MICROWAVE INTEGRATED PHASED ARRAY RECEIVERS IN SILICON

Thesis by

Xiang Guan

In Partial Fulfillment of the Requirements for the

degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2005

Defended September 22, 2005

Acknowledgement

First of all, I would like to express my sincere appreciation for my research advisor, Professor Ali Hajimiri, for his guidance and support over this five year journey. I would like to thank him for bringing me into the High-Speed Integrated Circuits Group at Caltech, where I have been able to work in my most favorite research field. Inspirations drawn from the fruitful and enlightening technical discussions with him have helped me overcome plenty of difficulties encountered in the design and experiments and have been crucial in making those projects successful. I am especially grateful for his constant encouragement, which will continue to promote my desire to do my best in my future career.

I would like to thank Professor David Rutledge and Dr. Sander Weinreb. I have learned a lot in the technical discussions with them. Their valuable suggestions as well as support in providing test equipments in the Caltech Microwave Laboratory have appreciably accelerated the progress of my research projects.

I feel very lucky to have worked with the current and previous members of Caltech High-Speed Integrated Circuits Group, Professor Hui Wu, Professor Donhee Ham, Professor Hossein Hashemi, Dr. Behnam Analui, Dr. Roberto Aparicio, Dr. Chris White, Arun Natarajan, Abbas Komijani, Jim Buckwalter, Ehsan Afshari, Ayding Babakhani, Yu-Jiu Wang, Hua Wang, Sam Mandegaran, Dr. Arjiang Hassibi, Dr. Ichiro Aoki, and Dr. Scott Kee, each of whom gave me a lot of help in many aspects and made my time at Caltech enjoyable.

I also appreciate the help that I received from many other members of the Caltech family, Dr. Dai Lu, Dr. Lawrence Cheung, Feiyu Wang, Guangxi Wang, Sanggeun Jeon, Rumi Chunara, Naveed Neer-Ansari, Kent Potter, Niklas Wadefalk, Ann Shen, Michelle Chen, Jim Endrizzi, and Victoria Lieding.

Special thanks to Professor Hossein Hashemi, my teammate in the project a 24-GHz integrated phased array receiver, which was accomplished via our seamless collaboration. His jokes helped me get through the tough project with joy.

I would also like to acknowledge my teammates in designing the integrated 77-GHz transceiver: Adyin Babakhani, Arun Natarajan, Abbas Komijani and Yu-Jiu Wang.

I am thankful to the committee members in my candidacy exam: Professor David Rutledge, Professor Shuki Bruck, Professor Changhuei Yang, Dr. Sander Weinreb, and Professor Ali Hajimiri, each of whom gave me important and insightful feedback that opened my thoughts to new visual angles.

I wish to thank my Ph.D defense members: Prof. David Rutledge, Prof. Changhuei Yang, Dr. Sander Weinreb, and Dr. Larry D'Addario for their time.

Special thanks to my long-time friend, Hui Wu, who helped me adjust to life in the U.S. during my first year at Caltech.

I am profoundly grateful for my parents and my grandparents, who always love me, believe in me, and support me, no matter where I am.

Finally, I wish to express my appreciation for my wife, Yi Shen, whose unconditional support and caring is an indispensable source of my strength.

Abstract

Microwave integrated systems in silicon provide a low cost, low power and high yield solution for wideband data communication, radar, and many other applications. Phased-array systems are capable of steering the radiation beam by electronic means, emulating the behavior of a directional antenna. This dissertation is dedicated to presenting various techniques to implement microwave integrated phased-array receivers in silicon-based technologies in the context of three design examples.

A 24-GHz 0.18-μm complementary metal oxide semiconductor (CMOS) front-end was demonstrated. The front-end consists of a low noise amplifier (LNA) and a mixer. The LNA utilizes a novel topology common-gate with resistive feedthrough to obtain low-noise performance. The entire front-end achieves a 7.7dB noise figure and a 27.5dB power gain.

A fully integrated 8-element 24-GHz silicon germanium (SiGe) phased array receiver was implemented. The receiver uses two-step downconversion and local oscillator (LO) phase shifting with 4-bit resolution. The signal is combined at the 4.8-GHz intermediate frequency. The 16 phases of 19.2-GHz LO signal are generated with a voltage controlled oscillator (VCO) and symmetrically distributed to the phase selectors at all path. Appropriate phase sequence is applied to the phase distribution transmission lines to minimize mismatch. An integrated frequency synthesizer locks the 19.2-GHz VCO output to a 75-MHz external reference. Measured array patterns show a peak-to-null ratio of more than 20dB and a beam steering range covering all signal incident angles.

An integrated 4-element 77-GHz SiGe wideband phased-array transceiver was implemented. Two-step conversion is used at both the receiver and the transmitter. A differential phase of 52 GHz is generated by the VCO and distributed to all RF paths at

the transmitter and receiver. The phase shifting is performed at the LO ports of the RF mixers using continuous analog phase shifters. The quadrature signal of the second LO frequency is generated by dividing the VCO frequency by a factor of 2 using a cross-coupled injection-locked frequency divider. The signal combining is performed at IF with an active combining amplifier. The receiver achieves a 41dB gain at 80 GHz with 3 GHz of bandwidth. The 52-GHz-to-50MHz frequency divider chain obtains 7% locking range.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The demand for high-speed data communication motivates wireless system to operate at higher frequencies where larger bandwidth is available. According to Shannon's theorem, the channel capacity (C) characterized by the highest data rate of reliable transmission in bits per second (bps), is given by [1]

$$C = B \times \log_2(1 + S/N) \tag{1.1}$$

which indicates two fundamental factors setting the upper bound on the information transmission speed: the channel bandwidth B and the link signal-to-noise ratio (SNR) S/N. While the improvement of S/N is subject to various natural and implementation limitations, increasing B looks like a direct way to enhance achievable data-rate.

Wireless consumer applications utilizing the spectrum below 10 GHz have experienced explosive growth over the last decade, due to both the market demand and the advance of silicon-based technologies such as complementary metal-oxide semiconductor (CMOS) and silicon-germanium (SiGe) bipolar CMOS (BiCMOS) technologies, making low price and compact wireless mobile device a reality. One obstacle to utilize the frequency range above 10 GHz for wide-spread consumer applications is the high cost associate with current solutions using compound semiconductors, such as gallium arsenide (GaAs) and indium phosphide (InP). Compared to compound semiconductor-based technologies, silicon-based technologies provide significant advantages in a higher level of integration on a single chip, thereby reducing cost and power dissipation. Today's most advanced CMOS and SiGe BiCMOS processes offer transistors with transition and maximum oscillation frequency (f_T, f_{max}) comparable to the compound semiconductor transistors [2], making possible the silicon-based integrated system and new applications using the microwave or millimeter-wave spectrum. Meanwhile, many new design

challenges have been introduced at such high frequencies due to the realities of silicon-based technologies such as lossy substrates, low breakdown voltages, low-Q passives, long interconnect parasitics, and high frequency coupling issues [3]. New design techniques need to be devised to deal with those problems.

One promising silicon-based microwave integrated system is a phased array transceiver. Phased arrays constitute a special class of multiple antenna systems that enable beam and null forming in various directions with electronic methods. This electronic steering makes it possible to take advantage of the antenna gain and directionality while eliminating the need for continuous mechanical reorientation of the antenna. Additionally, multiple-antenna systems alleviate the requirements for individual active devices used in the array and make the system more robust to individual component failure. Operating at high frequencies reduces the required element size and inter-element spacing in an antenna array.

This dissertation will present three works investigating the feasibly and performance of microwave and millimeter-wave integrated phased array receivers in silicon-based technologies. Various innovations developed along the way will be revealed in detail together with measurement verifications.

1.1 Organization

After reviewing the receiver fundamentals, the basic operations of phased array will be introduced in Chapter 2. We will then discuss the advantages, architectures, and applications of phased arrays in detail.

Chapter 3 will present our first step in this adventure, a 24-GHz CMOS front-end. A novel low noise amplifier (LNA) topology, common-gate with resistive feedthrough (CGRF), is developed to obtain low-noise performance at an operation frequency comparable to f_T.of the transistor. The advantages of this topology compared to traditional ones will be explained via thorough theoretical analysis. Measurement results demonstrate the first 24-GHz 0.18-μm CMOS front-end with noise figure less than 8dB.

A fully integrated 8-element 24-GHz SiGe phased array receiver will be presented in Chapter 4. We will extensively address many aspects of this system, such as system architecture, a 24-GHz SiGe LNA, signal combining, a 19.2-GHz integrated PLL, multiphase distribution, a 24-GHz test setup, etc. Measurement results demonstrate the spatial selectivity and beam forming capability of the array as well a the high-performance receiver and frequency synthesizer.

In Chapter 5, we will describe a 77-GHz integrated SiGe wideband phased-array transceiver. The design, implementation, and measurement of the receiver signal path and a 52-GHz-to-25-MHz frequency divider chain will be presented where the important innovations include an active signal combining technique and a crossed-coupled quadrature injection-locked frequency divider. Finally, a summary of the highlights and some recommendations for future work will be given in Chapter 6.

Chapter 2

Fundamentals of Single-Path and Multi-Path Receiver

The objective of this chapter is to provide a theoretical foundation for discussions in the following chapters and a review of the existing technologies. The basic concepts in wireless radio reception and single-path receiver are reviewed in Section 2.1. Section 2.2 describes the principles, advantages, and applications of phased array systems, the implementation of which is the main theme of this dissertation.

2.1 Wireless Radio Reception

Electromagnetic (EM) waves have been used to transmit information over air since Guglielmo Marconi invented the world's first radio system in 1897. After more than one hundred years of evolution, the wireless communication systems have become tremendously complex, intelligent, and versatile. However, the essential obstacles for achieving fast and reliable information transmission remain the same: noise and interference.

2.1.1 Noise

In his essay *On Noise*, Arthur Schopenhauer wrote "Noise is a torture to all intellectual people." Certainly circuit designers are among those suffering because we are constantly combating with electronic noise that blurs the signals and causes erroneous or even failed information transmission.

2.1.1.1 Noise Sources in Circuits

Noise in electronic systems arises from the random fluctuation in current flows and takes the form of thermal noise, shot noise, and flicker noise. Thermal noise originates from the random thermal motion of the carrier charges. The most common instance of thermal noise is resistor noise. If one measures the AC voltage across a resistor, a random voltage fluctuation of $v_n(t)$ with zero mean and Gaussian amplitude distribution is observed. This noisy resistor can be represented with a noiseless resistor in series with a noise voltage source v_n as shown in Figure 2.1 (a). Since $v_n(t)$ is a stationary random variable, it is characterized by its power spectrum density (PSD), which is given by

$$\frac{\overline{v_n^2(f)}}{\Delta f} = 4kTR \tag{2.1}$$

where R is the resistance value, T is the absolute temperature in Kelvin, k is Boltzmann's constant equal to $1.38\times10^{-23} Joules/Kelvin$. Equivalently, as shown in Figure 2.1 (b), the noisy resistor can be modeled by a noiseless resistor with a current noise generator i_n, whose PSD is given by

$$\frac{\overline{i_n^2(f)}}{\Delta f} = \frac{4kT}{R} \tag{2.2}$$

The maximum amount of noise power a resistor can pass to the load is delivered when

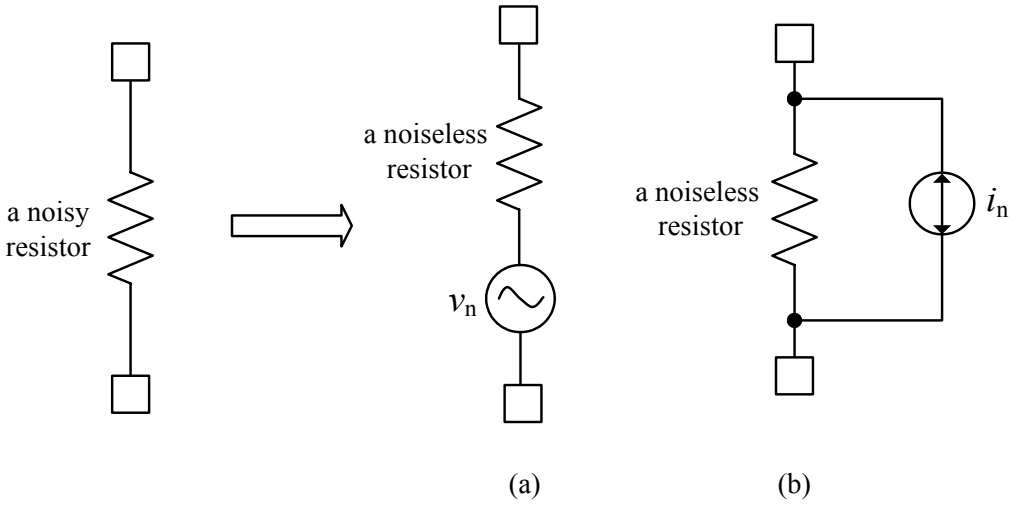


Figure 2.1: Resistor noise model: (a) equivalent voltage (b) equivalent current

the load impedance is matched, which is given by

$$P_{av} = kBT \tag{2.3}$$

where B is the noise bandwidth of the measurement. It is worth noticing that this available power is independent of the resistance value. In a receiving system the receiver input is often matched to the source impedance to get maximum signal power but also obtains the maximum noise power.

Equations (2.1) to (2.3) indicate that the resistor thermal noise has a flat spread spectrum, and hence is called white noise. In fact, resistor thermal noise does have a bandwidth that prevents infinite noise power. The -3dB bandwidth of the resistor thermal noise is on the order of 1 terahertz [4], therefore, in the frequency range of our interests, it can be treated as purely white. Thermal noise was first measured and clearly explained by Johnson [5] and Nyquist [6] in 1928, and therefore is referred to as Johnson noise or Nyquist noise. Thermal noise also exists in the conductance channel of a transistor such as field effect transistor (FET).

The current in a p-n junction barrier of a transistor or a diode is formed by discrete and independent carrier charges. Sampling the instantaneous number of charges crossing the junction with sensitive equipment, one can notice that it has a random variation with a Poisson distribution. This current variation is named shot noise, whose power spectrum is also white with power density [7]:

$$\frac{\overline{i_n^2(\Delta f)}}{\Delta f} = 2qI_{dc} \tag{2.4}$$

where q is the electron charge (1.6×10^{-19} coulomb) and I_{dc} is the direct current flowing through the junction. Unlike thermal noise, shot noise power density is independent of temperature.

Flicker noise is believed to be caused by the defects at the interface of different materials in a semiconductor device such as SiO_2/Si interface in metal-oxide-silicon field effect transistor (MOSFET) and SiGe/Si interface in SiGe heterojunction bipolar transistor (HBT). These defects give rise to extra energy states that can randomly trap and release carrier charges, producing current variations. The power spectral density of the flicker noise is reversely proportional to the device size and frequency. Hence it is also called 1/f noise or pink noise. Due to its frequency dependence, flicker noise is

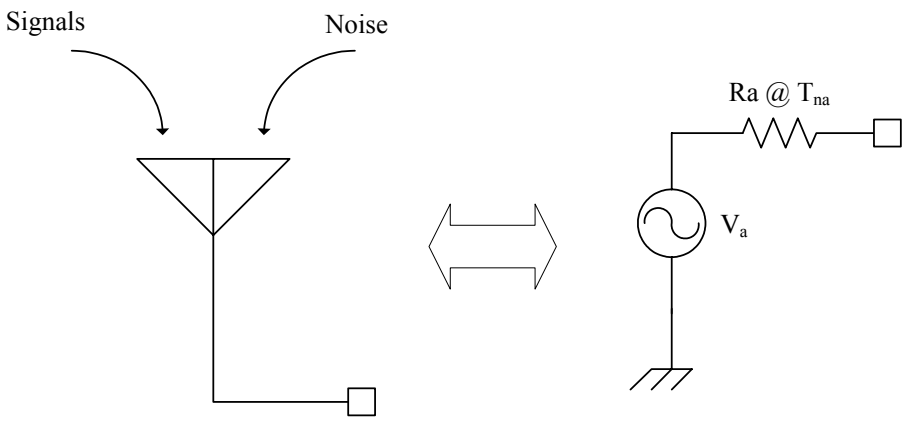


Figure 2.2: Antenna noise model

usually negligible in radio frequency (RF) circuits but can dominant the output noise power in baseband circuits up to a few hundred kilohertz.

2.1.1.2 Antenna Noise

An antenna at the receiver input not only picks up the desired signal carried by electromagnetic waves in air but also many forms of noise including broadband "black body" radiations from all the objects in space. The total noise power collected by the antenna is an integral over its spatially selective receiving pattern and depends on the physical temperature of the black body objects.

The noise performance of the antenna is quantified by using the equivalent model shown in Figure 2.2 [4]. V_a represents the signal collected by the antenna, R_a is a hypothetical resistance equal to the output impedance of the antenna, which is commonly 50Ω in wireless receiving system, and T_{na} is termed the antenna noise temperature, which is the absolute temperature at which R_a generates the same amount of noise power as the total noise power collected by the antenna. The available noise power from the antenna is given by

$$P_{n,av} = kBT_{na} \tag{2.5}$$

2.1.1.3 Correlated and Uncorrelated Noise

The total output noise power of an electronic system is the summed effect of all noise sources. Unlike deterministic signals, which are simply treated with the superposition principle, the calculation of total noise power due to various noise sources is different.

Considering two noise vectors X_1 and X_2, the average power of their summation is given by

$$P_{sum} = \overline{\left|(X_1 + X_2)^2\right|} \tag{2.6}$$

$$= \overline{(X_1 + X_2)(X_1^* + X_2^*)} \tag{2.7}$$

$$= \overline{|X_1|^2} + \overline{|X_2|^2} + \overline{X_1 X_2^*} + \overline{X_1^* X_2} \tag{2.8}$$

$$= P_{X_1} + P_{X_2} + 2\,\mathrm{Re}[c]\sqrt{P_{X_1}.P_{X_2}} \tag{2.9}$$

where c is termed *correlation coefficient* and defined as

$$c = \frac{\overline{X_1 X_2^*}}{\sqrt{\overline{X_1^2}\,\overline{X_2^2}}} \tag{2.10}$$

which is a measure of the similarity of two random processes. If c=0, X_1 and X_2 are uncorrelated and the total noise power is the summation of the individual noise power of each noise source. If c=1, X_1 and X_2 are fully correlated. In other cases, X_1 and X_2 are partially correlated. Usually, the noises originated from independent physical sources are uncorrelated, such as the noise generated in different circuit components. The noises generated by the same physical source can be fully or partially correlated, such as the channel noise and gate noise in a field effect transistor (FET). It is noteworthy that c can be a complex number if the correlation between two noise variables is related to their relative phases.

2.1.1.4 Noise Factor

The noise performance of the receiver is measured with noise factor (F) , defined at a specified frequency as the ratio of the output noise power per unit bandwidth to the output noise power engendered by the source [8]. In most wireless receiving systems, the source impedance is 50Ω and F is defined at the standard noise temperature T_o of 290K. The noise factor expressed in decibel form is called the noise figure (NF).

Assuming the antenna noise temperature is 290K at the input of a single path receiver, it can be derived that F is the ratio of the receiver's signal-to-noise ratio (SNR) at the output to that at the input, which can be expressed in dB format as follows

$$SNR_{out,dB} = SNR_{in,dB} - NF \tag{2.11}$$

(2.11) indicates that *NF* represents the amount of SNR degraded after the signal is processed by the receiver. It is worth noting that in astronomic receivers, the antenna is pointed at much colder objects (or much hotter if it is looking at the sun), so the antenna noise temperature is much lower than 290K. In this case, the input/output SNR ratio is given by

$$\frac{SNR_{in}}{SNR_{out}} = 1 + (F-1)\frac{T_o}{T_{na}} \tag{2.12}$$

This ratio can be much higher than F if $T_{na} \gg T_o$.

2.1.1.5 Noise in Cascade System

Consider one generic electronic system (Figure 2.3) composed of several blocks in cascade, i.e., the output of one stage feeds the input of the next. The i^{th} stage exhibits an available power gain G_i and a noise factor F_i. Assuming that all stages are matched to the system characteristic impedance, the overall noise factor of the system is determined by the gain and noise factor of each stage via

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \ldots + \frac{F_n - 1}{G_1 G_2 G_3 \ldots G_{n-1}} \tag{2.13}$$

Equation (2.13) is known as Friis's formula [9], which indicates that the noise factor of the first stage is most critical to the system noise performance because the noise due to each cascade stage is suppressed by the available power gain preceding it.

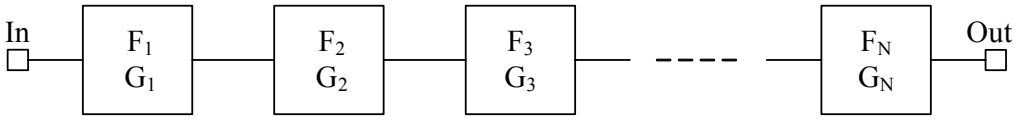


Figure 2.3: Cascade system

2.1.1.6 Noise in Frequency Translation

The receiver usually translates the radio-frequency (RF) signal to lower frequencies in order to facilitate signal processing. When frequency translation is involved, noise characterization is more complicated than in linear systems. To understand this, consider an ideal noiseless receiver using two separate LOs to downconvert RF signals to baseband via an intermediate frequency (IF) stage.

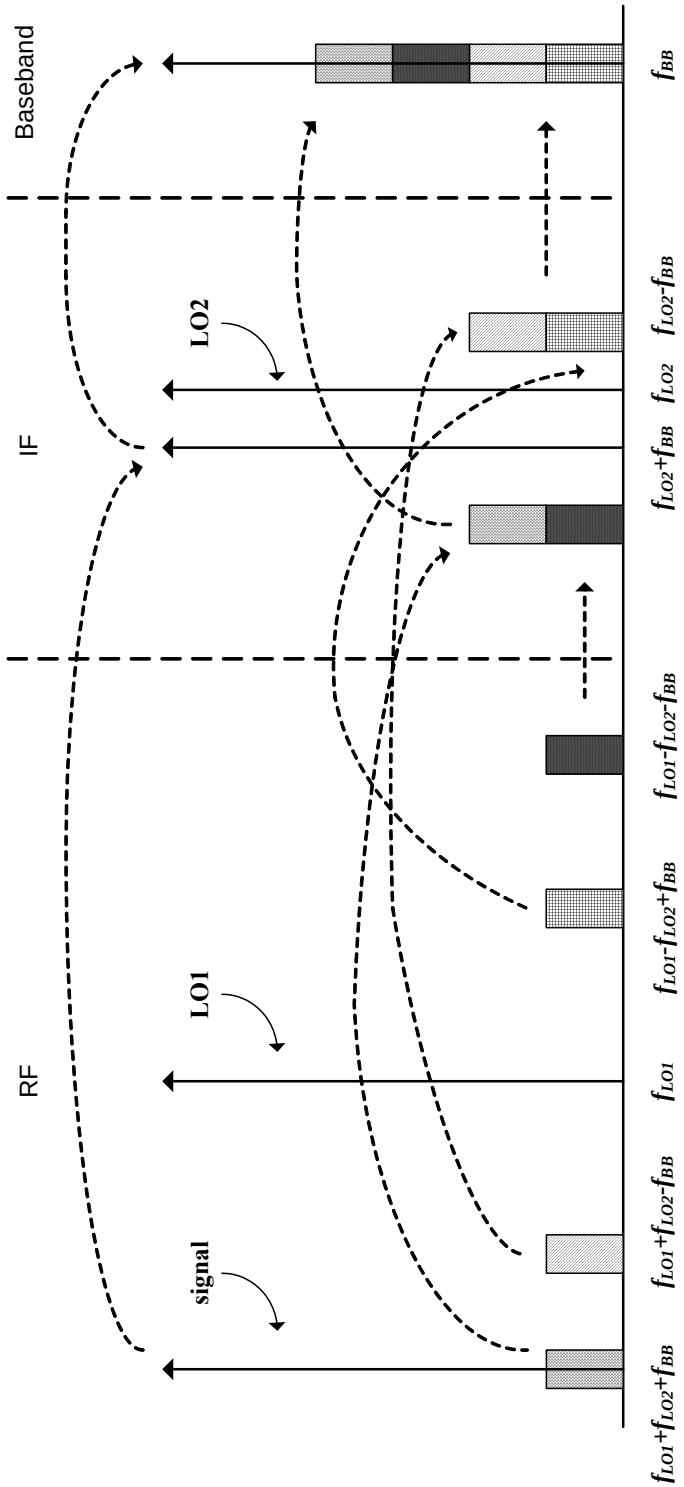


Figure 2.4: Noise translation in a two-step downconversion receiver

Assuming that no frequency selection is performed by the receiver and a unit conversion gain, at each step of the downconversion process, it is not only the noise at the signal band but also at the image band that is folded on top of the downconverted signal. Assuming that a single-tone RF signal resides at $f_{LO1} + f_{LO2} + f_{BB}$, where f_{LO1} and f_{LO2} are the first and second local oscillator (LO) frequency respectively, and f_{BB} is the baseband frequency, as shown in Figure 2.4, the RF noise in four bands, given by $f_{LO1} + f_{LO2} + f_{BB}$, $f_{LO1} + f_{LO2} - f_{BB}$, $f_{LO1} - f_{LO2} + f_{BB}$, and $f_{LO1} - f_{LO2} - f_{BB}$, respectively, reaches f_{BB} via the downconversion process. Therefore, although the receiver itself is noiseless, the SNR_{out} degrades by 6dB compared to SNR_{in}. Moreover, the LO signal is often a strong signal at the frequency translation stage. Even if the LO signal is a pure single tone, it plays a role of square wave. Hence, the odd harmonics of f_{LO} translate the noise at higher frequencies to signal band, further deteriorating the output SNR.

To clarify the confusion about the noise performance of a frequency translation system, two sets of definitions for noise factor are used. For the first definition, the source noise refers to those in the same frequency band of the signal only, which is called the single-side band (SSB) noise figure. For the second definition, the source noise refers to those in all the image bands and for a single frequency translation device, it is called double-side band noise figure (DSB). The SSB NF is always larger than the DSB NF, and the difference depends on the frequency selectivity of the receiver.

2.1.2 Linearity

Any unwanted signal fed into a receiver is called interference. Most interference comes from the signals intended for other users or other applications. The interference power can be orders of magnitude higher than the desired signal power and corrupt the signal if the linearity of the receiver is poor.

Any real receiver is a nonlinear system that responses linearly only if the input signal is sufficiently small. When the input signal increases beyond some extent, the nonlinear behaviors of the receiver become evident and are represented in gain compression and intermodulation products (IP) above noise floor.

One figure of merit for receiver linearity is the gain compression point. Theoretically, the receiver's output power increases linearly with the injected input power regardless of the input power level, as shown in Figure 2.5 [4] by the dashed line. The solid line in Figure 2.5 depicts a typical input/output transfer function of a real receiver. It can be seen that around P_{in}=0, the real I/O curve can be approximated with the straight line. As P_{in} increases, P_{out} gradually deviates from the linear curve and is eventually saturated. The point at which P_{out} is 1dB lower than its linear theoretical value is called the 1-dB compression point (CP1dB). The importance of this point is that it indicates where the receiver starts to leave the linear region and the intermodulation becomes serious problem. The receiver also generates spurs at the harmonics of the signal frequency when the gain goes into compression.

The most important specification of a receiver's linearity is the third-order interception point. Consider two closely spaced interferences at f_1 and f_2 in the vicinity of signal band, where the strongest interference commonly originates. When the interference power is high enough, the receiver generates noticeable spurs at $\pm nf_1 \pm mf_2$ due to intermodulation, where *n* and *m* are integers including zero. Two of these IPs, located at $2f_1 - f_2$ and $2f_2 - f_1$, are particularly threatening to the received signal because they can fall into the signal band and become impossible to eliminate by filtering. In general the power of the $(n+m)^{th}$ IP increases with a slope of (n+m)dB/1dB at the response to the increase of input interference power. Figure 2.6 shows the typical curves of the main tone and the third-order intermodulation power as a function of P_{in}. The third-order interception point is obtained by extrapolating the main-tone output at the slope of 1dB/1dB and the third-order IP curve at 3dB/1dB from the low input power level until they intersect with each other, as shown in Figure 2.6. The x-coordinate of the intersection point is called the input referred third-order interception point (IIP_3), and the y-coordinate is called the output referred third-order interception point (OIP_3).

In a cascaded system as shown in Figure 2.3, the overall IIP_3 of the system is given by

$$\frac{1}{IIP_3} = \frac{1}{IIP_{3,1}} + \frac{G_1}{IIP_{3,2}} + \frac{G_1 G_2}{IIP_{3,3}} + \ldots \frac{G_1 G_2 G_3 \ldots G_{N-1}}{IIP_{3,N}} \tag{2.14}$$

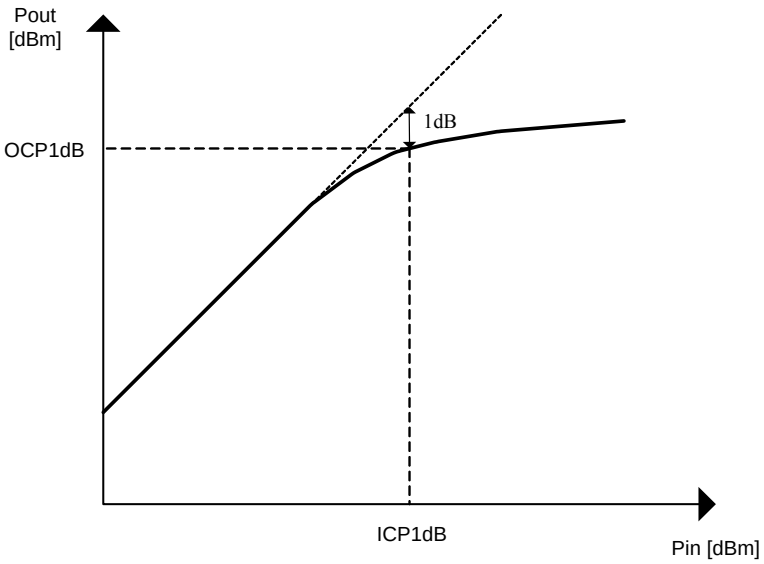


Figure 2.5: Receiver linearity – single-tone test

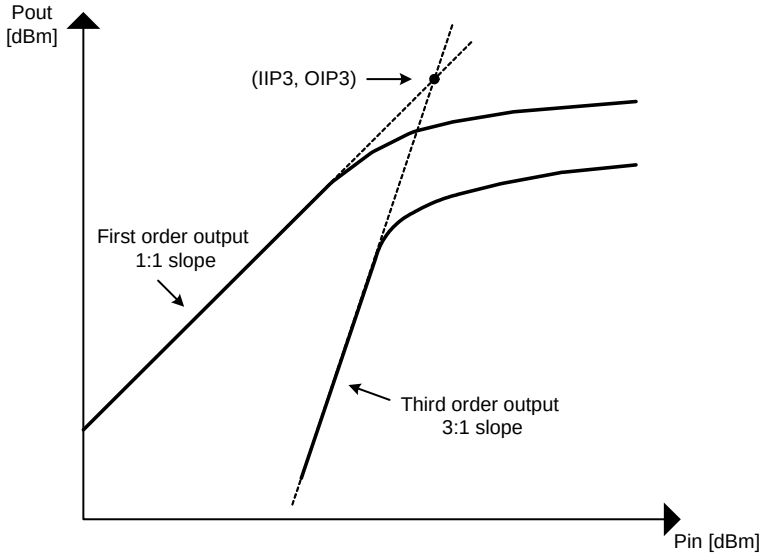


Figure 2.6: Receiver linearity – two-tone test

It can be seen from (2.14) that in a cascade system the linearity requirements on the blocks at the back-end are more stringent because their effects on the overall system are "magnified" by the preceding gain.

2.1.3 Dynamic Range

Dynamic range (DR) is defined as the ratio of the maximum input power level that the circuit can tolerate to the minimum input power level that the circuits can properly detect [10]. DR specifies how well the system can handle signals with various power levels.

The lower bound of the dynamic range is set by the receiver sensitivity, defined as the lowest input signal power a receiver can appropriately process. To calculate the receiver sensitivity, one starts from the maximum bit error rate (BER) the data transmission can tolerate. To achieve this BER, the receiver must provide a minimum SNR_{out} to the subsequent demodulator. Therefore, a minimum SNR_{in} must be achieved at the receiver input, which is given by

$$SNR_{in,\min,dB} = SNR_{out,\min,dB} + NF \tag{2.15}$$

Assuming the receiver input is impedance matched to the antenna, the noise power delivered to the receiver is

$$P_{n,in} = kBT_{na} \tag{2.16}$$

If the antenna noise temperature T_{na} is 290K, the receiver sensitivity can be obtained from (2.15) and (2.16) as

$$P_{\mathrm{s,in,min,dBm}} = NF - 174\mathrm{dBm} + 10\log(B) + SNR_{\mathrm{out,min,dB}} \tag{2.17}$$

where -174dBm comes from $10\log(KT_o)$.

The upper limit of the dynamic range has various definitions that result in different bounds [4], but all are related to the linearity of the receiver. For instance, the most common definition, the spur-free dynamic range (SFDR), defines the maximum allowed input signal power as the one causing the minimum intermodulation product equal to the output noise power. From Figure 2.6, this input power level can be easily solved by using the graphical method, which is given by

$$P_{in,\max,dBm} = \frac{2}{3} IIP_{3,dBm} + \frac{1}{3}(NF - 174 + 10\log B) \tag{2.18}$$

From (2.17) and (2.18), the receiver dynamic range can be found:

$$P_{in,\max,dBm} = \frac{2}{3}(IIP_{3,dBm} - NF + 174 - 10\log B) - SNR_{out,\min,dB} \qquad (2.19)$$

The receiving system often contains devices having adjustable gain at various stages. When the gain setting is changed, the dynamic range of the receiver is shifted up or down so that the overall dynamic range is improved, a process analogous to changing transmission gears in automobiles to provide a wide range of output speed.

2.1.4 Single-Path Receiver Architecture

In over one hundred years of development, many receiver architectures have been proposed and demonstrated for different requirements of various wireless applications. The choice of receiver architecture considers performance, cost, complexity, power, integrity, and flexibility. A detailed discussion and comparison about each single-path architecture is beyond the scope of this dissertation and can be found in [10][11]. Here our discussion is focused on the two most common architectures, superheterodyne receiver and direct-conversion receiver, to show the general criteria and trade-offs at the system level.

Figure 2.7 depicts the block diagram of a generic heterodyne receiver [12]. The EM power picked up by the antenna is first pre-selected by an RF filter to reject the out-of-band interference and partially suppress the image signal. The RF filter must exhibit a low loss since it is directly added to the overall receiver noise figure. The LNA amplifies the signal power and provides the necessary gain for suppressing the noise of the subsequent blocks. An image rejecting filter is inserted between the LNA and the mixer to further attenuate the image interferences. The RF signal and its image are separated by $2f_{\text{if}}$ in frequency domain. If f_{if} is large enough, the RF filter and the tuned LNA may afford sufficient attenuation to the image, eliminating the need for IR filter. However, a high IF increases the quality factor requirement for the channel selection filter. Therefore, the choice of IF is a trade-off between channel selection and image rejection.

The mixer downconverts the RF signal to IF. The LO port of the mixer is usually driven by a frequency synthesizer that generates a tunable LO frequency. The receiver

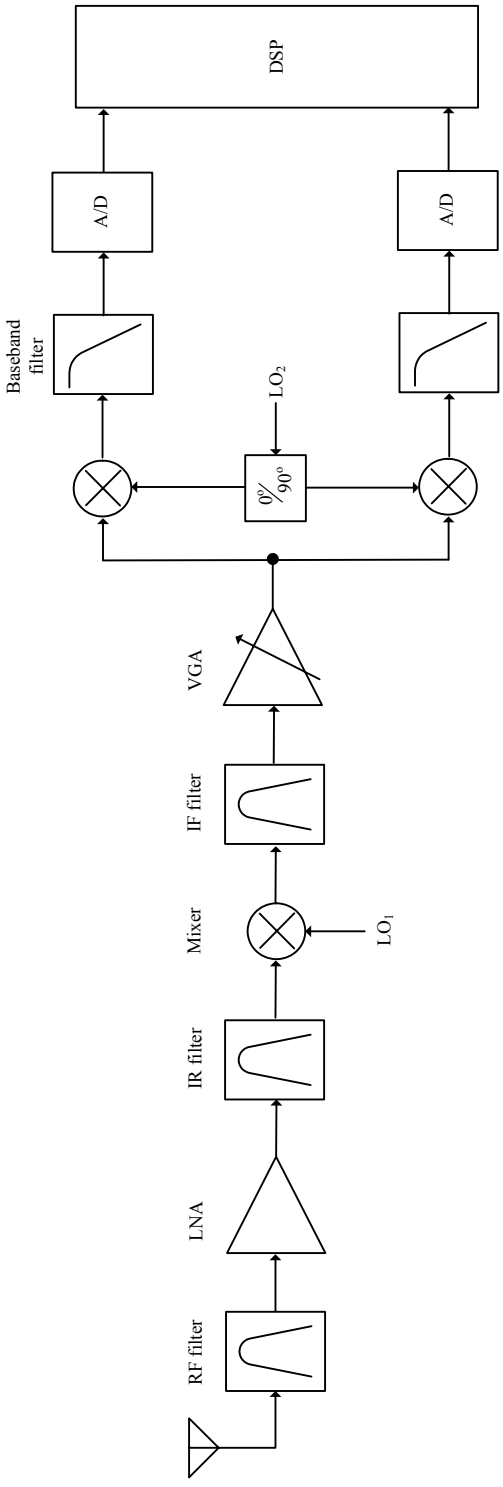


Figure 2.7: A generic superheterodyne receiver

may need to cope with RF signals at different channels in a time-division fashion. A tunable LO translates RF signal at different channels to the same IF frequency so that a fixed channel selection filter can be used. A variable gain amplifier (VGA) prevents the subsequent circuits saturated by a large input. If DR of the input power is very high, a VGA can also be employed at the RF front-end and baseband to achieve more tuning capability.

Quadrature paths are often employed to translate the signal from IF to baseband, i.e., the LO signals driving the in-phase path (I path) mixer and quadrature path (Q path) mixer differ by 90^{o}. This is because in the bandwidth-efficient modulation scheme, the signal spectrum is asymmetric around the carrier frequency. When downconverted to baseband, the information carried in the upper-side band will be irreversibly lost in those of the lower-side band. The solution to this problem is to separate the signal into two elements differing in phase and treat the two elements together as a complex variable, whose frequency spectrum is not necessarily symmetric to dc, so that the asymmetric information can be preserved.

The downconverted signal is further amplified, filtered, and transformed to the digital domain by using an analog-to-digital converter (ADC), from where much more complex and versatile functions can be performed by digital signal processing (DSP).

Because interference rejection and gain control can be performed at various stages of the downconversion path, the superheterodyne receiver achieves superior performance to other architectures with respect to selectivity, sensitivity, and dynamic range. Since being invented by Edwin Howard Armstrong in 1918 [13], the superheterodyne receiver has served the vast majority of the commercial wireless receivers to date.

The main drawback of the superheterodyne receiver is that when implemented in integrated circuits, it requires external IR and IF filters such as the surface acoustic wave (SAW) or ceramic filter, since the quality factor of integrated filters is limited by the substrate and ohmic loss. To drive the off-chip component via package parasitics, the LNA and mixer demand more power. Most importantly, more external components are used, lowering the cost efficiency of the whole system.

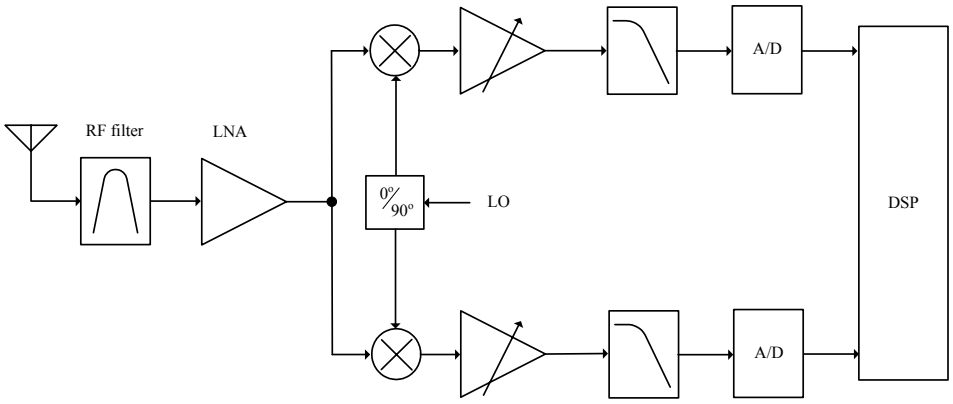


Figure 2.8: A generic homodyne receiver

Two modified superheterodyne architectures have been proposed for integrated implementation: wideband-IF receiver [14] and low-IF receiver [15]. Both architectures choose to separate the signal to I and Q path at the first downconversion instead of the second downconversion, circumventing the image problem. However, the number of IF components is doubled, as well as the power consumption.

Figure 2.8 shows the block diagram of a direct-conversion receiver [16], also known as homodyne, or zero-IF, receiver. The direct-conversion receiver employs only one frequency translation step by setting the LO frequency equal to the RF carrier frequency. This architecture minimizes the number of external components by eliminating the IF stage and using quadrature path instead of IR filter to suppress images, hence it is more amenable to monolithic implementation than the superheterodyne receiver. A reduced number of building blocks and no off-chip components can lead to a low system power consumption. Due to those advantages, direct-conversion topology has been more and more popular in modern integrated communication systems.

However, to design a direct-conversion receiver one needs to carefully address several important problems which are less serious in the heterodyne receiver. One of those problems is LO-to-RF leakage. The LO power is leaked to the RF port through parasitic components, EM coupling or substrate, and mixes with the main LO tone, creating a DC offset, which is troublesome to remove in narrow-band modulation. For wideband modulation such as WCDMA, this DC offset is removed by using a base-

band high-pass filter with a low cut-off frequency, which has little impact on the signal quality. Another problem caused by LO-to-RF leakage is that the leaked LO signal can intermodulate with some strong interferences (for example, in WCDMA receivers, the powerful T_X signal leaks into the receiver [17]) creating in-band distortions which are difficult to eliminate. I/Q mismatch is another serious consideration in a homodyne receiver. The phase and amplitude mismatches in the I and Q paths corrupt the signal by distorting the signal constellation. Although quadrature downconversion is also employed at the last downconversion stage in superheterodyne receiver, the I/Q mismatch is a less severe issue in this case because the low frequency mixer is less sensitive to parasitic mismatches. In addition, the direct-conversion receiver is more vulnerable to second order distortion and flicker noise in the circuits.

In short, there is no receiver architecture globally advantageous to all the others. The optimum choice is determined under certain specifications and applications.

2.1.5 Frequency Synthesizer

A pure, accurate, stable, and tunable LO signal is another key factor for high performance communication system. The LO signal at gigahertz ranges is commonly generated by using a voltage controlled oscillator (VCO). However, the VCO output frequency has poor accuracy and varies with temperature. It has to be locked to a stable frequency source, such as a crystal temperature compensated oscillator, with a working frequency usually below 100MHz and frequency error below a few parts per million. The device that defines the relation of the output frequency to the reference frequency is called frequency synthesizer.

The frequency synthesizer has to achieve a sufficient tuning range and switching time as required by the specified communication system. Most importantly, it has to provide a pure output spectrum that most closely resembles an ideal impulse at the desired frequency, i.e., the spurs at the offset frequency should be low and the skirt around the main tone should be as narrow as possible. The quality of the main tone is quantified by phase noise. The frequency synthesizer output can be mathematically expressed as

$$v(t) = (A + A_n(t))\cos(\omega_{LO}t + \phi_n(t)) \tag{2.20}$$

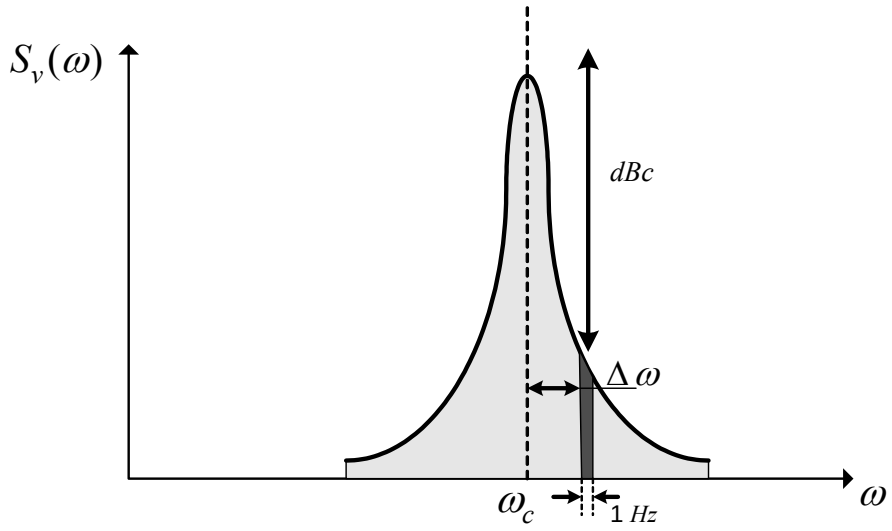


Figure 2.9: LO spectrum and phase noise definition

where A and ω_{LO} are the amplitude and radian frequency of the main tone, respectively, and $A_n(t)$ and $\phi_n(t)$ are called amplitude noise and phase noise, respective, which represent the random disturbance at the output arising from the circuit and reference noise. Since the LO acts as a large signal at the switching device, an accurate zero-crossing time is critical to the receiver performance while it is insensitive to amplitude noise. Therefore, amplitude noise is not a significant concern in frequency synthesizer design. On the other hand, phase noise has to be minimized because it changes the zero-crossing time, downconverts nearby interferences into signal band, and integrates the noise around RF signal. The measure of phase noise is defined in the unit of dBc/Hz as the noise power per unit bandwidth at an offset $\Delta\omega$ with respect to ω_{LO}, normalized to the total carrier power under the spectrum, as shown in Figure 2.9 [18]. The phase noise specification is determined by the minimum signal-to-interference-plus-noise ratio (SINR) and the block level at the offset frequency.

In integrated implementation, the frequency synthesizer is often realized with a phase locked loop (PLL). If the output frequency is an integer multiple of the reference frequency, it is called an integer-N frequency synthesizer [19]. The integer-N frequency synthesizer has a simple configuration, as shown in Figure 2.10. The output frequency is adjusted by programming the frequency division ratio in the

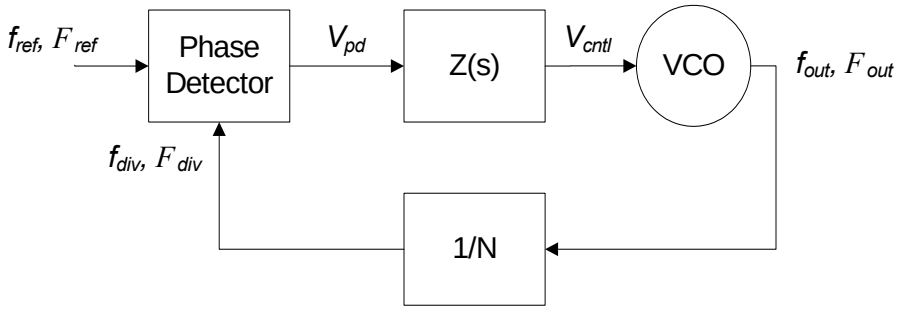


Figure 2.10: A PLL-based frequency synthesizer

feedback path. The main drawback of integer-N topology is that the resolution of the output frequency is limited to the reference source. Therefore, if a fine resolution is required, a high division ratio multiplies the reference phase noise at the output and a narrow bandwidth associated with low reference frequency increases the settling time.

