Copyright © 2001, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

A DESIGN METHODOLOGY FOR HIGHLY-INTEGRATED LOW-POWER RECEIVERS FOR WIRELESS COMMUNICATIONS

by

Dennis Gee-Wai Yee

Memorandum No. UCB/ERL M01/17

1 May 2001

A DESIGN METHODOLOGY FOR HIGHLY-INTEGRATED LOW-POWER RECEIVERS FOR WIRELESS COMMUNICATIONS

by

Dennis Gee-Wai Yee

Memorandum No. UCB/ERL M01/17

1 May 2001

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

A Design Methodology for Highly-Integrated Low-Power Receivers for Wireless Communications

by

Dennis Gee-Wai Yee

B.S. (University of California, Berkeley) 1994
M.S. (University of California, Berkeley) 1996

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering-Electrical Engineering and Computer Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Robert W. Brodersen, Chair
Professor Paul R. Gray
Professor Paul K. Wright

Spring 2001

The dissertation of Dennis Gee-Wai Yee is approved:

Robert W. Brodersen 2/9/01

Chair Date

Paul R. Gray 2/9/01

Date

Paul K. Wright 3/23/2001

Date

University of California, Berkeley

Spring 2001

A Design Methodology for Highly-Integrated Low-Power Receivers for Wireless Communications

Copyright 2001

by

Dennis Gee-Wai Yee

Abstract

A Design Methodology for Highly-Integrated Low-Power Receivers for Wireless Communications

by

Dennis Gee-Wai Yee

Doctor of Philosophy in Engineering-Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Robert W. Brodersen, Chair

Due to its potential to offer ubiquitous information access, wireless connectivity is playing an increasingly significant role in communications systems. The success of future wireless systems will depend heavily on their ability to provide high capacity while maintaining low cost, small form factor, and low power consumption in the portable devices. However, many existing commercial transceivers are expensive, consist of a large number of discrete components, and exhibit moderate to high levels of power consumption. One possible explanation for these inefficient solutions lies in the historically unilateral relationship between system designers and hardware designers. An efficient solution requires a design strategy which tightly incorporates implementation issues throughout the process of defining the system specifications.

This thesis describes a design methodology which facilitates the evaluation of tradeoffs between implementation issues and overall system performance, focusing primarily on the receiver as an example. First, system-level specifications, such as modulation scheme and signal bandwidth, strongly influence the choice of receiver architecture, which in turn, has ramifications on the achievable power consumption and integration level. When system-level specifications are determined without considering their impact on receiver architecture selection, single-chip solutions may be very difficult to achieve or just simply infeasible. Based on system-level considerations, guidelines are presented for the

selection of receiver architectures, including the heterodyne, direct-conversion, image-reject, and low-IF topologies.

Second, the rapid improvements in digital CMOS technology provide an opportunity to use advanced digital signal processing algorithms which in the past were considered too complex to implement in the mobile device. These algorithms promise significant increases in system performance but their performance may ultimately be limited by analog circuit impairments, such as noise and distortion. This thesis describes the detrimental effects of a number of these impairments and presents a system-level simulation framework which facilitates the direct evaluation of these effects on the performance of digital communications algorithms. The simulation framework is implemented in Simulink, which offers compatibility with MATLAB, a simulation tool already widely used for the development and evaluation of communications algorithms. This simulation framework relies on baseband-equivalent models for all of the RF building blocks in order to avoid simulation at the carrier frequency, resulting in faster simulation times.

These strategies are then applied to the design of a high-speed wireless downlink for an indoor picocellular system. The system provides an aggregate data rate of 50 Mb/s with a transmission bandwidth of 32.5 MHz and a carrier frequency of 2 GHz. The wide bandwidth of the desired signal facilitates the use of a direct-conversion architecture. A receiver prototype is implemented to meet the specifications determined from the system-level simulations. A power-efficient solution is achieved by taking advantage of the relaxed specifications as well as by using low-power circuit implementation techniques. This receiver prototype includes the low-noise amplifier, frequency synthesizer, mixers, baseband amplifiers and filters, and analog-to-digital converters, all implemented on a single chip with a power dissipation of about 100 mW.

Robert W. Brodersen, Chair

To Mom and Dad

Table of Contents

List of Figures

List of Tables

Acknowledgments

The work described in this thesis could not have been accomplished without the help and support of others. Foremost, I would like to thank my research advisor Professor Robert Brodersen for his guidance during the past seven years. I also want to thank you, along with Professor Jan Rabaey, for providing such a wonderful research environment at the BWRC for all of us. I would like to thank Professor Paul Gray and Professor Paul Wright for reviewing this thesis and for providing helpful comments. I would also like to thank both of you along with Professor Robert Meyer for providing advice and guidance during various stages of this project. Finally, I am very grateful to all of you along with all of the other faculty members at UC Berkeley for providing me with such a high-quality education at both the undergraduate and graduate levels.

I am greatly indebted to the members of the Baseband Analog and RF (BARF) group: Sayf Alalusi, Sam Blackman, Chinh Doan, Brian Limketkai, Ian O'Donnell, David Sobel, and Johan Vanderhaegen. Chinh, Dave, and Brian made significant contributions to the receiver hardware design, while Chinh and Dave, along with Johan, also made significant contributions to the implementation of the Simulink simulation framework. Thanks Chinh for doing a great job on the PLL and for helping out on all different aspects of the project. Thanks Dave for doing a great job on the ADC and the SCRBARF test board and for expanding my vocabulary with words like absquatulate, callipygian, and nugatory. Thanks Brian for doing a great job on the mixer and for giving me a ride when I really needed one. Thanks Johan for doing a great job on modeling the base-station transmitter in Simulink and for all the delicious Belgian chocolates. Thanks Sayf for designing the SCRRX test board. Thanks Ian for being a great project partner in EE225C, EE241, and EE290Y. And thanks Sam for all the great stock tips. You BARF dudes are the best!

Craig Teuscher has helped me to better understand all aspects of digital communications. There are a lot of excellent communications engineers out there, but you are one of the few who are experts in both communications engineering as well as the art of communicating. I am indeed lucky to have worked with someone who has such an innate ability to clearly explain even the most difficult concepts.

I would like to thank Tom Burd, Chris Rudell, Jeff Ou, Sekhar Narayanaswami, Jeff Weldon, Andy Abo, Keith Onodera, Jeff Gilbert, and Ali Niknejad for all their help and advice over the many years of graduate school. Thank you all for providing me with invaluable feedback for papers and practice talks and for just being a bunch of cool dudes. Thanks Tom and Jeff G. for always providing a non-RF perspective which improved the papers and talks tremendously. Also, thanks Jeff G. for making our whirlwind trip through Italy such a memorable experience. Thanks Chris and Keith for helping me with EE140 when I was still an undergrad. I would also like to thank Chris and Jeff O. for all the edifying discussions on various aspects of RF design. Thanks Ali for ASITIC and for all the helpful discussions about inductors. Thanks Sek for being a great project partner in EE231 and for all the lighthearted discussions about Cal basketball and football. Go Bears!

I am also grateful to the original members of the RF group: Lapoe Lynn, Ian O'Donnell, Jim Peroulas, Sam Sheng, and Kevin Stone. In particular, I am greatly indebted to Sam Sheng both for being the best TA I have ever had and for mentoring me during my first couple of years in graduate school.

I have also had the pleasure of working with many other past and present graduate students including but not limited to Ada Poon, Andy Klein, Henry Jen, Chris Taylor, George Chien, Marlene Wan, Varghese George, Ning Zhang, Arthur Abnous, Li Lin, Martin Tsai, Rhett Davis, Josie Ammer, Heather Bowers, Roy Sutton, Trevor Pering, David Lidsky, Srenik Mehta, Shankar Narayanaswamy, and Arnie Feldman. Thanks Ada, Andy, and Ning for all the helpful discussions on various topics in communications. Also, thanks Ada for always being so nice and cheery, and always remember to respect your grandpa!

The students at the BWRC are very fortunate to have the opportunity to interact with many industrial and academic visitors. I have had the pleasure of working with Ken Kundert, Bob Frye, Andrei Vladimirescu, Marco Sabitini, Isabelle Telliez, and Bhusan Gupta. Thanks Ken for all the help with SpectreRF. And thanks Marco, Isabelle, and Bhusan for all the help with getting our chips taped out.

The students at the BWRC are also very fortunate to have support from such friendly and efficient administrative and technical staff members, including Tom Boot, Elise Mills, Kahren Kim, Gary Kelson, Kevin Zimmerman, Brian Richards, Fred Burghardt, and Sue Mellers. I want to thank all of you as well as Peggye Brown in Cory Hall for help in purchasing parts for test boards, arranging itineraries for travel to conferences and workshops, and for just keeping the BWRC running so smoothly.

I am grateful to Professor Rinaldo Castello at the University of Pavia and Professor Andrea Baschirotto at the University of Lecce along with all of your students for being such gracious hosts during my trip to Italy. Danilo Gerna, who was my partner for the EECS foosball tournament when he was a visiting scholar at Berkeley, was also a very gracious host during my stay in Pavia.

I would also like to thank everyone who has helped to make the stress of graduate school much more tolerable. Working out at the RSF has helped to relieve the stress associated with the trials and tribulations of graduate school. My past and present workout buddies include: Jeff "Hanz" Gilbert, Sayf Alalusi, Tom Burd, Ian O'Donnell, Chinh Doan, Chris Taylor, Arnie Feldman, and Dave Lidsky.

I have also enjoyed mountain biking and more recently, road biking with Brian Limketkai, Sayf Alalusi, Ian O'Donnell, Tom Burd, Chris Rudell, Jeff Weldon, and Danilo Gerna. Together we have enjoyed many of the awesome bike routes in the Berkeley area. I also want to thank Gabe Moy for answering all of my questions about bikes.

Another pastime which I picked up during graduate school is golf. Those who have witnessed by dreadful swing include Keith Onodera, Andy Abo, Tom Burd, George

Chien, Chinh Doan, Henry Jen, Andy Klein, and Jeff Weldon. It can be difficult for a bunch of poor graduate students to support an expensive pastime like golf; so I am very grateful for Tilden's College Night, when a round of golf is only $10 for Cal students. Also, thanks Jeff for helping me with my swing and for being so patient.

Finally, I would like to thank my family and friends for their support and encouragement. I have many fond memories of my high school and college years with friends like Burton Lee, Milton Yee, Calvin Low, Ivan Choi, Norman Chan, Wayne Yeung, Nick Mitchell, Garvin Leon, Albert Yee, John Balk, Moses Kim, Ben Shieh, Paul Chou, and Jim Young.

I am also very fortunate to have great friends like Albert Lee, Richard Ng, and Bryant Woo. Thanks for always being there for me, through the good times and the bad.

I would also like to thank my Aunt and Uncle and all of my cousins along with their wonderful families: Randy and Kelly; Phillis, Randy, Kelsey, Mitchell, and Kristen; Andy, Ginger, Ryan, and Nicole; and Debbie, Richard, and Adam. Thanks Andy for all the help and advice during my first couple of years at Cal.

Yvette, thank you for your encouragement, your patience, and your understanding. And above all, thank you for making the past eleven years so rich and so fulfilling.

Davis, I am truly fortunate to have a big brother who always watches over me and so selflessly shares everything with me. Thank you for giving me the benefit of all of your experiences and for teaching me the meaning of friendship, loyalty, courage, and perseverance.

Mom and Dad, both of you have sacrificed so much for Davis and me. Your unwavering support has been a continual source of strength for both of us. Thank you for instilling in us the importance of a good education and for encouraging the both of us to pursue our dreams. The accomplishments which I have achieved belong as much to the both of you as they do to me. I dedicate this work to the both of you with all my love.

Chapter 1

Introduction

1.1 Motivation

Due to its potential to offer ubiquitous information access, wireless connectivity is continuing to play an increasingly significant role in communications systems. The proliferation of wireless technologies is already evident in the success of modern paging and cellular telephony applications. Although wireless connectivity is inherent to the functionality of these devices, wireless connectivity is still absent from many portable devices such as laptops and personal digital assistants (PDAs). Future communications systems will offer new wireless services for devices such as laptops and PDAs as well as expand on the existing wireless capabilities of devices such as cellular telephones and pagers. These applications include Internet access, video teleconferencing, high-fidelity audio, and other high-speed services.

Wireless connectivity is not limited to only portable devices, but can also be used for applications which currently rely on tethered connections, including local area networks (LANs) and local loop applications such as Integrated Services Digital Network (ISDN) and Digital Subscriber Line (DSL), both of which rely on copper twisted pair, as well as

cable applications, which rely on a combination of fiber optic and coaxial cables. Bluetooth is one example of a wireless standard which is targeted at applications which currently rely on wires [1]. Bluetooth aims to use wireless connectivity to replace cables such as the those connecting peripheral devices to a computer. Emerging wireless solutions for LAN applications include the IEEE 802.11a and 802.11b standards at 5 GHz and 2.4 GHz, respectively, in the United States [2] as well as the ETSI HIPERLAN standards in Europe [3]. Several wireless solutions have also been proposed for local loop applications including the Local Multipoint Distribution Service (LMDS) operating at 28 GHz and the Multichannel Multipoint Distribution Service (MMDS) operating at 2.5 GHz [4].

For systems designed to provide wireless connectivity to mobile devices, the success of these systems will depend heavily on their ability to provide high capacity while maintaining low cost, small form factor, and low power consumption in the portable units. However, many existing commercial transceivers are expensive, consist of a large number of discrete components, and exhibit moderate to high levels of power consumption. One possible explanation for these inefficient solutions lies in the historically unilateral relationship between system designers and hardware designers: first system designers develop standards concentrating mainly on communications issues, and then hardware designers must implement a solution to meet these standards. An efficient solution requires cooperation between both system and hardware designers as well as a design strategy which tightly incorporates implementation issues throughout the process of defining the system specifications.

1.2 Research Goals

The goal of this research is to establish a design framework to evaluate tradeoffs between implementation issues and system performance. This framework will focus on one of the key units in an indoor wireless system: the receiver in the mobile device. In order to achieve a single-chip solution with low power consumption, three design strategies are proposed. First, at the system level, implementation issues must be considered even during the earliest stages of system definition. Clearly this approach is not possible for

systems which have already been defined. However, selecting system features which allow for relaxed hardware requirements is critical for achieving single-chip, low-power receiver implementations in future wireless systems. In addition, the allocation of unlicensed spectra in the Industrial, Scientific, and Medical (ISM) bands at 900 MHz and 2.4 GHz and the Unlicensed National Information Infrastructure (U-NII) band at 5 GHz provides opportunities for development of custom wireless systems.

Second, efficient implementations require careful evaluation of the effects of analog receiver impairments on the performance of digital communications algorithms. The rapid improvements in digital CMOS technology facilitate the integration of increasingly more functionality onto a single chip. In particular, advanced signal processing algorithms are very amenable to low-power digital design techniques and promise increased capacity along with higher data rates [5]. However, the performance of these algorithms may ultimately be limited by analog circuit impairments, such as noise, distortion, and mismatch. By accounting for analog impairments during the earliest stages of algorithm exploration, it may be possible to relax some of the analog hardware requirements without necessarily sacrificing overall system performance.

Third, low-power circuit implementation techniques are required to minimize the power consumption of the analog circuits. Despite efforts to simplify the analog hardware, the analog section of the receiver can still dominate the overall receiver power consumption.

1.3 Thesis Organization

Chapter 2 provides an overview of various receiver architectures. The choice of receiver architecture affects both the power consumption and the level of integration. The chapter describes the key features of the heterodyne, direct-conversion, image-reject, and low-IF architectures and presents some design guidelines for receiver architecture selection.

Chapter 3 describes the effects of receiver front-end impairments, such as noise, distortion, and mismatch, which can potentially degrade the performance of digital communications systems. A QPSK signal constellation is used to demonstrate the detrimental effects of many of these impairments.

Chapter 4 first provides an overview of the conventional approach of using link budget calculations to determine the allowable levels of receiver impairments and then describes a system-level simulation framework that includes models for the analog impairments described in Chapter 3. This simulation framework is implemented in Simulink and facilitates the direct evaluation of the effects of analog impairments on the performance of digital communications algorithms.

Chapter 5 describes the design of a high-speed wireless downlink for an indoor picocellular system. System specifications are chosen in order to facilitate the use of a direct-conversion architecture as well as to relax many of the performance requirements of the analog hardware without significantly degrading the overall system performance.

Chapter 6 describes the implementation of the receiver prototype including the low-noise amplifier, frequency synthesizer, mixers, baseband amplifiers and filters, and analog-to-digital converters, focusing primarily on design choices which result in the most power-efficient implementation. All of these components are integrated onto a single-chip, and a power-efficient solution is achieved by taking advantage of these relaxed requirements along with low-power circuit implementation techniques.

Chapter 7 presents the simulated and measured performance results of the receiver prototype, and Chapter 8 concludes with a summary as well as suggestions for future research.

Chapter 2

Receiver Architectures

2.1 Introduction

One of the key components of portable devices used in wireless communications systems is the receiver, which senses an incoming signal and extracts the desired information. Since the Federal Communications Commission (FCC) regulates the frequencies at which signals can be transmitted, the incoming signal is typically centered at a frequency which is much larger than the bandwidth of the desired signal. For the ideal case illustrated in Fig. 2.1, the radio-frequency (RF) front-end of the receiver simply translates the incoming signal from a carrier frequency, f_c, down to baseband.

Unfortunately, in a real wireless transmission environment, the received signal is almost

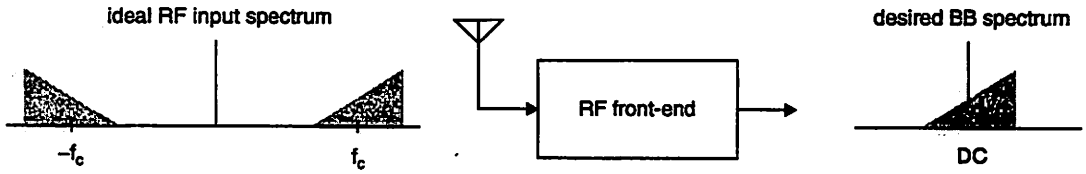


Figure 2.1: RF front-end for an ideal RF input signal.

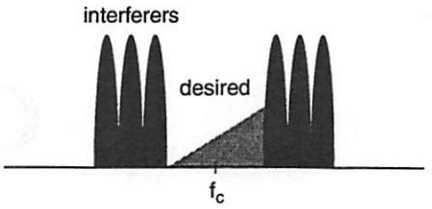


Figure 2.2: RF input signal with a weak desired signal and strong adjacent interferers.

always far from ideal. The signal which reaches the receiver can be very weak because of attenuation by objects which obstruct the transmission path between the transmitter and receiver or simply because of the loss due to spatial separation between the transmitter and receiver. In addition, the received signal can include unwanted signals along with the desired one. These unwanted signals, or interferers, can be significantly stronger than the desired signal as illustrated in Fig. 2.2.

Due to the limited amount of attenuation achievable by practical filter designs as well as the noise and distortion introduced by circuits used to implement the RF front-end, the design of a highly-integrated, low-power receiver becomes increasingly challenging when the received signal consists of a very weak desired signal in the presence of strong adjacent interferers. Two metrics which are used to evaluate receiver performance are sensitivity and selectivity. A receiver with high sensitivity can correctly process a very weak desired signal whereas a receiver with high selectivity can correctly process a desired signal in the presence of very strong interferers at adjacent frequencies. The required sensitivity and selectivity of a receiver are highly dependent on the specifications of the underlying communications system. In order to meet the sensitivity and selectivity requirements of a particular system while facilitating a highly-integrated, low-power implementation, the architecture used for the receiver must be carefully considered.

This chapter provides an overview of various receiver architectures, starting with the heterodyne architecture, which, unfortunately, is not very amenable to high levels of integration, and followed by an overview of a few other receiver architectures which are more conducive to single-chip implementations.

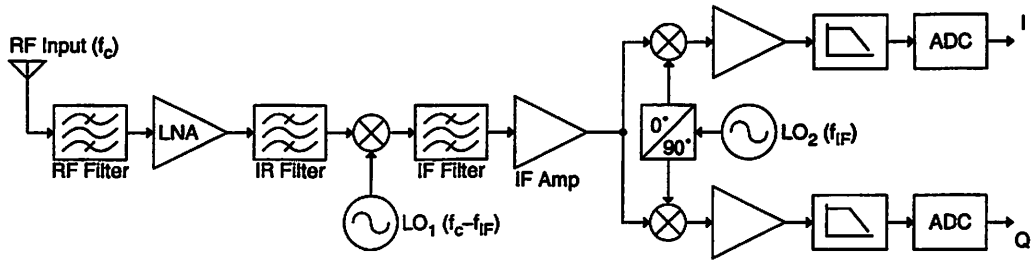


Figure 2.3: Heterodyne architecture block diagram.

2.2 Heterodyne Architecture

The heterodyne architecture (also called the superheterodyne architecture) is probably the most commonly used architecture in current commercial receiver implementations. In this architecture the received signal is converted to baseband in multiple frequency translation steps. A block diagram of the heterodyne architecture with two frequency translation steps is illustrated in Fig. 2.3. In this architecture, the signal received at the antenna first passes through an RF filter before being amplified by a low-noise amplifier (LNA). The signal is then filtered by an image-reject (IR) filter before being frequency translated to an intermediate frequency (IF) by the first local oscillator (LO). At the intermediate frequency the signal is further filtered and amplified before being frequency translated to baseband along parallel in-phase (I) and quadrature (Q) signal paths by the second LO. At baseband, the signal is further amplified and filtered before being converted to a digital signal by the analog-to-digital converter (ADC).

2.2.1 The Image Problem

The selection of the intermediate frequency in this architecture is directly related to the image problem. In Fig. 2.4, the desired signal centered at the carrier frequency f_c is frequency translated to the intermediate frequency f_{IF} by an LO located at the frequency $f_c - f_{IF}$. However, the signal centered at the image frequency $f_c - 2f_{IF}$ is also frequency translated to f_{IF}. Since the image signal can be much stronger than the desired signal, the image signal must be sufficiently attenuated before frequency translation.

The choice of f_{IF} depends on the characteristics of practical filter implementations. For a typical ceramic filter [6] (Fig. 2.5), the amount of attenuation increases at frequencies

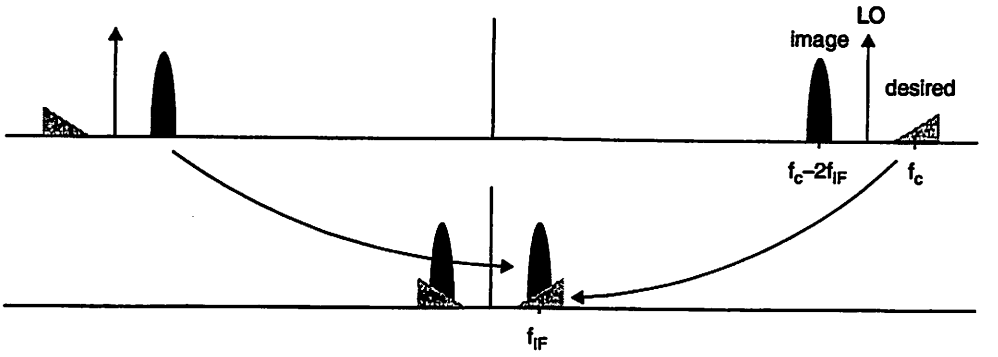


Figure 2.4: The image problem.

farther away from the center frequency f_0. Consequently, in order to achieve a large amount of image signal attenuation, it is preferable to select a high intermediate frequency so that the image signal is far away from the center frequency of the filter. However, a high intermediate frequency also increases the design challenges in the IF filtering and amplification circuits. Consequently, the choice of f_{IF} must be based on the following tradeoffs:

- a high intermediate frequency results in maximum image signal attenuation from the IR filter, while
- a low intermediate frequency results in relaxed IF filtering and amplification requirements.

2.2.2 Implementation

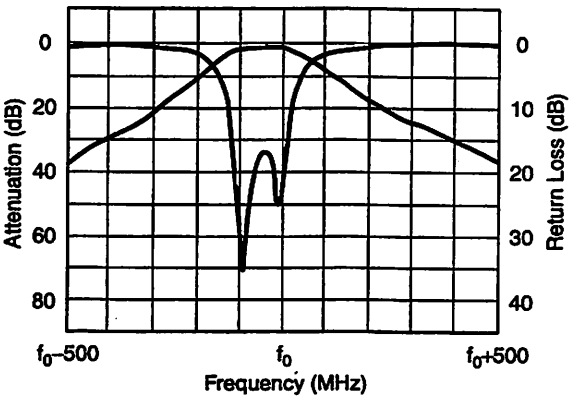


Figure 2.5: Typical ceramic filter characteristic.

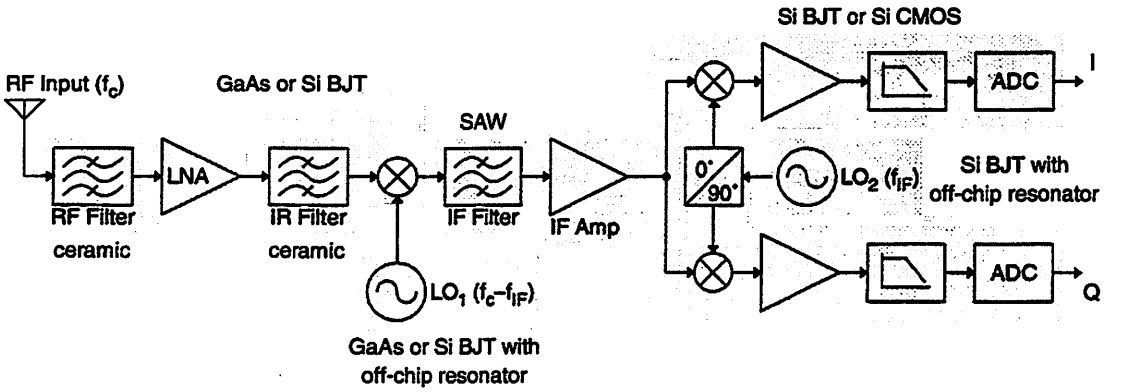


Figure 2.6: Heterodyne architecture implementation.

The heterodyne architecture is commonly used in current commercial receiver implementations because of its excellent sensitivity and selectivity performance. This excellent performance is achieved by using the best technologies to implement the various components. For example, the RF and IF filters are typically implemented using ceramic filter technology while the IR filter is typically implemented using surface acoustic wave (SAW) technology. The remaining components are implemented using an assortment of gallium arsenide, silicon bipolar, and silicon CMOS technologies. As illustrated in Fig. 2.6, a typical implementation consists of a large number of components, implemented in multiple technologies.

Since small form factor and low power consumption are two critical design goals in the design of portable units, the heterodyne architecture is inadequate and other receiver architectures which are more amenable to highly-integrated, low-power implementations must be considered for future wireless communications systems. These architectures include:

1. the direct-conversion architecture;
2. the image-reject architecture; and
3. the low-IF architecture.

2.3 Direct-Conversion Architecture

Rather than frequency translating the received signal to an intermediate frequency, the direct-conversion receiver architecture downconverts the received signal directly to

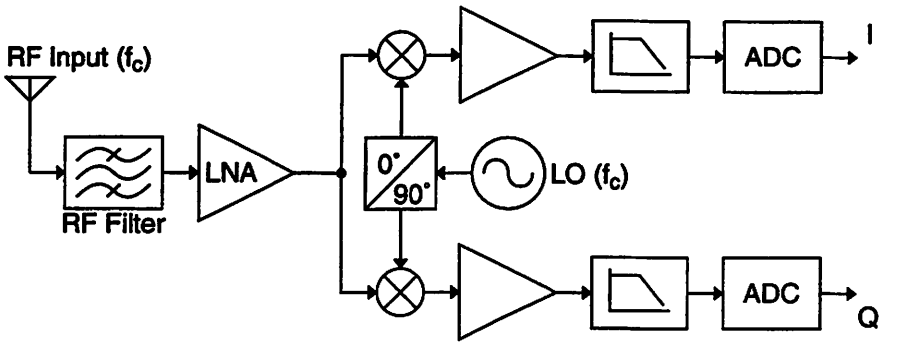


Figure 2.7: Direct-conversion architecture block diagram.

baseband, and consequently, image rejection is no longer necessary. A block diagram of the direct-conversion architecture is illustrated in Fig. 2.7. The RF signal appearing at the antenna is filtered and amplified before being downconverted to baseband along parallel I and Q signal paths. The frequency translation is performed using two mixers and an LO fixed at the carrier frequency and operating in quadrature. The I and Q baseband signals are then amplified and low-pass filtered prior to analog-to-digital conversion. Because the RF signal is converted directly to baseband, this architecture eliminates all intermediate-frequency components and their associated design challenges, including the image-reject problem. Moreover, all of the remaining analog components can be integrated onto a single chip using a single technology such as silicon CMOS, with the exception of the antenna and the RF filter [7]–[9]. However, two practical considerations have limited the use of the direct-conversion architecture: dc offsets and flicker noise.

2.3.1 DC Offsets

Implementations based on the direct-conversion architecture are particularly sensitive to dc offsets since the desired signal is downconverted directly to baseband. DC offsets are problematic for two reasons. First, dc offsets can saturate the baseband circuits, such as amplifiers and filters. Second, even if the baseband circuits do not saturate, dc offsets, if uncorrected, degrade the bit-error rate (BER) performance of the system.

There are three primary sources of dc offsets: LO self-mixing, even-order distortion, and baseband circuit mismatch. As illustrated in Fig. 2.8, the LO signal can couple to the RF signal path, for example, through the input of the LNA or through the RF port of the

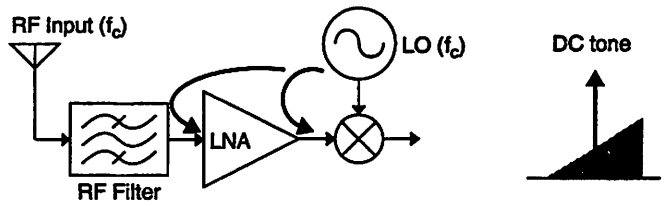


Figure 2.8: LO self-mixing.

mixer. The LO signal then mixes with itself, creating a dc offset. LO self-mixing can be represented as the product of two sinusoids at the same frequency f_c,

$$S_1 \cos(2\pi f_c t) \times S_2 \cos(2\pi f_c t) = \frac{S_1 S_2}{2}\{1 + \cos[2\pi(2f_c)t]\} \tag{2.1}$$

which results in a dc component as well as a sinusoidal component at $2f_c$.

A second source of dc offsets is even-order harmonic distortion. The transfer function of the analog front-end can be represented by the following equation:

$$s_o = a_1 s_i + a_2 s_i^2 + a_3 s_i^3 + \ldots \tag{2.2}$$

where the first term represents the gain and the remaining terms represent the nonlinear behavior. In the case of second-order harmonic distortion, for an input sinusoid $s_i = S_1 \cos(2\pi f_1 t)$, the resulting output signal is

$$s_o = \frac{a_2 S_1^2}{2}\{1 + \cos[2\pi(2f_1)t]\} \tag{2.3}$$

which again consists of a dc component and a sinusoidal component at $2f_1$. Although this example considers only second-order distortion, in fact, all even-order distortion terms result in a dc offset component. In addition, as seen in (2.3), the amount of dc offset resulting from even-order distortion depends on the input signal amplitude S_1, which varies over time. Consequently, the cancellation of dc offsets can be challenging because of its time-varying nature.

In addition to dc offsets, even-order distortion can also result in other low-frequency components which can be detrimental to the performance of direct-conversion receivers.

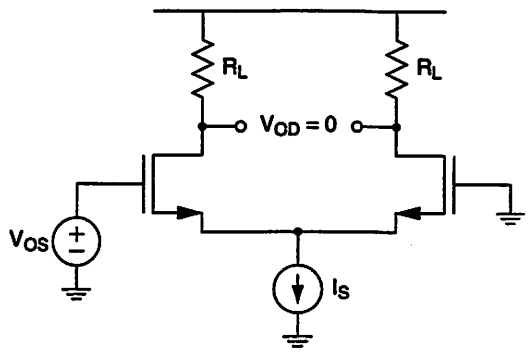


Figure 2.9: Input offset voltage for a differential CMOS amplifier.

In the case of second-order intermodulation distortion, for an input signal $s_i = S_1 \cos(2\pi f_1 t) + S_2 \cos(2\pi f_2 t)$, the resulting output signal is

$$s_o = a_2 S_1 S_2 \{\cos[2\pi(f_1 - f_2)t] + \cos[2\pi(f_1 + f_2)t]\} \tag{2.4}$$

which consists of sinusoidal components at frequencies $f_1 - f_2$ and $f_1 + f_2$. If the frequency separation between the two input sinusoids is small, then second-order intermodulation distortion results in a low-frequency component, $f_1 - f_2$, which can potentially corrupt the desired baseband signal. Since even-order intermodulation distortion depends on the input signal amplitudes, S_1 and S_2, the amplitude of the resulting low-frequency components are also time-varying.

Finally, a third source of dc offsets is baseband circuit mismatch. If differential circuits are used to implement the baseband amplifiers and filters, device mismatches give rise to dc offsets. For the differential CMOS amplifier illustrated in Fig. 2.9, the input offset voltage is [10]

$$V_{OS} = \Delta V_t + \frac{(V_{GS} - V_t)}{2}\left[\left(\frac{-\Delta R_L}{R_L}\right) - \left(\frac{\Delta(W/L)}{(W/L)}\right)\right]. \tag{2.5}$$

The dc offset in this case is related to the mismatch in the transistor threshold voltage, the mismatch in the transistor geometry, and the mismatch in the load resistors. Moreover, dc offsets arising from circuit mismatches are also dependent on temperature variations.

2.3.2 Flicker Noise

Flicker noise can also degrade the performance of direct-conversion receivers. The power spectral density of the input-referred voltage noise for a MOS transistor consists of a flicker noise component which is inversely proportional to frequency and a thermal noise component:

$$\frac{\overline{v_i^2}}{\Delta f} = \frac{K_f}{WLC_{ox}f} + 4kT\frac{2}{3}\frac{1}{g_m}. \tag{2.6}$$

The noise power spectral density is plotted as a function of frequency in Fig. 2.10. For high frequencies, the thermal noise component dominates, but for low frequencies, the flicker noise component is stronger. In fact, since the power spectral density of flicker noise is inversely proportional to frequency, the flicker noise component can be quite large near dc. Since the desired signal is frequency translated directly to baseband, flicker noise can be particularly problematic for direct-conversion architectures.

2.4 Image-Reject Architecture

In the heterodyne architecture, the image problem arises from the use of a real sinusoidal LO signal to frequency translate the input signal to an intermediate frequency. More specifically, the Fourier transform of a real sinusoidal LO signal, $\cos(2\pi f_{LO1}t)$, consists of a negative frequency component at $-f_{LO1}$ and a positive frequency component at $+f_{LO1}$:

$$\cos(2\pi f_{LO1}t) \longleftrightarrow \frac{1}{2}[\delta(f - f_{LO1}) + \delta(f + f_{LO1})]. \tag{2.7}$$

During the frequency translation process, the negative frequency component of the LO signal downconverts the positive frequency component of the desired signal to f_{IF}, while the positive frequency component of the LO signal downconverts the negative frequency

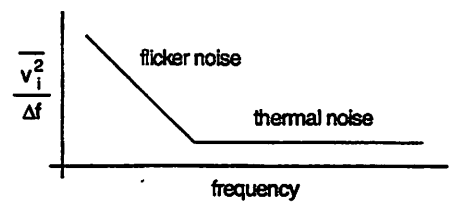


Figure 2.10: Noise power spectral density for MOS transistors.

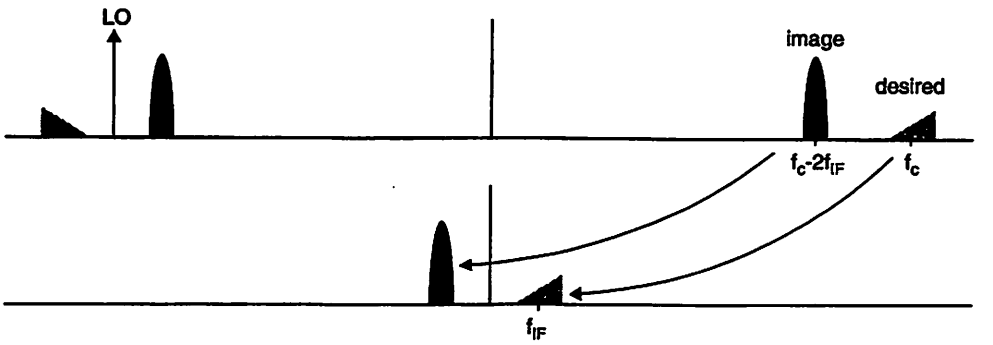


Figure 2.11: Frequency translation to f_{IF} using a complex sinusoidal LO signal.

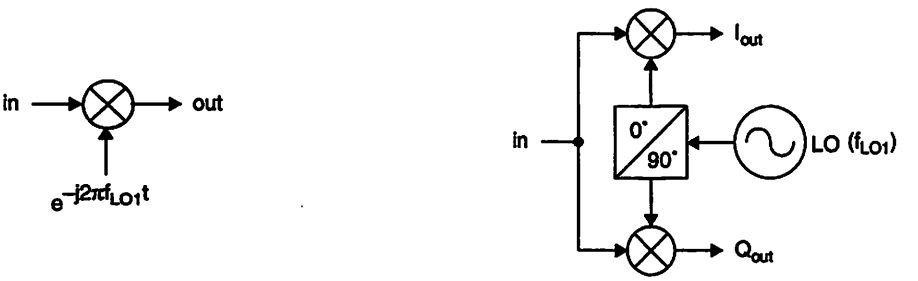


Figure 2.12: Complex mixing. (a) Concept. (b) Implementation.

component of the image signal also to f_{IF} (Fig. 2.4).

The preceding description of the mechanism behind the image problem suggests a potential solution: use a complex sinusoidal LO signal to downconvert the input signal to the intermediate frequency. The Fourier transform of a complex sinusoidal LO signal, $e^{-j2\pi f_{LO1}t}$, consists of only a negative frequency component at $-f_{LO1}$:

$$e^{-j2\pi f_{LO1}t} \longleftrightarrow \delta(f + f_{LO1}). \tag{2.8}$$

When the input signal is multiplied by the LO signal, only the positive frequency components of the input signal are translated to the intermediate frequency as illustrated in Fig. 2.11. Thus, the image problem is avoided.

The complex mixing function described above can be implemented using the image-reject mixer illustrated in Fig. 2.12. The image-reject mixer consists of a pair of real mixers driven by an LO operating in quadrature. This complex mixing function serves as

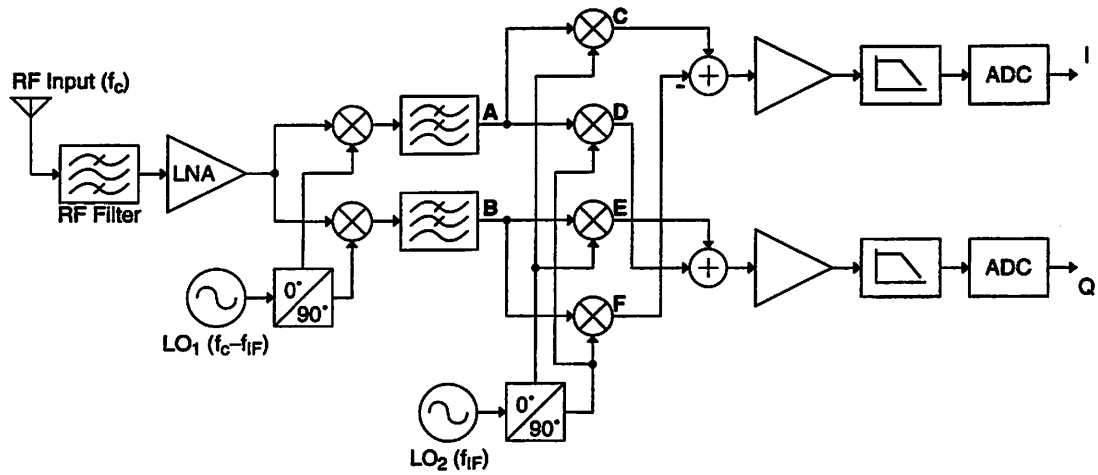


Figure 2.13: Image-reject (Weaver) architecture block diagram.

the basis for the image-reject receiver architecture, also called the Weaver architecture [11], illustrated in Fig. 2.13. Suppose that the signal appearing at the antenna is

$$s(t) = s_{desired}(t) + s_{image}(t) \tag{2.9}$$

where $s_{desired}(t)$ is the desired signal and $s_{image}(t)$ is the undesired image signal. The desired signal can be expressed as

$$s_{desired}(t) = I(t)\cos(2\pi f_c t) + Q(t)\sin(2\pi f_c t) \tag{2.10}$$

where $I(t)$ and $Q(t)$ are the desired baseband I and Q signals, and the image signal can be expressed as

$$s_{image}(t) = I_{image}(t)\cos[2\pi(f_c - 2f_{IF})t] + Q_{image}(t)\sin[2\pi(f_c - 2f_{IF})t]. \tag{2.11}$$

The received signal $s(t)$ is filtered and amplified before being downconverted to the intermediate frequency along two signal paths using two mixers driven by I and Q LO signals at $f_c - f_{IF}$. At the intermediate frequency, each of the two signals is low-pass filtered in order to remove the mixing components at $2f_c - f_{IF}$. At points A and B, the signals are, respectively,

$$s_A(t) = \frac{1}{2}\{[I(t) + I_{image}(t)]\cos(2\pi f_{IF}t) + [Q(t) - Q_{image}(t)]\sin(2\pi f_{IF}t)\} \tag{2.12}$$

$$s_B(t) = \frac{1}{2}\{[Q(t) + Q_{image}(t)]\cos(2\pi f_{IF}t) + [-I(t) + I_{image}(t)]\sin(2\pi f_{IF}t)\}. \tag{2.13}$$

Each of the signals, $s_A(t)$ and $s_B(t)$, are then downconverted to baseband along two signal paths using two mixers driven by I and Q LO signals at f_{IF}. The signals at points C, D, E, and F are, respectively,

$$s_C(t) = \frac{1}{4}[I(t) + I_{image}(t)] \tag{2.14}$$

$$s_D(t) = \frac{1}{4}[Q(t) - Q_{image}(t)] \tag{2.15}$$

$$s_E(t) = \frac{1}{4}[Q(t) + Q_{image}(t)] \tag{2.16}$$

$$s_F(t) = \frac{1}{4}[-I(t) + I_{image}(t)]. \tag{2.17}$$

Subtracting (2.17) from (2.14) gives the desired baseband signal $I(t)$ while eliminating the unwanted image signal $I_{image}(t)$. Similarly, the sum of (2.15) and (2.16) gives the desired baseband signal $Q(t)$ while eliminating the unwanted image signal $Q_{image}(t)$.

2.4.1 Practical Considerations

Unfortunately, in practice, the amount of image rejection achievable by implementations based on this architecture is limited by the gain mismatch between the different signal paths of the receiver as well as by the quadrature phase mismatch between the I and Q signals in the two local oscillators. The image rejection is given by [11]

$$R(dB) = 10\log\left[\frac{1+(1+\Delta A)^2 + 2(1+\Delta A)\cos(\Delta\phi_{LO1} + \Delta\phi_{LO2})}{1+(1+\Delta A)^2 - 2(1+\Delta A)\cos(\Delta\phi_{LO1} - \Delta\phi_{LO2})}\right] \tag{2.18}$$

where ΔA is the gain error, and $\Delta\phi_{LO1}$ and $\Delta\phi_{LO2}$ are the phase errors in the first and second local oscillators, respectively. For ΔA = 3%, $\Delta\phi_{LO1}$ = 2°, and $\Delta\phi_{LO2}$ = 0°, the Weaver architecture achieves an image rejection of 32.8 dB. Integrated circuit implementations typically achieve 25 – 40 dB of image rejection. If additional image rejection is required, a high intermediate frequency can be used so that the image signal is far away from the center frequency of the RF filter, which then provides additional attenuation of the image signal.

Finally, the Weaver architecture is also susceptible to dc offsets and flicker noise. Self-mixing due to the second LO can occur during downconversion of the received signal from the intermediate frequency to baseband. And similar to the direct-conversion architecture, even-order distortion and baseband circuit mismatch can also result in dc offsets in the Weaver architecture. In addition, since the desired signal is frequency translated to baseband prior to analog-to-digital conversion, flicker noise from the baseband analog amplifiers and filters can potentially corrupt the desired signal.

2.5 Low-IF Architecture

One way to avoid the problems associated with dc offsets and flicker noise is to perform analog-to-digital conversion at the intermediate frequency. By using digital circuit techniques to downconvert the desired signal from the intermediate frequency to baseband as well as to perform the subsequent amplification and filtering, impairments associated with analog implementation techniques can be avoided. However, due to conversion speed limitations in analog-to-digital converters, this approach is limited to low intermediate frequencies, and consequently, a third receiver architecture which is based on this technique is called the low-IF architecture [12] (Fig. 2.14).

Since the received signal is downconverted to an intermediate frequency, the low-IF architecture must also contend with the image problem. Consequently, the same image-reject mixing technique used in the Weaver architecture must also be used in the low-IF architecture. As with the Weaver architecture, the amount of image rejection achievable by integrated circuit implementations based on the low-IF architecture is limited by gain and phase mismatches. Moreover, the RF filter can not be used to provide additional image rejection since the use of a low intermediate frequency places the image signal close to the desired signal and within the passband of the RF filter. Consequently the low-IF architecture is limited to applications which have relaxed image-rejection requirements at small frequency offsets from the desired signal.

2.5.1 Digital Frequency Translation to Baseband

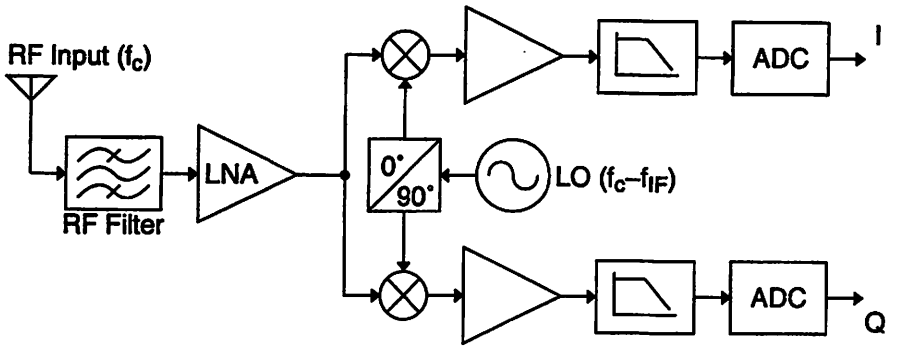


Figure 2.14: Low-IF architecture block diagram.

In the low-IF architecture, additional digital circuitry is required to perform the downconversion from the intermediate frequency to baseband. In the example illustrated in Fig. 2.15, the in-phase and quadrature LO signals are generated digitally, for example, using a read-only memory (ROM) look-up table, and digital frequency translation is performed using four multipliers. The multiplier outputs are then combined using two adders to provide cancellation of the image signal.

Finally, one particular relationship between the intermediate frequency and the ADC sampling frequency results in tremendous simplifications in the hardware required for digital frequency translation. Suppose the ADC sampling frequency, f_s, is exactly four

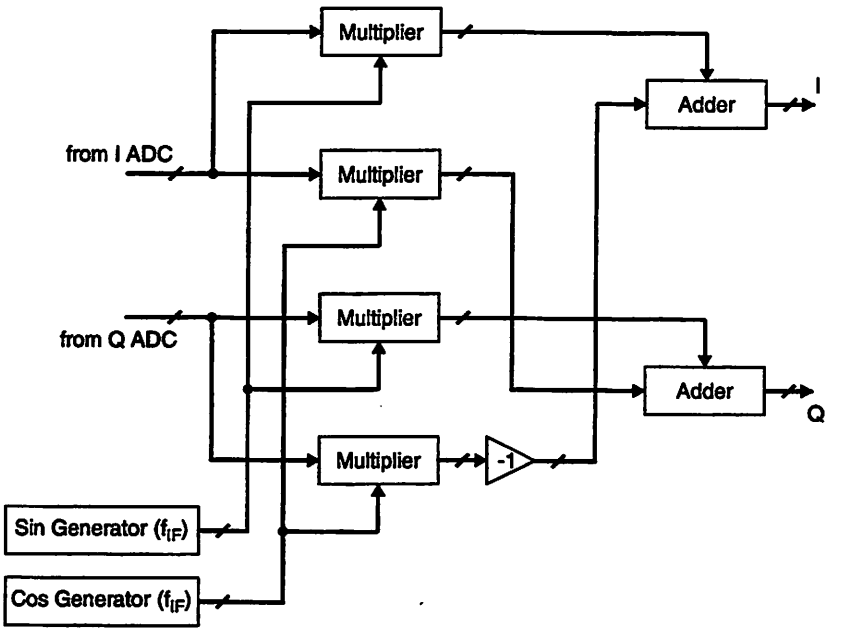


Figure 2.15: Digital frequency translation block diagram.

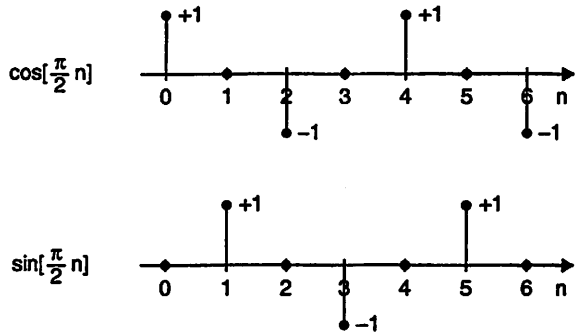


Figure 2.16: LO signals for $f_s = 4f_{IF}$.

times the intermediate frequency, f_{IF}:

$$f_s = 4f_{IF}. \tag{2.19}$$

Then the in-phase and quadrature LO signals are, respectively,

$$\cos(2\pi f_{IF}t) = \cos(2\pi \frac{f_s}{4}t) \rightarrow \cos[\frac{\pi}{2}n] \tag{2.20}$$

$$\sin(2\pi f_{IF}t) = \sin(2\pi \frac{f_s}{4}t) \rightarrow \sin[\frac{\pi}{2}n]. \tag{2.21}$$

As illustrated in Fig. 2.16, a cosine wave of frequency f_{IF} sampled at $4f_{IF}$ yields the sequence $\{\ldots +1, 0, -1, 0, \ldots\}$ while a sine wave of frequency f_{IF} sampled at $4f_{IF}$ yields the sequence $\{\ldots 0, +1, 0, -1, \ldots\}$. Consequently, frequency translation to baseband can be accomplished by simply deinterleaving each of the ADC outputs into two data streams and then toggling the sign of every other data sample. By choosing $f_s = 4f_{IF}$, the digital hardware becomes trivial, eliminating the need for multipliers as well as circuits to generate the digital in-phase and quadrature LO signals.

2.6 Receiver Architecture Selection

The direct-conversion architecture, the image-reject architecture, and the low-IF architecture are all potential candidates for highly-integrated receiver implementations. The hardware requirements for the three receiver architectures are summarized in Table 2.1. The number of LNAs, mixers, LOs, and ADCs are listed for each architecture.

The operating frequency for each LO as well as the minimum sampling rate, f_s, for each ADC are also indicated in the table.

2.6.1 Direct-Conversion Architecture

A major advantage of the direct-conversion architecture is its simplicity. Although, implementations based on the direct-conversion architecture require a minimal number of RF circuit components, they must also contend with dc offsets and flicker noise. The direct-conversion architecture has been used for paging applications which use frequency-shift keying (FSK) as the modulation scheme [13]. In this case, a high-pass filter can be used to eliminate dc offsets since FSK signals have little frequency content near dc. In [13], high-pass filters, which are implemented using 330-pF on-chip capacitors, provide dc blocking up to 150 Hz. Although high-pass filtering is a simple and effective way of eliminating dc offsets, this technique may not be feasible if prohibitively large capacitors or resistors are required. In particular, for signals which occupy a narrow bandwidth, high-pass filters with very low corner frequencies are required, which in turn require very large on-chip capacitors and resistors. In [13], even though the 330 pF on-chip capacitors are implemented using high-density dielectric capacitors (1 nF/mm^2), they still occupy over half of the 18-mm^2 chip area. Moreover, although techniques such as autozeroing and chopper stabilization (Section 5.3.5) are effective in suppressing DC offsets and flicker noise, these techniques also introduce additional complexity in implementing the baseband circuits. Consequently, the direct-conversion architecture is best suited for applications which use a wideband signaling

	direct-conversion	image-reject	low-IF
LNAs	1	1	1
mixers	2	6	2
LOs	1 @f_c	1 @$f_c - f_{IF}$ 1 @f_{IF}	1 @$f_c - f_{IF}$
ADCs	2 @$f_s > 2f_{sig}$	2 @$f_s > 2f_{sig}$	2 @$f_s > 2(f_{sig} + f_{IF})$
other	dc offsets flicker noise	dc offsets flicker noise	digital frequency translation

Table 2.1: Comparison of receiver hardware requirements.

scheme. In this case, a very low cutoff frequency is not required and dc offsets can be removed without requiring prohibitively large capacitors or resistors.

2.6.2 Image-Reject Architecture

A major drawback of the image-reject Weaver architecture is that it requires a large number of RF circuit blocks, including six mixers and two local oscillators. In addition, implementations based on the Weaver architecture must also deal with dc offsets and flicker noise. Nevertheless, this architecture still has its advantages. One major advantage of the Weaver architecture is that it facilitates integration of the frequency synthesizer [11]. In a frequency-division multiple access (FDMA) system, the receiver must be able to perform channel selection by tuning the LO to different frequencies. For example, in the heterodyne architecture, channel selection is performed by the first LO, which frequency translates the desired signal to the fixed center frequency of the IF filter (Fig. 2.3). For a narrowband system, this LO must be able to tune to closely-spaced frequency channels, and consequently, implementations based on a phase-locked loop (PLL) require a low reference frequency. However, in order to ensure loop stability, the PLL bandwidth is limited to one tenth of the reference frequency. Such a narrow bandwidth provides very little attenuation of the phase noise in the voltage-controlled oscillator (VCO). Consequently, in a heterodyne architecture, the VCO is usually implemented using discrete-component high-quality inductors and varactor diodes in order to achieve very low phase noise.

In a completely integrated approach, however, a very low phase noise VCO is difficult to implement due to the lack of high-quality on-chip passive components, especially when using a standard digital CMOS process. In [11], a receiver for the Digital European Cordless Telephone (DECT) system is implemented based on the image-reject Weaver architecture. Since this architecture does not rely on fixed-frequency IF filters for channel selection, the first LO is fixed at 1.7 GHz and the second LO, with a tuning range of 181 – 197 MHz, is used to downconvert the desired signal to baseband. Since the first LO is fixed at 1.7 GHz, it can be implemented using a PLL with a wide loop bandwidth, which significantly reduces the phase noise contributed by an integrated VCO. In

addition, since the second LO operates at a much lower frequency, it can be implemented with good phase noise performance even with low-quality on-chip passive components. Consequently, for communications systems with very stringent phase noise requirements, the image-reject Weaver architecture with a fixed-frequency first LO is highly amenable to single-chip receiver implementations.

2.6.3 Low-IF Architecture

In the low-IF architecture, if the ADC sampling frequency is four times the intermediate frequency, then the additional digital circuitry needed to downconvert the desired signal to baseband becomes trivial. In this case, the low-IF architecture may appear to be a better alternative to the direct-conversion architecture since it also requires a minimal number of RF circuit components but avoids the problems associated with dc offsets and flicker noise. However, implementations based on the low-IF architecture may not be feasible under certain conditions. First, if the bandwidth of the desired signal is very large, then a very high ADC sampling rate is required. In order to avoid aliasing, the sampling frequency of the ADCs must be at least twice the highest frequency component of the input signal. For the low-IF architecture, the minimum ADC sampling frequency is

$$f_s > 2(f_{IF} + f_{sig}) \tag{2.22}$$

where f_{IF} is the intermediate frequency and f_{sig} is the single-sided bandwidth of the desired baseband signal. Since f_{IF} must be greater than f_{sig} in order to avoid problems associated with dc offsets and flicker noise, a wideband signaling scheme would require a prohibitively fast ADC sampling frequency. Consequently, the low-IF architecture is best suited for applications which use a narrowband signaling scheme.

The low-IF architecture also requires relaxed image-reject requirements at small frequency offsets from the desired signal. This requirement, however, is not prohibitively restrictive since many communications systems provide for relaxed interferer levels in nearby frequency channels. For example, in [14], a GSM (Global System Mobile) receiver is implemented based on the low-IF architecture. GSM is a European digital cellular system which uses a narrowband signaling scheme with a single-sided baseband

bandwidth of 100 kHz. The receiver described in [14] uses an intermediate frequency of 200 kHz and requires only 32-dB image rejection. In contrast, the DECT receiver described in [11], which is based on the Weaver architecture, uses an intermediate frequency in the range 181 – 197 MHz and requires more than 70-dB image rejection.

2.6.4 Receiver Architecture Selection Guidelines

As seen from the examples presented above, system-level specifications, such as modulation scheme, signal bandwidth, and interference rejection requirements, strongly influence the choice of receiver architecture. In some cases, these system-level specifications are determined without considering implementation issues, and

	advantages	disadvantages	design guidelines
heterodyne	• excellent sensitivity and selectivity performance	• large number of discrete components	• use this architecture when all else fails
direct-conversion	• minimal number of RF components	• dc offsets and flicker noise	• for wideband signaling schemes, dc offsets can be removed with a high-pass filter using on-chip capacitors and resistors
image-reject	• facilitates integration of low phase noise LO	• large number of RF components • dc offsets and flicker noise • image-rejection is limited by gain and phase mismatches	• fixed-frequency first LO facilitates the use of a wideband PLL for VCO phase noise suppression • tunable low-frequency second LO for channel selection
low-IF	• minimal number of RF components • avoids problems associated with dc offsets and flicker noise	• ADC sampling frequency must be at least $f_{IF} + f_{sig}$ • image-rejection is limited by gain and phase mismatches	• narrowband signaling schemes relax ADC sampling requirement • requires relaxed image-rejection requirements at small frequency offsets from the desired signal • if $f_s = 4f_{IF}$, the digital downconversion circuitry becomes trivial

Table 2.2: Receiver architecture summary and guidelines.

consequently, single-chip implementations are very difficult to achieve or just simply infeasible. With the allocation of unlicensed spectra in the 900-MHz, 2.4-GHz, and 5-GHz frequency bands, many new wireless communications systems will be designed for custom applications. These custom applications present an opportunity to design systems which are more amenable to highly-integrated low-power CMOS receiver implementations. The advantages and disadvantages of the heterodyne architecture, direct-conversion architecture, the image-reject architecture, and the low-IF architecture are summarized in Table 2.2. In addition, a few general design guidelines are also provided.

Chapter 3

Receiver Impairments

3.1 Introduction

Wireless communications systems are continuing to take advantage of the exponential improvements in mainstream CMOS technology by integrating increasingly more functionality onto a single chip. Advanced communications algorithms, which in the past were considered too complex to implement due to the stringent power consumption and form-factor restrictions of mobile devices, are now being considered for future wireless systems. These algorithms are very amenable to low-power digital design techniques and promise orders of magnitude improvement in spectral efficiency, resulting in increased capacity and higher data rates. However, one of the most critical components of any wireless system which may ultimately limit the performance of these communications algorithms is the receiver front-end. Analog front-end impairments, such as noise, distortion and mismatch, may affect algorithm performance by degrading the integrity of the desired signal. This chapter describes the various impairments associated with analog front-ends as well as their consequences on a quadrature phase-shift keying (QPSK) signal constellation [15].

3.2 Quadrature Phase-Shift Keying Modulation

In digital communications systems, the binary data must be mapped to a set of corresponding signal waveforms for transmission. In a phase-shift keying (PSK) signaling scheme, the data is modulated on the phase of the carrier and the corresponding signal waveforms are represented as [16]

$$\begin{aligned} s_m(t) &= g(t)\cos\left[2\pi f_c t + \frac{2\pi}{M}(m-1)\right], \quad m = 1,2,\ldots,M, \quad 0 \le t \le T \\ &= g(t)\cos\left[\frac{2\pi}{M}(m-1)\right]\cos(2\pi f_c t) - g(t)\sin\left[\frac{2\pi}{M}(m-1)\right]\sin(2\pi f_c t) \end{aligned} \tag{3.1}$$

where $g(t)$ is a signal pulse which shapes the spectrum of the transmitted signal, f_c is the carrier frequency, M is the number of symbols, and T is the symbol period. As seen in (3.1) the signal waveforms can also be represented as the linear combination of two quadrature carriers, $\cos(2\pi f_c t)$ and $\sin(2\pi f_c t)$. The amplitude $\cos[2\pi(m-1)/M]$ is typically referred to as the in-phase (I) data while the amplitude $\sin[2\pi(m-1)/M]$ is typically referred to as the quadrature (Q) data. The mapping of the data bits to the phase of the carrier can be represented in a constellation diagram. The constellation diagram for a four-phase PSK ($M = 4$) signaling scheme with an initial phase value of $\pi/4$ is illustrated in Fig. 3.1. Four-phase PSK or quadrature phase-shift keying (QPSK) is a very popular modulation scheme used in wireless communications.

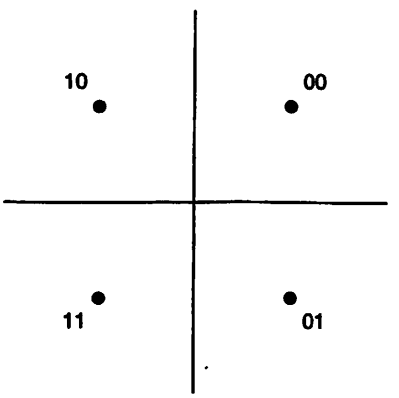


Figure 3.1: QPSK constellation diagram.

When a QPSK signal is corrupted by additive white Gaussian noise (AWGN), the probability of error is [16]

$$\begin{aligned} P_b &= \frac{1}{2}\left\{1-\left[1-Q\left(\sqrt{\frac{d^2}{2N_0}}\right)\right]^2\right\} \\ &= Q\left(\sqrt{\frac{d^2}{2N_0}}\right)-\frac{1}{2}Q^2\left(\sqrt{\frac{d^2}{2N_0}}\right) \\ &\approx Q\left(\sqrt{\frac{d^2}{2N_0}}\right) \end{aligned} \tag{3.2}$$

where d is the distance between adjacent symbols, N_0 is the noise power spectral density, and $Q(x)$ is defined as

$$\begin{aligned} Q(x) &= \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt, \quad x \geq 0 \\ &= \frac{1}{2}\text{erfc}\left(\frac{x}{\sqrt{2}}\right). \end{aligned} \tag{3.3}$$

In the case of ideal QPSK, the distance between adjacent signals is just

$$d = 2\sqrt{E_b} \tag{3.4}$$

where E_b is the energy per bit. Consequently, (3.2) becomes

$$P_b = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{3.5}$$

where E_b/N_0 is the SNR. The bit-error probability is plotted as a function of SNR in Fig. 3.2. From the figure, an SNR of 8.4 dB is required to achieve a BER of 10^{-4}.

3.3 Receiver Noise

Two types of electrical noise which are particularly important in CMOS implementations are thermal noise and flicker noise. Thermal noise arises from the random thermal motion

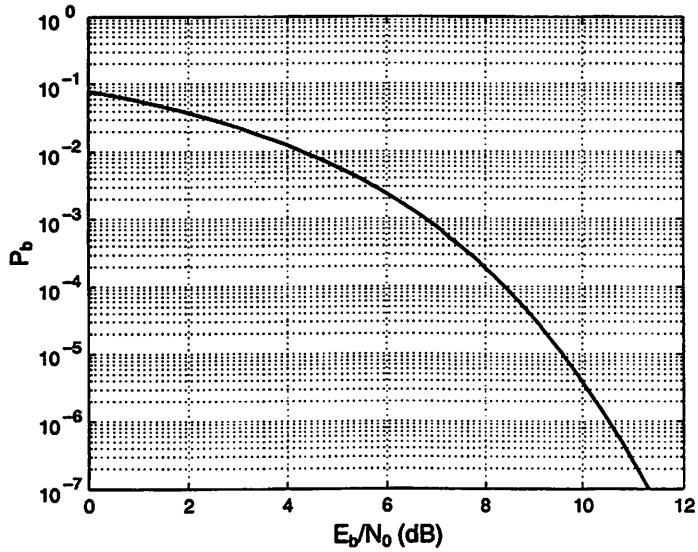


Figure 3.2: Probability of error for QPSK.

of electrons and is present in any passive resistor. The noise associated with a resistance R has a mean-square voltage

$$\overline{v^2} = 4kTR\Delta f \tag{3.6}$$

where k is Boltzmann's constant, T is temperature and Δf is bandwidth in hertz. Since the voltage spectral density is constant over frequency, thermal noise is white. In addition, its amplitude has a Gaussian probability distribution. At the receiver, thermal noise is present in the antenna as well as in the other passive and active devices which are used to implement the analog front-end. The thermal noise in the antenna is associated with the antenna's radiation resistance and originates from the black-body radiation in the transmission environment, while the thermal noise in active devices is associated with the resistive channel of the underlying CMOS transistors. Transistor thermal noise can be represented by a noise generator between the device drain and source terminals with mean-square current

$$\overline{i_d^2} = 4kT\gamma g_{do}\Delta f \tag{3.7}$$

where g_{do} is the drain-source conductance when the drain-source voltage, V_{DS}, is 0 V and γ is a constant for a particular technology. Thermal noise is an especially important design consideration in the LNA but must also be considered in the design of other RF

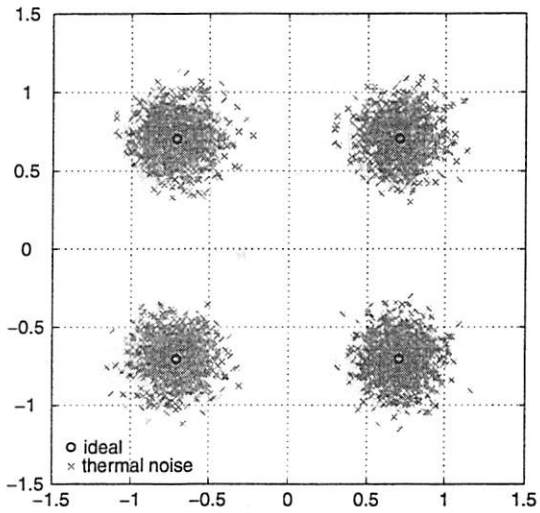


Figure 3.3: Effect of thermal noise on a QPSK constellation.

blocks such as mixers as well as baseband components such as amplifiers and filters. The effect of thermal noise on a QPSK constellation is illustrated in Fig. 3.3 and the probability of error is given by (3.5).

Another type of noise which may adversely affect the performance of digital communications systems is flicker noise. Flicker noise occurs in CMOS transistors and can be modeled by a noise generator between the device drain and source terminals with mean-square current

$$\overline{i_d^2} = K\frac{I_D^a}{f}\Delta f \tag{3.8}$$

where K is a constant for a particular device, I_D is the drain bias current and a is constant for a particular technology. The flicker noise current spectral density is inversely proportional to frequency and its amplitude is often not Gaussian [10]. Flicker noise is especially problematic in receivers which frequency translate the desired signal to dc prior to analog-to-digital conversion, since the flicker noise spectral density can be quite large at low frequencies. The problem is exacerbated in CMOS implementations since flicker noise performance is significantly worse for CMOS transistors than for devices in other technologies such as silicon bipolar. The effect of flicker noise on a QPSK constellation is illustrated in Fig. 3.4. Although the effect on the constellation is similar to

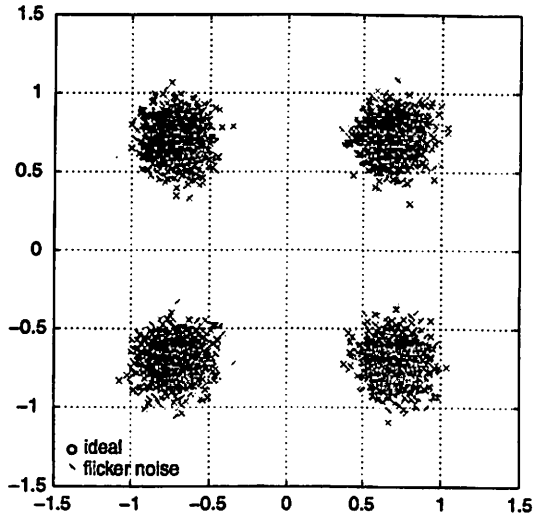


Figure 3.4: Effect of flicker noise on a QPSK constellation.

that of thermal noise, unfortunately, the probability of error in this case is difficult to determine since flicker noise is not AWGN.

3.4 Gain Mismatch

Another analog impairment which degrades performance in digital communications systems is gain mismatch along the different receiver signal paths. For example, in the direct-conversion architecture illustrated in Fig. 2.7, the received signal is downconverted to baseband along parallel I and Q signal paths. The gain along these two signal paths should be identical, but in practical implementations, may be different due to circuit mismatches. Gain mismatch in a direct-conversion receiver can be modeled by the block diagram illustrated in Fig. 3.5, where A is the average gain and α is the difference in gain

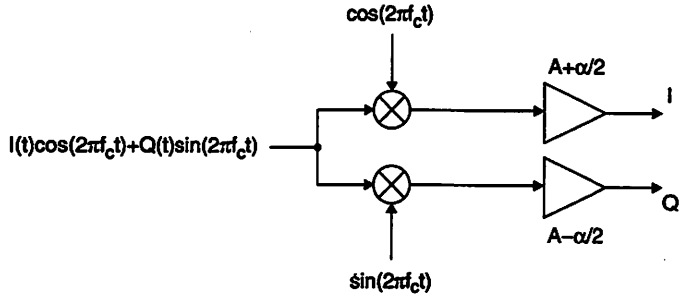


Figure 3.5: I and Q gain mismatch in a direct-conversion receiver.

along the I and Q signal paths. Thus, with gain mismatch, the output of the I signal path is given by

$$I(t)\frac{A}{2}\left(1+\frac{\alpha}{2A}\right) \tag{3.9}$$

while the output of the Q signal path is given by

$$Q(t)\frac{A}{2}\left(1-\frac{\alpha}{2A}\right). \tag{3.10}$$

The effect of gain mismatch on a QPSK constellation is illustrated in Fig. 3.6.

When a QPSK signal with gain mismatch is corrupted by AWGN, the probability of error can be approximated by averaging the error probabilities for two binary antipodal signals separated by distances $d_0(1+\alpha/2A)$ and $d_0(1-\alpha/2A)$:

$$\begin{aligned} P_b &\approx \frac{1}{2}\left\{Q\left[\left(1+\frac{\alpha}{2A}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]+Q\left[\left(1-\frac{\alpha}{2A}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]\right\} \\ &\approx \frac{1}{2}\left\{Q\left[\left(1+\frac{\alpha}{2A}\right)\sqrt{\frac{2E_b}{N_0}}\right]+Q\left[\left(1-\frac{\alpha}{2A}\right)\sqrt{\frac{2E_b}{N_0}}\right]\right\}. \end{aligned} \tag{3.11}$$

The error probabilities as a function of SNR for various amounts of gain mismatch are plotted in Fig. 3.7. As seen in Fig. 3.7, the BER degradation is minimal for a gain

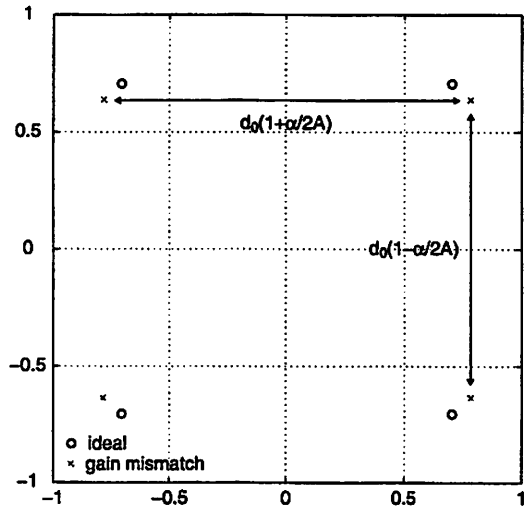


Figure 3.6: Effect of gain mismatch on a QPSK constellation.

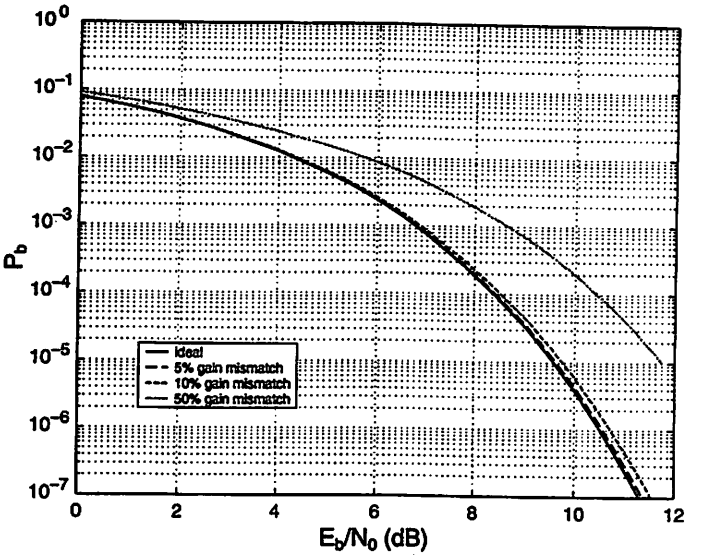


Figure 3.7: Probability of error for QPSK with gain mismatch.

mismatch of 5% or less, which is indeed achievable in highly-integrated receiver implementations. Consequently, the direct-conversion architecture is relatively insensitive to gain mismatch.

The image-reject and low-IF architectures, however, are much more sensitive to gain mismatch. In these two architectures, gain mismatch limits the amount of image rejection, as already discussed in Section 2.4.1.

