e^{j(\psi_i - \phi_i)}$$

To maximize the magnitude of H, ϕ_i is chosen equal to ψ_i, in which case $|H|^2$ reaches its maximum value of $|H_{opt}|^2 = N^2$. In practice, however, factors such as quantization, clock jitter, and other sources of noise make it virtually impossible to realize the desired phase-shifts. Most of these errors are unpredictable and time varying, and are best modeled with random variables:

$$\hat{\psi}_i = \phi_i + \delta_i$$

We will assume that $\delta_i \sim U[-\delta_{max}, \delta_{max}]$, where $0 \leq \delta_{max} \leq 180^\circ$ is an upper bound on the amplitude of phase deviation[3]. Furthermore, we assume that the errors are identically and independently distributed (*i.i.d.*) across different antennas. In this case the channel response becomes:

$$\hat{H}_{opt} = \sum_{i=1}^{N} e^{j\delta_i} = \sum_{i=1}^{N} \cos(\delta_i) + j \sum_{i=1}^{N} \sin(\delta_i)$$

We wish to characterize the effect of the phase errors on the square magnitude of the channel response. To simplify the analysis, we introduce two new random variables $X_i = \cos(\delta_i)$ and $Y_i = \sin(\delta_i)$ and compute the following expectations:

$$\mu_X = E[X_i] = E[\cos(\delta_i)] = \frac{1}{2\delta_{max}} \int_{-\delta_{max}}^{\delta_{max}} \cos(x)dx$$

$$= \frac{1}{\delta_{max}} \int_{0}^{\delta_{max}} \cos(x)dx = \frac{\sin(\delta_{max})}{\delta_{max}}$$

$$\mu_{X^2} = E[X_i^2] = \frac{1}{2\delta_{max}} \int_{-\delta_{max}}^{\delta_{max}} \cos^2(x)dx = \frac{1}{\delta_{max}} \int_{0}^{\delta_{max}} \cos^2(x)dx$$

$$= \frac{1}{2\delta_{max}} \int_{0}^{\delta_{max}} (1 + \cos(2x))dx = \frac{1}{2} + \frac{\sin(2\delta_{max})}{4\delta_{max}}$$

$$\mu_Y = E[Y_i] = E[\sin(\delta_i)] = 0 \quad \text{(by symmetry)}$$

$$\mu_{Y^2} = E[Y_i^2] = \frac{1}{2\delta_{max}} \int_{-\delta_{max}}^{\delta_{max}} \sin^2(x)dx = \frac{1}{\delta_{max}} \int_{0}^{\delta_{max}} \sin^2(x)dx$$

$$= \frac{1}{2\delta_{max}} \int_{0}^{\delta_{max}} (1 - \cos(2x))dx = \frac{1}{2} - \frac{\sin(2\delta_{max})}{4\delta_{max}}$$

[2]The channel response is identical in the scenario with multiple transmitters and a single receive antenna.

[3]We assume a uniform distribution to simplify the calculations. Note that no assumptions were made regarding the geometry of the array or the direction of arrival, so the result holds for an arbitrary array.

Now, we can rewrite the expression for the channel response as:

$$\hat{H}_{opt} = \sum_{i=1}^{N} X_i + j \sum_{i=1}^{N} Y_i$$

$$\Rightarrow |\hat{H}_{opt}|^2 = \left(\sum_{i=1}^{N} X_i \right)^2 + \left(\sum_{i=1}^{N} Y_i \right)^2 = \sum_{k=1}^{N} \sum_{l=1}^{N} (X_k X_l + Y_k Y_l)$$

$$\Rightarrow E[|\hat{H}_{opt}|^2|] = \sum_{k=1}^{N} \sum_{l=1}^{N} (E[X_k X_l] + E[Y_k Y_l])$$

$$E[X_k X_l] = \begin{cases} E[X_k]E[X_l] = \mu_X^2 & \text{when } k \neq l \text{ (using independence)} \\ E[X_k^2] = \mu_{X^2} & \text{when } k = l \end{cases}$$

$$E[Y_k Y_l] = \begin{cases} E[Y_k]E[Y_l] = \mu_Y^2 & \text{when } k \neq l \text{ (using independence)} \\ E[Y_k^2] = \mu_{Y^2} & \text{when } k = l \end{cases}$$

$$\Rightarrow E[|\hat{H}_{opt}|^2|] = (N^2 - N)(\mu_X^2 + \mu_Y^2) + N(\mu_{X^2} + \mu_{Y^2})$$

$$= (N^2 - N) \left(\frac{\sin^2(\delta_{max})}{\delta_{max}^2} \right) + N = (N^2) \left(\frac{\sin^2(\delta_{max})}{\delta_{max}^2} \right) + N \left(1 - \frac{\sin^2(\delta_{max})}{\delta_{max}^2} \right)$$

If we normalize $E[|\hat{H}_{opt}|^2|]$ by dividing by the maximum value $|H_{opt}|^2 = N^2$, we obtain:

$$\Phi_N(\delta_{max}) = \frac{E[|\hat{H}_{opt}|^2|]}{N^2} = \frac{\sin^2(\delta_{max})}{\delta_{max}^2} + \frac{1}{N} \left(1 - \frac{\sin^2(\delta_{max})}{\delta_{max}^2} \right)$$

$$\Rightarrow \Phi(\delta_{max}) = \lim_{N \to \infty} \Phi_N(\delta_{max}) = \frac{\sin^2(\delta_{max})}{\delta_{max}^2}$$

Figure 2.2(a) shows a plot of $\Phi(\delta_{max})$ for $0 \leq \delta_{max} \leq 180^\circ$. Figure 2.2(b) shows the same function in dB scale. Figure 2.2(b) also shows that the calculated array gain closely matches simulation results. Figure 2.2(c) shows that the actual distribution of the phase errors has little impact on the loss in array gain. Notice that using a single bit of phase resolution corresponds to $\delta_{max} = 90^\circ = \frac{\pi}{2}$, and $\Phi(\frac{\pi}{2}) = (\frac{2}{\pi})^2 \approx .4 \approx -3.9dB$! We expect the bound to become tighter as N increases, due to the law of large numbers. The graphs in Figure 2.3 show the loss in array gain as a result of quantizing the phase to one and two bits. Figures 2.3(g,h) show that quantization does not increase the width of the main lobe. Also, notice that when $\delta_{max} = 180^\circ$, which corresponds to completely randomizing the phase of each antenna, the normalized array response $\Phi_N(\pi) = \frac{1}{N}$, which reduces the array gain to that of an omni-directional antenna. So a simple way of creating an omni-like beampattern without reducing the radiated power is to choose the phases randomly[4].

[4]With omni-directional antennas, the absolute, non-normalized power of the signals adds.

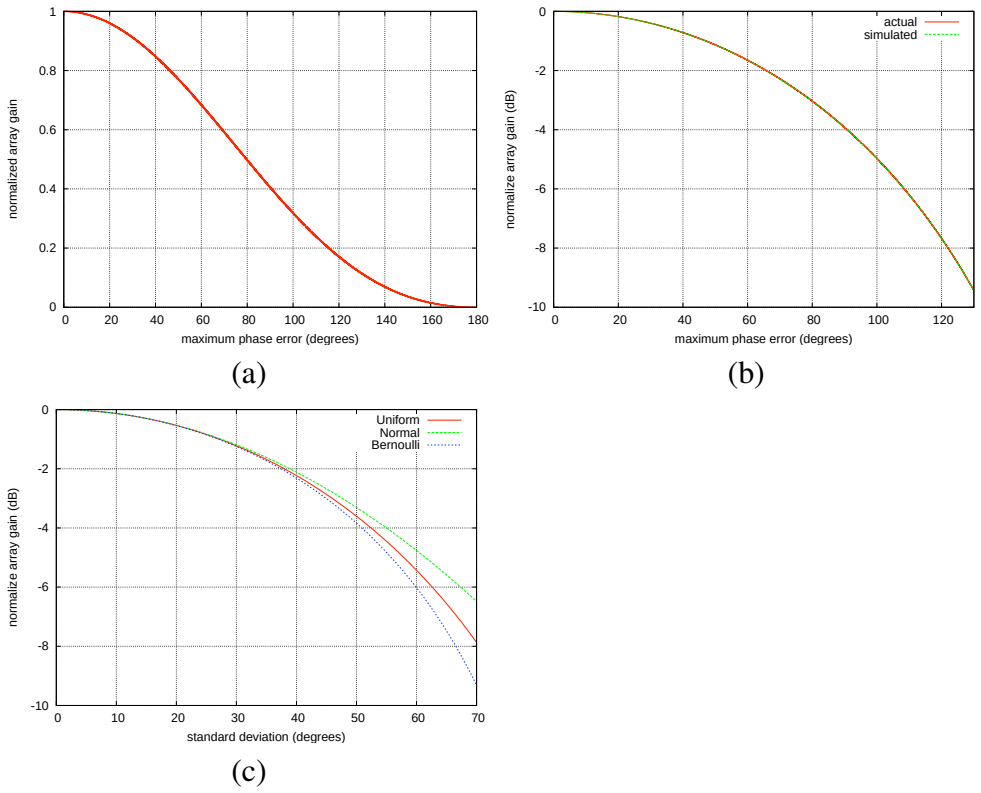


Figure 2.2: (a) The normalized array gain $\Phi(\delta_{max})$ as a function of the maximum phase error δ_{max}. (b) The normalized array gain in dB scale, $10\log\Phi(\delta_{max})$. The plot shows both the calculated gain and the simulated gain for a 10000 element array. (c) The simulated gain (dB) for a 10000 element array with different phase error distributions. For a uniform distribution, the standard deviation is $\sigma_\delta = \delta_{max}/\sqrt{3}$.

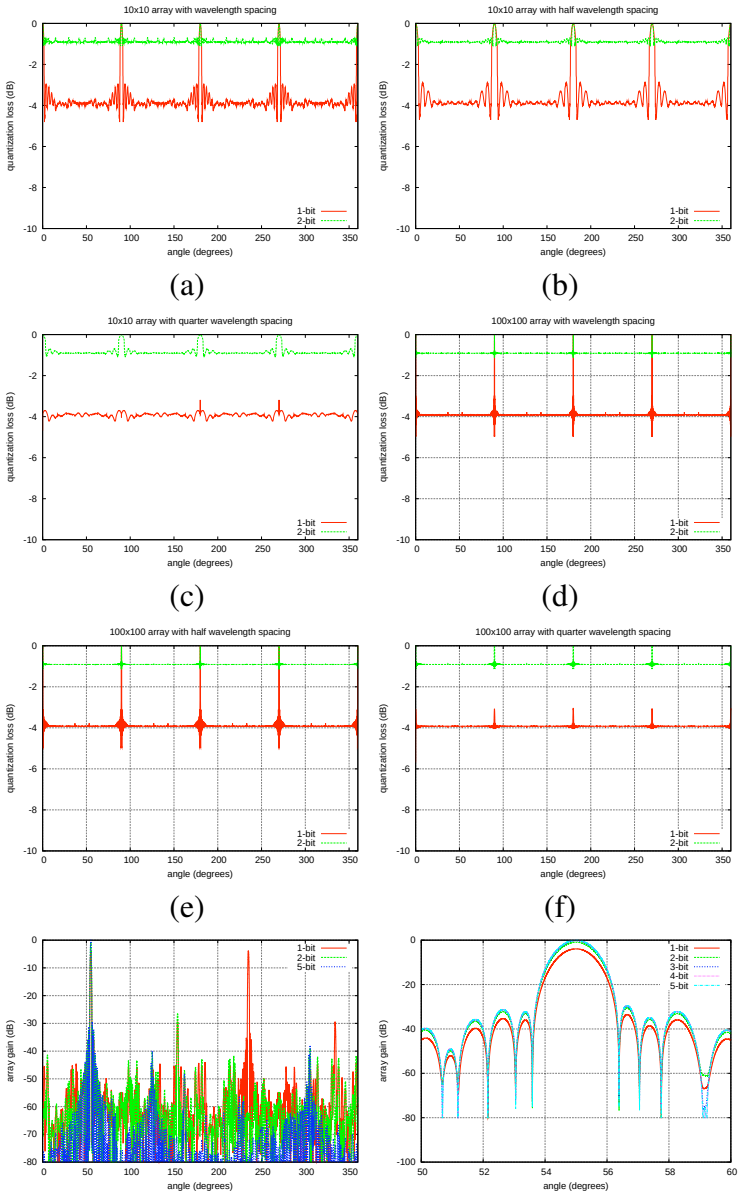


Figure 2.3: (a)-(f) Array loss as a result of phase error for different size arrays. (g)-(h) 2-dimensional horizontal beampattern of a 100x100 array with $\lambda/2$ spacing (steered towards 55 degrees) for different phase resolutions.

A second method of proving a lower bound on the array gain is by using the mean of the random variable, which is often easier to compute, instead of the mean of the square of the random variable. By Jensen's inequality, the square of the mean of a random variable is less than or equal to the mean of its square:

$$E[X]^2 \leq E[X^2]$$

for any random variable X. More generally:

$$f(E[X]) \leq E[f(X)] \quad \text{when } f(\cdot) \text{ is a convex function.}$$

Using this fact, and the expected value of the channel response, we see that:

$$E[\hat{H}_{opt}] = E\left[\sum_{i=1}^{N} X_i + j\sum_{i=1}^{N} Y_i\right] = N\mu_X = N\frac{\sin(\delta_{max})}{\delta_{max}}$$

$$\Rightarrow E[|\hat{H}_{opt}|^2] \geq E[\hat{H}_{opt}]^2 = N^2\left(\frac{\sin(\delta_{max})}{\delta_{max}}\right)^2$$

$$\Phi_N(\delta_{max}) = \frac{E[|\hat{H}_{opt}|^2|]}{N^2} \geq \left(\frac{\sin(\delta_{max})}{\delta_{max}}\right)^2 = \Phi(\delta_{max})$$

2.3 Beam-nulling

In addition to maximizing SNR by steering the direction of the beam towards desired locations, many communication systems are faced with unwanted interference. In most of these systems, simply steering the direction of the peak is not sufficient to suppress large interferers and signal jammers. Other techniques, such as null-steering and side lobe suppression, are required to provide the necessary rejection of interfering signals.

Adaptive systems that require precise control of the locations of the nulls and side lobe levels need to adjust both the phase and amplitude response of each antenna element. In this case, we need to account for both phase and amplitude errors. Analyzing the combined effect of phase and amplitude errors is easier when we consider the problem in the spatial domain where the optimal complex beamforming weights and channel responses can be represented as complex vectors in the N-dimensional Euclidean space, where N is the number of antennas in the array. Let us assume that we have $K+1$ vectors: a desired vector $\mathbf{h}_d$ corresponding to the direction of the desired signal[5], and K interfering vectors $\mathbf{h}_i \ \forall_{1\leq i\leq K}$ corresponding to the directions of K interfering signals.

$$\mathbf{h}_d = [\alpha_{1d}e^{j\beta_{1d}}, \ldots, \alpha_{Nd}e^{j\beta_{Nd}}]^\top$$

$$\mathbf{h}_i = [\alpha_{1i}e^{j\beta_{1i}}, \ldots, \alpha_{Ni}e^{j\beta_{Ni}}]^\top \quad \forall_{1\leq i\leq K}$$

[5]We will denote scalars in lower case, vectors in bold lower case, and matrices in bold upper case.

The incoming signal at the input of the array $\mathbf{y}[n]$ is the sum of the desired signal and interference and noise:

$$\mathbf{y}[n] = \mathbf{h}_d d[n] + \sum_{i=1}^{K} \mathbf{h}_i d_i[n] + \mathbf{v}[n]$$

where $d[n]$ is the desired signal, $d_i[n]$ is interfering signal i, and $\mathbf{v}[n]$ is the white noise vector at the receiver (the variance of each component of $\mathbf{v}[n]$ is σ_v^2). For simplicity, we shall assume that the desired and interfering signals have the same power. Using beamforming weights $\mathbf{w}$ (without loss of generality, we can restrict $|\mathbf{w}| = 1$), the signal at the output of the array will be $\mathbf{w}^H \mathbf{y}[n]$. The output signal to noise plus interference ratio (SINR) is given by:

$$\text{SINR}_{out} = \frac{|\mathbf{w}^H \mathbf{h}_d|^2}{|\sum_{i=1}^{K} \mathbf{w}^H \mathbf{h}_i|^2 + \sigma_v^2}$$

where $(\cdot)^H$ denotes the complex conjugate transpose. Let $\mathbf{H}_I = [\mathbf{h}_1, \ldots, \mathbf{h}_K]$ be the matrix whose columns are the interference vectors. Complete interference rejection can be achieved by choosing a beamforming weight vector $\mathbf{w}$ that is the projection of the desired vector $\mathbf{h}_d$ onto the subspace orthogonal to the column space of $\mathbf{H}_I$ (or the null-space of $\mathbf{H}_I^\top$, which is also known as the left nullspace of $\mathbf{H}_I$), as described in [40]:

$$\mathbf{w}_{opt} = \mathbf{w}_{projection} = \mathbf{h}_d - \mathbf{H}_I (\mathbf{H}_I^H \mathbf{H}_I)^{-1} \mathbf{H}_I^H \mathbf{h}_d$$

We can see that rejecting all the interfering signals is only possible when the left nullspace is non-empty. This is guaranteed when $K < N$. The projection-based beamformer does not take noise into account. In general, maximizing the output SINR does not necessarily require complete interference rejection; reducing the interference to the noise level may be sufficient. Optimizing the output SINR leads to the Minimum Variance Distortionless Response (MVDR) beamformer [25]. If we define the noise+interference correlation matrix $\mathbf{R}_{N+I}$ as:

$$\mathbf{R}_{N+I} = \sum_{i=1}^{K} \mathbf{h}_i^H \mathbf{h}_i + \sigma_v^2 I_N$$

where I_N is the $N \times N$ identity matrix, then the output SINR can be maximized by choosing $\mathbf{w}_{opt}$:

$$\mathbf{w}_{opt} = \mathbf{w}_{\text{MVDR}} = \frac{\mathbf{R}_{N+I}^{-1} \mathbf{h}_d}{\mathbf{h}_d^H \mathbf{R}_{N+I}^{-1} \mathbf{h}_d}$$

The denominator is a normalizing factor. When the interference power is much larger than the noise power, both projection and MVDR yield virtually identical results. In practice, however, phase and amplitude errors degrade the performance of both beamformers[6]. We will assume that

[6]The errors can also result from uncertainties about the channel responses for both desired and interfering signals. Up to this point, we have assumed perfect knowledge of the channels.

an optimum beamformer $\mathbf{w}_{opt}$ is computed using projection, and $\hat{\mathbf{w}}_{opt}$ takes into account both phase and amplitude errors:

$$\mathbf{w}_{opt} = [\alpha_{1w}e^{j\beta_{1w}}, \ldots, \alpha_{Nw}e^{j\beta_{Nw}}]^\top$$

$$\hat{\mathbf{w}}_{opt} = [\alpha_{1w}(1+\varepsilon_1)e^{j(\beta_{1w}+\delta_1)}, \ldots, \alpha_{Nw}(1+\varepsilon_N)e^{j(\beta_{Nw}+\delta_N)}]^\top$$

where $\varepsilon_i \, \forall_{1\leq i\leq N}$ are $i.i.d.$ zero mean real random variables with variance $E[\varepsilon_i^2] = \sigma_\varepsilon^2$, and $\delta_i \, \forall_{1\leq i\leq N}$ are $i.i.d$ zero mean real random variables with variance $E[\delta_i^2] = \sigma_\delta^2$. We also assume that the phase and amplitude errors are independent of each other. Furthermore, we scale the weights so that $\mathbf{w}_{opt}$ has unit norm (i.e. $\sum_{i=1}^N \alpha_{iw}^2 = 1$).

The phase and amplitude errors result in $\hat{\mathbf{w}}_{opt}$ deviating from $\mathbf{w}_{opt}$ by an angle θ. Note that θ does not necessarily correspond to a physical angle or direction. This deviation will result in a reduction in the signal strength in the desired direction as well as an increase in the interference power, since $\hat{\mathbf{w}}_{opt}$ will no longer be orthogonal the interference subspace. The desired power is proportional to $\cos(\theta)$, and the increase in interference (leakage) is proportional to $\sin(\theta)$ (see Figure 2.4). For small angles θ, we can use the standard approximations[7] $\sin(\theta) \approx \theta$ and $\cos(\theta) \approx 1$. Thus, we can characterize the effect of phase and amplitude errors on beam nulls by considering how the mean square angle $\sigma_\theta^2(\sigma_\delta, \sigma_\varepsilon, N) = E[\theta^2]$ behaves as a function of σ_δ, σ_ε, and N.

If we assume that the phase and amplitude variations are small, and given that $\mathbf{w}_{opt}$ is unit norm, then we can approximate the error angle θ with the error vector:

$$\blacksquare\mathbf{w} = \mathbf{w}_{opt} - \hat{\mathbf{w}}_{opt} = [\alpha_{1w}e^{j\beta_{1w}}(1-(1+\varepsilon_1)e^{j\delta_1}), \ldots, \alpha_{Nw}e^{j\beta_{Nw}}(1-(1+\varepsilon_N)e^{j\delta_N})]^\top$$

We can further simplify the above expression using the approximations $\cos(\delta_i) \approx 1$, $\sin(\delta_i) \approx \delta_i$, and $\delta_i\varepsilon_i \approx 0$.

$$\begin{aligned}\blacksquare\mathbf{w} &= [\alpha_{1w}e^{j\beta_{1w}}(1-1-\varepsilon_1 - j\delta_1), \ldots, \alpha_{Nw}e^{j\beta_{Nw}}(1-1-\varepsilon_N - j\delta_N)]^\top \\ &= [-\alpha_{1w}e^{j\beta_{1w}}(\varepsilon_1 + j\delta_1), \ldots, -\alpha_{Nw}e^{j\beta_{Nw}}(\varepsilon_N + j\delta_N)]^\top \\ &\Rightarrow |\blacksquare\mathbf{w}|^2 = (\blacksquare\mathbf{w})^H(\blacksquare\mathbf{w}) = \sum_{i=1}^N \alpha_{iw}^2(\varepsilon_i^2 + \delta_i^2)\end{aligned}$$

By taking the expectation of this expression:

$$\sigma_\theta^2 \approx E[|\blacksquare\mathbf{w}|^2] = E\left[\sum_{i=1}^N \alpha_{iw}^2(\varepsilon_i^2+\delta_i^2)\right] = \sum_{i=1}^N \alpha_{iw}^2(E[\varepsilon_i^2]+E[\delta_i^2])$$

$$= \sum_{i=1}^N \alpha_{iw}^2(\sigma_\varepsilon^2+\sigma_\delta^2) = (\sigma_\varepsilon^2+\sigma_\delta^2)\sum_{i=1}^N \alpha_{iw}^2 = \sigma_\varepsilon^2+\sigma_\delta^2$$

[7] This explains why nulls are more sensitive than peaks to phase and amplitude errors, since $\sin(\theta)$ changes more rapidly than $\cos(\theta)$ when θ is small.

As we can see, the mean square error angle σ_θ^2 is equal to the sum of the mean square phase error σ_δ^2 and the mean square amplitude error σ_ε^2. The key conclusion that we draw from this result is that the angle error is independent of N, the number of antennas[8].

The simulation results shown in Figure 2.5 verify this result. Figure 2.5(a) shows a linear relationship between $10\log(\sigma_\varepsilon^2+\sigma_\delta^2)$ (x-axis) and $10\log(\sigma_\theta^2)$ (y-axis), with slope equal to 1. Figure 2.5(b) shows that the relationship between $10\log(\sigma_\delta^2)$ and $10\log(\sigma_\theta^2)$, when the amplitude errors ε_i are set to 0, is also linear with slope equal to 1, which demonstrates that the phase and amplitude errors contribute equally to the overall error angle θ. In both Figures 2.5(a) and 2.5(b), we simulated a 100 element array. We repeated the same experiment for a 1000 element array and the results are identical, as shown in Figures 2.5(c) and 2.5(d). Figure 2.6 also shows identical results, where we plot the interference rejection as a function of phase and amplitude errors for different array sizes. This shows that the number of antenna elements has no effect on the mean square error angle σ_θ^2 or the interference rejection. This means that the depth of beam nulls is limited by gain and phase accuracy, and is independent of the size of the array N and the number of interferers K, as long as $N > K$.

2.4 Conclusion

Adaptive antenna arrays are a key component in many modern communication systems. They are used to both increase the gain in the direction of a desired signal as well as to reject interfering signals. However, when these adaptive arrays are implemented, a variety of practical considerations will cause the actual antenna weights to differ from the optimal weights, which in turn degrades the performance of the array.

In this chapter, we analyzed the performance loss due to phase and amplitude errors in the weights. We began by considering a beamforming system, which maximizes the gain in a desired direction. We derived an expression for the loss in gain due to uniform phase errors, and provided simulations that validate this result. Then, we considered a beam-nulling system, which rejects interfering signals. We analyzed the effect of uniform amplitude and phase errors, and again provide numerical simulations. We showed that the interference rejection is a function of the errors in the weights, and is independent of the number of antennas, assuming that there are more antennas than interferers.

[8]The power leakage into the interference subspace is independent of the number of antennas. However, increasing the number of antennas can still increase the output SINR (peak to null ratio) by increasing the power gain of the desired signal. Increasing the number of antennas also increases the degrees freedom necessary to null more interferers.

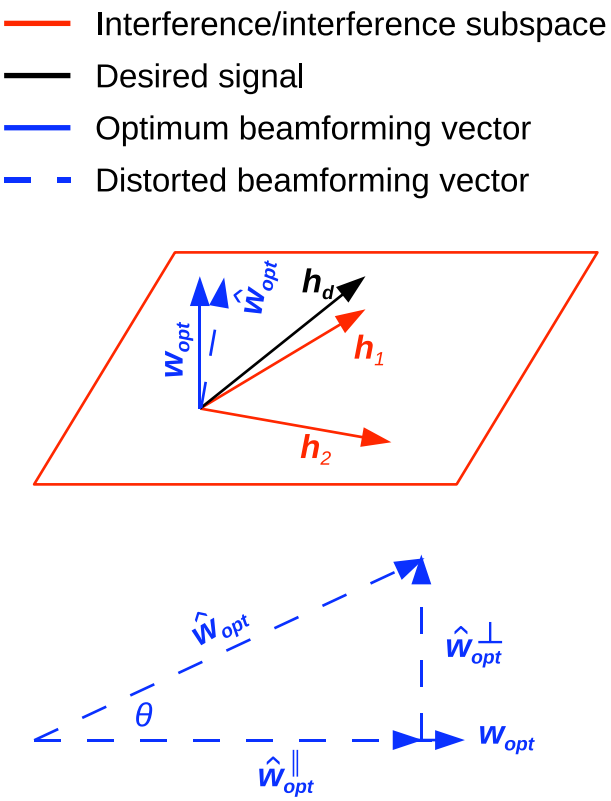


Figure 2.4: The optimum beamforming vector $\mathbf{w}_{opt}$ can be viewed as a projection of the desired signal onto the subspace orthogonal to the interference subspace. The distorted beamforming vector $\hat{\mathbf{w}}_{opt}$ can decomposed into two orthogonal components: $\hat{\mathbf{w}}_{opt} = \hat{\mathbf{w}}_{opt}^{\perp} + \hat{\mathbf{w}}_{opt}^{\parallel}$. $\hat{\mathbf{w}}_{opt}^{\parallel}$, which is parallel to $\mathbf{w}_{opt}$, represents the potential loss in beamforming gain, and is proportional to $\cos(\theta)$. $\hat{\mathbf{w}}_{opt}^{\perp}$, which is orthogonal to $\mathbf{w}_{opt}$, represents the potential leakage into the interference subspace, and is proportional to $\sin(\theta)$.