This problem can be alleviated by using fractional-N architecture [20] where the output frequency can be varied at a fraction of the reference frequency. The architecture of the fractional-N synthesizer is similar to integer-N synthesizer, except that the divide-by-n frequency divider is replaced with a dual-modulus divide-by-n or divides-by-(n+1) frequency divider. By varying the percentage of time the frequency divider spends at the two divider values, the averaged VCO output frequency can be changed with a very fine granularity. Compared to the integer-N synthesizer, the fraction-N synthesizer can utilize a higher reference frequency, implying enhanced phase noise suppression and faster setting time. However, it requires a large scale of additional circuits to reduce spurious outputs at the fractional offset.

The design of the state-of-art PLL will be further elaborated upon in Chapter 4.

2.2 Phased Array Systems

2.2.1 Omnidirectional and Directional Communication

Omni-directional communication has been extensively used in various applications due to the insensitivity of orientation and location. Unfortunately, such systems suffer from several shortcomings [21]. As shown in Figure 2.11, the transmitter radiates electromagnetic power in all directions, and only a small fraction reaches the intended

receiver. Thus, for a given receiver sensitivity, a substantially higher power needs to be radiated by an omni-directional transmitter. Not only is a major fraction of this power wasted, but it also adds interference to other users. With dramatically expanding wireless applications and a rapidly enlarged number of users in each application, achievable data-rates in currently deployed wireless communication networks have become more interference-limited than noise-limited [22][23], wherein an increase in transmit power for all users enhances the interference level as well, producing no net benefit for the system capacity. Moreover, modern mobile stations such as cell phones or wireless LAN terminals are often serving in urban or office environments. The transmitted signal can be scattered by various objects such as terrain, walls, trees, vehicles, and people, creating multiple channel paths. The pockets of signal arriving at the receiver via different propagation paths are varied in amplitude and phase and can be added destructively. At certain points the receiver may receive zero signal even though the average transmitted signal power is high. This effect is called “fading” in communication theory and is the primary reason why a cell phone losses a signal during a conversation [22]. Fading is an even more serious problem when moving into high frequencies, because when the receiver is moving it constantly passes the peaks and nulls of the fading effect; the distance between adjacent peak and null is proportional to the carrier wavelength. For instance, at 77GHz the wavelength in air is below half of a centimeter. To obtain reliable data transmission it is imperative for an omnidirectional receiver to be equipped with adaptive gain control and ultra-fast switching time, which is difficult to achieve. In addition, the multi-path propagation also causes inter-symbol interference (ISI), which further impairs the signal quality and limits the maximum achievable data rate.

Limited by the interference, fading, and delay spread, it has become more and more difficult to improve the system capacity per unit bandwidth in an omnidirectional communication scheme. Fortunately, such problems can be mitigated by utilizing space dimension in a directional communication, as shown in Figure 2.12. In a directional communication system, power is only transmitted in the desirable direction(s) and is received from the intended source(s). This is commonly achieved by using directional antennas (e.g., a parabolic dish) that provide antenna gain for certain directions and attenuation in others. Due to the passive nature of the antenna and the conservation of energy, the antenna gain and its directionality go together; a

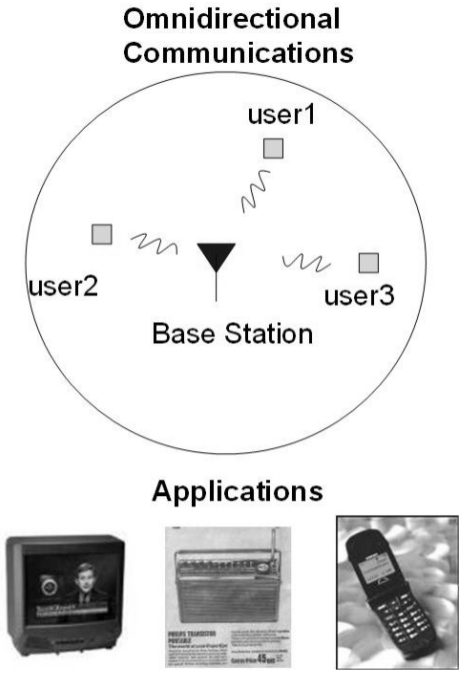


Figure 2.11: Omnidirectional communication scheme

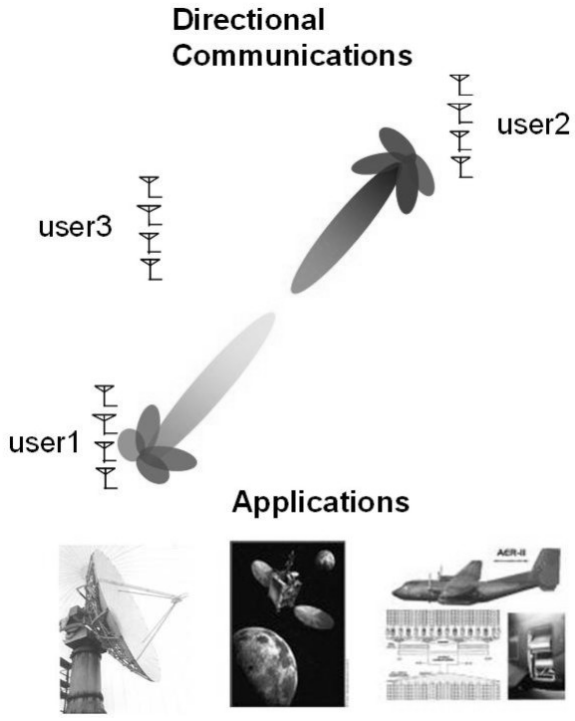


Figure 2.12: Directional communication scheme

higher gain corresponds to a narrower beam width. Directional antennas are used when the relative location and orientation of neither the transmitter nor the receiver change quickly or frequently and are known in advance. For example, this is the case in fixed-point microwave links and satellite receivers. The additional antenna gain at the transmitter and/or receiver can substantially improve SINR and thereby increase the effective channel capacity. However, a single directional antenna is not well suited for portable applications, where its orientation needs to be changed quickly and constantly via mechanical means.

Fortunately, multiple antenna systems can be used to imitate a directional antenna whose bearing can be controlled electronically with no need for mechanical movement [24]-[29]. This electronic steering makes it possible to take advantage of the antenna gain and directionality while eliminating the need for continuous mechanical reorientation of the antenna. Additionally, multiple antenna systems alleviate the requirements for individual active devices used in the array and make the system more robust to individual component failure.

2.2.2 Operation Principles of Phased Array Systems

Multiple antenna systems can be employed on either the receive side (signal-input multiple-output: MIMO), the transmit side (multiple-input single-output: SIMO), or both ends (multiple-input multiple-output: MIMO) [30]. One type of multiple antenna system is to utilize antenna space diversity to create an independent channel path and combine the received signal in an optimum way using space-time processing [23][31][32]. The algorithm and implementation of a MIMO system based on this principle has intrigued a large volume of research and industrial effort in last decade. This technique is easy to implement in the base station of a mobile communication system [33]. However, such a system is not favorable for a mobile unit since it requires the antenna separation on the order of a magnitude higher than the wavelength to obtain a low channel correlation coefficient, and a comparatively higher power due to little hardware shared [34], which conflicts with the compact and low-power requirements of the mobile devices.

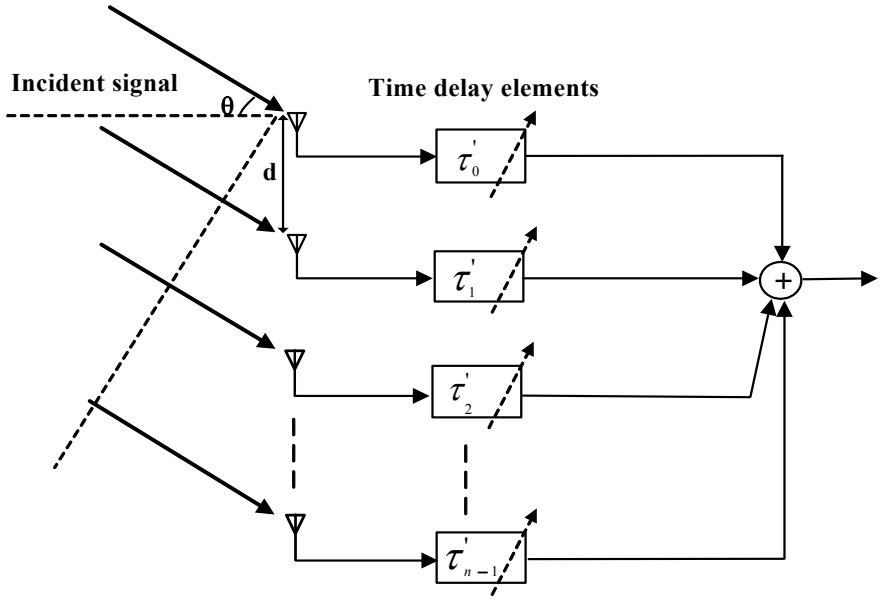


Figure 2.13: A generic phased-array architecture

The phased array is a special type of multiple antenna system. A phased array receiver consists of several signal paths, each connected to a separate antenna. Generally, radiated signal arrives at spatially-separated antenna elements at different times. An ideal phased-array compensates the time delay difference between the elements and combines the signals coherently to enhance the reception from the desired direction(s) while rejecting emissions from other directions. We will use a one-dimensional *n*-element linear array as an example to illustrate the principle as shown in Figure 2.13. We will discuss only the receiver case in this paper, but similar concepts are applicable to the transmitter due to reciprocity.

For a plane wave, the signal arrives at each antenna element with a progressive time delay τ at each antenna. This delay difference between two adjacent elements is related to their distance (*d*) and the signal angle of incidence with respect to the normal, θ, by

$$c\tau = d\sin\theta \tag{2.21}$$

where *c* is the speed of light. In general, the signal arriving at the first antenna element is given by

$$S_0(t) = A(t)\cos[\omega_c t + \varphi(t)] \tag{2.22}$$

where $A(t)$ and $\varphi(t)$ are the amplitude and phase of the signal and ω_c is the carrier frequency. The signal received by the kth element can be expressed as

$$S_k(t) = S_0(t - k\tau) = A(t - k\tau)\cos[\omega_c t - k\omega_c \tau + \varphi(t - k\tau)] \tag{2.23}$$

The equal spacing of the antenna elements is reflected in (2.23) as a progressive phase difference $\omega_c\tau$ and a progressive time delay τ in $A(t)$ and $\varphi(t)$. Adjustable time delay elements ($\tau_n^{'}$) can compensate the signal delay and phase difference simultaneously, as shown in Figure 2.13. The combined signal $S_{sum}(t)$ can be expressed as

$$S_{sum}(t) = \sum_{k=0}^{n-1} S_k(t - \tau_k^{'}) = \sum_{n=0}^{n-1} A(t - k\tau - \tau_k^{'})\cos[\omega_c t - \omega_c \tau_k^{'} - k\omega_c \tau + \varphi(t - k\tau - \tau_k^{'})] \tag{2.24}$$

For $\tau_k^{'} = -k\tau$ the total output power signal is given by:

$$S_{sum}(t) = nA(t)\cos[\omega_c t + \varphi(t)] \tag{2.25}$$

The most straightforward way to obtain this time delay is by using broadband adjustable delay elements in the RF path. However, adjustable time delays at RF are challenging to integrate due to many non-ideal effects such as loss, noise, and nonlinearity. While an ideal delay can compensate the arrival time differences at all frequencies, in narrowband applications it can be approximated via other means. For a narrow band signal, $A(t)$ and $\varphi(t)$ change slowly relative to the carrier frequency, i.e., when $\tau << \tau_{\text{modulate}}$ we have

$$A(t) \approx A(t - k\tau) \tag{2.26}$$

$$\varphi(t) \approx \varphi(t - k\tau) \tag{2.27}$$

Therefore, we only need to compensate for the progressive phase difference $\omega_c\tau$ in (2.23). The time delay element can be replaced by a phase shifter which provides a phase-shift of ϕ_k to the kth path. To add the signal coherently, ϕ_k should be given by

$$\phi_k = k\omega_c \tau \tag{2.28}$$

Unlike the wideband case, phase compensation for the narrowband signal can be made at various locations in the receiving chain, i.e., RF, LO, IF, analog baseband, or digital domain.

2.2.3 Spatial Filtering and Processing

One important advantage of a phased-array is its ability to significantly attenuate the incident interference power from other directions, even by using omnidirectional

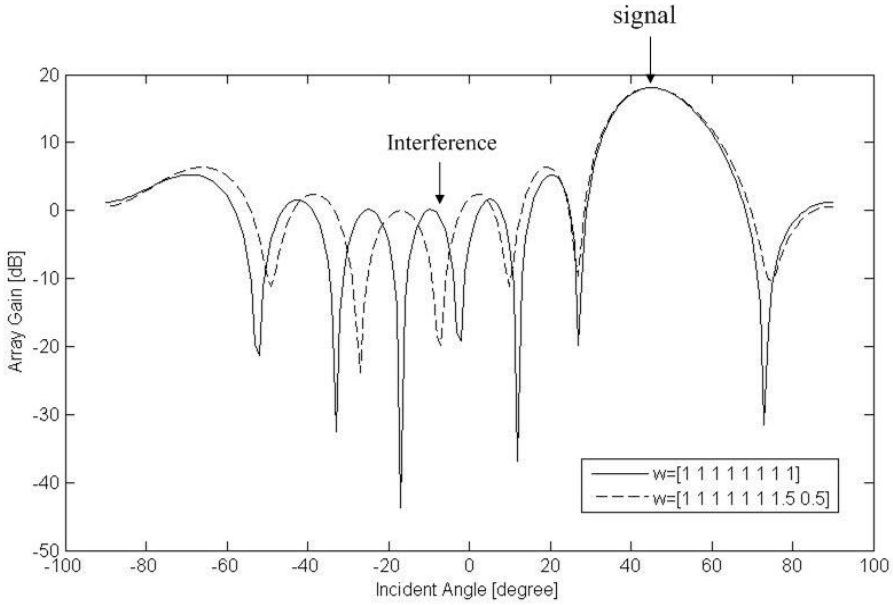


Figure 2.14: Pattern of the array factor of an eight-element array with isotropic antenna elements and $d = \lambda / 2$

antenna elements. The received or radiated pattern of an array is obtained by multiplying the received pattern of a single antenna element by an array factor, assuming identical current distribution in each antenna element. The array factor for a linear 8-element array is plotted against the incident angle in Figure 2.14 intended for a 45^{o} signal angle of incidence. The plot is for a narrowband signal and an antenna spacing of $d = \lambda / 2$, where λ is the wavelength. It can be seen that the signals incident from other angles are significantly suppressed. This function is often referred to as space filtering. Additionally, in phased-array systems the signal power in each path can also be weighted to adjust the null positions or to obtain a lower side-lobe level [27][28]. For example, the dashed line in Figure 2.14 shows the array factor when the signal magnitude of eight receiving paths are weighted by the vector w=[1 1 1 1 1 1 1.5 0.5]. The received signal power from the desired direction remains the same. If a dominant interference comes from the direction signified by the arrow, it is attenuated by more than 20dB by applying different weights. This process is often referred to as space processing.

It is also worth noting that the array factor is a function of array geometry. The antenna elements of an array can be arranged in different spatial forms such as line, two-dimentional rectangle, co-centric circles, or conformal to the surface of a three-

dimensional object, obtaining various spatial filtering properties [27]. For a simple example, if the antenna separation is larger than half of the wavelength in a linear array, a narrower beamwidth and a finer beam steering resolution can be achieved in space corresponding to particular phase compensating resolution in the receiver. However, the array pattern will exhibit multiple-beams.

2.2.4 SNR Improvement

For a given modulation scheme, a maximum acceptable BER translates to a minimum SNR at the baseband output of the receiver (input of the demodulator). For a given receiver sensitivity, the output SNR sets an upper limit on the noise figure of the receiver. In the case of a single path receiver

$$SNR_{out,\mathrm{dB}}=SNR_{in,\mathrm{dB}}-NF \tag{2.29}$$

which cannot be directly applied to multi-port systems such as phased-arrays. Consider the *n*-path phased-array system, shown in Figure 2.15. Since the input signals are added coherently,

$$S_{out}= n^2 G_1 G_2 S_{in} \tag{2.30}$$

The antenna's noise contribution is primarily determined by the temperature of the object(s) it is pointed at. When antenna noise sources are uncorrelated, such as in an indoor environment, the output total noise power is given by

$$N_{out}=n(N_{in}+N_1)G_1G_2+N_2G_2 \tag{2.31}$$

Thus, compared to the output SNR of a single-path receiver, the output SNR of the array is improved by a factor between n and n^2 depending on the noise and gain contribution of different stages. The array noise factor can be expressed as

$$F = \frac{n(N_{in}+N_1)G_1G_2+N_2G_2}{nN_{in}G_1G_2} \tag{2.32}$$

$$= n\frac{SNR_{in}}{SNR_{out}} \tag{2.33}$$

which shows the SNR at the phased-array output can be even smaller than SNR at the input if n>F. For a given NF, an n-array receiver improves the sensitivity by 10log(n) in dB compared to a single-path receiver. For instance, an 8-path phased-array can improves receiver sensitivity by 9dB.

The noise factor of a phased array is affected by array weighting. It can be derived that

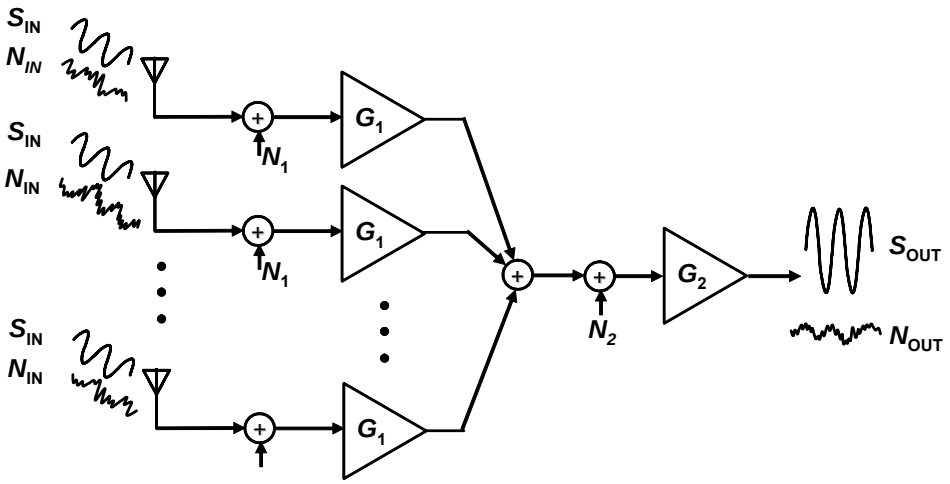


Figure 2.15: SNR improvement by the phased array

$$F = \frac{\left(\sum w_k\right)^2}{\sum w_k^2} \frac{SNR_{in}}{SNR_{out}} \tag{2.34}$$

where w_k is the weight of the k^{th} path.

We next investigate how the LO phase noise affects the array performance. The LO output signal is given by (2.20). In a phased array implementation, the LO phase noise can be decomposed into

$$\phi_{n,k}(t) = \phi_{n,u,k}(t) + \phi_{n,c,k}(t) \tag{2.35}$$

where k is the path index; $\phi_{n,u,k}(t)$ originates from the common LO components of all paths such as a core PLL and thus is fully correlated among the paths; $\phi_{n,c,k}(t)$ arises from the individual LO components of each path, such as the local LO limiters or phase shifters, and thus is uncorrelated to one another. Assuming the weighting is uniform, the combined signal can be approximated with

$$S_{sum}(t) \approx nA(t)\cos[\omega_c t + \phi(t) + \frac{\sum_{i=1}^{n}\phi_{n,u,k}(t)}{n} + \frac{\sum_{i=1}^{n}\phi_{n,c,k}(t)}{n}] \tag{2.36}$$

It can be seen from (2.36) that the array does not enhance phase noise compared to single-path. Since $\phi_{n,u,k}(t)$ is uncorrelated to one another, its average power at the array output is suppressed by a factor of n. On the contrary, the average power of $\phi_{n,c,k}(t)$ represents itself at the array output with no attenuation.

2.2.5 Phased Array Architectures

As a single-path receiver, the phased array receiver can be realized using various down conversion schemes such as superheterodyne, direct conversion, wide-band IF, low-IF, etc., involving similar trade-offs in signal-path design. Phased array receiver can also be classified by where the delay compensation is performed, namely, passive RF, active RF, IF, analog baseband, digital domain, or LO path.

Figure 2.16 shows the passive RF phase shifting architecture. Passive phase shifters or time delay elements directly follow the antenna elements. Their outputs, the phase or delay compensated signals, are summed via a combining network fed into a single LNA input. The true time domain compensation, resulting in broadband frequency response, can only be achieved by using time delay elements before the first frequency translation. Such time delay can be realized using transmission lines whose effective length can be adjusted electronically [28]. A single-path receiver can be readily employed in this architecture. Since the signal combining process enhances the signal level and tremendously attenuates the interference, the noise and linearity requirements of the LNA are greatly relaxed, allowing them to trade off with other system performance. The main drawback of this architecture is that the loss of the phase shifters and combining network directly degrades the receiver sensitivity, and so they are limited to waveguide type in practical implementation, which are bulky, heavy, and expensive, prohibiting wide-spread usage. Another limitation of passive phase shifting is the lack of amplitude control.

Figure 2.17 illustrates the active RF phase shifting architecture. In this architecture, the phase shifters or delay elements are introduced after the LNAs. The multiple LNAs increase the system power consumption. However, thanks to the LNA gain, the phase shifters do not need to be optimized for low loss. The amplitude control can also be realized using RF VGA. The space processing at RF relaxes the DR requirements of the mixer and the subsequent blocks. The design challenge is to create compact, linear, wideband and relatively low-loss RF phase shifters, which are difficult to realize in integrated implementation.

The IF phase shifting architecture is shown in Figure 2.18. After the LO signals with identical phases mix with the RF signals, only carrier phases can be compensated

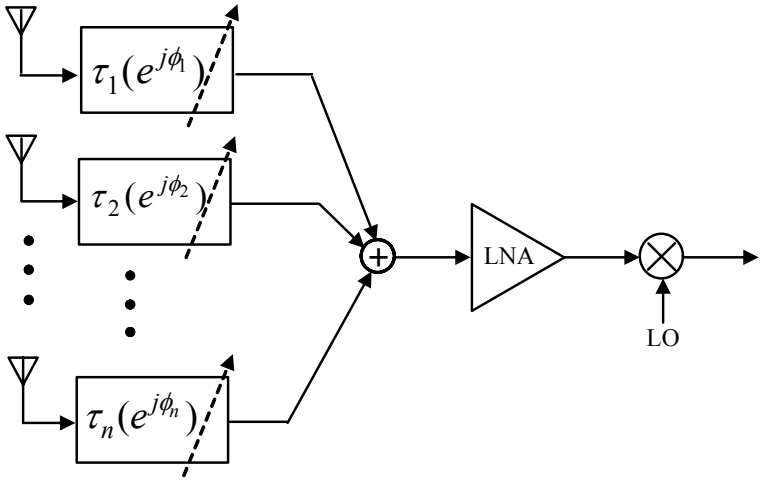


Figure 2.16: Passive RF phase shifting architecture

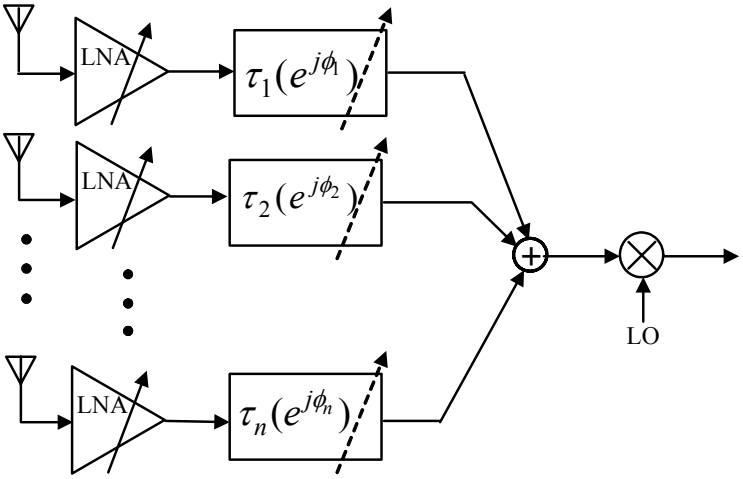


Figure 2.17: Active RF phase shifting architecture

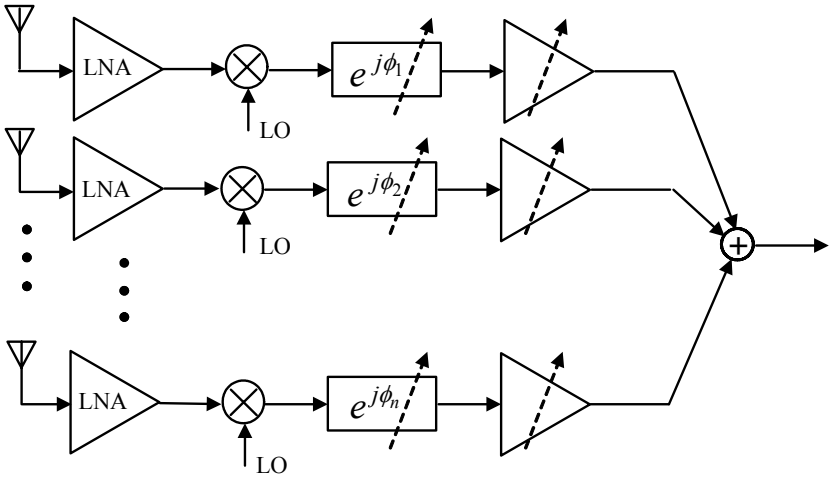


Figure 2.18: IF or baseband phase shifting architecture

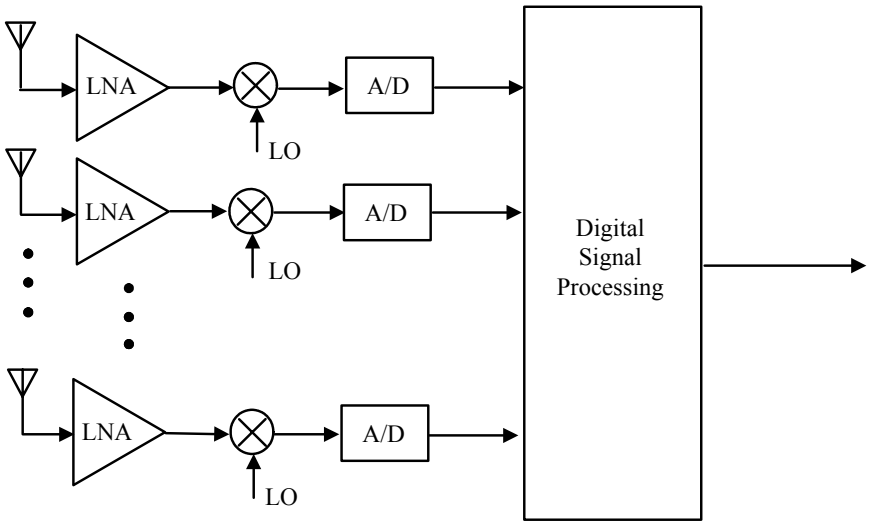


Figure 2.19 Digital phase shifting architecture

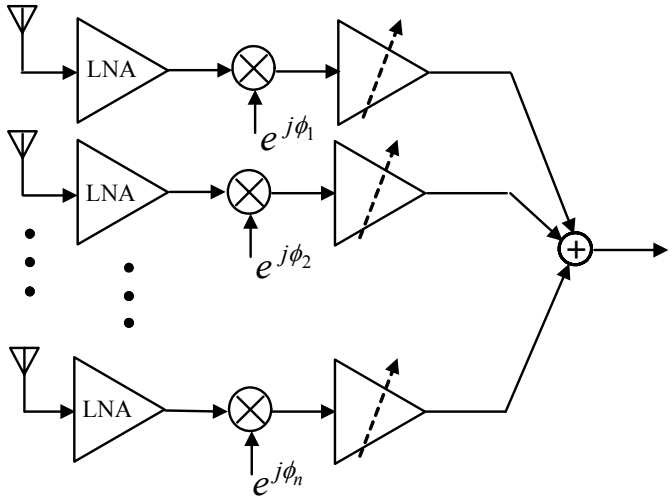


Figure 2.20: LO phase shifting architecture

correctly. The time delay compensation at IF will give rise to an unbalanced phase in each path owing to the mixed LO signals. Hence the IF phase shifting is only suited for narrow band modulation. When the phase shifting stage moves towards the back-end of the receiver, fewer unshared components increase the overall system noise and power consumption. Moreover, the unshared blocks before the signal combining experience the same SINR so that they need to provide the same DR as those in single-path receiver, implying an additional increase in power. Compared to the RF phase shifter, the IF phase shifter exhibits lower loss and lower power consumption due to the lower operation frequency. However, the dimensions of the passive devices, i.e., the inductors and capacitors, used in phase shifters are generally reversely proportional to operation frequency. Therefore, the IF phase shifting consumes more valuable silicon area than the RF phase shifting. The same tradeoff applies to analog baseband phase shifting.

Taking advantage of the large amount of transistors provided by CMOS technology, the amplitude and phase control can be performed in digital domain as shown in Figure 2.19, referred to as digital phase shifting architecture. Using a digital signal processor (DSP), the space processing can be performed with various algorithms, suggesting the most versatile topology. However, each block in the single-path receiver has to be multiplied in this array implementation, including the power hungry ADC, which might make it exceed the power budget of a portable device and

the most noisy implementation among all the architectures. Another serious design challenge is the high-speed high-throughput data I/O of the DSP, which is currently a bottleneck to the achievable bandwidth of this configuration.

An alternative approach for an integrated implementation of such a system is to perform the phase shifting in the LO path in this so called LO phase shifting architecture as shown in Figure 2.20. The possible amplitude control can be realized by employing the variable amplifiers at the RF or IF stages. If different downconversion mixers are driven with LO signals of different phases, we can achieve the phase shifting at the LO and approximate the delay elements over a limited bandwidth. This architecture is advantageous in that amplitude noise and mismatches at the LO path do not deteriorate the receiver sensitivity and spatial selectivity directly. Moreover, this architecture is particularly attractive for silicon-based integrated systems due to the possibility of accurate multiple phase generation and distribution [35].

2.2.6 Applications

The largest commercial potent of phased array lies in communications. For example, phased arrays are typically used in AM broadcast stations to favor signal coverage in the city of license while minimizing interference to other areas [36]. Phased-array based satellite TV systems are also available in the current commercial market. Compared to traditional parabolic dish systems, the phased-array implementation is more robust to environmental changes [37] such as wind, rain or snow, and easier to be mounted on roofs. Moreover, the adaptive beamforming enables satellite program to be delivered to mobile objects such as planes and vehicles [38].

For consumer mobile data transmission, voice service, or multimedia service, the additional gain and spatial filtering properties of phased array can be utilized to: 1) increase system capacity; 2) extend coverage range; 3) mitigate impairments caused by multi-path effects; and 4) provide user location information [39].

The benefits of phased array for enhancing signal qualities in wireless communications have been proved by field experiments. For instance, in [40], a 4-element phased array receiver with adaptive beamforming is tested with over 250

experiments in rural, suburban, and urban channels with two mutually interfering transmitters. The measurement results demonstrate 30 to 50dB SINR improvements in rural, line-of-sight scenarios and over 20dB SINR improvements in urban and suburban outdoor, non line-of-sight, peer-to-peer scenarios. In indoor environments, phased-array receivers operating in current wireless communication bands such as 2.4GHz and 5GHz are less attractive because the desired signal at those frequencies are more evenly distributed in space dimensions due to multiple scattering of the walls and ceilings, where equalization technique might be a more appropriate choice to maximize received signal power. However, an investigation on 60GHz indoor wireless channels using ray-tracing algorithm [41] shows microwave wireless channels exhibit different properties compared to low-GHz channels due to the significant attenuation to the ultra-high frequency signal by the building materials and air. Simulation for a typical office environment shows that the received 60-GHz signal power is more concentrated in one direction. Using a directional transmitter and receiver with 30^{o} beam width, a delay spread of less than 10ns and a k-factor (ratio of the power in dominant signal component to the sum of that in the random multi-path component) of more than 7dB are achieved at 90% of the locations, compared to delay spread greater than 23ns and a k-factor of less than 5dB in 50% of the locations when isotropic transmitter and receiver are used. Considering that the array gain compensates the added path loss introduced at these high frequencies and that the high operating frequencies reduces the dimension of the antenna array, making it possible to be used in hand-held terminals, we can predict that phased array is a critical technique to realize microwave consumer wireless communications, one of the contemporary research frontiers.

The phased array concept has been widely used in radar systems which emit continuous-wave or pulse signals at certain directions and obtain the information of distant objects by analyses of the reflected waves. Radar is a fundamental apparatus for surveillance, object tracking, remote sensing, projectile guidance, and synthetic imaging. The electronic scanning of the beam of phased array radar is orders of magnitude faster than the traditional radar rotated by mechanic motors.

Vehicular radar has been developed for decades and is being installed on high-end luxury sedans at the moment [42]. As shown in Figure 2.21, radar sensors mounted

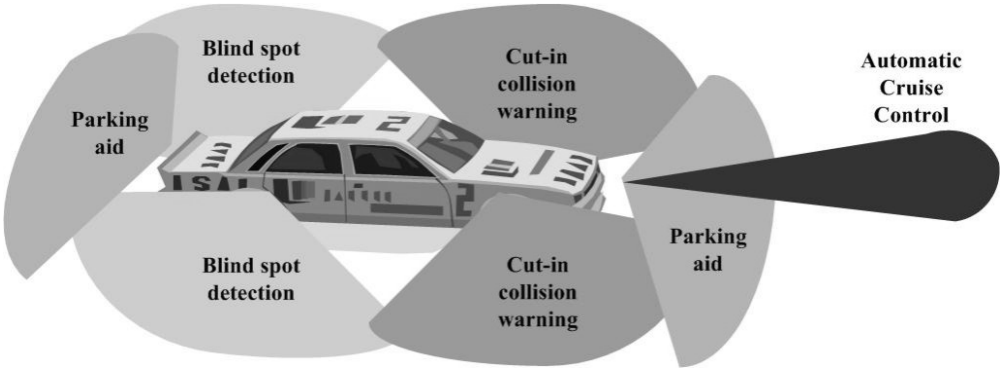


Figure 2.21: Automotive radar sensors provides multiple driving-aid functions

around the car can provide multiple driving-aid functions such as automatic cruise control (ACC), parking aid, blind spot detection, and side collision warning [43]. High resolution radar systems with advanced image processing can further enable objects classification, roadside detection, and lane predition [44]. Ultimately, autonomous driving is possible by combining short-range radar, global positioning techniques, and wireless communications. Phased arrays can provide the narrow beam and low sidelobe requirements of the automotive radar [43] together with compact or even conformal antennas which are 'invisible' to consumers having aesthetic judgments. Developing phased arrays operating at 24GHz or 77GHz frequency bands allocated by Federal Communications Commission (FCC) for vehicular radar applications is an intense research topic at the moment [42]-[47].

Radio astronomy is another important application area of phased array. The next generation radio telescope demands sensitivity one or two orders of magnitude lower than current telescopes in use, requiring a total collecting aperture of approximately one square kilometer [48]. Instead of using an ultra-giant single parabolic antenna, such a system can be implemented with an array of more than one-hundred million small antenna elements, providing additional benefits such as adaptive radio-interferences rejection.

Biomedication is an emerging yet promising application of phased array. In [49], a microwave imaging method is proposed using phased array to detect early-stage breast cancer. The antenna array placed at the breast surface emits the wideband

impulses sequentially by each antenna. The beaming-forming is employed at the receiver to focus the backscattered signal from the malignant tumor and compensate for the frequency-dependent propagation effect. The signal reflection is primarily due to the dielectric discontinuity at the edge of the malignant tumors and the normal breast tissue. The relevant contrast is an order of magnitude higher for microwave than for X-ray or ultrasound [50], suggesting a much higher detection probability. Microwave imaging is also a much cheaper solution than other current alternatives such as magnetic resonance imaging (MRI) and is less harmful to the patients than X-ray. In [51], a hyperthermia system is presented using a conformal phased array to treat tumors in human limbs. The array consists of 8 dipole radiators mounted on a cylindrical surface, focusing EM waves to the tumor inside the limb to heat it to a higher temperature than surrounding tissues. The thermal pattern can be varied by adjusting the amplitude and phase of each antenna element. Tumors heated repeatedly to higher temperature sometimes exhibits regression and necrosis.

Phased array electronic systems can also be applied to fields where the information carrier is not EM waves, such as ultrasound imaging in biomedication [52] or sonar system for underwater applications [39].

In summary, phased array provides us with various ways to explore the space dimension and take advantage of space diversity conveniently using electronic methods. Its potential application range is only limited by the imagination of the engineers.

2.2.7 Integrated Phased Array System in Silicon

Phased array techniques have existed for decades, with tremendous research and industrial efforts resulting in a large number of implementations. However, the practical application of phased array is still limited by its high cost. Although the development of monolithic microwave integrated circuits (MMIC) and III-V compound transistors has lowered the cost of active arrays by orders of magnitude compared to the traditional passive arrays [28], its price is still prohibitively expensive for vast-volume consumer products. Taking advantage of the advents in silicon-based integration providing millions of transistors with continuously increasing density and speed, a fully-integrated high performance phased array system

in silicon can be a key enabler for wide-spread consumer applications such as microwave wideband wireless communications and automotive radars. Integration of a complete phased array system in silicon results in substantial improvements in cost, size, and reliability. At the same time, it provides numerous opportunities to perform on-chip signal processing and conditioning without having to go off-chip, leading to additional savings in cost and power.

2.3 Chapter Summary

In this chapter, two basic problems which must be addressed in a receiver design, noise and interference, were discussed in the context of receiver specifications. Single-path receiver architectures were described and compared at both signal path and LO path (frequency synthesizer). A special type of multi-path receiver, known as a phased array system, was introduced with mathematical derivations for its signal combining process and SNR improvements. Tradeoffs in diverse phased array architectures were discussed. The benefits and applications of phased array in various fields such as communications, radar, radio astronomy, biomedication, and sonar, were extensively reviewed. Finally, a vignette into the possibilities of a silicon-based fully integrated phased array was offered.

Our exploration of microwave silicon-based integrated phased array receivers will be presented in the subsequent chapters.

Chapter 3

A 24-GHz CMOS Front-End

In this chapter, a 24-GHz CMOS front-end employing novel LNA topology is reported. The project motivations and goals are briefly introduced in Section 3.1. The basic theories of a linear noisy twoport and the high-frequency model of MOSFET are reviewed in Section 3.2, based on which a novel LNA topology common-gate with resistive feedthrough is introduced and analyzed. The LNA performance in terms of noise, gain, input matching, power dissipation, and stability are addressed in detail. Section 3.3 describes the circuit design of the front-end, followed by measurement results in Section 3.4. Section 3.5 summarizes the chapter with conclusion.

3.1 Introductions

3.1.1 Motivations

Most of today's wireless schemes for consumer applications are centered around 2.4 and 5GHz frequency ranges. However, the growing demand for higher data rates motivates integrated circuits to move toward higher frequencies where significantly larger bandwidth is available. Furthermore, wireless transmissions using higher carrier frequencies reduce the size of common resonate-based antenna and their spacing in a multiple antenna scheme, making phased-array antenna systems practical for portable applications.

The industrial, scientific, and medical (ISM) band at 24GHz is a good candidate for broadband wireless communications. For indoor environments, the walls and ceilings provide more isolation to 24GHz signals than to low-GHz signals [53] , increasing the possibility of frequency reuse, enhancing the information security, and reducing the interference to other users. Furthermore, an FCC ruling released in 2002 opened the 22GHz ~ 29GHz frequency band for ultra-wideband (UWB) vehicular radar applications [53] Consequently, research on 24-GHz range wireless technologies

has accelerated, demonstrating various building blocks and single-path receivers at this frequency [54]-[59].

The rapid evolution of the wireless communication world has resulted in a tremendous amount of activities involving building high-performance RF circuits in various technologies. Among many contenders, CMOS technology is particularly attractive for its low cost and high level of integration, offering digital circuits composed of a huge number of transistors which can be used to perform various digital signal processing options. Therefore, CMOS technology is a promising candidate for building a fully-integrated phased array system. That the answer is yes or no depends on whether the high-performance CMOS front-end at very high frequencies can be implemented. In the last decade CMOS has been demonstrated to be a viable medium for implementing RF circuits for applications in the low-GHz range [14][60]-[62]. However, a good performance or even the possibility of CMOS tranceivers for applications over 20GHz has not been seriously investigated prior to this work. The above consideration motivates this design effort to develop a CMOS receiver front-end (LNA+mixer) operating at frequencies around the 24GHz range as the first step towards a fully integrated multi-channel receiver for a phased array system.

3.1.2 System Block Diagram

A simplified block diagram of one receiving channel is shown in Figure 3.1. The low noise amplifier (LNA) and the first downconversion mixer are present in vast majority of the wireless receiver systems. The LNA boosts the power level of the radio-frequency (RF) signal picked up by the antenna and the succeeding mixer translates the RF signal to lower frequencies. Depending on the frequency downconversion schemes, the intermediate frequency (IF) stage is optional. In homodyne topology, the signal is translate from RF directly to baseband. On the other hand, in heterodyne receivers the RF signal is shifted to baseband through multiple intermediate stages. This work is comprised of the LNA and the first mixer, which are essential blocks in both heterodyne and homodyne architectures and most difficult to implement in silicon because they operate at the highest frequencies of the receiver chain. In this work, intermediate frequency is chosen to be 5GHz for large image rejection without external filters.

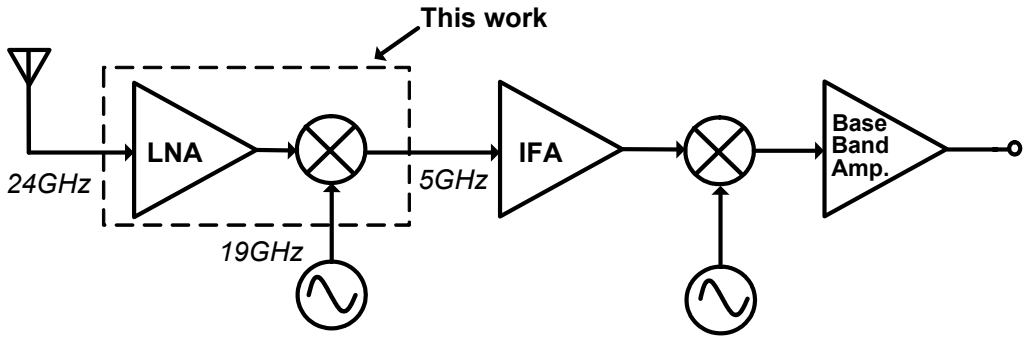


Figure 3.1: Reciver block diagram

3.2 Common-Gate with Resistive Feedthrough LNA

The input stage of the LNA sets the limit on the sensitivity of the receiver. Therefore, low noise is one of the most important design goals. Unfortunately the noise figure increases with frequency, primarily due to lower gain at high frequencies. The input stage also needs to achieve a sufficient gain to suppress the noise of the following stages and good linearity to handle out-of-band interference while providing well-defined input impedance, which is normally 50Ω, as required by the preceding block such as antenna, filter or duplexer. In this section a novel LNA topology common-gate with resistive feedthrough is introduced, which can achieve a lower noise figure at very high frequencies compared to the trbaditional LNA topologies.

3.2.1 Basics of Twoport Noise Analysis

The circuit unit performing signal processing such as amplification and filtering can usually be represented by a linear noisy two port as shown in Figure 3.2 (a). The noise generated inside a twoport is characterized at any specific frequency by the noise factor, *F*, or noise figure, *NF*.

Based on Thevenin's theorem, a twoport containing internal noise sources can be separated into a noise-free twoport with two external noise generators. One example of such equivalent circuits is shown in Figure 3.2 (b), where the internal noise sources are represented by a voltage noise source v_n adding in series with the input voltage and a current noise source i_n flowing in parallel with the input current. The

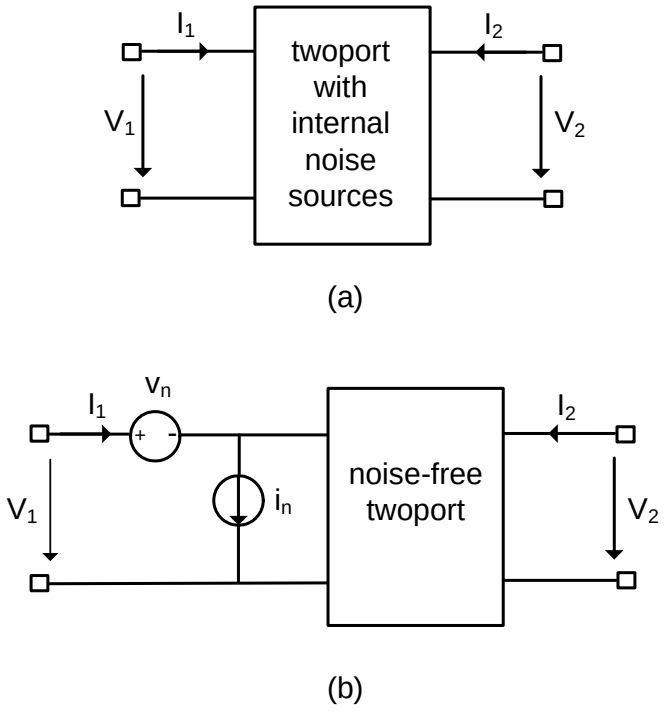


Figure 3.2: (a) A linear noisy twoport (b) An equivalent twoport

correlation between v_n and i_n is characterized by correlation admittance Y_r, which is given by

$$Y_r = G_r + jB_r \tag{3.1}$$

$$= \frac{\overline{i_n v_n^*}}{\overline{v_n^2}} \tag{3.2}$$

where G_r and B_r are the real and imaginary parts of Y_r, respectively. The four parameters, v_n, i_n, G_r, and B_r completely describe the noise performance of a twoport. The noise factor F for all input terminations can be directly derived from these four parameters and the signal source admittance Y_s. However, it is more convenient to express F with another set of four parameters, $F_{\min}$, R_n, G_o, and B_o, via [64] [65]

$$F = F_{\min} + \frac{R_n}{G_s}\left[(G_o - G_s)^2 + (B_o - B_s)^2\right] \tag{3.3}$$

where G_s and B_s are real and imaginary parts of Y_s, respectively. $F_{\min}$ is the lowest achievable noise factor of the twoport by adjusting Y_s. This minimum noise factor is

obtained when Y_s is set to the optimum source admittance $Y_o = G_o + jB_o$. The equivalent noise resistance R_n characterizes the sensitivity of F to the distance between Y_s and Y_o.

The four noise parameters in (3.3) can be determined by v_n, i_n, G_r, and B_r by the following transformation rules [64] [65]

$$R_n = \frac{\overline{v_n^2}}{4kT\Delta f} \tag{3.4}$$

$$G_o = \left[\frac{G_u + R_n G_r^2}{R_n} \right]^{1/2} \tag{3.5}$$

where

$$G_u = \frac{\overline{|i_n - Y_r v_n|^2}}{4kT\Delta f} \tag{3.6}$$

$$B_o = -B_r \tag{3.7}$$

$$F_{\min} = 1 + 2R_n(G_r + G_o) \tag{3.8}$$

Sometimes it is also convenient to express F in terms of optimum noise impedance $Z_o = R_o + jX_o$ and signal source impedance $Z_s = R_s + jX_s$ as

$$F = F_{\min} + \frac{R_n}{R_s}\left[\frac{(R_o - R_s)^2 + (X_o - X_s)^2}{R_o^2 + X_o^2} \right] \tag{3.9}$$

We are going to use (3.4) to (3.8) in the next subsection to characterize the noise properties of an intrinsic CMOS transistor.

