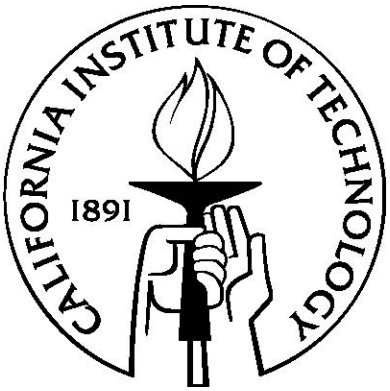
DIRECT ANTENNA MODULATION (DAM) FOR ON-CHIP MM-WAVE TRANSCEIVERS

Thesis by

Aydin Babakhani

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy



California Institute of Technology
Pasadena, California

2008

(Defended May 28, 2008)

To My Parents

Acknowledgement

First and foremost, I would like to thank my thesis supervisor, Professor Ali Hajimiri, for being the best mentor that anyone could possibly ask for. I have always been amazed with the time and effort that he spends on training students. More than the technical advice he gave, his guidance in matters beyond the academic has been and will continue to be vital to any success that I have achieved at caltech and beyond. In addition, I have continually found his enthusiasm and demeanor to be a source of inspiration. His influence in my life makes it impossible for me to adequately express my respect and gratitude in words.

I also wish to express my deepest gratitude to my co-advisor, Professor David Rutledge. Despite his ever-increasingly busy schedule, he always made the time to give me invaluable technical advice that proved extremely useful to my research. In addition to the academic guidance he provided, I wish to thank him for the enormous positive impact that he has had in my life.

I am also grateful to Professor Sander Weinreb for his technical support and assistance over the course of my PhD. In particular his technical input during the phased array project proved instrumental to its success and as a result, this thesis as well.

I would like to thank Professor Ali Hajimiri, Professor David Rutledge, Professor Sander Weinreb, and Professor Babak Hassibi for serving on my candidacy and thesis committees. I also wish to thank Professor Jonas Zmuidzinas, and Dr. Charles Elachi for

serving on my candidacy committee, and Professor Azita Emami for serving on my thesis committee.

I am grateful to all members of Caltech High Speed Integrated Circuits and Caltech Millimeter-Wave IC groups. I am particularly indebted to Dr. Abbas Komijani, Professor Ehsan Afshari, Dr. Arun Natarajan, Dr. Xiang Guan, and Dr. Behnam Analui for their mentoring during the early part of my Ph.D. I would like to specially thank Yu-Jiu Wang for helping me in designing of the Direct Antenna Modulation project and Joseph Bardin for helping in the measurement of the phased-array chip. I wish to thank my colleagues Professor James Buckwalter, Professor Arjang Hassibi, Hua Wang, Florian Bohn, Edward Keehr, Juhwan Yoo, Sanggeun Jeon, Kaushik Sengupta, Jay Chen, and Steven Bowers for their support.

I also thank Michelle Chen, Jim Endrizzi, Dale Yee, Naveed Near-Ansari, John Lilley, Niklas Wadefalk, Hamdi Mani, Alireza Ghaffari, Linda Dozsa, Tanya Owen, Carol Sosnowski, Gary Waters, Lynn Hein, Janet Couch, Shirley Slattery, and Kent Potter for their valuable assistance during my time at Caltech.

I am very fortunate in my friends at Caltech- Professor Arash Yavari, Yan Chen, Shafigh Shirinfar, Dr. Shervin Taghavi, Dr. Arash Kheradvar, Dr. Maziar Motahari, Dr. Amir Sadjadpour, Dr. Farzad Parvaresh, Sormeh Shadbakht, and Ali Vakili.

Finally, I thank my parents and my brother for their love and patience.

Abstract

In the last few decades the puissant desire to miniaturize the digital circuits to achieve higher speed and larger density has shaped the evolution of the silicon-based technologies. This development opens a new era in the field of millimeter-wave electronics in which many low-cost high-yield silicon-based transistors can be used on a single chip to enable creation of novel architectures with unique properties not achievable with old processes. In addition to this high level of integration capability, the die size of comparable or even larger than the wave-length makes it possible to integrate antennas, transceivers, and digital circuitry all on a single silicon die.

It is important to realize that although smaller parasitic capacitors and shorter transistor channels have improved f_T and f_{max} of transistors, extremely thin metal layers, highly doped substrates, and low breakdown voltage transistors have severely affected the performance of analog and RF building blocks. For example, thin metal layers have increased the loss and lowered the quality factor of the building blocks such as capacitors and inductors and low breakdown voltage transistors have made the power generation quite challenging. Additionally, if not carefully designed, small wave-lengths in the millimeter-wave range may cause unintended radiation by on-chip components. In this dissertation, we address these issues in design of millimeter-wave silicon-based single-chip phased-array transceivers with integrated antennas. We also introduce the technique of Direct Antenna Modulation (DAM) and implement two proof-of-concept chips operating at 60 GHz.

We will present the receiver and the on-chip antenna sections of a fully integrated 77 GHz four-element phased-array transceiver with on-chip antennas in silicon. The receiver section of the chip includes the complete down-conversion path comprising low-noise amplifier (LNA), frequency synthesizer, phase rotators, combining amplifiers, and on-chip dipole antennas. The signal combining is performed using a novel distributed active combining amplifier at an IF of 26 GHz. In the LO path, the output of the 52 GHz VCO is routed to different elements and can be phase shifted locally by the phase rotators. A silicon lens on the backside is used to reduce the loss due to the surface-wave power of the silicon substrate. Our measurements show a single-element LNA gain of 23 dB and a noise figure of 6.0 dB. Each of the four receive paths has a gain of 37 dB and a noise figure of 8.0 dB. Each on-chip antenna has a gain of +8 dBi.

A direct antenna modulation (DAM) technique is also introduced, where the radiated far-field signal is modulated by time-varying changes in the antenna near-field electromagnetic (EM) boundary conditions. This enables the transmitter to send data in a direction-dependent fashion producing a secure communication link. The transmitter architecture makes it possible to use narrow-band highly-efficient switching power amplifiers to transmit wideband constant and non-constant envelope modulated signals. Theoretically, these systems are capable of transmitting independent data in multiple directions at full-rate concurrently using a single transmitter. Direct antenna modulation (DAM) can be performed by using either switches or varactors. Two proof-of-concept DAM transmitters operating at 60GHz using switches and varactors are demonstrated in silicon proving the feasibility of this approach.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"*Imagination is more important than knowledge*."

"*The world we have made as a result of the level of thinking we have done thus far creates problems we cannot solve at the same level of thinking at which we created them*."

-Albert Einstein (1879–1955)

Are today's advancements in the field of wireless communication a result of human knowledge or imagination? With remarkable improvements in computational power and the existence of sophisticated design tools, brute force, as opposed to creative approaches appears to have become the weapon of choice in overcoming challenging problems in the field of wireless communication. The prodigious market for digital consumer electronics which is the main driving force behind the advancements in process technologies has resulted in transistors with higher speeds and larger densities. The same puissant force has significantly increased available computational power, making conventional circuit design more efficient and reliable through the use of new tools which exploit this computational power.

During my five years of graduate studies at Caltech, I have been a witness to this rapid evolution of process technology. In 2003 (when I started my graduate studies) a common process technology available to research groups in academia was 0.18 μm CMOS, while

today, most of them have access to 65 nm CMOS process. A factor of three improvement in size of the transistors during these five years happened due to the enormous global investment in silicon-based technologies making faster transistors, which we as circuit designers had nothing to do with. During the same period of time, design software, electromagnetic simulators, and computational engines became faster and faster (a typical personal computer in 2003 was a Pentium 4 with 500MB of RAM while today's typical PC is a dual-core with 4GB of RAM). It seems that, sitting in a well-equipped research group and just repeating existing designs in faster processes, it is possible to improve the specifications of wireless blocks and publish new papers. At first glance, it seems that this magical sheer force of transistor scaling will do everything for us, there is no need for creation or imagination, and we just need to rely on our experience or knowledge. But this is not true in reality.

First of all, the evolution of process technology has not always been beneficial to analog and RF designers. Although smaller parasitic capacitors and shorter transistor channels have improved f_T and f_{max} of the transistors, extremely thin metal layers, highly doped substrates, and low breakdown voltage transistors have severely affected the performance of analog and RF building blocks. For example, thin metal layers have increased the loss and lowered the quality factor of the basic circuit elements such as capacitors and inductors. Low breakdown voltage transistors have made power generation and stacking of transistors extremely challenging.

Second, having access to so many high-frequency transistors has opened the door to a plethora of new applications not even accessible to compound transistors due to limited

yield. The ability to integrate many mm-wave transistors with high yield on a single chip provides a new frontier for integrating very complicated systems with previously unimaginable levels of complexity, including fully integrated single-chip phased arrays. Single-chip phased-array systems, which are discussed in Chapters 4 and 6, impose new design challenges, including accurate on-chip phase generation and distribution, as well as finding suitable phased-array architecture for specific applications (for example, RF phase-shift versus LO phase-shift). Furthermore, tuning and calibration capability achieved by co-integration of digital circuitry can be used to significantly improve the performance of the critical analog/RF elements. Also the integration of advanced base-band and signal processing elements enables the realization of novel system architectures for existing and emerging mm-wave applications.

Thirdly and most importantly, the rules of the wireless communications game are changing. The age in which analog designers could remain ignorant of the electromagnetic and antenna design is gone. At low frequencies (few GHz), RF input and output of the transceivers can be connected to on-chip pads and on-chip wafer probing can be done to complete the testing of a transceiver. At low frequencies, analog designers are not responsible for designing the antennas because the PCB-based standard antennas can be used. In this case wirebond or flip-chip technology can be utilized to transfer the on-chip RF signal to the PCB-based antenna. On the contrary, in mm-wave frequencies, analog designers need to completely understand the details of the antenna design. At these frequencies, a 1 mm long wirebond which has about 1 nH of inductance introduces a large mismatch between PCB-based antennas and on-chip pads (1 nH at 60 GHz translates to 377 Ω imaginary impedance). A 1 mm wirebond at 60GHz can become an

undesired radiating element as well. As wavelengths become smaller at higher frequencies it is possible to physically fit a half-wavelength antenna on a reasonably sized silicon die. Unfortunately, however, placing an antenna on a silicon die does not come without its own set of challenges. In order to implement efficient on-chip antennas at these frequencies, these challenges must be addressed. For instance, at mm-wave frequencies, the high dielectric constant of silicon (ε_r= 11.7), and substrate thickness comparable to the wavelength couple most of the antenna radiated power into the silicon substrate in the form of substrate modes which will be discussed in Chapter 2. In addition, the inevitable high doping levels of silicon substrates used to implement active devices result in a lossy substrate (resistivity of about 10 Ω-cm) and wastes most of the substrate coupled power as heat. Thus, as we move into this new frontier, everything is getting more complicated and it is the responsibility of the analog designer to take care of the details of antenna design as well. The marriage of analog/RF and antenna design is all but inevitable.

Finally, all of these new capabilities make it possible to fundamentally change the architectures of wireless systems. Almost everyone in this field knows that with his invention of the super-heterodyne receiver in 1918, Armstrong introduced the idea of modulating the signal at low frequencies, or baseband, and up-converting it to the RF frequency. Since that time, there have been many breakthroughs in related technologies, including the invention of the transistor itself in 1947. However, there have been few fundamental changes in transceiver architectures; most of today's high performance systems still use ideas based on the heterodyne or homodyne architectures. In Chapter 5, we will see that by implementing on-chip antennas and reflectors, as well as digitally

controlling the scattering properties of these reflectors, we can directly modulate the antennas and implement a wireless transmitter with unique desirable characteristics not possessed by conventional (heterodyne or homodyne) schemes.

To conclude the introduction, it seems that at this state of technology revolution we are only limited by our own creativity and imagination.

1.1 Organization

In Chapter 2, the fundamental problems involved in designing on-chip antennas are discussed. Substrate-modes as an important concept in designing the on-chip antennas is reviewed and an optical representation of them is introduced. In this chapter several on-chip antenna configurations are proposed and their pros and cons are compared.

Chapter 3 covers the main applications of the mm-wave systems including communication, radar, and imaging.

In Chapter 4 the fundamentals of the phased-array systems are discussed. The first fully integrated silicon-based millimeter-wave (77 GHz) phased-array transceiver with on-chip antennas is introduced, and details of the design and measurement of this chip are included.

In Chapter 5, a fundamentally new technique for implementing wireless communication systems, Direct Antenna Modulation (DAM), is introduced. The unique characteristics of this new system are discussed and two proof-of-concept chips operating at 60 GHz are designed and measured.

Chapter 6 covers the design and measurement of a silicon-based scalable 2×2 60-GHz phased-array transmitter with on-chip antennas.

Chapter 2

On-Chip Antennas

2.1 Introduction

An antenna converts electrical power in the circuit domain to electromagnetic wave radiations in a propagation medium and vice versa. The radiated energy appears as loss if looked at from a pure circuit domain perspective and is thus modeled as a resistance[1]. In addition to this so-called "radiation resistance" which is essential to the antenna operation, a second resistive part is required to model the physical energy loss in the non-ideal metals and the dielectrics. For an antenna excited with a current source, loss and radiated power can be calculated as,

$$P_{rad} = \frac{R_{rad} I^2}{2}$$
$$P_{loss} = \frac{R_{loss} I^2}{2} \qquad (2.1)$$

where P_{rad} is the radiated power, P_{loss} is the lost power, R_{rad} is the radiation resistance, R_{loss} is the loss resistance, and I is the antenna current. Obviously, high loss resistance wastes power and lowers the overall efficiency. In fact, radiation efficiency is directly related to the ratio of loss and radiation resistances. By knowing these two values, radiation efficiency can be calculated as,

[1] Often a reactive part is also used to account for the resulting phase difference between antenna's voltage and current.

$$\eta_{rad} = \frac{R_{rad}}{R_{rad} + R_{loss}} \tag{2.2}$$

We will focus on important antenna parameters such as gain and efficiency and compare several antenna configurations suitable for silicon on-chip implementation based on these parameters.

One of the important concepts involved in designing silicon-based on-chip antennas is the substrate modes. Due to silicon's high dielectric constant and die's comparable thickness to the wavelength at millimeter wave frequencies; most of the electromagnetic power will be coupled to the substrate modes in unshielded structures.

Concept of the substrate modes has been studied extensively in literature and is public domain material. Kogelnik [1] and Rutledge [2] have done extensive research in this area. Section 2.2 uses their approach in explaining the concept.

2.2 Substrate Modes

A complete analysis of substrate modes in planar substrates requires the use of Summerfield integrals and asymptotic methods. This section does not cover these advanced topics but utilizes the optical theory of dielectric waveguides to explain how substrate modes are generated and how to calculate the amount of power coupled to these modes.

One of the important parameters used in analyzing the substrate modes is the effective guide thickness. Effective guide thickness depends on the mode of propagation and can

be used to calculate the power coupled to each mode. To calculate the effective guide thickness let us start with reviewing the propagation of an optical wave in a simple planar dielectric waveguide with a cross section shown in Figure 2.1.

Figure 2.1 Dielectric Waveguide [1]

We assume that dielectric 1 and 3 are both air and dielectric 2 is silicon. Reflection and transmission coefficients can be calculated separately for each interface. An interface of two lossless, isotropic, homogenous dielectric media of refractive index n_1 and n_2 is shown in Figure 2.2.

From Snell's law,

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 . \tag{2.3}$$

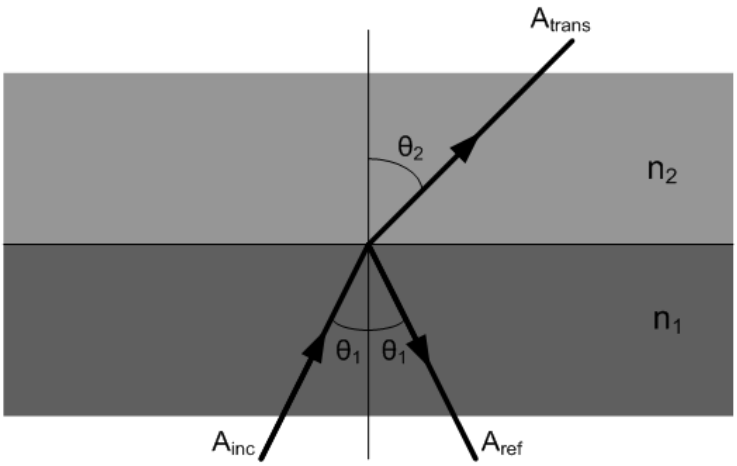


Figure 2.2 Ray optics of reflection and transmission

The incident, transmitted, and reflected waves have complex amplitude of A_{inc}, A_{ref}, and A_{trans}, as shown in Figure 2.2. Assuming a complex reflection coefficient of R and transmission coefficient of Γ, we have the following equations:

$$A_{ref} = R \cdot A_{inc} \tag{2.4}$$

$$A_{trans} = \Gamma \cdot A_{inc} \ . \tag{2.5}$$

The reflection and transmission coefficients depend on the incident angle and the polarization of the light. These coefficients can be calculated by applying boundary conditions for the tangential components of electric field and magnetic field of an incident planar wave. For the transverse electric (TE) polarization in which electric field is perpendicular to the direction of the incident wave (wave normal) and normal to the interface, the reflection coefficient is given by:

$$R_{TE} = \frac{(n_1 \cos\theta_1 - n_2 \cos\theta_2)}{(n_1 \cos\theta_1 + n_2 \cos\theta_2)} = \frac{(n_1 \cos\theta_1 - \sqrt{n_2^2 - n_1^2 \sin^2\theta_1})}{(n_1 \cos\theta_1 + \sqrt{n_2^2 - n_1^2 \sin^2\theta_1})} \ . \tag{2.6}$$

For the transverse magnetic (TM) polarization, the magnetic field is perpendicular to the direction of the incident wave (wave normal) and normal to the interface. In this case the reflection coefficient is given by:

$$R_{TM} = \frac{(n_2 \cos\theta_1 - n_1 \cos\theta_2)}{(n_2 \cos\theta_1 + n_1 \cos\theta_2)} = \frac{(n_2^2 \cos\theta_1 - n_1\sqrt{n_2^2 - n_1^2 \sin^2\theta_1})}{(n_2^2 \cos\theta_1 + n_1\sqrt{n_2^2 - n_1^2 \sin^2\theta_1})} \ . \tag{2.7}$$

Based on the equations (2.6) and (2.7), for incident angles smaller than the critical angle ($\theta_1 < \theta_c$) we have $|R| < 1$ and R is a real number. For incident angles greater than the critical angle ($\theta_1 > \theta_c$) we have $|R| = 1$ and R is a complex number. The critical angle, θ_c, is given by

$$\sin\theta_c = \frac{n_2}{n_1} \ . \tag{2.8}$$

For incident angles greater than the critical angle ($\theta_1 > \theta_c$), R is given by

$$R = e^{2j\phi} \tag{2.9}$$

where ϕ is different for TE and TM modes:

$$\tan\phi_{TE} = \frac{\sqrt{n_1^2 \sin^2\theta_1 - n_2^2}}{n_1 \cos\theta_1} \tag{2.10}$$

$$\tan\phi_{TM} = \frac{n_1^2}{n_2^2}\frac{\sqrt{n_1^2 \sin^2\theta_1 - n_2^2}}{n_1 \cos\theta_1} \ . \tag{2.11}$$

Figure 2.3 shows the ϕ_{TE} and ϕ_{TM} for the interface between air and silicon. In this case, $n_2 = 1$ and $n_1 = 11.7^{0.5} = 3.4$.

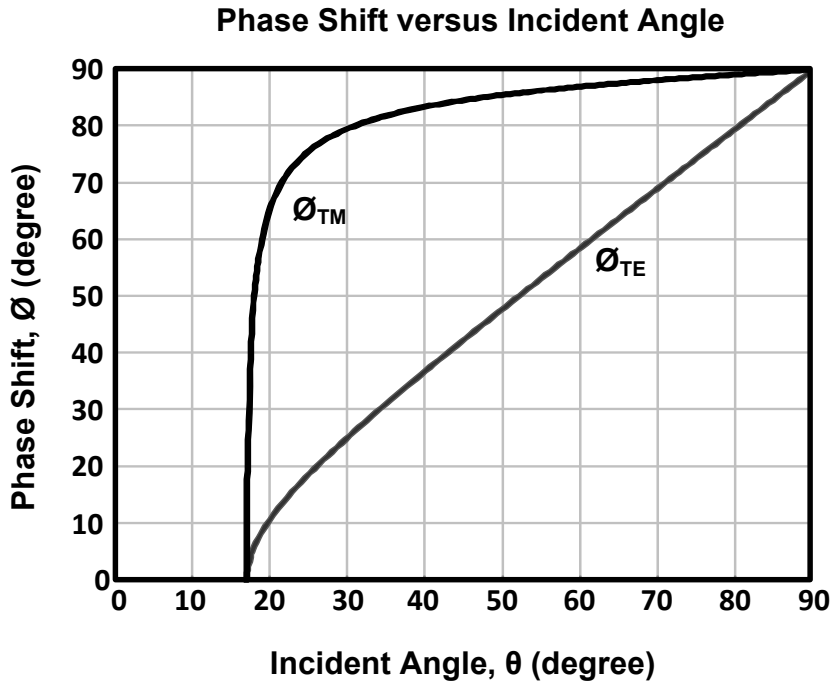


Figure 2.3 ϕ_{TE} and ϕ_{TM} versus incident angle ($n_1 = 3.4$ and $n_2 = 1$)

As is shown in Figure 2.3, the phase shift increases from 0 at the critical angle to 90° at the grazing incidence ($\theta_1 = 90°$). It increases with infinite slope at $\theta_1 = \theta_c$.

The above discussion helps us understand how a guided mode is created. Imagine the dielectric waveguide of Figure 2.4 in which we assume $n_2 > n_1, n_3$. In this case we will have two critical angles θ_{21}, θ_{23} and we will have a guided mode if $\theta_{inc} > \theta_{21}, \theta_{23}$. Guided waves will travel in zig-zag fashion, as shown in Figure 2.4. The fields of these guided waves vary as

$$\exp[-jkn_2(\pm x\cos\theta + z\sin\theta)] \tag{2.12}$$

where

$$k = \frac{2\pi}{\lambda} = \frac{\omega}{c} \ . \tag{2.13}$$

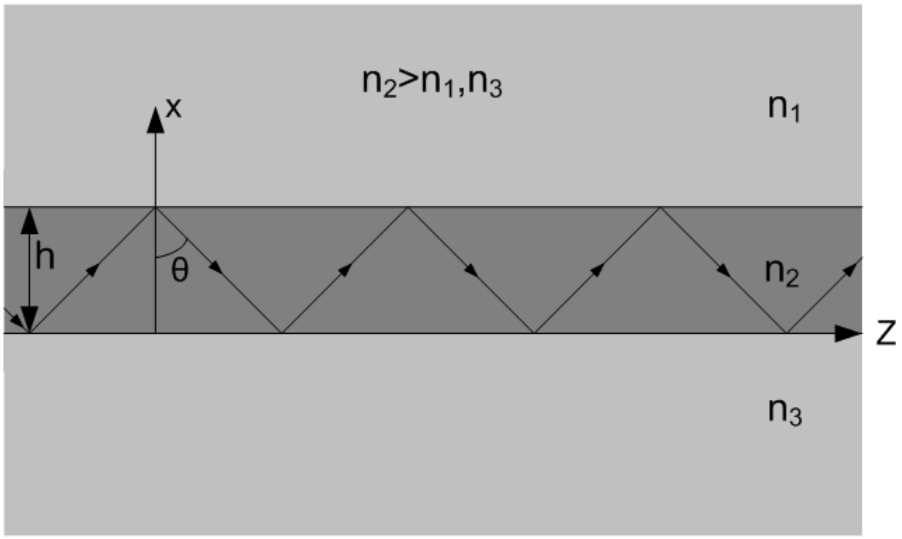


Figure 2.4 Guided wave

The effective propagation constant in the *Z*-direction is given by

$$\beta = \omega / v_p = k n_2 \sin\theta \ . \tag{2.14}$$

We realize that only discrete set of angles lead to a self-consistent picture that corresponds to what we call the "guided modes". For a fixed *z*, let us add up the phase shifts that occur as we move up from the lower film boundary ($x = 0$) with one wave to the other boundary ($x = h$) and then back down again with the reflected wave to where we started from. For guided modes, the sum of all these phase shifts must be a multiple of 2π. This means:

$$2kn_2 h\cos\theta - 2\phi_1 - 2\phi_3 = 2m\pi \tag{2.15}$$

where $-2\varphi_1$ and $-2\varphi_3$ are phase shifts due to the reflections from dielectrics 1 and 3 and *m* is the mode number.

The Goos-Hanchen Shift and the Effective Guide Thickness [1]

Here we compare the phase-shift due to the total reflection for two rays with z-component wave vectors of β-$\Delta\beta$ *and* β+$\Delta\beta$. In this case, we are dealing with total complex amplitude of $A_{inc,total}$ given by:

$$A_{inc,total} = A_{inc,\beta-\Delta\beta} + A_{inc,\beta-\Delta\beta} = e^{-j(\beta-\Delta\beta)z} + e^{-j(\beta+\Delta\beta)z} = 2\cos(\Delta\beta z)e^{-j\beta z} \tag{2.16}$$

and $A_{ref,total}$ can be calculated by (for small $\Delta\beta$)

$$A_{ref,total} = A_{ref,\beta-\Delta\beta} + A_{ref,\beta+\Delta\beta} = e^{-j((\beta-\Delta\beta)z-2\phi_{\beta-\Delta\beta})} + e^{-j((\beta+\Delta\beta)z-2\phi_{\beta+\Delta\beta})}$$

$$= (e^{j(\Delta\beta z-2\Delta\phi)} + e^{-j(\Delta\beta z-2\Delta\phi)})e^{-j(\beta z-2\phi)} = \cos(\Delta\beta(z-2z_1))e^{-j(\beta z-2\phi)} \tag{2.17}$$

where

$$z_1 = \frac{d\phi}{d\beta} \tag{2.18}$$

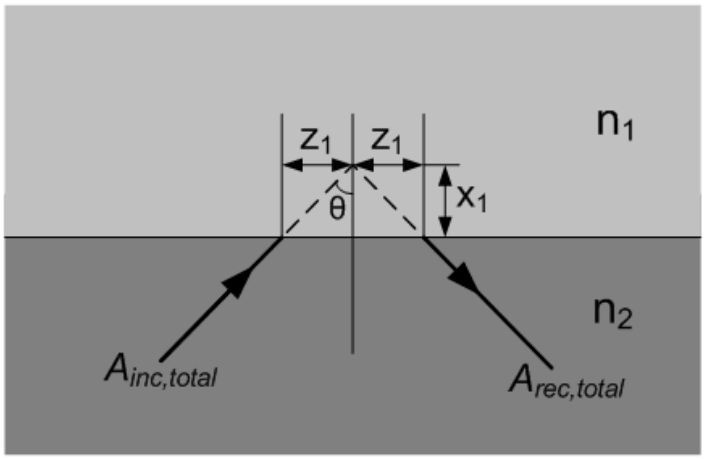


Figure 2.5 Lateral shift of the wave packet in reflection

where z_1 represents the lateral shift of the wave packet as shown in Figure 2.5. From equations (2.10), (2.11), and (2.14), we can derive the following expressions for the TE and TM modes [1]:

$$kz_1 = (N^2 - n_1^2)^{-1/2} \tan\theta \tag{2.19}$$

$$kz_1 = (N^2 - n_1^2)^{-1/2} \tan\theta / (\frac{N^2}{n_1^2} + \frac{N^2}{n_2^2} - 1) \tag{2.20}$$

where N is the "effective guide index" given by:

$$N = \beta / k = n_2 \sin\theta \ . \tag{2.21}$$

Based on Figure 2.5, later shift of z_1 is translated to the penetration depth of x_1 given by:

$$x_1 = z_1 / \tan\theta \ . \tag{2.22}$$

Based on the electromagnetic theory of waveguides, x_1 is closely related to the decay constant of the evanescent fields in the substrate [1].

Now let us study a configuration similar to Figure 2.6 in which $n_2 > n_1, n_3$. In this case a guided wave can be generated inside the dielectric 2. From Figure 2.6 we can calculate the effective guide thickness:

$$h_{eff} = h + x_1 + x_3 \ . \tag{2.23}$$

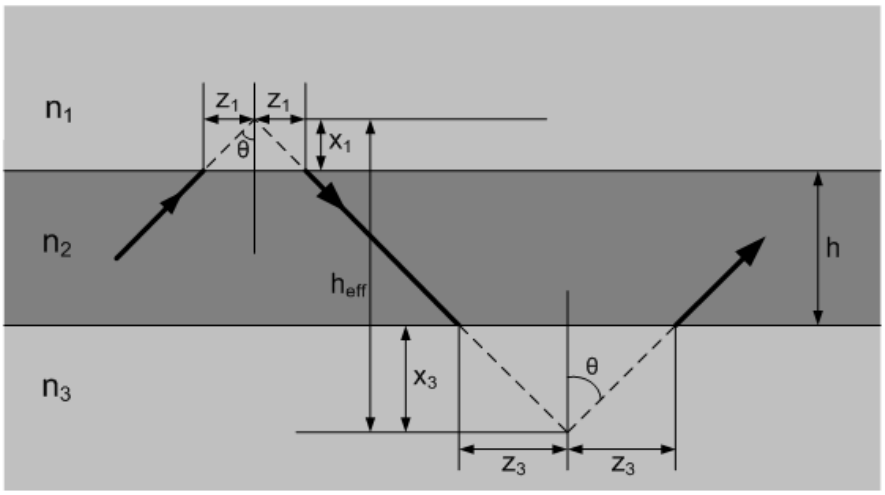


Figure 2.6 Three dielectrics n_1, n_2 and n_3 ($n_2 > n_1, n_3$)

If we assume the material associated with n_1 and n_3 is air and solve the complete electromagnetic problem inside the waveguide, we will have the following results [2]:

for TE modes: $$z_0 = \frac{1}{\alpha} \tag{2.24}$$

for TM modes: $$z_0 = (\frac{1}{\alpha})(\frac{\beta^2}{k_d^2} + \frac{\beta^2}{k_0^2} - 1)^{-1} \tag{2.25}$$

where z_0 is the effective penetration depth, β is the guide propagation constant, and

$$\alpha = \sqrt{\beta^2 - k_0^2} \tag{2.26}$$

$$k_0 = \frac{2\pi}{\lambda_0} \tag{2.27}$$

$$k_d = \frac{2\pi}{\lambda_d} \tag{2.28}$$

where λ_0 is the wavelength in air and λ_d is the wavelength in the dielectric.

For slabs without a ground plane on either side, we can write

$$h_{eff} = h + 2z_0 \tag{2.29}$$

where h is the physical thickness of the substrate. For slabs with a ground plane on one side, we can write

$$h_{eff} = h + z_0 \ . \tag{2.30}$$

And for slabs with ground planes on both sides,

$$h_{eff} = h \ . \tag{2.31}$$

TEM mode is the only special case in which we have [2]

$$h_{eff} = 2h \ . \tag{2.32}$$

We study the h_{eff} because we can use it to relate h_{eff} to the power per width in all of the above modes by the following equation [2],

$$P = \frac{EHh_{eff}}{4} \tag{2.33}$$

where E and H are the maximum transverse electrical and magnetic fields amplitudes. Figure 2.7 shows how h_{eff} varies for the different modes and substrate thicknesses when ε_r=4 [2]. For ungrounded substrate and grounded substrate h_{eff} has a minimum where the surface-wave power is largest. For very thin dielectric guides, h_{eff} becomes very large and surface-wave losses go to zero. For parallel-plate guides, h_{eff} goes to zero as the guide becomes thin, and the losses are large. The guide modes can be excited depending on ε_r, h, and whether or not there are ground planes on each side of the substrate.

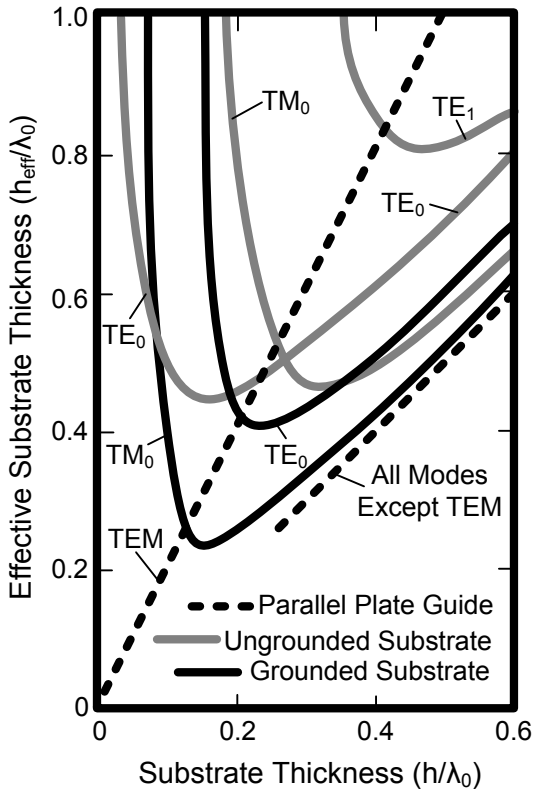


Figure 2.7 h_{eff} versus substrate thickness for different guide modes (ε_r = 4) [2]

Figure 2.7 also shows that by increasing the substrate thickness a greater number of substrate modes get excited. A silicon chip placed on a metallic substrate (the most common configuration) shapes a grounded substrate. As seen in Figure 2.7, for this configuration the TM_0 mode gets excited before the TE_0 mode. In the next section, we will review different possible solutions for implementing on-chip antennas in standard silicon processes and discuss their pros and cons.

2.3 On-Chip Antennas and Silicon Processes

In a standard silicon process the substrate resistivity is typically between 1 Ω.cm and 10 Ω.cm. The main reason for this low resistivity is the high level of doping used in these processes. Due to this low resistivity of the silicon substrate, most of the energy coupled to the substrate modes gets wasted as heat and reduces the overall antenna efficiency. The remaining part of the energy coupled to these substrates modes reaches the edge of the chip, couples to the air, and interferes with the main beam of the antenna. In a well-designed on-chip antenna we would like to minimize these substrate modes to increase the antenna efficiency and achieve a predictable antenna pattern. In the following section we compare different options in designing on-chip antennas and discuss their pros and cons.

A. Radiating from topside without ground shield

The most obvious choice for on-chip antennas is to implement them as metal lines on top of the substrate and radiate upward into the air. In this subsection, we show why this may not be an effective solution by looking at a dipole antenna placed at the boundary of semi-infinite regions of air and dielectric (Figure 2.8).

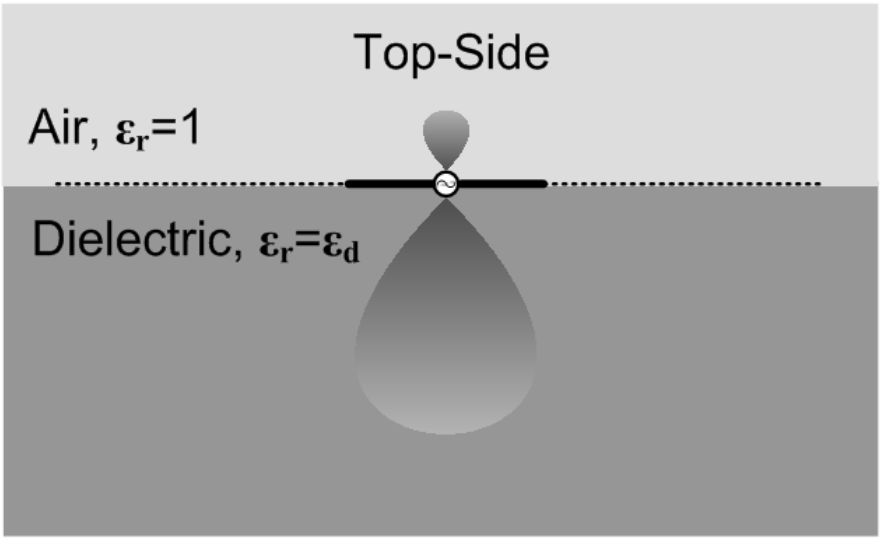


Figure 2.8 Radiating from top side without any ground shield

Although this over-simplified configuration does not correspond to the practical setting, it guides us to better understand the effects of silicon high dielectric constant on antenna radiation pattern and efficiency. For a dipole antenna seeing the vacuum (or air) on one side and a dielectric on the other side, the ratio of the power coupled into air to the total radiated power is approximated[2] by [2],

$$\frac{P_{air}}{P_{total}} = \frac{1}{\varepsilon_d^{3/2}} \tag{2.34}$$

[2] Within a factor of 2

where P_{air} is the radiated power into air, P_{total} is the total radiated power, and ε_d is the dielectric constant. From this formula for silicon dielectric ($\varepsilon_d \sim 11.7$) a very small portion of the power radiates into the air (about 3%) and the rest of it couples into silicon. This demonstrates that without any mechanism to reroute the power coupled into silicon substrate, it is not possible to implement a high-efficiency antenna on silicon this way.

B. Radiating from topside with on-chip ground shield

Another possible option is to incorporate an on-chip ground shield and try to reflect the radiated energy upward, thus preventing it from coupling into silicon, as shown in Figure 2.9.

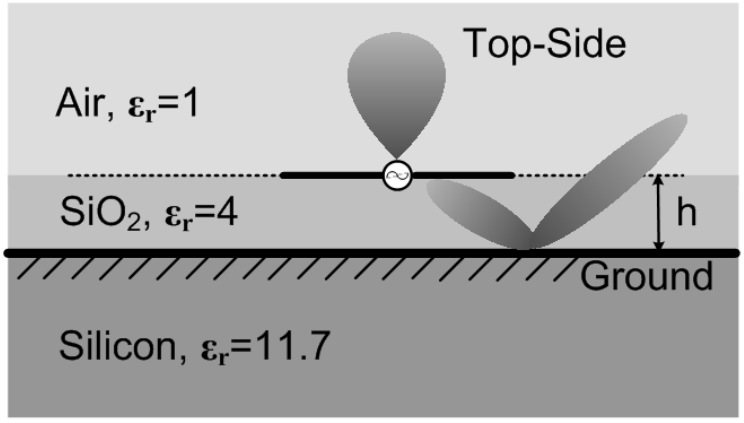


Figure 2.9 Radiating from top side with an on-chip ground shield

In the case of an on-chip ground shield, the on-chip antenna and the ground shield have to be placed inside the SiO_2 due to process limitations. For such a configuration the distance of the antenna and the ground shield affects the antenna-ground coupling and determines the radiation resistance. Unfortunately the distance between the bottom of the

top metal layer and the top of the lowest metal layer rarely exceeds 15 μm in today's process technologies. This small antenna-ground spacing causes a strong coupling between the antenna and the ground layer which lowers the all-important radiation resistance. Figure 2.10 shows the results of the electromagnetic simulations of a copper dipole antenna placed over a metal ground plane with a SiO_2 dielectric of thickness, *h*, sandwiched in between [3]. The dipole dimensions are 4 μm × 20 μm × 1150 μm and a moment-based EM simulator, IE3D [9], is used to perform the simulations. The dipole-length is equal to a length of a resonant dipole at 77 GHz which is placed in the boundary of semi-infinite regions of air and SiO_2. Based on this simulation, for a spacing of 15 μm between the antenna and the ground layer, the radiation resistance is very small (0.1 Ω) hence the total resistance is dominated by the ohmic loss of the copper, resulting in a radiation efficiency of less than 5%.

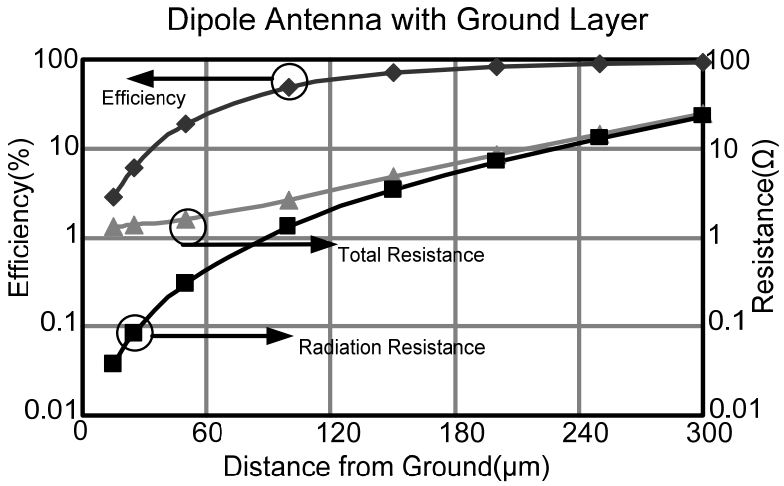


Figure 2.10 Dipole radiation resistance and efficiency

An option to increase the efficiency of the antenna seems to be the implementation of an off-chip ground shield to increase the distance between the antenna and its ground layer. We will discuss this case in the next subsection.

C. Radiating from topside with off-chip ground shield

As shown in Figure 2.11, an off-chip ground shield can be placed underneath the silicon substrate. In this case the silicon substrate thickness is much larger than the SiO_2 layer and effectively we are dealing with a high dielectric constant substrate ($\varepsilon = 11.7$). Unfortunately, because of the high dielectric constant of silicon and the large substrate thickness (100 μm or more) most of the power couples into substrate modes.

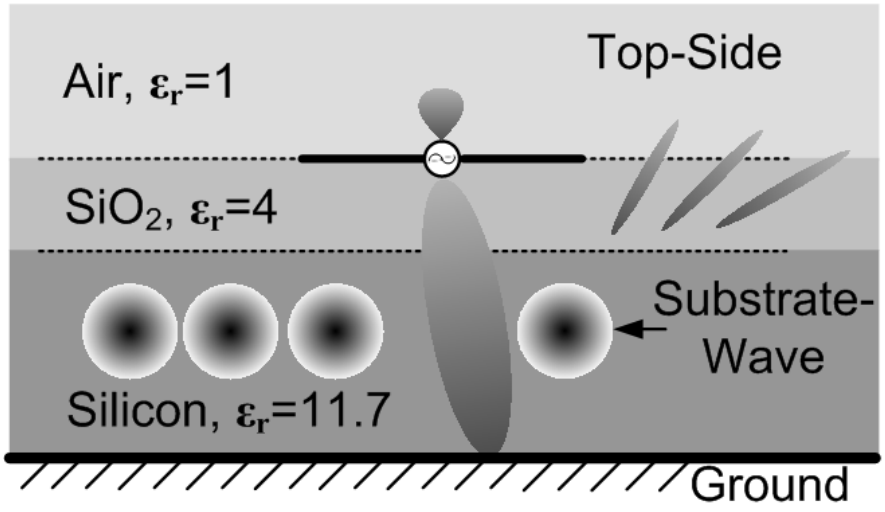


Figure 2.11 Radiating from top side with an off-chip ground shield

If we assume the thickness of SiO_2 is negligible compared to that of silicon, then the substrate modes' power can be numerically calculated. Based on these results, normalized radiated power and surface-wave power (substrate modes' power) are plotted

in Figure 2.12 and Figure 2.13, respectively. These quantities are normalized to dipole's free space radiated power given by [2]:

$$P_0 = \omega^2 \mu_0^{3/2} \varepsilon_0^{1/2} I^2 d^2 / 12\pi \tag{2.35}$$

where I is the current, ω is the angular frequency, and d is the effective length of the dipole. As is shown in Figure 2.12, at 77 GHz the maximum radiated power, which is around 1.3 P_0, occurs at the silicon substrate thickness of 290 μm (h = 0.075λ_0). However, at this substrate thickness, the power in all the surface wave modes is more than 3.5 P_0 (Figure 2.13), which indicates that even in the case of lossless silicon substrate, the power wasted in the surface wave modes is 2.7 times greater than the useful radiated power. It is important to realize that for a lossy and finite-dimensional substrate, the surface-wave power is either dissipated due to the substrate conductivity or radiated from the edge of the chip, and that often results in an undesirable radiation pattern.

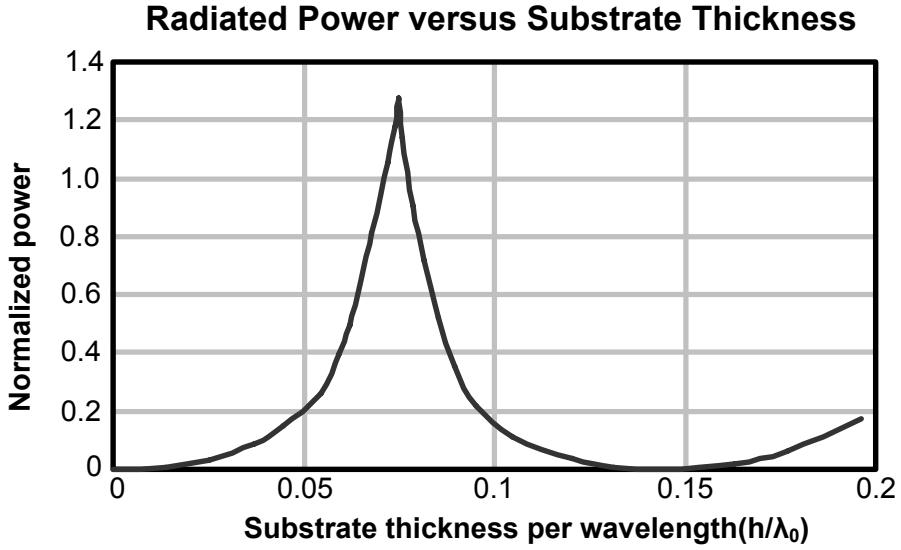


Figure 2.12 Normalized radiated power for a grounded dipole [2]

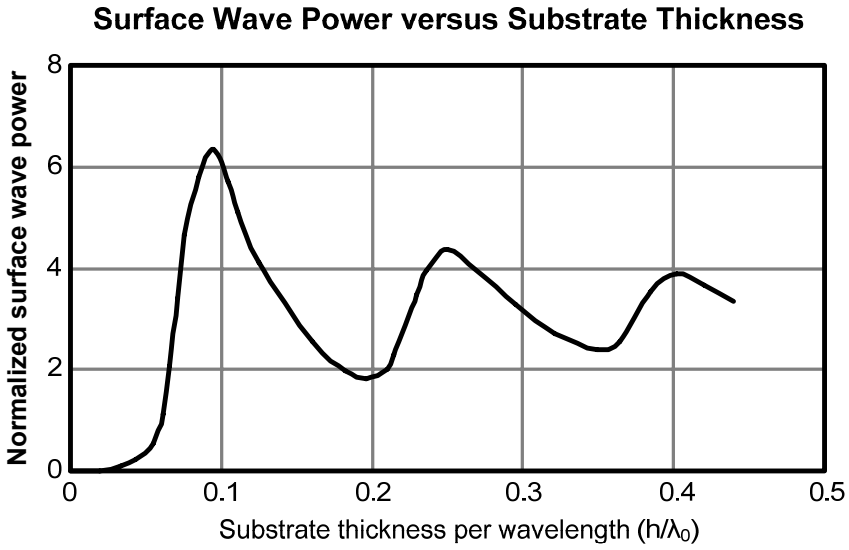


Figure 2.13 Surface wave power for a grounded dipole [2]

D. Radiating from the planar back side

Following the prior discussion we attempt to determine what happens if we remove any ground shield and radiate from the backside of the chip (see Figure 2.14). In this case, based on numerical calculations [2], the normalized radiated and surface-wave powers are plotted in Figure 2.15 and Figure 2.16, respectively. At 77 GHz the total radiated power, the sum of the power radiated from the air side and the substrate side, peaks at the silicon substrate thickness of 580 μm ($h = 0.15\lambda_0$) and approaches P_0. At this substrate thickness, the total surface-wave power is around 3.4 P_0 which is 3.4 times greater than the power radiated from the air side and the substrate side combined.