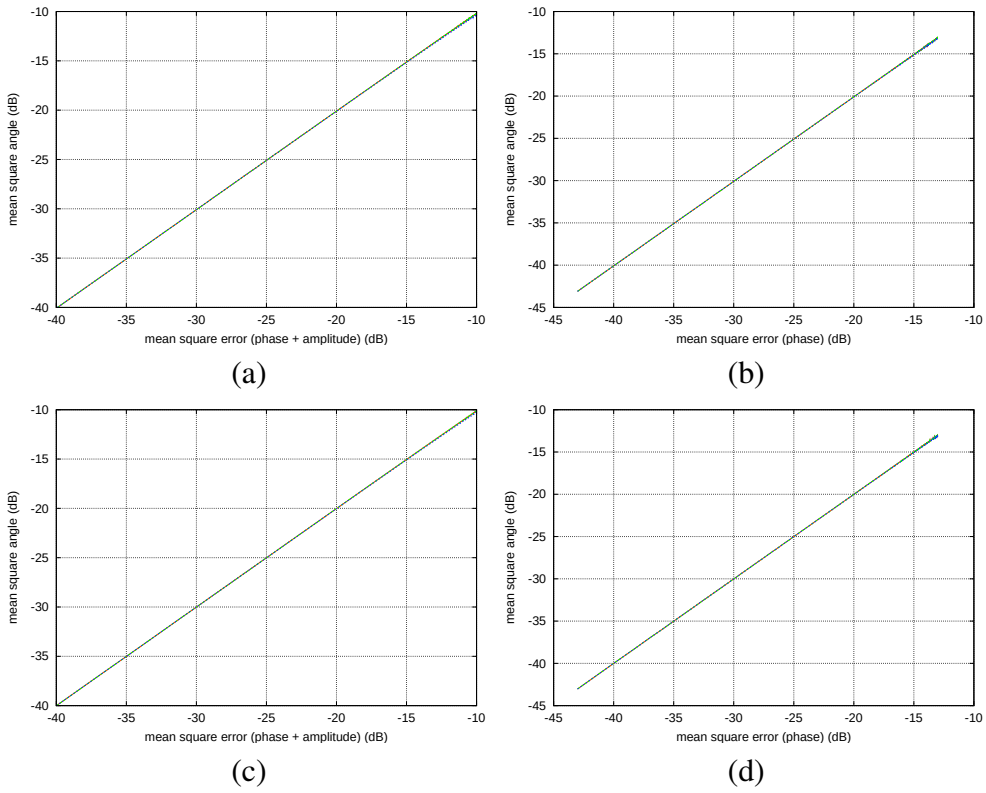


Figure 2.5: Simulated relationship between phase and amplitude errors and the mean square error angle: (a),(c) $10\log(\sigma_\varepsilon^2 + \sigma_\delta^2)$ on the x-axis, $10\log(\sigma_\theta^2)$ on the y-axis. (b),(d) $10\log(\sigma_\delta^2)$ on the x-axis, $10\log(\sigma_\theta^2)$ on the y-axis. In (a) and (b), the simulated array had 100 elements. In (c) and (d), the simulated array had 1000 elements.

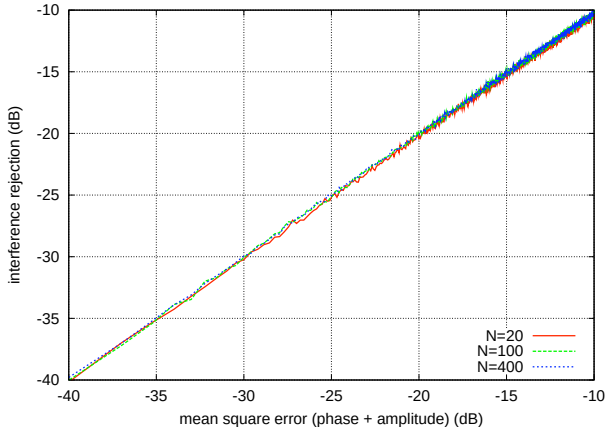


Figure 2.6: Simulated power leakage (interference rejection) as a function of phase and amplitude errors: $10\log(\sigma_\varepsilon^2 + \sigma_\delta^2)$ on the x-axis versus the interference rejection in dB on y-axis. Interference rejection (IR) is defined as the ratio of the interferer power after beam-nulling to the interferer power before beam-nulling. Before beam-nulling, we choose the beamforming vector $\mathbf{w}_{before}$ along the direction of the desired signal $\mathbf{h}_d$ (i.e. $\mathbf{w}_{before} = \frac{\mathbf{h}_d}{|\mathbf{h}_d|}$). In this case, the input power at the receiver from interferer $\mathbf{h}_i$ will be $|\mathbf{w}_{before}^H \mathbf{h}_i|^2$. After beam-nulling, we choose the beamforming vector $\mathbf{w}_{after}$ as the projection of the desired vector onto the subspace orthogonal to the interference subspace. In this case, the input power at the receiver from interferer $\mathbf{h}_i$ after beam-nulling (and phase and amplitude distortion) will be $|\hat{\mathbf{w}}_{after}^H \mathbf{h}_i|^2$. Thus, on the y-axis, the interference rejection $\text{IR} = 20\log \frac{|\hat{\mathbf{w}}_{after}^H \mathbf{h}_i|}{|\mathbf{w}_{before}^H \mathbf{h}_i|}$. The relationship is plotted for several array sizes. Both $\mathbf{h}_d$ and $\mathbf{h}_i$ are complex random vectors whose components (both real and imaginary parts) are sampled independently from a standard Gaussian distribution.

Chapter 3

Circuit Design and Technology Considerations

3.1 Introduction

Replacing a traditional single antenna with a phased-antenna array results in a tunable antenna pattern. Suitable gain and phase vectors are applied to each elements in order to shape the array pattern. This can be employed at the transmitter to efficiently transmit the power to the desired direction and/or at the receiver to spatially filter out unwanted signals. The array gain is proportional to the number of elements, and output power and receiver sensitivity improve as the array gets larger at the expense of having a narrower beam. If the whole wafer is dedicated to demonstrate a large phased array system at mm-wave frequencies (for example 90GHz) we will have:

$$f = 90GHz \Rightarrow \lambda = 3mm \tag{3.1}$$

$$Wafer\ size = 300mm \Rightarrow N = \frac{\pi R^2}{(\lambda/2)^2} = 4\pi(R/\lambda)^2 \sim 25000 \tag{3.2}$$

This large number of elements results in 44dB of array gain (antenna directivity). There are 25,000 integrated radiators which means that the effective radiated output power is 25,000 times more than the case of a single radiator connected to an antenna with 44dB of directivity. In the other words, the total gain will be 88dB with respect to a single element connected to an omni-directional antenna. Assuming each radiator is capable of transmitting 10mW of output power [15] the Effective Isotropic Radiated Power (EIRP) will come to:

$$EIRP = P_{TX} \cdot N \cdot G \sim 6MW \tag{3.3}$$

In reality on-chip antennas have a 10% efficiency which reduces the effective output radiated power to 600KW. This enormous power level opens up new applications and opportunities for the

silicon technology. Estimating the efficiency of each transceiver to be 10%, a dc power of about

$$P_{DC} = 10 \cdot 10mW \cdot 25000 = 2.5KW \tag{3.4}$$

should be supplied. This raises additional concerns about the feasibility and reliability of providing the DC power to the chip and removing the dissipated heat off from it. Fortunately advances in the TSV (Thru-Silicon-Via) techniques help. This technology allows metalization on the back side of the wafer as another means to access the active devices and leaves the front side of the wafer to be devoted to radiating elements. As a result of the high number of radiating elements, the beam is ultra narrow and spanned over a sub-degree angle which makes the system more sensitive but also greatly enhances the resolution for identifying objects.

3.2 Phased Array Architectures

Phase shifting can be done in either digital signal processing domain (Fig. 3.1), LO frequency and clock distribution network (Fig. 3.2) or RF path (Fig. 3.3), with advantages and disadvantages associated with each method in terms of die area, power consumption, dynamic range and programmability.

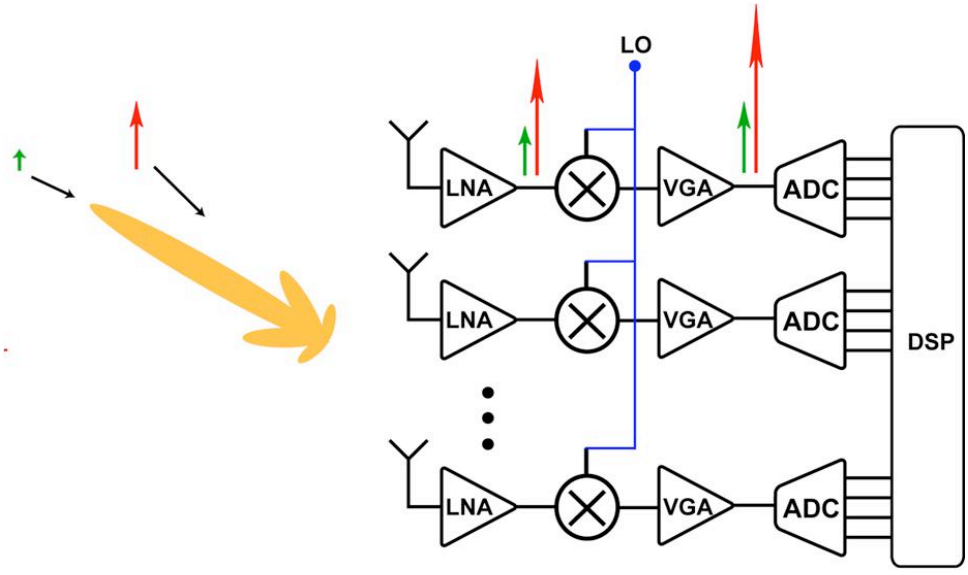


Figure 3.1: A digital phase shifting architecture.

Providing the phase shift in the digital domain makes the system highly flexible and exploits high-speed DSPs to run complex high resolutions algorithms in digital domain (Fig. 3.1). How-

ever in a digital beamformer, all building blocks are replicated for each path, making the system demanding in terms of area and power consumption. Moreover since the spatial filtering is done in the DSP, all the building blocks prior to that should have enough dynamic range to cope with the large blocker levels.

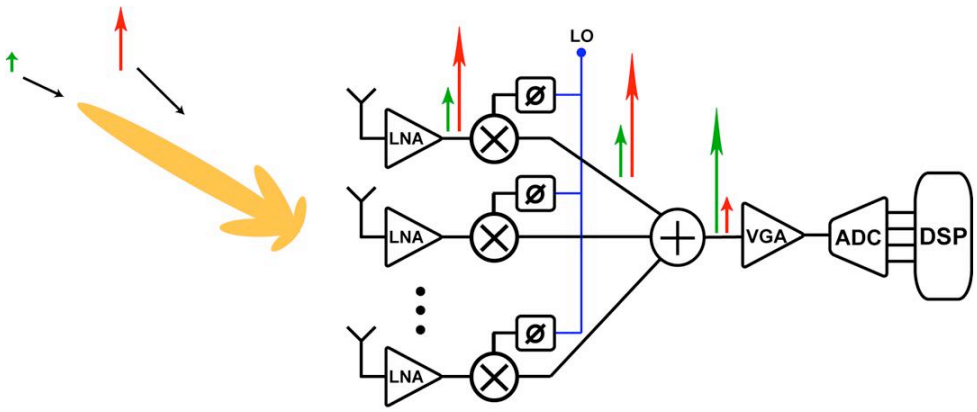


Figure 3.2: An LO phase shifting architecture.

Phase shifting in the LO and clock distribution network (Fig. 3.2) reduces the number of ADCs and baseband circuitry and relaxes the dynamic range requirements. The spatial filtering after the signal combining attenuates interfering signals, which in general have a different angle of arrival, but moves the bottleneck of the system to the LO distribution network, which should be highly symmetric to provide the exact desired phase shift. The LO distribution network also burns a lot of power in buffer stages to provide a strong enough LO signal to each of the many mixers in the array.

RF phase shifting offers the simplest system (Fig. 3.3) in terms of lowest component count and power consumption, at the expense of designing challenging RF phase shifting elements and including their nonidealites (loss, noise, nonlinearity) directly in the RF path. In the proposed wafer scale radio, the number of elements is extremely large, and therefore to reduce the complexity and component count of the system, RF phase shifting is the best candidate. As presented in following sections, novel architectures along with advances in the performance of silicon technology promises the realization of adding phase shifting elements directly to the RF path of the signal. Although providing the phase shift at RF is less flexible than the digital beamformer in terms of programmability, nonetheless simple signal processing such as windowing and spatial spectral filtering can be done by manipulating the gain of each path on top of adjusting its phase, hence having a building block that alters both the gain and phase is highly desirable.

As depicted in Fig. 3.4 a triangular window function decreases the gain of the main lobe by 6dB and broadens the beamwidth, but it greatly reduces the sidelobe levels and minimizes the

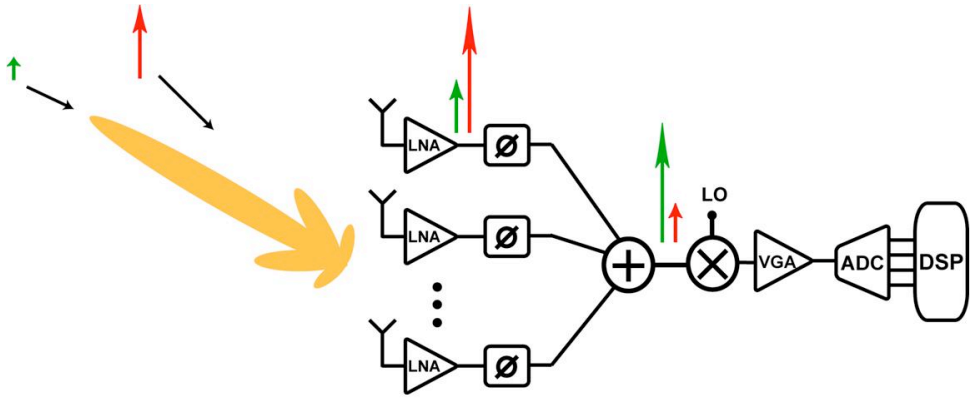


Figure 3.3: An RF phase shifting architecture.

effect of signals coming from unwanted directions. By using more complicated gain and phase vectors, null steering can be done to cope with very strong jammers.

Having signals from all 25,000 elements being phase shifted and combined in the RF domain raises severe issues about the loss of the signal combiner/divider and the poor overall flexibility of the system. Therefore the final solution will be a combination of digital and RF phase shifting where the total area is divided into reticles of sub-arrays. In each subarray, beamforming is achieved by RF phase shifting and then the sub-arrays are connected and synchronized via the relatively lower clock frequency and will be programmed digitally. Beams of each sub-array can be unique or can be combined with the beam(s) of neighboring arrays to create a stronger beam. Multiple target tracking at different directions and different frequency bands will be controlled by the high-speed DSPs that program the sub-arrays.

3.3 True Time Delay Elements Versus Phase Shifters

As shown in the Fig. 3.5, the wavefront reaches each antenna element of a phased array with a dely shift of $\frac{d}{c}\sin\theta$, and in order to steer the angle of look towards the direction of incoming signal, this delay should be compensated before combining signals from each path. Delays correspond to linear phase shift in the frequency domain. A linear phase shift can be approximated as constant phase shift over a narrow bandwidth.

$$\Delta\varphi = \omega\Delta t \sim \omega_o\Delta t = k_o d\sin\theta, d = \lambda/2 \Rightarrow \Delta\varphi = \pi\sin\theta \tag{3.5}$$

As the bandwidth increases, the delay-phase approximation fails to be accurate. For small arrays (4-8 elements) delay-phase approximation works up to 20% of fractional bandwidth, but as the number of array elements get larger, the aperture size becomes larger and the array will be more

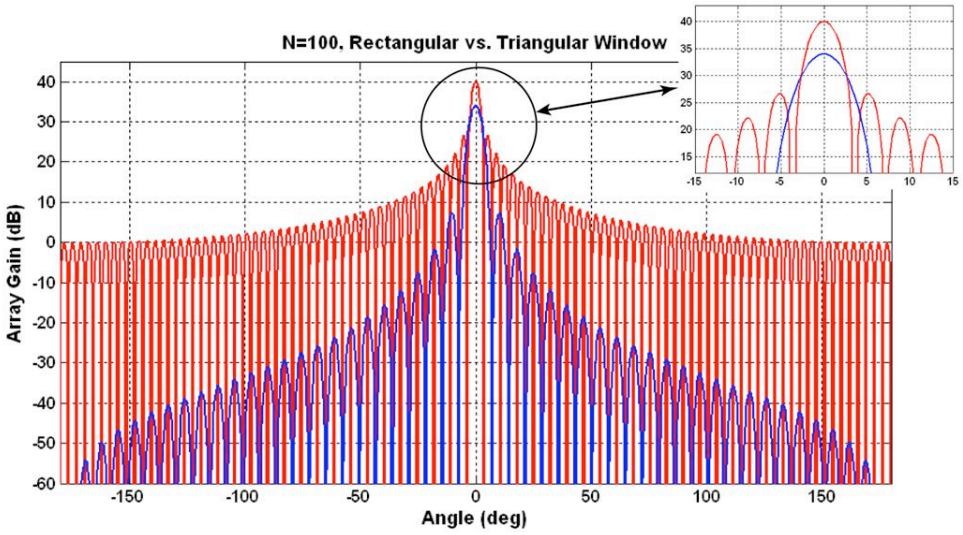


Figure 3.4: Windowing can improve the array factor of a phased-array by reducing the side-lobe levels.

directive. For a linear N-element array

$$D = (N-1)\lambda/2 \;\Rightarrow\; Beamwidth = \lambda/D = \frac{2}{N-1} \tag{3.6}$$

Which shows that the beam gets narrower as N increases. Having constant phase shift corresponds to variable group delay over the band, which correspondingly results in different direction of look over the bandwidth

$$\theta_{new} = \arcsin\left(\frac{\omega \pm \Delta\omega}{\omega_o}\sin\theta_o\right) \tag{3.7}$$

For large arrays with extremely small beamwidth, the array pattern overlaps for beams operating at the band edges can be minimal as depicted in Fig. 3.6. Due to the nonlinear function of $\Delta\varphi = \pi \sin\theta$, the situation worsens as the direction of main beam moves from broadside towards end-fire. All above observations show that for a wafer-size array true time delay elements are needed instead of phase shifters.

3.4 True Time Delay Elements

Traditionally delay elements are implemented via switched transmission line networks. These switched passive structures have a high bandwidth and low insertion loss, but they also occupy a large footprint. In a wafer-scale radio it is desired that the size of the array be dominated by

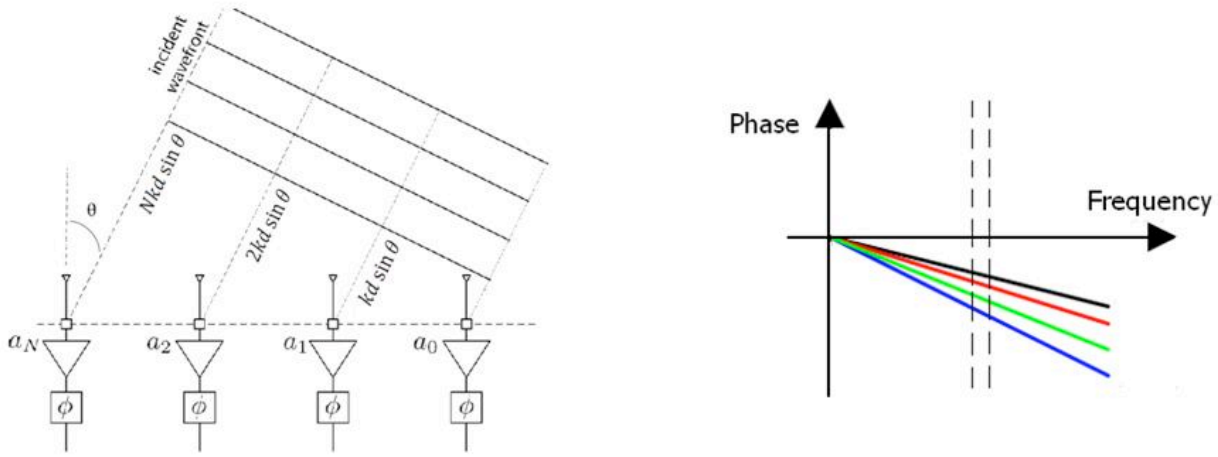


Figure 3.5: The wavefront of a plane wave impinges at an angle to a phased-array. All signals can be summed in phase if a uniform time delay is applied to the array elements.

antenna elements and the electronics be small enough to fit in the inter-antenna spacing. As the electronic circuitry becomes larger and comparable to the antenna size, fewer number of antennas will fit on the wafer and the array gain will be reduced.

To decrease the size of delay elements, synthesized (LC) transmission lines have been used in the literature. By this technique the size of the array will be greatly reduced. To tune the delay of a section of the artificial transmission line $T_D = \sqrt{LC}$, the capacitance of the line is modified by the varactor loading. However as capacitance variation also changes the charactristic impedance of the line $Z_o = \sqrt{\frac{L}{C}}$, and the delay variation is limited (usually up to 20%) in order not to violate matching requirements.

In our work we developed a technique where both the capacitance and inductance of the line are adjustable. As a result, the Z_o of the line will be kept relatively constant as the delay of each section is greatly varied. The inductance of a single loop is a function of its geometry and the permeability of surrounding materials that are both fixed after the IC fabrication and will not provide means for the inductance tuning. As inductance of a loop is defined by total magnetic flux passing through the loop divided by the loop current ($L = \frac{\varphi}{i}$), another nearby loop current can be manipulated to alter the flux in the desired main loop. Therefore a transformer can be used where its secondary current is a multiplicative copy of the primary current and the total effective inductance seen at the primary is calculated by

$$i_2 = n \cdot i_1 \;\Rightarrow\; L_{effective} = L_1 + n \cdot M \tag{3.8}$$

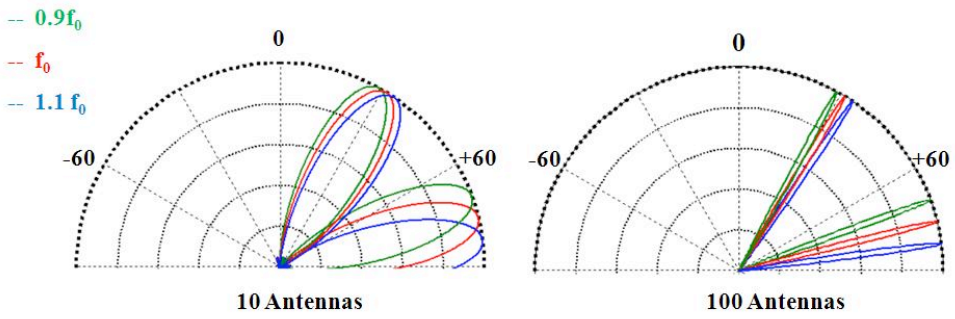


Figure 3.6: Simulated beamwidth for a 10 element and 100 element array at the center of the band (where the phase shift is ideal) versus 10% away from the center. A large array shows a high sensitivity as a function of frequency.

For a 1:1 transformer

$$L_1 = L_2 = L \Rightarrow M = k \cdot L \ , \ L_{effective} = (1 + n \cdot k)L \tag{3.9}$$

Fig. 3.7 demonstrates a transformer embedded in a switching network that provides the capability of reversing the current direction in the secondary ($n = \pm 1$). Therefore the effective inductance seen by each loop can be either $L_{low} = (1-k)L$ or $L_{high} = (1+k)L$. If the capacitance ratio matches this inductance ratio, $\frac{C_{max}}{C_{min}} \frac{L_{max}}{L_{min}} = \frac{1+k}{1-k}$, then the delay variation will be $\frac{T_{high}}{T_{low}} = \frac{1+k}{1-k}$.

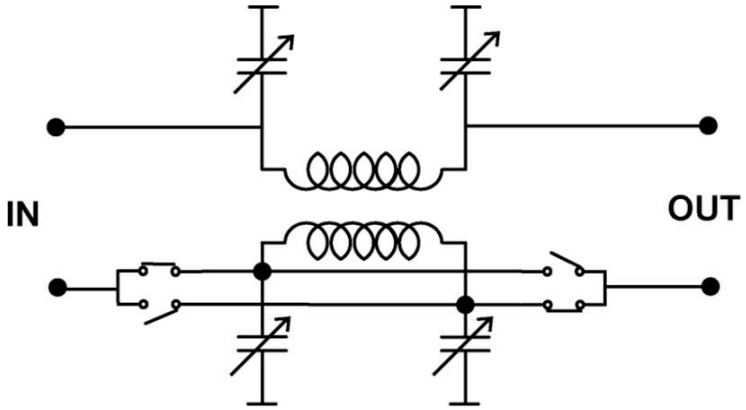


Figure 3.7: The proposed switched-transformer delay element.

For $k = 0.5$, which is easily achievable in integrated transformers on silicon, a delay ratio of $\frac{T_{high}}{T_{low}} = 3$ will be achieved (which is much larger than the 20% traditionally obtained) while maintaining matching requirements over the bandwidth $Z_{high} = Z_{low}$.

Series voltage switching mandates placing a MOS switch in series with the transformer that adds to the loss of the network. To make the loss negligible, switches become prohibitively large and their parasitic capacitances limit the bandwidth. Simulation verified by measurement results show that the series switching network is functional up to 10GHz in 90nm CMOS (Fig. 3.8).

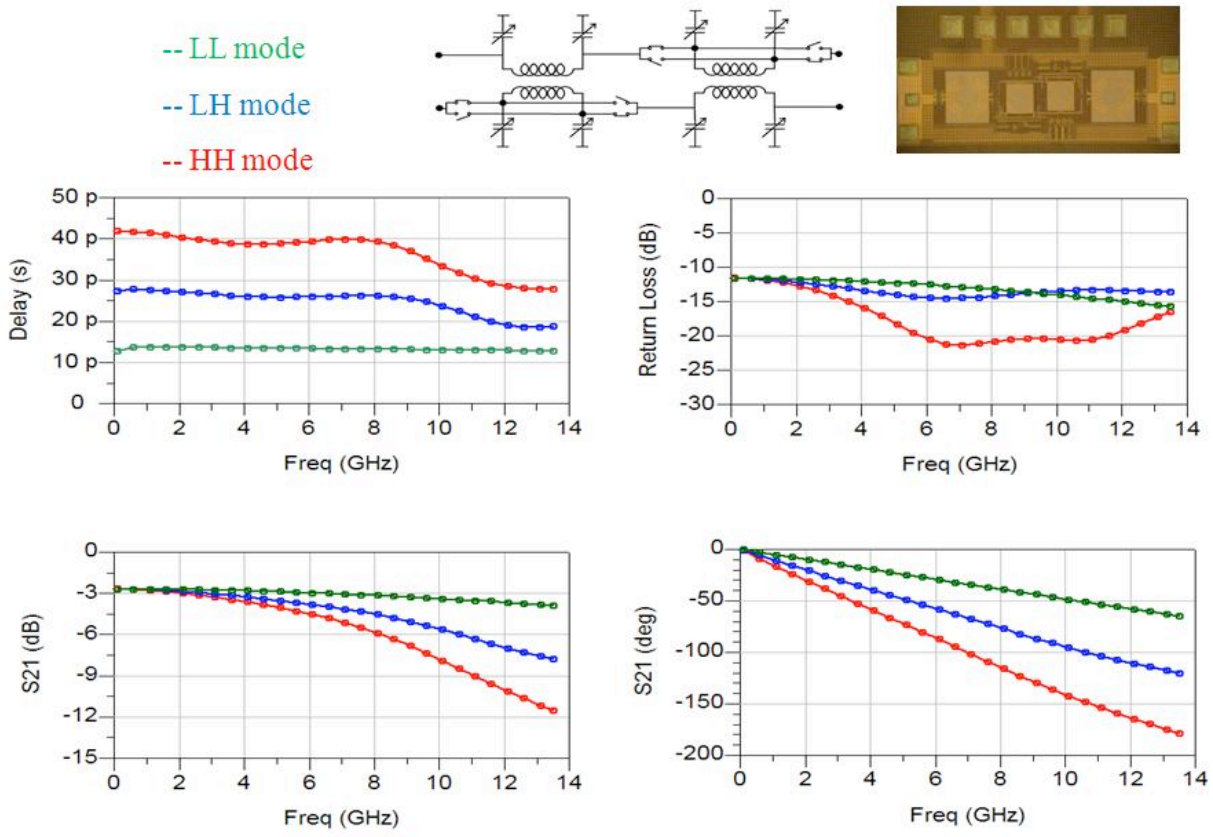


Figure 3.8: Measurement results for a switched transformer based delay element.