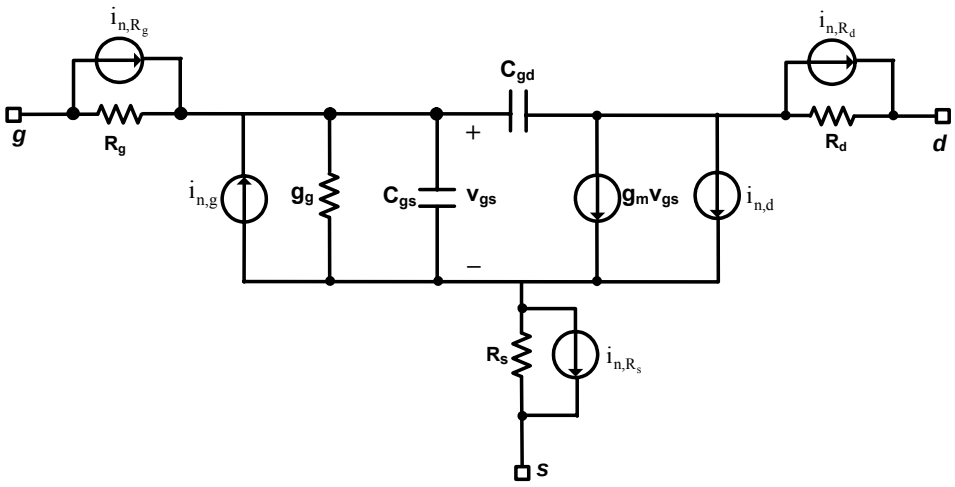


Figure 3.3: Small-signal equivalent circuits of MOSFET

3.2.2 Noise Model of MOSFET

The small signal equivalent circuits of MOSFET including noise sources is shown in Figure 3.3, the resistive MOSFET channel has a thermal noise $\overline{i_{n,d}^2}$ with power spectral density given by

$$\frac{\overline{i_{n,d}^2}}{\Delta f} = 4kT\gamma g_{d0} \tag{3.10}$$

where k is the Bolzmann constant, T is the absolute temperature in Kelvin, γ is the channel thermal noise coefficient, and g_{d0} is the channel conductance at zero drain-to-source voltage. For long channel devices $g_{d0} = g_m$, and for short channel devices g_{d0} is larger than g_m.

At high frequencies, the coupling between channel and gate is due to a distributed RC network, which results in a real part of the gate admittance g_g. In the pinch-off region, g_g is related to the radian frequency ω, gate-source capacitor C_{gs} and g_{d0} through [66]

$$g_g = \frac{{C_{gs}}^2 \omega^2}{5 g_{d0}} \tag{3.11}$$

This conductance has a thermal noise $\overline{i_{n,g}^2}$ associate with it, which is called induced gate noise. The power spectral density of $\overline{i_{n,g}^2}$ is given by [66]

$$\frac{\overline{i_{n,g}^2}}{\Delta f} = 4kT\delta g_g \tag{3.12}$$

where δ is the gate noise coefficient. Since $\overline{i_{n,d}^2}$ and $\overline{i_{n,g}^2}$ are originated from the same noise source, they are partially correlated with a complex correlation coefficient *c* given by

$$c = \frac{\overline{i_{n,g} i_{n,d}^*}}{\sqrt{\overline{i_{n,g}^2}\,\overline{i_{n,d}^2}}} \tag{3.13}$$

For long-channel CMOS devices operating at pinch-off and strong inversion, the values of γ, δ, and *c* are given by 2/3, 4/3, and 0.395j, respectively [66]. The noise characteristics for short channel CMOS devices have been investigated, and it is found that γ and δ tend to increase with the decrease of channel length [68]-[72]. The typical values of γ, δ, and *c* for 0.18-μm MOSFET are 2, 4, and 0.4j respectively according to [67][70].

The parasitic ohmic resistance at each node also contributes thermal noise, as shown in Figure 3.3. In addition, the transistor also suffers from the noise coupled from substrate.

3.2.3 Noise Parameters of MOSFET

In this subsection we are going to derive the four noise parameters of MOSFETs based on the model illustrated in Figure 3.3. First we need to make some reasonable assumptions to simplify the analysis. In this subsection we ignore the thermal noise of the stray resistance R_g, R_s, and R_d, for those resistance are generally very small in multi-finger transistors using minimum finger width so that their contribution to total output noise power is negligible compared to that of $i_{n,d}$ and $i_{n,g}$. We also temporarily ignore g_g because g_g is in parallel with C_{gs}. By reformatting (3.11) and using

$$\omega_T = \frac{g_m}{C_{gs}} \tag{3.14}$$

where ω_T is the transistor cut-off frequency, we obtain

$$g_g \approx \frac{1}{5}\frac{\omega}{\omega_T}\omega C_{gs} \tag{3.15}$$

which indicates that g_g is much smaller that ωC_{gs} when ω is well below ω_T and can be regarded as a second-order effect. Previous publications confirm accurate analytical results can be obtained in good agreement with the measurements [67] without account for C_{gd} due to its value being much smaller than C_{gs}. Therefore, here we assume the common-source transistor is unilateral with zero C_{gd}. For a similar reason we also neglect the back-gate transconductance g_{mb}. In our derivation, we use the equation $g_m = g_{d0}$, which is true for long-channel MOSFET. For short-channel device g_{d0} is larger than g_m. The inequality between g_{d0} and g_m as well as the effects of g_{mb} and g_g will be taken into account later in a more sophisticated analysis.

With the above assumptions, we calculate the MOSFET input noise voltage v_n, the equivalent input noise current i_n, and their correlation parameter $Y_r = G_r + jB_r$. There are two general configurations of the input transistor, common-source and common-gate, as shown in Figure 3.3.4 (a) and (b) respectively. Firstly by analysis of the common-source stage we have

$$v_n = -\frac{i_{n,d}}{g_m} \tag{3.16}$$

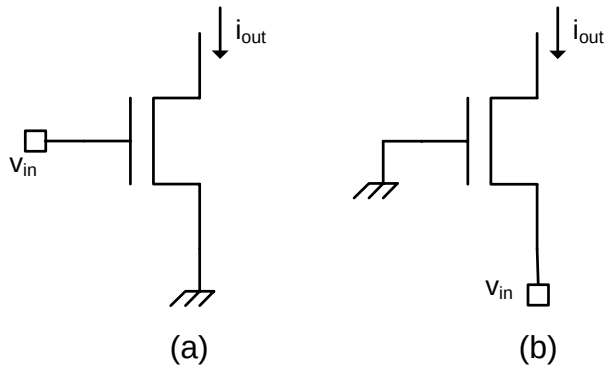


Figure 3.4: Transistor configuration (a) common-source (b) common-gate

$$i_n = -\frac{j\omega C_{gs}}{g_m} i_{n,d} - i_{n,g} \tag{3.17}$$

$$G_r = 0 \tag{3.18}$$

$$B_r = \omega C_{gs}(1 + |c|\sqrt{\frac{\delta}{5\gamma}}) \tag{3.19}$$

With the aid of (3.4) to (3.8), we obtain the four noise parameters [67]

$$R_n = \frac{\gamma}{g_m} \tag{3.20}$$

$$G_o = \omega C_{gs}\sqrt{\frac{\delta}{5\gamma}(1 - |c|^2)} \tag{3.21}$$

$$B_o = -\omega C_{gs}(1 + |c|\sqrt{\frac{\delta}{5\gamma}}) \tag{3.22}$$

$$F_{\min} = 1 + 2\left(\frac{\omega}{\omega_T}\right)\sqrt{\frac{\gamma\delta}{5}(1 - |c|^2)} \tag{3.23}$$

Similar analysis is applied to the common-gate stage. Interestingly, it is found that the four noise parameters of common-gate configuration are exactly identical to those of common-source.

Now we can draw the following conclusions:

1. The noise properties of the amplifier are the same when the signal source is applied at either gate or source of the transistor.
2. $F_{\min}$ increases linearly with the ratio of the operation frequency and the transistor f_T.
3. From (3.21), (3.21), and (3.22) we can express optimum source impedance $Z_o = 1/(G_o + jB_o)$ as

$$Z_o = \frac{1}{\omega C_{gs}}\beta_1 \tag{3.24}$$

where β_1 is a complex constant for a specified process whose value is only related to γ, δ and $|c|$. Therefore, Z_o is inversely proportional to transistor width and operation frequency for certain process.

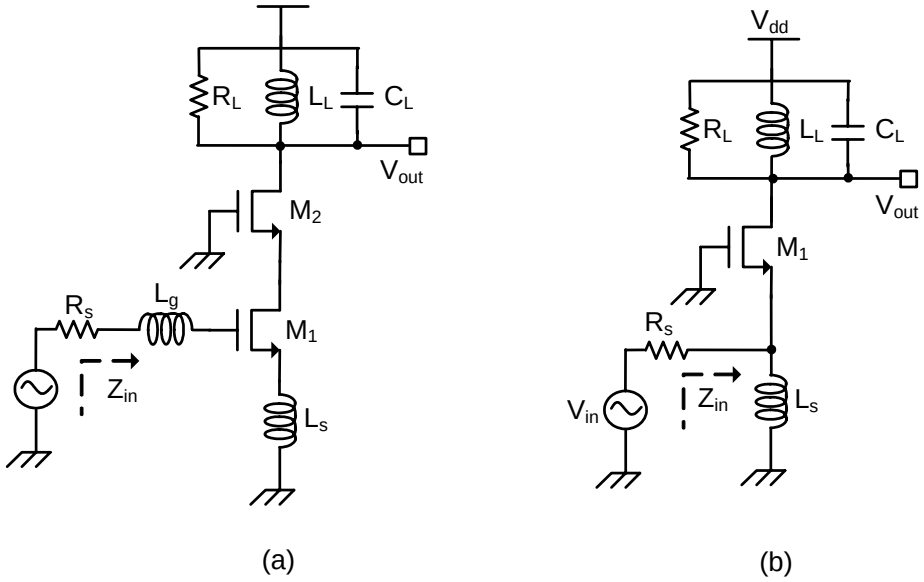


Figure 3.5 LNA topologies (a) common-source with inductive degeneration (b) common-gate

3.2.4 Common-Source and Common-Gate LNA

Although the noise properties of common-source and common-gate LNA are identical with the same signal source, their input impedance Z_{in} is significantly different, i.e., Z_{in} of common-source is mostly imaginary, but Z_{in} of common-gate has a real part given by $1/g_m$. As we know, the LNA is not only needed to achieve a lowest noise figure, but also to be input-matched to the source impedance to avoid the loss of signal power. Such requirements lead to a major difference in common-source and common-gate LNA design, which is going be elaborated in this subsection.

The common-source with inductive degeneration stage as shown in Figure 3.5 (a) has been commonly used in CMOS LNA implementations. Many previous works [14][60][61] [72]-[77] show that this stage can achieve good performance at low GHz bands. The source inductance L_s introduces a real part to the impedance Z_{in} looking into the gate, which is given by

$$\mathrm{Re}[Z_{in}] = \frac{g_m}{C_{gs}} L_s \tag{3.25}$$

This real impedance is used to match the amplifier's input impedance to source impedance, which is usually 50Ω. Furthermore, there is a well-known design

procedure [78] intended for achieving noise matching and power matching simultaneously. This design flow is summarized as follows.

1. The drain current density associated with the lowest $F_{\min}$ is determined first.
2. The transistor size is scaled by varying the number of gate fingers while maintaining constant finger width and current density until R_o reaches R_s.
3. The source degeneration inductance L_s is added to bring the real part of the amplifier's input impedance Z_{in} to R_s. It has been verified that L_s has negligible impacts on $F_{\min}$ and R_o of the amplifier [78].
4. The gate inductance L_g is added to neutralize Im[Z_{in}] at the operation frequency.

By this means, Z_{in} and R_o of the amplifier are matched to R_s. However, $F = F_{\min}$ is only obtained here if X_o is zero, which is not true due to the correlation between $i_{n,d}$ and $i_{n,g}$. Although common-source degeneration techniques can bring Z_o closer to Z_s while maintaining a low $F_{\min}$, power matching and noise matching cannot be perfectly achieved at the same time. Tradeoff between noise figure and signal power transfer are inevitably involved in the common-source LNA design.

It is also instructive to investigate the LNA performance under perfect input power matching conditions. For the common-source stage in Figure 3.5 (a), assuming $i_{n,d}$ is the dominant internal noise source, the effective transconductance G_m and the noise factor of the stage can be approximated with [67]

$$G_{m,CS} \approx \frac{1}{2R_s}\frac{\omega_T}{\omega_o} \tag{3.26}$$

$$F_{CS} \approx 1 + \gamma g_{d0} R_s \left(\frac{\omega_o}{\omega_T}\right)^2 \tag{3.27}$$

where ω_o is the operation frequency of the LNA. The dependence of (3.26) and (3.27) on ω_o / ω_T indicates that an inductively degenerated common-source LNA is well suited for applications where ω_o is well below ω_t. However, the performance of this topology degrades substantially when ω_o becomes comparable to ω_t.

In contrast, in the common-gate LNA shown in Figure 3.5 (b), the resistive part looking into the source of the transistor is used to match the input to R_s. In the MOSFET case, this impedance is given by $1/g_m$. The source inductance L_s is used to resonate out the capacitance seen at the source at the working frequency. Obviously, g_m is fixed to $1/R_s$ for the purpose of power matching. Scaling of the transistor size with constant g_m will either result in a low ω_T and thus a high $F_{\min}$, or a Z_o far away from R_s. Therefore, compared to common-source with an inductive degeneration technique, the common-gate stage is lack of the flexibility of adjusting transistor size to bring Z_o closer to R_s while maintaining a low $F_{\min}$. In other words, the goals of power matching and a low noise figure strongly conflict with each other in a common-gate stage.

At perfect power matching and taking $i_{n,d}$ into account only, the effective transconductance and noise factor of the common-gate stage can be expressed as

$$G_{m,CG} = \frac{1}{2R_s} \tag{3.28}$$

$$F_{CG} \approx 1 + \gamma \tag{3.29}$$

Equations (3.28) and (3.29) are independent of frequency, indicating the performance of the common-gate stage degrades more gracefully with the increase of working frequency . However, the achievable noise figure at the power matching condition is far above $F_{\min}$, which disqualifies common-gate as an optimal design.

3.2.5 Common-Gate with Resistive Feedthrough LNA

In this subsection we are going to introduce a novel LNA input stage, common-gate with resistive feedthrough (CGRF) [57][63], which provides a low noise figure by lowering $F_{\min}$, R_n of the traditional common-gate LNA and reducing $|Y_o - Y_s|$ at the power matching condition.

The schematic of the CGRF architecture is illustrated in Figure 3.6. In this topology, a feedthrough resistor R_f is added to the traditional common-gate stage in parallel with the input transistor. C_p is a large capacitor for isolating dc level. R_L is the resistive load at the drain of M_1 owing to the finite quality factor Q of the resonant

load. For input matching and channel selection, both parasitic capacitances at source and drain of M_1 should be absorbed into the LC tank and resonated out at working frequency, i.e.,

$$\omega_o = \frac{1}{\sqrt{L_s C_s}} = \frac{1}{\sqrt{L_L C_L}} \tag{3.30}$$

where C_s and C_L are the capacitance seen at source and drain respectively of the input transistor. Figure 3.7 shows the small signal equivalent circuit of the CGRF stage at resonance, including major noise sources.

The idea of adding a feedthrough resistor originates from the observation that $i_{n,d}$ and R_f form a closed loop. If R_f is small compared to R_L, it will attract a substantial amount of $i_{n,d}$ flowing only inside this loop instead of going to the output, so that the total output noise power is reduced.

R_f can introduce additional thermal noise into the circuits. However, R_f can be formed by the transistor gate-drain resistance r_{ds} alone or a parallel combination of r_{ds} and an external resistance R_p, as shown in Figure 3.7. r_{ds} is a small-signal equivalent resistance and thus it is noise-free. For analyzing the worst-case scenario, we assume R_f is purely formed by a real resistor R_p associated with a noise current generator $i_{n,f}$, where

$$\frac{\overline{i_{n,f}^2}}{\Delta f} = 4kTR_f \tag{3.31}$$

Temporarily ignoring i_{n,R_L}, g_{mb}, and g_g, we obtain the noise parameters of the CGRF stage :

$$v_n = \frac{1}{g_m + 1/R_f} i_{n,d} + \frac{1}{g_m + 1/R_f} i_{n,f} \tag{3.32}$$

$$i_n = i_{n,g} + \frac{j\omega C_{gs}}{g_m + 1/R_f} i_{n,d} + \frac{j\omega C_{gs}}{g_m + 1/R_f} i_{n,f} \tag{3.33}$$

$$G_r = 0 \tag{3.34}$$

$$B_r = \omega C_{gs}\left(1 + \frac{\gamma + \gamma g_m R_f}{1 + \gamma g_m R_f}|c|\sqrt{\frac{\delta}{5\gamma}}\right) \tag{3.35}$$

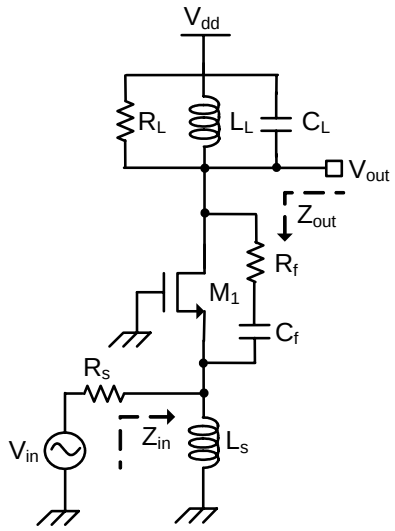


Figure 3.6: Common-gate with resistive feedthrough LNA

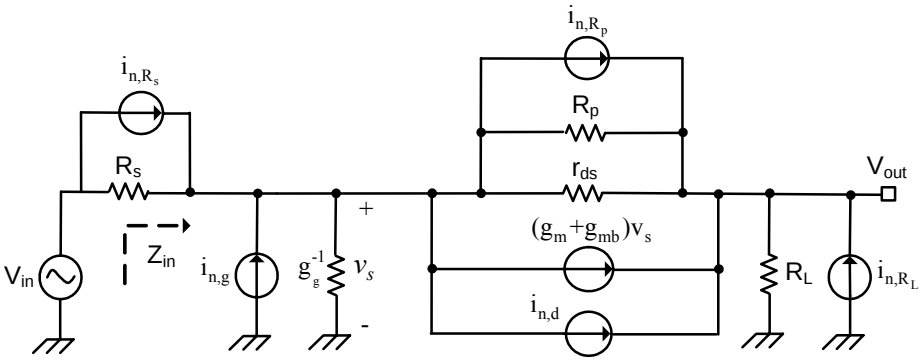


Figure 3.7: Small-signal circuits of CGRF stage

$$R_n = \frac{\gamma g_m + 1/R_f}{(g_m + 1/R_f)^2} \tag{3.36}$$

$$G_o = \omega C_{gs}\sqrt{\frac{\delta}{5\gamma}\left(1-|c|^2\frac{\gamma g_m R_f}{1+\gamma g_m R_f}\right)\frac{\gamma(1+g_m R_f)^2}{g_m R_f(1+\gamma g_m R_f)}} \tag{3.37}$$

$$B_o = -B_r \tag{3.38}$$

$$F_{\min} = 1 + 2\left(\frac{\omega_o}{\omega_t}\right)\sqrt{\frac{\delta\gamma}{5}\left(1-|c|^2\frac{\gamma g_m R_f}{1+\gamma g_m R_f}\right)\frac{g_m R_f(1+\gamma g_m R_f)}{\gamma(1+g_m R_f)^2}} \tag{3.39}$$

When R_f approaches infinity, Equation (3.36) to (3.39) converge to (3.20) to (3.23) as expected. On the other hand, when R_f decreases toward zero, $F_{\min}$ is lowered to one and R_n reduces to zero. With $|c| \approx 0.4$ and $\gamma \geq 2/3$, both $F_{\min}$ and R_n monotonously decrease with R_f. The noise parameters normalized to their values of the traditional COMMON-GATEstage ($g_m R_f = \infty$) are plotted in Figure 3.8 as a function of $g_m R_f$. Figure 3.8 (a) is for long-channel devices with $|c| \approx 0.4, \gamma = 2/3$ and $\delta = 4/3$. Fig .5.8 (b) is for short-channel devices with $|c| \approx 0.4, \gamma = 2$ and $\delta = 4$. It can be observed that when $g_m R_f \leq 1$, $F_{\min}$ decreases rapidly toward one. However, when $g_m R_f \geq 10$, $F_{\min}$ only reduces slightly compared to that of the traditional COMMON-GATEstage. Similar trends are exhibited for the remaining three noise parameters too, with the exception that G_o is decreasing instead of increasing with $g_m R_f$. In practical circuits design, $g_m R_f >> 1$ usually holds, which is going to be explained later in this section. Therefore, we can use Equation (3.20) to (3.23) to approximate the noise parameters of the CGRF stage.

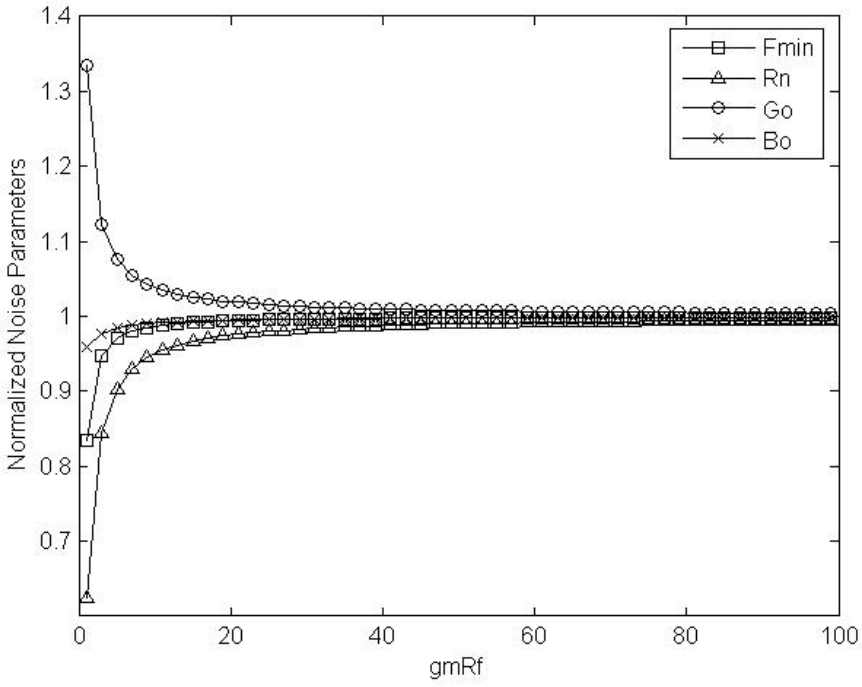
Now we investigate the power matching condition. Analysis of the circuits in Figure 3.7 (a) yields that Z_{in} at resonance is related to g_m, R_f, and R_L via

$$Z_{in}(\omega_o) = \frac{R_f + R_L}{1 + g_m R_f} \tag{3.40}$$

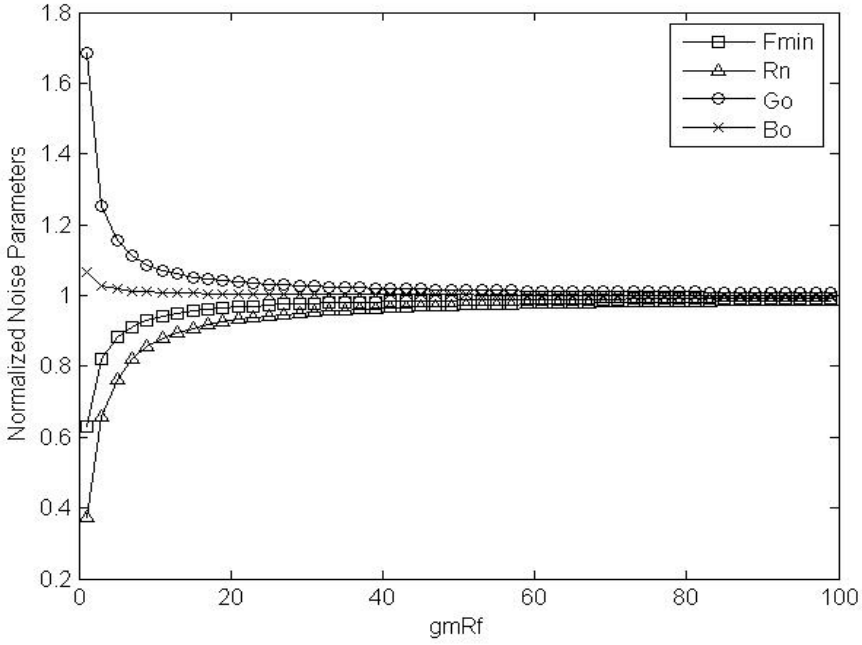
In power matching this impedance equals to R_s and the effective transconductance of CGRF stage is given by

$$G_{m,CGRF} = \frac{1}{2R_s} \tag{3.41}$$

It is noted that $G_{m,CGRF}$ is equal to $G_{m,CG}$ and independent of g_m and R_f. This is because when $Z_{in} = R_s$, the input current is always given by $v_{in}/(2R_s)$. This current is separated into two branches at the source, one branch flows through the transistor, and the other through R_f. These two branches of current then recombine at the drain output. Therefore, the variations of g_m and R_f subject to power matching constrains only change the current distribution between the transistor and R_f, but do not change the total output current.



(a)



(b)

Figure 3.8: Normalized noise parameters as a function of $g_m R_f$
(a)long channel (b) short channel

The voltage gain of the input-matched CGRF stage at resonance is given by

$$A_{V,CGRF} = \frac{R_L}{R_s} \tag{3.42}$$

The power matching constraints $Z_{in}(\omega_o) = R_s$ yields that

$$g_m R_f = \frac{R_f}{R_s} + A_{V,CGRF} - 1 \tag{3.43}$$

In general a practical amplifier desires a high voltage gain, resulting in $g_m R_f >> 1$.

The power matching condition can be simplified to

$$R_s = \frac{1}{g_m}\left(1 + \frac{R_L}{R_f}\right) \tag{3.44}$$

As we discussed earlier, when $g_m R_f >> 1$ the four noise parameters of the CGRF stage are insensitive to the change of R_f. However, (3.44) reveals that Z_{in} changes rapidly with R_f as long as $R_f < R_L$. These two observations remind us that the source inductor L_s in inductively degenerated common-source topology exhibits similar effects on noise parameters and input impedance. Therefore, a design procedure analogous to the one used for CS LNA can be used in designing CGRF stage, which is summarized below

1. The drain current density associated with the highest ω_T is determined.
2. Since G_o is proportional to C_{gs} at certain frequency, the transistor size is scaled by varying the number of gate fingers while maintaining constant finger width and current density, bringing G_o equal to G_s while maintaining the lowest $F_{\min}$.
3. A parallel resistor R_f is added to adjust $\mathrm{Re}[Y_{in}]$ to G_s. The appropriate value of R_f is a function of R_s, g_m, and R_L, subject to (3.44).
4. The source inductance L_s is added to adjust $\mathrm{Im}[Y_{in}]$. Power matching requires that L_s resonates out the capacitance seen at the source thus $\mathrm{Im}[Y_{in}] = 0$, resulting in

$$L_s = \frac{1}{\omega_o^2 C_{gs}} \tag{3.45}$$

One the other hand, according to (3.22), the noise matching prefers

$$L_s = \frac{1}{\omega_o^2 C_{gs}(1+|c|\sqrt{\frac{\delta}{\gamma}})} \tag{3.46}$$

which leads to S_{11} at approximately -13dB, indicating that $F_{\min}$ can be achieved with fair, though not perfect, power matching. In actual circuit design, the value of L_s could be somewhere in between (3.45) and (3.46) as a tradeoff between noise figure and power matching.

This design procedure minimizes $|Y_o - Y_s|$ while slightly reducing R_n and $F_{\min}$. Therefore, the noise factor of the CGRF stage is significantly less than that of the traditional COMMON-GATEstage. A lower R_n is also beneficial to minimize the sensitivity of the amplifier's noise performance to source impedance deviations due to the inaccurate modeling.

Compared to the conventional COMMON-GATEstage, the CGRF stage can achieve a substantially lower noise figure without sacrificing the gain and input power matching. The price comes with a higher power consumption. As can be seen in (3.44), compared to the COMMON-GATEstage a higher g_m is required to maintain the power matching. With constant ω_T, g_m is proportional to dc current. Therefore, CGRF technique provides a direct way to trade between noise and power.

3.2.6 Noise Factor Optimization under Power Matching Constraints

As we pointed out, because perfect noise matching and power matching can not be achieved simultaneously in either CS or CGRF LNA design, the final choice of the amplifier's input impedance is a trade-off between low noise performance and signal power transfer. In many applications, a good power matching with sufficient margin is mandatory, requiring LNA to be designed by optimizing F under power matching constraints instead of designing for $F_{\min}$. In this subsection we derive the expression for the noise factor of the CGRF stage with a perfect power match and compare to that of CS topology.

Amplifier's gain requirement results in $g_m R_f >> 1$, indicating $i_{n,d} >> i_{n,f}$. Hence we ignore $i_{n,f}$ as well as i_{n,R_L} in our analysis from now on. Instead, for a more

accurate result, we take g_g , g_{mb} , and the difference between g_m and g_{d0} into account. We define

$$\chi \equiv \frac{g_{mb}}{g_m} \tag{3.47}$$

$$\alpha \equiv \frac{g_m}{g_{d0}} \tag{3.48}$$

$$\eta(\omega_0) \equiv \frac{g_g}{g_m} \approx \frac{\alpha}{5}\left(\frac{\omega_0}{\omega_T}\right)^2 \tag{3.49}$$

A more accurate expression for the input impedance of the small-signal circuits in Figure 3.7 is given by

$$Z_{in} = \left(\frac{1 + g_m R_f (1+\chi)}{R_f + R_L} + \eta(\omega_0) g_m\right)^{-1} \tag{3.50}$$

The fact that $R_f /(R_f + R_L) < 1$ results in a lower-bound on g_m which is given by

$$g_m \geq \frac{1}{R_s(1+\eta(\omega_0))} \tag{3.51}$$

A value of g_m lower than this will make it impossible to achieve a perfect input match.

Assuming Z_{in} is perfectly matched to R_s, the effective transconductance of CGRF stage is given by

$$G_{m,CGRF} = \frac{1}{2R_s}\left(1 - g_m R_s \eta(\omega_0)\right) \tag{3.52}$$

which indicates a large g_m can degrade the gain. This is because the increase of g_m results in a larger g_g, making more signal loss through the gate. We choose the higher-bound of g_m as the value that causes 3dB G_m degradation from its low-frequency value, i.e.,

$$g_m \leq g_{m,-3dB} = \frac{1}{2R_s\eta(\omega_0)} \tag{3.53}$$

Input matching criterion and gain consideration set the limits on the design parameter g_m. The following discussion on noise figure is based on the assumption that the input is perfectly matched and the noise of the following stage is negligible. Therefore the expressions following are only valid for g_m inside the range specified by (3.51) and (3.53). A lower g_m will violate the input match assumption and a larger g_m can

tremendously degrade gain, making the noise contribution of the following stages significant.

At perfect power matching the following expression for F is yielded, as explained in Appendix 3.1:

$$F_{CGRF} \approx 1+\frac{\gamma}{\alpha}\left(\frac{1}{1+\chi}\right)^2\left(\frac{1}{g_m R_s}+\eta^2(\omega_0)g_m R_s+2\eta(\omega_0)\right)+\delta\eta(\omega_0)g_m R_S \qquad (3.54)$$

where the second term represents the contribution of channel thermal noise and the third term accounts for the contribution of induced gate noise. By equating the derivative of F_{CGRF} to zero and solving for g_m, we obtain an optimum g_m for the lowest noise figure, i.e.,

$$g_{m,CGRF,opt}=\frac{1}{R_s}\left(\frac{\delta\alpha}{\gamma}(1+\chi)^2\eta(\omega_0)+\eta^2(\omega_0)\right)^{-\frac{1}{2}} \qquad (3.55)$$

and the corresponding optimized F under power matching constraints given by

$$F_{CGRF,opt}=1+\frac{\gamma}{\alpha}\frac{2}{1+\chi}\eta(\omega_0)+\frac{2}{1+\chi}\sqrt{\frac{\gamma\delta}{\alpha}\eta(\omega_0)+\left(\frac{\gamma}{\alpha}\right)^2\left(\frac{1}{1+\chi}\right)^2\eta^2(\omega_0)} \qquad (3.56)$$

The noise factor of the inductively degenerated common-source stage under perfect power matching and its corresponding lowest value is given by [79]

$$F_{CS}=1+\left(\frac{\omega_0}{\omega_T}\right)\frac{\gamma}{\alpha}\left[\left(1+2|c|\sqrt{\frac{\delta\alpha^2}{5\gamma}}+\frac{\delta\alpha^2}{5\gamma}\right)\frac{1}{Q_s}+\frac{\delta\alpha^2}{5\gamma}Q_s+\frac{2\alpha}{5}\frac{\omega_0}{\omega_T}(1+c\sqrt{\frac{\delta\alpha^2}{5\gamma}})\right] \qquad (3.57)$$

where

$$Q_s=\frac{1}{g_m R_S}\left(\frac{\omega_T}{\omega_0}\right) \qquad (3.58)$$

By similar procedure, we obtain the lowest value for F_{CS} when it is optimized under power matching constraints

$$F_{CS,opt}=1+\gamma\left[\sqrt{\frac{4\delta}{5\gamma}\left(1+2|c|\sqrt{\frac{\delta\alpha^2}{5\gamma}}+\frac{\delta\alpha^2}{5\gamma}\right)}\left(\frac{\omega_0}{\omega_T}\right)+\frac{2}{5}\left(1+c\sqrt{\frac{\delta\alpha^2}{5\gamma}}\right)\left(\frac{\omega_0}{\omega_T}\right)^2\right] \qquad (3.59)$$

The difference between $F_{CS,opt}$ and $F_{CGRF,opt}$ can be evaluated by subtracting (3.56). from (3.57). Assuming $\alpha=1$ and $\chi=0$, this difference is given by

$$F_{CS,opt} - F_{CGRF,opt} = \sqrt{\frac{4\gamma\delta}{5}}\left(\sqrt{\left(1+2|c|\sqrt{\frac{\delta}{5\gamma}}+\frac{\delta}{5\gamma}\right)}-1\right)\left(\frac{\omega_0}{\omega_T}\right)+\frac{2c}{5}\sqrt{\frac{\gamma\delta}{5}}\left(\frac{\omega_0}{\omega_T}\right)^2 \quad (3.60)$$

With (3.60), it is evident that CGRF LNA is advantageous over common-source LNA in terms of minimum achievable noise figure. And most importantly, this improvement increases with frequency.

Power dissipation is another important factor in LNA design. To illustrate the trade-off between power and noise figure, the noise figure is plotted against g_m in Figure 3.9 for both common-source and CGRF stages at different ω_0 to ω_T ratio. For CGRF curves the g_m is constrained in the range specified by (3.51) and (3.53). To first order estimation, assuming the MOSFET is biased at a fixed ω_T, a larger g_m is directly related to a larger transistor width and larger power dissipation. Figure 3.10 shows that the optimum g_m of common-source topology is less than that of a CGRF one. Figure 3.10 also illustrates that at high power level CGRF stage achieves a lower noise figure than common-source one, but at low power level it is opposite. Therefore, the choice of topology depends on both the noise specification and the power budget. At very high frequencies, where low-noise is a principle challenge and can not be met by using common-source LNA due to its theoretical limitation, CGRF LNA provides a way to design towards lower noise figure at the price of more power consumption.

Low-noise requirements lead to our choice of CGRF in this work. However, we also take power consumption into consideration and a trade-off will be shown in the next section.

3.2.7 Stability

Since in CGRF stage R_f acts as a positive feedback, the stability issue needs to be carefully addressed. Considering the input transistor with feedthrough resistor as a two-port network shown in Figure 3.10, it is a sufficient condition to prevent oscillation that the real parts of both impedances seen looking into the source and the drain of the input transistor are positive. It is easy to show that Re[Z_{in}] and Re[Z_{out}] can be expressed as

$$\mathrm{Re}[Z_{in}] = \frac{R_f + \mathrm{Re}[Z_d]}{1 + g_m R_f (1+\chi)} \quad (3.61)$$

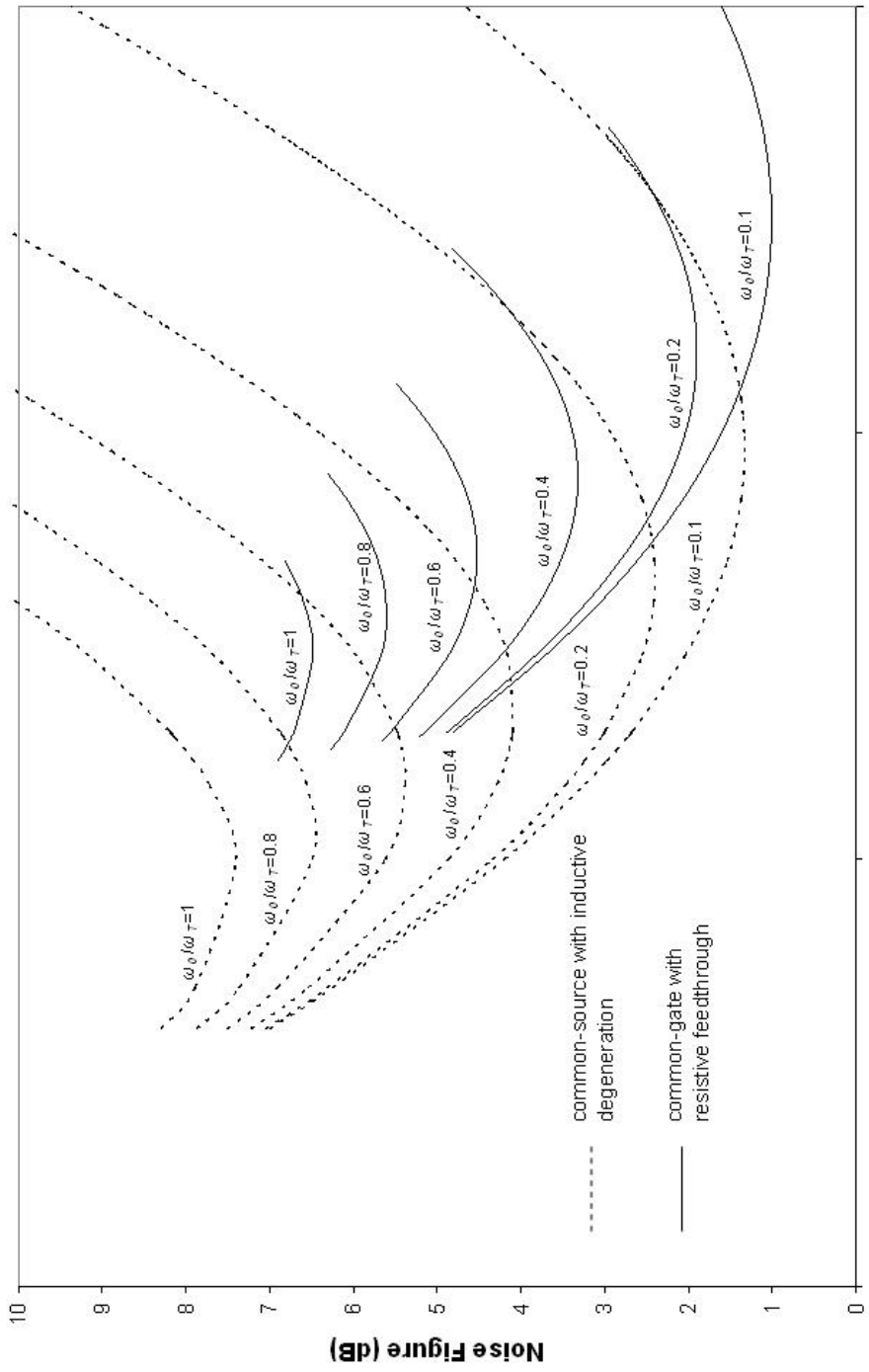


Figure 3.9: Noise figure of CS and CGRF LNA under power matching constraints

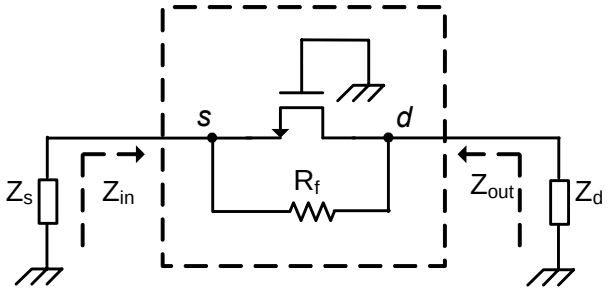


Figure 3.10: Two-port configuration

$$\mathrm{Re}[Z_{out}] = R_f + (g_m R_f + 1)\mathrm{Re}[Z_s] \tag{3.62}$$

where Z_s and Z_d are the load impedances at source and drain respectively. (3.61) and (3.62) indicate that as long as Re[Z_d] and Re[Z_s] are positive, which is true for any passive termination, the stability of the CGRF stage is guaranteed.

3.3 Circuits Implementation

3.3.1 Neutralizing Substrate Effects

The analysis in the previous section ignores all substrate effects. However, in the 24-GHz range capacitive coupling and resistive loss through the substrate have a tremendous influence on the circuits performance. A simplified substrate network model for a MOSFET is shown in Figure 3.11 [80]. Simulation results show that the capacitive coupling between drain and source through this network harms stability and noise figure. A shunt inductor L_p in a series with a large bypass capacitor C_p can

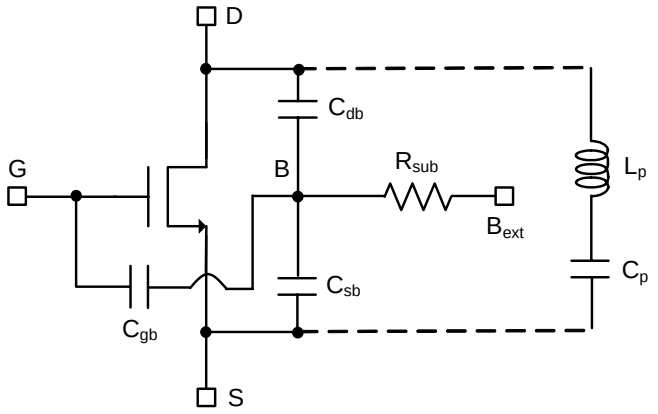


Figure 3.11 Reducing substrate coupling by using parallel inductor

be added, as shown in Figure 3.11, to resonate the equivalent capacitance between drain and source so that the substrate effects are reduced. The series resistance of L_p can be converted to an equivalent parallel resistance, which affects the performance of the LNA as a feedthough resistor. In this case, the feedthrough resistance can be expressed as

$$R_f = Q\omega_0 L_p \,||\, r_{ds} \tag{3.63}$$

where Q is the quality factor of L_p.

3.3.2 Schematics of the Front-End

Figure 3.12 shows the 24-GHz CMOS LNA. It consists of three stages: The first stage employs common-gate with resistive feedthrough topology, where the shunt inductor L_2 resonates the capacitive coupling while forming a feedthrough resistance given by (3.61) between the drain and source of M_1. The capacitor C_1 bypasses the gate of M_1 to ground at high frequencies. The second and third stages are both common-source with inductive degeneration amplifiers used to enhance the overall gain. AC coupling is employed between the stages.

The peak f_T of the 0.18-μm CMOS device used at the 1.5 V bias is about 60 GHz. To achieve the minimum noise figure at 24 GHz, the optimum g_{m1} is estimated to be about 80 mS by using (3.55). To reduce the power consumption we choose g_{m1} to be 40 mS in this design. We also reduce the $\left(V_{gs} - V_t\right)$ by half from its value for peak f_T, which is more power efficient, resulting in a current decrease of more than 50%. The f_T however is only reduced by about 10%. Finally, M_1 is biased at 8 mA with 54GHz f_T. The second and third stages consume 4 mA each.

The model suggested in [67] is used to simulate the noise performance, including the effect of gate noise. In the simulation we use γ=2 and δ=4. The simulation result shows that the noise figure of the 3-stage LNA is 5.7 dB and that of the first stage is 4.6 dB. This includes the contributions of all parasitic noise sources. The noise figure of the first stage associated with $\overline{i_{n,d}^2}$ and $\overline{i_{n,g}^2}$ only is 3.9dB.

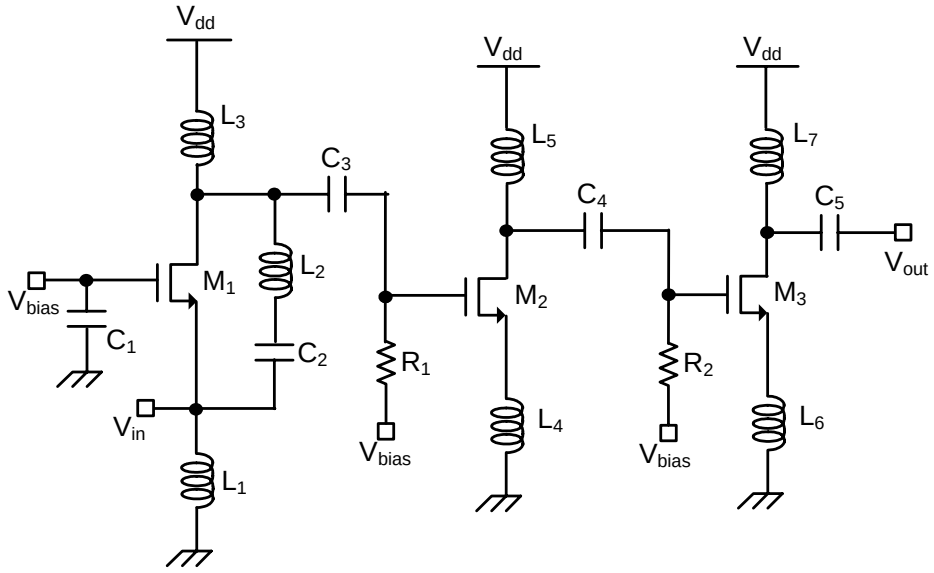


Figure 3.12: Three-stage LNA

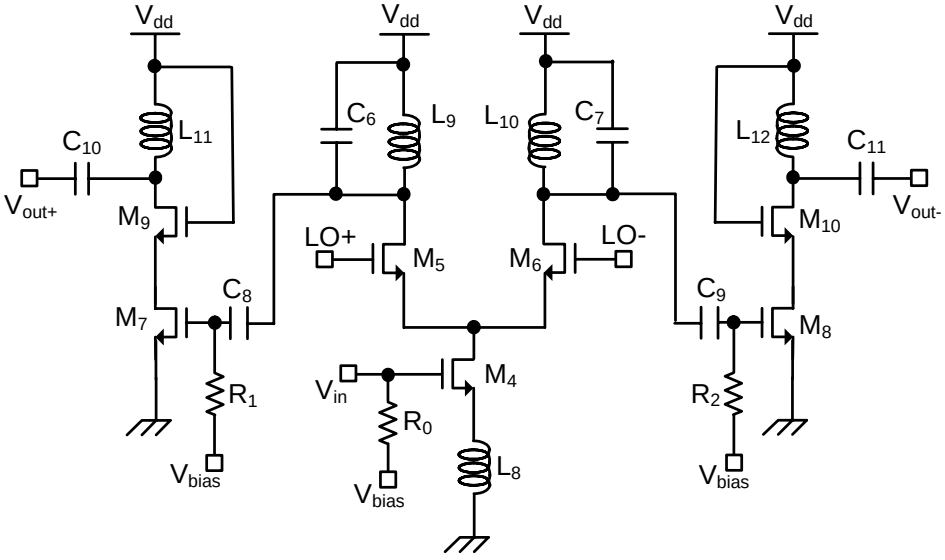


Figure 3.13: Downconversion mixer

We also simulated the common-source with inductive degeneration LNA of Figure 3.5(a) using the same model. The contribution of M_2 noise to the overall noise figure is significant because of the reduced gain of M_1 at high frequencies. The minimum simulated noise figure by using a single stage common-source with inductive degeneration LNA is more than 6 dB.