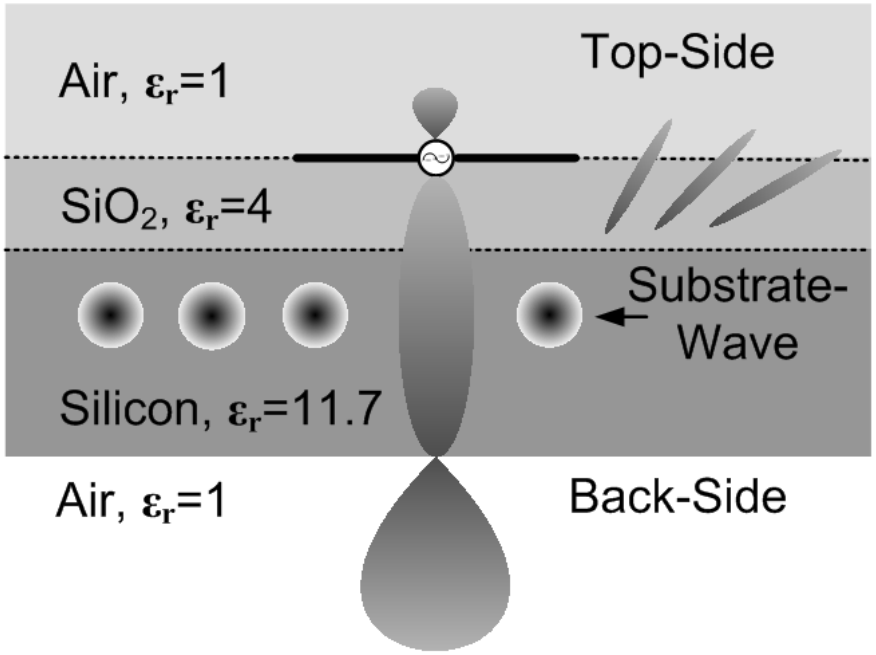


Figure 2.14 Radiating from planar back side

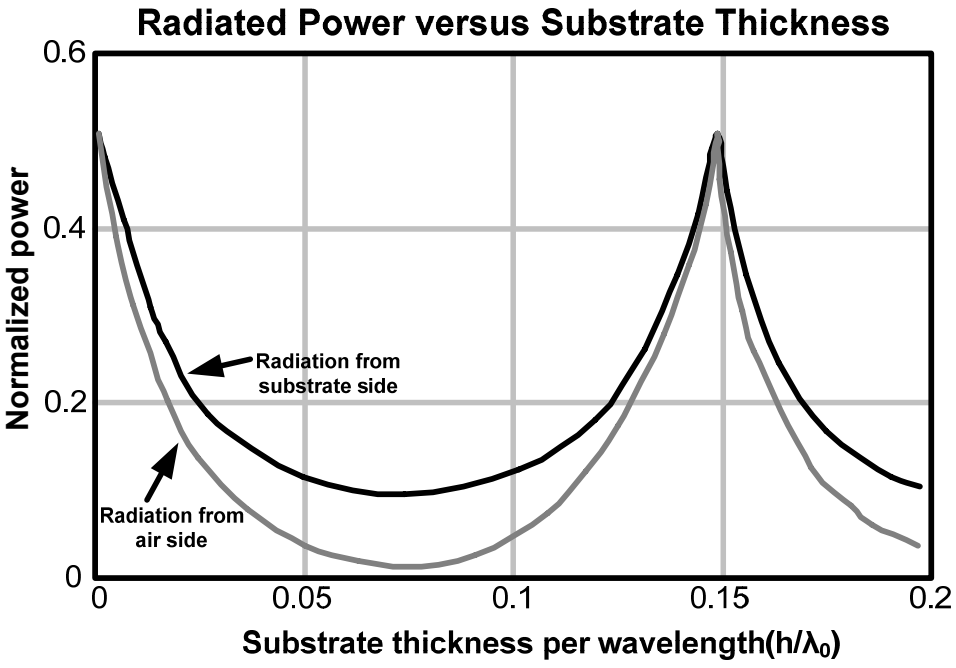


Figure 2.15 Normalized radiated power for an ungrounded dipole [2]

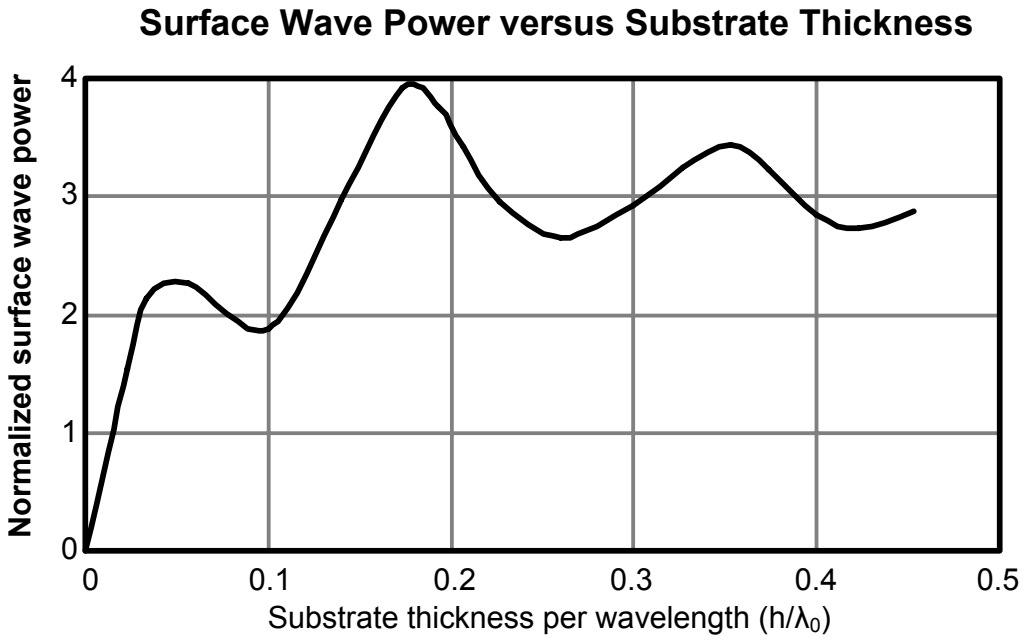


Figure 2.16 Surface wave power for an ungrounded dipole [2]

E. Radiating from the back side using a dielectric lens

Fortunately the amount of the total power absorbed into surface-waves depends on the geometry of the substrate. A hemispherical silicon lens with a matching layer can convert the surface-wave power to a useful radiated power [2]–[8]. This configuration is illustrated in Figure 2.17. A quarter-wave-length matching layer can be used to match the silicon impedance (Z_{si} = 110 Ω) to air impedance (Z_{air} = 377 Ω) [2].

In most of the designs discussed in the subsequent chapters we have used the silicon lens without the matching layer due to the fabrication limitations. Details of these designs are explained in Chapters 3 and 4.

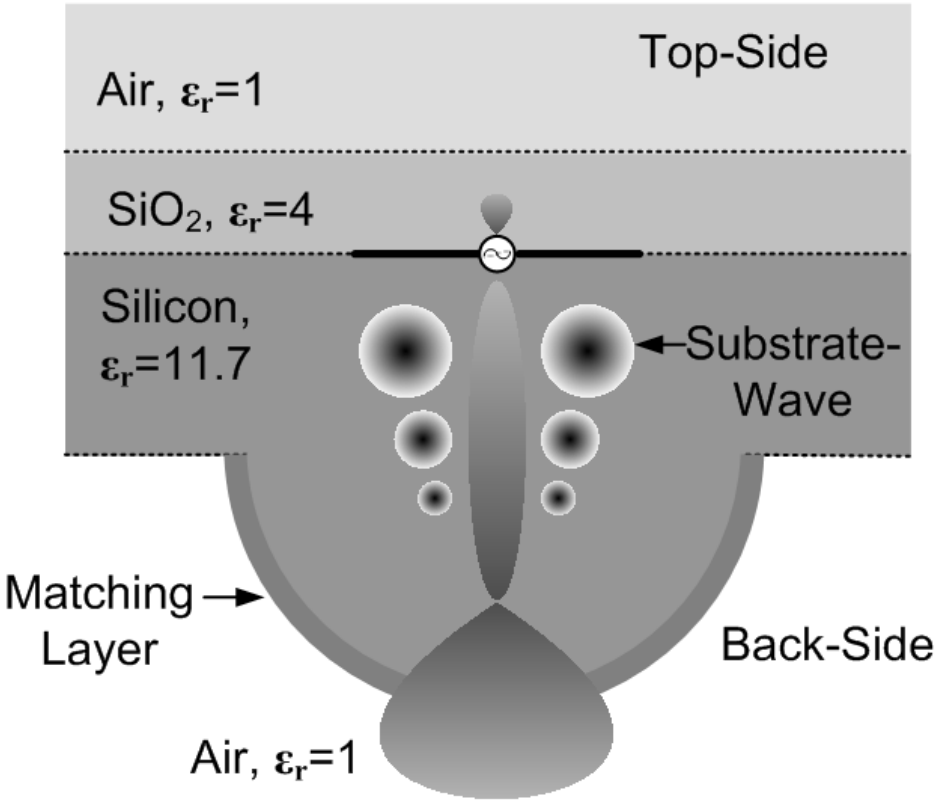


Figure 2.17 Radiating from back side using a silicon lens

2.4 Chapter Summary

In this chapter we reviewed the concept of the substrate modes. Substrate modes significantly affect the antenna performance including its gain, efficiency, and impedance. In a well-designed on-chip antenna structure, these modes should be carefully analyzed and understood. We have reviewed different on-chip antenna configurations and compared their pros and cons. A silicon lens is used to minimize the substrate modes and convert them to useful radiating modes.

Chapter 3

Millimeter-Wave Applications

3.1 Introduction

In this chapter the main applications of the millimeter wave transceivers are discussed. These include wireless communication, automotive radar, and imaging (for medical and security). One of the important characteristics of the millimeter waves which makes them suitable for a particular applications is the level of attenuation in different environments. Amazingly, the plot of the attenuation (by atmospheric gases) versus frequency is a highly variant non-monotonous curve, as shown in Figure 3.1 with a red color. As shown in this figure, in the millimeter-wave band (which is shown in the green area), the attenuation is around 15 dB/km at 60 GHz (due to the absorption of Oxygen molecules O_2), and less than 0.5 dB/km at 94 GHz. This makes the 60 GHz band suitable for indoor communication and the 94 GHz band suitable for ground to space communication. In the sub-millimeter range, the attenuation level reaches to 1000 dB/km, and in the infrared range it goes down to 0.05 dB/km.

Beside the attenuation, the other factor determining the applications for a specific frequency in the millimeter-wave range is the availability of low-cost active devices in that particular frequency. With today's low-cost CMOS and SiGe technologies, it is possible to implement active devices to amplify signals with frequency of less than

200GHz. In the following sections, we will review some of these applications in more detail.

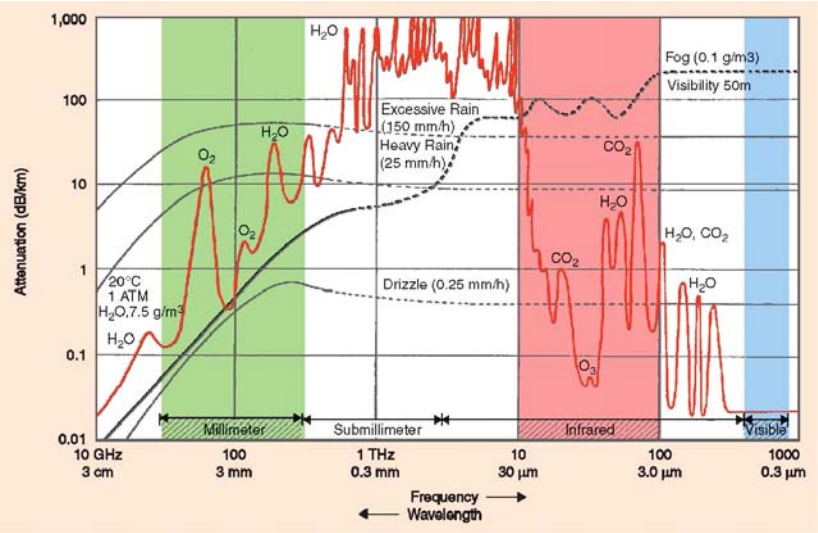


Figure 3.1 Attenuation of millimeter waves by atmospheric gases (red curve), rain, and fog [10]

3.2 Wireless Communication

Among several frequency bands in the millimeter wave range approved by Federal Communications Commission (FCC), 24-24.25GHz[3], 57-64GHz, and 92-95GHz bands are chosen to be used for communication purposes.

24–24.25 GHz [11]— In this band, point-to-multipoint systems, omnidirectional applications, and multiple co-located intentional radiators transmitting the same

[3] Although technically speaking the millimeter frequency range starts at 30GHz the behavior and general considerations for 24GHz systems are close enough to 30GHz to be considered in that category for the purposes of our discussions.

information are not allowed. Fixed, point-to-point operation is permitted in the 24.05–24.25 GHz band subject to the following conditions:

1) The field strength of emissions in this band shall not exceed 2500 millivolt/meter[4].

2) The frequency tolerance of the carrier signal shall be maintained within + 0.001% of the operating frequency over a temperature variation of –20°C to +50°C at normal supply voltage.

3) Antenna gain must be at least 33 dBi. Alternatively, the main lobe beam-width must not exceed 3.5 degrees. The beam-width limit shall apply to both the azimuth and elevation planes.

57–64 GHz— In 2001, the Federal Communications Commission (FCC) allocated 7 GHz in the 57–64 GHz band for unlicensed use. Table 3.1 shows the frequency band and the output power in different regions [12].

Region	Output power	Other considerations
Australia (59.4–62.9 GHz) Canada and USA (57–64 GHz)	10 mW into antenna, 500 mW peak	150 W peak EIRP min. BW = 100 MHz
Japan (59–66 GHz)	10 mW into antenna	47 dBi max. ant. Gain
Europe (57–66 GHz)	+57 dBm EIRP	min. BW = 500 MHz

Table 3.1 Spectrum and Emission Power [12]

[4] Field strength limits are specified at a distance of 3 meters.

FCC sets the following rules for the 57–64 GHz band [11]:

1) The band cannot be used for equipment used on aircraft or satellites and field disturbance sensors, including vehicle radar systems.

2) Emission levels shall not exceed 9 $\mu W/cm^2$ (EIRP of 10.2 W), as measured 3 meters from the radiating structure.

Standardization— One of the completed standards for high-speed radio communication which supports data rates of up to 4 Gbps is WirelessHD [13]. WirelessHD is designed to replace the wires in the High Definition Multimedia Interface (HDMI) with radio links, and is designed to handle high-definition television (HDTV) video streams between AV equipment. The target is defined as handling full HD (1080 p) video without high-efficiency coding. Existing technologies such as wireless local area network (LAN; 20 Mbps to 30 Mbps) and Ultra Wide-Band (UWB; about 200 Mbps) cannot handle 1080 p without using high-efficiency coding.

3.3 Automotive Radar

Automobile radar operating in the 77 GHz frequency band is one such application as the 76~77 GHz band has been allocated for this purpose in many countries around the world [3]. Also, the Electronic Communications Committee (ECC) within the European Conference of Postal and Telecommunications Administrations has allocated the 77~81 GHz window for automotive UWB short-range radar [14]. Compared to the radar bands at lower frequencies, such as the 24 GHz band, operating at the 77 GHz band is more compatible with other applications in the nearby frequency spectrum.

One of the applications of 76–77 GHz band is the Automotive Cruise Control (ACC) for vehicles. ACC systems are used to relieve the driver of part of his task of keeping distance and warn him in critical situations, thus making driving less strenuous, especially in flowing traffic. ACC can be activated typically at speeds of 30 km/h to 180 or 200 km/h. ACC systems are usually mounted in the radiator-grill or front bumper, and operate in the 77 GHz band.

FCC has approved unlimited use of the 24 GHz band for short range automotive applications but in Europe, The European Commission approved the decision on allocation of the 24 GHz frequency band for automotive short-range radar from only 2005 until 2013. From mid-2013 new cars have to be equipped with SRR sensors which operate in the frequency range between 77–81 GHz (79 GHz band).

Short Range Radar (SRR) can be used to monitor the surroundings of cars and gather useful information for safety and comfort applications. Radar appears to be the best sensor principle, because alternatives like video, laser, and ultrasound may have difficulties under bad weather conditions, when they are needed most. Additionally radar offers the vehicle manufacturers a stylistic advantage of mounting behind a plastic bumper that can be considered nearly transparent to the radar signal without requiring specific cut-outs or similar accommodations.

Some of the applications of the SRR are the following [16]:

1) ACC support with stop & go functionality

2) Collision warning

3) Collision mitigation

4) Blind spot monitoring

5) Parking aid (forward and reverse)

6) Lane change assistant

7) Rear crash collision warning

Especially the combination of LRR and SRR provides valuable data for advanced driver assistance systems (ADAS). Figure 3.2 illustrates some of these applications.

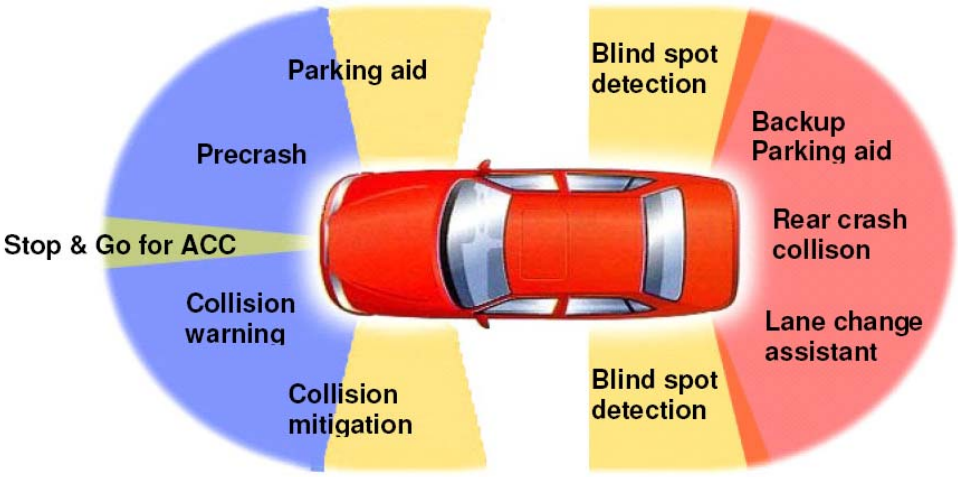


Figure 3.2 Applications of automotive radar [15]

3.4 Medical and Security Imaging

Passive and active techniques can be used in millimeter-wave imaging. In the passive millimeter-wave (PMMW) imaging, the receiver works as an infrared camera measuring the black body radiation in the millimeter-wave range. Active millimeter-wave imaging systems transmit mm-wave signals to an object and receive the scattered signals. The scattered data can be used to form the mm-wave picture of the object.

MM-wave systems are able to form images during the day or night; in clear weather or in low-visibility conditions, such as haze, fog, clouds, smoke, or sandstorms; and even through clothing. In the far IR and sub-millimeter-wave regime, significant attenuation occurs from water vapor. Conversely, in the millimeter-wave regime, there are propagation windows at 35, 94, 140, and 220 GHz, where the attenuation is relatively modest in both clear air and fog. Even taking into account the much higher blackbody radiation at IR and visible frequencies, millimeter waves give the strongest signals in fog when propagated over distances of useful interest (Figure 3.3). In fact, millimeter-wave radiation is attenuated millions of times less in clouds, fog, smoke, snow, and sandstorms than visual or IR radiation. This is the critical advantage of PMMW imagery [10].

Figure 3.4 compares the quality of the image of a runway taken by a visible and passive microwave camera. As seen in the picture, the quality of the image taken with visible camera is significantly deteriorated in foggy weather, whereas the quality of the image taken by a passive microwave camera is not changed significantly.

It is important to realize that by increasing the aperture size, the quality of the image is improved. Figure 3.5 compares three images taken by a passive scanning system with different aperture sizes. As expected, the quality of the scanning system with larger aperture size is better than the one taken by a smaller aperture.

One of the important characteristics of the millimeter imaging (especially active imaging) is the penetration. This can be used in several applications including concealed weapon detection for security purposes and tumor detection for medical purposes. Figure 3.6 shows a picture taken by a passive camera and Figure 3.7 shows the one taken by an active camera. In both cases, it is clear that the person in the picture is carrying a weapon. Figure 3.8 shows one of the medical applications. In this figure, a picture of a breast tumor taken by an active camera is shown.

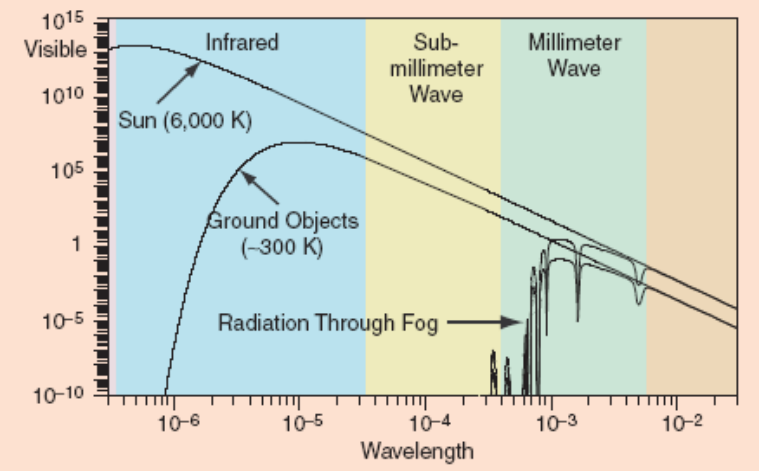


Figure 3.3 The effect of fog on the blackbody radiation intensity of the sun (6,000 K) and a ground object (~ 300 K) as a function of wavelength. Curves are shown for both objects without fog and with the effect of 1 km of fog. [10]

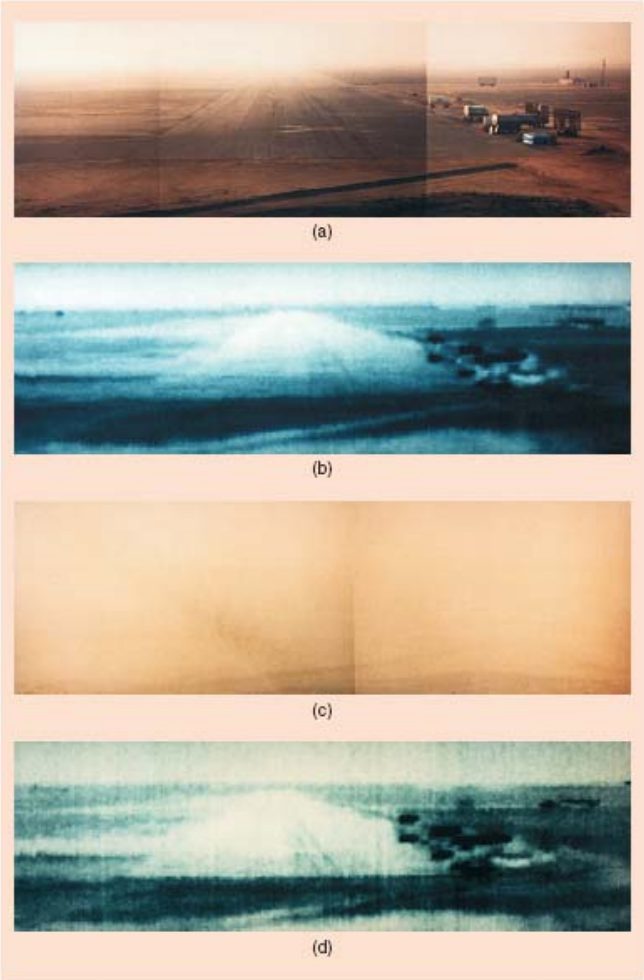


Figure 3.4 PMMW images of a runway: (a) and (c) show visible images in clear and foggy weather; (b) and (d) show the corresponding PMMW images. [10]

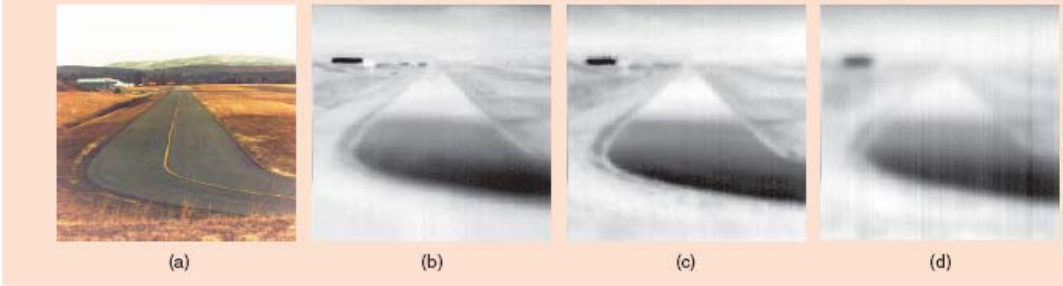


Figure 3.5 The airport scene in visible light (a) with varying aperture sizes for the 94 GHz PMMW scanning system: (b) 48 in, (c) 24 in, and (d) 12 in [10]

Figure 3.6 Concealed weapon detection with PMMW [16]

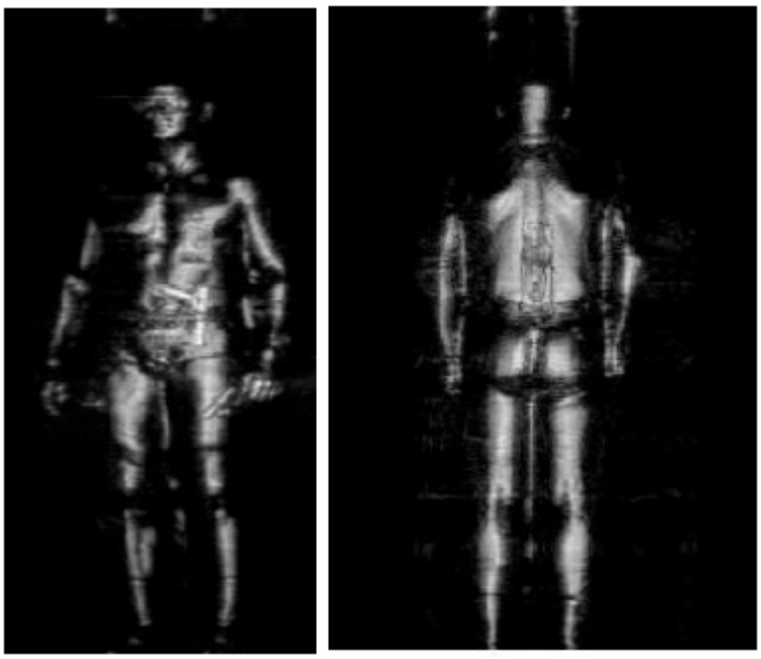


Figure 3.7 Concealed weapon detection with active imaging [16]

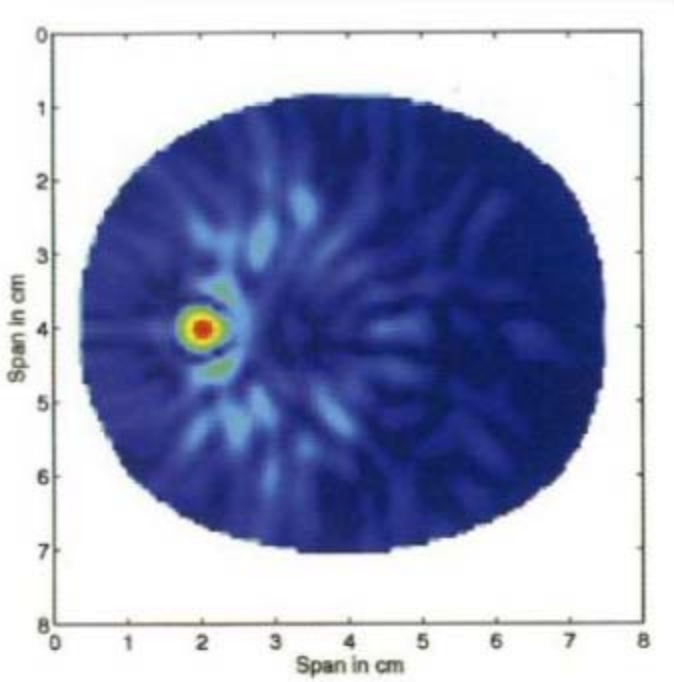


Figure 3.8 Active microwave detection of breast tumor [16]

3.5 Chapter Summary

In this chapter, we reviewed some of the millimeter waves' applications. These include high-speed wireless communication, accurate radar systems, and security and medical imaging. Penetration and ability to perform in bad weather are the unique characteristics of these systems, which are extremely important for radar and imaging applications.

Chapter 4

A 77 GHz 4-Channel Phased-Array Transceiver with On-Chip Antennas in Silicon

4.1 Introduction

Silicon technology has entered the realm of millimeter wave frequencies by the sheer force of transistor scaling, unmatched levels of integration, low cost, and high yield. This has opened the door to a plethora of new applications, formerly accessible only to compound semiconductors. Although compound semiconductor technologies, such as GaAs and InP, provide better single device performance, they do not provide the same large-scale integration capabilities in terms of yield and cost. High yield integration of silicon based technologies lowers the cost of the production and at the same time makes it possible to integrate billions of transistors on the same die to form RF systems enabling an unprecedented level of functionality and flexibility. Tuning and calibration capability achieved by co-integration of digital circuitry improves the performance of the critical analog/RF elements and makes it possible to use silicon devices which have traditionally had poor modeling quality. More importantly, integration of advanced base-band and signal processing elements enables the realization of novel system architectures for existing and emerging mm-wave applications.

To be able to compete with compound technologies in terms of performance, some key problems need to be addressed. Low breakdown voltage transistors in silicon-based

processes make power generation more challenging and poor metal conductivity makes it less efficient. Lossy silicon substrate and low conductivity interconnect make the integration of passive elements, such as inductors, capacitors, and transmission lines, less effective by significantly reducing the quality factor (Q) of these passive structures. While GaAs and InP-based technologies provide relatively high substrate resistance (10^7 to 10^9 Ω-cm), the silicon bulk used in today's standard processes limits the substrate resistivity to less than 10 Ω-cm, mainly due to the high level of doping required to avoid transistor latch-up.

These problems limit the efficiency and maximum output power of the PAs on the transmitter side, increases the noise figure (NF) of the LNA, and thus lowers its sensitivity and gain on the receiver side. However, higher levels of integration enable implementation of low-cost highly integrated phased arrays which can alleviate both of these problems. Phased arrays which are discussed in Section 4.1.1 increase the transmission Effective Isotropically Radiated Power (EIRP) and improve the system Signal-to-Noise Ratio (SNR) on the receiving side.

In this chapter, we will present the first fully integrated W-band on-chip radar which includes a four-channel phased-array transmitter, four-channel phased-array receiver, frequency synthesizer, phase shifters, and on-chip antennas [3] , [17].

4.1.1 Phased-Array Systems

Multiple antennas with appropriate phase and amplitude adjustments can be used to shape the effective radiation pattern of the whole antenna array. Figure 4.1 shows N linear receiving antenna elements located on the x-axis. Assuming an individual antenna gain of $G_k(\theta)$, phase delay of $\phi_k(\theta)$, and amplitude adjustment of $A_k(\theta)$ for antenna k, the array pattern of $\Psi(\theta)$ on the x-y plane can be calculated as

$$\Psi(\theta) = \sum_{k=1}^{N} G_k(\theta) A_k(\theta) e^{-j(\phi_k + 2\pi \frac{x_k \sin\theta}{\lambda})} \tag{4.1}$$

where x_k is the location of the antenna k and λ is the free-space wavelength.

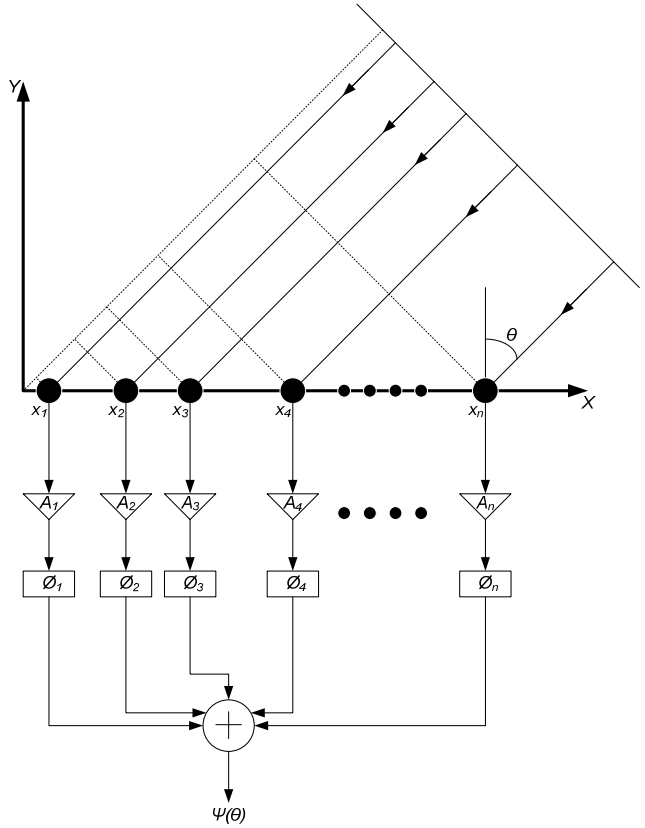


Figure 4.1 Receiving antenna array

Due to the reciprocity theorem of linear time independent systems (LTI), the transmitting array pattern can also be calculated by equation (4.1). A sample of a transmitting array is illustrated in Figure 4.2.

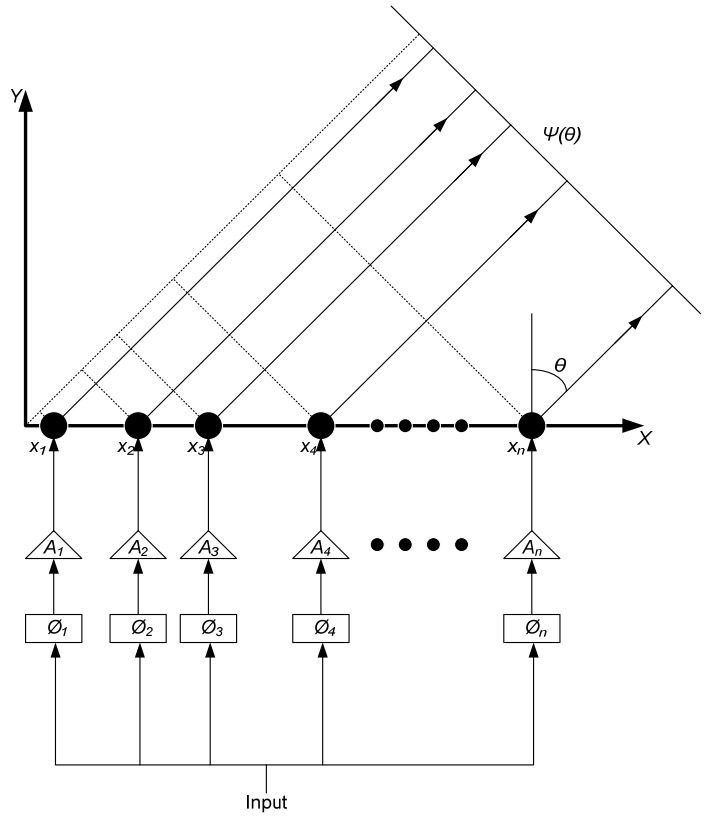


Figure 4.2 Transmitting antenna array

In most of the conventional phased arrays, the spacing between adjacent antenna elements is chosen to be about $\lambda/2$, gain factor (A_k) to be constant across elements, and $\phi_k - \phi_{k-1} = \phi_0 =$ constant. With this choice of parameters, and assuming an isotropic single element pattern for each antenna, there will be a single main lobe with minimized side lobes. In this case, the magnitude of the array pattern can be calculated by

$$|\Psi(\theta)| = \frac{\sin[N(\pi \sin\theta + \varphi_0)/2]}{\sin[(\pi \sin\theta + \varphi_0)/2]} \tag{4.2}$$

and power gain in dB will be

$$G_p(\theta) = 10 \cdot \log_{10}(|\Psi(\theta)|^2) \ . \tag{4.3}$$

It is important to mention that the advantage of phased-array receivers is not limited to the capability of shaping the array pattern. On the receiver side, a phased array system can improve the signal-to-noise ratio (SNR) of the receiver. This is because of the fact that signals at each element combine coherently but noise combines incoherently [18]. Thus, the signal power increases by a factor of N^2 while noise amplifies by a factor of N, hence the SNR improves by a factor of N ($10 \cdot \log_{10}(N)$ in dB). For example, for an array size of $N = 8$, the SNR improves by a factor of 9dB.

On the transmitter side, phased array can alleviate the problems involved in designing a high power PA. The EIRP of a single PA connected to an isotropic antenna while transmitting power P can be calculated as

$$EIRP_s = P \ . \tag{4.4}$$

However, the EIRP of an array of PAs each connected to an isotropic antenna and each transmitting power (P/N) can be calculated as

$$EIRP_{array} = N^2(P/N) = NP \ . \tag{4.5}$$

From the above equations, we see that to achieve the same *EIRP*, each PA in the array has to transmit N^2 times less power than the single PA connected to an isotropic antenna.

4.2 A 77 GHz Phased-Array Transceiver

In this section, we will go through the details of the design, layout, and measurement of the 77 GHz transceiver.

Process— The chip is implemented in the IBM 130 nm Silicon Germanium BiCMOS process. This process has substrate resistivity of about 10 Ω.cm and five levels of copper metal layers, including M_1, M_2, M_3, M_4, each with a thickness of 0.32 μm, and M_5, with a thickness of 0.55 μm, as well as two levels of aluminum with thicknesses of 1.25 μm and 4 μm. The process also provides 120 nm NPN bipolar transistors with peak f_T of 200 GHz.

4.2.1 Receiver Block Diagram

The 77 GHz 4-element phased-array transceiver integrates multiple signal transmission elements, receiving elements, signal distribution and combination, LO signal generation and distribution, phase shifters, and 77 GHz antennas on a single silicon die [3] , [17]. In this section, we describe the block diagram of the 77 GHz 4-element phased-array receiver. The transmitter and LO generation circuits are discussed in [17].

Figure 4.3 shows the architecture of the receiver chip. It adopts a two-step down-conversion scheme with an RF frequency of 76~81 GHz and an IF in the 25~27 GHz. This frequency plan enables us to generate the first and second LO signals off of a single frequency synthesizer. It is also noteworthy to mention that the IF is located in the 22~29 GHz radar band, hence a dual-mode radar chip can be potentially developed by bypassing

the RF front-end. Each receiver front-end consists of an on-chip dipole antenna coupled to a dielectric lens, an LNA, an RF mixer, and a dual-mode IF amplifier. The gain of the IF amplifier can be varied by 15 dB using a single digital control bit.

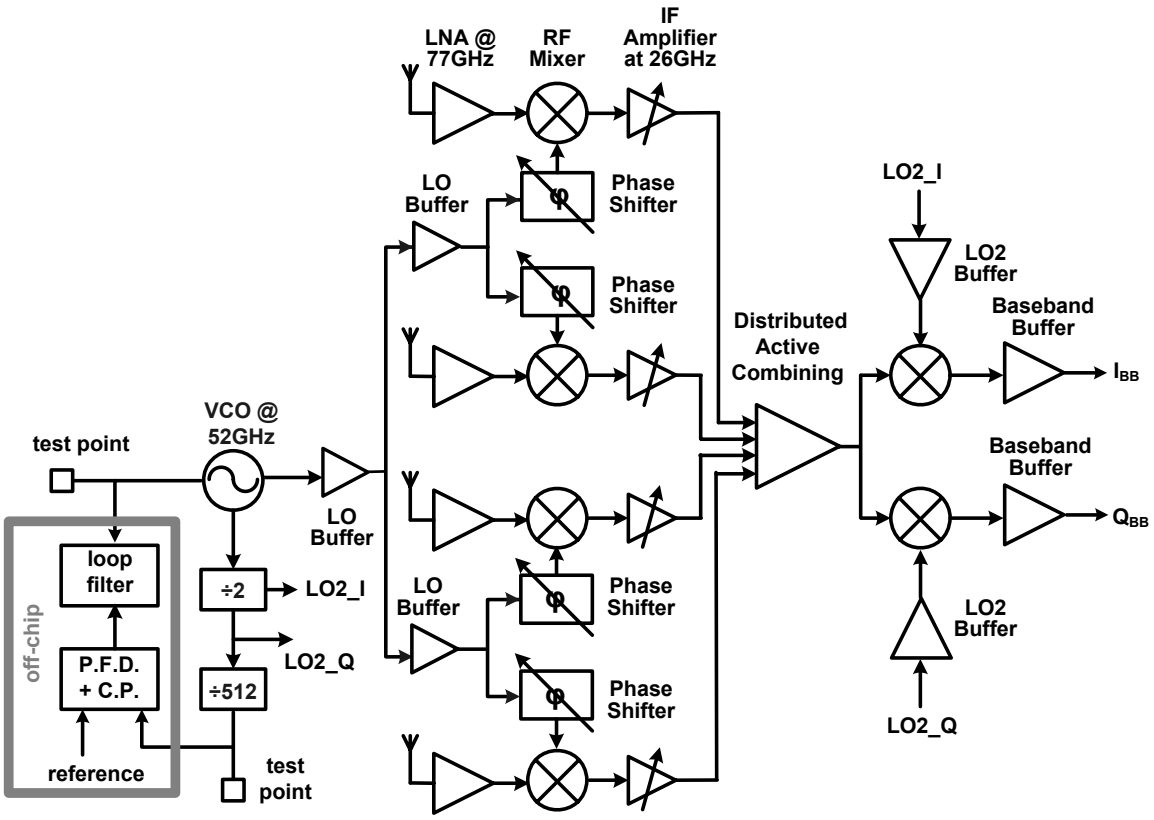


Figure 4.3 77 GHz phased-array receiver system architecture [3] , [17][5]

4.2.2 The 77GHz On-chip Dipole Antenna Design

In Chapter 2, the concept of the substrate modes is introduced and the power coupled into the substrate modes is calculated for several configurations, including the planar substrates and the substrates with hemispherical lens. As mentioned in Chapter 2, the planar substrate behaves as a dielectric waveguide and converts the radiated power into

[5]Designed by Aydin Babakhani, Xian Guan, Abbas Komijani, and Arun Natarajan.

the substrate modes. In this section we will review the electromagnetic simulation results for on-chip antennas operating at 77 GHz.

Figure 4.4 shows an on-chip dipole antenna with a length of 688 µm placed on a planar silicon substrate. The simulation is performed for four different cases including doped (lossy) grounded, doped (lossy) ungrounded, un-doped (non-loss) grounded, and un-doped (non-loss) ungrounded silicon substrates.

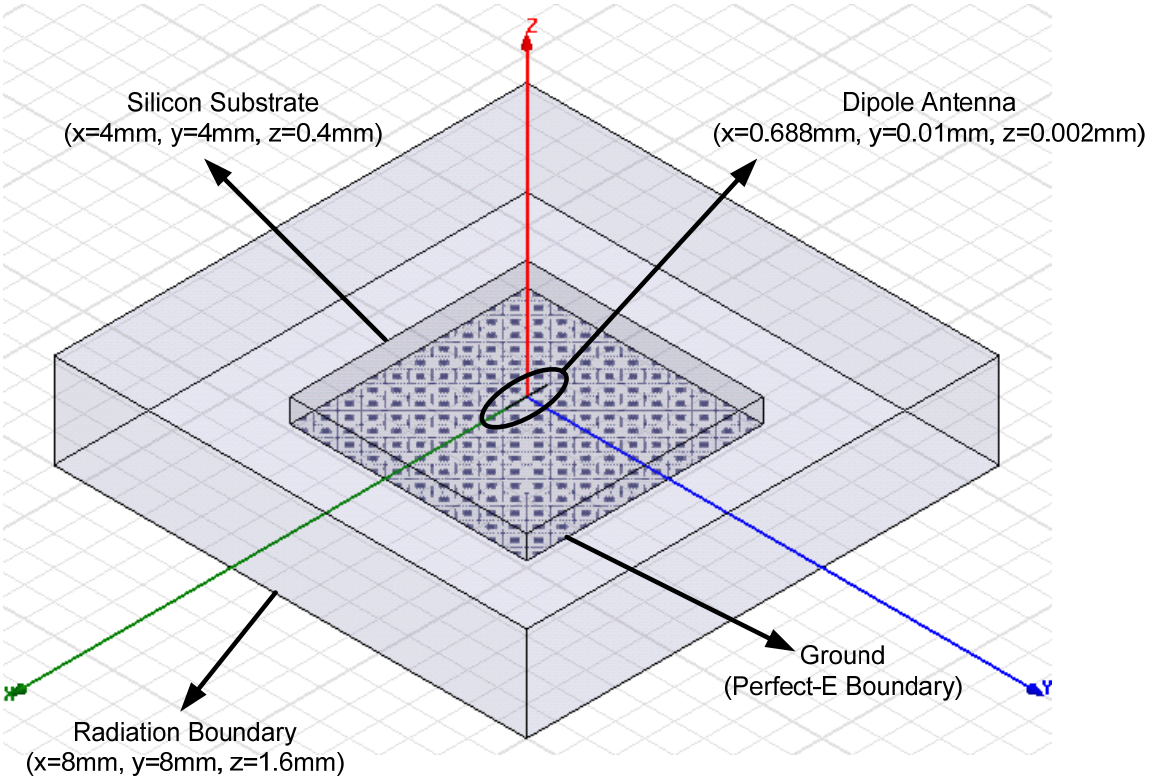


Figure 4.4 An on-chip dipole antenna on a 400 µm thick silicon substrate. The bottom face of the silicon substrate is a perfect electric conductor.