3.5 Quadrature Phase Mismatch

Quadrature phase mismatch in the LO signals also degrades the performance of digital communications systems. For example, in the direct-conversion architecture, frequency

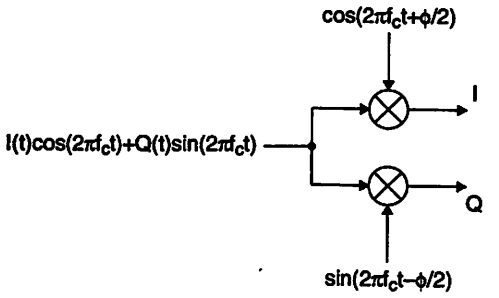


Figure 3.8: Quadrature phase mismatch in a direct-conversion receiver.

translation is performed using two mixers and an LO fixed at the carrier frequency and operating in quadrature. In practical implementations, the phases of the LO signals may deviate from quadrature due to circuit mismatches. Quadrature phase mismatch in a direct-conversion receiver can be modeled by the block diagram illustrated in Fig. 3.8, where ϕ is the deviation from quadrature in the two LO signals. Thus, with quadrature phase mismatch, the output of the I signal path is given by

$$\frac{1}{2}\left[I(t)\cos\left(\frac{\phi}{2}\right)-Q(t)\sin\left(\frac{\phi}{2}\right)\right] \tag{3.12}$$

while the output of the Q signal path is given by

$$\frac{1}{2}\left[Q(t)\cos\left(\frac{\phi}{2}\right)-I(t)\sin\left(\frac{\phi}{2}\right)\right]. \tag{3.13}$$

The effect of quadrature phase mismatch on a QPSK constellation is illustrated in Fig. 3.9.

When a QPSK signal with quadrature phase mismatch is corrupted by AWGN, the probability of error can be approximated by averaging the error probabilities for the symbols in each of the four quadrants of the constellation diagram. Each of the symbols in the first and third quadrants are located at a distance $(d_0/2)[\cos(\phi/2)-\sin(\phi/2)]$ from

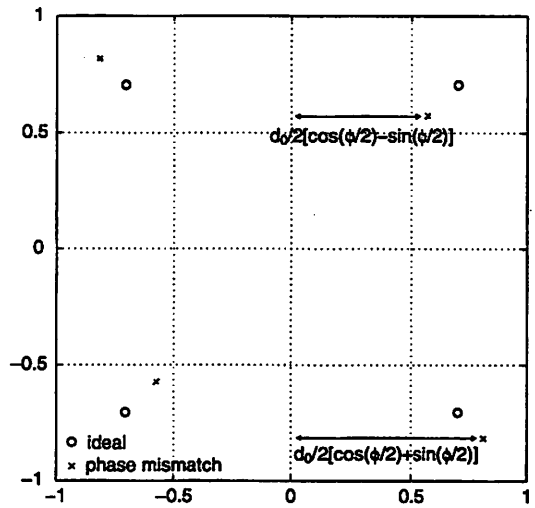


Figure 3.9: Effect of quadrature phase mismatch on a QPSK constellation.

the decision boundaries and the error probability for these two symbols is approximately

$$P_{b(I,III)} \approx \frac{1}{2}\left\{Q\left[\left(\cos\frac{\phi}{2}-\sin\frac{\phi}{2}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]+Q\left[\left(\cos\frac{\phi}{2}-\sin\frac{\phi}{2}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]\right\}$$
$$\approx Q\left[\left(\cos\frac{\phi}{2}-\sin\frac{\phi}{2}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]. \tag{3.14}$$

Similarly, each of the symbols in the second and fourth quadrants are located at a distance $(d_0/2)[\cos(\phi/2)+\sin(\phi/2)]$ from the decision boundaries and the error probability for these two symbols is approximately

$$P_{b(II,IV)} \approx \frac{1}{2}\left\{Q\left[\left(\cos\frac{\phi}{2}+\sin\frac{\phi}{2}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]+Q\left[\left(\cos\frac{\phi}{2}+\sin\frac{\phi}{2}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]\right\}$$
$$\approx Q\left[\left(\cos\frac{\phi}{2}+\sin\frac{\phi}{2}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]. \tag{3.15}$$

Consequently, the overall probability of error is approximately

$$P_b \approx \frac{1}{2}[P_{b(I,III)}+P_{b(II,IV)}]$$
$$\approx \frac{1}{2}\left\{Q\left[\left(\cos\frac{\phi}{2}-\sin\frac{\phi}{2}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]+Q\left[\left(\cos\frac{\phi}{2}+\sin\frac{\phi}{2}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]\right\} \tag{3.16}$$
$$\approx \frac{1}{2}\left\{Q\left[\left(\cos\frac{\phi}{2}-\sin\frac{\phi}{2}\right)\sqrt{\frac{2E_b}{N_0}}\right]+Q\left[\left(\cos\frac{\phi}{2}+\sin\frac{\phi}{2}\right)\sqrt{\frac{2E_b}{N_0}}\right]\right\}.$$

The error probabilities as a function of SNR for various amounts of gain mismatch are plotted in Fig. 3.10. As seen in Fig. 3.10, the BER degradation is minimal for a phase mismatch of 5° or less, which is indeed achievable in highly-integrated receiver implementations. Consequently, the direct-conversion architecture is also relatively insensitive to quadrature phase mismatch.

As in the case of gain mismatch, the image-reject and low-IF architectures, are also much more sensitive to quadrature phase mismatch. In these two architectures, phase mismatch limits the amount of image rejection, as already discussed in Section 2.4.1.

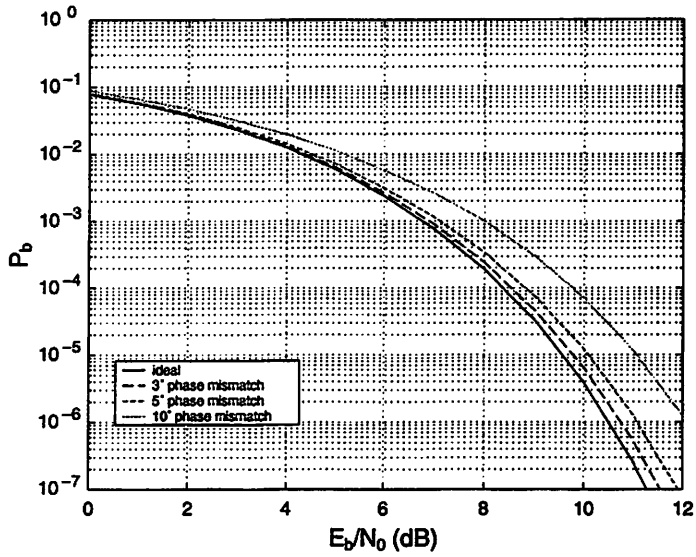


Figure 3.10: Probability of error for QPSK with quadrature phase mismatch.

3.6 Frequency Offset

The local oscillators in the transmitter and receiver are based on accurate frequency references. However, due to their physical separation, different frequency references are used in the transmitter and receiver. The frequency stability of these references are limited to $50-100$ ppm for low-cost crystal references [17] or $1-50$ ppm for high-stability references [6], where the frequency stability is defined as

$$\text{frequency stability}\,[\text{ppm}] \equiv \frac{\Delta f\,[\text{Hz}]}{f_c\,[\text{MHz}]} \tag{3.17}$$

where f_c is the nominal frequency and Δf is the frequency offset. The use of different

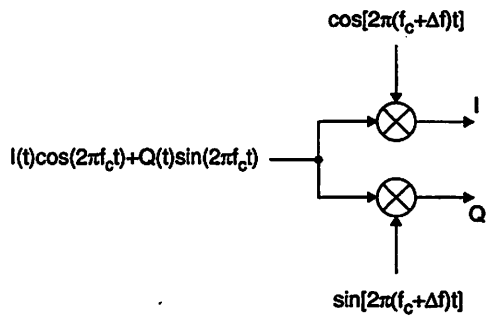


Figure 3.11: Frequency translation using I and Q LO signals with frequency offset.

frequency references in the transmitter and receiver introduces a frequency offset between the local oscillators, which can be modeled by the block diagram in Fig. 3.11. The output of the I signal path is

$$\frac{1}{2}[I(t)\cos(2\pi\Delta ft) - Q(t)\sin(2\pi\Delta ft)] \tag{3.18}$$

while the output of the Q signal path is

$$\frac{1}{2}[Q(t)\sin(2\pi\Delta ft) + I(t)\cos(2\pi\Delta ft)]. \tag{3.19}$$

Consequently, frequency offset results in a spinning constellation, as illustrated in Fig. 3.12. If the frequency offset is sufficiently small relative to the signal bandwidth, a differentially-encoded signaling scheme such as differential phase-shift keying (DPSK) can be used [18]. In this case, the data is encoded in the transitions between symbols so that only the phase difference between successive symbols is needed for signal demodulation. Unfortunately, differentially-encoded modulation schemes require a larger SNR for the same BER performance as modulation schemes which use absolute phase encoding. For example, at large SNR, four-phase DPSK requires about 2.3 dB additional SNR for the same BER performance as four-phase PSK [16].

If coherent demodulation is preferred or if the frequency offset is large relative to the

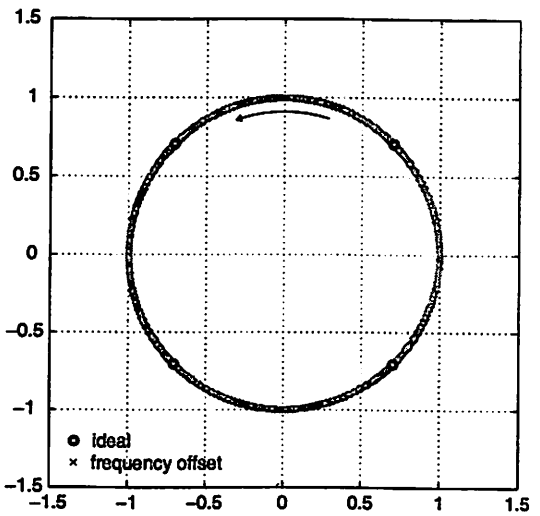


Figure 3.12: Effect of frequency offset on a QPSK constellation.

(a) (b)

Figure 3.13: Frequency spectra of LO signals. (a) Ideal. (b) With phase noise.

signal bandwidth, then frequency offset compensation is required. Several techniques for frequency estimation and compensation are discussed in [19].

3.7 LO Phase Noise

Both thermal noise and flicker noise contribute to the nonideal behavior of the LO called phase noise. The frequency spectra of LO signals with and without phase noise are illustrated in Fig. 3.13. At the receiver, the desired signal can be quite weak and may be accompanied by very strong interfering signals at adjacent frequencies. If the interfering signals are not sufficiently attenuated, they can potentially corrupt the desired signal through reciprocal mixing, as illustrated in Fig. 3.14. Fig. 3.15 illustrates the constellation diagram for a QPSK signal when the desired signal is accompanied by a single-tone interferer at an adjacent frequency and is downconverted by an LO signal with phase noise.

Even if interfering signals are not present, the close-in phase noise of the LO degrades the receiver performance by corrupting the information contained in the phase of the carrier.

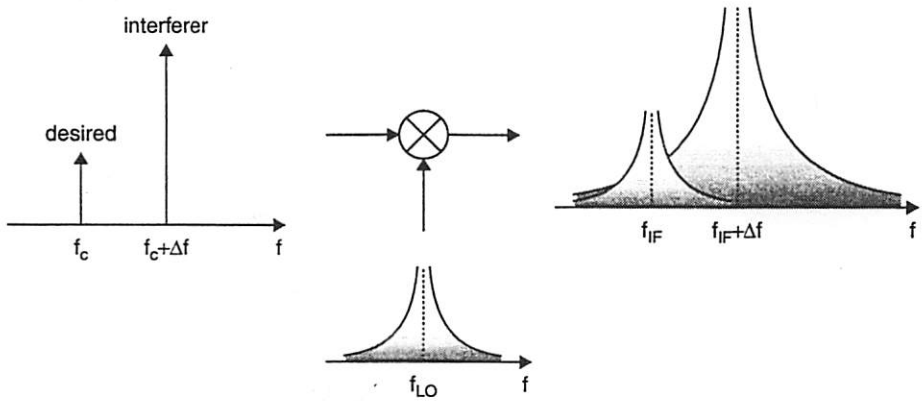


Figure 3.14: Reciprocal mixing.

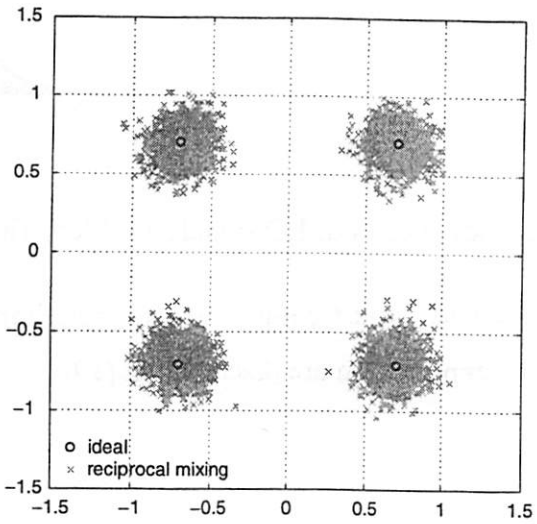


Figure 3.15: Effect of reciprocal mixing on a QPSK constellation.

Frequency translation using in-phase and quadrature LO signals with phase noise $\phi(t)$ can be modeled by the block diagram in Fig. 3.16. The output of the I signal path is

$$\frac{1}{2}[I(t)\cos\phi(t) - Q(t)\sin\phi(t)] \tag{3.20}$$

while the output of the Q signal path is

$$\frac{1}{2}[Q(t)\sin\phi(t) + I(t)\cos\phi(t)]. \tag{3.21}$$

Consequently, close-in phase noise results in a time-varying phase-offset in the received symbols. The effect of close-in phase on a QPSK constellation is illustrated in Fig. 3.17, where the amount of constellation rotation is equal to the magnitude of $\phi(t)$.

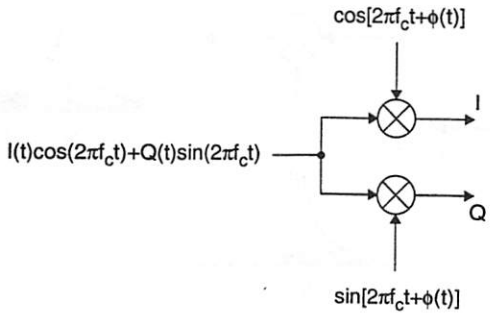


Figure 3.16: Frequency translation using I and Q LO signals with phase noise.

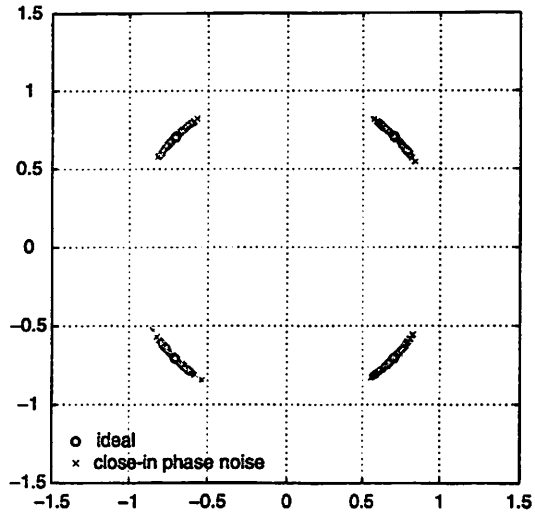


Figure 3.17: Effect of close-in phase noise on a QPSK constellation.

3.8 Receiver Distortion

Another source of performance degradation in a receiver is distortion. In a nonlinear system, the output can contain frequency components which are not present in the input signal. A nonlinear system can be described by the following transfer function:

$$s_o = a_1 s_i + a_2 s_i^2 + a_3 s_i^3 + \ldots \tag{3.22}$$

where s_i and s_o are the input and output signals, respectively. Harmonic distortion occurs when a single sinusoid is applied to a nonlinear system. In this case, the output signal consists of frequency components which are integer multiples of the input frequency, f_1, as illustrated in Fig. 3.18.

Intermodulation distortion occurs when two sinusoids of different frequencies, f_1 and f_2, are applied to a nonlinear system. In the case of second-order intermodulation distortion,

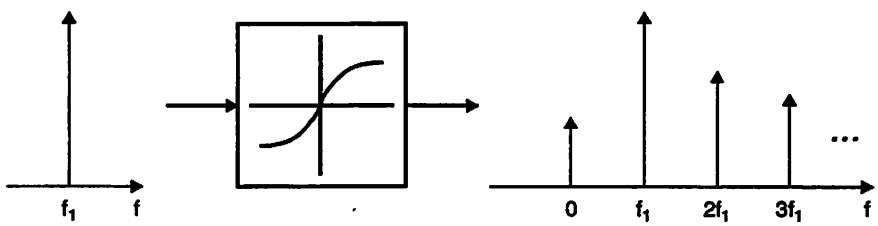


Figure 3.18: Harmonic distortion.

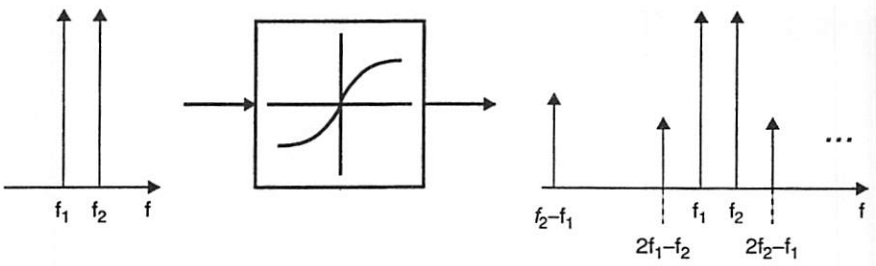


Figure 3.19: Intermodulation distortion.

the output signal consists of frequency components at $f_2 - f_1$ and $f_1 + f_2$, while in the case of third-order intermodulation distortion, the output signal consists of frequency components at $2f_1 - f_2$ and $2f_2 - f_1$ (Fig. 3.19).

As an example of how distortion can degrade the performance in the receiver, consider a received signal which consists of a weak desired signal accompanied by two large fixed-amplitude sinusoidal interferers at adjacent frequencies, f_1 and f_2, as illustrated in Fig. 3.20. The third-order intermodulation product at $2f_1 - f_2$ from the two interferers corrupts the desired signal and results in the constellation diagram illustrated in Fig. 3.21.

In addition, as already mentioned in Section 2.3.1, even-order distortion is particularly troublesome in direct-conversion receivers since it gives rise to signal-dependent dc offsets and low-frequency components, all of which can potentially corrupt the desired signal.

3.9 Filtering

The received signal usually contains the desired signal as well as undesired signals at

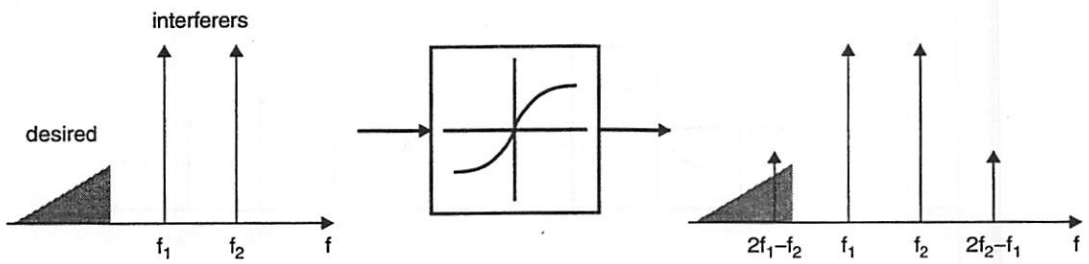


Figure 3.20: Third-order intermodulation distortion.

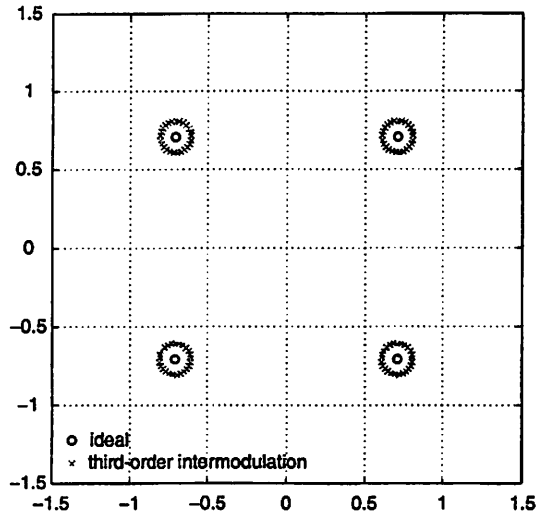


Figure 3.21: Effect of intermodulation distortion on a QPSK constellation.

adjacent frequencies. In the receiver, filters are used to pass the desired signal while rejecting the unwanted frequency components. Several types of filters are commonly used to approximate an ideal low-pass filter response, including the Butterworth and Chebyshev filters [20].

The magnitude response of a Butterworth filter is given by

$$|H(j\omega)| = \frac{1}{\sqrt{1+\varepsilon^2\left(\frac{\omega}{\omega_p}\right)^{2N}}} \tag{3.23}$$

where N is the filter order, ω_p is the passband edge, and ε determines the magnitude variation in the passband. At $\omega = \omega_p$, the magnitude is

$$|H(j\omega)| = \frac{1}{\sqrt{1+\varepsilon^2}}. \tag{3.24}$$

The magnitude responses for several Butterworth filters with $\varepsilon = 1$ are illustrated in Fig. 3.22. A Butterworth filter provides a maximally flat response and the degree of flatness in the passband increases as the filter order increases.

Another commonly used filter is the Chebyshev filter, which exhibits an equiripple response in the passband. The magnitude response of a Chebyshev filter is given by

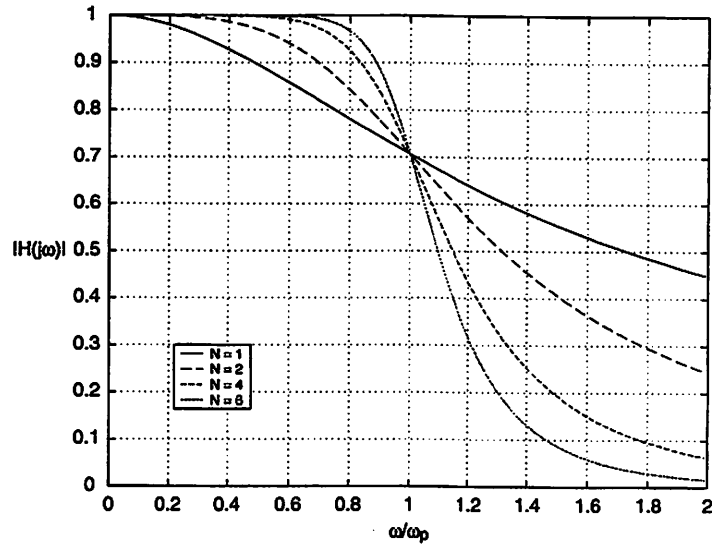


Figure 3.22: Butterworth filter magnitude responses with $\varepsilon = 1$.

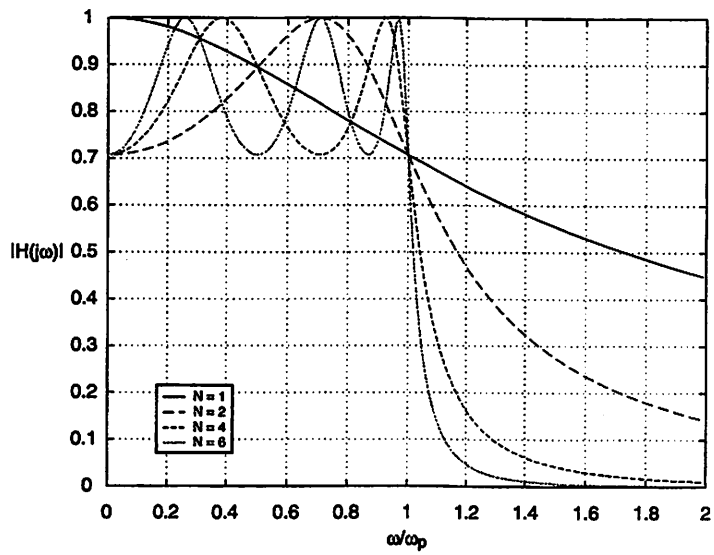


Figure 3.23: Chebyshev filter magnitude responses with $\varepsilon = 1$.

$$|H(j\omega)| = \begin{cases} \dfrac{1}{\sqrt{1+\varepsilon^2 \cos^2\left[N\cos^{-1}\left(\dfrac{\omega}{\omega_p}\right)\right]}}, & 0 \le \omega \le \omega_p \\ \dfrac{1}{\sqrt{1+\varepsilon^2 \cosh^2\left[N\cosh^{-1}\left(\dfrac{\omega}{\omega_p}\right)\right]}}, & \omega \ge \omega_p \end{cases} \tag{3.25}$$

where N is the filter order, ω_p is the passband edge, and ε determines the magnitude variation in the passband. At $\omega = \omega_p$, the magnitude is

$$|H(j\omega)| = \frac{1}{\sqrt{1+\varepsilon^2}}. \tag{3.26}$$

The magnitude responses for several Chebyshev filters with $\varepsilon = 1$ are illustrated in Fig. 3.23. For the same order and the same passband variation, a Chebyshev filter provides greater stopband attenuation than a Butterworth filter. Alternatively, a lower-order Chebyshev filter can achieve the same stopband attenuation as a higher-order Butterworth filter. However, the Chebyshev filter achieves its superior stopband attenuation at the price of increased nonlinearity in the phase response. The phase responses for third-order Butterworth and Chebyshev filters are illustrated Fig. 3.24.

The effect on receiver performance must also be considered when selecting a particular filter response. The effect of a fourth-order Butterworth filter, with $\varepsilon = 1$, on a QPSK signal along with the corresponding eye diagram are illustrated in Fig. 3.25. This Butterworth filter degrades performance slightly when using a QPSK signaling scheme. As seen from the constellation diagram, this Butterworth response affects the I and Q signal amplitudes but leaves their phase relationship unchanged. The eye diagram offers a

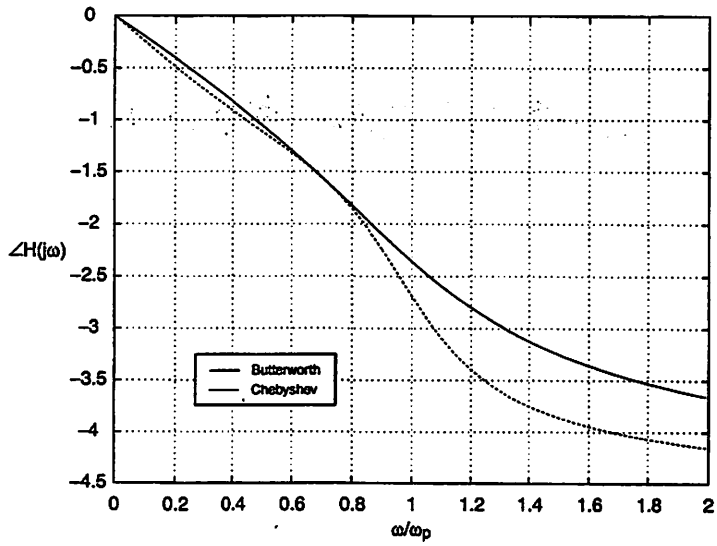


Figure 3.24: Phase responses for third-order Butterworth and Chebyshev filters.

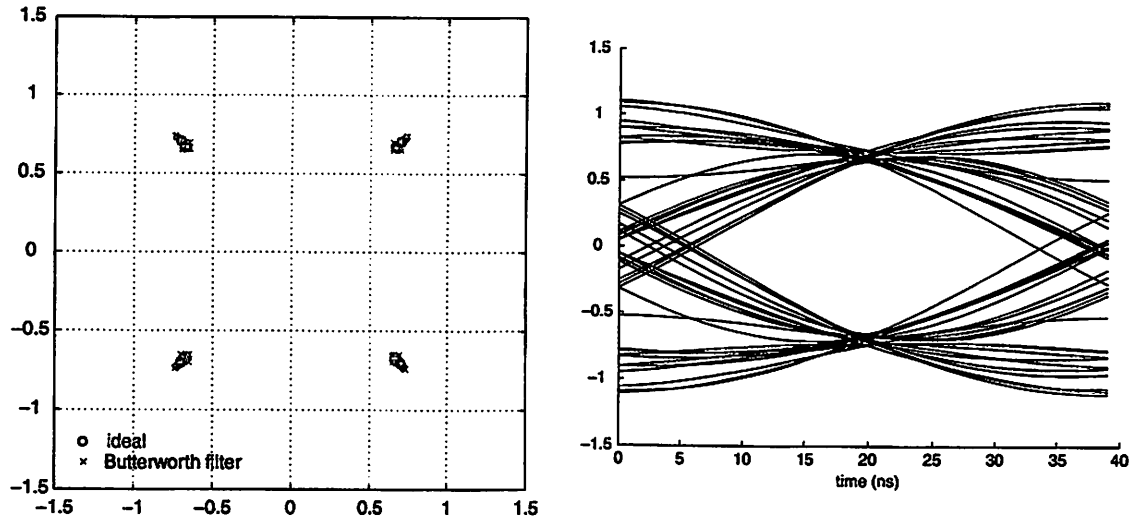


Figure 3.25: Butterworth filter. (a) Effect on a QPSK constellation. (b) Eye diagram.

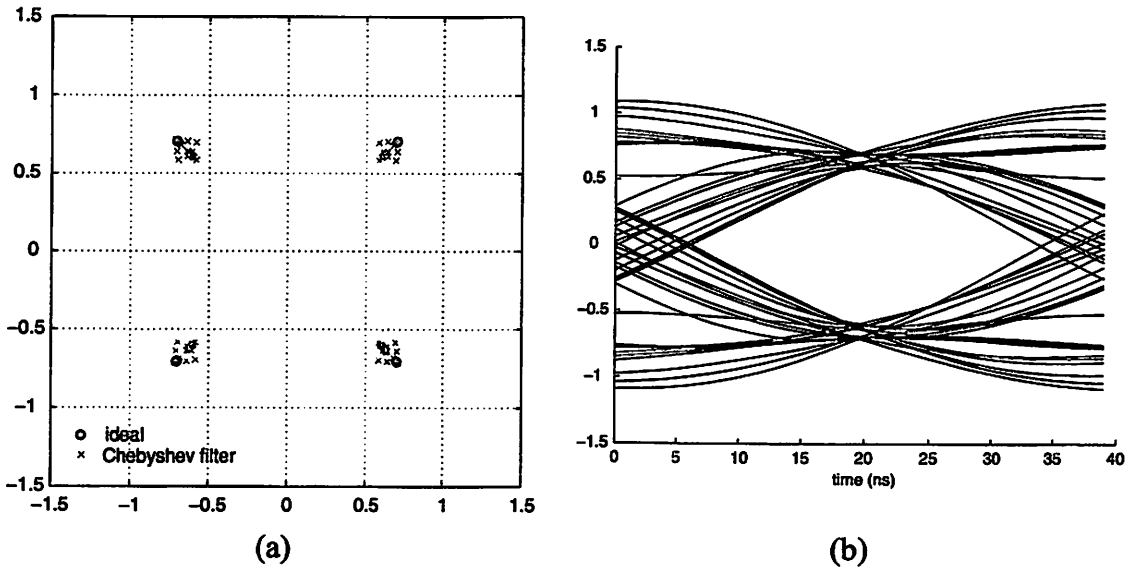


Figure 3.26: Chebyshev filter. (a) Effect on a QPSK constellation. (b) Eye diagram.

different perspective, revealing that this Butterworth response introduces a small amount of intersymbol interference (ISI), which is consistent with the amplitude deviation seen in the constellation diagram.

The effect of a fourth-order Chebyshev filter, with a maximum passband ripple of 1 dB, on a QPSK signal and the corresponding eye diagram are illustrated in Fig. 3.26. The

constellation diagram reveals that the nonlinear phase response of the Chebyshev filter alters the phase relationship between the I and Q signals.

Consequently, when selecting a filter response, in addition to achieving sufficient attenuation of undesired signals, the filter must also maintain the integrity of the desired signal. In addition to the Butterworth and Chebyshev filters, other potential candidates for filtering in the receiver include the Bessel, inverse Chebyshev, and elliptic filters [20].

3.10 DC Offsets

Both the direct-conversion and image-reject architectures are sensitive to dc offsets since in both architectures, the desired signal is downconverted to baseband prior to analog-to-digital conversion. DC offsets are problematic for two reasons. First, dc offsets can saturate the baseband circuits, such as amplifiers and filters, and consequently, the receiver ceases to function properly. Second, even if the baseband circuits do not saturate, dc offsets, if uncorrected, degrade the performance of the system. Fig. 3.27 illustrates the effect of dc offsets on a QPSK constellation, where Δd_I and Δd_Q are the dc offsets in the I and Q channels, respectively.

When a QPSK signal with dc offsets is corrupted by AWGN, the probability of error can be approximated by averaging the error probabilities for two binary antipodal signals,

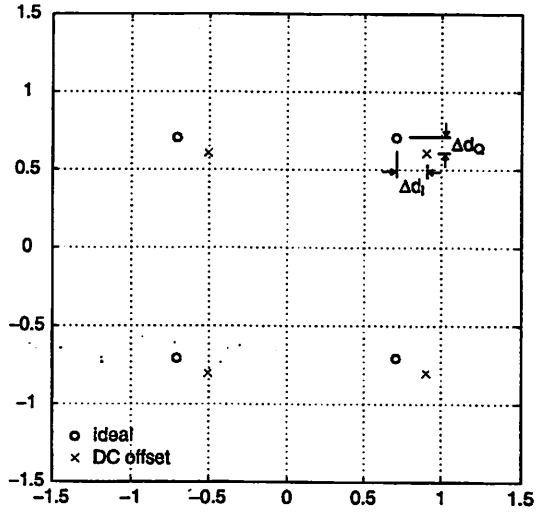


Figure 3.27: Effect of dc offsets on a QPSK constellation.

each with a dc offset. For the I component, the dc offset is Δd_I, and the two symbols are located at distances $(d_0/2)(1-2\Delta d_I/d_0)$ and $(d_0/2)(1+2\Delta d_I/d_0)$ from the decision boundary. Consequently, the probability of error for the I component is approximately

$$P_{b(I)} \approx \frac{1}{2}\left\{Q\left[\left(1-\frac{2\Delta d_I}{d_0}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]+Q\left[\left(1+\frac{2\Delta d_I}{d_0}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]\right\}. \tag{3.27}$$

Similarly, for the Q component, the dc offset is Δd_Q, and the two symbols are located at distances $(d_0/2)(1-2\Delta d_Q/d_0)$ and $(d_0/2)(1+2\Delta d_Q/d_0)$ from the decision boundary. The probability of error for the Q component is approximately

$$P_{b(Q)} \approx \frac{1}{2}\left\{Q\left[\left(1-\frac{2\Delta d_Q}{d_0}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]+Q\left[\left(1+\frac{2\Delta d_Q}{d_0}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]\right\}. \tag{3.28}$$

Thus, the overall error probability is approximately

$$\begin{aligned} P_b &\approx \frac{1}{2}[P_{b(I)}+P_{b(Q)}] \\ &\approx \frac{1}{4}\left\{Q\left[\left(1-\frac{2\Delta d_I}{d_0}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]+Q\left[\left(1+\frac{2\Delta d_I}{d_0}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]+\right. \\ &\quad \left. Q\left[\left(1-\frac{2\Delta d_Q}{d_0}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]+Q\left[\left(1+\frac{2\Delta d_Q}{d_0}\right)\sqrt{\frac{d_0^2}{2N_0}}\right]\right\} \\ &\approx \frac{1}{4}\left\{Q\left[\left(1-\frac{2\Delta d_I}{d_0}\right)\sqrt{\frac{2E_b}{N_0}}\right]+Q\left[\left(1+\frac{2\Delta d_I}{d_0}\right)\sqrt{\frac{2E_b}{N_0}}\right]+\right. \\ &\quad \left. Q\left[\left(1-\frac{2\Delta d_Q}{d_0}\right)\sqrt{\frac{2E_b}{N_0}}\right]+Q\left[\left(1+\frac{2\Delta d_Q}{d_0}\right)\sqrt{\frac{2E_b}{N_0}}\right]\right\}. \end{aligned} \tag{3.29}$$

The error probabilities as a function of SNR for various amounts of dc offset are plotted in Fig. 3.28. For dc offsets greater than 1%, the BER degradation can be significant.

3.11 ADC Quantization

In many receivers, an ADC converts the received signal to a digital signal for further processing. In order to avoid destructive aliasing, the ADC must sample the input signal

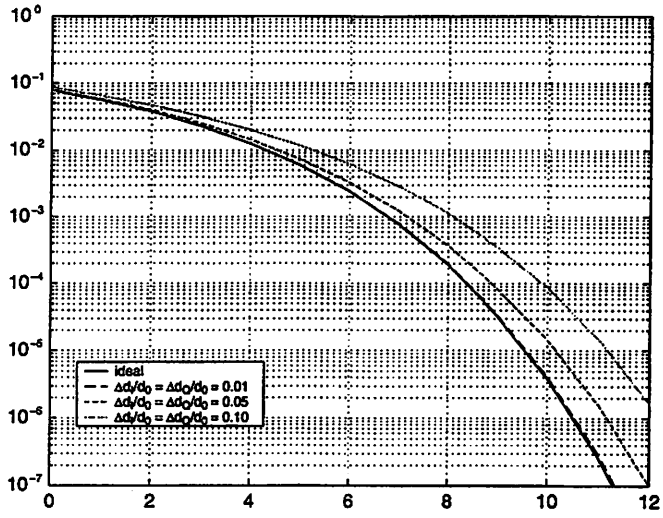


Figure 3.28: Probability of error for QPSK with dc offsets.

at the Nyquist sampling rate, $f_s \geq 2f_{sig}$, where f_{sig} is the highest frequency contained in the input signal. The ADC quantizes each sample to one of 2^R amplitude levels, where R is the number of bits used to represent each sample. The finite resolution of the quantizer results in an error, q_n, in the digital output signal,

$$q_n = \bar{x}_n - x_n \tag{3.30}$$

where x_n is the analog input signal and $\bar{x}_n$ is the digital output signal. For a uniform quantizer, the error is statistically characterized by the uniform probability density function [21]

$$p(q) = \frac{1}{\Delta}, \quad -\frac{1}{2}\Delta \leq q \leq \frac{1}{2}\Delta \tag{3.31}$$

where Δ is the step size of the quantizer. The step size Δ is

$$\Delta = \frac{V_{FS}}{2^R} \tag{3.32}$$

where V_{FS} is the full-scale level of the ADC, and the mean square value of the quantization error is

$$E(q^2) = \frac{1}{12}\Delta^2 = \frac{1}{12} \times \frac{V_{FS}^2}{2^{2R}} = 20\log V_{FS} - 6R - 10.8 \text{ dB}. \tag{3.33}$$

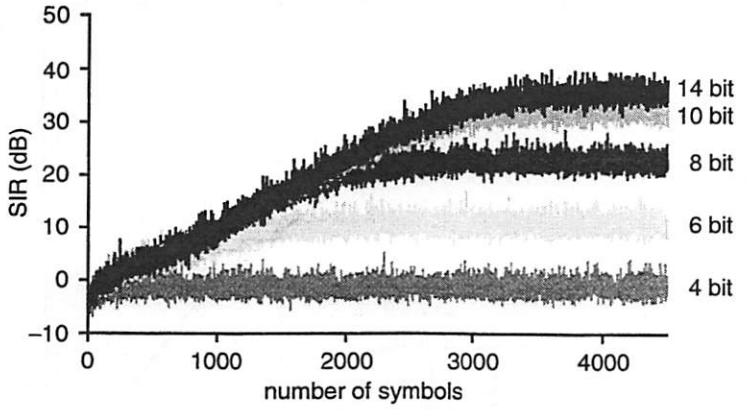


Figure 3.29: Effect of ADC resolution on an adaptive MUD algorithm [22].

Consequently the quantization noise decreases by 6 dB for each additional bit used in the quantizer. The amount of quantization noise introduced by the ADC affects the performance of the algorithms implemented in the subsequent digital circuitry. For example, Fig. 3.29 illustrates the effect of ADC resolution on an adaptive multiuser detection (MUD) algorithm [22]. Clearly, a larger number of bits improves the signal-to-interference ratio (SIR).

3.12 Summary

Receiver front-end impairments, such as noise, distortion, and mismatch, can potentially degrade the performance of digital communications systems. This chapter described the effects of many of these impairments on a QPSK signal constellation. By including models for all of the analog impairments described in this chapter, the simulation framework described in the next chapter facilitates the direct evaluation of the effects of these impairments on the performance of digital communications algorithms.

Chapter 4

System-Level Simulation Framework

4.1 Introduction

All of the analog front-end impairments described in Chapter 3 can potentially degrade the performance of digital communications systems. In a conventional approach, link budget calculations based on a few system-level specifications are performed to determine the allowable levels of receiver impairments, such as noise and distortion [23]. This approach can be advantageous since it abstracts the analog hardware design from the detailed system-level specifications. However, the abstraction offered by this approach can also be a drawback since it does not offer the ability to more closely evaluate the effects of the analog front-end impairments on the performance of digital communications algorithms. This chapter describes a system-level simulation framework which allows for complete end-to-end simulations of communications systems. This framework includes models for the nonideal behavior of analog front-end components and facilitates the exploration of tradeoffs between analog impairments and overall system performance.

4.2 Receiver Performance Calculations

In a conventional approach, the allowable levels of receiver impairments are determined from a few system-level specifications, such as required BER, reference sensitivity, and worst-case out-of-band blocker levels. This method of determining the performance requirements of the analog front-end components is based strictly on numerical calculations. The following sections provide a brief overview of receiver performance calculations for receiver noise and distortion requirements.

4.2.1 Noise Calculations

In RF circuit design, the receiver noise performance is typically characterized in terms of noise factor or noise figure. The noise factor, F, of an RF circuit component is defined as

$$F \equiv \frac{SNR_{in}}{SNR_{out}} \tag{4.1}$$

where SNR_{in} and SNR_{out} are the signal-to-noise ratios at the input and output, respectively. The noise figure is simply the noise factor expressed in decibels, $10\log(F)$. Although it is the preferred metric for noise performance in RF circuit design, noise figure is typically not a system-level specification. However, the noise figure can be calculated based on the reference sensitivity, the bandwidth of the desired signal, and the required BER.

One of the primary system-level specifications is the reference sensitivity, which is defined as the minimum signal level which the receiver must be able to correctly detect. The input SNR, SNR_{in} in (4.1), is simply the reference sensitivity divided by the noise power at the receiver input. The noise at the input of the receiver originates from the black-body radiation in the transmission environment and is modeled by a radiation resistance, R_s, associated with the antenna. Consequently, the noise power which appears at the receiver input is determined by the voltage divider between the receiver input resistance and the antenna source resistance (Fig. 4.1):

$$P_{noise} = \frac{\overline{v_{in}^2}}{R_{in}} = \frac{R_{in}R_s}{(R_{in}+R_s)^2} 4kT\Delta f \,. \tag{4.2}$$

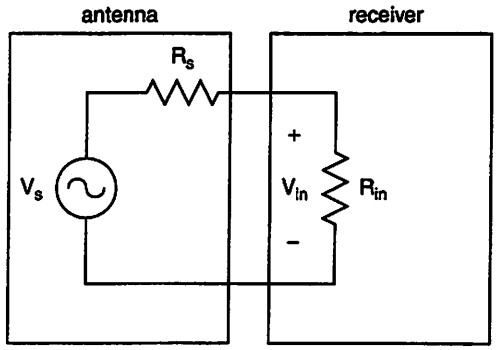


Figure 4.1: Circuit model of antenna with receiver input.

Most receivers are designed for maximum power transfer from the antenna to the input of the receiver, in which case, $R_{in} = R_s$ and (4.2) becomes

$$P_{noise} = kT\Delta f \,. \tag{4.3}$$

In RF circuit design, power levels are commonly referred to in decibels referenced to 1 mW, or 0 dBm. At 300 K, the noise power in dBm is

$$\begin{aligned} P_{noise}\,[\text{dBm}] &= 10\log(1.38\times10^{-23}\ \text{J/K}\times300\,\text{K}\times10^{3}\ \text{mW/W}) + 10\log\Delta f \\ &= -173.8 + 10\log\Delta f \end{aligned} \tag{4.4}$$

where Δf is the signal bandwidth in hertz. Consequently, the input SNR in decibels is just

$$\begin{aligned} SNR_{in}\,[\text{dB}] &= P_{sig}\,[\text{dBm}] - P_{noise}\,[\text{dBm}] \\ &= P_{sig}\,[\text{dBm}] - 10\log(\Delta f\,[\text{Hz}]) + 173.8 \end{aligned} \tag{4.5}$$

where P_{sig} is the reference sensitivity.

The other parameter required to calculate the receiver noise figure is the output SNR, SNR_{out} in (4.1). Although the output SNR is typically not explicitly specified at the system-level, it can be determined from the maximum tolerable BER. A reliable communications link is guaranteed when the specified BER is achieved. The BER for a particular communications system is related to the SNR of the received signal, and the exact relationship depends on a number of factors including the modulation scheme as well as the specific algorithm used for signal detection. For example, for a QPSK

modulation scheme, the BER as a function of SNR is given in (3.5) and illustrated in Fig. 3.2. Consequently, the required noise figure can be expressed as

$$\begin{aligned} NF\,[\text{dB}] &= SNR_{in}\,[\text{dB}] - SNR_{out}\,[\text{dB}] \\ &= P_{sig}\,[\text{dBm}] - 10\log(\Delta f\,[\text{Hz}]) - SNR_{out}[\text{dB}] + 173.8 \end{aligned} \tag{4.6}$$

where P_{sig} is the reference sensitivity, Δf is the bandwidth of the desired signal, and SNR_{out} is the SNR corresponding to the maximum tolerable BER.

Once the required noise performance of the receiver is determined, the noise budget must be partitioned between the various receiver building blocks. For cascaded receiver stages, the total noise factor is given by the Friis equation [24]:

$$F_{tot} = F_1 + \frac{F_2 - 1}{G_1} + \ldots + \frac{F_N - 1}{\prod_{i=1}^{N} G_i} \tag{4.7}$$

where F_i and G_i are the noise factor and power gain, respectively, of the i^{th} receiver stage.

Finally, before concluding this section, it is worthwhile to clarify the distinction between single-sideband (SSB) noise figure and double-sideband (DSB) noise figure [25]. Fig. 4.2 depicts an RF input signal which is corrupted by AWGN. The power and bandwidth of the RF signal are P_{sig} and W, respectively, while the PSD of the AWGN is $N/2$, and consequently, the input SNR is

$$SNR_{in} = \frac{P_{sig}}{N \times W}. \tag{4.8}$$

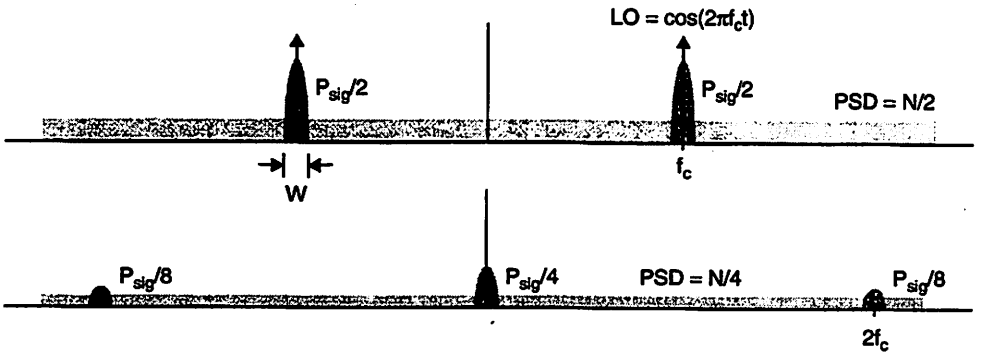


Figure 4.2: Frequency translation of RF signal directly to baseband.

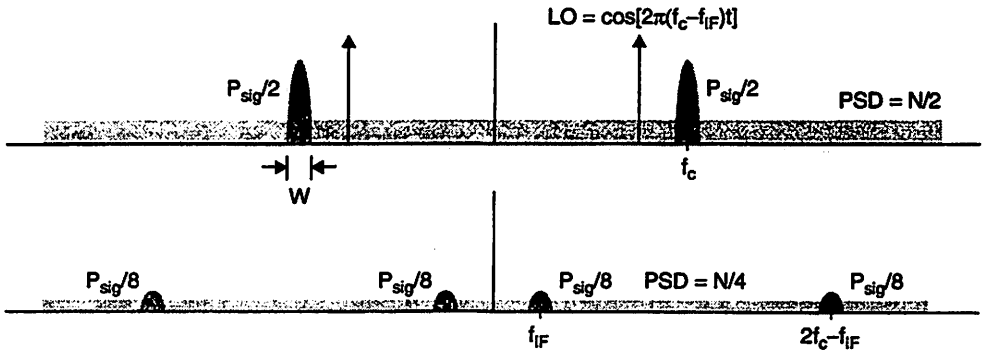


Figure 4.3: Frequency translation of RF signal to an intermediate frequency.

If the RF signal is frequency translated directly to baseband by an LO signal $\cos(2\pi f_c t)$, as illustrated in Fig. 4.2, the output SNR of the resulting baseband signal is

$$SNR_{out} = \frac{P_{sig}}{N \times W}. \tag{4.9}$$

In this case, the noise figure determined by (4.8) and (4.9) is referred to as the DSB noise-figure. However, if the RF signal is frequency translated to an intermediate frequency by an LO signal $\cos[2\pi(f_c - f_{IF})t]$, as illustrated in Fig. 4.3, the output SNR of the resulting IF signal is

$$SNR_{out} = \frac{P_{sig}}{2N \times W}. \tag{4.10}$$

In this case, the noise figure determined by (4.8) and (4.10) is referred to as the SSB noise figure and is 3 dB higher than the DSB noise figure:

$$NF_{SSB} = NF_{DSB} + 3\,\text{dB}. \tag{4.11}$$

4.2.2 Distortion Calculations

In RF circuit design, the distortion performance of the receiver is typically characterized by several measured parameters such as 1-dB compression point and intermodulation intercept point. These parameters are commonly modeled using approaches based on either power series or Volterra series [26]. The latter approach provides a very accurate characterization of nonlinearities in circuits with memory, i.e., circuits with capacitors

and inductors. However, calculations using Volterra series are rather complex, even when using computer simulation techniques. On the other hand, calculations using power series are much more tractable at the expense of only providing an accurate description of distortion in circuits that are memoryless. A system with second-order and third-order nonlinearities can be described by the following power series:

$$s_o = a_1 s_i + a_2 s_i^2 + a_3 s_i^3 \tag{4.12}$$

where s_i and s_o are the input and output signals, respectively. For an input signal, $s_i = A\cos(2\pi ft)$, the output signal is

$$s_o = \frac{a_2 A^2}{2} + \left(a_1 A + \frac{3a_3 A^3}{4}\right)\cos(2\pi ft) + \frac{a_2 A^2}{2}\cos[2\pi(2f)t] + \frac{a_3 A^3}{4}\cos[2\pi(3f)t]. \tag{4.13}$$

As seen in (4.13), third-order distortion alters the gain of the fundamental component, $\cos(2\pi ft)$. For large input signal amplitudes A,

$$\frac{3a_3 A^2}{4} \approx a_1 . \tag{4.14}$$

Since a_3 is typically negative, the gain of the fundamental component decreases, or compresses, for large input signal amplitudes. The 1-dB compression point is defined as the input signal level which causes the gain of the fundamental to drop by 1 dB, as illustrated in Fig. 4.4. The 1-dB compression point of the receiver should be larger than the strongest anticipated input signal. The 1-dB compression point can be determined from (4.13):

$$20\log\left|a_1 + \frac{3a_3 A_{comp}^2}{4}\right| = 20\log|a_1| - 1\,\text{dB}. \tag{4.15}$$

Solving for A_{comp} gives

$$A_{comp} = \sqrt{0.145\left|\frac{a_1}{a_3}\right|}. \tag{4.16}$$

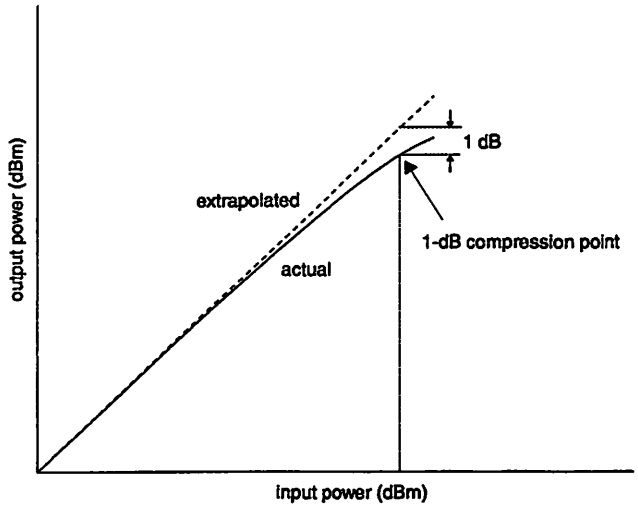


Figure 4.4: 1-dB compression point.

Another metric of a receiver's distortion performance is intermodulation intercept point. As already discussed in Section 3.8, intermodulation distortion occurs when two sinusoids of different frequencies, f_1 and f_2, are applied to a nonlinear system. In the case of third-order intermodulation distortion, the output signal consists of frequency components at $2f_1 - f_2$ and $2f_2 - f_1$. For the transfer function in (4.12), if the input signal consists of two sinusoids, $A_1 \cos(2\pi f_1 t)$ and $A_2 \cos(2\pi f_2 t)$, then the third-order intermodulation products at $2f_1 - f_2$ and $2f_2 - f_1$ are

$$\frac{3a_3 A_1^2 A_2}{4}\cos[2\pi(2f_1 - f_2)t] + \frac{3a_3 A_1 A_2^2}{4}\cos[2\pi(2f_2 - f_1)t]. \tag{4.17}$$

In RF circuit design, the third-order intermodulation performance is characterized by the third-order intermodulation intercept point (IP_3), which is the point where the amplitudes of the intermodulation products and the fundamental components are equal when two equal amplitude sinusoids are applied at the input. The input amplitude at which the intercept point occurs can be determined from (4.17):

$$\frac{3 \mid a_3 \mid A_{IP3}^3}{4} = \mid a_1 \mid A_{IP3}. \tag{4.18}$$

Solving for A_{IP3} gives

$$A_{IP3} = \sqrt{\frac{4}{3}\left|\frac{a_1}{a_3}\right|}. \tag{4.19}$$

Equation (4.19) can also be expressed in terms of the 1-dB compression point in (4.16):

$$A_{IP3} = \sqrt{\frac{4}{3} \times \frac{1}{0.145}} A_{comp} = 3.03 A_{comp}. \tag{4.20}$$

In practice, the higher-order distortion terms become significant at the intercept point, and consequently, the IP_3 measurement must be performed with sufficiently small input amplitudes so that the contribution from the higher-order distortion terms is negligible. The power levels of the third-order intermodulation products and the fundamental components are plotted for a few input power levels, and the intercept point is determined by the intersection of the two lines extrapolated from these points, as illustrated in Fig. 4.5.

The receiver's required third-order intermodulation distortion requirement can be determined from the system-level specifications for the anticipated levels of out-of-band interferers, the reference sensitivity, and the desired BER. As described in Section 4.2.1, the maximum BER specification corresponds to a minimum SNR which depends on various factors, including the type of modulation scheme and the specific algorithm used for signal detection. Consequently, at the output of the receiver, the third-order

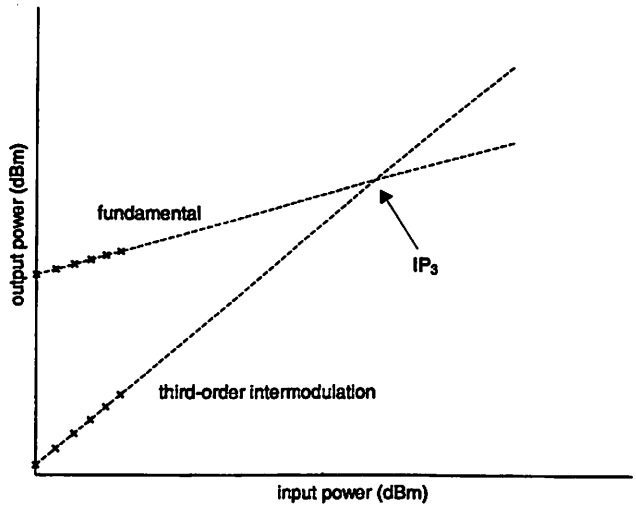


Figure 4.5: Third-order intermodulation intercept point.

intermodulation products due to the out-of-band signals which fall within the bandwidth of the desired signal must be small enough to still maintain the minimum SNR:

$$\frac{G \times P_{sig}}{P_{IM3}} \geq SNR \tag{4.21}$$

where $G = a_1^2$ is the power gain of the receiver, P_{sig} is the power of the desired signal at receiver input, or the reference sensitivity, and P_{IM3} is the power of the intermodulation product at the receiver output. In this case, the intermodulation product is assumed to be uncorrelated with the desired signal and equivalent to AWGN. From (4.17), the power of the intermodulation product is

$$P_{IM3} = \frac{9}{16} a_3^2 P_{int}^3 \tag{4.22}$$

where P_{int} is the power of each of the out-of-band interferers. Substituting (4.22) into (4.21) gives

$$\frac{16}{9} \frac{a_1^2}{a_3^2} \frac{P_{sig}}{P_{int}^3} \geq SNR. \tag{4.23}$$

But from (4.19), the input power level at the third-order intermodulation intercept point is

$$P_{IIP3} = \frac{4}{3} \left| \frac{a_1}{a_3} \right| \tag{4.24}$$

so (4.23) becomes

$$\frac{P_{sig} P_{IIP3}^2}{P_{int}^3} \geq SNR. \tag{4.25}$$

Consequently, the input power level corresponding to the required third-order intermodulation intercept point is

$$P_{IIP3} \geq P_{int} \sqrt{\frac{SNR \times P_{int}}{P_{sig}}} \tag{4.26}$$

or

$$P_{IIP3}\,[\text{dBm}] \geq P_{int}\,[\text{dBm}] - \frac{1}{2}(SNR\,[\text{dB}] + P_{int}\,[\text{dBm}] - P_{sig}\,[\text{dBm}]). \qquad (4.27)$$

Once the required IP_3 performance of the receiver is determined, the distortion budget must be partitioned between the various receiver building blocks. For cascaded receiver stages, if the distortion contribution from the various stages are all uncorrelated, then the input power level corresponding to the third-order intermodulation intercept point can be determined from the following expression

$$\frac{1}{P_{IIP3}} = \frac{1}{P_{IIP3(1)}} + \frac{G_1}{P_{IIP3(2)}} + \ldots + \frac{\prod_{i=1}^{N-1} G_i}{P_{IIP3(N)}} \qquad (4.28)$$

where $P_{IIP3(i)}$ is the input power at the intercept point of the i^{th} stage and G_i is the power gain of the i^{th} stage.

4.2.3 Summary

In a conventional approach, a few system-level specifications, such as maximum BER and reference sensitivity, are converted to other metrics more commonly used in RF circuit design, such as noise figure and intermodulation intercept point, using equations such as (4.6) and (4.27). These requirements are budgeted between the various blocks which make up the receiver using additional equations such as (4.7) and (4.28), and spreadsheets are commonly used in order to facilitate such calculations. Although this approach is relatively straightforward and simple, it lacks the ability to model the detailed interactions between the analog front-end impairments and the digital back-end algorithms. Because of its simplicity, the equation-based method still serves as a good starting point. However, a simulation framework which is capable of end-to-end simulations can provide a much more accurate assessment of the effects of receiver impairments on system-level performance. For such an approach, the simulation environment must be simple and capable of rapid simulation speeds. Consequently, behavioral models, rather than circuit models, should be used for the receiver components. The following sections provide detailed descriptions of a system-level

simulation framework which can be used to explore the effects of analog front-end impairments as well as to evaluate overall receiver performance.

4.3 Baseband-Equivalent Models

Even with the use of behavioral models, a modeling framework which simulates the analog front-end at the carrier frequency is unacceptable in terms of speed. For such a simulation, the maximum step size must be based on the carrier frequency, whereas the total number of steps must be based on the symbol rate. Since the carrier frequency is typically much higher than the symbol rate in wireless communications systems, such a simulation would be prohibitively slow.

In order to decrease the simulation time, the simulation framework relies on baseband-equivalent models for the analog RF building blocks, such as amplifiers, mixers, and oscillators. The method is similar to envelop simulation techniques used in some RF circuit-level simulators [27]. The baseband-equivalent models for the various RF building blocks are based on the following expression which can be used to represent any real signal along the RF signal path:

$$s(t) = s_{DC}(t) + \sum_{n=1}^{N} [s_{In}(t)\cos(n\omega_c t) + s_{Qn}(t)\sin(n\omega_c t)] \tag{4.29}$$

where the bandwidths of $s_{DC}(t)$, $s_{In}(t)$, and $s_{Qn}(t)$ are assumed to be much less than ω_c. For example,

$$s(t) = s_{I1}(t)\cos(\omega_c t) + s_{Q1}(t)\sin(\omega_c t) \tag{4.30}$$

represents an ideal RF signal where $s_{I1}(t)$ and $s_{Q1}(t)$ are the baseband I and Q signals, respectively, and ω_c is the carrier frequency. An example spectrum of $s(t)$ in (4.29) is illustrated in Fig. 4.6. Baseband-equivalent models are derived by letting the input and output signals of the various RF building blocks take the same form as (4.29). Then the dependence on ω_c is eliminated by keeping track of only the time-varying coefficients, $s_{DC}(t)$, $s_{In}(t)$, and $s_{Qn}(t)$. The baseband-equivalent models for RF amplifiers, mixers, and oscillators are derived in the following sections.

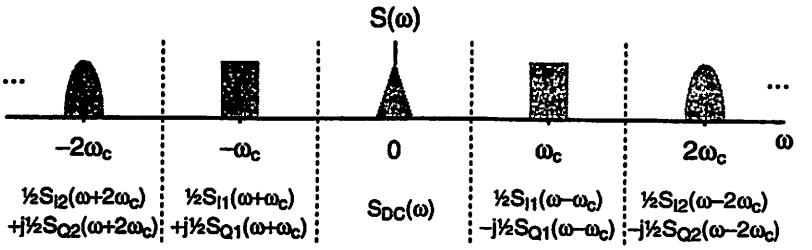


Figure 4.6: Example spectrum of $s(t)$ in (4.29).

4.3.1 RF Amplifiers

The transfer function of any RF gain block such as the LNA can be represented by the following relationship:

$$x_o(t) = \sum_{n=0}^{N} a_n x_i^n(t) \tag{4.31}$$

where $x_i(t)$ and $x_o(t)$ are the input and output signals, respectively. The coefficients a_0 and a_1 represent dc offset and gain, respectively, while the remaining coefficients a_n represent the nonlinear behavior of the RF gain block. Then let $x_i(t)$ and $x_o(t)$ have the same form as (4.29):

$$x_i(t) = x_{iDC}(t) + \sum_{n=1}^{N} [x_{iIn}(t)\cos(n\omega_c t) + x_{iQn}(t)\sin(n\omega_c t)] \tag{4.32}$$

$$x_o(t) = x_{oDC}(t) + \sum_{n=1}^{N} [x_{oIn}(t)\cos(n\omega_c t) + x_{oQn}(t)\sin(n\omega_c t)] \tag{4.33}$$

and solve for the output coefficients $x_{oDC}(t)$, $x_{oIn}(t)$, and $x_{oQn}(t)$ in terms of the input coefficients $x_{iDC}(t)$, $x_{iIn}(t)$, and $x_{iQn}(t)$. The results for $N = 3$ are given in Appendix A. By pre-computing the relationship between these time-varying coefficients instead of keeping track of the actual signals, $x_o(t)$ and $x_i(t)$, the dependence on the carrier frequency is removed.

4.3.2 Local Oscillators

The baseband-equivalent model for oscillators is also derived in a similar fashion by letting the oscillator output signal have the same form as (4.29):

$$y_{LO}(t) = y_{LODC}(t) + \sum_{n=1}^{N}[y_{LOIn}(t)\cos(n\omega_c t) + y_{LOQn}(t)\sin(n\omega_c t)]. \quad (4.34)$$

By specifying appropriate functions for the time-varying coefficients $y_{LODC}(t)$, $y_{LOIn}(t)$, and $y_{LOQn}(t)$, (4.34) accounts for local oscillator impairments such as quadrature phase mismatch, frequency offset, and phase noise. For example, for an oscillator with frequency offset $\Delta\omega$ and phase offset $\Delta\phi$, the output signal can be expressed as

$$\begin{aligned} y_{LO}(t) &= \cos[(\omega_c + \Delta\omega)t + \Delta\phi] \\ &= \cos(\Delta\omega t + \Delta\phi)\cos(\omega_c t) - \sin(\Delta\omega t + \Delta\phi)\sin(\omega_c t) \\ &= y_{LOI1}(t)\cos(\omega_c t) + y_{LOQ1}(t)\sin(\omega_c t) \end{aligned} \quad (4.35)$$

where $y_{LOI1}(t) = \cos(\Delta\omega t + \Delta\phi)$ and $y_{LOQ1}(t) = -\sin(\Delta\omega t + \Delta\phi)$.

4.3.3 Mixers

A similar method is used to derive the baseband-equivalent model for mixers in direct-conversion receivers. The transfer function of a mixer can be represented by

$$y_o(t) = y_i(t) \times y_{LO}(t) \quad (4.36)$$

where $y_i(t)$, $y_{LO}(t)$, and $y_o(t)$ are the input, oscillator, and output signals, respectively. Then let $y_i(t)$, $y_{LO}(t)$, and $y_o(t)$ have the same form as (4.29):

$$y_i(t) = y_{iDC}(t) + \sum_{n=1}^{N}[y_{iIn}(t)\cos(n\omega_c t) + y_{iQn}(t)\sin(n\omega_c t)] \quad (4.37)$$

$$y_{LO}(t) = y_{LODC}(t) + \sum_{n=1}^{N}[y_{LOIn}(t)\cos(n\omega_c t) + y_{LOQn}(t)\sin(n\omega_c t)] \quad (4.38)$$

$$y_o(t) = y_{oDC}(t) + \sum_{n=1}^{N}[y_{oIn}(t)\cos(n\omega_c t) + y_{oQn}(t)\sin(n\omega_c t)] \quad (4.39)$$

and solve for the output coefficients $y_{oDC}(t)$, $y_{oIn}(t)$, and $y_{oQn}(t)$ in terms of the input coefficients $y_{iDC}(t)$, $y_{iIn}(t)$, $y_{iQn}(t)$, $y_{LODC}(t)$, $y_{LOIn}(t)$, and $y_{LOQn}(t)$. Additional reductions in simulation complexity can be achieved by keeping track of only the baseband component $y_{oDC}(t)$. This approximation is valid for large carrier frequencies since the frequency

components at ω_c and greater are significantly attenuated by baseband filtering. The results for $N = 3$ are given in Appendix A.

The derivation of this baseband-equivalent model is based on (4.29), in which all of the harmonic components are assumed to have bandwidths much less than ω_c. In the case of direct-conversion receivers, application of this model is straightforward since the LO and the desired signal are at the same frequency, ω_c, and all the mixing products fall on multiples of ω_c. However, in the case of heterodyne receivers, this baseband-equivalent model must be used with caution since the harmonic components of the output may overlap. For example, consider an input signal centered at the carrier frequency ω_c

$$y_i(t) = y_{iI1}(t)\cos(\omega_c t) \tag{4.40}$$

and an LO signal at $\omega_c - \omega_{IF}$

$$\begin{aligned} y_{LO} &= \cos[(\omega_c - \omega_{IF})t] \\ &= y_{LOI1}(t)\cos(\omega_c t) + y_{LOQ1}(t)\sin(\omega_c t) \end{aligned} \tag{4.41}$$

where

$$y_{LOI1}(t) = \cos(\omega_{IF} t) \tag{4.42}$$

$$y_{LOQ1}(t) = \sin(\omega_{IF} t). \tag{4.43}$$

From Table A.3, the mixer output signal is

$$y_o(t) = y_{oDC}(t) + y_{oI2}(t)\cos(2\omega_c t) + y_{oQ2}(t)\sin(2\omega_c t) \tag{4.44}$$

where

$$y_{oDC}(t) = \frac{1}{2} y_{iI1}(t)\cos(\omega_{IF} t) \tag{4.45}$$

$$y_{oI2}(t) = \frac{1}{2} y_{iI1}(t)\cos(\omega_{IF} t) \tag{4.46}$$

$$y_{oQ2}(t) = \frac{1}{2} y_{iI1}(t)\sin(\omega_{IF} t). \tag{4.47}$$

In low-IF receivers, the intermediate frequency, ω_{IF}, is typically much smaller than the carrier frequency ω_c. In this case, the harmonic components do not overlap, and consequently, the same baseband-equivalent model used for mixers in direct-conversion receivers can be used for mixers in low-IF receivers. However, if $\omega_{IF} > \omega_c/2$, then the spectral content of $y_{oDC}(t)$ in (4.45) actually extends beyond $\omega_c/2$, which violates the bandwidth constraints for the coefficients in (4.29). Likewise, both $y_{oI2}(t)$ in (4.46) and $y_{oQ2}(t)$ in (4.47) have positive frequency components which extend beyond $2\omega_c \pm \omega_c/2$. Consequently, a different baseband-equivalent model is used for intermediate-frequency mixers in heterodyne architectures, such as the image-reject architecture. In this case, $y_i(t)$ in (4.36) is still represented by (4.37), but $y_{LO}(t)$ in (4.36) is instead represented by the following expression

$$\begin{aligned} y_{LO}(t) &= y_{LODC}(t) + \sum_{n=1}^{N} \{y_{LOIn}(t)\cos[n(\omega_c - \omega_{IF})t] + y_{LOQn}(t)\sin[n(\omega_c - \omega_{IF})t]\} \\ &= y_{LODC}(t) + \sum_{n=1}^{N} \{[y_{LOIn}(t)\cos(n\omega_{IF}t) - y_{LOQn}(t)\sin(n\omega_{IF}t)]\cos(n\omega_c t) + \\ &\qquad [y_{LOIn}(t)\sin(n\omega_{IF}t) + y_{LOQn}(t)\cos(n\omega_{IF}t)]\cos(n\omega_c t)\}. \end{aligned} \tag{4.48}$$

Equation (4.48) differs from (4.38) in that the dependence on the intermediate frequency, ω_{IF}, is not included in the time-varying coefficients. Similarly, $y_o(t)$ in (4.36) is represented by

$$\begin{aligned} y_o(t) = y_{oDC}(t) &+ \sum_{n=1}^{N} [y_{oIIFn}(t)\cos(n\omega_{IF}t) + y_{oQIFn}(t)\sin(n\omega_{IF}t)] + \\ &\sum_{n=1}^{N} [y_{oIRFn}(t)\cos(n\omega_c t) + y_{oQRFn}(t)\sin(n\omega_c t)] + \\ &\sum_{n=1}^{N}\sum_{m=1}^{N} \{y_{oIPnm}(t)\cos[(n\omega_c + m\omega_{IF})t] + y_{oQPnm}(t)\sin[(n\omega_c + m\omega_{IF})t]\} + \\ &\sum_{n=1}^{N}\sum_{m=1}^{N} \{y_{oINnm}(t)\cos[(n\omega_c - m\omega_{IF})t] + y_{oQNnm}(t)\sin[(n\omega_c - m\omega_{IF})t]\}. \end{aligned} \tag{4.49}$$

The baseband-equivalent model is derived by solving for the output coefficients of $y_o(t)$ in terms of the input coefficients of $y_i(t)$ and $y_{LO}(t)$. Although this model seems quite complex, reductions in simulation complexity can be achieved by keeping track of only the relevant output coefficients. In heterodyne receivers, the components far from ω_{IF} are

significantly attenuated by filtering, and consequently, it is only necessary to keep track of the components near ω_{IF}. The results for $N = 3$ are given in Appendix A.

4.3.4 Summary

By letting the input and output signals of the various RF building blocks take the same form as (4.29) for direct-conversion and low-IF receivers or (4.49) for heterodyne receivers and pre-computing the relationship between the time-varying coefficients, the dependence on the carrier frequency is eliminated. By specifying appropriate functions for the time-varying coefficients, these baseband-equivalent models account for many circuit impairments in the RF components, including distortion, phase noise, quadrature phase mismatch, frequency offset, and dc offsets. Moreover, the maximum step size of the simulation is now determined by the maximum frequency component in the time-varying coefficients rather than the carrier frequency. The additional simulation complexity of this technique depends on the number of harmonics, N, which must be chosen in order to accurately model the effects of analog circuit impairments such as distortion. For typical wireless communications systems, the number of harmonics should be chosen to be at least three.

4.4 Simulation Framework Implementation

A simulation framework which allows for detailed, yet rapid, simulations of the analog front-end is implemented in Simulink, a graphical simulation environment built on top of MATLAB. Implementing the simulation framework in Simulink offers compatibility of the analog front-end simulations with MATLAB, which is already widely used for development and evaluation of communications algorithms. Consequently, using this simulation framework allows for complete end-to-end simulations of communications systems, including all analog, mixed-signal, and digital components. Such a framework facilitates the exploration of analog and digital design tradeoffs, leading to solutions which can potentially achieve lower power consumption and higher levels of integration.

Since behavioral models offer fast simulation speed, the simulation environment described here includes behavioral models for all of the analog front-end components.