Following the LNA, the mixer shown in Figure 3.13 is used to downconvert a 24 GHz RF signal to 5 GHz IF. The core is a conventional single-balanced Gilbert cell. The RF input applies at the gate of M_4 which is used as a transconductance amplifier. The linearity of this transconductance amplifier is improved by using a source degeneration inductor L_8. L_8 also adjusts the input impedance seen looking into the gate of M_4 in order to improve the input matching at the LNA-mixer interface. The M_4 is biased at 4mA dc current.

The chopping function is accomplished by the M_2~M_3 mixing cell, and a 1.6 V peak-to-peak differential LO signal is applied. Cascode amplifiers following differential mixing cell are used to drive the 50-Ω loads. The output-match is accomplished by the LC impedance transforming network.

3.3.3 Layout Issues

The circuit was designed and fabricated using 0.18-μm CMOS transistors. The process offers 6 metal layers with two top layers of 1-μm thick copper. L_4 and L_6 in the LNA and L_8 in the mixer are slab inductors with an inductance around 0.1nH, all other inductors are spirals. All inductors are modeled by using electromagnetic (EM) simulation tool, such as Sonnet and ASITIC.

Long metal lines are used inevitably as interstage connections. The models of those metal lines are extracted from electromagnetic simulation and put back into circuits to examine their effects and adjust the design accordingly in post layout simulation.

Shielded pads [81] are employed at both RF and IF ports. Grounded Metal1 underneath the pads prevents loss of signal power and noise generation associated with the substrate resistance. Ground rings are placed around each transistor at

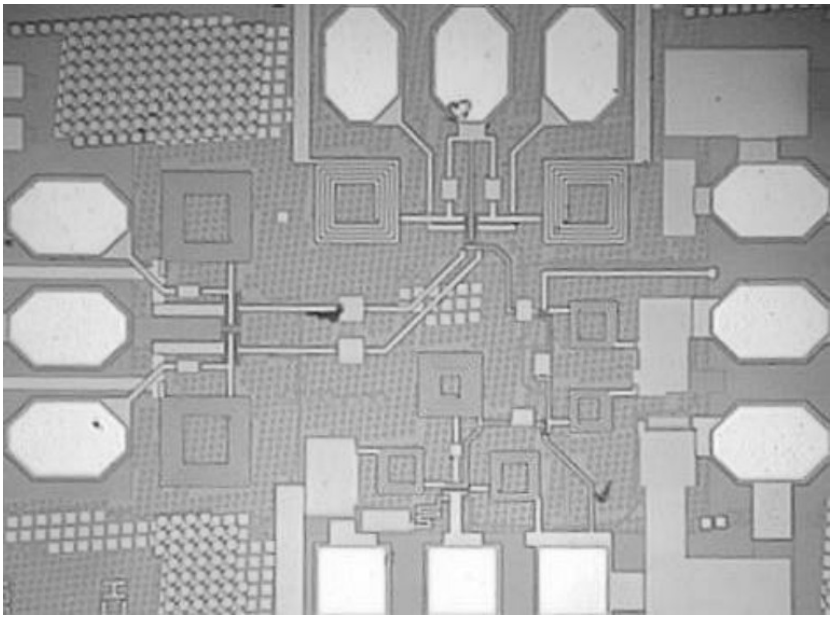


Figure 3.14: Die micrograph of the 24GHz CMOS front-end

minimum distance to reduce the substrate loss. Separated Vdd pads are assigned to the LNA, mixer and bias circuits. Large on-chip bypass capacitors with various sizes are placed between each Vdd and ground.

The die micrograph is shown in Figure 3.14. The size of the chip is 0.8 x 0.9 mm^2 including large area occupied by wide ground rings and pads. The size of the core cell is only 0.4 x 0.5 mm^2.

3.4 Experimental Results

The front-end is tested by probing the input, output, and LO ports. The power and ground pads are wire-bonded to the testing board. The differential 19GHz LO signal is provided by a signal generator and a 180^0 power splitter. Firstly, the reflection coefficients at RF and IF ports are tested by using a network analyzer. Then HP noise figure test set is employed for conversion gain and NF measurement.

Figure 3.15 shows the measured input and output reflection coefficients S_{11} and S_{22}. The RF input and the IF output are well matched at the respective frequencies. The measurement shows that a 27.5dB maximum power gain appears for an RF of 21.8 GHz and an IF of 4.9 GHz. Figure 3.16 shows the measured power gain and

extracted voltage gain with a 16.9-GHz LO frequency. The peak voltage gain is 8dB higher than the peak power gain instead of 6dB owing to the imperfect power matching. The frequency offset from the 24 GHz is likely due to inaccurate modeling of MOS transistor and planar inductor at high frequencies. The LNA achieves a 28 dB maximum voltage gain and 15 dB power gain. The mixer followed further enhances the signal power by 13 dB. Because of the imperfect conjugate-matching at the LNA- mixer interface, the overall power gain of the front-end is slightly lower than the sum of the individual power gain of the two blocks.

The measured noise figure is shown in Figure 3.18. A minimum noise figure of 7.7 dB is achieved for the combined LNA and mixer at 22.08 GHz. The individual noise figure of the LNA and the mixer are 6 dB and 17.5 dB respectively. The noise figure of the first CGRF stage is extracted to be 4.8dB. (3.54) calculates the noise figure to be 4.1dB. We came back to the simulation and found that the remaining 0.7dB can be attributed to the thermal noise of the parasitic resistance and substrate noise.

Figure 3.17 reports measured large signal nonlinearity. The input-referred -1dB compression point of the front-end appears at -23 dBm. The -1dB compression point of the LNA and the mixer alone are -8 dBm and -8.3 dBm, respectively. The image rejection of the front-end is -31 dB. This performance is achieved because of the large IF and the multi stage nature of the LNA. The overall current consumption of the front-end including output buffers is 43 mA, while 23 mA are consumed by the output buffers. The LNA and the mixer draw 16 mA and 4 mA, respectively from a 1.5-V supply voltage.

The measured performance of the front-end and the de-embedded LNA performance are summarized in Table 3.1. A comparison of the LNA in this work and the previously reported works is given in Table 3.2. Our work presented in this chapter was first published in 2002. The LNA performance is better than previously reported CMOS LNA about 15GHz in terms of power and noise. A 24GHz CMOS LNA [84] is reported in June 2004, presenting a slightly lower noise figure but much higher power consumption. As we discussed, the design of the CGRF LNA involves a tradeoff between noise and power consumption. If more dc current is used, a lower noise figure can be expected from our design.

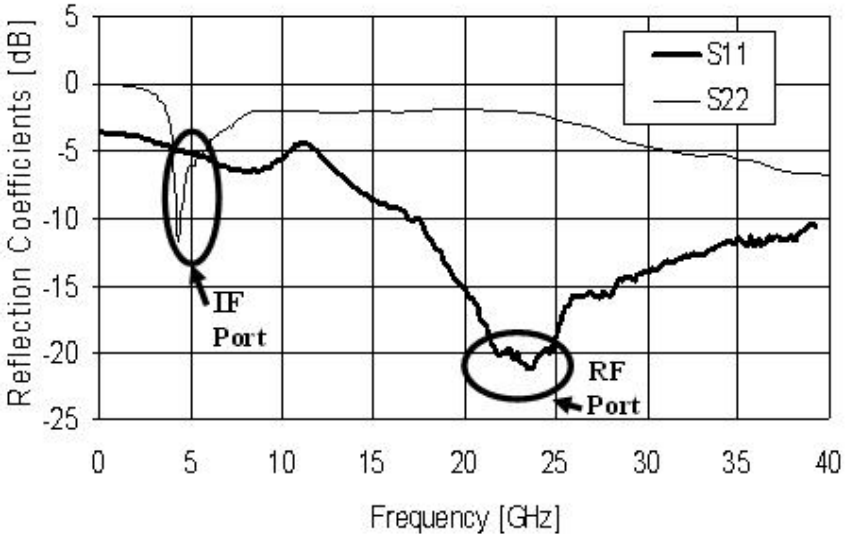


Figure 3.15: Input and output reflection coefficient

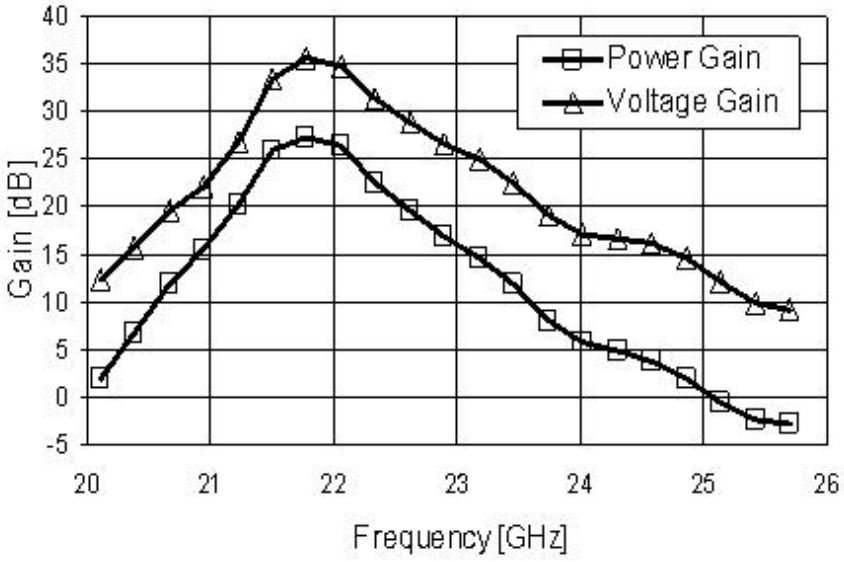


Figure 3.16: Voltage gain and power gain of the front-end

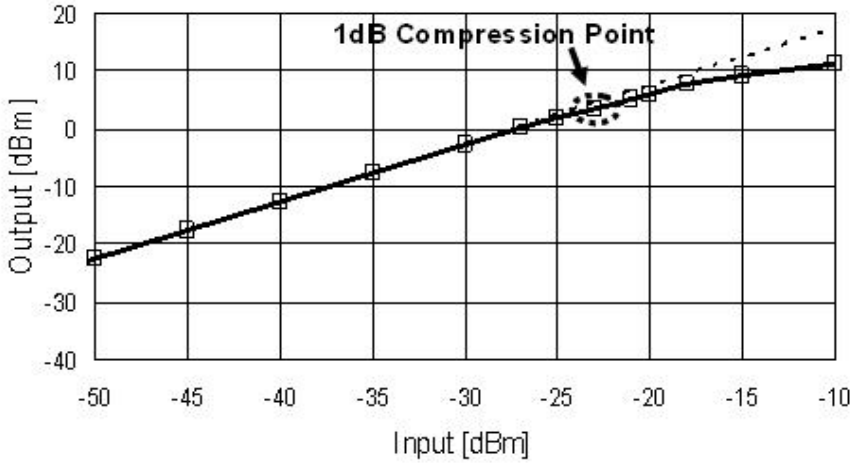


Figure 3.17: Large-signal nonlinearity

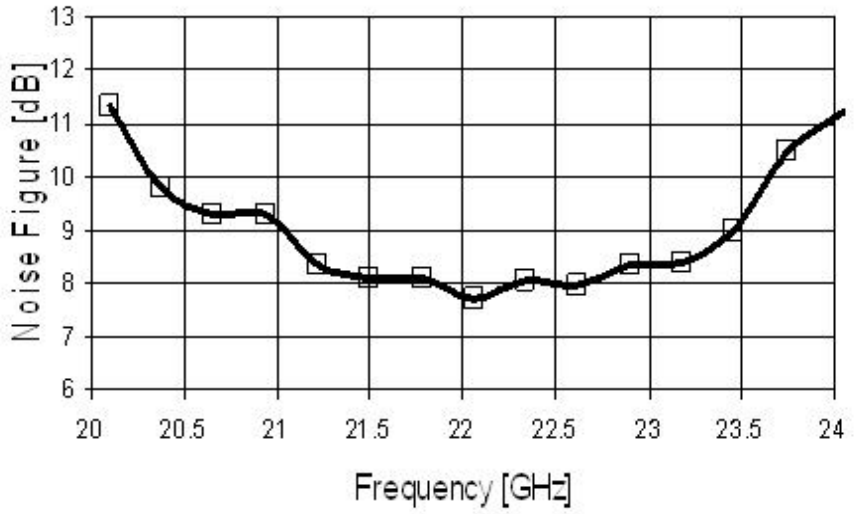


Figure 3.18: Overall noise figure of the front-end

Parameters	Measured Performance
S_{11}	-21dB
S_{22}	-10dB
Frequency of Maximum Gain	21.8 GHz
Maximum Power Gain	27.5dB
Maximum Voltage Gain	35.7dB
LNA Power Gain	15dB
LNA Noise Figure	6.0 dB
Overall Noise Figure	7.7dB
LNA Current Consumption	16mA
Mixer Current Consumption	4mA
Overall -1dB Compression Point	-23dBm
Image Rejection	31dB
Supply Voltage	1.5 V
Die Area	0.8 x 0.9 mm^2(0.4 x 0.5mm^2 core)

Table 3.1: Summary of the measurement performance of the 24-GHz CMOS front-end

Author	Year	Tech-nology	Topo-logy	Center Frequency	Power Gain	Noise Figure	Current Consumption
B. Floyd *et al* [77]	2001	0.18µm CMOS	CS	14.4GHz	21dB	8dB	18.6mA
B. Floyd *et al* [83]	2002	0.1µm CMOS SOI	CG	24GHz	7.6dB	10dB	53mA
This Work	2002	0.18µm CMOS	CGRF	22GHz	15dB	6dB	16mA
K. Yu *et al* [84]	2004	0.18µm CMOS	CS	24GHz	12.9dB	5.6dB	30mA

Table 3.2 LNA performance comparison

b

3.5 Chapter Summary

In this chapter, a novel LNA topology, common-gate with resistive feedthrough, is introduced. A detailed analysis of this topology based on classic noisy linear two-port theory and high-frequency MOSFETs model is given. The equations for its gain, noise parameters, noise figure, and the lowest noise figure at perfect power matching are derived.

By introducing a feedthrough resistor R_f much bigger than $1/g_m$ between the drain and source of the transistor, $F_{\min}$ and R_n of the common-gate amplifier reduce slightly and G_o and B_o change little. Based on this observation, an optimization procedure is devised to achieve noise matching and a fairly low input return loss simulaneously. Obliged to power matching, the gain of CGRF stage is independent of R_f and g_m. The benefit of a lower noise figure comes with a price of higher current consumption.

The CGRF topology and inductively degenerated common-source topology are compared based on analytical results. It has been illustrated that the GGRF stage can achieve a considerably lower noise figure at ultra high frequencies where low noise should be the primary consideration prior to power consumption.

The first 24-GHz CMOS front-end has been implemented. The CGRF topology is employed in the LNA input stage. The LNA-plus-mixer combination achieves a total power gain of 27.5dB and an overall noise figure of 7.7dB. The LNA achieves a 6dB noise figure and 15dB power gain, while consuming 16mA from a 1.5V power supply. The LNA performance corresponds well to the theoretical prediction. The LNA performance is superior to the previously reported CMOS LNA operating above 15GHz. This work demonstrates that CMOS technology is a viable candidate for building fully-integrated receivers at frequencies higher than 20 GHz.

Appendix 3.1: Derivation of (3.54) to (3.59)

By neglecting nondominant noise sources $i_{n,f}$ and i_{n,R_L}, the noise factor of CGRF LNA can be expressed as,

$$F_{CGRF} \approx 1 + \frac{S_{out,i_{n,d}}(\omega_0)}{S_{out,i_{n,R_s}}(\omega_0)} + \frac{S_{out,i_{n,g}}(\omega_0)}{S_{out,i_{n,R_s}}(\omega_0)} \tag{3.A.1.1}$$

where $S_{out,source}(\omega_0)$ denotes the power spectral density of the output noise current flowing through R_L caused by the referred noise source.

From nodal analysis of the circuits in Figure3.7, we can draw that the input matching condition is

$$R_s = \left(\frac{1 + g_m R_f (1+\chi)}{R_f + R_L} + \eta(\omega_0) g_m \right)^{-1} \tag{3.A.1.2}$$

where $\eta(\omega_0)$ is defined in (3.49). Assuming $g_m R_f >> 1$, by reformatting (3.A.1.2) we obtain the following useful expression,

$$\frac{R_f + R_L}{R_f} = (1+\chi) \frac{g_m R_S}{1 - g_m R_s \eta(\omega_0)} \tag{3.A.1.3}$$

The output noise current flowing through R_L produced by $i_{n,d}$ can be expressed as

$$i_{out,i_{n,d}} = i_{n,d} \left(1 + \frac{g_m (1+\chi)}{\frac{1}{R_s} + \eta(\omega_0) g_m} + \frac{R_L + R_s \,||\, (\eta(\omega_0) g_m)^{-1}}{R_f} \right)^{-1} \tag{3.A.1.4}$$

$$\approx i_{n,d} \left(\frac{g_m (1+\chi)}{\frac{1}{R_s} + \eta(\omega_0) g_m} + \frac{R_f + R_L}{R_f} \right)^{-1} \tag{3.A.1.5}$$

Substituting (3.A.1.3) into .(3.A.1.5), $i_{out,i_{n,d}}$ can be re-expressed in terms of $g_m R_s$ as

$$i_{out,i_{n,d}} = i_{n,d} \frac{1 - (g_m R_s \eta(\omega_0))^2}{2 g_m R_S (1+\chi)} \tag{3.A.1.6}$$

If input is matched, the effective transconductance G_m of the stage is given by

$$G_m = \frac{1}{2R_s}\frac{\frac{1}{R_s} - g_g}{\frac{1}{R_s}} = \frac{1}{2R_s}\left(1 - g_m R_s \eta(\omega_0)\right) \tag{3.A.1.7}$$

The output noise current contributed by i_{n,R_s} is given by

$$i_{out,i_{n,R_s}} = i_{n,R_s} R_s G_m = \frac{i_{n,R_s}}{2}(1 - g_m R_S \eta(\omega_0)) \tag{3.A.1.8}$$

therefore

$$\frac{S_{out,i_{n,d}}(\omega_0)}{S_{out,i_{n,R_s}}(\omega_0)} = \frac{4kT\gamma\frac{g_m}{\alpha}\left[\frac{1-(g_m R_s \eta(\omega_0))^2}{2g_m R_s(1+\chi)}\right]^2}{\frac{4kT}{R_s}\left[\frac{1-g_m R_s \eta(\omega_0)}{2}\right]^2} \tag{3.A.1.9}$$

$$= \frac{\gamma}{\alpha}\frac{1}{(1+\chi)^2}\left(\frac{1}{g_m R_s} + \eta^2(\omega_0) g_m R_s + 2\eta(\omega_0)\right) \tag{3.A.1.10}$$

In Figure 3.7, $i_{n,g}$ and i_{n,R_s} are applied between source and ground in parallel, therefore

$$\frac{S_{out,i_{n,g}}(\omega_0)}{S_{out,i_{n,R_s}}(\omega_0)} = \frac{S_{i_{n,d}}(\omega_0)}{S_{i_{n,R_s}}(\omega_0)} = \frac{4KT\delta\eta(\omega_0) g_m}{\frac{4KT}{R_s}} = \delta\eta(\omega_0) g_m R_s \tag{3.A.1.11}$$

Substituting (3.A.1.10) and (3.A.1.11) into (3.A.1.1), we obtain

$$F_{CGRF} = 1 + \frac{\gamma}{\alpha}\frac{1}{(1+\chi)^2}\left(\frac{1}{g_m R_s} + \eta^2(\omega_0) g_m R_s + 2\eta(\omega_0)\right) + \delta\eta(\omega_0) g_m R_s \tag{3.A.1.12}$$

Equating the derivative of (3.A.1.12) in terms of g_m to zero, an optimum g_m is solved for minimum F_{CGRF}, i.e.,

$$g_{m,CGRF,opt} = \frac{1}{R_s}\left(\frac{\delta\alpha}{\gamma}(1+\chi)^2\eta(\omega_0) + \eta^2(\omega_0)\right)^{-\frac{1}{2}} \tag{3.A.1.13}$$

and the minimum noise factor is given by

$$F_{\min,CGRF} = 1 + \frac{\gamma}{\alpha}\frac{2}{1+\chi}\eta(\omega_0) + \frac{2}{1+\chi}\sqrt{\frac{\gamma\delta}{\alpha}\eta(\omega_0) + \left(\frac{\gamma}{\alpha}\right)^2\left(\frac{1}{1+\chi}\right)^2\eta^2(\omega_0)} \tag{3.A.1.14}$$

For traditional common-gate LNA, where $R_f = \infty$, at input matching condition

$$g_m R_s = \left(1 + \chi + \eta(\omega_0)\right)^{-1} \tag{3.A.1.15}$$

By substituting (3.A.1.15) into (3.A.1.7), we obtain that the G_m of conventional common-gate LNA is given by

$$G_{m,CG} = \frac{1}{2R_s}\left(1+\frac{\eta(\omega_0)}{(1+\chi)}\right)^{-1} \tag{3.A.1.16}$$

Substituting (3.A.1.15) into (3.A.1.12) and performing some simple mathematical manipulation, we obtain the noise factor expression of the conventional common-gate LNA, i.e.,

$$F_{CG} = 1+\frac{\gamma}{\alpha}\frac{1}{1+\chi}\frac{\left(1+\chi+2\eta(\omega_0)\right)^2}{1+\chi+\eta(\omega_0)}+\delta\frac{\eta(\omega_0)}{1+\chi+\eta(\omega_0)} \tag{3.A.1.17}$$

Appendix 3.2: Impacts of the Feedthrough Resistor on the Performance of a Common-Gate Amplifier in Terms of NF, Gain, S_{11} and their Tradeoff

When a LNA is used in a cascade system, both the LNA noise figure and gain affect the system noise performance. Therefore, a minimum noise figure of LNA does not necessarily result in the lowest system noise. The noise measure M is used as a figure of merit for the LNA performance by taking the tradeoff between gain and noise figure into account, which is defined as [82]

$$M = \frac{F-1}{1-1/G_A} \tag{3.A.2.1}$$

where F is the noise factor and G_A is the available power gain. The impacts of the feedthough resistor in Figure 3.6 on a common-gate amplifier (not necessarily matched) noise figure, gain, and their tradeoff are discussed in this appendix.

For first order analysis, we assume the channel thermal noise and the thermal noise of R_f are the dominant noise sources in the CGRF LNA, as shown in Figure 3.6. We also assume that $g_m = g_{d0}$ and $g_{mb} = 0$. When R_f is set to an arbitrary value, the LNA is not necessarily matched and its noise factor at the operation frequency can be expressed as

$$F = 1 + \frac{(1+\gamma g_m R_f)R_f}{(1+g_m R_f)^2 R_s} \tag{3.A.2.2}$$

In (3.A.2.2), G_A is defined as the ratio between the available power from the amplifier outputs and the available power from the source. However, using the definition of G_A is confusing in this case because conjugate matching can not be achieved in the CGRF architecture. Therefore, we use the definition of transducer gain G_T instead, which is referred to as the ratio between the effectively delivered power to the load (R_L) and the power obtained from the source. For CGRF LNA, G_T is given by

$$G_T = \frac{R_L\left(1+g_m R_f\right)}{R_f + R_L} \tag{3.A.2.3}$$

We define a figure of merit

$$M_o = \frac{F-1}{1-1/G_T} \tag{3.A.2.4}$$

as a measure for the tradeoff between noise and gain performance. In addition, the S_{11} of the CGRF stage is given by

$$S_{11} = \frac{R_f + R_L - R_s - g_m R_s R_f}{R_f + R_L + R_s + g_m R_s R_f} \tag{3.A.2.5}$$

The impacts of R_f on F, G_T, M_o, and S_{11} are investigated using a typical numerical example, where g_m=80mS, R_s=50Ω, and R_L=500Ω. To show the tradeoffs involved in various devices, both $\gamma = 2/3$ and $\gamma = 2$ are used to calculate the noise performance. The F, G_T, M_o, and S_{11} are plotted against R_f, as shown in Figure (3.A.2.1) to (3.A.2.4) respectively. It can be seen that reducing R_f decreases both F and G_T. However, whether M_o increases or decreases with R_f depends on the transistor noise properties. The feedthrough resistor provides a way to adjust input impedance with little impact on M_o, allowing for optimizing the transistor for best noise performance under power matching constraints.

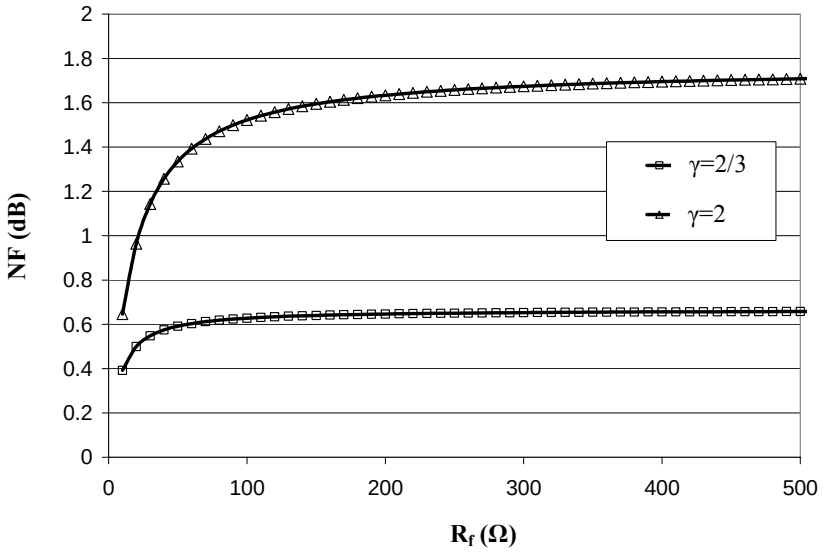


Figure 3.A.2.1: The NF of the CGRF LNA as a function of R_f (g_m=80mS, R_s=50Ω, and R_L=500Ω)

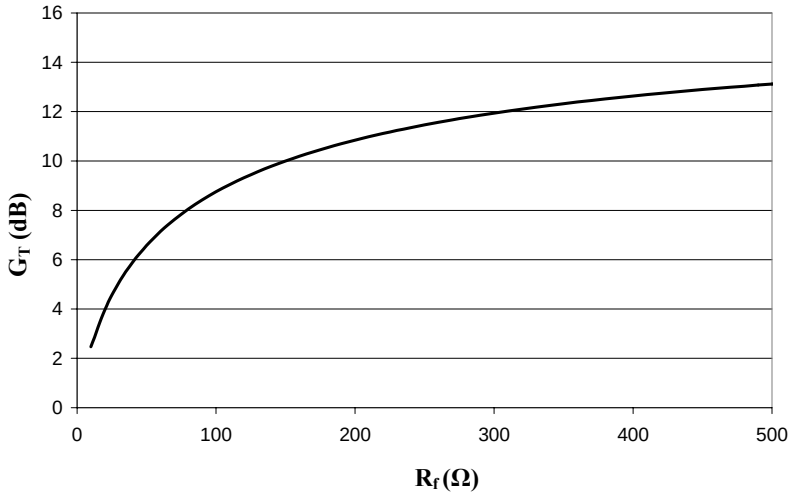


Figure 3.A.2.2: The G_T of the CGRF LNA as a function of R_f (g_m=80mS, R_s=50Ω, and R_L=500Ω)

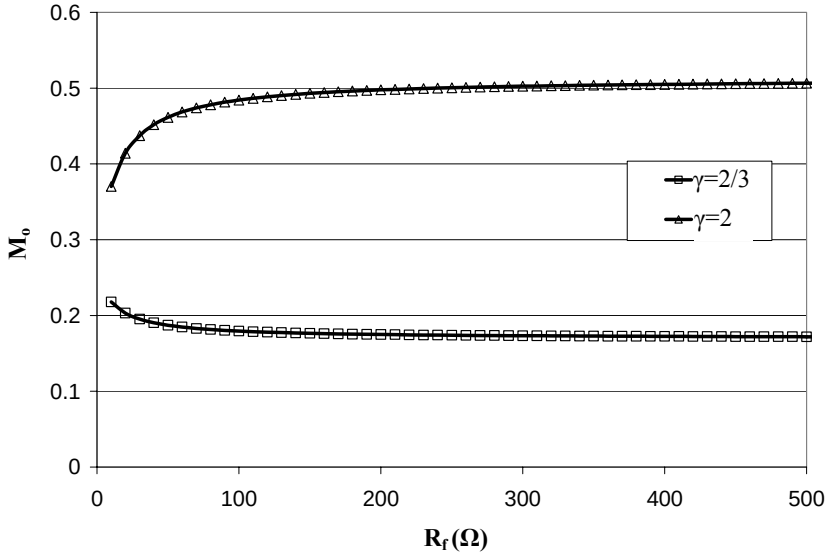


Figure 3.A.2.3: The G_T of the CGRF LNA as a function of R_f (g_m=80mS, R_s=50Ω, and R_L=500Ω)

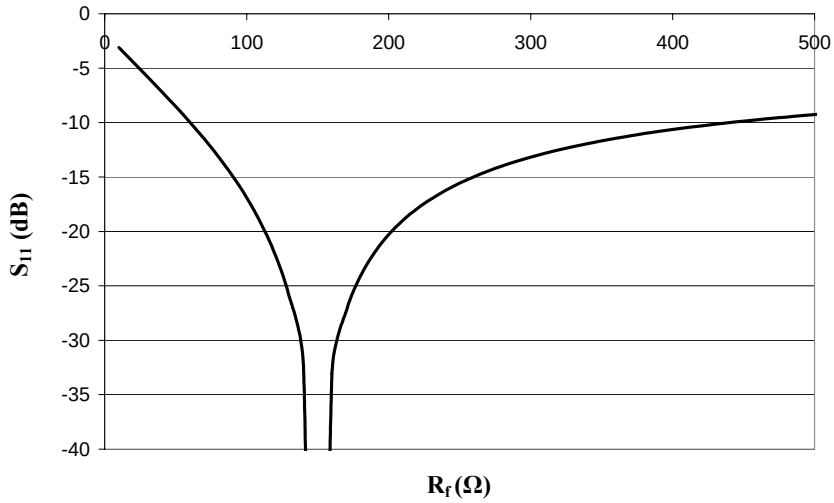


Figure 3.A.2.4: The S_{11} of the CGRF LNA as a function of R_f (g_m=80mS, R_s=50Ω, and R_L=500Ω)

Chapter 4

A Fully-Integrated 8-element 24-GHz Phased-Array Receiver in Silicon[1]

After demonstrating a 24-GHz CMOS front-end with good performance, a fully-integrated 24-GHz phased array receiver in silicon-based technologies was in sight. This chapter presents the first fully-integrated 24-GHz phased array receiver in silicon. Although integration in CMOS is feasible, we attempt to use SiGe process in this work mainly for low power considerations. Section 4.1 introduces the system-level architecture and considerations. The implementation issues of the signal path, the LO phase generation, and the phase distribution are covered by subsequent sections 4.2 to 4.4. The experimental results are shown and discussed in Section 4.5. Finally, a chapter summary is given in Section 4.6.

4.1 System Architecture

LO phase shifting architecture is adopted in this work because the receiver is less sensitive to the amplitude variations at the LO port of the mixer, circumventing the lossy and noisy RF phase shifters at signal path. To avoid problems involved in direct-conversion architectures such as large DC offset and flicker noise, a two-step downconversion heterodyne architecture is employed.

4.1.1 Top Level Block Diagram

Figure 4.1 shows the block diagram of the 24-GHz 8-element phased-array receiver [85][86]. The receiver uses a two-step downconversion with an IF of 4.8GHz, allowing both LO frequencies to be generated using a single synthesizer loop. A single oscillator core generates 16 discrete phases providing 4-bits (22.5^{o}) of raw phase resolution. A set of 8 phase-selectors (i.e., analog phase multiplexer) apply the

[1] The 24-GHz phased-array receiver is a joint work done by Xiang Guan and Hossein Hashemi. The VCO, frequency divider chain and phase selectors were designed entirely by Hossein Hashemi.

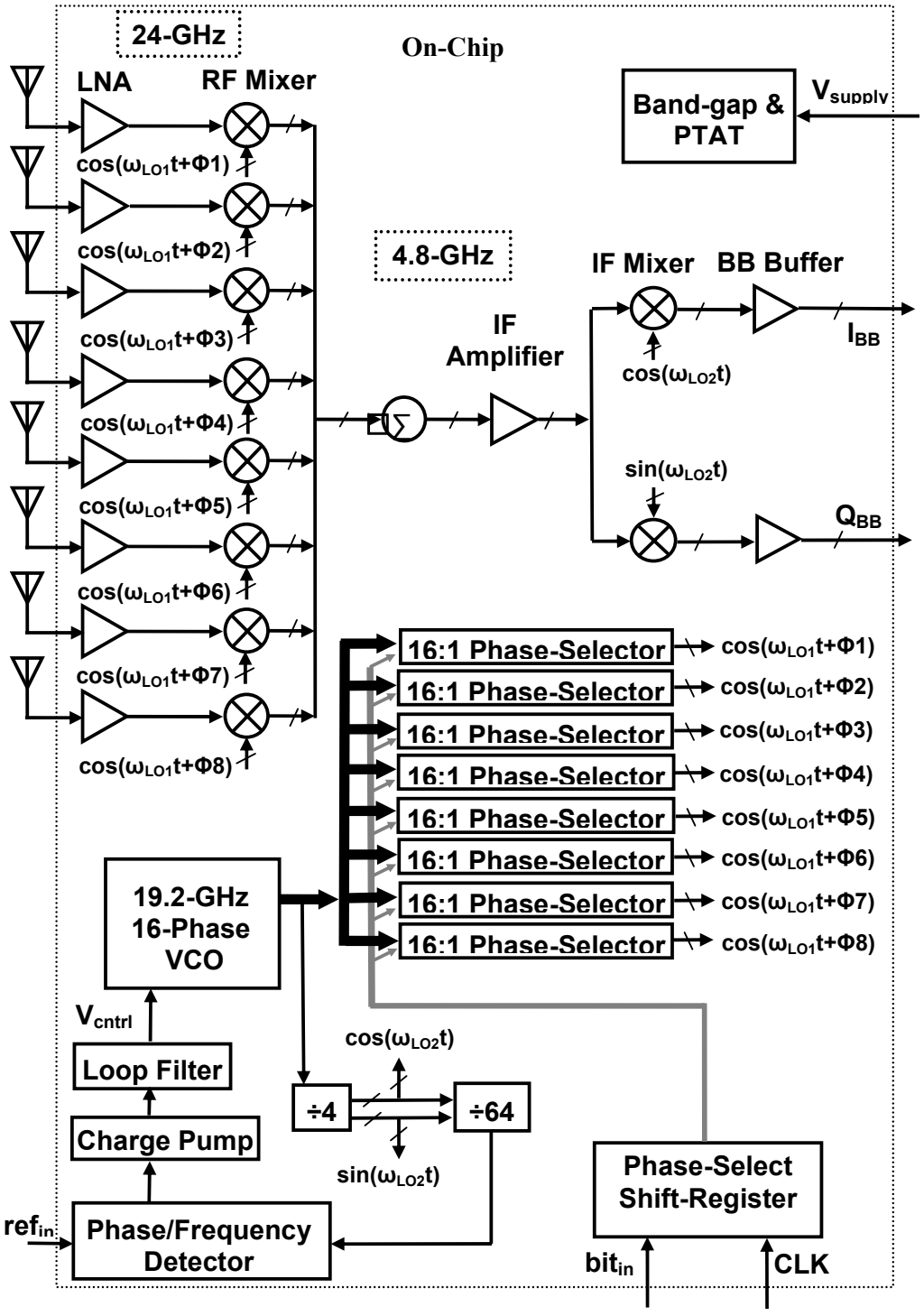


Figure 4.1 System Architecutre

appropriate phase of the LO to the corresponding RF mixer for each path independently. The operating state of the chip including phase-selection information (beam-steering angle) is serially loaded into an on-chip shift-register using a standard serial interface. The image at 14.4GHz is attenuated by the front-end's narrowband transfer function, *i.e.*, antenna and LNA.

Each of the eight RF front-ends consists of two inductively degenerated common-emitter LNA stages followed by a double-balanced Gilbert-type mixer. The input of the first LNA is matched to 50Ω and the subsequent blocks of the front-end are power matched for maximum power transfer. The output of all eight mixers are combined in current domain and terminated to a tuned load at the IF. The combined signal is further amplified by an IF amplifier and downconverted to baseband by a pair of double-balanced Gilbert-type mixers driven by *I* and *Q* signals generated by the divide-by-4 block. Two baseband differential buffers drive the I and Q outputs. On-chip proportional-to-absolute-temperature (PTAT) and band-gap references generate the bias currents and voltages, respectively.

4.1.2 Array Pattern

The simulated 16 corresponding array patterns are shown in Figure 4.2, for omni directional antenna elements with a spacing of λ/2. Figure 4.2 illustrates that the system is capable of steering the beam from -90^{o} to $+90^{o}$ and a steering step size of 7.2^{o} at the normal direction. It can be noticed that the beam width and steering step is minimum at the broadside and maximum when the beam is steered to $\pm 90^{o}$.

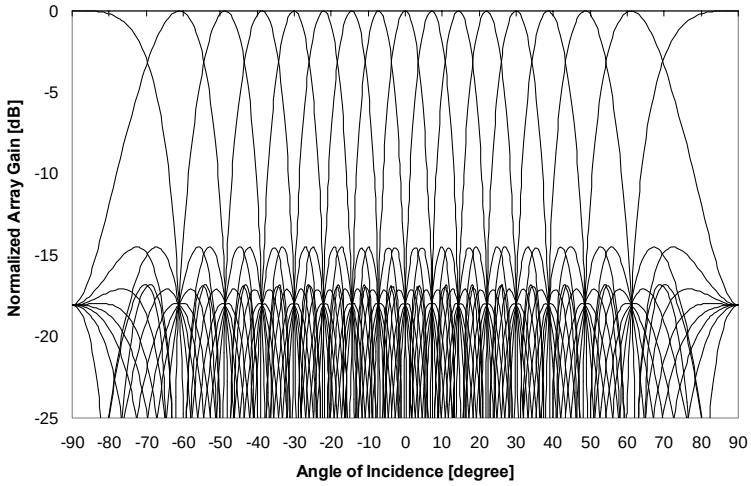


Figure 4.2: Array patterns of 16 different LO phase setttings

4.2 Signal Path

The signal path of the receiver is comprised of LNAs, RF mixers, a signal-combining structure, an IF amplifier, IF quadrature mixers, baseband amplifiers, output buffers, and bias references. Since there is no well-defined wireless communication standard at the 24-GHz band so far, the primary target of this design is to prove the concept. A "good-performance receiver" is defined as one providing a comparable single-path noise figure, linearity, and image rejection to those silicon-based systems operating at a low-GHz range, despite the fact that the benefits of array can actually relax those specifications.

4.2.1 A 24-GHz SiGe Low Noise Amplifer

As we discussed in Chapter 3, the choice of the LNA topology depends on the operation frequency, process, and power budget. When ω_o / ω_T is small, the inductively degenerated common-source LNA can achieve a sufficiently low noise figure with reasonable power dissipation. Although the CGRF stage can achieve a lower noise figure, the improvement is tiny at the ω_o / ω_T ratio and the bias current for optimum noise is high. In this process, the peak f_T of SiGe heterojunction bipolar transistor (HBT) is 120GHz, which is much higher than 24GHz. Therefore, inductively degenerated common-source topology is adopted in this work. As in Chapter 3, we begin the discussion with the transistor noise model and noise parameters and then show the design procedure to achieve simultaneous power and noise matching.

4.2.1.1 Noise Model of SiGe Heterojunction Bipolar Transistor

The small signal and noise equivalent circuit of a SiGe HBT device is similar to that of the traditional silicon-based BJT, as shown in Figure 4.3. Primary noise sources in a SiGe HBT include the collector shot noise $\overline{i_{cn}^2}$, the base shot noise $\overline{i_{bn}^2}$, the base resistor thermal noise $\overline{V_{bn}^2}$, and the emitter resistor thermal noise $\overline{V_{en}^2}$. The mean square value of those noise sources can be expressed by the following equations, respectively:

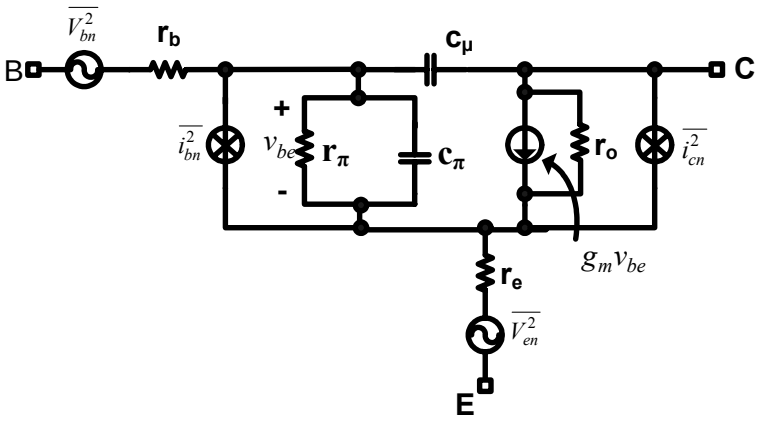


Figure 4.3: Small-signal and noise equivalent circuits of SiGe HBT

$$\overline{i_{cn}^2} = 2qI_c\Delta f \tag{4.1}$$

$$\overline{i_{bn}^2} = 2qI_b\Delta f \tag{4.2}$$

$$\overline{v_{bn}^2} = 4kTr_b\Delta f \tag{4.3}$$

$$\overline{v_{en}^2} = 4kTr_e\Delta f \tag{4.4}$$

where I_c and I_b are the collector and base dc current, respectively, r_b and r_e are the parasitic terminal resistance at the base and the emitter, respectively, and Δf is the bandwidth.

4.2.1.2 Noise Parameters of HBT

The parameters of an HBT in a designer's choice are its lateral dimension and bias current. It is important to understand how the transistor noise performance change with the design parameters.

In Figure 4.3, the i_{bn} and i_{cn} can be assumed uncorrelated up to the frequencies approaching $f_T/2$ [87]. In this case the noise parameters of the transistor can be approximately expressed by the following equations [88][89]:

$$R_n \cong \frac{n^2 V_T}{2I_C} + (r_b + r_e) \tag{4.5}$$

$$\mathrm{Re}[Z_{opt}] \cong R_n \times \frac{\sqrt{\frac{I_C}{2V_T}(r_b + r_e)(\frac{f^2}{f_T^2} + \frac{1}{\beta_0}) + \frac{n^2}{4\beta_0}}}{\frac{I_C}{2V_T}(r_b + r_e)(\frac{f^2}{f_T^2} + \frac{1}{\beta_0}) + \frac{n^2}{4}(\frac{f^2}{f_T^2} + \frac{1}{\beta_0})} \tag{4.6}$$

$$\mathrm{Im}[Z_{opt}] = R_n \times \frac{n/4}{\frac{I_C}{2V_T}(r_b + r_e)(\frac{f^2}{f_T^2} + \frac{1}{\beta_0}) + \frac{n^2}{4}(\frac{f^2}{f_T^2} + \frac{1}{\beta_0})} \tag{4.7}$$

$$F_{\min} \approx 1 + \frac{n}{\beta_o} + \sqrt{\frac{2I_C}{V_T}(r_b + r_e)(\frac{f^2}{f_T^2} + \frac{1}{\beta_0}) + \frac{n^2}{\beta_0}} \tag{4.8}$$

where I_C is the dc bias current, r_b and r_e are the ohmic resistance at base and emitter, respectively, β_0 is the dc current gain, f is the operation frequency, and n is the collector current ideality factor approximately equal to 1.

All noise parameters vary nonlinearly with the emitter width w_e. [88]. It has been verified that the $F_{\min}$ increases with w_e, and hence a transistor of minimum emitter width is desired. With a fixed device width and neglecting fringe effects, I_c and r_b+r_e can be expressed as

$$I_C = J_C \times l_e \tag{4.9}$$

$$r_b + r_e = (r_b + r_e)_u / l_e \tag{4.10}$$

where l_e is the emitter length, J_c is the dc current per unit l_e, and $(r_b+r_e)_u$ is the sum of base and emitter resistance per unit l_e. β_0 , n, and $(r_b + r_e)_u$ can be considered constants as a function of J_c and w_e. Therefore, Equation (4.5) ~ (4.8) indicate that R_n, Re[Z_{opt}], and Im[Z_{opt}] scale linearly with the inverse of l_e, while $F_{\min}$ stays constant to the first order with fixed current density and emitter width.

4.2.1.3 Input Stage Design Procedure

A well known procedure for bipolar LNA design is used to achieve optimum noise matching and power matching simultaneously, which is detailed below with emphasis on its differences with CMOS LNA design [88][89]

1. Determine J_c associated with the lowest $F_{\min}$. The r_b of a bipolar transistor is much bigger than the r_g of MOSFET. The thermal noise of r_b dominants at a low J_c. On the other hand, shot noise prevails at a high J_c. This interaction results in

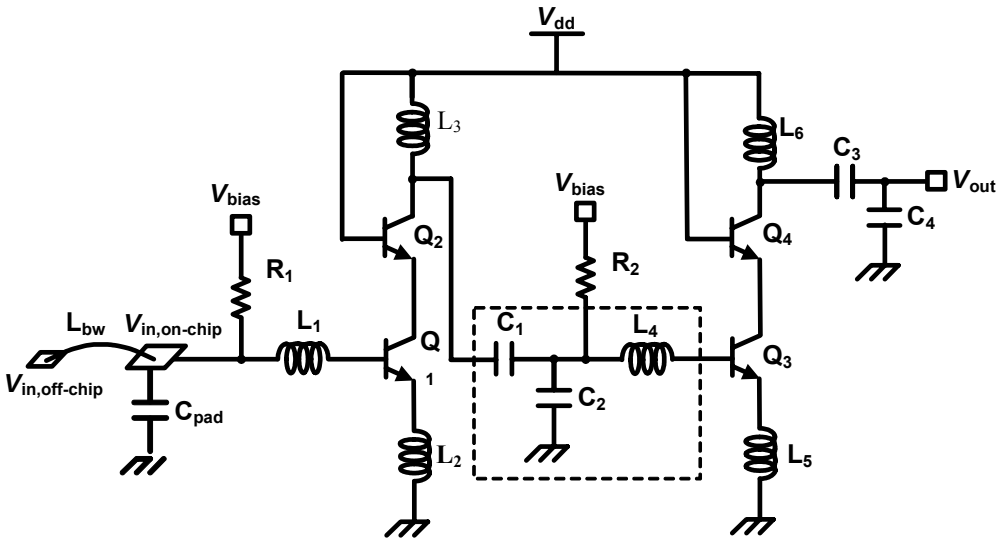


Figure 4.4: A 2-stage 24-GHz LNA

the optimum J_c of a bipolar device smaller than its value at peak f_T. In contrast, the optimum J_c of MOSFET usually corresponds to its peak f_T.