To improve the operation frequency, a CML-like current mode switching network is adopted (Fig. 3.9). Having parallel transistors on the secondary side of the transformer to provide the switching function converts the structure to a differential cascode amplifier with a switchable artificial transmission line section between the top and bottom transistors. Hence a variable delay amplifier(VDA) is achieved. As the signal passes through this building block, it gets amplified and also a desired delay will be applied to it.

Simulation result of such an structure in 0.13μm IBM 8HP-SiGe technology shows that a cascade of 3 of these amplifiers provide 24dB of gain in a 30GHz bandwidth around 60GHz while provideing 12ps of delay in 4ps steps.

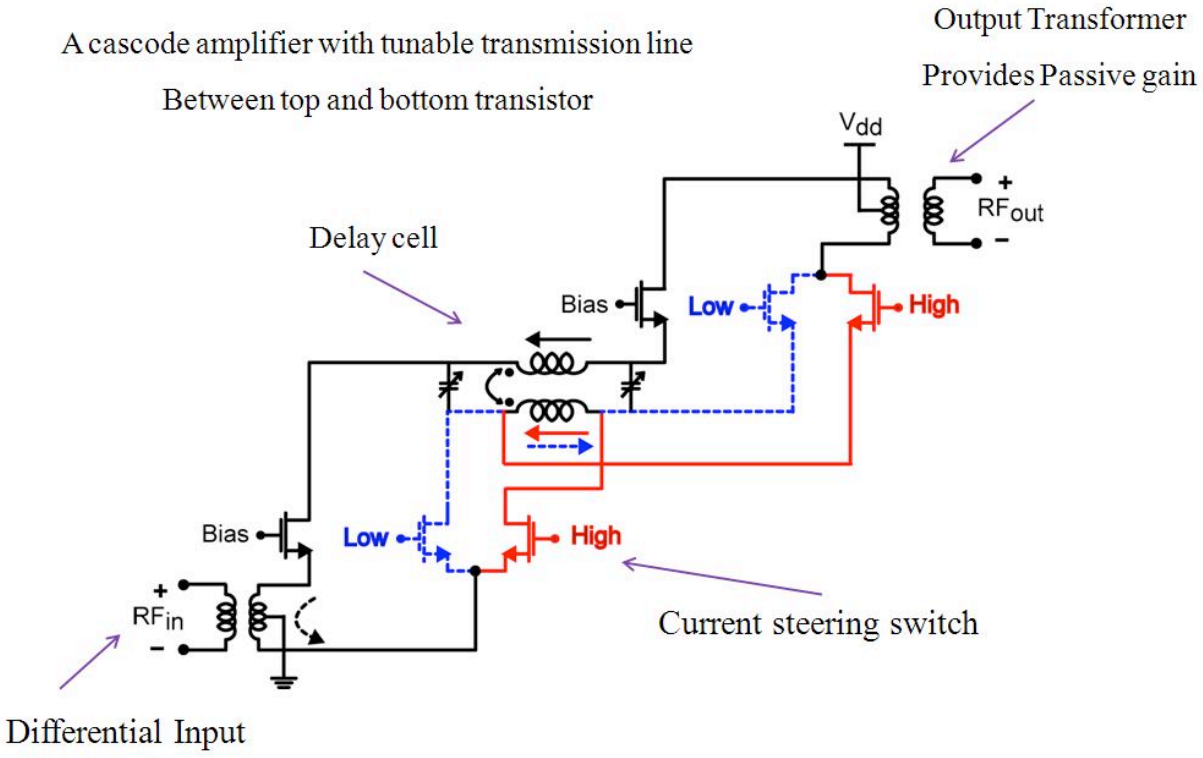


Figure 3.9: Schematic of proposed Variable Delay Amplifier (VDA) architecture.

3.5 Conclusion

Issues and benefits of large phased array systems were discussed. Different schemes for a phase shifting system were studied and a combination of RF and digital phase shifting is proposed. To be applicable to non-narrowband and large arrays, true time delays should be used rather than phase shifting approximations. Different tunable delay mechanism were explored and in order not to limit the size of the array by the delay elements, artificial transmission lines with tuning capability on both the capacitance and inductance of the line were studied which resulted into a new variable delay amplifier architecture with a promising simulated performance.

3.6 Wideband Distributed Power Amplifier Design Based on Device Size and Output T-Line Impedance Tapering

3.7 Introduction

Wideband amplifiers are critical building blocks for future wafer-scale radio front-ends. Depending on the application, bandwidth of 80GHz or even higher is desired. In the world of analog designs, amplifier bandwidth is very limited due to the existence of parasitic capacitors. One popular way of doing high frequency design is to use tuned amplifiers, where parasitic capacitors are resonated out by inductors at desired frequency, but these amplifiers are inherently narrowband, especially when the Q of the resonance network is large. On the other hand, the cutoff frequency of an artificial transmission line can be very high, often in excess of 100GHz, as long as the distributed inductance and capacitance are small in each section. Therefore by absorbing the transistor parasitic capacitance into artificial transmission lines, distributed amplifiers can be built with much higher bandwidth. In recent years, a number of silicon based DAs (Distributed Amplifiers) have been reported with bandwidth approaching 100GHz and decent power gain. Yet one major problem of these DAs is the low power efficiency, preventing them to be used for power amplification. In this project, circuit techniques and design considerations are investigated in order to maximize the DA efficiency. A simultaneous device size and output T-Line impedance tapering technique is proposed to improve the efficiency of these extremely wideband structures without gain or bandwidth degradation.

3.8 Distributed Power Amplifier Design

3.8.1 Concept of device size and output T-Line impedance tapering

In a conventional distributed amplifier, the input signal travels along the input transmission line and gets amplified by the gain cells. The amplified signals then add constructively when they travel towards the load. Because of this nature, one can easily notice that the largest voltage swing occurs at the last gain stage since it is the summation of voltage swings of all previous stages. This means only the last stage experiences maximum allowed voltage swing, usually defined by the supply voltage, under saturation power level while previous gain stages never reach that level. Since power is the product of voltage and current, if more voltage swing is utilized in the previous stages, less current swing is needed to produce the same power, which in turn means less bias current is needed for them. To achieve this, the output T-Line characteristic impedance should gradually increase from the load to the termination resistor but the gain cell device size as well as the cell bias current should decrease in the same direction. The basic structure is illustrated in Fig. 3.10. Since less current is used to produce the same output power level, the overall efficiency can be improved. To prove the concept, two types of distributed power amplifiers are designed,

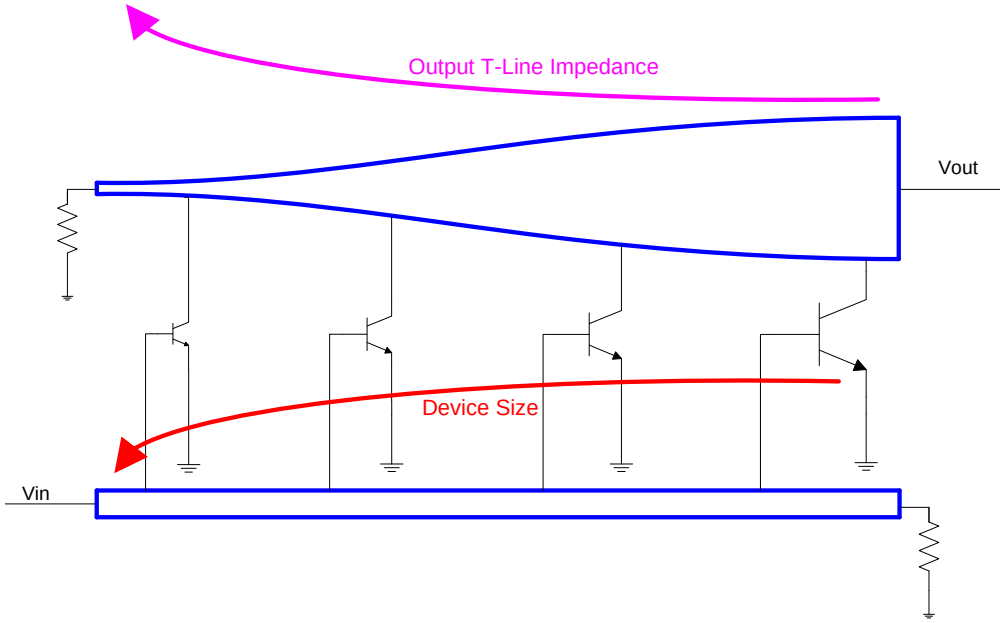


Figure 3.10: Tapered distributed power amplifier structure.

one based on cascode gain cells while the other based on common-emitter gain cells. Both of them are designed and simulated in IBMs 0.13μm SiGe BiCMOS process. The following two sub-sections will discuss the design procedure in detail.

3.8.2 General design guidelines for distributed amplifier gain cells

For a distributed amplifier, it is desirable to design gain cells with small input conductance and susceptance. Small input conductance reduces the shunt loss of the input transmission line while small susceptance reduces the length of the input transmission line which in turn reduces the series loss. Unfortunately, in this technology, both the input conductance and susceptance of a bipolar transistor are very large. At 60GHz, for instance, the input conductance is nearly 25mS while the equivalent input capacitance is around 80fF. This will contribute significant high frequency loss on the input transmission line which eventually limits the bandwidth. One way to deal with this problem is to use resistive emitter degeneration. Both the conductance and susceptance decrease when the degeneration resistor value increases. In addition, adding degeneration resistor effectively prevents the thermal run-away problem of bipolar transistors. With the help of emitter degeneration, the 3-dB bandwidth of the DA can reach 80GHz. To further enhance the bandwidth, high frequency zeros can be added to the transfer function and it can be easily achieved by adding degeneration capacitors in parallel with degeneration resistors. However, it is important to

point out that adding large emitter degeneration capacitors results in negative input impedance and therefore can cause potential instability. To ensure a stable operation, a capacitor value of 20fF is chosen. Since the capacitance value is relatively small, it can be implemented by overlapping two bottom metal layers of the process.

3.8.3 Cascode based distributed power amplifier

One benefit of using cascode gain cell is that it has better reverse isolation than other structures. As a result, input and output t-lines can be designed independently. Since it is a power and efficiency oriented design, the bias voltage and current are determined by the desired output saturation power level as well as the optimum load impedance. Without using an impedance transformation network at the output, which is in nature a narrow band circuit, the optimum load impedance is 25Ω, and this is because the output current can travel in both forward and reverse directions and effectively the gain cell is loaded by two 50Ω resistors in parallel. Also, the number of stages is optimized for minimum transmission line loss. Both input and output transmission lines are implemented in microstrip structure.

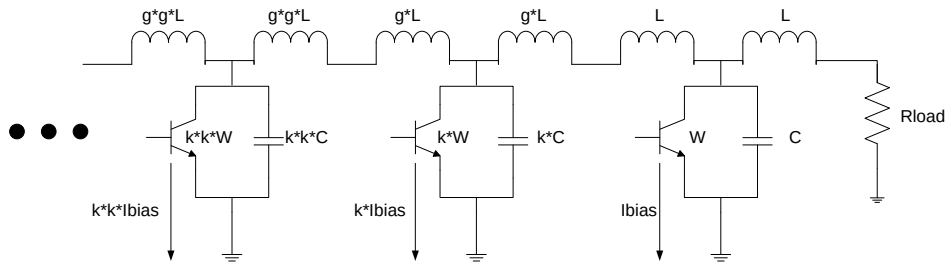


Figure 3.11: Lumped element model of a distributed amplifier.

To facilitate the design of the tapered transmission line, lumped elements are used to model the distributed amplifier, as shown in Fig. 3.11. Though it is a first order analysis, it provides sufficient insight into how to optimize the voltage and current scaling factor between gain stages. Assuming the output transmission line inductance scales by a factor of g from stage to stage and device size scales by a factor of k, the impedance seen by each stage can be expressed in terms of the load impedance. The power generated by each stage is equal to the current square multiplied by the impedance. Then the relation between power generated by one stage of a tapered DA and the power generated by one stage of the corresponding uniform DA (no scaling between stages) can be determined. In order to maintain the same output power as a uniform DA, k^3g this must be equal to one, which tells one how to simultaneously scale the output impedance and the device size.

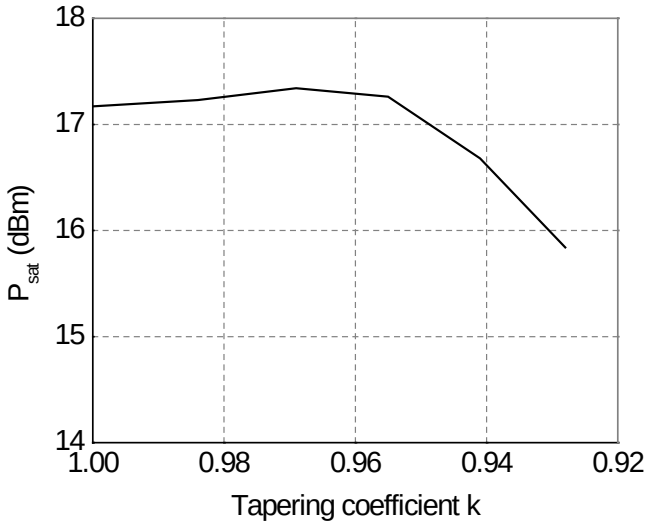


Figure 3.12: Output saturation power as a function of tapering coefficient k.

Fig. 3.12 and Fig. 3.13 plot the output saturation power and peak drain efficiency as a function of the tapering coefficient k. It can be seen that when k decreases, which means increased tapering between stages, the output power remains relatively constant but the efficiency increases as a result of reduced bias current. However, if k goes too small, the output power eventually starts to roll off and so does the efficiency. This is mainly because the mismatch between stages becomes significant and the output signal experiences large reflections when it travels along the output T-Line. The optimum tapering coefficient is around 0.955 for an 8-stage DA. In practice, it is hard to achieve very high loaded T-Line impedance, so the actual tapering coefficient is slighter greater the optimum value. The complete schematic is shown in Fig. 3.14.

3.8.4 Common-emitter based distributed power amplifier

In spite of better reverse isolation which makes design easier, cascode gain cells have lower efficiency since the voltage swing is limited by the knee voltage of the cascode device. In comparison, common-emitter gain cells usually provide better efficiency. However, common-emitters are notorious for poor reverse isolation, therefore circuit techniques are needed to overcome the problem. One way to deal with the poor stability is to use cross-coupled capacitors which can neutralize the differential pair. This also reduces the input capacitance since the Miller effect is canceled out to the first order. In addition to that, the output T-Lines are implemented in coupled strip lines without ground plane. This provides much higher odd mode characteristic impedance than standard microstrip lines and much better common-mode rejection since the only return current path for

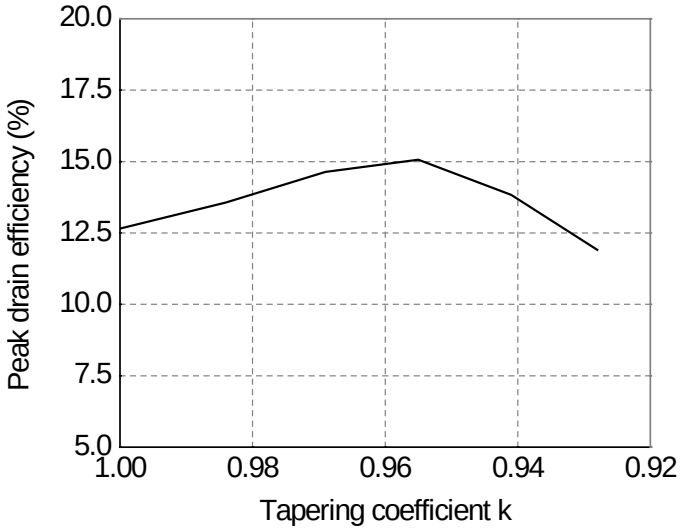


Figure 3.13: Peak drain efficiency as a function of tapering coefficient k.

even mode signals is the substrate. The scaling factors between stages are determined based on the same analysis presented in the previous subsection and the optimum tapering coefficient k is 0.94 for a 6-stage common-emitter DA. The complete schematic is shown in Fig. 3.15.

3.9 Simulated Performance

3.9.1 Cascode based distributed power amplifier performance

Fig. 3.16 shows the simulated S parameters of the cascode distributed power amplifier. The small signal gain is 10.3dB and the 3-dB bandwidth is 110GHz. The amplifier is unconditionally stable at any frequency. Fig. 3.17 and Fig. 3.18 plot the simulated large signal performance of the power amplifier. In order to make comparison, a uniform distributed power amplifier is also designed using the same technology, the performance of which is plotted in dashed curves. In terms of output saturation power, the tapered DA matches the uniform counterpart reasonably well at any given frequency, which proves that the tapering concept works effectively. The tapered DA has relatively higher output P1dB at high frequency, and it is mainly because smaller devices are used in the front and less distortion is introduced. In terms of drain efficiency, as a result of reduced total bias current, the tapered DA performs better than the uniform DA at any frequency. The peak drain efficiency is greater than 10% up to 90GHz.

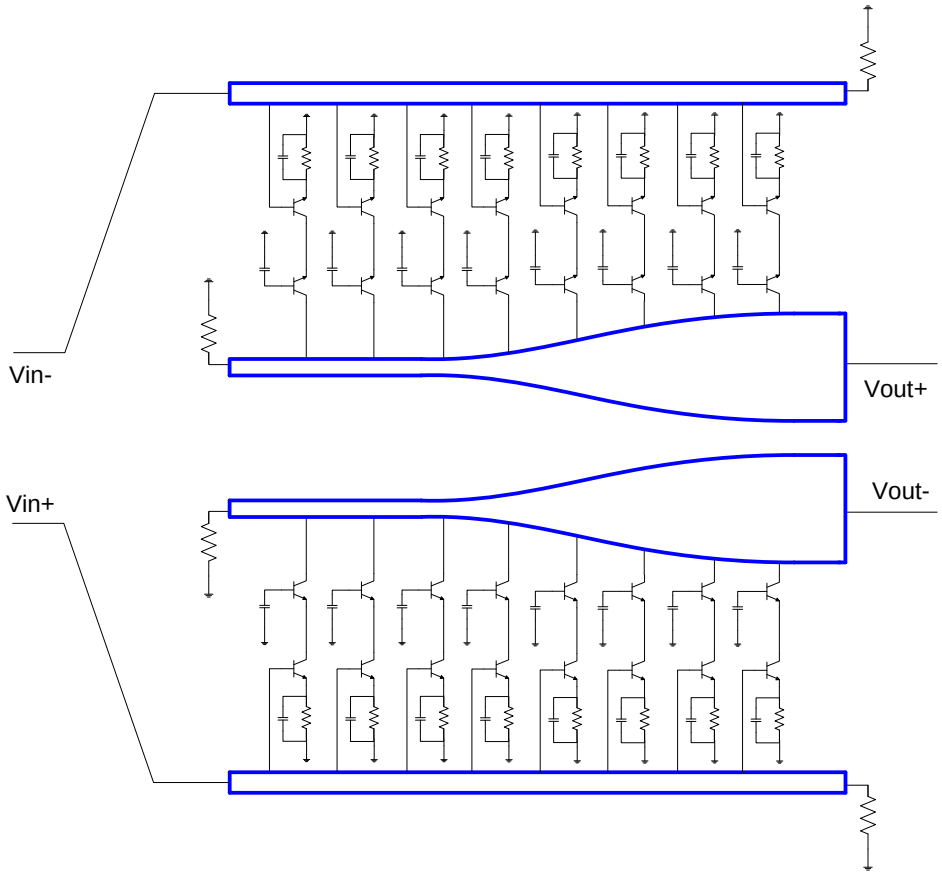


Figure 3.14: Schematic of the tapered cascode distributed power amplifier.

3.9.2 Common-emitter based distributed power amplifier performance

Fig. 3.19 shows the simulated S parameters of the common-emitter distributed power amplifier. The small signal gain is 7.4dB and the 3-dB bandwidth is 113GHz and the amplifier is unconditionally stable. Fig. 3.20 and Fig. 3.21 show the simulated large signal performance. Again, a uniform counterpart is designed for comparison and it can be clearly seen that in term of output saturation power, the two DAs are very close. But the peak efficiency of the tapered DA is better than the uniform one at any given frequency. The peak efficiency is greater than 17% up to 90GHz. Compared to cascode DAs, the efficiency of the common-emitter DA is significantly enhanced.

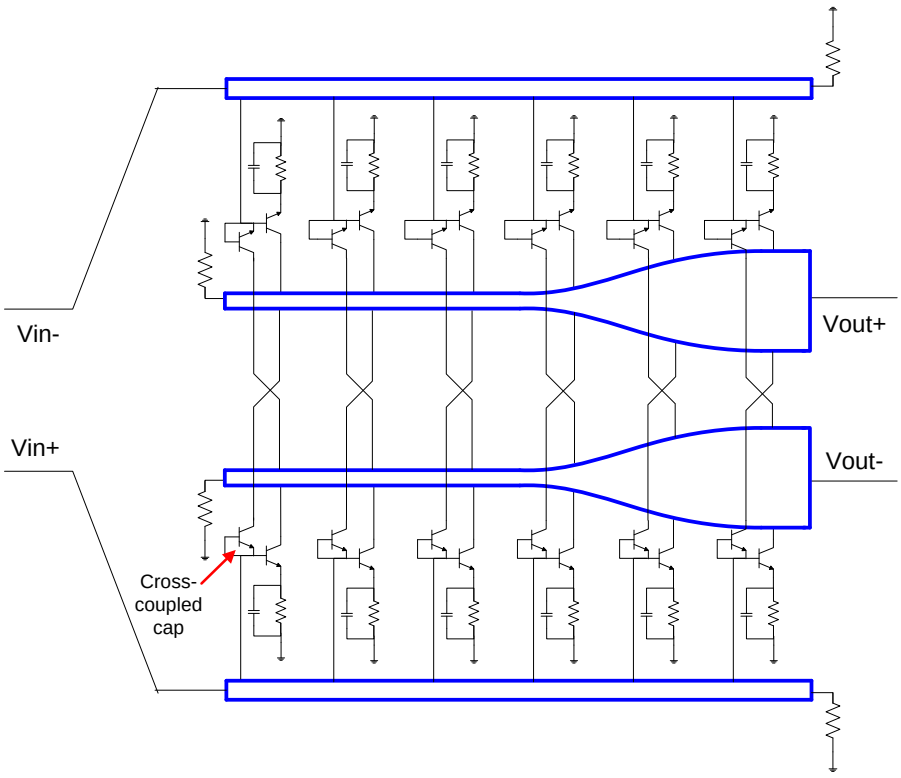


Figure 3.15: Schematic of the tapered common-emitter distributed power amplifier.

3.10 Conclusion

In conclusion, distributed power amplifiers with bandwidth in excess of 100-GHz have been demonstrated. These types of amplifiers pave the road to future integration of extremely wideband systems such as wafer-scale radios. Systematic approach for power and efficiency optimization has been analyzed. In particular, a device size and output T-Line impedance tapering technique has been proposed to enhance the power efficiency of distributed amplifiers, overcoming a major problem with traditional DAs. The concept is verified by simulating both cascode and common-emitter distributed power amplifiers in IBMs SiGe BiCMOS process.

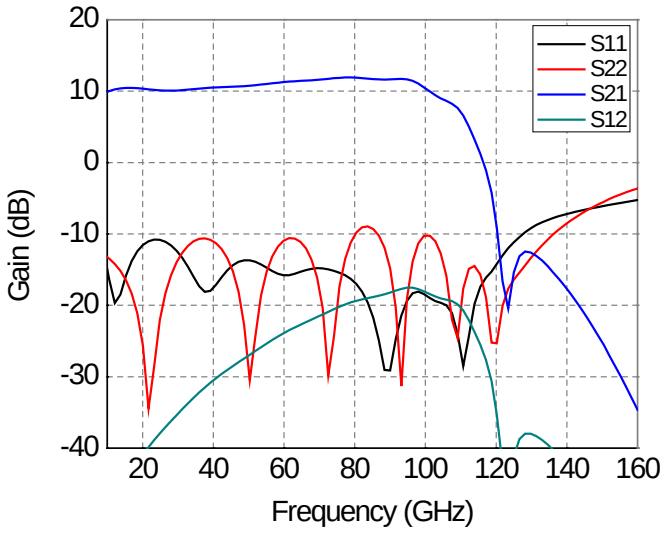


Figure 3.16: Simulated S parameter of the cascode distributed power amplifier.

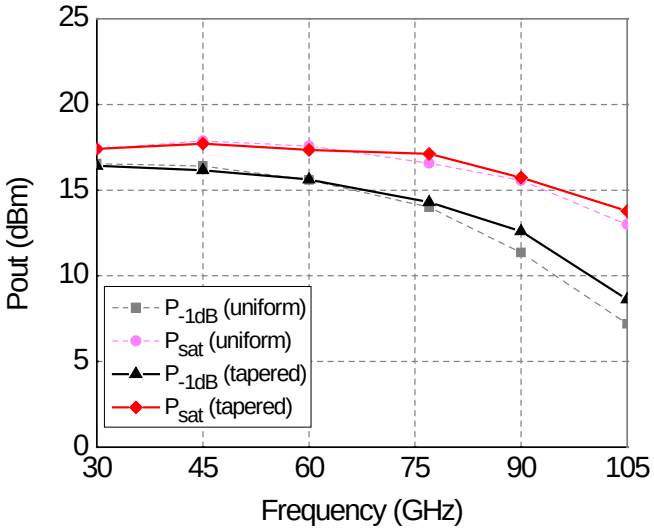


Figure 3.17: Simulated output power of cascode distributed power amplifiers as a function of frequency.

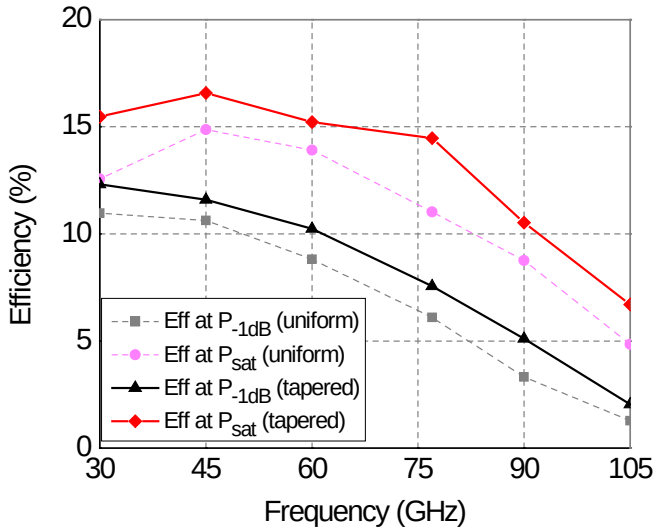


Figure 3.18: Simulated drain efficiency of cascode distributed power amplifiers as a function of frequency.