The behavioral models for all high-frequency components, such as RF amplifiers, mixers, and oscillators, are based on the baseband-equivalent models described in Section 4.3. In addition, structural models are also available for some of the more complex receiver components, such as phase-locked loops and sigma-delta analog-to-digital converters. Although structural models require longer simulation times, they offer increased accuracy when used in system-level simulations. Finally, the simulation framework described here also supports the conventional design approach in which the performance requirements of the analog front-end components are determined by the numerical calculations described in Section 4.2. However, instead of using spreadsheets to facilitate such calculations, equations such as (4.7) and (4.28) are directly incorporated into the models of the various receiver building blocks. Consequently, the simulation framework can provide overall receiver performance metrics, such as noise figure and input IP_3, based on the specifications of the individual cascaded components. The following sections provide additional detail about various aspects of this system-level simulation framework.

4.4.1 Thermal Noise

Thermal noise introduced by the various components in the analog front-end is modeled by the Band-Limited White Noise block in Simulink. A few precautions should be exercised when specifying the simulation parameters for this block. First, the noise power parameter is used to specify the desired double-sided power spectral density (PSD) of the white noise. The double-sided PSD is half of the single-sided PSD, which is more commonly used in circuit analysis. For example, the single-sided PSD of the voltage noise introduced by a resistor is

$$\frac{\overline{v^2}}{\Delta f} = 4kTR\,. \tag{4.50}$$

When using the Band-Limited White Noise block to model the noise introduced by a resistor, the noise power parameter should be set to the double-sided PSD, $2kTR$.

Next, the sample time parameter is used to specify the correlation time of the noise. The inverse of this parameter is just the noise bandwidth. Ideally, white noise should have infinite bandwidth or a correlation time of zero. In order to approximate the wideband

nature of the noise, the noise bandwidth should be much greater than the largest frequency component in the system. In other words, the sample time parameter should be chosen to be much smaller than the shortest time constant in the system. The following condition on the sample time, t_c, is recommended for good results [28]:

$$t_c \approx \frac{1}{100}\frac{2\pi}{f_{\max}}. \tag{4.51}$$

Clearly, this constraint can increase the simulation time significantly. In fact, the speed of the simulation framework described here is limited by the ability to accurately model the behavior of broadband white noise. One way to reduce the simulation time is simply to relax the noise sampling time constraint. Simulations were performed to evaluate the effect of noise sampling time on accuracy by feeding broadband white noise with a single-sided PSD of 1 V^2/Hz through a third-order Butterworth low-pass filter, the cutoff frequency of which determines the smallest time constant of the system. The magnitude response of a third-order Butterworth filter is

$$|H(f)|^2 = \frac{1}{1+(f/f_c)^6} \tag{4.52}$$

and the variance of the filtered output noise is

$$P = \int_{-\infty}^{\infty} \frac{df}{1+(f/f_c)^6} = \frac{2}{3}\pi f_c. \tag{4.53}$$

These simulation results are summarized in Table 4.1. As expected, the simulated output noise variance is very accurate when the noise bandwidth is much larger than the filter bandwidth. However, the results are still fairly accurate for a noise bandwidth which is only five to ten times larger than the filter bandwidth. Thus, if simulation speed is critical, then the constraint on the noise sampling time in (4.51) can be relaxed, but should be no more than about one-fifth the shortest time constant in the system.

Finally, the seed parameter in the Band-Limited White Noise block sets the starting seed for the random number generator. Different seeds should be used for different noise sources in order to keep the various noise sources uncorrelated. Setting the seed

noise sampling time	noise bandwidth	filter bandwidth	expected output noise variance	simulated output noise variance
1 ns	1 GHz	1 MHz	63.2 dB	63.2 dB
1 ns	1 GHz	10 MHz	73.2 dB	73.2 dB
1 ns	1 GHz	20 MHz	76.2 dB	76.2 dB
1 ns	1 GHz	30 MHz	78.0 dB	78.0 dB
1 ns	1 GHz	40 MHz	79.2 dB	79.2 dB
1 ns	1 GHz	50 MHz	80.2 dB	80.2 dB
10 ns	100 MHz	1 MHz	63.2 dB	63.2 dB
10 ns	100 MHz	10 MHz	73.2 dB	73.0 dB
10 ns	100 MHz	20 MHz	76.2 dB	75.9 dB
10 ns	100 MHz	30 MHz	78.0 dB	77.3 dB
10 ns	100 MHz	40 MHz	79.2 dB	78.2 dB
10 ns	100 MHz	50 MHz	80.2 dB	78.7 dB

Table 4.1: Effect of noise sampling time on simulation accuracy.

parameter to a random number guarantees that all the noise sources in the same simulation are independent.

4.4.2 Flicker Noise

The power spectral density of the input-referred voltage noise for a MOS transistor consists of a flicker noise component and a thermal noise component and is given by (2.6) and repeated here

$$\frac{\overline{v_i^2}}{\Delta f} = \frac{K_f}{WLC_{ox}f} + 4kT\frac{2}{3}\frac{1}{g_m}. \tag{4.54}$$

Flicker noise is particularly important in the direct-conversion and Weaver architectures since in both cases, the desired signal is frequency translated to baseband prior to analog-to-digital conversion. In addition, it can also be a dominant contributor to phase noise in the LO. In order to examine the effects of flicker noise on overall system performance, a flicker noise model is included in the simulation framework. Flicker noise is generated by passing white noise through a filter with magnitude response [29]

$$|H(f)| = \frac{1}{\sqrt{f}}. \tag{4.55}$$

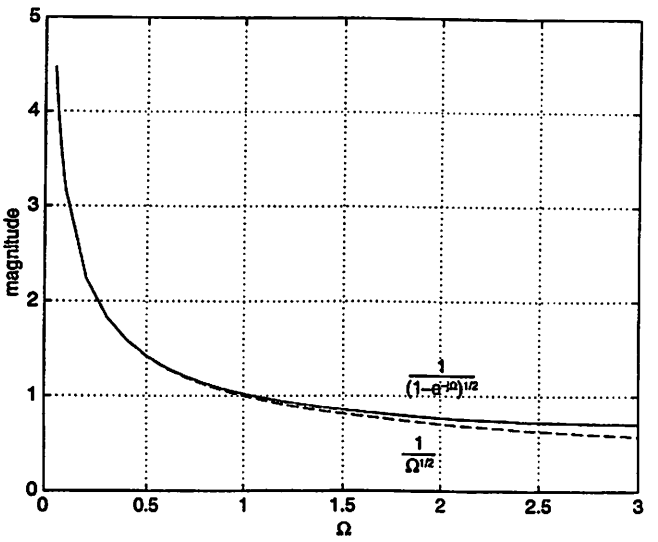


Figure 4.7: Magnitude response of $H(e^{j\Omega})$ in (4.57).

Such a filter can be approximated by the discrete-time transfer function [30]

$$H(z) = \frac{1}{(1-z^{-1})^{1/2}}. \tag{4.56}$$

The frequency response of $H(z)$ is given by

$$H(e^{j\Omega}) = \frac{1}{(1-e^{-j\Omega})^{1/2}} \tag{4.57}$$

and the magnitude of $H(e^{j\Omega})$ is plotted in Fig. 4.7 and matches well with the magnitude response of $1/\Omega^{1/2}$. The denominator in (4.56) can be expanded in a power series, giving

$$H(z) = \frac{1}{1-\frac{1}{2}z^{-1}-\frac{1}{2}\left(1-\frac{1}{2}\right)\frac{1}{2!}z^{-2}-\ldots}. \tag{4.58}$$

The coefficients of the power series in the denominator of (4.58) are given by the following iterative equations:

$$\begin{aligned} a_1 &= 1 \\ a_k &= \left(k-\frac{5}{2}\right)\frac{a_{k-1}}{k-1}. \end{aligned} \tag{4.59}$$

```
a = zeros(1,num_taps);
a(1) = 1;
for i = 2:num_taps
   a(i) = (i–2.5)*a(i–1)/(i–1);
end
```

Figure 4.8: Filter initialization code corresponding to (4.59).

In Simulink, this filter is implemented using the Direct-Form II Transpose Filter block and the denominator coefficients are determined by the initialization code illustrated in Fig. 4.8. The amount of flicker noise is determined by specifying the power spectral density, PSD_a, of the noise at the flicker noise corner frequency, f_a, which is the frequency at which the flicker noise and thermal noise asymptotes intersect. The Band-Limited White Noise block is then used to generate white noise with power spectral density equal to $2\pi f_a PSD_a$. The output of the Band-Limited White Noise block is then passed through the Direct-Form II Transpose Filter block, and the resulting power spectral density is

$$PSD = 2\pi f_a PSD_a \times \frac{1}{2\pi f} = PSD_a \frac{f_a}{f}. \tag{4.60}$$

The accuracy of the filter can be increased by increasing the number of terms in the power series expansion of the denominator in (4.58). Of course, this increased accuracy comes at the cost of increased simulation time. The power spectral density corresponding to a filter with 1000 taps is illustrated in Fig. 4.9. As seen in Fig. 4.9, the simulated power

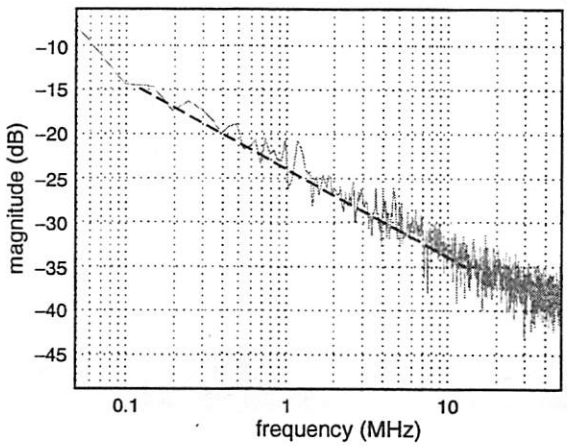


Figure 4.9: Simulated flicker noise power spectral density.

spectral density matches well with the expected $1/f$ behavior.

4.4.3 RF Amplifiers

The baseband-equivalent model for RF amplifiers described in Section 4.3.1 is implemented in Simulink using S-functions. S-functions may be written in either the MATLAB or C programming languages. For this simulation framework, all the S-functions are written in C in order to maintain compatibility with the Real-Time Workshop. Compiling Simulink designs using the Real-Time Workshop offers simulation speed improvements of up to ten times. Unfortunately, the Real-Time Workshop cannot be used to compile Simulink designs which contain S-functions written in the MATLAB programming language.

The transfer function of an RF gain block is described by (4.31), and in this simulation framework, the baseband-equivalent model is implemented for $N = 3$. The behavior of each RF gain block is specified by the coefficients of the equation

$$x_o(t) = a_0 + a_1 x_i(t) + a_2 x_i^2(t) + a_3 x_i^3(t) \tag{4.61}$$

as well as by the block's noise performance. Since $N = 3$, the input and output of an RF gain block each consists of seven signals representing the time-varying coefficients of the equation

$$s(t) = s_{DC}(t) + \sum_{n=1}^{3} [s_{In}(t)\cos(2\pi n f_c t) + s_{Qn}(t)\sin(2\pi n f_c t)]. \tag{4.62}$$

In order to model broadband white noise, noise sources with the appropriate power spectral densities must be added to each of the seven input signals. In Fig. 4.10, the total double-sided noise PSD is partitioned into multiple non-overlapping components. The

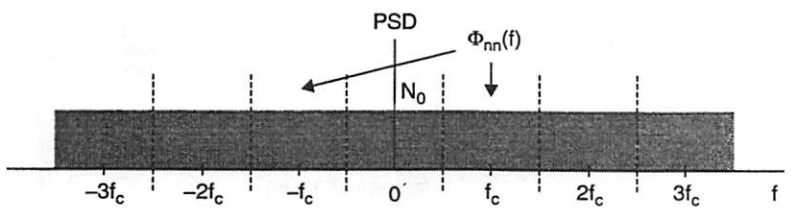


Figure 4.10: Partitioning of broadband white noise.

component from $-f_c/2 < f < f_c/2$ has a double-sided PSD equal to N_0. The components centered at $-f_c$ and f_c can be represented by the following expression:

$$n(t) = x(t)\cos(2\pi f_c t) - y(t)\sin(2\pi f_c t) \tag{4.63}$$

where $n(t)$ is a wide-sense stationary stochastic process with zero mean and PSD $\Phi_{nn}(f)$, and $x(t)$ and $y(t)$ are the I and Q components of $n(t)$, respectively, both of which are band-limited from $-f_c/2 < f < f_c/2$. Since $n(t)$ is zero mean, $x(t)$ and $y(t)$ are also zero mean. In addition, since $n(t)$ is stationary, the autocorrelation and cross-correlation functions of $x(t)$ and $y(t)$ satisfy the following properties:

$$\phi_{xx}(\tau) = \phi_{yy}(\tau) \tag{4.64}$$

$$\phi_{xy}(\tau) = -\phi_{yx}(\tau). \tag{4.65}$$

Moreover, since the PSD for band-pass white noise is symmetric about $f = 0$, $\phi_{yx}(\tau) = 0$. Consequently, the autocorrelation function of $n(t)$ is

$$\phi_{nn}(\tau) = \phi_{xx}(\tau)\cos(2\pi f_c \tau) \tag{4.66}$$

and the PSD of $n(t)$ is

$$\Phi_{nn}(f) = \frac{1}{2}[\Phi_{xx}(f - f_c) + \Phi_{xx}(f + f_c)]. \tag{4.67}$$

From Fig. 4.10, $\Phi_{nn}(f) = N_0$ for $-3f_c/2 < f < -f_c/2$ and $f_c/2 < f < 3f_c/2$, and therefore, the power spectral densities of $x(t)$ and y(t) are

$$\Phi_{xx}(f) = \Phi_{yy}(f) = 2N_0. \tag{4.68}$$

Consequently, if the noise component centered at dc is modeled by a white noise source with a double-sided PSD equal to N_0, then the I and Q noise components centered at f_c, $2f_c$, and $3f_c$ are each modeled by an independent white noise source with a double-sided PSD equal to $2N_0$.

4.4.4 Mixers

The baseband-equivalent model for direct-conversion and low-IF mixers described in Section 4.3.3 is implemented for $N = 3$ using S-functions written in C. Similar to the model for RF amplifiers, the gain and distortion performance of each mixer is specified by the coefficients in (4.61). In fact, the same baseband-equivalent model used for RF amplifiers is used to model the gain and distortion performance of each mixer. In this implementation, the gain and distortion introduced by the mixer is assumed to occur before the frequency translation process. Finally, similar to the RF amplifier blocks, the broadband white noise introduced by the mixer is modeled by adding noise sources with the appropriate power spectral densities to each of the seven input signals. If the noise component centered at dc is modeled by a white noise source with a double-sided PSD equal to N_0, then the I and Q noise components centered at f_c, $2f_c$, and $3f_c$ are each modeled by an independent white noise source with a double-sided PSD equal to $2N_0$.

4.4.5 Local Oscillators

Many receivers use in-phase and quadrature oscillator signals for frequency translation:

$$y_{LOI} = A_I \cos\left[2\pi(f_c + \Delta f)t + \frac{\phi_0}{2} + \phi_n(t)\right] \tag{4.69}$$

$$y_{LOQ} = A_Q \sin\left[2\pi(f_c + \Delta f)t - \frac{\phi_0}{2} + \phi_n(t)\right] \tag{4.70}$$

where A_I and A_Q are the amplitudes of the in-phase and quadrature oscillator signals, respectively, f_c is the oscillation frequency, Δf is the frequency offset, ϕ_0 is the quadrature phase mismatch and $\phi_n(t)$ is the phase noise. Equations (4.69) and (4.70) can be expressed in the form of (4.29):

$$\begin{aligned} y_{LOI} = {} & A_I \cos\left[2\pi\Delta f t + \frac{\phi_0}{2} + \phi_n(t)\right]\cos(2\pi f_c t) \\ & - A_I \sin\left[2\pi\Delta f t + \frac{\phi_0}{2} + \phi_n(t)\right]\sin(2\pi f_c t) \end{aligned} \tag{4.71}$$

$$y_{LOQ} = A_Q \sin\left[2\pi\Delta f t - \frac{\phi_0}{2} + \phi_n(t)\right]\cos(2\pi f_c t) + A_Q \cos\left[2\pi\Delta f t - \frac{\phi_0}{2} + \phi_n(t)\right]\sin(2\pi f_c t). \quad (4.72)$$

This baseband-equivalent model is implemented using built-in Simulink blocks [31].

The phase noise performance of the oscillator can be represented by either a simple behavioral model or a more complex structural model [31]. In the simple behavioral model, at small frequency offsets from the center frequency, the power spectral density of the phase noise is assumed to be constant, while at large frequency offsets, it is assumed to fall off at f^{-2}. This phase noise profile is generated by passing white noise through a first-order low-pass filter. An example spectrum is illustrated in Fig. 4.11. In this case, the phase noise is specified to be –70 dBc/Hz at low frequencies and –100 dBc/Hz at a 100-kHz frequency offset. This behavioral model provides a good first-order approximation of the phase noise performance, while being simple enough to provide fast simulation times.

A more complex phase noise model is also available is this simulation framework. In many receivers, the local oscillator signal is generated from a very accurate reference frequency using a phase-locked loop (PLL) as illustrated in Fig. 4.12. The phase detector (PD) compares the phases the two input signals, and its output passes through a loop filter (LF) before being applied to the voltage-controlled oscillator (VCO). A divider reduces

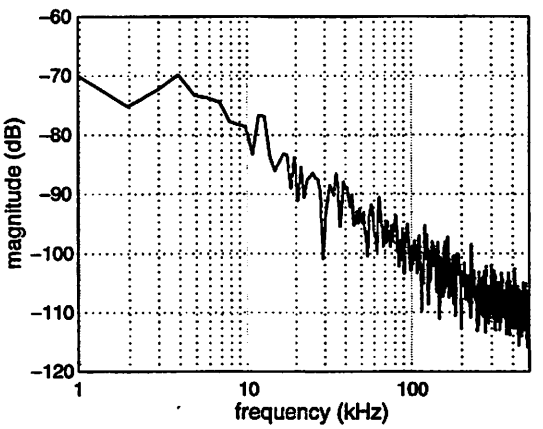


Figure 4.11: Example spectrum of phase noise generated by simple behavioral model.

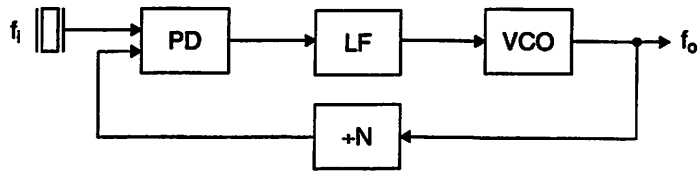


Figure 4.12: PLL block diagram.

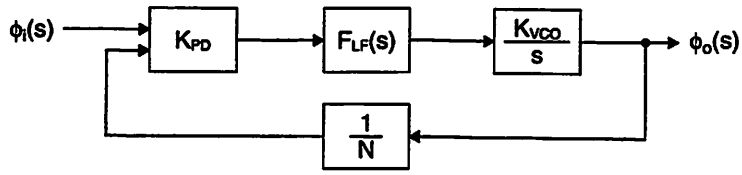


Figure 4.13: Linear model of PLL.

the frequency of the oscillator signal by a factor N and this signal is fed back to the PD to complete the loop. When the loop is locked, $f_o = Nf_i$. Under this condition, the PLL can be modeled as a linear system (Fig. 4.13). From this linear model, the noise contribution from each of the PLL components can be determined. Fig. 4.14 illustrates the linear model along with all relevant noise sources. The noise source ϕ_{n1} models the noise contributions from the PD, reference oscillator, and divider, while ϕ_{n2} models the noise from the LF and ϕ_{n3} models the noise from the VCO. The total noise at the PLL output is

$$\phi_o = \phi_{o1} + \phi_{o2} + \phi_{o3} \tag{4.73}$$

where

$$\phi_{o1} = \frac{K_{PD}K_{VCO}\dfrac{F_{LF}(s)}{s}}{1+\dfrac{K_{PD}K_{VCO}}{N}\dfrac{F_{LF}(s)}{s}}\phi_{n1} \tag{4.74}$$

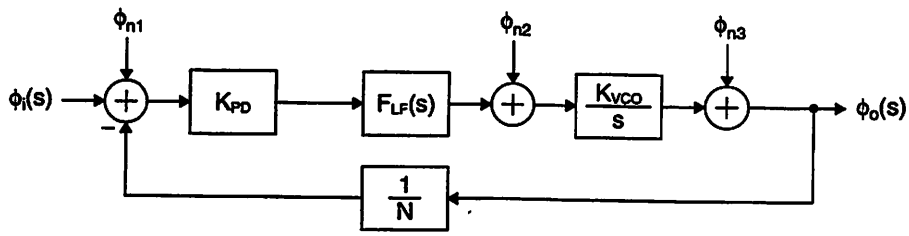


Figure 4.14: Linear model of PLL with noise sources.

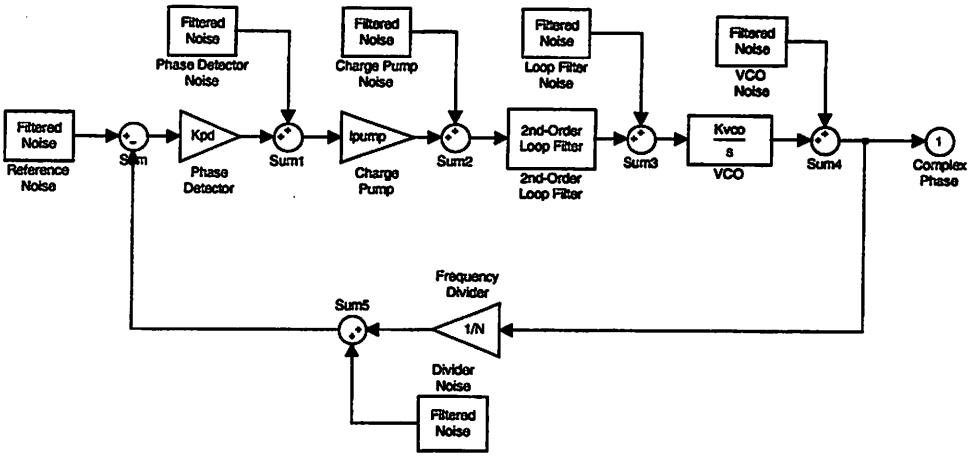


Figure 4.15: Simulink implementation of complex phase noise model.

$$\phi_{o2} = \frac{K_{VCO}\dfrac{1}{s}}{1+\dfrac{K_{PD}K_{VCO}}{N}\dfrac{F_{LF}(s)}{s}}\phi_{n2} \tag{4.75}$$

$$\phi_{o3} = \frac{1}{1+\dfrac{K_{PD}K_{VCO}}{N}\dfrac{F_{LF}(s)}{s}}\phi_{n3}. \tag{4.76}$$

This phase noise model is implemented for a PLL with a second-order loop filter using built-in Simulink blocks and is illustrated in Fig. 4.15 [31]. In this model, the phase noise performance is specified by PLL parameters such as PD gain, VCO gain, and divider ratio N, as well as by circuit parameters such as the resistor and capacitor values used to implement the LF. Because of its accuracy, this model is very useful as an aid in exploring the tradeoffs between various PLL parameters. Finally, although this model provides a very accurate representation of the PLL phase noise, its complexity results in longer simulation times. Consequently, this phase noise model should be avoided in system-level simulations unless very accurate results are required.

4.4.6 Baseband Amplifiers and Filters

Simulink provides a comprehensive library of blocks which can be used to model baseband amplifiers and filters. In particular, the Transfer Fcn block can be used to

implement any linear transfer function by specifying the coefficients of the numerator and denominator polynomials:

$$H(s) = \frac{a_1 s^{n-1} + a_2 s^{n-2} + a_3 s^{n-3} + \ldots + a_n}{b_1 s^{m-1} + b_2 s^{m-2} + b_3 s^{m-3} + \ldots + b_m}. \tag{4.77}$$

Alternative, the Zero-Pole block can be used to implement any linear transfer function by specifying the poles and zeros:

$$H(s) = K \frac{(s - z_1)(s - z_2)(s - z_3)\ldots(s - z_n)}{(s - p_1)(s - p_2)(s - p_3)\ldots(s - p_n)}. \tag{4.78}$$

Finally, Simulink also provides an Analog Filter Design block in the Filter Designs library of the DSP Blockset which can be used to model Butterworth, Chebyshev Type I, Chebyshev Type II, and Elliptic filter responses. Filters can be specified as low-pass, high-pass, band-pass, or band-stop. For a Bessel response, the simulation framework provides a custom Simulink block based on the *besself* MATLAB function. The numerator and denominator coefficients of an N^{th}-order Bessel filter with cutoff frequency ω are determined by the following MATLAB command:

$$[num, den] = besself(N, \omega). \tag{4.79}$$

Once the filter coefficients are determined, the Bessel filter response is implemented using the Transfer Fnc block.

4.4.7 Analog-to-Digital Converters

A simple ADC behavioral model implemented using Simulink blocks is illustrated in Fig. 4.16. In this model, the analog input signal is sampled at the ADC sampling rate by the Zero-Order Hold block and the sampled value is quantized to one of 2^R amplitude levels by the Quantizer block, where R is the ADC resolution. In addition, the Saturation

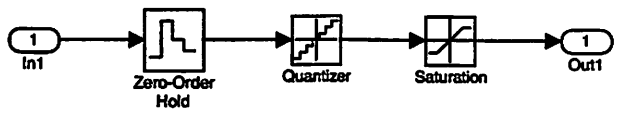


Figure 4.16: Simple ADC behavioral model.

block models the clipping which can occur when the input signal amplitude exceeds the full-scale voltage of the ADC. This simple behavioral model results in rapid simulation times and is adequate for the preliminary evaluation of performance degradation resulting from ADC impairments.

A more accurate evaluation of the effects of ADC impairments requires a more complex structural model. Since there are many types of ADCs, developing a comprehensive library of ADC structural models is impractical. However, despite requiring different structural models, the modeling issues for the different types of ADCs are similar since they are based on similar building blocks, such as switched-capacitor circuits and comparators. The structural model of a first-order sigma-delta (ΣΔ) converter is described below [32], [33] as an example and similar techniques can be used to develop structural models for other types of ADCs.

The block diagram of a first-order ΣΔ converter is illustrated in Fig. 4.17. The integrator is implemented as a single-ended switched-capacitor circuit illustrated in Fig. 4.18, and when the gain of the operational transconductance amplifier (OTA) is large, the transfer function is given by

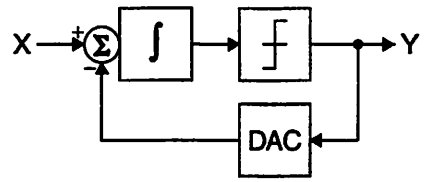


Figure 4.17: First-order ΣΔ converter.

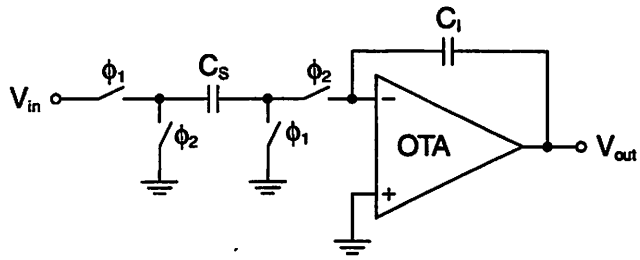


Figure 4.18: Switched-capacitor integrator.

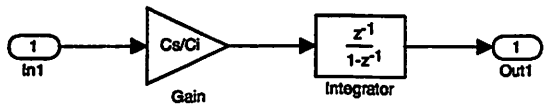


Figure 4.19: Model of ideal integrator.

$$\frac{V_{out}}{V_{in}} = H(z) = \frac{C_S}{C_I}\frac{z^{-1}}{1-z^{-1}}. \tag{4.80}$$

In this ideal case, the integrator can be modeled in Simulink using the Gain and Discrete Filter blocks as illustrated in Fig. 4.19.

In reality, the gain the of the OTA is limited. In this case, the OTA transfer function is given by

$$\frac{V_{out}}{V_{in}} = H(z) = \frac{C_S}{C_I}\frac{\beta z^{-1}}{1-(1-\varepsilon)z^{-1}} \tag{4.81}$$

where

$$\beta = \frac{1}{1+\frac{C_S + C_I}{AC_I}} \tag{4.82}$$

$$1-\varepsilon = \frac{1}{1+\frac{C_S}{AC_I}} \tag{4.83}$$

and A is OTA gain. Consequently, finite OTA gain reduces the forward gain of the integrator by β in addition to shifting the pole of $H(z)$ inside the unit circle. An integrator with finite OTA gain is modeled in Simulink as illustrated in Fig. 4.20.

In a practical implementation, integrator noise also degrades the performance of the ΣΔ converter. These noise sources include thermal noise from the sampling switches as well

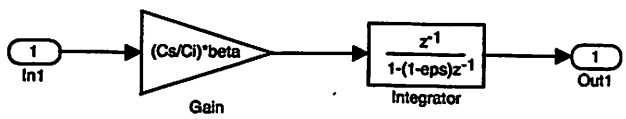


Figure 4.20: Model of integrator with finite OTA gain.

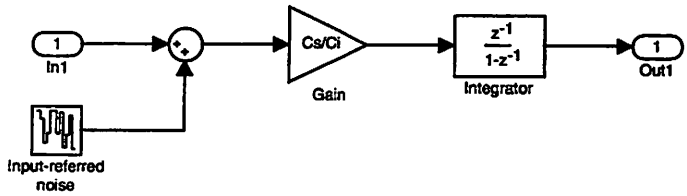


Figure 4.21: Model of integrator with noise.

as thermal noise and flicker noise from the OTA. An integrator with equivalent input-referred thermal and flicker noise sources is modeled in Simulink as illustrated in Fig. 4.21.

The performance of the ΣΔ converter is also affected by the nonlinear transfer function of the integrator. In this case, the transient response of the switched-capacitor integrator is illustrated in Fig. 4.22. The nonlinear transfer function can be expressed as [34]

$$v_o(v_i)=\begin{cases} -Kv_i+(K+1)v_i(1-e^{-n_\tau}), & |(K+1)v_i|\le\dfrac{SRN}{n_\tau} \\ v_i-\mathrm{sgn}(v_i)\dfrac{SRN}{n_\tau}e^{\left(\frac{|(K+1)v_i|n_\tau}{SRN}-n_\tau-1\right)}, & \dfrac{SRN}{n_\tau}<|(K+1)v_i|\le SRN\left(\dfrac{1}{n_\tau}-1\right) \\ -Kv_i+\mathrm{sgn}(v_i)SRN, & SRN\left(\dfrac{1}{n_\tau}-1\right)<|(K+1)v_i| \end{cases} \quad (4.84)$$

where K accounts for the charge feed-through, n_τ is the number of time constants occurring during the settling period, and SRN is the normalized slew rate. The nonlinear transfer function of the integrator is modeled in Simulink using the Look-Up Table block (Fig. 4.23).

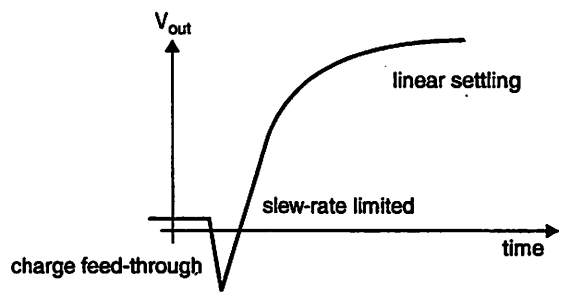


Figure 4.22: Switched-capacitor integrator transient response.

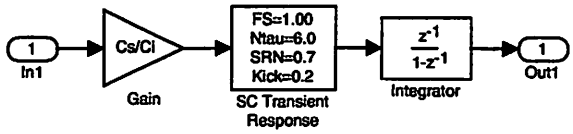


Figure 4.23: Model of integrator with a nonlinear transfer function.

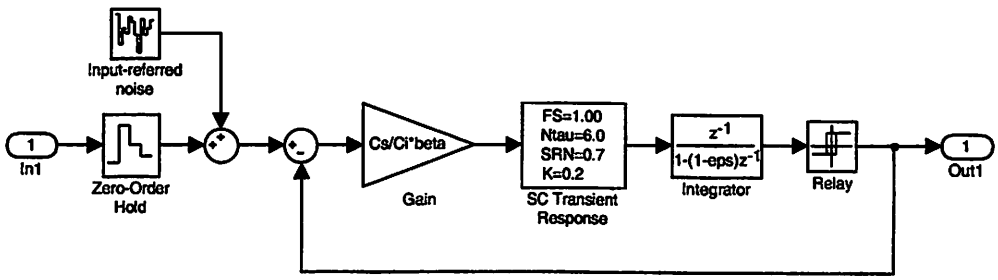


Figure 4.24: Structural model of first-order ΣΔ converter.

Finally, the performance of the ΣΔ converter is also affect by comparator offset or hysteresis, which is modeled by the Relay block in Simulink. The complete structural model for the first-order sigma delta converter is illustrated in Fig. 4.24.

4.4.8 Receiver Performance Metrics

Finally, the simulation framework provides overall receiver performance metrics for the analog front-end, including total gain, noise figure, input IP_2, and input IP_3, based on the specifications of the individual cascaded components. The noise figure and input IP_3 are calculated based on (4.7) and (4.28), respectively, while the input IP_2 is calculated from

$$\frac{1}{P_{IIP2}} = \frac{1}{P_{IIP2(1)}} + \frac{G_1}{P_{IIP2(2)}} + \ldots + \frac{\prod_{i=1}^{N-1} G_i}{P_{IIP2(N)}} \tag{4.85}$$

where $P_{IIP2(i)}$ is the input power at the second-order intermodulation intercept point of the i^{th} stage and G_i is the power gain of the i^{th} stage.

4.5 Summary

The simulation framework presented in this chapter facilitates the exploration of tradeoffs between analog front-end impairments and system-level performance. This framework is implemented in Simulink, a graphical simulation environment built on top of MATLAB.

The use of baseband-equivalent models for all RF building blocks in addition to compiling Simulink designs using the Real-Time Workshop result in fast end-to-end simulations of entire receiver systems.

Chapter 5

A High-Speed Wireless Downlink

5.1 Introduction

Next-generation wireless systems will have to support new applications which require higher and higher data rates. These applications include wireless internet as well as other multimedia applications including full-motion video and real-time high-fidelity audio. This chapter describes a wideband code-division multiple access (WCDMA) system which is designed to operate in an indoor picocellular environment [22], [35]. Within each cell, a single base station supports aggregate data rate of 50 Mb/s. The system bandwidth is 32.5 MHz and operates at a carrier frequency of 2 GHz.

5.2 Base-Station Transmitter

A block diagram of the digital baseband section of the base-station transmitter is illustrated in Fig. 5.1. Two different systems are proposed. In the first case, each cell supports up to 15 binary data channels, each with a data rate of up to 3.33 Mb/s [22]. In the second case, each cell supports up to 31 binary data channels, each with a data rate of up to 1.61 Mb/s [36]. In both cases, however, the aggregate data rate is 50 Mb/s and the total transmission bandwidth is 32.5 MHz. Each of the K binary data channels is mapped

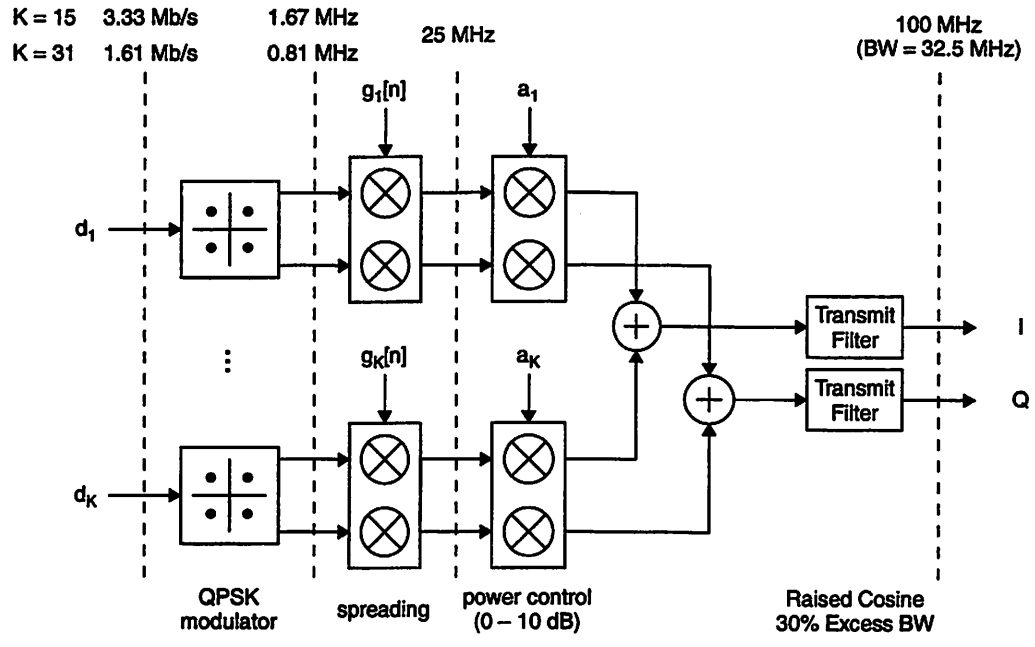


Figure 5.1: Block diagram of base-station transmitter (digital baseband section).

to a four-point QPSK constellation, as described in Section 3.2. Although higher-order constellations result in increased spectral efficiency, they also require higher SNRs for the same BER. The proposed system achieves a spectral efficiency in excess of 1.5 b/s/Hz. The output from each of the QPSK modulators consists of I and Q data streams each operating at 1.67 MHz for $K = 15$ or at 0.81 MHz for $K = 31$.

5.2.1 Multiple Access Method and Power Control

Code-division multiple access (CDMA) is used as the multiple-access scheme for the proposed system. In CDMA, each of the K data channels is assigned a distinct signature sequence, or a spreading code, and all data channels are transmitted simultaneously using the entire frequency band. Since CDMA is a direct-sequence spread-spectrum signaling technique, this multiple-access strategy is robust against frequency-selective narrowband fading which results from multipath propagation [16]. In addition, the processing gain of spread-spectrum signals provides some immunity against narrowband interference.

Each of the K data channels is assigned a discrete-time signature waveform $g_k[n]$ with unit energy. This spreading code is superimposed on both the I and Q binary antipodal

data streams at the output of the QPSK modulator. For $K = 15$, the spreading codes are based on a 4-bit maximal-length shift-register (MLSR) sequence, e.g.,

$$\{-1,-1,-1,+1,-1,-1,+1,+1,-1,+1,-1,+1,+1,+1,+1\}. \tag{5.1}$$

In this case, 15 distinct codes are formed by using different phase offsets of the MLSR sequence in (5.1). For $K = 31$, the spreading codes are based on a 5-bit MLSR sequence, e.g.,

$$\begin{aligned}&\{+1,+1,+1,+1,+1,-1,-1,+1,+1,-1,+1,-1,-1,+1,-1,\\&-1,-1,-1,+1,-1,+1,-1,+1,+1,+1,-1,+1,+1,-1,-1,-1\}.\end{aligned} \tag{5.2}$$

Similarly, the 31 distinct codes are formed by using different phase offsets of the MSLR sequence in (5.2). For both $K = 15$ and $K = 31$, after spreading, the data rate of each of the I and Q data streams for a particular channel is 25 MHz.

Optional power control levels which are adjustable from 0 – 10 dB are provided for each data channel after spreading. Power control can be used to compensate for signal attenuation due to shadowing as well as to provide variable quality of service (QOS) for different data channels.

All of the I data streams from the K data channels are then combined to form a single I data channel, while all of the Q data streams are combined to form a single Q data channel. These I and Q signals can be expressed as, respectively,

$$s_I[n] = \frac{1}{\sqrt{2}} \sum_{k=1}^{K} a_k b_{Ik} g_k[n] \tag{5.3}$$

$$s_Q[n] = \frac{1}{\sqrt{2}} \sum_{k=1}^{K} a_k b_{Qk} g_k[n] \tag{5.4}$$

where the subscript k denotes the k^{th} data channel, a_k sets the power level, $g_k[n]$ is the spreading code, and b_{Ik} and b_{Qk} are the binary antipodal data streams at the output of the QPSK modulator.

5.2.2 Pulse Shaping

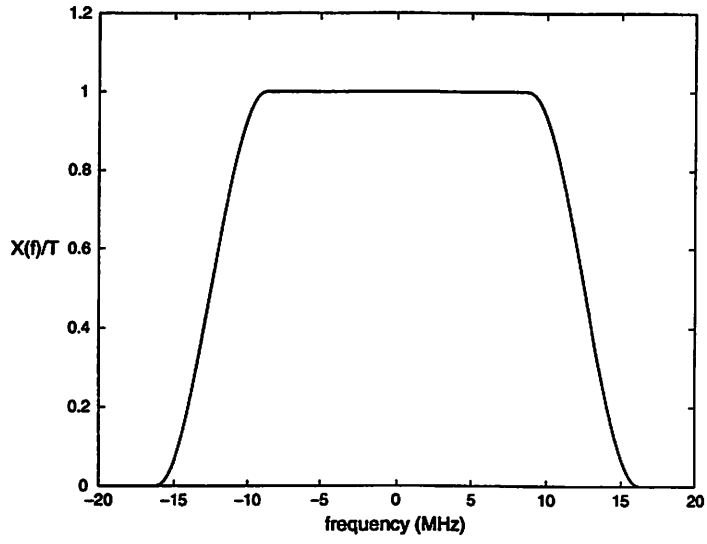


Figure 5.2: Frequency response of raised-cosine pulse ($1/T = 25$ MHz, $\beta = 0.3$).

After all of the I data streams and all of the Q data streams are combined to form a single I channel and a single Q channel, each of these channels is shaped by a raised-cosine filter with an excess bandwidth of 30%. This pulse-shaping filter limits the double-sided bandwidth of each of the I and Q signals to 32.5 MHz. In addition, the raised-cosine filter satisfies the Nyquist condition for zero intersymbol interference (ISI) [16]. A raised-cosine spectrum is described in the time domain as

$$x(t) = \frac{\sin(\pi t/T)}{\pi t/T}\frac{\cos(\pi\beta t/T)}{1-4\beta^2 t^2/T^2} \tag{5.5}$$

where β is the rolloff factor or excess bandwidth. The frequency characteristic of a raised-cosine pulse is given by

$$X(f) = \begin{cases} T, & 0 \le |f| \le \frac{1-\beta}{2T} \\ \frac{T}{2}\left\{1+\cos\left[\frac{\pi T}{\beta}\left(|f|-\frac{1-\beta}{2T}\right)\right]\right\}, & \frac{1-\beta}{2T} \le |f| \le \frac{1+\beta}{2T} \\ 0, & |f| > \frac{1+\beta}{2T} \end{cases}. \tag{5.6}$$

$X(f)$ is plotted for $1/T = 25$ MHz and $\beta = 0.3$ in Fig. 5.2. In the proposed system, the input I and Q data streams are oversampled by a factor of four [18] so the data rate of the outputs from the I and Q raised-cosine filters is 100 MHz [36].

5.2.3 Pilot Channel/Symbol

In order to facilitate receiver functions such as timing recovery, automatic gain control, and channel estimation, the base station transmits either a pilot tone or pilot symbols. In the former case the base station reserves one of the K data channels for transmitting a pilot tone, which is simply one of the K distinct signature waveforms. Although an entire channel is dedicated to transmitting the pilot tone, it is shared by all of the active users in the cell. Since the pilot tone is continuously available to the mobile receivers, this approach is advantageous when the transmission channel is changing relatively quickly. This technique was proposed originally for the $K = 15$ case [22].

In an alternative approach, pilot symbols are periodically inserted along with the data in each of the K data channels. In this approach, the mobile receivers perform functions such as timing recovery and channel estimation during the transmission of the pilot symbols. The information acquired during the training period is then used to recover the actual data which is transmitted during subsequent time slots. The frequency of transmission of the pilot symbols depends on how rapidly the channel changes. Clearly, this approach is advantageous when the transmission channel is changing relatively slowly since fewer pilot symbols are required. This technique was proposed for the $K = 31$ case [36].

5.2.4 Mobility Support: Picocells

The achievable integration level and power consumption of a receiver are directly related to the system sensitivity and selectivity requirements, and consequently, choosing system features which allow for relaxed requirements is critical for achieving a single-chip, low-power receiver implementation. Using a picocellular system architecture provides a large coverage area, and the sensitivity requirement is relaxed by restricting transmission distances within each cell to less than 5 m with an aggregate transmit power of 0 dBm at each base station.

For hexagonal cells, the cellular reuse factor is restricted to [37]

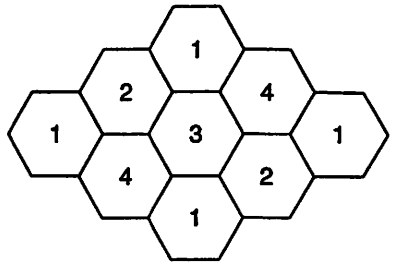


Figure 5.3: Cellular reuse pattern with $N = 4$.

$$N = i^2 + ij + j^2 = 3, 4, 7, 9, 12, \ldots \tag{5.7}$$

where i and j are integers. In the case of $K = 31$, there are six unique 5-bit MLSR sequences [16]. Consequently, a cellular reuse pattern with $N = 4$ (Fig. 5.3) can be used, with each of the four cells transmitting at the same carrier frequency but using a different MSLR code.

In the case of $K = 15$, there are only two unique 4-bit MLSR sequences, so unfortunately, a similar cellular reuse strategy is not possible. In this case, if a cellular architecture based on frequency reuse is employed, then the reuse factor should be small, e.g., $K = 3$, in order to minimize the total system bandwidth.

In both cases, cell site planning should take into account the physical characteristics of the coverage area in order to minimize the amount of interference between cells, resulting in relaxed selectivity requirements. In particular, the proposed system should take advantage of indoor partitions such as floors, ceilings, and walls in order to increase the isolation between adjacent cells. Moreover, the base stations are restricted to a total output power of 0 dBm or 1 mW. This relatively low output power is sufficient to cover a small cell with a 5-m radius while minimizing the amount of interference introduced in neighboring cells.

5.2.5 Analog Front-End

A block diagram of the analog front-end in the base-station transmitter is illustrated in Fig. 5.4 [36]. Two 8-bit 100-MHz digital-to-analog converters (DACs) convert the I and Q digital data streams from the output of the raised-cosine filters to analog signals [38].

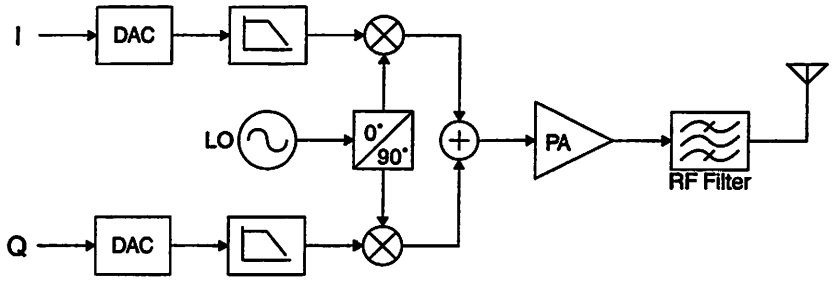


Figure 5.4: Block diagram of base-station transmitter (analog front-end).

The DAC creates unwanted images of the desired signal at multiples of the DAC sampling frequency. The transfer function introduced by the zero-order hold operation performed by the DAC provides some attenuation of the unwanted images:

$$|H(j\omega)| = \frac{\sin\left(\frac{\omega T}{2}\right)}{\frac{\omega}{2}} \tag{5.8}$$

where T is the sampling period of the DAC. A subsequent second-order low-pass Butterworth filter provides additional attenuation of out-of-band spectral energy. The Butterworth response provides a good compromise between maximally flat gain and linear phase response.

Translation of the baseband I and Q signals to the 2-GHz carrier frequency is based on a direct-conversion architecture [39]. The frequency translation is performed using two mixers and an LO fixed at the carrier frequency and operating in quadrature. At the outputs of the two mixers, the I and Q signals are combined before being amplified by a power amplifier (PA). Next, a band-pass filter centered at the carrier frequency further attenuates out-of-band spectral energy before the signal is transmitted through the antenna.

One potential drawback of a direct-conversion transmitter architecture is LO pulling [39]. LO pulling occurs when the output signal from the PA corrupts the LO signal, either by shifting the LO frequency or by deteriorating its spectral purity. LO pulling usually occurs when the output signal from the PA is both close in frequency and comparable in size to the LO signal. Since small form factor is not as critical in the base-station

transmitter, physically separating the LO from the rest of the transmitter is one way of avoiding LO pulling. For example, implementing the LO on a separate chip increases the isolation between the PA and the LO circuitry and helps to prevent the PA output signal from corrupting the LO signal. For a single-chip solution, implementing the LO signal as the product of two lower frequency sinusoids is a second way of avoiding LO pulling. In this case, the PA output signal does not affect the two oscillators used to generate the LO signal which operate at frequencies much lower than the carrier frequency.

Finally, the base stations are restricted to a total output power of 1 mW. Not only does this specification minimize the amount of interference introduced in neighboring cells, but it also significantly relaxes the design requirements of the PA. A linear PA can be used and low power consumption can still be achieved despite the low efficiency of linear amplifiers due to the low output power requirement of 1 mW.

5.3 Mobile Receiver

The design requirements for the mobile receiver are much more stringent than those for the base-station transmitter. The portable nature of the mobile unit imposes strict requirements on the form factor and power consumption of the receiver. The following sections describe the performance requirements and the architecture of the receiver beginning with a brief discussion of wireless propagation models.

5.3.1 Propagation Models

In free space, the radiation power density at the receiver is [40]

$$S_R = \frac{G_T P_T}{4\pi d^2} \tag{5.9}$$

where G_T is the directivity gain of the transmitting antenna, P_T is the transmit power, and d is the distance between the transmitter and the receiver. If the receiving antenna has an aperture A, the received power is

$$P_R = A \times S_R. \tag{5.10}$$

However, the gain G and aperture A of all antennas are related by

$$A = \frac{\lambda^2}{4\pi} G. \tag{5.11}$$

Substituting (5.9) and (5.11) into (5.10) gives

$$P_R = G_R G_T P_T \left(\frac{\lambda}{4\pi d}\right)^2 \tag{5.12}$$

where G_R is the gain of the receiving antenna. Equation (5.12) can be rewritten as

$$P_R\,[\text{dBm}] = P_T\,[\text{dBm}] - L\,[\text{dB}] \tag{5.13}$$

where P_R and P_T are the receive and transmit powers in dBm, respectively, and L is the propagation loss in dB:

$$L\,[\text{dB}] = 32.44 + 20\log f\,[\text{GHz}] + 20\log d\,[\text{m}] - G_R\,[\text{dB}] - G_T\,[\text{dB}]. \tag{5.14}$$

For example, for transmitting and receiving antenna gains of 0 dB, the propagation loss of a 2-GHz signal over 5 m is about 52 dB in free space.

In a real wireless transmission environment, the propagation loss seldom behaves as indicated by (5.12). For indoor wireless propagation, several loss models have been proposed. One model which is commonly used is similar to (5.12) but assumes that the received power is inversely proportional to d^n [41]:

$$P_R = G_R G_T P_T \left(\frac{\lambda}{4\pi}\right)^2 \frac{1}{d^n}. \tag{5.15}$$

In this case the power index n is determined empirically and is typically greater than two, accounting for losses due to walls, ceilings, floors and other objects. A second model which is commonly used assumes the same path-loss model described by (5.12) but includes an additional loss factor α [42], [43]:

$$P_R = \frac{G_R G_T P_T}{\alpha} \left(\frac{\lambda}{4\pi d}\right)^2. \tag{5.16}$$

In this case, the loss factor α, also determined empirically, accounts for the signal attenuation due to shadowing by various objects in an indoor transmission environment.

A third model, which is the one used in this work, is a combination of the first two models and has demonstrated a better fit to experimental measurements [44], [45]:

$$P_R = \frac{G_R G_T P_T}{\alpha}\left(\frac{\lambda}{4\pi}\right)^2 \frac{1}{d^n}. \tag{5.17}$$

Equation (5.17) can be rewritten as

$$P_R\,[\text{dBm}] = P_T\,[\text{dBm}] - L\,[\text{dB}] \tag{5.18}$$

where

$$L\,[\text{dB}] = 10\log\left[\frac{1}{G_R G_T}\left(\frac{4\pi}{\lambda}\right)^2 d^n\right] + \alpha\,[\text{dB}]. \tag{5.19}$$

For the proposed system, it is assumed that $G_R = G_T = 1$, $n = 3$, and that the worst-case shadowing loss, α, is 10 dB.

5.3.2 Receiver Sensitivity

The weakest signal expected to appear at the receiver input determines the sensitivity requirement of the receiver. For the case of $K = 15$ data channels, two different scenarios can potentially result in very weak received signals. In the first scenario, all portable units are located at the edge of the cell and experience the worst case shadowing loss of 10 dB. The total path loss is given by (5.19):

$$L = 10\log\left\{\left[\frac{4\pi(2\times10^9)}{3\times10^8}\right]^2 5^3\right\} + 10 = 69.43\,\text{dB}. \tag{5.20}$$

In this case, the 1-mW transmit power from the base station is equally shared between the 15 data channels, so the transmit power for an individual data channel is

$$P_T = \frac{1}{15}\times 1\,\text{mW} = 0.067\,\text{mW} = -11.76\,\text{dBm}. \tag{5.21}$$

Consequently, for this first scenario, the receiver sensitivity must be better than –81.2 dBm.

In the second scenario, all portable units are located at the edge of the cell and all but one experience the worst case shadowing loss of 10 dB. In order to compensate for the shadowing loss experienced by the 14 portable units, the base-station transmitter increases the power of each of the 14 data channels by 10 dB. Since the total transmit power is limited to 1 mW, the power allocated to the single data channel which does not experience any shadowing decreases correspondingly. For this data channel, the transmit power is ten times less than the power of each of the other 14 data channels. Consequently, the transmit power for the single data channel which does not experience any shadowing is determined by the following equation:

$$\begin{aligned} P_T + 14 \times 10 P_T &= 1\,\text{mW} \\ P_T &= \frac{1}{141}\,\text{mW} \text{ or } -21.5\,\text{dBm}. \end{aligned} \tag{5.22}$$

For this data channel, the total path loss at a distance 5 m away from the base station is 59.43 dB, so the receiver sensitivity must be better than –80.9 dBm. For both scenarios described above, the receiver sensitivity requirements are virtually identical.

The receiver sensitivity requirement for the case of $K = 31$ data channels can be determined in a similar way. When all the portable units are located at the edge of the cell and experience the worst case shadowing loss of 10 dB, the total path loss is still 69.43 dB. However, the 1-mW transmit power from the base station is now equally shared between the 31 data channels, so the transmit power for an individual data channel is

$$P_T = \frac{1}{31} \times 1\,\text{mW} = 0.032\,\text{mW} = -14.91\,\text{dBm}. \tag{5.23}$$

Consequently, the receiver sensitivity must be better than –84.3 dBm.

For $K = 31$, when all portable units are located at the edge of the cell and all but one experience the worst case shadowing loss of 10 dB, the transmit power for the single data channel which does not experience any shadowing is determined by the following equation:

$$P_T + 30 \times 10 P_T = 1\,\text{mW}$$
$$P_T = \frac{1}{301}\,\text{mW or} - 24.8\,\text{dBm}. \tag{5.24}$$

For this data channel, the total path loss at a distance 5 m away from the base station is still 59.43 dB, so the receiver sensitivity must be better than –84.2 dBm. Again, for both scenarios described above, the receiver sensitivity requirements are virtually identical.

5.3.3 Receiver Processing Gain

Since spread-spectrum signals provide processing gain, which enhances the SNR of the received signal after data recovery, CDMA systems may appear to have an advantage over frequency-division multiple access (FDMA) or time-division multiple access (TDMA) systems. Although the SNR of the received signal is indeed enhanced by the processing gain in CDMA systems, these systems in fact do not have increased noise tolerance when compared to FDMA or TDMA systems. This section resolves this common misconception and explains the implications of processing gain on receiver noise performance.

Consider the CDMA and FDMA systems depicted in Fig. 5.5. Both systems support up to K data channels and are limited to a total transmission bandwidth of f_T and a total transmit power of P_T. In the FDMA system, both the total transmission bandwidth and

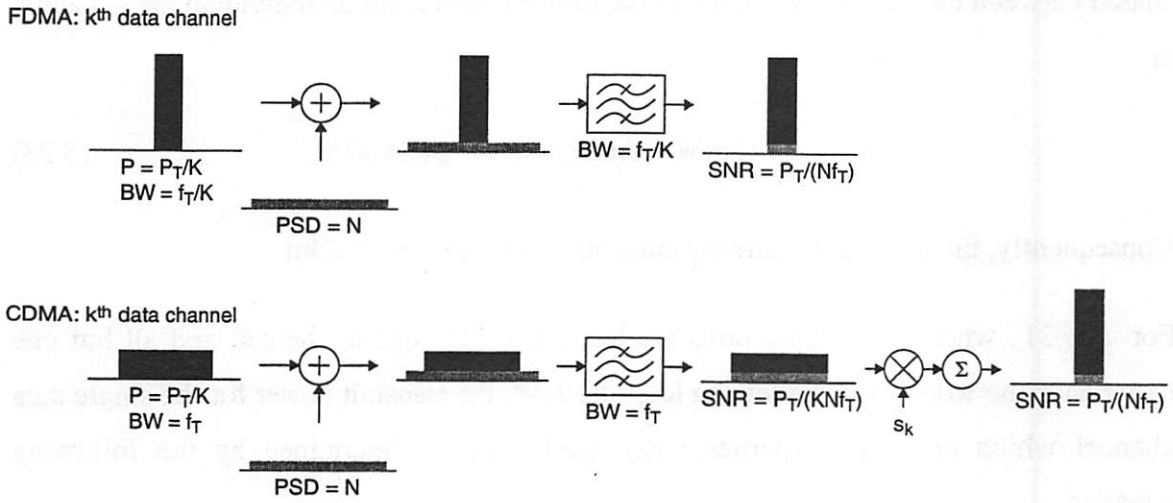


Figure 5.5: Comparison of FDMA and CDMA systems.

the total transmit power are equally divided between the K data channels, so the bandwidth and transmit power of each data channel are f_T / K and P_T / K, respectively. After transmission, assume that the signal is neither amplified nor attenuated but is corrupted by AWGN with a single-sided power spectral density of N. At the receiver, an ideal band-pass filter with bandwidth f_T / K selects the k^{th} data channel. The received signal power is P_T / K while the received noise power is Nf_T / K, and consequently, the SNR for the FDMA system is

$$SNR = \frac{P_T}{K} \times \frac{K}{Nf_T} = \frac{P_T}{Nf_T}. \tag{5.25}$$

In the CDMA system, the total transmit power is equally divided between the K data channels but since the distinct signature sequences are mutually orthogonal, each data channel can transmit over the entire system bandwidth. In this case the bandwidth and transmit power of each data channel are f_T and P_T / K, respectively. As in the FDMA system, assume that the transmitted CDMA signal is neither amplified nor attenuated but is corrupted by AWGN with a single-sided power spectral density of N. At the receiver, an ideal band-pass filter limits the bandwidth of the received signal to f_T. At this point, the signal power of the k^{th} data channel is P_T / K while the received noise power is Nf_T, and consequently, the SNR of the k^{th} data channel is

$$SNR = \frac{P_T}{K} \times \frac{1}{Nf_T} = \frac{P_T}{KNf_T}. \tag{5.26}$$

The k^{th} data channel is selected by multiplying the received signal by the k^{th} signature sequence. After data recovery, the signal power of the k^{th} data channel is still P_T / K but the signal bandwidth is reduced to f_T / K. However, the noise power spectral density is still N since the noise and the k^{th} signature sequence are uncorrelated. Consequently, in the CDMA system, the SNR of the k^{th} data channel after data recovery is

$$SNR = \frac{P_T}{K} \times \frac{K}{Nf_T} = \frac{P_T}{Nf_T}. \tag{5.27}$$

After data recovery, the SNR of the k^{th} data channel increases by the processing gain K. However, as expected, the SNR in the CDMA system is exactly the same as the SNR in the FDMA system, and indeed, the CDMA system does not have an increased noise tolerance. Although the SNR in the CDMA system does increase by the processing gain after data recovery, this merely compensates for the lower SNR at the receiver input resulting from the larger noise bandwidth.

Although the processing gain does not appear to provide the CDMA system any advantage over the FDMA system in the above example, spread-spectrum techniques do provide some advantages. In particular, CDMA signals, which are spread across the entire system bandwidth, are particularly robust against frequency-selective narrowband fading due to multipath propagation.

5.3.4 Receiver Architecture

Due to its simplicity and potential for high integration, the direct-conversion architecture (Fig. 2.7) is the most promising candidate for implementing the receiver for the proposed system [46], [47]. As discussed in Sections 2.3.1 and 2.3.2, receiver implementations based on the direct-conversion architecture must contend with dc offsets and flicker noise. Although the low-IF architecture has the advantage of avoiding problems associated with dc offsets and flicker noise, the wide bandwidth of the desired signal precludes the use of this architecture, since digitizing the IF signal would require a prohibitively fast ADC sampling frequency. However, as already discussed in Section 2.6.1, the wide bandwidth of the desired signal in the proposed system makes the direct-conversion architecture a very attractive approach since on-chip high-pass filtering can be used as a very simple and effective way of eliminating dc offsets and low-frequency flicker noise. The following two sections will review the different techniques for mitigating the problems associated with dc offsets and flicker noise.

5.3.5 Flicker-Noise Suppression

In the direct-conversion architecture, the flicker noise introduced by the baseband circuits as well as by the mixer in some cases can corrupt the potentially weak desired signal.

Since the power spectral density of the input-referred flicker noise of an MOS transistor is inversely proportional to the device dimensions as given by (2.6), flicker noise can be minimized by using large transistor sizes. When circuit speed requirements limit the amount that devices sizes can be increased as well as when the flicker noise performance is not well controlled, which is the case in many CMOS processes, circuit techniques such as autozeroing and chopper stabilization can be used to suppress the flicker noise [48]. The autozero technique is typically implemented using a two-phase clock. During the first phase, the circuit with flicker noise is disconnected from the signal path and its flicker noise is sampled and stored. During the second phase, the circuit is reconnected to the signal path and the stored value of flicker noise is subtracted from the desired signal. In this technique, the previous value of the noise rather than its current value is subtracted from the desired signal, and consequently, this technique is effective only when the noise varies slowly relative to the frequency at which it is sampled. Thus, the autozero technique essentially high-pass filters this noise and is effective at eliminating low-frequency flicker noise but not broadband thermal noise. Since this technique requires that the circuit with flicker noise be disconnected from the signal path, it may not be compatible with continuous-time applications but is well suited for sampled-data applications based on switched-capacitor circuit implementations.

Chopper stabilization is a second technique which can be used to suppress flicker noise. Fig. 5.6 illustrates the chopper stabilization technique applied to an amplifier with flicker noise. In this approach, the input signal is multiplied by a periodic waveform $m_1(t)$ with frequency f_{chop}, which translates the input signal to a higher frequency where flicker noise is negligible. The resulting signal is then amplified at this higher frequency before being translated back down to baseband. This final frequency translation is performed by

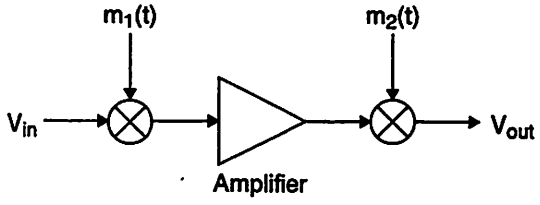


Figure 5.6: Chopper stabilization.

multiplying the amplifier output signal by a second periodic waveform $m_2(t)$ also with frequency f_{chop}.

Although both autozeroing and chopper stabilization are effective in suppressing flicker noise, both techniques also introduce additional complexity in implementing the baseband circuits. Consequently, the proposed implementation relies solely on large transistor sizes to minimize the flicker noise contribution of the baseband circuits. In addition, on-chip high-pass filtering is used to eliminate the low-frequency flicker noise introduced by the downconversion mixers.

5.3.6 DC-Offset Compensation

One very simple and effective way of eliminating dc offsets is through capacitive coupling or high-pass filtering. Since practical single-chip implementations prohibit the use of very large capacitors and resistors, this method is feasible only for systems with large signal bandwidths. However, even if the signal bandwidth is large, this method still results in some BER degradation, since capacitive coupling or high-pass filtering removes low-frequency signal energy along with dc offsets. Fig 5.7 illustrates the effects of high-pass filtering on an otherwise ideal signal constellation for the proposed system with $K = 15$. For simplicity, the transmitted signal consists of a single data channel while

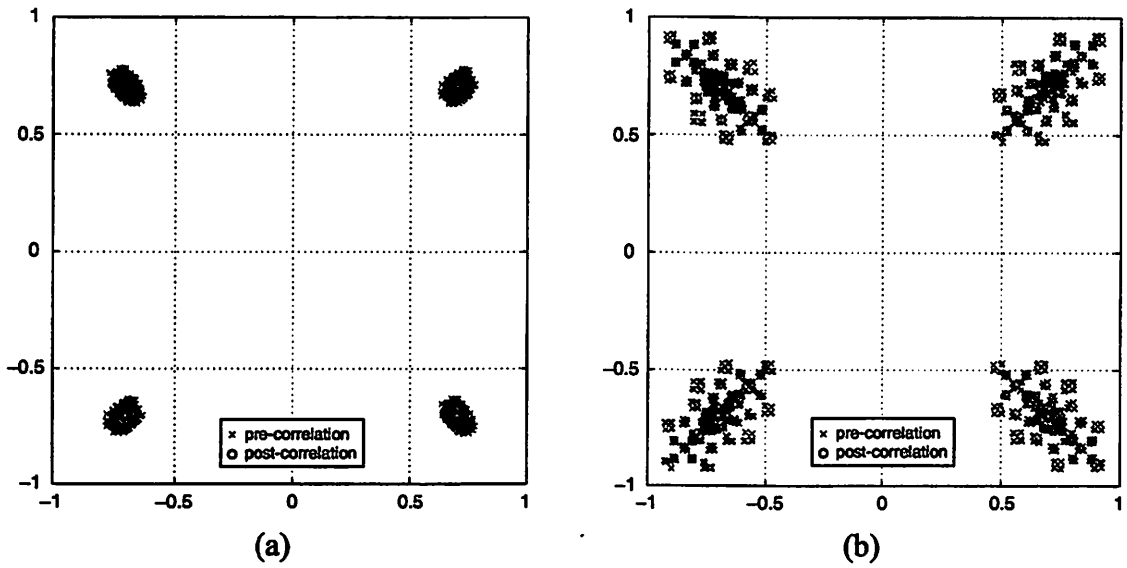


Figure 5.7: Effect of high-pass filtering. (a) 100 kHz. (b) 500 kHz.

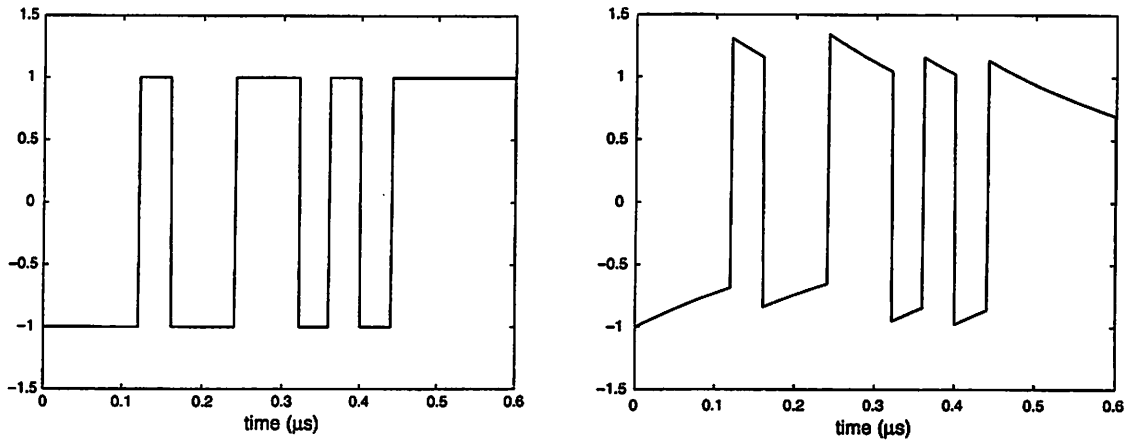


Figure 5.8: (a) $K = 15$ spreading code. (b) After 500-kHz high-pass filter.

data recovery is performed using a single-user correlator (Section 5.3.12). As seen from the signal constellations before data recovery, high-pass filtering results in ISI. In particular, the 500-kHz high-pass filter results in significantly more ISI than the 100-kHz filter. In both cases, however, after data recovery, the signal constellations are close to ideal.

In order to gain a better understanding of the effect of high-pass filtering on the performance of the proposed system, consider the $K = 15$ spreading code in (5.1). For this sequence, the discrete-time autocorrelation function is

$$\phi[m] = \sum_{n=1}^{15} s[n]s[n+m] = \begin{cases} 15, & m = 0 \\ -1, & m \neq 0 \end{cases} \tag{5.28}$$

and the average power, P_{ave}, is $\phi[0] = 15$. Fig. 5.8 illustrates the detrimental effects of filtering this sequence with a 500-kHz high-pass filter. However, correlating this signal with the original spreading sequence results in $P_{ave} \approx 14$, a degradation in signal power of only about 0.27 dB. Consequently, after correlation, the effect of the ISI introduced by the high-pass filter is reduced, as evident in the signal constellations illustrated in Fig. 5.7.

In order to approximate the SNR degradation resulting from high-pass filtering, first consider a received signal with a single-sided bandwidth of f_{sig} as depicted in Fig. 5.9.

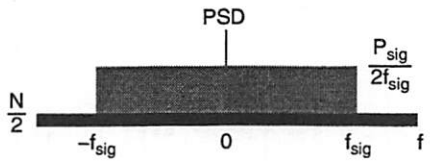


Figure 5.9: Signal and noise power spectral densities.

The signal has a constant PSD of $P_{sig}/2f_{sig}$ over this frequency range, so the total power is P_{sig}. This signal is corrupted by AWGN with PSD $N/2$, and the resulting SNR is given by

$$SNR = \frac{P_{sig}}{Nf_{sig}}. \tag{5.29}$$

After data recover, the signal power and the noise PSD remain unchanged, but the signal bandwidth decreases by the spreading factor K. Thus the resulting SNR is

$$SNR = \frac{KP_{sig}}{Nf_{sig}}. \tag{5.30}$$

Next suppose that a first-order high-pass filter is used to remove the dc offsets from the received signal. The transfer function of such a filter is given by

$$H(f) = \frac{1}{1 - j\frac{f_{HPF}}{f}} \tag{5.31}$$

where f_{HPF} is the corner frequency of the high-pass filter. The high-pass filter removes the low-frequency content of both the desired signal as well as the noise, and consequently, the SNR before and after high-pass filtering are identical. After data recovery, the signal power remains unchanged and is given by

$$P_{HPF} = \int_{-f_{sig}}^{f_{sig}} |H(f)|^2 \frac{P_{sig}}{2f_{sig}} df = P_{sig}\left[1 - \frac{f_{HPF}}{f_{sig}} \arctan\left(\frac{f_{sig}}{f_{HPF}}\right)\right] \tag{5.32}$$

while the signal bandwidth decreases by the spreading factor K. In addition, the correlation process whitens the filtered noise, and after data recovery, the noise PSD is $N/2$. The resulting SNR is

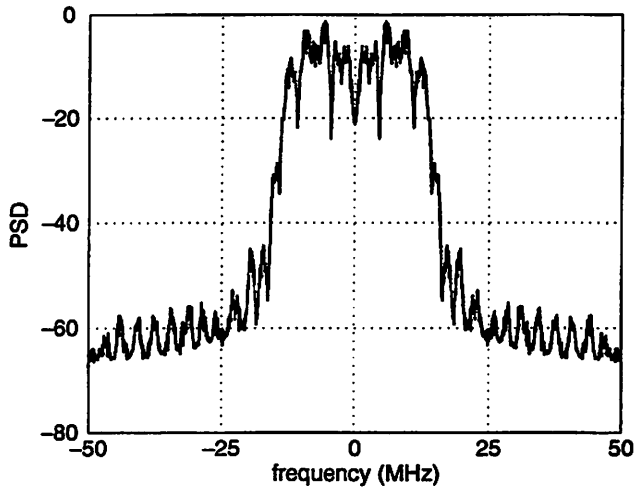


Figure 5.10: Simulated transmit spectrum.

$$SNR = \frac{KP_{HPF}}{Nf_{sig}} = \frac{KP_{sig}}{Nf_{sig}}\left[1 - \frac{f_{HPF}}{f_{sig}}\arctan\left(\frac{f_{sig}}{f_{HPF}}\right)\right] \tag{5.33}$$

and the SNR degradation resulting from high-pass filtering is

$$L = 1 - \frac{f_{HPF}}{f_{sig}}\arctan\left(\frac{f_{sig}}{f_{HPF}}\right). \tag{5.34}$$

For $f_{sig} = 25\,\text{MHz}$ and $f_{HPF} = 500\,\text{kHz}$, the SNR degradation is about 0.14 dB. From simulations, the SNR degradation in this case is approximately 0.1 dB. The calculated value is slightly pessimistic since the received signal actually does not have a constant PSD. The $K = 15$ sequence in (5.1) has very little spectral content near dc with an average value of about 0.067. The simulated PSD is illustrated in Fig. 5.10. Consequently, by using spreading codes with little or no spectral content near dc, the SNR degradation after high-pass filtering can be minimized. Moreover, for wideband signals, the resulting degradation in spectral efficiency due to such coding is insignificant.