Figure 4.5 shows the magnitude of the electric field on the E- and H-planes of the dipole antenna placed on an un-doped (non-loss) grounded substrate. Figure 4.6 shows the simulation result for a doped (lossy) substrate.

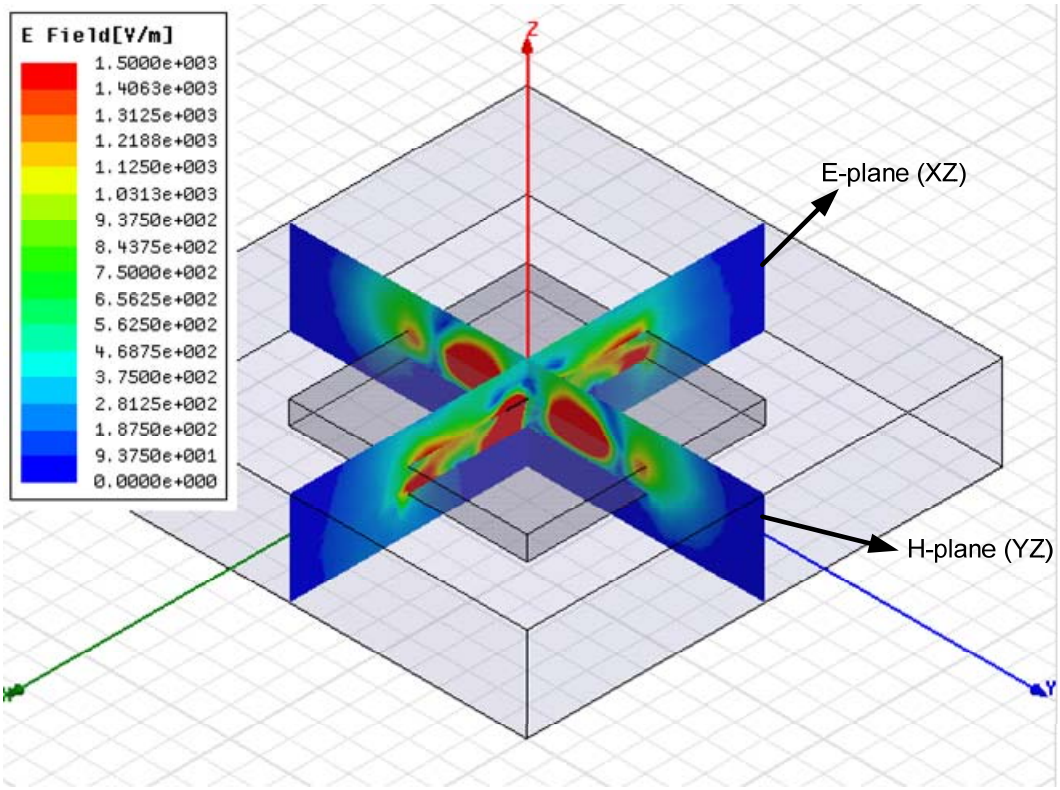


Figure 4.5 Magnitude of the electric field for an un-doped (non-loss) grounded substrate

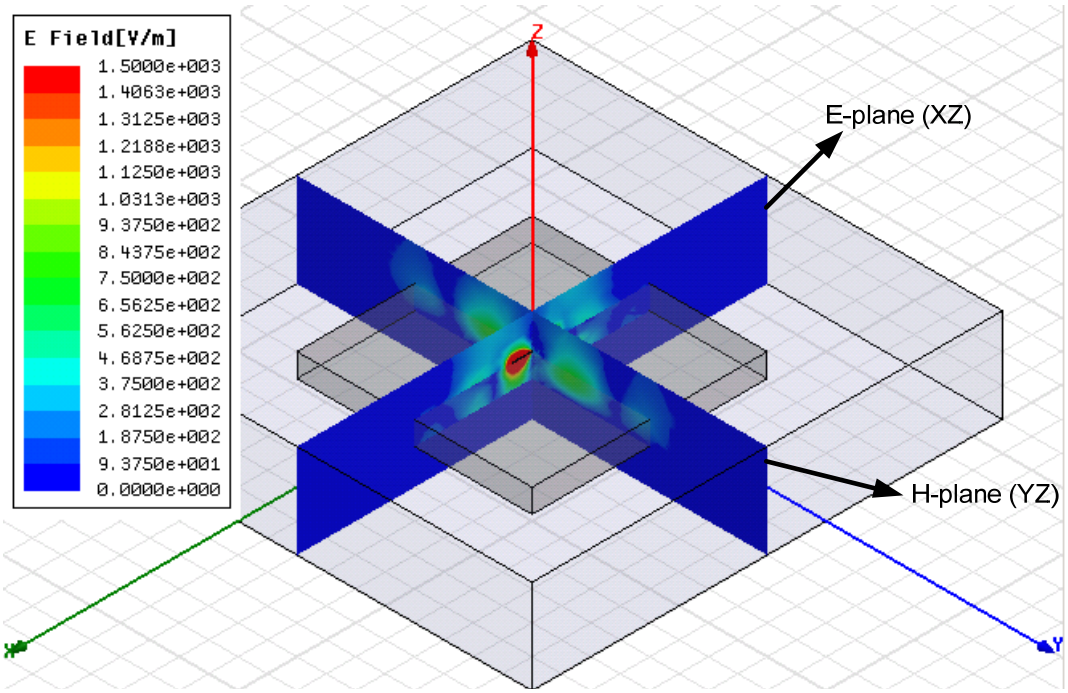


Figure 4.6 Magnitude of the electric field for a doped (lossy) grounded substrate

As seen in Figure 4.7, Figure 4.8, Figure 4.9, and Figure 4.10, most of the dipole power is coupled to the substrate modes. The substrate waves travel inside the silicon slab and reach to the edges of the chip resulting in an undesirable radiation pattern. In the case of a lossy substrate (Figure 4.8 and Figure 4.10), most of the substrate modes' power is wasted as heat. That is why the strength of the electric field in Figure 4.8 and Figure 4.10 is weaker than that in Figure 4.7 and Figure 4.9.

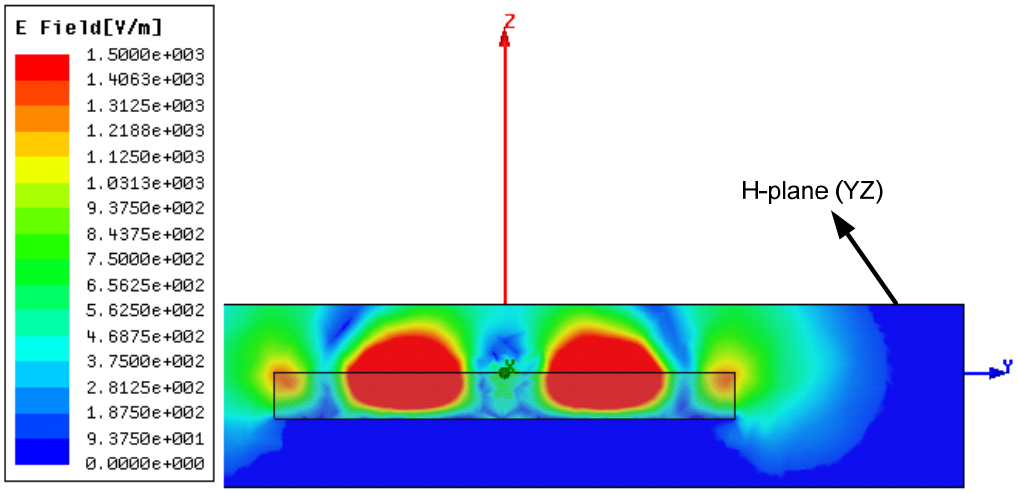


Figure 4.7 Magnitude of electric field for an un-doped (non-loss) grounded substrate (H-plane)

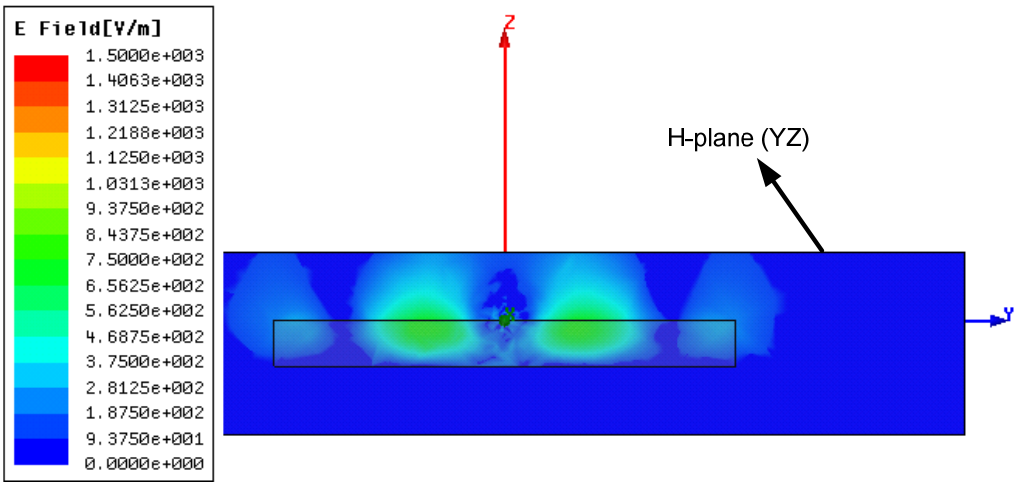


Figure 4.8 Magnitude of the electric field a doped (lossy) grounded substrate (H-plane)

Figure 4.7 and Figure 4.8 show the magnitude of the electric field on the H-plane (YZ) while Figure 4.9 and Figure 4.10 show the magnitude of the electric field on the E-plane (XZ). In all of these simulations, the bottom face of the silicon substrate is assumed to be a perfect electric conductor (grounded substrate).

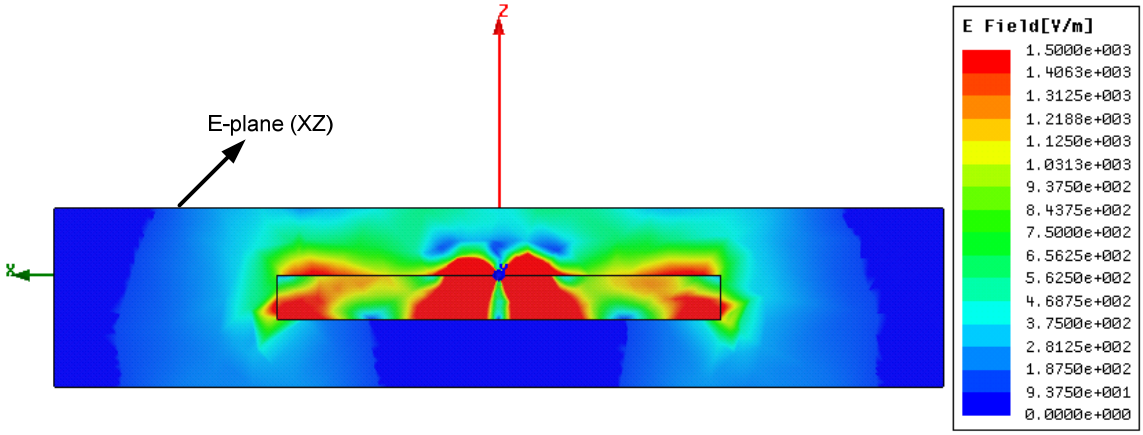


Figure 4.9 Magnitude of the electric field for an un-doped (non-loss) grounded substrate (E-plane)

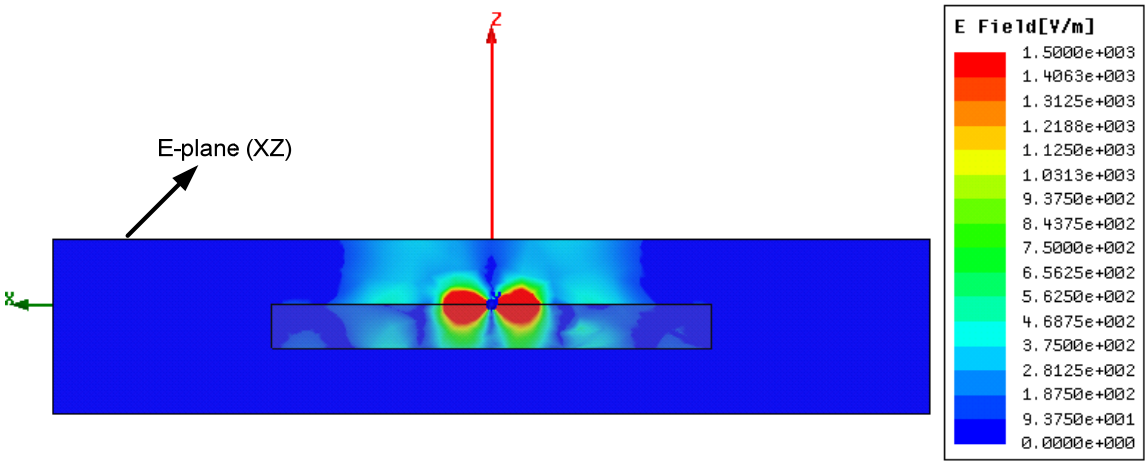


Figure 4.10 Magnitude of the electric field for a doped (lossy) grounded substrate (E-plane)

Figure 4.11 to Figure 4.16 show the simulation results for an ungrounded silicon substrate. In these cases the perfect-E boundary is removed from the bottom face of the substrate.

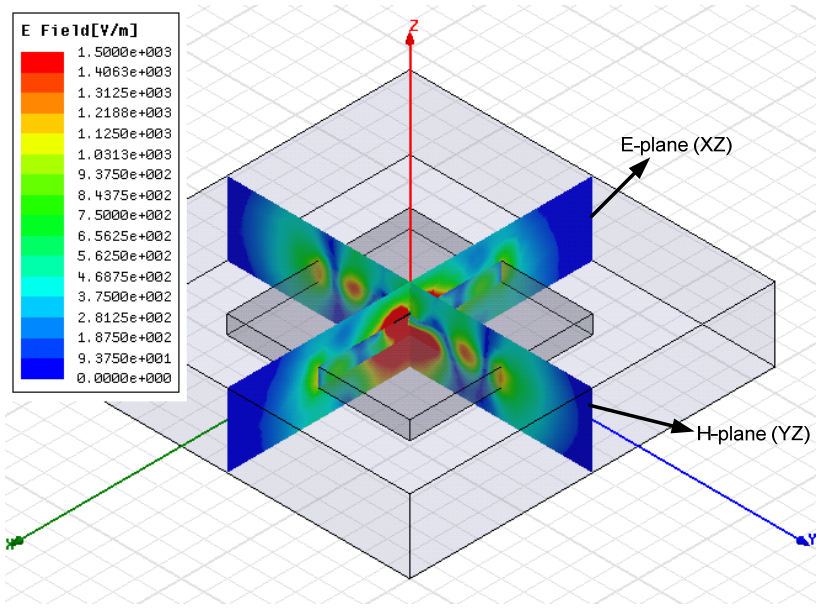


Figure 4.11 Magnitude of the electric field for an un-doped (non-loss) ungrounded substrate

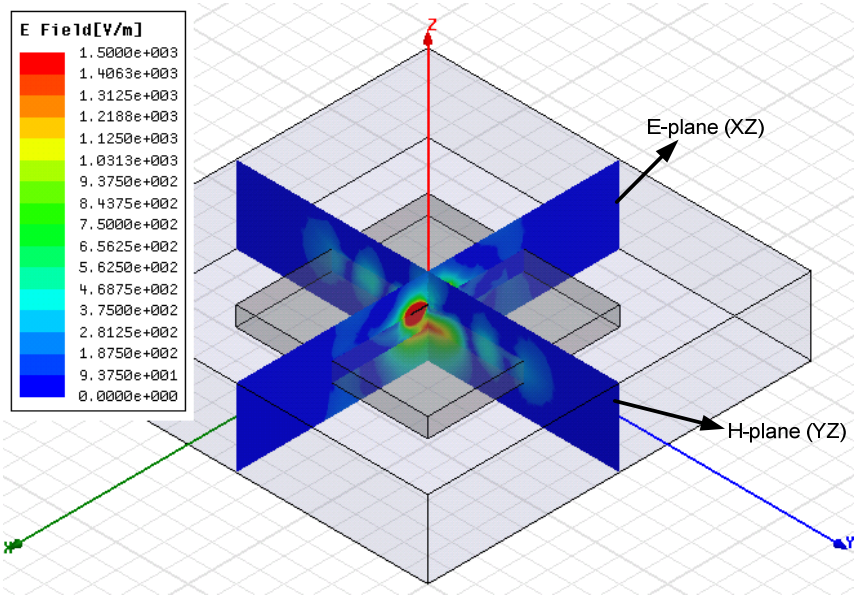


Figure 4.12 Magnitude of the electric field for a doped (lossy) ungrounded substrate

As seen in Figure 4.13 to Figure 4.16, most of the power is radiated from the back side (bottom) of the substrate instead of the top side. In this case, a significant portion of the power is coupled to the substrate modes. This power travels to the chip edge, couples into the air, and interferes with the power radiated from the bottom side, resulting in an undesired radiation pattern.

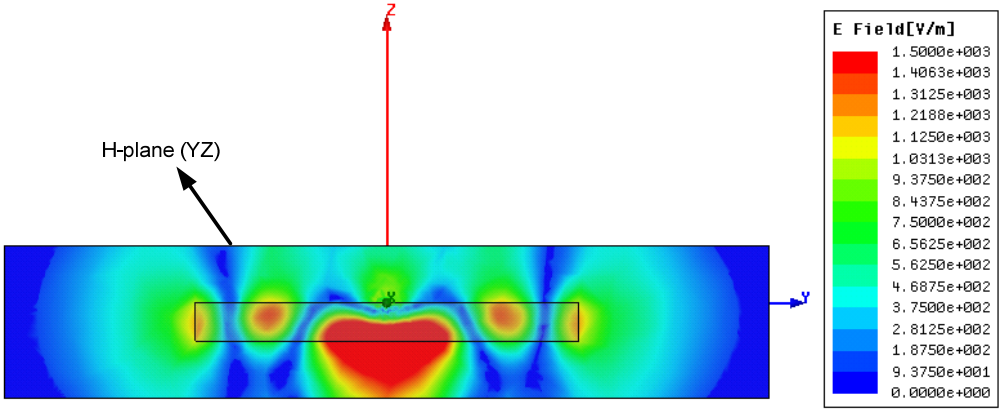


Figure 4.13 Magnitude of the electric field for an un-doped (non-loss) ungrounded substrate (H-plane)

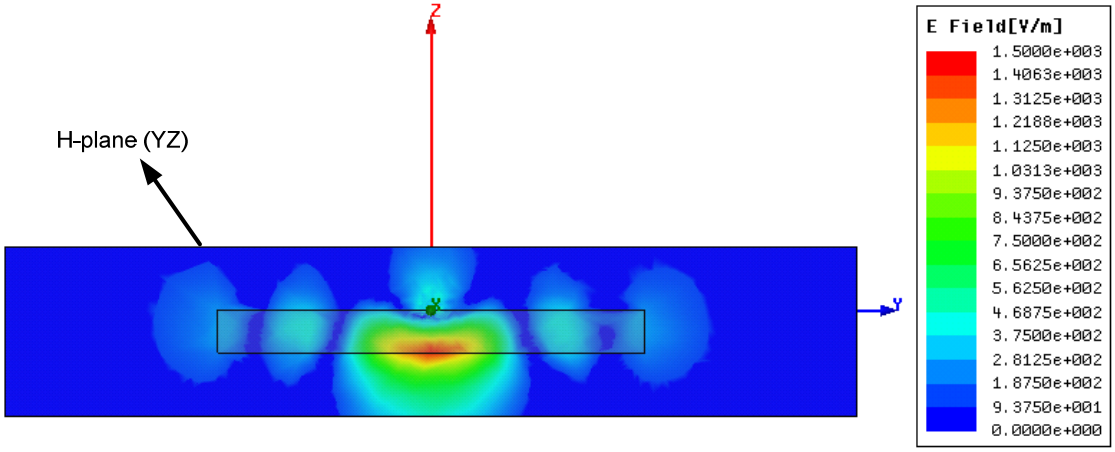


Figure 4.14 Magnitude of the electric field for a doped (lossy) ungrounded substrate (H-plane)

Figure 4.14 and Figure 4.16 show the simulation results for a doped (lossy) substrate. In these cases, most of the power coupled into the substrate modes is wasted as heat.

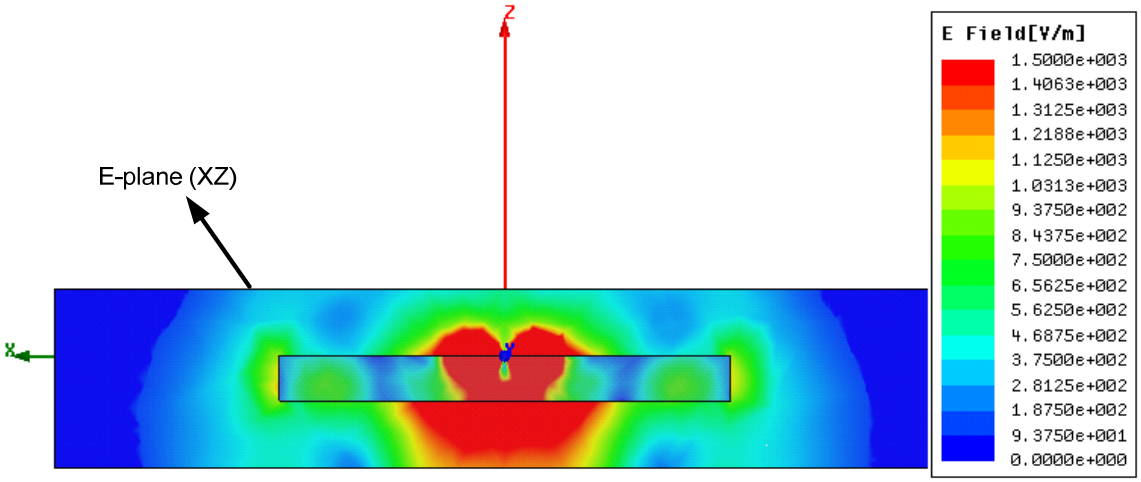


Figure 4.15 Magnitude of the electric field for an un-doped (non-loss) ungrounded substrate (E-plane)

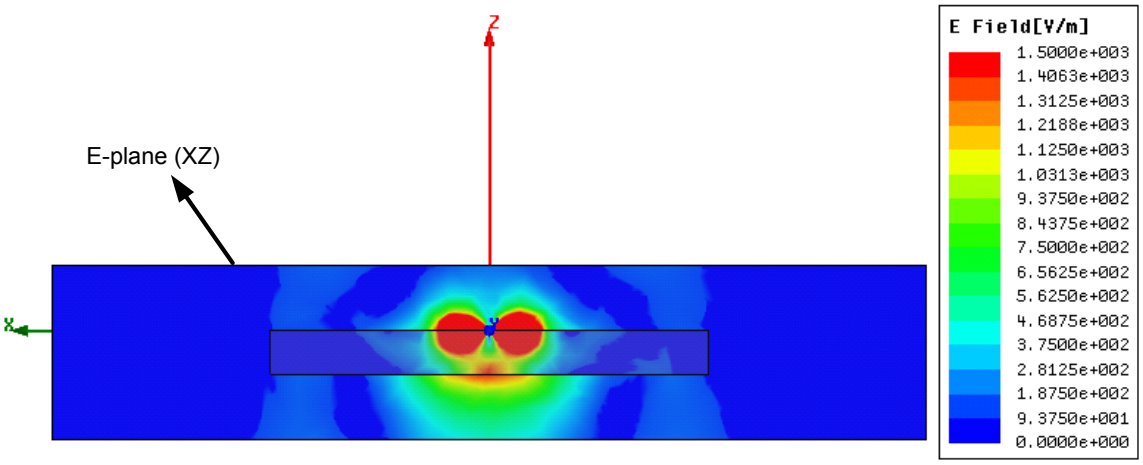


Figure 4.16 Magnitude of the electric field for a doped (lossy) ungrounded substrate (E-plane)

As mentioned in Chapter 2, the amount of the total power coupled into the substrate modes depends on the substrate geometry. By placing an un-doped (non-loss) hemispherical silicon lens on the backside of the substrate (Figure 4.17), most of the substrate modes' power will travel inside the silicon lens and couple into the air. Due to the impedance mismatch between the silicon and air, about 30% of the power is reflected from the silicon-air boundary. Figure 4.18 to Figure 4.20 show the magnitude of the electric field inside the silicon substrate and the silicon lens. The magnitude of the electric field on the H-plane is shown in Figure 4.19, whereas the magnitude of the electric field on the E-plane is shown in Figure 4.20.

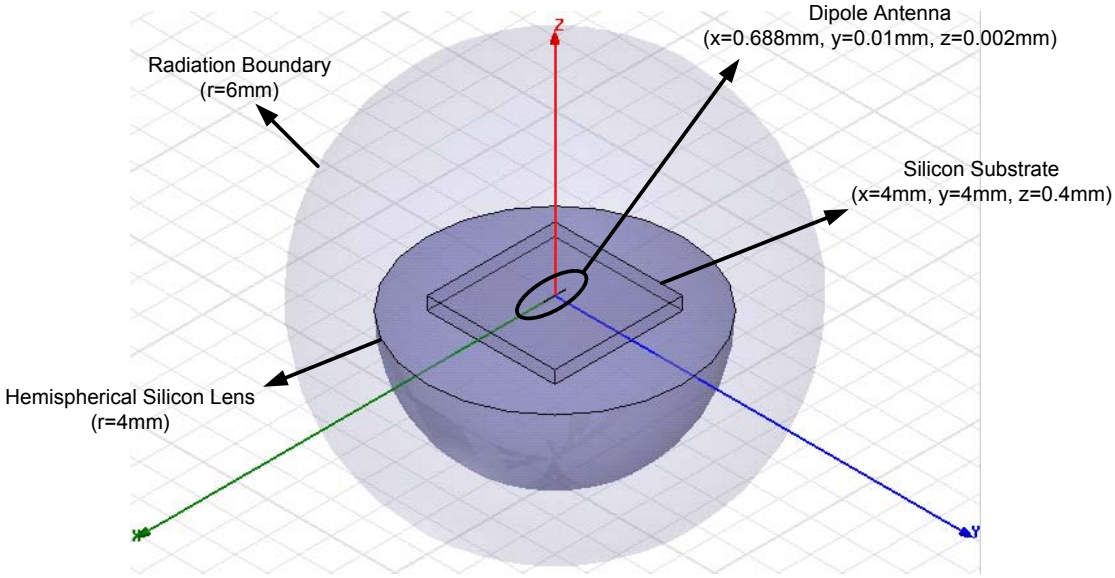


Figure 4.17 An on-chip dipole antenna on a 400 µm thick un-doped silicon substrate which is placed on a hemispherical silicon lens (lens and substrate are both un-doped)

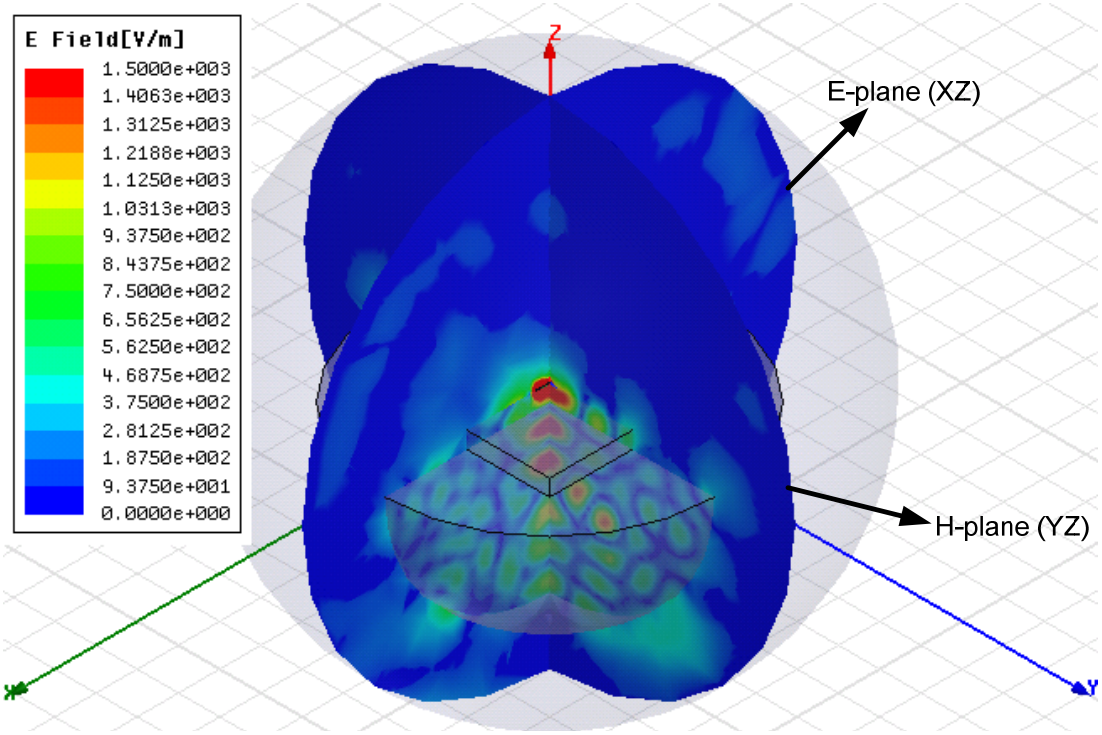


Figure 4.18 Magnitude of the electric field in the substrate and the silicon lens

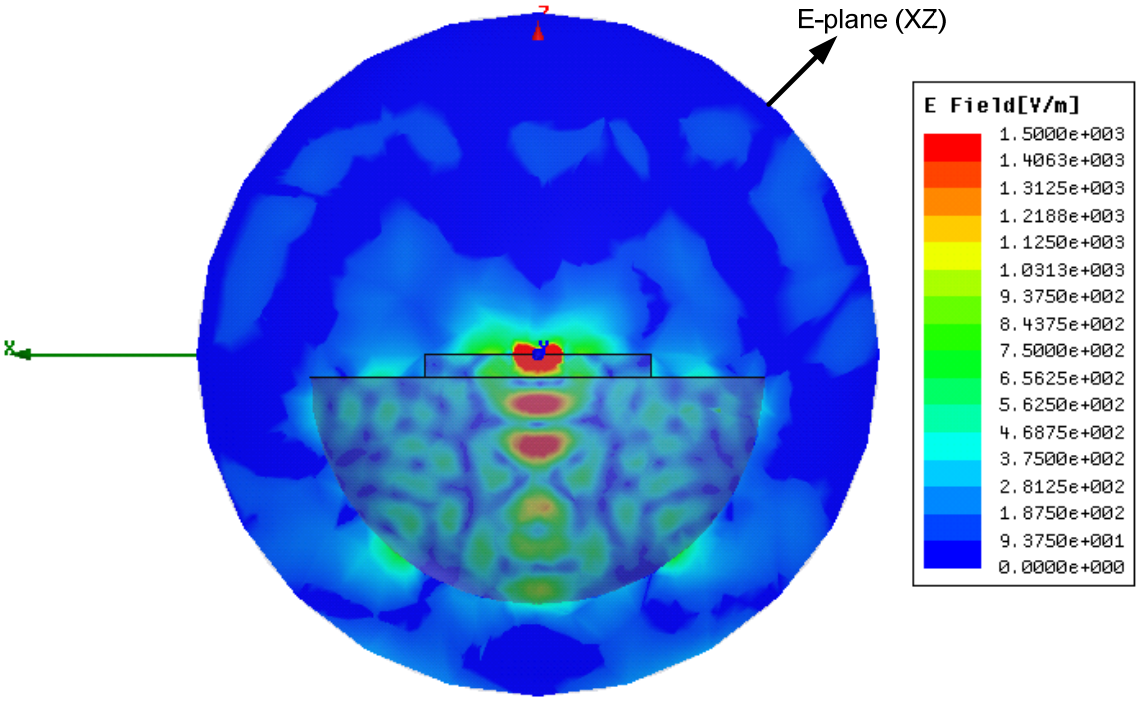


Figure 4.19 Magnitude of the electric field on the E-plane

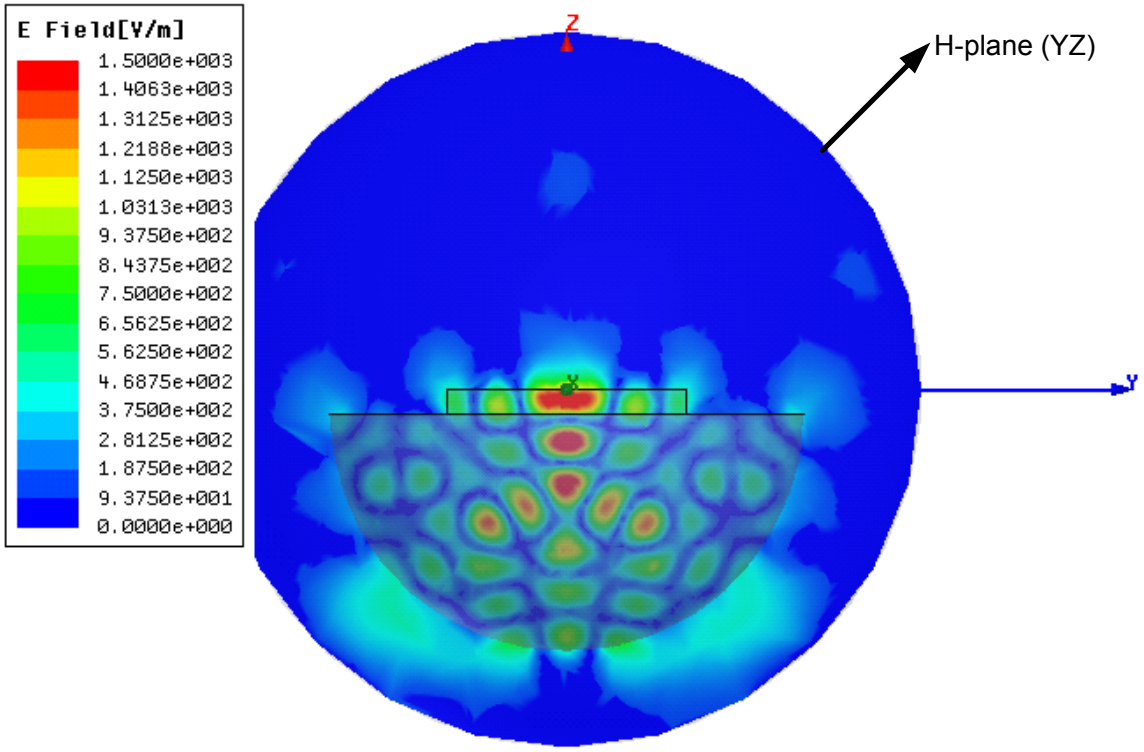


Figure 4.20 Magnitude of the electric field on the H-plane

In the real implementation, antennas are fabricated by using bottom metal layers to minimize the distance to the substrate. A parallel combination of three bottom metal layers maintains high antenna metal conductivity. To further reduce the substrate loss, the silicon chip is thinned down to 100 μm. This minimizes the path length through which the radiated wave travels inside the lossy doped substrate. Due to layout limitations in our design, antennas are placed at the edge of the chip and a slab of un-doped silicon is abutted to the substrate to maintain a uniform dielectric constant substrate underneath the antenna (Figure 4.21). For mechanical stability, a 500 μm thick un-doped silicon wafer is placed underneath the chip and the silicon lens is mounted on the back side seen in Figure 4.21. All of the low frequency connections are brought to the chip by board metal traces

and wire-bond connections. As this setup is highly compatible with flip-chip technology, all of these low frequency signals can be carried by flip-chip connections as well.

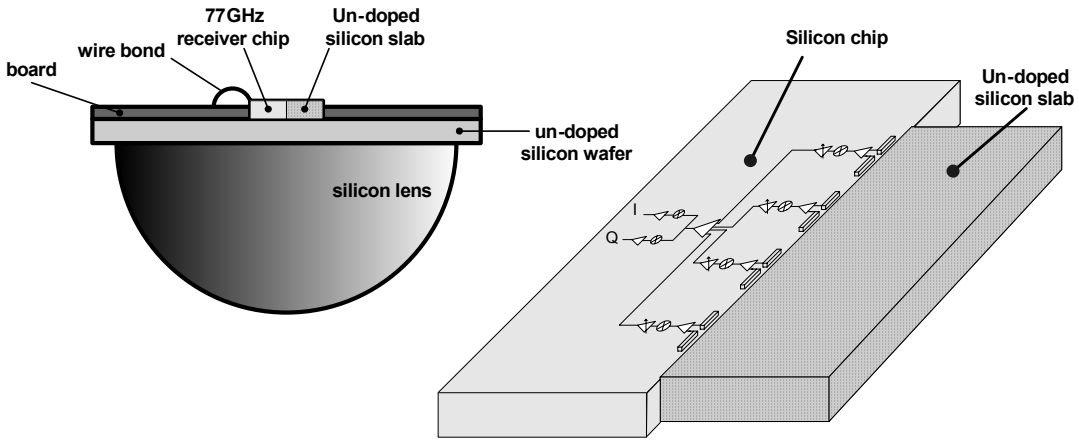


Figure 4.21 Board setup configuration

4.2.3 Receiver Circuits Schematics

77 GHz LNA Design— A differential two-stage LNA is designed and implemented. The LNA amplifies a differential signal received at the port of the on-chip dipole antenna and couples the differential signal to the down-converter mixer. One of the main challenges of the design is achieving a relatively high Common-Mode Rejection Ratio (CMRR) at millimeter-wave frequencies. The limitation is coming from the parasitic capacitance of the current source in the differential pair of the LNA. Figure 4.22 shows this parasitic capacitance.

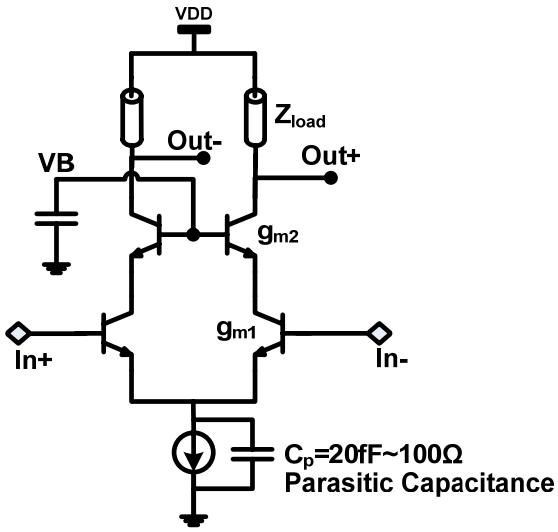


Figure 4.22 A differential pair in the LNA

The CMRR of the LNA can be calculated as:

$$CMRR = |1 + g_{m1}(jC_p\omega)^{-1}| \ . \tag{4.6}$$

At 77 GHz, a parasitic capacitance of 20 fF is translated to an impedance of about 103 Ω. This low impedance of the current source significantly limits the CMRR by increasing the common-mode gain without changing the differential gain. To alleviate this problem, the architecture of the conventional design needs to be modified in a way to increase the common-mode gain without changing the differential gain. In our design, we have reduced the common-mode gain with the following methods:

I— Instead of using two single-ended transmission lines, a differential coupled-wire transmission line is designed to carry the differential signal. This differential T-line introduces a differential impedance of 100 Ω to keep a good match for the differential

signal, but introduces a large mismatch for the common-mode signal. Figure 4.23 illustrates this technique.

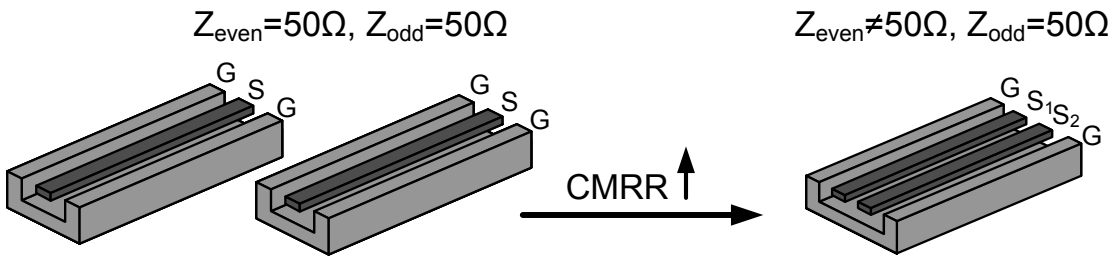


Figure 4.23 Coupled-wire transmission line

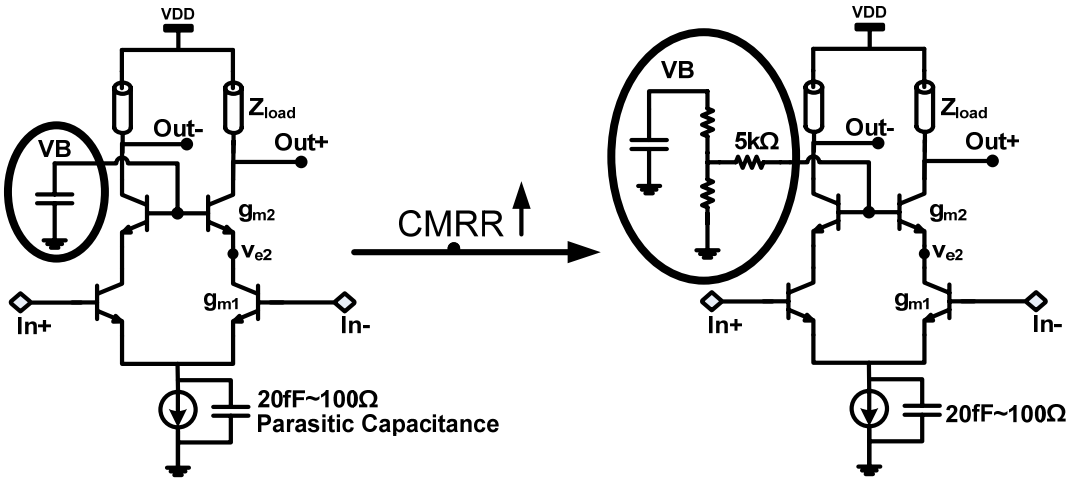


Figure 4.24 Cascode-node biasing

II— To further minimize the common mode signal and increase the CMRR, the cascode node of the differential pair is not directly connected to a bias voltage. As is shown in Figure 4.24, this bias voltage goes through a series resistor to increase the base resistance of the differential pair for the common-mode signal. In this case, the common-mode voltage applied to the emitter of the differential pair gets divided between the base-source junction of the transistor and the series resistor shown in Figure 4.24. Consequently, this common-mode voltage is incapable of generating a strong common-

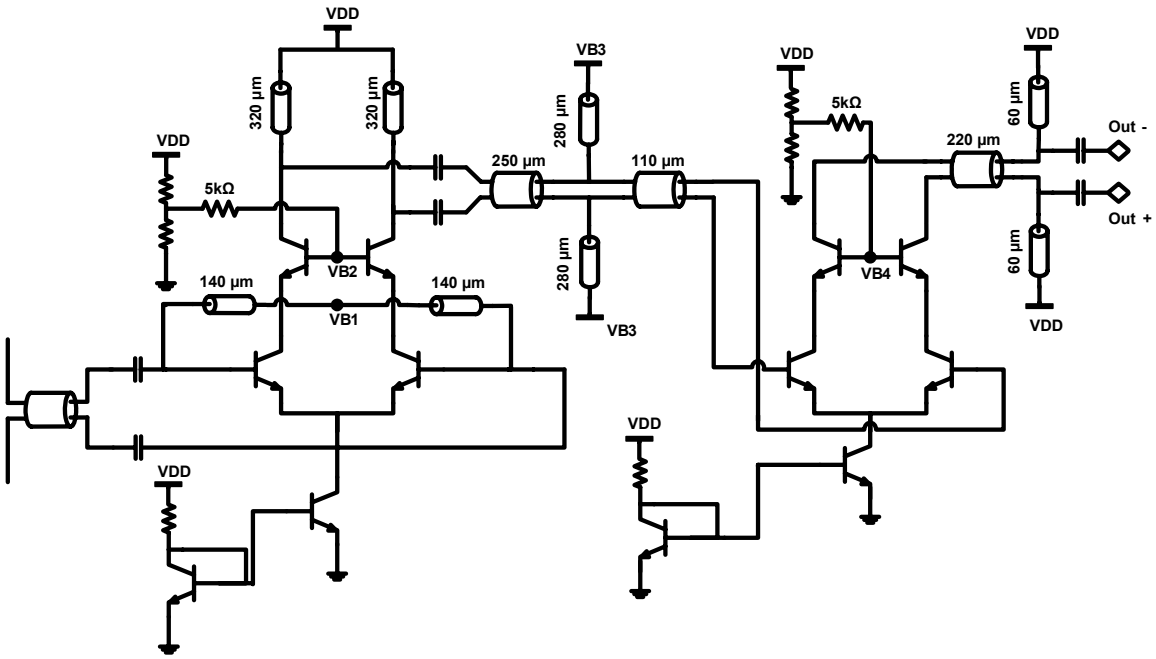
mode current at the cascode transistor. The following equations show the relation between the current of the cascode transistor and the common-mode voltage of the emitter node:

$$i_{c2} = -\frac{g_{m2}}{1 + g_{m2} R_b / \beta_2} v_{e2} \quad \text{(with the series base resistance(R}_b\text{))} \tag{4.7}$$

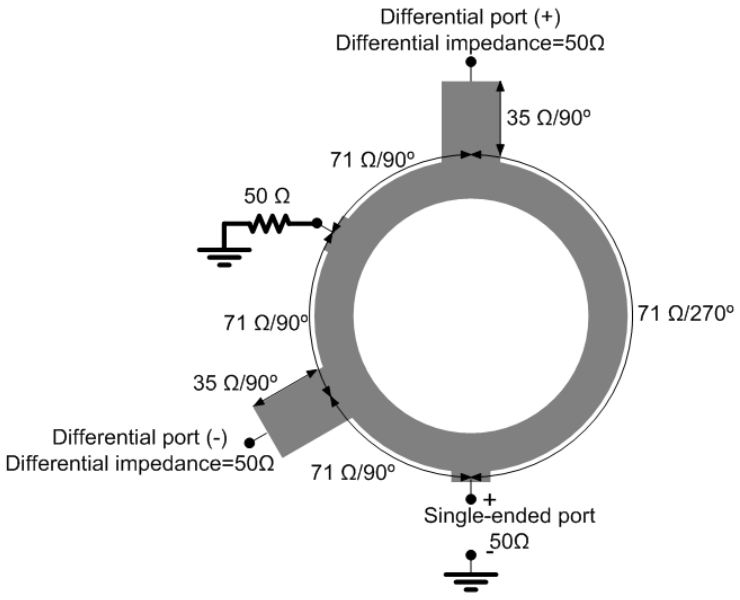
$$i_{c2} = -g_{m2} v_{e2} \quad \text{(with-out the series base resistance)} . \tag{4.8}$$

By using the above methods, the CMRR of the LNA is increased by a factor of 20 dB. The schematic of the two-stage LNA is shown in Figure 4.25(a). This differential cascode LNA driven by a differential dipole antenna uses shunt and series t-lines for impedance matching. The differential input impedance of the LNA and its differential output impedance are designed to be 50 Ω and 100 Ω respectively at 77 GHz. To bias the drain of the cascode transistors and base of the differential pairs, the stub transmission line is used. The V_{DD} voltage is brought off-chip by using wire bonds, and bias voltage is generated by using a diode-connected transistor, as shown in Figure 4.25(a). To minimize the effect of the wirebond, on-chip MIM capacitors are used at the end of the stub transmission lines to introduce short impedance to ground at 77 GHz. The size of each capacitor is about 1 pF. To be able to independently bias the two stages, AC-coupling MIM capacitors are used, as shown in Figure 4.25(a). In order to characterize LNA performance independently, a single-ended-to-differential converter (balun) following a λ/4 t-line is placed at the front of the standalone LNA test structure, as shown in Figure 4.25(b). This combined structure converts the differential 50 Ω input impedance of the

LNA to single-ended 50 Ω impedance which can be easily driven by a single-ended 50 Ω waveguide probe.