2. The emitter length is adjusted with a constant J_c until the optimum source resistance R_o is 50Ω at the operating frequency.

3. The emitter inductor L_E is added to match the real part of the input impedance Z_{in} to 50Ω.

4. The inductor L_B is added at the base to resonate out the reactance seen into the base of the transistor. Unlike MOSFET, the noise sources in a bipolar transistor can be treated as uncorrelated ones. It can be shown that L_B not only neutralizes the input reactance, but also brings the optimum noise reactance to zero. Therefore, noise matching and power matching are achieved simultaneously.

4.2.1.4 LNA Implementation

The schematic of the two-stage 24-GHz LNA is shown in Figure 4.4. The optimization results in a 4mA dc current for each stage and an emitter degeneration inductance of 0.2nH. The cascode transistor Q_2 is used to improve reverse isolation. At 24GHz the load inductance L_3 and, thereby, the achievable gain of a single stage are limited by the large collector-substrate capacitance of Q_2. Simulation results show that the power gain achievable by a single stage is not sufficient to suppress the noise

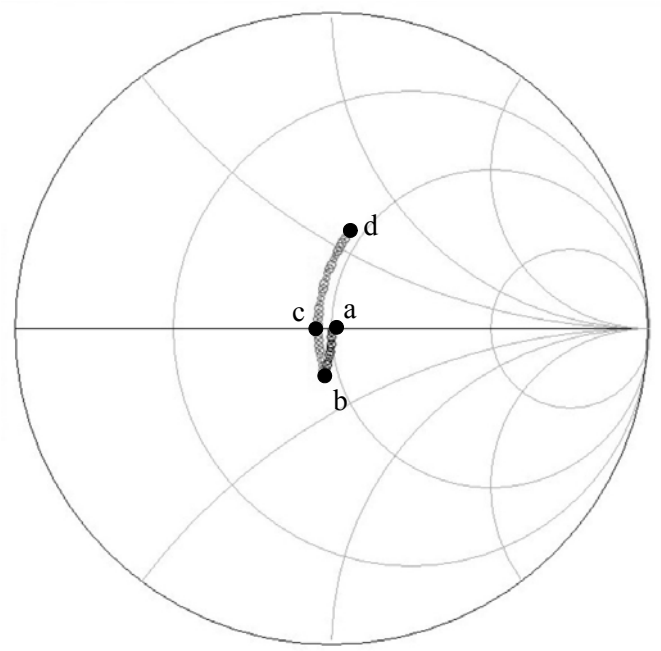


Figure 4.5: Effects of bond pad and bond wire to LNA input impedance

of the subsequent mixer, and so an identical second stage is added to enhance the signal level.

At 24GHz, the input pad and bond wire have considerable effects on the input reflection coefficient of the LNA. The LNA is designed to be well-matched to 50 Ω (S_{11} less than -10dB) looking into L_1. The smith chart in Figure 4.5 illustrates the variation of Z_{in} by the bond pad capacitance C_{pad} and bond wire inductance L_{bw}. EM simulations show that the capacitance of a 75µm x 75µm pad with Metal 1 shielding is around 40fF. As C_{pad} increases from 0 to 40fF, Z_{in} moves from point *a* to *b* along the curve in Figure 4.5, corresponding to a S_{11} of -16dB. The bond wire inductor pulls Z_{in} from *b* towards *d*. Point *c* corresponds to an L_{bw} of 0.1nH, where an optimum S_{11} of -25dB is achieved. When L_{bw} is further increased to 0.3nH (point *d*), S_{11} reaches -10dB. Therefore, 0.3nH is the maximum bond wire inductance that can be tolerated by the specification.

The V_{dd} and ground lines of the LNA are bypassed on a chip with a metal-insulator-metal (MIM) capacitor resonating at 24GHz to realize a low impedance supply. All the inductors used in this LNA are between 0.2nH ~ 0.5nH. To save the silicon area, spiral inductors are used, although slab inductors provide higher quality

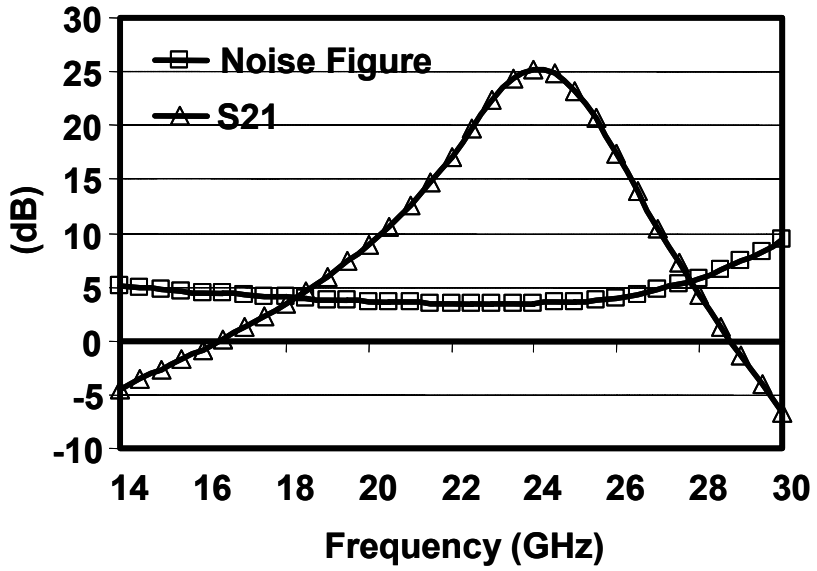


Figure 4.6: LNA simulation results

factors. All spirals and interconnections are modeled by electromagnetic simulations using IE3D.

The RF input pad is shielded by Metal 1 to minimize substrate loss [81]. The size of the RF input pad is 75µm x 75µm, which is smaller than the other pads to reduce the parasitic capacitance.

Simulation results as shown in Figure 4.6 indicates that an approximately 25dB gain and a noise figure of less than 5dB can be expected from this 2-stage LNA.

4.2.1.5 Impedance Matching Network

The impedance matching network is widely used in discrete microwave systems to maximize signal power transfer. However, it is rarely employed between on-chip blocks operating at low GHz range due to a large area cost and signal loss caused by additional inductors with low quality factors. The current IC technologies provide thick top metal for implementing on-chip inductors with relative high quality factors. The required inductor value as well as size reduce with the increase of frequency. Therefore, on-chip matching becomes a plausible technique at the 24-GHz range. In this work, T networks are used between two LNA stages as well as LNA and a mixer to maximize signal power transfer, as shown in Figure 4.4.

At very high frequencies such as 24GHz, the interactions between blocks are enormous. Any change in the following blocks may change the gain, center frequency, and even input impedance of the preceding blocks, making a more complex design process and possibly leading to sub-optimal results. This problem is mitigated by using impedance matching. When one block changes, we only need to accordingly change the impedance matching network associated with it so that the performance of the adjacent blocks won't be affected, ensuring each block can be designed and optimized independently. Furthermore, the optimization process is also eased by absorbing the effects of long wires interconnecting blocks into the matching network.

Since the first stage is optimized for low noise, the same design is used as the second stage. A capacitive divider of C_1 and C_2 transforms the output impedance of the first stage to 50Ω, which is also the optimum impedance for second stage in terms of power and noise. The capacitance of C_1 and C_2 are chosen to be 80fF and 160fF, respectively, as a trade-off between large load inductance and accuracy. L_4 has an inductance of 0.2nH and occupies a 50μm x 50um silicon area. A first order estimation of the loss through this impedance matching network is given by

$$loss \approx \frac{\omega_o L_4}{Q_{L_4} R_{in}} \tag{4.11}$$

EM simulations show a Q_{L4} of 15. The loss at 24GHz and 50Ω input is calculated to be 0.17dB. Circuit simulation shows a signal loss of 0.25dB through this network. The additional loss is caused by the other interconnection wires and imperfect matching.

Alternatively, if we couple the first stage directly to the second stage without using a matching network, a capacitive reactance at the input of the second stage will significantly off-tune the first stage. The inductance of L_3 needs to be reduced from 0.45nH to 0.25nH to adjust the center frequency back to 24GHz, therefore reducing the overall gain by 7dB and increasing the noise figure by roughly 1dB.

4.2.2 A 24GHz Downconverter and IF Combining Structure

Compared to a low noise amplifier, a mixer usually has a higher noise figure due to noise contribution from the switching cells. Meanwhile, the mixer needs to operate linearly for a larger input swing. In our IF-combining phased array architecture, the

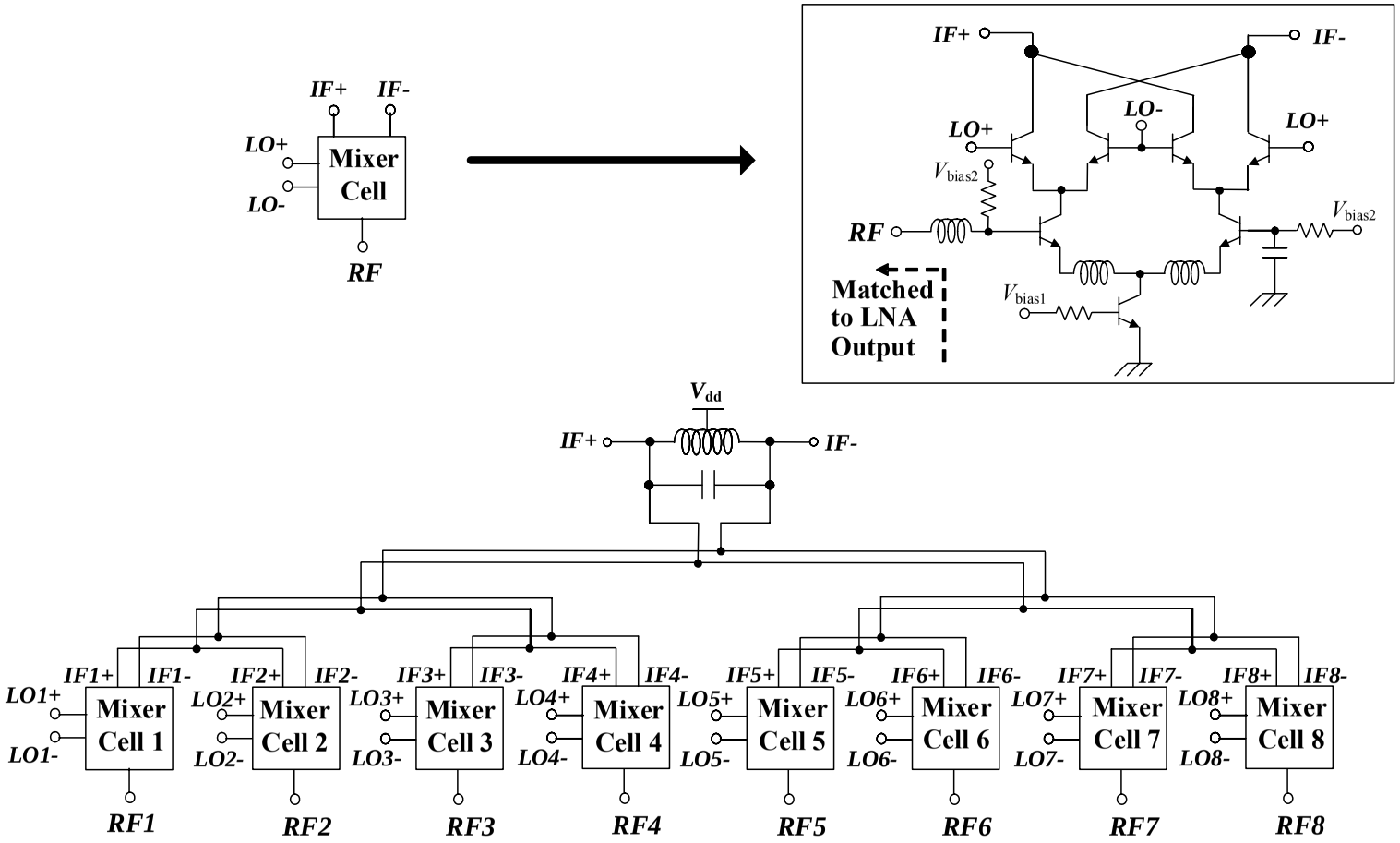


Figure 4.7: RF mixer and IF signal combining

power budget of the mixer is especially stringent because multiple mixers need to operate simultaneously.

Although the LNA architecture is single-ended as we discussed in the last section, it is advantageous to build the remaining circuits differentially to suppress the common-mode noise coupled from power supply, substrate, and adjacent passive components and cancel the even-order harmonics generated in each branch.

Gilbert-type double-balanced multipliers are used to downconvert the single-ended 24-GHz RF signal to a differential signal at 4.8 GHz, as shown in Figure 4.7. The input of the mixer is power matched to the LNA output through an impedance-transforming network. Inductive emitter degeneration is used to improve mixer linearity. To convert single-ended signals to differential ones, one branch of the input differential pair is bypassed to ac ground by a large on-chip capacitor.

A dc bias current of 1.25 mA is chosen for each mixing cell as a reasonable tradeoff between power dissipation, linearity, and noise figure. Simulation shows that each mixing cell achieves a conversion transconductance of 6.5mA/V. The downconverted IF signal is combined in current domain through a symmetric binary tree and terminated to a tuned load at 4.8 GHz, as shown in Figure 4.7.

Simulation shows each mixer cell exhibits a noise figure of 11.3dB. With a two-stage LNA gain of more than 20dB, the noise figure of the front-end is dominant by LNA noise.

The binary tree structure acts as a current combiner. The total geometric length from each input port to the output port is roughly 1.5mm. At 4.8GHz microwave network and transmission line theories must be applied to analyze this tree structure. The current excitation $i_1 \sim i_8$ can be decomposed into 4 modes as shown below

$$i_1 = \frac{i_1 - i_2}{2} + \frac{\sum_{k=1}^{2} i_k - \sum_{k=3}^{4} i_k}{4} + \frac{\sum_{k=1}^{4} i_k - \sum_{k=5}^{8} i_k}{8} + \frac{\sum_{k=1}^{8} i_k}{8} \tag{4.12}$$

$$i_2 = \frac{i_2 - i_1}{2} + \frac{\sum_{k=1}^{2} i_k - \sum_{k=3}^{4} i_k}{4} + \frac{\sum_{k=1}^{4} i_k - \sum_{k=5}^{8} i_k}{8} + \frac{\sum_{k=1}^{8} i_k}{8} \tag{4.13}$$

$$i_3 = \frac{i_3 - i_4}{2} + \frac{\sum_{k=3}^{4} i_k - \sum_{k=1}^{2} i_k}{4} + \frac{\sum_{k=1}^{4} i_k - \sum_{k=5}^{8} i_k}{8} + \frac{\sum_{k=1}^{8} i_k}{8} \tag{4.14}$$

$$i_4 = \frac{i_4 - i_1}{2} + \frac{\sum_{k=3}^{4} i_k - \sum_{k=1}^{2} i_k}{4} + \frac{\sum_{k=1}^{4} i_k - \sum_{k=5}^{8} i_k}{8} + \frac{\sum_{k=1}^{8} i_k}{8} \tag{4.15}$$

$$i_5 = \frac{i_5 - i_6}{2} + \frac{\sum_{k=5}^{6} i_k - \sum_{k=7}^{8} i_k}{4} + \frac{\sum_{k=5}^{8} i_k - \sum_{k=1}^{4} i_k}{8} + \frac{\sum_{k=1}^{8} i_k}{8} \tag{4.16}$$

$$i_6 = \frac{i_6 - i_5}{2} + \frac{\sum_{k=5}^{6} i_k - \sum_{k=7}^{8} i_k}{4} + \frac{\sum_{k=5}^{8} i_k - \sum_{k=1}^{4} i_k}{8} + \frac{\sum_{k=1}^{8} i_k}{8} \tag{4.17}$$

$$i_7 = \frac{i_7 - i_8}{2} + \frac{\sum_{k=7}^{8} i_k - \sum_{k=5}^{6} i_k}{4} + \frac{\sum_{k=5}^{8} i_k - \sum_{k=1}^{4} i_k}{8} + \frac{\sum_{k=1}^{8} i_k}{8} \tag{4.18}$$

$$i_8 = \frac{i_8 - i_7}{2} + \frac{\sum_{k=7}^{8} i_k - \sum_{k=5}^{6} i_k}{4} + \frac{\sum_{k=5}^{8} i_k - \sum_{k=1}^{4} i_k}{8} + \frac{\sum_{k=1}^{8} i_k}{8} \tag{4.19}$$

Mode 1 is comprised of the 1st term in each equation from (4.12) to (4.19). Similarly, Mode 2 to 4 consists of 2nd, 3rd, and 4th terms in each equation, respectively. Mode 1, 2, and 3 are all odd modes producing zero output current due to the symmetry of the tree. Mode 4 is the only even mode where all input ports have identical excitations. In this mode, the symmetry of the tree ensures the isolation between input ports. The impedance matching at each T-junction is desired for maximum power transfer.

At first, the structure illustrated in Figure 4.8 was considered where the transmission line impedance is scaled down by a factor of 2 after each combination. Each input port of the tree is fed by the transistor drain current and thereby sees a high source impedance. Let us assume that the source impedance R_s is much higher than the input impedance of the tree so that $i_1 \sim i_8$ acts as an ideal current source. It is easy to prove that the impedance matching is achieved at each T-junction and if the transmission line thermal loss is negligible, the output voltage and power are given by

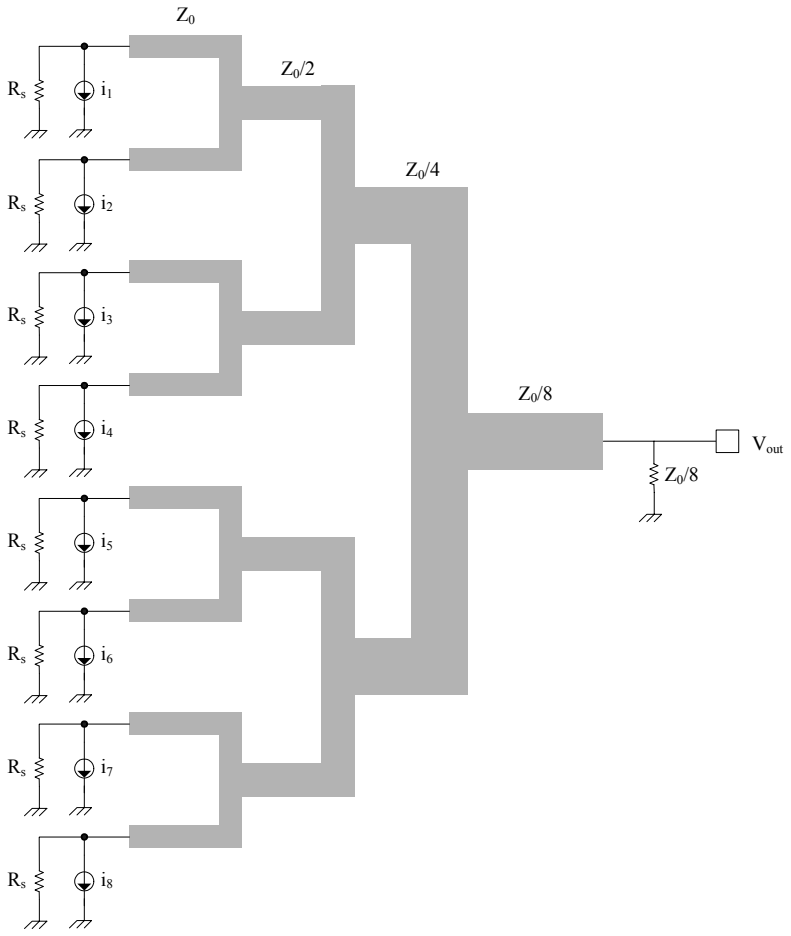


Figure 4.8: A passive current combining structure

$$V_{out} = \sum_{k=1}^{8} i_k \cdot Z_0 / 8 \tag{4.20}$$

$$P_{out} = \left(\sum_{k=1}^{8} i_k \right)^2 \cdot Z_0 / 8 \tag{4.21}$$

Unfortunately, the achievable transmission line impedance on chip in this process is less than 100Ω due to the limitation of the dielectric thickness and minimum width of the metal wire. Therefore, the voltage gain achieved via this combining structure is too small to qualify it as a valid candidate. Another difficulty is that in this structure the transmission line impedance has to reduce by half at each level. The achievable ratio of maximum and minimum transmission line impedance on chip is on the order of 10, so it is impractical to use it in phased array systems with more elements.

The Wilkinson type of power combiner [90] is also considered. The advantage of the Wilkinson power combiner is that characteristic impedance at all ports are identical and thus the transmission line impedance does not need to scale with levels. Hence, the signal combining tree can be extended to arbitrary number of elements. However, the Wilkinson power combiner requires a quarter-wavelength impedance transformer in each T-junction, which is too big to implement on-chip at 5-GHz frequency range.

In this work we use identical transmission lines in all levels of the structure and a large load resistance at the output. The network suffers from the power loss by reflection at each T-junction however, simulations incorporating the transmission line model show that a single-path down-converison gain higher than 3dB can still be achieved.

In Chapter 5 we will introduce an active signal combining structure which overcomes many of the above problems.

4.2.3. IF Circuitry

The IF amplifier is the first block after signal combining. The noise contribution of such blocks in overall noise figure is not only suppressed by the single-path gain of the front-end, but also by the array gain. The interference arriving at the input of the IF amplifier has been attenuated by the spatially selective array. Therefore, both noise

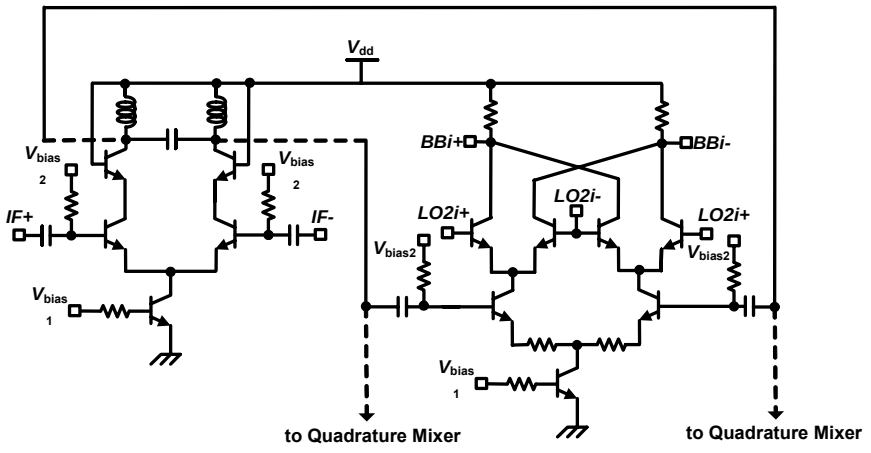


Figure 4.9: 4.8-GHz amplifier and mixer

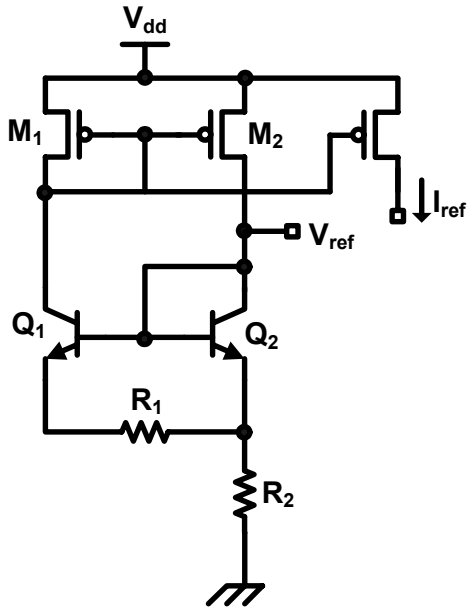


Figure 4.10: A bandgap and PTAT reference

and linearity requirements of the IF amplifier and subsequent blocks are relaxed.

The amplified IF signal is further downconverted to baseband using quadrature paths to recover the in-phase (I) and quadrature-phase (Q) component of the signal. The schematics of the IF amplifier and mixer are shown in Figure 4.9.

4.2.4 Bandgap and PTAT References

All current and voltage biases of the signal-path are regulated by the on-chip bias references. In this work, those references are generated by using the "bandgap" technique to accommodate temperature and supply variations.

The schematics of the bandgap references are presented in Figure 4.10. M_1 and M_2 form a current mirror defining the collector current ratio between Q_1 and Q_2. In this design we set the same current in two branches and denote it as I_C. Q_1 and Q_2 are built by using identical unit transistor, but the number of unit transistors in Q_1 is n times larger than that in Q_2. Obviously, the output voltage V_{ref} can be expressed as

$$V_{ref} = V_{be2} + 2I_C R_2 \tag{4.22}$$

where V_{be2} is the base-emitter voltage of Q_2, whose temperature coefficient is usually negative. On the other hand, the derivative of $I_C R_2$ with respect to absolute

temperature T is positive as proved later. By choosing appropriate components values, we can set $\partial V_{ref} / \partial T = 0$. To elaborate, the derivation of the expression for $\partial V_{ref} / \partial T$ is given step by step as follows

$$V_{be2} = V_T \ln(\frac{I_C}{I_S}) \tag{4.23}$$

$$V_{be1} = V_T \ln(\frac{I_C}{nI_S}) \tag{4.24}$$

where I_s is the reverse saturation current of Q_1 and Q_2.

$$I_C = \frac{V_{be2} - V_{be1}}{R_1} \tag{4.25}$$

$$= \frac{V_T}{R_1} \ln(n) \tag{4.26}$$

Because R_2 is usually formed by a serial or parallel connection of multiple resistors identical to R_1, R_2/R_1 remains constant, although the absolute values of R_1 and R_2 fluctuate with temperature. Therefore,

$$\frac{\partial(I_C R_2)}{\partial T} = \frac{V_T}{T} \frac{R_2}{R_1} \ln(n) \tag{4.27}$$

Using. (4.23), we can write

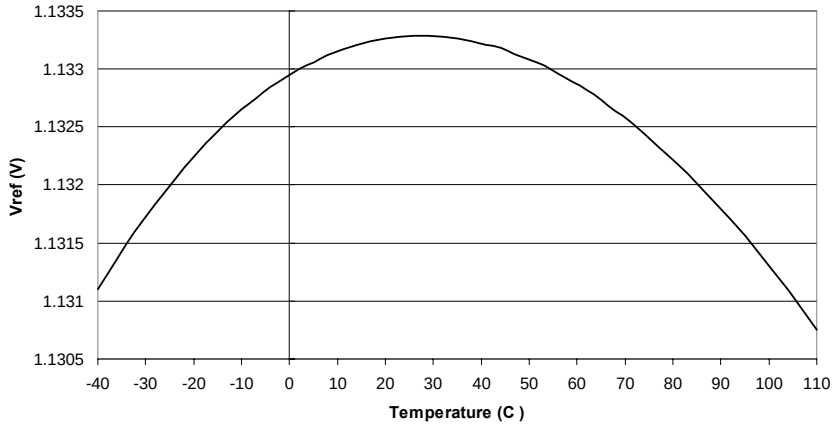
$$\frac{\partial V_{be2}}{\partial T} = \frac{\partial V_T}{\partial T} \ln(\frac{I_C}{I_S}) + \frac{V_T}{I_C} \frac{\partial I_C}{\partial T} - \frac{V_T}{I_S} \frac{\partial I_S}{\partial T} \tag{4.28}$$

$$\frac{V_T}{I_S} \frac{\partial I_S}{\partial T} = (4+m) \frac{V_T}{T} + \frac{E_g}{KT^2} V_T \quad [91] \tag{4.29}$$

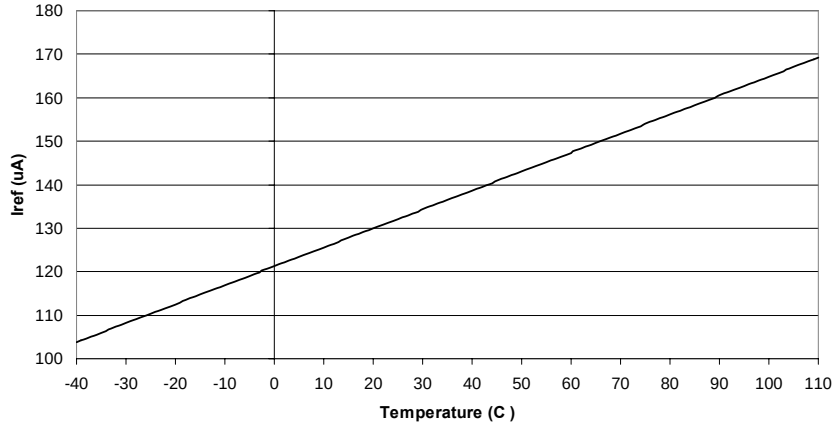
where m is a constant roughly equal to -1.5, and E_g is the bandgap energy of silicon approximately equal to 1.12eV (this is why it is called "bandgap reference"). Substituting (4.29) into (4.28) and with some simple manipulations, we obtain that the condition for zero temperature coefficient is given by

$$\frac{\partial V_{ref}}{\partial T} = \frac{V_{be} - (3+m)V_T - E_g / q}{T} - \frac{V_T}{R_1} \frac{\partial R_1}{\partial T} + 2 \frac{V_T}{T} \frac{R_2}{R_1} \ln(n) = 0 \tag{4.30}$$

In the design process, the values of n and R_1 are firstly selected according to the desired bias current, and then R_2 is scaled as an integer multiple of R_1 to minimize $\partial V_{ref} / \partial T$. In this design, n=2 and R_2/R_1=10 are employed.



(a)



(b)

Figure 4.11 Simulated result of bandgap reference (a) voltage reference (b) current reference

In many applications, it is desirable to have a transistor gain independent of the temperature, i.e., a constant g_m over the temperature range of operation. For bipolar transistor, g_m is related to I_C and T via

$$g_m = \frac{qI_C}{KT} \tag{4.31}$$

Therefore, to obtain a constant g_m we need a current bias proportional to the absolute temperature. It is noted that the bias current of Q_2 in Figure 4.10 fits this requirement to the first order, as indicated by (4.26). Therefore, the bandgap circuit can generate both a temperature-independent voltage and a PTAT current reference. Please note that both V_{ref} and I_{ref} are independent of the supply voltage.

Figure 4.11 (a) and (b) show the simulated V_{ref} and I_{ref} as a function of temperature from -40^{o} to 110^{o}. It demonstrates the V_{ref} has a zero temperature coefficient around room temperature and the variation across the interested temperature range is only 0.2%. It also shows that I_{ref} scales linearly with the temperature as intended.

4.3 Local Oscillator Path – PLL Design and Phase Generation

The local oscillator path of the fully-integrated 24-GHz phased array receiver provides 19.2-GHz and 4.8-GHz signals that are applied to the LO ports of the down-conversion mixers. The first LO frequency is generated by an on-chip VCO whose output frequency is locked to an external reference via an integrated phased-lock loop. The I and Q of the second LO frequency are created by dividing the VCO frequency by a factor of 4. In addition to the general functions of a frequency synthesizer in a single-path wireless transceiver, the LO path in this design also creates multiple phases at 19.2-GHz for the requirements of phase-shifting. The symmetric phase generation and distribution are crucial to maintaining a high spatial selectivity of the array pattern. This section focuses on PLL design and phase generation technique. The issues of on-chip phase distribution are going to be addressed in next section.

4.3.1 PLL Basics

To downconvert the RF signals at multiple channels to baseband, wireless communication systems require one (in homodyne architecture) or multiple (in heterodyne architecture) internal signal sources with tunable, stable, and accurate output frequencies. Such signal sources are commonly generated by using a negative feedback loop which fixes the phase and thereby the frequency relation of a high-frequency oscillator output to a stable and accurate low-frequency reference. Such a negative feedback loop is called PLL.

PLL is an indispensable component in various advanced electronic systems. Besides wireless transceivers, it is also used for clock generation in microprocessors and clock-and-data recovery in optical communications systems.

Designing a high-performance PLL is not an easy task. First, PLL is commonly modeled as a linear system at the locked condition, which is inherently inaccurate due to fact that PLL is a nonlinear time-variant system. When there is a large frequency jump, such a model is not even applicable making the analyses of frequency pull-in range and transient response difficult. Secondly, the simulation of PLL takes a very long time and sometimes is not even practical if the design period is short because the time constants involved in a PLL simulation can be varied by multiple orders of magnitude [92]. Thirdly, the impurities of the PLL output spectrum, such as phase-noise, reference feedthrough, and additional spurs can significantly deteriorate the performance of the communication systems. Therefore fully optimized PLLs are generally desired. In addition, to reduce the cost of mass production, the modern integrated wireless system requires a PLL-based frequency synthesizer be implemented on the same chip with the transmitter and receiver, which introduces new problem such as frequency pulling by the power amplifier. Furthermore, in addition to the signal path, the design of ultra high-frequency integrated PLLs in silicon suffers from a lower transistor gain, significant passive loss, and more substrate noise. The design of such PLLs requires a comprehensive knowledge of microwave integration at the system level, the transistor level and the physics level.

The block diagram of a common PLL has been depicted in Figure 2.10. For the simplicity of notation, the dependence of each variable on time t is not explicitly denoted in the figure. The VCO provides output of the PLL. The instantaneous output frequency of VCO, $f_{\text{out}}(t)$, is depend on the voltage of its control input V_{cntl} (t) by the following equation

$$f_{out}(t) = f_o + 2\pi K_{vco} V_{cntl}(t) \tag{4.32}$$

where f_{o} is the VCO free running frequency, and K_{vco} is called "VCO gain" which specifies the sensitivity of the VCO output frequency to the control voltage. The phase of the PLL output $\Phi_{out}(t)$ is the integration of the instantaneous VCO frequency over t, which is given by

$$\Phi_{out}(t) = f_o t + 2\pi K_{vco} \int V_{cntl}(t) dt \tag{4.33}$$

The first term is dropped due to its independence of the loop operation, hence it is modified to

$$\Phi_{out}(t) = 2\pi K_{vco} \int V_{cntl}(t) dt \tag{4.34}$$

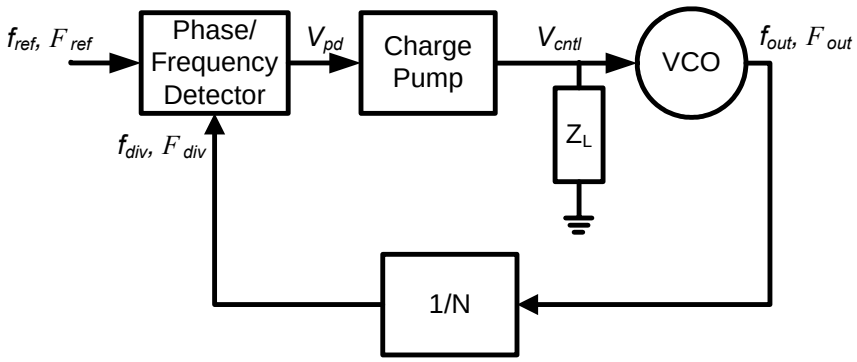


Figure 4.12: Block diagram of a generic charge pump PLL

and its Laplace transform can be expressed as

$$\Phi_{out}(s) = \frac{2\pi K_{vco} V_{cntl}(s)}{s} \tag{4.35}$$

The frequency divider, denoted by *1/N*, divides the VCO output frequency by a factor of *N*. Meanwhile, at locked condition, the phase of the divider output is related to the VCO output phase by

$$\Phi_{out,div}(t) = \frac{\Phi_{out}(t)}{N} \tag{4.36}$$

The phase detector compares the phase of a reference signal and the divider output and generates an output voltage V_{pd} proportional to the phase difference. The voltage of V_{pd} can be expressed as,

$$V_{pd}(t) = K_{pd}(\Phi_{ref}(t) - \Phi_{div}(t)) \tag{4.37}$$

A subsequent loop filter with transfer function *H(s)* ideally removes all high frequency distortion in V_{pd}(t) and provides a dc voltage of V_{cntl}(t) that corresponds to the desired oscillating frequency. The certain combination of VCO, PD, and the loop filter should make sure the polarity of the feedback is negative. The overall transfer function of PLL is given by

$$H(s) = \frac{2\pi K_{pd} K_{vco} Z(s)}{s + 2\pi K_{pd} K_{vco} Z(s)/N} \tag{4.38}$$

PLL can be implemented in various architectures. Among those architectures, charge pump type PLL is particularly attractive due to the following advantages:

1. Ideally there is zero phase error between input and output at locked status.

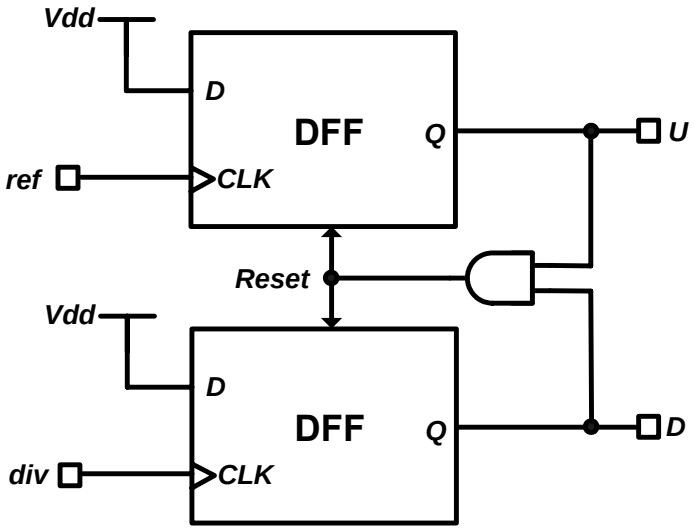


Figure 4.13: Phase/frequency detector

2. Locking range is only limited by the frequency tuning range of VCO in most cases.

3. There is less output noise power contributed by phase/frequency detector.

The architecture of a charge-pump PLL is presented in Figure 4.12 [92]. The phase/frequency detector generates a pulse signal whose duty cycle is proportional to the phase difference of the reference and divider output. The pulse signal switches on/off the current pump or sink of charge pump to adjust $V_{cntl}(t)$. The loop bandwidth and phase margin are significantly affected by the charge pump current I_{cp} and the loop filter impedance $Z_L(s)$. It can be shown that the charge pump loop transfer function is given by

$$H(s) = \frac{I_{cp} K_{vco} Z_L(s)}{s + I_{cp} K_{vco} Z_L(s) / N} \tag{4.39}$$

The design of each block in charge-pump PLL will be discussed in detail in the following subsections.

4.3.2 Phase/Frequency Detector

A common implementation of the digital phase-frequency detector by using D-type flip-flop (DFF) is shown in Figure 4.13. Assuming the rising edge of the reference

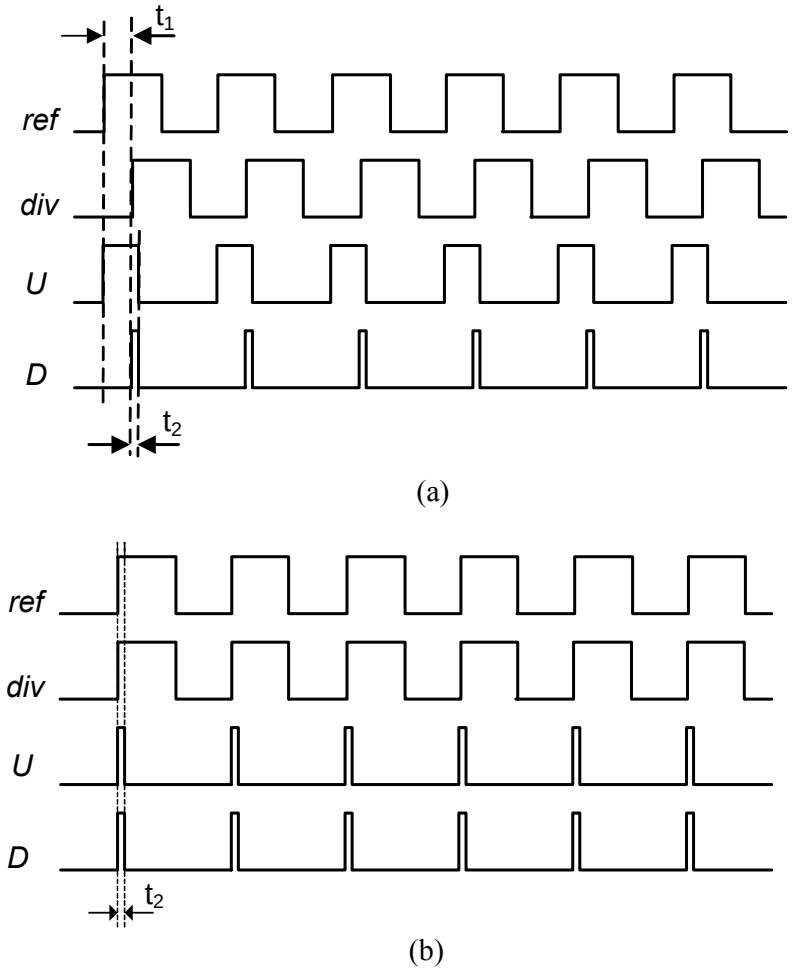


Figure 4.14: Output waveforms of PFD (a) $\Delta\Phi \neq 0$ (b) $\Delta\Phi = 0$

signal *ref* occurs first, the output signal *U* will be triggered active while *D* remains zero. When the rising edge of the divider output *div* arrives, the *D* signal will be triggered active. Subsequently, an AND gate immediately generates a reset signal, bringing both *U* and *D* back to zero. Obviously, at static state, *U* and *D* can be both zero and either of them can be "1" alone. However, they can not be "1" simultaneously.

The Figure 4.14 (a) shows the typical output waveform of the PFD if *ref* and *div* have the identical frequency but a constant phase difference. The *D* signal is active for a short period of time t_2 per each reference cycle T_{ref} due to the delay of the AND gate and the resetting time of the DFF. The *U* signal is active for a total period of t_1+t_2, where

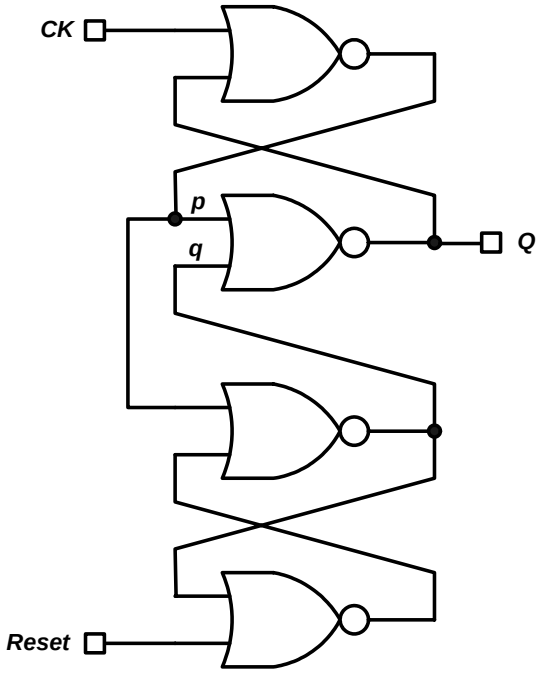


Figure 4.15: Implementation of DFF in Figure 4.13

$$t_1 = \frac{\Delta\phi}{2\pi} T_{ref} \tag{4.40}$$

Please note that even if *ref* and *div* are exactly in phase, *U* and *D* are not constantly low but active for a short period of t_2 per each reference cycle as shown in Figure 4.14 (b). This characteristic may lead to a spurious signal at the PLL output if asymmetry is exhibited in the source and sink current of the charge pump.

Because at locked condition the duty cycle of *U* and *D* are very low, the phase noise contribution of PFD to the PLL output is negligible, which is a significant advantage of the PLL using PFD/CP combination to those using continuous-time mixer as the phase detector.

The implementation of each D-type flip-flop in this work is shown in Figure 4.15. Since the "D" input of DFF is fixed at logic 1, it is hidden in this schematic. A major consideration in this PFD design is the maximum operation frequency f_{max}. When *ref* and *div* exceed a certain frequency, a transition edge can be missed within the duty-cycle of the reset signal. Consequently, in the following reference cycle the PLL will pull-out the VCO frequency instead of pull-in this reference. In this case the PLL can

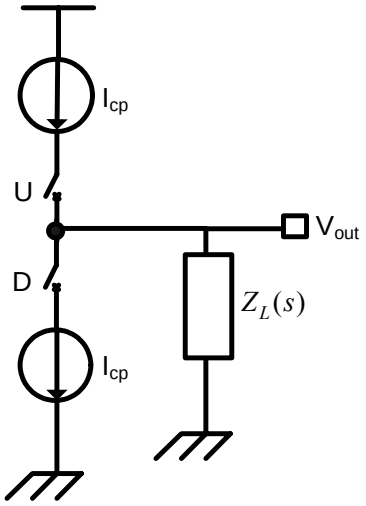


Figure 4.16: A generic charge pump

never get locked. Particularly in digital PFD, the limitation of operation frequency is caused by the delay of logic gates. For instance, for PFD using DFF shown in Figure 4.15, the criteria of proper operation is that when node *p* is set to low by the rising edge of *ck*, the reset signal *q* has returned to zero. Let us assume each NOR gate in Figure 4.15 and the AND gate in Figure 4.13 have an identical delay of τ, an analysis of the logic propagation in this PFD yields

$$f_{\max} = \frac{1}{16\tau} \tag{4.41}$$

Simulations show that each logic gate causes a delay on the order of 120ps, and the maximum operation frequency of PFD is 500MHz. The results are in good agreement with (4.41). For this particular design, 19.2-GHz VCO and a dividing ratio of 256 require the PFD to operate at 75MHz. The PFD is over designed for possible usage in other systems in the future.

4.3.3 Charge Pump

The basic block diagram of a charge pump [93] is shown in Figure 4.16. The PFD output *U* and *D* turn on/off the switches in the charge pump so that the load $Z_L(s)$ will be charged or discharged accordingly with a constant current I_{cp}. Since the duty cycle of *U* and *D* is proportional to the phase difference $\Delta\Phi$ between the reference and the divider output, the average output current $\overline{I_{cp}}$ is given by

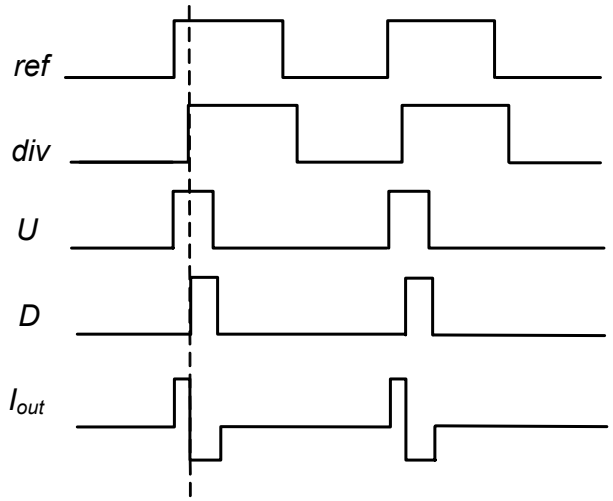


Figure 4.17: PFD and chargepump I/O waveforms when current mismatch exists

$$\overline{I_{cp}} = \frac{I_{cp}\Delta\Phi}{2\pi} \tag{4.42}$$

If the loop bandwidth is much smaller than the reference clock period, the PLL can be approximated with a continuous system by using the average value per phase comparison cycle [10]. In this case, the gain of the PFD/CP combination can be approximated as

$$K_{pd} = \frac{I_{cp}}{2\pi} \tag{4.43}$$

In general there are infinite sets of I_{cp}, K_{vco} and $Z_L(s)$ that can be used to implement a functional PLL. A smaller I_{cp} results in a low power design. However, a higher I_{cp} can reduce the resistance value of the loop filter so that its contribution to the PLL output phase noise is minimized [92]. In this design, I_{cp} is chosen to be 2.5mA as a trade-off. The corresponding K_{pd} is approximately 0.4mA/rad.