As discussed in Section 2.3.1, dc offsets originate from multiple sources, including LO self-mixing, even-order distortion, and systematic offsets in the baseband circuits. If dc offsets are removed immediately prior to analog-to-digital conversion, the dc offsets from earlier stages can still saturate the subsequent baseband stages prior to dc-offset removal

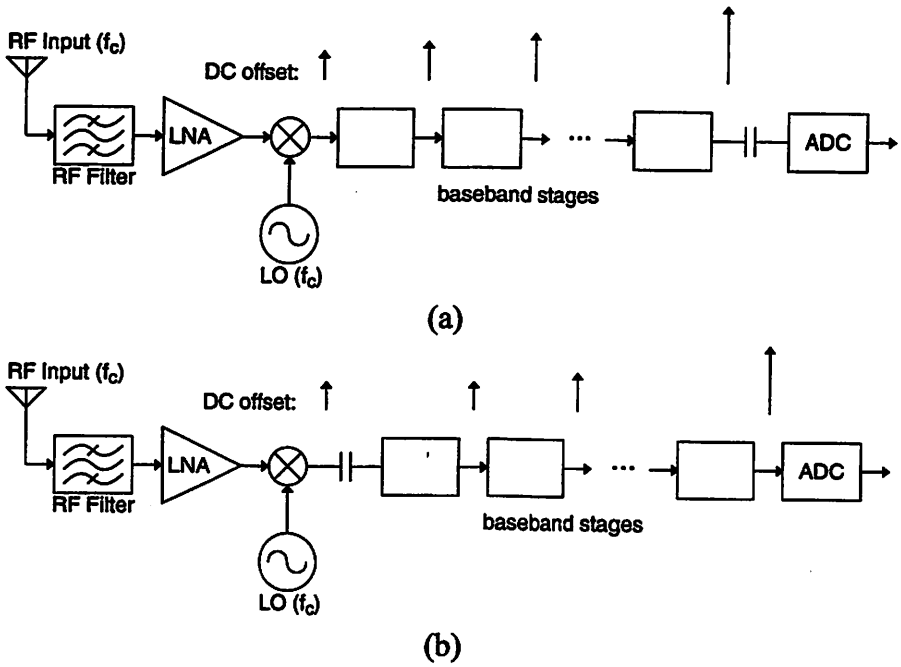


Figure 5.11: DC offset removal. (a) Before ADC. (b) After mixer.

(Fig. 5.11a). Alternatively, if dc offsets are removed immediately after frequency translation to baseband, the dc offsets caused by subsequent baseband stages can still be problematic (Fig. 5.11b). In the latter approach, the dc offsets introduced by the subsequent baseband stages can be minimized by using the same techniques used to reduce flicker noise, such as using large transistor dimensions, autozeroing, and chopper stabilization (Section 5.3.5). Alternatively, additional coupling capacitors or high-pass filters can also be used to remove the dc offsets in these subsequent stages.

In the proposed implementation, a high-pass filter is located immediately after the mixer along each of the I and Q signal paths in order to remove dc offsets. This filter is effective in removing dc offsets caused by LO self-mixing but does not address the dc offset problem in subsequent baseband stages. The proposed implementation relies on large transistor dimensions as well as layout techniques which improve transistor matching in order to minimize the dc offsets introduced by these baseband stages.

Finally, although capacitive coupling or high-pass filtering using on-chip capacitors and resistors are simple and effective ways of eliminating dc offsets for wideband systems, this technique is not feasible for narrowband systems based on the direct-conversion

architecture. For these systems, the narrow bandwidth of the desired signal requires very large capacitance and resistance values for removal of the dc offsets. One alternative is to use off-chip components for these passive structures. However, this approach is inconsistent with the goal of a highly-integrated implementation. An overview of some the techniques used for dc-offset cancellation in direct-conversion narrowband receivers is provided in Appendix B.

5.3.7 Receiver Noise Figure

The required noise figure is determined from the receiver sensitivity as given by (4.6), which is repeated here for convenience:

$$NF\,[\mathrm{dB}] = P_{sig}\,[\mathrm{dBm}] - 10\log(\Delta f\,[\mathrm{Hz}]) - SNR_{out}[\mathrm{dB}] + 173.8\,. \qquad (5.35)$$

In order to guarantee an average BER of 10^{-4}, an SNR of approximately 15 dB is required [22]. This specification assumes a multipath transmission channel with QPSK modulation. If data recovery is performed in the digital section of the receiver, the required SNR at the output of the analog section is relaxed by an amount equal to the processing gain. The processing gains for $K = 15$ and $K = 31$ are 11.76 dB and 14.91 dB, respectively. For a system bandwidth of 32.5 MHz, the noise figure must be better than 14.2 dB in both cases. The noise figure requirements for both cases are virtually identical since the larger processing gain in the $K = 31$ case offsets its more stringent sensitivity requirement.

5.3.8 ADC Performance

The proposed direct-conversion receiver requires two ADCs, one for each of the I and Q baseband channels. The partitioning of receiver functions between the analog and digital sections directly impacts the performance requirements of the ADCs [49]. Two possible configurations are illustrated in Fig. 5.12. The I and Q channels are identical, so for simplicity, only one channel is shown. In the architecture depicted in Fig. 5.12a, data recovery is performed in the digital section after analog-to-digital conversion. In this case, the sampling rate of the ADC must be at least 25 MHz in order to avoid destructive aliasing. In the architecture depicted in Fig. 5.12b, data recovery is performed in the

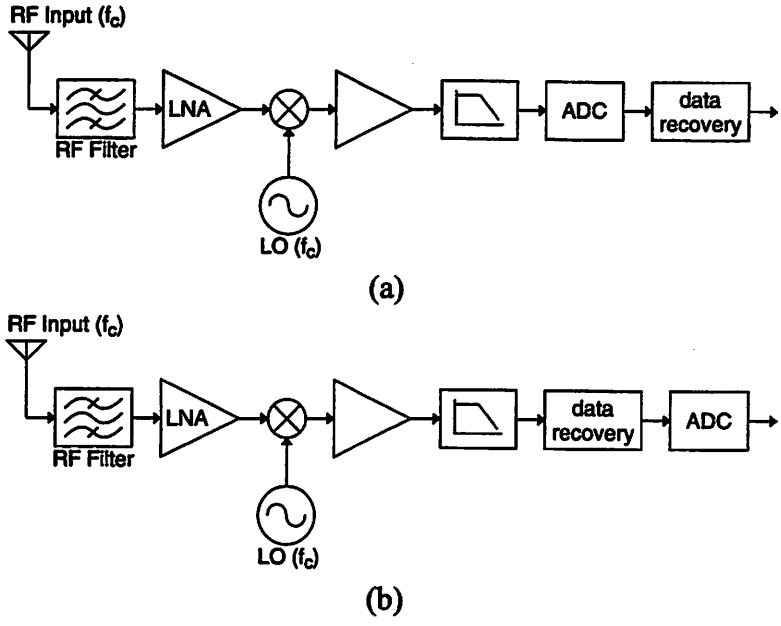


Figure 5.12: (a) ADC before data recovery. (b) ADC after data recovery.

analog section prior to analog-to-digital conversion. In this case, the Nyquist sampling rate requirement of the ADC is reduced to the symbol rate of 1.67 MHz for $K = 15$ or 0.81 MHz for $K = 31$. Determining the resolution requirements of the ADCs is not as straightforward since these requirements depend heavily on the specific algorithms used for data recovery. For example, for CDMA systems which rely on single-user techniques (Section 5.3.12) for data recovery, the architecture depicted in Fig. 5.12a typically requires approximately four bits of resolution in the ADC [18], [50], while the architecture depicted in Fig. 5.12b requires only one bit of resolution in the ADC [49]. For this example, based on the ADC requirements alone, the architecture depicted in Fig. 5.12b appears to be the obvious choice, since the ADC has both a lower sampling rate requirement as well as a lower resolution requirement.

However, a fair comparison of the two architectures must also take into account the implementation of the data recovery algorithm. For CDMA systems which rely on single-user techniques for data recovery, the power consumption of an analog implementation of the data recovery algorithm is about the same as that of a digital implementation [49], and consequently, the architecture depicted in Fig. 5.12b still appears to be the obvious choice. Nevertheless, the architecture depicted in Fig. 5.12a may actually be more

attractive for several reasons. First, a digital implementation of the data recovery algorithm provides increased design flexibility. Since digital implementations can take advantage of circuit synthesis techniques, much faster design times are possible. Second, the continued scaling of CMOS technology results in significant improvements in the speed, size, and power consumption of digital circuits. In contrast, the scaling of CMOS technology has actually hindered the design of analog circuits, mainly because of decreasing supply voltages [51]. Finally, for receivers which rely on more advanced techniques for data recovery, a digital implementation may be the only feasible alternative. In order to increase system performance, receivers are beginning to incorporate more advanced algorithms for timing synchronization [19] and data detection, such as multiuser techniques [52]. While the complexity of these algorithms along with the decreasing supply voltages of CMOS processes result in very challenging analog implementations, these algorithms are actually very well suited to low-power digital implementation techniques [5]. As a result, the benefits of a digital implementation of the data recovery block may actually outweigh the disadvantages of implementing an ADC with higher speed and resolution requirements. Consequently, the proposed direct-conversion receiver is based on the configuration depicted in Fig. 5.12a. The remainder of this section focuses on the performance requirements of the I and Q ADCs in the proposed architecture.

In order to avoid destructive aliasing, the sampling rate of each of the I and Q ADCs must be at least 25 MHz. Although the single-sided bandwidth of the desired signal is actually 16.25 MHz, a minimum sampling rate of 25 MHz rather than 32.5 MHz is actually sufficient to avoid aliasing, despite the forebodings of the Nyquist Sampling Theorem, which states that a signal with single-sided bandwidth f_B is uniquely represented by samples taken at the Nyquist frequency, $f_s \geq 2f_B$ [16]. In the proposed system, the bandwidth expansion results from the raised-cosine pulse-shaping filter with 30% excess bandwidth, which satisfies the Nyquist criterion for zero ISI [16]. The equivalent low-pass transmitted signal can be expressed as

$$y(t) = \sum_{n=0}^{\infty} I_n x(t - nT) \tag{5.36}$$

where $\{I_n\}$ represents the discrete symbol sequence and $x(t)$ is the raised-cosine pulse given by (5.5). The single-sided bandwidth of $y(t)$ is $(1+\beta)/2T$. If $y(t)$ is sampled at $f_s = 1/T$ then the resulting sequence is

$$\begin{aligned} y(kT) &= \sum_{n=0}^{\infty} I_n x[(k-n)T], \quad k = 0,1,2,\ldots \\ &= \sum_{n=0}^{\infty} I_n \frac{\sin[\pi(k-n)]}{\pi(k-n)} \frac{\cos[\pi\beta(k-n)]}{1-4\beta^2(k-n)^2} \\ &= I_k \end{aligned} \tag{5.37}$$

since

$$\sum_{\substack{n=0 \\ n\neq k}}^{\infty} I_n \frac{\sin[\pi(k-n)]}{\pi(k-n)} \frac{\cos[\pi\beta(k-n)]}{1-4\beta^2(k-n)^2} = 0. \tag{5.38}$$

Consequently, when the transmitted symbol sequence is shaped by a raised-cosine pulse resulting in a single-sided bandwidth of $(1+\beta)/2T$, a minimum sampling rate of $1/T$ rather than $(1+\beta)/2T$ is required in order to recover the desired symbol sequence. For the proposed system, the minimum ADC sampling rate requirement is 25 MHz. However, the sampling rate may actually be greater than 25 MHz since oversampling the received signal facilitates digital timing recovery.

Next, the resolution requirement of each of the I and Q ADCs is determined [22]. The mean square value of the quantization error introduced by the ADC is given by (3.32) and is repeated here:

$$\sigma_e^2 = \frac{1}{12} \times \frac{V_{FS}^2}{2^{2R}}. \tag{5.39}$$

In the proposed system, each of the received baseband I and Q signals can be represented as

$$r[n] = \frac{1}{\sqrt{2}} \sum_{k=1}^{K} a_k b_k g_k[n] \tag{5.40}$$

where $a_k/\sqrt{2}$, b_k, and $g_k[n]$ are the amplitude, bit sequence, and spreading code, respectively, of the k^{th} data channel. The signal $r[n]$ may be approximated as a Gaussian random variable with zero mean and variance

$$\sigma_r^2 = \frac{1}{2}\sum_{k=1}^{K} a_k^2 . \tag{5.41}$$

If the signal amplitude has a Gaussian distribution, then only 0.064% of the samples have an amplitude greater than $4\sigma_r$ [21]. Thus, by setting

$$\frac{V_{FS}}{2} = 4\sigma_r \tag{5.42}$$

(5.39) becomes

$$\sigma_e^2 = \frac{1}{12} \times \frac{\frac{64}{2}\sum_{k=1}^{K} a_k^2}{2^{2R}} = \frac{8}{3} \times \frac{\sum_{k=1}^{K} a_k^2}{2^{2R}} . \tag{5.43}$$

For the k^{th} user, the output SNR is

$$\begin{aligned} SNR &= \frac{a_k^2}{2\sigma_e^2} \\ &= \frac{3}{16} \times 2^{2R} \times \frac{a_k^2}{\sum_{k=1}^{K} a_k^2} \\ &= -7.27 + 6.02R + 10\log\frac{a_k^2}{\sum_{k=1}^{K} a_k^2} \text{ dB.} \end{aligned} \tag{5.44}$$

In order to guarantee an average BER of 10^{-4}, an SNR of approximately 15 dB is required [22]. Assuming that the receiver noise is dominated by thermal noise, then the SNR due to only quantization noise must be much better than 15 dB. Since data recovery is performed in the digital section of the receiver, the required SNR at the output of the ADC is relaxed by an amount equal to the processing gain. For an SNR of 25 dB due to quantization noise alone, the ADC resolution requirement is 6 bits for both $K = 15$ and $K = 31$ when all data channels equally share the total transmit power. If power control is employed at the base station, then the ADC resolution requirement is more stringent due

to the increased dynamic range of the received signal. If the transmit power of the k^{th} data channel is ten times less than the power of each of the other $K-1$ data channels, then the SNR of the k^{th} data channel after data recovery is

$$SNR = -7.27 + 6.02R + 10\log\frac{1}{1+10(K-1)} + 10\log K. \tag{5.45}$$

In this case, for an SNR of 25 dB, the ADC resolution requirement is 7 bits for both $K = 15$ and $K = 31$.

5.3.9 Receiver Gain

The receiver gain requirements are determined by the minimum and maximum signal levels expected to appear at the receiver input. Since the proposed system is designed to operate over short distances, out-of-band interference is assumed to negligible. The minimum gain requirement is determined by the largest in-band signal appearing at the receiver input. Assuming that the minimum separation between the transmitter and the receiver is 1 m and that no shadowing losses occur, the minimum path loss is given by

$$L = 10\log\left\{\left[\frac{4\pi(2\times10^9)}{3\times10^8}\right]^2 1^3\right\} = 38.46\,\text{dB}. \tag{5.46}$$

For a transmit power of 0 dBm, the maximum received power is –38.46 dBm. Assuming a 1-V swing for the baseband circuits, the minimum receiver gain is

$$G_{min} = 13.01\,\text{dBm} - (-38.46\,\text{dBm}) = 51.47\,\text{dB}. \tag{5.47}$$

Similarly, the maximum gain requirement is determined by the smallest in-band signal appearing at the receiver input. Assuming that the maximum separation between the transmitter and the receiver is 5 m and that the maximum shadowing loss is 10 dB, the maximum path loss is 69.43 dB as given by (5.20). For a transmit power of 0 dBm, the minimum received power is –69.43 dBm. Again, assuming a 1-V swing for the baseband circuits, the maximum receiver gain is

$$G_{max} = 13.01\,\text{dBm} - (-69.43\,\text{dBm}) = 82.44\,\text{dB}. \tag{5.48}$$

Consequently, the receiver must have a dynamic range of at least 31 dB. One approach is to increase the ADC resolution in order to accommodate this dynamic range. However, doing so exacerbates the already stringent design requirements of the ADC. Alternatively, an automatic gain control (AGC) loop can be used to adjust the gain of the receiver depending on the received signal power.

5.3.10 Receiver Distortion

The distortion performance of a receiver is determined by the anticipated levels of out-of-band interferers relative to the level of the desired signal. For the proposed system, specifications for worst-case out-of-band blocker levels are not available, and thus, the receiver intermodulation distortion requirements cannot be precisely determined. However, the desired signal in the proposed system is generally stronger than potential out-of-band interferers due to the short transmission distances. Consequently, the intermodulation distortion performance of the proposed receiver is not particularly stringent.

Finally, since the received signal can be as large as –38.46 dBm as described in Section 5.3.9, the 1-dB compression point of the receiver must be better than –38.46 dBm.

5.3.11 Receiver AGC Loop

In the proposed system, the received signal can be as small as –69.43 dBm or as large as –38.46 dBm. Without gain control, a weak received signal will not take full advantage of the dynamic range of the ADC. Moreover, a very strong signal will saturate the analog circuits in the receiver. For the proposed system, an AGC loop with a dynamic range of 31 dB is used to adjust the amplitude of the received signal.

Since the AGC loop must set the proper gain before the ADC in order to take advantage of the full dynamic range of the ADC, many AGC loops are designed using only analog circuits. In this case, the gain of an analog variable-gain amplifier (VGA) is adjusted based on the signal amplitude, which is determined using an analog peak detector circuit

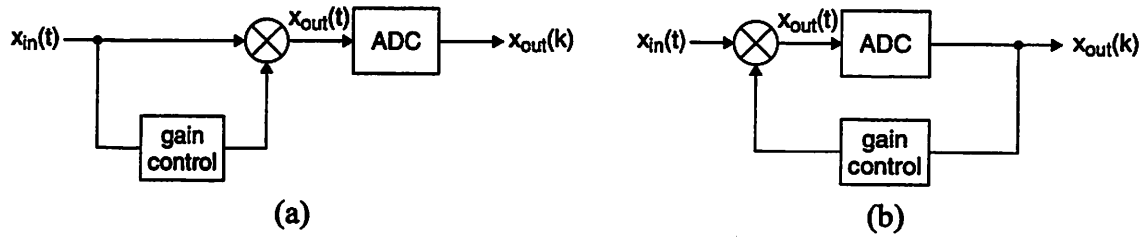


Figure 5.13: AGC architectures. (a) Feedforward. (b) Feedback.

[53]. For the proposed system, the AGC loop is partitioned between analog and digital circuits. The gain control algorithm is implemented in the digital section and then the proper gain setting is fed back to the analog VGA.

Two types of AGC loops are illustrated in Fig. 5.13. The feedforward AGC generally converges faster than the feedback AGC. Moreover, feedforward loops generally do not have stability problems. However, the feedforward architecture is not very amenable to a mixed-signal implementation. The multiplier or VGA must precede the ADC in order to set the proper signal amplitude at the ADC input. Consequently, in order for the gain-control algorithm to be performed digitally in the feedforward AGC, an additional ADC is required at the input of the gain-control block. In the feedback AGC, only a single ADC is required. Consequently, the AGC loop for the proposed system is based on the feedback architecture. By designing the VGA to have discrete gain settings, the digital signal can control the VGA directly without the need for a DAC. Since the AGC loop for the proposed system must have a dynamic range of 31 dB, the gain can be set by activating various combinations of five amplifiers with gains of 1 dB, 2 dB, 4 dB, 8 dB, and 16 dB.

In order to determine the correct gain, an estimate of the received amplitude is required. The amplitude estimate should be robust even in the presence of receiver impairments such as circuit noise, frequency offset, and distortion. For the proposed system, the estimate is based on both the I and Q data, which in the ideal case, form a four-point constellation centered and symmetric about the origin. Three methods of amplitude estimation are [54]:

1. $\sqrt{I^2+Q^2}$ (5.49)

2. $|I|+|Q|$ (5.50)

3. $\max(|I|,|Q|)+\frac{1}{2}\min(|I|,|Q|)$. (5.51)

If the receiver is noiseless and the only impairment is the frequency offset between the transmitter and receiver oscillators, then the constellation will rotate along a circle as illustrated in Fig. 3.12. The I and Q data can be represented as, respectively,

$$I = I_o \cos(\Delta\omega t) \tag{5.52}$$

$$Q = Q_o \sin(\Delta\omega t) \tag{5.53}$$

where $\Delta\omega$ is the frequency offset and $|I_o| = |Q_o|$ for an ideal constellation. Assuming that $|I_o| = |Q_o| = 1$, (5.49) becomes

$$\sqrt{I^2+Q^2} = 1. \tag{5.54}$$

Similarly, (5.50) and (5.51) become, respectively,

$$|I|+|Q| = |\cos(\Delta\omega t)|+|\sin(\Delta\omega t)| \tag{5.55}$$

$$\max(|I|,|Q|)+\tfrac{1}{2}\min(|I|,|Q|) = \max[|\cos(\Delta\omega t)|,|\sin(\Delta\omega t)|]+ \tfrac{1}{2}\min[|\cos(\Delta\omega t)|,|\sin(\Delta\omega t)|]. \tag{5.56}$$

These three estimates are plotted in Fig. 5.14. The amplitude estimate given by (5.54) is constant over time and the average value is equal to one. A low-pass filter can be used to reduce the variation in the estimates given by (5.55) and (5.56). If the low-pass filter bandwidth is sufficiently small, then the output of the low-pass filter is simply the average value of the input. The average values of (5.55) and (5.56) are, respectively,

$$\frac{8}{2\pi} \approx 1.27 \tag{5.57}$$

$$\frac{2+\sqrt{2}}{\pi} \approx 1.09. \tag{5.58}$$

For practical low-pass filter implementations, some variation still exists, and consequently, (5.49) provides the best amplitude estimate when the receiver is noiseless and the only impairment is frequency offset.

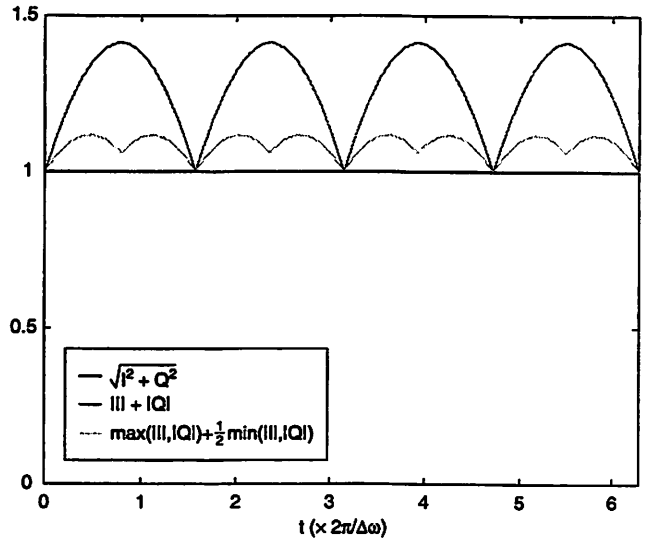


Figure 5.14: Amplitude estimates with frequency offset $\Delta\omega$.

When the noise power is significantly larger than the signal power, then the performance of all three estimates is comparable. In this case I and Q are given by, respectively,

$$I = I_o \cos(\Delta\omega t) + N_I \approx N_I \tag{5.59}$$

$$Q = Q_o \sin(\Delta\omega t) + N_Q \approx N_Q . \tag{5.60}$$

Assuming that the noise components are both independent Gaussian random variables with zero mean and variance, σ_N^2, the expected value of the estimate given by (5.49) is

$$E[\sqrt{I^2 + Q^2}] = E[V] \tag{5.61}$$

where $V = \sqrt{I^2 + Q^2}$ is a Rayleigh random variable [55] with mean and variance given by, respectively,

$$E[V] = \sqrt{\frac{\pi}{2}}\sigma_N \tag{5.62}$$

$$VAR[V] = \left(2 - \frac{\pi}{2}\right)\sigma_N^2 . \tag{5.63}$$

The expected value of the estimate given by (5.50) is

$$E[|I| + |Q|] = 2\int_{-\infty}^{\infty} |x| \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{x^2}{2\sigma^2}} dx = 2\sqrt{\frac{2}{\pi}}\sigma_N \tag{5.64}$$

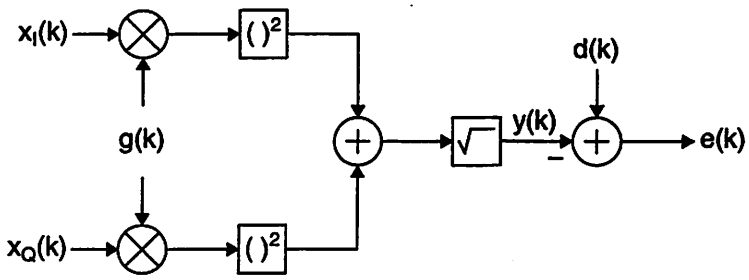


Figure 5.15: Equivalent combiner for feedback AGC loop based on (5.49).

while the expected value of the estimate given by (5.51) is

$$E\left[\max(|I|,|Q|)+\frac{1}{2}\min(|I|,|Q|)\right]=\frac{3}{2}\sqrt{\frac{2}{\pi}}\sigma_N. \tag{5.65}$$

The gain-control algorithm can be implemented using an adaptive least mean squares (LMS) algorithm [56]. The equivalent combiner for a feedback AGC loop based on (5.49) for amplitude estimation is illustrated in Fig. 5.15. The error signal $e(k)$ is given by

$$e(k)=d(k)-g(k)\sqrt{x_I^2(k)+x_Q^2(k)}. \tag{5.66}$$

Adaptation using the stochastic gradient descent method results in the following update equation:

$$g(k+1)=g(k)-\mu\frac{\partial}{\partial g(k)}\left\{\frac{1}{2}e^2(k)\right\}=g(k)-\mu e(k)\frac{\partial}{\partial g(k)}\{e(k)\} \tag{5.67}$$

where μ is the step size. Taking the partial derivative of $e(k)$ with respect to $g(k)$, (5.67) becomes

$$g(k+1)=g(k)+\mu e(k)\sqrt{x_I^2(k)+x_Q^2(k)}. \tag{5.68}$$

The stability criterion for this algorithm is determined by first setting

$$d(k)=\phi\sqrt{x_I^2(k)+x_Q^2(k)}. \tag{5.69}$$

The prediction error is then

$$e(k)=d(k)-y(k)=[\phi-g(k)]\sqrt{x_I^2(k)+x_Q^2(k)}. \tag{5.70}$$

The parameter error vector update is

$$\begin{aligned}\bar{g}(k+1) &= \phi - g(k+1) \\ &= \phi - g(k) - \mu[\phi - g(k)][x_I^2(k) + x_Q^2(k)] \\ &= \bar{g}(k)\{1 - \mu[x_I^2(k) + x_Q^2(k)]\}\end{aligned} \tag{5.71}$$

and the summed squared parameter error increment is

$$\bar{g}^2(k+1) - \bar{g}^2(k) = -\mu\bar{g}^2(k)[x_I^2(k) + x_Q^2(k)]\{2 - \mu[x_I^2(k) + x_Q^2(k)]\}. \tag{5.72}$$

If the algorithm converges, then the parameter error vector update at time $k+1$ must be less than that at time k. Consequently, the summed squared parameter error increment must always be negative. For a positive step size μ, the following relationship must be satisfied if the algorithm converges:

$$0 < \mu < \frac{2}{x_I^2(k) + x_Q^2(k)}. \tag{5.73}$$

Implementation of the update equation in (5.68) requires the calculation of $\sqrt{x_I^2(k) + x_Q^2(k)}$, which may be achieved by dividing $y(k)$ in Fig. 5.15 by $g(k)$. This division operation may be eliminated by using the sign-data algorithm [57] instead of the stochastic gradient descent method described above. Adaptation using the sign-data algorithm results in the following update equation:

$$g(k+1) = g(k) + \mu e(k)\,\mathrm{sgn}[\sqrt{x_I^2(k) + x_Q^2(k)}]. \tag{5.74}$$

Since

$$\sqrt{x_I^2(k) + x_Q^2(k)} > 0 \tag{5.75}$$

(5.74) becomes

$$g(k+1) = g(k) + \mu e(k). \tag{5.76}$$

The stability criterion is determined by rewriting (5.76) as

$$g(k+1) = g(k) + \frac{\mu}{\sqrt{x_I^2(k) + x_Q^2(k)}} e(k) \sqrt{x_I^2(k) + x_Q^2(k)}\,. \tag{5.77}$$

Comparing (5.77) with the update equation for the stochastic gradient descent algorithm in (5.68), the sign-data algorithm is stable if

$$0 < \mu < \frac{2}{\sqrt{x_I^2(k) + x_Q^2(k)}}. \tag{5.78}$$

The computational complexity can be further reduced by eliminating the square-root operation. In this case, the square of the signal amplitude, instead of the signal amplitude itself, is estimated:

$$I^2 + Q^2. \tag{5.79}$$

The equivalent combiner is illustrated in Fig. 5.16. The error signal $e(k)$ is given by

$$e(k) = d(k) - y(k) = d(k) - g^2(k)[x_I^2(k) + x_Q^2(k)]. \tag{5.80}$$

Unfortunately, the error signal is not linear in $g(k)$, and consequently, the techniques used to analyze the stochastic gradient descent and sign-data algorithms cannot be applied in this case. An update equation which results in a simple implementation is given by

$$g(k+1) = g(k) + \mu e(k). \tag{5.81}$$

The performance and stability of this algorithm is beyond the scope of this discussion and the interested reader is referred to [58] for more details. From simulations, the performance of this algorithm is very similar to that of the sign-data algorithm using

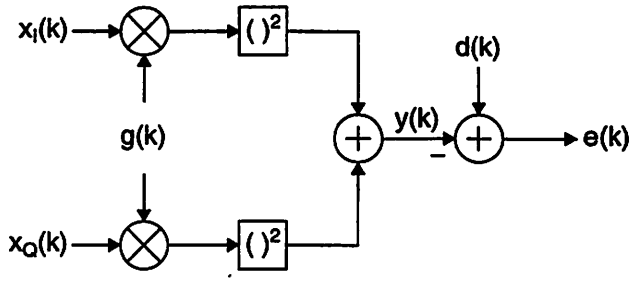


Figure 5.16: Equivalent combiner based on $I^2 + Q^2$.

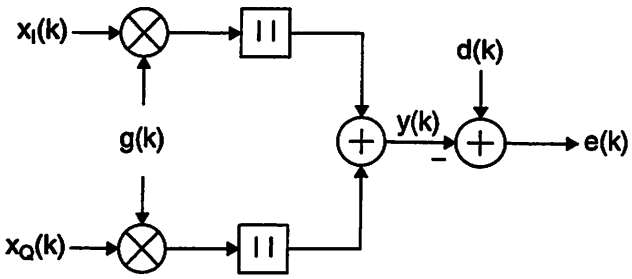


Figure 5.17: Equivalent combiner for feedback AGC loop based on (5.50).

(5.49) for amplitude estimation.

The equivalent combiner for a feedback AGC loop using (5.50) for amplitude estimation is illustrated in Fig. 5.17. The error signal $e(k)$ is given by

$$e(k) = d(k) - [|\, g(k)x_I(k)\,| + |\, g(k)x_Q(k)\,|]. \tag{5.82}$$

Since $g(k) > 0$, (5.82) becomes

$$e(k) = d(k) - g(k)[|\, x_I(k)\,| + |\, x_Q(k)\,|]. \tag{5.83}$$

Adaptation using the stochastic gradient descent method results in the following update equation:

$$g(k+1) = g(k) + \mu e(k)[|\, x_I(k)\,| + |\, x_Q(k)\,|]. \tag{5.84}$$

And the stability criterion is given by

$$0 < \mu < \frac{2}{[|\, x_I(k)\,| + |\, x_Q(k)\,|]^2}. \tag{5.85}$$

In this case, the computational complexity can also be reduced by using the sign-data algorithm rather than the stochastic gradient descent method. Adaptation using the sign-data algorithm results in the following update equation:

$$g(k+1) = g(k) + \mu e(k)\,\mathrm{sgn}[|\, x_I(k)\,| + |\, x_Q(k)\,|]. \tag{5.86}$$

Since

$$|x_I(k)| + |x_Q(k)| > 0 \tag{5.87}$$

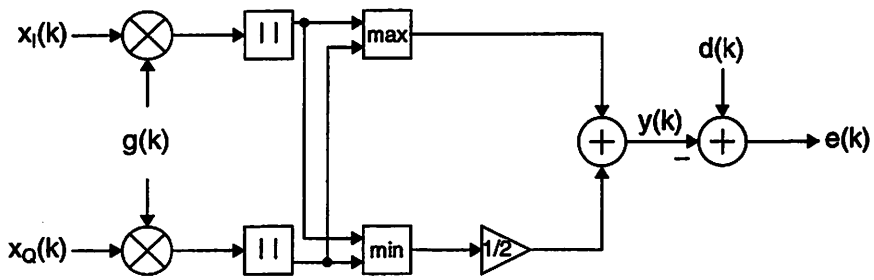


Figure 5.18: Equivalent combiner for feedback AGC loop based on (5.51).

(5.86) becomes

$$g(k+1) = g(k) + \mu e(k). \tag{5.88}$$

This algorithm is stable if

$$0 < \mu < \frac{2}{|x_I(k)| + |x_Q(k)|}. \tag{5.89}$$

Finally, the equivalent combiner for a feedback AGC loop using (5.51) for amplitude estimation is illustrated in Fig. 5.18. The error signal $e(k)$ is given by

$$\begin{aligned} e(k) = d(k) - \{&\max[|g(k)x_I(k)|, |g(k)x_Q(k)|] + \\ &\frac{1}{2}\min[|g(k)x_I(k)|, |g(k)x_Q(k)|]\}. \end{aligned} \tag{5.90}$$

Since $g(k) > 0$, (5.90) becomes

$$e(k) = d(k) = g(k)\left\{\max[|x_I(k)|, |x_Q(k)|] + \frac{1}{2}\min[|x_I(k)|, |x_Q(k)|]\right\}. \tag{5.91}$$

Adaptation using the stochastic gradient descent method results in the following update equation:

$$g(k+1) = g(k) + \mu e(k)\left\{\max[|x_I(k)|, |x_Q(k)|] + \frac{1}{2}\min[|x_I(k)|, |x_Q(k)|]\right\}. \tag{5.92}$$

And the stability criterion is given by

$$0 < \mu < \frac{2}{\left\{\max[|x_I(k)|, |x_Q(k)|] + \frac{1}{2}\min[|x_I(k)|, |x_Q(k)|]\right\}^2}. \tag{5.93}$$

Again, the computational complexity can be reduced by using the sign-data algorithm rather than the stochastic gradient descent method. Adaptation using the sign-data algorithm results in the following update equation:

$$g(k+1) = g(k) + \mu e(k)\,\mathrm{sgn}\left\{\max[|x_I(k)|,|x_Q(k)|] + \frac{1}{2}\min[|x_I(k)|,|x_Q(k)|]\right\}. \quad (5.94)$$

Since

$$\max[|x_I(k)|,|x_Q(k)|] + \frac{1}{2}\min[|x_I(k)|,|x_Q(k)|] > 0 \quad (5.95)$$

(5.94) becomes

$$g(k+1) = g(k) + \mu e(k). \quad (5.96)$$

This algorithm is stable if

$$0 < \mu < \frac{2}{\max[|x_I(k)|,|x_Q(k)|] + \frac{1}{2}\min[|x_I(k)|,|x_Q(k)|]}. \quad (5.97)$$

The sign-data LMS algorithm based on (5.79) for amplitude estimation offers excellent performance under both high SNR and low SNR conditions. However, implementation of this algorithm requires two squaring circuits as illustrated in Fig. 5.16. In contrast, implementation of the sign-data LMS algorithm based on (5.50) is much simpler, requiring two absolute value circuits instead of two squaring circuits. Moreover, if a sign-magnitude number representation is used, the absolute value operation is trivial. However, this algorithm performs poorly under high SNR conditions since amplitude estimates using (5.50) result in large variations as illustrated in Fig. 5.14. The sign-data LMS algorithm based on (5.51) for amplitude estimation provides a good compromise between good performance and ease of implementation. Implementation of this algorithm requires just two absolute value circuits and two comparators, while multiplication by 1/2 can be accomplished by a simple shift operation. Moreover, this algorithm offers good performance under high SNR conditions. The variations in the amplitude estimates using (5.51) as illustrated in Fig. 5.14 can be minimized by passing the estimates through a low-pass filter.

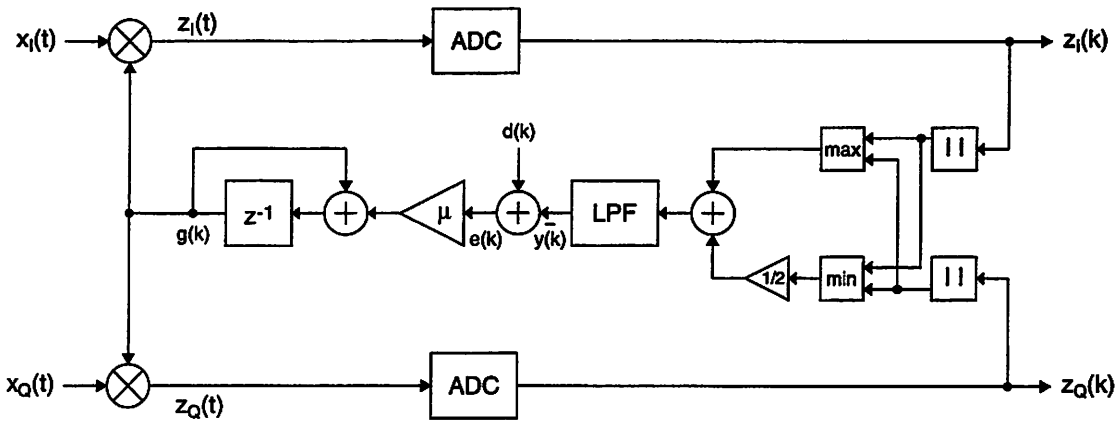


Figure 5.19: AGC loop based on sign-data LMS algorithm using (5.51).

An AGC loop based on the sign-data LMS algorithm using (5.51) for amplitude estimation is illustrated in Fig. 5.19. The AGC loop converges without the need for timing recovery since the amplitude estimation algorithm accounts for frequency offsets between the transmitter and receiver oscillators. The performance of this AGC loop is evaluated using Simulink. The ADCs are modeled by the simple behavioral model depicted in Fig. 4.16. Each of the ADCs samples the input signal at 25 MHz and quantizes it to 8 bits. Amplitude estimation based on (5.51) is performed after the I and Q ADCs and the digital update signal $g(k)$ controls both the I and Q VGAs. For simplicity, each of the VGAs is modeled as a multiplier.

Simulations reveal a potential problem with the AGC loop depicted in Fig. 5.19. The time constant of the sign-data LMS algorithm is inversely proportional to the step size and the amplitude estimate [57]

$$\tau \propto \frac{1}{\mu\left\{\max[|x_I(k)|,|x_Q(k)|]+\frac{1}{2}\min[|x_I(k)|,|x_Q(k)|]\right\}}. \tag{5.98}$$

For a fixed step size, the time constant depends on the amplitude of the input signal, and consequently, a step size which results in fast convergence for weak input signals may be too large for strong input signals, resulting in convergence noise, or perhaps even worse, the loop may become unstable. Conversely, a step size which results in fast convergence for strong input signals may be too small for weak input signals, resulting in very slow

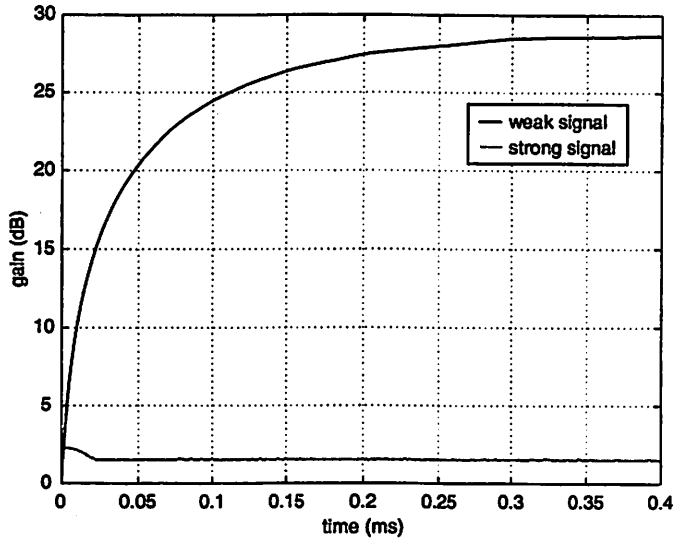


Figure 5.20: Performance of AGC loop depicted in Fig. 5.19.

convergence. As illustrated in Fig. 5.20, a step size of $\mu = 0.005$ results in fast convergence for large input signals but very slow convergence for small input signals. This simulation takes into account all receiver impairments including a receiver noise figure of approximately 13 dB and a frequency error of 50 ppm.

One way of speeding up the convergence time for all signal amplitudes is to use the normalized LMS algorithm [57]. The update equation for this algorithm is

$$g(k+1) = g(k) + \frac{\mu e(k)}{\max[|x_I(k)|,|x_Q(k)|] + \frac{1}{2}\min[|x_I(k)|,|x_Q(k)|]} \tag{5.99}$$

and the time constant is independent of the amplitude estimate:

$$\tau \propto \frac{1}{\mu}. \tag{5.100}$$

Consequently, for a fixed step size, this algorithm results in the same convergence time for all input signal amplitudes. Unfortunately, implementation of the update equation in (5.99) requires division by the amplitude estimate.

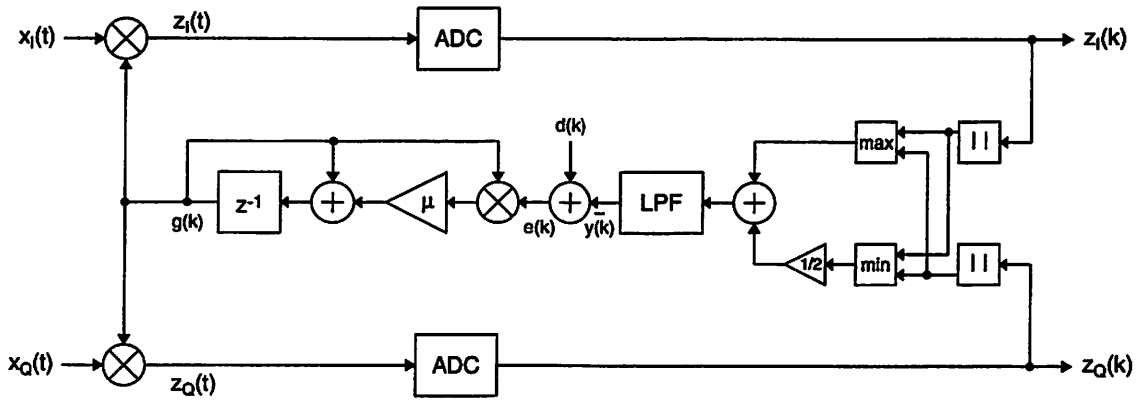


Figure 5.21: AGC loop based on the update equation in (5.101).

Fig. 5.21 illustrates the AGC loop for the proposed system. This AGC loop is based on an adaptive algorithm which converges rapidly for all signal amplitudes. The update equation for this algorithm is

$$g(k+1) = g(k) + \mu g(k) e(k). \tag{5.101}$$

A loose bound for the stability of this algorithm is

$$0 < \mu < \frac{2}{g(k)\left\{\max[|x_I(k)|,|x_Q(k)|] + \frac{1}{2}\min[|x_I(k)|,|x_Q(k)|]\right\}} \tag{5.102}$$

and the time constant is

$$\tau \propto \frac{1}{\mu g(k)\left\{\max[|x_I(k)|,|x_Q(k)|] + \frac{1}{2}\min[|x_I(k)|,|x_Q(k)|]\right\}} \approx \frac{1}{\mu d(k)}. \tag{5.103}$$

The time constant for this algorithm is only weakly dependent on the input signal amplitude. Although the implementation of this algorithm requires an additional multiplier, the benefit in performance is significant. As illustrated in Fig. 5.22, for the same step size of $\mu = 0.005$, the AGC loop converges much more rapidly for weak input signals without affecting the convergence performance for strong input signals.

5.3.12 Multiuser Detection

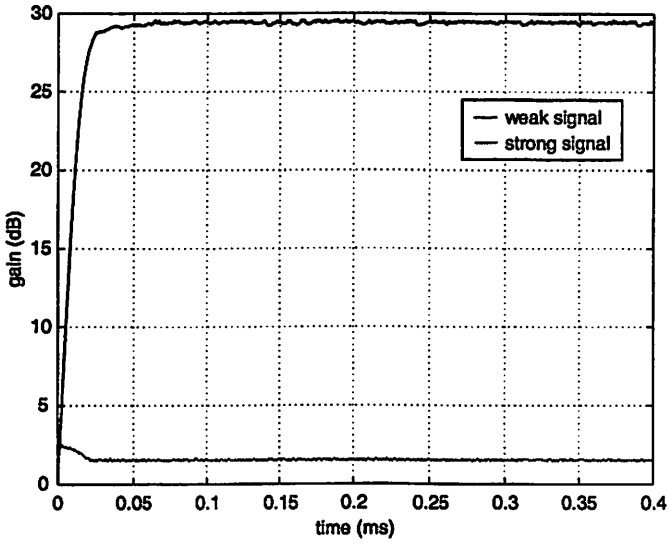


Figure 5.22: Performance of AGC loop depicted in Fig. 5.21.

In the proposed system, data recovery is performed using multiuser detection [22], [52]. This section provides a brief overview of detection algorithms for CDMA systems before describing the adaptive multiuser detection algorithm used in the proposed system. The interested reader is referred to [22] for a more detailed discussion.

In the proposed system, each of the baseband I and Q signals at the transmitter can be represented as

$$s[n] = \frac{1}{\sqrt{2}} \sum_{k=1}^{K} a_k b_k g_k[n] \tag{5.104}$$

where $a_k/\sqrt{2}$, b_k, and $g_k[n]$ are the amplitude, bit sequence, and spreading code, respectively, of the k^{th} data channel. Assuming that the signal is corrupted by AWGN with PSD $N/2$ during transmission and that perfect timing synchronization and gain control are maintained at the receiver, then each of the baseband I and Q signals at the ADC outputs can be represented as

$$r[m] = \frac{1}{\sqrt{2}} \sum_{j=1}^{K} a_j b_j g_j[m] + n[m] \tag{5.105}$$

where $n[m]$ is a discrete-time random process representing the sampled noise.

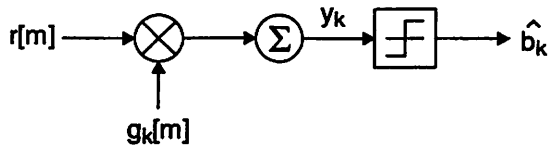


Figure 5.23: Single-user detector.

The conventional single-user detector used for data recovery in CDMA systems is illustrated in Fig. 5.23. This approach is commonly used for data recovery in the mobile receivers of many CDMA systems, including the IS-95 standard for digital cellular telephony, because of its simplicity and ease of implementation. In this approach, the k^{th} data channel is recovered by multiplying the received signal $r[m]$ with the k^{th} signature sequence and accumulating the result over K samples:

$$\begin{aligned} y_k &= \sum_{m=1}^{K} r[m]g_k[m] \\ &= \frac{1}{\sqrt{2}}\sum_{m=1}^{K}\sum_{j=1}^{K} a_j b_j g_j[m]g_k[m] + \sum_{m=1}^{K} n[m]g_k[m] \\ &= \frac{1}{\sqrt{2}} a_k b_k \sum_{m=1}^{K} g_k[m]g_k[m] + \frac{1}{\sqrt{2}}\sum_{m=1}^{K}\sum_{\substack{j=1 \\ j\neq k}}^{K} a_j b_j g_j[m]g_k[m] + \sum_{m=1}^{K} n[m]g_k[m]. \end{aligned} \tag{5.106}$$

If the signature sequences are orthogonal,

$$\sum_{m=1}^{K}\sum_{j=1}^{K} g_j[m]g_k[m] = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases} \tag{5.107}$$

then (5.106) becomes

$$y_k = \frac{1}{\sqrt{2}} a_k b_k + \sum_{m=1}^{K} n[m]g_k[m]. \tag{5.108}$$

In this case, system performance is noise-limited and the single-user detector is the optimum detector. However, when the signature sequences are not orthogonal, as in the case of the proposed system, the second term in (5.106) can be significant. For the MLSR signature sequences used in the proposed system,

$$\sum_{m=1}^{K}\sum_{j=1}^{K} g_j[m]g_k[m] = \begin{cases} 1, & j = k \\ -1/K, & j \neq k \end{cases} \tag{5.109}$$

and (5.106) becomes

$$y_k = \frac{1}{\sqrt{2}} a_k b_k - \frac{1}{\sqrt{2K}} \sum_{\substack{j=1 \\ j \neq k}}^{K} a_j b_j + \sum_{m=1}^{K} n[m] g_k[m]. \tag{5.110}$$

The interference from the other data channels can be significant, especially if any of the amplitudes a_j are sufficiently larger than the amplitude of the desired data channel. In this case, system performance is limited by multiple-access interference (MAI) and the single-user detector is no longer the optimum detector.

The decorrelating detector is the first of two linear multiuser detectors described in this section for performing data recovery in CDMA systems (Fig. 5.24). The detector output is

$$\begin{aligned} \hat{\mathbf{b}} &= \text{sgn}(\mathbf{z}) \\ &= \text{sgn}(\mathbf{R}^{-1}\mathbf{y}). \end{aligned} \tag{5.111}$$

The vector **y** represents the output from the K correlators and is given by

$$\mathbf{y} = \mathbf{RAb} + \mathbf{n} \tag{5.112}$$

where **R** is the cross-correlation matrix of the signature sequences with the entries of **R** given by

$$\mathbf{R}(i,j) = \frac{1}{K} \sum_{m=1}^{K} g_i[m] g_j[m] \tag{5.113}$$

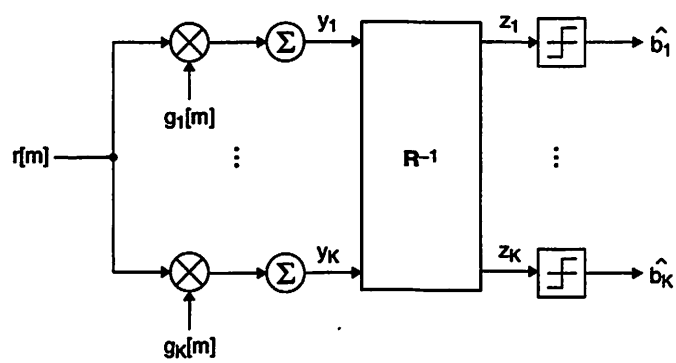


Figure 5.24: Decorrelating detector.

A is a $K \times K$ diagonal matrix of the amplitudes,

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & a_K \end{bmatrix} \tag{5.114}$$

b is a vector of the binary antipodal data streams, and **n** is a noise vector. The detector output is

$$\begin{aligned} \hat{\mathbf{b}} &= \text{sgn}(\mathbf{R}^{-1}\mathbf{y}) \\ &= \text{sgn}(\mathbf{R}^{-1}\mathbf{RAb} + \mathbf{R}^{-1}\mathbf{n}) \\ &= \text{sgn}(\mathbf{Ab} + \mathbf{R}^{-1}\mathbf{n}). \end{aligned} \tag{5.115}$$

Consequently, the decorrelating detector eliminates MAI, and when the system performance is interference-limited, the decorrelating detector is the optimum detector. However, since $\mathbf{R}^{-1}(k,k) \geq 1$, the decorrelating detector eliminates MAI at the expense of noise enhancement, and when system performance is noise-limited, the decorrelating detector is no longer the optimum detector.

The decorrelating detector depicted in Fig. 5.24 recovers all K data channels. However, in many cases, only a single data channel needs to be recovered at the mobile receiver. Fig. 5.25 illustrates an implementation of the decorrelating detector which recovers only the k^{th} data channel. This detector is very similar to the single-user detector illustrated in Fig. 5.23. However, in this case, the received signal $r[m]$ is multiplied by a modified signature sequence $h_k[m]$:

$$h_k[m] = \sum_{i=1}^{K} \mathbf{R}^{-1}(k,i) g_i[m]. \tag{5.116}$$

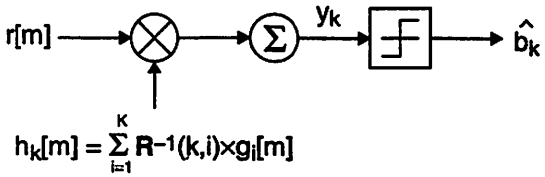


Figure 5.25: Decorrelating detector for a single data channel.

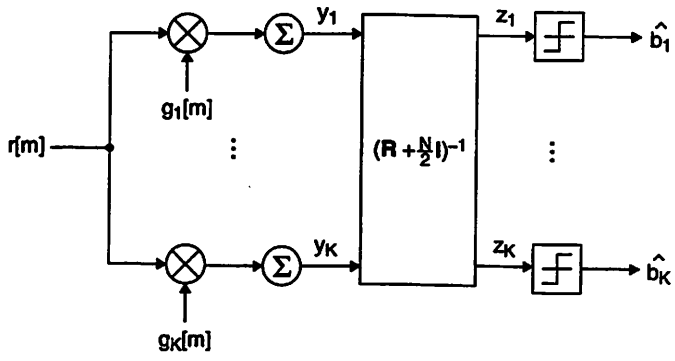


Figure 5.26: MMSE detector.

Finally, a second linear multiuser detector for performing data recovery in CDMA systems is the minimum mean-square error (MMSE) detector depicted in Fig. 5.26. In this approach, the algorithm attempts to minimize the mean-square error between $\hat{\mathbf{b}}$ and $\mathbf{z}$. The detector output is

$$\begin{aligned}\hat{\mathbf{b}} &= \text{sgn}(\mathbf{z}) \\ &= \text{sgn}\left[\left(\mathbf{R}+\frac{N}{2}\mathbf{I}\right)^{-1}\mathbf{y}\right]\end{aligned} \tag{5.117}$$

where $\mathbf{R}$ is the cross-correlation matrix of the signature sequences, the vector $\mathbf{y}$ represents the output from the K correlators, $N/2$ is the PSD of the AWGN, and $\mathbf{I}$ is the identity matrix. The performance of the MMSE detector approaches that of the decorrelating detector when $N \to 0$, while its performance approaches that of the single-user detector when $N \to \infty$. Consequently, the MMSE detector provides a good compromise between MAI suppression and noise enhancement when the system is neither interference-limited

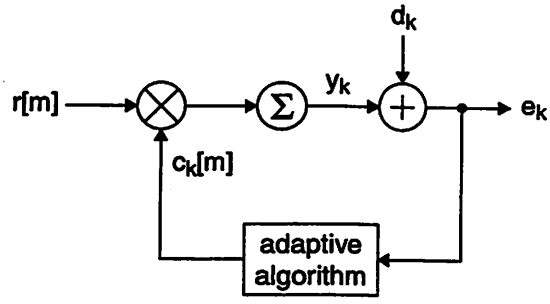


Figure 5.27: Adaptive MMSE detector for a single data channel.

nor noise-limited.

Data recovery in the proposed system is performed by an adaptive MMSE detector which recovers only a single data channel as illustrated in Fig. 5.27. The received signal $r[m]$ is multiplied by an adaptive sequence $c_k[m]$. Adaptation of $c_k[m]$ in the MMSE detector is achieved through the LMS algorithm. Additional details about the adaptive MMSE detector used in the proposed system is described in [22].

5.3.13 Summary

A block diagram of the proposed receiver is illustrated in Fig. 5.28. The receiver is based on a direct-conversion architecture and dc offsets are eliminated by high-pass filtering the I and Q signals immediately after translating the RF signal down to baseband. Due to the wide bandwidth of the desired signal, the corner frequency of each of the high-pass filters can be as high as 500 kHz, and consequently, these filters can be implemented using on-chip passive structures exclusively. DC offsets and flicker noise in the subsequent baseband stages are minimized by using large transistor dimensions. The carrier frequency and system bandwidth are 2 GHz and 32.5 MHz, respectively, and the required noise figure must be better than 14.2 dB in order to guarantee an average BER of 10^{-4}, which corresponds to an SNR of approximately 15 dB after data recovery. The minimum

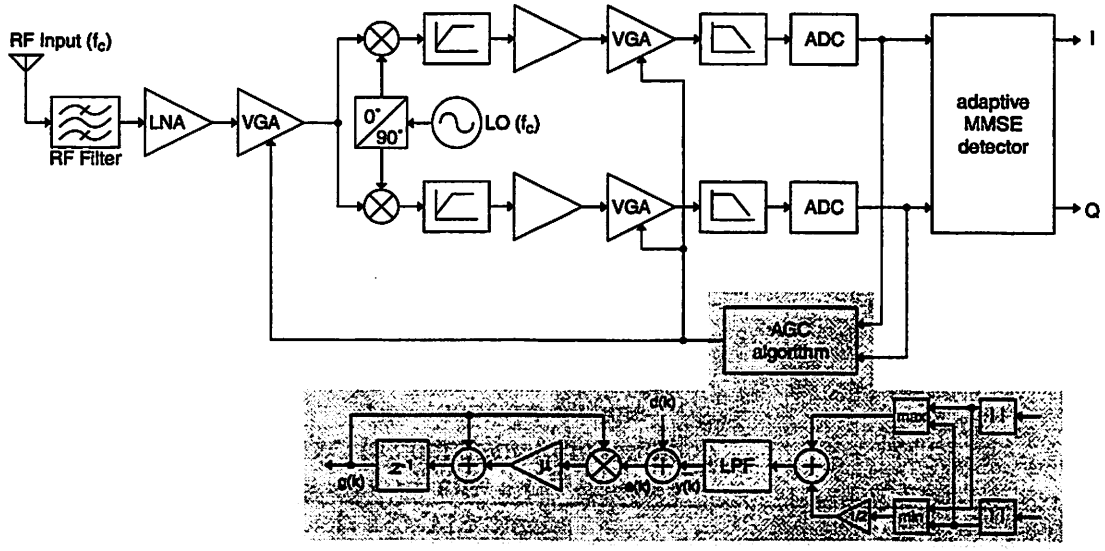


Figure 5.28: Proposed direct-conversion receiver.

and maximum gain requirements are about 51 dB and 82 dB, respectively, and an AGC loop with a dynamic range of 31 dB is used to adjust the amplitude of the received signal. The AGC loop is based on a feedback architecture and is partitioned between analog and digital circuits. The gain control algorithm is implemented in the digital section and then the proper gain setting is fed back to the analog VGA. By designing the VGA to have discrete gain settings, the digital signal can control the VGA directly without the need for a DAC. The sampling rate and resolution of each of the I and Q ADCs must be at least 25 MHz and 7 bits, respectively, while data recovery is performed using multiuser techniques. An adaptive MMSE detector provides a good compromise between MAI suppression and noise enhancement. The receiver specifications are summarized in Table 5.1.

The specifications summarized in Table. 5.1 serve as a starting point for designing the analog front-end of the receiver. However, before circuit design can begin, additional information is required. For example, the noise and gain requirements must be partitioned

	$K=15$	$K=31$
receiver architecture	direct conversion: high-pass filtering for dc offsets; large transistor sizes for dc offsets and flicker noise	
carrier frequency	2 GHz	
system bandwidth	32.5 MHz	
sensitivity	−81.2 dBm	−84.3 dBm
processing gain	11.76 dB	14.91 dB
noise figure	14.2 dB	
distortion	$P_{-1dB} > -38.46$ dBm	
gain	minimum: 51.47 maximum: 82.44	
ADC	Nyquist rate: 25 MHz resolution: 7 bits	
high-pass filter	corner frequency: < 500 kHz	
AGC loop	dynamic range: 31 dB amplitude estimate: $\max(\|I\|,\|Q\|)+\frac{1}{2}\min(\|I\|,\|Q\|)$ update equation: $g(k+1)=g(k)+\mu g(k)e(k)$ digital implementation of AGC algorithm; analog VGA with discrete gain settings	
data recovery	adaptive MMSE detector	

Table 5.1: Summary of receiver specifications.

between the various receiver blocks, such as the LNA and the mixer. In addition, all of the analog front-end impairments described in Chapter 3 can potentially degrade the performance of the MMSE multiuser detection algorithm. The next section determines the effects of these analog impairments using the system-level simulation framework described in Chapter 4.

5.4 System Simulation

The adaptive MMSE multiuser detection algorithm described in Section 5.3.12 provides a good compromise between MAI suppression and noise enhancement. However, the performance of this algorithm may be compromised by the analog impairments introduced by the receiver front-end. In addition to noise, these impairments also include receiver distortion, gain mismatch, quadrature phase mismatch, and LO phase noise. The system-level simulation framework described in Chapter 4 is used to explore the tradeoffs between these analog impairments and overall system performance. The system downlink is simulated in Simulink and a top-level schematic is illustrated in Fig. 5.29.

5.4.1 Base-Station Transmitter

The base-station transmitter consists of a digital section and an analog section. Since this research focuses primarily on the design and implementation of the receiver, the simulation does not include any transmitter impairments. The digital section of the transmitter implements the QPSK modulation, signal spreading, power control, and pulse shaping described in Sections 5.2.1 and 5.2.2. For this particular simulation, the base-station transmitter supports up to 15 channels, one of which is a pilot channel as

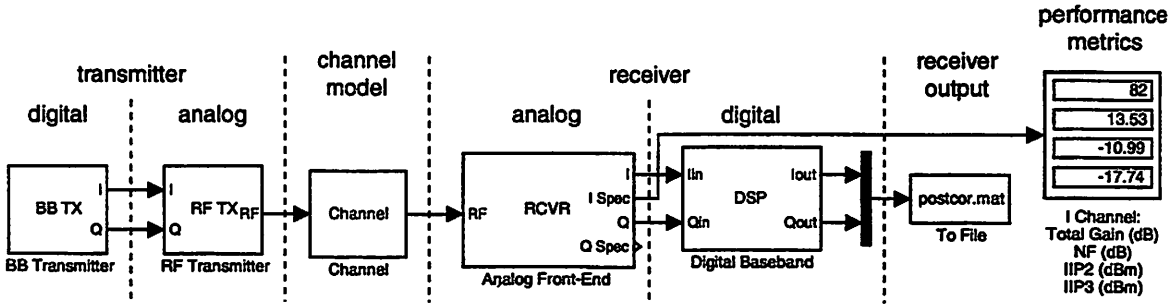


Figure 5.29: Top-level schematic of system downlink simulation.

described in Section 5.2.3.

As described in Section 5.2.5, digital-to-analog conversion and translation of the baseband I and Q signals to the 2-GHz carrier frequency are implemented in the analog section of the transmitter. The signal appearing at the output of an ideal transmitter is

$$s(t) = s_{I1}(t)\cos(\omega_c t) + s_{Q1}(t)\sin(\omega_c t) \tag{5.118}$$

where $s_{I1}(t)$ and $s_{Q1}(t)$ are the baseband I and Q signals and ω_c is the carrier frequency. In order to decrease the simulation time, the simulation framework relies on baseband-equivalent behavioral models for the receiver RF components. The inputs to these baseband-equivalent models are the time-varying coefficients of the equation

$$s(t) = s_{DC}(t) + \sum_{n=1}^{3}[s_{In}(t)\cos(n\omega_c t) + s_{Qn}(t)\sin(n\omega_c t)]. \tag{5.119}$$

Consequently, the outputs of the transmitter block in the simulation are simply $s_{I1}(t)$ and $s_{Q1}(t)$, and frequency translation to the carrier frequency is unnecessary. Finally, the analog section of the transmitter also restricts the total transmit power to 1 mW or 0 dBm.

5.4.2 Channel Model

For this simulation, the channel block only models the attenuation due to free-space propagation and shadowing as described in Section 5.3.1. However, a more complex model which includes other effects such as multipath propagation can be easily incorporated into the channel block. For example, if the transmitted signal is given by (5.118), then multipath propagation results in the following signal appearing at the receiver:

$$\begin{aligned} r(t) &= \sum_n \alpha_n(t) s(t-\tau_n) \\ &= \sum_n \alpha_n(t)\{s_{I1}(t-\tau_n)\cos[\omega_c(t-\tau_n)] + s_{Q1}(t-\tau_n)\sin[\omega_c(t-\tau_n)]\} \end{aligned} \tag{5.120}$$

where $\alpha_n(t)$ and τ_n are the attenuation factor and propagation delay, respectively, for the signal received on the n^{th} path. A baseband-equivalent model for multipath propagation

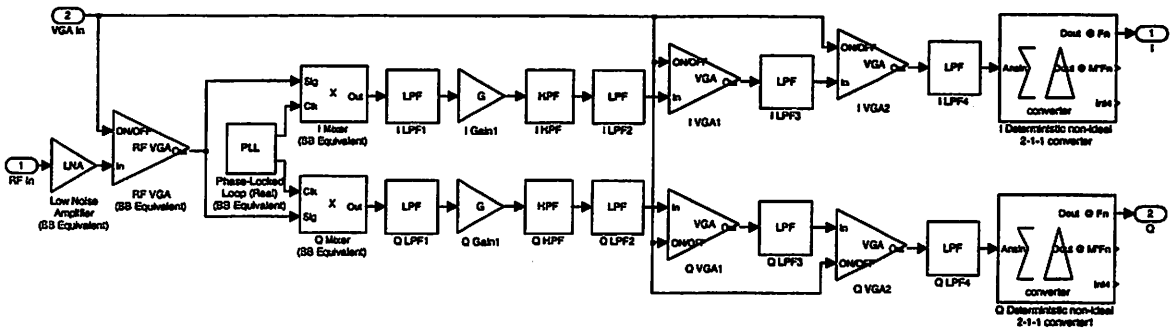


Figure 5.30: Receiver front-end schematic.

which can be incorporated into the proposed simulation framework is derived by expressing (5.120) in the form of (4.29):

$$r(t) = \sum_n \alpha_n(t)[\cos(\omega_c\tau_n)s_{I1}(t-\tau_n) - \sin(\omega_c\tau_n)s_{Q1}(t-\tau_n)]\cos(\omega_c t) + \alpha_n(t)[\cos(\omega_c\tau_n)s_{Q1}(t-\tau_n) + \sin(\omega_c\tau_n)s_{I1}(t-\tau_n)]\sin(\omega_c t). \quad (5.121)$$

5.4.3 Mobile Receiver

The mobile receiver consists of an analog section and a digital section. The analog section models the front-end of the direct-conversion receiver and is illustrated in Fig. 5.30. Simulation of the high-frequency components, such as the RF amplifiers, the PLL, and the I and Q mixers, relies on the baseband-equivalent behavioral models described in Section 4.3. The simulation also models all of the amplification and filtering in the baseband portion of the analog front-end and includes a first-order high-pass filter for eliminating dc offsets. In addition, the analog section includes a structural model for the two I and Q ADCs, each of which is a 7-bit 25-MS/s ΣΔ converter operating at 200 MHz.

The digital section of the receiver performs data recovery and includes a structural model of the adaptive multiuser detection (MUD) algorithm described in Section 5.3.12. For this simulation, the multiuser detector has a diversity order of two, providing increased robustness against fading due to multipath propagation.

5.4.4 Simulation Outputs

Overall system performance is determined by evaluating the I and Q outputs from the digital section of the receiver. With a small target BER of 10^{-4}, determining the BER directly would require simulating a large number of data bits, resulting in very long simulation times. A much better approach is to infer the average BER from an estimate of the output SNR, which requires the simulation of much fewer data bits. Finally, the simulation also provides conventional receiver performance metrics for the analog front-end, including total gain, noise figure, input IP_2, and input IP_3.

For this simulation, the transmitter output signal consists of ten equal-power data channels, including the pilot channel. The receiver specifications for this simulation are summarized in Table 5.2. The overall cascaded double-sideband (DSB) noise figure of the receiver is 13.5 dB. Since the noise performance of the receiver is most critical when the received signal is very weak, the transmitted signal experiences the worst-case channel attenuation of 69.43 dB and the receiver gain is set to the maximum level of 82 dB. This simulation accounts for other receiver impairments, including a gain mismatch of 4% between the I and Q signal paths, a quadrature phase mismatch of 2.5°, as well as PLL phase noise. PLL phase noise is modeled by the simple behavioral model described in Section 4.4.5 and is specified to be –80 dBc/Hz at a 100-kHz offset. The overall cascaded input IP_2 and input IP_3 are –11.0 dBm and –17.7 dBm, respectively, while the 1-dB compression point of the receiver is estimated from the specification for

center frequency	2 GHz
noise figure (DSB)	13.5 dB
gain	82 dB
I/Q gain mismatch	4%
PLL phase noise	–80 dBc/Hz @ 100 kHz
I/Q phase mismatch	2.5°
IIP_2	–11.0 dBm
IIP_3	–17.7 dBm
P_{-1dB}	–27.3 dBm
HPF corner frequency	500 kHz
ADC	7-bit, 25-MS/s ΣΔ
data recovery	adaptive MUD with second-order diversity

Table 5.2: Receiver specifications for system-level simulation.

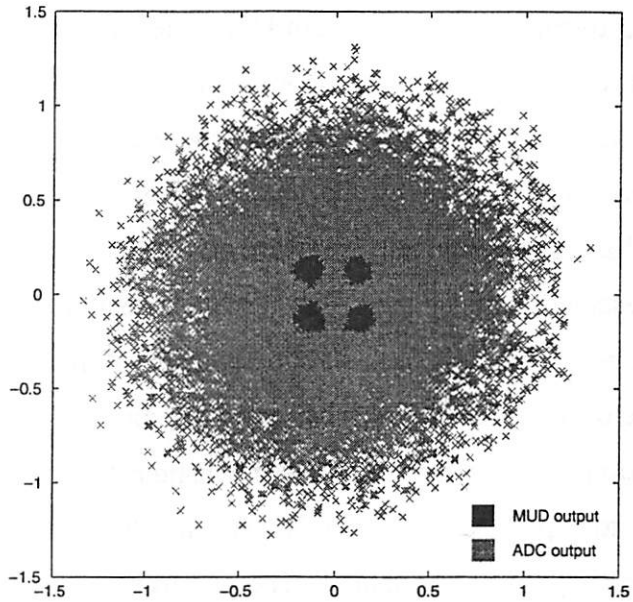


Figure 5.31: Constellation diagrams from system-level simulation.

input IP_3 as given by (4.20). The simulation includes a pair of first-order, 500-kHz high-pass filters for dc-offset removal as well as a pair of 7-bit, 25-MS/s ΣΔ modulators for analog-to-digital conversion.

Fig. 5.31 illustrates the constellation diagrams for the I and Q signals at the output of the ΣΔ ADCs and at the output of the multiuser detector. The SNR of the I and Q data from the output of the multiuser detector is approximately 15 dB, which corresponds to an average BER of 10^{-4} for this system. Thus, the target BER is achieved despite the relaxed receiver performance specifications as indicated by the constellation diagram for the I and Q data from the output of the ADCs. Indeed, most of the receiver specifications listed in Table 5.2 can be easily achieved in a highly-integrated CMOS implementation. Although a couple of the specifications, such as the maximum gain and the ADC requirements, are not as easily achievable, by applying low-power design techniques for these receiver circuits, the proposed system is still quite amenable to a low-power single-chip solution.

5.5 Summary

The exponential improvements in mainstream CMOS technology are clearly facilitating the implementation of advanced digital communications algorithms. However, the potential performance improvements may not be realized if these algorithms are very sensitive to impairments introduced by the analog front-end of the receiver. A high-performance WCDMA system which is relatively insensitive to analog front-end impairments was presented in this chapter. The system is designed to be used in an indoor picocellular environment, and each base station supports as many as 15 data channels, each with a data rate of up to 3.33 Mb/s. The design of this system relied heavily on the system-level simulation environment described in Chapter 4. This simulation framework allows the designer to rapidly and efficiently evaluate the affects of analog front-end impairments on overall system performance. The proposed system relies on an adaptive MUD algorithm for data recovery and the analog front-end of the receiver is based on a very simple direct-conversion architecture. Most of the receiver specifications are not very stringent and can be easily achieved in a low-power, highly-integrated CMOS implementation.

Chapter 6

Receiver Prototype

6.1 Introduction

By adhering to a design strategy which tightly incorporates implementation issues at the system level, many of the analog hardware requirements are relaxed while still achieving excellent overall system performance. A power-efficient solution is achieved by taking advantage of these relaxed requirements along with low-power circuit implementation techniques. The direct-conversion receiver is integrated onto a single chip and implements all analog receiver functions except for variable gain amplification (Fig. 6.1). All circuits on this chip use a 2.5-V supply, and a fully-differential signal path is used to mitigate the coupling between different receiver components. The LNA is capacitively coupled to the RF ports of the I and Q mixers, while the frequency synthesizer connects directly to the LO ports. Along each baseband signal path, a high-pass filter is used to eliminate dc offsets, while large transistor sizes are used to minimize the flicker noise contribution of the baseband circuits. In addition, the baseband signal paths provide moderate amplification as well as low-pass filtering before digitization of the I and Q signals. This direct-conversion receiver was fabricated in a 0.25-μm, single-poly, 6-metal CMOS process. The rest of this chapter describes the design and implementation of each

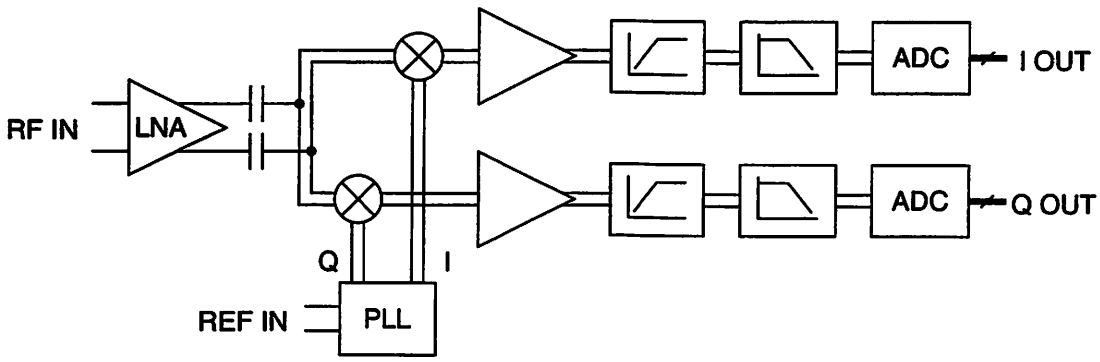


Figure 6.1: Block diagram of receiver prototype.

circuit block, focusing primarily on design choices which result in the most power-efficient implementation.