(a)



(b)

Figure 4.25 (a) 77 GHz LNA schematic. (b) Schematic of the 77 GHz balun

In the millimeter frequencies, it is very important to carefully calculate all of the parasitic capacitances and inductances. At 77 GHz, a 20 fF parasitic capacitance is equivalent to an imaginary impedance of –100Ω which can significantly affect the matching. At this frequency, a 100 pH parasitic inductance translates to an imaginary impedance of 50 Ω. To extract the parasitic capacitance of the transistors and metal layers, we have used Cadence [20], but to simulate the transmission line structures, IE3D [9] is used. The S-parameters of the EM-structures extracted from IE3D are inserted in Cadence and the complete LNA including the parasitic elements is simulated using Cadence.

Figure 4.26 shows the IE3D layout of the inter-stage transmission lines and Figure 4.27 shows the current density on these lines. Six ports are defined at the input, output, and biasing nodes of the inter-stage transmission lines. The 6×6 *S*-parameter matrix generated by IE3D is imported into Cadence. In Figure 4.26, ports 1 and 2 are the input ports (output of the 1st stage), ports 3 and 4 are the output ports (input of the 2nd stage), and ports 5 and 6 are used to supply the bias voltage (V_{b3} in Figure 4.25(a)).

The IE3D layout and the current density of the output transmission lines are shown in Figure 4.28 and Figure 4.29, respectively. In Figure 4.28, ports 1 and 2 are the input ports (ouput of the second stage), ports 3 and 4 are the output ports (output of the LNA), and ports 5 and 6 are connected to V_{DD}. MIM capacitors are used to connect ports 5 and 6 to the ground. For DC signals, these nodes are connected to V_{DD} but for RF signals, these nodes are shorted to the ground. Output of the LNA, ports 3 and 4 in Figure 4.28, which

has a differential impedance of 100 Ω, is matched to the 100 Ω differential input impedance of the mixer.

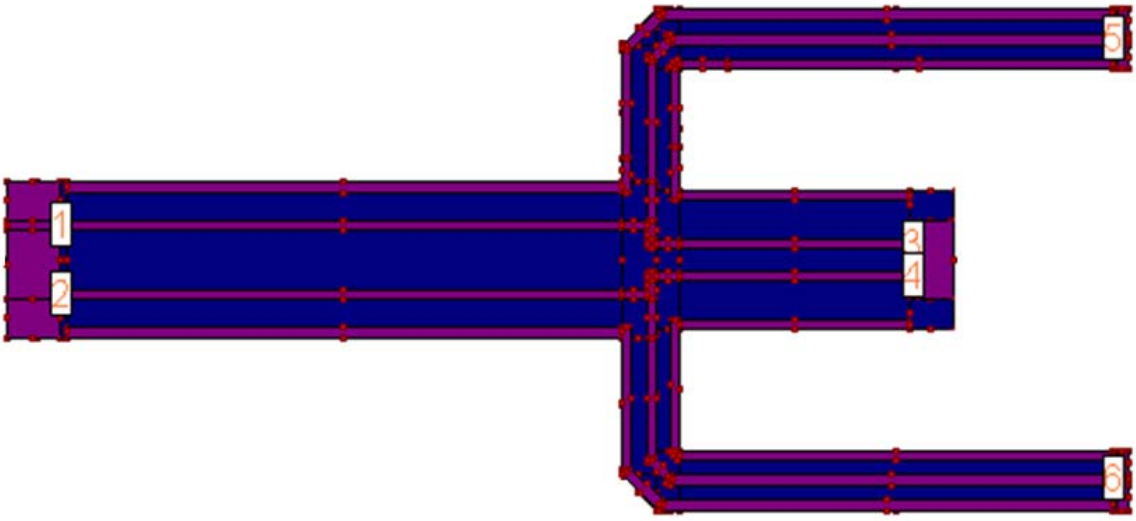


Figure 4.26 IE3D layout of the inter-stage transmission lines

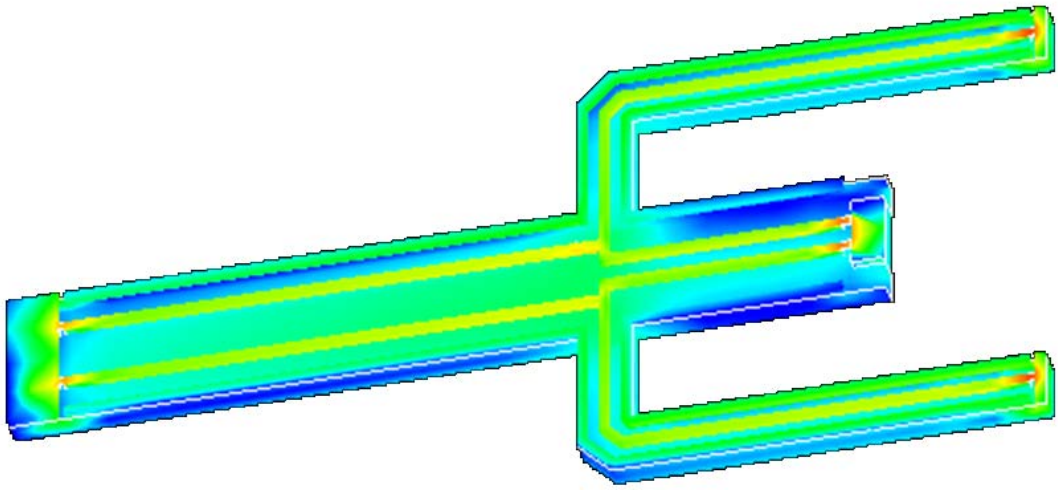


Figure 4.27 Current density on the inter-stage transmission lines

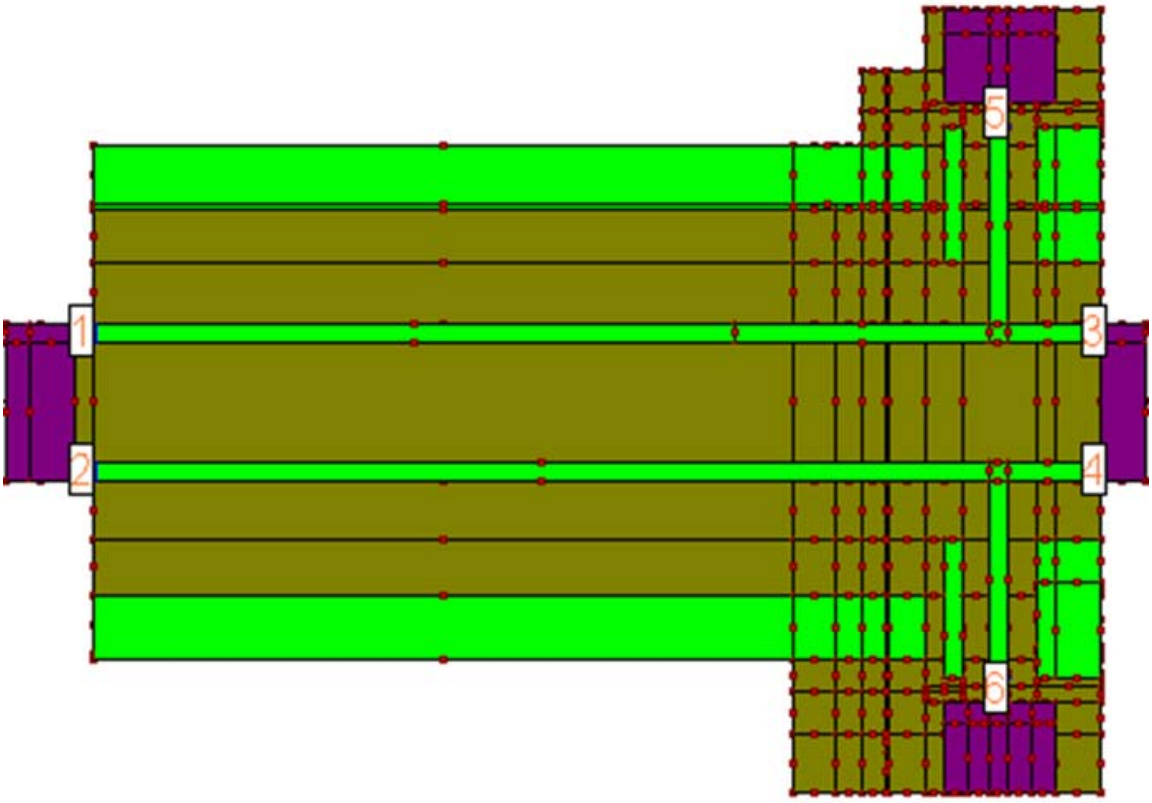


Figure 4.28 IE3D layout of the output transmission lines

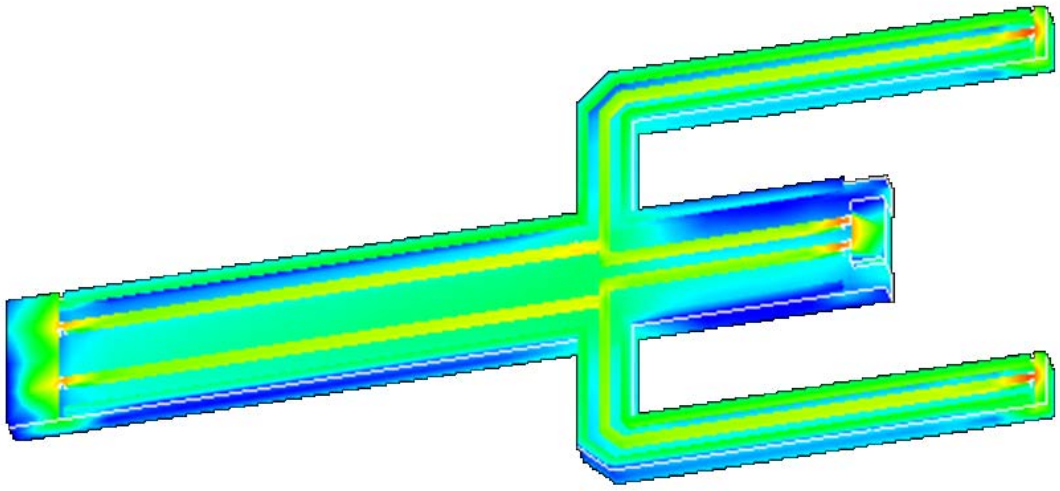


Figure 4.29 Current density on the output transmission lines

Figure 4.30 shows the IE3D layout of the on-chip dipole antenna and the input pads. The LNA is modeled as a large sheet of ground layer. The input of the LNA is connected to the on-chip dipole antenna as well as the input pads. To characterize the performance of the receiver using wafer probing, laser trimming is used to disconnect the dipole antenna from the LNA input and the on-chip pads. To measure the receiver performance with on-chip dipole antennas, laser trimming is used to disconnect the on-chip pads.

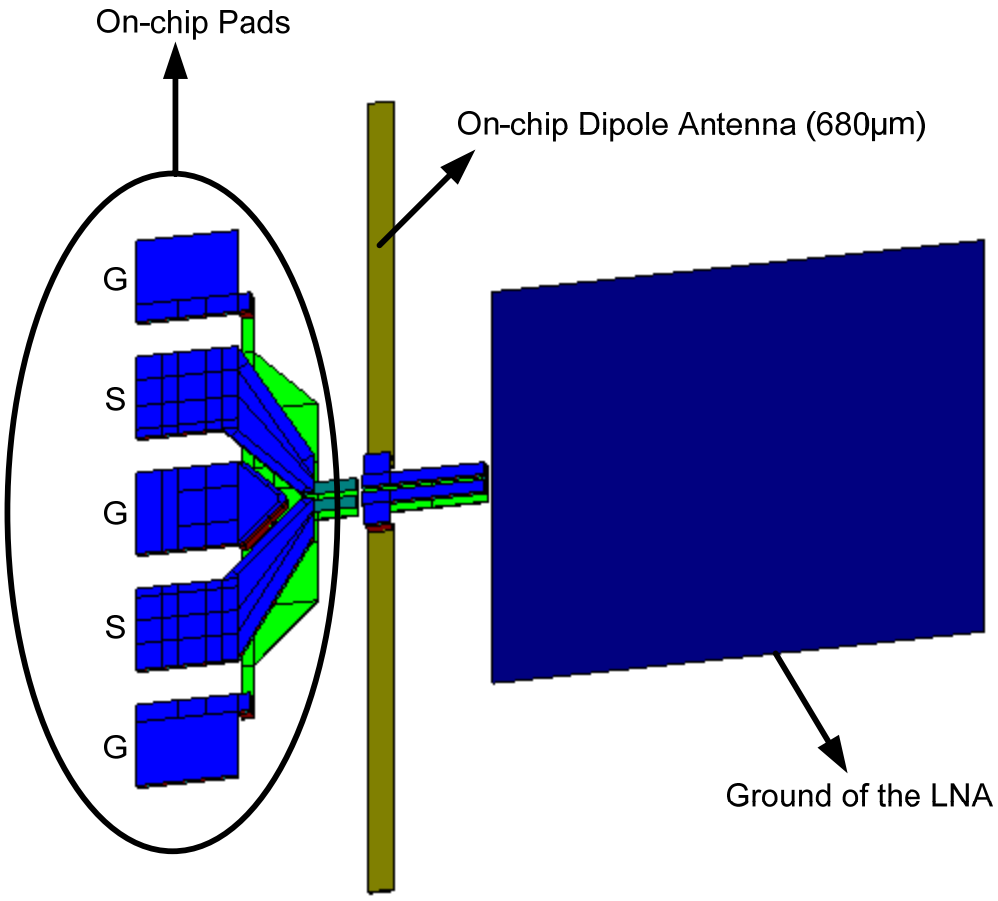


Figure 4.30 Dipole antenna with on-chip pads

Figure 4.31 shows the simulated input conductance of the antenna for the configuration shown in Figure 4.30. In this simulation, the on-chip pads are included in the measurement but they are not electrically connected to the dipole input. Based on this simulation, at 80 GHz, the on-chip antenna has a real input impedance of about 40 Ω.

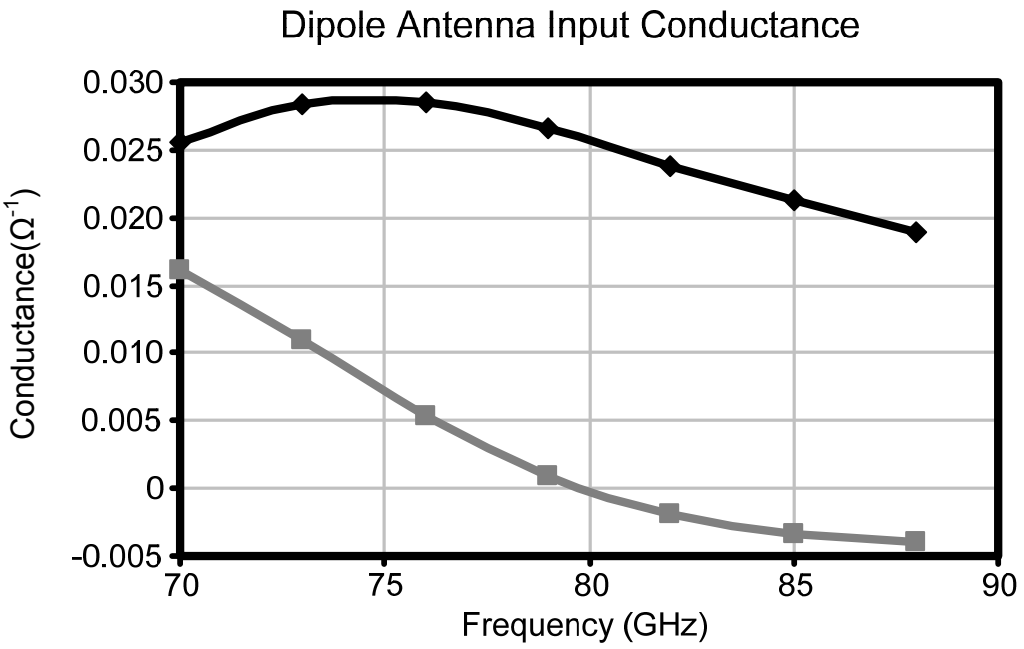


Figure 4.31 Input conductance of the dipole antenna

Distributed Active Combiner[6] **[19]—** The 4-path 26-GHz signals are combined through a distributed active combining amplifier, as shown in Figure 4.32. Emitter resistive degeneration is implemented at the input transconductors to improve the linearity, and accordingly the dynamic range, of the system. The current outputs of the transconductors are routed to the combining node via a symmetric two-stage binary structure. A pair of cascode transistors is inserted at each combining junction to isolate the input and output ports, thereby improving the overall stability of the amplifier. The

[6]Designed by Dr. Xian Guan.

total length of each routing transmission line, T_1, is 340 μm and that of T_2 is roughly 2.55 mm. Both T_1 and T_2 use a differential T-line structure with ground and side metal shields to minimize the substrate loss and cross coupling. Matched transmission line terminations are obtained by choosing the appropriate bias current so that the conductance g_m of the cascade transistors is matched to the real t-line conductance. For the operating frequency of ω_0, the imaginary part of the emitter-base admittance, $j\omega_0 C_\pi$, is much smaller than g_m if the transistor transition frequency ω_T is much higher than ω_0. Therefore, a matched termination can be achieved even without additional passive tuning. In this work, we dedicated 1 mA dc bias current to each branch. T_1 is designed to exhibit 64 Ω odd-mode impedance, while T_2 has an odd impedance of 32 Ω. Simulations show that the return loss is better than 10 dB at the terminations of the transmission lines at both levels. The differential output of the amplifier is loaded with an LC tank with parallel resistors to improve the bandwidth.

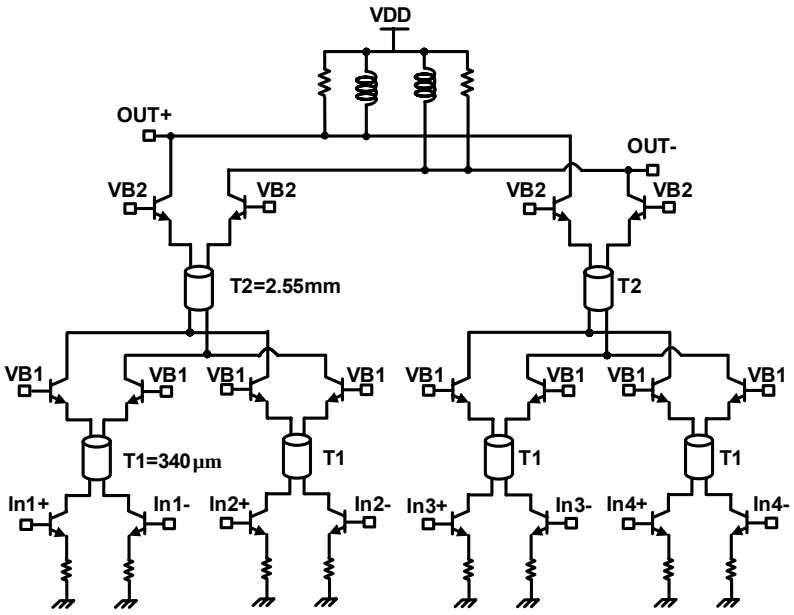


Figure 4.32 A 26 GHz combining amplifier [19]

Down-converter Mixer[7][19]— A pair of double-balanced mixers driven by quadrature LO signals are used to perform frequency translation from 26 GHz to baseband, one of which is shown in Figure 4.33. The 26 GHz signals are coupled into the mixer transconductance stage through 0.9 pF MIM capacitors. The input differential pair is degenerated with 30 Ω resistors at the emitter to improve linearity.

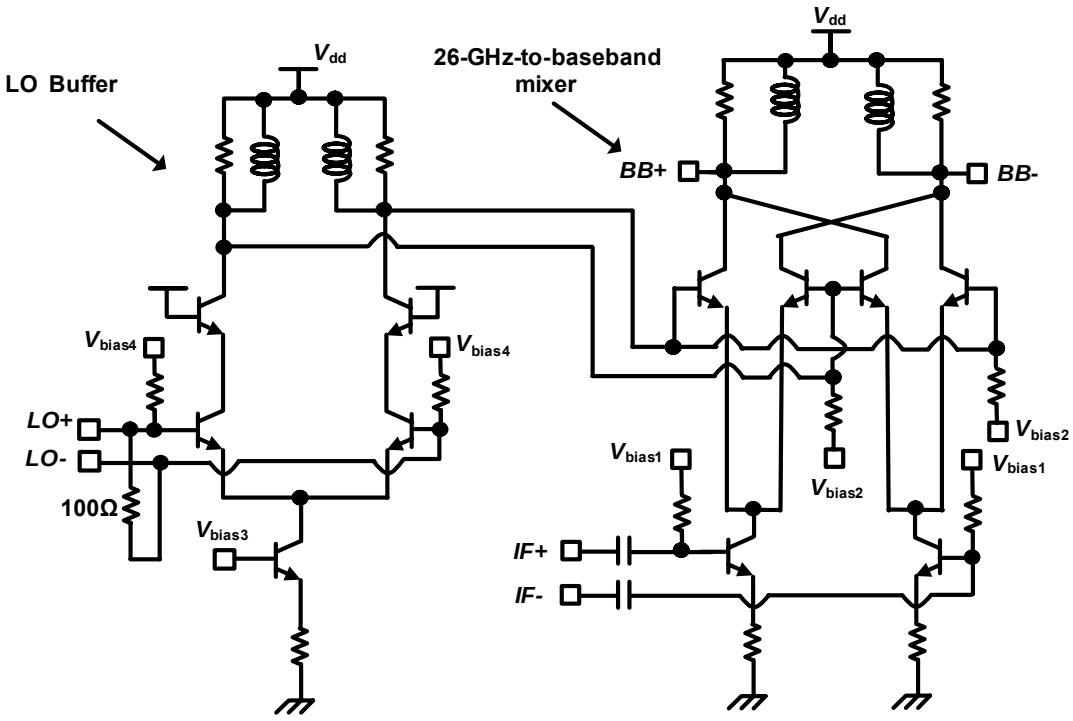


Figure 4.33 26-GHz-to-baseband mixer and 26 GHz LO buffer [19]

The LO port of the mixer is fed by a 26 GHz buffer, which is used to compensate the LO signal loss through the distribution network, ensuring the differential LO amplitude applied to the mixer is larger than 200 mV so that the mixer gain is saturated. The input matching of the LO buffer is provided by an explicit 100 Ω resistor directly connected between the differential inputs. Although a matching network composed of inductors and

[7]Designed by Dr. Xian Guan.

capacitors can provide additional voltage gain, this solution is prohibited by the limited silicon area. The LO buffer is loaded with 0.6 nH spiral inductors and 320 Ω de-Q resistors, providing a gain of 15 dB. With 280 Ω load resistor, the second mixer achieves 6 dB conversion gain and 8 GHz IF-referred bandwidth. The mixer core consumes 4 mA dc current and the LO buffer drains 1 mA.

An on-chip differential voltage-controlled oscillator generates the first LO signals at 52 GHz, which are symmetrically distributed to each RF mixer using differential transmission lines (T-line). A network of LO buffers is used for LO signal distribution to compensate for the T-line loss and to ensure hard switching of the mixers. The continuous analog phase shifting is performed locally at each RF mixer by an analog phase rotator, which realizes continuous beam steering while accurately compensating for the phase and amplitude deviations. The quadrature second LO is obtained by dividing the first LO frequency by 2. A frequency divider chain is used to further divide the second LO frequency down to 50 MHz to be locked to an external reference frequency.

4.2.4 Transmission-Line-Based Design

The conductor-backed coplanar waveguide (CBCPW) structure, shown in Figure 4.34, is used for impedance matching. The use of vias to connect back and side ground planes eliminates unwanted parallel-plate modes. Figure 4.35 shows the magnetic field distribution in the transmission line, simulated with Ansoft HFSS 3-D field solver. The characteristic impedance of the transmission line in this simulation is 50 Ω. The bottom plate carries very little current (small tangential component of the magnetic field) while the side shield carries most of the return current.

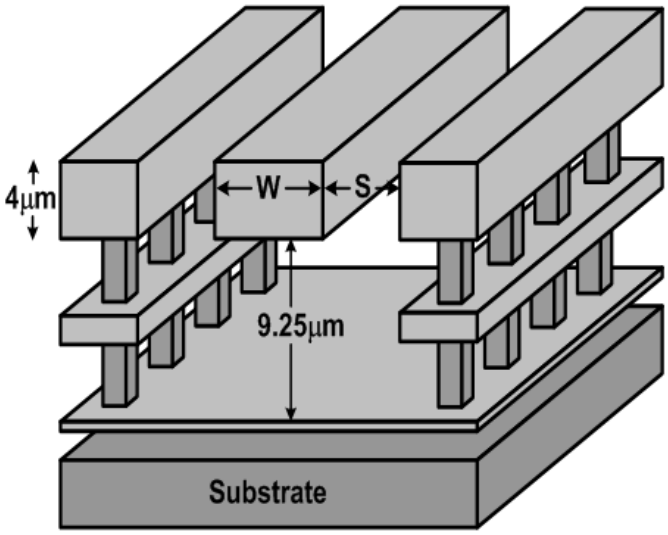


Figure 4.34 On-chip transmission line

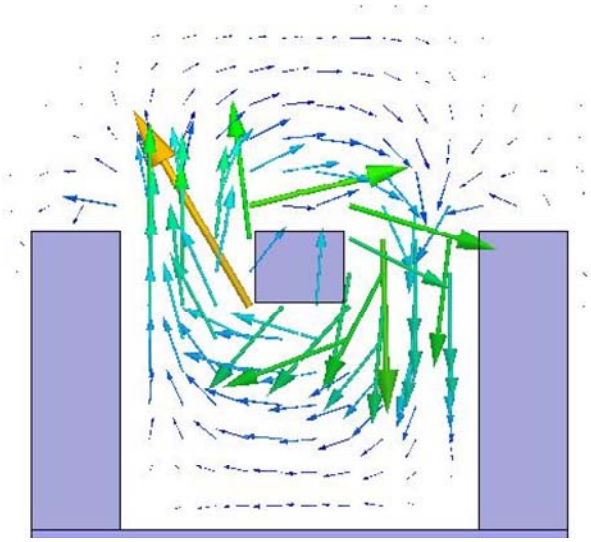


Figure 4.35 Magnetic field distribution in the tub transmission line

The tub shape reduces surface wave propagation in the silicon substrate, improving isolation between lines. Figure 4.36 shows the isolation between two adjacent 50 Ω lines versus their center-to-center spacing, simulated using IE3D [9]. The lines are implemented using the top three metals of the process. The side shields increase isolation by more than 20 dB. The coupling in the secondary line is larger in the direction opposite

to the wave direction of the primary line. There is a trade-off between the isolation of lines and their insertion loss. Since the side-shield increases unit length capacitance, in order to keep the characteristic impedance constant, the width of the line should be reduced. This increases the loss of the transmission line. The 50 Ω line with shield has a loss of 0.75 dB/mm and a quality factor of 10 at 77 GHz.

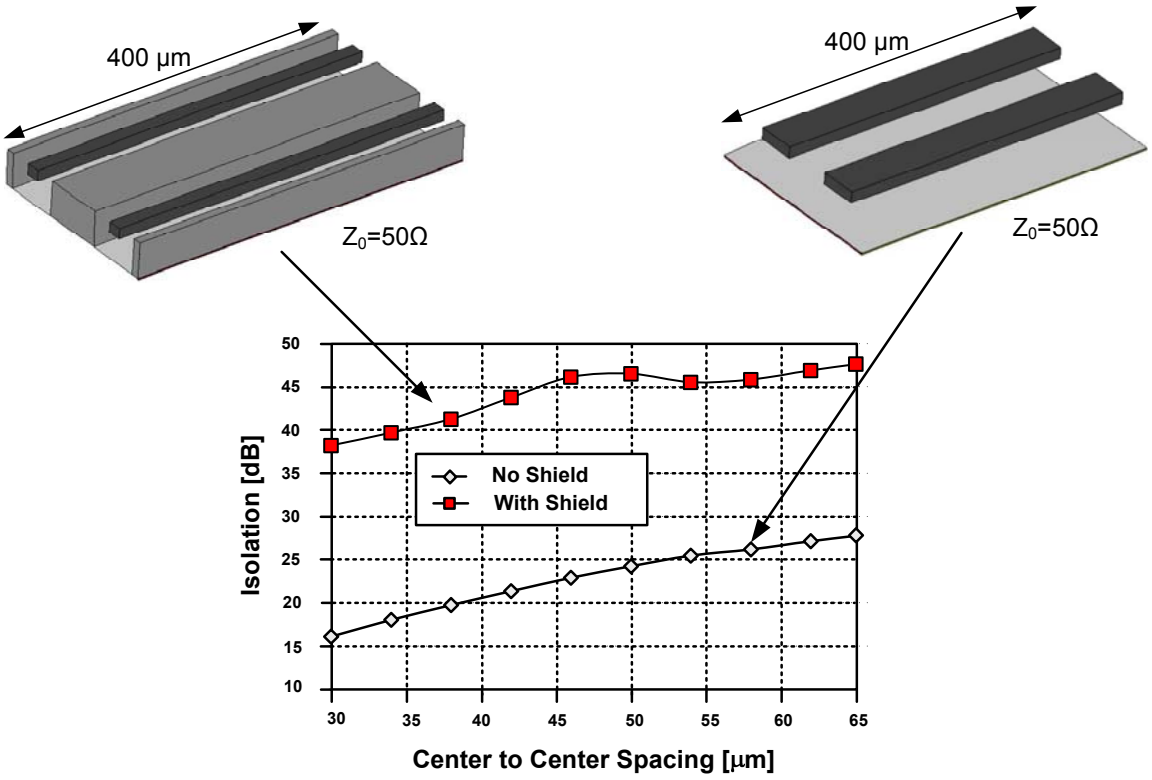


Figure 4.36 Isolation between two lines (with shield and without shield)

4.2.5 LNA Layout

Figure 4.37 shows the standalone LNA layout. Parasitic capacitances of the transistors are carefully calculated. Transmission lines and MIM capacitors are used for matching purposes. Due to the low quality factor of inductors at these frequencies and their coupling to the substrate, no inductor is used. LNA transmission lines are meandered to minimize the area and ease the floor-planning of the whole transceiver chip.

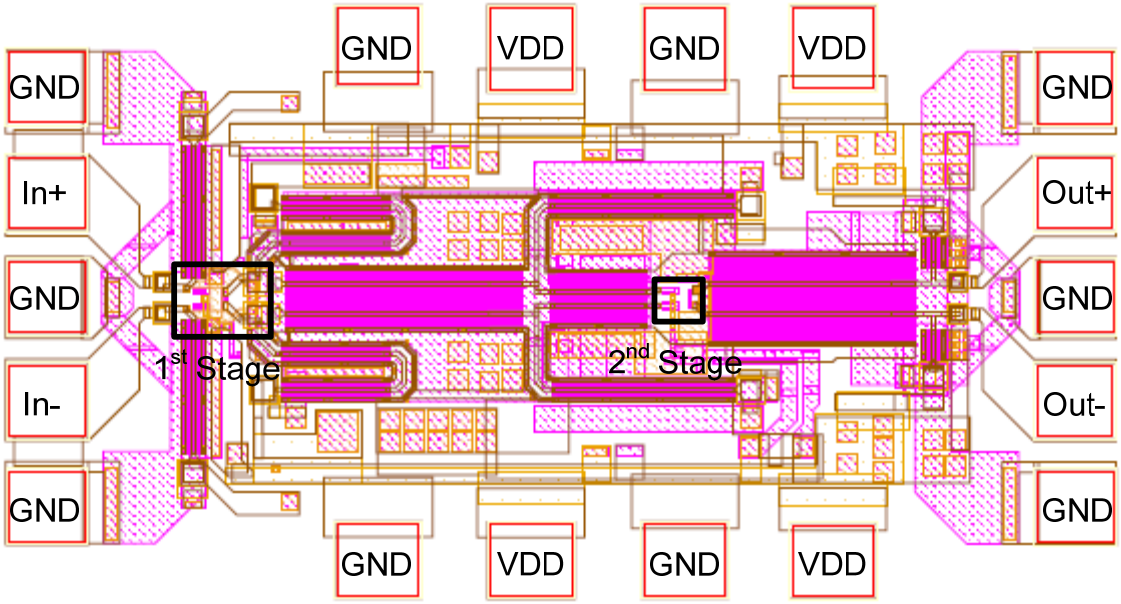


Figure 4.37 77 GHz LNA layout

The 1st and 2nd stages of the LNA are shown in Figure 4.37. Coupled-wire transmission lines carry the differential signal from the first stage to the second one, and single-wire transmission lines behave as stub lines maintaining the impedance match.

4.2.6 System Layout

Layout of the system is shown in Figure 4.38. The system is designed to enable an independent test of the transmitter and receiver. Biasing pins of the transmitter are located at the receiver as well as at the top and bottom sides of the chip. In the setup used for transmitter measurement, there are no wire-bonds on the transmitter side of the chip and high-frequency probes can easily couple the W-band signal to the pads on the transmitter side of the chip. On the other side, biasing pins of the receiver are located at the transmitter, as well as the top and bottom of the chip (Figure 4.38). This eases the receiver measurement in which there are no wire bonds on the receiver side of the chip and high-frequency probes can couple the signal to the pads on the receiver side.

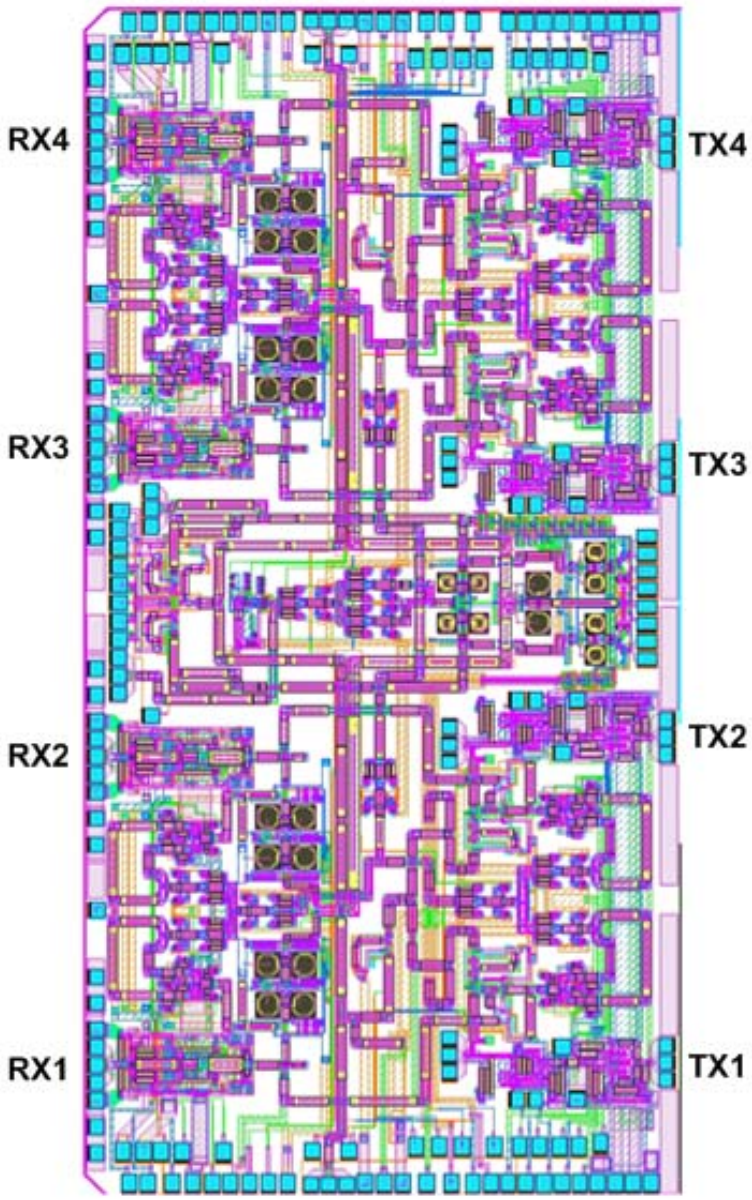


Figure 4.38 77 GHz system layout

4.3 Measurement of the 77 GHz Phased Array Receiver

The 77 GHz phased array is fabricated in a 130 nm SiGe BiCMOS process with a f_T of 200 GHz for SiGe HBT devices. The receiver section occupies roughly 9 mm^2 of the 6.8mm × 3.8mm total chip. Figure 4.39 shows the die micrograph.

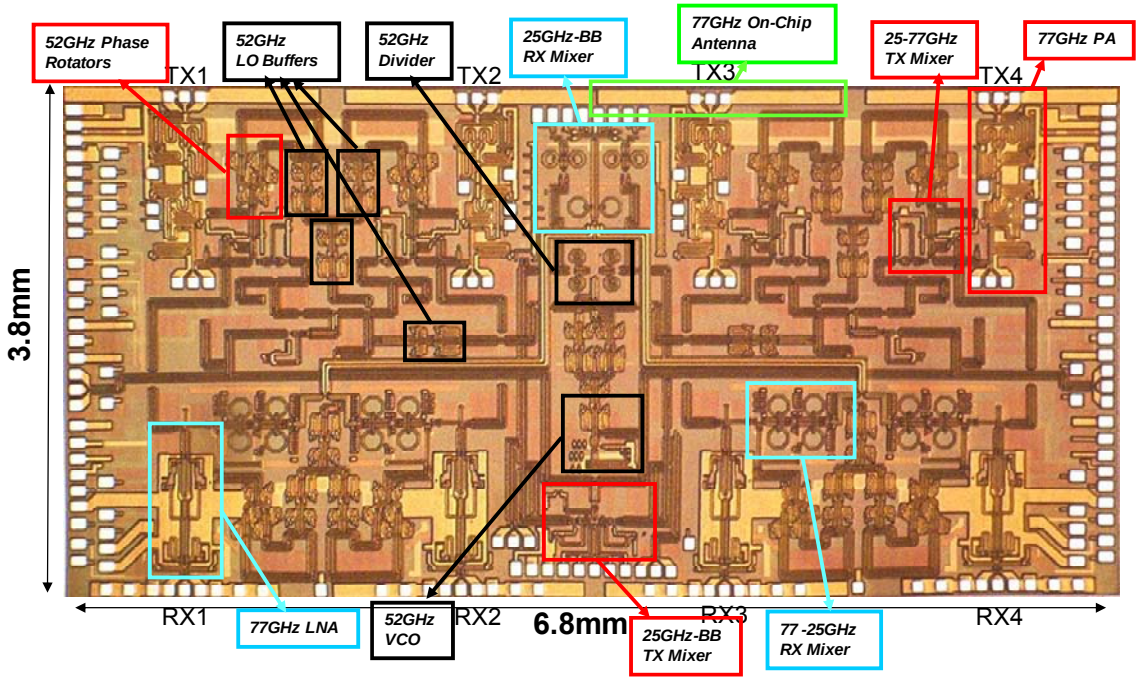


Figure 4.39 77 GHz chip micrograph

To accurately characterize the receiver performance, a stand-alone LNA with integrated balun is measured. One of the important parameters necessary for accurate de-embedding of the stand-alone LNA measurements is the loss of the balun and the following λ/4 t-line. Two identical baluns including the matching transmission lines are designed and connected together at their differential nodes, as shown in Figure 4.40.

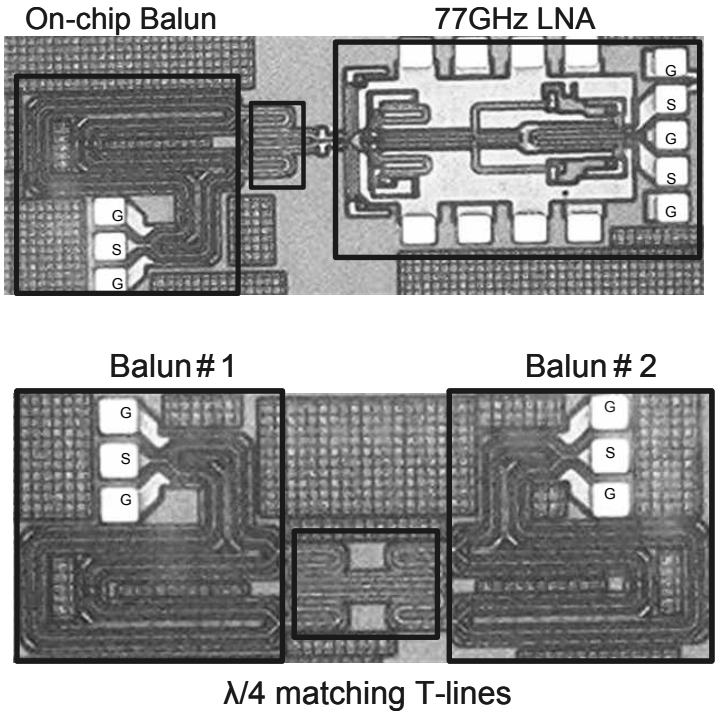


Figure 4.40 77 GHz LNA with on-chip balun

The loss of the two series identical structures is expected to be twice that of a single one. The measured balun loss and the LNA performance versus frequency are shown in Figure 4.41 and Figure 4.42, respectively.