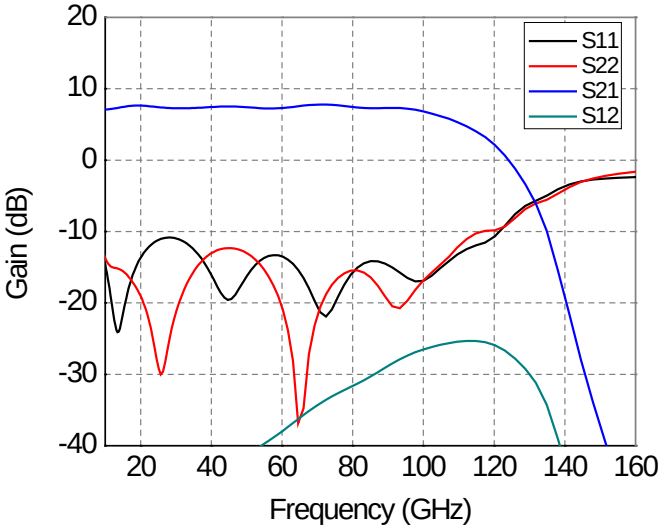


Figure 3.19: Simulated S parameter of the common-emitter distributed power amplifier.

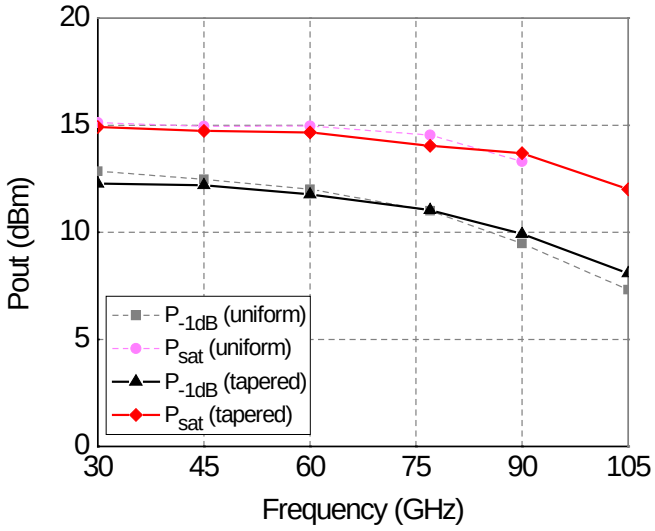


Figure 3.20: Simulated output power of common-emitter distributed power amplifiers as a function of frequency.

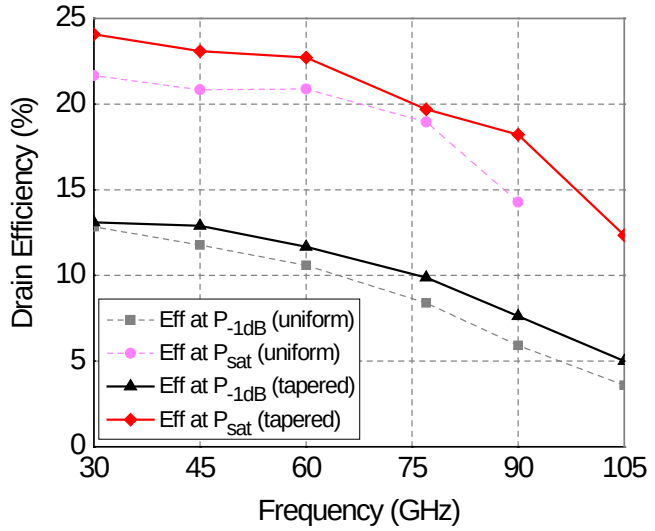


Figure 3.21: Simulated drain efficiency of common-emitter distributed power amplifiers as a function of frequency.

Chapter 4

On-Chip Antenna and Phased-Array Performance

4.1 Introduction

Integrated antennas on silicon is a promising technology at millimeter-wave frequencies. The size of antenna unit element could be designed comparable to traditional bond-wire pads. Therefore integrated antenna could be a cost-effective solution compared to conventional packaging. Moreover, packaging the antenna with transceivers causes large insertion-loss at millimeter-wave range. Moreover, fully integrated system in a single chip provides extra design flexibilities by co-designing antenna with transceivers to achieve the broader space coverage, wide bandwidgh, and better beam shaping characteristics.

The main challenge of the integrated antennas on the silicon substrate is the low radiation efficiency which is generally less than 10% [38]. There are two main reasons for such a low radiation efficiency. One is the conduction loss owing to low resistivity of the silicon substrate, and another is the surface wave mode excitation caused by the thick silicon substrate with a high permittivity [31]. In order to mitigate the effect of low resistivity, the high resistive (HR) substrate has been investigated [11][28]. The SOI substrate is a good example of the HR substrate which makes it possible to achieve a fully integrated transceiver with on-chip antenna having relatively high radiation efficiency. However, SOI substrate is not a cost effective choice for on-chip antenna. One interesting way to get a HR substrate using a low resistive silicon substrate is the proton implantation method using cyclotron ion source [13]. However, this method could cause damage to the semiconductor active layers. Another approach to achieve a high radiation efficiency is to use MEMS technology. Using MEMS, a lossy silicon substrate is substituted with a dielectric membrane to mitigate both conductive loss and surface wave excitation [2]. Another exotic example using MEMS technology is a patch antenna with air substrate which achieves radiation efficiency of 94% [30]. However, MEMS technology requires various extra processes other than a conventional CMOS technology, which prohibits full integration on chip with CMOS ICs. Moreover, the

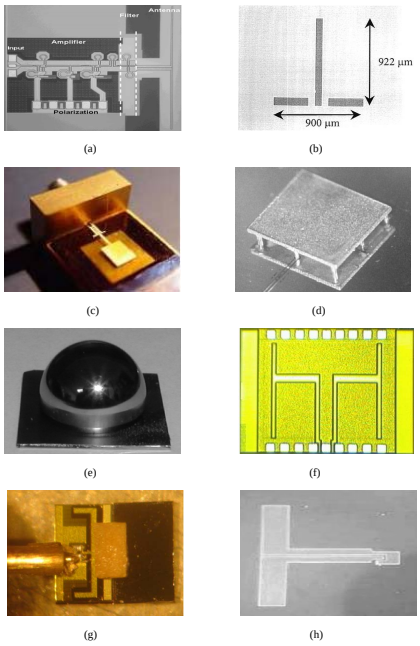


Figure 4.1: Various approach to improve radiation efficiency-HR substrate (a-[31], b-[13]), MEMS technology (c-[2], d-[30]), substrate/superstrate dielectric lens (e-[39], f-[4]), Dielectric resonator (g-[12]), (h) [38] is a typical example of the native antennas on a lossy Si substrate-differential dipole antenna which has radiation efficiency around 10%.

structural robustness is usually sacrificed with the improved radiation efficiency.

In terms of surface wave excitation, the substrate dielectric lens is widely used in millimeter-wave applications where the substrate is thick enough to cause multimode surface wave excitation [36, 9, 39]. The dielectric lens confines the electric field which prohibits surface wave excitation. Moreover the dielectric lens can improve the antenna gain by confining the field to a certain dircection.The main disadvantage of using a dielectric lens is large lens size as well as a process difficulty in dielectric lens fabrication. In order to overcome the fabrication difficulty for the substrate lens, a superstrate dielectric layers over the on-chip antenna has been introduced [4]. Another type of antenna which avoids the surface wave excitation is a dielectric resonator antenna [12]. For this type of antenna, most of the electric fields are confined within a high dielectric resonator that the conduction loss in a lossy silicon substrate is mitigated. However, this type of antenna has critical limitation in terms of fabrication which requires an extra-process and a precise alignment of the dielectric resonator.

Table 4.1 summarizes several different approaches to improve the radiation efficiency (see also Fig. 4.1). From the table, we conclude that two main factors for low radiation efficiency could be mitigated by applying the wafer-thinning technology. This approach takes advantage of the

Table 4.1: Reported design approaches for the radiation efficiency improvement.

		Advantages	Disadvantages
Native Antennas		- Low cost / Ease of fabrication - Best for full integration	- Low radiation efficiency - Limited antenna gain - Main-lobe distortion due to substrate modes in mmW range
Modified Antennas	**Technology**		
	HR Si substrate	- Very good for full integration - Moderate radiation efficiency	-Needs SOI substrate - Mainlobe distortion (needs thinning) - Possible damage on IC for the proton implantation method
	MEMS	- High radiation efficiency	- Bad structural robustness - Bad for full integration
	Substrate Dielectric Lens	- Widely used in mmW range - Applicable to various types of antennas - High directivity - High radiation efficiency	- High cost - Large lens size - Difficult in fabrication - Extra process
	Super-strate Dielectric Lens	- Moderate radiation efficiency - Good for full integration	- Extra process - Not yet widely used
	Dielectric Resonator Antenna	- High radiation efficiency - Good for full integration	- Extra process - Not yet widely used

native antenna design which is cost effective and easy for integration with transceivers in a single chip. Moreover, wafer-thinning is required for an advanced packaging technology to improve the insertion loss owing to wire-bonding.

4.2 Antenna Design Considerations

4.2.1 Antenna Radiation Efficiency

The input impedance of an antenna is computed from $Z_{in} = V/I(0) = R_A + jX_A$, where $I(0)$ is the current value at the input terminals. Antenna radiation efficiency is given by

$$\eta = \frac{R_{rr}}{R_A} \tag{4.1}$$

where $R_A = R_r + R_\Omega$ is the real part of antenna input impedance. When the substrate thickness is thin enough, the resonant input resistance is approximated as radiation resonant resistance R_{rr}. However, when the surface wave is excited, the surface wave resonant resistance R_{rs} has to be considered, i.e., $R_r = R_{rs} + R_{rr}$ since R_{rs}, and R_{rr} are directly proportional to the power coupled into the substrate as guided modes and to the power radiated in space, respectively [5]. Therefore, antenna radiation efficiency can be expressed as follows:

$$\eta = \frac{R_{rr}}{R_{rr} + R_{rs} + R_\Omega} \tag{4.2}$$

In order to improve radiation efficiency, both R_Ω and R_{rs} has to be minimized. Based on this expression, HR Si substrate technology is categorized to the group of reducing R_Ω, and substrate/superstrate dielectric lens technology belongs to the group of reducing R_{rs}. The MEMS technology and wafer-thinning technology reduces both R_Ω and R_{rs}. Therefore wafer-thinning approach could be a good candidate for the on-chip antenna on a lossy silicon substrate without modifying antenna structures.

4.2.2 Wafer-Thinning Technology

Wafer thinning for advanced packaging methods has gained importance as demand has increased for memory cards, portable computing systems, multiple chip packages (MCPs), and other applications that require thin integrated circuits (ICs) [37]. Although, it is a significant challenge in thinning wafers to final thicknesses less than 100 μm, semiconductor packaging groups have developed several thinning technology to achieve thin-wafer less than 100 μm with high reliability. Table 4.2 lists widely used wafer-thinning technology to relieve stress which causes die-cracking and die-breaking. In terms of antenna performance, radiation efficiency can be effectively improved by thinning the lossy and high dielectric substrate. Moreover, an antenna with thinned wafer has less radiation pattern distortion as surface wave excitation is minimized. Surface waves causes distortion (ripple) in the radiation pattern.

Table 4.2: Summary of the widely used wafer-thinning technology.

Technology	Description	Advantage	Disadvantage
Mechanical Grinding	-Two-step process : Coarse grinding (5um/sec) Fine grinding (<1um/sec)	-Fastest processing -Low cost	-Largest damage -Defect structures remains -Rough thickness tolerance -Cannot applicable for very thin wafer thinning
Chemical Mechanical Polishing (CMP)	-Polishing based on buffered silica slurries (a few um/min.)	-Very flat surfaces -Total thickness variation (TTV) is low	-Slowest thinning rate - Applicable for 200um or more thickness depending on the wafer size
Wet-Etching (Spin-etching)	-Front surface has to be protected -Mixture of HF and HNO3	-Very flat surfaces comparable to CMP -Mass production	-MCL(minority carrier lifetime) is much higher than dry-etching, comparable to CMP
Atmospheric Downstream Plasma-Dry Chemical Etching (ADP-DCE)	-Uses Ar/CF4 Plasma (20 um/min) -Uniformity <2% for 20um etching	-Faster than CMP -Lesser damage than mechanical grinding	-Electrically active defects form near the plasma-etched surface

4.2.3 Antenna Unit Element Design

In order to integrate an antenna on a silicon chip, a good radiation efficiency is one of the critical issues. Another issue is the size of the radiator. Considering silicon integrated circuit process, planar antennas such as patch, dipole or monopole, and slot could be good candidates as a radiating element. Among them, widely used circular disk monopole antenna and two types of slot antennas (folded slot/bow-tie slot) will be discussed in detail.

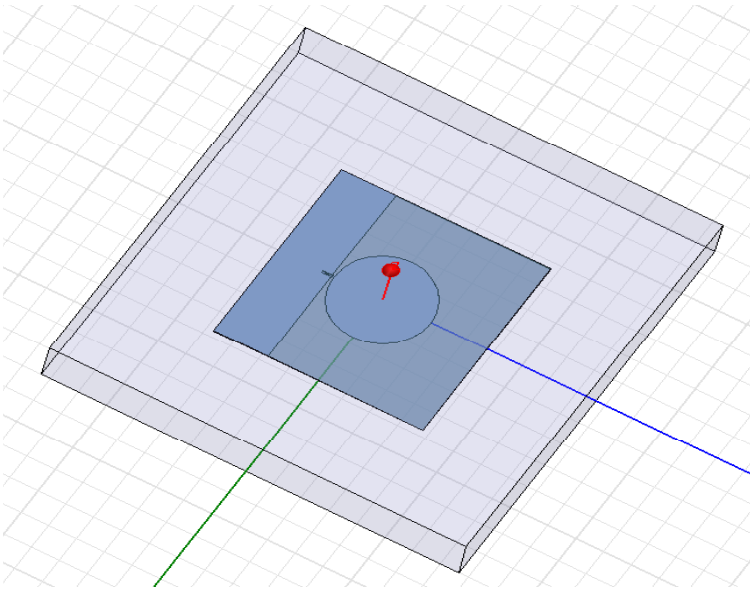


Figure 4.2: Designed circular disk monopole Antenna (diameter=700 μm).

Circular Disk Monopole Antenna

The circular disk monopole antenna is widely used for ultra-wideband application owing to its low group delay variation with wide bandwidth [24]. It has been demonstrated that the optimal design of this type of antenna can achieve an ultra wide bandwidth with satisfactory radiation properties. The performance of this type of antenna and its characteristics in the frequency domain is mostly dependent on the feed gap, the width of the ground plane and the dimension of the disc. The first resonant frequency is directly associated with the dimension of the circular disc because the current is mainly distributed along the edge of the disc [24]. A circular monopole antenna was designed using a 0.13 μm BiCMOS process which has oxide thickness of 12 μm. The top metal layer with thickness of 3 μm (Metal 6) was used for antenna structure. The width of the ground plane was chosen to be 400 μm. Fig. 4.2 shows designed circular monopole antenna with diameter of 700 μm.

Folded Slot Antenna

In terms of array, a slot antenna is more practical than a printed dipole antenna because slots couple to the dominant TM_0 mode along the perpendicular direction to their axis (broadside), which can alleviate the complexity of the required feed network. Moreover, many slots could be integrated in the same ground plane, therefore, it can be easily integrated with amplifiers with coplanar waveguide (CPW) transmission-lines. The folded slot antenna is another popular type of antenna which can be easily fed by CPW. Usually, the circumference of the folded-slot is designed to be approximately equal to one guided wavelength (λ_g) [42]. It can be fed with a CPW allowing for easy integration of three-terminal devices or MMICs for microwave amplification and reception. Basically, folded antenna has the characteristic of a broad bandwidth of frequency and a radiation pattern with maximum radiation at the broadside. From Babinets principle [1], the input impedance of complementary antennas can be calculated by

$$Z_{slot} = \frac{376.7^2}{4Z_{dipole}} \approx 500\Omega \tag{4.3}$$

In order to reduce the Z_{slot} and to realize an appropriate matching network, several stubs can be included in the slot, the complement of the N-element dipole antenna ($Z_{in,N} = N^2 Z_{dipole}$). Therefore, an N-element slot antenna input impedance can be reduced by,

$$Z_{in,N} = \frac{Z_{slot}}{N^2} \tag{4.4}$$

Therefore when $N = 2$, around 100Ω of the input impedance could be realized. Further reduction of the antenna impedance is performed by controlling the stub size. Fig. 4.3 shows a designed folded slot antenna at 94 GHz. In terms of the antenna bandwidth, the Chu-Harrington and McLean Limits relate quality factor Q (inverse fractional bandwidth) of an ideal, perfectly efficient antenna to its size denoted by the radius $r_{\lambda,C}$ of the boundary sphere as follows [27]:

$$Q = \frac{1 + 2(2\pi r_{\lambda,C})^2}{(2\pi r_{\lambda,C})^3 (1 + (2\pi r_{\lambda,C})^2)} \approx \frac{f_c}{BW} \tag{4.5}$$

where $f_c = \sqrt{f_L f_H}$, and boundary sphere radius in units of wavelengths at the center frequency, $r_{\lambda,C}$.

Therefore, by increasing the size of a slot, the physical aperture of the antenna is increased which results in increase of antenna bandwidth. The folded slot could also achieve a wideband characteristic if the slot becomes wide in shape. Fig. 4.4 shows designed "fat" folded slot antenna. Comparing previously designed slot size which was 700 μm $\times$ 30 μm, the slot shape widens to 515 μm $\times$ 515 μm to increase the aperture of the antenna. From HFSS simulation, its bandwidth increased from 10% to 50%. Therefore bandwidth has to be compromised with area consumption of the substrate. In order to achieve wider band matching characteristic, an internal stub at the center of the rectangular slot can be tuned [18].

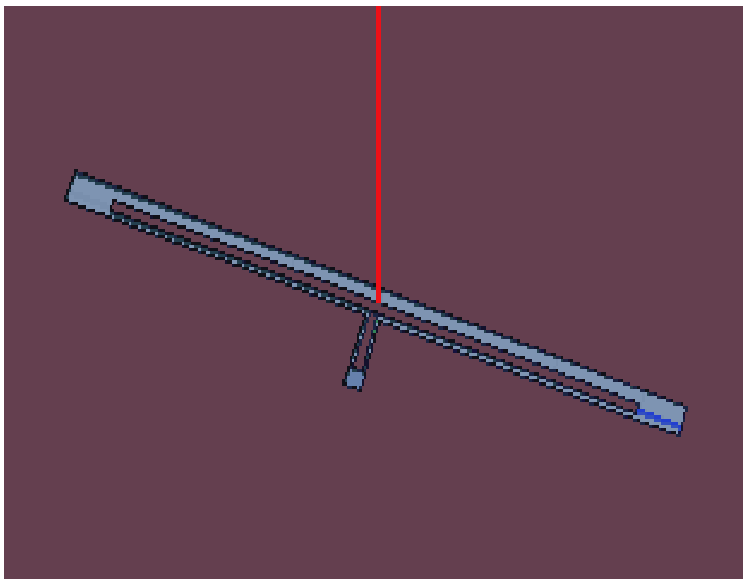


Figure 4.3: Designed folded slot antenna (slot size : 700 μm $\times$ 30 μm).

Bow-tie Slot Antenna

The bow-tie slot antenna is a dual of the bow-tie dipole antenna which is in the category of the frequency independent antenna. Therefore this antenna can have wide input matching characteristic. Because the slot is formed in the ground plane, this antenna could be easily fed by CPW. Fig. 4.5 shows the designed bowtie slot antenna with stub [17]. The design was based on the a generic 90nm CMOS process with an oxide layer thickness of 6 μm. The top metal layer (Metal 7) was used as the antenna element. The stub was used to increase the bandwidth, which lowers the antenna input impedance to be around 50Ω.

4.2.4 Antenna Array Considerations

When a multiple number of antennas are used as an array, each unit element is affected by mutual coupling of other antenna elements surrounding it. The mutual coupling distorts the radiation pattern of unit elements, and changes the input impedance characteristics. Moreover, it causes blind-spots for the phased array system [32]. Therefore mutual coupling effects must to be considered early in the design process. Array pattern ($F(\theta,\phi)$) approximation under infinite array condition is as follows:

$$F(\theta,\phi) = g_{ae}(\theta,\phi) f(\theta,\phi) \tag{4.6}$$

where $g_{ae}(\theta,\phi)$ is an average active-element pattern, and $f(\theta,\phi)$ is array factor caused by the antenna array. Fig. 4.6 shows 3$\times$3 arrays of folded slot antennas. From the simulation of HFSS,

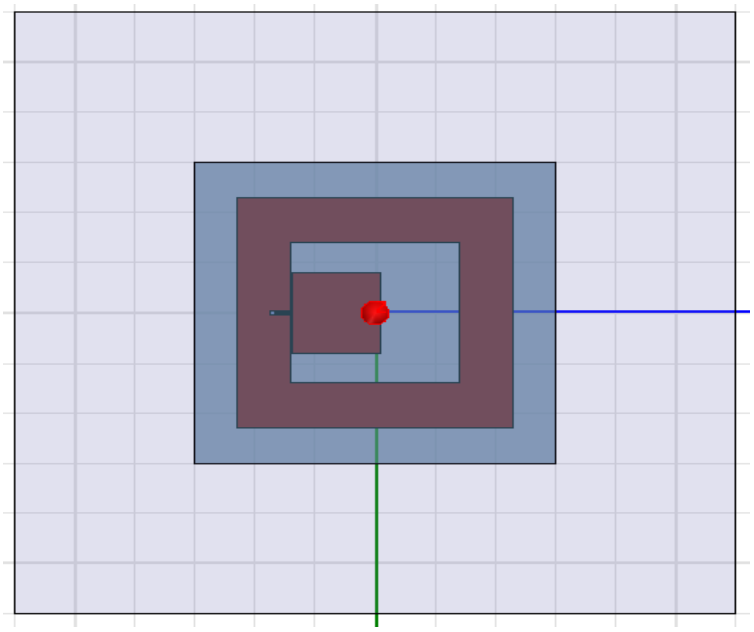


Figure 4.4: Fat folded slot antenna (slot size : 515 μm $\times$ 515 μm).

the radiation pattern was slightly changed from that of single folded slot antenna. This simulation setup considers every effect in terms of EM phenomenon except edge effects. However, it is impractical to run a 3-D EM simulation for a large array considering simulation time consumption.

When the number of unit element becomes large enough, unit elements couple to each other, and the effective radiation patterns become symmetric. For a large enough array, we can assume the infinite array approximation. The infinite array environment was modeled using Master-Slave boundary setup in HFSS with PML (Perfect Matching Layer) boundary condition [7]. The Master-Slave setup replicates the fields considering only the phase difference with the assumption of the infinite array environment. In wafer scale radio applications, more than hundreds of antenna elements would be integrated in a single wafer. Therefore, it is reasonable to apply the infinite array approximation. However, infinite array approximation does not reflect edge effects which will affect on the input impedance of the antennas around edge as well as antenna side-lobes.

In a phased array, the input impedance (scan impedance) changes with respect to array scan angle. When the scan reflection coefficient (Γ_s) has magnitude of unity for a certain scan angle, the angle is called a "blind spot". At the blind spot, the scan element pattern is zero, and all input power is converted to surface wave power at the blind spot. This blind spot cannot be found from either an average active-element pattern or array factor from the array. Therefore, it has to be characterized using a proper EM simulation or in an analytical way. In order to include this effect, the average active element pattern can be expressed as follows [33]

$$g_{ae}(\theta_0, \phi_0) \approx g_i(\theta_0, \phi_0)(1 - |\Gamma_s(\theta_0, \phi_0)|^2) \tag{4.7}$$

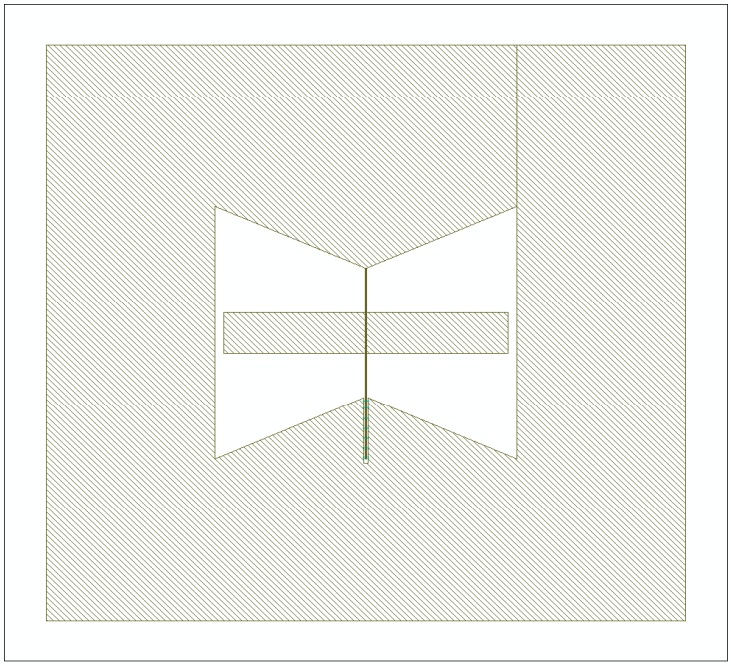


Figure 4.5: Designed bow-tie slot antenna.

where $|\Gamma_s(\theta_0, \phi_0)| = 1$ when the scan angle (θ_0) is at the blindness angle. It is known that the active element pattern for the infinite planar printed dipole array has the scan blindness at an angle of about 46 degrees. Another consideration in a large array is that there is the lack of surface wave loss for the infinite array except at the blindness angle [32]. Therefore, for a large antenna array, ohmic loss becomes important factor. From this aspect, the wafer-thinning technique which physically eliminates large portions of the lossy silicon substrate is an appropriate approach in improving radiation efficiency. However, further investigation is required as to how surface wave effects would change the blindness spot characteristics. For a small antenna array, the antenna mutual coupling affects on input impedance of each element as well as radiation pattern is captured. Therefore, antenna elements around the edge of the wafer have to be designed separately.

4.3 Simulation Results and Discussions

In 3-D electromagnetic simulation, HFSS ver10/ver11were used for the simulations given this section. The silicon dielectric loss tangent was not considered, and the conductivity of the silicon substrate was set to be 10 mS/m.

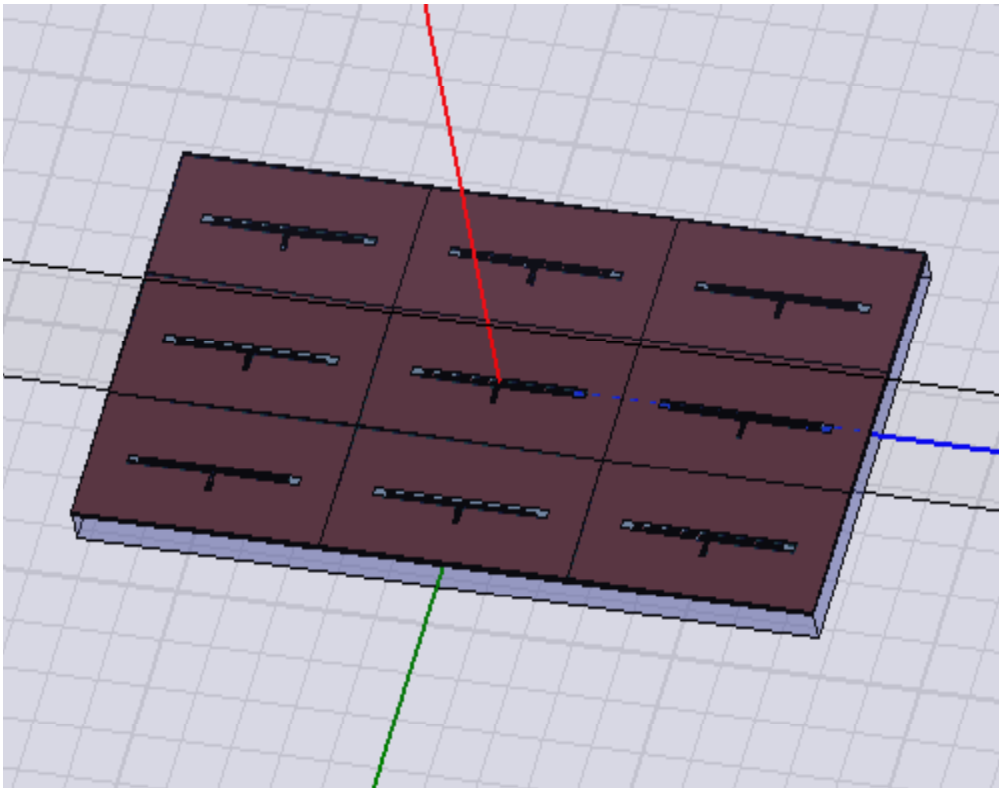


Figure 4.6: 3×3 array folded slot antenna.