As we discussed in Subsection 4.3.2, even in in-lock situation, *U* and *D* will turn on both switches in a charge pump for a short period of time. If the source current and sink current in the charge pump are identical, no current will follow to the loop filter during this period. However, if there is a mismatch between the two currents, in the locked situation one signal will turn on longer than the other to compensate for this mismatch, as shown in Figure 4.17, producing spurs at the PLL output with an offset

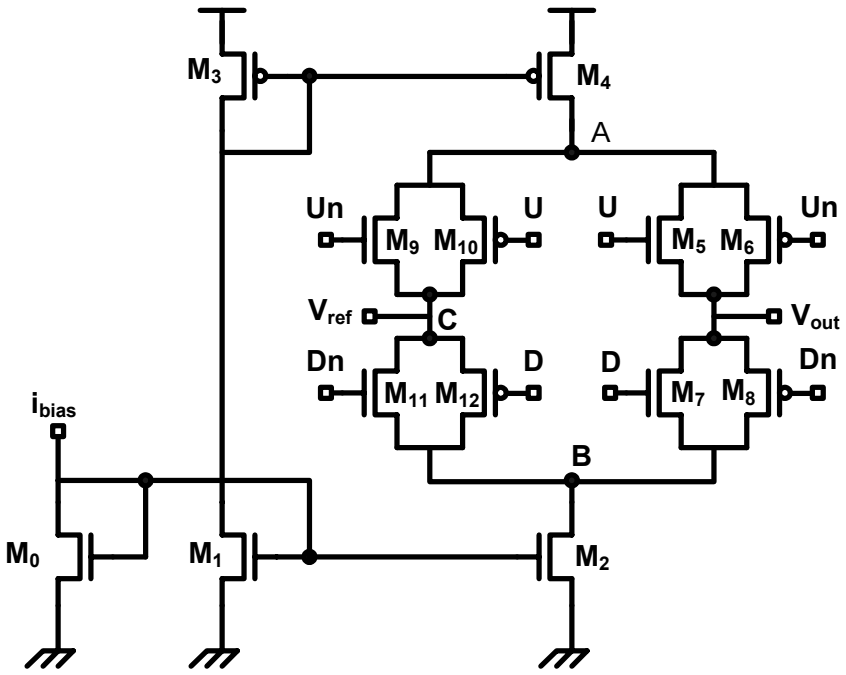


Figure 4.18: A multi-switch charge pump

at the multiples of the reference frequencies. Therefore, special efforts are required to minimize the source and sink current mismatches.

The circuit implementation of the charge pump is shown in Figure 4.18 [94]. The sink current is regulated by M_0~M_2. M_3 and M_4 mirror the sink current to the upper part which is serving as the source. To improve the matching, a long channel MOSFET with a channel length of 1μm is used for M_0~M_4 to increase the output impedance of the current mirror. The upper and bottom switches are implemented by using complementary switch pairs M_5 and M_6, M_7 and M_8, respectively to minimize clock feedthrough. A parallel branch M_9~M_{12} is used for the following reasons: Without M_9~M_{12}, when *U* is low, the voltage of node *A* will be pulled to V_{dd}. Hence, at the moment when *U* is set to high, M_4 is in the triode region so that the source current is not equal to I_{cp}. The same phenomenon happens to the sink current flowing through node *B*. This defection will change the loop parameters from the designed value, limit the maximum frequency the charge pump can be used, and increase the mismatch. To mitigate this problem, a second branch comprising of M_9~M_{12} is used, as shown in Figure 4.18. The mid-point *C* is biased at the nominal value of the VCO control voltage. When *U* or *D* is low, a complementary signal will turn on the

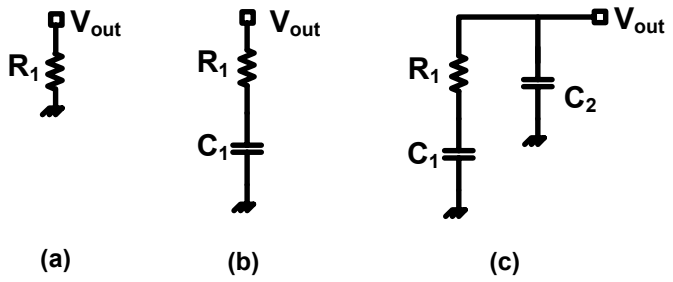


Figure 4.19: Examples of the loop filter (a) single resistor (b) 1st-order RC filter (c)2nd-order RC filter

corresponding branch on the left so that the voltages at *A* and *B* will remain relatively constant at the switching moment.

4.3.4 Loop Filter

In practice, the load of the charge pump can not be a single capacitor, because such a system has two poles at the origin and inherently unstable [92]. A loop filter including resistors has to be used to achieve stable operation, which requires that the transfer function of the PLL exhibit appropriate bandwidth and sufficient phase margin. In typical designs, a rule-of-thumb is to choose a loop bandwidth that is approximately one-tenth of the reference frequency.

The loop filter can be built by using an active or passive filter. The active filter can decouple the VCO control-terminal load from the loop filter but comes with the price of higher power consumption [93]. The passive filter is adopted in this design for low-power consideration. A few examples of the passive filter architectures are given in Figure 4.19, all of which can lead to a stable system if appropriate components values are used. For the filters in Figure 4.19 (a) and (b), a voltage jump at the node V_{out} will happen at each transition point of the charge pump, resulting in significant spurs at the PLL output at the offset of harmonics of the reference frequency [91]. A second order filter, as show in Figure 4.19 (c), is usually required to alleviate this problem.

The load impedance of the filter in Figure (c) is given by

$$Z_L(s) = \frac{1}{s(C_1 + C_2)} \frac{1 + sR_1C_1}{1 + sR_1 \frac{C_1C_2}{C_1 + C_2}} \tag{4.44}$$

The loop dynamic is determined by I_{cp}, K_{vco}, and the filter components. The final choices of those parameters are $K_{vco}=2.1GHz/V$, $I_{cp}=2.5mA$, $R_1=256k\Omega$, $C_1=30pF$, and $C_2=8.28pF$. The calculated loop bandwidth is 5.9MHz, and the phase margin is approximately 40^o.

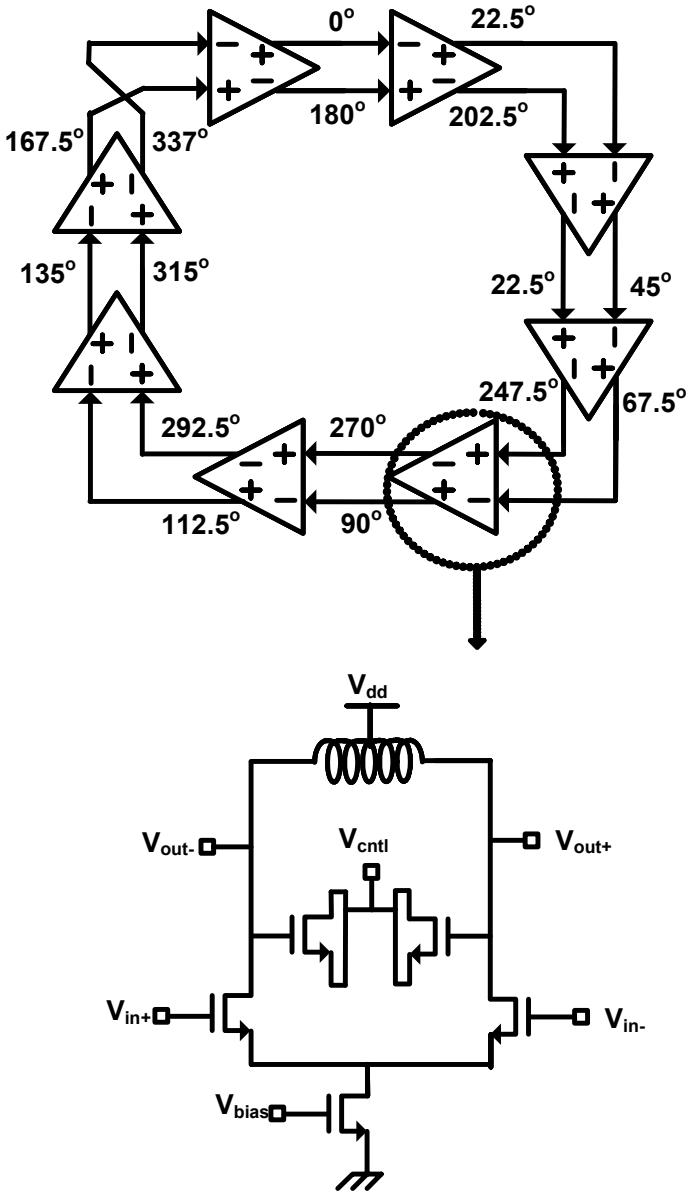


Figure 4.20: 16-phase CMOS VCO

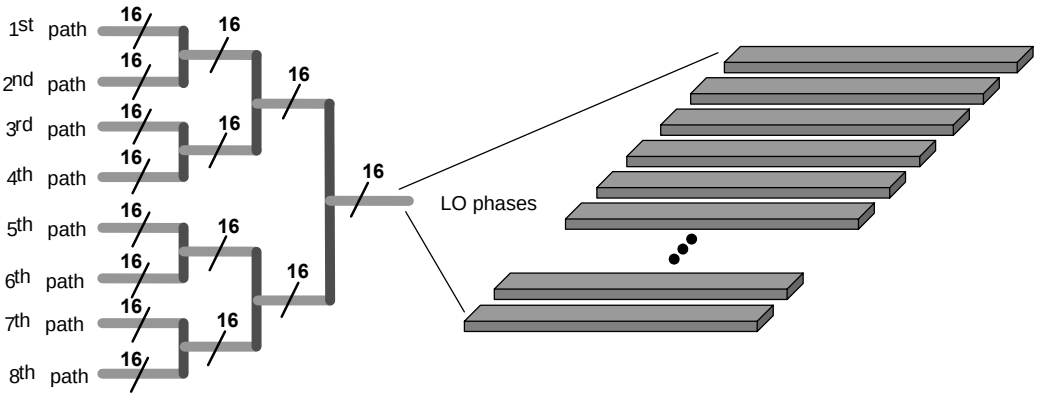


Figure 4.21: Phase distribution binary tree

4.3.5 VCO and Frequency Divider

The VCO and frequency divider were designed entirely by Hossein Hashemi. For the purpose of the completeness of this dissertation, those blocks are briefly introduced in the subsection.

A ring connection of 8 differential CMOS tuned amplifiers forms the 19.2GHz VCO, as shown in Figure 5.10 [95], which generates 16 phases evenly allocated between 0^o~360^o used to provide phase shifting at the LO path with a 4-bit resolution.

Digital frequency dividers with cross-coupled D-type flip-flops using emitter coupled logic are employed in all divide-by-2 blocks.

4.4 Local Oscillator path - Phase Distribution

The 16 phases at 19.2GHz generated by the core oscillator need to be fed into local phase selectors at 8 paths with equal amplitudes and delays. The deviations of amplitudes or relative phases can significantly deteriorate the spatial selectivity of the array pattern [96][97]. This section addresses the issue of symmetric phase distribution in detail.

4.4.1 Binary Tree Structure

The binary tree structure shown in Figure 4.21 is used to deliver the 16 phases to 8 phase selectors symmetrically. Inside each bus is a transmission line array comprised

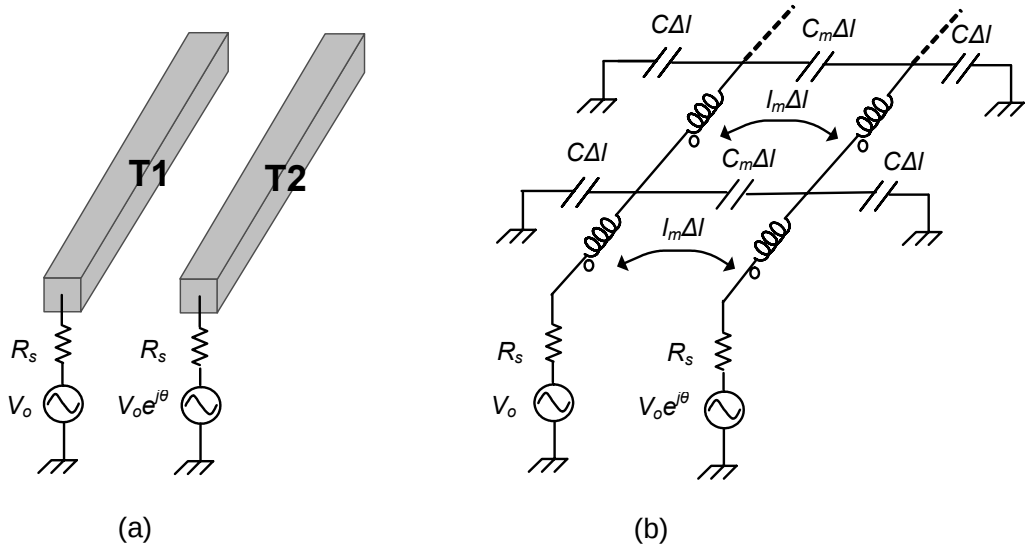


Figure 4.22: Two coupled transmission lines (a) basic structure (b) lumped model

of 16 top-metal wires carrying different phases. Special attention has been paid to ensure identical geometric length of each phase route.

Despite the global symmetry of the binary structure, the discontinuity at the edge and the unwanted electromagnetic coupling between the metal wires can produce mismatches in both amplitude and relative phase, which is going to be discussed in following subsections.

4.4.2 Coupling Effects of Two Parallel Transmission Lines

Consider two identical lossless transmission lines T_1 and T_2 running in parallel and driven by two signal sources V_o and $V_o e^{j\theta}$, respectively, as shown in Figure 4.22 (a). The equivalent lumped model of this transmission line pair is shown in Figure 4.22 (b), where c, l, l_m, and c_m are per-unit-length capacitance to ground, inductance, mutual inductance and coupling capacitance, respectively.

The transmission line impedance is defined as the voltage-to-current ratio at the driving port if its length is infinite [90]. The impedance of each transmission line in a coupled pair not only depends on its geometric and physical properties, but also on their relative phase θ [90]. If $\theta=0^o$ (even-mode excitation), the voltages on both lines have even symmetry along the center line, effectively making the coupling capacitor

between the two lines c_m open-circuited. Meanwhile, the l and l_m retard the current variation in the same direction. Hence the characteristic impedance of each line is given by

$$Z_{even} = \sqrt{\frac{l + l_m}{c}} \tag{4.45}$$

On the other hand, if $\theta=180^o$ (odd-mode excitation) the voltages on both lines have odd symmetry along the center line, setting the center line as a virtual ground. Therefore, each line sees an effective per-unit-length capacitance to ground of $c+2c_m$. At the same time, the l and l_m pull the current in the opposite direction. Therefore, the characteristic impedance of each line is given by

$$Z_{odd} = \sqrt{\frac{l - l_m}{c + 2c_m}} \tag{4.46}$$

In the general case, when θ is not 0^o or 180^o, the traveling wave can be decomposed as a linear combination of even and odd modes, where even-mode excitation can be expressed as

$$V_{1,even} = V_{2,even} = \frac{V_o + V_o e^{j\theta}}{2} \tag{4.47}$$

and odd mode excitation can be expressed as

$$V_{1,odd} = -V_{2,odd} = \frac{V_o - V_o e^{j\theta}}{2} \tag{4.48}$$

We can write,

$$V_1 = V_{1,even} + V_{1,odd} \tag{4.49}$$

$$I_1 = \frac{V_{1,even}}{Z_{even}} + \frac{V_{1,odd}}{Z_{odd}} \tag{4.50}$$

$$V_2 = V_{2,even} + V_{2,odd} \tag{4.51}$$

$$I_2 = \frac{V_{2,even}}{Z_{even}} + \frac{V_{2,odd}}{Z_{odd}} \tag{4.52}$$

$$Z_{o1} = \frac{V_1}{I_1} \tag{4.53}$$

$$Z_{o2} = \frac{V_2}{I_2} \tag{4.54}$$

Using (4.59) to (4.54), it can be derived that the transmission line impedance Z_{o1} and Z_{o2} can be expressed in terms of c, l, l_m, and c_m and θ as shown below

$$|Z_{o1}| = |Z_{o2}| = \frac{\sqrt{2} Z_{even} Z_{odd}}{\sqrt{Z_{odd}^2 (1+\cos\theta) + Z_{even}^2 (1-\cos\theta)}} \tag{4.55}$$

$$\angle Z_{o1} = -\angle Z_{o2} = \arctan \frac{Z_{even} \sin\theta - Z_{odd} \sin\theta}{Z_{odd} (1+\cos\theta) + Z_{even} (1-\cos\theta)} \tag{4.56}$$

It can be seen that Z_{o1} and Z_{o2} form a complex conjugate pair, which are equal unless θ is 0^o or 180^o. Therefore, such a phase distribution pair will cause an unbalanced load at the driving amplifier's outputs, resulting in unmatched phase deviations.

4.4.3 EM Coupling inside a Transmission Line Array

In the transmission line array shown in Figure 4.21, we need to take into account not only the coupling between adjacent wires, but also the crosstalk between nonadjacent wires [98]. EM simulations using IE3D are performed on an array of 16 on-chip transmission lines, as shown in Figure 4.23. In our design, each line is 4μm thick, 5μm wide, and 200μm long, with a 5μm edge-to-edge spacing. These lines are 12μm above the silicon substrate. Figure 4.24 shows the extracted mutual inductance and coupling capacitance normalized to the inductance l and capacitance c respectively. It illustrates that the capacitive coupling is negligible between nonadjacent lines because the electric field is shielded, it also shows that the magnetic coupling is significant and the mutual inductance decreases very slowly and extends beyond multiple lines, increasing the asymmetry inside a finite array.

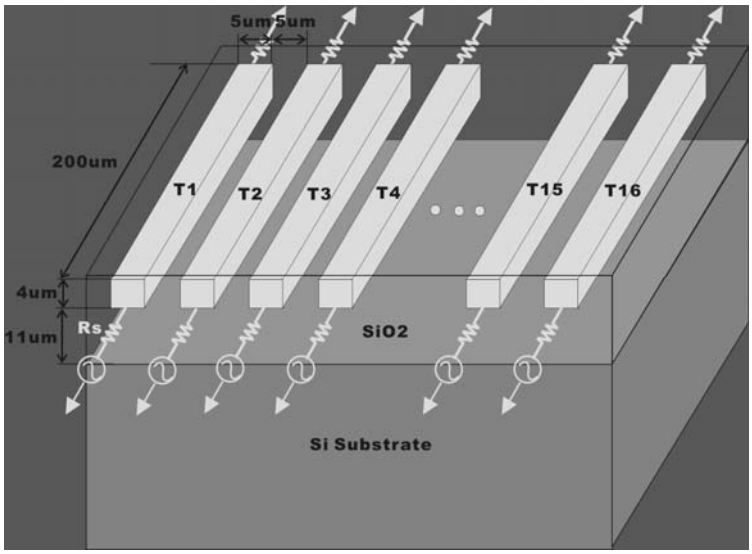


Figure 4.23: Transmission line arrays on silicon substrate

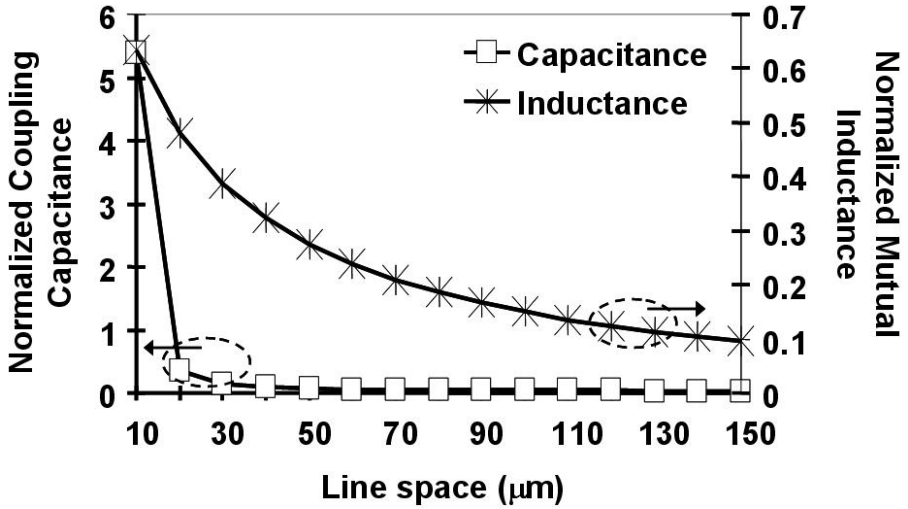


Figure 4.24: EM crosstalk inside a transmission line array

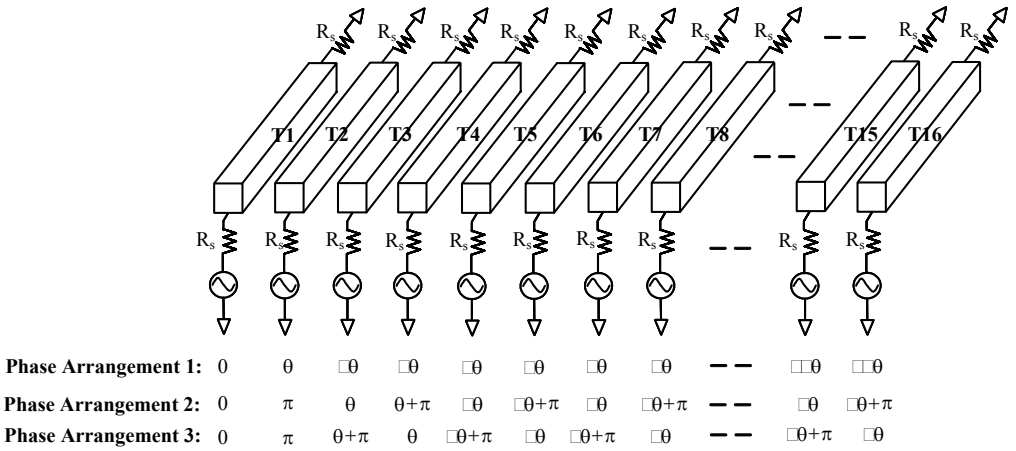


Figure 4.25: Three phase arrangments

4.4.4 Transmission Line Properties in Various Phase Sequences

Due to the EM crosstalk between wires, the transmission line impedance and matching properties in an array not only depends on its geometric and physical characteristics, but also on the phase sequence allocated. Figure 4.25 shows three different phase arrangements in a transmission line bus carrying multiple phases. If the array has an infinite number of lines, arrangement 1 provides the best symmetry. Considering a differential length of line *dz*, we see that

$$\frac{\partial V}{\partial z} = -(l + 2\sum_{k=1}^{\infty} l_{mk} \cos k\theta)\frac{\partial I}{\partial t} \tag{4.57}$$

$$\frac{\partial I}{\partial z} = -(c + 2\sum_{k=1}^{\infty} c_{mk} \cos k\theta)\frac{\partial V}{\partial t} \tag{4.58}$$

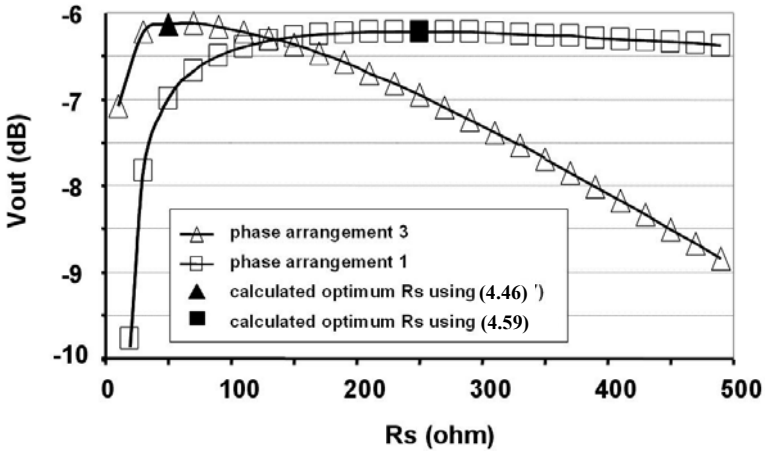
By applying the similar procedure in [90], it can be derived that the transmission line impedance in arrangement 1 is given by

$$Z_o = \sqrt{\frac{l + 2\sum_{k=1}^{\infty} l_{mk} \cos k\theta}{c + 2\sum_{k=1}^{\infty} c_{mk}(1 - \cos k\theta)}} \tag{4.59}$$

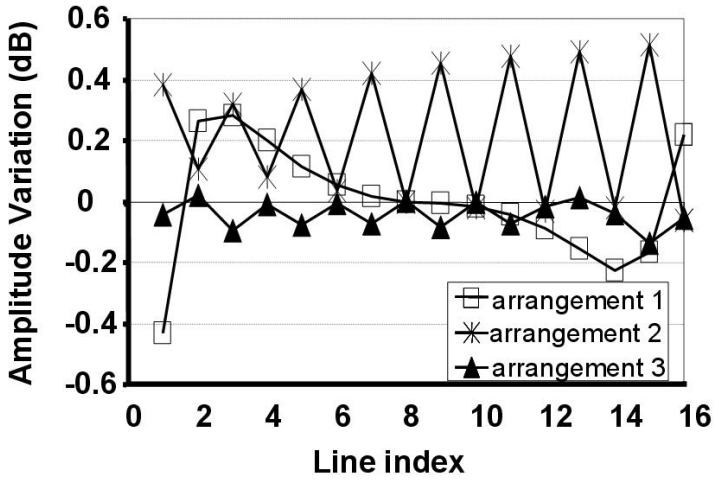
where l_{mk} and c_{mk} are the mutual inductance and coupling capacitance between two lines with a phase difference of $k\theta$. However, in a finite array the discontinuity at the edge and the inductive crosstalk between nonadjacent lines can produce significant mismatch at the outputs of arrangement 1.

According to Ampere's law, placing differential phase pairs as shown in arrangements 2 and 3 can minimize magnetic coupling. If θ is small (θ=22.5^o in this work) arrangement 3 has better phase and amplitude matching characteristics than the other two. This is because in 3 the adjacent lines of two different pairs are closer in phase so that the capacitive coupling between them is minimized. For a small θ, the characteristic impedance of the transmission lines in arrangement 3 can be approximated by the odd-mode impedance given by (4.46).

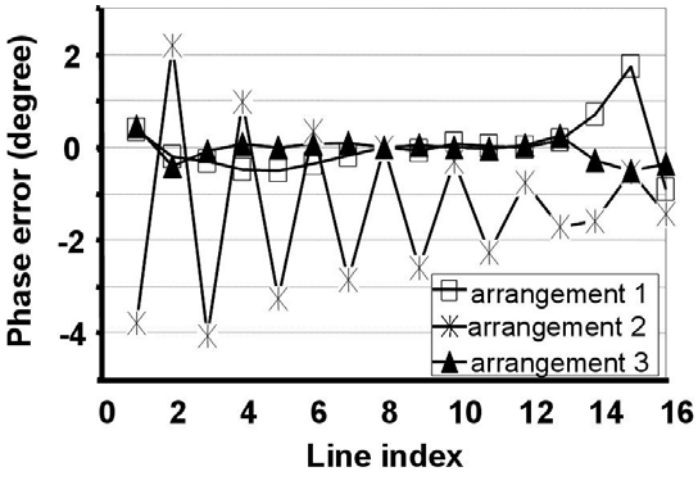
To compare these three proposed phase arrangements, the results of the EM simulations were employed in Agilent ADS. Each of the three arrays is driven by 16 evenly-spaced phases of a 19.2GHz sinusoid. The transmission lines see a resistance R_s at both input and output ports. Figure 5.16 (a) illustrates the voltage at the output port of the central wire as a function of R_s. It verifies that using resistance values estimated by (4.59) and (4.46) results in maximum V_{out} for arrangements 1 and 3, respectively. Figure 5.16 (b) and (c) shows the magnitude and phase variations, respectively, of the voltages at the 16 output ports for 3 arrangements, it can be seen that arrangement 3 exhibits less mismatch, and hence is adopted in our 24-GHz phased array receiver.



(a)



(b)



(c)

Figure 4.26: EM simulation results (a) transmission line impedance (b) amplitude variations (c) phase variations

4.5 Experimental Results

4.5.1 Implementation

The phased array receiver is implemented in an IBM 7HP SiGe BiCMOS technology providing a bipolar f_T of 120 GHz and f_{max} of 100GHz, together with 0.18μm CMOS transistors [99]. It offers five metal layers with a 4μm-thick top analog metal used for on-chip spiral inductors as well as transmission lines routing the high-frequency signals. The other features of the process include MIM capacitors, MOS varactors, and various types of diffusion and polysilicon resistors. The substrate resistance of the process is approximately 8Ω.cm. The die micrograph of the phased-array receiver is shown in Figure 4.27. The 8 RF front-ends are placed in parallel on the left hand side. The multi-phase VCO and frequency synthesizer are located on the right hand side. Phase distribution transmission lines and phase selectors can be seen in the middle. At the bottom side are IF, baseband, and bias circuitry. The size of the chip is 3.3 x 3.5 mm^2.

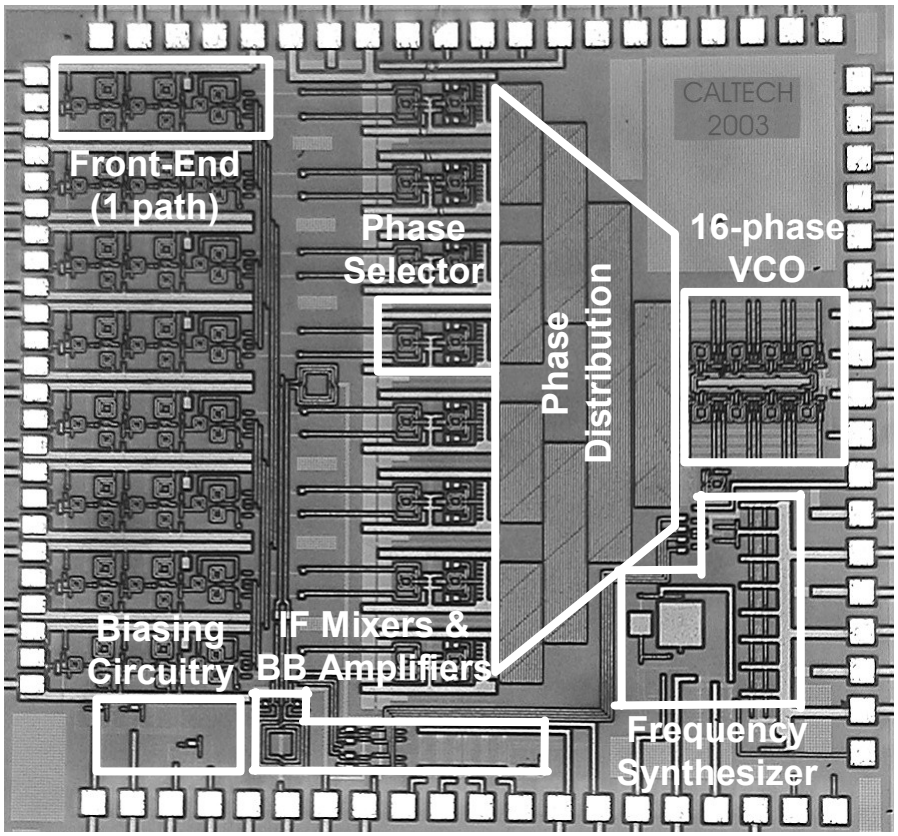


Figure 4.27: Die Micrograph

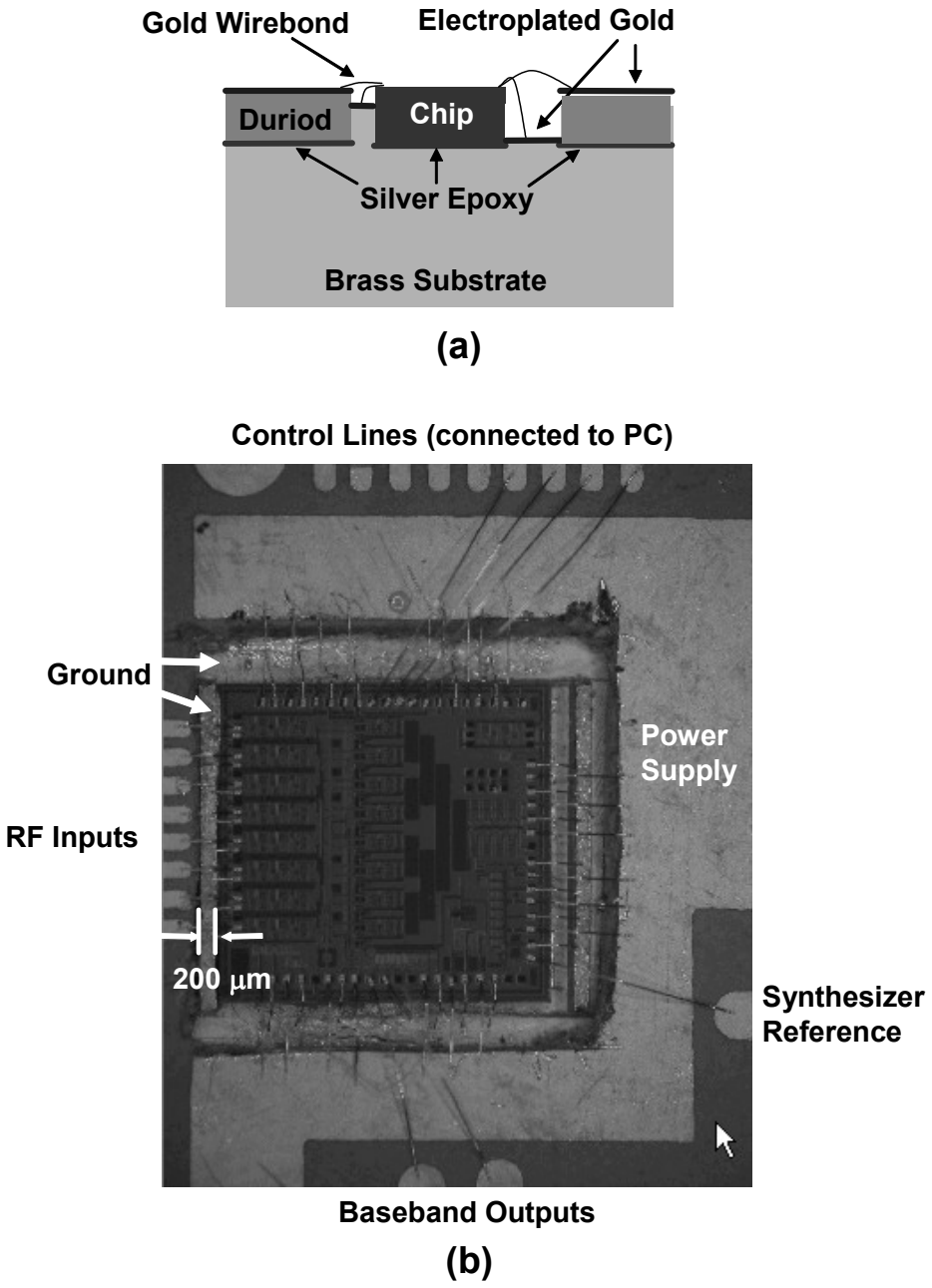


Figure 4.28: Test package

4.5.2 Test Package

The test board connects power supply, 24-GHz RF inputs, frequency synthesizer references, analog baseband outputs, and digital controls for programming the phase shifting status to the receiver chip. To carry the 24-GHz signal, the test board is fabricated on Rogers 5880 high-frequency duroid laminate. The die and test board are mounted on a brass platform using silver epoxy, as shown in Figure 4.28. The brass substrate serves as a high-efficiency ground for microwave signals. The thickness of the employed Duroid board is chosen to be 10mil, approximately the same height as the chip, this minimizes signal bond wire length and curvature. A 3.5mm-long brass step with width and height of 200μm is built along the RF side of the chip. The ground pads for the RF circuitry are wire-bonded to the top surface of this step to minimize the length of the ground bond wire. The inputs of every path are symmetrically wire-bonded to 50Ω transmission lines on board. All signal and bias lines are fed with standard SMA connectors attached to the brass membrane. It is noteworthy that this configuration facilitates the integration of the planar antenna on the same package.

4.5.3 Receiver Measurement Results

To characterize VCO and the frequency synthesizer without affecting the symmetry of multi-phase generation, a coil is used to pick up the near-field high-frequency signal the chip generates and feed it into a K-band amplifier. The spectrum analyzer is used to observe the amplifier output at the interested frequencies. The free running VCO achieves a phase noise of -103dBc/Hz at a 1MHz offset as shown in Figure 4.29 and a tuning range of 2.1GHz [100]. The frequency synthesizer is locked from 18.4 ~ 20.4GHz with settling time of less than 50μs. As predicted, the locking range of the charge-pump PLL is primarily limited to the VCO tuning range. Figure 4.30 (a) shows the output spectrum of the frequency synthesizer locked at 19.2GHz, demonstrating reference suppression better than 35dB. Figure 4.30 (b) shows the measured phase noise of the synthesizer output at different frequencies. In this measurement the signal generator HP8643A provides the 75-MHz reference singal. It can be seen in Figure 4.30 (b) that the in-band phase noise of the synthesizer output is $10\log_{10}(256)$ dB larger the phase noise of the reference signal, which indicates that the

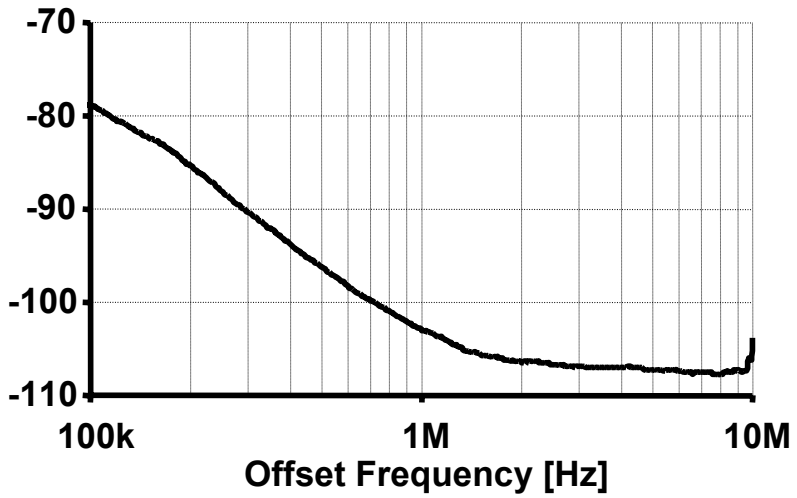


Figure 4.29: Phase-noise of free running VCO

frequency synthesizer phase noise is limited by the reference noise in this setup. A substantially lower phase noise level can be expected if crystal-type reference is used.

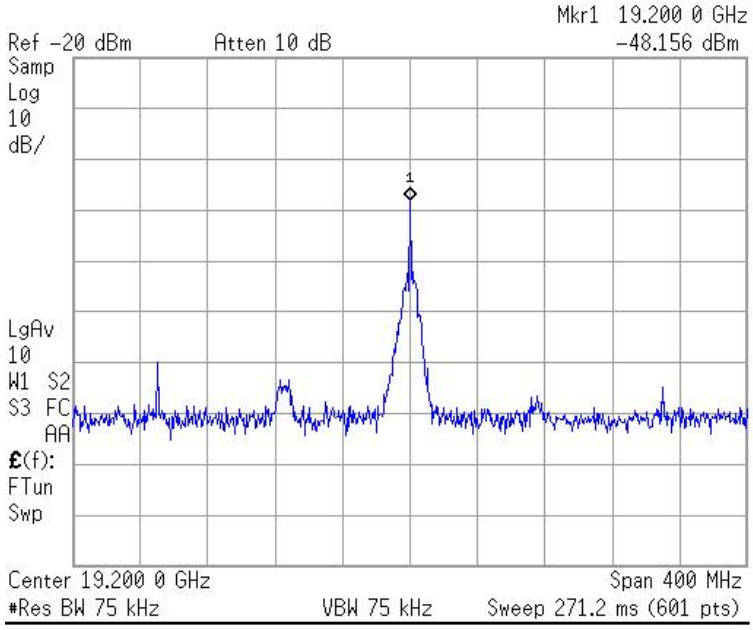
The input reflection coefficients S_{11} at 24-GHz RF ports are characterized both on chip and at the SMA connectors of the RF inputs on board. The receiver demonstrates good input matching properties at the frequency range of interest in both cases, as shown in Figure 4.31.

Figure 4.32 depicts the gain of a single path as a function of the input frequency, showing a 43dB peak gain at 23GHz and a 35dB on-chip image rejection. The image signals will be further attenuated by narrow band antennas. A 3dB gain variation is observed among all paths. Figures 4.33 and 4.34 show the measured nonlinearity of a single path. The input-referred 1dB compression point is observed at -27dBm, and the input-referred intercept point of the third-order distortion is -11.5dBm. The receiver noise figure as a function of input frequency is shown in Figure 4.35. A DSB noise figure of 7.4dB is measured over the signal bandwidth of 250MHz.

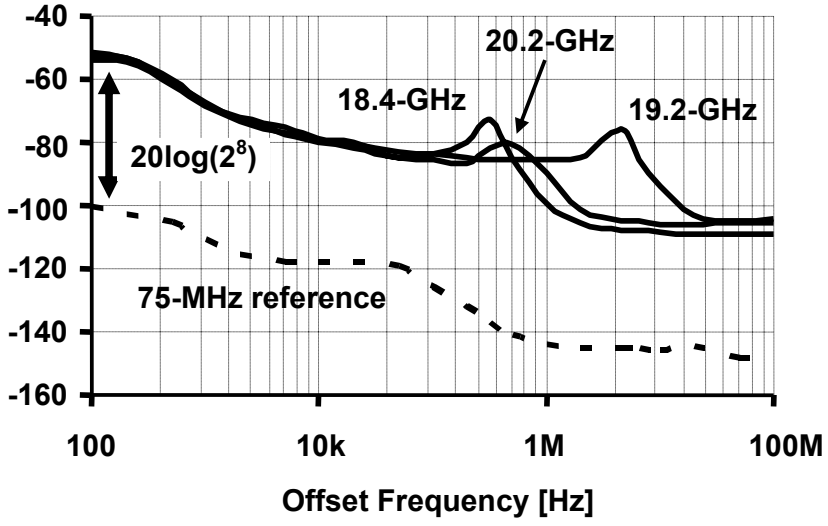
Figure 4.36 shows the on-chip isolation between different paths. The signal is fed to the fifth path only. The phase selector of each path is turned on alternatively to measure the output power caused by coupling. When all phase selectors are off, the system has a -27dB signal leakage (normalized to single-path receiver gain). The coupling is lower than -20dB in all paths. The strongest coupling is seen between

adjacent paths, e.g. the fourth and fifth paths as expected. However, when the phase selector at the fourth path is turned off and the one at the sixth path is turned on, a significantly lower output power is observed which may due to the coexisting coupling and leakage canceling each other. The couplings between non-adjacent paths are either close to or lower than the leakage level.

The array performance is assessed using the setup shown in Figure 4.37. An artificial wave front is generated by feeding the RF inputs to each receiver path via power-splitters and adjustable phase-shifters, as shown in Figure 4.37. This way, the array performance is measured independently of the antenna properties. Figure 4.38 and Figure 4.39 show the measured array patterns at different LO-phase settings for two and four-path operations, respectively. Figure 4.38 and Figure 4.39 clearly demonstrate the spatial selectivity of the phase-array receiver and its steering of the beam over the entire 180^0 range by LO phase programming. The difference between the peak and the null is 10-20dB in all cases. This value is mostly limited by the mismatch in different paths and can be significantly improved with a gain control block in each receiver path for future implementations. The measured performance is summarized in Table 4.1.