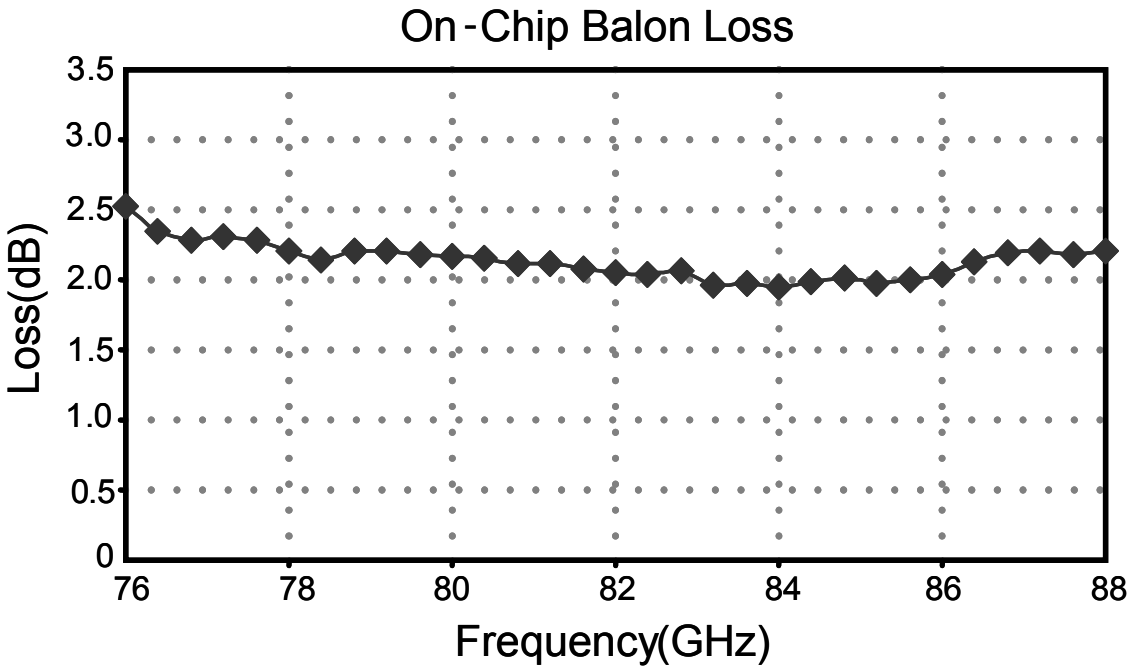


Figure 4.41 Measured on-chip balun loss

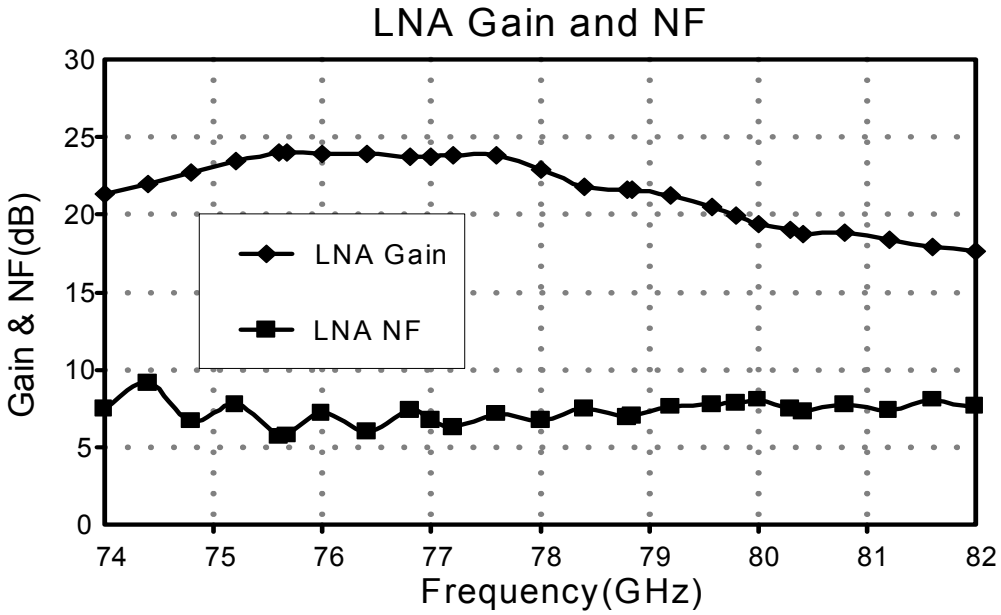


Figure 4.42 LNA gain and noise figure

Standalone LNA peak gain of 23.8 dB is measured at 77 GHz with a 3 dB bandwidth of more than 6 GHz, while the lowest noise figure of 5.7 dB is measured at 75.7 GHz. The LNA consumes 17.5 mA from a 3.5 V supply.

Figure 4.43 illustrates the test setup for measuring the receiver gain. The input signal at 77 GHz range is provided by a Spacek frequency quadrupler capable of delivering output frequency from 60–90 GHz. The input of the frequency quadrupler is supplied by signal generator working up to 26.5 GHz. The power of the input signal can be adjusted by a variable linear attenuator. A WR-12 planar wafer probe is used to feed the single-ended signal to the LNA input. The external connections between W-band components are built using WR-12 waveguides. The microwave input power is calibrated up to the probe tip using Agilent E4418B power meter with a W-band power sensor. An exclusive OR (XOR) logic gate acting as a phase detector and a first-order RC low-pass filter complete the PLL which locks the phase and frequency of the 52 GHz VCO to a 50 MHz reference provided by signal generator. The baseband outputs are characterized using Agilent 4448A spectrum analyzer. The same setup is also used for receiver noise figure measurement except the RF inputs are replaced with a W-band noise source.

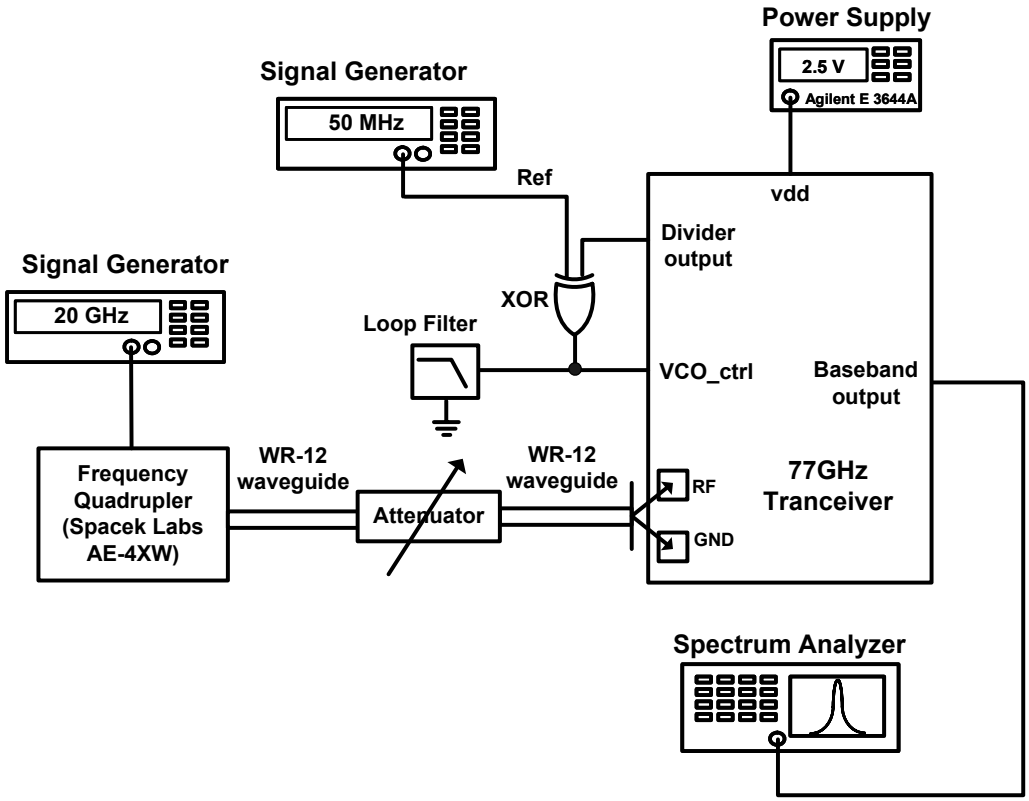


Figure 4.43 Receiver test setup

The stand-alone electrical performance of the receiver is characterized after laser trimming the antennas. A 37 dB single-path receiver gain (Figure 4.44) is measured at 79.8 GHz with a 2 GHz bandwidth, corresponding to an inferred array gain of 49 dB. The minimum receiver noise figure is measured to be 8 dB (at 78.8 GHz).

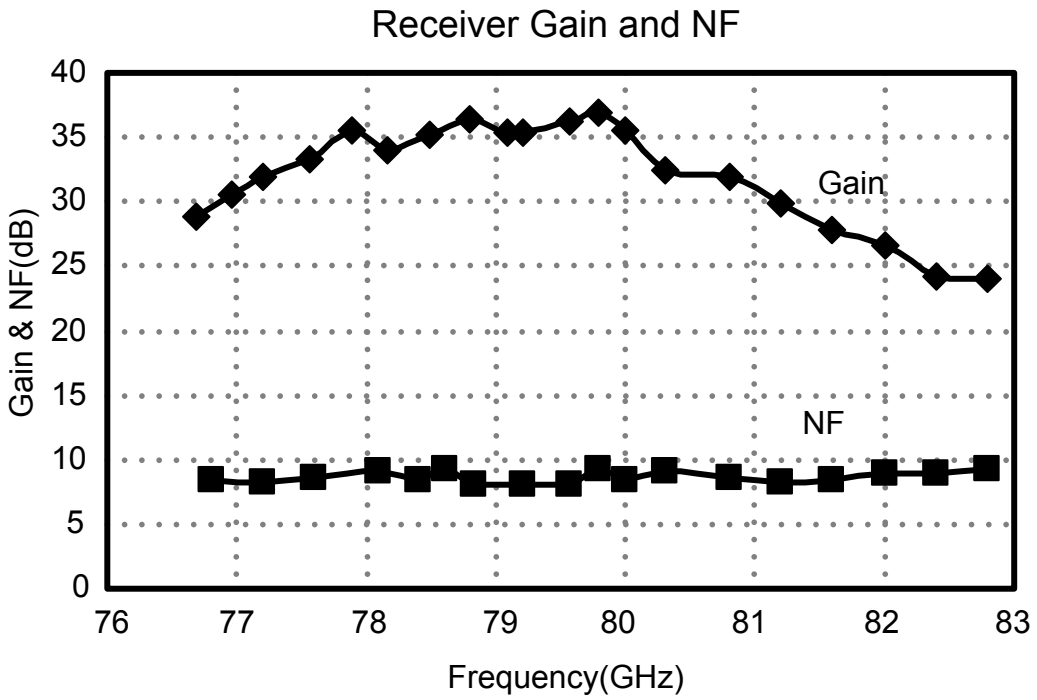


Figure 4.44 Single-path receiver gain and noise figure

Antenna Measurement— The chip micrograph and a magnified picture of the layout of the phased array transceiver with integrated antennas and the pads are shown in Figure 4.45.

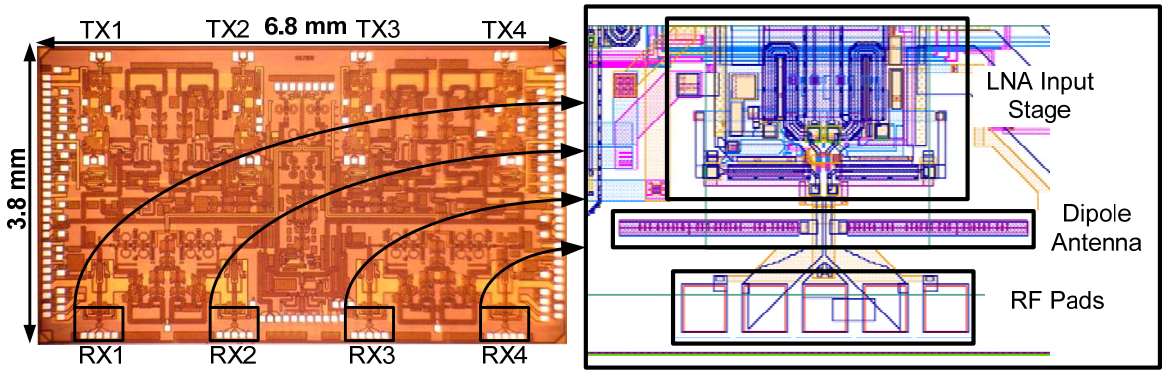


Figure 4.45 Chip micrograph and integrated antennas

A 2-axis spherical far field measurement technique is utilized to measure the radiation pattern while a W-band horn antenna is used to irradiate the integrated dipoles. The 3-D

measured patterns of two middle antennas are shown in Figure 4.46 where $-40° < \theta < 40°$ and $0° < \varphi < 180°$. In this plot X=Sin(θ)×Cos(φ), Y= Sin(θ)×Sin(φ), and Z=Gain(dB). Maximum peak gain of about +8 dB is achieved in this measurement. As seen in Figure 4.46 due to the off-axis properties, [7], the peaks of two antennas occur at two different directions.

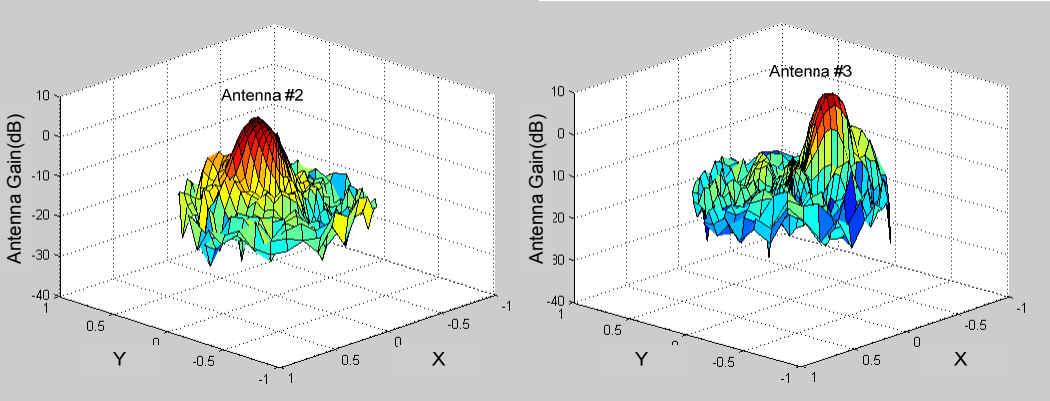


Figure 4.46 Radiation pattern (X=Sin(θ)×Cos(φ), Y= Sin(θ)×Sin(φ), and Z=Gain(dB))

Figure 4.47 shows the E-plane pattern of the two middle antennas. The effective radius and extension length of the lens are 12.7 mm and 1.6 mm respectively (Figure 4.48). Due to the off-axis properties, these two peaks occur at different angles with spacing of about 25°. The theoretical value [7] of this spacing is 24° which is very similar to what we measured. Based on the measured results, the lens improves the antenna gain by a factor of 10 dB.

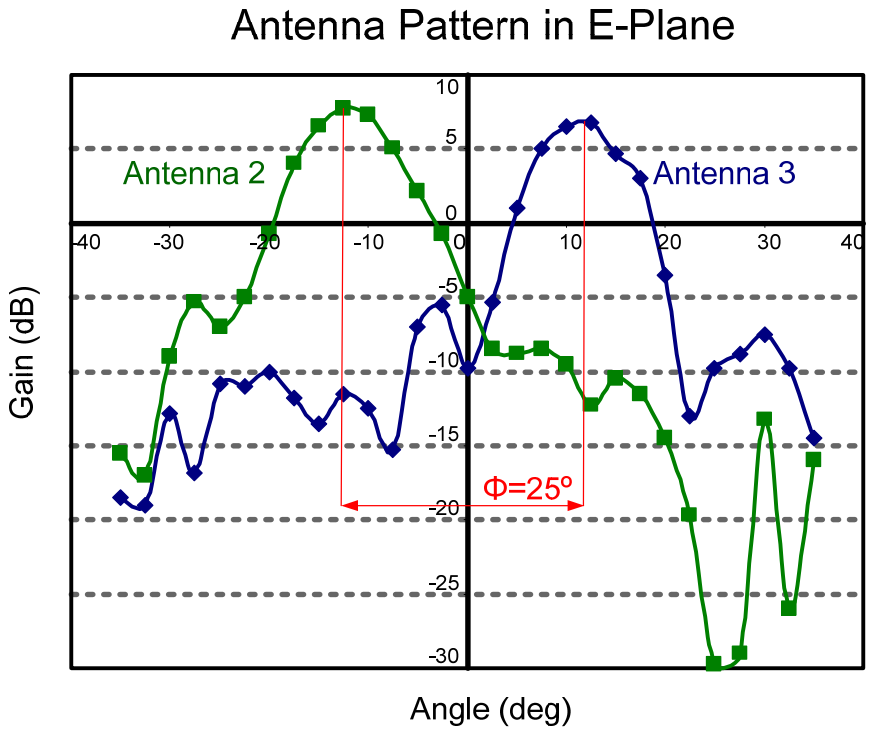


Figure 4.47 E-plane pattern of two middle antennas

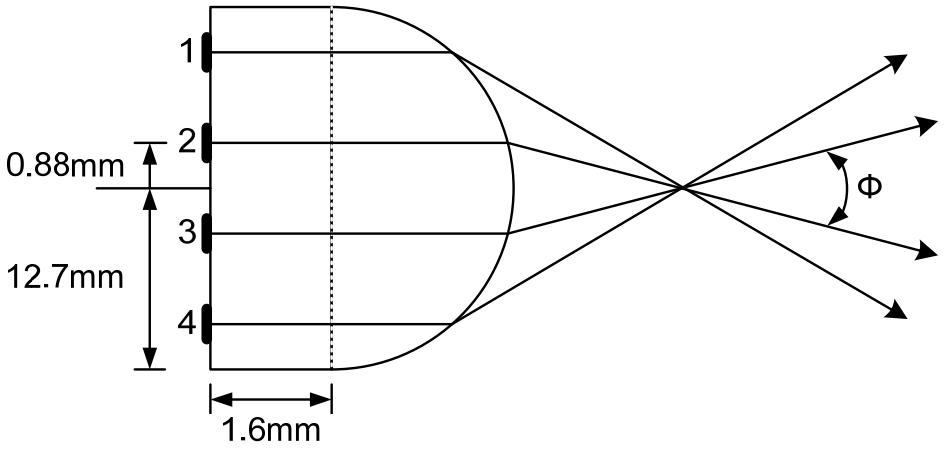


Figure 4.48 Lens dimensions

4.4 Chapter Summary

In this chapter, phased array systems are reviewed and some of the important characteristics of these systems are explained. The details of the design, layout, and measurement the first fully integrated 77 GHz 4-channel phased-array transceiver with on-chip antennas is presented. In the receiver, a two-step down-conversion scheme is used with a single VCO. The signal combining is performed using a novel distributed active combining amplifier at IF. In the LO path, a cross-coupled quadrature injection locked frequency divider (QILFD) divides the 52 GHz VCO frequency by a factor of 2 and is followed by a divide-by-512 divider chain. Conversion gain of more than 37 dB, 2 GHz band-width, and 8 dB NF are achieved. W-band integrated dipole antennas and a four-channel phased transceiver are implemented in IBM 130 nm silicon germanium BiCMOS process. The chip requires no high-frequency electrical connection to the outside world. It includes the complete receiver, transmitter, signal generation blocks, phase shifters, and on-chip dipole antennas. A hemispherical silicon lens with a diameter of about one inch is used to minimize the substrate modes. Receiver gain is measured separately using on-chip probing. A maximum antenna gain of about +8 dB is achieved in this measurement.

Chapter 5

Transmitter Architectures Based on Direct Antenna Modulation (DAM)

5.1 Introduction

In the late 1800s and early 1900s, wireless transmission of information started its journey. During this period, inventors such as David E. Hughes, Heinrich Hertz, Nikola Tesla, Guglielmo Marconi, Reginald Fessenden, and Edwin H. Armstrong had to work with long wavelengths due to the lack of high frequency amplifiers [21], [22]. At frequencies smaller than 10 GHz, the size of an efficient antenna is much larger than the size of the active devices used. The practical issues involved in implementing such large antennas limited the number of effective solutions for designing wireless transceivers. Among these solutions, we can mention the invention of the regenerative, heterodyne, and homodyne receivers [21] , [22]. With his invention of super-heterodyne receiver in 1918, Armstrong introduced the idea of modulating the signal at low frequencies, or baseband, and up-converting it to the RF frequency. Since that time, there have been many breakthroughs in related technologies including the invention of the transistor itself in 1947. However, there have been few fundamental changes in transceiver architectures despite the availability of revolutionary supporting technologies; most of today's high-performance systems still use ideas based on the heterodyne or homodyne architectures. Today's silicon technologies provide transistors with unity-current-gain frequencies (f_T) of greater than 200 GHz, which make it possible to implement mm-wave integrated

transceivers on a single chip [23] – [29]. At these frequencies, wavelengths are comparable to the die size and orders of magnitude smaller than those used by Armstrong to implement his first heterodyne receiver (for instance, a simple half-wavelength dipole antenna is 2.5 mm long at 60 GHz). Access to faster transistors and the ability to implement antennas on the same die as analog, RF, and digital circuits is motivation enough to reevaluate classical receiver architectures and investigate new and fundamentally different architectures that deal with system problems across multiple levels of abstraction.

Figure 5.1 shows the block diagram of a conventional direct-conversion transmitter. In this architecture, the in-phase (I) and quadrature (Q) signals are modulated at base-band and then up-converted to radio frequency (RF). The modulated RF signal goes through a power amplifier (PA) which drives the antenna. It is noteworthy that in this conventional transmitter architecture, the *already modulated* signal couples to the antenna. As seen in Figure 5.1, in a mostly line-of-sight scheme, a receiver sitting in a side lobe of the antenna receives the same information as the receiver located at the antenna's main beam. The only differences between the receiver data at different directions are the signal power and a time shift. Therefore, given a high-sensitivity receiver it would be possible for a receiver in an unintended direction to eavesdrop on the communication.

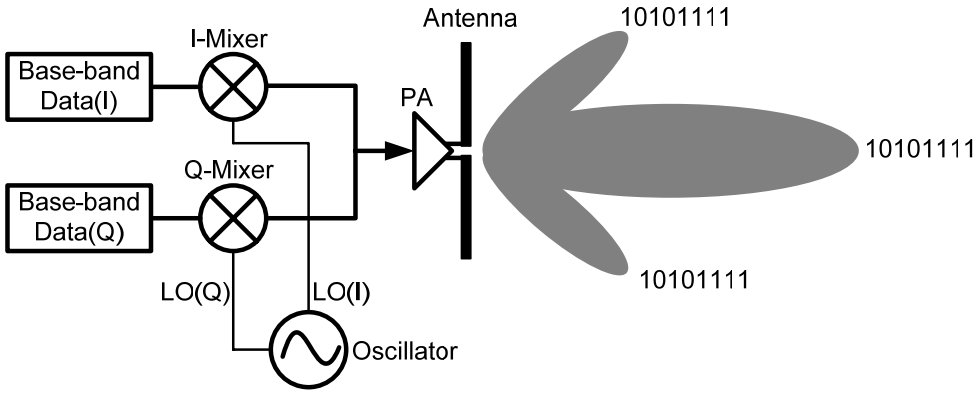


Figure 5.1 Conventional transmitter architecture

We will see in the rest of this chapter how DAM can be used to overcome the security challenge using a direction-dependent information transmission. In Section 5.2 we will discuss the new direct antenna modulation (DAM) approach. In Sections 5.3 and 5.4 we will discuss the switch-based DAM, while in Section 5.5 we will investigate the varactor-based version. We will review the circuit level details in Section 5.6.

5.2 Concept of Direct Antenna Modulation

Figure 5.2 illustrates modulation at base band. In the absence of multipath, any change at the base band appears in the desired direction as well as the undesired direction (the only difference is the power level and a delay), as illustrated symbolically by the move from point 1 to point 2. This is because the signal is already modulated before the antenna and because the antenna pattern does not change quickly.

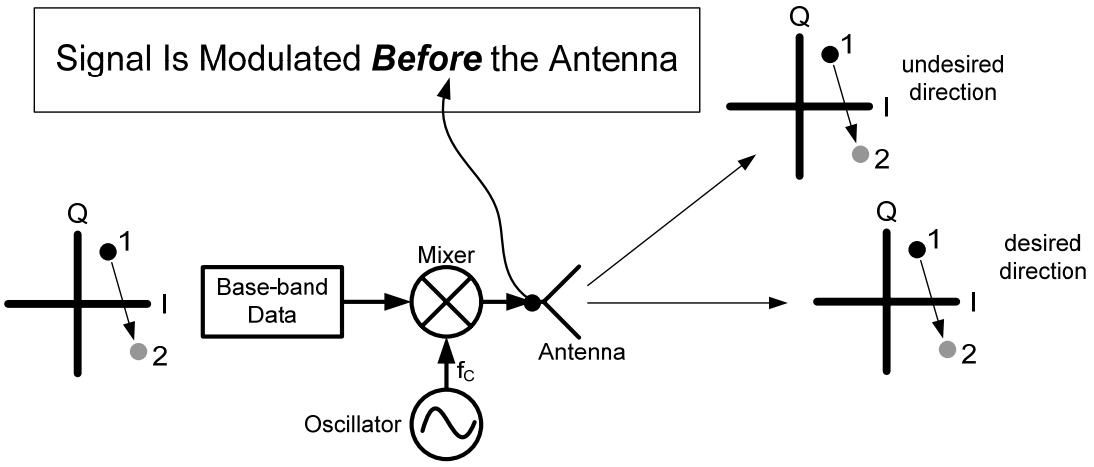


Figure 5.2 Modulation at base band

The proposed direct antenna modulation (DAM) technique, illustrated in Figure 5.3, is fundamentally different from the one shown in Figure 5.2. The DAM transmitter utilizes only a locked RF signal source and a PA to drive the antenna. In this scheme we modulate the phase and amplitude by changing the antenna characteristics and hence its pattern at the symbol transmission rate. As a result, we are able to transmit different signals independently to the desired and the undesired directions, unlike the conventional architectures. It is important to realize that we need to change the antenna characteristics, its near field and far field, *at the rate of the symbol rate* in order to properly modulate the signal. Next, we introduce two alternative techniques to vary the antenna boundary conditions thereby changing the phase and amplitude of the antenna far-field pattern.

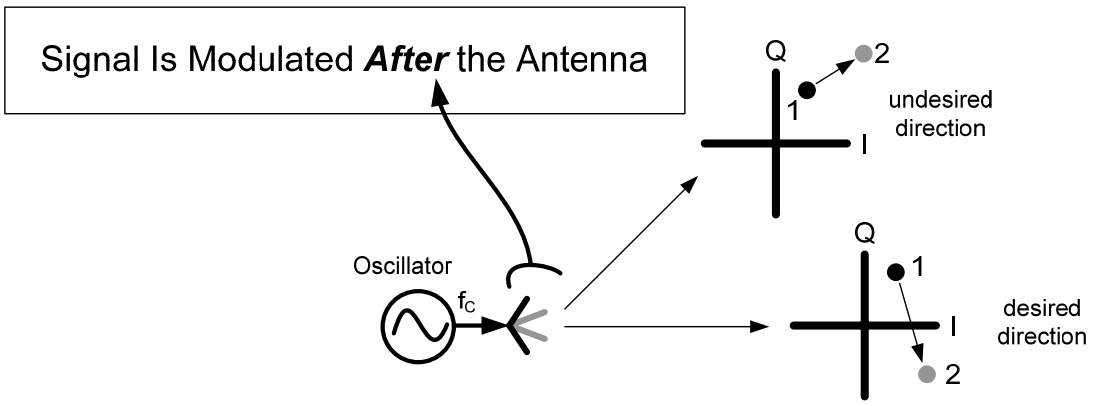


Figure 5.3 Modulation after the antenna

5.3 Switch-Based DAM

A toy example of this technique is shown in Figure 5.4. On the left side of Figure 5.4, we show a dipole antenna with an adjacent reflector. The reflector is composed of two metal pieces connected with an ideal switch. Let us say the main dipole antenna radiates a signal $A_0 \cos(\omega t + \varphi_0)$ in the z-direction, normal to the plane (bore-sight). Some part of the main signal couples to the adjacent reflector in the near field of the antenna, causing the reflector to scatter a signal $A_1 \cos(\omega t + \varphi_1)$ in the z direction. By closing the switch, we can change the reflector's effective length and scattering properties, which cause the reflected signal to have a different phase and amplitude, $A_2 \cos(\omega t + \varphi_2)$. In these two cases the far-field signal in the z-direction can be calculated as,

Open switch: $A_0 \cos(\omega t + \varphi_0) + A_1 \cos(\omega t + \varphi_1) = A' \cos(\omega t + \varphi')$ (5.1)

Closed switch: $A_0 \cos(\omega t + \varphi_0) + A_2 \cos(\omega t + \varphi_2) = A'' \cos(\omega t + \varphi'')$. (5.2)

The real and imaginary parts of the combined signal in the far-field are shown on the signal constellation diagram (Figure 5.4). The example above shows how a simple switch can change the characteristics of the reflector and hence modulate the signal.

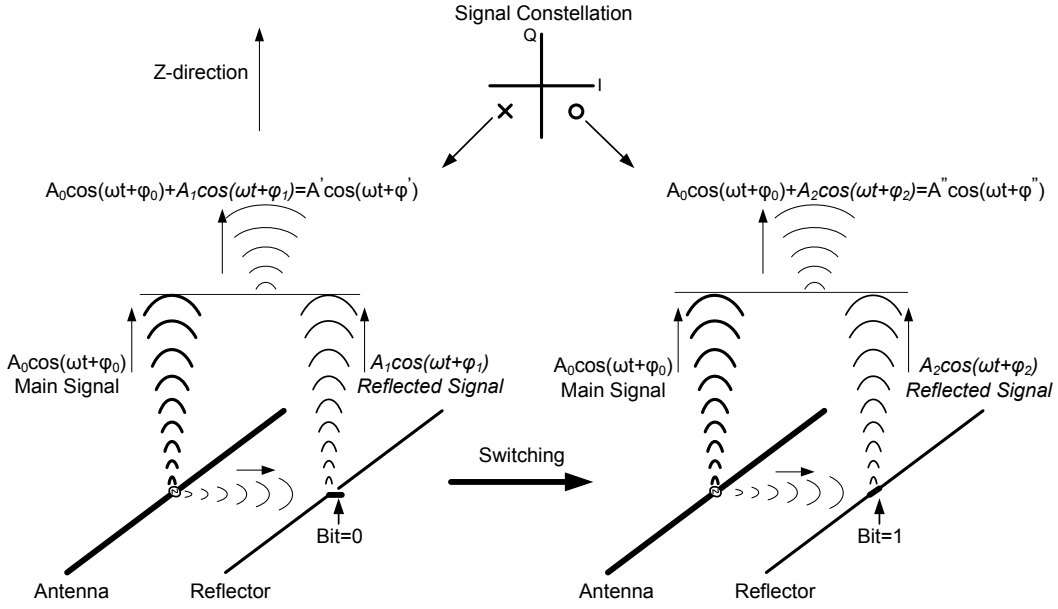


Figure 5.4 Signal modulation using switches on the reflectors

Consider a practical system in which more than two constellation points are required in order to transmit the signal at an acceptable bit-rate. The number of points on the constellation diagram can be increased by introducing multiple reflectors, each with multiple switches, as shown in Figure 5.5. For N total switches, 2^N constellation points can be generated. Thus, if we have a sufficiently large number of switches, it is possible to generate a very large number of constellation points. This is illustrated in Figure 5.6, where 10,000 random switching combinations are simulated and real and imaginary parts of the induced voltage on a receiving dipole antenna located at the far field are plotted. Five reflectors are placed at each side of the antenna and 9 ideal switches are placed on

each reflector. Based on this simulation, it is possible to cover all of the four quadrants on the signal constellation diagram.

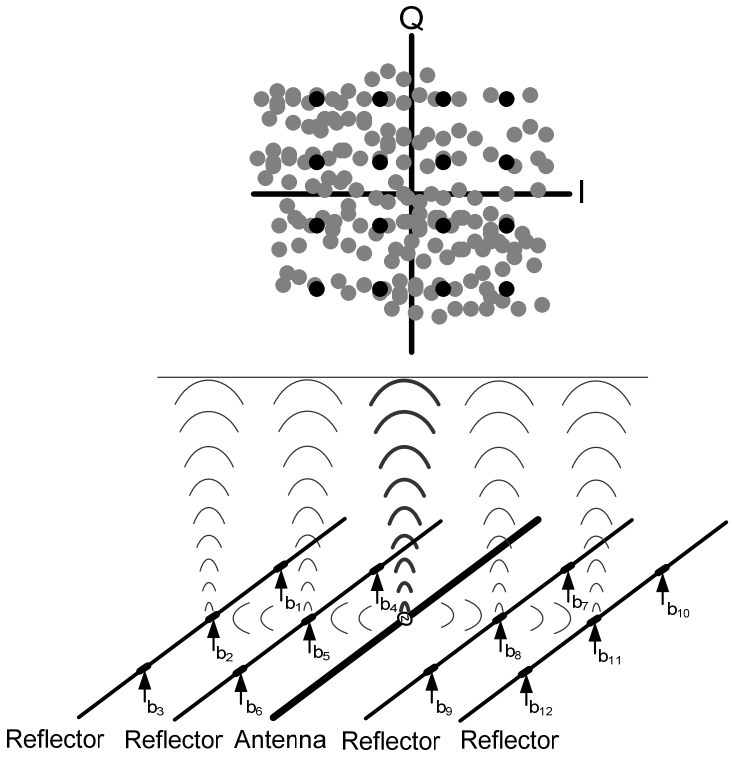


Figure 5.5 Arbitrary signal modulation

It should be noted that only the carrier signal goes through the PA, so there is no need to design a broad-band PA. A locked oscillator generates an un-modulated sinusoidal signal that drives the PA. As a result, this system is capable of transmitting broad-band information while using a narrow-band PA. This system also can utilize highly efficient switching PA in transmission of constant and non-constant envelope-modulated signals. In a silicon implementation of the switch-based DAM scheme, the switches can be implemented using small-feature-size MOS transistors.

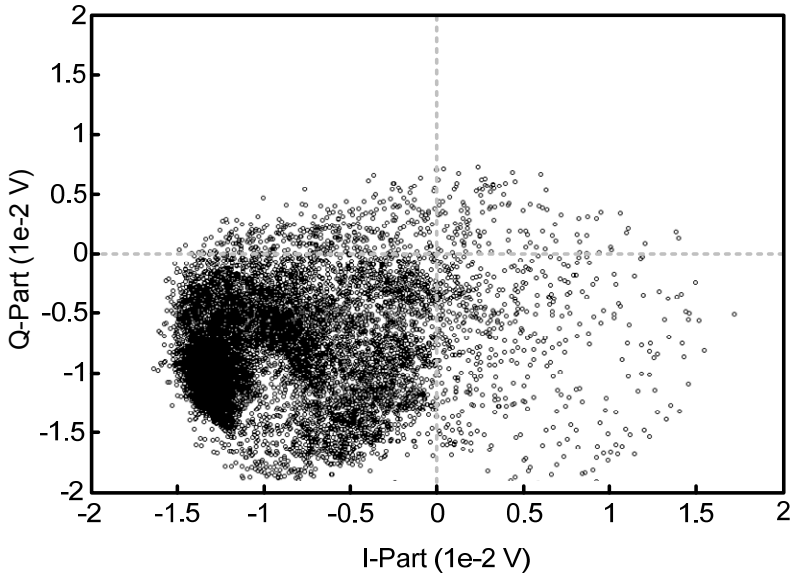


Figure 5.6 Simulation results of the switch-based DAM transmitter (10,000 points)

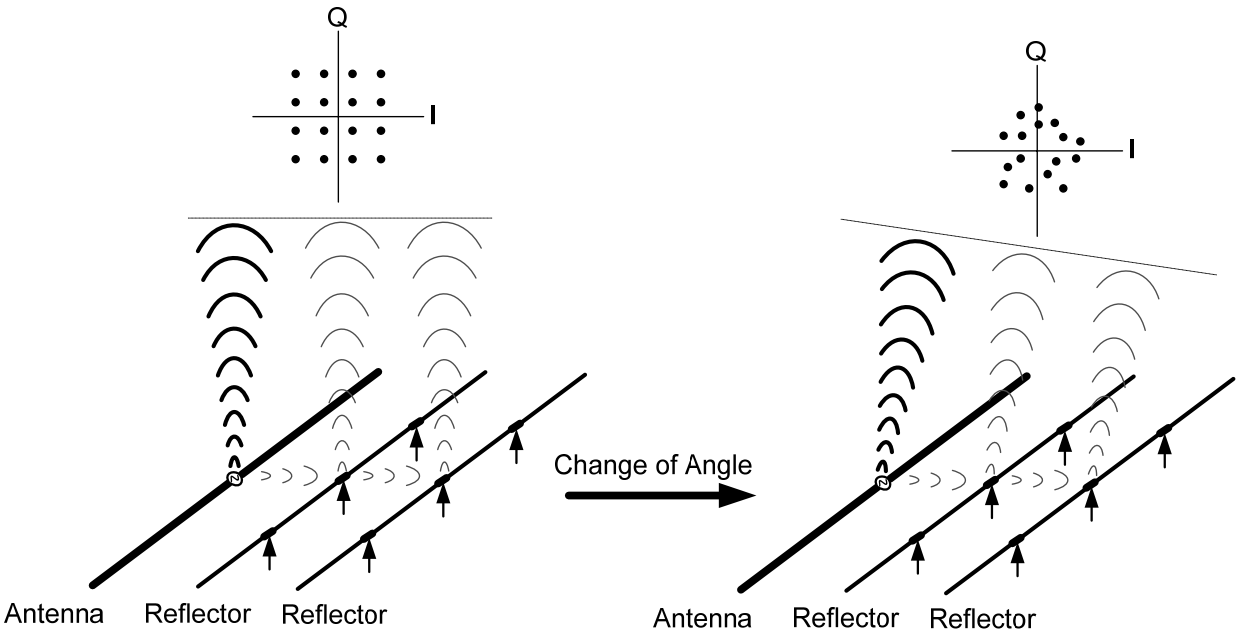


Figure 5.7 Communication security

Secure communication link— As mentioned earlier, one of the unique characteristics of this system is its ability to transmit independent signals in different directions, as depicted in Figure 5.7. To see how this is achieved, assume a set of switching combinations has

been found which generates a 16 QAM at the bore-sight (Figure 5.7). If we look at the modulation points at a direction different than the bore-sight, we will see that the constellation points are translated from their original locations to seemingly random locations, causing a scrambled set of points on the signal constellation diagram. This happens because the scattering properties of the reflectors and hence the phase and amplitude of the reflected signal vary with angle. The scrambling property prevents undesired receivers from properly demodulating the signal. At larger angles, some of these constellation points move to adjacent cells (Figure 5.8) and introduce error. Figure 5.8 shows the simulated error rate versus angle in the E-plane (parallel to dipole) and the H-plane (normal to dipole) of an on-chip dipole antenna. In this simulation, 210 equally spaced constellation points at the bore-sight are selected and viewed at different angles on the E- and H-planes of the antenna. In the H-plane, error rate rises rapidly and reaches ~ 50% at an offset angle of 2–3 degrees. In the E-plane, error rate reaches ~ 50% at an offset angle of 6–7 degrees off from the bore-sight. Receivers located at angles +/- 1° can completely recover the modulated signal without any error in the absence of noise and other channel non-idealities. In other words, receivers located within the *information beam-width* of the antenna can properly recover the signal. Thus, it is necessary to define the information beam-width in addition to the power beam-width for such systems. The radiation pattern beam-width represents the *power* directionality of the beam while the information beam-width refers to the *information* directionality of the beam.

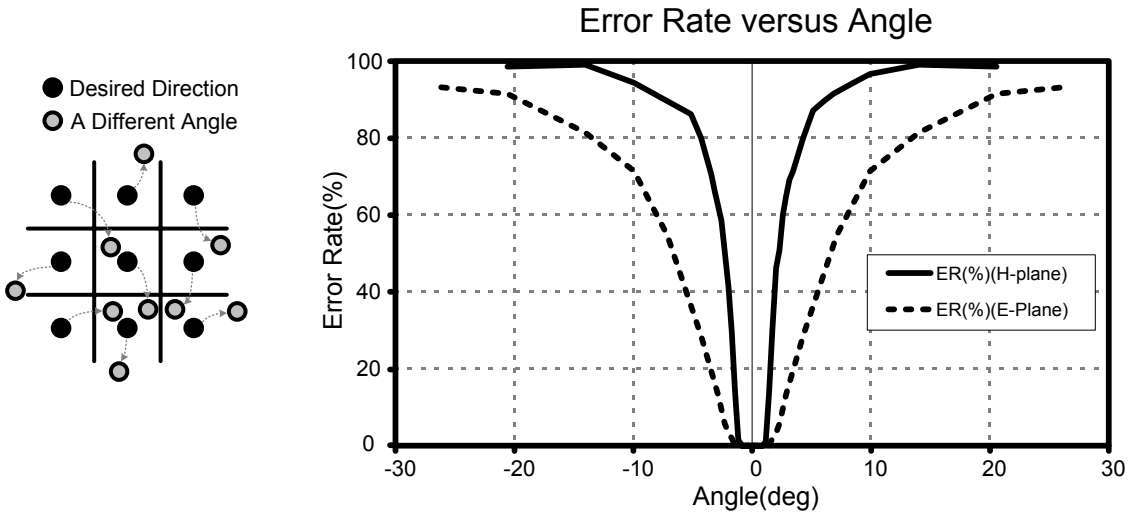


Figure 5.8 Information beam-width

It is noteworthy that each complete DAM system including all the reflectors and switches can serve as a single element in a phased-array made of such elements. The phased array allows us to create a narrow power beam-width in addition to the narrow information beam-width achieved via the DAM transmitter, as illustrated in Figure 5.9.

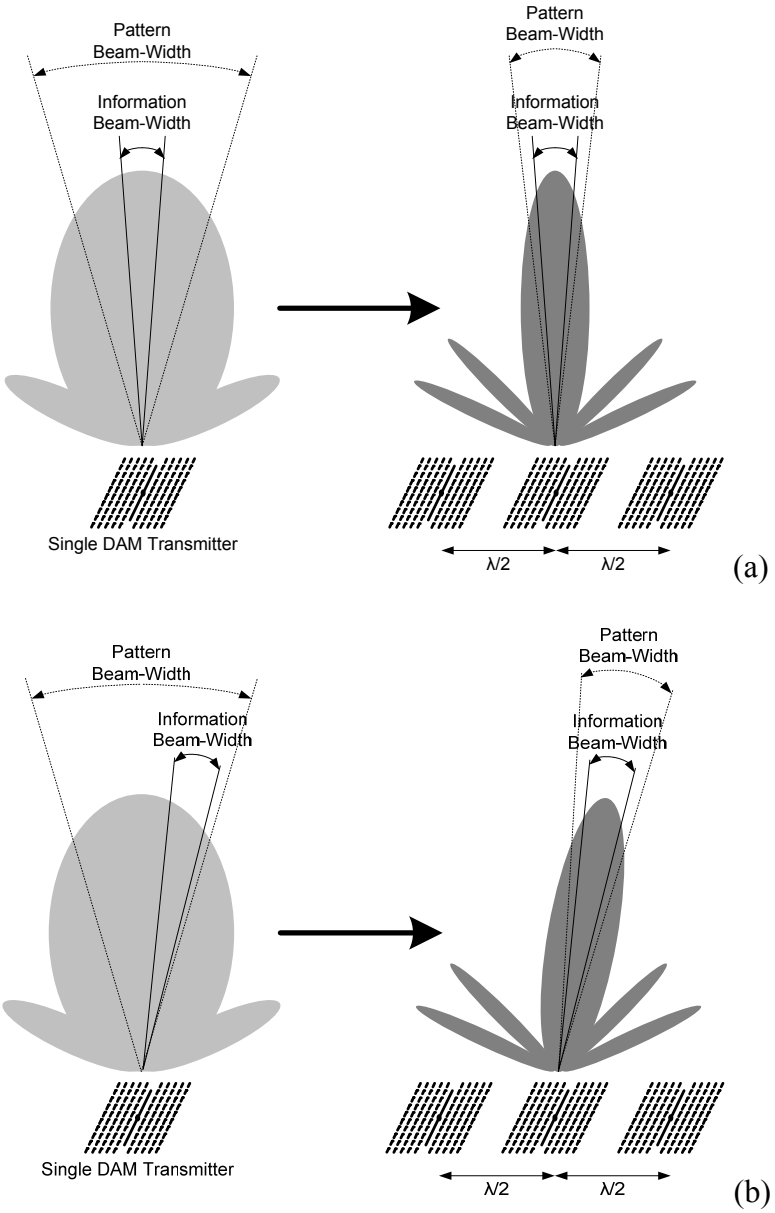


Figure 5.9 DAM transmitter in a phased-array configuration

Redundancy and added security— It was mentioned that by adding each switch the total number of the switching combinations is doubled (2^N combinations for N switches). In our switch-based DAM transmitter prototype design, we have ten reflectors each with nine switches, totaling $N = 90$ switches. This results in 2^{90}~10^{27} combinations. Obviously,

it is not necessary to use all of these combinations, but this large number of combinations creates so much redundancy that can be utilized in a productive way.

It is possible to generate a desired phase and amplitude in a given direction using so many different switch combinations. For a large enough number of combinations, it is possible to find different switch combinations that produce the same point in a given direction while generating widely scattered points in a different direction (Figure 5.10). This property of the system allows the transmission of a set of pre-defined modulation points in a desired direction while simultaneously randomly changing the pattern of the constellation points in the undesired directions, thus making it even more challenging for unintended receivers to find a one-to-one mapping between the received signal at the desired direction and the undesired directions.

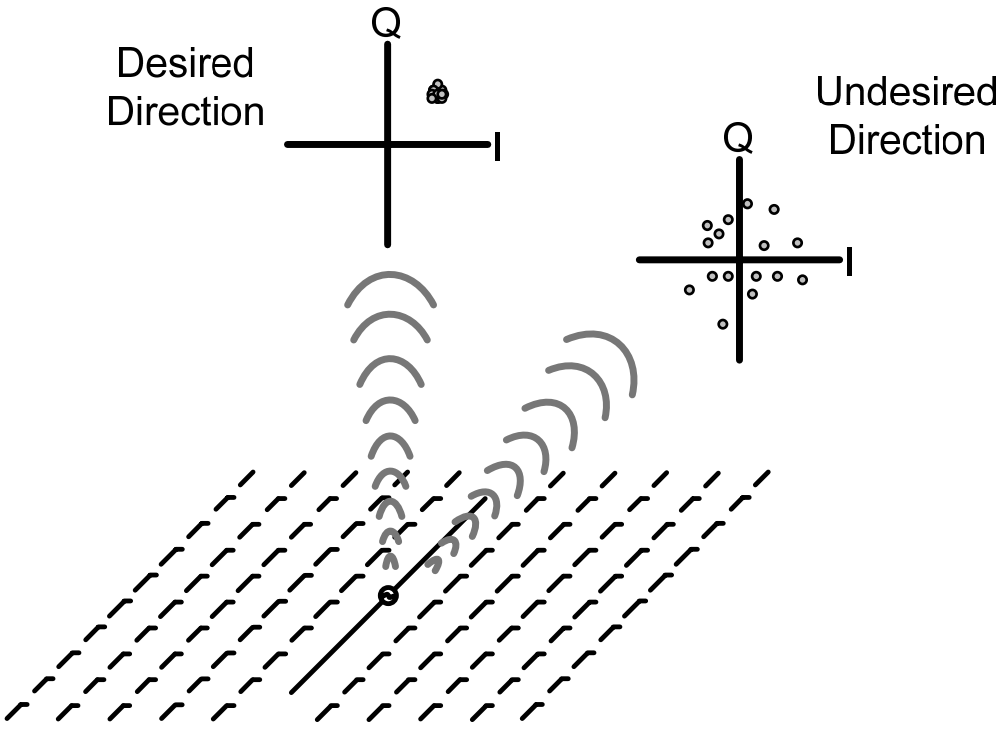


Figure 5.10 Enhancing security by leveraging redundancy

Redundancy and multiple beam transmission— The aforementioned redundancy may be used to transmit two or more independent streams of information to two or more different directions simultaneously using a single transmitter at its full rate. For a reasonably chosen set of symbols, a large enough number of switches, under the right set of EM conditions and switch properties, it is possible to find a switch combination that generates two arbitrarily selected symbols in two sufficiently different directions. Under these circumstances it will be possible to transmit two independent streams of symbols in two directions at full rate without having to resort to time-, frequency-, or code-division multiple-access approaches. In fact, it can be viewed as an electromagnetic-domain multiple-access scheme. Figure 5.11 conceptualizes this functionality, where the transmitter sends a 16 PSK modulated signal to the right side and a 16 QAM signal to the left side.