4.3.1 Circular monopole antenna

A circular monopole antenna was designed based on a 0.13 μm BiCMOS process (ST-Microelectronics B9MW) which has oxide thickness of 12 μm. The top metal layer with thickness of 3 μm (Metal 6) was used for antenna structure. Fig. 4.7 shows the simulated radiation pattern of HFSS v10 for the designed circular mono-pole antenna. When the substrate is thinned up to 50 μm, the radiation efficiency was 75.4%. For 100 μm of substrate thickness, the radiation efficiency was 63%. The radiation pattern has typical omni-directional pattern which has an isotropic pattern in the H-plane. The peak directivity of this antenna was 2.24 dBi. The designed circular disk monopole antenna provides more than 40 GHz bandwidth for 50Ω input matching. The diameter of the disk was 700 μm. One of main drawbacks of this antenna is the effect of ground plane underneath of the disk. This ground plane affects not only on the single antenna elements, but also makes it difficult to employ it as a unit element for an antenna array. This occurs since a unit antenna is affected by the ground plane images around it, caused by other neighboring elements.

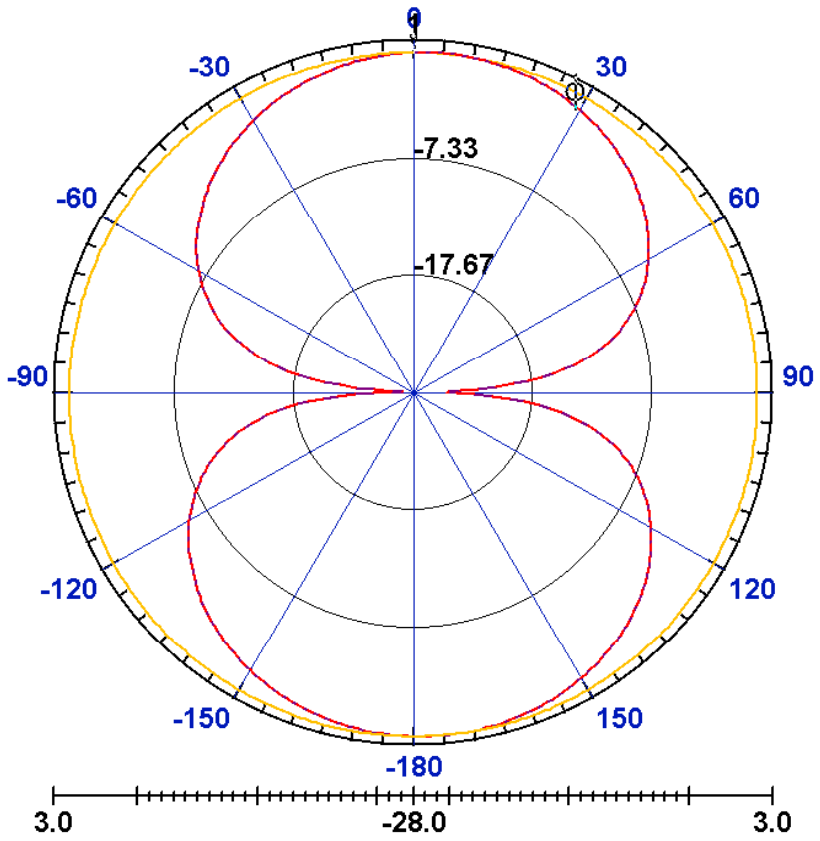


Figure 4.7: The radiation pattern of the circular disk monopole antenna.

4.3.2 Simulation Results of Folded Slot Antenna

For the designed folded slot antenna, the slot length is 700 μm and its width is 30 μm. The structure is shown in Fig. 4.3. The length of the stub is 600 μm with the width of 10 μm, the gap in the bottom is 6.6 μm and the gap on top is 13.4 μm. Fig. 4.8 shows the radiation pattern of the designed antenna. The radiation pattern has the broadside characteristic and the maximum directivity is 4.98 dBi in the direction of the substrate and 3.43 dBi toward air. Fig. 4.9 presents radiation efficiency depending on the substrate thickness.

The radiation efficiency is 57% when substrate thickness is set at 100 μm. When h¿200 μm, there was steep efficiency degradation which is due to the surface wave mode excitation. Fig. 4.9 presents the radiation efficiency of the folded slot antenna depending on substrate thickness.

As shown in the Fig. 4.10, the designed folded slot antenna has a bandwidth of 10%. As described in section 4.2.3, the slot size should be increased to achieve wider frequency bandwidth. The designed folded slot antenna provides more than 50 GHz bandwidth when the input impedance

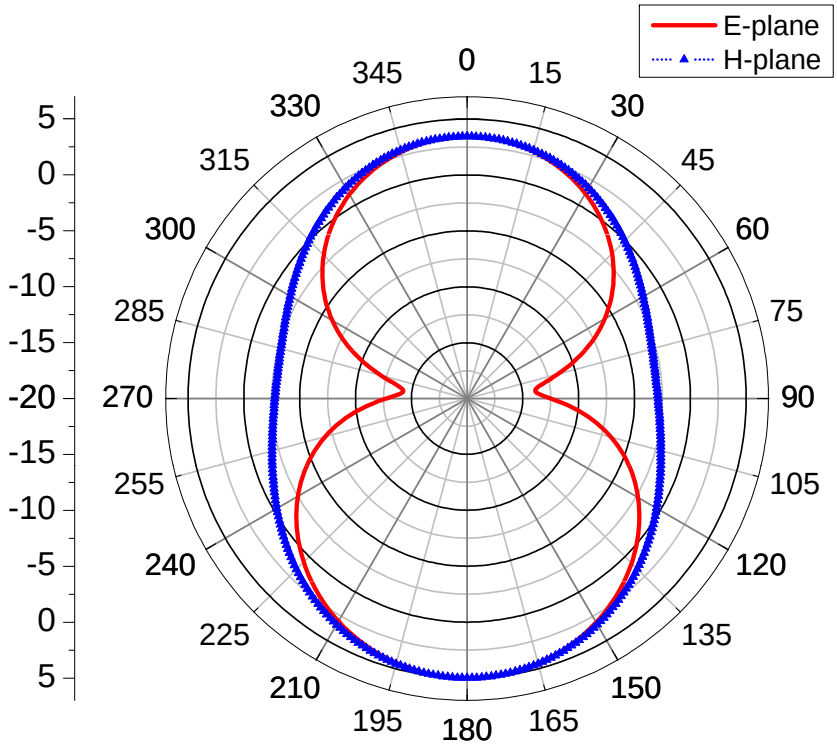


Figure 4.8: The radiation pattern of the circular disk monopole antenna.

of the signal source is 40Ω from the Fig. 4.11. As shown in the Fig. 4.4, the designed "fat" version of the folded slot antenna has larger aperture. When substrate thickness is 100 μm, the simulated radiation efficiency for the fat version is 65.1%, which is slightly better than that of circular disk monopole antenna. The three dimensional radiation pattern is shown in Fig. 4.12. The peak directivity of this antenna is 3.51 dBi.

4.3.3 Simulation Results of Bow-tie Slot Antenna

A bow-tie slot antenna was designed based on a 90nm CMOS process which has oxide layer thickness of 6 μm. The top metal layer of 0.9 μm (Metal 7) was used as the antenna element. Fig. 4.13 shows radiation pattern of the designed bow-tie slot antenna. Owing to wafer-thinning, the radiation pattern was almost symmetric for top and bottom. Considering 50 Ω input impedance of the radiating source, the center frequency is 90 GHz and bandwidth is 17 GHz. The radiation efficiency is 52% when substrate thickness is 50 μm, and 46% when the thickness is 100 μm. The peak directivity of the antenna was 5.4dBi.

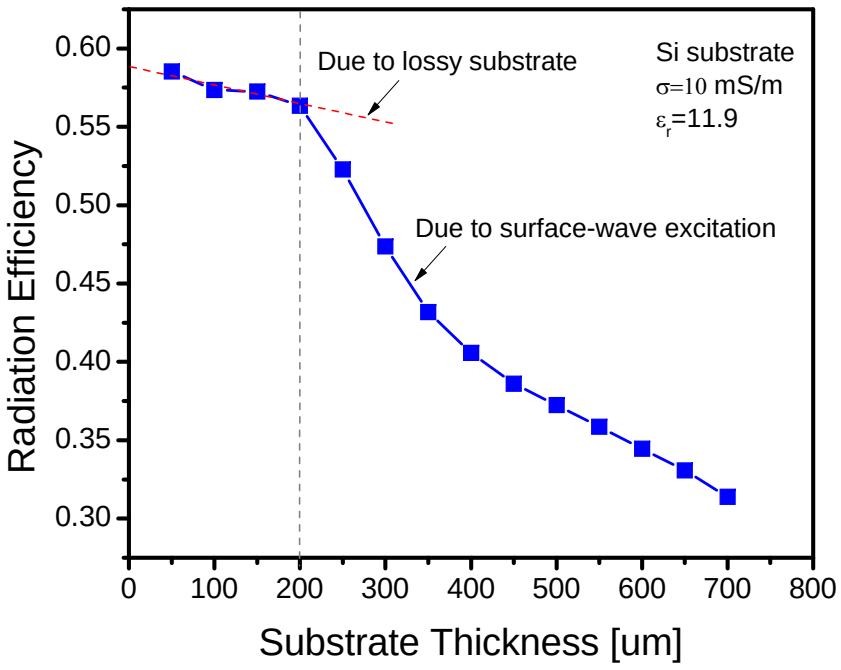


Figure 4.9: The radiation efficiency of the folded slot antenna depending on substrate thickness.

4.3.4 Simulation Results for Antenna Array

In order to simulate a large array, the infinite array environment was modeled using Master-Slave boundary setup in HFSS with PML (Perfect Matching Layer) boundary condition. Fig. 4.14 shows the boundary condition set-up for active element. From the simulation, directivity difference for the active antenna element between two cases was around 1 dB, and the main difference occurs around the null. Therefore, the infinite array approximation is an effective tool to examine large arrays despite the error in neglecting the substrate edge-effect. Fig. 4.15 compares the radiation pattern of the folded slot antenna extracted from HFSS. The left radiation pattern is for a single antenna element among the 3×3 array, the right pattern is extracted using Master-Slave setup in HFSS.

The same simulation setup was applied to the bow-tie slot antenna case with substrate thickness of 100 μm which is shown in Fig. 4.16. The radiation pattern of the active element is more symmetric. The Master-Slave setup replicates the fields considering only the phase difference with the assumption of the infinite array environment. As shown in the Fig. 4.17, this array has 20 dB of antenna gain, and the side-lobe level (SLL) was 13.4 dB, which follows from a uniform array with uniform power distribution setup. The radiation efficiency was degraded for all the three antenna arrays to around 40%.

Using the Master-Slave setup, the phase difference between each boundary was used as a scan

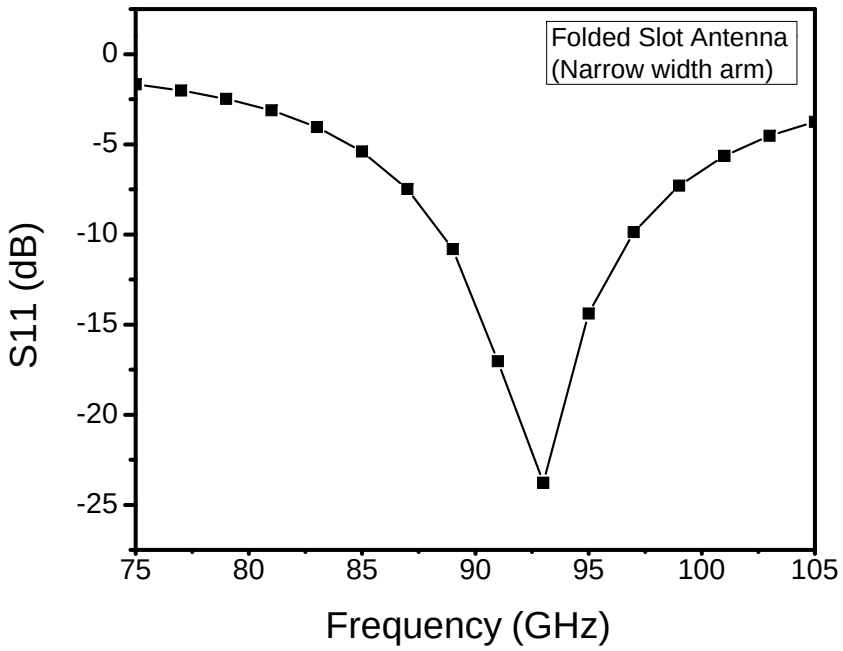


Figure 4.10: S11 of the folded slot antenna.

angle input. Impedance layers with 377 Ω were also inserted on top of each PML layers. To simulate the vector property of the scan angle, the impedance was scaled depending on θ_0, given by $376.7\cos(\theta_0)$. Fig. 4.18 shows S11 of the infinite folded slot antenna array. Because it is a complementary of the printed dipole antenna, the blindness spot is expected to be at the same place. The resulting blindness spot was around 46 degrees which is similar to the theoretical value.

4.4 Conclusion

We found that the wafer-thinning technology is an attractive way of improving the antenna radiation efficiency considering cost and the realizability of an antenna array structure on the wafer. We designed three different types of antennas to investigate the feasibility of the fully integrated antenna on a lossy silicon wafer. The radiation efficiency could be more than 70% when the silicon substrate is thinned up to 50 μm for the designed antenna structures. In order to design the antenna array, we considered mutual coupling under the infinite array approximation. The blindness spot for the infinite array using HFSS Master/Slave boundary condition corresponded well with the reported analytical result [32]. The simulation results showed that the fully integrated antenna could be possible on a silicon wafer with radiation efficiency more than 40% for 100 μm of substrate thickness. Considering 3-4 dB insertion loss caused by off-chip wire-bonding and elec-

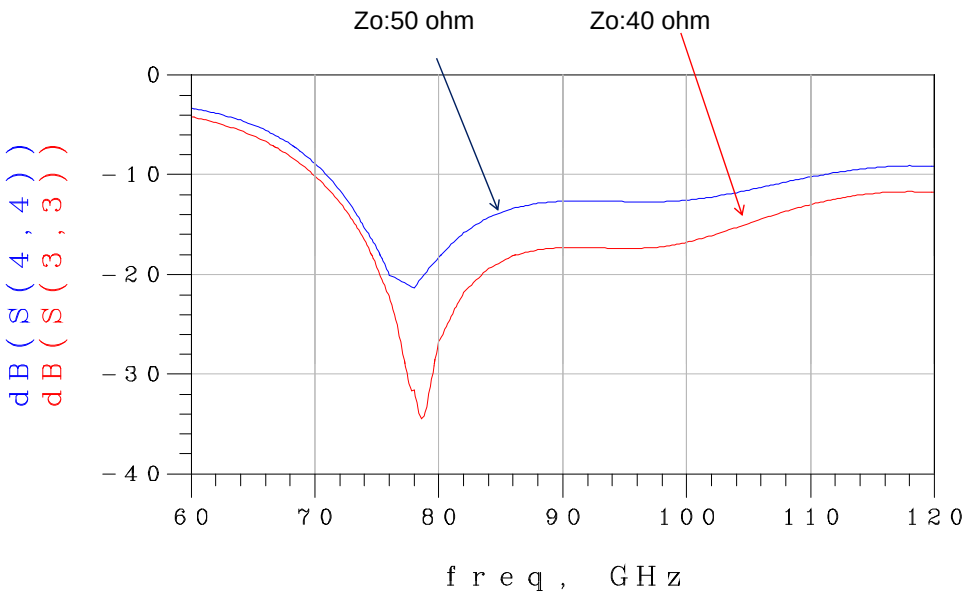


Figure 4.11: S11 of the “fat” folded slot antenna.

trostatic discharge (ESD) protection circuits, on-wafer integrated antenna would be advantageous and obviate the need for mm-wave packaging and testing.

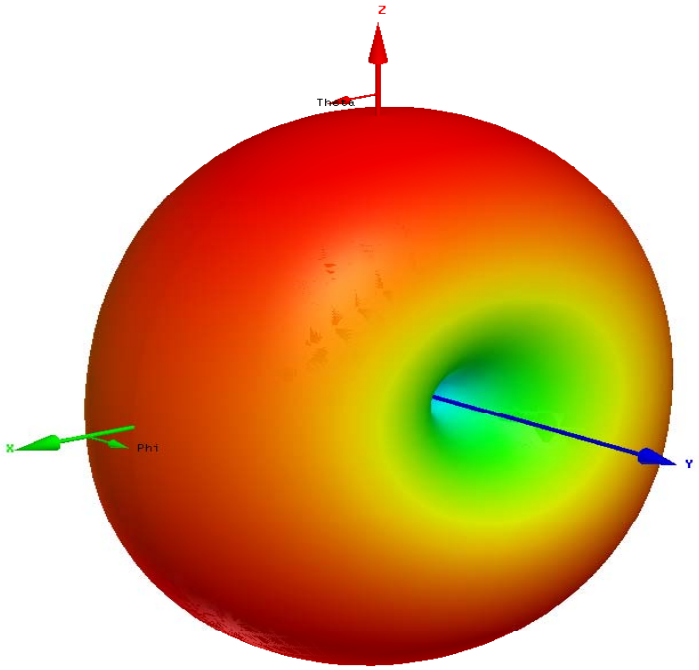


Figure 4.12: 3-D Radiation pattern of the “fat” folded slot antenna.

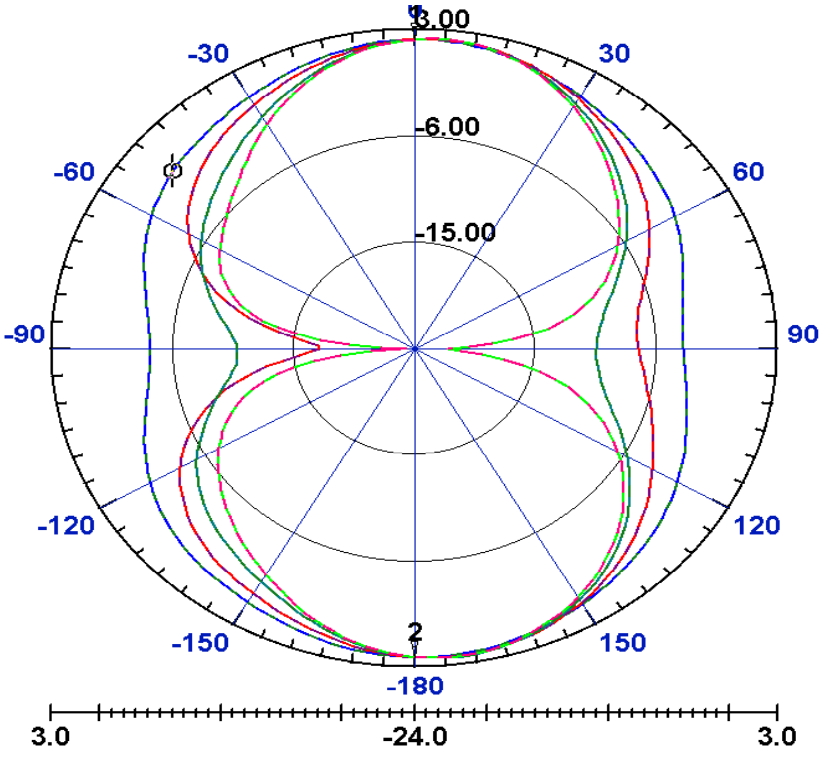


Figure 4.13: 3-D Radiation pattern of the “fat” folded slot antenna.

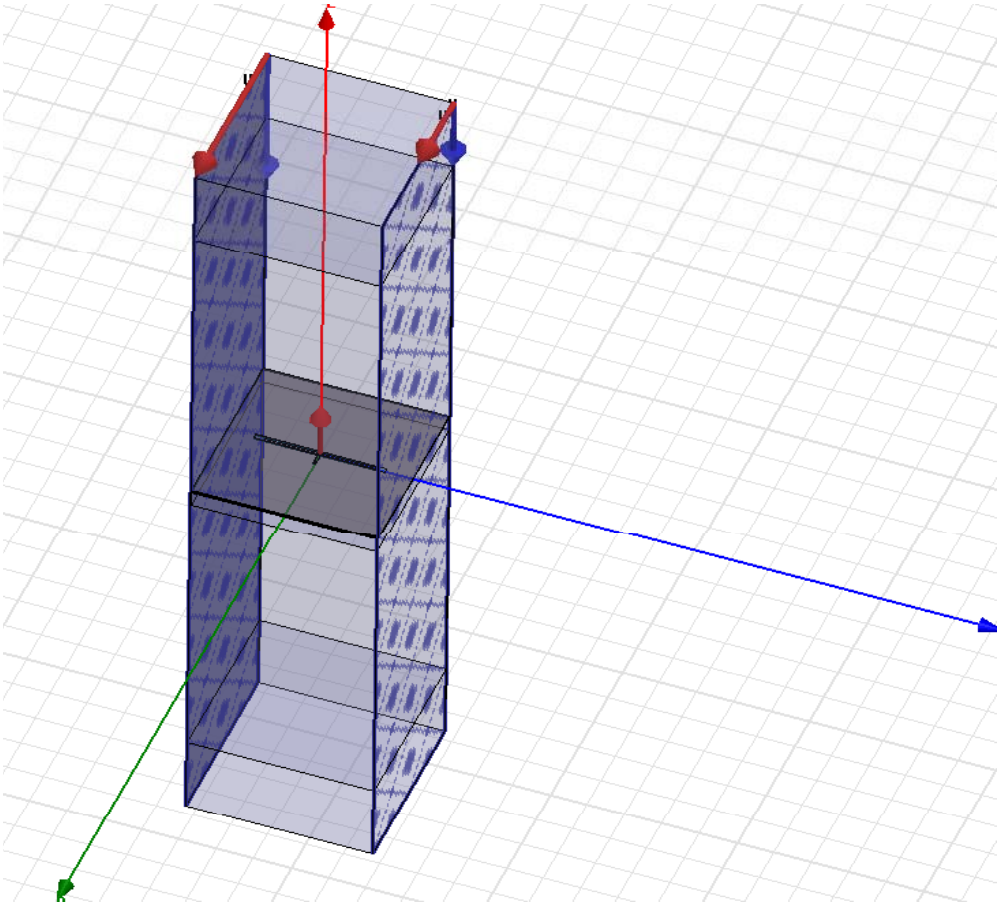


Figure 4.14: Master/Slave boundary setup for an active element in infinite array condition.

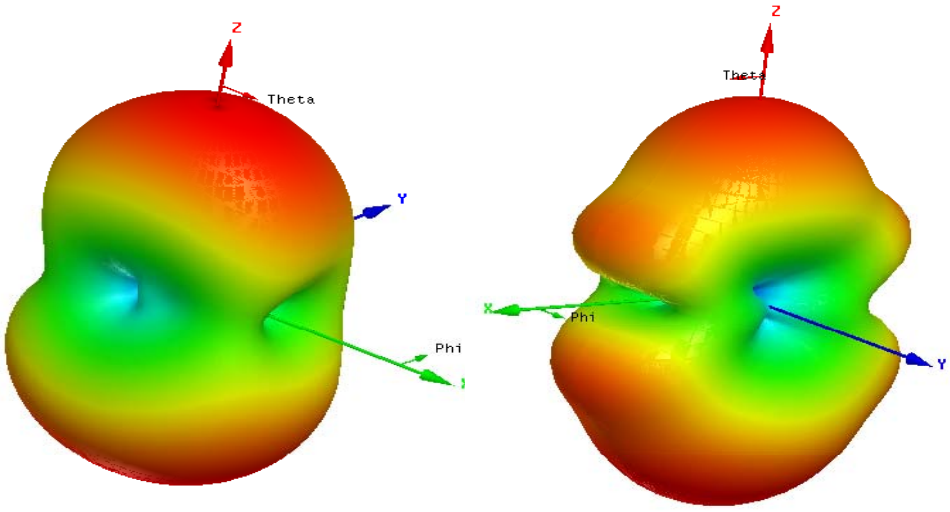


Figure 4.15: Radiation pattern of the active element comparison between full EM simulation (left) and infinite array approximation (right) for 3×3 folded slot antenna array.

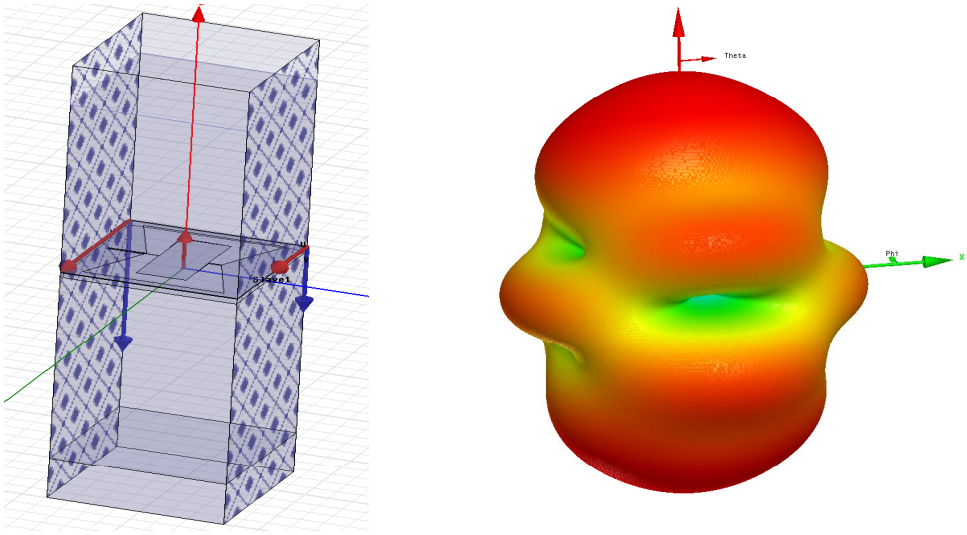


Figure 4.16: Master/Slave boundary setup for an active element in infinite array condition.

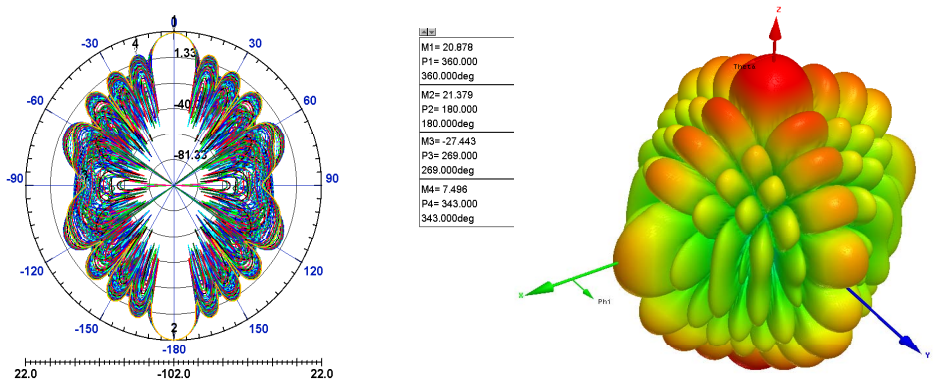


Figure 4.17: Radiation pattern for 10×10 bow-tie slot antenna array in the infinite array condition and its radiation pattern.