6.2 Low-Noise Amplifier

The LNA is one of the first components along the received signal path and its design must be considered in conjunction with the components which precede it, including the antenna and RF filter (Fig. 6.2). The antenna receives electromagnetic waves from the wireless transmission environment, and although the antenna usually has a tuned frequency response, the signal at its output consists of the desired signal as well as potentially strong out-of-band interferers. The RF filter immediately following the antenna helps to attenuate these out-of-band signals, while the subsequent LNA amplifies the received signal. In addition, the noise contribution of the LNA must be sufficiently low so as not to corrupt the potentially weak desired signal.

Traditional implementations require that the impedances at each of the component interfaces be 50 Ω and the usual explanation for this requirement is the desire for

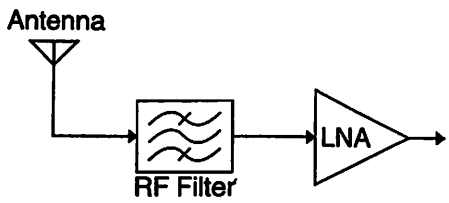


Figure 6.2: Antenna, RF filter, and LNA.

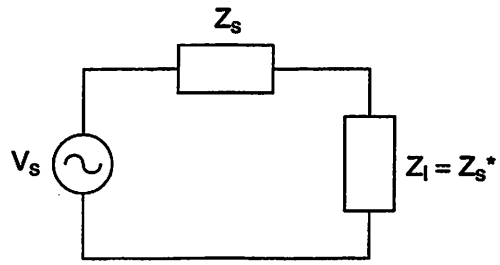


Figure 6.3: Conjugate impedance matching for maximum power transfer.

maximum power transfer. More precisely, maximum power transfer requires conjugate impedance matching between the source and load as illustrated in Fig. 6.3, and the 50-Ω requirement is a legacy from microwave designs using coaxial cables, where the 50-Ω interface resistance is a compromise between the 30-Ω resistance for maximum power handling and the 77-Ω resistance for minimum loss [25], [59]. Integrated-circuit implementations have already abandoned this antiquated requirement, and more recently, the 50-Ω requirement at the interface between external and on-chip components, e.g., between the external RF filter and the on-chip LNA, has also come under intense scrutiny. In order to clarify the need for a well-defined LNA input impedance, a review of microwave filter design follows.

6.2.1 Microwave Filter Design

The insertion loss method is a very common approach in microwave filter design, where the filter response is characterized by its insertion loss, or power loss ratio [59]:

$$P_L = \frac{1}{1-|\Gamma(\omega)|^2} \tag{6.1}$$

where $\Gamma(\omega)$ is the reflection coefficient as a function of frequency ω. Since $|\Gamma(\omega)|^2$ is an even function of ω, it can be expressed as a polynomial in ω^2:

$$|\Gamma(\omega)|^2 = \frac{M(\omega^2)}{M(\omega^2)+N(\omega^2)}. \tag{6.2}$$

Consequently, (6.1) may be expressed as

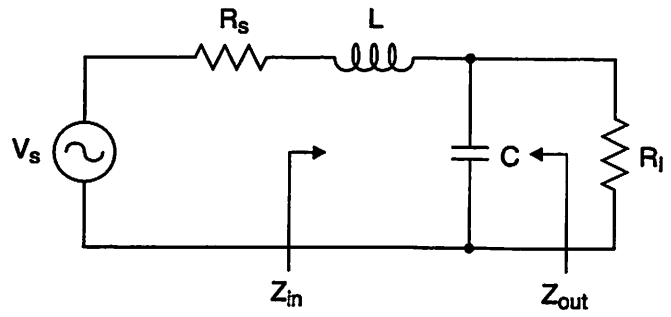


Figure 6.4: Circuit for second-order Butterworth low-pass filter.

$$P_L = 1 + \frac{M(\omega^2)}{N(\omega^2)}. \tag{6.3}$$

The power loss ratio may be specified for various filter responses. For example, the power loss ratio for a Butterworth low-pass filter response is given by

$$P_{L,Butterworth} = 1 + \varepsilon^2 \left(\frac{\omega}{\omega_c} \right)^{2N} \tag{6.4}$$

where ε determines the magnitude variation in the passband, ω_c is the passband edge, and N is the filter order. Consider the design of a second-order Butterworth low-pass filter based on a single LC section illustrated in Fig. 6.4. In this case, the filter is doubly terminated with a load resistor R_l at the output and a source resistor R_s at the input. If the –3-dB frequency is ω_c, then the desired power loss ratio is

$$P_{L,Butterworth} = 1 + \left(\frac{\omega}{\omega_c} \right)^4. \tag{6.5}$$

The input and output impedances of the filter are, respectively,

$$Z_{in} = j\omega L + \frac{R_l}{1 + j\omega R_l C} \tag{6.6}$$

$$Z_{out} = \frac{R_s + j\omega L}{1 - \omega^2 LC + j\omega R_s C} \tag{6.7}$$

and the reflection coefficient is

$$\Gamma(\omega) = \frac{Z_{in} - R_s}{Z_{in} + R_s}. \tag{6.8}$$

Substituting (6.6) and (6.8) into (6.1),

$$P_L = \left(\frac{1}{2} + \frac{R_l}{4R_s} + \frac{R_s}{4R_l}\right) + \left(\frac{L^2}{4R_lR_s} - \frac{R_lLC}{2R_s} + \frac{1}{4}R_sR_lC^2\right)\omega^2 + \frac{R_lL^2C^2}{4R_s}\omega^4. \tag{6.9}$$

By equating (6.5) and (6.9), the following component values are required in order to achieve the desired second-order Butterworth low-pass frequency response:

$$R_l = R_s \tag{6.10}$$

$$L = \frac{\sqrt{2}R_s}{\omega_c} \tag{6.11}$$

$$C = \frac{\sqrt{2}}{R_s\omega_c}. \tag{6.12}$$

In this case, the load resistance must equal the source resistance in order to achieve the desired filter response, i.e., the filter will not function properly unless this condition is satisfied. A similar analysis can be applied to other filter responses. For Bessel and odd-order Chebyshev responses, the load and source resistances must also be equal, while for an even-order Chebyshev response, the load and source resistances are related but unequal [59].

In Fig. 6.2, the antenna and LNA present source and load impedances, respectively, to the RF filter. In order to design the RF filter for a particular frequency response, the source and load impedances of the antenna and LNA, respectively, must be known a priori. Since RF filters are usually designed independently from the antenna and the LNA, a standard impedance must be chosen. Commercially-available filters are typically designed assuming 50-Ω source and load impedances, and consequently, deviating from 50 Ω results in poor and unpredictable RF filter performance.

Since the receiver prototype is intended to be used with a commercially-available RF filter, the on-chip LNA must be designed to have a 50-Ω input impedance. Future designs will rely on both a custom RF filter and a custom antenna so that the 50-Ω input

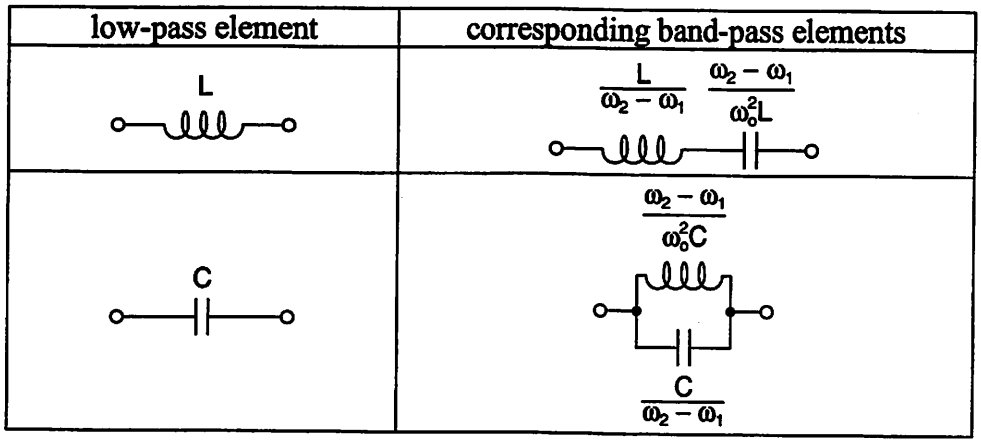


Figure 6.5: Transformation of a low-pass response to a band-pass response.

impedance requirement of the LNA can be removed. The implications of removing this constraint are explored in Appendix C.

Finally, the RF filter should actually have a band-pass response. A low-pass response can be transformed to a band-pass response by applying the transformations illustrated in Fig. 6.5.

6.2.2 LNA Performance Metrics

The main function of the LNA is to amplify the potentially weak desired signal without corrupting it through mechanisms such as noise or distortion. Since the linearity performance of receivers is usually limited by components following the LNA, such as the mixer, the distortion performance of the LNA is usually not very stringent. However, the noise contribution of the LNA must be sufficiently low so as not to corrupt the potentially weak desired signal at its input.

The noise performance of an LNA may be characterized by a couple of different metrics: noise factor and noise measure. Noise factor or noise figure is the most common metric and is a measure of how much the LNA degrades the SNR of the received signal. Noise factor is defined as

$$F \equiv \frac{SNR_{in}}{SNR_{out}} \tag{6.13}$$

where SNR_{in} and SNR_{out} are the SNRs at the input and output, respectively, of the LNA. A related metric, noise figure, is simply the noise factor expressed in decibels, $10\log(F)$. One potential drawback of this metric is that it does not account for amplification. For example, an ideal wire is obviously not a very good LNA since it does not provide any amplification, although it does have an excellent noise figure of zero. Consequently, specifying the noise factor or noise figure of an LNA is meaningless without also specifying its gain.

A metric less commonly used to characterize the noise performance of an LNA is noise measure. Noise measure accounts for both the noise and gain of the LNA and is defined as [60]

$$M \equiv \frac{F-1}{1-1/G} \tag{6.14}$$

where F and G are the noise factor and power gain, respectively, of the LNA. In the case of an ideal wire, the noise measure is infinite and is consistent with our notion that a wire in not a very good LNA. The power gains of practical LNA topologies are sufficiently large, e.g., $G > 10$, so that noise measure and noise factor become equivalent metrics. Consequently, only noise factor is evaluated for the LNA topologies described later in Section 6.2.5.

6.2.3 Transistor Noise Model

Regardless of which metric is preferable, the key design goal is to minimize the noise figure while maximizing the gain of the LNA. Since the noise performance of the LNA is so critical, an accurate transistor noise model is essential. In particular, the measured thermal noise in short-channel MOS devices is greater than the amount predicted by long-channel theory [61]–[63]. This section begins with a review of the long-channel MOS noise model followed by a discussion of some recently proposed noise models for short-channel MOS devices.

Long-Channel MOS Noise Model. Since the channel material of an MOS device is resistive, the drain current exhibits thermal noise. For an MOS transistor operating in

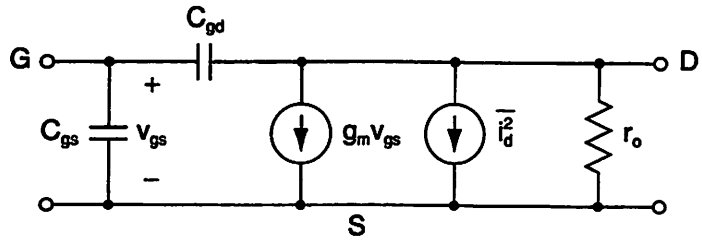


Figure 6.6: MOS small-signal equivalent circuit with noise generators.

strong inversion, the small-signal equivalent circuit with noise generators is illustrated in Fig. 6.6. According to long-channel theory, the power spectral density of the drain current noise may be expressed as [10]

$$S_{I_D}(f) = \frac{\overline{i_d^2}}{\Delta f} = 4kT\frac{2}{3}g_m \tag{6.15}$$

where g_m is the device transconductance. Equation (6.15) assumes that the device is operating at frequencies well above the flicker noise corner frequency so that flicker noise may be neglected. This model significantly underestimates the actual noise present in short-channel MOS devices. Recently, more accurate noise models have been proposed for short-channel MOS devices. Two proposed mechanisms resulting in the observed excess thermal noise include high-field effects [65] and induced gate current noise [64], [65].

MOS Noise Model including High-Field Effects. The first modification to the traditional MOS noise model is an increased drain current noise resulting from high-field effects in short-channel devices. In this case, the power spectral density of the drain current noise is expressed instead as

$$S_{I_D} = 4kT\gamma g_{do} \tag{6.16}$$

where γ is a bias-dependent parameter used to account for the increased drain current noise and g_{do} is the zero-bias drain conductance of the device. For long-channel devices in strong inversion, γ is equal to 2/3 and g_{do} is equal to the device transconductance, g_m,

$$g_{do} = g_m = \mu C_{ox}\frac{W}{L}(V_{GS} - V_t) \tag{6.17}$$

so (6.16) reduces to (6.15). For short-channel devices, γ may be as high as two to three and may be attributed to hot electron effects [65]. Under high electric fields, the temperature of electrons in the channel can rise above that of the lattice, resulting in an increase in the drain current noise. In this case, the power spectral density of the drain current noise is given by [65]

$$S_{I_D} = \frac{4kT}{L^2 I_D} \int_0^{V_d} \frac{T_e}{T} g^2(V) dV \tag{6.18}$$

where $g(V)$ is the channel conductance at a given point along the channel, V is the corresponding voltage, and T_e and T are the electron and lattice temperatures, respectively. Equating (6.16) and (6.18) results in the following expression for γ:

$$\gamma = \frac{1}{g_{do} L^2 I_D} \int_0^{V_d} \frac{T_e}{T} g^2(V) dV . \tag{6.19}$$

Including the effects of mobility degradation, the drain current I_D is given by

$$\begin{aligned} I_D &= \mu_{eff} C_{ox} W (V_{GS} - V_t - V) \frac{E(y)}{1 + \dfrac{E(y)}{E_{sat}}} \\ &= \left[\mu_{eff} C_{ox} W (V_{GS} - V_t - V) - \frac{I_D}{E_{sat}} \right] \frac{dV}{dy}. \end{aligned} \tag{6.20}$$

However, I_D may also be expressed in terms of $g(V)$:

$$I_D = g(V) E(y) = g(V) \frac{dV}{dy}. \tag{6.21}$$

Equating (6.20) and (6.21) results in the following expression for $g(V)$:

$$g(V) = \mu_{eff} C_{ox} W (V_{GS} - V_t - V) - \frac{I_D}{E_{sat}}. \tag{6.22}$$

Next, in order to evaluate the integral in (6.19), an expression for T_e/T is also required. Unfortunately, the exact dependence of electron temperature on electric field strength is unknown. For this calculation, it is assumed that

$$\frac{T_e}{T} = \left[1 + \frac{E(y)}{E_{sat}}\right]^2 = \left[1 + \frac{I_D}{g(V)E_{sat}}\right]^2. \tag{6.23}$$

Substituting (6.22) and (6.23) into (6.19) gives

$$\gamma = \frac{1}{g_{do}L^2 I_D}\int_0^{V_D} \mu_{eff}^2 C_{ox}^2 W^2 (V_{GS} - V_t - V)^2 dV = \frac{\mu_{eff}^2 C_{ox}^2 W^2 V_D}{g_{do}L^2 I_D}\left[(V_{GS} - V_t)^2 - (V_{GS} - V_t)V_D + \frac{1}{3}V_D^2\right]. \tag{6.24}$$

In strong inversion [66],

$$I_D = I_{Dsat} = \frac{1}{2}\mu_{eff} C_{ox} W E_{sat} \frac{(V_{GS} - V_t)^2}{(V_{GS} - V_t) + E_{sat}L} \tag{6.25}$$

$$V_D = V_{Dsat} = \frac{(V_{GS} - V_t)E_{sat}L}{(V_{GS} - V_t) + E_{sat}L}. \tag{6.26}$$

In addition, for short-channel devices, g_{do} is given by

$$g_{do} = \frac{d}{dV_D}\left\{\frac{\mu_{eff} C_{ox}\frac{W}{L}}{1 + \frac{V_D}{E_{sat}L}}\left[(V_{GS} - V_t)V_D - \frac{1}{2}V_D^2\right]\right\}\Bigg|_{V_D=0} = \mu_{eff} C_{ox}\frac{W}{L}(V_{GS} - V_t). \tag{6.27}$$

Substituting (6.25), (6.26), and (6.27) into (6.24) results in the following expression for γ:

$$\gamma = \frac{1}{[(V_{GS} - V_t) + E_{sat}L]^2}\left[\frac{2}{3}(E_{sat}L)^2 + 2(V_{GS} - V_t)E_{sat}L + 2(V_{GS} - V_t)^2\right]. \tag{6.28}$$

For long-channel devices, $E_{sat}L$ is much larger than $V_{GS} - V_t$, and (6.28) reduces to $\gamma = 2/3$. However, in the limit that $E_{sat}L$ is much smaller than $V_{GS} - V_t$, (6.28) reduces to $\gamma = 2$.

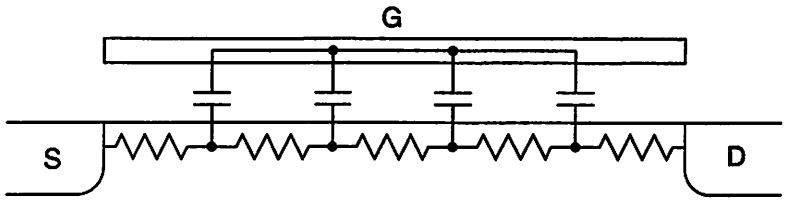


Figure 6.7: Distributed gate capacitance and channel resistance at high frequencies.

Induced Gate Current Noise. At high frequencies, induced gate noise becomes significant, which arises from the distributed nature of the device as illustrated in Fig. 6.7. In this case, the gate admittance consists of an additional conductive component [65]:

$$Y_g = j\omega C_{gs} + g_g \tag{6.29}$$

where C_{gs} and g_g are given by, respectively,

$$C_{gs} = \frac{2}{3} WLC_{ox} \tag{6.30}$$

$$g_g = \frac{\omega^2 C_{gs}^2}{5 g_{do}}. \tag{6.31}$$

Since g_g is a physical resistance, it has an associated noise current with power spectral density given by

$$S_{I_G} = \frac{\overline{i_g^2}}{\Delta f} = 4kT\delta g_g \tag{6.32}$$

where δ is 4/3 for long-channel devices. For short-channel devices where high-field effects may be significant, the power spectral density of the gate current noise is given by [65]

$$S_{I_G} = \frac{4kT\omega^2 C_{ox}^2 W^2}{I_D^3} \int_0^{V_D} \frac{T_e}{T} g^2(V)(V_a - V)^2 dV \tag{6.33}$$

where

$$V_a = V_{Dsat} + \frac{V_{Dsat}^2}{2E_{sat}L} - \frac{(V_{GS} - V_t)\frac{V_{Dsat}}{2} - \frac{V_{Dsat}^2}{6}}{V_{GS} - V_t - \frac{V_{Dsat}}{2}} \left(1 + \frac{V_{Dsat}}{E_{sat}L}\right). \tag{6.34}$$

Equating (6.32) and (6.33) results in the following expression for δ:

$$\delta = \frac{45 g_{do}}{4L^2 I_D^3} \int_0^{V_D} \frac{T_e}{T} g^2(V)(V_a - V)^2 dV . \tag{6.35}$$

Assuming that T_e/T is given by (6.23), (6.35) becomes

$$\delta = \frac{1}{[(V_{GS} - V_t) + E_{sat}L]^4} \left[\frac{4}{3}(E_{sat}L)^4 + \frac{17}{2}(E_{sat}L)^3 (V_{GS} - V_t) + 23(E_{sat}L)^2 (V_{GS} - V_t)^2 + \frac{45}{2} E_{sat}L (V_{GS} - V_t)^3 + \frac{15}{2}(V_{GS} - V_t)^4 \right]. \tag{6.36}$$

For long-channel devices, $E_{sat}L$ is much larger than $V_{GS} - V_t$, and (6.36) reduces to $\delta = 4/3$, as expected. However, in the limit that $E_{sat}L$ is much smaller than $V_{GS} - V_t$, (6.36) reduces to $\delta = 15/2$, which is more than five times larger than the long-channel limit. Consequently, for short-channel devices operating at very high frequencies, induced gate current noise can be quite detrimental to low noise performance.

Finally, since the induced gate current noise originates from the distributed nature of the gate capacitance and the channel resistance, it is partially correlated with the drain current noise with a correlation coefficient given by

$$c = \frac{\overline{i_g i_d^*}}{\sqrt{\overline{i_g i_g^*} \times \overline{i_d i_d^*}}} \tag{6.37}$$

where

$$\overline{i_d i_d^*} = 4kT\gamma g_{do} \Delta f \tag{6.38}$$

$$\overline{i_g i_g^*} = 4kT\delta \frac{\omega^2 C_{gs}^2}{5 g_{do}} \Delta f \tag{6.39}$$

$$\overline{i_g i_d^*} = 4kT\varepsilon j\omega C_{gs} \Delta f . \tag{6.40}$$

Consequently, (6.37) may be expressed as

$$c = j\sqrt{\frac{5}{\gamma\delta}}\varepsilon \tag{6.41}$$

where γ and δ are given by (6.19) and (6.35), respectively. Equating (6.40) with

$$\overline{i_g i_d^*} = 4kT\Delta f \frac{j\omega C_{ox} W}{L I_D^2} \int_0^{V_D} \frac{T_e}{T} g^2(V)(V_a - V)dV \tag{6.42}$$

where V_a is given by (6.34), results in the following expression for ε:

$$\varepsilon = \frac{3}{2L^2 I_D^2} \int_0^{V_D} \frac{T_e}{T} g^2(V)(V_a - V)dV . \tag{6.43}$$

Assuming that T_e/T is given by (6.23), ε becomes

$$\varepsilon = \frac{(E_{sat}L)^2}{[(V_{GS} - V_t) + E_{sat}L]^3} \left[\frac{1}{6} E_{sat}L + \frac{1}{2}(V_{GS} - V_t) \right]. \tag{6.44}$$

For long-channel devices, $E_{sat}L$ is much larger than $V_{GS} - V_t$, and (6.44) reduces to $\varepsilon = 1/6$. Consequently, the correlation coefficient in (6.41) is

$$c \leq j0.395 \tag{6.45}$$

where the equality holds for long-channel devices.

For an MOS transistor operating in strong inversion, the revised small-signal equivalent circuit with noise generators is illustrated in Fig. 6.8. When using this model for noise calculations, the correlation between the gate current noise and the drain current noise must be taken into account.

6.2.4 Matching for Minimum Noise Figure

Minimizing the noise figure is one of the key design goals when designing an LNA. For a

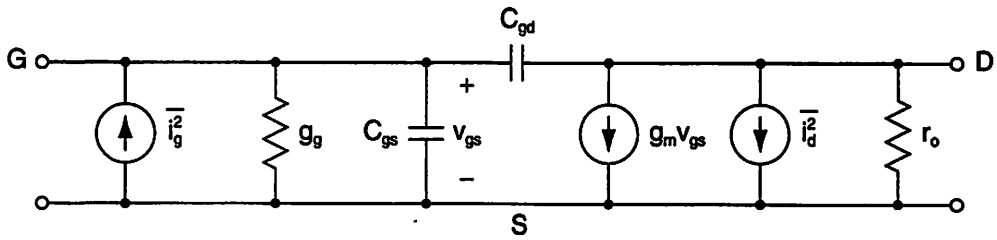
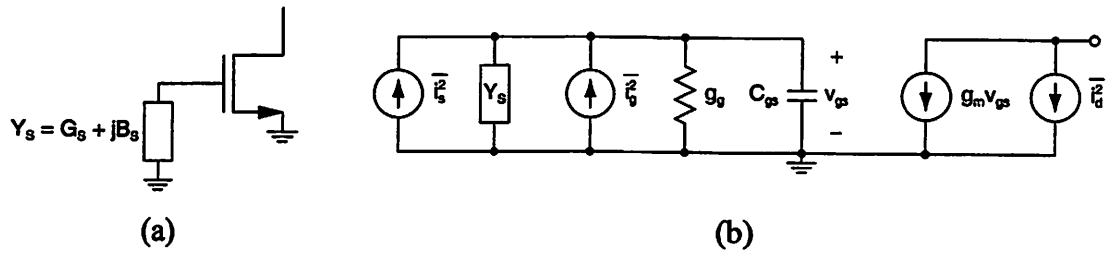


Figure 6.8: Revised MOS small-signal equivalent circuit with noise generators.



(a) (b)

Figure 6.9: (a) Common-source amplifier. (b) Small-signal equivalent circuit.

particular LNA topology, an optimum source admittance exists for minimum noise figure. A common-source transistor amplifier driven by a source admittance $Y_s = G_s + jB_s$ is illustrated in Fig. 6.9a and the corresponding small-signal equivalent circuit with noise generators is illustrated in Fig. 6.9b. For this example, C_{gd} and r_o are ignored in order to simplify the calculation, while the thermal noise current due to the conductive component of the source admittance is given by

$$\overline{i_s^2} = 4kTG_s\Delta f\,. \tag{6.46}$$

The current noise components appearing at the drain due to i_s, i_g, and i_d are given by, respectively,

$$i_{os} = \frac{g_m i_s}{g_g + G_s + j(\omega C_{gs} + B_s)} \tag{6.47}$$

$$i_{og} = \frac{g_m i_g}{g_g + G_s + j(\omega C_{gs} + B_s)} \tag{6.48}$$

$$i_{od} = i_d \tag{6.49}$$

and the noise factor is

$$F = 1 + \frac{VAR[i_{og} + i_{od}]}{\overline{i_{os}^2}} \tag{6.50}$$

where for two zero-mean random variables, A and B, the sum of the two random variables has variance

$$VAR[A+B] = VAR[A] + VAR[B] + E[AB^*] + E[A^*B]\,. \tag{6.51}$$

This formulation is necessary because of the correlation between the gate current noise and the drain current noise. Consequently, the noise factor is

$$F = 1 + \frac{\overline{i_{og}^2} + \overline{i_{od}^2} + \overline{i_{og} i_{od}^*} + \overline{i_{og}^* i_{od}}}{\overline{i_{os}^2}} \tag{6.52}$$

where

$$\overline{i_{od}^2} = \overline{i_d^2} \tag{6.53}$$

$$\overline{i_{og}^2} = \frac{g_m^2 \overline{i_g^2}}{(g_g + G_s)^2 + (\omega C_{gs} + B_s)^2} \tag{6.54}$$

$$\overline{i_{os}^2} = \frac{g_m^2 \overline{i_s^2}}{(g_g + G_s)^2 + (\omega C_{gs} + B_s)^2} \tag{6.55}$$

$$\overline{i_{og} i_{od}^*} = \frac{g_m \overline{i_g i_d^*}}{g_g + G_s + j(\omega C_{gs} + B_s)} \tag{6.56}$$

$$\overline{i_{og}^* i_{od}} = \frac{g_m \overline{i_g^* i_d}}{g_g + G_s - j(\omega C_{gs} + B_s)}. \tag{6.57}$$

From (6.37),

$$\overline{i_g i_d^*} = \left(\overline{i_g^* i_d}\right)^* = c\sqrt{\overline{i_g i_g^*} \times \overline{i_d i_d^*}} = j\,|c|\sqrt{\overline{i_g i_g^*} \times \overline{i_d i_d^*}}\,. \tag{6.58}$$

The noise factor is then

$$F = 1 + \frac{\alpha\delta\omega^2 C_{gs}^2}{5 g_m G_s} + \frac{\gamma}{\alpha g_m G_s}[(g_g + G_s)^2 + (\omega C_{gs} + B_s)^2] + 2\,|c|\sqrt{\frac{\delta\gamma}{5}}\frac{\omega C_{gs}}{g_m G_s}(\omega C_{gs} + B_s) \tag{6.59}$$

where α is defined as

$$\alpha \equiv \frac{g_m}{g_{do}} = \frac{E_{sat} L[(V_{GS} - V_t) + 2E_{sat} L]}{2[(V_{GS} - V_t) + E_{sat} L]^2} \le 1. \tag{6.60}$$

The noise factor is minimum when G_s and B_s are, respectively,

$$G_s = G_{sopt} = \sqrt{g_g^2 + \alpha^2 \omega^2 C_{gs}^2 \frac{\delta}{5\gamma}(1-|c|^2)} \tag{6.61}$$

$$B_s = B_{sopt} = -\omega C_{gs}\left(1+\alpha|c|\sqrt{\frac{\delta}{5\gamma}}\right) \tag{6.62}$$

and the corresponding noise factor is

$$F_{min} = 1+2\sqrt{\frac{\delta\gamma}{5}}\frac{\omega}{\omega_T}\sqrt{(1-|c|)^2+\frac{\gamma}{5\delta}\left(\frac{\omega}{\omega_T}\right)^2}+\frac{2\gamma}{5}\left(\frac{\omega}{\omega_T}\right)^2 \tag{6.63}$$

where

$$\omega_T = \frac{g_m}{C_{gs}}. \tag{6.64}$$

For $\omega << \omega_T$, (6.63) becomes

$$\begin{aligned} F_{min} &= 1+2\sqrt{\frac{\delta\gamma}{5}(1-|c|^2)}\frac{\omega}{\omega_T} \\ &= 1+2.32\frac{\omega}{\omega_T} \end{aligned} \tag{6.65}$$

where the latter equality assumes that $\gamma = 2$, $\delta = 4$, and $|c| = 0.395$.

Minimum noise figure for the common-source transistor amplifier is achieved with an optimum source conductance and susceptance given by (6.61) and (6.62), respectively, and consequently, for minimum noise figure, the RF filter which precedes this amplifier should have an output admittance equal to the required optimum noise admittance. At the same time, maximum power transfer from the RF filter to the amplifier requires a conjugate match between the output admittance of the RF filter and the input admittance of the amplifier, which is given by

$$Y_{in} = g_g + j\omega C_{gs}. \tag{6.66}$$

A simultaneous noise and power match requires that $Y_s = Y_{in}^*$, and comparing (6.61) and (6.62) with (6.66) reveals that such a match is impossible for this amplifier topology. For

a power match alone, the required source admittance is $Y_s = g_g - j\omega C_{gs}$ and the corresponding noise factor is

$$F = 1 + \delta + \frac{4\gamma}{5}\left(\frac{\omega}{\omega_T}\right)^2 \geq 5 \text{ or } 6.99\,\text{dB} \tag{6.67}$$

where the latter inequality assumes that $\omega << \omega_T$ and $\delta = 4$. Under this condition of maximum power transfer, the noise performance of the common-source transistor amplifier is rather poor.

6.2.5 LNA Topologies

Since the receiver prototype is intended to be used with a commercially-available RF filter, the on-chip LNA must be designed to have a 50-Ω input impedance. In this section, several potential LNA topologies are analyzed, including the common-source, common-gate, common-source with inductive degeneration, and local shunt feedback topologies.

Common-Source LNA. In order to achieve a 50-Ω input resistance, a slight modification is made to the common-source LNA already analyzed in Section 6.2.4. A shunt inductor is added to the input of the LNA in order to tune out the gate capacitance of the transistor (Fig. 6.10), where

$$B_s = -\frac{1}{\omega L_s} \tag{6.68}$$

and L_s is chosen such that

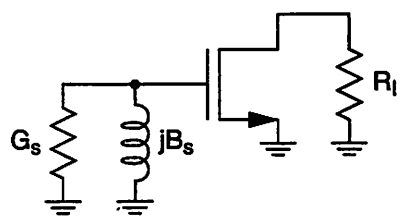


Figure 6.10: Common-source LNA with tuned input.

$$\omega^2 = \frac{1}{L_s C_{gs}}. \tag{6.69}$$

Under this condition, the input conductance of the LNA is $Y_{in} = g_g$ and a 50-Ω input impedance is achieved by setting $Y_{in} = g_g = 1/(50\,\Omega)$. Also, the output conductance of the RF filter is $G_s = 1/(50\,\Omega)$ and the corresponding noise factor is given by (6.67). Alternatively, a broadband input match to 50 Ω may be achieved by eliminating the shunt inductor at the LNA input and selecting

$$\begin{aligned} g_g &>> \omega C_{gs} \\ \omega &>> \frac{5\omega_T}{\alpha}. \end{aligned} \tag{6.70}$$

Under this condition, the noise factor is given by (6.67):

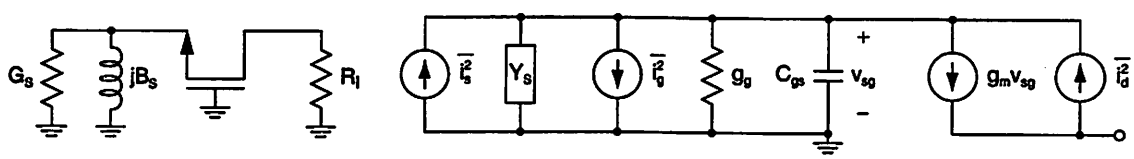
$$\begin{aligned} F &= 1 + \delta + \frac{4\gamma}{5}\left(\frac{\omega}{\omega_T}\right)^2 \\ &\geq 1 + \delta + \frac{20\gamma}{\alpha^2} \\ &\geq 67.5 \text{ or } 18.3\,\text{dB} \end{aligned} \tag{6.71}$$

where the latter inequality assumes that $\alpha = 0.8$, $\delta = 4$, and $\gamma = 2$. Consequently, a broadband input match is achieved at the expense of very poor noise performance.

Finally, the voltage gain of this LNA is

$$A_v = -g_m R_l \tag{6.72}$$

where R_l is the load resistance at the output of the LNA.



(a) (b)

Figure 6.11: (a) Common-gate LNA. (b) Small-signal equivalent circuit.

Common-Gate LNA. A second topology which may be used to achieve a 50-Ω input resistance is the common-gate LNA illustrated in Fig. 6.11a. From the small-signal equivalent circuit with noise generators illustrated in Fig. 6.11b, the noise factor is

$$F = 1 + \frac{\alpha\delta\omega^2 C_{gs}^2}{5 g_m G_s} + \frac{\gamma}{\alpha g_m G_s}[(g_g + G_s)^2 + (\omega C_{gs} + B_s)^2] + 2|c|\sqrt{\frac{\delta\gamma}{5}}\frac{\omega C_{gs}}{g_m G_s}(\omega C_{gs} + B_s)^2 \tag{6.73}$$

which is identical to the noise factor of the common-source LNA in (6.59). Consequently, the minimum noise factor for the common-gate LNA is also given by (6.65) and the corresponding source conductance and susceptance are given by (6.61) and (6.62), respectively.

In order to achieve a 50-Ω LNA input resistance, a shunt inductor is added at the input of the LNA in order to tune out the gate capacitance of the transistor (Fig. 6.11a). Under this condition, the input conductance of the LNA is $Y_{in} = g_g + g_m$ and a 50-Ω input resistance is achieved by setting $Y_{in} = g_g + g_m = 1/(50\,\Omega)$. Also, the output conductance of the RF filter is $G_s = 1/(50\,\Omega)$ and the corresponding noise factor is given by

$$\begin{aligned} F &= 1 + \frac{\frac{\gamma}{\alpha} + \frac{\delta\gamma}{5}\left(\frac{\omega}{\omega_T}\right)^2}{1 + \frac{\alpha}{5}\left(\frac{\omega}{\omega_T}\right)^2} + \frac{4\gamma}{5}\left(\frac{\omega}{\omega_T}\right)^2 \\ &\geq 1 + \frac{\gamma}{\alpha} = 3.5 \text{ or } 5.44\,\text{dB} \end{aligned} \tag{6.74}$$

where the latter inequality assumes that $\omega << \omega_T$, $\gamma = 2$, and $\alpha = 0.8$. Under this condition, the noise performance of the common-gate transistor amplifier is rather poor. Alternatively, a broadband input match to 50 Ω may be achieved by eliminating the shunt inductor at the LNA input and selecting

$$\begin{aligned} \omega C_{gs} &<< g_m \\ \omega &<< \omega_T. \end{aligned} \tag{6.75}$$

Under this condition, g_g is also much less than g_m, and consequently, the input admittance reduces to $Y_{in} = g_m$ and the noise factor is

$$\begin{aligned} F &= 1 + \frac{\gamma}{\alpha} + \frac{4\gamma}{5}\left(\frac{\omega}{\omega_T}\right)^2 \\ &\geq 1 + \frac{\gamma}{\alpha} = 3.5 \text{ or } 5.44\,\text{dB} \end{aligned} \tag{6.76}$$

where the latter inequality assumes that $\omega << \omega_T$, $\gamma = 2$, and $\alpha = 0.8$.

Finally, the voltage gain of the common-gate LNA is

$$A_v = g_m R_l \tag{6.77}$$

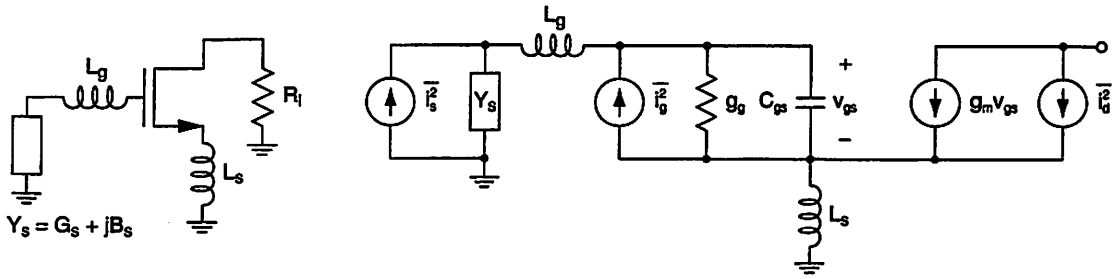
where R_l is the load resistance at the output of the LNA. When the LNA input resistance is matched to 50 Ω, the voltage gain is $A_v = g_m R_l = 0.02 R_l$.

Common-Source LNA with Inductive Degeneration. A third topology which may be used to achieve a 50-Ω input resistance is the common-source LNA with inductive degeneration illustrated in Fig. 6.12a [11], [64], while the corresponding small-signal equivalent circuit illustrated in Fig. 6.12b. In order to simplify the calculations, g_g is ignored. This approximation is valid when

$$\begin{aligned} g_g &<< \omega C_{gs} \\ \frac{\omega^2 C_{gs}^2}{5 g_{do}} &<< \omega C_{gs} \\ \omega &<< \frac{5\omega_T}{\alpha} \end{aligned} \tag{6.78}$$

where α is defined in (6.60). Indeed this condition is easily met in most designs where the devices are designed to operate at frequencies much less than ω_T. The input impedance of this LNA is

$$Z_{in} = \frac{g_m L_s}{C_{gs}} - j\frac{1 - \omega^2 C_{gs}(L_g + L_s)}{\omega C_{gs}}. \tag{6.79}$$



(a) (b)

Figure 6.12: (a) Common-source LNA with inductive degeneration. (b) Small-signal equivalent circuit.

A 50-Ω input resistance is achieved by equating the real part of Z_{in} to 50 Ω and then selecting the values of inductors L_g and L_s to tune out the imaginary part of Z_{in}:

$$\frac{g_m L_s}{C_{gs}} = 50\,\Omega \tag{6.80}$$

$$L_g + L_s = \frac{1}{\omega^2 C_{gs}}. \tag{6.81}$$

The noise factor of this LNA is

$$\begin{aligned} F = 1 + {} & \frac{\alpha\delta\omega^2 C_{gs}^2}{5 g_m G_s}\{[1-\omega B_s(L_g+L_s)]^2 + \omega^2 G_s^2 (L_g+L_s)^2\} + \\ & \frac{\gamma}{\alpha g_m G_s}\{G_s^2[1-\omega^2 C_{gs}(L_g+L_s)]^2 + [\omega C_{gs} + B_s[1-\omega^2 C_{gs}(L_g+L_s)]]^2\} + \\ & 2|c|\sqrt{\frac{\delta\gamma}{5}}\frac{\omega C_{gs}}{g_m G_s}\{[1-\omega B_s(L_g+L_s)][\omega C_{gs} + B_s[1-\omega^2 C_{gs}(L_g+L_s)]] - \\ & \omega(L_g+L_s)G_s^2[1-\omega^2 C_{gs}(L_g+L_s)]\}. \end{aligned} \tag{6.82}$$

Substituting (6.81) into (6.82) results in the following expression for noise factor:

$$\begin{aligned} F = 1 + {} & \frac{\alpha\delta\omega^2 C_{gs}^2}{5 g_m G_s}\left[\frac{G_s^2}{\omega^2 C_{gs}^2} + \left(1 - \frac{B_s}{\omega C_{gs}}\right)^2\right] + \frac{\gamma\omega^2 C_{gs}^2}{\alpha g_m G_s} + \\ & 2|c|\sqrt{\frac{\delta\gamma}{5}}\frac{\omega^2 C_{gs}^2}{g_m G_s}\left(1 - \frac{B_s}{\omega C_{gs}}\right). \end{aligned} \tag{6.83}$$

The noise factor is minimum when G_s and B_s are, respectively,

$$G_s = G_{sopt} = \frac{\omega C_{gs}}{\alpha}\sqrt{\frac{5\gamma}{\delta}(1-|c|^2)} \tag{6.84}$$

$$B_s = B_{sopt} = \omega C_{gs}\left(1+\frac{|c|}{\alpha}\sqrt{\frac{5\gamma}{\delta}}\right) \tag{6.85}$$

and the corresponding noise factor is

$$F_{min} = 1+2\sqrt{\frac{\delta\gamma}{5}(1-|c|)^2}\,\frac{\omega}{\omega_T} \tag{6.86}$$

which is identical to the minimum noise factor given in (6.65) for the common-source and common-gate topologies. As in those two cases, a simultaneous noise and power match is impossible for the common-source LNA with inductive degeneration.

When the source admittance is purely real, $B_s = 0$ and the corresponding noise factor is

$$F = 1+\frac{\alpha\delta\omega^2 C_{gs}^2}{5g_m G_s}\left[1+\frac{G_s^2}{\omega^2 C_{gs}^2}\right]+\frac{\gamma\omega^2 C_{gs}^2}{\alpha g_m G_s}+2|c|\sqrt{\frac{\delta\gamma}{5}}\frac{\omega^2 C_{gs}^2}{g_m G_s}. \tag{6.87}$$

In this case, the noise factor is minimum when

$$G_s = G_{sopt} = \omega C_{gs}\sqrt{1+\frac{5\gamma}{\alpha^2\delta}+\frac{2|c|}{\alpha}\sqrt{\frac{5\gamma}{\delta}}} \tag{6.88}$$

and the corresponding noise factor is

$$\begin{aligned} F_{min} &= 1+2\sqrt{\frac{\alpha\delta}{5}\left(\frac{\gamma}{\alpha}+\frac{\alpha\delta}{5}+2|c|\sqrt{\frac{\delta\gamma}{5}}\right)}\,\frac{\omega}{\omega_T} \\ &= 1+3.26\frac{\omega}{\omega_T} \end{aligned} \tag{6.89}$$

where the latter equality assumes that $\alpha = 0.8$, $\delta = 4$, $\gamma = 2$, and $c = j0.395$.

The above approach identifies the optimum source conductance for a MOS device with a fixed geometry at a particular bias point. However, for integrated-circuit implementations, the device geometry is actually a design variable, while the source

conductance is fixed. Thus, a more appropriate design goal is to determine the optimum device geometry for a particular bias current and a fixed source conductance, e.g., 50 Ω. For short-channel devices, the drain current may be expressed as a function of transconductance and gate capacitance:

$$I_D = \sqrt{\frac{x}{x-y}}\left(x-\frac{y}{2}\right)-x \tag{6.90}$$

where x and y are given by

$$x = \frac{3}{2}\mu_{eff} C_{gs} E_{sat}^2 \tag{6.91}$$

$$y = 2g_m E_{sat} L. \tag{6.92}$$

Solving for g_m in (6.90) and substituting the result into (6.87) results in an expression for noise factor which depends only on C_{gs}, G_s, and I_D. The optimum value of C_{gs} can be determined by differentiating the noise factor with respect to C_{gs}, and the corresponding device width is given by

$$W = \frac{3C_{gs}}{2LC_{ox}}. \tag{6.93}$$

Unfortunately, for short-channel devices, the relationship between I_D, g_m, and C_{gs} is rather complicated, and the resulting equations are too complex to provide any insight into the design process. Although the long-channel equations are invalid, the results derived from these equations can still provide some rough design guidelines. For long-channel devices, the transconductance is given by

$$\begin{aligned} g_m &= \sqrt{2I_D\mu_{eff}C_{ox}\frac{W}{L}} \\ &= \frac{\sqrt{3I_D\mu_{eff}C_{gs}}}{L}. \end{aligned} \tag{6.94}$$

Substituting (6.94) into (6.87) results in an expression for noise factor that is minimum when

$$C_{gs} = C_{gsopt} = \frac{G_s}{\omega}\sqrt{\frac{\frac{\alpha\delta}{5}}{3\left(\frac{\alpha\delta}{5}+\frac{\gamma}{\alpha}+2|c|\sqrt{\frac{\delta\gamma}{5}}\right)}} \tag{6.95}$$

and the corresponding minimum noise factor is

$$F_{min} = 1+\frac{4}{3}L\sqrt{\frac{\omega G_s}{\mu_{eff} I_D}}\left(\frac{\alpha\delta}{5}\right)^{\frac{3}{4}}\left[\frac{1}{3}\left(\frac{\alpha\delta}{5}+\frac{\gamma}{\alpha}+2|c|\sqrt{\frac{\delta\gamma}{5}}\right)\right]^{\frac{1}{4}}. \tag{6.96}$$

Equation (6.96) is plotted as a function of I_D in Fig. 6.13 along with the result derived from the short-channel equations. The noise performance predicted by (6.96) agrees well with the short-channel result and is only slightly optimistic at high bias currents. Once C_{gs} and g_m are determined, the remaining design equations are:

$$L_s = \frac{C_{gs}}{g_m G_s} \tag{6.97}$$

$$L_g = \frac{1}{\omega^2 C_{gs}} - L_s. \tag{6.98}$$

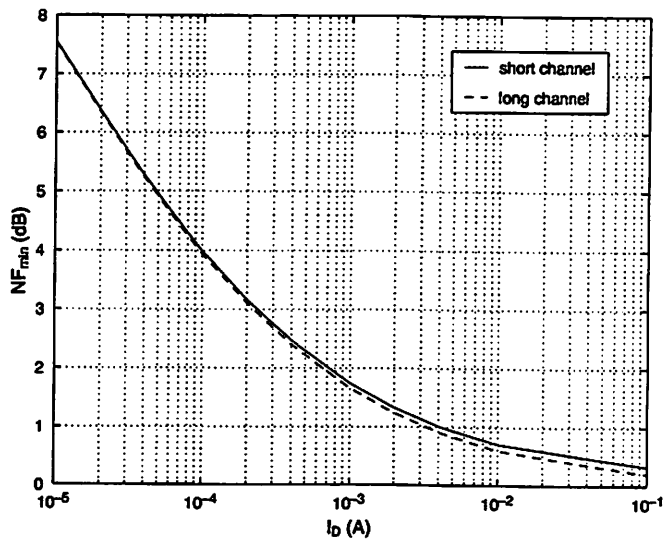


Figure 6.13: NF_{min} versus I_D based on short-channel and long-channel equations for the inductively-degenerated LNA topology. (α= 0.8, δ= 4, γ= 2, $c = j0.395$, G_s = 1/(50 Ω), $\omega = 2\pi 2\times10^9$ rad/s, $L = L_{eff}$ = 0.18 μm (L_{drawn} = 0.25 μm), μ_{eff} = 400 cm^2/Vs, E_{sat} = 5×10^4 V/cm.)

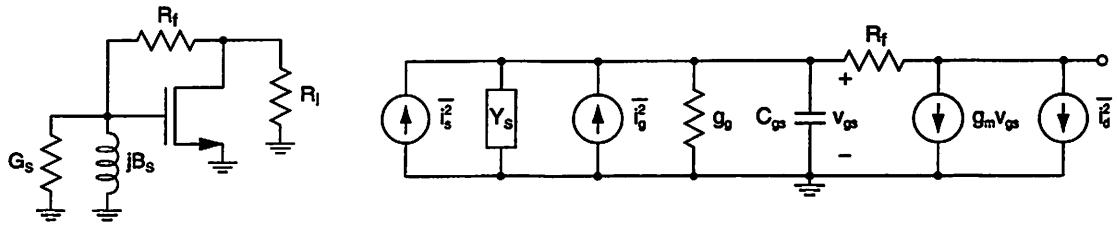


Figure 6.14: (a) Local shunt feedback LNA. (b) Small-signal equivalent circuit.

Finally, the voltage gain of this LNA is

$$A_v = \frac{R_l}{j\omega L_s\left[1+\dfrac{g_g(L_g+L_s)}{g_m L_s}\right]} = \frac{R_l}{j\omega L_s\left[1+\dfrac{\alpha C_{gs}}{5g_m^2 L_s}\right]} \tag{6.99}$$

where R_l is the load resistance at the output of the LNA.

Local shunt feedback LNA. A fourth topology which may be used to achieve a 50-Ω input resistance is based on a single transistor with local shunt feedback as illustrated in Fig. 6.14a [67]. The input impedance of this LNA is

$$Z_{in} = \frac{R_f + R_l}{1+g_m R_l + (R_f+R_l)(g_g + j\omega C_{gs})}. \tag{6.100}$$

A narrowband input match may be achieved by adding a shunt inductor at the input of the LNA in order to tune out the gate capacitance of the transistor. In this case, the input impedance is

$$Z_{in} = \frac{R_f + R_l}{1+g_m R_l + g_g(R_f+R_l)}. \tag{6.101}$$

Alternatively, a broadband input match may be achieved by eliminating the shunt inductor at the LNA input and selecting

$$\begin{aligned} \omega C_{gs}(R_f+R_l) &<< g_m R_l \\ \omega &<< \omega_T \frac{R_l}{R_f+R_l}. \end{aligned} \tag{6.102}$$

Under this condition, $g_g(R_f + R_l)$ is also much less than $g_m R_l$, and consequently, the input impedance reduces to

$$Z_{in} = \frac{R_f + R_l}{1 + g_m R_l}. \tag{6.103}$$

In the case of a broadband input match, both g_g and C_{gs} may be ignored, and the noise factor is

$$F = 1 + \frac{\alpha\delta\omega^2 C_{gs}^2}{5 g_m G_s} + \frac{\gamma g_m[(1 + R_f G_s)^2 + (R_f B_s)^2]}{\alpha G_s (1 - g_m R_f)^2} + \frac{R_f[(g_m + G_s)^2 + B_s^2]}{G_s(1 - g_m R_f)^2} - 2|c|\sqrt{\frac{\delta\gamma}{5}}\frac{\omega C_{gs} R_f B_s}{G_s(1 - g_m R_f)}. \tag{6.104}$$

For $g_m R_f >> 1$, the minimum noise factor is achieved when

$$G_s = G_{sopt} = \sqrt{\frac{\alpha g_m}{\gamma R_f} + \frac{\alpha^2\delta\omega^2 C_{gs}^2(1 - |c|^2)}{5\gamma}} \tag{6.105}$$

$$B_s = B_{sopt} = -\alpha|c|\sqrt{\frac{\delta}{5\gamma}}\omega C_{gs} \tag{6.106}$$

and the corresponding noise factor is

$$F_{min} = 1 + 2\sqrt{\frac{\delta\gamma}{5}(1 - |c|^2)}\frac{\omega}{\omega_T} \tag{6.107}$$

which is identical to the minimum noise factor for the other LNA topologies. As in those cases, a simultaneous noise and power match is also impossible for this LNA.

When the source admittance is purely real, the noise factor becomes

$$F = 1 + \frac{\alpha\delta\omega^2 C_{gs}^2}{5 g_m G_s} + \frac{\gamma g_m (1 + R_f G_s)^2}{\alpha G_s (1 - g_m R_f)^2} + \frac{R_f (g_m + G_s)^2}{G_s (1 - g_m R_f)^2}. \tag{6.108}$$

The minimum noise factor is achieved when .

$$G_s = G_{sopt} = \sqrt{\frac{5g_m^2(\gamma+\alpha g_m R_f)+\alpha^2\delta\omega^2 C_{gs}^2(1-g_m R_f)^2}{5g_m R_f(\alpha+\gamma g_m R_f)}} \tag{6.109}$$

and the corresponding noise factor is

$$F_{min} = 1+\frac{2(\alpha+\gamma)g_m R_f}{\alpha(1-g_m R_f)^2}+ \frac{2\sqrt{g_m R_f(\alpha+\gamma g_m R_f)[5g_m^2(\gamma+\alpha g_m R_f)+\alpha^2\delta\omega^2 C_{gs}^2(1-g_m R_f)^2]}}{\sqrt{5}\alpha g_m(1-g_m R_f)^2}. \tag{6.110}$$

By assuming $g_m R_f >> 1$, (6.108) – (6.110) may be further simplified. In this case, the noise factor becomes

$$F = 1+\frac{\alpha\delta\omega^2 C_{gs}^2}{5g_m G_s}+\frac{\gamma G_s}{\alpha g_m}+\frac{1}{R_f G_s}. \tag{6.111}$$

The minimum noise factor is achieved when

$$G_s = G_{sopt} = \sqrt{\frac{\alpha g_m}{\gamma R_f}+\frac{\alpha^2\delta\omega^2 C_{gs}^2}{5\gamma}} \tag{6.112}$$

and the corresponding noise factor is

$$\begin{aligned} F_{min} &= 1+2\sqrt{\frac{\delta\gamma}{5}}\frac{\omega}{\omega_T} \\ &= 1+2.53\frac{\omega}{\omega_T} \end{aligned} \tag{6.113}$$

where the latter equality assumes that $\delta = 4$ and $\gamma = 2$.

The above approach identifies the optimum source conductance for a MOS device with a fixed geometry at a particular bias point. However, for integrated-circuit implementations, the device geometry is actually a design variable, while the source conductance is fixed. Thus, a more appropriate design goal is to determine the optimum device geometry for a particular bias current and a fixed source conductance, e.g., 50 Ω. In this case, using the long-channel expression for g_m from (6.94) still results in very

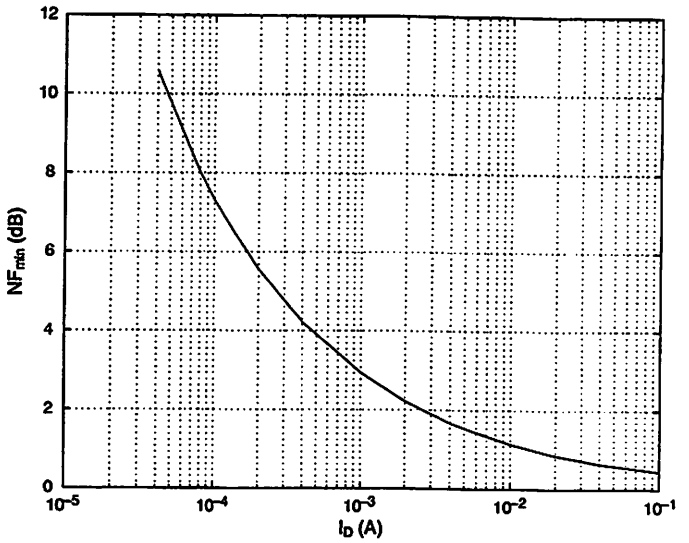


Figure 6.15: NF_{min} versus I_D for the LNA topology with local shunt feedback. ($\alpha = 0.8$, $\delta = 4$, $\gamma = 2$, $c = j0.395$, $G_s = 1/(50\,\Omega)$, $\omega = 2\pi 2\times 10^9$ rad/s, $L = L_{eff} = 0.18$ μm ($L_{drawn} = 0.25$ μm), $\mu_{eff} = 400$ cm^2/Vs, $E_{sat} = 5\times 10^4$ V/cm, $R_l = 350\,\Omega$.)

complicated expressions for C_{gsopt} and the corresponding F_{min}, which are not reported here. However, the minimum noise figure based on the short-channel equations is plotted in Fig. 6.15. This result is derived by substituting expressions for g_m and R_f into (6.108). An expression for g_m is derived from (6.90), while an expression for R_f is derived by setting the source conductance equal to the LNA input conductance in (6.103):

$$G_s = \frac{1 + g_m R_l}{R_f + R_l}. \tag{6.114}$$

Finally, in the case of a broadband input match, the voltage gain of this LNA is

$$A_v = (1 - g_m R_f)\frac{R_l}{R_f + R_l} \tag{6.115}$$

where R_l is the load resistance at the output of the LNA. For $g_m R_f >> 1$, the voltage gain becomes

$$A_v = -g_m (R_f \parallel R_l). \tag{6.116}$$

Summary of LNA Performance. The performance of the four LNA topologies are summarized in Table 6.1. When the device geometry and bias current are fixed, all four

	common-source	common-gate	common-source with inductive degeneration	local shunt feedback
F_{min}	$=1+2\sqrt{\frac{\delta\gamma}{5}(1-\|c\|^2)}\frac{\omega}{\omega_T}$ $=1+2.32\frac{\omega}{\omega_T}$	$=1+2\sqrt{\frac{\delta\gamma}{5}(1-\|c\|^2)}\frac{\omega}{\omega_T}$ $=1+2.32\frac{\omega}{\omega_T}$	$=1+2\sqrt{\frac{\delta\gamma}{5}(1-\|c\|^2)}\frac{\omega}{\omega_T}$ $=1+2.32\frac{\omega}{\omega_T}$	$=1+2\sqrt{\frac{\delta\gamma}{5}(1-\|c\|^2)}\frac{\omega}{\omega_T}$ $=1+2.32\frac{\omega}{\omega_T}$
G_{sopt}	$\sqrt{\frac{\alpha^2\delta(1-\|c\|^2)\omega^2C_{gs}^2}{5\gamma}}$	$\sqrt{\frac{\alpha^2\delta(1-\|c\|^2)\omega^2C_{gs}^2}{5\gamma}}$	$\frac{\omega C_{gs}}{\alpha}\sqrt{\frac{5\gamma}{\delta}(1-\|c\|^2)}$	$\sqrt{\frac{\alpha g_m}{\gamma R_f}+\frac{\alpha^2\delta(1-\|c\|^2)\omega^2C_{gs}^2}{5\gamma}}$
B_{sopt}	$-\omega C_{gs}\left(1+\alpha\|c\|\sqrt{\frac{\delta}{5\gamma}}\right)$	$-\omega C_{gs}\left(1+\alpha\|c\|\sqrt{\frac{\delta}{5\gamma}}\right)$	$\omega C_{gs}\left(1+\frac{\|c\|}{\alpha}\sqrt{\frac{5\gamma}{\delta}}\right)$	$-\omega C_{gs}\alpha\|c\|\sqrt{\frac{\delta}{5\gamma}}$
Z_{in}	$\frac{1}{g_g}$	$\frac{1}{g_g+g_m}$	$\frac{g_m L_s}{C_{gs}}$	$\frac{R_f+R_l}{1+g_m R_l}$
$F_{50\Omega}$	$\geq 1+\delta$ ≥ 5 or 6.99 dB	$\geq 1+\frac{\gamma}{\alpha}$ ≥ 3.5 or 5.44 dB	≥ 1 or 0 dB	≥ 1 or 0 dB
A_v	$-g_m R_l$	$g_m R_l$	$=-j\frac{R_l}{\omega L_s}$ $=-j\frac{G_s g_m R_l}{\omega C_{gs}}$	$-g_m(R_f \parallel R_l)$

Table 6.1: Summary of LNA topologies ($\alpha = 0.8$, $\delta = 4$, $\gamma = 2$, $c = j0.395$).

LNA topologies have the same minimum noise factor, although the optimum source admittance varies. In all four cases, an impedance match which simultaneously provides minimum noise factor and maximum power transfer is impossible.

When the source impedance is purely real and fixed at 50 Ω, a power match results in minimum achievable noise figures of 6.99 dB and 5.44 dB for the common-source and common-gate topologies, respectively. On the other hand, the minimum noise figures of the common-source LNA with inductive degeneration and the LNA with local shunt feedback both have an asymptotic limit of 0 dB. Consequently, for applications with very stringent noise requirements, the common-source topology with inductive degeneration and the topology with local shunt feedback are the best candidates for low-noise amplification. However, for applications with relaxed noise figure requirements, all four topologies are viable options. In this case, the selection should be based on other criteria, such as feasibility of integration or power consumption, which will be discussed next.

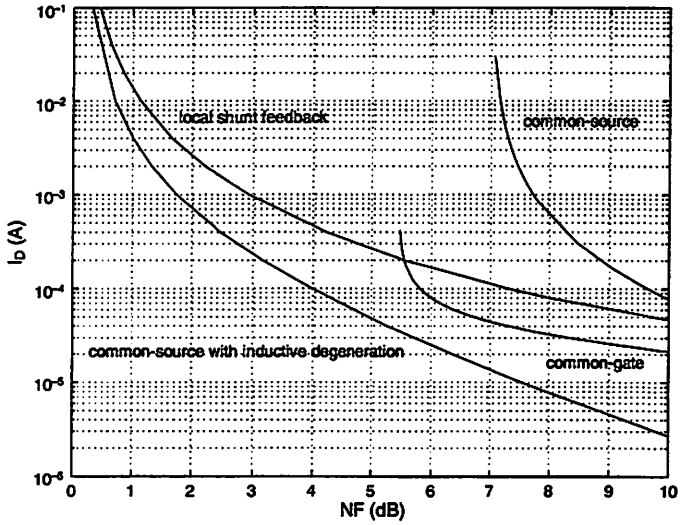
Integration and Power Consumption Issues. The topologies which rely on inductors for narrowband input matching are less feasible for integration. These topologies include the common-source LNA and the common-source LNA with inductive degeneration. Although the common-source topology may also be designed for a broadband input match, the noise factor in this case is in excess of 18.3 dB, which is intolerable even for applications with very relaxed noise requirements. For topologies with a narrowband input match, an inductor is added to the input of the LNA in order to tune out the gate capacitance of the transistor. Although an ideal inductor is noiseless, practical realizations introduce noise due to the series resistance of the inductor. Since this inductor appears at the input of the LNA, excessive noise from the inductor series resistance may not be tolerable. Many standard digital CMOS technologies rely on low-resistivity silicon substrates, and the quality factors of on-chip spiral inductors implemented on these substrates tend to be rather poor, resulting in increased noise. Consequently, off-chip inductors must be used for these topologies. Nevertheless, small form factor is still achievable for implementations based on these topologies by implementing the off-chip inductors as bond wires or by incorporating them into the package.

For applications with relaxed noise figure requirements, a second criteria for topology selection is power consumption. For the common source topology with a narrowband input match to 50 Ω, the noise figure is given by (6.67) and $g_g = G_s = 1/(50\ \Omega)$. The corresponding expressions for g_m and C_{gs} are, respectively,

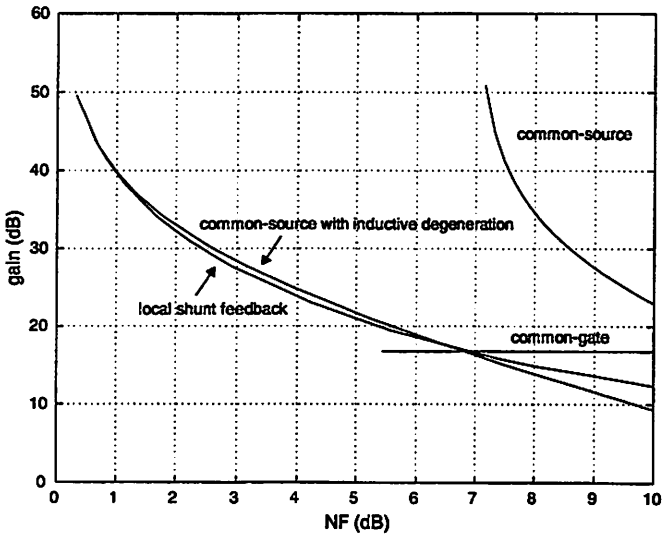
$$g_m = \frac{4\gamma G_s}{\alpha(F-1-\delta)} \tag{6.117}$$

$$C_{gs} = \frac{5G_s}{\alpha\omega}\sqrt{\frac{4\gamma}{5(F-1-\delta)}}. \tag{6.118}$$

In this case, the drain current is given by (6.90), where g_m and C_{gs} are given by (6.117) and (6.118), respectively. The drain current is plotted as a function of noise figure in Fig. 6.16a assuming $\alpha = 0.8$, $\delta = 4$, $\gamma = 2$, $c = j0.395$, $G_s = 1/(50\ \Omega)$, $\omega = 2\pi 2\times10^9$ rad/s, $L = L_{eff} = 0.18$ μm ($L_{drawn} = 0.25$ μm), $\mu_{eff} = 400$ cm^2/Vs, and $E_{sat} = 5\times10^4$ V/cm. Also, the



(a)



(b)

Figure 6.16: (a) Drain current versus noise figure. (b) Gain versus noise figure. ($\alpha = 0.8$, $\delta = 4$, $\gamma = 2$, $c = j0.395$, $G_s = 1/(50\ \Omega)$, $\omega = 2\pi 2\times 10^9$ rad/s, $L = L_{eff} = 0.18$ μm ($L_{drawn} = 0.25$ μm), $\mu_{eff} = 400$ cm^2/Vs, $E_{sat} = 5\times 10^4$ V/cm, $R_l = 350\ \Omega$.)

gain is plotted versus noise figure in Fig. 6.16b with the additional assumption that $R_l = 350\ \Omega$.

For the common-gate topology with a broadband input match to 50 Ω, the noise figure is given by (6.76) and $g_m = G_s = 1/(50\ \Omega)$. The corresponding gate capacitance is

$$C_{gs} = \frac{G_s}{\omega}\sqrt{\frac{5}{\alpha\delta}\left(F - 1 - \frac{\gamma}{\alpha}\right)}. \tag{6.119}$$

The drain current and gain are plotted as a functions of noise figure in Fig. 6.16.

When the inductively-degenerated common-source topology is matched to 50 Ω, the expression for the minimum noise factor and the corresponding expressions for g_m and C_{gs} are rather complicated as already discussed in Section 6.2.5. Although the results are not explicitly reported here, the drain current and gain are plotted versus noise figure in Fig. 6.16.

Finally, when the topology with local shunt feedback is matched to 50 Ω, the expression for the minimum noise factor and the corresponding expressions for g_m and C_{gs} are also rather complicated. Although the results are not explicitly reported here, the drain current and gain are plotted versus noise figure in Fig. 6.16.

The inductively-degenerated common-source LNA provides the best overall performance. For the same noise figure performance, the common-source LNA with inductive degeneration requires the least amount of current, and consequently, is the lowest power solution. Also, its gain performance is comparable to that of the LNA with local shunt feedback. For relaxed noise figure requirements, the common-source topology provides the highest gain but also consumes significantly more power than the other LNA topologies for the same noise performance.

For the receiver prototype described here, low power consumption is the most important design consideration. The common-source LNA with inductive degeneration is a good candidate since it has the lowest power consumption while providing adequate gain. The drain current and gain are plotted as a functions of noise figure in Fig. 6.16, while C_{gs}, g_m, L_s, and L_g are plotted versus noise figure in Fig. 6.17. In summary, the following guidelines facilitate the design of the inductively-degenerated common-source LNA for a particular noise figure requirement:

1. For a given noise figure requirement, I_D is plotted in Fig. 6.16.
2. For this noise figure requirement, C_{gs} is plotted in Fig. 6.17a, and the corresponding device width is given by $W = 3C_{gs}/(2LC_{ox})$.
3. The corresponding device transconductance is determined from (6.90).

4. The values of L_s and L_g are then determined by (6.97) and (6.98), respectively.

6.2.6 Inductively-Degenerated Differential LNA

The LNA for this prototype is implemented using an inductively-degenerated differential amplifier topology illustrated in Fig. 6.18 [11]. Since the LNA is integrated onto the same chip along with other receiver components, noise introduced by other circuits can couple to the LNA through the supply or the substrate. The common-source LNA with inductive degeneration described in Section 6.2.5 is implemented in a differential configuration in order to improve its common-mode rejection, resulting in increased robustness against

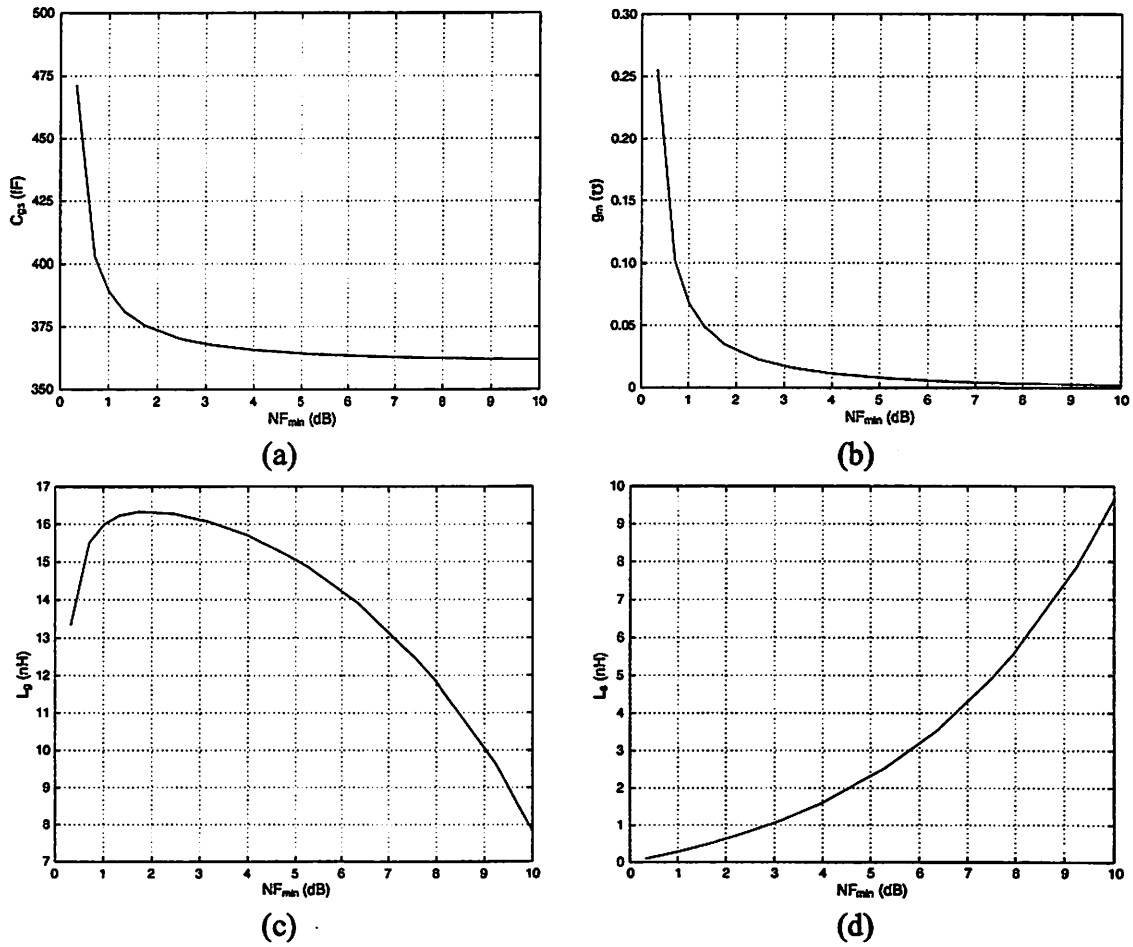


Figure 6.17: Designing for minimum noise figure in the common-source LNA with inductive degeneration. (a) C_{gs}. (b) g_m. (c) L_g. (d) L_s. ($\alpha = 0.8$, $\delta = 4$, $\gamma = 2$, $c = j0.395$, $G_s = 1/(50\ \Omega)$, $\omega = 2\pi 2\times10^9$ rad/s, $L = L_{eff} = 0.18$ µm ($L_{drawn} = 0.25$ µm), $\mu_{eff} = 400$ cm^2/Vs, $E_{sat} = 5\times10^4$ V/cm, $R_l = 350\ \Omega$.)

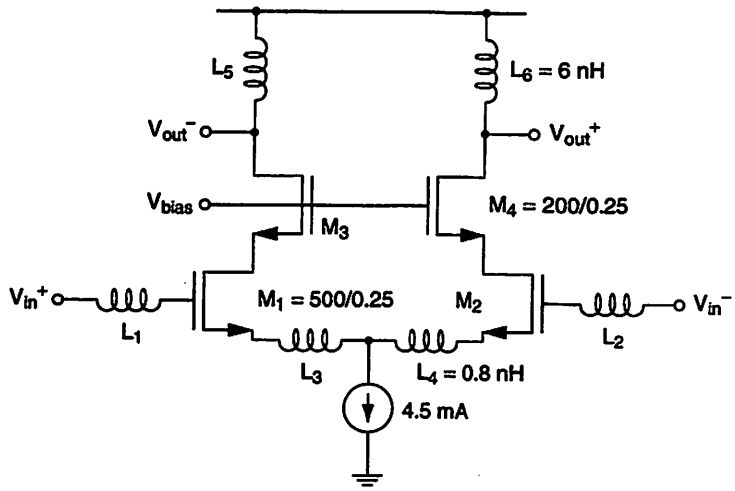


Figure 6.18: Inductively-degenerated differential LNA.

noise coupling.