(a)



(b)

Figure: 4.30: PLL measurement results (a) Output Spectrum (b) Phase Noise

`

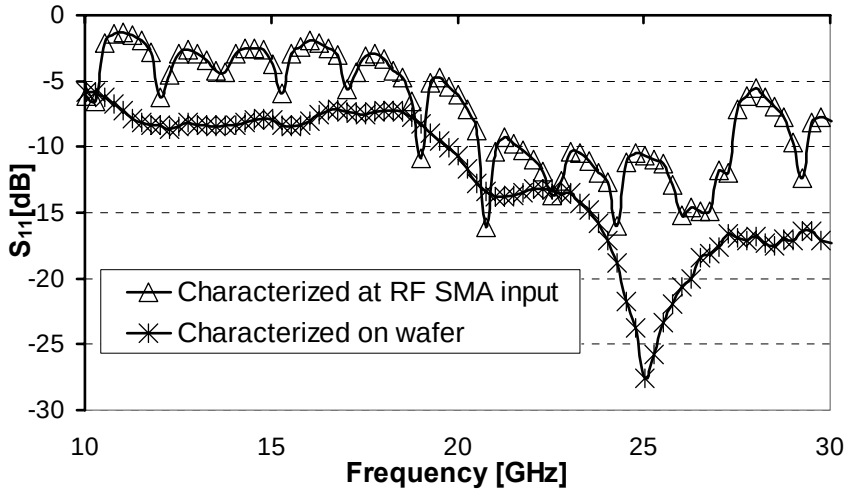


Figure 4.31: RF input reflection coefficient

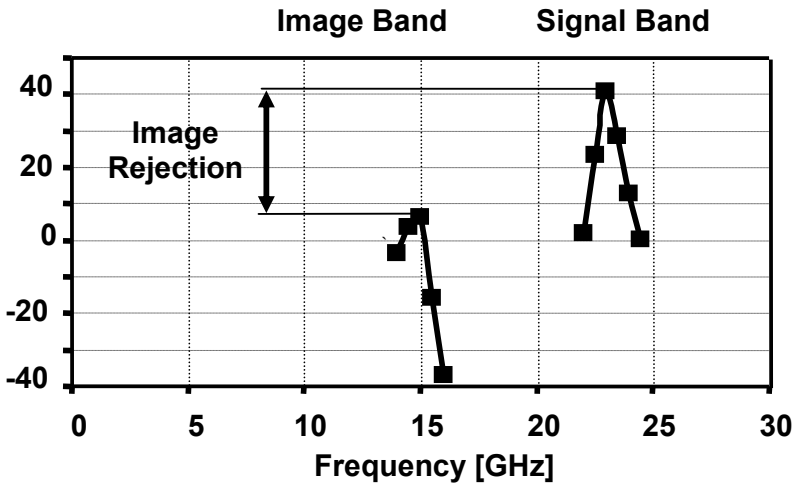


Figure 4.32: Single-path receiver gain

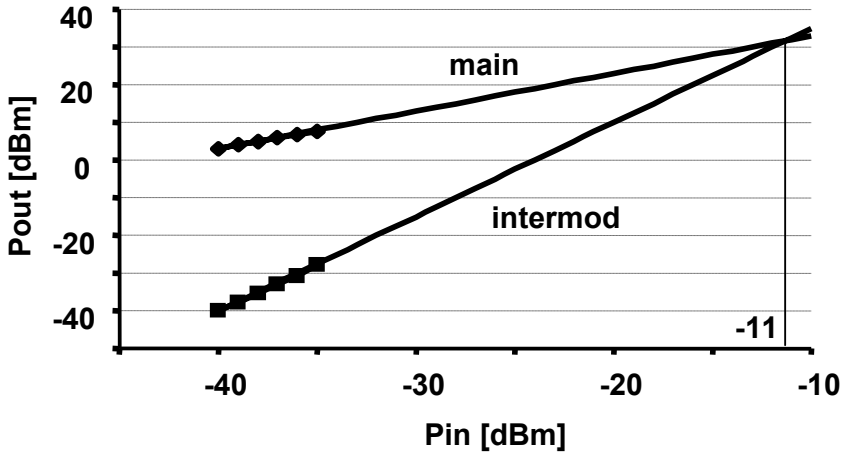


Figure 4.33: Two-tone measurement

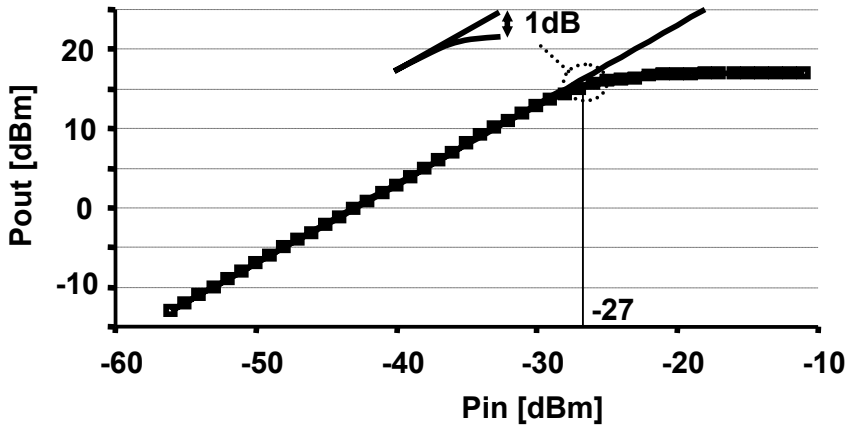


Figure 4.34 Gain compression

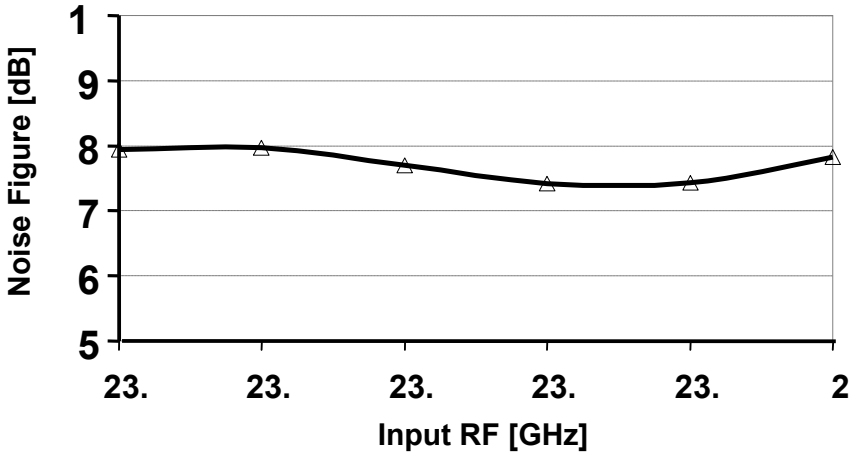


Figure 4.35: Single-path noise figure

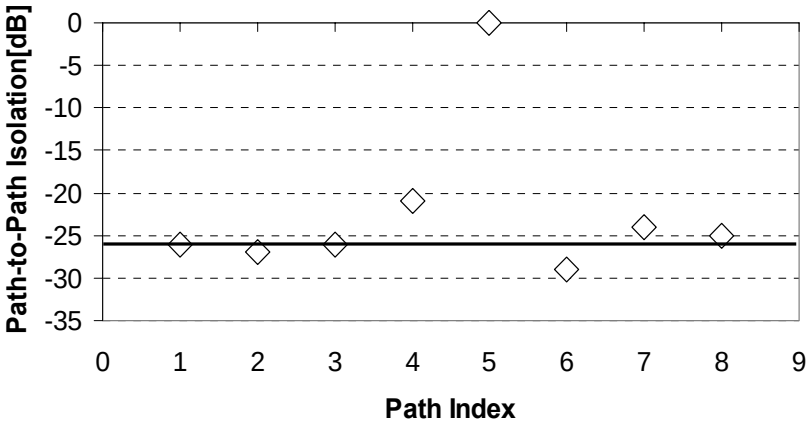


Figure 4.36: On-chip path-to-path isolation

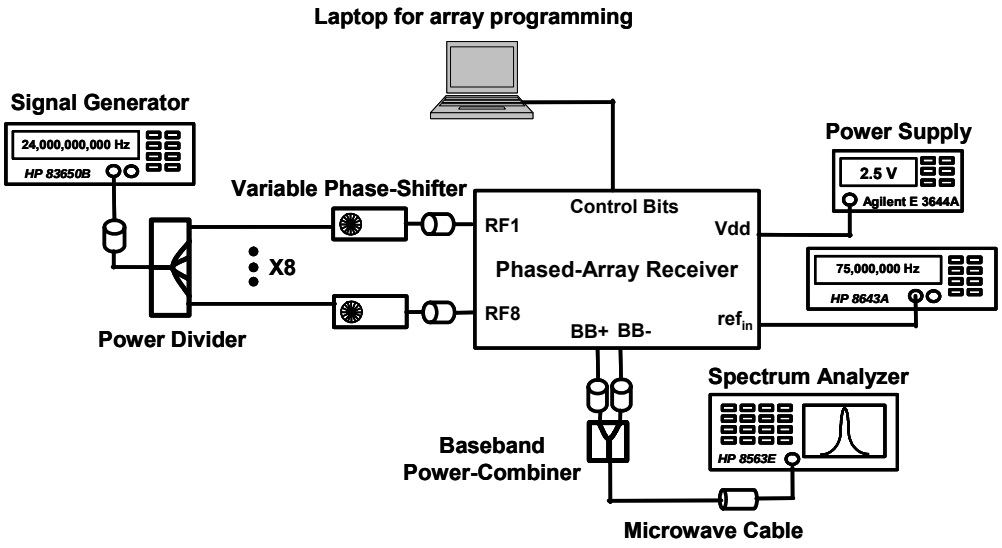


Figure 4.37: Test setup for characterizing array performance

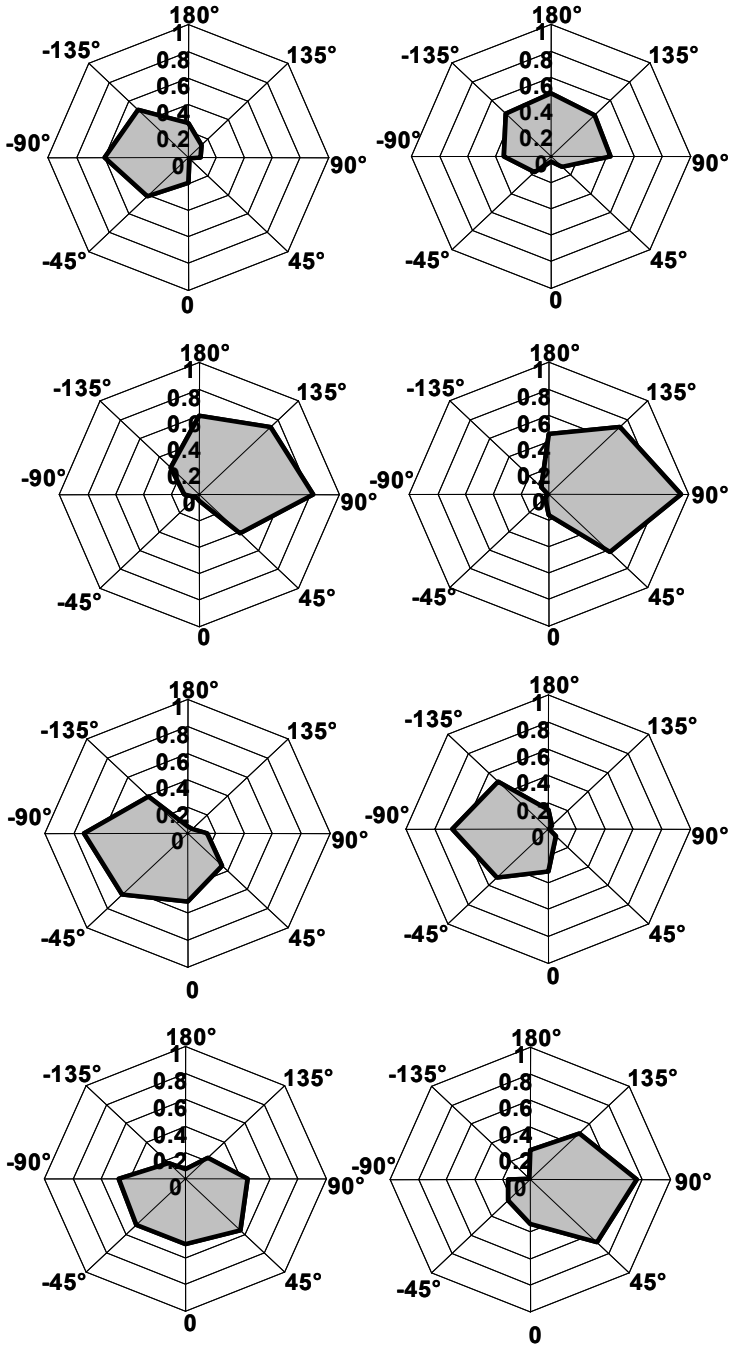


Figure 4.38: Normalized two-path array gain as a function of input phase difference at eight different LO settings

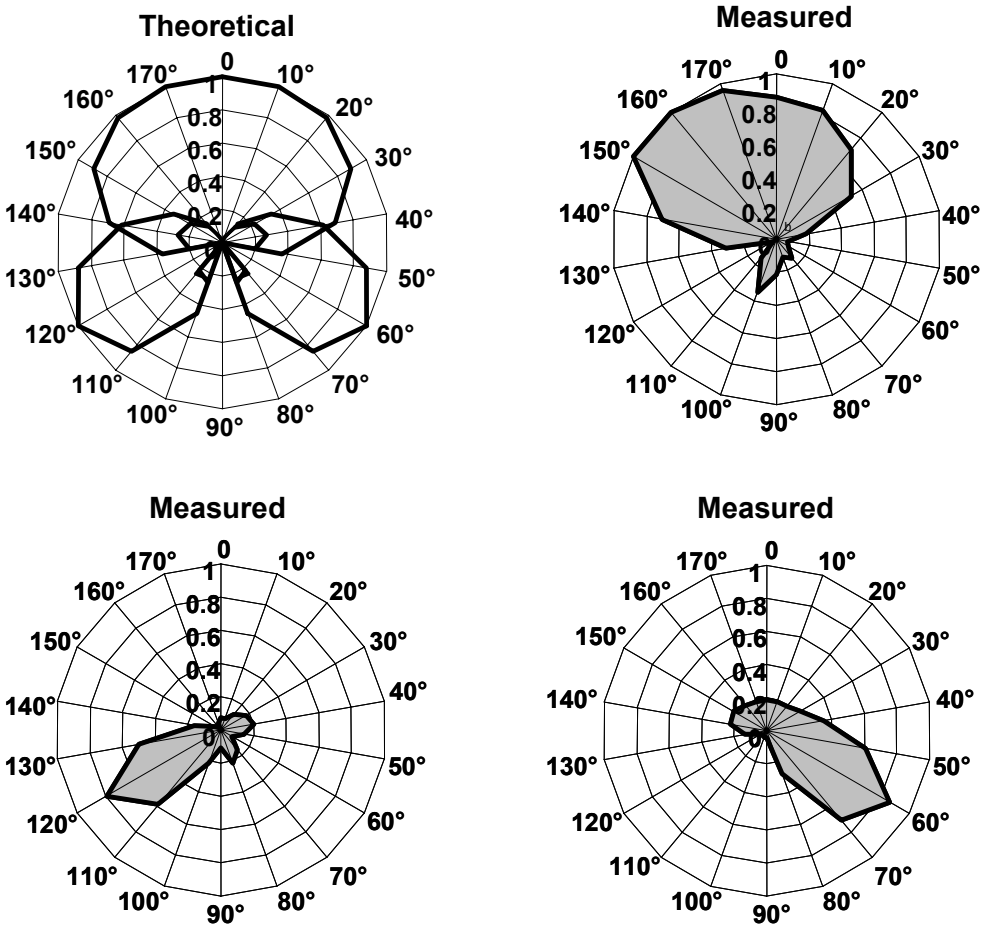


Figure 4.39: Normalized four-path array gain as a function of incident angle at three different LO settings compared to theoretical results

Signal Path Performance (per path)	
Peak Gain	43dB
Noise-Figure	7.4dB
Input-Referred 1dB Compression Point	-27dBm
Input-Referred 3rd-Order Intercept Point	-11.5dBm (2 tones 5MHz apart)
On-chip Image Rejection	35dB
S11	< -10dB
LO Path Performance	
Synthesizer locking range	2GHz
Synthesizer bandwidth	7MHz
Synthesizer settling time	< 50μs
VCO phase noise	-103dBc/Hz @ 1MHz offset
Complete Receiver Performance (8 paths)	
Total Array Gain	61dB
SNR Improvement	9dB
Phase-shifting Resolution	11.25°
Beam-forming Peak-to-Null Ratio	20dB (measured for 4 paths)
Power Dissipation @ 2.5V	364mA 287mA (w/o biasing and baseband buffers)
Technology	SiGe, 120GHz HBT 0.18μm CMOS
Die Area	3.5mm x 3.3mm

Table 4.1: Summary of the measurement performance of the 24-GHz phased array receiver

4.6 Chapter Summary

A silicon-based fully integrated 24GHz 8-element phased-array receiver is demonstrated for the first time. The system architecture, receiver signal-path, frequency generation circuitry, and phase distribution technique have been addressed in this chapter. Each signal path achieves a gain of 43dB, a noise figure of 7.4dB, and an IIP3 of -11dBm. The 8-path array achieves an inferred array gain of 61dB and a peak-to-null ratio of 20dB, improving the signal-to-noise ratio at the output by 9dB.

Chapter 5

A 77-GHz Fully-Integrated SiGe Phased-Array Tranceiver[1]

Encouraged by the successful implementation of the 24-GHz phased array system in silicon-based technologies, we shift our research interests to fully-integrated phased array systems operating at even higher frequencies. This chapter presents a 77-GHz 4-element transmitter-receiver chip integrated in a SiGe process providing a f_T of 200GHz for HBT. Section 7.1 introduces the motivations and design challenges. Section 7.2 describes the system architecture. The circuit design is detailed in Section 7.3, following by Section 7.4 presenting experimental results. The chapter is summarized in Section 7.5. [2]

5.1 Introduction

The concept and application of automotive radar were introduced in Section 2.2.6. The operation frequency approved by the FCC for such applications include the 22-to-29-GHz range for ultra-wide band (UWB) short-range radar [101], and 76-to-77-GHz for frequency-modulated continuous wave (FMCW), or pulse-Doppler radar suitable for long-range operation [44]. In addition, Electronic Communications Committee (ECC) within the European Conference of Postal and Telecommunications Administrations (CEPT) has granted a 77-to-81-GHz window for automotive UWB short-range radar since 2005 [103].

In comparison with the 24-GHz band, the 77-GHz band operation provides the following advantages: 1) Operating at a higher frequency results in reduced antenna size and compact package. In particular, the wavelength at 77GHz on silicon is at the

[2]The 77-GHz phased array transceiver work is a joint work done by: Xiang Guan, Aydin Babakhani, Abbas Komijani and Arun Natarajan.

same order of chip size, making an on-chip antenna which can significantly reduced the cost of packaging and eliminating the associated parasitic effects a possibility. 2) Despite of the different frequency allocation policies in various districts, using the 77-GHz band for automotive radar application is a global trend, while 24GHz UWB band is not available in every country. 3) A global concern for utilizing 24GHz for consumer radio location is that it can potentially degrade meteorological and related environmental activities currently using the 23.6-to-24-GHz range which is very sensitive to interferences [104]. In contrast, operation at 77GHz is more compatible with other applications using the same frequency spectrum [105].

The concept of single beam autonomous cruise control (ACC) radar has existed for several decades, and systems with proposed functionalities have been commercially available in premium-class vehicles. However, the cost of such systems using traditional technologies such as discreet microwave module or MMIC is still significantly beyond the price that an average customer is willing to pay. A silicon-based integrated phased array solution can potentially provide a low-cost, high-yield solution required by any type of mass production. By integrating the microwave front-end, analog signal processing, digital signal processing, and frequency generation on the same chip, the costly assembling process is dramatically simplified and the reduced number of off-chip components implies a lower power consumption of the system.

Although the current efforts at the 77-GHz range are focused on automotive radar, the 77-GHz phased array can potentially be used for other applications, such as short-range surveillance, microwave imaging, and ultra high-speed data transmission. The objective of this project is to demonstrate a general purpose fully-integrated phased array transceiver operating at 76 – 81 GHz that can be used in both wireless communication and short range radar. The design challenges of such systems include accurately modeling the components and parasitic at microwave range, routing the microwave signal over high-loss silicon substrate, finding appropriate methods to perform signal combining, signal distribution and phase shifting, achieving a low noise performance at receiver and providing sufficient W-band output power at transmitter, implementing ultra-high speed frequency generating blocks such as VCO and frequency divider, and realizing highly efficient on-chip antennas.

5.2 System Architecture

The 77GHz 4-element phased-array transmitter-receiver chip integrates the completed signal transmission paths, reception paths, signal distribution and combination, LO signal generation and distribution, phase shifting elements, and 77-GHz antenna on a single silicon die. Figure 5.1 illustrates the system block diagram.

In the transmitting path, quadrature upconversion is used to transfer a signal from baseband to 26-GHz IF while rejecting image interference. The IF signal is symmetrically divided into four radiating paths via binary distribution structure consisting of IF buffers and transmission lines. The RF mixer in each path upconverts the signal to 77-GHz using LO frequency at 52 GHz. The carrier phase shift due to the propagation delay is compensated at the LO port of each RF mixer using an analog phase shifter. Finally, the signal power in each path is boosted to the desired level by a 77-GHz PA and radiated off with an on-chip dipole antenna.

The receiver uses a frequency translation plan opposite of the transmitter's so that they can share the same frequency generation circuitry. Each RF front-end consists of an on-chip dipole antenna, LNA, mixer, and IF amplifier. The phase shifting is performed at the LO port of the mixer at 52-GHz with an analog phase shifter. By switching the digital control bit, the gain of the IF amplifier can be varied by 15dB so that the system dynamic range is enhanced. The 26GHz signals are combined using a symmetric active combining amplifier. The combined signal is further downconverted using a quadrature IF-to-baseband mixer.

The first LO signal at 52-GHz is generated using a voltage-controlled oscillator. To reduce the VCO power and area-cost of the LO distribution, only a differential phase is generated by the core oscillator and distributed across the chip. The transmission line loss is compensated by the inter-link LO buffers. The continuous analog phase shifting is performed locally at each path by an analog phase shifter, allowing continuous beam steering capability and accurate compensation of the phase and amplitude mismatch between each path caused by asymmetry in phase distribution

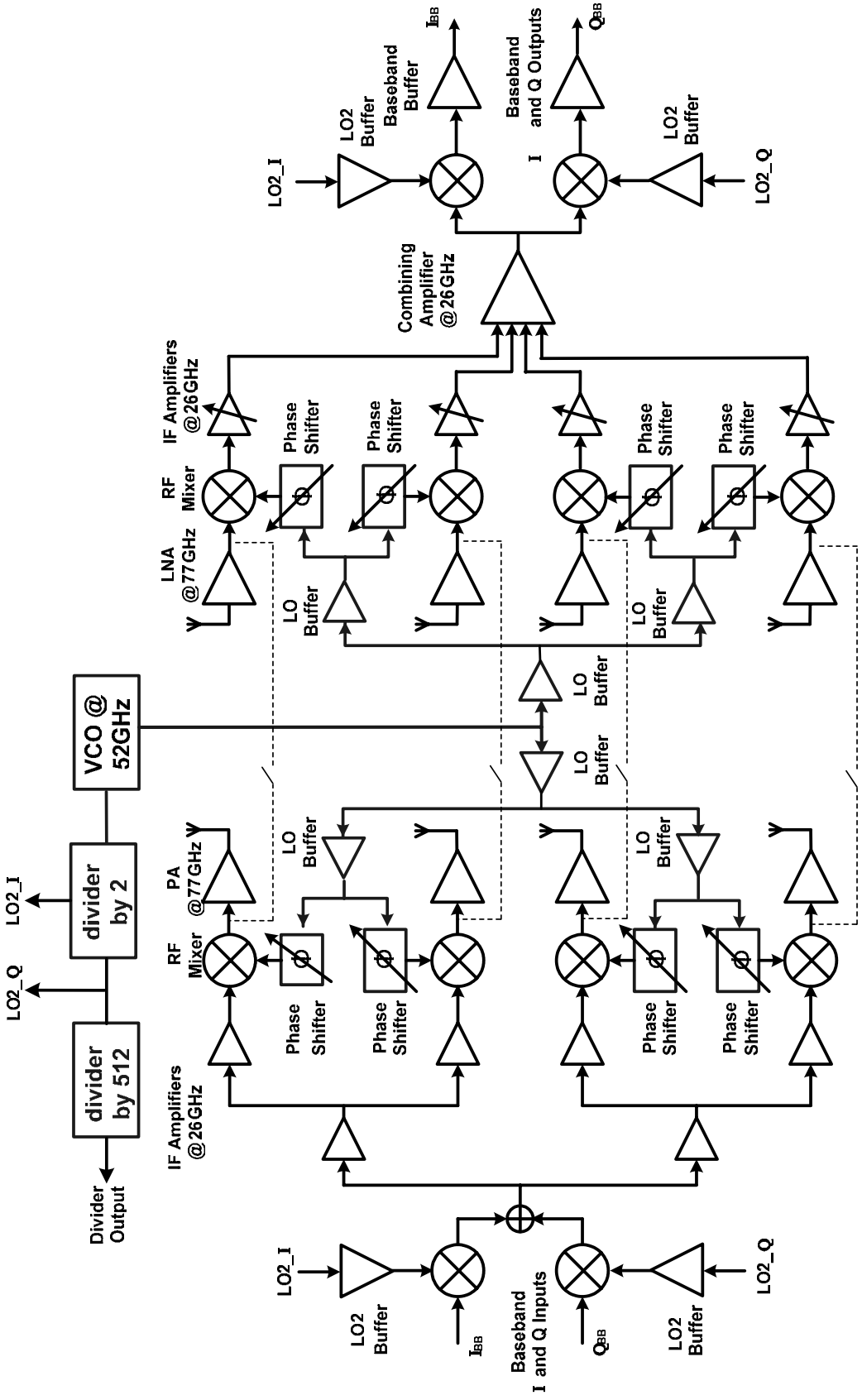


Figure 5.1: A fully-integrated 77-GHz phased-array transmitter-receiver

and antenna elements. The quadrature phase of the 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. Ideally, a fully-integrated PLL can be implemented to lock the VCO phase to a 50MHz external reference. Due to the time constraints, in the first prototype of this transceiver the PLL is completed using an off-chip phase detector and loop filter.

A loop-back mode is also created on-chip, directly connecting the output of the RF mixer in the transmitter to the input of the RF mixer in the receiver in each path. When the chip is switched to this mode, a four-input-four-output upconversion-downconversion link is formed which can be used to perform baseband-to-baseband measurement with no requirements of microwave equipments. This measurement is particularly convenient and informative in evaluating the array pattern, beam steering, and data-rate capabilities of the system.

It is noteworthy that the frequency plan of this system allows for the development of a dual-mode automotive radar system in the future. The first mode is operating at a 76-to-81 GHz radar band, and the second mode is operating at 22-to-29GHz radar band by bypassing the RF input to IF input of the system. Thus, this general-use system can utilize both radar bands for diverse applications subject to various specifications.

The 77-GHz transmitter-receiver chip was co-designed with A. Babakhani, A. Natarajan, and A. Komijani. This chapter is mainly focused on the circuitry designed by the author, which includes the completed frequency downconversion path from 77-GHz mixer to the baseband and the 52-GHz-to-50-MHz frequency divider chain.

5.3 Circuits Design

5.3.1 A 77-to-50-GHz Mixer

A double-balanced Gilbert-type mixer is employed in each RF path to downcovert a 76~81GHz RF signal to 52GHz IF, as depicted in Figure 5.2. To maximize signal power transfer and ease the measurement of the individual block separately, the differential output of the LNA and RF input of the mixer are both matched to 100 Ω

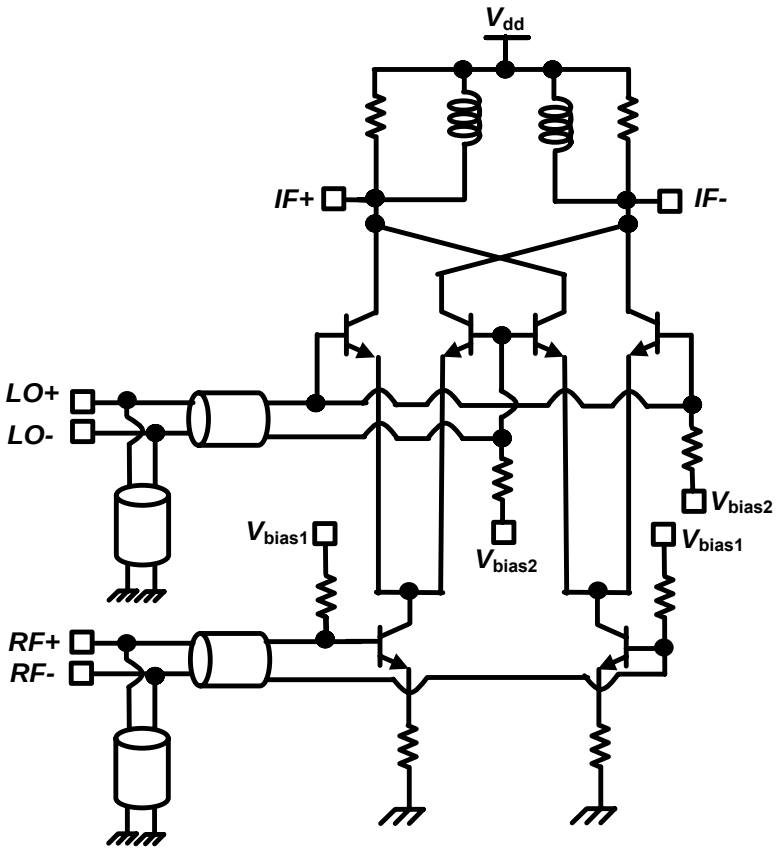


Figure 5.2: 77-to-26-GHz Mixer

as well as the LO port of the mixer. The impedance matching at both the RF and LO ports of the mixer is realized using transmission line stub tuning [90]. Simulation shows an input return loss of -20dB at RF port and -11dB at LO port. Since the differential resistance seen into the base of the RF and LO differential pairs are larger than 100Ω, a voltage gain of 3.5dB is achieved via the passive RF input matching network, and a 3.8dB gain is achieved via LO matching network. To save the chip area, resistive emitter degeneration instead of inductive degeneration is used to enhance the linearity. The common-node of the degeneration resistors are connected to the ground instead of a tail current source for better linearity.

We targeted at a minimum 3-GHz bandwidth of the whole receiver path. The receiver consists of 5 gain stages. Consider a fifth order low-pass system whose transfer function is given by

$$H(s) = \frac{1}{(1+\frac{j\omega}{\omega_p})^5} \tag{5.1}$$

where ω_p is the -3-dB bandwidth of each gain stage. It can be derived that the -3-dB bandwidth of this system is approximately $0.4\omega_p$. Therefore, roughly 8 GHz of bandwidth is desirable of each gain stage. We choose a 10-GHz bandwidth specification for this mixer for sufficient margin.

The bandwidth of this mixer is primarily determined by the quality factor Q of the resonant load at 26 GHz. The -3-dB bandwidth BW of the load impedance is related to Q via

$$BW = \frac{\omega_o}{Q} \tag{5.2}$$

where ω_o is the resonant frequency. The load impedance Z_L at resonance is given by

$$Z_L \approx \omega_o LQ \tag{5.3}$$

Therefore, the choice of Q is a trade-off between bandwidth and gain. To achieve the desired bandwidth, a maximum Q of 3.5 is allowed, according to which we choose 0.4nH inductance and 250Ω de-Q resistance to form the load with the capacitance of the transistor parasitic and input impedance of the subsequent stage. Simulation shows a 5-dB voltage gain is achieved.

5.3.2 A 26-GHz Two-Mode Amplifier

A differential resistively degenerated cascode is used as the 26-GHz amplifer, as shown in Figure 5.3. A differential current-bleeding branch consisting of Q_2 and Q_3 is added. The dc bias voltage at the base of Q_2 and Q_3 can be toggled between two values by digital switches, corresponding to a high-gain and a low-gain mode of the amplifier. In high-gain mode, Q_2 and Q_3 are off. In low-gain mode, the gain normalized to its high gain value is approximately given by

$$\frac{A_{v,low}}{A_{v,high}} = \left(1 + \frac{A_2}{A_4}\exp\left(\frac{V_{B,Q_2} - V_{dd}}{V_T}\right)\right)^{-1} \tag{5.4}$$

where A_2, A_4 are the emitter area of Q_2 and Q_4, respectively. In this design, V_{B,Q_2} at low-gain mode is set to V_{dd}, and A_2 / A_4 is fixed at 11/3. Equation (5.4) predicts a

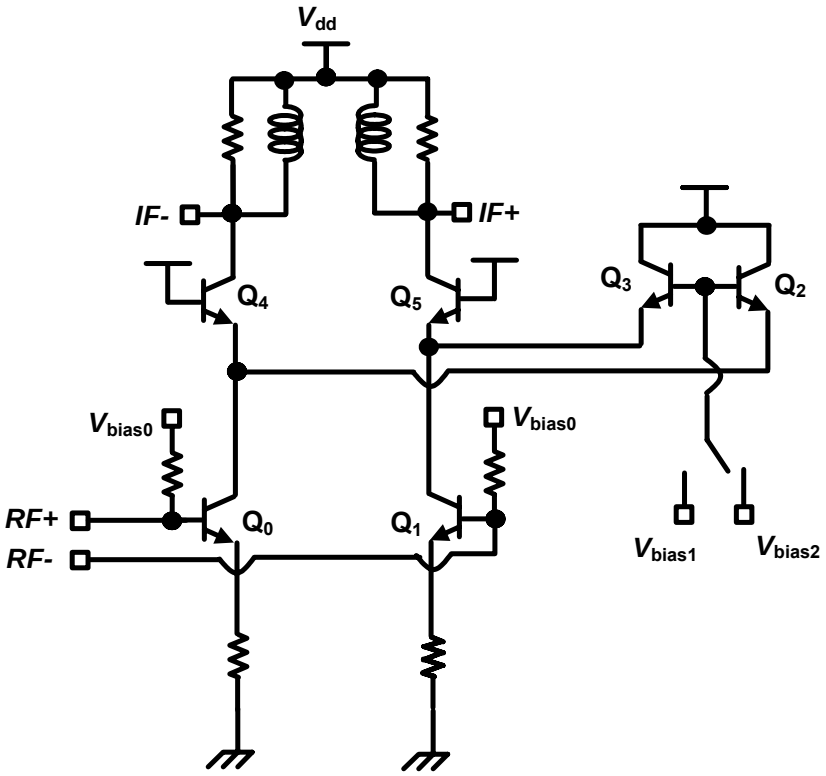


Figure 5.3 26-GHz two-gain mode amplifer

13.5dB gain variation between the two modes. The simulation result shows a 15dB gain variation. The 1.5dB discrepancy is due to the loss through parasitic capacitance at the cascode node.

It is noteworthy that a digital-to-analog converter (DAC) can be used to choose the bias voltage of Q_2 so that a variable gain 26-GHz amplifier with finer resolution can be implemented.

5.3.3 A 26-GHz Signal Combining Amplifier

The 4-path 26-GHz signals were combined through an active combining amplifier, as shown in Figure 5.4. The differential transconductors with resistive degeneration convert the 24-GHz signal from voltage domain to current domain. The current output of each transconductor is symmetrically routed to the combining node via a two-stage binary structure. A pair of cascode transistors is inserted at each combining junction, isolating the input ports and output ports. The total length of each routing

transmission line T_1 is 340μm and that of T_2 is roughly 2.55mm. Both T_1 and T_2 use differential transmission line structures with ground and side metal shield to minimize substrate loss and cross coupling. The differential output of the amplifier is loaded with an LC tank. Since the parasitic capacitance and the load capacitance at outputs are quite large at 24GHz, no additional capacitor is added to the tank. Two de-Q resistors are added in parallel with the LC tank to provide large bandwidth and reduce the sensitivity of the gain to the parasitic capacitance value. Compared to the passive combining structure discussed in Section 4.2.3, the active signal combining provides the following advantages:

1. The active input stages compensates for the transmission line loss.
2. The output ports are isolated to the transmission line impedance by cascode transistors. For the proposed passive combining structure in Figure 4.8, the transmission line impedance should be scaled down by half at each level as a requirement of matching. In addition, the output load impedance should be matched to the impedance of the transmission line at the top level, which is very low. Since the current remains the same, the signal voltage is not amplified but attenuated. For example, for an eight-path design the voltage loss is 9dB from any input port to the output port. On the contrary, in the active combining structure in Figure 5.4, the cascode transistors isolate the input and output ports of each T-shaped combining junction. At the outputs, the impedance seen into the collector of the cascode transistor is very high, therefore, a high impedance load can be directly applied as desired by a voltage amplifier.
3. To remove the LO feedthrough, upconverted signals and harmonic distortions, a tuned load is desired at the output. An accurate center frequency of this tuned load is one of the most important design considerations. For the passive combining structure in Figure 4.8, the impedance seen into its output port depends on the source impedance, the geometric properties of the transmission lines, and the EM coupling between them, which is very difficult and time-consuming to model accurately. In contrast, for active combining structure, thanks to the isolation provided by the cascode transistors, the parasitic capacitance at the output is only comprised of the collector-base junction capacitance C_μ and the collector-substrate capacitance C_{cs}, which are generally

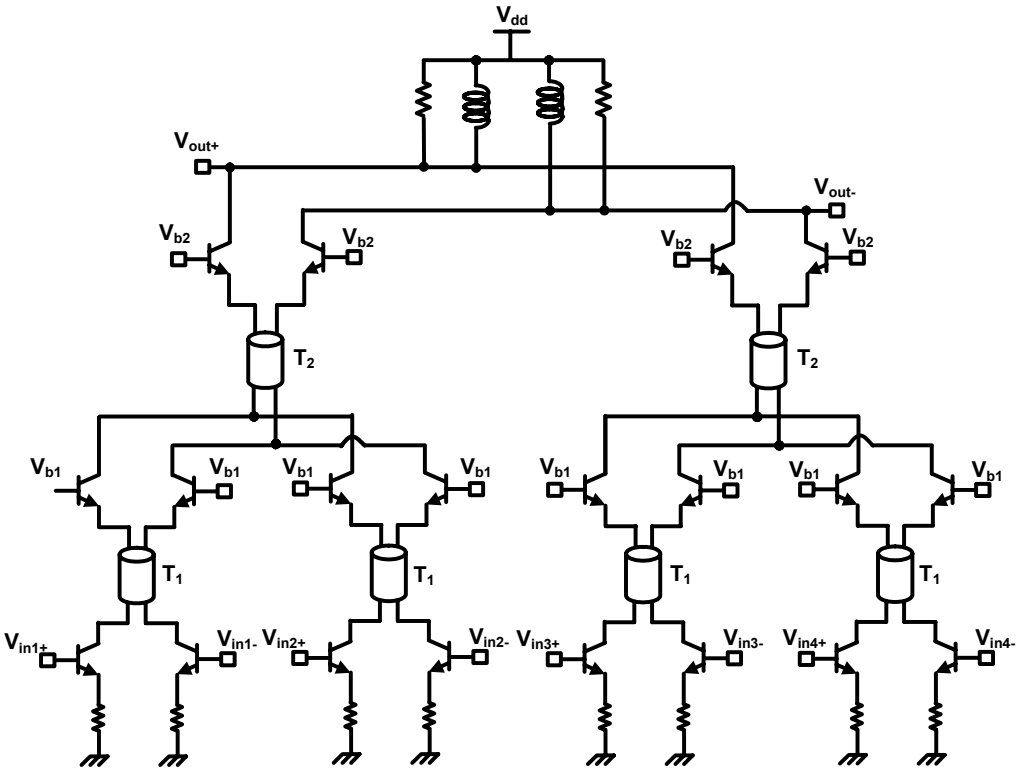


Figure 5.4: A 26-GHz 4-element signal combining amplifier

well-modeled and verified by the foundry, facilitating the design of the resonant tank.

4. The transmission line termination is provided by the emitter-base impedance of the cascode transistors. The transconductance g_m of the bipolar transistor is given by

$$g_m = \frac{I_C}{V_T}$$

where I_C is the collector bias current, which can be chosen so that $1/g_m$ is matched to the transmission line impedance. The emitter-base admittance $j\omega_o c_\pi$ is much smaller than $1/g_m$ if the transition frequency ω_T is much higher than ω_o. Therefore, good matching can be achieved even without additional passive tuning. It is noteworthy that the dc current is doubled after each combination, hence the transmission line impedance need to be reduced by half accordingly. In this work, 1mA dc bias current is applied at each

branch. Simulations show that the input return loss is below -10dB at the emitters of the cascode transistors at both level.

5. Multi-level cascode transistors significantly improve the reverse isolation and thereby the stability of the 24GHz amplifier. If the number of cascode levels is limited by the supply voltage headroom, the cascode transistors can be employed at the final output level, while the rest of the part uses passive structure. All benefits still exist except additional stability improvement.

5.3.4 IF-to-Baseband Mixer and Buffer

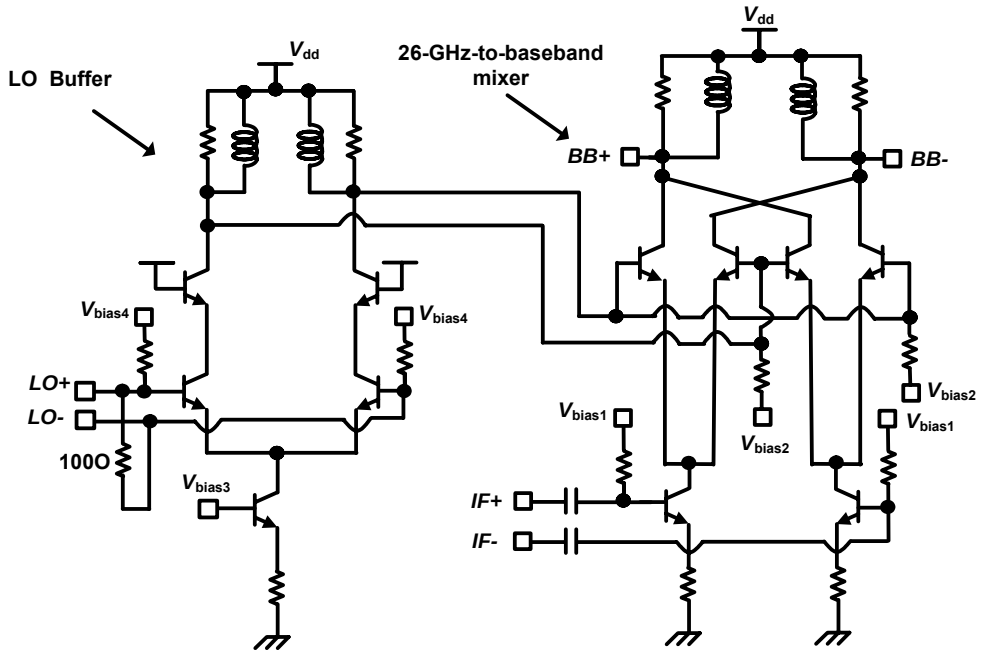


Figure 5.5: 26-GHz-to-baseband mixer and 26-GHz LO buffer

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 5.5. The 26-GHz signals are coupled into the mixer transconductance stage though 0.9pF MIM capacitors. The input differential pair is degenerated with 30Ω resistors at the emitter to improve linearity.

The LO port of the mixer is fed by a 26GHz 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 200mV so that the mixer switches completely. The input matching of the LO buffer is provided by a 100Ω resistor directly connected between the differential inputs. Although a matching network composed of inductors and capacitors can provide additional voltage gain, this solution is prohibited by the limited silicon area. The LO buffer is loaded with 0.6nH spiral inductors and 320Ω de-Q resistors, providing a gain of 15dB. With a 280Ω load resistor, the second mixer achieves a 6dB conversion gain and 8-GHz IF-referred bandwidth. The mixer core consumes 4mA dc current and the LO buffer drains 1mA. An emitter follower consuming 7mA dc current is used at each baseband output to drive 50Ω load, as shown in Figure 5.6.

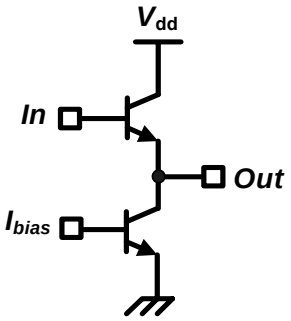


Figure 5.6: Baseband output buffer

5.3.5 A 52-GHz-to-50-MHz Frequency Divider Chain

In integrated systems, two types of frequency dividers are commonly used, namely a digital frequency divider and a injection-locked frequency divider. The digital frequency divider, as shown in Figure 5.7, consists of two D-type flip flops (DFF) connected as a ring. To achieve fast operation, DFF is commonly implemented with emitter coupled logic (ECL). This type of frequency divider can achieve a wide dividing range, consume less silicon area, and facilitate cascading. Moreover, the quadrature signals are inherently generated at the outputs due to the symmetry of two DFF. However, a high-speed digital frequency divider usually consumes a large amount of current and is thus less attractive for low-power design.

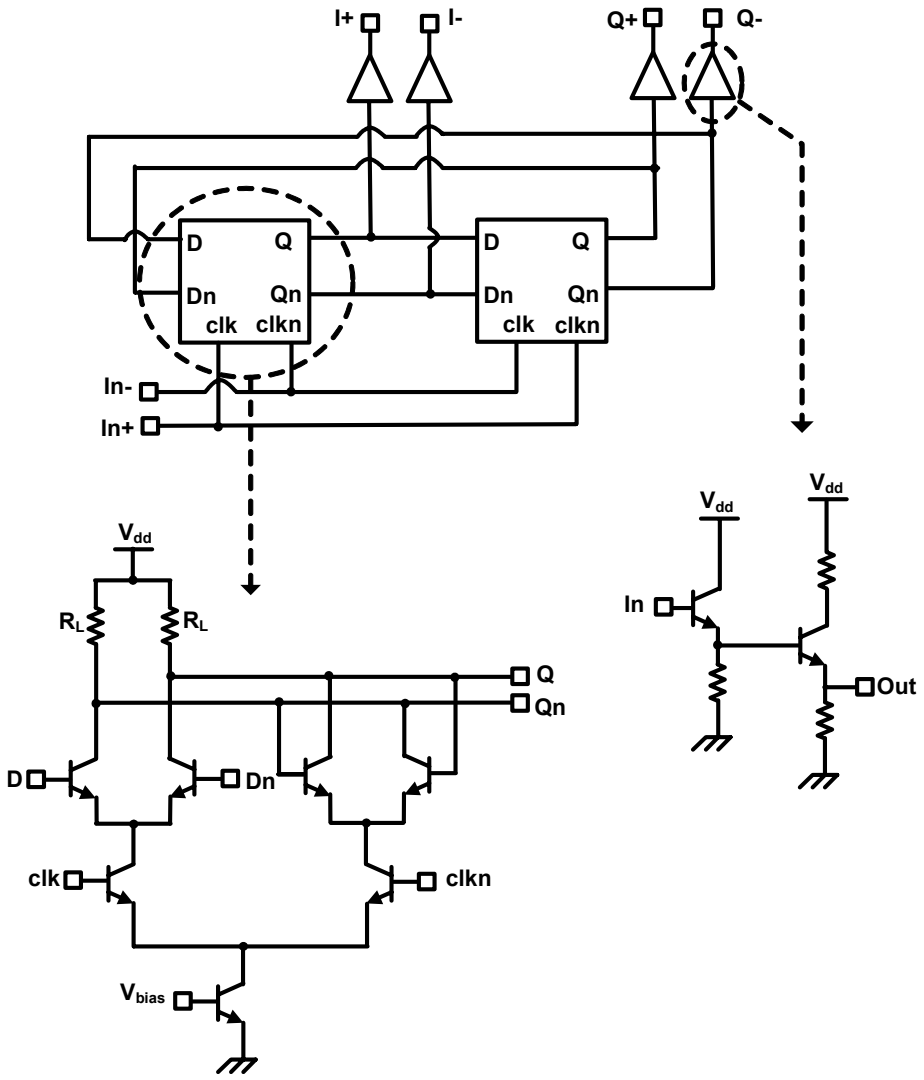


Figure 5.7: A digital frequency divider using emitter coupled logic DFF

The alternative is to use an injection-locked frequency divider (ILFD), as shown in Figure 5.8 (a), which appropriately operates only in a limited frequency range but consumes much less power than its digital counterpart. The ILFD is basically an oscillator whose oscillation frequency is locked to the frequency of the injected signal by the nonlinear feedback mechanism in an oscillating circuit. The design challenges for ILFD are to accurately locate the narrow locking range at the desired band and maximize the locking range within the power budget.