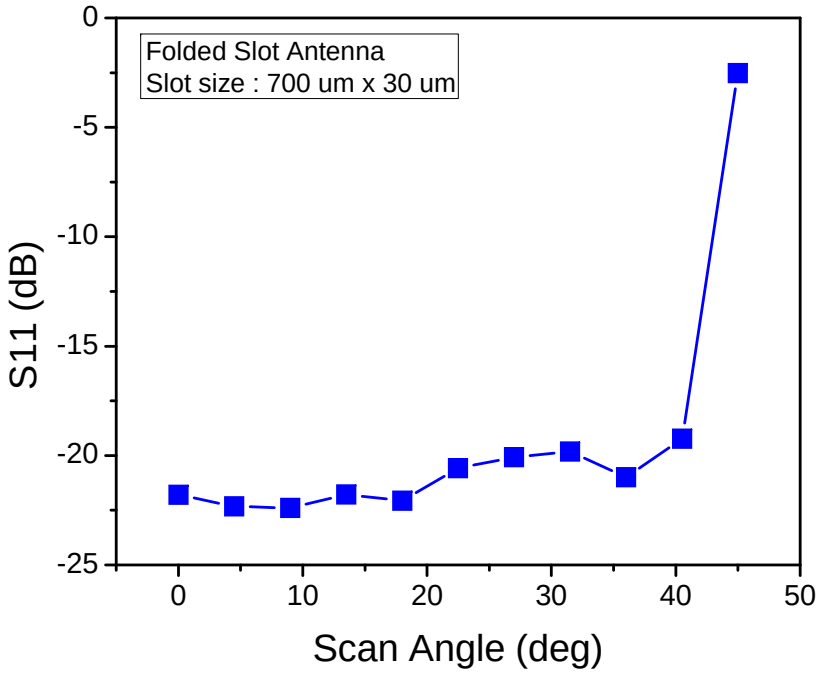


Figure 4.18: S11 with respect to scan angle for the folded slot antenna with infinite array condition.

Chapter 5

Clock Distribution and Synchronization

5.1 System Synchronization Design Considerations

5.1.1 Introduction

Power on three clock networks (H-tree, distributed PLLs [20], coupled oscillators) versus the scaling of the chip size is investigated for synchronous digital systems. In [20], the relations between the clock uncertainty and the parameters in H-tree and distributed-PLLs clock network are given. It was shown that the distributed-PLL clock network is superior in terms the scaling of technology nodes. However, to determine which clock network is preferable in a wafer-scale system, the clock power versus the scaling of the chip size is estimated and compared, given a specified clock uncertainty. It is concluded that the distributed clock networks consume much less power than traditional H-tree network as chip size scales. Moreover, a coupled-oscillator network is promising when applied to large-scale synchronous digital systems.

5.1.2 Methods, Assumptions, and Procedures

Technology

The power of each clock network is normalized to the case where a 1-GHz clock is distributed to a total of 555-pF load capacitance over a chip size of 2 cm $\times$ 2 cm in a 90-nm CMOS technology, while each clock network needs to meet a clock uncertainty of 115 ps (11.5% of the clock cycle) across the chip. A 1-μm wide wire on Metal-9 has sheet resistance R_w=0.046$\Omega/\square$ and capacitance of C_w=0.03 fF/μm. The equivalent resistance $R_\square$ and input capacitance C_g for an inverter is 25.57 k$\Omega/\square$ and 1.17 fF/μm, respectively.

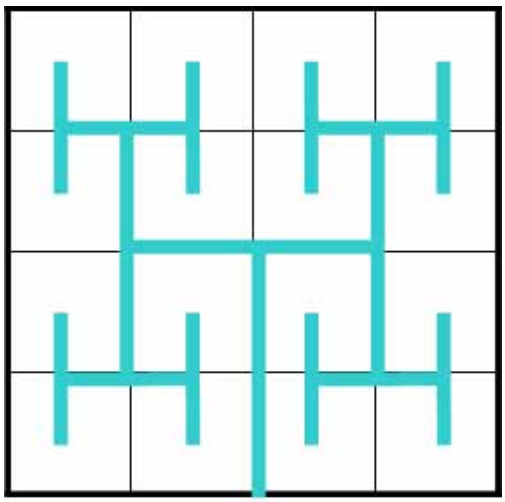


Figure 5.1: H-Tree clock network with level $n = 4$.

Model of Clock Uncertainty and Clock Power

H-Tree

Fig. 5.1 shows an H-Tree clock network with level $n = 4$. An n-level tree would results in $2n$ tiles, and the length from the root to the leaf L is (chip dimension) $\times$ $(1.5(1/2)n/2)$. We have assumed that the H-Tree is optimally buffered [35] and the width of the tree halves after each branching. The clock uncertainty comes from the uncertainty from the root to the leaves of the H-tree and the delay mismatches from the leaves to the flip-flops within the tile, called the internal skew $t_{sk,int}$ [20]. The delay uncertainty from the root to the leaves due to the variability in the buffered segments and time-varying noise coupled from signal transitions of the nearby wires (called jitter t_{jitter}) is linearly proportional to L [20], which is determined by the incorporated variability model. The internal skew is inversely proportional to the square of the length from the leaves to the clock loads or the area of the tile. The clock uncertainty Δt then can be formulated as

$$\Delta t = \alpha t_{H-tree} + t_{jitter} + t_{sk,int} \tag{5.1}$$

where α is assumed to be 0.1. For an 8-level tree, L = 2.875 cm. The length of the optimal buffer segment and the optimal buffer size are calculated to be 2.4 mm and 73 fF, which gives the delay from the root to the leaves to be 270 ps. So the clock uncertainty in the first term is due to the uncertainty in the delay from the root to the leaves is 27 ps. To estimate the jitter, it is assumed that up to 5% of the capacitance of any wire may transition during the time a clock edge propagates [20]. Given the wire capacitance in each segment to be 72 fF and the input and output capacitance of the inverter buffer to be 73 fF (assume $\gamma = 1$), the variation in delay is $\pm$ (5% $\times$ 72)/(72+72+73)= $\pm$1.6%. So the total variation would be 3.2% or 8.6 ps. The maximum

internal skew is $0.046 \times 625\ \mu$m $\times$ 514.7 fF = 15.5 ps. So the total clock uncertainty is 51 ps. The total clock power is contributed by the total wire capacitance of the tree and the buffers, which is proportional to $2n/2$.

With the relations between clock uncertainty/power and the depth of the tree, the required tree depth needed to meet the total clock uncertainty for different chip sizes can be calculated.

$$\Delta t' = (\alpha t_{H-tree_0} + t_{jitter_0}) \cdot s \cdot \frac{1.5 - 2^{n'/2}}{1.5 - 2^{n_0/2}} + t_{sk,int_0} \cdot s^2 \cdot \frac{2^{-n'/2}}{2^{-n_0/2}} < 115\text{ps} \tag{5.2}$$

where $n_0 = 8$ (which is the minimum level to meet this constraint for $s = 1$), $t_{H-tree_0} = 27$ ps, $t_{jitter_0} = 8.64$ps, $t_{sk,int_0} = 15.5$ps, and s is the scaling factor of the chip size.

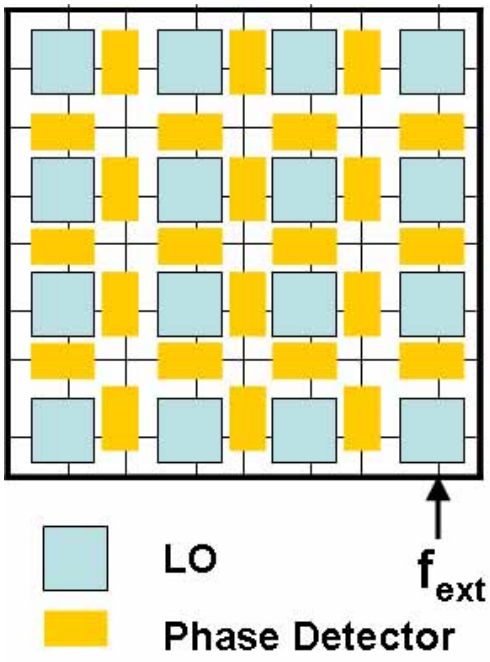


Figure 5.2: Distributed-PLL clock network with 16 tiles.

Distributed PLLs

Fig. 5.2 shows the distributed-PLL network, where each tile has its own local oscillator and there is a phase detector in each boundary [20]. Ideally, the incorporated PLL is a Type-II feedback system and hence the phases between each local oscillator are matched. However, every boundary between tiles introduces some skew because of mismatch in the phase detector. As the number of tiles increases, the number of boundaries and hence the clock uncertainty across the chip increase. In addition, there are still internal skews within tiles.

The clock uncertainty and power for a distributed-PLL network with $n \times n$ tiles is formulated as follows

$$\Delta t = n^2 t_{sk,PD} + \underbrace{t_{sk,int}}_{\propto n^{-1}} \tag{5.3}$$

$$P = n^2 P_{osc} + \left[\frac{(n-2)^2}{4} + 4(n-1) \right] \underbrace{P_{interconnect}}_{\propto (s/n)^2} \tag{5.4}$$

where $t_{sk,PD} = 0.05$ ps is the skew due to the mismatch in phase detectors, P_{osc} the power of an oscillator, and $P_{interconnect}$ is the power on the interconnect. Similar to the H-Tree, given the relation between clock uncertainty/power and the number of tiles, the required number of tiles n for different chip sizes can be calculated.

$$\Delta t' = n'^2 t_{sk,PD} + s^2 \cdot \frac{n_0^2}{n'^2} t_{sk,int_0} < 115\text{ps} \tag{5.5}$$

where $t_{sk,PD} = 0.05$ ps, n_0 =9 (which is the minimum n to meet this constraint for $s = 1$), the resulted $t_{sk,int}$ is 0.046 $\times$ 1111 μm $\times$ 1.7 pF = 64 ps and the total clock uncertainty is 91.6 ps. The clock power for $s = 1$ is normalized to be the same as that in the H-Tree for the ease of comparison.

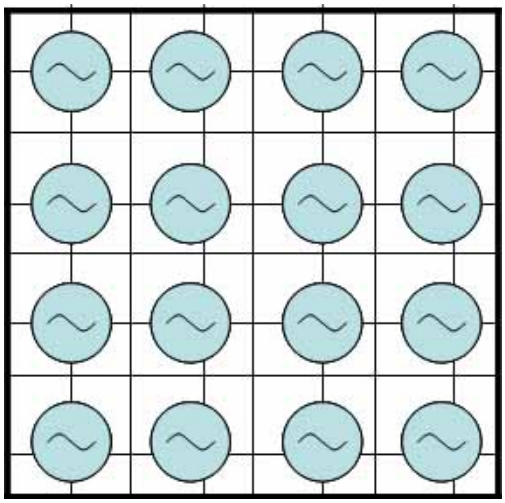


Figure 5.3: Coupled-oscillator clock network with 16 tiles.

Coupled-Oscillator Network

Fig. 5.3 shows a coupled-oscillator network, which is a natural distributed PLLs network only to have a non-linear Type-I characteristic. Specifically, in the steady-state there are phase mismatches

between oscillators according to the differences in their free-running frequencies. The relation is described by the Kuramoto model

$$\frac{d\theta_i}{dt} = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i) \tag{5.6}$$

where θ_i is the phase of each oscillator, ω_i the difference between the free-running frequency in each oscillator and the average frequency, K the coupling strength, N the number of oscillators. Given the standard deviation of the free-running frequency, the standard deviation of the phase mismatch can be found by setting $d\theta_i/dt = 0$ (in the steady state) and inverting the sine operation. To simplify the analysis, we approximated the phase mismatch as $\Delta\theta$ with $\sin(\Delta\theta) = (\sum\sin(\theta_j - \theta_i))/N$. Hence the clock uncertainty is

$$\Delta t_{sk,OSC} = \frac{T}{2\pi}\sin^{-1}\Delta\theta = \frac{T}{2\pi}\sin^{-1}\left(\frac{\Delta\omega}{K}\right) \tag{5.7}$$

where $\Delta\omega$ is the standard deviation of the free-running frequency. If $\Delta\omega$ is independent of the chip size, the clock uncertainty and power for a coupled-oscillator clock network with $n \times n$ tiles is

$$\Delta t = t_{sk,OSC} + \underbrace{t_{sk,int}}_{\propto n^{-2}} \tag{5.8}$$

$$P = n^2 P_{osc} + \left[\frac{(n-2)^2}{4} + 4(n-1)\right]\underbrace{P_{interconnect}}_{\propto (s/n)^2} \tag{5.9}$$

Since the clock uncertainty between the coupled oscillators does not scale with the chip size, n is linearly proportional to s. For $t_{sk,OSC} = 48$ps, $n_0 = 10$ would give $t_{sk,int} = 0.046 \times 1000\mu\text{m} \times 1.4$ pF = 64 ps to meet the constraint of the total clock uncertainty. Similarly, the power for $s = 1$ is normalized to that for the H-Tree.

5.1.3 Results and Discussions

Fig. 5.4 shows the power versus the scaling of the chip size for the above three cases. The power for the distributed-PLL and coupled-oscillator case is normalized to that of the H-Tree, which is expressed in terms of the total clock load (550 pF). The power of the H-Tree grows exponentially with the chip size while the other two distributed networks show much slower trend. Even if the power is not a limiting factor, the chip size cannot keep increasing for the H-Tree and the distributed PLLs since the constraint on clock uncertainty can no longer be met. However, a coupled-oscillator network is not limited by the performance constraint as chip size scales. We assume the variation of the oscillation frequency does not depend on the chip size, which needs further verification with simulation tools.

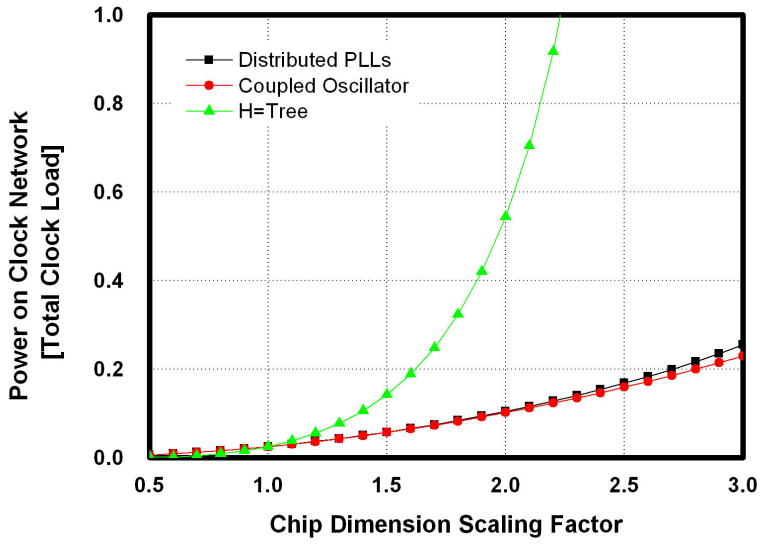


Figure 5.4: Power for the three clock networks versus chip dimension scaling factor.

The trend of clock power as chip size scales is investigated for three cases. Distributed clock networks show significantly lower power than a traditional clock tree. Moreover, coupled-oscillator clock network is not performance-limited as chip size scales. Therefore, it is favorable to apply distributed clock network such as coupled oscillators in a wafer-scale system.

5.2 Clock Distribution

The distributed radio on the whole wafer is composed by thousands of individual transceivers, creating a large flexible phased array. Synchronization of these transceivers is mandatory for achieving good performance of the beam forming and required directivity. The purpose of this task is to study the clock distribution and synchronization at the wafer scale. The distributed clock would be used both for baseband data synchronization at few GS/s and as a reference for the 90GHz LO PLLs.

In the first part, the main criterions for the clock distribution scheme will be derived from the constraints of the system. From a panel of architectures, a clock distribution using coupled standing-wave oscillators has shown to be promising for this large-scale application. The design flow and simulation tools will be introduced and transmission line model will be described in this preliminary part.

In a second section, the coupled standing-wave oscillators architecture will be reviewed and a design methodology will be exposed for reducing both the clock skew between clusters and the

global power consumption. Issues, such as tolerance to process variations, clock buffer design and locking range, will be addressed for this particular implementation to evaluate if this scheme is suitable for a wafer-scale clock distribution with good synchronization. Finally, as integration is of obvious interest in this project, this clock distribution scheme should be co-designed with the antenna pattern and perturbations of this scheme on the radiation pattern must be investigated. We propose here a clock distribution pattern that fits the antenna pattern and is well suited for the cluster-based distribution as clock skew between them is reduced. From the defined pattern, simulations of coupled effects at large-scale must be investigated to ensure that there is no locking issues between clusters.

This study has shown that the coupled standing-wave oscillators are very good candidates for the clock distribution at the wafer-scale in terms of acceptable synchronization and reduced power consumption. Although the simulation results seem promising, further work would demonstrate the feasibility of the proposed co-designed clock distribution pattern on a small scale before being able to address the large-scale locking issues that can appear on a whole wafer.

5.3 Methods, Assumptions, and Procedures

5.3.1 Assumptions for the wafer-scale system

For this study, we will assume that the reticles over the wafer can be stitched to each others with a good alignment to ensure communication between them for clocks, control signals and data networks. This is largely done in image sensor chip design where area larger than one reticle is needed [23][8]. Furthermore, each reticle, that may contain several transceivers, would be identical. That will define one of the biggest constraints for the clock distribution.

Another assumption is that power is supplied to all reticles the same way. As antennas would be integrated on top of the wafer, power can be supplied either from the sides of the wafer or through the backside using through-substrate interconnects (TSI). The first option does not seem to be feasible due to different travel lengths between reticles in the middle and at the periphery of the wafer. A through substrate power delivery is more suitable and would be assumed here. In the continuation of this project, this system level issue must be addressed and opens a good research field on 3D integration, correlated with the wafer-scale project.

5.3.2 Criterions for the clock distribution and architecture choice

The main criterions for choosing a convenient distribution scheme are the ability to be tile-able and repeatable and to provide low skew. Furthermore, the power consumption, depending on frequency of operation and granularity of the scheme, is another important criteria. The following architectures for clock distibution has been identified and discussed:

- H clock trees or grids. The clock tree implementation suffers from the fact that it is not

easily patternable and presents large clock skew. A grid can help reduce clock skew but increase power consumption during transients and has a slower edge rate.

- Clock trees with feedbacks [19]. Feedbacks can be included to compensate for the clock tree delays but, on a whole wafer, it appears hard to compare nodes placed far away for each other.

- Bi-directonal signaling [34]. A bidirectional transmission line travels all over the wafer and average time extractors are used to pick up a synchronous clock at any point. This approach provides low skew and clever floorplan can be used to reach all reticles. The disadvantage is that it requires no defects on the wafer reticles. If a reticle does not work properly, then the line is cut and disconnects all following reticles.

- Rotary travelling-wave oscillators [43]. A wave travels along a differential transmission line having edges inverted and connected together. The phase varies linearly over the transmission line. By coupling another oscillator at the right location, the right phase of the signal can be extracted to lock the oscillators together. This scheme appears suitable for a wafer-scale distribution.

- Coupled standing-wave oscillators [29][6]. Contrary of travelling waves, the standing-wave structure just reflects the waves travelling along its lines, creating a constant phase signal, independent of the position. Nevertheless, the amplitude will vary from one position to another. This architecture would require compensation of the line losses.

- Coupled ring oscillators [22][14]. Ring oscillators would be designed locally on each cluster and then connected to their neighbors to lock them together. The obvious limitation is the large length required between each cluster that adds some latency on the coupling scheme and thus increases the clock skew.

- Distrbuted PLLs [21]. VCOs are distributed on each cluster and phase detectors are implemented at the boundary between two clusters. The PLLs then corrects for the phase and frequency mismatch between two VCO cores. The same limitation as coupled ring oscillators is found as the length between the VCOs and the phase detectors would be significant.

Rotary travelling-wave and coupled standing-wave oscillators seems suitable for a wafer-scale clock distribution, as they inherently cover large areas and can be easily physically coupled to their closest neighbors. Standing-wave oscillators are particularly interesting due to the fact that the phase of the oscillation remains constant all over the transmission lines, what is greatly suitable to coupled structures. Traveling-wave oscillators rotate the phase and require special care to be coupled together. For this reason, the focus will be made on the study of coupled standing-wave oscillators.

5.3.3 Design Flow and simulation tools

For the evaluation of the coupled standing-wave oscillators potential, several tools have been used together. For the line model, ADS Momentum electromagnetic simulator has been used to extract the S-parameters of a transmission line in a nanoscale CMOS back-end process and to create a *RLCG* model of a small portion of the line. Cadence simulation tools have been used to simulate the transmission lines together with the compensation cells and to evaluate the performance in terms of power consumption, clock skew and amplitude swing. A theoretical study has been performed using Matlab to define a design methodology for sizing the parameters relating to the desired performance and to validate the Cadence simulation results.

In further steps of the design, electromagnetic simulators would be used for simulating the co-integrated antennas with the coupled standing-wave oscillators and large-scale numerical simulations, such as Spice++ simulator, would be used to detect any locking issue between oscillators at the wafer-scale.

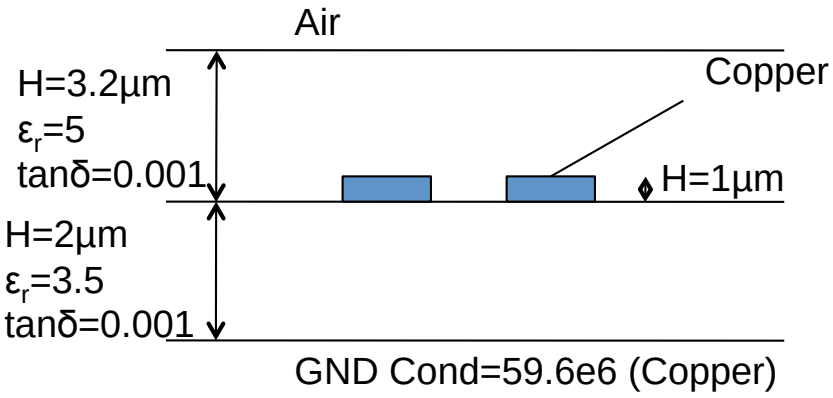


Figure 5.5: Model for the transmission line in a top metal layer with infinite ground.

5.3.4 Transmission Line model

Coupled standing-wave oscillators use transmission lines in a nanoscale CMOS backend that have to be modeled. A simple model for transmission lines in a top metal layer has been made in Momentum for deriving the *RLCG* parameters of the line. The model used is described in Fig. 5.5 and uses a low metal layer as a ground. This ground is modeled with an infinite thickness. S-parameters for 12μm wide and 4μm space differential lines is extracted. Then frequency-dependent *RLCG* parameters are obtained using the following formulas [16][26] and are imported into Cadence circuit simulation tools as a mtline element.

First, the S-parameters must be converted into ABCD parameters using:

$$A = \frac{(1+S_{11})(1-S_{22})+S_{12}S_{21}}{2S_{21}} \tag{5.10}$$

$$B = Z_{0,port}\frac{(1+S_{11})(1+S_{22})-S_{12}S_{21}}{2S_{21}} \tag{5.11}$$

$$C = \frac{1}{Z_{0,port}}\frac{(1-S_{11})(1-S_{22})-S_{12}S_{21}}{2S_{21}} \tag{5.12}$$

$$D = \frac{(1-S_{11})(1+S_{22})+S_{12}S_{21}}{2S_{21}} \tag{5.13}$$

Next, Z_0 and γ are computed from the ABCD parameters using:

$$Z_0 = \sqrt{\frac{B}{C}} \tag{5.14}$$

$$\gamma = \frac{arccosh(A)}{l} \tag{5.15}$$

Finally, *RLCG* parameters are extracted from Z_0 and γ using:

$$R = \Re(Z_0\gamma) \tag{5.16}$$

$$L = \frac{\Im(Z_0\gamma)}{\omega} \tag{5.17}$$

$$G = \Re(\gamma/Z_0) \tag{5.18}$$

$$C = \frac{\Im(\gamma/Z_0)}{\omega} \tag{5.19}$$

As an example, the *RLCG* parameters at 10GHz are R = 3.7495 kΩ/m, L = 297.75 nH/m, G = 8.5257 mS/m and C = 142.88 pF/m. This model uses an infinite ground and will be used for analysis purpose only. For a more realistic model, a finite ground thickness must be introduced and more sophisticated structures, such as a slotted ground, should be included.

5.4 Results and Discussions

5.4.1 Standing-wave oscillator description

A standing-wave oscillator is basically a $\lambda/2$ differential transmission line shorted at both ends. If we consider an ideal transmission line, by injecting current at the center of this line, a wave propagates to the edge, reflects onto the shorted node and comes back to where it comes from, creating a standing wave on the line. The particularity of this standing wave is that its phase is the

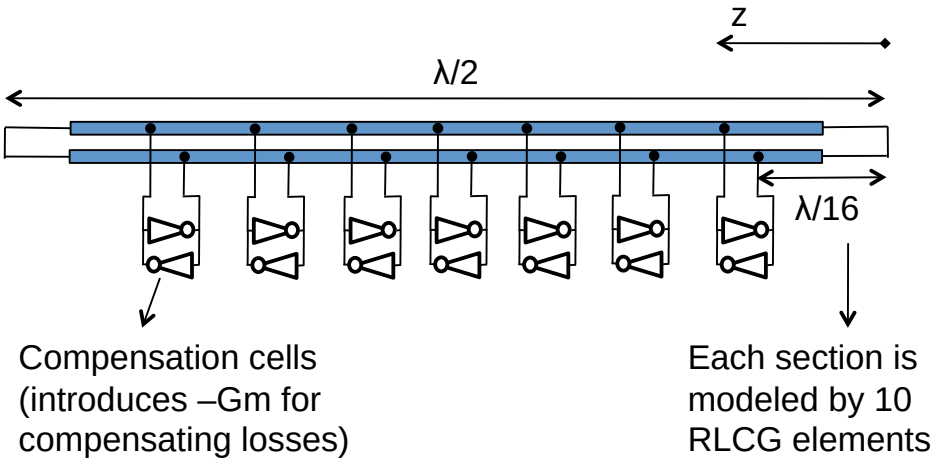


Figure 5.6: Standing-wave oscillator description.

same at any point of the line. Nevertheless, the amplitude of the clock is maximum at the center and minimum at the edges.

Transmission lines implemented on CMOS back-ends are lossy. This has to be taken into consideration by distributing compensation cells along the lines. These compensation cells should present a negative conductance, such as cross-coupled inverters. Lets describe the behavior of the voltage along the line to quantify the required compensation. A line model with distributed coupled inverters is presented on Fig. 5.6. The voltage on the standing-wave oscillator is $V(z) = V_0(e^{\gamma z} + \Gamma e^{-\gamma z})$, with γ, the propagation constant and Γ, the reflection coefficient. As edges are shortened, $\Gamma = -1$. That leads to:

$$V(z) = 2V_0 \sinh(\gamma z) \tag{5.20}$$

If a lossless transmission line is considered, then $\gamma = j\beta$ and:

$$V(z) = -j2V_0 \sin(\beta z) \tag{5.21}$$

. That means that the phase is constant and that the amplitude is rising from 0 at the edge to $2V_0$ at the center of the transmission line with a sine shape.