Although the design guidelines for the differential topology are similar to those for the single-ended case, there are a few minor differences. First the input impedance of the differential LNA is given by

$$Z_{in} = 2\left[\frac{g_m L_s}{C_{gs}} - j\frac{1-\omega^2 C_{gs}(L_g + L_s)}{\omega C_{gs}}\right] \tag{6.120}$$

where $g_m = g_{m1} = g_{m2}$, $C_{gs} = C_{gs1} = C_{gs2}$, $L_s = L_3 = L_4$, and $L_g = L_1 = L_2$. In this case the input impedance is twice that of the single-ended topology. For this receiver prototype, an external balun is used to convert the single-ended RF input signal to a differential signal for the LNA. The unbalanced impedance of commercially-available baluns is 50 Ω in order to provide a match to the preceding RF filter, while the balanced impedance can be either 50 Ω, 100 Ω, or 200 Ω [6]. An inductively-degenerated differential LNA is depicted in Fig. 6.19 along with the equivalent half-circuit. Comparing the half-circuit in Fig. 6.19b with the single-ended circuit in Fig. 6.12a, the two circuits are equivalent for $G_s' = G_s/2$ and $I_D' = 2I_D$. The drain current of the differential LNA is plotted in Fig. 6.20 versus noise figure for $G_s' = 1/(50\,\Omega)$, $1/(100\,\Omega)$, and $1/(200\,\Omega)$. For the same noise figure performance, the differential LNA matched to the lowest source conductance consumes the least amount of current. Conversely, for the

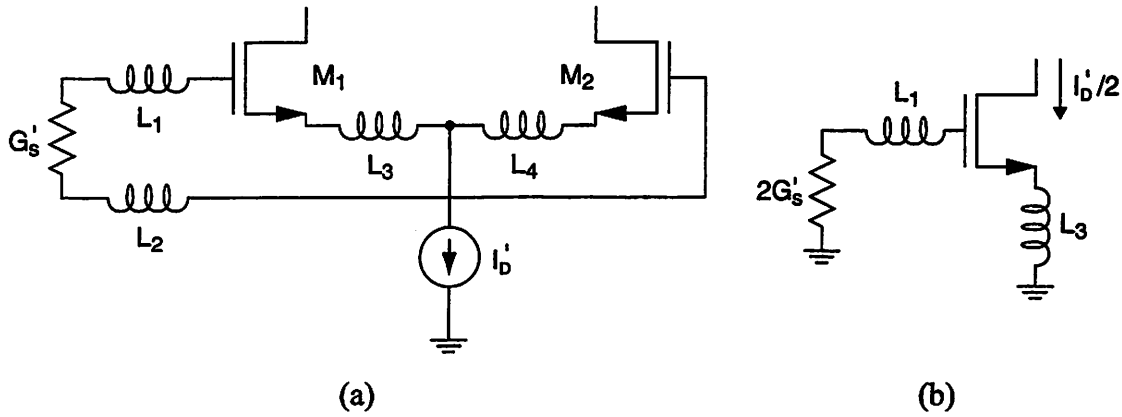


Figure 6.19: (a) Differential LNA with source conductance G_s' and tail current I_D'. (b) Equivalent half-circuit.

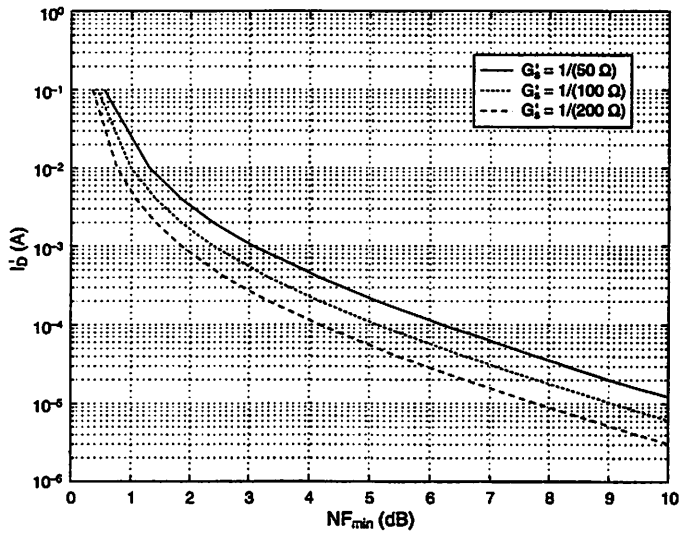


Figure 6.20: Drain current versus noise figure of the inductively-degenerated differential LNA for several values of source conductance.

same current consumption, the differential LNA matched to the lowest source conductance has the best noise figure [68]. Unfortunately, the impact of the source conductance on the noise figure was not fully appreciated at the time, and the LNA used in this receiver prototype was designed to match a balanced resistance of 50 Ω instead of 200 Ω.

In the receiver prototype, the dimensions of each of the transistors M_1 and M_2 is 500 μm × 0.25 μm. The LNA is biased at 4.5 mA and is powered by a separate 2.5-V supply,

which helps to isolate the LNA from the potentially noisy supplies of the other receiver circuits. In the layout, each of the input transistors is partitioned into ten blocks of ten fingers of 5 μm × 0.25 μm devices. Fingering the devices helps to reduce the resistance associated with the polysilicon gate [64], [69]. Keeping this resistance small is critical since the noise associated with this resistance appears directly at the input of the LNA. In addition, substrate contacts are placed generously around and between the ten blocks in order to reduce the substrate resistance near the LNA. The noise voltage associated with this resistance modulates the bulk of the transistor and introduces an additional component to the drain current noise,

$$\overline{i_{d,sub}^2} = 4kTR_{sub}g_{mb}^2\Delta f \tag{6.121}$$

which can degrade the noise performance of the LNA [25], [70], [71]. The layout of the input transistors M_1 and M_2 is illustrated in Fig. 6.21.

The noise factor of the differential LNA in Fig. 6.18 is also affected by the cascode transistors M_3 and M_4. While these cascode transistors slightly degrade the noise performance of the LNA, they provide increased isolation, reducing the amount of LO leakage from the mixer to the receiver input. In direct-conversion receivers, the frequency

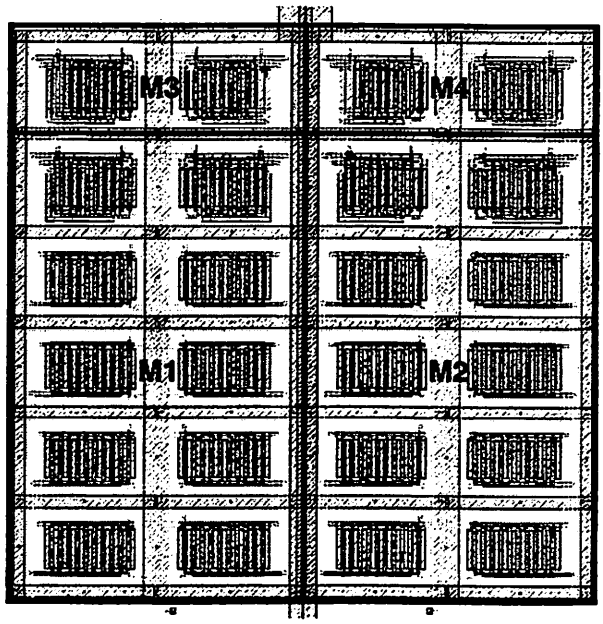


Figure 6.21: Layout of LNA transistors $M_1 - M_4$.

of the leakage signal is the same as that of the desired signal, and consequently, is radiated from the antenna without any attenuation from the RF filter. This leakage signal is problematic since it can potentially interfere with other systems operating in the same frequency band. In addition to reducing the amount of LO leakage to the receiver input, the cascode transistors also reduce the influence of the gate-drain overlap capacitances of M_1 and M_2 on the LNA input impedance [68]. Including the gate-drain overlap capacitance, the input impedance of the equivalent half-circuit may be determined by the small-signal equivalent circuit illustrated in Fig. 6.22:

$$Z_{in} = \frac{g_m L_s + R_l C_{gd} - \omega^2 g_m L_g L_s C_{gd} - \omega^2 R_l (L_g + L_s) C_{gd} C_{gs}}{C_{gs} + C_{gd}(1 + g_m R_l - \omega^2 L_s C_{gs} + j\omega g_m L_s + j\omega R_l C_{gs})} + \frac{1 - \omega^2 (L_g + L_s) C_{gs} - \omega^2 L_g C_{gd} - \omega^2 g_m R_l (L_g + L_s) C_{gd} + \omega^4 L_g L_s C_{gd} C_{gs}}{j\omega [C_{gs} + C_{gd}(1 + g_m R_l - \omega^2 L_s C_{gs} + j\omega g_m L_s + j\omega R_l C_{gs})]}. \tag{6.122}$$

Assuming that

$$\omega^2 = \frac{1}{(L_g + L_s) C_{gs}} \tag{6.123}$$

(6.122) becomes

$$Z_{in} = \frac{g_m L_s \left[1 - \dfrac{L_g C_{gd}}{(L_g + L_s) C_{gs}}\right]}{C_{gs} + C_{gd}\left[1 - \dfrac{L_s}{L_g + L_s} + g_m R_l + j\omega (g_m L_s + R_l C_{gs})\right]} - \frac{\dfrac{C_{gd}}{C_{gs}}\left[\dfrac{L_g}{L_g + L_s} + g_m R_l + \dfrac{L_g L_s}{(L_g + L_s)^2}\right]}{j\omega \left\{C_{gs} + C_{gd}\left[1 - \dfrac{L_s}{L_g + L_s} + g_m R_l + j\omega (g_m L_s + R_l C_{gs})\right]\right\}}. \tag{6.124}$$

Since C_{gd} is typically much less than C_{gs}, (6.124) reduces to $g_m L_s / C_{gs}$ as long as R_l is small. For the differential LNA in Fig. 6.18, the cascode transistors M_3 and M_4 present a small load resistance ($R_l = 1/g_{m3} = 1/g_{m4}$) to each of the input devices M_1 and M_2, thus reducing the effect of the gate-drain overlap capacitances of M_1 and M_2 on the input impedance. Consequently, for increased insensitivity to the effects of the gate-drain

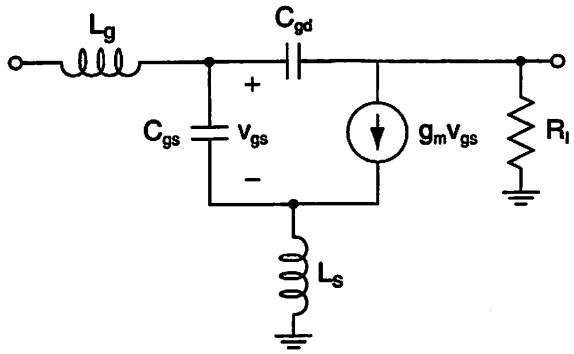
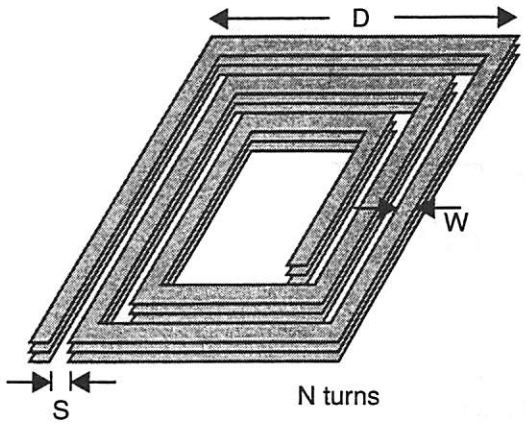


Figure 6.22: Small-signal equivalent circuit of the inductively-degenerated LNA including gate-drain overlap capacitance.

overlap capacitances, the transconductance of each of the cascode transistors should be large. However, a large transconductance also increases the noise contribution of each of the cascode devices. The widths of the cascode transistors should be chosen as a compromise between these two opposing factors. In the receiver prototype, the dimensions of each of the transistors M_3 and M_4 is 100 μm × 0.25 μm.

Inductors $L_3 - L_6$ are realized as on-chip spiral inductors. A test chip (RFTRIPLED) was fabricated in order to evaluate the performance of various inductor test structures and the results are reported in Appendix D. The spiral inductors implemented in the receiver prototype use the top three layers of metal all shorted together using a large number of vias in order to reduce the series resistance, thus improving the inductor quality factor. Separate test structures were fabricated in order to characterize the performance of the inductors actually used in the receiver prototype. The geometries of inductors $L_3 - L_6$ are summarized in Fig. 6.23. At 2 GHz, the measured quality factors of inductors L_3 and L_5 are 3.9 and 3.6, respectively. These low quality factors are typical for on-chip spiral inductors implemented on low-resistivity silicon substrates. The low quality factors of these on-chip spiral inductors are not detrimental to the low noise performance of the amplifier. Indeed, the series resistance of each of the inductors L_3 and L_4 must be kept small since the noise associated with each of these resistances contributes directly to the LNA noise figure. However, despite the low quality factor, the series resistance of each of the source inductors is actually quite small due to the low inductance values. The series resistance may be estimated by the following expression for quality factor:



	L_3, L_4	L_5, L_6
D (μm)	150	250
W (μm)	19	13
S (μm)	2.5	2
N	2.5	5.5
L (nH)	0.8	6
Q @ 2 GHz	3.9	3.6

Figure 6.23: Summary of on-chip spiral inductors.

$$Q = \frac{\omega L}{R}. \tag{6.125}$$

From (6.125), the series resistance of each of the inductors L_3 and L_4 is only 2.6 Ω at 2 GHz. The noise associated with the series resistance of each of the load inductors L_5 and L_6 also contributes to the LNA noise figure. In this case, the mean-square value of the input-referred voltage noise is approximately

$$\overline{v_{i,R_l}^2} = 4kTR_l\left(\frac{L_s}{L_l}\right)^2 \Delta f \tag{6.126}$$

where L_s is the source inductance and L_l and R_l are the inductance and series resistance, respectively, of the load inductor. Since L_s is much less that L_l, the noise contribution due to the series resistance of each of the load inductors is minimal with an equivalent noise resistance of only 0.37 Ω at 2 GHz.

Each of the LNA input bond pads consists of the top three layers of metal, all shorted together, while a fourth lower layer of metal acts as a shield [72]. However, rather than connecting the shield to ground, the shield instead is connected to the source terminal of the input transistor so that the pad capacitance appears in parallel with the gate-source capacitance of the input transistor as illustrated in Fig. 6.24 [73]. This technique significantly reduces the effect of the pad capacitances on the input matching. Including the pad capacitances, the input impedance of the amplifier can be described by

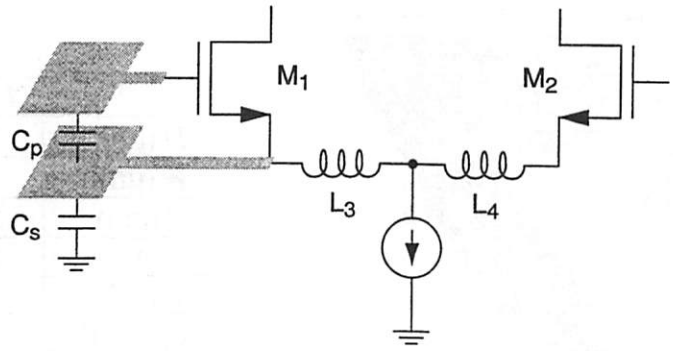


Figure 6.24: Interface between LNA and input pads.

$$Z_{in} = 2\left[\frac{g_{m1}L_3}{C_T(1-\omega^2 C_s L_3)} - j\frac{1-\omega^2 C_T\left(L_1 + \dfrac{L_3}{1-\omega^2 C_s L_3}\right)}{\omega C_T}\right] \tag{6.127}$$

where C_T is the parallel combination of the transistor gate-source capacitance and the pad capacitance C_p, while C_s is the capacitance between the pad shield and the substrate. In this case, input matching is achieved by equating the real part of Z_{in} to 50 Ω and then selecting the values of inductors L_1 and L_2 to tune out the imaginary part of Z_{in} at the carrier frequency:

$$L_1 = L_2 = \frac{1}{\omega^2 C_T} - \frac{L_3}{1-\omega^2 C_s L_3}. \tag{6.128}$$

On-chip spiral inductors should be avoided when implementing inductors L_1 and L_2. These inductors appear directly at the LNA input and any series resistance can potentially degrade the noise performance of the LNA. In particular, for low noise performance, the gate inductance must be relatively large, and implementation as an on-chip spiral inductor results in a relatively large series resistance. Instead, the inductors L_1 and L_2 are realized using the input bond wires, which have quality factors in excess of 20 and provide an inductance of about 1 nH/mm. The LNA input bond pads are recessed about 300 μm from the edge of the chip in order to accommodate longer input bond wires, and consequently, higher inductance values. The layout of the LNA is illustrated in Fig. 6.25.

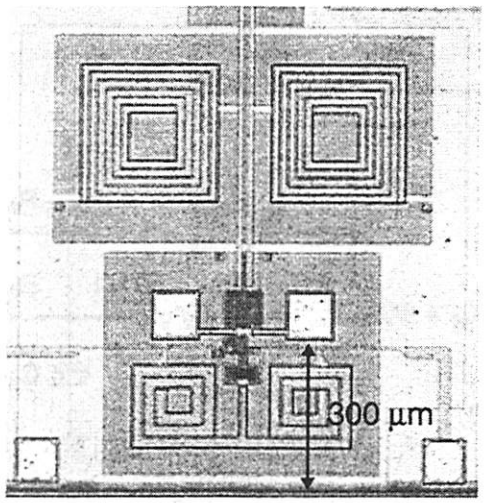


Figure 6.25: LNA layout.

Inductors L_5 and L_6 provide a tuned response at the output of the LNA. At resonance, the load impedance is

$$Z_l = \frac{1 + j\frac{\omega L_l}{R_l}}{j\omega C_l} = \frac{1 + jQ}{j\omega C_l} \approx \frac{Q}{\omega C_l} \tag{6.129}$$

where C_l is the load capacitance, and L_l and R_l are the inductance and series resistance, respectively, of the load inductor. Larger voltage gains are achieved by using larger load impedances. However, the load impedance is limited by two opposing factors. First, for the same resonance frequency, larger load impedances are achieved by using smaller capacitance and larger inductance values. However, on-chip spiral inductors with larger inductance values tend to have lower quality factors. In the receiver prototype, the LNA load impedance is about 290 Ω.

The output of the LNA is connected to the input of the subsequent mixer through a pair of 2.3-pF coupling capacitors. Large coupling capacitors are used in order to minimize the signal attenuation resulting from the capacitive voltage divider formed by the coupling capacitors and the input capacitances of the mixer. The coupling capacitors are implemented as sandwich structures using the top four layers of metal [74]. In order to

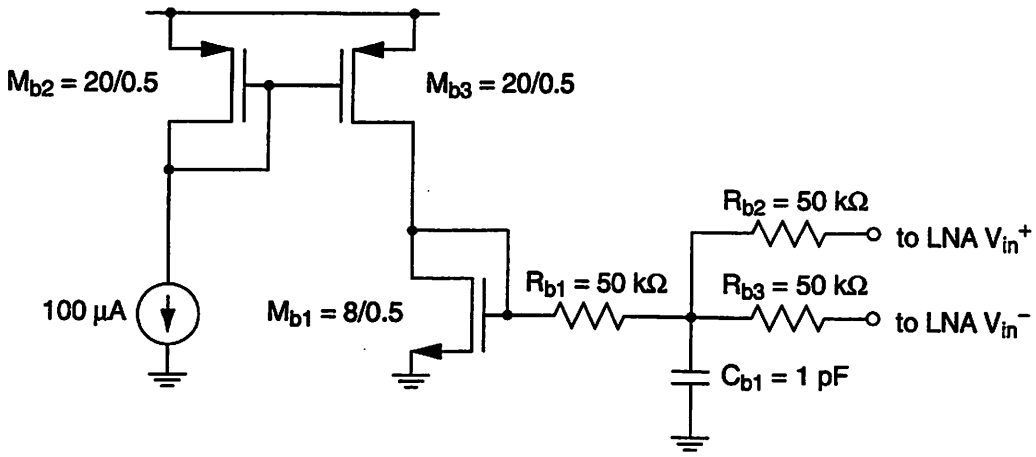
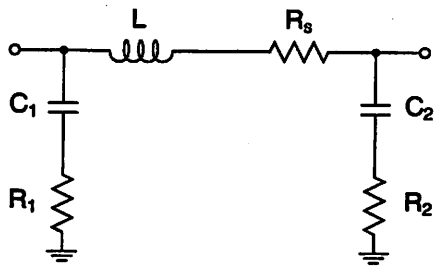


Figure 6.26: Bias circuit for the LNA input.

reduce the effects of voltage division, the parasitic bottom plate capacitances are placed at the LNA output, where all the parasitic capacitances are tuned out by load inductors L_5 and L_6.

Finally, the bias circuit for the LNA input is illustrated in Fig. 6.26 [18]. The bias voltage at the LNA input is determined by the gate voltage of transistor M_{b1}. Each of the R-C-R networks acts as a choke for high-frequency signals without degrading the noise performance of the LNA.

Simulation Results. The LNA was simulated in SpectreRF with device models based on the Philips MOS Model 9, which includes the effects of induced gate current noise [75]. Due to the tuned nature of the LNA, accurate simulation results require inclusion of all parasitic capacitances. A netlist for the LNA including the input bond pads was extracted



(a)

	L_3, L_4	L_5, L_6
L	0.8 nH	6 nH
R_s	2.5 Ω	20 Ω
C_1	75 fF	180 fF
C_2	75 fF	180 fF
R_1	20 Ω	10 Ω
R_2	20 Ω	10 Ω

(b)

Figure 6.27: (a) Circuit model for on-chip spiral inductors. (b) Component values at 2 GHz.

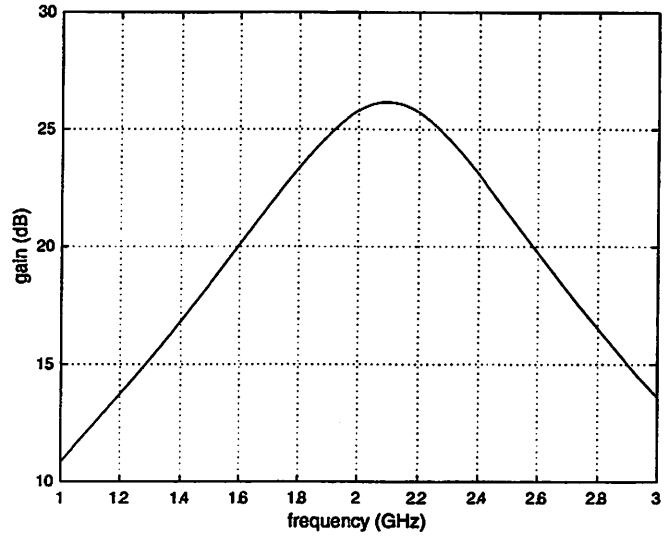


Figure 6.28: Simulated LNA gain.

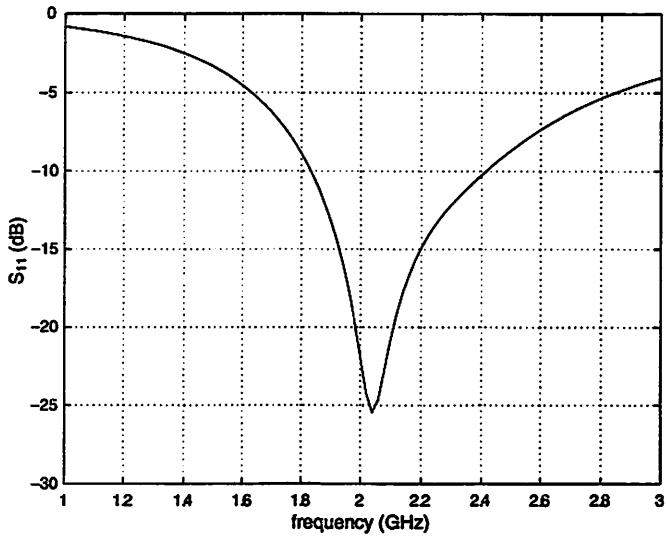


Figure 6.29: Simulated LNA S_{11}.

from the layout for use in simulations. In addition, the circuit illustrated in Fig. 6.27 was used to model the on-chip spiral inductors.

The simulated gain of the LNA is illustrated in Fig. 6.28. The peak gain is 26 dB and occurs at approximately 2.1 GHz, while the gain at 2 GHz is about 25.8 dB. The simulated S_{11} of the LNA is illustrated in Fig. 6.29. S_{11} is the reflection coefficient seen looking into the input of the LNA [59]:

	output noise (V^2/Hz)	contribution
R_s (50 Ω)	6.23×10^{-8}	47.2%
M_1, M_2	4.17×10^{-8}	31.6%
M_3, M_4	1.69×10^{-8}	12.8%
L_5, L_6	6.41×10^{-9}	4.8%
L_3, L_4	3.60×10^{-9}	2.7%
total	1.32×10^{-7}	100%

Table 6.2: Summary of LNA noise performance from simulation.

$$S_{11} = \frac{V_1^-}{V_1^+} = \frac{Z_{in} - Z_o}{Z_{in} + Z_o} \tag{6.130}$$

where V_1^+ and V_1^- are the voltage amplitudes of the incident and reflected waves, respectively, and Z_{in} and Z_o are the source and LNA input impedances, respectively. Consequently, S_{11} is a measure of how well the LNA input impedance is matched to the source impedance, where $S_{11} = 0$ for a perfect match. This simulation was performed with a pair of ideal 4.5-nH gate inductors to complete the input tuning. The S_{11} at 2 GHz is about –22 dB.

The simulated noise figure of the LNA is about 3.26 dB and the dominate noise contributors are summarized in Table 6.2. The simulated noise figure due to M_1 and M_2 alone is about 2.23 dB and agrees well with the 2.18-dB noise figure calculated from (6.87) using the simulated values of g_m and C_{gs} and assuming that $\alpha = 0.8$, $\delta = 4$, $\gamma = 2$, $c = j0.395$, $G_s = 1/(50\ \Omega)$, and $\omega = 2\pi 2\times 10^9$ rad/s. The simulated values of g_m and C_{gs} are 0.03 Ω^{-1} and 453 fF, respectively, where C_{gs} consists of the parallel combination of the 165-fF pad capacitance and the 288-fF gate-source capacitance of transistor M_1. Although the simulated noise figure agrees well with the performance predicted by (6.87), the noise figure of the LNA can actually be improved slightly. As illustrated in Fig. 6.20 for $G_s' = 1/(50\,\Omega)$, the minimum achievable noise figure due to M_1 and M_2 alone is less than 2 dB for a bias current of 4.5 mA. The simulated performance of the LNA is summarized in Table 6.3.

6.3 Frequency Synthesizer

I_D'		4.5 mA
V_{DD}		2.5 V
G_s'		1/(50 Ω)
M_1, M_2	W/L	500 μm × 0.25 μm
	g_m	0.03 Ω^{-1}
	C_{gs}	288 fF
M_3, M_4	W/L	100 μm × 0.25 μm
	g_m	0.017 Ω^{-1}
L_1, L_2	L	4.5 nH
	Q	∞ @ 2 GHz
L_3, L_4	L	0.8 nH
	Q	4 @ 2 GHz
L_5, L_6	L	6 nH
	Q	3.77 @ 2 GHz
input pads	C_p	165 fF
	C_s	165 fF
input bias		827 mV
gain		25.8 dB @ 2 GHz
S_{11}		−21.9 dB @ 2 GHz
NF		3.26 dB

Table 6.3: Summary of LNA simulation.

An excellent and detailed description of the design and implementation of the frequency synthesizer is provided in [31]. This section offers a brief overview of the design along with a discussion of some of the design choices made for low power consumption. The 2-GHz I and Q LO signals are generated using a fully-differential, wide-bandwidth PLL illustrated in Fig. 6.30. The 2-GHz output signal from the VCO is divided by ten to produce the 200-MHz sampling clock for the ADC. This 200-MHz signal is then divided by four and locked to an external 50-MHz reference signal. The PLL has a nominal loop

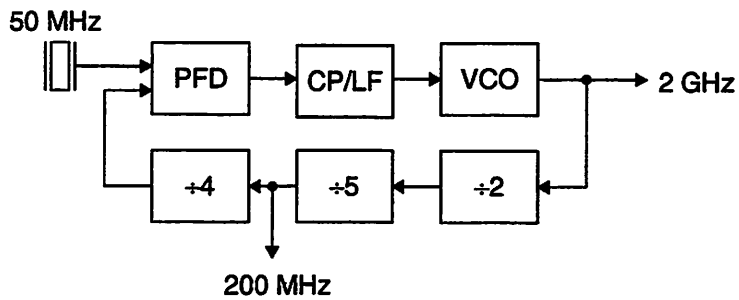


Figure 6.30: Block diagram of phase-locked loop frequency synthesizer.

bandwidth of 3 MHz.

6.3.1 Voltage-Controlled Oscillator

The phase noise performance of an oscillator at an offset frequency $\Delta\omega$ from the center frequency ω_0 can be described by [76]

$$L(\Delta\omega) = 10\log\left\{\frac{2FkT}{P_{sig}}\left[1+\left(\frac{\omega_0}{2Q\Delta\omega}\right)^2\right]\left(1+\frac{\Delta\omega_{1/f^3}}{\Delta\omega}\right)\right\} \tag{6.131}$$

where F is an excess noise factor, P_{sig} is the output power of the oscillator, Q is the quality factor of the tank, and $\Delta\omega$ is the corner frequency between the $1/f^2$ and $1/f^3$ regions. From (6.131), the oscillator phase noise can be improved by increasing P_{sig} as well as by increasing Q. Unfortunately, the lack of high-quality on-chip passive components exacerbates the difficulty of integrating low-power VCOs.

Two commonly used VCO topologies are the LC-tuned oscillator and the ring oscillator.

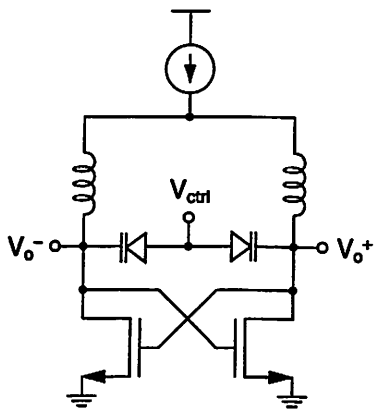


Figure 6.31: Example of a differential LC-tuned VCO.

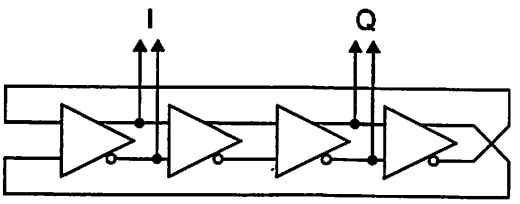


Figure 6.32: Example of a ring-oscillator VCO.

	LC-tuned VCO	ring-oscillator VCO
quality factor	~4	1 – 1.5
area	large	small
quadrature generation	requires additional circuits	inherent

Table 6.4: Comparison of LC-tuned and ring-oscillator VCOs.

An example of a differential LC-tuned VCO is illustrated in Fig. 6.31, while an example of a ring-oscillator VCO is illustrated in Fig. 6.32. The advantages and disadvantages of each of these oscillator topologies are summarized in Table. 6.4. The phase noise performance of LC-tuned oscillators is typically much better than that of ring oscillators due to the higher tank quality factor. Although the quality factor of on-chip spiral inductors is limited to about four for processes which rely on low-resisitivity silicon substrates (Appendix D), the equivalent quality factor of ring oscillators is even less with values ranging from 1 to 1.5 [77]. Nevertheless, for applications with relaxed phase noise requirements, the ring-oscillator VCO has two major advantages. First, the area of the ring-oscillator VCO is much less than that of the LC-tuned VCO, which requires large on-chip spiral inductors for implementing the tank. Second, the ring-oscillator VCO inherently provides the I and Q outputs required for quadrature demodulation. Although the power consumption of the LC-tuned oscillator alone is probably less than that of a ring oscillator for the same phase noise performance, the additional power consumption required for quadrature generation can be significant.

Due to the relaxed phase-noise requirement of the application described here, the VCO in this receiver prototype is implemented as a four-stage ring oscillator illustrated in Fig. 6.33. Additional circuits for quadrature generation are not required, resulting in substantial savings in power consumption. The I and Q outputs are connected directly to the mixer LO ports, while a third output is connected to the frequency divider. Dummy

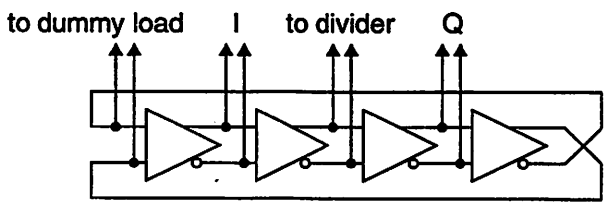


Figure 6.33: Four-stage ring-oscillator VCO.

divider circuits are connected to the remaining VCO output in order to provide load matching for improved quadrature accuracy. Since all users transmit simultaneously in the same frequency band, the frequency synthesizer needs to generate only a single carrier frequency and can be implemented using a wide-bandwidth PLL. The wide loop bandwidth of the PLL suppresses the close-in phase noise of the ring-oscillator VCO [78], thus improving the overall phase noise performance of the frequency synthesizer.

6.3.2 Other Design Considerations

A deadzone-free phase-frequency detector (PFD) is used to increase the pull-in range of the PLL. The PFD outputs rail-to-rail signals to drive the current steering charge pump, reducing the leakage currents that cause spurious tones to appear at the synthesizer output. A passive, second-order loop filter is chosen in order to minimize noise as well as power consumption. In addition, because of the wide loop bandwidth, the passive components used to implement the loop filter do not require significant die area.

Minimizing the amount of signal coupling between the LO and the rest of the receiver is critical when designing frequency synthesizers for highly-integrated implementations. Unwanted signal coupling can occur through numerous mechanisms, including substrate current injection, capacitive coupling to long interconnects, and power supply bounce. In this prototype, the amount of coupling is reduced by implementing the digital portions of the PLL using either source-coupled logic (SCL) or differential cascode voltage switch logic (DCVSL) [79]. Fig. 6.34 illustrates inverter implementations based on these two

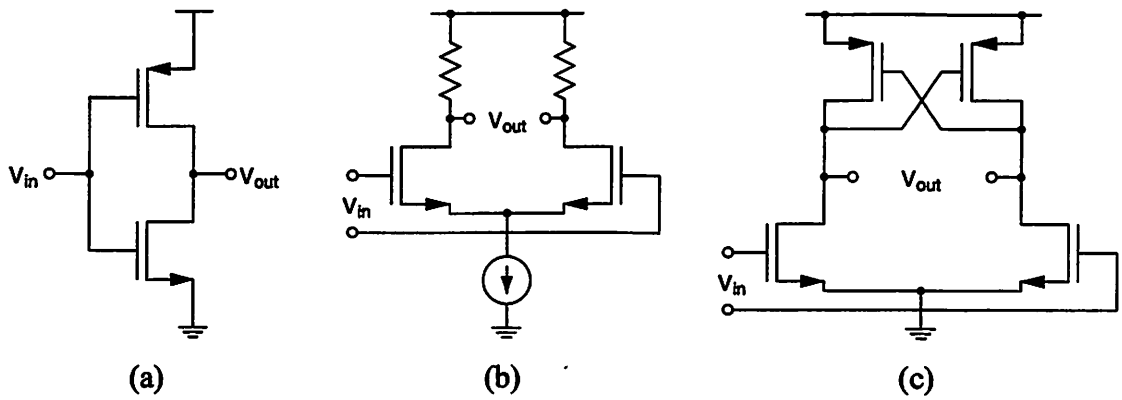


Figure 6.34: Inverter implementations. (a) Static CMOS. (b) SCL. (c) DCVSL.

logic styles along with static CMOS. Substrate current injection caused by charging and discharging capacitors to the substrate can be canceled to first-order by using fully-differential or pseudo-differential circuit topologies, such as SCL or DCVSL, respectively. These two logic styles are also more robust against common-mode noise than single-ended implementations, such as static CMOS. Furthermore, the amount of supply bounce during transitions can be significantly reduced by using SCL, which drains constant current from the supply.

Although SCL circuits consume static power, this logic style is used to realize the VCO and the high-frequency stages of the divider in order to minimize the amount of high-frequency current that can potentially couple into sensitive circuit components such as the LNA. Implementation of the last stage of the divider and the PFD is based on DCVSL, which eliminates static power consumption, but unlike static CMOS, also lessens the substrate current noise. In addition, a separate 2.5-V supply is used for these two blocks in order to further increase the isolation between the digital and analog circuit components. In terms of power consumption, the use of different logic styles is feasible since the overhead required to convert between SCL and DCVSL logic levels is minimal.

6.4 Mixer

An excellent and thorough description of the design and implementation of the I and Q mixers is provided in [74]. This section offers a brief overview of the design along with a discussion of some of the design choices made for low power consumption. The main function of the mixer is to frequency translate the desired RF signal to baseband without corrupting it through mechanisms such as noise or distortion. Two different types of mixers are passive mixers and active mixers. An example of a CMOS passive mixer is illustrated in Fig. 6.35 [68], while an example of a CMOS active mixer is illustrated in Fig. 6.36. The following criteria must be considered when selecting a mixer topology for a particular application: linearity, noise, conversion gain, power consumption, and port isolation.

6.4.1 Passive Mixers

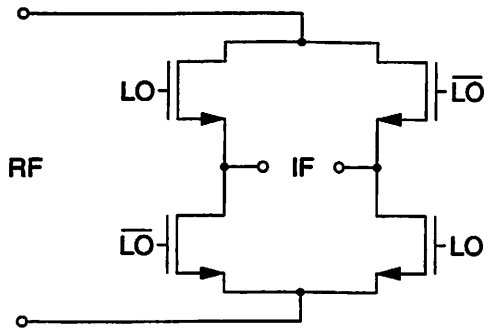


Figure 6.35: CMOS double-balanced passive mixer.

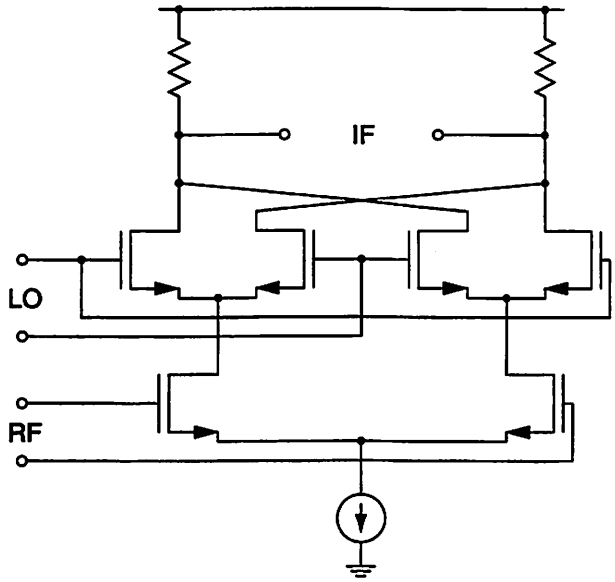


Figure 6.36: CMOS double-balanced current-commutating mixer.

Passive mixers do not provide any conversion gain, and in fact, actually result in conversion loss. In direct-conversion receivers, the desired signal can still be relatively weak at the input to the mixer, and this conversion loss imposes more stringent noise requirements in the subsequent baseband circuits. Although passive mixers introduce thermal noise, they do not introduce any flicker noise due to the absence of current, and thus passive mixers are potentially attractive for use in direct-conversion receivers.

Low power consumption and excellent linearity performance [80] are two additional advantages of passive mixers. However, both of these advantages are negated if an additional amplifier is needed to compensate for the conversion loss of the passive mixer.

Finally, the port-to-port isolation of passive mixers is poor. In particular, the LO signal can couple to the RF port through the gate capacitances of the switches. Poor LO-to-RF isolation is unacceptable in direct-conversion receivers, which are susceptible to problems such as LO radiation from the antenna as well as DC offsets resulting from LO self-mixing.

6.4.2 Active Mixers

A popular active mixer is the double-balanced current-commutating mixer (Fig. 6.36) based on the Gilbert cell multiplier [81]. Unlike passive mixers, active mixers actually provide conversion gain, which relaxes the noise and gain requirements in the subsequent baseband circuits. However, active mixers contribute both thermal noise and flicker noise, which may be problematic in direct-conversion receivers.

As already mentioned in Section 6.4.1, the power consumption and linearity of passive mixers alone are generally superior to that of active mixers. However, a fair comparison must also include the power consumption and linearity of any additional amplifiers to compensate for the lack of conversion gain in passive mixers. In this case the, the choice between an active mixer and a passive mixer is not as clear-cut.

Finally, the port-to-port isolation of current-commutating mixers is generally better than that of passive mixers. In particular, the path between the LO and RF ports is separated by two transistors rather than just one.

6.4.3 Mixer Implementation

For this receiver prototype, the RF signal is frequency translated to baseband along parallel I and Q signal paths using a pair of double-balanced current-commutating mixers. The I and Q mixers are both based on the same topology illustrated in Fig. 6.37. An active mixer is selected because of its ability to provide conversion gain, and a double-balanced configuration is chosen in order to increase the mixer's immunity to common-mode variations, including substrate noise introduced by the digital sections of the receiver. In addition, this topology provides excellent isolation between the LO and

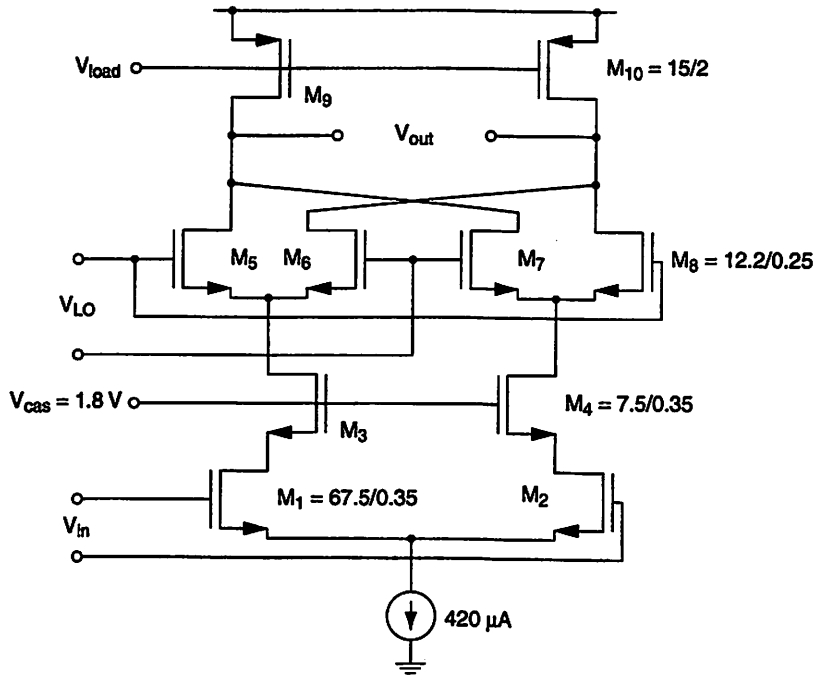


Figure 6.37: Mixer topology used in the receiver prototype.

RF ports of the mixer, which is further improved by adding cascode transistors M_3 and M_4.

The gates of the switching transistors $M_5 - M_8$ are connected directly to the I and Q outputs of the VCO. By dc coupling the LO signals to the input of the mixer and by locating these transistors immediately adjacent to the VCO in the layout, the capacitive loading on the VCO is minimized. Consequently, the need for clock buffers is avoided, resulting in substantial power savings. The sizes of transistors $M_5 - M_8$ are chosen as a compromise between flicker noise performance of the mixer and power consumption in the VCO. Since the VCO outputs are directly connected to these transistors, small device dimensions are desirable in order to reduce the capacitive loading on the VCO. On the other hand, large device dimensions are desirable for improved flicker noise performance. Although the receiver prototype is based on a direct-conversion architecture, the wide bandwidth of the desired signal reduces the impact of flicker noise on the performance of this system. Consequently, the mixer's flicker noise requirement is relaxed in favor of reduced power consumption in the frequency synthesizer.

Finally, the load devices M_9 and M_{10} are biased in the linear region and their resistances can be changed by adjusting their gate bias voltages to provide variable gain capability [11].

6.5 Baseband Amplification and Filtering

Each of the I and Q baseband signals must be amplified and filtered before subsequent analog-to-digital conversion. Although the VGAs are not implemented in this receiver prototype, the baseband section still must provide a moderate amount of fixed gain since the gain from the LNA and mixer alone is not sufficient to meet the minimum gain requirement of this system. The baseband section of this direct-conversion receiver also provides high-pass filtering for dc-offset removal. Moreover, low-pass filtering provides rejection of out-of-band interferers as well as anti-alias filtering for the subsequent ADCs.

6.5.1 Low-Pass Filtering

For the receiver prototype, a Sallen and Key section is used for low-pass filtering. Alternative approaches include switched-capacitor techniques as well as other continuous-time techniques such as MOSFET-*C* or transconductance-*C* filters. In general, a continuous-time approach based on Sallen and Key sections is more appropriate for applications with relaxed filtering requirements and results in very simple, low-power implementations. On the other hand, switched-capacitor techniques and continuous-time techniques based on MOSFET-*C* or transconductance-*C* filters are more appropriate for applications with increased selectivity requirements. A brief discussion of each of these analog filtering techniques follows.

Sallen and Key Filter. A block diagram of a Sallen and Key section based on an amplifier with gain K is illustrated in Fig. 6.38 [82], [21]. The voltage transfer function is given by

$$\frac{V_{out}}{V_{in}} = \frac{a_0 K}{s^2 + a_1 s + a_0} \tag{6.132}$$

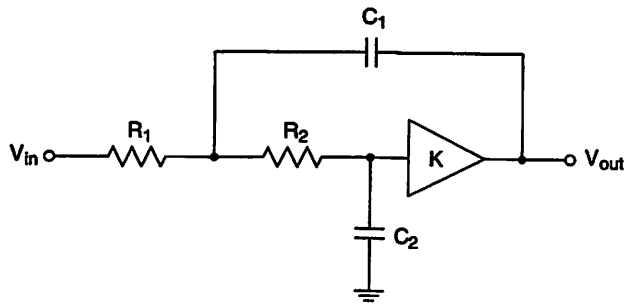


Figure 6.38: Sallen and Key filter block diagram.

where a_0 and a_1 are given by, respectively,

$$a_0 = \frac{1}{R_1 R_2 C_1 C_2} \tag{6.133}$$

$$a_1 = \frac{1}{R_1 C_1} + \frac{1}{R_2 C_1} + \frac{1}{R_2 C_2}(1-K). \tag{6.134}$$

Thus, the Sallen and Key section can be used to realize a second-order response using only a single active gain element. In fact, for $K = 1$, the active gain element can be implemented simply as a voltage follower. Although the Sallen and Key section lends itself to very simple circuit implementations, the response of this filter is particularly sensitive to component variations for poles with high quality factors [83], where the quality factor of a complex pole pair is a measure of the distance of the poles from the imaginary axis in the s plane as illustrated in Fig. 6.39. Thus, this approach may be problematic when used to implement higher-order filter responses which require complex

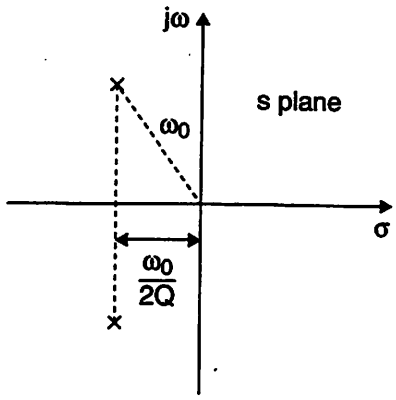


Figure 6.39: Quality factor of a complex pole pair.

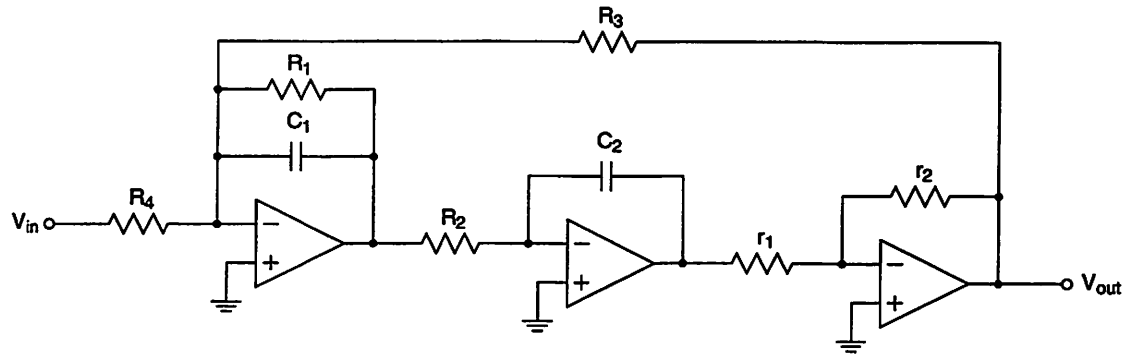


Figure 6.40: Tow-Thomas biquad.

pole pairs with high quality factors.

MOSFET-C and Transconductance-C Filters. Lower sensitivities to component variations may be achieved for higher-order filters by implementing a second-order transfer function using additional active gain elements, such as the Tow-Thomas biquad illustrated in Fig. 6.40 [20], [83], [84]. The voltage transfer function is given by

$$\frac{V_{out}}{V_{in}} = -\frac{\dfrac{r_2}{r_1}\dfrac{1}{R_2R_4C_1C_2}}{s^2 + \dfrac{1}{R_1C_1}s + \dfrac{r_2}{r_1}\dfrac{1}{R_2R_3C_1C_2}}. \tag{6.135}$$

The key building block for this biquad is the active *RC* integrator illustrated Fig. 6.41a which has the following transfer function:

$$\frac{V_{out}}{V_{in}} = -\frac{1}{sRC}. \tag{6.136}$$

Implementations of this *RC* integrator based on continuous-time techniques include the

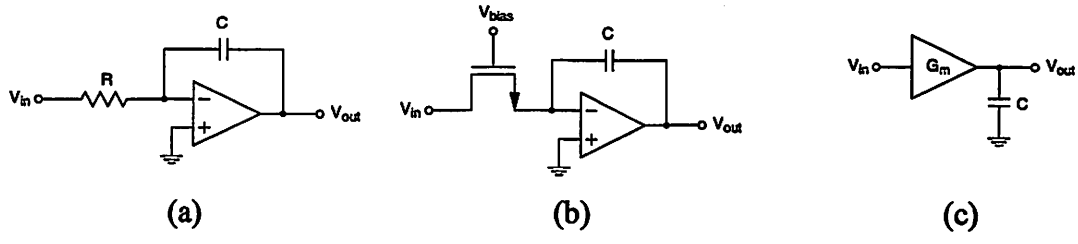


Figure 6.41: Integrators. (a) *RC*. (b) MOSFET-*C*. (c) Transconductance-*C*.

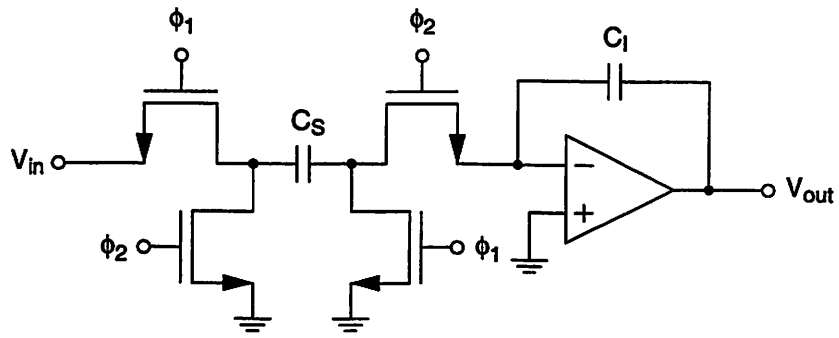


Figure 6.42: Switched-capacitor integrator.

MOSFET-*C* integrator (Fig. 6.41b) [85] and the transconductance-*C* integrator (Fig. 6.41c) [86], [87]. In the former case, the resistance *R* is implemented using an MOS transistor operating in the linear region, and the resistance can be adjusted through the transistor gate bias voltage, V_{bias}, resulting in the ability to compensate for process variations. Alternatively, the *RC* integrator can be implemented using a transconductance element, resulting in the following integrator transfer function.

$$\frac{V_{out}}{V_{in}} = \frac{G_m}{sC}. \tag{6.137}$$

In this case, the effective resistance can be adjusted through the bias current of the transconductance element.

Switched-Capacitor Filter. The *RC* integrator can also be implemented using switched-capacitor techniques as illustrated in Fig. 6.42 [88], [89]. The MOS switches are driven by a nonoverlapping clock with frequency *f* and phases ϕ_1 and ϕ_2. In this case, the time constant of the integrator is

$$\tau = \frac{C_I}{fC_S} \tag{6.138}$$

and is dependent only on the clock frequency and the ratio of the integrating and sampling capacitors. Consequently, filters based on this technique are very robust against process variations and do not require additional tuning.

Comparison of Filter Implementation Techniques. Sallen and Key sections are very amenable to simple, low-power circuit implementations. However, the response of high-order filters based on this technique is particularly sensitive to component variations, and thus, the use of Sallen and Key sections is most appropriate for applications with relaxed filtering requirements.

For higher-order filter responses, lower sensitivity to component variations may be achieved by using switched-capacitor techniques as well as continuous-time techniques based on MOSFET-*C* and transconductance-*C* integrators. The main tradeoff for this reduced sensitivity is increased power consumption. For MOSFET-*C* and transconductance-*C* filters, tuning is still required in order to achieve very accurate filter responses. On the other hand, switched-capacitor filters are very robust against process variations, and consequently, very accurate filter responses can be achieved without the need for tuning. Finally, continuous-time techniques based on MOSFET-*C* and transconductance-*C* integrators are more appropriate for applications with higher bandwidth requirements, while switched-capacitor techniques are more appropriate for applications with lower bandwidth requirements.

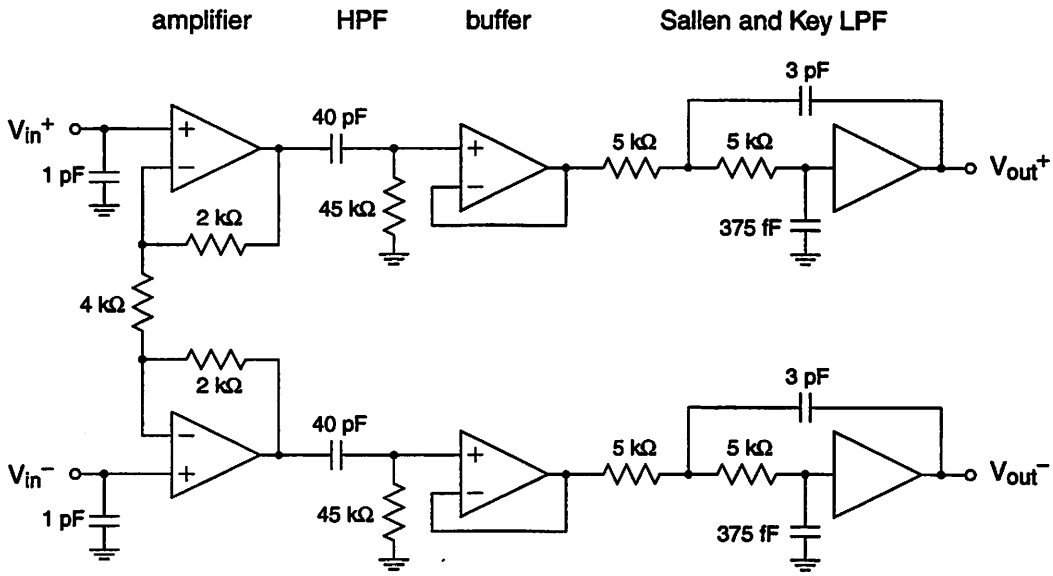


Figure 6.43: Block diagram of baseband amplification and filtering.

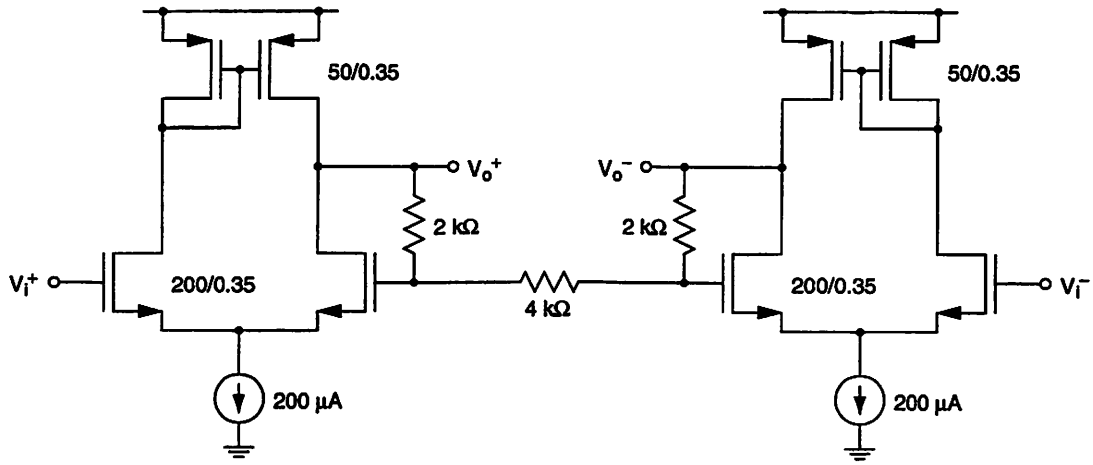


Figure 6.44: Noninverting amplifier schematic.

6.6 Implementation of Baseband Amplifiers and Filters

The I and Q baseband sections are each based on the block diagram illustrated in Fig. 6.43. All of the baseband circuits are implemented using large transistor sizes in order to reduce the amount of flicker noise. Immediately after frequency translation, shunt 1-pF capacitors in combination with the mixer output impedance provide first-order low-pass filtering of each of the baseband I and Q signals. A noninverting amplifier then provides moderate gain in order to reduce the impact of noise contributed by subsequent stages. The circuit schematic of this amplifier is illustrated in Fig. 6.44. The input of this amplifier is dc coupled to the mixer output, which sets the common-mode bias voltage at a nominal value of 1.9 V.

Each of the baseband signals then passes through a first-order high-pass filter, which removes dc offsets and flicker noise from previous receiver stages. System-level simulations reveal that the SNR degradation is less than 0.5 dB for a high-pass corner frequency of up to 500 kHz. However, a much lower corner frequency is implemented in order to further reduce the SNR degradation as well as to account for process variations. Each filter is realized using on-chip passive structures, which include a pair of 40-pF capacitors and a pair of 45-kΩ resistors, placing the high-pass corner frequency at about 90 kHz. The on-chip resistors and capacitors are implemented using unsalicided n+-poly and poly/n-well structures, respectively. The poly/n-well structure operating in

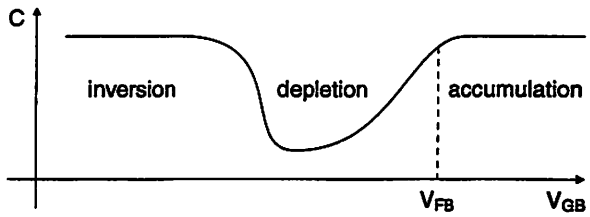


Figure 6.45: Capacitance of poly/n-well structure versus bias voltage.

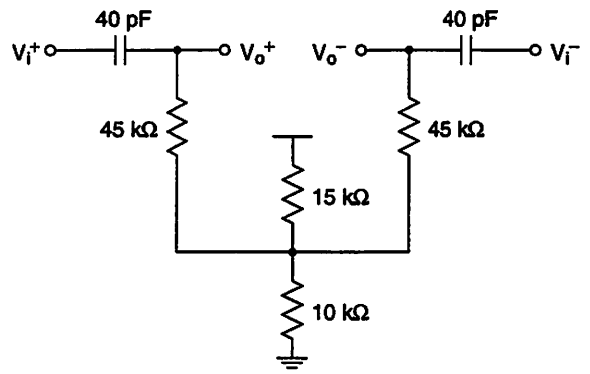


Figure 6.46: High-pass filter schematic.

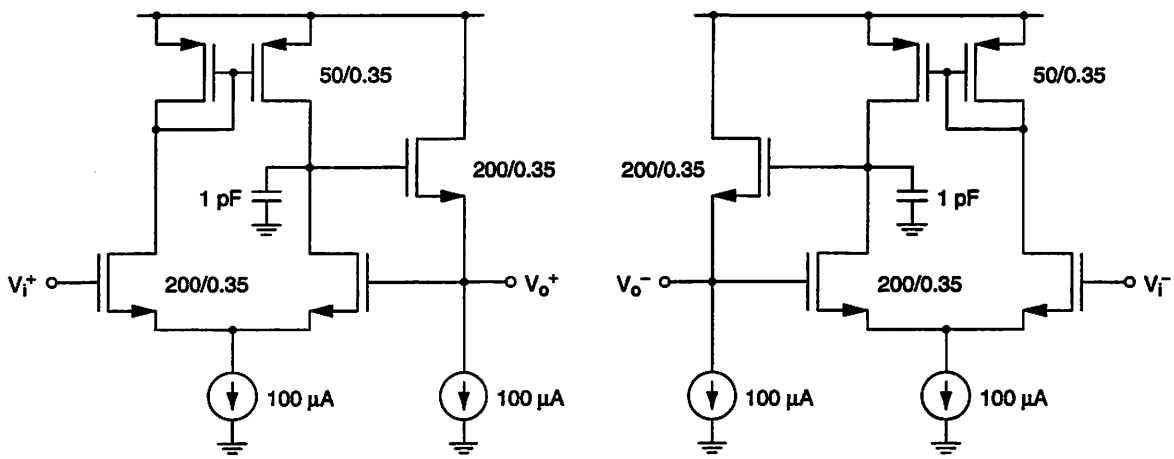


Figure 6.47: Buffer schematic.

accumulation as illustrated in Fig. 6.45 [90] offers a large capacitive density of about 6 fF/µm^2 when biased above the flat-band voltage, V_{FB}, of about 90 mV. The circuit schematic of the high-pass filter is illustrated in Fig. 6.46. The resistor string sets the common-mode bias voltage at the filter output. Large resistors are used in order to minimize the power consumed by the bias string. The high-pass filter is followed by the unity-gain buffer illustrated in Fig. 6.47.

pole location ($\times\omega_0$)	quality factor
-1.0	0.5
$-0.5 \pm j0.8660254$	1.0

Table 6.5: Poles for a third-order Butterworth low-pass frequency response.

Next, each of the baseband signals passes through a second-order Sallen and Key low-pass filter. This low-pass filter provides attenuation of out-of-band interferers as well as anti-alias filtering for the subsequent ADC. Although implementations based on continuous-time filtering techniques are susceptible to variations in component values due to process variations, such variations are not as critical for this particular application due to the relaxed selectivity requirements. Moreover, since the baseband filtering is followed by a ΣΔ ADC operating at 200 MHz, the anti-alias filtering requirements are also much less stringent. The main requirement of the low-pass filter is that the corner frequency should be no less than the 16.25-MHz single-sided bandwidth of the desired signal.

The poles of the Sallen and Key filter in combination with the pole at the mixer output provide an overall third-order Butterworth low-pass frequency response. The maximally flat gain and very linear phase response of the Butterworth filter result in very little signal distortion. When normalized to the corner frequency, ω_0, the poles of a third-order Butterworth low-pass response are given by the roots of the equation [20]

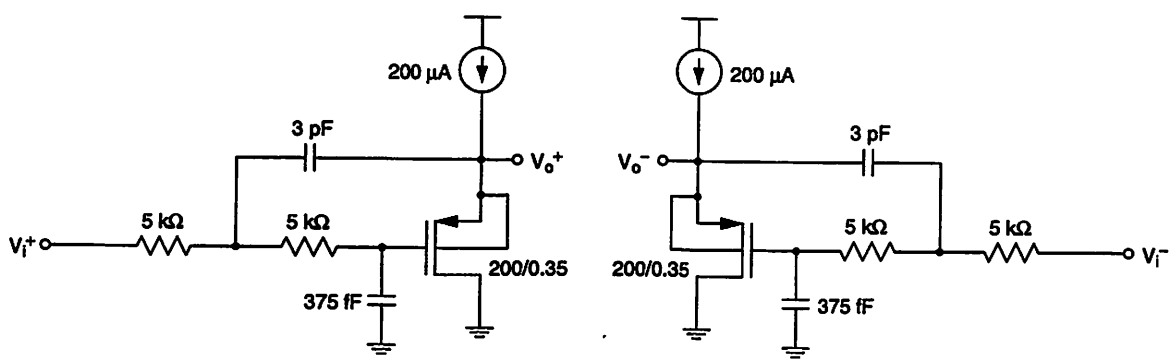


Figure 6.48: Sallen and Key filter circuit schematic.

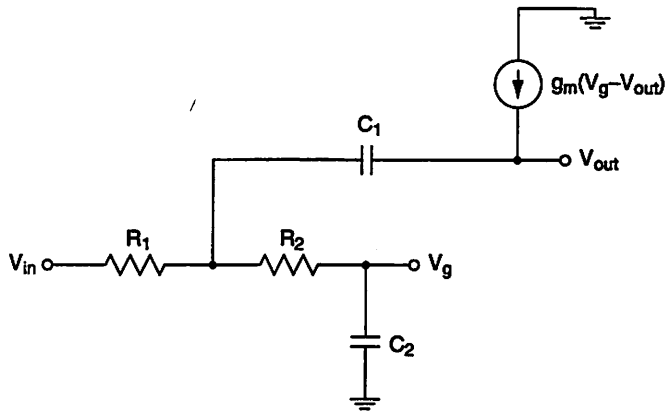


Figure 6.49: Equivalent half-circuit of Sallen and Key filter

$$\begin{aligned} s^3 + 2s^2 + 2s + 1 &= 0 \\ (s+1)(s^2 + s + 1) &= 0 \end{aligned} \tag{6.139}$$

and the pole locations are summarized in Table 6.5. The circuit schematic of the Sallen and Key filter is illustrated in Fig. 6.48. The gain element of the Sallen and Key filter is implemented using a PMOS voltage follower [91]. In this process the source and bulk nodes of PMOS devices can be connected together in order to eliminate the body effect. In this case, the small-signal model of the equivalent half-circuit is illustrated in Fig. 6.49 and the gain of the voltage follower is

$$K = \frac{V_{out}}{V_g} = 1 + \frac{R_2 C_1 C_2 s^2}{C_1 s + g_m}. \tag{6.140}$$

Substituting (6.140) into (6.132) results in the following expression for the voltage transfer function of this filter:

$$\frac{V_{out}}{V_{in}} = \frac{\dfrac{R_2 C_1 C_2}{g_m} s^2 + \dfrac{C_1}{g_m} s + 1}{\left[\dfrac{(R_1 + R_2) C_1 C_2}{g_m} + R_1 R_2 C_1 C_2\right] s^2 + \left[\dfrac{C_1}{g_m} + (R_1 + R_2) C_2\right] s + 1}. \tag{6.141}$$

The passive component values are determined by setting

$$\frac{(R_1 + R_2) C_1 C_2}{g_m} + R_1 R_2 C_1 C_2 = \frac{1}{\omega_0^2} \tag{6.142}$$

$$\frac{C_1}{g_m} + (R_1 + R_2)C_2 = \frac{1}{\omega_0}. \tag{6.143}$$

For $R_1 = R_2 = 5\ \text{k}\Omega$, $\omega_0 = 2\pi(16.25\,\text{MHz})$, and $1/g_m = 372\ \Omega$, the corresponding values for C_1 and C_2 are 4 pF and 830 fF, respectively. However, in order to account for process variations as well as parasitic capacitances, the nominal values of C_1 and C_2 were chosen to be 3 pF and 375 fF, respectively. For these selected component values, simulations over process corners indicate that the –3-dB corner frequency varies between 17 MHz and 22 MHz. The resistors and capacitors are implemented using unsalicided n+-poly and poly/n-well structures, respectively.

The transfer function in (6.141) also contains zeros which can affect the desired Butterworth frequency response. The zero locations are given by

$$s = -\frac{1}{2R_2C_2}\left(1 \pm j\sqrt{4g_mR_2\frac{C_2}{C_1} - 1}\right). \tag{6.144}$$

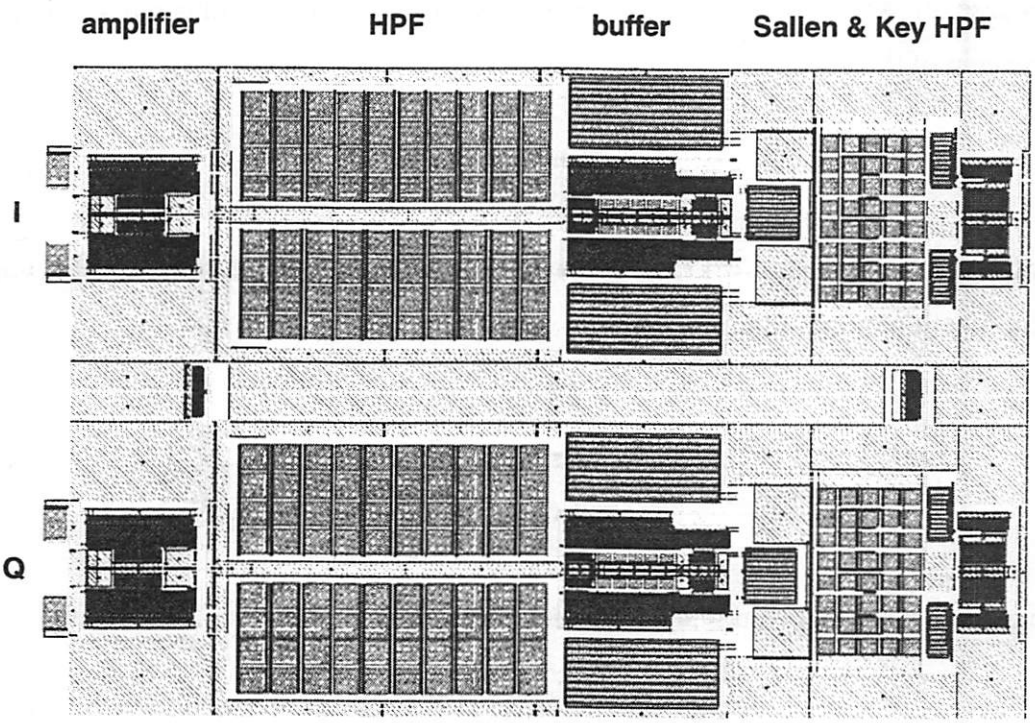


Figure 6.50: Layout of baseband amplifiers and filters.

For $R_1 = R_2 = 5$ kΩ, $\omega_0 = 2\pi(16.25\,\text{MHz})$, $1/g_m = 372$ Ω, $C_1 = 4$ pF, and $C_2 = 830$ fF, the zero locations are $s = -1.2 \times 10^8 \pm j3.8 \times 10^8$. The magnitude of these zeros is approximately 63 MHz and is sufficiently greater than the corner frequency of 16.25 MHz so as not to adversely affect the desired Butterworth frequency response. Nevertheless, these zeros can still be problematic since they decrease the out-of-band attenuation of the filter. However, the inevitable presence of higher-order poles mitigates the severity of this problem.

Finally, the layout of the baseband amplifiers and filters is illustrated in Fig. 6.50. The symmetric layout helps to improve the matching between the I and Q sections.

6.7 Analog-to-Digital Converter

An excellent and detailed description of the design and implementation of the I and Q ADCs is provided in [32]. This section offers a brief overview of the design along with a discussion of some of the design choices made for low power consumption. The ADC requirements are determined in Section 5.3.8 and summarized in Table. 5.1. For this system, each of the ADCs must have a Nyquist rate of at least 25 MHz and a resolution of at least 7 bits.

6.7.1 Pipeline Architecture

For these specifications, one possible approach of implementing the ADC is to use a pipeline architecture illustrated in Fig. 6.51 [92]. In this architecture, each of the N stages samples the signal from the previous stage and quantizes it to B bits. The quantized signal is then subtracted from the input signal and the result is amplified by an interstage gain block before being sampled by the next stage. This architecture is particularly amenable to low-power implementations. First, for an $N \times B$-bit ADC, this architecture requires $N \times 2^B$ comparators compared to $2^{N \times B}$ comparators for a flash architecture. Second and more importantly, the circuit requirements of subsequent stages are relaxed, and capacitive scaling techniques can be used, resulting in significant power savings. Finally,

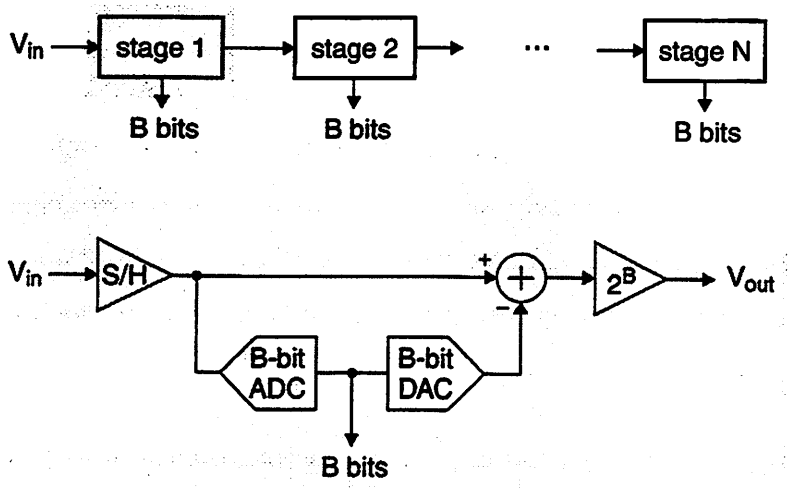


Figure 6.51: Pipeline ADC architecture.

one potential disadvantage of the pipeline ADC architecture, as with any pipeline architecture, is latency.

Several low-power ADCs have been implemented based on the pipeline architecture with capacitive scaling, including a 10-bit, 20-MS/s ADC consuming 35 mW [93] and a 10-bit, 40-MS/s ADC consuming 28 mW [94],[95]. The former was implemented in a 1.2-μm process and the latter was implemented in a 0.6-μm process.

Although the pipeline architecture is a promising approach for implementing a low-power ADC with the minimum requirements specified for this system, other factors must be considered in order to minimize the overall power consumption of the entire receiver. In particular, one of the most critical receiver functions is timing recovery, and a more efficient receiver implementation may be achieved when the ADC is designed in conjunction with the timing recovery algorithm.