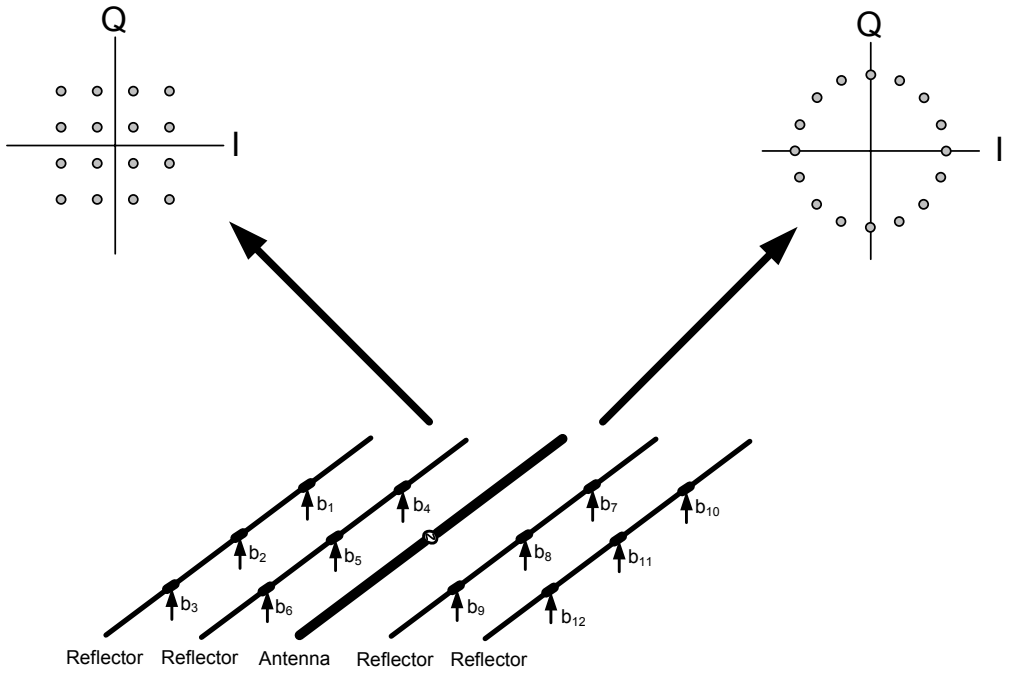


Figure 5.11 Multiple beam transmission by using a single transmitter

Spectral control— In most communication transmitters, not only should one adhere to the specifications on the transmitted signal in the frequency band of interest, but it is also necessary to control the out-of-band EM emission radiated by the antenna. This is often done through spectral control at base band (e.g., via pulse shaping) which needs to be maintained through the transmit chain (including the PA). This requirement generally translates to a tight specification on the linearity of the transmit path, in particular the PA. It is important for any alternative signal modulation scheme to offer an effective method to control this out-of-band emission of the signal.

The spectral control requirement can be translated to trajectory control in the I-Q plane of the signals. A very large number of combinations (e.g., 10^{27} in our prototype) allows coverage of most of the constellation points. This full coverage enables the control of out-of-band emissions as illustrated in Figure 5.12. As shown in this figure, instead of directly moving a point on the signal constellation diagram from A to B and generating high levels of out-of-band emissions, the intermediate steps of 1, 2, and 3 can be taken, thereby reducing unwanted out-of-band emissions. In general, as long as we have enough redundancy to have a good coverage of the signal constellation, it will be possible to render any trajectory in the I-Q plane using multiple intermediate points to achieve spectral control.

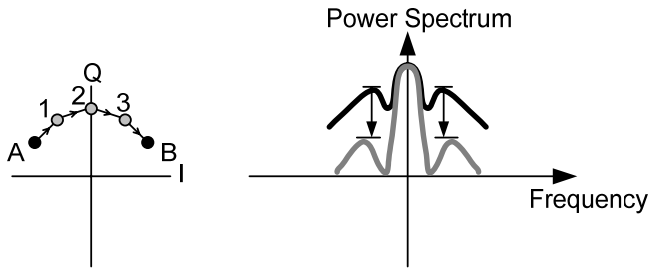


Figure 5.12 Spectral control

5.4 A Switch-Based DAM 60 GHz Transmitter Architecture

As a proof of concept, a 60 GHz transmitter has been implemented in the IBM 130 nm BiCMOS SiGe (8HP) process. Figure 5.13 shows the block diagram of this transmitter [29]. As shown in this figure, a PA driven by a locked oscillator sends a differential un-modulated sinusoidal signal to a shielded differential transmission line which carries the signal to the on-chip dipole antenna. The dipole has a length of 835 μm and a width of 20 μm. The differential transmission line uses top metal layer M_7 for the signal line, M_6 for the bottom ground line, and M_7 for the side ground shield. The on-chip dipole antennas and reflectors are all implemented on lower metal layers M_1, M_2, and M_3. Five reflectors are placed at each side of the antenna and 9 switches are placed on each reflector resulting in a total number of 90 switches. The antenna and reflectors occupy an area of 1.3×1.5 mm^2. As shown in Figure 5.13, the base-band data goes through a digital control unit which controls the state of the switches.

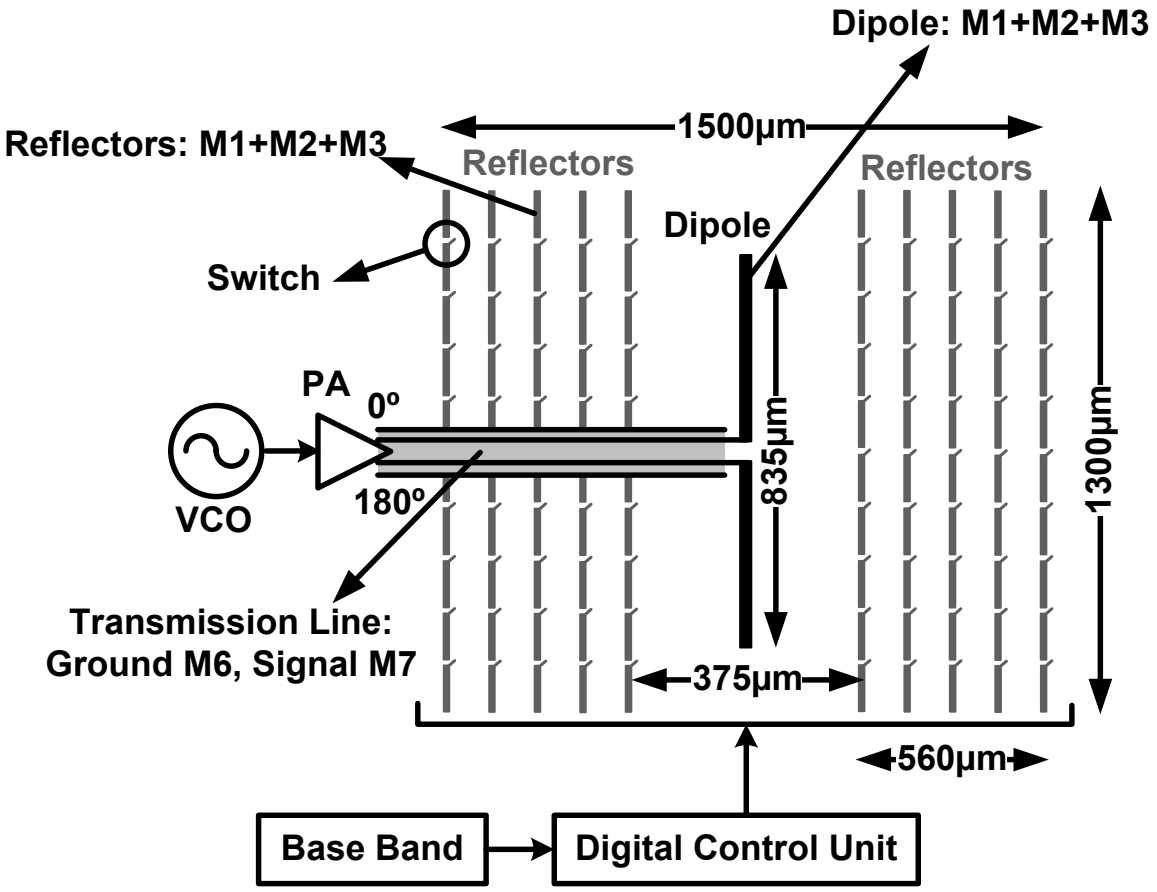


Figure 5.13 Details of the switch-based DAM transmitter architecture

In the switch-based DAM architecture, one of the important parameters that needs to be optimized is the distance of the reflectors from the antenna. If the reflectors are placed too far from the main antenna, the reflected signal will be very weak and it will not change the phase and amplitude of the main signal significantly. On the other hand, if we place reflectors too close to the antenna the size of the aperture will be very small and the system will not be able to excite a sufficient number of radiation modes to transmit independent information to several directions simultaneously. In this design, the location of the reflectors and their distance to the main antenna have been optimized. In addition, an optional coarse control unit was implemented as shown in Figure 5.14. This optional

unit can be used as a quadrant selector with control signals *A* and *B* taking values 1 and -1. In this mode of operation, switches on the reflectors can be used to generate the constellation points inside each quadrant and the optional coarse control unit can be used to choose the quadrant on the signal constellation diagram.

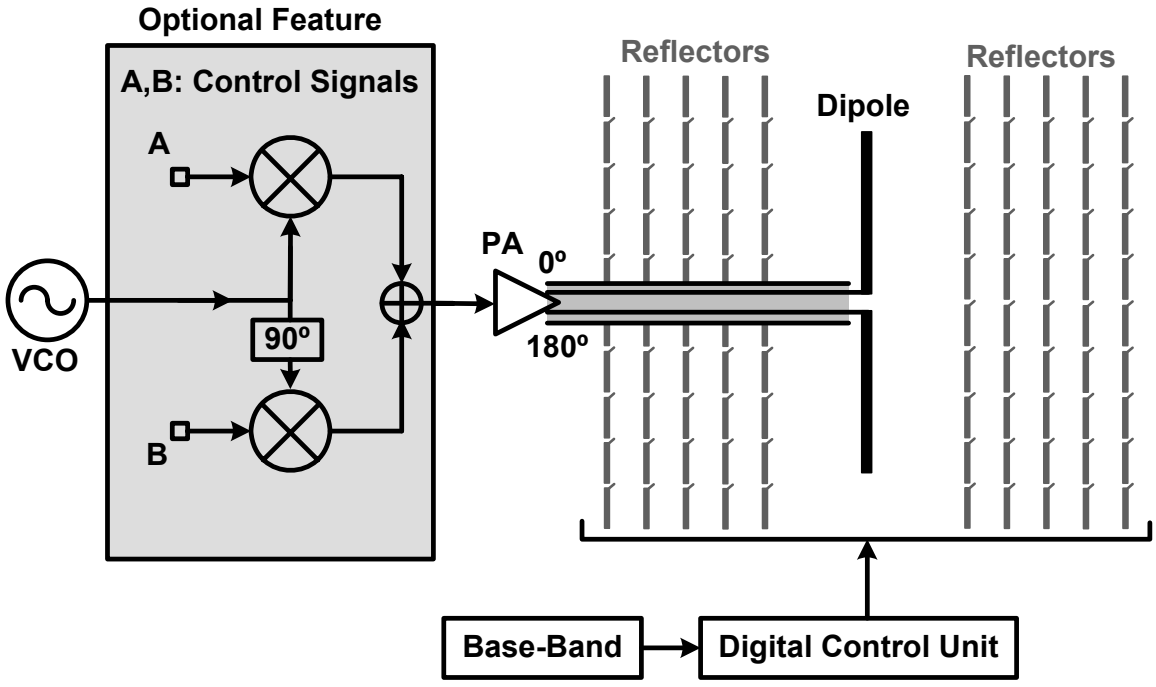


Figure 5.14 Optional coarse control unit

Figure 5.15 shows the design of the switch itself. In order to achieve a low on impedance (1–5 Ω) between the drain and the source of the switch in its closed state ($v_{g\text{-}ds}$ = 1.2V), the switch size cannot be small. This results in large gate-drain, C_{gd}, and gate-source, C_{gs}, capacitances. The NMOS switch used in this design has an effective width of 150 μm, as shown in Figure 5.15. To resonate out the switch capacitance and achieve high impedance in the open state ($v_{g\text{-}ds}$ = 0 V), a circular shielded transmission line behaving as an inductor is placed between the drain and the source of the switch. This

transmission line resonates with the open-switch capacitance at 60 GHz resulting in a maximum impedance of 70 Ω. The circular transmission line has a diameter of 60 μm and is implemented on metal layers M_2, M_3, and M_4.

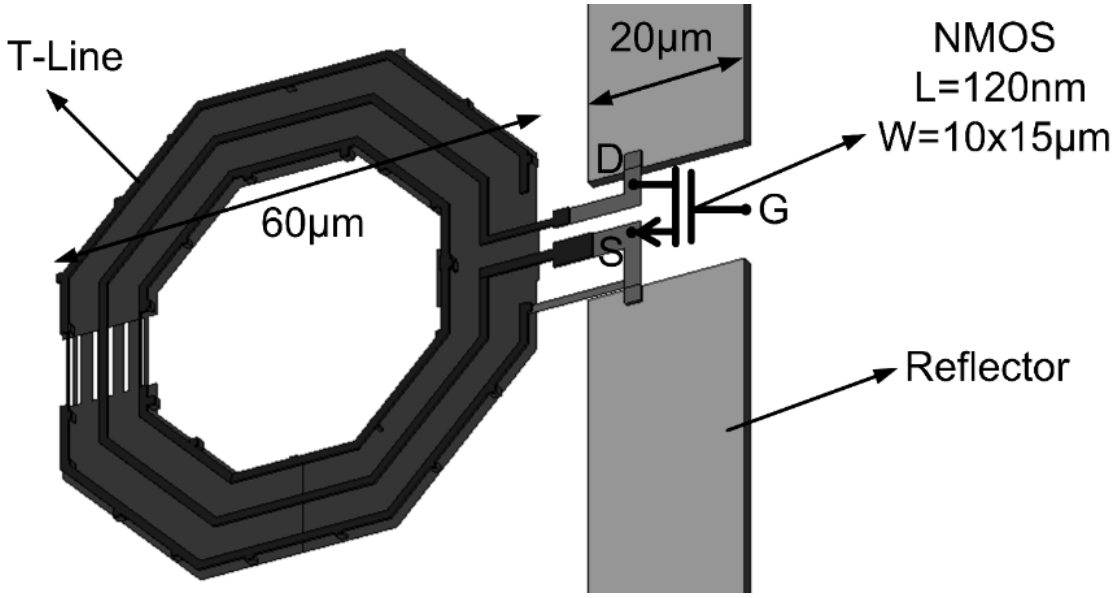


Figure 5.15 A 60 GHz resonant NMOS switch

One of the main issues which has to be addressed is the transient response of the switches and the reflectors. This determines the speed of far-field changes and the maximum symbol modulation rate. The limitation comes from two independent factors. The first one is related to the finite time which the wave needs to travel between the antenna and the reflectors, and the second is related to the transient response of the switch. Figure 5.16 shows the simulation results of the far field's transient response to a change in the switching combination when ideal switches are used. As shown in this figure, the far field takes less than 200 ps to adapt to a new switching combination. This simulation shows that the effective transient response of the whole system is limited by the transient time of the switch itself, as the switch response is usually longer than 200 ps.

By designing a switch with a transient response of about 800 ps we can achieve a symbol rate of better than 1 GS/s.

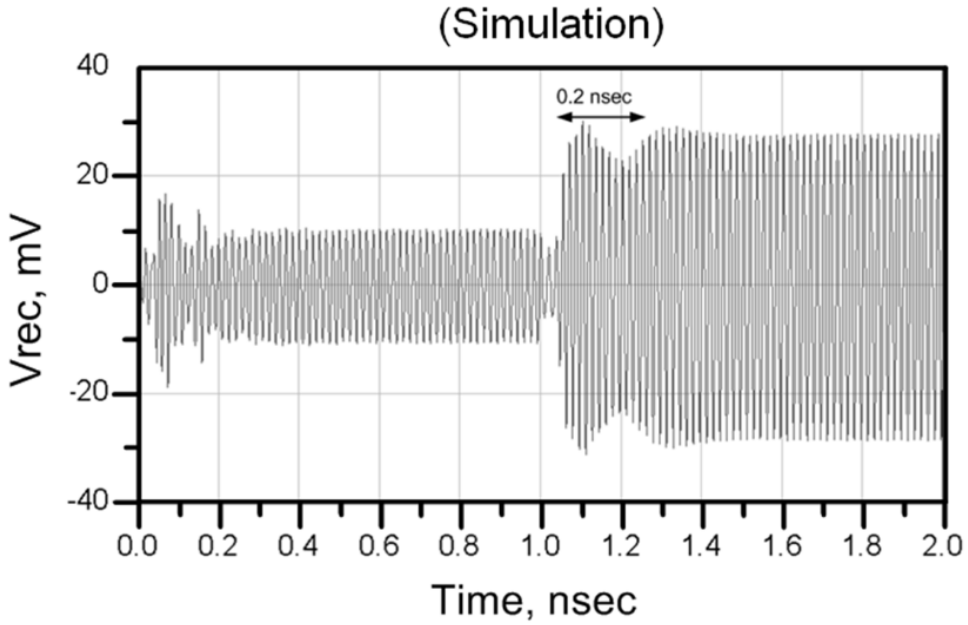


Figure 5.16 Transient response of the far-field

One of the important disadvantages of silicon-based on-chip antennas is the low antenna efficiency. This is a result of two factors; silicon's high dielectric constant (11.7), and substrate's low resistivity (1–10 Ω·cm). The high level of doping required to fabricate active circuits limits the silicon substrate's resistivity. Silicon's high dielectric constant and its large substrate thickness (200–300 μm) couple most of the dipole output power into substrate-modes in unshielded structures as shown in Figure 5.17 [3]. If we use an on-chip ground shield to isolate the on-chip antenna from the lossy substrate, the radiation efficiency will be very small (around 1%). In standard silicon processes the distance between on-chip metal layers rarely exceeds 15 μm. A ground layer at this distance, which is much smaller than the wave-length in mm-wave frequencies (2.5 mm

wavelength in SiO_2 at 60 GHz), shorts the antenna by introducing a negative image current very close to the antenna and hence reduces both the radiation resistance and the efficiency. On the other hand, if we do not use an on-chip ground shield, the silicon substrate behaves as a dielectric waveguide, generates the substrate modes, and leads the power to the chip edges resulting in an undesirable pattern. Due to the silicon substrate's low resistivity, most of the power that couples into substrate-modes disappears as heat reducing the overall antenna efficiency [3].

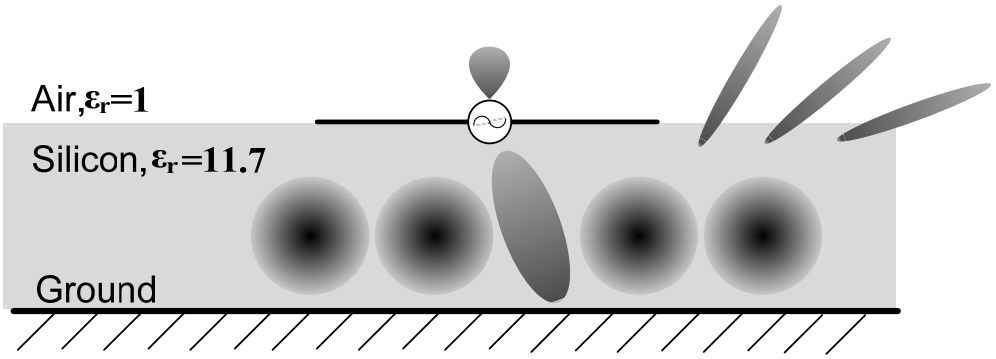


Figure 5.17 Substrate modes [3]

The amount of total power coupled into the substrate modes depends on the substrate's geometry. If an un-doped silicon hemispherical lens (or a dielectric lens with a dielectric constant similar to silicon's) is attached at the backside of the substrate as shown in Figure 5.18, antenna efficiencies up to 10% can be acheived. A Silicon lens takes substrate modes and converts them into useful radiated power [2] - [8]. Due to the impedance mismatch between the silicon and the air ($Z_{si} = \sqrt{\frac{\mu_0}{\varepsilon_{si}}} = 110\Omega$ versus $Z_{air} = \sqrt{\frac{\mu_0}{\varepsilon_0}} = 377\Omega$) about 30% of the radiated power will be reflected from the silicon-

air boundary, but this problem can be resolved by using a quarter-wavelength matching layer on the boundary of the silicon lens and the air [2] - [8]. In this design a hemispherical silicon lens with a diameter of about 1 inch is used to couple the power to the air.

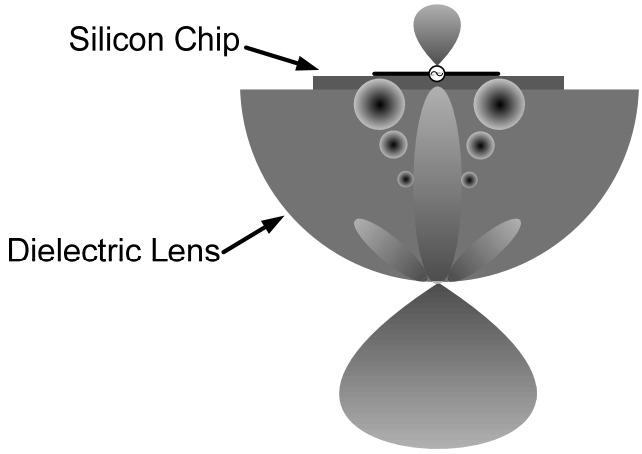


Figure 5.18 Silicon lens

Figure 5.19 shows the micrograph of the chip prototype, which is implemented in IBM 8HP 130 nm SiGe BiCMOS process. In addition to the dipole and the reflectors which occupy an area of 1.3 mm × 1.5 mm, an optional coarse control unit is designed which includes the base-band amplifiers, the up-converter mixers, and the PA. We have also designed a 60 GHz receiver with on-chip dipole antenna. The receiver is composed of a 60 GHz LNA, down-converter mixers, and base-band amplifiers. The LO signal is generated on-chip by using a V-band VCO. An injection-locked divider and divider chain are used to divide down the VCO signal by a factor of 1024.

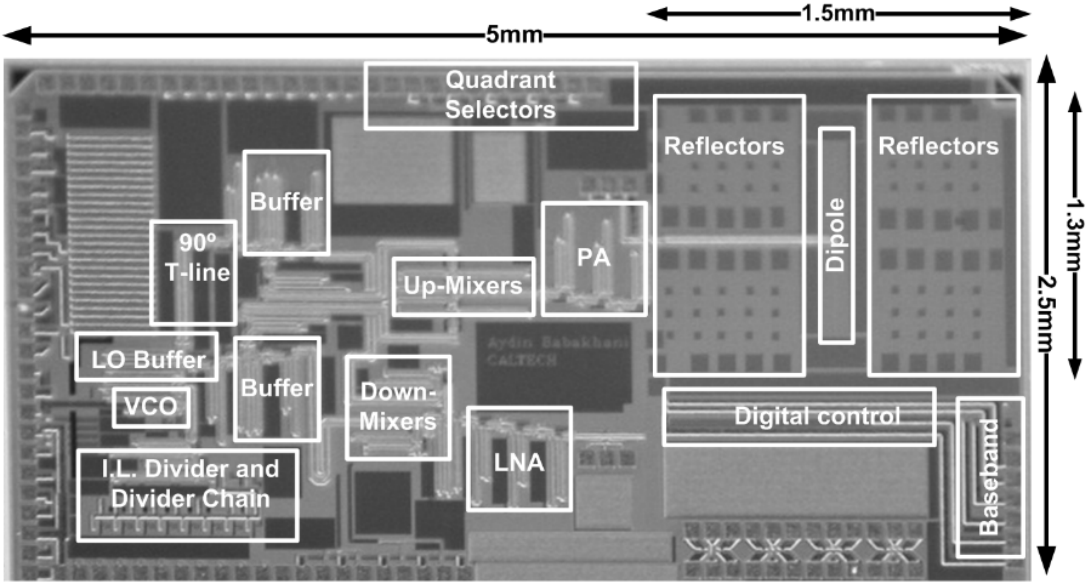


Figure 5.19 Switch-based DAM chip micrograph

5.5 Varactor-Based DAM

An alternative to the switch-based DAM approach presented earlier is to use a varactor-based approach to change the antenna characteristics and modulate the signal at the far-field. Varactors can be used to change the effective capacitive load on the reflectors. It is noteworthy that the controlling is still done in a digital fashion by switching the capacitance of the varactor between C_0 and C_1, which are the capacitance values for a digital control value of 0 or 1, respectively.

Figure 5.20 shows the block diagram of the varactor-based DAM system. The differential signal goes through a transmission line and drives the dipole antenna. The varactor-based DAM system uses 90 differential varactors on 10 reflectors as shown in Figure 5.20. In this design, varactors are implemented by using a series combination of two NMOS transistors as shown in Figure 5.20 and Figure 5.21. A binary control voltage

("1" = 1.2 V, " 0 " = 0 V) is applied to the gate of the NMOS transistors. Depending on the control voltage, gate to drain-source voltage of the varactors, $V_{g\text{-}ds}$, can be either +0.6 V or -0.6 V. Each NMOS transistor has a size of 3 × 16 × 5 µm × 240 nm. The whole system occupies an area of 2.1 mm × 2.4 mm. A micrograph of the varactor-based DAM chip is shown in Figure 5.22, where it is implemented in the 130 nm IBM 8HP SiGe BiCMOS process.

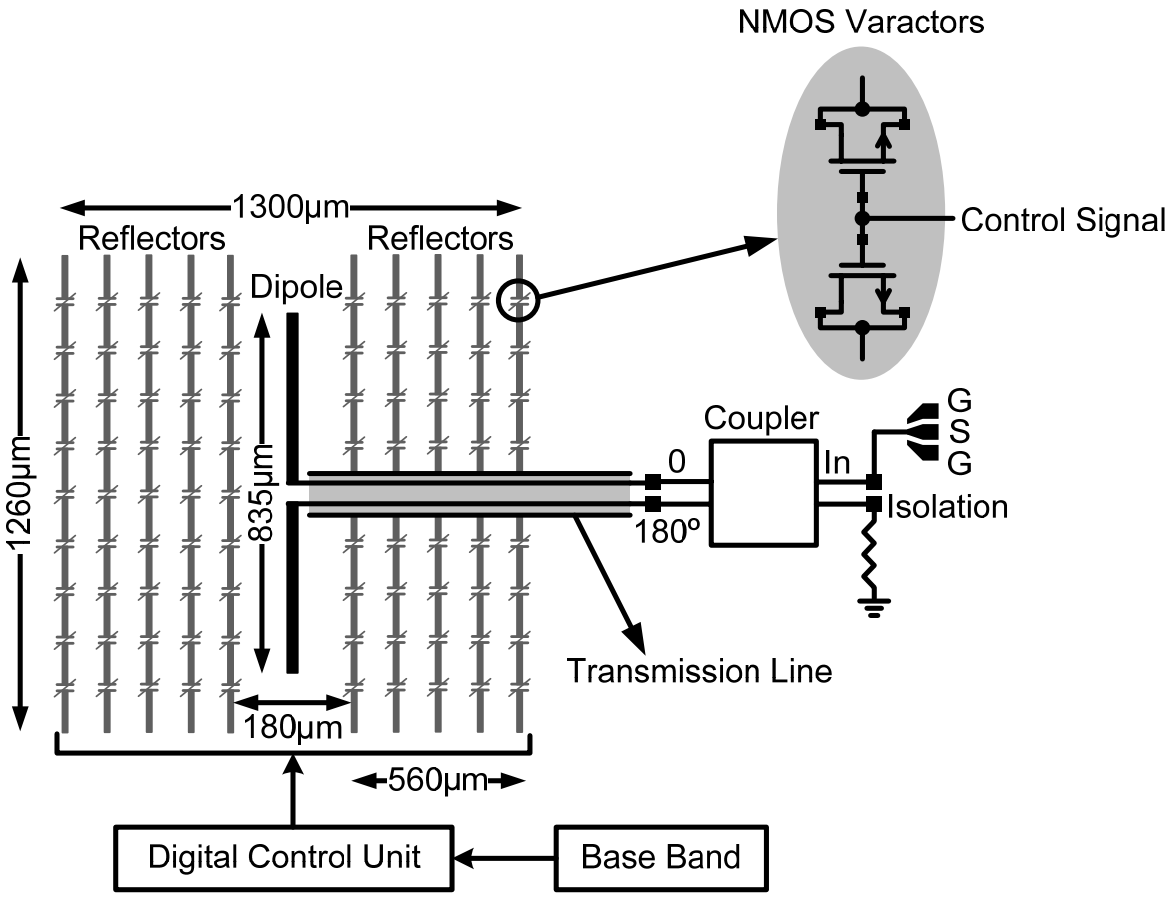


Figure 5.20 Varactor-based system's block diagram (varactor-based DAM)

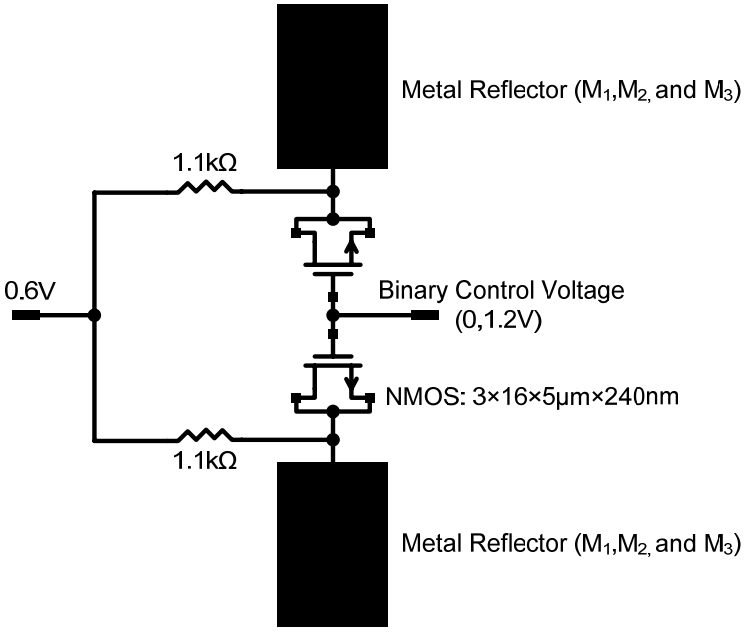


Figure 5.21 Design of the varactors and reflectors

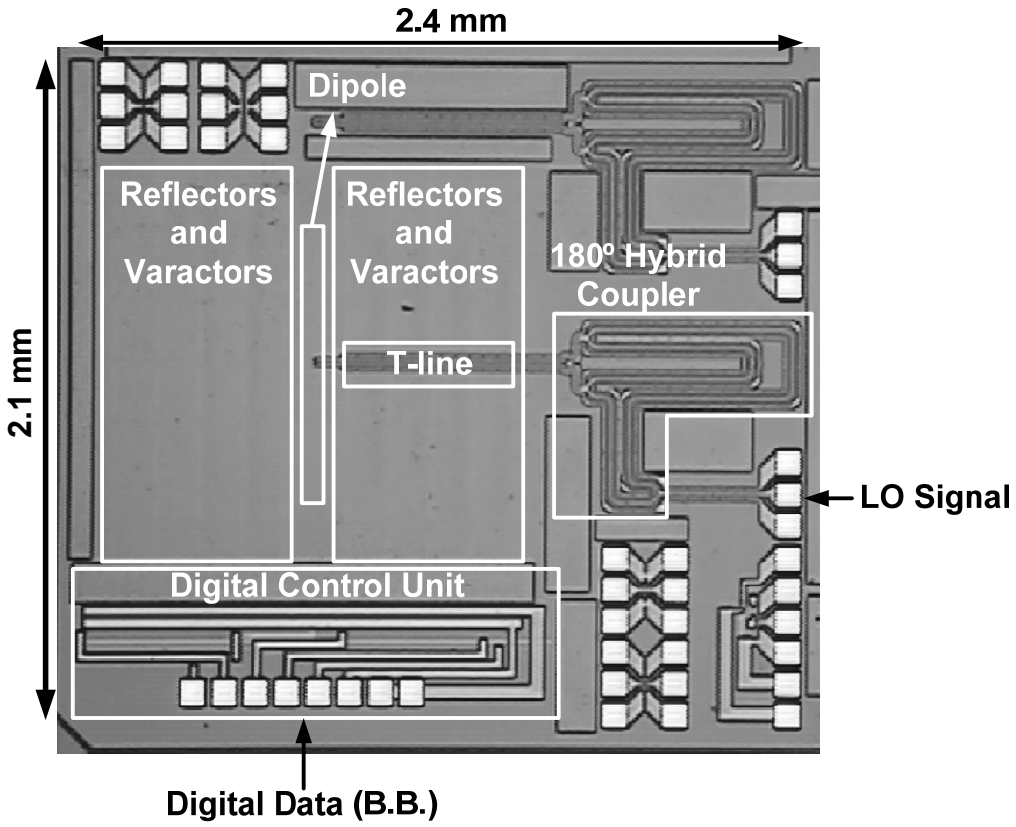


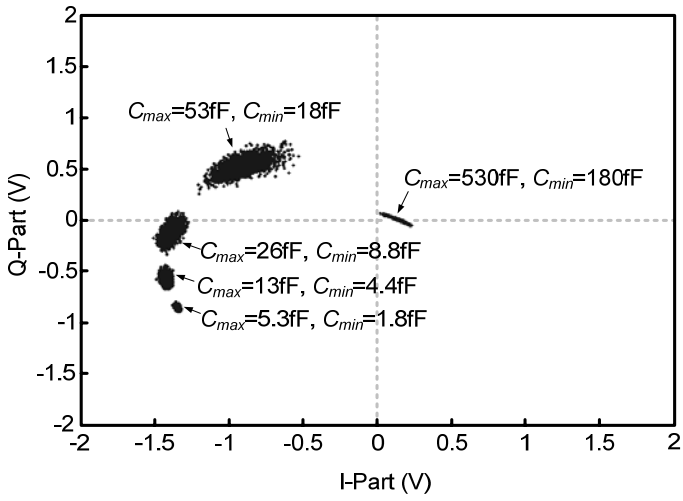
Figure 5.22 Varactor-based DAM's chip micrograph

One of the important parameters that we would like to maximize is the amount of coverage of the signal constellation diagram. In other words, the reflected signals should be strong enough to change the phase and amplitude of the main signal to cover the four quadrants in the signal constellation diagram without affecting the impedance which the PA sees at the antenna port to maintain a high PA efficiency. To maximize the coverage of the signal constellation diagram, we choose capacitances C_0 and C_1 close to the C_{min} and C_{max} of the varactor.

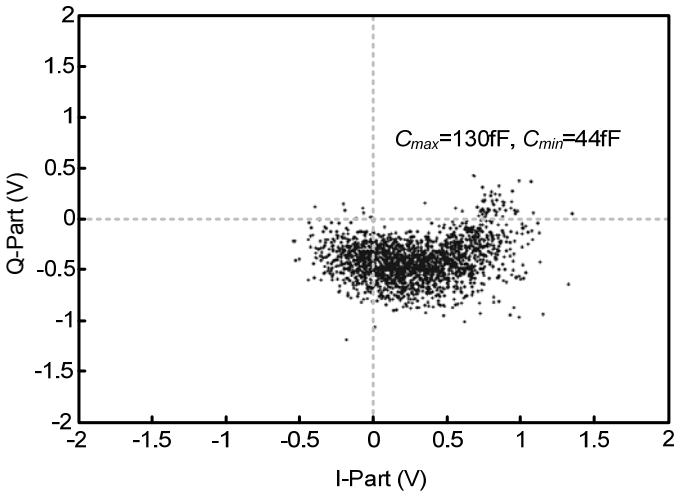
Usually the ratio of C_{max}/C_{min} of a varactor is a fixed value and is independent of the varactor size. In the process technology used to implement our prototype, the ratio of C_{max}/C_{min} is around 3–4. If we choose very small values for C_{max} and C_{min} then the effective values of the varactors' impedance will be very high in the frequency of interest. For example at f_c = 60 GHz, C_{max} = 4 fF and C_{min} = 1 fF will be equivalent to $|Zc_{max}|=(2\pi f_c C_{max})^{-1}$ = 663 Ω and $|Zc_{min}|=(2\pi f_c C_{min})^{-1}$ = 2.65 kΩ. Due to the high impedance of these varactors, the induced current on the reflectors will be negligible compared to the current of the main antenna itself and the reflected signal will be too weak to change the phase and amplitude of the main signal. In other words, the varactors will behave as open circuits in both cases of C_{min} and C_{max}, and the phase and amplitude of the far field will not change by varying the capacitance of the varactor. On the other hand, if we choose very large values for C_{max} and C_{min}, then the effective values of the varactors' impedance will be very low in the frequency of interest. For example C_{max} = 4 pF and C_{min} = 1 pF will result in $|Zc_{max}|=(2\pi f_c C_{max})^{-1}$ = 0.663 Ω and $|Zc_{min}|=(2\pi f_c C_{min})^{-1}$ = 2.65 Ω at f_c = 60 GHz. At these values, varactors behave as short circuits for both values

of C_{max} and C_{min} and hence the amplitude and phase of far-field will not change by varying the capacitance of the varactors.

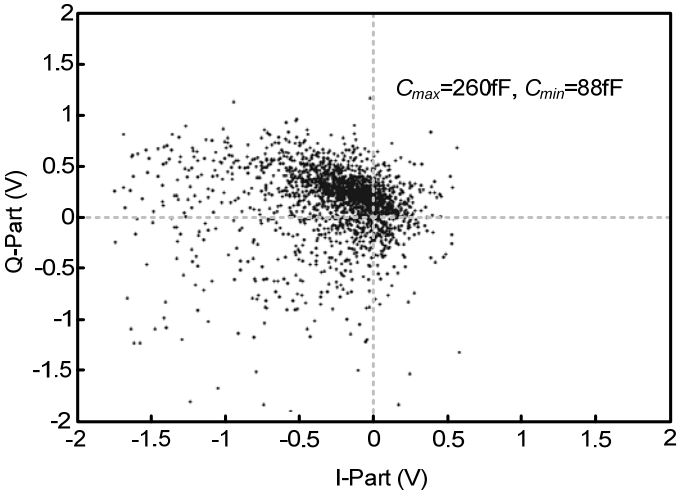
Figure 5.23 shows how signal constellation coverage changes with the values of C_{max} and C_{min}. In this simulation we simulated 2000 random points, and for each point, for each switching combination, we plotted the real and imaginary parts of the voltage of a dipole antenna located at the far-field. At each point a Matlab program assigns a random binary value for the capacitance C_{var} of each varactor, $C_{var} \in \{C_{max}, C_{min}\}$, and plots the real and imaginary parts of the dipole's voltage located at the bore-sight. As shown in Figure 5.23 (a), for very small varactor sizes and very large ones the spread of the points on the signal constellation diagram is very limited. As we mentioned before, this is because of the fact that very small varactors behave as open circuits and very large ones behave as short circuits at 60 GHz. Figure 5.23 (b) and (c) show the spread of the constellation points for varactor sizes (C_{max}= 130 fF, C_{min} = 44 fF) and (C_{max} = 260 fF, C_{min} = 88 fF). As shown in Figure5.23 (c), with varactor size of (C_{max} = 260 fF, C_{min} = 88 fF) it is easily possible to cover all of the four quadrants with binary controlling of the varactors ("1" = C_{max}, "0" = C_{min}). It is important to mention that in the above simulations ideal varactors (infinite Q) are used. Simulations results show that varactors with smaller quality factor reduce the coverage of the signal constellation diagram. Figure 5.24 shows the simulation results for varactors' size of (C_{max} = 260 fF, C_{min} = 88 fF) with quality factors of 1, 4, and 10. Based on these results, the coverage of the signal constellation diagram is a strong function of the varactors' quality factor, Q. In this simulation a silicon lens is used to minimize the substrate modes and efficiently couple the signal to the air.



(a)



(b)



(c)

Figure 5.23 Signal constellation coverage for different varactor sizes (ideal varactors)

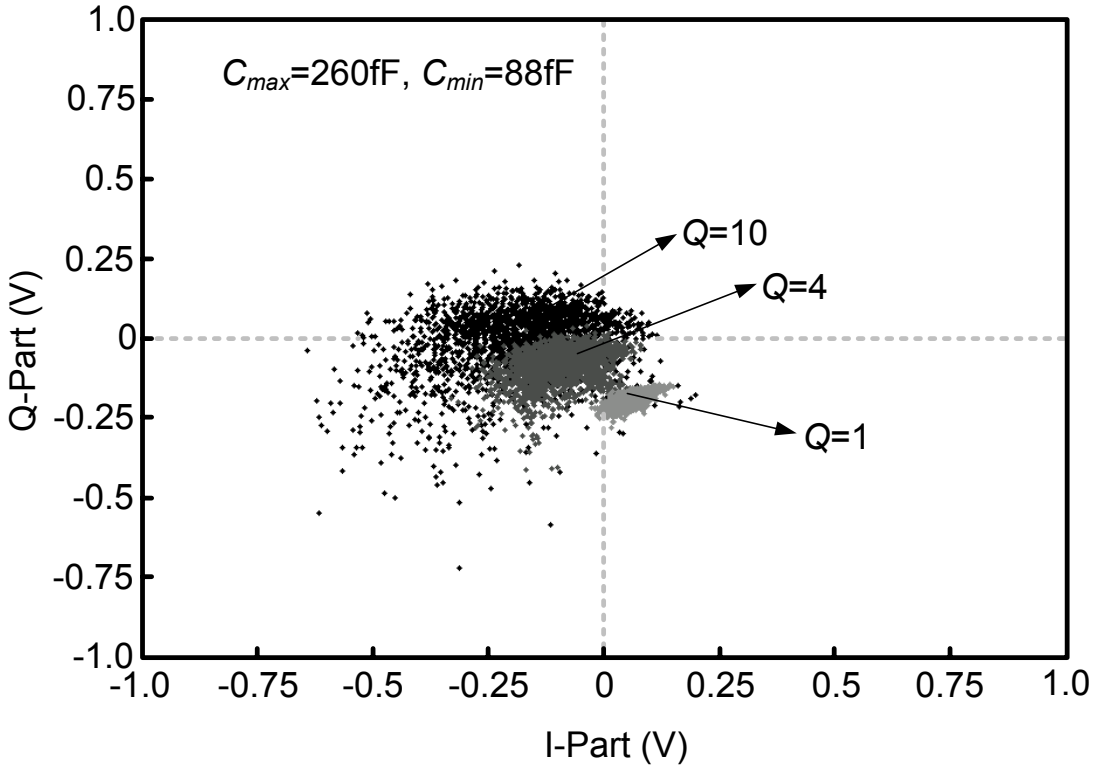


Figure 5.24 Signal constellation coverage for different varactor's quality factors

5.6 Schematic of the Blocks Used in the Switch-Based DAM Transmitter

V-band amplifier— Figure 5.25 shows the schematic of the V-band amplifier which is composed of three differential stages. Stub-tuning and coupled-wire differential transmission lines are used for matching purposes. To isolate the differential transmission line from the neighboring blocks, adjacent ground shields as well as bottom ground shields are used [3] , [17]. To bias the differential pair, a diode-connected transistor with a series resistor is designed as shown in Figure 5.25. The base of this transistor is connected to a differential transmission line with a length of 225 μm. This transmission line carries the biasing DC voltage to the base nodes of the transistors in the differential

pair. At RF frequencies, the line behaves as a stub line and, in conjunction with a 40 μm series differential transmission line, provides a power match between the output of the up-converter mixer and the input of the V-band amplifier. In this design, 150 fF MIM capacitors are placed at the output of each differential pair isolating the DC bias of each stage from the following one. The 3-stage amplifier provides up to +7 dBm output power.