For a lossy transmission line, an attenuation constant is added to the propagation constant and will degrade the phase of the standing wave. Thus, $\gamma = \alpha + j\beta$ and can also be expressed as:

$$\gamma = \sqrt{(R + jL\omega)(G + jC\omega)} \tag{5.22}$$

when considering the *RLCG* distributed parameters of the line. A low-loss approximation can be used for linking the *RLCG* parameters to α and β. This approximation assumes that

$$\frac{G}{C\omega} \ll 1 \tag{5.23}$$

$$\frac{R}{L\omega} \ll 1 \tag{5.24}$$

and is valid for high frequencies (over 1GHz). This leads to:

$$\alpha = \frac{1}{2}(\frac{R}{Z_0} + GZ_0) \tag{5.25}$$

$$\beta = \omega\sqrt{LC} \tag{5.26}$$

, where $Z_0 \approx \sqrt{\frac{B}{C}}$.

By computing γ with a real part into the voltage equation, Fig. 5.7 is generated and describes the skew related to the clock period versus the position on the transmission line. We observe that using a higher clock frequency leads to a lower skew between the edges and the center of the lines.

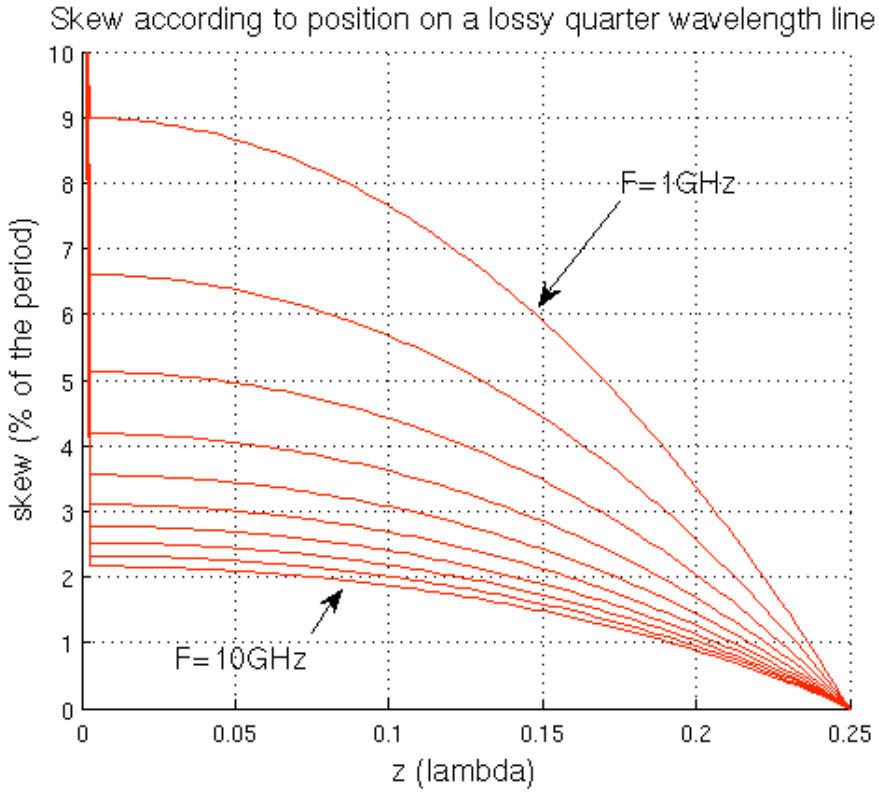


Figure 5.7: Skew related to the clock period versus according to the position on the transmission line and for various oscillation frequencies.

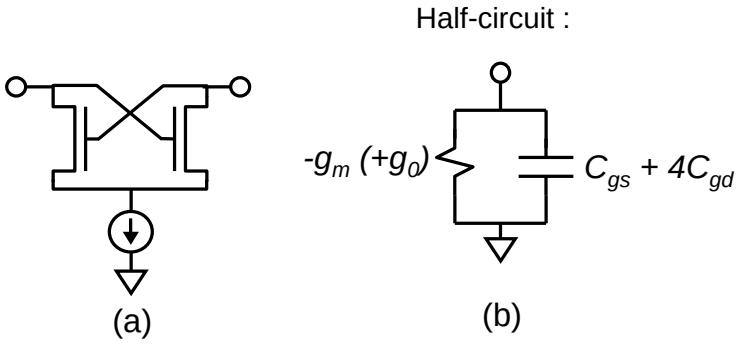


Figure 5.8: Compensation cell based on a cross-coupled NMOS pair (a) and associated equivalent circuit (b).

5.4.2 Compensation of the coupled standing-wave oscillator

Compensation cells are distributed along the transmission line and introduce a negative conductance to compensate for the losses of the line. The compensation cells will be designed as cross-coupled NMOS transistors driven by a current source (Fig. 5.8a). The current will be supplied at the edges of the line by connecting the nodes at Vdd. As the total resistance of the line will not be high, the DC points all over the line would roughly be V_{dd} with a few mV difference. The compensation cells introduce a $-Gm$ conductance and an additional C_m parallel capacitance, as shown in Fig. 5.8b. The C_m will be equal to $C_{gs}+4C_{gd}$ and the total conductance to $-G_m+G_0$, this latter value being neglected in front of G_m. By including G_m and C_m, the components of the propagation constant become:

$$\alpha = \frac{1}{2}(R\sqrt{\frac{C+Cm}{L}} + (G-Gm)\sqrt{\frac{L}{C+Cm}}) \tag{5.27}$$

$$\beta = \omega\sqrt{L(C+Cm)} \tag{5.28}$$

The compensation would be effective if $\alpha = 0$ what leads to:

$$G_m = \frac{R}{{Z_0}^2} + G \tag{5.29}$$

Nevertheless, C_m will reduce the effect of $-G_m$ and will shorten the line for the same oscillation frequency. In a real system, we would choose G_m to be equal to 1.5 to 2 times this minimum value for starting and maintaining the oscillation with an acceptable signal swing.

A methodology has been defined to correctly size the transistor width and the required current to meet the desired specifications. NMOS transistors in a 65nm process have been characterized using f_t, which gives the ratio between G_m and C_m of the cross-coupled pair. It is defined here as:

$$f_t = \frac{G_m}{2\pi(Cgs+4Cgd)} \tag{5.30}$$

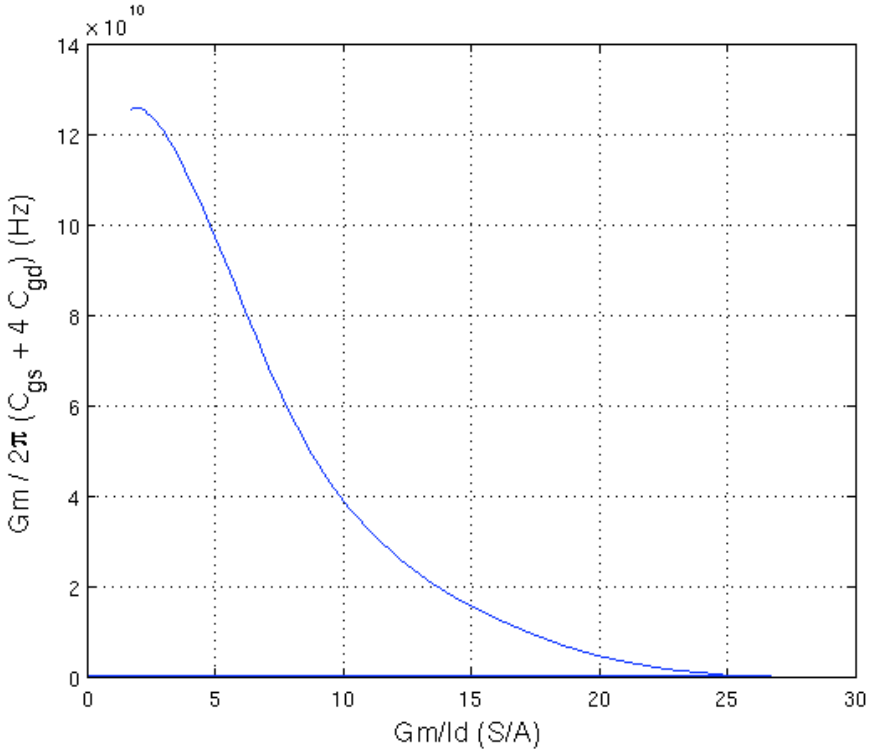


Figure 5.9: f_t versus G_m/I_d for 65nm NMOS transistors.

Fig. 5.9 presents the relation between ft and G_m/I_d for a minimum length and 1μm wide transistor. This helps us relate the current and the transistor width with the G_m and C_m design parameters. Thus, f_t can be taken as the main parameter to analyze the compensation cell.

Using Matlab to compute the transmission line and compensation cells formulas, the required G_m value per meter is plotted versus f_t on Fig. 5.10a for various frequencies of operation from 2GHz to 20GHz. For this design, we consider 2 times the minimum G_m that compensates exactly for the losses. The current consumption per unit meter is also plotted on Fig. 5.10b. The length of a λ line is computed from the added capacitance and is plotted on Fig. 5.10d. Finally, the power consumption for a whole wafer can be derived from the current and the length of the line. The power consumption is dependent on the pattern chosen for the coupled standing-wave oscillators. For this computation, a simple $\lambda/2$ x $\lambda/2$ pattern is chosen, although the pitch would be smaller for the real pattern. The power consumption, for a 1V V_{dd}, is then:

$$P_{tot} = \frac{4Id}{\lambda} \times \frac{\pi d^2}{4} = \frac{Id \times \pi d^2}{\lambda} \tag{5.31}$$

, where d is the diameter of the wafer and I_d is the current in A/m. Fig. 5.10c depicts the power consumption versus f_t for various oscillation frequencies. From these curves, the whole system is

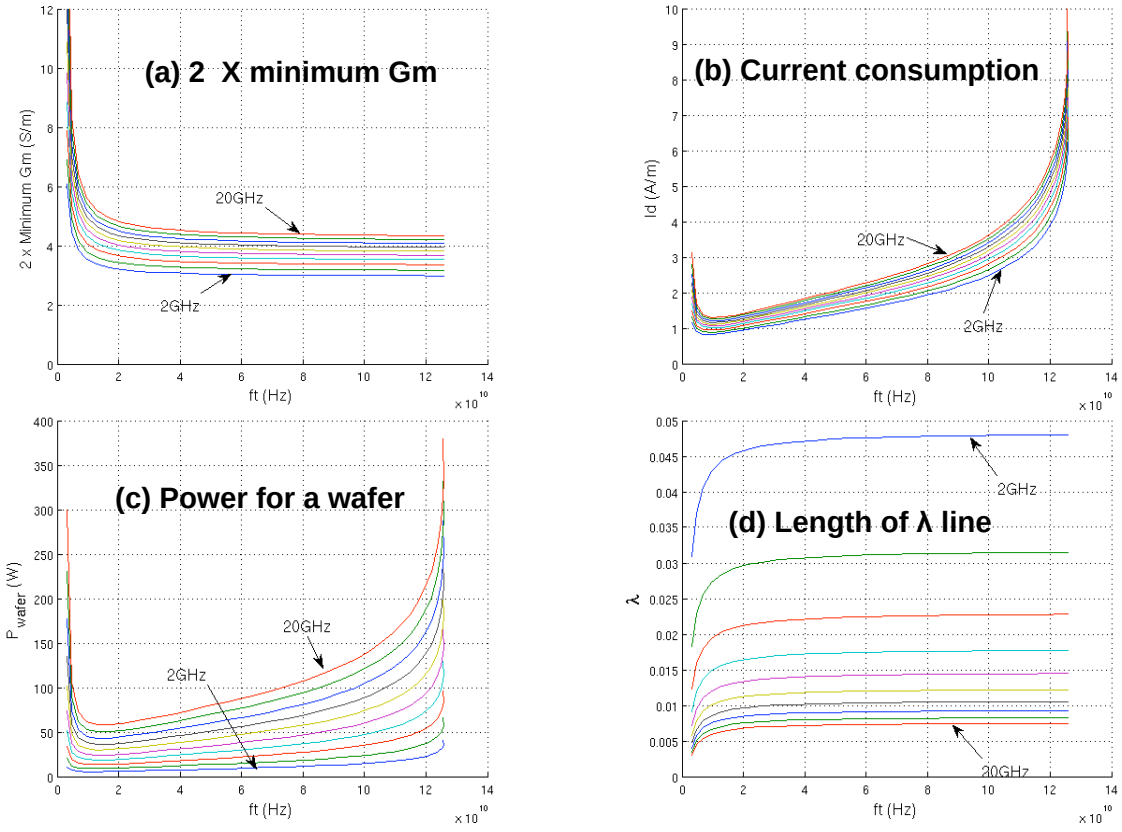


Figure 5.10: $2G_{m,min}$ (a), current consumption (b), power consumption (c) and length of line (d) versus f_t for various oscillation frequencies.

characterized by choosing an f_t value. For choosing the design point, the swing of the oscillation must be considered. The amplitude is roughly equal to the tail total current I_{tail} multiplied by the equivalent parallel resistance of the tank R_{eq}. As an example, for this particular configuration, a 2mA current by compensation cell is translated to a swing of about 750 to 800mVpp. Then, this value can be plotted on the graph from Fig. 5.10b to derive the f_t value for the chosen oscillation frequency. Fig. 5.11 describes this choice for a 10GHz oscillation frequency. Table 5.1 sums up the designed values for this particular operation point with 12μm wide transmission lines spaced by 4μm and oscillating at 10GHz. Note that the wafer power is also computed for a λ/8 x λ/8 pattern.

When a real current source is simulated instead of the ideal current source, as described on Fig. 5.12, the skew between the center and other positions on the line is degraded. This is mainly due to the fact that the current drawn by the sources is dependent of the voltage swing, which is in fact dependent of the position on the line.

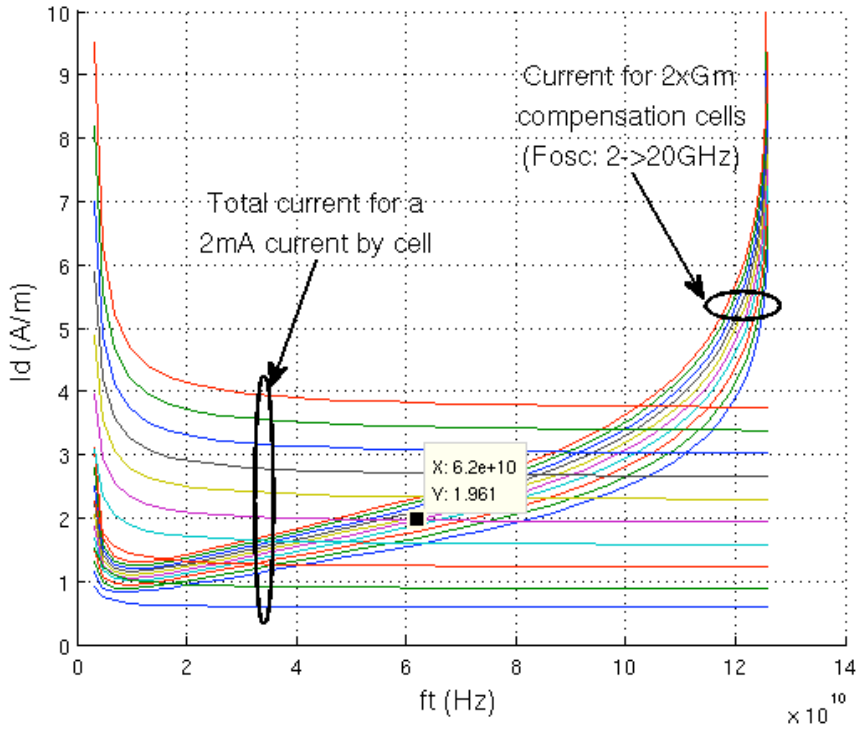


Figure 5.11: Current consumption versus f_t and design point for a 10GHz standing-wave oscillator.

5.4.3 Design optimization

To optimize the design in terms of power, we have stated before that the signal swing $Amp \approx I_{tail} \times R_{eq}$. That means that to reduce the consumed current for the same signal swing, the equivalent parallel impedance must be increased. Compensation cells have a very small effect on the impedance seen from the middle of the line, Z_{in}, which is a strong function of the width and space of the transmission lines. *RLCG* parameters are plotted on Fig. 5.13 versus width W and space S of the lines. Next, Q and Z_{in} are computed by using:

$$Q = \frac{\omega L}{R} \tag{5.32}$$

$$Z_{in} = \frac{L}{CR} \tag{5.33}$$

Only the inductive Q is considered here, as the model of the line is simple and does not include a finite ground plane for the low metal layer, thus leading to a very small G value. Q and Z_{in} plots are depicted on Fig. 5.14. From these plots, the design point chosen to minimize the power

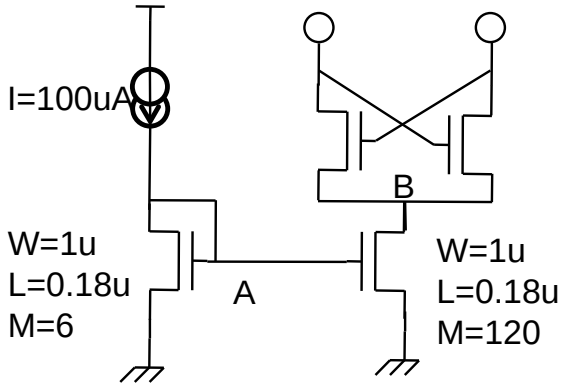


Figure 5.12: Compensation cell including the mirror-based current source.

Table 5.1: Design parameters for a 12μm wide 4μm space standing-wave oscillator at 10GHz

	Ideal current source	Real current source
λ/2 line length	7.12mm	
Cross-coupled transistors width	11.1μm	
Cross-coupled pairs current	2mA	
Wafer power for λ/2 x λ/2 pattern	39W	
Wafer power for λ/8 x λ/8 pattern	310W	
Amplitude at the center	824mVpp	862mVpp
Skew (between center and λ/8 point)	0.9ps	2.6ps
Skew (between center and λ/16 point)	1.7ps	4ps

consumption is a 6μm width and a 20μm space. The quality factor of the transmission line should therefore be equal to 4. Note that another trade-off might be chosen if a finite ground thickness is modeled. Moreover, a quick study has shown that using a slotted ground would be beneficial as it increases Z_{in} while keeping the Q factor almost constant. When implemented, the transmission line with slotted ground should be carefully modeled to take into account all effects in the simulation. The model has not been detailed so far, because the main purpose of this study was to draw a picture of possible achievements and improvements.

Table 5.2 sums up the designed parameters for the chosen 6μm wide, 20μm space transmission lines. We can observe that the power consumption is a quarter less than previous results for the same performance of the standing-wave oscillator. Furthermore, the consumption can be further decreased, at the expense of the swing amplitude, by decreasing the compensation negative conductance to $1.5G_{m,min}$. This saves an additional third of the power and decrease the skew, but the amplitude of the oscillation is slightly lower. The skew decrease is easily explained by the fact

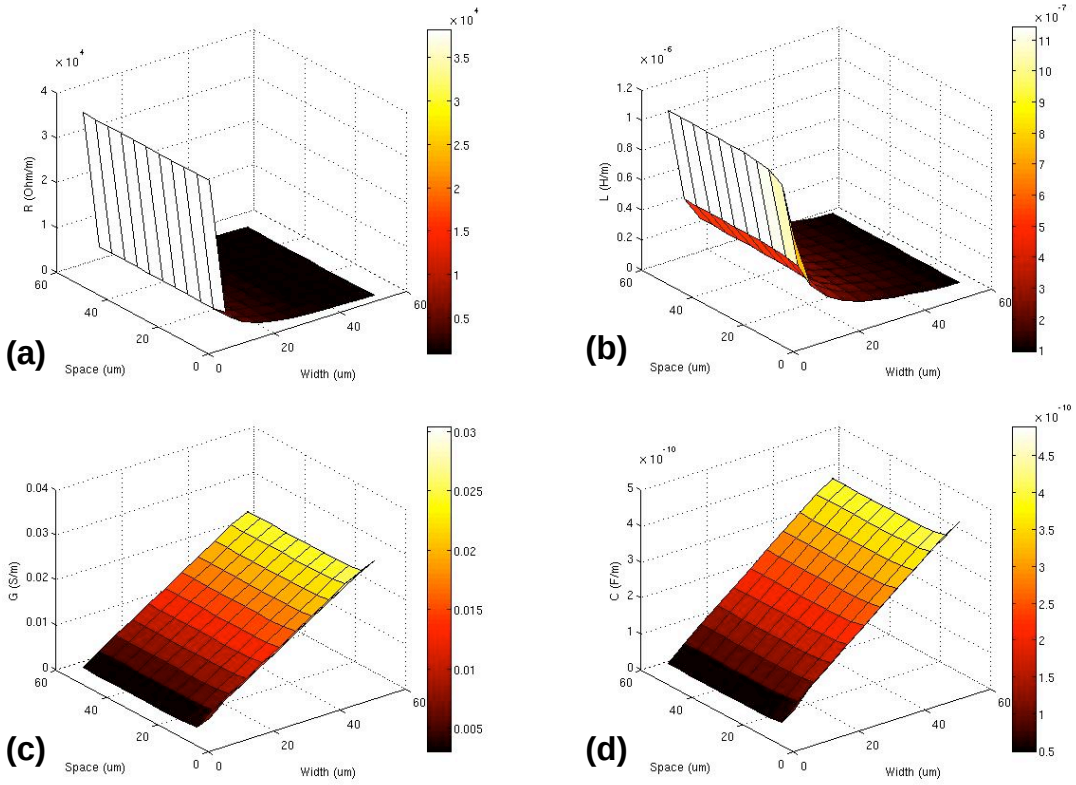


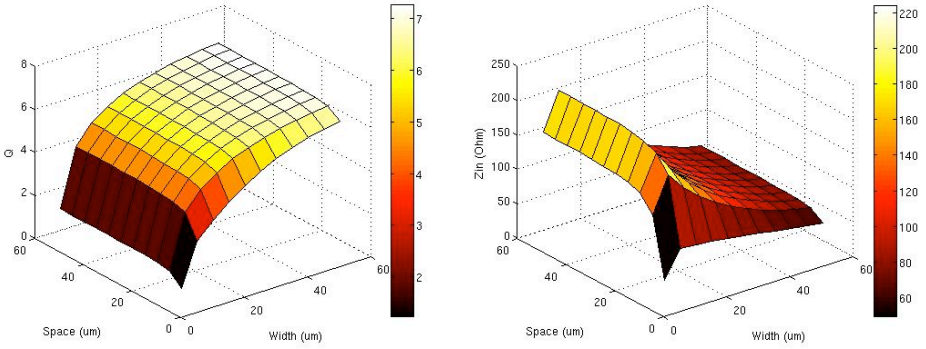
Figure 5.13: R (a), L (b), G (c) and C (d) parameters of the transmission line versus width and space.

that compensation cells add an additional G_m that compensates the losses, but as the G_m introduced is higher than the $G_{m,min}$, then extra "losses" are introduced and increase the skew. In fact, if $G_m = 2 \times G_{m,min}$, then the skew would be equal to the case in which there is no compensation (although in the present case the oscillation is maintained with an acceptable level). Decreasing G_m to $1.5G_{m,min}$ appears to be a good trade-off between power, swing and skew.

5.4.4 Tolerance of the standing-wave oscillator to PVT variations

To evaluate the tolerance of the standing-wave oscillator to variations, the previous model has been used and parameters has been changed one at a time. The influence on oscillation frequency, amplitude swing and clock skew has been studied. First, the physical dimensions of the transmission line are modified (width, length, space, thickness of metal layer and height of the dielectric). Next, the process, voltage and temperature are modified. Finally, the worst-case in all deviations is considered, since running extensive simulations with all corners is too time consuming.

The influence of physical parameters on frequency, amplitude and skew is shown in Table 5.3.



(a) (b)

Figure 5.14: Q (a) and Z_{in} at the center of the line (b) versus width and space.

Table 5.2: Design parameters for a 6μm wide, 20μm space standing-wave oscillator at 10GHz

	2 x Gmmin	1.5 x Gmmin
$\lambda/2$ line length	6.725mm	6.875mm
Cross-coupled transistors width	7.4μm	5.7μm
Cross-coupled pairs current	1.34mA	1mA
Wafer power for $\lambda/2$ x $\lambda/2$ pattern	29W	21W
Wafer power for $\lambda/8$ x $\lambda/8$ pattern	232W	170W
Amplitude at the center	860mVpp	650mVpp
Skew (between center and $\lambda/8$ point)	3ps	1.1ps
Skew (between center and $\lambda/16$ point)	4.7ps	1.9ps

The length of the line has the strongest influence on the frequency as a 1% change in length translates in a 1% change in frequency. The amplitude is affected by all deviations and the skew is mostly influenced by the height of the dielectric with about a 0.1ps variation per percentage of height variation. As accuracy of process is absolute, the relative variations are obtained by taking into account a 40nm imprecision on the metal layer drawing, a 0.15μm imprecision on the metal layer thickness and a 0.3μm imprecision on dielectric height, which are typical values for nanoscale CMOS processes. This leads to the relative variations of Table 5.3. It is notable that vertical variations have a higher relative influence than horizontal variations. Finally, absolute variations are mainly driven by the metal layer thickness and dielectric height. Taking all these variations into account leads to a frequency variation of roughly 70MHz, an amplitude variation of 200mV and a skew deviation of 1.5ps. Frequency and skew deviations seem low, while amplitude is largely affected by physical parameters.

Table 5.3: Variations of frequency, amplitude and clock skew related to physical parameters of the transmission lines

	Width	Length	Space	Thickness	Height
ΔFrequency (MHz/%)	7.4	-100	2.9	-2.6	-1.5
ΔAmplitude (mv/%)	-3	-6.75	2.15	6.2	8.95
ΔSkew (fs/%)	-15-20	-30-40	15-20	35-50	70-100
Relative variations	0.3%	0.0005%	1%	15%	12%
ΔFrequency (MHz)	2.22	-0.05	2.9	-39	-18
ΔAmplitude (mV)	-1	-0.003	2.15	93	107
Δskew (fs)	-6	-0.02	20	750	1200

Process variations on transistors of the cross-coupled pairs, voltage and temperature are analyzed the same way. The effects of these variations are small on all design parameters. Table 5.4 sums up these variations. The process has been varied from SS to FF. The temperature has been swept from 27°C to 80°C and the voltage by ±200mV.

We can conclude that metal thickness and dielectric thickness variations dominate the global variations on frequency, amplitude and clock skew and that the standing-wave oscillator architecture is quite tolerant to all types of variations encountered on a whole wafer. Only 100MHz frequency offsets could be observed for 10GHz oscillators on different reticles, what seems to be good enough to synchronize these oscillators at a large scale. The clock skew variation would be less than 2ps, considering the center of each oscillator, what is definitely a good value for synchronizing several transceivers. However, the amplitude variation is quite large, around 300mV for a 1V supply, and would contribute to further mismatch between clusters. In order to counteract this effect, a feedback system able to track the voltage swing and modify the bias current to correct the variation must be investigated.

Table 5.4: Variations of frequency, amplitude and clock skew related to physical and PVT variations.

	Physical param.	Process (SS or FF)	Temperature	Vdd
ΔFrequency	70MHz	10MHz	10MHz	4MHz
ΔAmplitude	200mV	10mV	30mV	60mV
ΔSkew	1.5ps	0.3ps	0.2ps	0.1ps

5.4.5 Clock buffer design and evaluation

As clocks are used both for 90GHz PLL reference and for the baseband circuitry, a buffer is useful in the latter case to square the standing-wave signal. If the clock is picked at different positions on the line (i.e. center, 25%, 50% and 75%), the amplitude of the input signal will be different. Then a buffer having an input amplitude-independent propagation delay is needed. Moreover, the buffer must be tolerant to PVT variations.