6.7.2 Timing Recovery Considerations

In order to achieve an even more efficient receiver implementation, receiver functions must be carefully partitioned between the analog and digital hardware. Despite the relatively high Nyquist rate requirement of the ADC, a ΣΔ modulator is used for analog-to-digital conversion in this receiver. This ΣΔ ADC is designed in conjunction with an

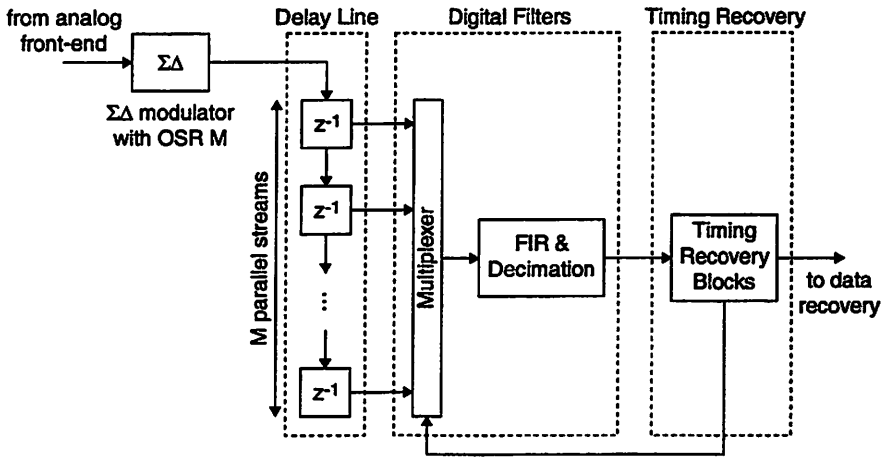


Figure 6.52: ΣΔ-assisted timing recovery scheme.

all-digital timing recovery algorithm, and when both the analog and digital hardware are considered, this approach results in a very efficient implementation.

The use of different frequency references in the base-station transmitter and the mobile receiver introduces a frequency offset between the two LOs, and consequently, frequency estimation and compensation must be performed at the receiver before the data can be recovered. In order to provide adequate granularity for digital timing recovery, receivers typically oversample the input signal by at least two times the Nyquist rate [19]. Since oversampling is an inherent property of ΣΔ modulators, a ΣΔ ADC is an attractive approach for analog-to-digital conversion [96]. A block diagram of the proposed timing recovery algorithm for this receiver is illustrated in Fig. 6.52. The output of the ΣΔ modulator before decimation is a low-resolution, oversampled version of the data signal. Timing recovery can be performed by properly adjusting the phase of the digital decimation filter following the modulator. By including a variable-length delay line before the decimation filter, the timing recovery block can control the effective phase of the ADC sampling instant to within the granularity provided by the oversampling ratio (OSR). In order to achieve the same timing granularity with a pipeline ADC, a converter with a sampling rate of OSR × 25 MHz would be required. Thus, when timing recovery issues are also considered, a pipeline ADC is no longer the definitive choice for low power consumption.

6.7.3 Sigma-Delta Analog-to-Digital Converter

In the receiver prototype, the baseband I and Q signals are digitized using a pair of 7-bit, 25-MS/s ΣΔ ADCs operating at 200 MHz. Since the high Nyquist rate of the baseband signals restricts the ΣΔ converter to a low OSR of 8, the required dynamic range is achieved by using a 2-1-1 cascade architecture with single-bit quantization in each stage (Fig. 6.53).

The integrators are implemented as fully-differential switched-capacitor circuits using folded-cascode operational amplifiers with NMOS input devices to maximize speed as illustrated in Fig. 6.54. The device sizes and bias points of each amplifier are optimized for minimum power consumption. Power consumption in the ΣΔ ADCs is further reduced by using capacitive scaling techniques [97]. However, the presence of parasitic capacitances limits the achievable power savings resulting from this approach. The capacitor values and bias current of each of the four integrators are summarized in Table 6.6.

6.8 Receiver Test Chips

A pair of test chips were fabricated to characterize the performance of the individual circuit components as well as the performance of the entire receiver. A micrograph of the first test chip (SCRRX) is illustrated in Fig. 6.55. This test chip includes the analog receiver consisting of the LNA, the mixer, the PLL, and the baseband circuits for amplification and filtering. Separate test circuits for the LNA, the mixer, and the PLL as well as a test circuit which consists of just the LNA and the mixer are also included. Test structures were also fabricated in order to facilitate the characterization of inductors used in the implementation of the LNA.

A micrograph of the second test chip (SCRBARF) is illustrated in Fig. 6.56. The complete direct-conversion receiver including the ADCs is fabricated on this test chip. The area of this chip is 5.0 mm × 5.2 mm including bond pads, but the circuits alone require only about 5 mm^2.

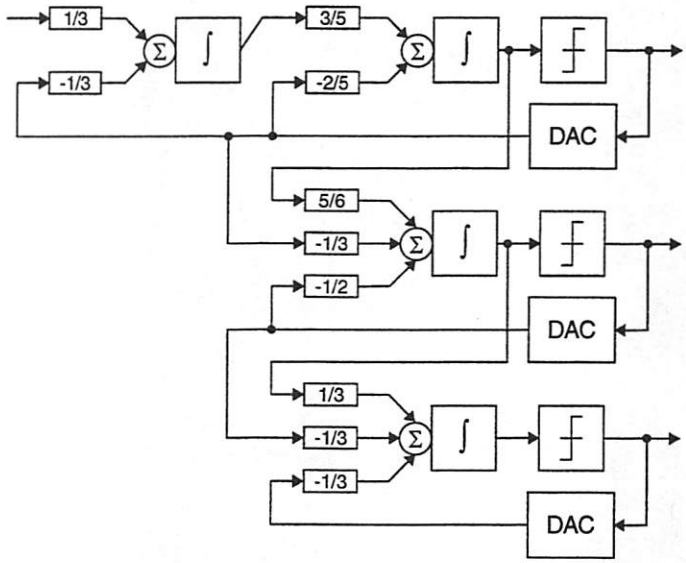


Figure 6.53: 2-1-1 cascade ΣΔ architecture.

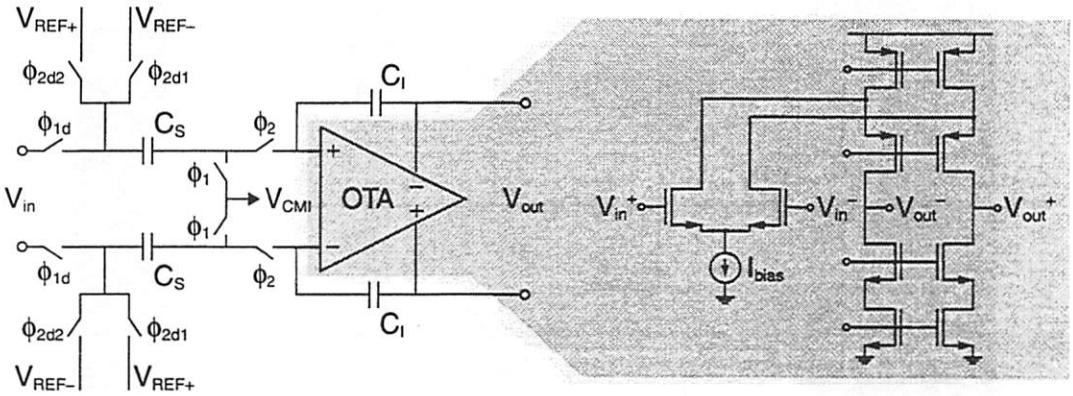


Figure 6.54: Switched-capacitor integrator.

Integrator	C_S (fF)	C_I (fF)	I_{bias} (μA)
1	90	270	270
2	90	150	270
3	87.5	105	240
4	75	150	138

Table 6.6: Capacitor values and bias currents for ΣΔ ADC.

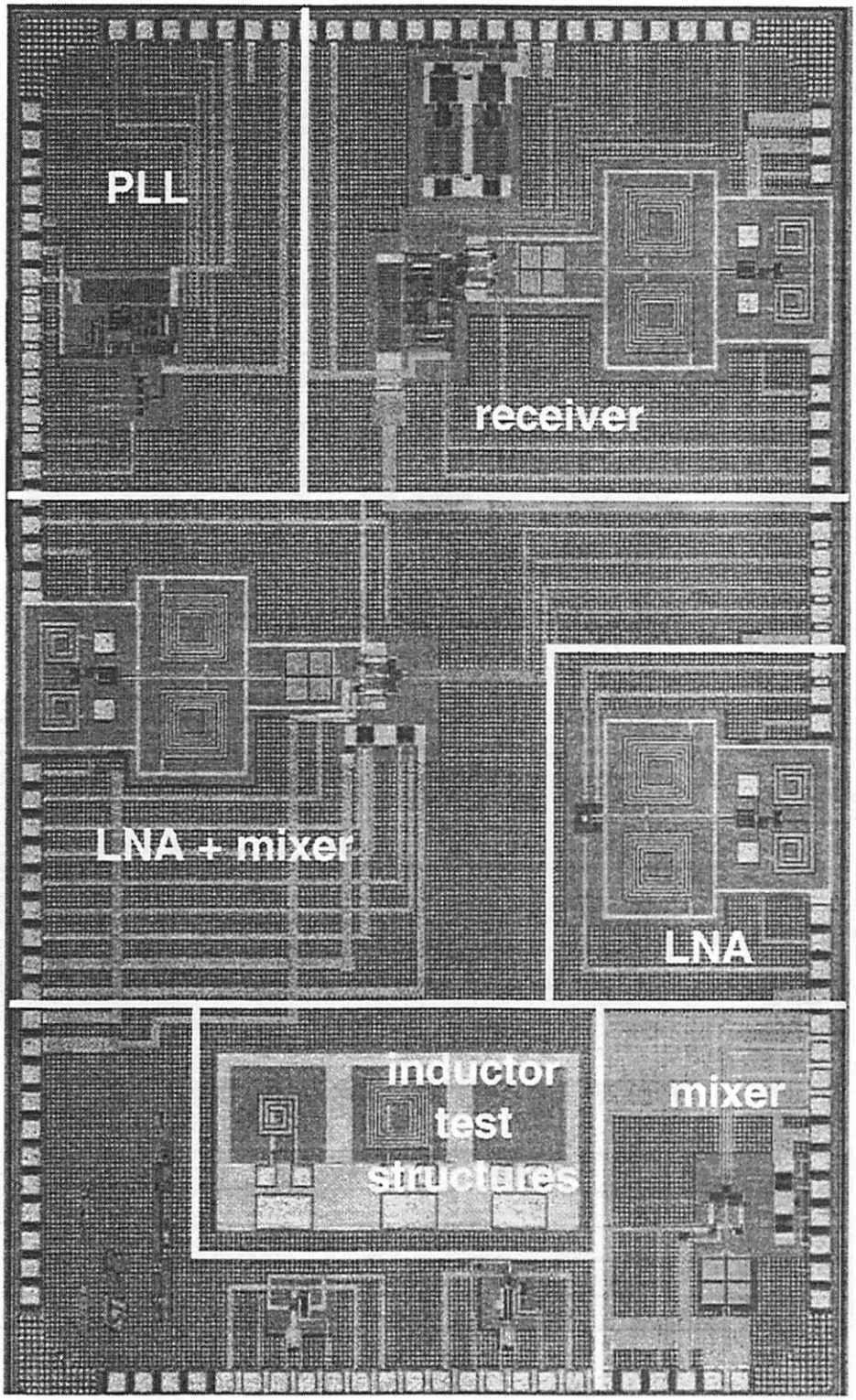


Figure 6.55: Test chip (SCRRX) micrograph.

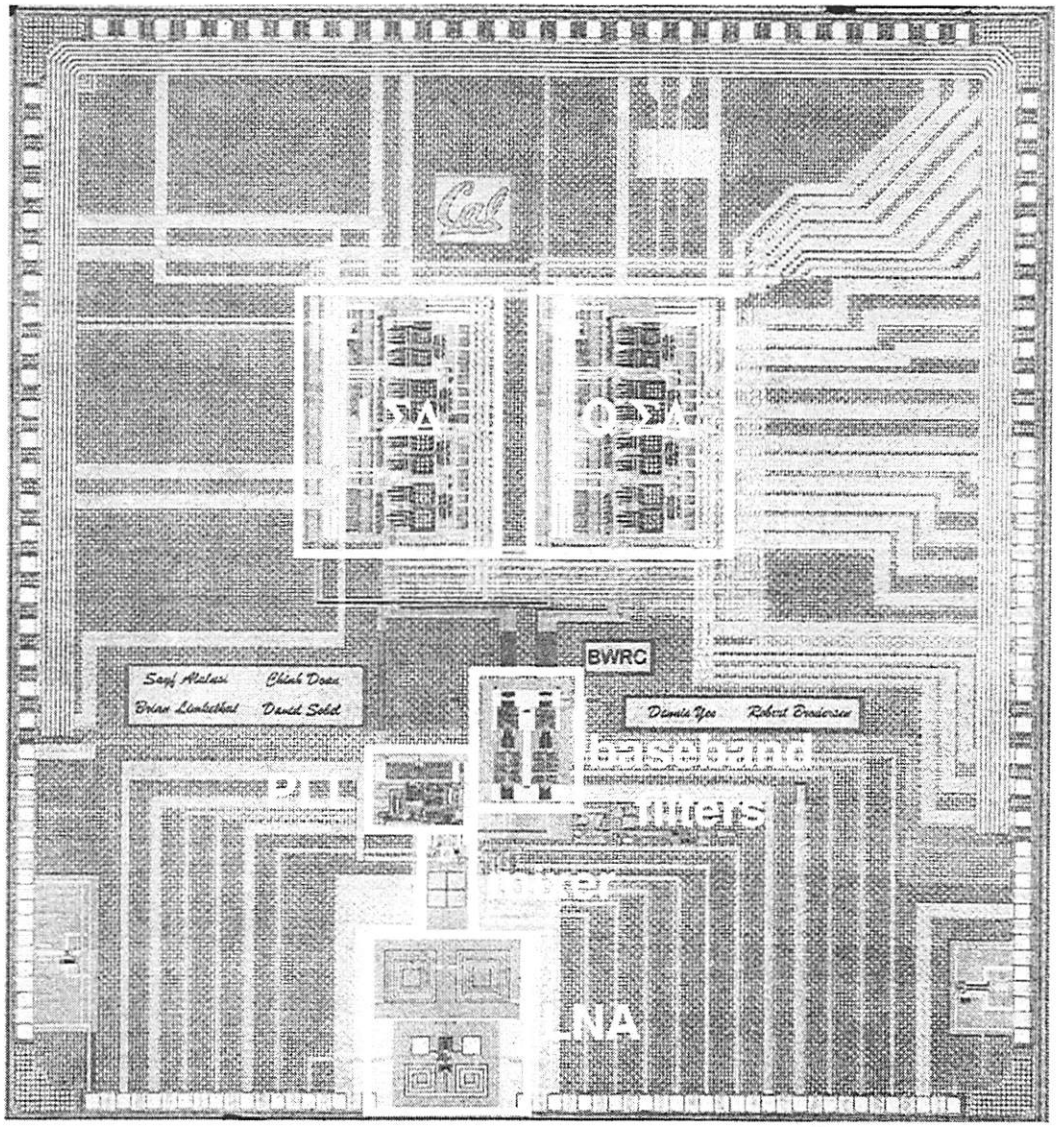


Figure 6.56: Test chip (SCRBARF) micrograph.

Chapter 7

Simulated Performance and Measurement Results

7.1 Simulated Performance

Extensive simulations were performed on the individual building blocks and many of these results are reported elsewhere. Simulation results for the LNA are reported in [98], the mixer in [74], the PLL in [31], and the ADC in [34]. In this section, the simulation results for the entire receiver are reported.

7.1.1 LNA/Mixer/Baseband Simulations

Most of the simulations were performed without the PLL and the ΣΔ ADCs since long simulation times are required when these blocks are included. Instead, only the LNA, the mixers, and the baseband amplifiers and filters are included. For these simulations, the LO signal is modeled as a 2-GHz noiseless sinusoid with an amplitude of 0.4 V. All simulations were performed using netlists extracted from the layout in order to include the effects of parasitic capacitances. In addition, all simulations were simulated across process corners: typical, fast, and slow. The gain at the output of the Sallen and Key filter is plotted versus frequency in Fig. 7.1, and a few characteristics of the frequency response are summarized in Table 7.1. The first-order high-pass filter and the third-order

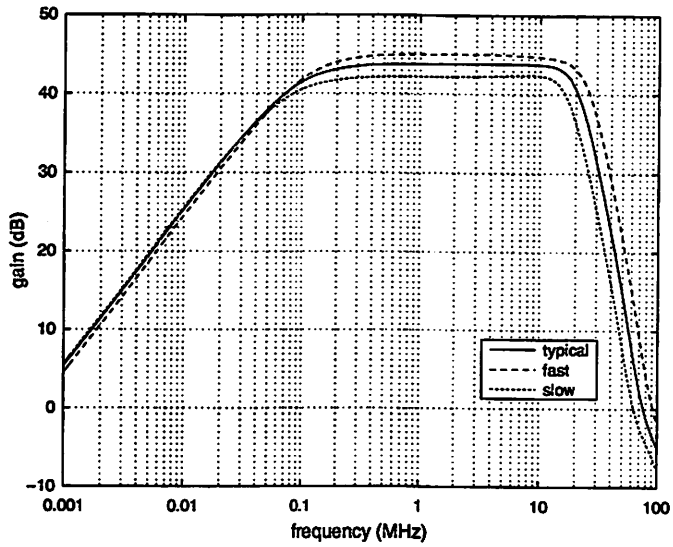


Figure 7.1: Simulated frequency response at the output of the Sallen and Key filter over process corners.

	typical	fast	slow
gain @ 1 MHz	43.8 dB	45.0 dB	42.2 dB
−3-dB frequency (HPF)	82.9 kHz	104.4 kHz	67.0 kHz
−3-dB frequency (LPF)	19.7 MHz	22.5 MHz	17.1 MHz

Table 7.1: Summary of receiver frequency response.

Butterworth low-pass filter responses are evident in the overall frequency response of the receiver.

For the typical process corner, the simulated frequency responses at the outputs of various receiver building blocks are illustrated in Fig. 7.2. At 1 MHz, the gains at the outputs of the mixer, the amplifier, the high-pass filter, and the Sallen and Key low-pass filter are 41.0 dB, 44.8 dB, 43.9 dB, and 43.8 dB, respectively.

The simulated noise performance of the receiver is summarized in Table 7.2. For the typical, fast, and slow process corners, the noise figures are 5.52 dB, 4.81 dB, and 6.26 dB, respectively.

The distortion performance of the receiver is summarized in Table 7.3. The −1-dB compression points are −34 dBm, −39 dBm, and −33 dBm for the typical, fast, and slow process corners.

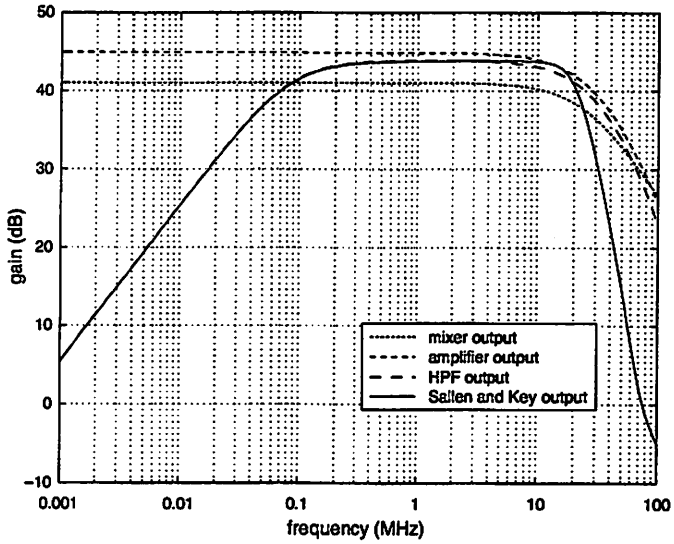


Figure 7.2: Simulated frequency responses at the outputs of various receiver components (typical process corner).

	typical	fast	slow
output noise due to source	448 μV	546 μV	352 μV
total output noise	845 μV	950 μV	722 μV
noise figure	5.52 dB	4.81 dB	6.26 dB

Table 7.2: Summary of simulated receiver noise performance (output noise integrated over 100 MHz).

	typical	fast	slow
−1-dB compression point	−34 dBm	−39 dBm	−33 dBm

Table 7.3: Summary of simulated receiver distortion performance.

Finally, a transient envelop simulation was performed using SpectreRF. The I and Q baseband transmit signals are first generated in Simulink (Fig. 7.3). Next, these signals are frequency translated in quadrature to RF using a behavioral model in SpectreRF. The spectra of the receiver output signals for input power levels of −43 dBm and −33 dBm are illustrated in Fig. 7.4. For an input power level of −33 dBm, the output signal is distorted due to compression.

7.1.2 LNA/Mixer/PLL/Baseband Simulation

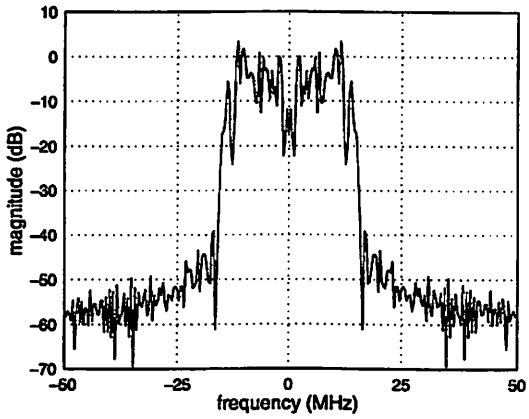


Figure 7.3: Baseband transmit spectrum from Simulink.

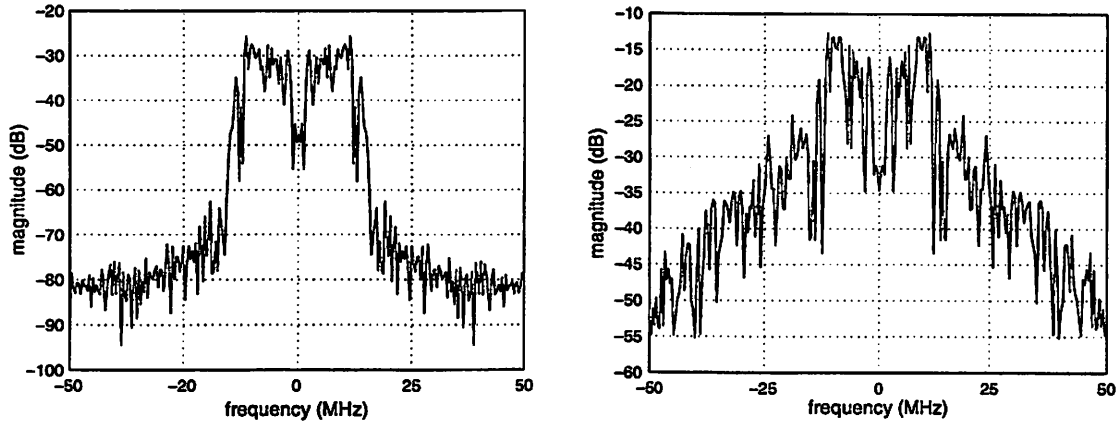


Figure 7.4: Spectrum of receiver output signal from transient envelop simulations. (a) –43-dBm input power. (b) –33-dBm input power.

When including the PLL along with the LNA, the mixers, and the baseband amplifiers and filters, the simulation times are much longer. Consequently, only a single simulation was performed in order to characterize the noise performance for the typical process corner. The simulated noise performance is summarized in Table 7.4. When including the PLL, the noise figure increases slightly to 5.7 dB compared to 5.52 dB without the PLL.

7.1.3 LNA/Mixer/PLL/Baseband/ADC Simulation

A transient simulation was performed on the entire receiver, including the LNA, the mixers, the PLL, the baseband amplifiers and filters, and the ΣΔ ADCs. A 2.01-GHz sinusoidal RF signal is applied to the receiver input and the 10-MHz output signals from

output noise due to source	429 μV
total output noise	827 μV
noise figure	5.7 dB

Table 7.4: Summary of simulated receiver noise performance including PLL for typical process corner (output noise integrated over 100 MHz).

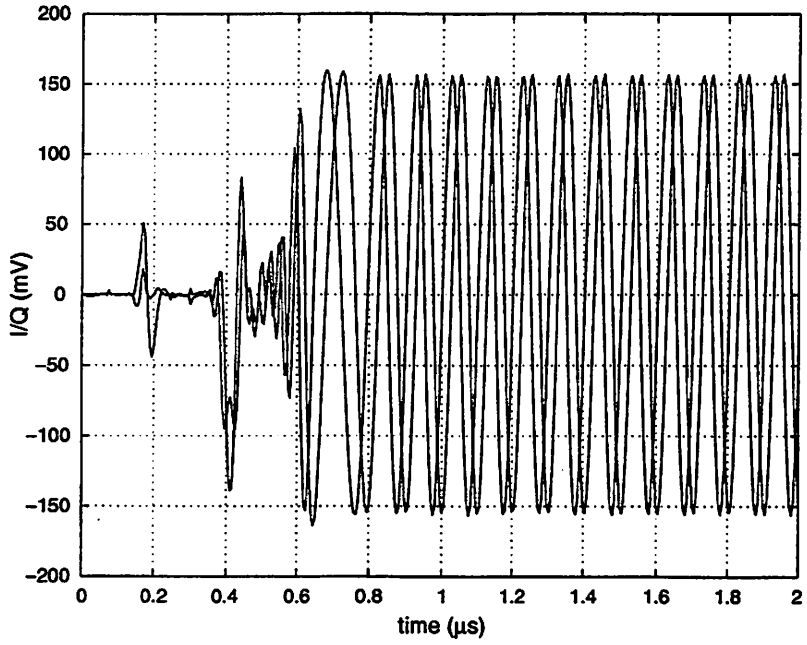


Figure 7.5: Simulated transient response.

each of the I and Q Sallen and Key filters is illustrated in Fig. 7.5. In addition, the digital output signals from each of the I and Q ΣΔ modulators were processed in MATLAB in order to verify the functionality of the entire receiver.

7.2 Measurement Results

The two receiver test chips, SCRRX and SCRBARF, are directly attached to the test boards using chip-on-board (COB) packaging technology. This technique offers reduced package parasitics since the chip pads are bonded directly to landing zones on the test board and the package is essentially eliminated. However, this benefit comes at the expense of decreased testing flexibility since the chips are not easily interchangeable.

Although initial testing verified the basic functionality of the various test structures on the SCRRX chip, two main factors prohibited extensive characterization of this chip. First, antenna protection diodes were not included on this chip, and the lack of these diodes resulted in inconsistent performance. These diodes are required at the gate nodes of transistors with long interconnects in order to prevent charge accumulation during the plasma processing steps [99]. Failure to do so results in unpredictable transistor threshold voltages, and unfortunately, the ramifications of not including these diodes were not fully appreciated at the time. Second, several design errors in the test board exacerbated the difficulty of fully characterizing the various test structures on the SCRRX test chip. All of these errors were corrected in the SCRBARF prototype chip and the corresponding test board, from which all of the results reported in this section were measured.

In order to ease testing, rather than iterating over different bond-wire lengths, a combination of the input bond wires and a pair of 1-nH chip inductors were used to complete the LNA input tuning. The receiver input provides an excellent match to 50 Ω with a measured S_{11} better than –30 dB (Fig. 7.6).

The measured frequency response at the output of each of the I and Q Sallen and Key filters is illustrated in Fig. 7.7. The high-pass and low-pass corner frequencies are about 100 kHz and 17 MHz, respectively. The receiver gain is 41 dB with less than 0.5-dB gain

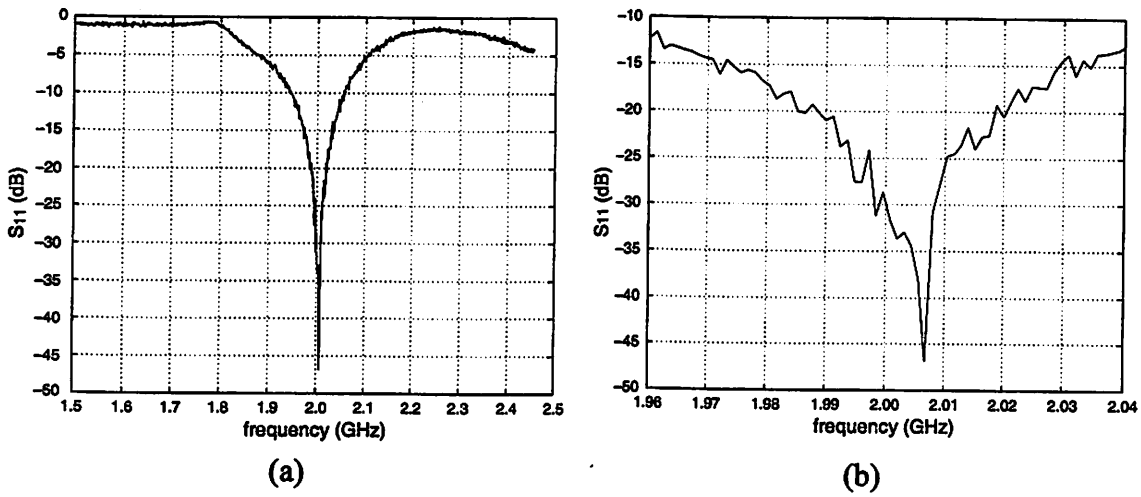


Figure 7.6: (a) Measured receiver S_{11}. (b) Zoomed view.

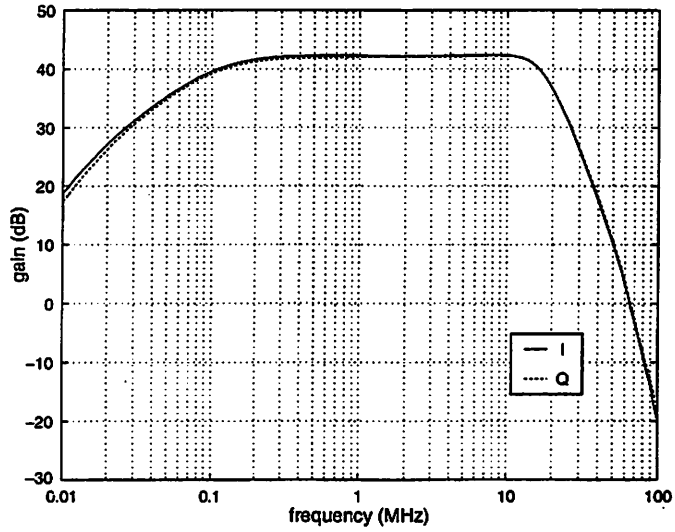


Figure 7.7: Measured frequency response at the output of each of the I and Q Sallen and Key filters.

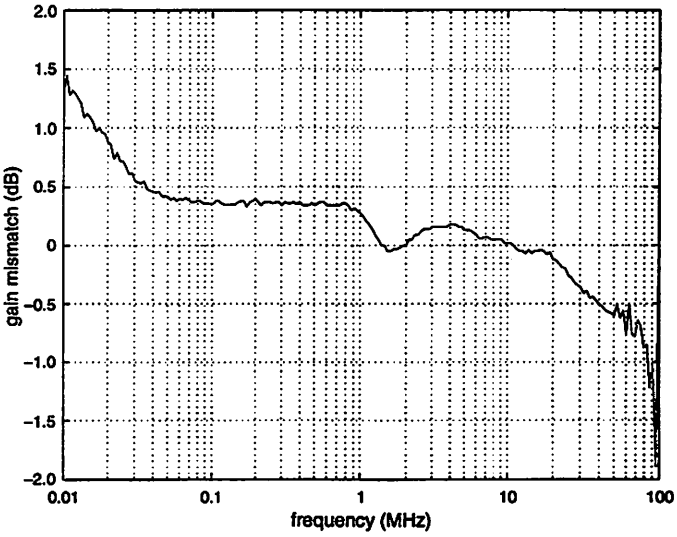


Figure 7.8: Measured I and Q gain mismatch.

mismatch between the I and Q paths over the –3-dB bandwidth (Fig. 7.8).

The noise performance of the receiver is measured at the output of the Sallen and Key filter. However, since the receiver output noise is below the noise floor of the spectrum analyzer, additional low-noise baseband amplifiers are used to amplify the receiver output noise. The noise performance of the receiver is then determined by first applying a noise source in the cold state at the receiver input and measuring the output noise on the spectrum analyzer [100]. In this case, the noise source produces a noise power of *kTB*.

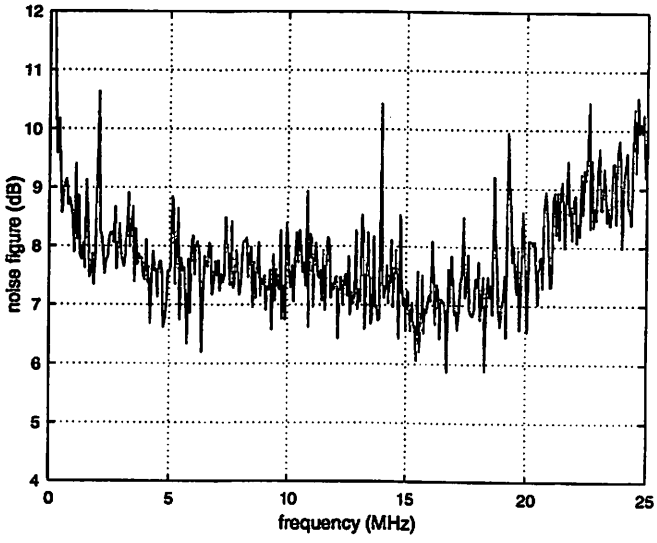


Figure 7.9: Measured receiver noise performance.

Next, the output noise is measured when a noise source in the hot state is applied at the receiver input. In this case, the noise source produces a noise power of $ENR \times kTB$, where ENR is the excess noise power ratio. The noise factor is given by

$$F = \frac{ENR - 1}{OPR - 1} \tag{7.1}$$

where OPR is the ratio between the measured output noise powers for the two cases. The noise figure of the receiver is plotted versus frequency in Fig. 7.9. The noise figure is less than 9 dB over the –3-dB bandwidth and includes approximately 1 dB of insertion loss from the external balun.

The measured distortion performance of the receiver is illustrated in Fig. 7.10. The input –1-dB compression point of the receiver is –31.1 dBm. The out-of-band IIP2 is measured by applying two sinusoids at the receiver input with offsets of 27 MHz and 37 MHz from the carrier frequency. In this case, the second-order intermodulation product appears at 10 MHz and the measured IIP2 is –6.7 dBm. The out-of-band IIP3 is measured by applying two sinusoids at the receiver input with offsets of 35 MHz and 60 MHz from the carrier frequency. In this case, the third-order intermodulation product also appears at 10 MHz and the measured IIP3 is –18.3 dBm.

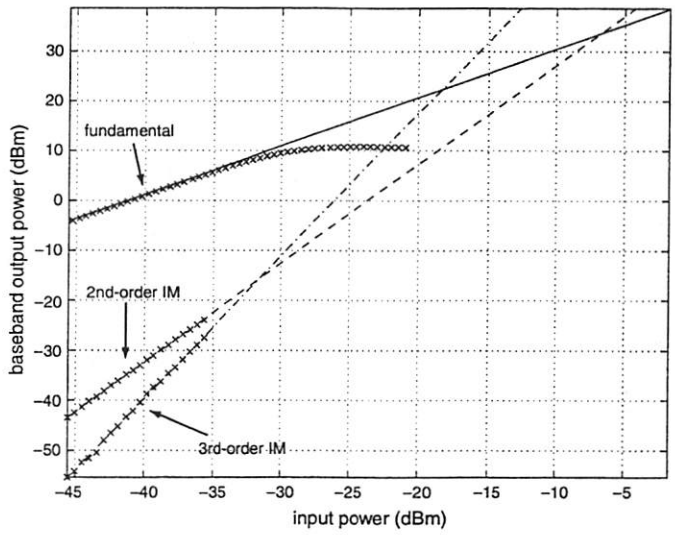


Figure 7.10: Measured receiver distortion performance.

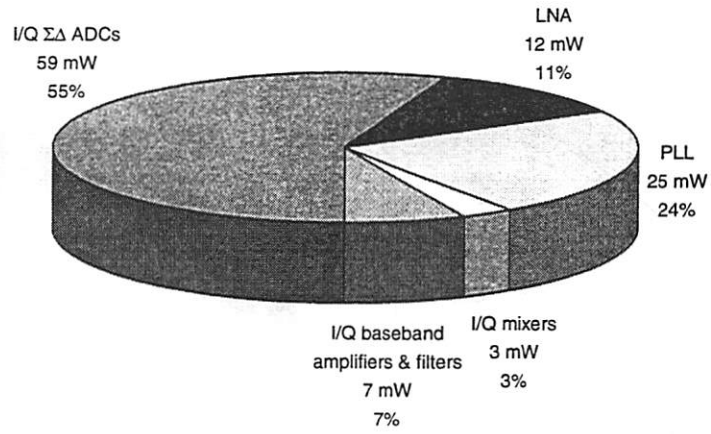


Figure 7.11: Measured receiver power consumption.

The measured LO-to-RF leakage is –81 dBm and the receiver's total power consumption is 106 mW. A breakdown of the power consumption is illustrated in Fig. 7.11.

The frequency synthesizer and the ΣΔ ADC were also characterized separately [31], [34]. The phase noise performance of the frequency synthesizer is illustrated in Fig. 7.12. The phase noise is –85 dBc/Hz at a 2.5-MHz offset. The ΣΔ ADC has a dynamic range of 42 dB and a peak SNDR of 40 dB when operating at a frequency of 200 MHz (Fig. 7.13a). The output spectrum when a –33-dBFS, 3.125-MHz sinusoid is applied to the input of the ΣΔ modulator is illustrated in Fig. 7.13b. The noise shaping of the ΣΔ modulator is evident in the measured output spectrum.

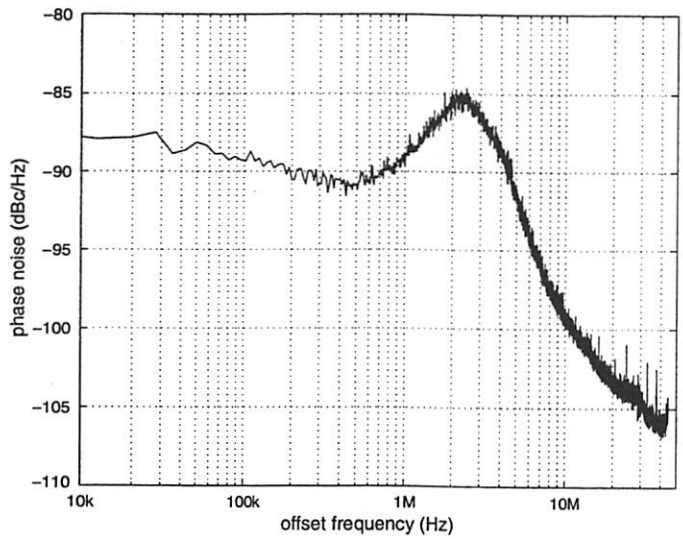
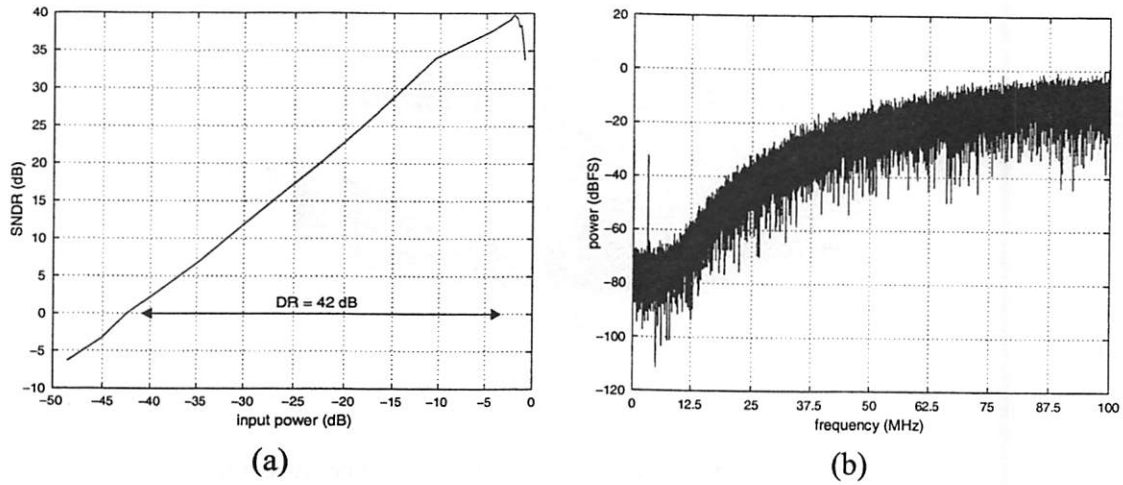


Figure 7.12: Measured phase noise performance of the frequency synthesizer.



(a) (b)

Figure 7.13: ΣΔ modulator operating at 200 MHz. (a) Measured dynamic range. (b) Measured output spectrum.

All of the receiver performance measurements are summarized in Table 7.5 along with the simulated results. Except for the phase noise of the PLL, all of the other measured receiver specifications either match or exceed the specifications used for the initial system-level simulation listed in Table 5.2. The system downlink was resimulated in Simulink using the measured performance specifications, and the SNR of the I and Q data from the output of the multiuser detector is still approximately 15 dB.

	simulated performance	measured performance
carrier frequency	2 GHz	2 GHz
noise figure (DSB)	4.81 – 6.26 dB	< 9 dB
S_{11}	< –20 dB	< –30 dB
voltage gain	42.2 – 45 dB	41 dB
I/Q gain mismatch	not simulated	< 2%
HPF corner frequency	67.0 – 104.4 kHz	100 kHz
LPF corner frequency	17.1 – 22.5 MHz	17 MHz
PLL phase noise	–92 – –85 dBc/Hz @ 2.5 MHz	–85 dBc/Hz @ 2.5 MHz
I/Q phase mismatch	not simulated	< 2.5°
IIP2	not simulated	–6.7 dBm
IIP3	not simulated	–18.3 dBm
–1-dB compression	–39 – –33 dBm	–31.1 dBm
ΣΔ dynamic range	46.4 dB @ 200 MHz	42 dB @ 200 MHz
ΣΔ SNDR	45 dB @ 200 MHz	40 dB @ 200 MHz
LO-to-RF leakage	not simulated	–81 dBm
power dissipation	not simulated	106 mW

Table 7.5: Summary of receiver performance measurements.

Finally, the receiver was tested using a modulated RF input signal. The test setup is illustrated in Fig. 7.14. The Simulink simulation framework is used to help verify the functionality of the receiver test chip. The same Simulink blocks that were used to evaluate the overall performance of the system are also used to generate the digital I and Q input signals for the base-station transmitter as well as to process the digital I and Q output signals from the receiver. The TX DSP block in Simulink performs QPSK modulation, signal spreading, and pulse shaping, and the 100-MHz output streams from this block, which consist of the combined data channels for ten users, are converted to baseband analog I and Q signals by the Arbitrary Waveform Generator. Next, the Vector Signal Generator translates the baseband I and Q signals to the 2-GHz carrier frequency, and this modulated RF signal is then applied to the receiver input. The digital I and Q output signals from the receiver are captured by the Logic Analysis System, and then the RX DSP block in Simulink performs timing and data recovery on the acquired data.

The constellation diagrams for the I and Q signals before and after data recovery using the adaptive MUD algorithm are illustrated in Fig. 7.15. These measured constellation

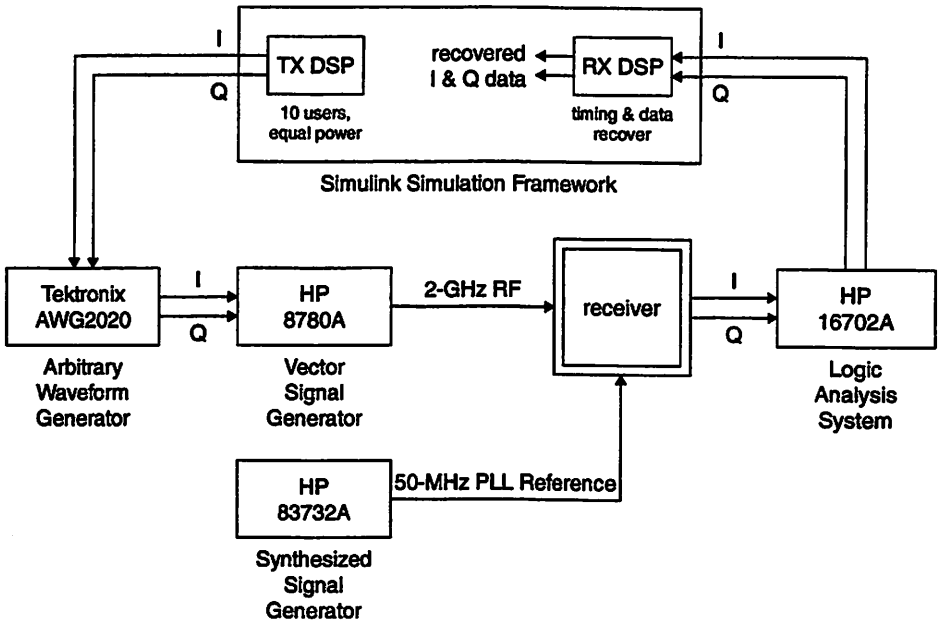


Figure 7.14: Test setup for receiver measurement with a modulated RF input signal.

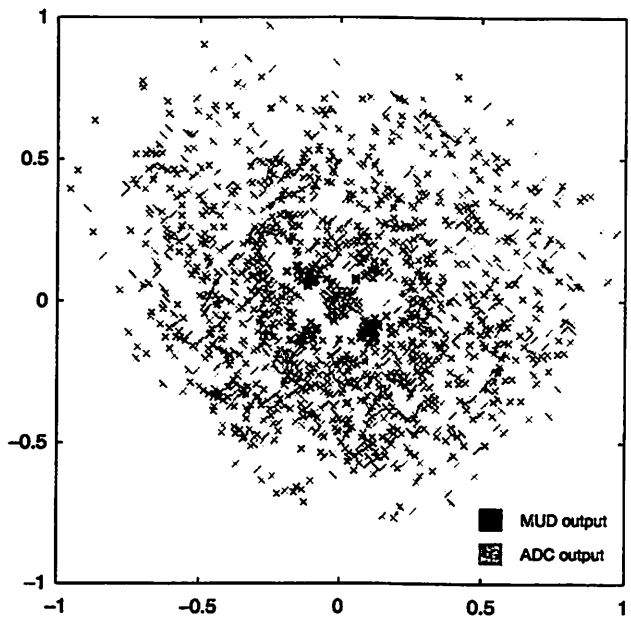


Figure 7.15: Measured constellation diagrams using a modulated RF input signal.

diagrams demonstrate that the system achieves good overall performance when a modulated RF signal is applied to the input of the receiver test chip.

7.3 Measurement Issues

7.3.1 Yield and Reliability

In the receiver prototype described in this thesis, electrostatic discharge (ESD) protection structures were not included in order to avoid additional parasitic capacitances which would be detrimental to the operation of the high-speed RF circuits. Unfortunately, the reliability of the prototype chips is significantly reduced without ESD protection structures [101], [102].

The yield of the receiver prototype chips was also affected by the lack of sufficient dummy structures. Dummy structures are required on the active, polysilicon, and metal layers for processes which rely on chemical mechanical polishing (CMP) for planarization. These dummy structures were not included around the receiver circuits since the effect of these structures on noise coupling was not well understood. Unfortunately, the yield of the prototype chips is significantly reduced without the dummy structures.

7.3.2 Packaging Technology

The receiver prototype chips were attached to the test boards using COB packaging technology. Although this technique offers reduced package parasitics, this benefit comes at the expense of decreased testing flexibility since the chips are not easily interchangeable. The inability to easily interchange prototype chips from the test boards was particularly problematic due to the low yield resulting from the lack of sufficient dummy structures. Alternatively, the ball grid array (BGA) package provides low parasitics without the disadvantage of decreased testing flexibility [103]. Moreover, this packaging technology can be used to include passive structures with high quality factors as an alternative to on-chip structures with much lower quality factors.

Chapter 8

Conclusion

8.1 Research Summary

The success of future wireless systems will depend heavily on their ability to provide high capacity while maintaining low cost, small form factor, and low power consumption in the portable devices. By tightly incorporating implementation issues throughout the process of defining the system specifications, an efficient solution can be achieved without necessarily sacrificing overall performance.

This thesis described a design methodology which facilitates the evaluation of tradeoffs between implementation issues and overall system performance, focusing primarily on the receiver as an example. First, system-level specifications, such as modulation scheme and signal bandwidth, strongly influence the choice of receiver architecture, which in turn, has ramifications on the achievable power consumption and integration level. When system-level specifications are determined without considering their impact on receiver architecture selection, single-chip solutions may be very difficult to achieve or just simply infeasible. Some selection guidelines were presented in Chapter 2 for the heterodyne, direct-conversion, image-reject, and low-IF receiver architectures.

Second, the rapid improvements in digital CMOS technology provide an opportunity to use advanced digital signal processing algorithms which in the past were considered too complex to implement in the mobile device. These algorithms promise significant increases in system performance but their performance may ultimately be limited by analog circuit impairments, such as noise and distortion. This thesis described the detrimental effects of a number of these impairments and presented a system-level simulation framework which facilitates the direct evaluation of these effects on the performance of digital communications algorithms. The simulation framework is implemented in Simulink, which offers compatibility with MATLAB, a simulation tool already widely used for the development and evaluation of communications algorithms. This simulation framework relies on baseband-equivalent models for all of the RF building blocks in order to avoid simulation at the carrier frequency, resulting in faster simulation times.

These strategies were then applied to the design of a high-speed wireless downlink for an indoor picocellular system. The system provides an aggregate data rate of 50 Mb/s with a transmission bandwidth of 32.5 MHz and a carrier frequency of 2 GHz. The wide bandwidth of the desired signal facilitates the use of a direct-conversion architecture. In this case, on-chip high-pass filtering can be used to remove dc offsets, and system-level simulations confirmed that the SNR degradation is less than 0.5 dB for a high-pass corner frequency of up to 500 kHz. Also, complete end-to-end simulations of the system downlink were performed in Simulink and revealed that the digital multiuser detection algorithm used for data recovery is relatively insensitive to analog hardware impairments.

Finally, a receiver prototype was implemented to meet the specifications determined from the system-level simulations. A power-efficient solution was achieved by taking advantage of the relaxed specifications as well as by using low-power circuit implementation techniques. This receiver prototype includes the low-noise amplifier, frequency synthesizer, mixers, baseband amplifiers and filters, and analog-to-digital converters, all implemented on a single chip with a power dissipation of about 100 mW.

8.2 Future Work

8.2.1 Bottom-Up Verification

The research presented in this thesis demonstrates that by tightly incorporating implementation issues throughout the process of system definition, a very efficient solution can be achieved without necessarily sacrificing overall performance. A top-down design approach based on Simulink was used to determine the tolerable levels of analog circuit impairments. However, once the receiver circuits are designed, it is necessary to verify that the system actually achieves the desired performance using these circuits.

For the receiver prototype described in this thesis, verification was performed by resimulating the entire system downlink in Simulink using both simulated results from SpectreRF and measured results from the receiver prototype. Verification was also performed using the transient envelop simulation capability of SpectreRF. Unfortunately, both of these verification approaches have their shortcomings. Although the former approach is quite straightforward, as with any approach based on behavioral models, the accuracy of the simulation results depends on the accuracy of the models. Unfortunately, the Simulink simulation framework described in this thesis relies on behavioral models which may not accurately represent the behavior of the actual circuits over all input power levels [104]. Finally, although transient envelop simulations in SpectreRF provide very accurate results, these simulations are also very slow since they rely on transistor-level models rather than on behavioral models.

8.2.2 Improved Behavioral Models

In the Simulink simulation framework described in this thesis, the behavioral models rely on a third-order power series in order to model circuit distortion (Appendix A). For some communications systems it may be necessary to include the effects of higher-order nonlinearities. Implementing the behavioral models so that the order of the power series is an input parameter rather than being fixed allows for increased flexibility.

Moreover, an approach based on power series provides an accurate description of distortion in circuits that are memoryless but may not be adequate for modeling distortion in circuits at high frequencies, where the effects of parasitic capacitances become

significant. Although an approach based on Volterra series provides much more accurate results in this case, calculations based on this approach are rather complex, even when using computer simulation techniques. Further investigation is needed in order to determine accurate yet efficient methods of modeling distortion in high-frequency circuits.

8.2.3 Single-Chip Integration

The receiver prototype presented in this thesis achieves a very high level of integration with a low number of off-chip components. All RF and analog baseband components are integrated onto a single chip, while the external components include an antenna, RF filter, crystal reference, a pair of chip inductors, and a pair of baluns. The off-chip inductors were used for convenience in order to avoid iterating over different bond wire lengths for the LNA input matching, and consequently, are not a serious impediment to eventual integration.

For this receiver prototype, a fully-differential on-chip signal path was used in order to mitigate the coupling between different receiver components. The use of a differential signal path necessitates the use of off-chip baluns to convert the single-ended off-chip signals to differential signals. One possible way of eliminating these baluns is to also use a fully-differential off-chip signal path, which requires further work in the areas of differential antenna design as well as differential RF filter design.

Furthermore, higher levels of integration may be achieved by taking advantage of packaging technologies such as the BGA package already mentioned in Section 7.3.2. This packaging technology can be used to include passive structures with high quality factors so that components such as the RF filter can be integrated seamlessly into the package.

Finally, in addition to using a fully-differential on-chip signal path, other techniques were used to mitigate the coupling between the different receiver components, including using separate supplies for the analog and digital circuits as well as implementing the digital sections of the PLL using logic styles, such as SCL and DCVSL, which result is less

substrate current injection. Despite these efforts, the consequences of integrating all of the digital baseband circuits, such as those required to implement the data and timing recovery algorithms, is unclear. In particular, further investigation is needed in order to determine the feasibility of integrating all of the analog and digital signal processing circuits onto a single chip.

Appendix A

Baseband-Equivalent Models

A.1 RF Amplifiers

The transfer function of any RF gain block can be represented by the following relationship:

$$x_o(t) = \sum_{n=0}^{N} a_n x_i^n(t) \tag{A.1}$$

where $x_i(t)$ and $x_o(t)$ are the input and output signals, respectively. Let $x_i(t)$ and $x_o(t)$ have the same form as (4.29):

$$x_i(t) = x_{iDC}(t) + \sum_{n=1}^{N} [x_{iIn}(t)\cos(n\omega_c t) + x_{iQn}(t)\sin(n\omega_c t)] \tag{A.2}$$

$$x_o(t) = x_{oDC}(t) + \sum_{n=1}^{N} [x_{oIn}(t)\cos(n\omega_c t) + x_{oQn}(t)\sin(n\omega_c t)]\,. \tag{A.3}$$

The output coefficients $x_{oDC}(t)$, $x_{oIn}(t)$, and $x_{oQn}(t)$ in terms of the input coefficients $x_{iDC}(t)$, $x_{iIn}(t)$, and $x_{iQn}(t)$ for $N = 3$ are given in Tables A.1 and A.2.

$x_{oDC}(t)$	$a_0 + a_1 x_{iDC} + a_2 x_{iDC}^2 + a_3 x_{iDC}^3 + \frac{1}{2} a_2 x_{iI1}^2 + \frac{3}{2} a_3 x_{iDC} x_{iI1}^2 + \frac{3}{4} a_3 x_{iI1}^2 x_{iI2} + \frac{1}{2} a_2 x_{iI2}^2 + \frac{3}{2} a_3 x_{iDC} x_{iI2}^2 + \frac{3}{2} a_3 x_{iI1} x_{iI2} x_{iI3} + \frac{1}{2} a_2 x_{iI3}^2 + \frac{3}{2} a_3 x_{iDC} x_{iI3}^2 + \frac{1}{2} a_2 x_{iQ1}^2 + \frac{3}{2} a_3 x_{iDC} x_{iQ1}^2 - \frac{3}{4} a_3 x_{iI2} x_{iQ1}^2 + \frac{3}{2} a_3 x_{iI1} x_{iQ1} x_{iQ2} - \frac{3}{2} a_3 x_{iI3} x_{iQ1} x_{iQ2} + \frac{1}{2} a_2 x_{iQ2}^2 + \frac{3}{2} a_3 x_{iDC} x_{iQ2}^2 + \frac{3}{2} a_3 x_{iI2} x_{iQ1} x_{iQ3} + \frac{3}{2} a_3 x_{iI1} x_{iQ2} x_{iQ3} + \frac{1}{2} a_2 x_{iQ3}^2 + \frac{3}{2} a_3 x_{iDC} x_{iQ3}^2$
$x_{oI1}(t)$	$a_1 x_{iI1} + 2 a_2 x_{iDC} x_{iI1} + 3 a_3 x_{iDC}^2 x_{iI1} + \frac{3}{4} a_3 x_{iI1}^3 + a_2 x_{iI1} x_{iI2} + 3 a_3 x_{iDC} x_{iI1} x_{iI2} + \frac{3}{2} a_3 x_{iI1} x_{iI2}^2 + \frac{3}{4} a_3 x_{iI1}^2 x_{iI3} + a_2 x_{iI2} x_{iI3} + 3 a_3 x_{iDC} x_{iI2} x_{iI3} + \frac{3}{4} a_3 x_{iI2}^2 x_{iI3} + \frac{3}{2} a_3 x_{iI1} x_{iI3}^2 + \frac{3}{4} a_3 x_{iI1} x_{iQ1}^2 - \frac{3}{4} a_3 x_{iI3} x_{iQ1}^2 + a_2 x_{iQ1} x_{iQ2} + 3 a_3 x_{iDC} x_{iQ1} x_{iQ2} + \frac{3}{2} a_3 x_{iI1} x_{iQ2}^2 - \frac{3}{4} a_3 x_{iI3} x_{iQ2}^2 + \frac{3}{2} a_3 x_{iI1} x_{iQ1} x_{iQ3} + a_2 x_{iQ2} x_{iQ3} + 3 a_3 x_{iDC} x_{iQ2} x_{iQ3} + \frac{3}{2} a_3 x_{iI2} x_{iQ2} x_{iQ3} + \frac{3}{2} a_3 x_{iI1} x_{iQ3}^2$
$x_{oI2}(t)$	$\frac{1}{2} a_2 x_{iI1}^2 + \frac{3}{2} a_3 x_{iDC} x_{iI1}^2 + a_1 x_{iI2} + 2 a_2 x_{iDC} x_{iI2} + 3 a_3 x_{iDC}^2 x_{iI2} + \frac{3}{2} a_3 x_{iI1}^2 x_{iI2} + \frac{3}{4} a_3 x_{iI2}^3 + a_2 x_{iI1} x_{iI3} + 3 a_3 x_{iDC} x_{iI1} x_{iI3} + \frac{3}{2} a_3 x_{iI1} x_{iI2} x_{iI3} + \frac{3}{2} a_3 x_{iI2} x_{iI3}^2 - \frac{1}{2} a_2 x_{iQ1}^2 - \frac{3}{2} a_3 x_{iDC} x_{iQ1}^2 + \frac{3}{2} a_3 x_{iI2} x_{iQ1}^2 + \frac{3}{2} a_3 x_{iI3} x_{iQ1} x_{iQ2} + \frac{3}{4} a_3 x_{iI2} x_{iQ2}^2 + a_2 x_{iQ1} x_{iQ3} + 3 a_3 x_{iDC} x_{iQ1} x_{iQ3} - \frac{3}{2} a_3 x_{iI2} x_{iQ1} x_{iQ3} + \frac{3}{2} a_3 x_{iI1} x_{iQ2} x_{iQ3} + \frac{3}{2} a_3 x_{iI2} x_{iQ3}^2$
$x_{oI3}(t)$	$\frac{1}{4} a_3 x_{iI1}^3 + a_2 x_{iI1} x_{iI2} + 3 a_3 x_{iDC} x_{iI1} x_{iI2} + \frac{3}{4} a_3 x_{iI1} x_{iI2}^2 + a_1 x_{iI3} + 2 a_2 x_{iDC} x_{iI3} + 3 a_3 x_{iDC}^2 x_{iI3} + \frac{3}{2} a_3 x_{iI1}^2 x_{iI3} + \frac{3}{2} a_3 x_{iI2}^2 x_{iI3} + \frac{3}{4} a_3 x_{iI3}^3 - \frac{3}{4} a_3 x_{iI1} x_{iQ1}^2 + \frac{3}{2} a_3 x_{iI3} x_{iQ1}^2 - a_2 x_{iQ1} x_{iQ2} - 3 a_3 x_{iDC} x_{iQ1} x_{iQ2} + \frac{3}{2} a_3 x_{iI2} x_{iQ1} x_{iQ2} - \frac{3}{4} a_3 x_{iI1} x_{iQ2}^2 + \frac{3}{2} a_3 x_{iI3} x_{iQ2}^2 + \frac{3}{4} a_3 x_{iI3} x_{iQ3}^2$

Table A.1: Baseband-equivalent model for RF amplifiers (I).

$x_{oQ1}(t)$	$a_1 x_{iQ1} + 2a_2 x_{iDC} x_{iQ1} + 3a_3 x_{iDC}^2 x_{iQ1} + \frac{3}{4} a_3 x_{iI1}^2 x_{iQ1} - a_2 x_{iI2} x_{iQ1} - 3a_3 x_{iDC} x_{iI2} x_{iQ1} + \frac{3}{2} a_3 x_{iI2}^2 x_{iQ1} - \frac{3}{2} a_3 x_{iI1} x_{iI3} x_{iQ1} + \frac{3}{2} a_3 x_{iI3}^2 x_{iQ1} + \frac{3}{4} a_3 x_{iQ1}^3 + a_2 x_{iI1} x_{iQ2} + 3a_3 x_{iDC} x_{iI1} x_{iQ2} - a_2 x_{iI3} x_{iQ2} - 3a_3 x_{iDC} x_{iI3} x_{iQ2} + \frac{3}{2} a_3 x_{iI2} x_{iI3} x_{iQ2} + \frac{3}{2} a_3 x_{iQ1} x_{iQ2}^2 + \frac{3}{4} a_3 x_{iI1}^2 x_{iQ3} + a_2 x_{iI2} x_{iQ3} + 3a_3 x_{iDC} x_{iI2} x_{iQ3} - \frac{3}{4} a_3 x_{iI2}^2 x_{iQ3} - \frac{3}{4} a_3 x_{iQ1}^2 x_{iQ3} + \frac{3}{4} a_3 x_{iQ2}^2 x_{iQ3} + \frac{3}{2} a_3 x_{iQ1} x_{iQ3}^2$
$x_{oQ2}(t)$	$a_2 x_{iI1} x_{iQ1} + 3a_3 x_{iDC} x_{iI1} x_{iQ1} - a_2 x_{iI3} x_{iQ1} - 3a_3 x_{iDC} x_{iI3} x_{iQ1} + \frac{3}{2} a_3 x_{iI2} x_{iI3} x_{iQ1} + a_1 x_{iQ2} + 2a_2 x_{iDC} x_{iQ2} + 3a_3 x_{iDC}^2 x_{iQ2} + \frac{3}{2} a_3 x_{iI1}^2 x_{iQ2} + \frac{3}{4} a_3 x_{iI2}^2 x_{iQ2} - \frac{3}{2} a_3 x_{iI1} x_{iI3} x_{iQ2} + \frac{3}{2} a_3 x_{iI3}^2 x_{iQ2} + \frac{3}{2} a_3 x_{iQ1}^2 x_{iQ2} + \frac{3}{4} a_3 x_{iQ2}^3 + a_2 x_{iI1} x_{iQ3} + 3a_3 x_{iDC} x_{iI1} x_{iQ3} + \frac{3}{2} a_3 x_{iI1} x_{iI2} x_{iQ3} + \frac{3}{2} a_3 x_{iQ1} x_{iQ2} x_{iQ3} + \frac{3}{2} a_3 x_{iQ2} x_{iQ3}^2$
$x_{oQ3}(t)$	$\frac{3}{4} a_3 x_{iI1}^2 x_{iQ1} + a_2 x_{iI2} x_{iQ1} + 3a_3 x_{iDC} x_{iI2} x_{iQ1} - \frac{3}{4} a_3 x_{iI2}^2 x_{iQ1} - \frac{1}{4} a_3 x_{iQ1}^3 + a_2 x_{iI1} x_{iQ2} + 3a_3 x_{iDC} x_{iI1} x_{iQ2} + \frac{3}{2} a_3 x_{iI1} x_{iI2} x_{iQ2} + \frac{3}{4} a_3 x_{iQ1}^2 x_{iQ2} + a_1 x_{iQ3} + 2a_2 x_{iDC} x_{iQ3} + 3a_3 x_{iDC}^2 x_{iQ3} + \frac{3}{2} a_3 x_{iI1}^2 x_{iQ3} + \frac{3}{2} a_3 x_{iI2}^2 x_{iQ3} + \frac{3}{4} a_3 x_{iI3}^2 x_{iQ3} + \frac{3}{2} a_3 x_{iQ1}^2 x_{iQ3} + \frac{3}{2} a_3 x_{iQ2}^2 x_{iQ3} + \frac{3}{4} a_3 x_{iQ3}^3$

Table A.2: Baseband-equivalent model for RF amplifiers (II).

A.2 Mixers

The transfer function of a mixer can be represented by

$$y_o(t) = y_i(t) \times y_{LO}(t) \tag{A.4}$$

where $y_i(t)$, $y_{LO}(t)$, and $y_o(t)$ are the input, oscillator, and output signals, respectively. For a mixers in direct-conversion and low-IF receivers, let $y_i(t)$, $y_{LO}(t)$, and $y_o(t)$ have the same form as (4.29):

$$y_i(t) = y_{iDC}(t) + \sum_{n=1}^{N} [y_{iIn}(t)\cos(n\omega_c t) + y_{iQn}(t)\sin(n\omega_c t)] \tag{A.5}$$

$$y_{LO}(t) = y_{LODC}(t) + \sum_{n=1}^{N} [y_{LOIn}(t)\cos(n\omega_c t) + y_{LOQn}(t)\sin(n\omega_c t)] \qquad \text{(A.6)}$$

$$y_o(t) = y_{oDC}(t) + \sum_{n=1}^{N} [y_{oIn}(t)\cos(n\omega_c t) + y_{oQn}(t)\sin(n\omega_c t)]. \qquad \text{(A.7)}$$

The output coefficients $y_{oDC}(t)$, $y_{oIn}(t)$, and $y_{oQn}(t)$ in terms of the input coefficients $y_{iDC}(t)$, $y_{iIn}(t)$, $y_{iQn}(t)$, $y_{LODC}(t)$, $y_{LOIn}(t)$, and $y_{LOQn}(t)$ for $N = 3$ are given in Table A.3.

$y_{oDC}(t)$	$y_{iDC}y_{LODC} + \frac{y_{iI1}y_{LOI1}}{2} + \frac{y_{iI2}y_{LOI2}}{2} + \frac{y_{iI3}y_{LOI3}}{2} + \frac{y_{iQ1}y_{LOQ1}}{2} + \frac{y_{iQ2}y_{LOQ2}}{2} + \frac{y_{iQ3}y_{LOQ3}}{2}$
$y_{oI1}(t)$	$y_{iI1}y_{LODC} + y_{iDC}y_{LOI1} + \frac{y_{iI2}y_{LOI1}}{2} + \frac{y_{iI1}y_{LOI2}}{2} + \frac{y_{iI3}y_{LOI2}}{2} + \frac{y_{iI2}y_{LOI3}}{2} + \frac{y_{iQ2}y_{LOQ1}}{2} + \frac{y_{iQ1}y_{LOQ2}}{2} + \frac{y_{iQ3}y_{LOQ2}}{2} + \frac{y_{iQ2}y_{LOQ3}}{2}$
$y_{oI2}(t)$	$y_{iI2}y_{LODC} + \frac{y_{iI1}y_{LOI1}}{2} + \frac{y_{iI3}y_{LOI1}}{2} + y_{iDC}y_{LOI2} + \frac{y_{iI1}y_{LOI3}}{2} - \frac{y_{iQ1}y_{LOQ1}}{2} + \frac{y_{iQ3}y_{LOQ1}}{2} + \frac{y_{iQ1}y_{LOQ3}}{2}$
$y_{oI3}(t)$	$y_{iI3}y_{LODC} + \frac{y_{iI2}y_{LOI1}}{2} + \frac{y_{iI1}y_{LOI2}}{2} + y_{iDC}y_{LOI3} - \frac{y_{iQ2}y_{LOQ1}}{2} - \frac{y_{iQ1}y_{LOQ2}}{2}$
$y_{oQ1}(t)$	$y_{iQ1}y_{LODC} + \frac{y_{iQ2}y_{LOI1}}{2} - \frac{y_{iQ1}y_{LOI2}}{2} + \frac{y_{iQ3}y_{LOI2}}{2} - \frac{y_{iQ2}y_{LOI3}}{2} + y_{iDC}y_{LOQ1} - \frac{y_{iI2}y_{LOQ1}}{2} + \frac{y_{iI1}y_{LOQ2}}{2} - \frac{y_{iI3}y_{LOQ2}}{2} + \frac{y_{iI2}y_{LOQ3}}{2}$
$y_{oQ2}(t)$	$y_{iQ2}y_{LODC} + \frac{y_{iQ1}y_{LOI1}}{2} + \frac{y_{iQ3}y_{LOI1}}{2} - \frac{y_{iQ1}y_{LOI3}}{2} + \frac{y_{iI1}y_{LOQ1}}{2} - \frac{y_{iI3}y_{LOQ1}}{2} + y_{iDC}y_{LOQ2} + \frac{y_{iI1}y_{LOQ3}}{2}$
$y_{oQ3}(t)$	$y_{iQ3}y_{LODC} + \frac{y_{iQ2}y_{LOI1}}{2} + \frac{y_{iQ1}y_{LOI2}}{2} + \frac{y_{iI2}y_{LOQ1}}{2} + \frac{y_{iI1}y_{LOQ2}}{2} + y_{iDC}y_{LOQ3}$

Table A.3: Baseband-equivalent model for mixers (direct-conversion and low-IF).

For mixers in heterodyne receivers, let $y_i(t)$ be given by (A.5) while $y_{LO}(t)$ and $y_o(t)$ are given by, respectively,

$$
\begin{aligned}
y_{LO}(t) &= y_{LODC}(t) + \sum_{n=1}^{N} \{y_{LOIn}(t)\cos[n(\omega_c - \omega_{IF})t] + y_{LOQn}(t)\sin[n(\omega_c - \omega_{IF})t]\} \\
&= y_{LODC}(t) + \sum_{n=1}^{N} \{[y_{LOIn}(t)\cos(n\omega_{IF}t) - y_{LOQn}(t)\sin(n\omega_{IF}t)]\cos(n\omega_c t) + \\
&\qquad [y_{LOIn}(t)\sin(n\omega_{IF}t) + y_{LOQn}(t)\cos(n\omega_{IF}t)]\sin(n\omega_c t)\}
\end{aligned} \tag{A.8}
$$

$$
\begin{aligned}
y_o(t) = y_{oDC}(t) &+ \sum_{n=1}^{N}[y_{oIIFn}(t)\cos(n\omega_{IF}t) + y_{oQIFn}(t)\sin(n\omega_{IF}t)] + \\
&\sum_{n=1}^{N}[y_{oIRFn}(t)\cos(n\omega_c t) + y_{oQRFn}(t)\sin(n\omega_c t)] + \\
&\sum_{n=1}^{N}\sum_{m=1}^{N}\{y_{oIPnm}(t)\cos[(n\omega_c + m\omega_{IF})t] + y_{oQPnm}(t)\sin[(n\omega_c + m\omega_{IF})t]\} + \\
&\sum_{n=1}^{N}\sum_{m=1}^{N}\{y_{oINnm}(t)\cos[(n\omega_c - m\omega_{IF})t] + y_{oQNnm}(t)\sin[(n\omega_c - m\omega_{IF})t]\}.
\end{aligned} \tag{A.9}
$$

The output coefficients of $y_o(t)$ in terms of the input coefficients of $y_i(t)$ and $y_{LO}(t)$ for $N = 3$ are given in Tables A.4 and A.5.

$y_{oDC}(t)$	$y_{iDC}y_{LODC}$		
$y_{oIIF1}(t)$	$\frac{y_{iI1}y_{LOI1}}{2} + \frac{y_{iQ1}y_{LOQ1}}{2}$	$y_{oQIF1}(t)$	$\frac{y_{iQ1}y_{LOI1}}{2} - \frac{y_{iI1}y_{LOQ1}}{2}$
$y_{oIIF2}(t)$	$\frac{y_{iI2}y_{LOI2}}{2} + \frac{y_{iQ2}y_{LOQ2}}{2}$	$y_{oQIF2}(t)$	$\frac{y_{iQ2}y_{LOI2}}{2} - \frac{y_{iI2}y_{LOQ2}}{2}$
$y_{oIIF3}(t)$	$\frac{y_{iI3}y_{LOI3}}{2} + \frac{y_{iQ3}y_{LOQ3}}{2}$	$y_{oQIF3}(t)$	$\frac{y_{iQ3}y_{LOI3}}{2} - \frac{y_{iI3}y_{LOQ3}}{2}$
$y_{oIRF1}(t)$	$y_{iI1}y_{LODC}$	$y_{oQRF1}(t)$	$y_{iQ1}y_{LODC}$
$y_{oIRF2}(t)$	$y_{iI2}y_{LODC}$	$y_{oQRF2}(t)$	$y_{iQ2}y_{LODC}$
$y_{oIRF3}(t)$	$y_{iI3}y_{LODC}$	$y_{oQRF3}(t)$	$y_{iQ3}y_{LODC}$
$y_{oIP11}(t)$	$\frac{y_{iI2}y_{LOI1}}{2} + \frac{y_{iQ2}y_{LOQ1}}{2}$	$y_{oQP11}(t)$	$\frac{y_{iQ2}y_{LOI1}}{2} - \frac{y_{iI2}y_{LOQ1}}{2}$
$y_{oIP12}(t)$	$\frac{y_{iI3}y_{LOI2}}{2} + \frac{y_{iQ3}y_{LOQ2}}{2}$	$y_{oQP12}(t)$	$\frac{y_{iQ3}y_{LOI2}}{2} - \frac{y_{iI3}y_{LOQ2}}{2}$
$y_{oIP13}(t)$	0	$y_{oQP13}(t)$	0

Table A.4: Baseband-equivalent model for IF mixers (I).