The 52GHz-to-50MHz frequency dividing is realized by cascading ten divide-by-2 blocks. Among them, the 52-to-26-GHz frequency divider operates at the highest

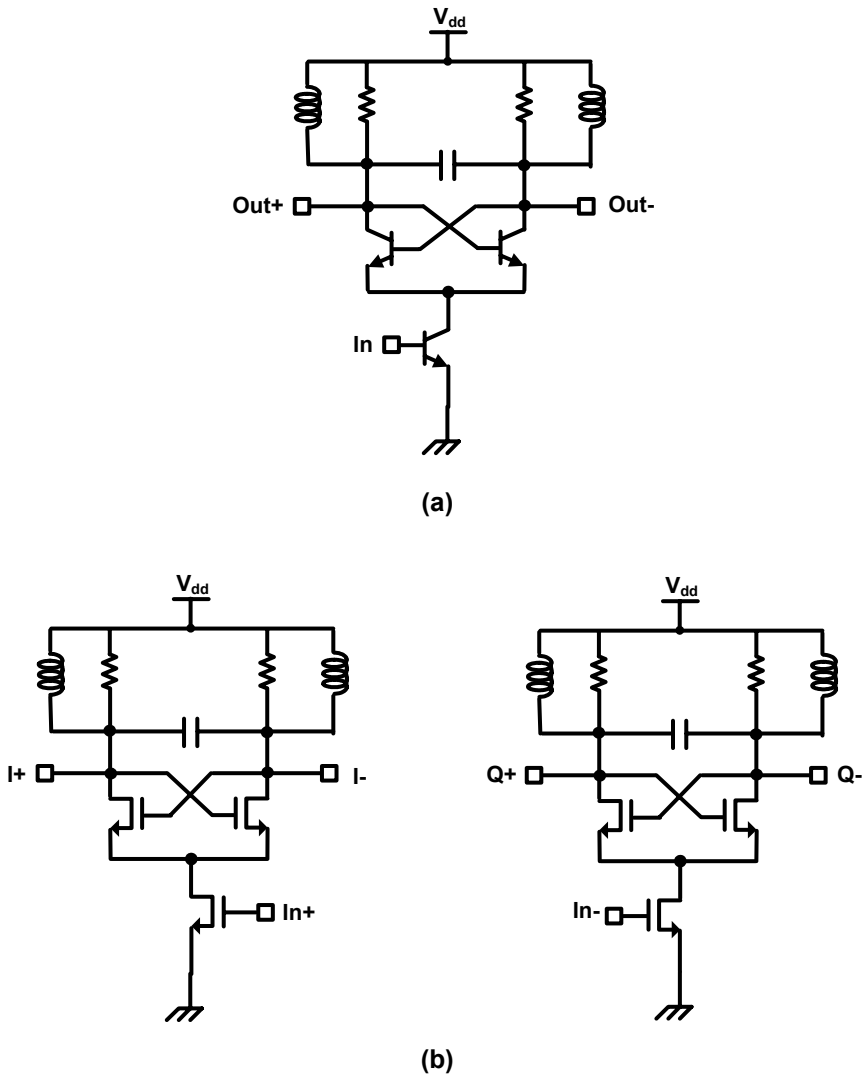


Figure 5.8: Injection locked technique (a) A differential injection-locked frequency divider (b) A quadrature injection-locked frequency divider proposed in [106]

frequency of the whole chain, and is thus the most difficult to implement in silicon. This divider not only needs to realize dividing function over sufficient frequency range but also provide quadrature signals at the outputs to drive I/Q IF-to-baseband mixers. We choose ILFD topology in this work for low power consideration. A simple ILFD topology shown in Figure 5.8 (a) does not provide the quadrature outputs. A quadrature ILFD (QILFD) as shown in Figure 5.8 (b) has been reported [106], which was implemented by driving two separated ILFD with opposite phases. However, due to the symmetry of the circuits, the signs of the differential outputs are not well defined.

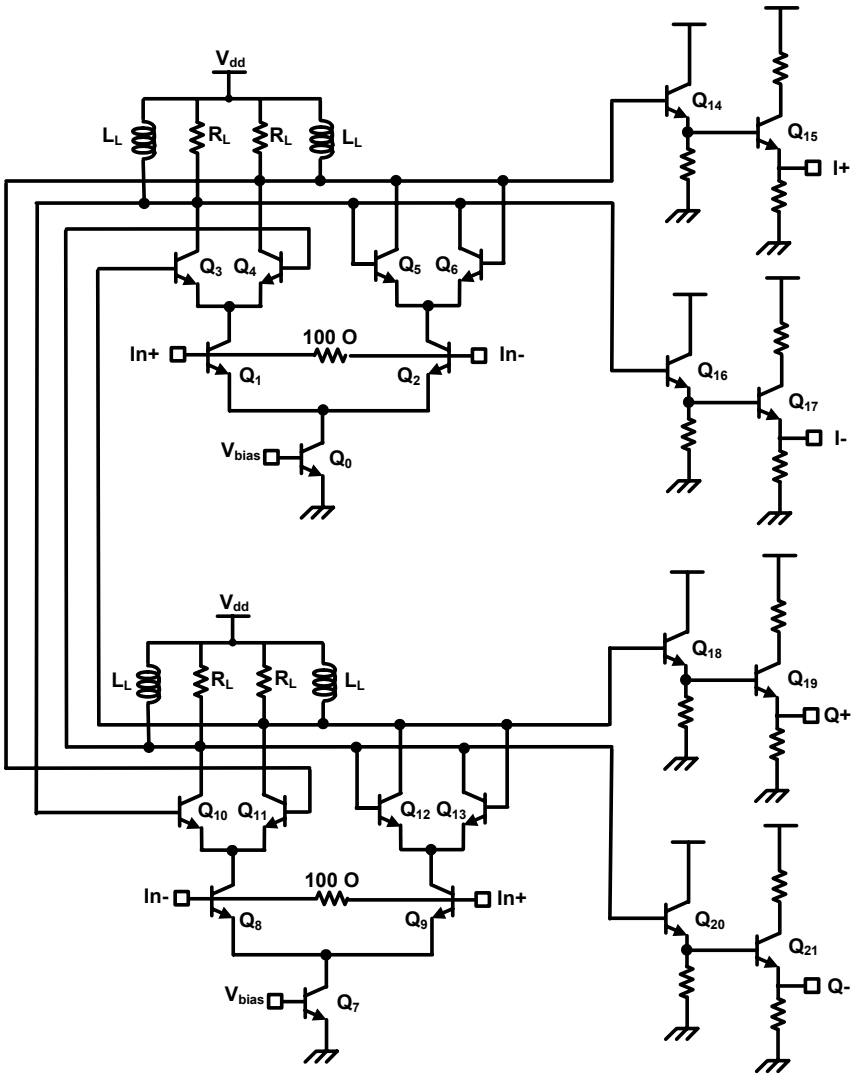


Figure 5.9: A cross-coupled quadrature frequency divider with output buffer

The topology of QILFD in this work is devised based on the architecture of a quadrature oscillator, as shown in Figure 5.9. The input signal is injected at the tail current of a basic cross-coupled quadrature oscillator formed by the transistors Q_0 ~ Q_{13}. Due to the symmetric cross coupling between M_3, M_4, M_{10}, and M_{11}, the relative phase of the four outputs are clearly defined. A 100Ω real resistor connected between the differential inputs provides impedance matching to the driving transmission line. A two-stage buffer is used at each output to provide the desired DC level for the cascaded blocks, isolation between the load and the oscillating core, and driving capabilities for the 50Ω load. The drain resistor at the second buffer stage is used to prevent the collector-base voltage of the output transistor exceeding the breakdown

threshold. The divider core consumes 3 mA in total, while each two-stage buffer consumes 7.2 mA.

Interestingly, this architecture can also be viewed as a digital frequency divider which is modified by replacing its resistive loads with the inductive loads. The digital and analog frequency dividing techniques are merged at very high frequencies. Since the driver of this divider is a 52-GHz VCO, providing sinusoidal instead of square-wave signals, and the divider is able to track a low-power input which does not completely switch the input transistors around the self-oscillation frequency, it is more proper to analyze this circuit using general ILFD theories than using the theories applied to digital frequency divider.

The locking range of ILFD is a function of the amplitude of the injected signal. Due to the symmetry, we only need to analyze the intermodulation components in the sum of drain current of M_3 and M_5. Let's define:

$$V_i(t) = V_{in+}(t) - V_{in-}(t) = V_i \cos(\omega_i t + \varphi) \tag{5.5}$$

$$V_{oI}(t) = V_{oI+}(t) - V_{oI-}(t) = V_o \sin(\omega_o t) \tag{5.6}$$

$$V_{oQ}(t) = V_{oQ+}(t) - V_{oQ-}(t) = V_o \cos(\omega_o t) \tag{5.7}$$

The ac current at the drain of M_3 can be expressed as

$$i_{d3}(t) = f\left(V_i(t), V_{oQ}(t)\right) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} K_{m,n} \cos(m\omega_i t + m\varphi)\cos(n\omega_o t) \tag{5.8}$$

where $K_{m,n}$ is the intermodulation coefficient of the m^{th} order harmonic of V_i(t) and the n^{th} order harmonic of V_{oQ}(t) [107]. Assuming the identical transistors are used for $M_3 \sim M_6$, the drain current of M_5 can be expressed as

$$\begin{aligned} i_{d3}(t) = f\left(-V_i(t), V_{oI}(t)\right) = &\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} K_{2m,n} \cos(2m\omega_i t + 2m\varphi)\sin(n\omega_o t) \\ &- \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} K_{2m+1,n} \cos\left((2m+1)\omega_o t + (2m+1)\varphi\right)\sin(n\omega_o t) \end{aligned} \tag{5.9}$$

For divide-by-2 ILFD, $\omega_i = 2\omega_o$. In addition, due to the bandpass selection of the resonant tank, only the frequency components around ω_o matter. Neglecting the intermodulation components beyond fourth order, we can write

$$i_{d3}(t) = K_{o,1}\cos(\omega_o t) + \frac{1}{2}K_{1,1}\cos(\omega_o t + \varphi) + \frac{1}{2}K_{1,3}\cos(\omega_o t - \varphi) \tag{5.10}$$

$$i_{d5}(t) = K_{o,1}\sin(\omega_o t) - \frac{1}{2}K_{1,1}\sin(\omega_o t + \varphi) - \frac{1}{2}K_{1,3}\sin(\omega_o t - \varphi) \qquad (5.11)$$

When $i_{d3}(t)$ and $i_{d5}(t)$ are added, the amplitude of each intermoduation product are enhanced by a factor of $\sqrt{2}$ because of the orthogonal phases, implying an increased locking range.

From the above analysis, we can see the crossed coupled pairs M_5, M_6 and M_{12}, M_{13} have two functions: 1) To generate a negative conductance at the load, which eases the self-oscillation at high frequencies. 2) To provide an orthogonal injection path for the input signal leading to an increased locking range. Although the configurations are same, their functions here differ from those when they are used in digital frequency divider to latch the output when the clock is off.

To save the silicon area, the rest of the divide-by-2 blocks are implemented by using traditional emitter-coupled digital frequency dividers as shown in Figure 5.7. The bias current of the divider core and the output buffers are successively reduced according to the operation frequency. The dc current of the 2nd, 3rd, and 4th divide-by-2 are 28mA, 14.4mA, and 13mA, respectively. Each of the remaining 6 dividers consumes 6.8mA. The whole 52-GHz-to-50-MHz divider chain consumes 130mA in total.

5.6 Experimental Results

The 77-GHz phased array transceiver was designed and fabricated by using IBM's 8HP SiGe BiCMOS process, providing an f_T of 200GHz for a HBT device and a 0.13-um CMOS transistor. The resistance of the silicon substrate in this process is approximately 13.5Ω.cm. Figure 5.10 shows the chip micrograph. It occupies an area of 6.8mm by 3.8mm.

Similar to the test setup of the 24GHz phased array, the 77-GHz chip and test board are mounted on the same brass substrate by using conductive adhesive. Gold bondwires are used to connect the power supply, ground, baseband inputs and outputs, divider outputs, and control terminal of the VCO on chip to the test board. Because the EM power is directly radiated and received by on-chip antennas, the microwave interface between the package and the chip is eliminated, allowing direct in-air test of

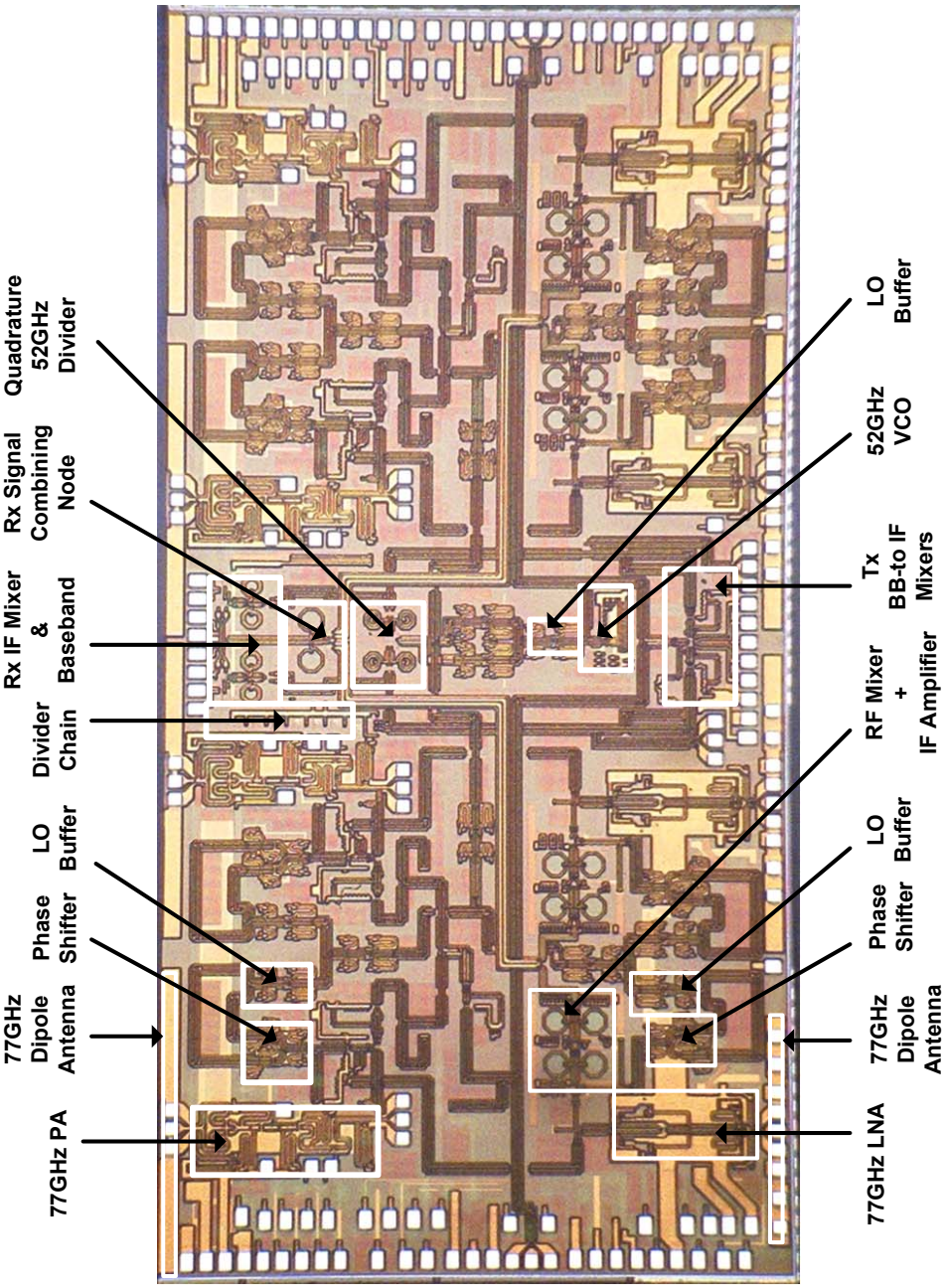


Figure 5.10: Die Micrograph of 77-GHz Transmitter-Receiver

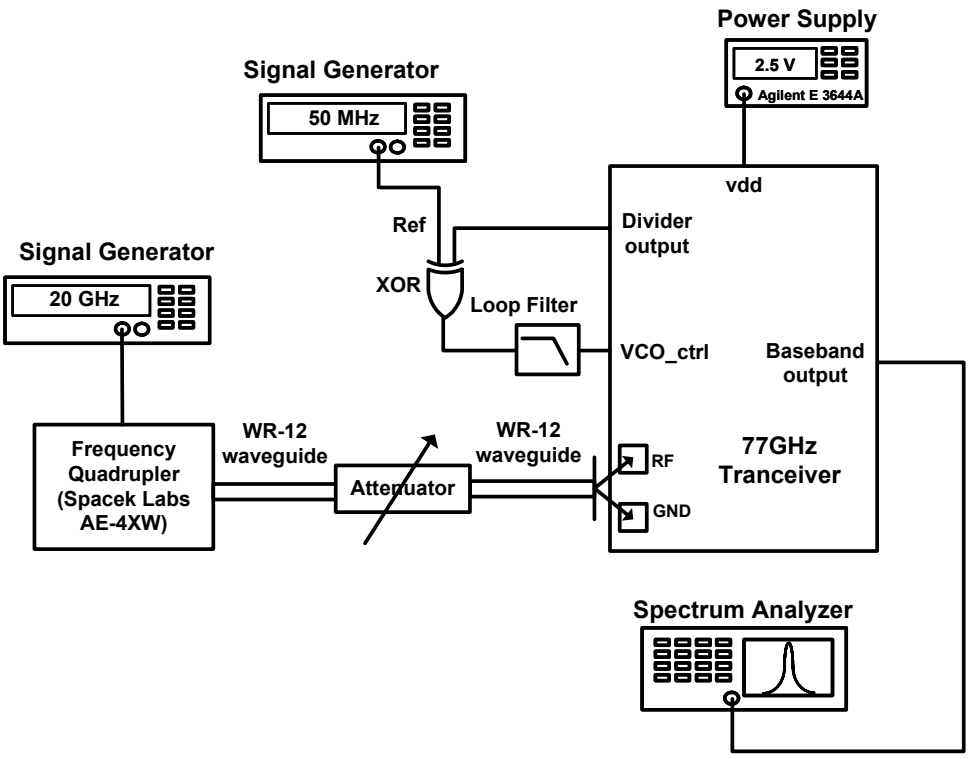


Figure 5.11: Receiver test setup

the transmitting and receiving pattern. The electronic performance of the receiver alone is characterized by cutting off the antenna and feeding the LNA input via a wafer-probe.

Figure 5.11 illustrates the test setup for measuring the receiver gain. The input signal at 77GHz range is provided by a frequency quadrupler (Spacek Labs AE-4XW) capable of delivering output frequency from 60-90 GHz. The input of the frequency quadrupler is supplied by an HP 83650B signal generator working up to 26.5GHz. 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 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 an Agilent E4418B power meter with a HPW8486A W-band power sensor. An exclusive OR (XOR) logic gate acting as a phase detector and a first order RC lowpass filter complete the PLL, which locks the phase and frequency of the 52-GHz VCO to a 50MHz reference provided by signal generator HP8643A. The baseband outputs are characterized using

an 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 Com NC5110 noise source.

Figure 5.12 shows the measured sensitivity curve of the frequency divider chain. The input of the first divider is driven by an Agilent E8257D signal generator. The input signal power is calibrated to the probe tips. The measurement results show the first frequency divider is self oscillating at 26.3 GHz. The tuning range shown in this curve is in fact limited by the available input power. When directly driven by the VCO buffer, the whole frequency divider chain can properly divides the VCO output frequency from 51.4-GHz to 55.5-GHz (7%) by a factor of 1024. The total frequency divider chain drains 143mA in total from a 2.5V supply.

Figure 5.13 shows the measured receiver gain (LNA+downconverter) as a function of the RF input frequency using LO signals $f_{LO1} + f_{LO2} = 78.87GHz$. Each downconversion path (including LNA) dissipates 60mA. A 41-dB single-path receiver gain is achieved at the center frequency of 80GHz with 3 GHz of bandwidth, and the inferred array gain is 53 dB. Figure 5.14 shows the DSB noise figure of a single-path receiver using the same LO frequencies. The lowest noise figure of 8.0dB is measured at 79GHz, and the average in-band noise figure is 8.6dB. Ideally the 4-element array improves SNR by 6dB

Some of the recent measurement results are summarized in Table 5.1.

Single-path receiver	
gain (single path)	41 dB
noise figure	8dB
Input-referred 1-dB compression point	-44dBm (high-gain) mode
current consumption (single-path)	60mA
LO frequency generation	
VCO tuning range	7 GHz
Frequency divider chain locking range	3.7 GHz (7%)
Divider chain current consumption	130 mA
Inferred array gain	53 dB
Inferred SNR improvement	6 dB
Supply Voltage	2.5V
Tranceiver die size	6.8 x 3.8 mm^2

Table 5.1: Summary of the recent measurement performance of the 77-GHz phased array transceiver (the receiver and the frequency synthesizer parts)

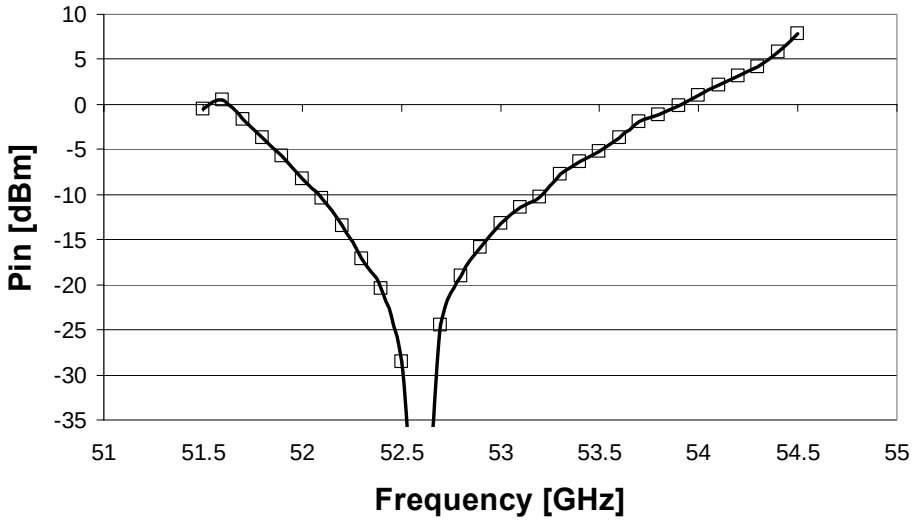


Figure 5.12: Divider chain sensitivity

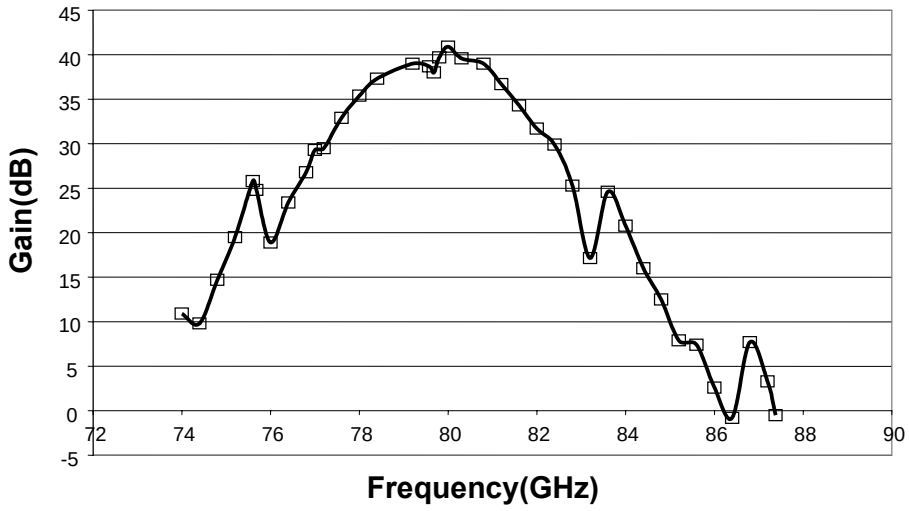


Figure 5.13: Receiver Gain

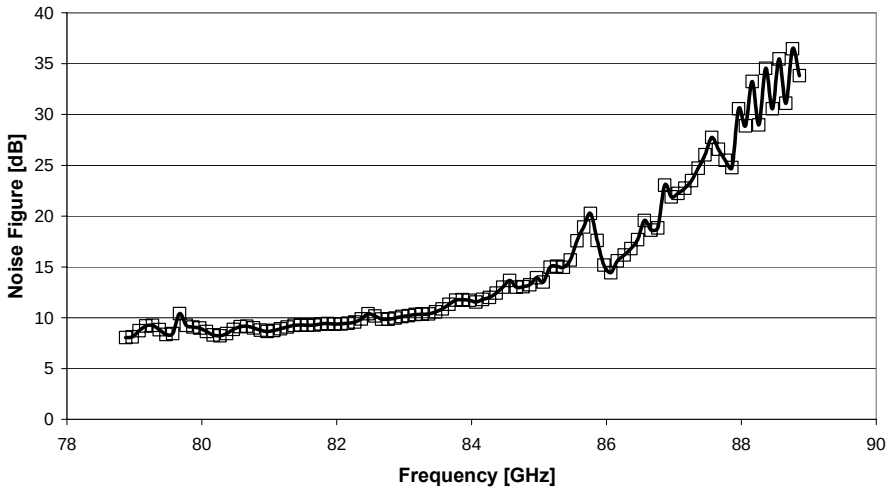


Figure 5.14 Receiver noise figure

5.5 Chapter Summary

The first integrated 77-GHz phased array transmitter-receiver system in silicon-based technology was implemented. LO phase shifting is performed at the LO ports of the RF mixers in both the transmitter and receiver. The receiver uses an active combining amplifier to sum the signals of 4-path. In a frequency generation path, a cross-coupled QILFD is designed to divide the VCO frequency by 2 while providing quadrature outputs. A wideband, low-noise, high-gain 77-GHz receiver in silicon has been demonstrated with measurement results. The measurement results also show a 7% locking range of the frequency divider chain.

Chapter 6

Conclusion

Phased array systems operating at microwave frequency range provide large bandwidth, compact antenna solution, spatial selectivity, and electronic beam steering that benefit high-speed data transmission and radar surveillance. This thesis explores various techniques to implement such systems in low-cost, high integration level, and high-yield silicon-based technologies. Three integrated receiver systems operating at the 24-GHz or 77-GHz range have been demonstrated in silicon for the first time. Some highlights of these works are summarized in following paragraphs.

A 24-GHz 0.18-μm CMOS front-end has been implemented using a novel LNA architecture, common-gate with resistive feedthrough. Theoretical analysis reveals that a large resistor in parallel with the common-gate transistor has little impact on its noise figure but can affect its input impedance significantly. An optimization procedure is developed based on this observation, leading to a better tradeoff between noise and power matching at the input. A thorough analysis on this new topology shows that it can obtain a lower noise figure at the perfect power matching compared to the traditional common-source LNA. Measurement results demonstrated a lower noise figure while consuming less power compared to the previously published works.

A fully-integrated 8-element phased-array receiver has been implemented in a SiGe HBT process providing a f_T of 120GHz for HBT and a 0.18-μm CMOS transistor. In this work, a LO phase shifting architecture is proposed using multi-phase VCO and an analog phase multiplexer to perform beam forming. Symmetric phase distribution is achieved by applying appropriate phase sequence in the transmission line array to minimize EM crosstalk. Multiple signal downversion paths, IF signal combining, and a fully integrated 19-GHz frequency synthesizer are demonstrated. Measured array patterns at various phase settings prove the spatial selectivity and beam steering capability of the system.

A 77-GHz wide-band phased array transceiver has been integrated in a SiGe HBT process providing a f_T of 200GHz for HBT and a 0.13-μm CMOS transistor. In this work, on-chip antenna is used for signal reception and radiation. The phase shifting is performed at LO ports of the RF mixers using continuous analog phase shifters. The author's efforts are focused on the receiver down-conversion path (excluding LNA) and a 52-GHz-to-25-MHz frequency divider chain. In the receiver part, signal combining is performed using an active combining amplifier. In the frequency synthesizer part, a novel cross-coupled quadrature injection-locked frequency divider is used to divide a 52-GHz VCO frequency by a factor of 2. Measurement results demonstrated a 41-dB gain and 3-GHz bandwidth for the receiver and a 7% locking range for the frequency divider chain.

In the very last paragraph of his seminal paper published in 1965 [108], Gordon Moore prophesied: "It is difficult to predict at the present time just how extensive the invasion of the microwave area by integrated electronics will be…The successful realization of such items such as *phased-array antennas*, for example, using a multiplicity of integrated microwave power sources, could completely revolutionize radar."

Forty years later, this prophecy was demonstrated for the first time by our work.

6.1 Recommendations for future work

Our work proves the feasibility of fully-integrated microwave phased-array receiver in silicon. Future trends would examine how to implement such system into products for specified applications, such as communication, radar, and microwave imaging, which demand more research efforts in system definition, circuit innovation and digital signal processing. To increase the number of array elements integrated on a single chip, more compact and lower power circuits, efficient signal combining and distribution methods, and system architectures that maximize circuits sharing need to be developed. Among various architectures, direct conversion phased array and a phased array with true time-domain compensation are particularly interesting for investigation.

Bibliography

[1] C. Shannon, “A Mathematical Theory of Communication,” *The Bell System Technical Journal*, vol. 27, pp. 379-423, 623-656, July-October 1948.

[2] A. J. Joseph, et al., “Status and direction of communication technologies – SiGe BiCMOS and RFCMOS,” *Proceedings of the IEEE*, vol. 93, no. 9, pp. 1539-1558, Sept. 2005.

[3] A. Hajimiri, et al., “Integrated phased array systems in silicon,” *Proceedings of the IEEE*, vol. 93, no. 9, pp. 1637-1655, Sept. 2005.

[4] K. McClaning and T.Vito, *Radio Receiver Design*, Noble Publishing Corp., 2000.

[5] J. B. Johnson, “Thermal agitation of electricity in conductors,” *Physical Review*, vol. 32, no. 7, pp. 97-109, July 1928.

[6] H. Nyquist, “Thermal agitation of electric charge in conductors,” *Physical Review*, vol. 32, no. 7, pp. 110-113, July 1928.

[7] P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*. John Wiley & Sons, 1993.

[8] H. A. Haus, “IRE standards on methods of measuring noise in linear twoports”, in *Proc. IRE*, vol. 48, pp. 60-68, Jan. 1960.

[9] H. T. Friis, “Noise figure of radio receiver,” *Proceedings of the IRE*, vol. 32, no.7, pp. 419-422, July 1944.

[10] B. Razavi, *RF Microelectronics*, Prentice Hall, 1998.

[11] D. K. Shaeffer and T. H. Lee, *Low-Power CMOS Radio Receiver*, Kluwer Academic Publishers, Boston, 1999.

[12] S. Mirabbasi and K. Martin, “Classical and modern receiver architectures,” *IEEE Commun. Mag*., vol. 38, no. 11, Nov. 2000, pp. 132-139.

[13] Edwin H. Armstrong, “A new system of short wave amplification,” *Proceedings of the IRE*, vol.9, pp. 3-27, Feb. 1921.

[14] J. C. Rudell et al., “A 1.9-GHz Wide-Band IF Double Conversion CMOS Receiver for Cordless Telephone Applications,” *IEEE J. Solid-State Circuits*, vol. 32, no. 12, Dec. 1997, pp. 2071–88.

[15] J. Crols and M. Steyaert, “Low-IF Topologies for High-Performance Analog Front Ends of Fully Integrated Receivers,” IEEE Trans. Circuits and Sys. II, vol. 45, no. 3, 1998, pp. 269–282.

[16] A. A. Abidi, “Direct-Conversion Radio Transceivers for Digital Communications,” *IEEE J. Solid-State Circuits*, vol. 30, no. 12, pp. 1399–1410, Dec. 1995,.

[17] S. K. Reynolds, *etc*., “A direct-conversion receiver IC for WCDMA mobile systems," *IEEE J. Solid-State Circuits*, vol. 38, no. 9, pp. 1555-1560, Sept. 2003.

[18] Ali Hajimiri, *Jitter and Phase Noise in Electrical Oscillators*, Ph.D. Dissertation, Stanford University, 1998.

[19] B. Razavi, ed., *Monolithic Phased-Locked Loops and Clock Recovery Circuits*, Piscataway, NJ: IEEE Press, 1996.

[20] B. G. Goldberg, “The evolution and maturity of fractional-N PLL Synthesizers,” *Microwave J*., pp. 124-134, Sept. 1996.

[21] X. Guan, H. Hashemi, A. Hajimiri, "A Fully Integrated 24-GHz Eight-Element Phased-Array Receiver in Silicon," *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 2311-20, Dec. 2004.

[22] G. Durgin, *Space-time Wireless Channels*, Prentice Hall, 2002

[23] A. Lozano*, et. al*. "Lifting the limits on high-speed wireless data access using antenna arrays," *IEEE Commun. Mag*., vol. 39, no. 9, Sept. 2001, pp. 156-162.

[24]V. Aulock and W. H., "Properties of phased arrays," *Proc. IRE*, vol. 48, pp. 1715–1728, Oct. 1960.

[25] R. C. Hansen, Ed., *Significant Phased Array Papers*, Artech House, Norwood, MA, 1973.

[26] R. S. Elliott*, Antenna Theory and Design*. Englewood Cliffs, NJ: Prentice-Hall, 1981.

[27] M. Golio, Ed., *The RF and Microwave Handbook, Session 6.9*, CRC Press, Boca Raton, FL, 2000.

[28] D. Parker and D. C. Zimmermann, "Phased arrays—Part I: Theory and architectures," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 678–687, Mar. 2002.

[29] D. Parker and D. C. Zimmermann,, "Phased-Arrays—Part II: Implementations, Applications, and Future Trends," *IEEE Trans. Microwave Theory Tech*., vol. 50, no. 3, pp. 688–698, Mar. 2002.

[30] A. Hajimiri, *et. al*., "Phased array system in silicon," *IEEE Commun. Mag*., vol. 42, no. 8, Aug. 2004, pp. 122-130

[31] R. G. Vaughan and J. B. Andersen, "Antenna diversity in mobile communications," *IEEE Trans. Vehic. Tech*., vol. 36, no. 4, Nov. 1987, pp. 149-72.

[32] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," IEEE JSAC, vol. 16, no. 8, Oct. 1998, pp. 1451-58.

[33] R. D. Murch and K. B. Letaief, "Antenna systems for broadband wireless access," *IEEE Commun. Mag*., vol. 40, no. 4, Apr. 2002, pp. 76-83

[34] C. H. Doan, *et. al*., "Design considerations for 60GHz CMOS radios," *IEEE Commun. Mag*., vol. 42, no. 12, Dec. 2004, pp. 132-140.

[35] X. Guan, H. Hashemi, A. Komijani, and A. Hajimiri, "Multiple Phase Generation and Distribution for a Fully-Integrated 24-GHz Phased-Array Receiver in Silicon," *Proc. of IEEE RFIC Symposium*, June 2004.

[36] "Phased Array", http://www.answers.com/topic/phased-array

[37] Y. Kawaguchi, *etc*., "Application of phased-array antenna technology to the 21-GHz broadcasting satellite for rain-attenuation compensation," *IEEE Int'l Conf. on Commn.*, May 2002, pp. 2962-2966.

[38] P. V. Brenna, "Low-cost phased array antenna for land-mobile satcom applications," *IEE Proc. on Microwaves, Antennas and Propogation*, vol. 138, no. 2, pp. 131-136, Apr. 1991.

[39] B. Allen and M. Ghavami, *Adaptive Array Systems – Fundamentals and Applications*, John Wiley and Sons, 2005.

[40] W. L. Stuzman, *etc*., "Recent results from smart antenna experiments – base station and handheld terminals," in *Proc. IEEE Radio Wireless Conf.*, 2000, pp. 139-142.

[41] M. R. Williamson, G. E. Athanasiadou and A. R. Nix, "Investigating the effects of antenna directivity on wireless indoor communication at 60GHz," *8th IEEE Int'l. Symp. PIMRC*, Sept. 1997, pp. 635-39.

[42] W. D. Jones, "Keeping cars from crashing," IEEE Spectrum, vo. 38, no. 9, pp. 40-45, Sept. 2001.

[43] I. Gresham, *etc*., "Ultra-wideband radar sensors for short-range vehicular applications," *IEEE Trans. Microwave Theory Tech.*, vol. 52, no.9, pp. 2105–2122, Sept. 2004.

[44] R. Schneider and J. Wenger, "System aspects for future automotive radar," *IEEE MTT-S Int'l Microwave Symp*., June 1999, pp. 293-296.

[45] R. Hermann, "Research activities in automotive radar," *The Fourth International Kharkov Symp. on Physics and Engineering of Millimeter and Sub-Millimeter Waves,* June 2001, pp. 48-51.

[46] I. Gresham, *etc*.," Ultra wide band 24GHz automotive radar front-end," *IEEE MTT-S Int'l Microwave Symp*., June 2003, pp. 369-372.

[47] H. P. Groll, "History of automotive anticollision radars and final experimental results of a MM-Wave car radar developed by the Technical University of Munich," *IEEE Aerospace and Electronic Systems Magazine*, vol. 12, no. 8, pp.15 – 19**,** Aug. 1997

[48] B. Smolders and G. Hampson, "Deterministic RF nulling in phased arrays for the next generation of radio telescopes," *IEEE Antennas and Propagation Magazine*, vol. 44, no. 4, pp. 12-22, Aug. 2002.

[49] E. J. Bond, *etc.*, "Microwave imaging via space-time beamforming for early detection of breast cancer," *IEEE Trans. Antennas Propagat*., vol. 51, no. 8, Aug. 2003.

[50] S. C. Hagness, *etc.*, "Two-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: fixed-focus and antenna-array sensors," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 12, Dec. 1998.

[51] P. Turner, "Mini-annular phased array for limb hyperthermia," *IEEE Trans. Microwave Theory Tech.*, vol. 34, no.5, pp. 508–513, May 1996.

[52] M. O'Donnell, "Phased array beam forming from a circular array: applications to imaging of coronary arteries," *IEEE Ultrasonics Symposium*, Dec. 1991, pp. 637-640.

[53] Federal Communications Commission, FCC 02-04, Section 15.515.15.521.

[54] D. Lu et al., "Investigation of indoor radio channel from 2.4GHz to 24GHz," in *IEEE AP-S Int. Symp. Digs*, pp. 134-137, June 2003.

[55] H. Hashemi, X. Guan, and A. Hajimiri, "A Fully Integrated 24GHz 8-path Phased-Array Receiver in Silicon," ISSCC Digest of Technical papers, vol.47,pp. 390-391 Feb. 2004.

[56] D. Lu, et al, "A 24-GHz Patch Array with a Power Amplifier/Low-Noise Amplifier MMIC," International Journal of Infrared and Millimeter Waves, vol.23, pp 693-704, May 2002.

[57] X. Guan and A. Hajimiri, "A 24-GHz CMOS Front-End," Proc. 28th ESSCIRC, pp. 155-158, Sept. 2002.

[58] E. Sonmez, et al, "A Single Chip 24GHz Receiver Front-End Using a Commercially Available SiGe HBT Foundry Process," IEEE RFIC Symposium, pp. 159-162, June 2002.

[59] I. Gresham, et al, "Ultra Wide Band 24GHz Automotive Radar Front-End," IEEE MTT-S Int. Microwave Symp. Dig., pp. 369-372, June 2003.

[60] S. Wu, B. Razavi, "A 900-MHz/1.8-GHz CMOS receiver for dual band applications", *IEEE J. Solid-State Circuits*, vol.33, pp. 2178-2185, Dec. 1998.

[61] H. Samavati, H. R. Rategh, T. H. Lee, "A 5-GHz CMOS wireless LAN receiver front end", *IEEE J. Solid-State Circuits*, vol.35, pp. 765-772, May 2000.

[62] H. Darabi, A. A. Abidi, "A 4.5-mW 900-MHz CMOS receiver for wireless paging", *IEEE J. Solid-State Circuits*, vol.35, pp. 1085-1096, Aug. 2000.

[63] X. Guan and A. Hajimiri, "A 24-GHz CMOS front-end," *IEEE J. Solid-State Circuits*, vol. 39, no. 2, pp. 368-373, Feb. 2003.

[64] Rothe and W. Dahlke, "Theory of noisy fourpoles," in *Proc. IRE*, vol. 44, pp. 811-818, June 1956.

[65] H. A. Haus, "Representation of noise in linear twoports," in *Proc. IRE*, vol. 48, pp. 69-74, Jan. 1960.

[66] A. van der Ziel, *Noise in Solid State Devices and Circuits*. New York: Wiley, 1986.

[67] D. K. Shaeffer, and T. H. Lee, *The Design and Implementation of Low-Power CMOS Radio Receivers*, Kluwer Academic Publishers, Boston, 1999.

[68] A. Abidi, "High-frequnecy noise measurements on FET's with small dimensions," *IEEE Trans. Electron Device*, vol. ED-33, pp.1801-1805, Nov. 1986.

[69] A. J. Scholten, H. J. Tromp, L. F. Tiemeijer, R. V. Langevelde, R. J. Havens, and P. W. H. de Vreede, "Accurate thermal noise model for deep-submicron CMOS," in *Proc. Int. Electron Deives Meetings*, Dec. 1999, pp. 155-158.

[70] J. -S. Goo, C. -H. Choi, F. Danneville, E. Morifuji, H. S. Momose, Z.Yu, H. Iwai, T. H. Lee, and R. W. Dutton, "An accurate and efficient high frequency noise simulation technique for deep submicron MOSFETs," *IEEE Trans. Electron Devices*, vol.43, pp.1950-1955, Nov. 1996.

[71] T. Manku, "RF simulations and physics of the channel noise parameters within MOS transistors," in *Proc. IEEE Custom Integrated Circuits Conf.* May 1999, pp. 369-372.

[72] C. Enz, "An MOS transistor model for RF IC design valid in all regions of operation", *IEEE Tran. Microwave Theory and Techniques*, Vol. 50, pp. 342-359, Jan. 2002.

[73] D. K. Shaeffer and T. H. Lee, "A 1.5 V 1.5 GHz CMOS low noise amplifier," in *IEEE Very Large Scale Integration Circuits Symp. Dig. Tech.Papers*, June 1996, pp. 32–33.

[74] Y.-C. Ho, M. Biyani, J. Colvin, C. Smihhisler, and K. O, "3 V low noise amplifier implemented using a 0.8μm CMOS process with three metal layers for 900 MHz operation," *Electron. Lett.*, vol. 32, no. 13, pp. 1191–1193, June 1996.

[75] P. Leroux, J. Janssens, and M. Steyaert, "A 0.8-dB NF ESD-protected 9-mW CMOS LNA operating at 1.23 GHz," *IEEE J. Solid-State Circuits*, vol. 37, pp. 760–765, June 2002.

[76] H. Hashemi and A. Hajimiri, "Concurrent multiband low-noise amplifiers—Theory, design, and applications," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 288–301, Jan. 2002.

[77] B. A. Floyd, C.-M. Hung, and K. K. O, "A 15-GHz wireless interconnect implemented in a 0.18-_m CMOS technology using integrated transmitters, receivers and antennas," in *IEEE Very Large Scale Integration Circuits Symp. Dig. Tech. Papers*, June 2001, pp. 155–158.

[78]T. Manku, "Microwave CMOS-device physics and design," *IEEE J. Solid-State Circuits*, vol. 34, pp. 277–285, March 1999.

[79] J. Chen, B. Shi, "Impact of intrinsic channel resistance on noise performance of CMOS LNA", *IEEE Electron Device Letters*, vol.23, pp. 34-36, Jan. 2002.

[80] S. F. Tin, A. A. Osman, K. Mayaram and C. Hu, "A simple subcircuit extension of BSIM3v3 Model for CMOS RF Design", *IEEE J. Solid-State Circuits*, vol.35, pp. 612-624, April 2000.

[81] A. Rofougaran, J. Y.-C. Chang, M. Rofougaran, and A. A. Abidi, "A 1-GHz CMOS RF front-end IC for a direct-conversion wireless receiver," *IEEE J. Solid-State Circuits*, vol. 31, pp. 880–889, July 1996.

[82] H. A. Haus and R. B. Adler, *Circuit Theory of Linear Noisy Networks*, John Wiley & Sons, 1959, New York.

[83] B. A. Floyd, L. Shi, Y. Taur, I. Lagnado, and K. K. O, "A 23.8-GHz SOI CMOS tuned amplifier," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 2193–2195, Sept. 2002.

[84] K. W. Yu, M. F. Chang, “CMOS K-band LNAs design counting both interconnect transmission line and RF pad parasitics,” *Proc. RFIC*, June 2004, pp. 101-105.

[85] H. Hashemi, X. Guan, and A. Hajimiri, “A Fully-Integrated 24-GHz 8-Channel Phase-Array Receiver in Silicon,” *Proc. of IEEE International Solid-State Circuits Conference*, pp. 390-391, Feb. 2004.

[86] X. Guan, H. Hashemi, and A. Hajimiri, "A Fully Integrated 24-GHz Eight-Element Phased-Array Receiver in Silicon," *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 2311-20, Dec. 2004.

[87] J. P. Roux, *et al*., “Small-signal and noise model extraction technique for heterojunction bipolar transistor at microwave frequencies,” *IEEE Trans. Microwave Theory Tech.*, vol. 43, no.2, pp. 293–298, Feb. 1995.

[88] S. P. Voinigescu, “A scalable high-frequency noise model for bipolar transistor with application to optimal transistor sizing for low-noise amplifier design,” *IEEE J. Solid-State Circuits*, vol. 32, no. 9, pp. 1430-1439, Sep. 1997.

[89] O. Shana’a, “Frequency-scalable SiGe bipolar RF front-end design,” *IEEE J. Solid-State Circuits*, vol. 36, no. 6, pp. 888-895, June 2001.

[90] D. M. Pozar, *Microwave Engineering*, John Wiley & Sons, 1998.

[91] B. Razavi, *RF Microelectronics*, Prentice Hall.

[92] C. S. Vaucher, *Architectures for RF Frequency Synthesizer*, Kluwer Academic Publishers, Boston, 2002.

[93] F. M. Gardner, “Charge-pump phase-lock loops,” *IEEE Tran. Commun*., vol. 28, no. 11, pp. 1849-1858, Nov. 1980.

[94] J. Craninckx, M. Steyaert, *Wireless CMOS Frequency Synthesizer Design*, Springer, 1998.

[95] J. Savoj and B Razavi, “A 10-Gb/s CMOS clock and data Recovery circuit with a half-rate binary phase-frequency detector,” *IEEE J. Solid-State Circuits*, vol. 38, no. 1, pp. 13-21, Jan. 2003.

[96] X. Guan, H. Hashemi, A. Komijani, and A. Hajimiri, “Multiple Phase Generation and Distribution for a Fully-Integrated 24-GHz Phased-Array Receiver in Silicon,” *Proc. of IEEE RFIC Symposium*, pp. 229-232, June 2004.

[97] H. Hashemi, X. Guan, A. Komijani, A. Hajimiri, "A 24-GHz SiGe Phased-Array Receiver-LO Phase-Shifting Approach," *IEEE Trans. Microwave Theory Tech.*, vol. 53, no. 2, pp. 614-626, Feb. 2005.

[98] C. K. Cheng, *et al*., *Interconnect Analysis and Synthesis*, John Wiley & Sons, New York, 2000.

[99] A. Joseph et al., “A 0.18µm BiCMOS Technology Featuring 120/100GHz (ft/fmax) HBT and ASIC-Compatible CMOS Using Copper Interconnect,” *BCTM Proceedings*, pp. 143-146, 2001.

[100] Hossein Hashemi, *Integrated Concurrent Multi-band Radios and Multiple-Antenna System*, Ph.D. Dissertation, California Institute of Technology, 2003.

[101] Federal Communications Commission, FCC 02-04, Section 15.515.15.521.

[102] J. Schoebel, *et al*., “Design considerations and technology assessment of phased-array antenna systems with RF MEMS for automotive radar application,” *IEEE Trans. Microwave Theory Tech.*, vol. 35, no.6, pp. 1968–1974, June 2005.

[103] “ECC decision of 19 March 2004 on the frequency band 77–81 GHz to be designated for the use of automotive short range radars,” Eur. Radiocommun. Office, Copenhagen, Denmark, (ECC/DEC/(04)03), 2004. Available: http://www.ero.dk.

[104] http://www.wmo.ch/web/www/TEM/SG-RFC04/EU-24G-WMO.doc

[105] “Compatibility of automotive collision warning short range radar operating at 79 GHz with radiocommunication services,” Eur. Radiocommun. Office, Copenhagen, Denmark, Oct., 2004. Available: http://www.ero.dk.

[106] A. Mazzanti, *et al*., "Injection locking LC dividers for low power quadrature generation," *Proc. Custom Integrated Circuits Conference*, Sept. 2003, pp. 563 – 566.

[107] H. R. Rategh and T. H. Lee, *Multi-GHz Frequency Synthesis & Division – Frequency Synthesizer Design for 5 GHz Wireless LAN systems*, Kluwer Academic Publishers, 2001.

[108] G. E. Moore, "Cramming more components onto integrated circuits," *Electron.*, vol. 38, no. 8, pp. 114–117, Apr. 1965.