For addressing the input amplitude-independent buffer, fixed dummy inverters are used to load the inverter chains to account for the sinusoidal amplitude distribution along the line. The dummy slows down the first inverter and equalizes the delay for the specified input amplitude. The dummy sizes are 0.94, 0.84 and 0.53, for the center, 25% and 50% of the line, respectively, as described on Fig. 5.15. The buffer chains are AC coupled to the line and the input load and output drive are equivalent for different buffers. The bias point is fixed to 0.5V.

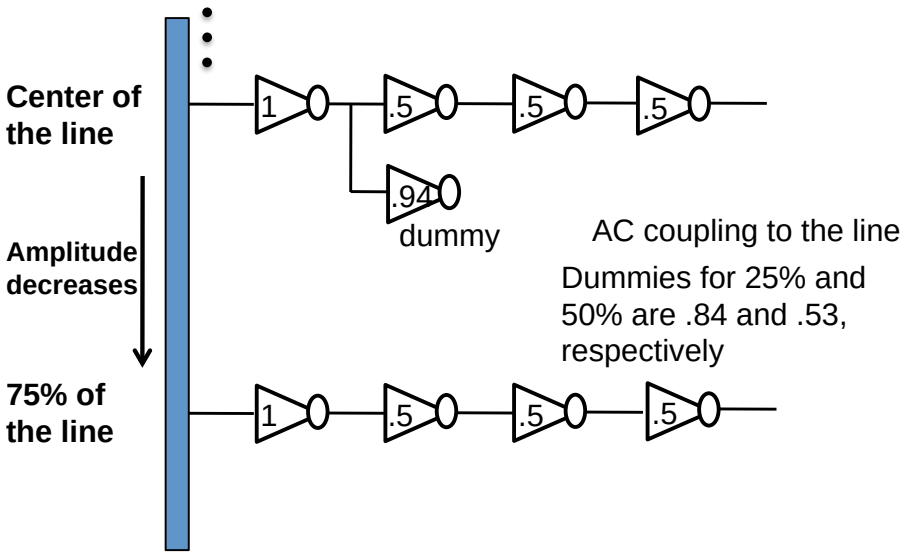


Figure 5.15: Clock buffer including dummy inverters for propagation delay equalization.

The propagation delay of the center buffer is plotted in Fig. 5.16a for process and temperature variations. It can be observed that the propagation delay may vary from 23 to 38ps, introducing a 15ps difference. Fig. 5.16b presents the skew for different buffers on the same line. The skew difference may be up to 2ps. The voltage influence on the propagation delay and skew is plotted on Fig. 5.17. The influence is on the same order than process and temperature. All together, the center buffer simulated at SS 80C 0.8V presents a 23ps difference compared to TT 27C 1V. The skew increases by 4ps. Thus, the global skew due to PVT variations over the wafer for a simple buffer is equal to 27ps, which is more than a quarter of the clock period when operating at 10GHz. As these variations are rather large, another architecture is needed to compensate for this. For example, a

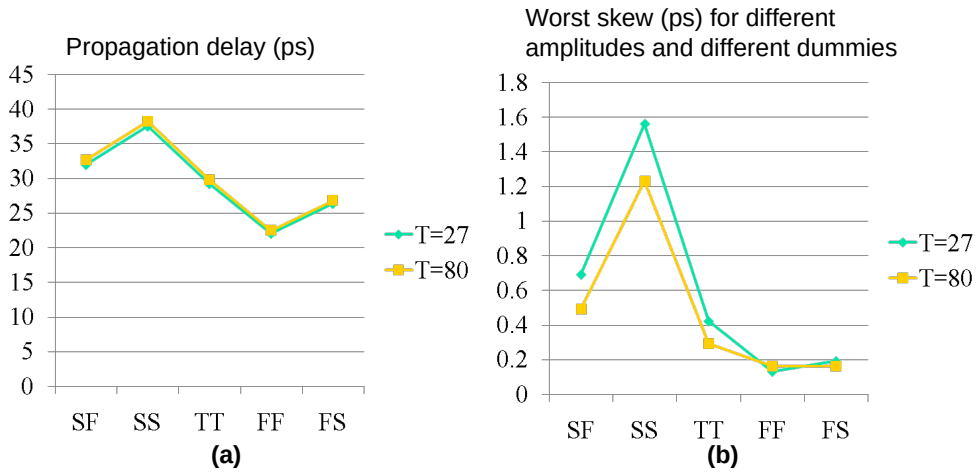


Figure 5.16: Buffer propagation delay (a) and worst skew (b) versus process and temperature corners.

DLL approach, as shown in Fig. 5.18, can help reduce the global skew difference. Moreover, an optimized pattern in which the clock is only picked at the same position on every line would be preferable for a global skew reduction.

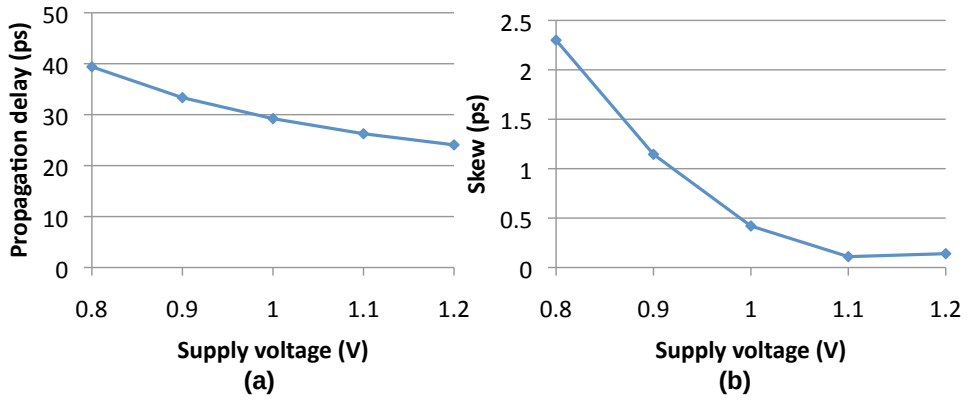


Figure 5.17: Buffer propagation delay (a) and worst skew (b) versus supply voltage.

5.4.6 Locking range of coupled standing-wave oscillators

Coupling two standing-wave oscillators can be achieved in many different ways, depending on the pattern chosen to cover the area. A simple coupling of 4 oscillators is shown in Fig. 5.19. All

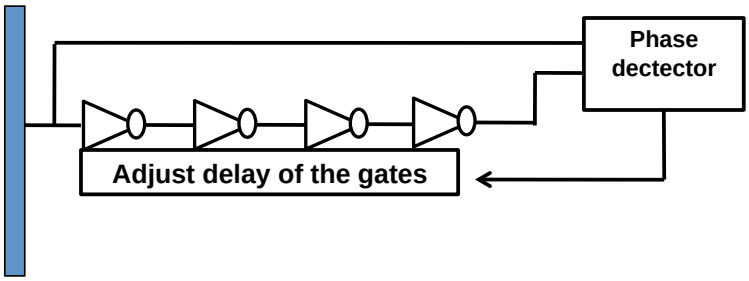


Figure 5.18: Proposed DLL-based approach for compensating the PVT variations on the clock buffers.

standing-wave oscillators are coupled at the exact same position to ensure a locking at the specified frequency. The coupling is very strong as it is accomplished by vias between two metal layers, providing only few ohms of interconnect resistance.

A quick analysis of the coupling behavior is provided here to show how strong is the coupling mechanism. The locking range will be analyzed. It roughly defines the maximum frequency deviation allowed for two oscillators to lock to each other. From Adlers equation, we can derive:

$$\Delta f_{lock} = \frac{f_0}{2Q}\frac{A_{inj}}{A} \tag{5.34}$$

with A_{inj} and A, the relative amplitude of the injecting signal and the oscillator signal, both at the center of the line, Q, the quality factor of the line and f_0, the oscillation frequency. As an example, with a Q of 4 and a coupling at the center of the line ($A_{inj} = A$), then the locking range is equal to 1250MHz for a 10GHz oscillation frequency. With the architecture previously described, the total single-ended injected current is equal to 4.7mA. This theoretical results has been validated in simulations, which show a locking for frequency differences of less than 1250MHz between two oscillators.

The locking range can be extended to other positions on the line:

$$\Delta\omega_{lock} = \frac{\omega_0}{2Q}\sin^2(2\pi\frac{z}{\lambda}) \tag{5.35}$$

with z being the position on the line and referenced at 0 at the edge. In fact the ratio between the injected signal and the oscillator signal at the center has a $\sin^2$ behavior due to the sinusoidal distribution of the amplitude and the coupling strength. That leads to a locking range that is dependent of the position as stated in Table 5.5. As expected, the coupling is stronger in the middle of the line and very weak at the edges. When compared to the frequency deviations due to PVT variations on a wafer, we are confident in being able to synchronize many standing-wave oscillators on the same wafer, although extensive simulations at a larger scale are needed to confirm this point.

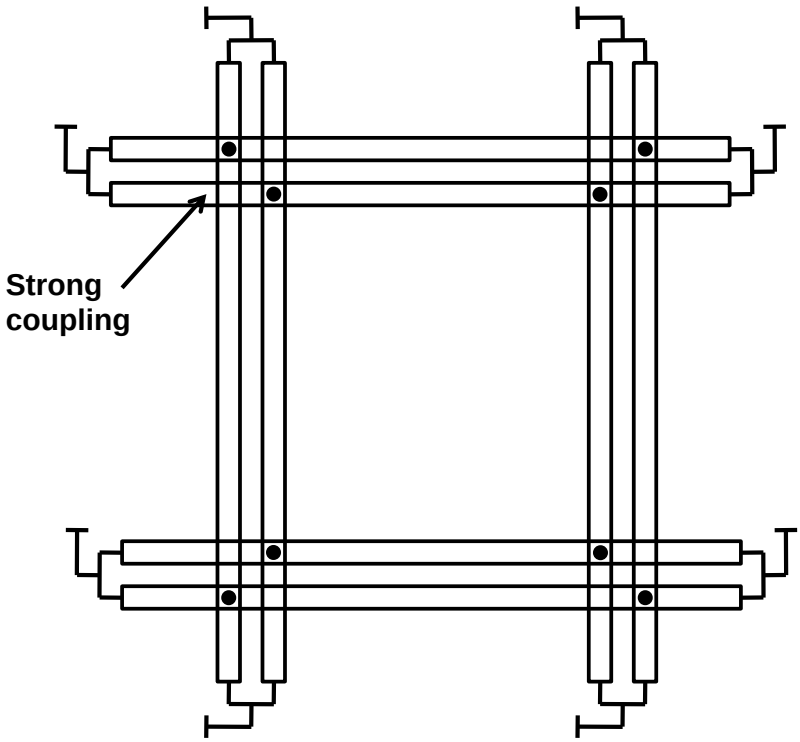


Figure 5.19: Example of coupling between 4 standing-wave oscillators.

5.4.7 Antenna co-design

For a good integration of the transceivers with the distribution and synchronization scheme, the clock distribution pattern must be co-designed with the antenna pattern. A typical slot antenna is represented in Fig. 5.20 and will create a large distributed array. The pitch of the array would be λantenna/2 and λantenna is defined by:

$$\lambda_{antenna} = \frac{c}{90GHz} \tag{5.36}$$

On the other hand, the length of the standing-wave transmission line will be λline/2 with λline being roughly equal to:

$$\lambda_{line} = \frac{c}{10GHz \times \sqrt{\varepsilon_r}} \tag{5.37}$$

This leads to a ratio between λline and λantenna of 4.5 to 5 for typical values of ε_4. This is true for a non-loaded line and as compensation cells are distributed on the line, then the line length would be shorter and the ratio would be roughly equal to 4. Thus, the easiest way to fit the clock distribution pattern with this antenna pattern is to follow the perimeter of the antenna.

Table 5.5: Locking range according to position on the line

Position	Center (z=λ/4)	z=3λ/16	z=λ/8	z=λ/16
Locking range	1.25GHz	1.06GHz	625MHz	180MHz

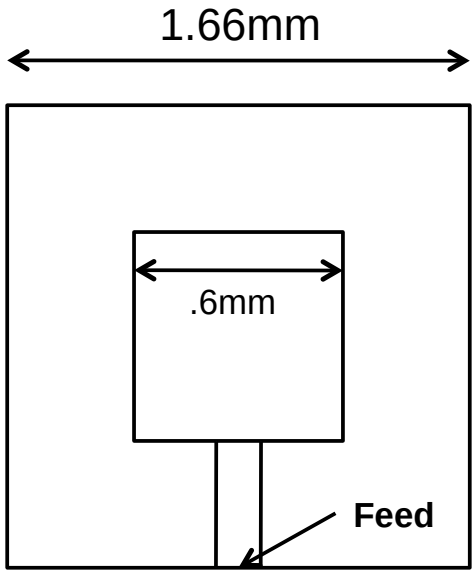


Figure 5.20: 90GHz slot antenna layout pattern.

Fig. 5.21 shows a proposed implementation of the clock distribution pattern in which the transmission lines follow the perimeter of each antenna. In this particular implementation, it is notable that the clock is taken at the exact same position (λ/4 away from the edges) on all the lines and this helps reduce the skew between clusters. The particularity of this scheme is that the coupling with the neighbors are either strong when coupled near the center or weak when coupled near the edge. This results in two columns of clusters being strongly coupled together, but weakly coupled to the adjacent two columns. Once again, this must be modeled and simulated to determine if this is an issue for correctly locking all the oscilators together.

Moreover, the influence of the clock distribution on the antenna radiation pattern is hard to predict and advanced electromagnetic simulations are required to evaluate the interactions between them. At first sight, the transmission lines would act as a shield ring between two antenna clusters and would be beneficial for the antenna radiation pattern. But the fact that current flows along these lines could definitely affects the radiation pattern. This must be investigated to ensure a good co-integration.

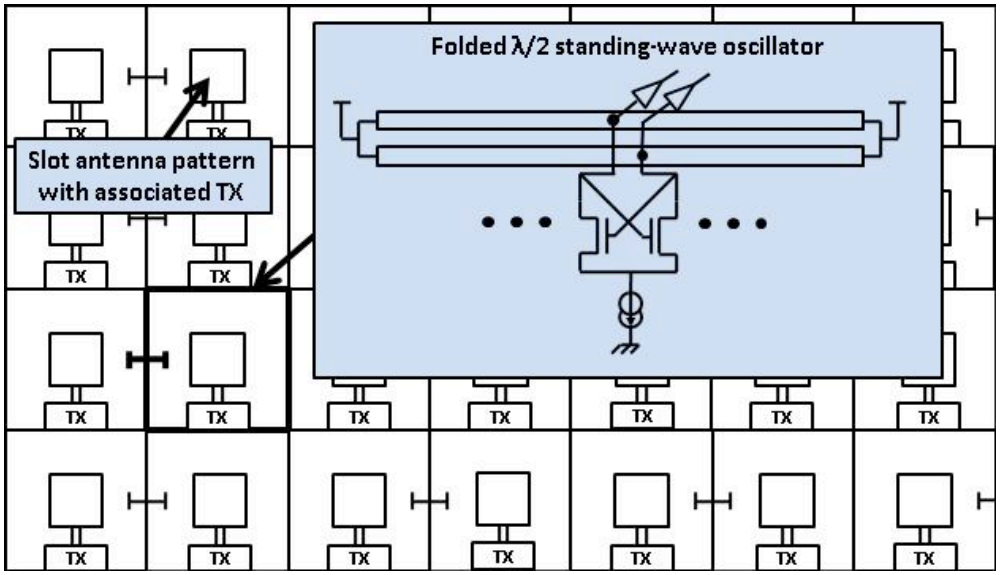


Figure 5.21: Proposed clock distribution pattern co-designed wit the antenna array.

The full simulation of coupled standing-wave oscillators using *RLCG* networks in Cadence is suitable for studying the coupling between up to 4 or 8 oscillators. If we would like to go further and study the locking behavior for a large-scale, other tools or methods may be used. This has been discussed with Prof. Jaijeet Roychowdhury's group at UC Berkeley and two approaches appear suitable. The first one is to use a reduction of the *RLCG* networks. That would help reduce the number of nodes and fasten the simulation. Thus, we would be able to simulate more elements using the same testbench in Cadence. But this does not solve the problem of very large-scale systems. The second and more robust solution is to define a macromodel from the behavior and characteristics of a single standing-wave oscillator and to model the coupling between two of them. We would then be able to use a higher-level simulator to understand the locking issues of a particular pattern and go deeper into the co-integration of the antenna clusters with the clock distribution scheme. We plan to collaborate with Prof. Jaijeet Roychowdhury's group to set up an appropriate behavioral model for the large-scale simulation of coupled standing-wave oscillators.

5.5 Conclusions

For synchronization of transceivers on a whole wafer, a clock distribution architecture based on coupled standing-wave oscillators has been demonstrated to be suitable for achieving both low clock skew between clusters and low power consumption. Based on a simple transmission line model, simulations of standing-wave oscillators have been performed to evaluate their performance and a design methodology has been defined for sizing the main parameters, according to

the desired performance.

Study of the influence of PVT variations on standing-wave oscillators has shown a good tolerance to variations. The main influence comes from the metal thickness and dielectric height variations. Worst-case deviations are low for frequency and clock skew (respectively 100MHz frequency deviation for a 10GHz clock and 2ps clock skew from one transmission line to another). Signal swing varies more with PVT variations and means to control the current drawn by the compensation cells are required. The clock buffer study has shown that this block is very sensitive to any type of variations and a DLL-based buffer needs to be designed.

Coupling study has demonstrated that the coupling strength is very strong when oscillators are coupled at the center and weaker when coupled near the edges. Nevertheless, the locking range seems to be sufficient enough to obtain a good locking behavior of many standing-wave oscillators. Co-design with the antenna array is mandatory to achieve a good integration of the global structure and a possible pattern has been proposed. The simulation results obtained in this study as well as the co-integration with the antenna need to be confirmed with a real chip design at a small-scale, in which realistic models for the lines would be considered and radiation pattern perturbations would be addressed.

Furthermore, this work has enabled collaborations on large-scale simulations of coupled standing-wave oscillators with Prof. Jaijeet Roychowdhury's group. Numerical simulation tools would help understand the coupling and locking behavior of this type of distributed oscillators on a large scale and would allow the design of an optimized clock distribution architecture for the distributed radio on a whole wafer.

Bibliography

[1] http://en.wikipedia.org/wiki/Babinet's_principle.

[2] ABDEL-AZIZ, M., GHALI, H., RAGAIE, H., HADDARA, H., LARIQUE, E., GUILLON, B., AND PONS, P. Design, implementation and measurement of 26.6 GHz patch antenna using MEMS technology. In *IEEE Antennas and Propagation Society International Symposium, 2003* (2003), vol. 1.

[3] ADABI, E., AND NIKNEJAD, A. Broadband variable passive delay elements based on an inductance multiplication technique. In *IEEE Radio Frequency Integrated Circuits Symposium, 2008. RFIC 2008* (2008), pp. 445–448.

[4] AHAMDI, M., AND SAFAVI-NAEINI, S. On-chip Antennas for 24, 60, and 77GHz Single Package Transceivers on Low Resistivity Silicon Substrate.

[5] ALEXOPOULOS, N., KATEHI, P., AND RUTLEDGE, D. Substrate optimization for integrated circuit antennas. In *Microwave Symposium Digest, MTT-S International* (1982), vol. 82, pp. 190–192.

[6] ANDRESS, W., AND HAM, D. Recent developments in standing-wave oscillator design: review. In *2004 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium, 2004. Digest of Papers* (2004), pp. 119–122.

[7] APPANNAGARRI, N., BARDI, I., EDLINGER, R., MANGES, J., VOGEL, M., CENDES, Z., AND HADDEN, J. Modeling phased array antennas in Ansoft HFSS. In *2000 IEEE International Conference on Phased Array Systems and Technology, 2000. Proceedings* (2000), pp. 323–326.

[8] AY, S., AND FOSSUM, E. A 76 x 77mm 2, 16.85 Million Pixel CMOS APS Image Sensor. In *VLSI Circuits, 2006. Digest of Technical Papers. 2006 Symposium on* (2006), pp. 19–20.

[9] BABAKHANI, A., GUAN, X., KOMIJANI, A., NATARAJAN, A., AND HAJIMIRI, A. A 77-GHz phased-array transceiver with on-chip antennas in silicon: Receiver and antennas. *IEEE Journal of Solid State Circuits 41*, 12 (2006), 2795.

[10] BALANIS, C. *Antenna Theory*. John Wiley & Sons, 2005.

[11] BARAKAT, M., NDAGIJIMANA, F., AND DELAVEAUD, C. On the design of 60 GHz integrated antennas on 0.13¿ m SOI technology. In *2007 IEEE International SOI Conference* (2007), pp. 117–118.

[12] BIJUMON, P., ANTAR, Y., FREUNDORFER, A., AND SAYER, M. Integrated dielectric resonator antennas for system on-chip applications. In *Microelectronics, 2007. ICM 2007. Internatonal Conference on* (2007), pp. 275–278.

[13] CHAN, K., CHIN, A., LIN, Y., CHANG, C., ZHU, C., LI, M., KWONG, D., MCALISTER, S., DUH, D., AND LIN, W. Integrated antennas on Si with over 100 GHz performance, fabricated using an optimized proton implantation process. *IEEE Microwave and Wireless Components Letters 13*, 11 (2003), 487–489.

[14] CHIEN, J., AND LU, L. Analysis and Design of Wideband Injection-Locked Ring Oscillators With Multiple-Input Injection. *IEEE Journal of Solid-State Circuits 42*, 9 (2007), 1906–1915.

[15] CHOWDHURY, D., REYNAERT, P., AND NIKNEJAD, A. A 60GHz 1V+ 12.3 dBm transformer-coupled wideband PA in 90nm CMOS. In *IEEE International Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers* (2008), pp. 560–635.

[16] EISENSTADT, W., AND EO, Y. S-parameter-based IC interconnect transmission linecharacterization. *IEEE Transactions on components, hybrids, and manufacturing technology 15*, 4 (1992), 483–490.

[17] ELDEK, A., ELSHERBENI, A., AND SMITH, C. Characteristics of bow-tie slot antenna with tapered tuning stubs for wideband operation. *Progress Electromagn Res (PIER) 49* (2004), 53–69.

[18] ELDEK, A., ELSHERBENI, A., SMITH, C., AND LEE, K. Wideband rectangular slot antenna for personal wireless communication systems. *IEEE Antennas and Propagation Magazine 44*, 5 (2002), 146–155.

[19] GEANNOPOULOS, G., AND DAI, X. An adaptive digital deskewing circuit for clock distributionnetworks. In *1998 IEEE International Solid-State Circuits Conference, 1998. Digest of Technical Papers* (1998), pp. 400–401.

[20] GUTNIK, V. *Analysis and characterization of random skew and jitter in a novel clock network*. PhD thesis, Massachusetts Institute of Technology, 2000.

[21] GUTNIK, V., AND CHANDRAKASAN, A. Active GHz clock network using distributed PLLs. *IEEE Journal of Solid-State Circuits 35*, 11 (2000), 1553–1560.

[22] HALL, L., CLEMENTS, M., LIU, W., AND BILBRO, G. Clock distribution using cooperative ring oscillators. In *Advanced Research in VLSI, 1997. Proceedings., Seventeenth Conference on* (1997), pp. 62–75.

[23] HURWITZ, J., PANAGHISTON, M., FINDLATER, K., HENDERSON, R., BAILEY, T., HOLMES, A., AND PAISLEY, B. A 35 mm film format CMOS image sensor for camera-back applications. In *2002 IEEE International Solid-State Circuits Conference, 2002. Digest of Technical Papers. ISSCC* (2002), vol. 1.

[24] LIANG, J., GUO, L., CHIAU, C., CHEN, X., AND PARINI, C. Study of CPW-fed circular disc monopole antenna for ultra wideband applications. *IEE Proceedings-Microwaves, Antennas and Propagation* (2005), 520–526.

[25] MANOLAKIS, D., INGLE, V., AND KOGON, S. *Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing*. McGraw-Hill, 2000.

[26] MARCU, C., AND NIKNEJAD, A. A 60 GHz high-Q tapered transmission line resonator in 90nm CMOS. In *2008 IEEE MTT-S International Microwave Symposium Digest* (2008), pp. 775–778.

[27] MCLEAN, J. A re-examination of the fundamental limits on the radiation Q ofelectrically small antennas. *IEEE Transactions on Antennas and Propagation 44*, 5 (1996).

[28] MONTUSCLAT, S., GIANESELLO, F., AND GLORIA, D. Silicon full integrated LNA, filter and antenna system beyond 40 GHz for MMW wireless communication links in advanced CMOS technologies. In *2006 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium* (2006), p. 4.

[29] O'MAHONY, F., YUE, C., HOROWITZ, M., AND WONG, S. A 10-GHz global clock distribution using coupled standing-wave oscillators. *IEEE Journal of Solid-State Circuits 38*, 11 (2003), 1813–1820.

[30] PAN, B., YOON, Y., PONCHAK, G., ALLEN, M., PAPAPOLYMEROU, J., AND TENTZERIS, M. Analysis and characterization of a high performance Ka-band surface micromachined elevated patch antenna. *micron 1*, 2j.

[31] POZAR, D. Considerations for millimeter wave printed antennas. *IEEE Transactions on antennas and propagation 31*, 5 (1983), 740–747.

[32] POZAR, D. Analysis of finite phased arrays of printed dipoles. *IEEE Transactions on Antennas and Propagation 33*, 10 (1985), 1045–1053.

[33] POZAR, D. The active element pattern. *IEEE Transactions on Antennas and Propagation 42*, 8 (1994), 1176–1178.

[34] PRODANOV, V., AND BANU, M. GHz Serial Passive Clock Distribution in VLSI Using Bidirectional Signaling. In *IEEE Custom Integrated Circuits Conference, 2006. CICC'06* (2006), pp. 285–288.

[35] RABAEY, J., CHANDRAKASAN, A., AND NIKOLIĆ, B. *Digital integrated circuits: a design perspective*. Prentice hall New Jersey, 1996.

[36] RAMAN, S., AND REBEIZ, G. Single-and dual-polarized millimeter-wave slot-ring antennas. *IEEE Transactions on Antennas and Propagation 44*, 11 (1996), 1438–1444.

[37] SANDIREDDY, S., JIANG, T., DEV, P., INC, A., AND BOISE, I. Advanced wafer thinning technologies to enable multichip packages. In *2005 IEEE Workshop on Microelectronics and Electron Devices, 2005. WMED'05* (2005), pp. 24–27.

[38] SHAMIM, A., POPPLEWELL, P., KARAM, V., ROY, L., ROGERS, J., AND PLETT, C. 5.2 GHz On-Chip Antenna/Inductor for Short Range Wireless Communication Applications. In *2006 IEEE International Workshop on Antenna Technology Small Antennas and Novel Metamaterials* (2006), pp. 213–216.

[39] SHIREEN, R., SHI, S., AND PRATHER, D. Lens coupled bow-tie antenna for millimeter wave operations. In *2007 IEEE Antennas and Propagation Society International Symposium* (2007), pp. 5055–5058.

[40] STRANG, G. *Linear Algebra and Its Applications*. Thomson Learning, 1988.

[41] TREES, H. V. *Optimum Array Processing*. Wiley-Interscience, 2002.

[42] WELLER, T., KATEHI, L., AND REBEIZ, G. Single and double folded-slot antennas on semi-infinite substrates. *IEEE Transactions on Antennas and Propagation 43*, 12 (1995), 1423–1428.

[43] WOOD, J., EDWARDS, T., AND LIPA, S. Rotary traveling-wave oscillator arrays: A new clock technology. *IEEE Journal of Solid-State Circuits 36*, 11 (2001), 1654–1665.