$y_{oIN11}(t)$	$y_{iDC}y_{LOI1}$	$y_{oQN11}(t)$	$y_{iDC}y_{LOQ1}$
$y_{oIN12}(t)$	$\frac{y_{iI1}y_{LOI2}}{2} + \frac{y_{iQ1}y_{LOQ2}}{2}$	$y_{oQN12}(t)$	$-\frac{1}{2}y_{iQ1}y_{LOI2} + \frac{y_{iI1}y_{LOQ2}}{2}$
$y_{oIN13}(t)$	$\frac{y_{iI2}y_{LOI3}}{2} + \frac{y_{iQ2}y_{LOQ3}}{2}$	$y_{oQN13}(t)$	$-\frac{1}{2}y_{iQ2}y_{LOI3} + \frac{y_{iI2}y_{LOQ3}}{2}$
$y_{oIP21}(t)$	$\frac{y_{iI3}y_{LOI1}}{2} + \frac{y_{iQ3}y_{LOQ1}}{2}$	$y_{oQP21}(t)$	$\frac{y_{iQ3}y_{LOI1}}{2} - \frac{y_{iI3}y_{LOQ1}}{2}$
$y_{oIP22}(t)$	0	$y_{oQP22}(t)$	0
$y_{oIP23}(t)$	0	$y_{oQP23}(t)$	0
$y_{oIN21}(t)$	$\frac{y_{iI1}y_{LOI1}}{2} - \frac{y_{iQ1}y_{LOQ1}}{2}$	$y_{oQN21}(t)$	$\frac{y_{iQ1}y_{LOI1}}{2} + \frac{y_{iI1}y_{LOQ1}}{2}$
$y_{oIN22}(t)$	$y_{iDC}y_{LOI2}$	$y_{oQN22}(t)$	$y_{iDC}y_{LOQ2}$
$y_{oIN23}(t)$	$\frac{y_{iI1}y_{LOI3}}{2} + \frac{y_{iQ1}y_{LOQ3}}{2}$	$y_{oQN23}(t)$	$-\frac{1}{2}y_{iQ1}y_{LOI3} + \frac{y_{iI1}y_{LOQ3}}{2}$
$y_{oIP31}(t)$	0	$y_{oQP31}(t)$	0
$y_{oIP32}(t)$	0	$y_{oQP32}(t)$	0
$y_{oIP33}(t)$	0	$y_{oQP33}(t)$	0
$y_{oIN31}(t)$	$\frac{y_{iI2}y_{LOI1}}{2} - \frac{y_{iQ2}y_{LOQ1}}{2}$	$y_{oQN31}(t)$	$\frac{y_{iQ2}y_{LOI1}}{2} + \frac{y_{iI2}y_{LOQ1}}{2}$
$y_{oIN32}(t)$	$\frac{y_{iI1}y_{LOI2}}{2} - \frac{y_{iQ1}y_{LOQ2}}{2}$	$y_{oQN32}(t)$	$\frac{y_{iQ1}y_{LOI2}}{2} + \frac{y_{iI1}y_{LOQ2}}{2}$
$y_{oIN33}(t)$	$y_{iDC}y_{LOI3}$	$y_{oQN33}(t)$	$y_{iDC}y_{LOQ3}$

Table A.5: Baseband-equivalent model for IF mixers (II).

Appendix B

DC-Offset Cancellation

B.1 Introduction

For systems with narrowband signals, dc offsets in a direct-conversion receiver can also be removed by using capacitive coupling or high-pass filtering. However, this technique requires very large capacitance and resistance values in order to remove as little low-frequency signal energy as possible. Consequently, an implementation using on-chip passive devices is not feasible unless very high-density structures are available in the process as in [13]. As an alternative, off-chip passive structures can be used to eliminate dc offsets. However, this approach is inconsistent with the goal of a highly-integrated implementation. The following sections provide an overview of alternative techniques used to eliminate dc offsets when a direct-conversion receiver architecture is used in narrowband systems.

B.2 Alternative DC-Offset Cancellation Techniques

In TDMA systems, data is received only during certain time intervals as illustrated in Fig. B.1. In this example, the TDMA system supports four data channels and each data

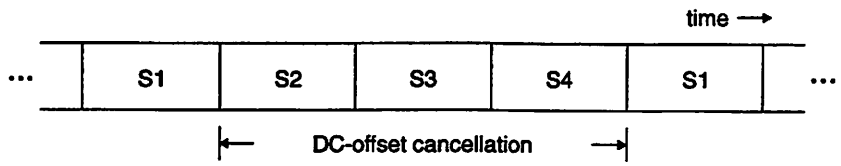


Figure B.1: TDMA time slots.

channel communicates only during one of the four time slots. Consequently, the receiver for the first data channel processes data only during time slots S1 and remains idle during time slots S2, S3, and S4. In this case, dc-offset cancellation can be performed during these idle time intervals [39], [105] in a manner very similar to the autozeroing technique described in Section 5.3.5. During the idle periods, the dc offset is stored, and then during the active time slot, the offset is subtracted from the received signal. Fig. B.2 illustrates two different approaches based on this technique. In one approach, the offset is stored on a capacitor [39], while in the second approach, the offset is stored digitally and then subtracted from the received signal using a DAC [11]. The drawback of this technique is that it relies on the offset not changing between the idle and active time slots.

In fact, dc offsets do vary over time as described in Section 2.3.1. In this case the dc offsets must be tracked even during the active time slots. A method which tracks the dc

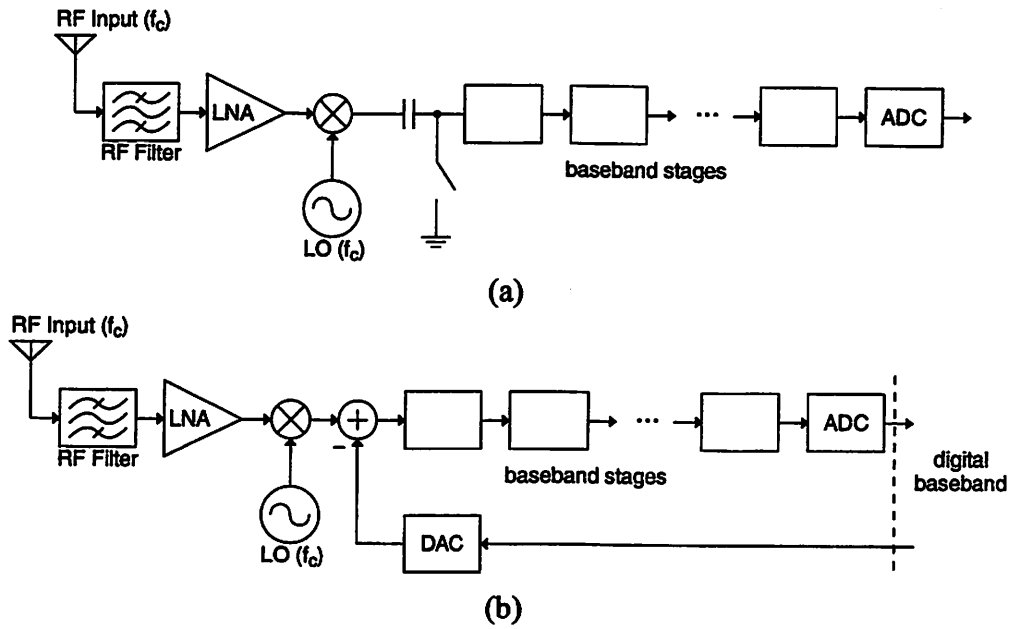


Figure B.2: DC-offset cancellation. (a) Capacitive storage. (b) Feedback DAC.

offsets while data is being received can be used for any system, not just TDMA systems. For systems which use signal constellations centered and symmetric about the origin, such as QPSK, the dc content of the received signal is ideally zero. For such systems, an averaging circuit, such as a low-pass filter, can be used to estimate the dc content of the received signal. The estimated dc offset can then be subtracted from the received signal.

This dc-offset cancellation scheme can be implemented using an adaptive LMS algorithm [56]. Consider the equivalent combiner for a feedback dc-offset correction loop illustrated in Fig. B.3, where $x(k)$ is the received signal, $g(k)$ is update signal, $y(k)$ is the estimated dc-offset, $d(k)$ is the desired value of the dc-offset, and $e(k)$ is the error signal. The error signal $e(k)$ is

$$e(k) = LPF\{x(k) - g(k)\} - d(k) = LPF\{x(k)\} - LPF\{g(k)\} - d(k). \qquad \text{(B.1)}$$

Adaptation using the stochastic gradient descent method [56] results in the following update equation:

$$g(k+1) = g(k) - \mu \frac{\partial}{\partial g(k)} \left\{ \frac{1}{2} e^2(k) \right\} = g(k) - \mu e(k) \frac{\partial}{\partial g(k)} \{e(k)\} \qquad \text{(B.2)}$$

where μ is the step size. Taking the partial derivative of $e(k)$ with respect to $g(k)$, (B.2) becomes

$$g(k+1) = g(k) + \mu e(k). \qquad \text{(B.3)}$$

The stability criterion for this algorithm is determined by first setting

$$d(k) = LPF\{x(k)\} - \phi. \qquad \text{(B.4)}$$

Then the prediction error is

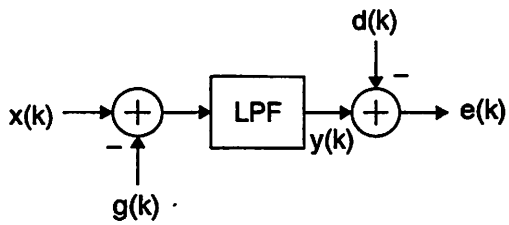


Figure B.3: Equivalent combiner for a feedback dc-offset correction loop.

$$e(k) = LPF\{x(k)\} - g(k) - d(k) = \phi - g(k). \tag{B.5}$$

The parameter error vector update is

$$\bar{g}(k+1) = \phi - g(k+1) = \phi - g(k) - \mu[\phi - g(k)] = \bar{g}(k)(1-\mu) \tag{B.6}$$

and the summed squared parameter error increment is

$$\bar{g}^2(k+1) - \bar{g}^2(k) = -\mu\bar{g}^2(k)(2-\mu). \tag{B.7}$$

If the algorithm converges, then the parameter error vector update at time $k+1$ must be less than the parameter error vector update at time k. Consequently, the summed squared parameter error increment must be negative. For a positive step size μ, the following relationship must be satisfied if the algorithm converges:

$$0 < \mu < 2. \tag{B.8}$$

Fig. B.4 illustrates one possible implementation of a dc-offset correction loop based on this adaptive LMS algorithm. A digital low-pass filter is used to estimate the dc content of the received signal and is followed by a digital implementation of the stochastic gradient descent LMS algorithm. The dc offset is then subtracted at the output of the mixer using a DAC. This approach is similar to the one illustrated in Fig. B.2b but can be used for any system which uses a signal constellation centered and symmetric about the origin and not just TDMA systems.

The effectiveness of this dc-offset cancellation loop is verified using Simulink. The simulation assumes a worst-case dc offset of 100 mV after the mixer due to LO self-

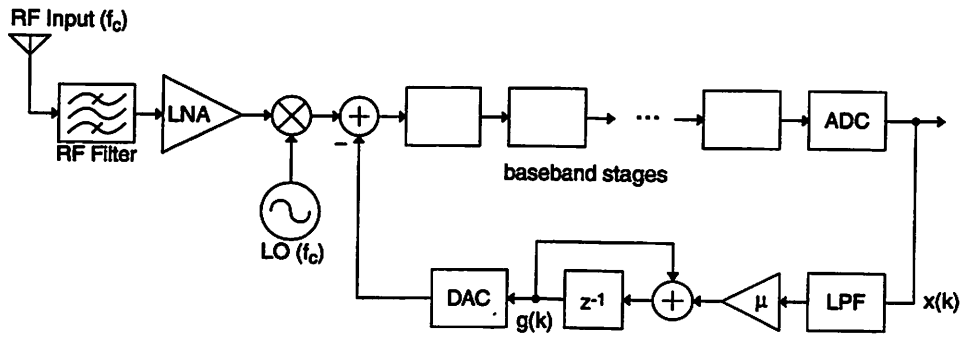


Figure B.4: Implementation of adaptive dc-offset correction loop.

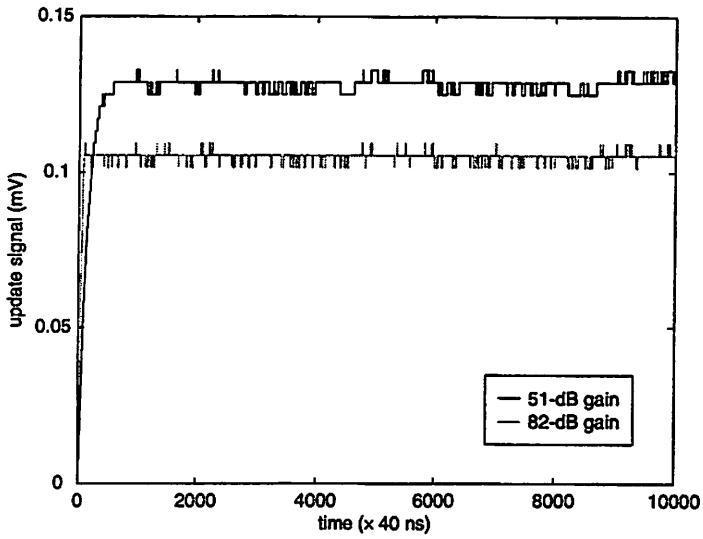


Figure B.5: Simulation of dc-offset cancellation loop.

mixing and a worst-case dc offset of 100 mV due to systematic offsets in the baseband circuits. The digital low-pass filter has a relative bandwidth of 0.1 and an 8-bit DAC is used to subtract the dc offset at the output of the mixer. The update signal from the dc-offset correction circuit is illustrated in Fig. B.5 for two different cases:

1. weak input signal and maximum gain (82 dB); and
2. strong input signal and minimum gain (51 dB).

The ripple in the update signal is due to the finite precision of the DAC. The step size used for the stochastic gradient descent LMS algorithm is 1/512, which was chosen as a compromise between convergence time and the amount of signal ripple after convergence.

Finally, in addition to the digital low-pass filter, the LMS algorithm described above also incorporates an implicit low-pass filtering operation. If the digital low-pass filter is omitted from the dc-offset correction loop, the update equation becomes

$$g(k) = g(k-1) + \mu[x(k-1) - g(k-1)]. \tag{B.9}$$

The z-transform of (B.9) is

$$G(z) = \frac{\mu z^{-1}}{1-(1-\mu)z^{-1}} X(z). \tag{B.10}$$

The pole is found by setting the denominator of (B.10) equal to zero:

$$z = 1 - \mu . \tag{B.11}$$

In order for this filter to be stable, the pole must be located inside the unit circle, or $0 < \mu < 2$, which is the same result derived in (B.8). Since the value of μ determines the cutoff frequency of the low-pass filter, a small value of μ is desirable. However, a small value of μ also results in a very long convergence time. Consequently, including the additional digital low-pass filter in the dc-offset correction loop allows the step size to be set independently from the low-pass filter corner frequency. However, for applications where a longer convergence time is tolerable, the digital low-pass filter may actually be omitted.

Appendix C

Why 50 Ω?

C.1 Introduction

Transmission line theory calls for conjugate impedance matching for maximize power transfer from the source to the load. In order to facilitate the independent design of different components, most microwave designs are based on a standard interface impedance of 50 Ω. Integrated-circuit implementations have already abandoned this antiquated requirement. For these implementations, the connections between on-chip components are typically much less than the signal wavelength so transmission line effects can be neglected. For example, for a 2-GHz signal, the wavelength is

$$\lambda = \frac{c}{f} = \frac{3\times10^8 \text{ m/s}}{2\times10^9 \text{ Hz}} = 0.15 \text{ m} \tag{C.1}$$

while the lengths of on-chip connections are typically no more than 1 mm. Consequently, integrated-circuit implementations do not need to adhere to the 50-Ω requirement.

Recently, the 50-Ω requirement at the interface between external and on-chip components has also come under intense scrutiny. This appendix addresses two questions which are at the heart of the controversy over the 50-Ω requirement. First, if maximum

power transfer is indeed the relevant design metric, then is 50 Ω the optimum interface impedance? And second, is maximum power transfer even the correct design goal?

C.2 Impedance Matching for Maximum Power Transfer

Conjugate impedance matching results in maximum power transfer from the source to the load. Under this condition, the reflection coefficient is not necessarily zero. The reflection coefficient Γ is defined as [59]

$$\Gamma = \frac{Z_s - Z_l}{Z_s + Z_l} \tag{C.2}$$

where Z_s and Z_l are the source and load impedances, respectively. Maximum power transfer requires $Z_l = Z_s^*$, while $\Gamma = 0$ requires $Z_l = Z_s$. For a complex source impedance, conjugate impedance matching does not eliminate reflections on the transmission line connecting the source and the load. However, if the source impedance is purely real, then the condition for maximum power transfer delivered to the load is identical to the condition for no reflections, which partially explains the choice of a standard 50-Ω interface impedance.

The choice of 50 Ω is also based on the use of coaxial cables, where the 50-Ω interface resistance is a compromise between the 30-Ω resistance for maximum power capacity and the 77-Ω resistance for minimum attenuation [25], [59]. The characteristic impedance of a coaxial line is

$$Z_0 = \frac{\eta}{2\pi} \ln \frac{b}{a} \tag{C.3}$$

while the attenuation due to finite conductivity is

$$\alpha_c = \frac{R_s}{2\eta \ln(b/a)} \left(\frac{1}{a} + \frac{1}{b} \right) \tag{C.4}$$

where R_s is the surface resistivity of the conductors, η is the intrinsic impedance of the dielectric material, and a and b are the radii of the inner and outer conductors, respectively. The attenuation is minimum when $x \ln x = 1 + x$, where $x = b/a$, and the

corresponding characteristic impedance is 77 Ω for $\eta = \eta_0 = 377\,\Omega$ in free-space. The power capacity of a coaxial line is given by

$$P_{max} = \frac{\pi a^2 E_d^2}{\eta_0} \ln \frac{b}{a} \tag{C.5}$$

where E_d is the electric field strength at breakdown. The power capacity is maximum when $\ln(b/a) = 1/2$ and the corresponding characteristic impedance is 30 Ω.

Despite the integration of increasingly more components onto a single chip, many highly-integrated transceivers still rely on an external antenna and an external RF filter. In this case, a coaxial cable usually connects the antenna to the filter, while a short board trace usually connects RF filter to the transceiver chip. The use of a coaxial cable between the antenna and the RF filter motivates the use of a 50-Ω interface impedance. In particular, one of the most important goals of the transmitter is to efficiently deliver as much signal power as possible to the transmission medium. Conventional designs usually also rely on a 50-Ω interface impedance between the output of the RF filter and the input of the transceiver. Commercially-available RF filters are typically designed assuming doubly-terminated source and load impedances of 50 Ω, and consequently, deviating from 50 Ω results in poor and unpredictable RF filter performance.

For a transceiver which relies on a custom RF filter, a 50-Ω interface impedance between the antenna and the RF filter is still a prudent choice when a coaxial cable connects the two components. However, the interface impedance between the RF filter and the transceiver chip no longer needs to be 50 Ω since a short board trace instead of a coaxial cable connects the two components. Nevertheless, the 50-Ω requirement between the antenna and the RF filter does impose some restrictions on the interface impedance between the RF filter and the transceiver chip. In Fig. C.1, the RF filter is represented by a transmission, or ABCD, matrix:

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix}. \tag{C.6}$$

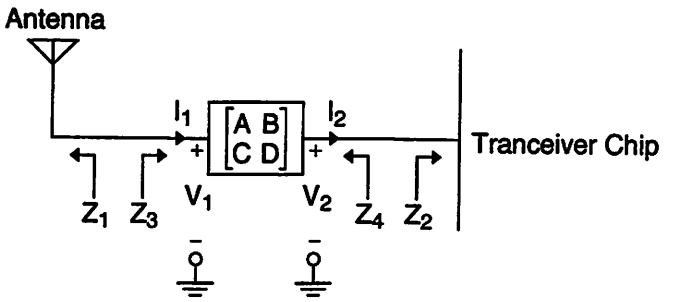


Figure C.1: Transmission matrix representation of the RF filter.

The input and output impedances of the RF filter are, respectively,

$$Z_3 = \frac{V_1}{I_1} = \frac{AV_2 + BI_2}{CV_2 + DI_2} = \frac{AZ_2 + B}{CZ_2 + D} \tag{C.7}$$

$$Z_4 = -\frac{V_2}{I_2} = -\frac{DV_1 - BI_1}{-CV_1 + AI_1} = \frac{DZ_1 + B}{CZ_1 + A}. \tag{C.8}$$

Next, assuming that the RF filter is reciprocal, i.e., no active devices, ferrites, or plasmas, results in the following constraint:

$$AD - BC = 1 \tag{C.9}$$

while assuming that the RF filter is lossless yields the following additional constraints:

$$A = A_R + jA_I = A_R \tag{C.10}$$

$$B = B_R + jB_I = jB_I \tag{C.11}$$

$$C = C_R + jC_I = jC_I \tag{C.12}$$

$$D = D_R + jD_I = D_R. \tag{C.13}$$

In order to achieve maximum power transfer between the antenna and the RF filter, a conjugate match is required:

$$Z_1 = Z_3^* = \left[\frac{A_R(R + jX) + jB_I}{jC_I(R + jX) + D_R}\right]^* = \frac{A_R R - j(B_I + A_R X)}{D_R - C_I X = jC_I R} \tag{C.14}$$

where $Z_2 = R + jX$. Substituting (C.14) into (C.8) results in the following expression for Z_4:

$$Z_4 = \frac{D_R \dfrac{A_R R - j(B_I + A_R X)}{D_R - C_I X - jC_I R} + jB_I}{jC_I \dfrac{A_R R - j(B_I + A_R X)}{D_R - C_I X - jC_I R} + A_R} = R - jX = Z_2^*. \qquad \text{(C.15)}$$

Hence, when a conjugate impedance match is required between the antenna and the RF filter, a conjugate match automatically results between the RF filter and the transceiver chip. In particular, for a 50-Ω match between the antenna and the RF filter, the interface impedance between the RF filter and the transceiver chip must also be purely real. However, this interface impedance does not have to be 50 Ω. For example, with fixed current consumption, a larger interface impedance results in better noise figure performance in the inductively-degenerated LNA as illustrated in Fig. 6.20. However, even for a custom RF filter implementation, the interface impedances cannot be chosen independently. For Butterworth, Bessel, and odd-order Chebyshev responses, the load and source resistances of the filter must be equal, while for an even-order Chebyshev response, the load and source resistances are related but not necessarily equal [59]. In the latter case, the relationship between the load and source resistances depends on the amount of passband ripple. An impedance transformation network may be placed between the RF filter and the transceiver to increase design flexibility at the expense of increased complexity.

C.3 Impedance Matching for Minimum Noise Figure

One of the most important goals of the transmitter is to efficiently deliver as much signal power as possible to the transmission medium. Consequently, impedance matching for maximum power transfer is a worthy design goal. On the other hand, one of the most important goals of the receiver is to amplify a potentially weak desired signal without corrupting it with noise. In this case, impedance matching for maximum power transfer is not as critical as matching for minimum noise figure. However, when the transmitter and receiver share the same RF filter and antenna, the interface impedance between the RF filter and the LNA is usually designed for maximum power transfer out of convenience rather than necessity.

Impedance matching for minimum noise figure is critical for very low noise applications, such as receivers for radio astronomy. Several LNA topologies are analyzed in Sections 6.2.4 and 6.2.5 and the optimum source admittance resulting in minimum noise figure for each of these topologies is summarized in Table 6.1. Minimum noise figure is achieved by designing the output admittance of the RF filter that precedes the LNA to be equal to the required optimum source admittance.

If the input admittance of the LNA is set equal to the conjugate of the optimum admittance for minimum noise figure, then a simultaneous match for minimum noise figure and maximum power transfer may be achieved. For all of the LNA topologies described in Sections 6.2.4 and 6.2.5, an impedance match which results in both minimum noise figure and maximum power transfer is impossible.

C.4 Impedance Matching for Maximum Voltage Transfer

With the increasing use of CMOS technology for RF applications, the traditional microwave approach of conjugate impedance matching for maximum power transfer has become quite controversial, especially for the receiver. One of the main goals of the RF section of the receiver is to amplify the voltage of the received signal for processing by the subsequent baseband section. In this case, designing the front-end components of the receiver, including the antenna, RF filter, and LNA, for maximum voltage transfer seems to be a more appropriate design goal.

In order to better understand the implications of impedance matching for either criterion, consider the following two design scenarios:

1. the source resistance is fixed but the designer has the freedom to select the load resistance;
2. both the source and load resistances are fixed but the designer has the freedom to select the network which connects the two resistances (Fig. C.2).

For the first scenario, selecting the load resistance R_l to be equal to the source resistance R_s results in maximum power transfer, $P_o = V_s^2/(4R_l)$, where V_s is the source voltage. In this case, the corresponding output voltage is $V_o = V_s/2$. In contrast, selecting $R_l \to \infty$

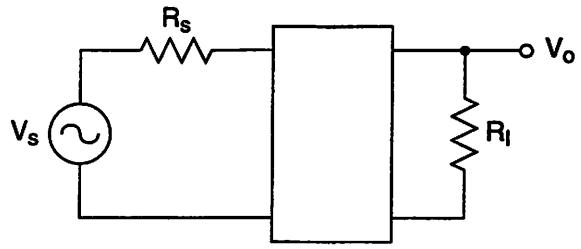


Figure C.2: R_s and R_l are fixed but the network connecting them is allowed to vary.

results in maximum voltage transfer, $V_o = V_s$, while the corresponding output power is zero. In the latter case, the power gain is zero but the voltage gain is a factor of two higher. In this scenario, the two design criteria have significantly different ramifications on the optimum value of R_l.

For the second scenario, there are numerous alternatives for connecting R_s and R_l. One approach is to simply connect the two resistances directly. In this case, the output voltage and output power are, respectively,

$$V_o = \frac{R_l}{R_s + R_l} V_s \tag{C.16}$$

$$P_o = \frac{R_l}{(R_s + R_l)^2} V_s^2. \tag{C.17}$$

If $R_s = 50\ \Omega$ and $R_l = 5\ \text{k}\Omega$, then $V_o = 0.99 V_s$ and $P_o = 1.96 \times 10^{-4} V_s^2$.

A second approach is to connect R_s and R_l together, and then, assuming that $R_l > R_s$, connect a third resistor R_p in parallel with R_l such that the equivalent resistance is $R_p \parallel R_l = R_s$. In this case, the output voltage and output power are $V_o = V_s/2$ and $P_o = V_s^2/[4(R_p \parallel R_l)]$. If $R_s = 50\ \Omega$ and $R_l = 5\ \text{k}\Omega$, then $V_o = 0.5 V_s$ and $P_o = 5 \times 10^{-3} V_s^2$. By connecting R_p in parallel with R_l, the equivalent resistance is matched to R_s, resulting in a higher output power but a lower output voltage than the corresponding values in the first approach.

A third approach is to connect the two resistances through a transformer as illustrated in Fig. C.3 [18]. In this case, the output voltage and output power are, respectively,

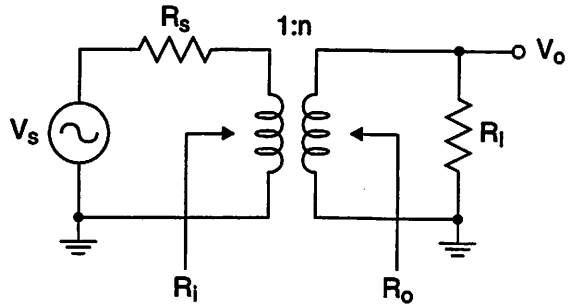


Figure C.3: R_s and R_l are connected through a transformer.

$$V_o = \frac{nR_l}{n^2 R_s + R_l} V_s \tag{C.18}$$

$$P_o = \frac{n^2}{(n^2 R_s + R_l)^2} V_s^2 \tag{C.19}$$

where n is the transformer turns ratio. The maximum output voltage and maximum output power are, respectively,

$$V_{omax} = \frac{1}{2}\sqrt{\frac{R_l}{R_s}} V_s \tag{C.20}$$

$$P_{omax} = \frac{V_s^2}{4R_s} \tag{C.21}$$

and in both cases, the corresponding turns ratio is

$$n_{opt} = \sqrt{\frac{R_l}{R_s}}. \tag{C.22}$$

In this case, if $R_s = 50\ \Omega$ and $R_l = 5\ \text{k}\Omega$, then $V_{omax} = 5V_s$ and $P_{omax} = 5\times10^{-3} V_s^2$, both of which are significantly larger than the corresponding values in the first approach. Although the output power is identical to that in the second approach, the output voltage is significantly higher. In this approach, the two design criteria result in identical values for the optimum transformer turns ratio, and thus, designing for maximum power transfer is equivalent to designing for maximum voltage transfer. In addition, for $n = n_{opt}$, the resistance R_i is equal to R_s, while the resistance R_o is equal to R_l. In other words, when

power transfer and voltage transfer are both maximum, R_i and R_o are matched to R_s and R_l, respectively.

The examples for the second scenario have profound implications for the design of the LNA in highly-integrated receiver implementations. For these implementations, R_s is the fixed driving-point resistance of the external antenna, while R_l is the fixed load resistance at the output of the integrated LNA, which is given by (6.129) for a tuned load. In this case, the designer has the freedom to choose the LNA topology which connects R_s and R_l. Indeed, selecting the LNA topology for maximum voltage transfer rather than maximum power transfer is commensurate with the overall goal of sufficiently amplifying the voltage of the received signal for processing by the subsequent baseband section. Based on this design criterion, the common-source and inductively-degenerated common-source topologies are analyzed below.

C.4.1 Common-Source LNA

When the source and load resistances are connected through a transistor in the common-source configuration as illustrated in Fig. C.4, the voltage gain is given by

$$\frac{V_o}{V_s} = -\frac{g_m R_l}{1 + j\omega R_s C_{gs}}. \tag{C.23}$$

From (6.59), assuming that $g_g << \omega C_{gs}$, the noise factor is

$$F = 1 + \frac{\gamma G_s}{\alpha g_m} + \left(\frac{\gamma}{\alpha} + \frac{\alpha\delta}{5} + 2|c|\sqrt{\frac{\delta\gamma}{5}}\right)\frac{\omega^2 C_{gs}^2}{g_m G_s}. \tag{C.24}$$

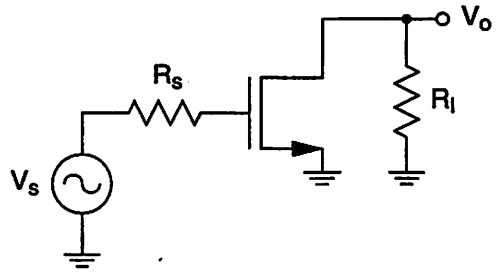


Figure C.4: Common-source LNA.

For simplicity, the long-channel expression for g_m in (6.94) is substituted into (C.24), and the noise factor is minimum when

$$C_{gs} = C_{gsopt} = \frac{G_s}{\omega}\sqrt{\frac{\frac{\gamma}{\alpha}}{3\left(\frac{\gamma}{\alpha}+\frac{\alpha\delta}{5}+2|c|\sqrt{\frac{\delta\gamma}{5}}\right)}} \tag{C.25}$$

and the corresponding minimum noise factor is

$$\begin{aligned} F_{min} &= 1+\frac{4}{3}L\sqrt{\frac{\omega G_s}{\mu_{eff} I_D}}\left(\frac{\gamma}{\alpha}\right)^{\frac{3}{4}}\left[\frac{1}{3}\left(\frac{\alpha\delta}{5}+\frac{\gamma}{\alpha}+2|c|\sqrt{\frac{\delta\gamma}{5}}\right)\right]^{\frac{1}{4}} \\ &= 1+\frac{0.0410}{\sqrt{I_D}} \end{aligned} \tag{C.26}$$

where the latter equality assumes that $\alpha = 0.8$, $\delta = 4$, $\gamma = 2$, $c = j0.395$, $G_s = 1/(50\ \Omega)$, $\omega = 2\pi 2\times10^9$ rad/s, $L = L_{eff} = 0.18$ μm ($L_{drawn} = 0.25$ μm), $\mu_{eff} = 400$ cm^2/Vs, and $E_{sat} = 5\times10^4$ V/cm. Under this condition, the voltage gain is

$$\frac{V_o}{V_s} = -\frac{1.63\sqrt{I_D}R_l}{1+j0.449} \tag{C.27}$$

and the magnitude is

$$\left|\frac{V_o}{V_s}\right| = 1.48\sqrt{I_D}R_l . \tag{C.28}$$

As discussed in Section 6.2.1, designing the RF filter for a particular frequency response requires knowledge of the source and load impedances of the antenna and LNA, respectively. A standard interface impedance, e.g., 50 Ω, allows RF filters to be designed independently from the antenna and the LNA based on doubly-terminated filter design techniques [20]. For the common-source LNA, the load to the RF filter is the gate-source capacitance of the MOS device. The RF filter is still preceded by an antenna, so the driving-point resistance of the antenna serves as the source impedance to the RF filter. The equivalent circuit for a second-order low-pass network is illustrated in Fig. C.5, and

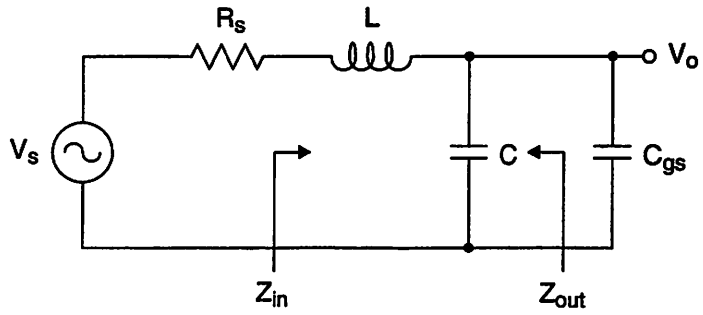


Figure C.5: Singly-terminated RF filter.

the design of this filter is based on the same approach used for singly-terminated filters [20]. The voltage transfer function is

$$\frac{V_o}{V_s} = \frac{1}{1-\omega^2 L(C+C_{gs})+j\omega R_s(C+C_{gs})} \tag{C.29}$$

and the following component values are required in order to achieve a second-order Butterworth low-pass frequency response:

$$L = \frac{R_s}{\sqrt{2}\,\omega_c} \tag{C.30}$$

$$C = \frac{\sqrt{2}}{R_s \omega_c} - C_{gs}. \tag{C.31}$$

Thus, the desired filter response may still be achieved in the case of a common-source LNA designed for maximum voltage transfer. However, a full reflection occurs at the interface between the antenna and the RF filter. This reflected signal is reradiated from the antenna and can potentially interfere with other receivers nearby. For low transmit power levels, e.g., 1 mW, the received signal is quite weak and the reradiated signal is even weaker, and thus, the impact of a full reflection occurring at the interface between the antenna and the RF filter is minimal.

C.4.2 Inductively-Degenerated Common-Source LNA

When the source and load resistances are connected through an inductively-degenerated common-source LNA as illustrated in Fig. C.6, the voltage gain is given by

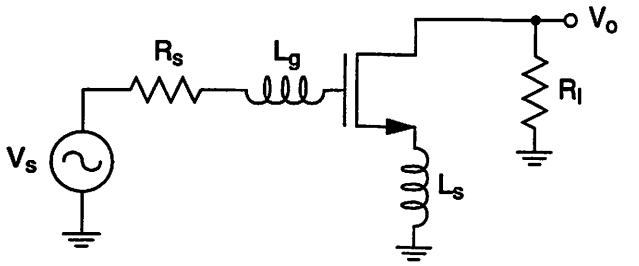


Figure C.6: Inductively-degenerated common-source LNA.

$$\left.\frac{V_o}{V_s}\right|_{\omega=\omega_c} = -\frac{g_m R_l}{j\omega_c(g_m L_s + R_s C_{gs})} \tag{C.32}$$

assuming that

$$\omega^2 = \omega_c^2 = \frac{1}{C_{gs}(L_g + L_s)}. \tag{C.33}$$

The maximum voltage gain is achieved when $L_s = 0$:

$$\left.\frac{V_o}{V_s}\right|_{L_s=0} = -\frac{g_m R_l}{j\omega R_s C_{gs}}. \tag{C.34}$$

Instead, if a power match is required at the LNA input,

$$R_s = \frac{g_m L_s}{C_{gs}} \tag{C.35}$$

and the voltage gain becomes

$$\left.\frac{V_o}{V_s}\right|_{g_m L_s = R_s C_{gs}} = -\frac{g_m R_l}{j2\omega R_s C_{gs}}. \tag{C.36}$$

Thus, designing for a power match at the LNA input degrades the voltage gain by a factor of two.

Relying on the long-channel expression for g_m in (6.94), the minimum noise figure is given by (6.96) and repeated here for convenience:

$$F_{min} = 1 + \frac{4}{3} L \sqrt{\frac{\omega G_s}{\mu_{eff} I_D}} \left(\frac{\alpha\delta}{5}\right)^{\frac{3}{4}} \left[\frac{1}{3}\left(\frac{\alpha\delta}{5} + \frac{\gamma}{\alpha} + 2|c|\sqrt{\frac{\delta\gamma}{5}}\right)\right]^{\frac{1}{4}} \qquad (C.37)$$
$$= 1 + \frac{0.0148}{\sqrt{I_D}}.$$

The corresponding voltage gain is

$$\left.\frac{V_o}{V_s}\right|_{L_s=0} = j5.10\sqrt{I_D} R_l . \qquad (C.38)$$

Thus, both the voltage gain and the noise performance of the inductively-degenerated common-source topology are superior. In addition, even if this LNA is designed for a power match at the input, the voltage gain is still superior to that of the common-source topology.

In the case of a power match, the LNA input resistance is equal to the driving-point resistance of the antenna. In this case, design of the preceding RF filter for a particular response is based on the approach used for doubly-terminated filters already described in Section 6.2.1. However, when $L_s = 0$, the input impedance of the LNA is zero at $\omega = \omega_c$, and the preceding RF filter is neither singly terminated nor doubly terminated. Thus, for applications which require an RF filter for increased selectivity, this latter approach may not be feasible. Nevertheless, even when the inductively-degenerated common-source LNA is designed for a power match, its voltage gain is superior to that of the common-source topology designed for maximum voltage transfer.

C.5 Antenna Circuit Model

A common circuit model for the components which precede the LNA, such as the antenna, is illustrated in Fig. C.7. In this case, the source resistance R_s is usually 50 Ω and represents the driving-point impedance of the antenna, while V_s represents the voltage of the received signal from the antenna. One potential pitfall of this model is the implied independence between the voltage V_s and the resistance R_s, which may lead to the incorrect conclusion that the input SNR can be arbitrarily improved by decreasing R_s for

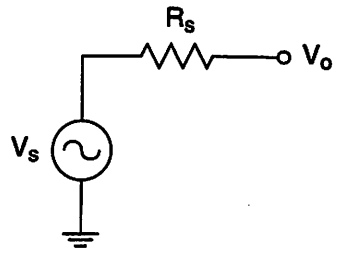


Figure C.7: Antenna circuit model.

a fixed transmit power level. The most important design goal for receivers is maximizing the output SNR:

$$SNR_o\,[\text{dB}] = SNR_i\,[\text{dB}] - NF\,[\text{dB}]. \tag{C.39}$$

In order to maximize the output SNR, the receiver noise figure should be minimized while the input SNR should be maximized. The open-circuit voltage of the antenna is

$$V_{oc} = V_s \tag{C.40}$$

and the noise voltage due to R_s is

$$V_n = \sqrt{4kTR_s\Delta f}\,. \tag{C.41}$$

Therefore, the input SNR is

$$SNR_i = \frac{V_{oc}^2}{V_n^2} = \frac{V_s^2}{4kTR_s\Delta f}. \tag{C.42}$$

If V_s and R_s are independent, then the input SNR may be increased by decreasing R_s. In reality, this is not the case and V_s actually depends on R_s. The vertical whip antenna is an antenna commonly used in mobile devices for cellular communications systems. For a vertical whip antenna over a group plane, the radiation resistance and series capacitance are given by [106], respectively,

$$R_r = 40\pi^2\left(\frac{h}{\lambda}\right)^2 \tag{C.43}$$

$$C_a\,[\text{pF}] = \frac{24.2h\,[\text{m}]}{\log\left(\frac{2h}{a}\right) - 0.7353} \tag{C.44}$$

where h is the antenna height, λ is the wavelength, and a is the antenna diameter. The resistance seen at the driving-point is typically larger than the radiation resistance due to additional losses such as that resulting from the physical resistance of the antenna. The open-circuit voltage of the antenna is given by

$$V_{oc} = V_s = Eh \tag{C.45}$$

where E is the electric field strength and h is the antenna height, and the input SNR is

$$SNR_i = \frac{V_{oc}^2}{4kTR_s\Delta f} \approx \frac{V_{oc}^2}{4kTR_r\Delta f} = \frac{(E\lambda)^2}{40\pi^2}. \tag{C.46}$$

Thus, the input SNR depends only on the electric field strength and the signal wavelength and cannot be arbitrarily increased by decreasing the driving-point resistance as suggested by (C.42). Consequently, if the transmit power level is fixed, then the only way to increase the output SNR is by decreasing the receiver noise figure.

C.6 Summary

Two questions were posed in the introduction to this appendix:

1. If maximum power transfer is indeed the relevant design metric, then is 50 Ω the optimum interface impedance?
2. Is maximum power transfer even the correct design goal?

As discussed in Section C.2, the 50-Ω requirement is a legacy from designs based on coaxial cables, for which 50 Ω is a good compromise between maximum power capacity and minimum attenuation. The use of a coaxial cable to connect the antenna to the RF filter in many transceivers still motivates the use of 50-Ω interface impedance. In particular, maximum power transfer is one of the most important design goals of the transmitter, which much efficiently deliver as much signal power as possible to the transmission medium. In addition, commercially-available RF filters are typically designed assuming doubly-terminated source and load impedances of 50 Ω, and consequently, deviating from 50 Ω results in poor and unpredictable RF filter performance.

For transceivers which rely on stripline or microstrip transmission lines rather than coaxial cables to connect the RF front-end components, maximum power transfer is still an appropriate design goal, particularly for the transmitter. Although the interface impedance no longer needs to be 50 Ω, deviating from a 50-Ω interface impedance dictates the use of a custom RF filter.

For broadcast applications which require a receiver rather than a transceiver, the design of the receiver is no longer constrained by the maximum power transfer requirement of the transmitter. For these applications, the receiver may be designed for other metrics such as minimum noise figure or maximum voltage transfer. Again, for both cases, the interface impedance no longer needs to be 50 Ω. However, deviating from a 50-Ω interface impedance again dictates the use of a custom RF filter.

Appendix D

Inductor Test Structures

D.1 Introduction

As described in Section 6.2.6, the inductively-degenerated differential LNA used in this receiver prototype relies on on-chip spiral inductors. A test chip (RFTRIPLED) was fabricated in order to evaluate the performance of various inductor structures. The process consists of six metal layers and a single polysilicon layer. The top two metal layers have a sheet resistance of 35 mΩ/□ while the first metal layer has a sheet resistance of 250 mΩ/□. The sheet resistance of the remaining metal layers is 55 mΩ/□. This process uses a low-resistivity 10-mΩ-cm silicon substrate with a 10-Ω-cm epitaxial layer. A total of seven different inductor structures were designed in ASITIC, a tool which provides rapid analysis, design, and optimization of inductors [107], [108]. The geometries of the seven inductors are summarized in Table D.1 and the parameters *D*, *W*, *S*, and *N* are defined in Fig. D.1. The first test structure is a planar spiral inductor implemented using only the top layer of metal, while the second test structure includes a patterned polysilicon ground shield [109]. This ground shield prevents the inductor electric field from penetrating the silicon substrate, which degrades the quality factor of the inductor, while patterning the shield prevents current flow which reduces the overall

inductor	description	D (μm)	W (μm)	S (μm)	N
1	m6	250	13.5	2	5.5
2	m6 with polysilicon shield	250	13.5	2	5.5
3	m5/m6 shunt	250	13.5	2	5.5
4	m4/m5/m6 shunt	250	13.5	2	5.5
5	m4/m5/m6 shunt with polysilicon shield	250	13.5	2	5.5
6	m5/m6 series	250	27	2	3.5
7	m2/m6 solenoid	250	13.5	2	15

Table D.1: Geometry of inductor test structures.

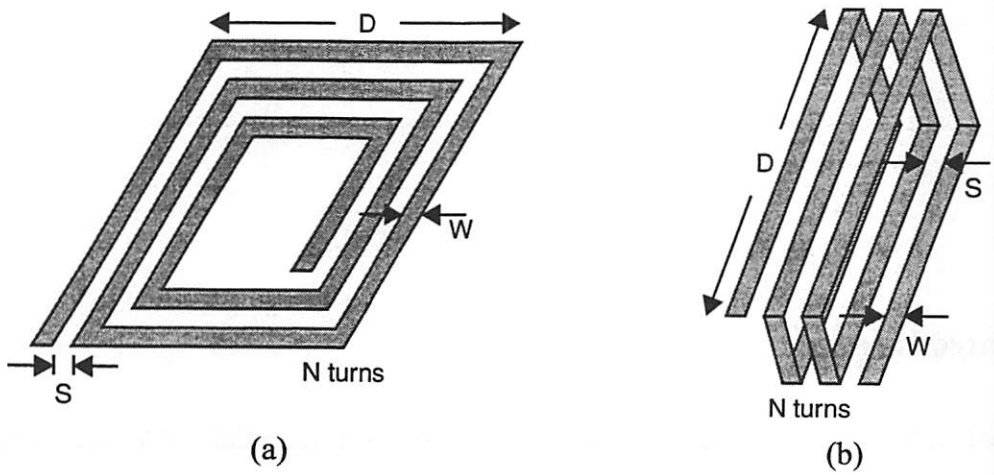


Figure D.1: Definition of geometric parameters. (a) Inductors 1 – 6. (b) Inductor 7.

inductance. One disadvantage of using a ground shield is that the capacitance between the inductor and ground increases, resulting in a lower self-resonance frequency.

Another source of degradation in the inductor quality factor is the resistance of the metal used to implement the inductor. The third and fourth test structures are implemented using the top two and top three layers of metal, respectively. These metal layers are all shorted together with numerous vias in order to reduce the series resistance of the inductor. As in the case of using a ground shield, using multiple metal layers also increases the quality factor of the inductor at the expense of lowering the self-resonance frequency. The fifth test structure combines both techniques, implementing the inductor using the top three layers of metal all shorted together as well as including a patterned polysilicon ground shield.

Finally, the sixth test structure is implemented using the top two layers of metal connected together in series, while the seventh inductor is a solenoid structure implemented using the second and sixth layers of metal as illustrated in Fig. D.1b.

D.2 ASITIC Simulation Results

ASITIC was used to simulate four of the inductor test structures. The equivalent circuit used to model the inductors is illustrated in Fig. 6.27. The simulated component values at 2 GHz as well as the self-resonance frequency of each of the inductors are summarized in Table D.2.

description	L (nH)	R_s (Ω)	C_1 (fF)	R_1 (Ω)	C_2 (fF)	R_2 (Ω)	f_s (GHz)	Q
m6	6.21	11.2	125	7.27	94.3	18.2	5.71	6.0, 6.2
m5/m6	5.97	7.57	158	8.61	122	14.5	5.18	8.2, 8.5
m4/m5/m6	5.79	6.77	200	8.49	160	14.3	4.68	8.4, 8.8
m5/m6 series	7.37	21.7	86.4	6.17	188	10.7	6.29	3.8, 3.2

Table D.2: Summary of inductor simulation results from ASITIC.

D.3 Test Chip Layout

The layout of the seven inductor test structures on the RFTRIPLED test chip is illustrated in Fig. D.2. All inductors are connected to the pad structure illustrated in Fig. D.3. The dimensions of the pad structure are designed to be compatible with GS-SG probes with a 125-μm pitch. The signal pads consist of the top three metal layers, all shorted together, while a fourth lower layer of metal (m2) acts as a ground shield [72]. The ground pad is implemented using the top five layers of metal, all shorted together. A ground ring consisting of the bottom three metal layers is placed around each of the inductor test structures at a distance of 50 μm, and all of these ground rings are connected to the substrate through a large number of substrate contacts.

A patterned polysilicon ground shield is positioned below two of the inductor test structures. The shields are connected to ground along the ground rings which surround the inductors. The layout of the polysilicon shield is illustrated in Fig. D.4.

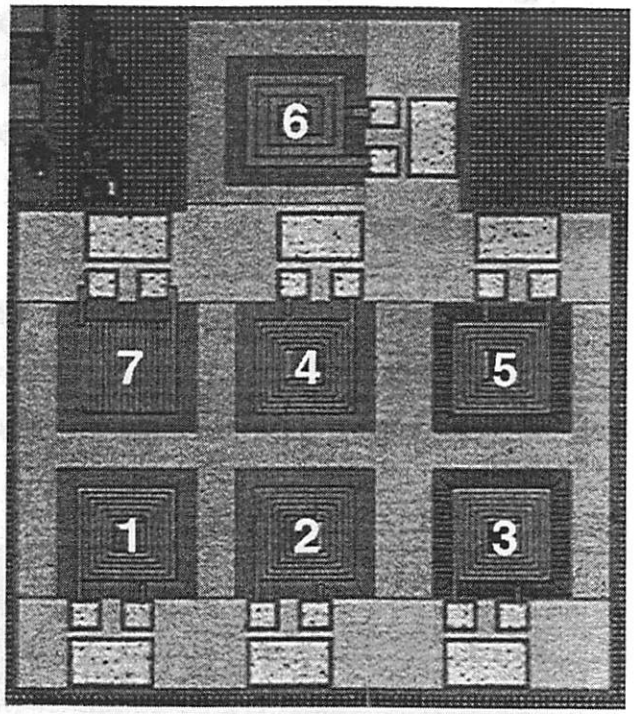


Figure D.2: Layout of inductor test structures.

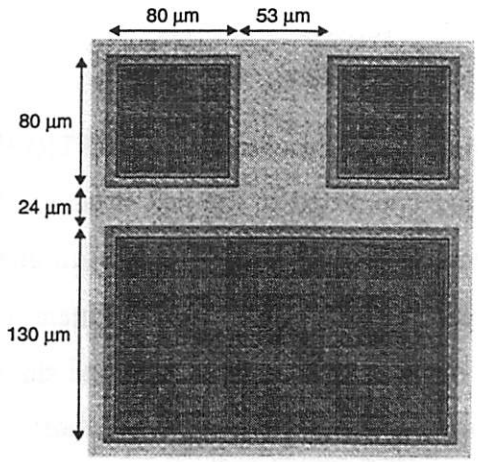


Figure D.3: Pad structure for inductors.

D.4 Measurement Results

The S parameters for each of the inductor test structures were measured using an HP8719C network analyzer. The component values at 2 GHz for the equivalent circuit illustrated in Fig. 6.27 were extracted from the measured S parameters and these results are summarized in Table D.3 along with the self-resonance frequency of each of the

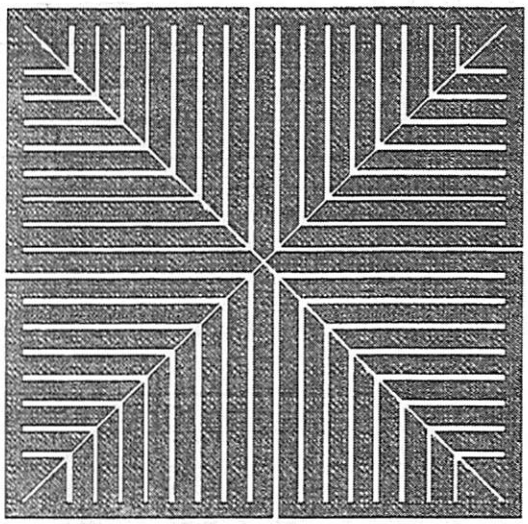


Figure D.4: Patterned polysilicon ground shield.

inductors. The dc resistances of some of the inductor test structures were also measured and are summarized in Table D.4.

The measured quality factors are significantly worse than the values predicted by ASITIC. Because this process uses a low-resistivity substrate, eddy currents play a significant role in limiting the achievable quality factor in this process. The version of ASITIC used for the initial simulations does not account for the effect of eddy currents flowing in the substrate. Although the most recent version of ASITIC does include eddy current effects, this version was not available at the time of initial simulations. The inductor test structures were simulated again in the latest version of ASITIC and these results are summarized in Table D.5. The quality factors predicted by the most recent

description	L (nH)	R_s (Ω)	C_1 (fF)	R_1 (Ω)	C_2 (fF)	R_2 (Ω)	f_s (GHz)	Q
m6	5.7	22.5	170	35	150	–18	5.7	2.5, 2.8
m6 with polysilicon shield	5.6	19.5	160	–8	160	8	5.5	3, 3.1
m5/m6	5.45	19.8	190	50	190	–40	5.2	2.6, 3
m4/m5/m6	5.25	15.4	220	19	235	–18	4.8	3.3, 3.6
m4/m5/m6 with polysilicon shield	5.25	15.9	246	–4	253	11	4.6	3.1, 3.3
m5/m6 series	9	82	270	16	130	–50	2.2	0.5, 1
m2/m6 solenoid	2	37	580	21	320	–13	5.4	0.25, 0.45

Table D.3: Summary of measured results from the inductor test chip.

description	R (Ω)
m6	11.8
m6 with polysilicon shield	11.6
m5/m6 shunt	7.3
m4/m5/m6 shunt	6.0
m4/m5/m6 shunt with polysilicon shield	6.4

Table D.4: Measured dc resistance of inductor test structures.

description	L (nH)	R_s (Ω)	C_1 (fF)	R_1 (Ω)	C_2 (fF)	R_2 (Ω)	f_s (GHz)	Q
m6	5.91	19.9	123	7.7	104	15.1	5.89	3.3, 3.3
m5/m6	5.61	16	154	6.74	131	15.4	5.40	3.7, 3.8
m4/m5/m6	5.41	15.2	194	6.91	167	15.2	4.89	3.7, 3.7
m5/m6 series	7.3	64.1	74.3	−16.3	225	11.1	6.7	1.25, 0.87

Table D.5: Summary of inductor simulation results from ASITIC version 3.19.00.

description	L (nH)	R_s (Ω)	C_1 (fF)	R_1 (Ω)	C_2 (fF)	R_2 (Ω)	f_s (GHz)	Q
m6	6.2	18	143	5	136	18		3.6, 3.6
m5/m6	5.9	14.2	178	8	168	12.5		4.1, 4.2

Table D.6: Summary of inductor simulation results from Momentum.

version of ASITIC are much closer to the measured values. A couple of the inductors were also simulated in Momentum, an electromagnetic simulation tool which is part of the Advanced Design System software from Agilent. The simulation results at 2 GHz are summarized in Table D.6.

From the measured results, the inductor implemented using the top three metal layers, all shorted together, offers the best quality factor. Consequently, this configuration is used to implement the on-chip spiral inductors for the inductively-degenerated differential LNA in the receiver prototype.

References

[1] Bluetooth, http://www.bluetooth.com.

[2] S. V. Saliga, "An introduction to IEEE 802.11 wireless LANs," *2000 IEEE Radio Frequency Integrated Circuits Symposium*, pp. 11–14, June 2000.

[3] European Telecommunications Standards Institute, http://www.etsi.org.

[4] H. D. Graves, "A detailed analysis of MMDS and LMDS," *1997 IEEE MTT-S Symposium on Technologies for Wireless Applications*, pp. 7–10, February 1997.

[5] N. Zhang, A. Poon, D. Tse, R. Brodersen, and S. Verdu, "Trade-offs of performance and single chip implementation of indoor wireless multi-access receivers," *IEEE Wireless Communications and Networking Conference*, pp. 226–230, September 1999.

[6] Murata Electronics North America, Inc., http://www.murata.com.

[7] A. Abidi, "Direct-conversion radio transceivers for digital communications," *IEEE Journal of Solid-State Circuits*, vol. 30, pp. 1399–1410, December 1995.

[8] C. Hull, J. Tham, and R. Chu, "A direct-conversion receiver for 900 MHz (ISM band) spread-spectrum digital cordless telephone," *IEEE Journal of Solid-State Circuits*, vol. 31, pp. 1955–1963, December 1996.

[9] T. Cho, E. Dukatz, M. Mack, D. Macnally, M. Marringa, S. Mehta, C. Nilson, L. Plouvier, and S. Rabii, "A single-chip CMOS direct-conversion transceiver for 900 MHz spread-spectrum digital cordless phones," *IEEE International Solid-State Circuits Conference*, pp. 228–229, February 1999.

[10] P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*. New York: John Wiley, 1993.

[11] J. C. Rudell, J.-J. Ou, T. B. Cho, G. Chien, F. Brianti, J. A. Weldon, and P. R. Gray, "A 1.9-GHz wide-band IF double conversion CMOS receiver for cordless telephone applications," *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 2071–88, December 1997.

[12] J. Crols and M. S. J. Steyaert, "A single-chip 900 MHz CMOS receiver front-end with a high performance low-IF topology," *IEEE Journal of Solid-State Circuits*, vol. 30, pp. 1483–1492, December 1995.

[13] J. F. Wilson, R. Youell, T. H. Richards, G. Luff, and R. Pilaski, "A single-chip VHF and UHF receiver for radio paging," *IEEE Journal of Solid-State Circuits*, vol. 26, pp. 1944–1950, December 1991.

[14] J. Crols and M. Steyaert, *CMOS Wireless Transceiver Design*. Boston: Kluwer Academic Publishers, 1997.

[15] B. Law and M. Groh, "Identifying RF-related impairments in full-service digital networks," *Microwave Journal*, vol. 39, pp. 88, 90, 92, 94, March 1996.

[16] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.

[17] Digi-Key Corporation, http://www.digikey.com.

[18] S. Sheng and R. Brodersen, *Low-Power CMOS Wireless Communications: A Wideband CDMA System Design*. Boston: Kluwer Academic Publishers, 1998.

[19] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. New York: John Wiley,1998.

[20] M. E. Van Valkenburg, *Analog Filter Design*. New York: Oxford University Press, 1982.

[21] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. New Jersey: Prentice-Hall, 1989.

[22] C. M. Teuscher, *Low Power Receiver Design for Portable RF Applications: Design and Implementation of an Adaptive Multiuser Detector for an Indoor, Wideband CDMA Application*, Ph.D. Dissertation, University of California, Berkeley, 1998.

[23] J. C. Rudell, J. A. Weldon, J.-J. Ou, L. Lin, and P. R. Gray, *An Integrated GSM/DECT Receiver: Design Specifications*, University of California, Berkeley, Electronics Research Laboratory, Memorandum No. UCB/ERL M97/82, 1997.

[24] H. T. Friis, "Noise Figure of Radio Receivers," *Proceedings of the IRE*, vol. 32, pp. 419–422, July 1944.

[25] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. New York: Cambridge University Press, 1998.

[26] R. G. Meyer, Class Notes, *EECS 242: Advanced Integrated Circuits for Communications*, University of California, Berkeley, Spring 1995.

[27] K. S. Kundert, "Introduction to RF simulation and its application," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 1298–1319, September 1999.

[28] The MathWorks, Inc., "Band-Limited White Noise," *Using Simulink Version 3*, pp. 8-18–8-19, January 1999.

[29] The MathWorks, Inc., "How do I generate pink noise in MATLAB or Simulink?," http://www.mathworks.com/support/solutions/data/8333.shtml.

[30] N. J. Kasdin, "Discrete simulation of colored noise and stochastic processes and $1/f^{\alpha}$ power law noise generation," *Proceedings of the IEEE*, vol. 83, pp. 802–827, May 1995.

[31] C. H. Doan, *Design and Implementation of a Highly-Integrated Low-Power CMOS Frequency Synthesizer for an Indoor Wireless Wideband-CDMA Direct-Conversion Receiver*, M.S. Project Report, University of California, Berkeley, 2000.

[32] D. A. Sobel, University of California, Berkeley, personal communication.

[33] S. Brigati, F. Francesconi, P. Malcovati, D. Tonietto, A. Baschirotto, and F. Maloberti, "Modeling sigma-delta modulator non-idealities in SIMULINK®" *Proceedings of the 1999 IEEE International Symposium on Circuits and Systems*, pp. 384–387, May 1999.

[34] D. A. Sobel, *A 25 Ms/s Nyquist-Rate Sigma-Delta Modulator for a Wideband CDMA Receiver*, M.S. Project Report, University of California, Berkeley, Electronics Research Laboratory, Memorandum No. UCB/ERL M00/32, May 2000.

[35] C. Teuscher, D. Yee, N. Zhang, and R. Brodersen, "Design of a wideband spread spectrum radio using adaptive multiuser detection," *Proceedings of the 1998 IEEE International Symposium on Circuits and Systems*, pp. 593–599, May 1998.

[36] J. P. Vanderhaegen, *Design and Implementation of a Wideband CDMA Basestation Transmitter*, M.S. Project Report, University of California, Berkeley, 1999.

[37] I. Katzela and M. Naghshineh, "Channel assignment schemes for cellular mobile telecommunication systems: a comprehensive survey," *IEEE Personal Communications*, vol. 3, pp. 10–31, June 1996.

[38] S. Blackman, *A Low Power, 8-bit, 200 MHz Digital-to-Analog Converter*, M.S. Project Report, University of California, Berkeley, Electronics Research Laboratory, Memorandum No. UCB/ERL M99/61, 1999.

[39] B. Razavi, *RF Microelectronics*. New Jersey: Prentice-Hall, 1998.

[40] R. C. V. Macario, *Cellular Radio: Principles and Design*. New York: McGraw-Hill, New York, 1997.

[41] S. E. Alexander, "Radio propagation within buildings at 900 MHz," *Electronics Letters*, vol. 18, pp. 913–914, October 1982.

[42] J.-F. Lafortune and M. Lecours, "Measurement and modeling of propagation losses in a building at 900 MHz," *IEEE Transactions on Vehicular Technology*, vol. 39, pp. 101–108, May 1990.

[43] G. K. Chan and F. A. Razaqpur, "Spectrum requirements of an indoor pico-cell radio system," *IEEE Transactions on Vehicular Technology*, vol. 44, pp. 24–30, February 1995.

[44] A. J. Motley and J. M. P. Keenan, "Personal communication radio coverage in buildings at 900 MHz and 1700 MHz," *Electronics Letters*, vol. 24, pp. 763–764, June 1988.

[45] A. M. D. Turkmani and A. F. de Toledo, "Modelling of radio transmissions into and within multistorey buildings at 900, 1800 and 2300 MHz," *IEE Proceedings I*, vol. 140, pp. 462–470, December 1993.

[46] B. Razavi, "A 2.4-GHz CMOS receiver for IEEE 802.11 wireless LANs," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 1382–1385, October 1999.

[47] A. Pärssinen, J. Jussila, J. Ryynänen, L. Sumanen, and K.Halonen, "A 2-GHz wide-band direct conversion receiver for WCDMA applications," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 1893–1903, December 1999.

[48] C. C. Enz and G. C. Temes, "Circuit techniques for reducing the effects of op-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization," *Proceedings of the IEEE*, vol. 84, pp. 1584–1614, November 1996.

[49] K. K. Onodera, *Low Power Techniques for High-Speed Wireless Systems*, Ph.D. Dissertation, University of California, Berkeley, 2000.

[50] QUALCOMM Incorporated, http://www.qualcomm.com.

[51] K. Bult, "Analog broadband communication circuits in pure digital deep sub-micron CMOS," *IEEE International Solid-State Circuits Conference*, pp. 76–77, February 1999.

[52] S. Verdu, *Multiuser Detection*. Cambridge: Cambridge University Press, 1998.

[53] R. G. Meyer and W. D. Mack, "Monolithic AGC loop for a 160 Mb/s transimpedance amplifier, *IEEE Journal of Solid-State Circuits*, vol. 31, pp. 1331–1335, September 1996.

[54] K. M. Cho, "Optimum gain control for A/D conversion using digitizing I/Q data in quadrature sampling," *IEEE Transactions on Aerospace and Electronic Systems*, vol.27, pp.178–181, January 1991.

[55] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering.* Massachusetts: Addison-Wesley, 1994.

[56] S. S. Haykin, *Adaptive Filter Theory.* New Jersey: Prentice-Hall, 1996.

[57] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation.* Boston: Kluwer Academic Publishers, 1997.

[58] G. Tavares and M. S. Piedade, "High performance algorithms for digital signal processing AGC," *IEEE International Symposium on Circuits and Systems*, pp. 1529–1532, May 1990.

[59] D. M. Pozar, *Microwave Engineering.* New York: John Wiley, 1998.

[60] G. D. Vendelin, A. M. Pavio, and U. L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques.* New York: John Wiley, 1990.

[61] A. A. Abidi, "High-frequency noise measurements on FETs with small dimensions," *IEEE Transactions on Electron Devices*, vol. ED-33, pp. 1801–1805, November 1986.

[62] R. P. Jindal, "Hot-electron effects on channel thermal noise in fine-line NMOS field-effect t ransistors," *IEEE Transactions on Electron Devices*, vol. ED-33, pp. 1395–1397, September 1986.

[63] B. Wang, J. R. Hellums, and C. G. Sodini, "MOSFET thermal noise modeling for analog integrated circuits," *IEEE Journal of Solid-State Circuits*, vol. 29, pp. 833–835, July 1994.

[64] D. K. Shaeffer and T. H. Lee, "A 1.5-V, 1.5-GHz CMOS low noise amplifier," *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 745–759, May 1997.

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

[66] S. Wolf, *Silicon Processing for the VLSI Era, Volume 3: The Submicron MOSFET.* Sunset Beach, California: Lattice Press, 1995.

[67] S. Sheng, L. Lynn, J. Peroulas, K. Stone, I. O'Donnell, and R. Brodersen, "A low-power CMOS chipset for spread spectrum communications," *IEEE International Solid-State Circuits Conference*, pp. 346–347, February 1996.

[68] A. R. Shahani, D. K. Shaeffer, and T. H. Lee, "A 12-mW wide dynamic range CMOS front-end for a portable GPS receiver," *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 2061–2070, December 1997.

[69] B. Razavi, R.-H. Yan, and K. F. Lee, "Impact of distributed gate resistance on the performance of MOS devices," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 41, pp. 750–754, November 1994.

[70] Q. Huang, P. Orsatti, and F. Piazza, "Broadband, 0.25μm CMOS LNAs with sub-2dB NF for GSM applications," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference*, pp. 67–70, May 1998.

[71] Q. Huang, P. Orsatti, and F. Piazza, "GSM transceiver front-end circuits in 0.25-μm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 292–303, March 1999.

[72] 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 Journal of Solid-State Circuits*, vol. 31, pp. 880–889, July 1996.

[73] J.-J. Ou, University of California, Berkeley, personal communication, 1999.

[74] B. Limketkai, *The Design and Implementation of a Downconversion Mixer for a Wideband CDMA Receiver*, M.S. Project Report, University of California, Berkeley, Electronics Research Laboratory, Memorandum No. UCB/ERL M99/66, December 1999.

[75] Philips Semiconductors, http://www.semiconductors.philips.com.

[76] D. B. Leeson, "A simple model of feedback oscillator noise spectrum," *Proceedings of the IEEE*, vol. 54, pp. 329–330, February 1966.

[77] B. Razavi, "A study of phase noise in CMOS oscillators," *IEEE Journal of Solid-State Circuits*, vol. 31, pp. 331–343, March 1996.

[78] L. Lin, L. Tee, and P. R. Gray, "A 1.4GHz differential low-noise CMOS frequency synthesizer using a wideband PLL architecture," *IEEE International Solid-State Circuits Conference*, pp. 204–205, February 2000.

[79] J. M. Rabaey, *Digital Integrated Circuits: A Design Perspective.* New Jersey: Prentice-Hall, 1996.

[80] J. Crols and S. J. Steyaert, "A 1.5 GHz highly linear CMOS downconversion mixer," *IEEE Journal of Solid-State Circuits*, vol. 30, pp. 736–742, July 1995.

[81] B. Gilbert, "A precise four-quadrant multiplier with subnanosecond response," *IEEE Journal of Solid-State Circuits*, vol. SC-3, pp. 365–373, December 1968.

[82] R. P. Sallen and E. L. Key, "A practical method of designing RC active filters," *IRE Transactions on Circuit Theory*, vol. CT-2, pp. 74–85, March 1955.

[83] M. S. Ghausi and K. R. Laker, *Modern Filter Design: Active RC and Switched Capacitor*. New Jersey: Prentice-Hall, 1981.

[84] A. S. Sedra and K. C. Smith, *Microelectronic Circuits*. Philadelphia: Saunders College Publishing, 1991.

[85] Y. Tsividis, M. Banu, and J. Khoury, "Continuous-time MOSFET-C filters in VLSI," *IEEE Journal of Solid-State Circuits*, vol. SC-21, pp. 15–29, February 1986.

[86] H. Khorramabadi and P. R. Gray, "High-frequency CMOS continuous-time filters," *IEEE Journal of Solid-State Circuits*, vol. SC-19, pp. 939–948, December 1984.

[87] C. A. Laber and P. R. Gray, "A 20-MHz sixth-order BiCMOS parasitic-insensitive continuous-time filter and second-order equalizer optimized for disk-drive read channels," *IEEE Journal of Solid-State Circuits*, vol. 28, pp. 462–470, April 1993.

[88] R. Gregorian, K. W. Martin, and G. C. Temes, "Switched-capacitor circuit design," *Proceedings of the IEEE*, vol. 71, pp. 941–966, August 1983.

[89] R. Castello and P. R. Gray, "Performance limitations in switched-capacitor filters," *IEEE Transactions on Circuits and Systems*, vol. CAS-32, pp. 865–876, September 1985.

[90] R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*. New York: John Wiley, 1986.

[91] G. T. Uehara, Circuit Techniques and Considerations for Implementation of High Speed CMOS Analog-to-Digital Interfaces for DSP-Based PRML Magnetic Disk Read Channels," Ph.D. Dissertation, University of California, Berkeley, Electronics Research Laboratory, Memorandum No. UCB/ERL M93/96, August 1993.

[92] T. B. Cho, *Low-Power Low-Voltage Analog-to-Digital Conversion Techniques using Pipelined Architectures*, Ph.D. Dissertation, University of California, Berkeley, Electronics Research Laboratory, Memorandum No. UCB/ERL M95/23, 1995.

[93] T. B. Cho and P. R. Gray, "A 10 b, 20 Msample/s, 35 mW pipeline A/D converter," ," *IEEE Journal of Solid-State Circuits*, vol. 30, pp. 166–172, March 1995.

[94] T. B. Cho, G. Chien, F. Brianti, and P. R. Gray, "A power-optimized CMOS baseband channel filter and ADC for cordless applications," *1996 Symposium on VSLI Circuits*, pp. 64–65, June 1996

[95] G. Chien, *High Speed, Low Power, Low Voltage Pipelined Analog-to-Digital Converter*, M.S. Project Report, University of California, Berkeley, Electronics Research Laboratory, Memorandum No. UCB/ERL M96/27, 1996.

[96] H.-K. Yang and M. Snelgrove, "Symbol timing recovery using oversampling techniques," *IEEE International Conference on Communications*, pp. 1296–1300, June 1996.

[97] A. R. Feldman, High-Speed, Low-Power Sigma-Delta Modulators for RF Baseband Channel Applications, Ph.D. Dissertation, University of California, Berkeley, Electronics Research Laboratory, Memorandum No. UCB/ERL M97/62, 1997.

[98] D. G. Yee, *A Design Methodology for Highly-Integrated Low-Power Receivers for Wireless Communications*, Ph.D. Dissertation, University of California, Berkeley, 2001.

[99] M. C. Chang, C. Y. Wu, G. L. Lin, I. C. Liao, Y. H. Lee, and C. N. Chen, "Degradation of MOS transistor characteristics by gate charging damage during plasma processing," *1993 International Symposium on VLSI Technology, Systems, and Applications*, pp. 320–324, May 1993.

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

[101] A. Z. Wang, "A new design for complete on-chip ESD protection," *Proceedings of the IEEE 2000 Custom Integrated Circuits Conference*, pp. 87–90, May 2000.

[102] M.-D. Ker, T.-Y. Chen, C.-Y. Wu, and H.-H. Chang, "ESD protection design on analog pin with very low input capacitance for RF or current-mode applications," *Twelfth Annual IEEE International ASIC/SOC Conference*, pp. 352–356, September 1999.

[103] Y. L. Low, Y. Degani, K. V. Guinn, T. D. Dudderrar, J. A. Gregus, and R. C. Frye, "RF flip-module BGA package," *IEEE Transactions on Advanced Packaging*, vol. 22, pp. 111–115, May 1999.

[104] L. Moult and J. E. Chen, "The K-model: RF IC modeling for communication systems simulation," IEE Colloquium Analog Signal Processing, pp. 11/1–11/8, October 1998.

[105] H. Yoshida, H. Tsurumi, and Y. Suzuki, "DC offset canceller in a direct conversion receiver for QPSK signal reception," *Proceedings of Ninth*

International Symposium on Personal, Indoor, and Mobile Radio Communications, pp. 1314–1318, September 1998.

[106] U. L. Rohde, J. C. Whitaker, and T. T. N. Bucher, *Communications Receivers: Principles and Design*. New York: McGraw-Hill, 1997.

[107] A. M. Niknejad and R. G. Meyer, "Analysis, design, and optimization of spiral inductors and transformers for Si RF IC's," *IEEE Journal of Solid-State Circuits*, vol. 33, pp. 1470–1481, October 1998.

[108] ASITIC, http://formosa.eecs.berkeley.edu/~niknejad/asitic.html.

[109] C. P. Yue and S. S. Wong, "On-chip spiral inductors with patterned ground shields for Si-based RF IC's," *IEEE Journal of Solid-State Circuits*, vol. 33, pp. 743–752, May 1998.