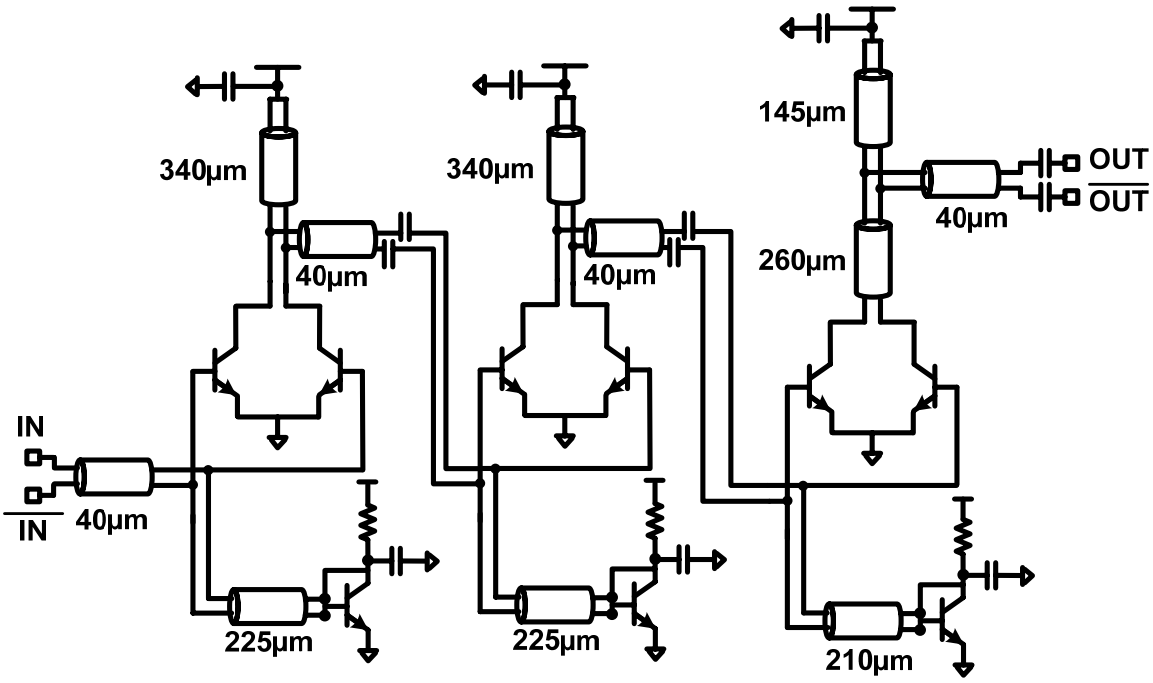


Figure 5.25 Schematic of the V-band power amplifier

Voltage controlled oscillator (VCO)– The schematic of the VCO is shown in Figure 5.26. A cross-coupled oscillator generates a V-band on-chip LO signal. The oscillation frequency is varied by changing the capacitance of the NMOS varactors which have a size of 15 μm, as shown in Figure 5.26. A differential coupled-wire transmission line is used to behave as an inductor resonating with the parasitic capacitances of the cross-coupled transistors and the capacitance of the NMOS varactors. The important trade-off

in designing the cross-coupled oscillator is the trade-off between the phase-noise and the tuning range. The minimum and maximum oscillation frequencies can be calculated by

$$f_{\min} = \frac{1}{2\pi\sqrt{L(C_{\text{var\_max}} + C_{parasitic})}} \tag{5.3}$$

$$f_{\max} = \frac{1}{2\pi\sqrt{L(C_{\text{var\_min}} + C_{parasitic})}} \ . \tag{5.4}$$

To achieve a low phase-noise LO, we need to use a large inductor which limits the value of the total capacitance, C_{total}=C_{var}+$C_{parasitic}$. Due to the fact that the parasitic capacitance of the transistor, $C_{parasitic}$, ultimately depends on the process and not much can be done to minimize it, the only way to reduce the size of the C_{total} is to reduce the size of the C_{var}. By making C_{var} smaller, the tuning range is reduced. To minimize the effective $C_{parasitic}$ at the collector node of the cross-coupled oscillator, instead of directly connecting the base of one cross-coupled transistor to the collector of another one, a capacitive divider is used, as shown in Figure 5.26. This capacitive divider minimizes the total parasitic capacitance seen at the collector node by minimizing the contribution of the parasitic capacitors connected to the base of the cross-coupled transistors. The capacitive divider also allows a higher voltage swing at the collector node of the transistor by isolating that node from the base node of the transistor. The high voltage swing improves the phase noise of the oscillator. To further reduce the parasitic capacitance at the collector node, the output signal is taken from the base node by connecting the load to the base node instead of the collector node. 91 fF AC-coupling MIM capacitors are used to connect the output of the oscillator to a following amplifier.

To bias the VCO, a current source and a diode-connected transistor are used as shown in Figure 5.26.

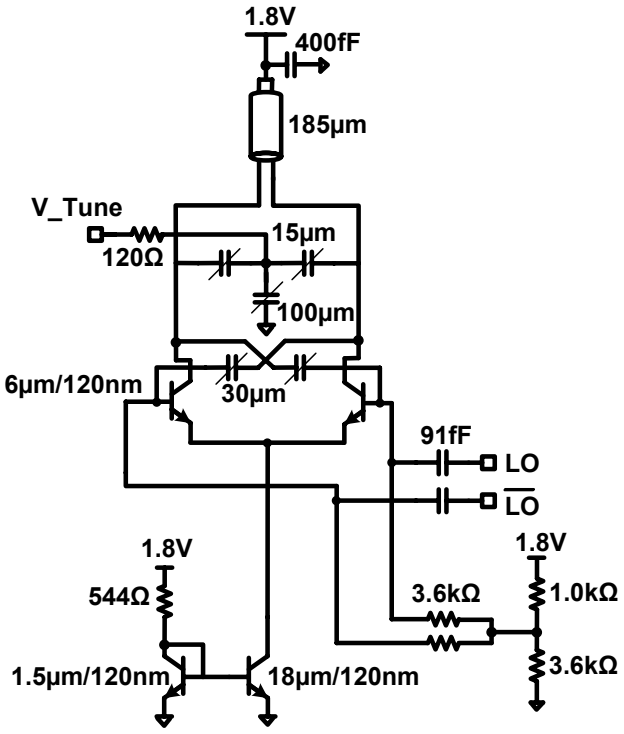


Figure 5.26 Schematic of the V-band VCO

Injection-locked divider— To divide down the LO signal and lock it to a reference signal, we have used an injection-locked divider in combination with a digital divider chain. The divider chain with the injection locked divider provides a dividing ratio of 1024, allowing an off-chip low-frequency signal to be locked with the on-chip LO signal. The design of the injection-locked oscillator is similar to that of the VCO but its load is tuned to provide a self-oscillation frequency of around $f_{LO}/2$. The LO signal is injected at the current source of the cross-coupled pair, as shown in Figure 5.27. A single-ended transmission line with length of 220 μm and an ac-coupling capacitor with size of 148 fF

are used to maintain a power match between the divider and the preceding stage. The preceding stage is an LO distributor amplifier which is discussed in the next paragraph.

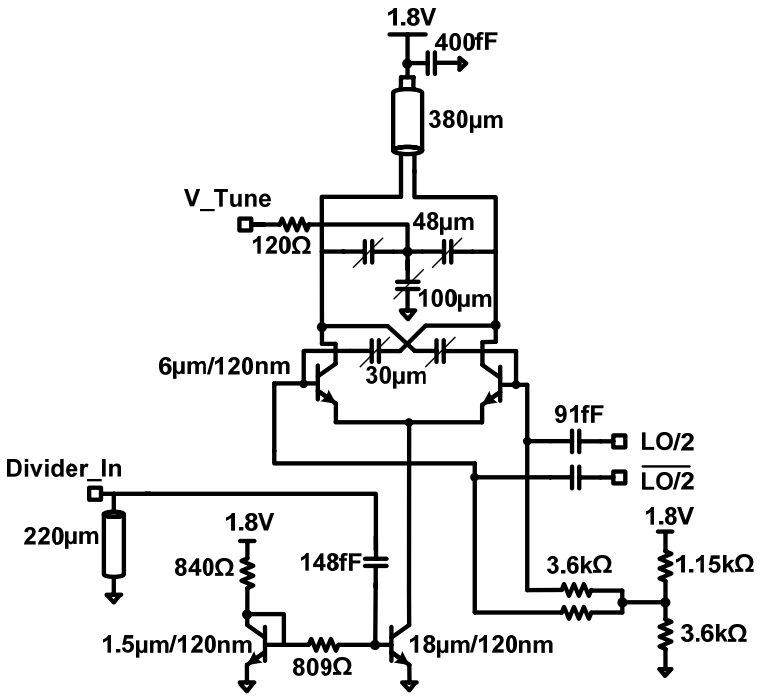


Figure 5.27 Schematic of the V-band injection locked divider

LO distributor amplifier— Several LO distributor amplifiers are used to provide LO signal to the injection-locked divider and up-convertor mixers in the transmitter. The schematic of the LO distributor amplifier is shown in Figure 5.28. The LO signal enters through a stub-matching network and is amplified by a cascode differential pair. To distribute the LO signal, the collector current of the input differential pair is divided equally between two identical cascode pairs. These pairs amplify the LO signal and use stub-tuning to efficiently deliver the LO power to the following stages.

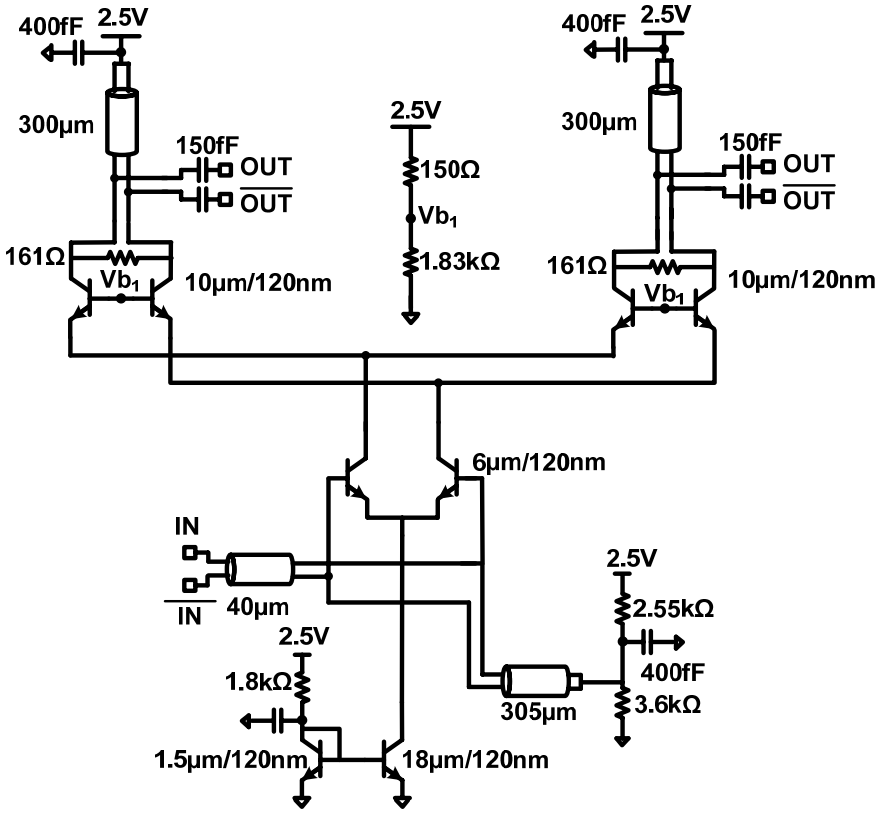


Figure 5.28 Schematic of the LO distributor amplifier

Up-convertor mixer— As an optional feature in our system, up-convertor mixers are designed and used in the optional coarse control unit. The coarse control unit can be used as a quadrant-selector on the signal constellation diagram. Figure 5.29 shows the schematic of the up-convertor mixer. Differential base-band or control signals are connected to the base nodes of the differential pair as shown in Figure 5.29. Stub-tuning is used at the RF and LO nodes to maintain a power match between the mixer and the adjacent blocks. To amplify the signal at the output of the mixer a differential cascode buffer amplifier is designed as shown in Figure 5.29. A 320 µm differential transmission line and 148 fF MIM ac-coupling capacitors are used to match the output of the mixer to the input of the buffer.

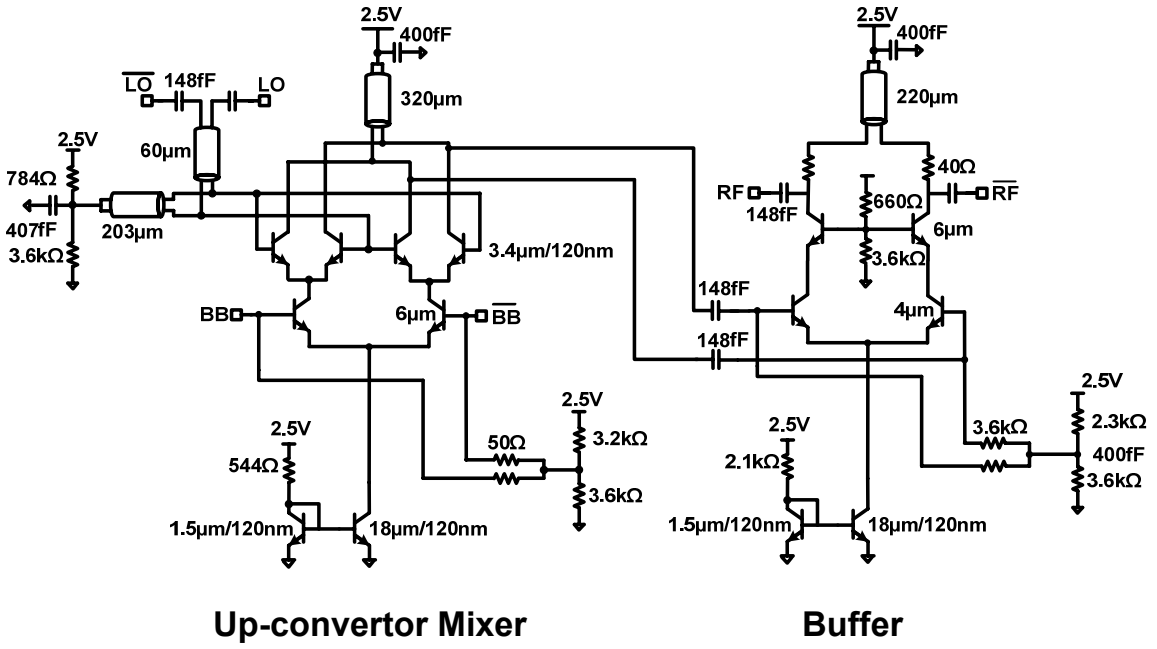


Figure 5.29 Up-converter mixer and buffer (optional feature)

5.7 Measurement Results

A block diagram of the measurement setup is shown in Figure 5.30. In the first measurement, the on-chip transmitter is disconnected (with laser trimming) from the antenna-reflectors combination and the electromagnetic structure itself is tested using an Agilent N5250A network analyzer. A LabView program [32] controls the state of the switches (switch-based DAM) and varactors (varactor-based DAM) through a data acquisition card. The data acquisition card sends the digital stream to the on-chip digital control unit and this unit programs the switches and the varactors. After sending the desired data streams and programming the chip, the LabView program communicates with the network analyzer through a GPIB card. One of the ports of the network analyzer sends the V-band signal through a 1.85 mm cable and connectors to the on-chip antenna,

and the other port uses a standard horn antenna at the receiving side to accurately measure the phase and amplitude of the S_{21}.

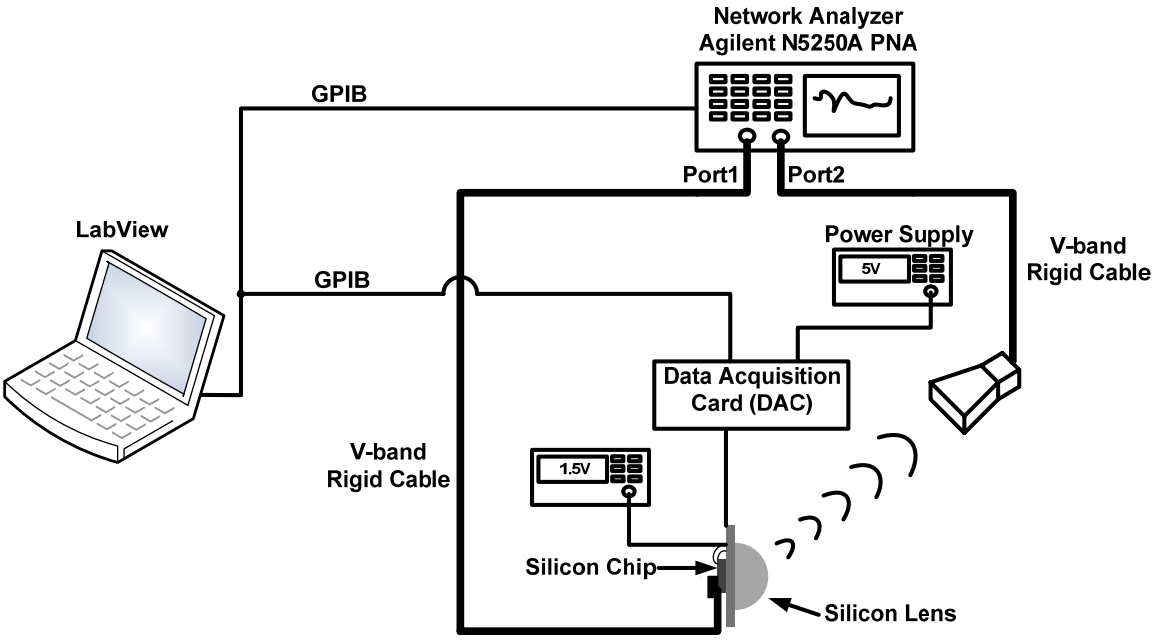


Figure 5.30 Measurement setup

By changing the switching combinations in the switch-based DAM chip we can accurately measure the variation in the phase and amplitude of the S_{21} and plot the real and imaginary parts of the S_{21} on the signal constellation diagram as shown in Figure 5.31. In this figure we have measured the variations of S_{21} in two different directions with an angular separation of approximately 90 degrees for the same set of switching combinations. We have measured about 2000 randomly chosen switching combinations and selected the ones which result in 20 equally spaced constellation points in the desired direction. As shown in Figure 5.31 the signal constellation points are completely scrambled in the undesired direction, proving the functionality of our system. To have a better look at the constellation points, we have used vertical color-coding to separate

these 20 points into several groups. This color-coding, and also the numbering of the constellation points, help us to better understand the scrambling nature of the system. To the best of our knowledge, this is the first published measurement result which proves the existence of a single transmitter capable of sending direction-dependent data.

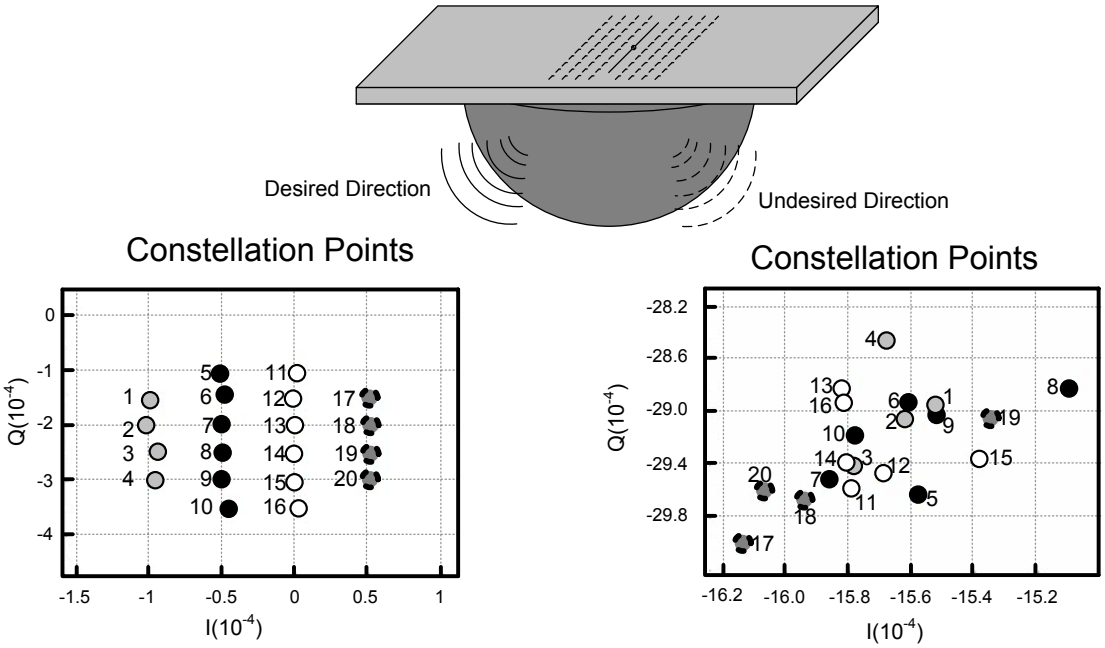


Figure 5.31 Measured constellation points of the switch-based DAM chip. In this measurement only the switches are used.

A picture of the measurement setup is shown in Figure 5.32. The optional quadrant-selector unit in conjunction with the reflector switching is used to cover the four quadrants on the signal constellation diagram. Figure 5.33 shows the measurement results of the full coverage of the four quadrants.

To measure the output power of the transmitter and characterize its linearity, the antenna is disconnected from the PA. A 1.85 mm coaxial wafer probe, an Agilent V-band power sensor (V8684A), and an Agilent V-band power (E4418B) meter are used to

measure the output power of the transmitter. The input signal is generated by an Agilent signal source (E8257D) and applied to the input of the transmitter through a wafer probe. Figure 5.34 shows the measured output power and gain versus the input power of the transmitter. A transmitter output power of +7 dBm, a small-signal gain of 33 dB, and a saturated gain of 25 dB are achieved.

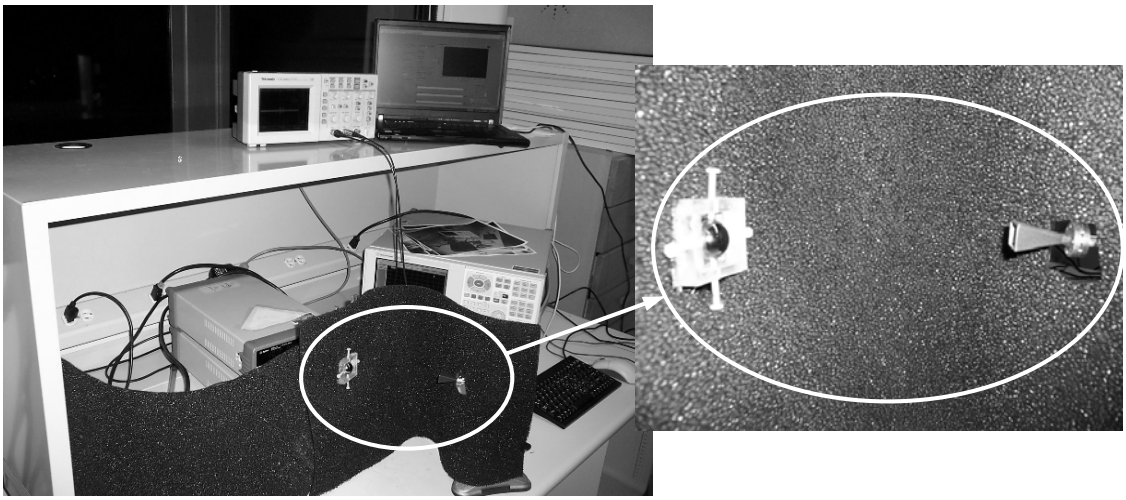


Figure 5.32 Picture of the measurement setup

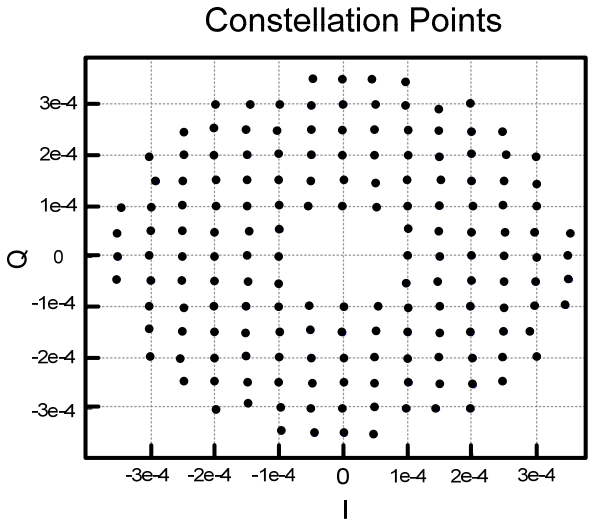


Figure 5.33 Measurement results of the four-quadrant coverage of the signal constellation space using the optional quadrant-selector and switches (switch-based DAM chip)

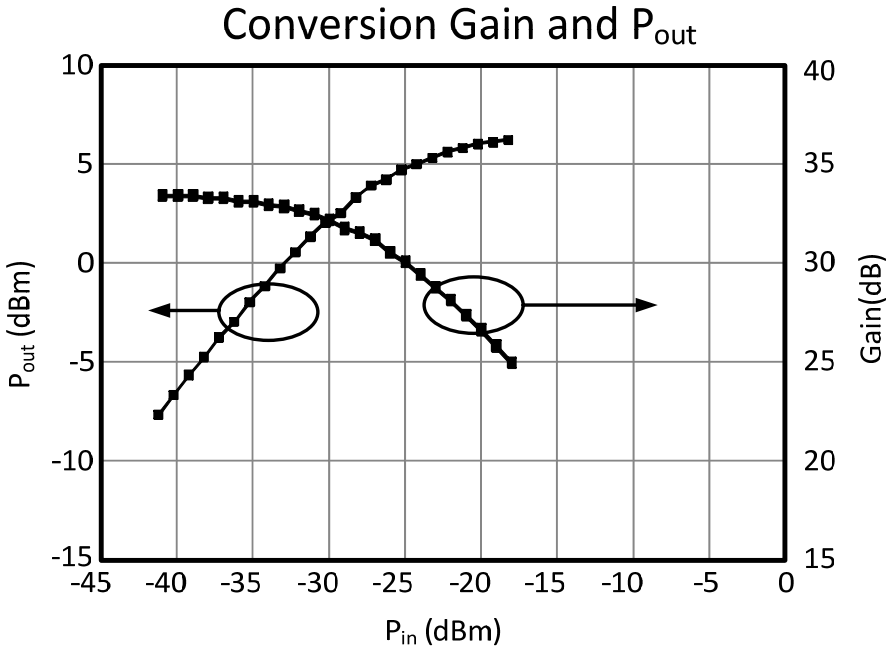


Figure 5.34 Conversion gain and p_{out} versus input power

In order to measure the VCO performance, a GSG test pad connecting to an on-chip LO signal is designed and used. An Agilent spectrum analyzer (E4448A) and a SpacekLabs off-chip down-convertor (GE-590) are used to measure the phase-noise and tuning range of the VCO. A phase-noise of -100 dBc at 10 MHz offset and a tuning range of more than 2.5 GHz are achieved in the measurement.

To measure the performance of the varactor-based DAM chip, a setup similar to Figure 5.30 is used. Figure 5.35 shows the measured constellation points at the received antenna. In this measurement the LabView program changes the pattern of the varactors by generating about 1400 random combinations and plots the variations of S_{21} measured by the Agilent N5250A network analyzer. This measurement proves the functionality of the varactor-based DAM system.

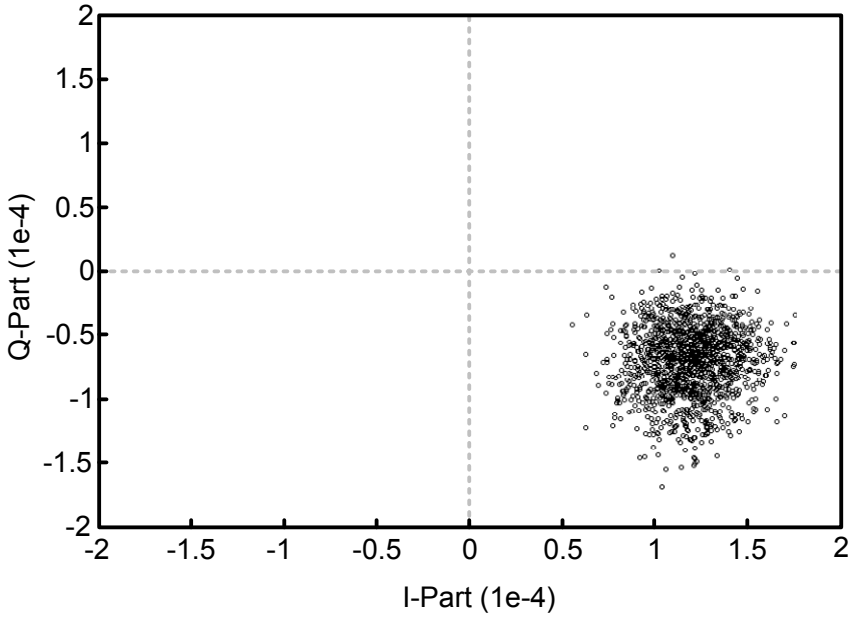


Figure 5.35 Measured constellation points in varactor-based DAM chip

5.8 Appendix

In this part we explain a simulation technique used in analyzing an electromagnetic structure with many switches or varactors. As mentioned in the discussion related to the DAM transmitter architecture, each switching combination defines a unique boundary condition around the antenna. In the switch-based DAM design, 90 switches are used where each assumes one of two states, on and off. This results in 2^{90} switching combinations which correspond to 2^{90} unique boundary conditions. While we have not simulated each unique boundary condition (as doing so is not particularly practical or time efficient, and fortunately not necessary), we have simulated 10^3 to 10^5 combinations and selected a small set of these combinations to implement an arbitrary digital modulation. If we use conventional techniques to treat each switching combination as an independent EM problem and use EM solvers such as IE3D [9] or HFSS [33] to calculate

the far-field for each switching combination, we will not be able to simulate a large enough number of the switching combinations due to the limitation of the simulation time. If we choose Method of Moment (MOM) and use IE3D to simulate the structure, each simulation takes about 5–10 min. To simulate the EM structure using finite element software, such as HFSS, each simulation can take up to 30–60 min. None of these techniques can be used to simulate 1000 points in a time-efficient manner.

To simulate this many points, we extracted a circuit model of the whole EM structure and used a combination of circuit software and Matlab code to run each simulation in 10–100 msec. Here we intend to discuss the details of this technique. Figure 5.36 shows three different EM problems. Problem 1 shows a transmitting antenna, such as dipole, with an arbitrarily shaped perfect electric conductor (PEC) adjacent to it. We have shown the PEC with a thick black curve. This PEC mimics the reflectors adjacent to the antenna. As shown in Figure 5.36, we have disconnected the PEC at a couple of locations and placed terminations with arbitrary impedances at the discontinuities. These impedances are used to model the switches and the varactors. We have also placed a receiving antenna at an arbitrary direction at the far field. This antenna is used to probe the far field at a specific angle.

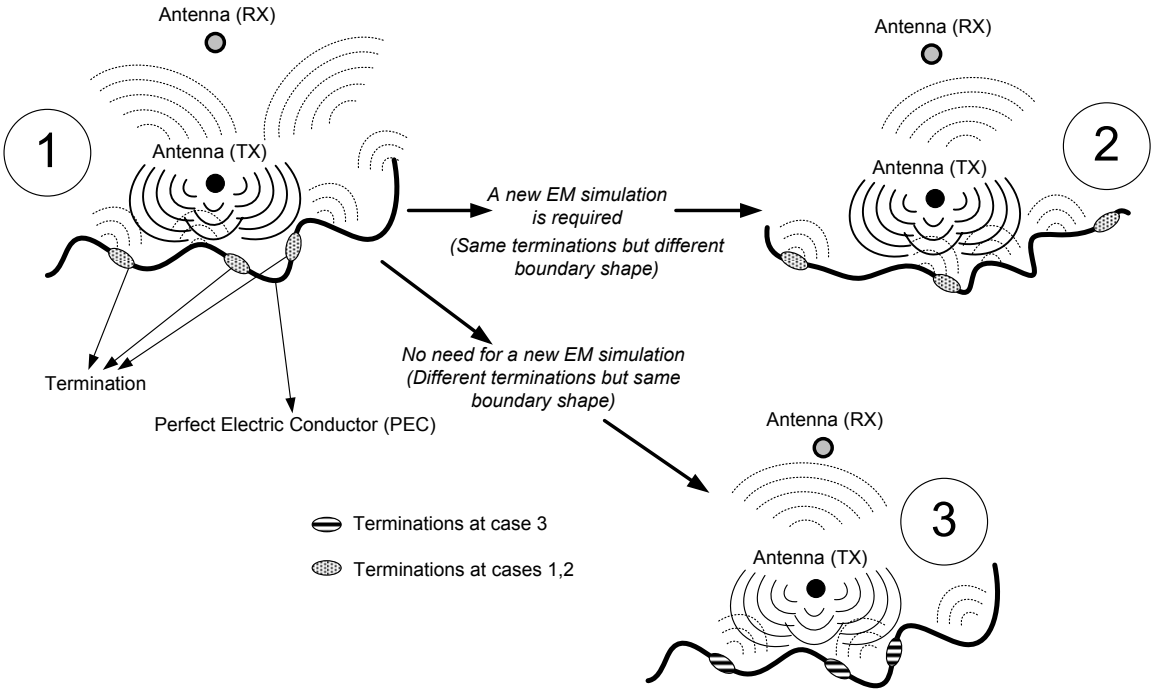


Figure 5.36 Comparison of two boundary value EM problems

In problem 2 we change the physical shape of the PEC without changing the termination impedances. To calculate the received voltage of the receiving antenna located at the far field, we have to run a new EM simulation and cannot use the EM simulation results of problem 1 to calculate the far-field pattern of problem 2. In problem 3 we vary the values of the termination impedances but keep the physical shape of the PEC unchanged. To calculate the far field in problem 3, there is no need to run a new EM simulation. In fact the simulation results of problem 1 can be used to calculate the voltage at the input port of the receiving antenna located at the far field. To do that, we replace the termination impedances in problem 1 with localized differential ports. We also place localized differential ports at the input ports of the transmitting and the receiving antennas. With one receiving antenna, one transmitting antenna, and N-2 terminations on the PEC, N independent ports are defined in problem 1. Now we can run the EM

simulation for problem 1 and extract an $N \times N$ *S*-parameter matrix containing all the information required to solve problem 3. To solve problem 3 we use the *S*-parameter matrix of problem 1 but we change impedance values of the terminations in the circuit model (Figure 5.37).

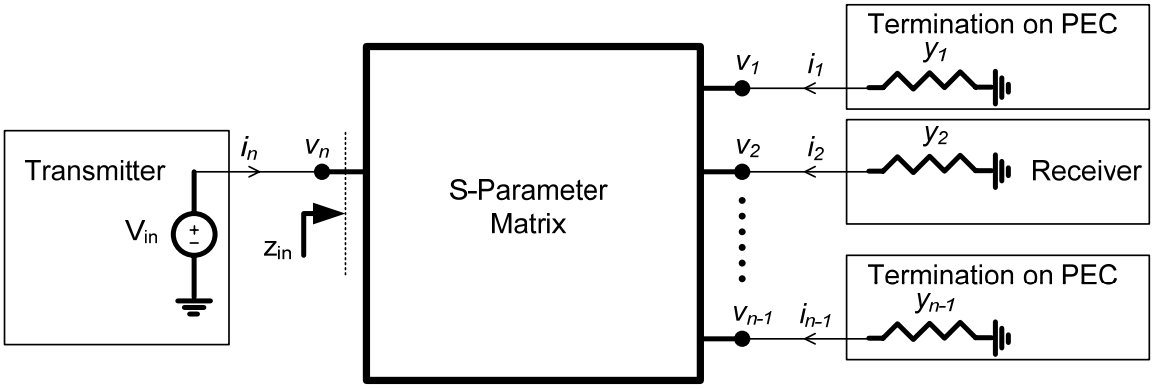


Figure 5.37 S-parameter matrix of problem #1

A Matlab code extracts the *S*-parameter data from an EM solver, changes the termination impedances or conductances, and computes the far field by calculating the voltage at the terminal of the receiving antenna.

From Figure 5.37, we can write the following equations,

$$\begin{bmatrix} i_1 & . & . & i_n \end{bmatrix} = \overline{Y}^T \cdot \begin{bmatrix} v_1 & . & . & v_n \end{bmatrix} \tag{5.5}$$

$$\begin{bmatrix} i_1 & . & . & i_n \end{bmatrix} = -\begin{bmatrix} v_1 & . & . & v_n \end{bmatrix} \cdot \begin{bmatrix} y_1 & 0 & 0 & 0 & 0 \\ 0 & . & 0 & 0 & 0 \\ 0 & 0 & . & 0 & 0 \\ 0 & 0 & 0 & y_{n-1} & 0 \\ 0 & 0 & 0 & 0 & -z_{in}^{-1} \end{bmatrix}. \tag{5.6}$$

In equations (5.5) and (5.6), $i_1,...,i_n$ represent the current values of the ports and $v_1,...,v_n$ represent the voltage values of the ports. Also y_1, $y_3,...,y_{n-1}$ represent the conductance of the terminations, y_2 represents the termination conductance of the receiving antenna, and z_{in} represents the impedance seen at the input port of the transmitting antenna. $\overline{Y}^T$ is the transposed *Y*-parameter matrix of the system and can be calculated from the *S*-parameter data by the following equation,

$$\overline{Y}^{-1} = Z_0(\bar{I} - \overline{S})^{-1}(\bar{I} + \overline{S}) \ . \tag{5.7}$$

In equation (5.7), Z_0 is the nominal impedance (50 Ω), $\bar{I}$ is the identity matrix, $\overline{S}$ is the *S*-parameter matrix, and $\overline{Y}^{-1}$ is the inverse of the *Y*-parameter matrix.

In equations (5.5), (5.6), and (5.7) we have 2*N* unknowns and 2*N* equations and hence have a unique nonzero solution. The unknown parameters are $v_1,...,v_{n-1},i_1,...,i_n,z_{in}$, and the known ones are $v_n,y_1,..,y_{n-1}$. v_n is the input voltage of the transmitting antenna which is a known parameter. Figure 5.38 shows the comparison between the conventional and the circuit-model-based optimization techniques in block flowchart form. In the conventional scheme, one EM simulation is required for each iteration. However, in the circuit-model-based technique, we only need to run a single EM simulation for the whole optimization problem. Due to the fact that most of the simulation time is consumed by the EM solver, the circuit-model-based scheme runs orders of magnitude faster than that of the conventional method.

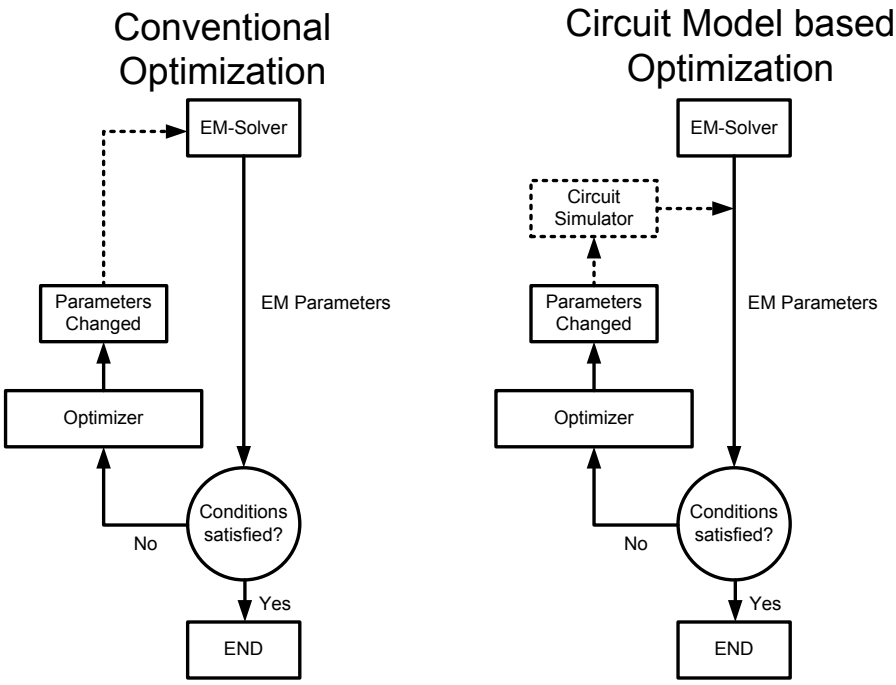


Figure 5.38 Comparison of the optimization techniques

To highlight the advantage of the above technique we have optimized a simple patch antenna by using these two techniques. Figure 5.39 shows a simple patch optimization problem. The edge-fed patch with a ground shield is composed of 9 square metal pieces. We have connected these metal squares by using 12 terminations, as shown in Figure 5.39.

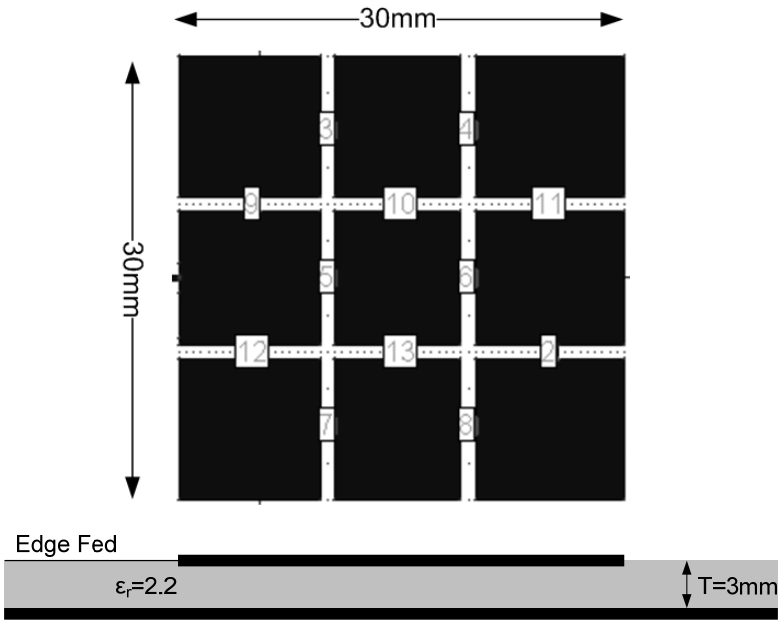


Figure 5.39 Patch optimization problem

The optimization goal is to maximize the patch bandwidth by finding optimum impedance values for the terminations. The range of the impedances is chosen based on the practical values of the surface mount inductors and capacitors. Our circuit-model-based Matlab code uses golden section search and parabolic interpolation to find an optimal solution for the impedance values of the terminations. Figure 5.40 shows the simulation results of the patch S_{11} versus frequency. By using one microprocessor we have achieved a maximum bandwidth of 210 MHz with simulation time of 85 seconds and 12986 iterations. A conventional optimization technique based on PSO/FDTD uses four processors and achieves a bandwidth of 150 MHz in 16 hours and 200 iterations [34]. This conventional technique optimizes the dimensions of the patch and the location of its coaxial feed point to maximize the bandwidth. This simple comparison proves the power of the circuit-model-based technique in optimizing the EM problems.

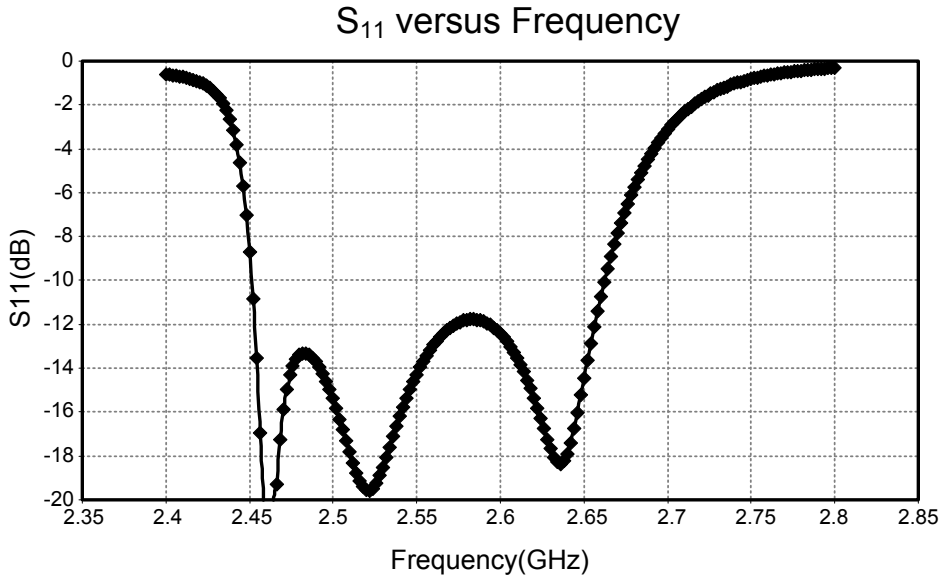


Figure 5.40 Optimized S_{11} versus frequency (patch problem)

5.9 Chapter Summary

In this chapter, the technique of direct antenna modulation (DAM) is introduced. Two different methods for implementing DAM systems are presented. The transmitter architectures based on DAM technique are capable of transmitting direction-dependent data. These systems can be used to enhance the security of the communication links, transmit independent data to multiple directions using a single transmitter, and utilize narrow-band highly efficient switching PA in transmission of wide-band constant and non-constant envelope-modulated signals. To support the idea, two 60-GHz proof-of-concept chips are implemented and measured.

Chapter 6

A Scalable 60 GHz Phased-Array Transmitter[8]

6.1 Introduction

This work demonstrates the application of a new circuit approach based on coupled-oscillator arrays for millimeter-wave phased-array designs in a standard SiGe integrated-circuit technology. To maintain phase coherence between phased-array elements, oscillators on different dies must be phase locked. While coupled phase-locked loop architectures can maintain the phase relationship between neighboring oscillators coupled oscillators are well suited for fully integrated phased-array circuits [35] – [37]. The wavelength at 60 GHz allows for several array elements to be located on a single die. Consequently, an oscillator can be placed on each die or at each transmit element. One particular advantage of an on-chip coupled-oscillator array is the distribution of the carrier frequency between different transmit elements. Global frequency distribution becomes increasingly difficult at millimeter frequencies. By co-locating an oscillator at each transmit stage, closer control of the delay mismatches between the local oscillator and the mixer stage is possible. Scalable architectures are particularly appropriate for creating large phased arrays by tiling several dies. In Figure 6.1 scalability is demonstrated for a phased-array transmitter [37]. The phased array requires phase coherence between all elements. Since each chip requires frequency generation,

[8] A joint project by A. Babakhani, J. F. Buckwalter, and A. Komijani. A. Babakhani designed the PA and on-chip antennas, J. F. Buckwalter designed the injection locked oscillators, and A. Komijani designed the phase shifters.

neighboring chips are injection locked to ensure phase coherence. If oscillators are located at each antenna element, injection locking can lock the oscillators both on-chip and between chips. This tiling approach is useful not only for phased arrays, but also for applications where the separation between antennas might be several wavelengths. Consequently, this design approach focuses on an injection-locking scheme appropriate for an on- and off-chip coupled-oscillator array. In this transmitter implementation, a two-element-by-two-element (2×2) array of transmitter cells is integrated on a single chip with antennas.

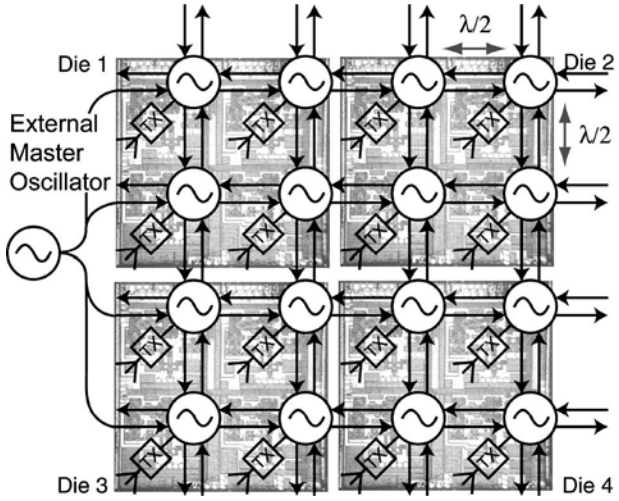


Figure 6.1 Coupled-oscillator scheme for fully integrated phased-array transmitter [37]

In this chapter, we will discuss the design and measurement of the 60 GHz power amplifier used in the coupled-oscillator phased-array transmitter. The design details of the injection-locked oscillator can be found in [37].

6.2 Power Amplifier Design

The 60 GHz power amplifier is composed of three differential stages. Figure 6.2 shows the output stage of the PA. Each output transistor has an effective width of 48 μm. To increase the stability of the PA, a parallel combination of a 300 Ω resistor and 0.4 pF capacitor is placed on the signal path just before the base of the transistor in the differential pair. At lower frequencies, the capacitor is almost open and a 300 Ω series resistor significantly reduces the gain of the PA. At 60 GHz, the PA gain is not affected by this parallel combination due to the small impedance of the capacitor at 60 GHz (6.6 Ω).

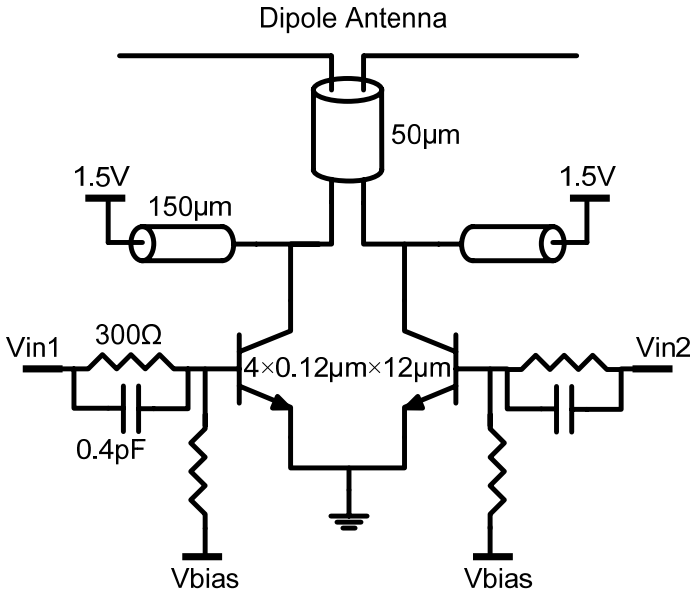


Figure 6.2 Output stage of the PA

Stub tuning is used at the output of the PA to match the output impedance of the PA to the differential 50 Ω impedance of the dipole antenna. Load-pull and source-pull

simulations are performed for each stage independently to find optimum load and source impedances for each stage. Figure 6.3 shows the load-pull simulation results of the output stage.

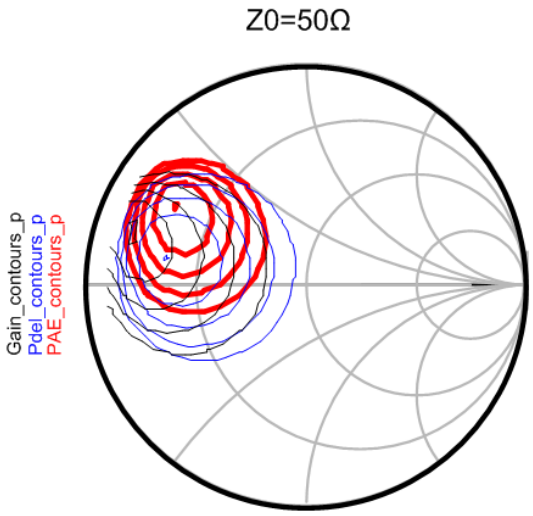


Figure 6.3 PA output stage load-pull

In the above figure, the red circles represent the PAE contours, the blue ones show the delivered power contours, and the black ones demonstrate the gain contours. At the chosen load impedance, simulated PAE of 30%, delivered power of 16 dBm, and power gain of 5 dB have been achieved. Figure 6.4 shows the simulated PAE, delivered power, and gain of the output stage versus input power at 60 GHz. To maximize the flexibility in the measurement, base bias voltage of each differential stage can be controlled independently through a separate pad. PA is designed to operate with a supply voltage of 1.5 V.

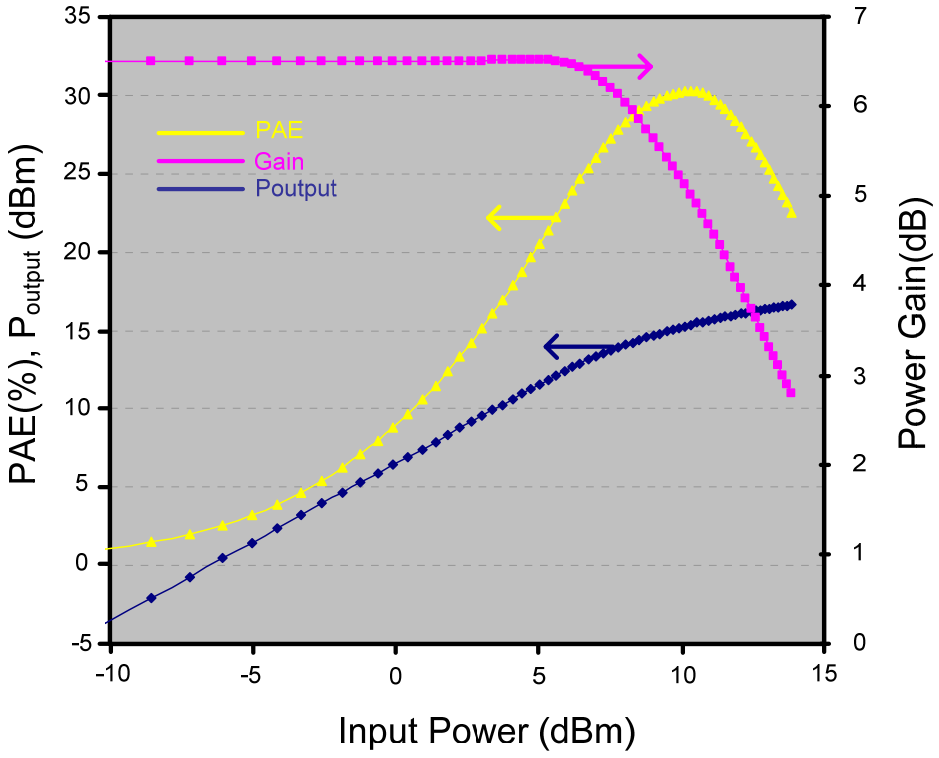


Figure 6.4 Simulated PAE, delivered power, and gain of the output stage

Figure 6.5 shows the schematics of the second stage of the PA. Each transistor of the differential pair has an effective width of 12 µm. Similar to the output stage; stub tuning is used to match the second stage to the output stage. The second stage generates an output power of 10 dBm and provides a differential impedance of 100 Ω at its input. Simulated P_{out}, PAE, and power gain of the second stage is shown in Figure 6.6. A maximum efficiency of 25% is achieved with input power of about 3.5 dBm. At this input level, simulation results show a power gain of 6 dB and delivered power of 9.5 dBm.

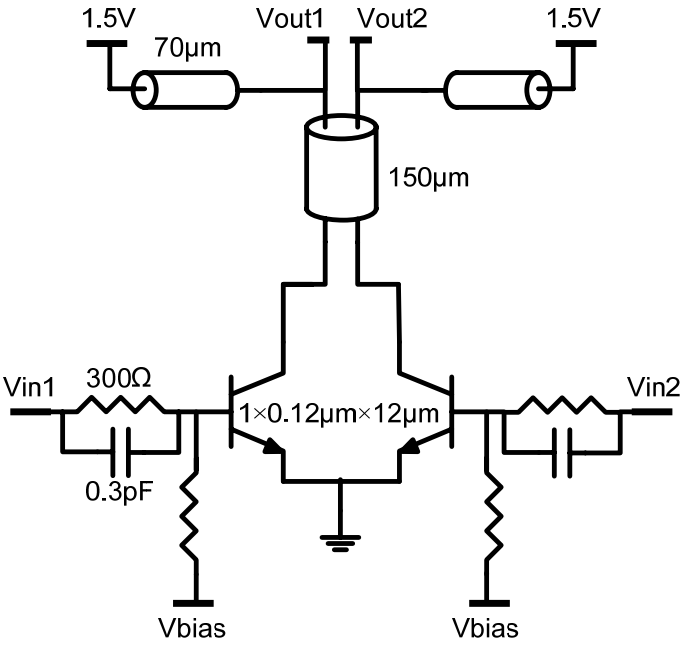


Figure 6.5 Second stage of the PA

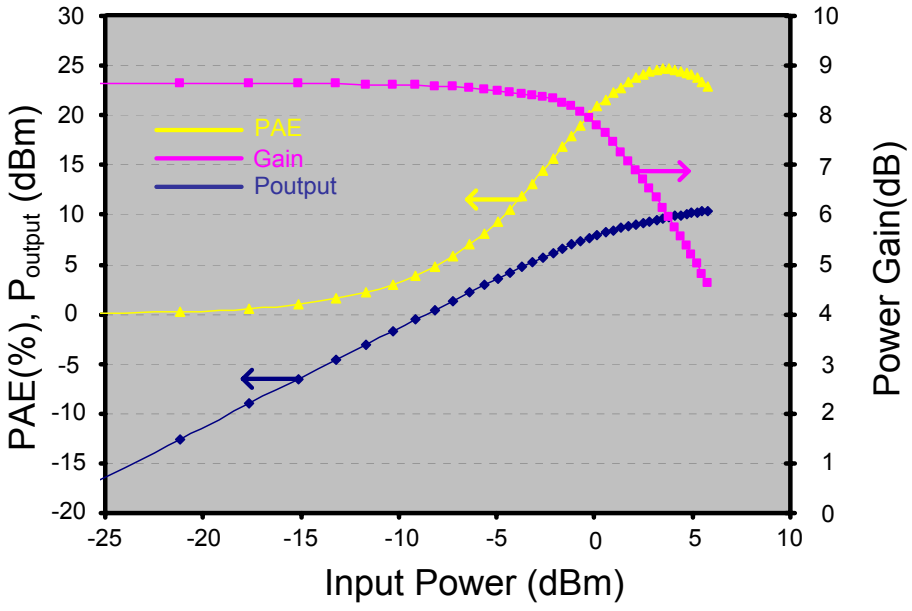


Figure 6.6 Simulated P_{out}, PAE, and power gain of the second stage

Finally, the schematic of the 1st stage is shown in Figure 6.7. In this case, the effective width of each transistor in the differential pair is about 6 μm. Differential input impedance of this stage is about 100 Ω.

Figure 6.8 shows the power gain, PAE, and delivered power versus input power for the combined three stages. At input power of -5 dBm, simulated output power, PAE, and power gain are about 15 dBm, 25%, and 19 dB, respectively.

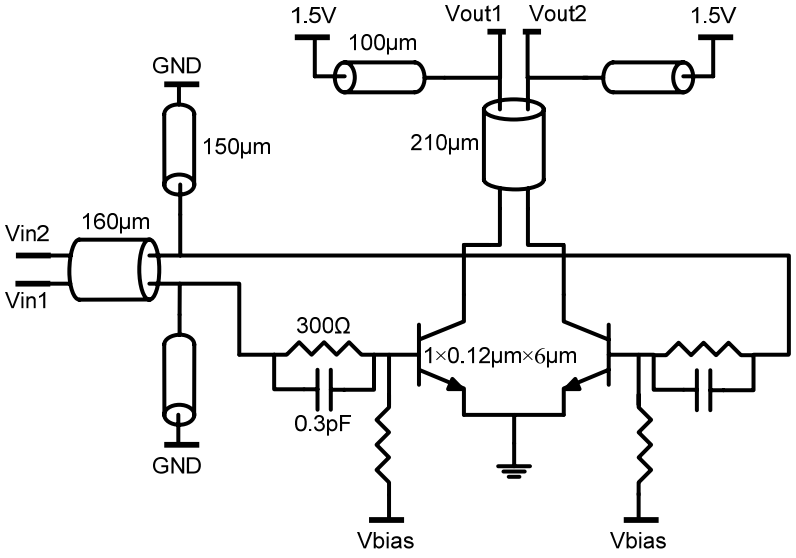


Figure 6.7 Schematic of the 1st stage

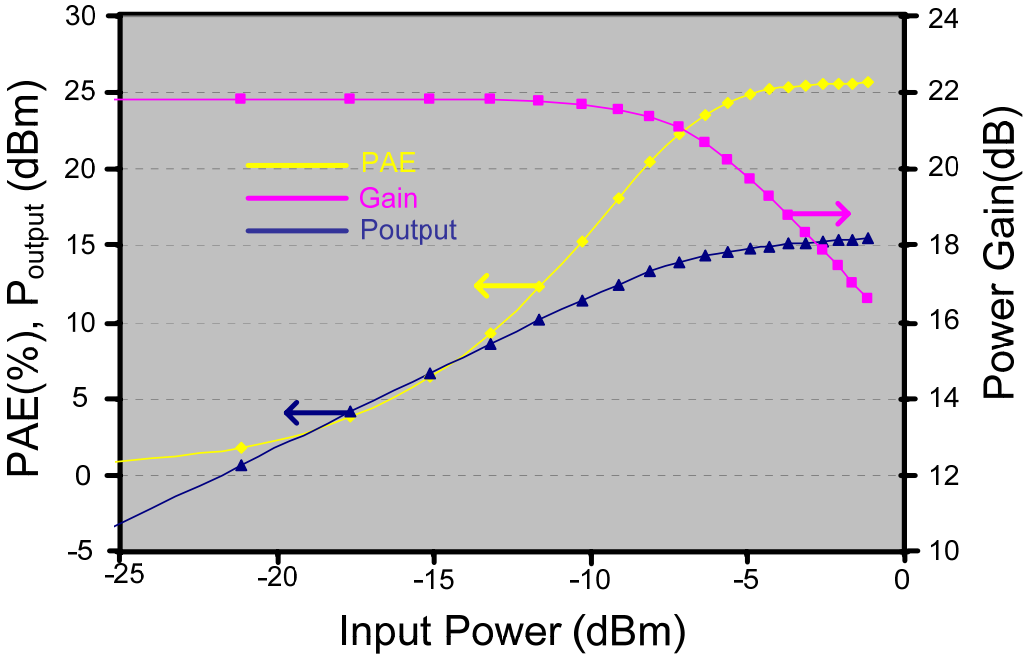


Figure 6.8 Simulated P_{out}, PAE, and power gain of the combined three stages

6.3 Power Amplifier Layout

Layout of the 60 GHz PA is shown in Figure 6.9. Transmission lines are implemented on metal layers AM (M7) and MQ (M5). Signal lines are made on AM and MQ is used as a bottom metal layer. To isolate the transmission line from other blocks, adjacent AM ground layers are also used to minimize the leakage of the signal power. Output of the PA is connected to a differential on-chip dipole antenna, as well as GSGSG pads. To characterize the PA parameters using differential probes, laser trimming is used to disconnect the PA from the on-chip antennas. In the measurement of the PA using the on-chip dipole antennas, laser trimming is used to disconnect the pads from the PA.

Transmission lines are meandered to minimize the PA area. Due to some layout floor-planning constraints, the first stage of the PA is rotated by 90 degrees, as shown in Figure 6.9.

Figure 6.10 shows the layout of the 60 GHz 2×2 scalable transmitter. Each transmitter includes an on-chip dipole antenna, a power amplifier, LO-signal phase rotators, an injection locked oscillator, and a DAC for programming the phase rotators.

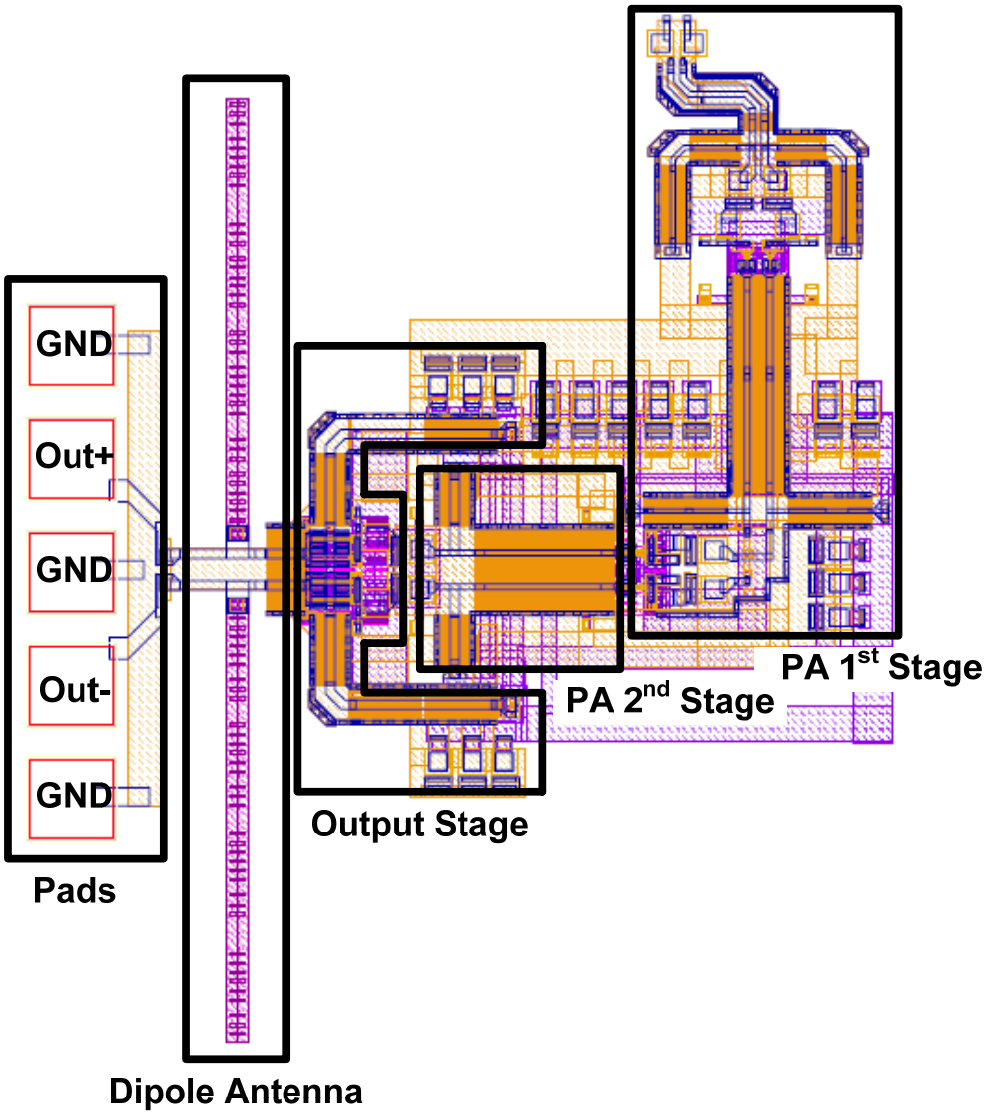


Figure 6.9 Layout of the 60 GHz PA

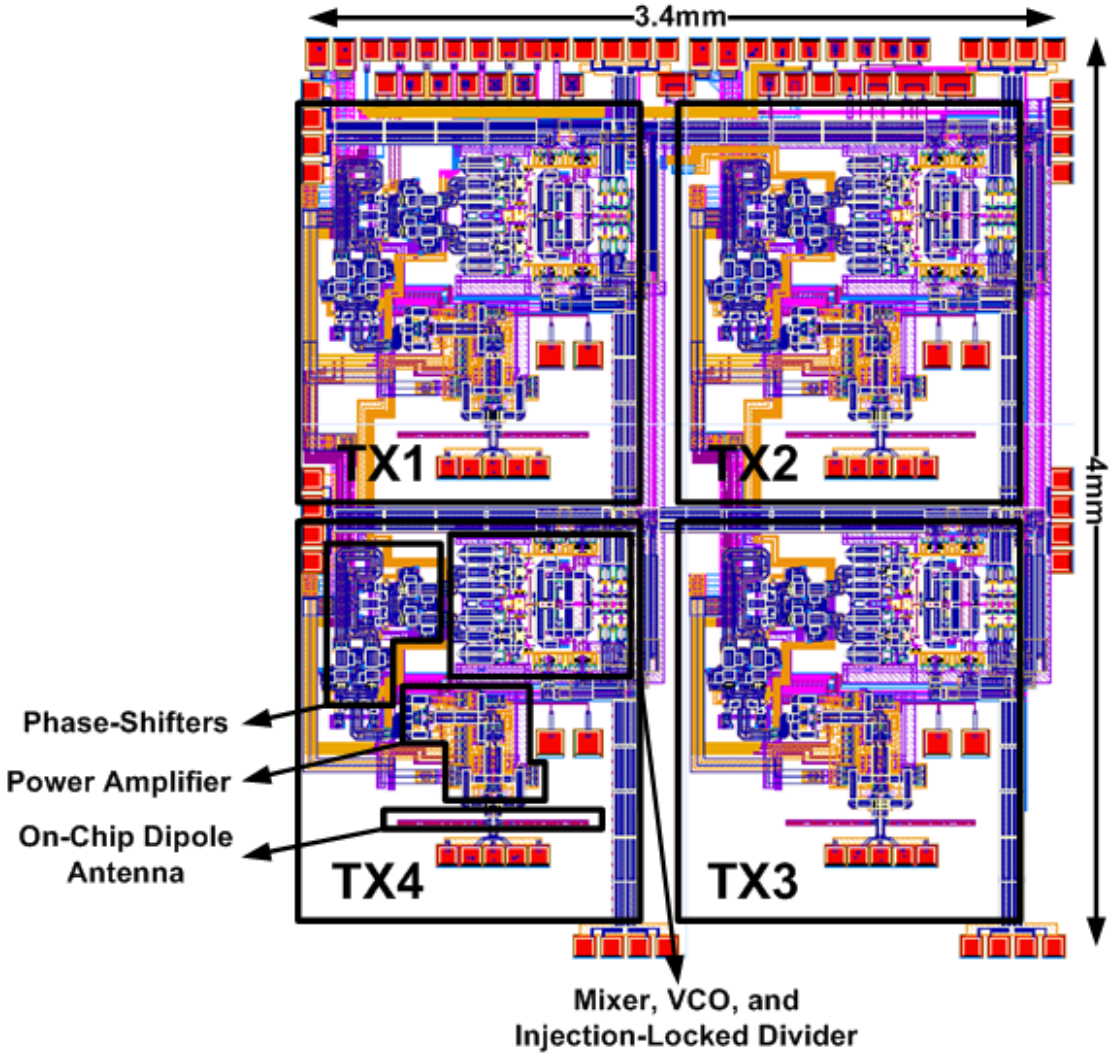


Figure 6.10 Complete layout of the 60 GHz 2×2 scalable transmitter.

6.4 60 GHz Balun and PA Test Structure

The V-band power amplifier used in the system has differential input and differential output pads. This makes the measurement of the PA very challenging. To be able to characterize PA performance independently, we have designed an on-chip 180° hybrid

balun which converts the differential signal to a single-ended one. The design of the balun is shown in Figure 6.11. In this figure, lengths of transmission lines are shown in degrees at 60 GHz. An on-chip 50 Ω resistor is used to terminate the isolated port.

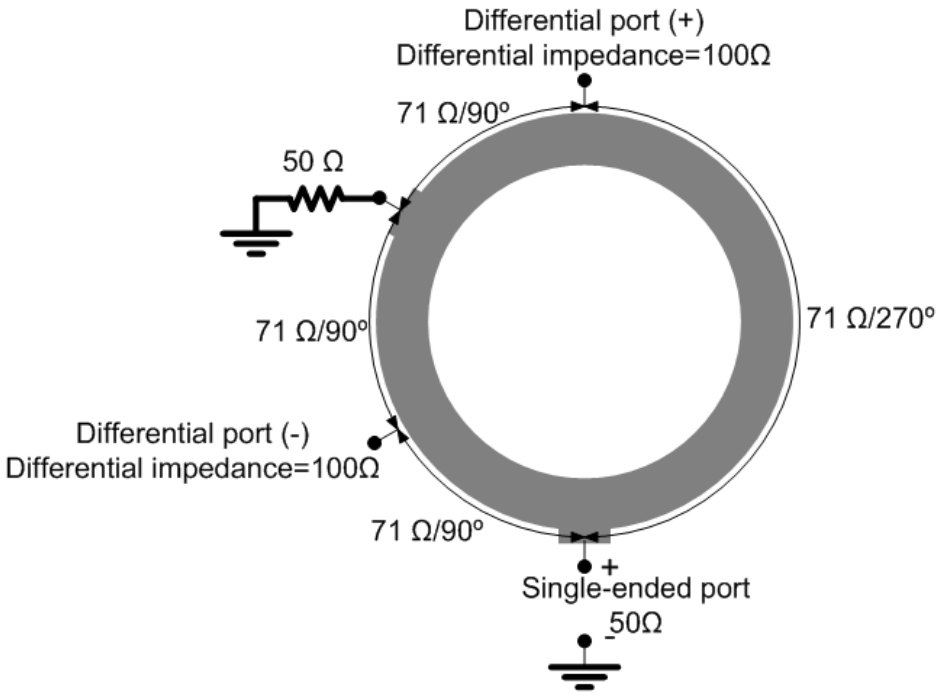


Figure 6.11 60 GHz 180° hybrid balun

Figure 6.12 shows the layout of the single PA connected to a hybrid 180° balun. The balun converts the differential input signal of the PA to a single-ended signal which can be easily driven with a single-ended V-band probe. Differential output of the PA is connected to a differential on-chip dipole antenna, as well as GSGSG pads. PA ouput stage is designed to drive the differential dipole antenna with an input impedance of 50 Ω. To be able to characterize the PA performance independent of the dipole antenna, laser trimming is used to disconnect the dipole antenna from the PA. In this case, an on-chip 25 Ω resistor is used to terminate one of the PA outputs and the other output is

connected to a single-ended V-band probe. By using these probes at the input and output of the PA, we can accurately measure the S-parameters of the PA.

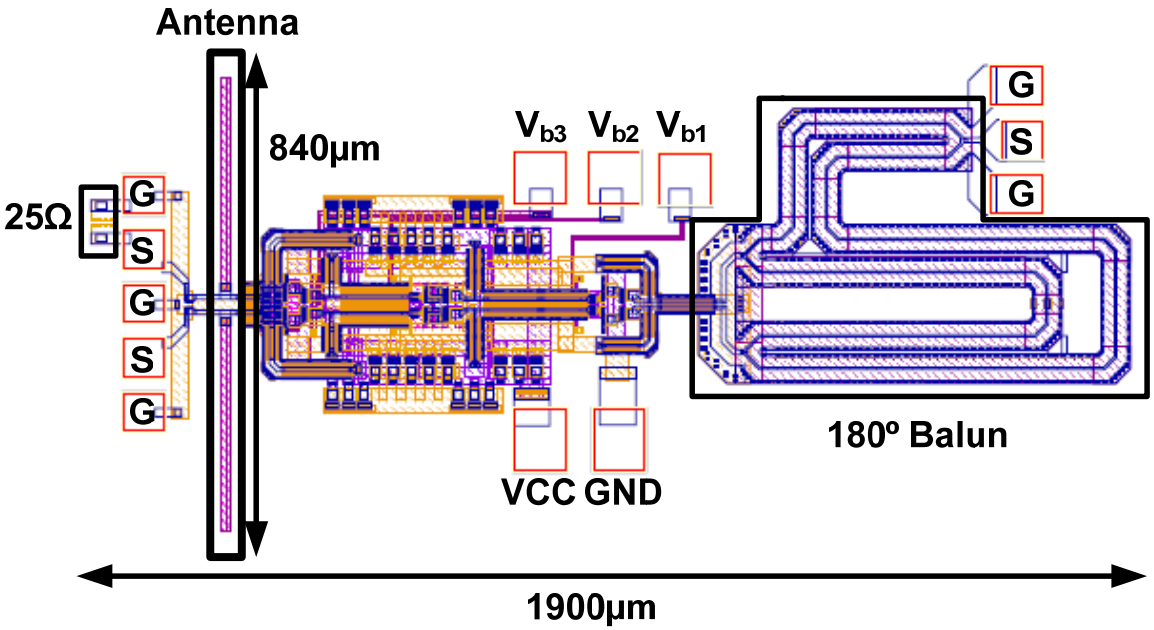


Figure 6.12 Stand-alone PA with input balun

One of the important parameters that needs to be measured is the effective loss of the balun. To characterize the balun performance itself, we have designed and connected two baluns side by side. In this combination, the differential ports of the balun are connected together and single ports are driven with single-ended V-band probes. Figure 6.13 shows the layout of these two baluns. Based on the measured results, the effective loss of each balun is about 2 dB. Figure 6.14 shows the measurement results of the standalone PA performance.

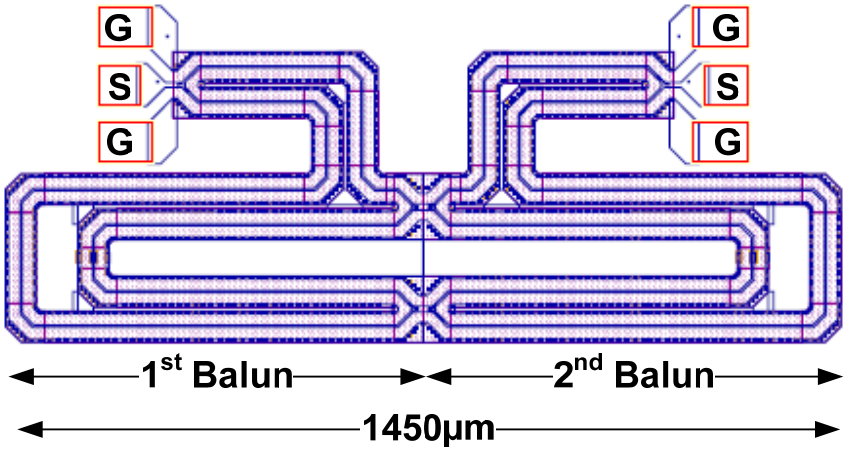


Figure 6.13 Layout of two connected baluns

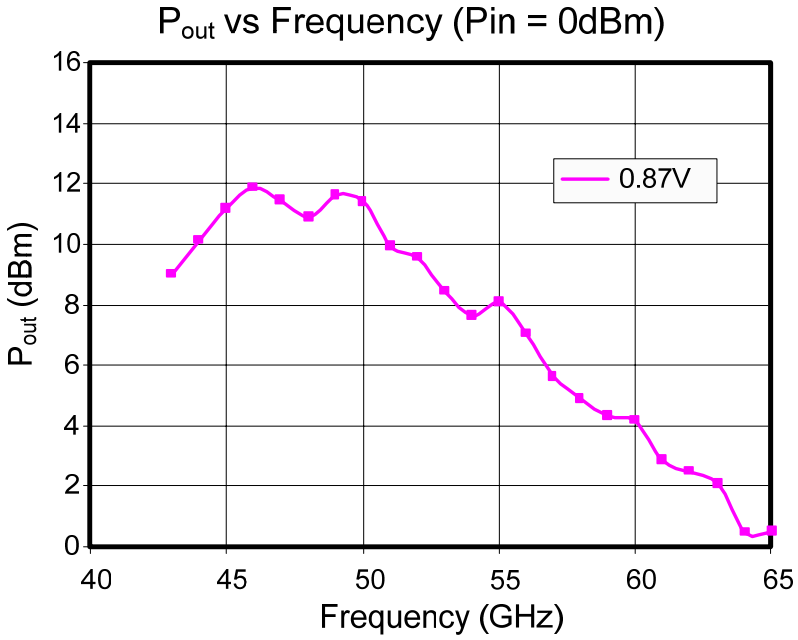


Figure 6.14 Measured power-versus-frequency of the stand-alone PA

The peak measured power using a probe is about 12 dBm. As mentioned before, the PA output is designed to drive an on-chip dipole antenna with input differential impedance of 50 Ω and single-ended impedance of 25 Ω. The single-ended probe used in

our measurement has an impedance of 50 Ω. Due to the mismatch between the probe impedance and the PA output, the measured peak power occurs between 46 GHz and 49 GHz instead of 60 GHz. At 50 GHz with 0.87 V bias and 0 dBm input power, a PAE of 21% has been achieved. To the best of our knowledge, this is the highest efficiency number ever achieved in a mm-wave PA implemented on a silicon substrate.

Figure 6.15 shows the die micrograph of the 60 GHz transceiver chip. The on-chip dipole antenna is used to feed a silicon hemispherical lens located at the back side of the chip. This lens couples the radiated power to the air.

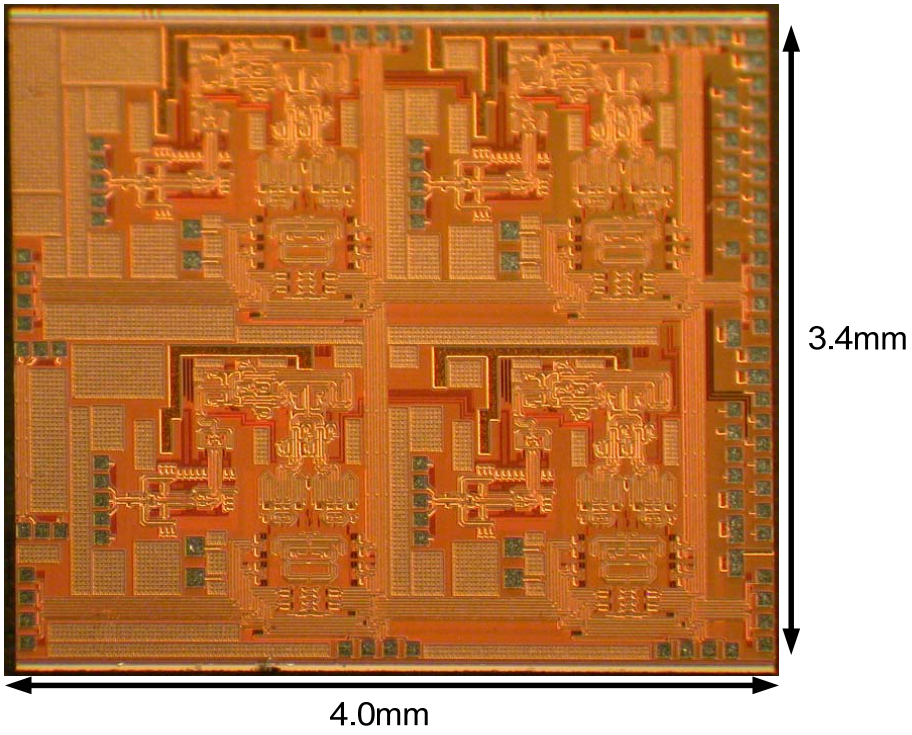


Figure 6.15 Die Micrograph of the 60 GHz scalable transmitter with on-chip antennas

6.5 Chapter Summary

In this chapter, the design of a 60 GHz scalable coupled transmitter phased-array is discussed. The technique used in this transmitter can be used to overcome the issues of global LO distribution in large arrays. The design and measurement results of a 60 GHz PA used in this scalable transmitter are presented. A 180° on-chip hybrid balun is used to measure the performance of the stand-alone PA. A hemispherical silicon lens is utilized to couple the power efficiently to the on-chip dipole antennas.

Chapter 7

Conclusion

In this dissertation, the design and implementation of several silicon-based millimeter transceivers with on-chip antennas are presented. These include the first fully integrated 77 GHz four-channel phased-array transceiver with on-chip antennas in silicon, a 60 GHz scalable 2×2 phased-array transmitter with on-chip dipole antennas in silicon, and fully integrated transmitter architectures based on direct antenna modulation (DAM).

We have reviewed the concept of the substrate modes and emphasized its importance in design of efficient on-chip antennas. In most of the measured chips, a silicon lens is used to minimize the power coupled into the substrate modes. In all of the process technologies discussed in this thesis, due to the fabrication problems and the necessity of using highly doped silicon, substrates become extremely lossy (substrate resistivity of 1–10 Ω.cm). To minimize the effect of the substrate loss, the silicon die is thinned down to 100 μm and an undoped silicon wafer is placed underneath the die to ensure the mechanical stability of the system.

In the receiver of the 77 GHz transceiver, a two-step down-conversion scheme is used with a single VCO. The signal combining is performed using a novel distributed active combining amplifier at IF. In the LO path, a cross-coupled quadrature injection locked frequency divider (QILFD) divides the 52 GHz VCO frequency by a factor of 2 and is followed by a divide-by-512 divider chain. Conversion gain of more than 37 dB, 2 GHz

BW, and 8 dB noise-figure are achieved. The on-chip antenna gain of 8 dB is achieved in the measurement.

The 60 GHz 2×2 transmitter chip uses an injection locked scheme to solve the global LO distribution problems in phased-arrays. Because of the unique configuration, multiple copies of this transmitter can be used to implement a scalable two-dimensional phased-array system. The V-band stand-alone PA provides a peak output power of 12 dBm and a PAE of 21%. To the best of our knowledge, this is the highest efficiency number ever achieved for a mm-wave PA implemented on a silicon substrate.

We have introduced the technique of direct antenna modulation (DAM). In the transmitter architectures based on DAM, the base-band data controls the state of the switches (or varactors) on the reflectors through a digital control unit. In these systems, because the modulated signals do not pass through the PA, a narrow-band highly efficient switching PA can be used to amplify the locked LO signal while the whole system is capable of transmitting wide band arbitrary modulated signals (including non-constant envelope signals). Because of the unique characteristic of these systems, it is possible to transmit independent signals to different directions using a single transmitter. This increases the security of the communication links by transmitting a properly modulated signal to a desired direction while scrambling the modulation points on the signal constellation diagram in the undesired directions. To prove the concept, two 60 GHz DAM chips are designed and tested. One of these chips uses the switches on the reflectors to change the scattering properties of the reflectors, change the phase and amplitude of the modulated signal in the far field, and hence modulate the far-field signal.

The second chip uses NMOS varactors on the reflectors to change the electrical properties of the reflectors and modulate the signal. In both of these chips 10 reflectors are used and 9 switches (or varactors) are placed on each reflector. We have reported successful measurement results and proved the concept.

As mentioned in the introduction, the fast progress in developing process technologies funded by the worldwide consumer electronic market opens a new plethora of opportunities for RF/microwave designers doing research in the millimeter and sub-millimeter wave frequency range. These opportunities carry their own set of challenges, which, if properly addressed, can open the door for creation of revolutionary technologies.

References

[1] H. Kogelnik, in Integrated Optics (edited by T. Tamir), ch. 2. Springer-Verlag, New York, 1979.

[2] D. B. Rutledge et al., "Integrated-Circuit Antennas," *Infrared and Millimeter-Waves*, pp. 1–90. New York, Academic, 1983.

[3] A. Babakhani, X. Guan, A. Komijani, A. Natarajan, and A. Hajimiri, "A 77 GHz phased array transceiver with on chip dipole antennas: Receiver and on-chip antennas," *IEEE J. Solid-State Circuits*, 41(12):2795–2806, December 2006.

[4] N. Engheta, and C. H. Papas, *Radio Sci.*17*,* 1557–1566, 1982.

[5] N. G. Alexopoulos, P.B. Katehi, and D. B. Rutledge, "Substrate Optimization for Integrated Circuit Antennas," *IEEE Trans. Microwave Theory and Techniques*, 83(7):550–557, July 1983.

[6] B. Chantraine-Barès, R. Sauleau*,*L. Le Coq, and K. Mahdjoubi, "A New Accurate Design Method for Millimeter-Wave Homogeneous Dielectric Substrate Lens Antennas of Arbitrary Shape," *IEEE Trans. Antennas and Propagation*, 53(3):1069–1082, March 2005.

[7] D.F. Filipovic, G.P. Gauthier, S. Raman and G.M. Rebeiz, "Off-axis properties of silicon and quartz dielectric lens antennas," *IEEE Trans. Antennas and Propagation*, 45(5): 760–766, May 1997.

[8] T. Nagatsuma et al, "Millimeter-wave photonic integrated circuit technologies for high-speed wireless communications applications," *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, pp. 448–449, February 2004.

[9] MoM-Based Electromagnetic Simulator (IE3D). http://www.zeland.com

[10] Larry Yujiri, Merit Shoucri, and Philip Moffa, "Passive mm-Wave Imaging" *IEEE Microwave Magazine,* pp.39-50, September 2003

[11] www.fcc.gov/oet/info/rules/part15

[12] N. Guo et al., "60-GHzMillimeter-Wave Radio: Principle, Technology, and New Results," *EURASIP Journal on Wireless Communications and Networking,* 7, 68253.

[13] www.wirelesshd.org

[14] "ECC Decision of 19 March 2004 on the Frequency Band 77–81 GHz to be Designated for the Use of Automotive Short Range Radars (ECC/DEC/(04)03)," Eur. Radiocommun. Office. Copenhagen, Denmark, http://www.ero.dk.

[15] K. M. Strohm et al., "Development of Future Short Range Radar Technology," *European Microwave Week 2005, Conference Proceedings, 3–7 October 2005, Paris, France*, pp. 165–168.

[16] A. Niknejad, "mm-Wave Technology: Communication and Medical Imaging," *CITRIS-Europe Research Symposium, 2007.*

[17] A. Natarajan, A. Komijani, X. Guan, A. Babakhani, and A. Hajimiri, "A 77 GHz phased array transmitter with local LO-path phase-shifting in silicon," *IEEE J. Solid-State Circuits*, 41(12):2807–2819, December 2006.

[18] A. Natarajan,"Millimeter-Wave Phased Arrays in Silicon," *Thesis.* California Institute of Technology, pp.19-24.

[19] Xian Guan, "Microwave Integrated Phase Array Receivers in Silicon," *Thesis.* California Institute of Technology*, 2005.*

[20] www.cadence.com

[21] F. J. Kahn, editor, *Documents of American Broadcasting,* fourth edition. Prentice-Hall,

Inc., 1984.

[22] L. W. Lichty and M. C. Topping., *American Broadcasting: A Source Book on the History of Radio and Television*. Hastings House, New York, 1975.

[23] X. Guan, H. Hashemi, and A. Hajimiri, "A fully integrated 24-GHz eight-element phased-array receiver in silicon," *IEEE J. Solid-State Circuits*, 39(12):2311–2320, December 2004.

[24] H. Hashemi, X. Guan, A. Komijani, and A. Hajimiri, "A 24-GHz SiGe phased-array receiver-LO phase shifting approach," *IEEE Trans. Microw.Theory Tech.*, 53(2):614–626, February 2005.

[25] A. Natarajan, A. Komijani, and A. Hajimiri, "A fully integrated 24-GHz phased-array transmitter in CMOS," *IEEE J. Solid-State Circuits*, 40(12):2502–2514, December 2005.

[26] B. A. Floyd, S. K. Reynolds, U. R. Pfeiffer, T. Zwick, T. Beukema, and B. Gaucher, "SiGe bipolar transceiver circuits operating at 60 GHz," *IEEE J. Solid-State Circuits*, 40(1):156–167, January 2005.

[27] B. Razavi, "A 60-GHz CMOS receiver front-end," *IEEE J. Solid-State Circuits*, 41(1):17–22, January 2005.

[28] C. H. Doan, S. Emami, A. M. Niknejad, and R. W. Brodersen, "Millimeter-wave CMOS design," *IEEE J. Solid-State Circuits*, 40(1):144–155, January 2005.

[29] A. Babakhani, D. B. Rutledge, and A. Hajimiri, "A Near-Field Modulation Technique Using Antenna Reflector Switching," *IEEE ISSCC Dig. Tech. Papers*, pp. 188–189, February 2008.

[30] M. J. M. van derVorst, P. J. I. de Maagt, and M. H. A. J. Herben, "Effect of internal reflections on the radiation properties and input admittance of integrated lens antennas,"

IEEE Trans. Microw. Theory Tech., 47(9):1696–1704, September 1999.

[31] P. Focardi, W. R. McGrath, and A. Neto, "Design guidelines for terahertz mixers and detectors," *Infrared and MillimeterWaves and 13th Int. Conf. Terahertz Electronics*, 2:624–625, September 19–23, 2005

[32] www.ni.com/labview

[33] www.ansoft.com/hfss

[34] N. Jin and Y. Rahmat-Samii, "Parallel Particle Swarm Optimization and Finite-Difference Time-Domain (PSO/FDTD) Algorithm for Multiband and Wide-Band Patch Antenna Designs," *IEEE Trans. Antennas Propagation*, 53(11):3459–3468, November. 2005.

[35] J. J. Lynch and R. A. York, "A mode-locked array of coupled phase-locked loops," IEEE Microw. Guided Wave Lett., 5(7):213–215, July 1995.

[36] J. F. Buckwalter, T. H. Heath, and R. A. York, "Synchronization design of a coupled phase-locked loop," *IEEE Trans. Microw. Theory Tech.*, 51(3):952–960, March 2003.

[37] J. F. Buckwalter, A. Babakhani, A. Komijani, and A. Hajimiri, "An Integrated Subharmonic Coupled-Oscillator Scheme for a 60 GHz Phased-Array Transmitter," *IEEE Trans. Microw. Theory Tech.*, 54(12):4271–4280, December 2006.