Energy-Aware Design of 2.4-GHz CMOS Receiver Front-Ends for WPANs

Aaron Do

A thesis submitted to The Nanyang Technological University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

2010
Abstract

The last decade has seen the near complete integration of the wireless transceiver, and the rise of CMOS as the choice technology in consumer-based wireless applications such as mobile phones and wireless local area network (WLAN). Full system integration continues to be a topic of interest in the research field in order to minimize both the cost and the form-factor of wireless transceivers. However, a new trend is emerging in RFIC System on Chip (SoC) design.

In the interests of longer battery life, ultra-low power design has recently become a hot topic for applications such as wireless personal area networks (WPAN), and wireless sensor nodes. The IEEE 802.15.4 standard has been specifically designed to cater to such applications, and transceivers which follow this standard have been designed to operate using less than 10 mA of DC current. These designs have relied on simplified circuit configurations to minimize power consumption. Despite their relative successes, we believe that significantly more power consumption can be saved both by further simplifying the circuit structures, and dynamically adjusting the performance of the receiver (RX) based on the quality of the received signal. The latter method is termed energy-aware design and is the main topic of this work. We propose to dynamically control the power consumption of an RX front end based on the real-time required noise figure (NF). We also discuss certain system level architectural decisions which can make full use of the proposed power reduction method. Based on our proposal, an energy-aware front-end, designed around the IEEE 802.15.4 standard, is demonstrated which can save up to 72% of its nominal power consumption under good channel conditions. An alternative design has been proven to require only 2.2 mW in the nominal state, and can achieve 69 % power savings under good channel conditions.
Acknowledgements

Firstly, I would like to express my gratitude to my supervisor Prof. Boon Chirn Chye for his continuous support and advice throughout my studies and the writing of this thesis.

I would also like to thank Prof. Yeo Kiat Seng, for his supervision in the EHS-II project, as well as Prof. Siek Liter for useful, and lighthearted, discussions, and Prof. Zhang Yue Ping, my Master’s Degree supervisor.

I also must acknowledge the help of friends and colleagues, particularly Ali, Vamshi, and Alper, and the support from Mr. Lim Wei Meng, and Mr. T.S. Wong in all of my fabrication and measurements. For loan of their equipment and help in the measurement I would like to thank Mike from Rohde and Schwarz and Yong Wee from Agilent. My gratitude also extends to the lab technicians of IC Design I, II, and CICS. I must also thank Mediatek Inc. for supporting the cost of several of the fabrications.

Of course, this work would not have been possible without the love and support – and good home cooking – of my wife and colleague, Mrs. Luo Sha. I also thank my mother, brothers and sisters for molding me into the person I am today. Lastly, I would also like to thank my father, Prof. Do Manh Anh, whose 28 years of guidance have – despite my failings – brought me to this point.
# Table of Contents

Abstract ................................................................................................................................. i
Acknowledgements ................................................................................................................ ii
Table of Contents .................................................................................................................. iii
List of Figures ........................................................................................................................ v
List of Tables .......................................................................................................................... x

Chapter 1. Introduction

1.1 Motivation ..................................................................................................................... 1
1.2 Objectives and Major Contributions ............................................................................. 2
1.3 Organization of the Thesis ........................................................................................... 2

Chapter 2. Fundamentals of RFIC Design

2.1 Introduction ................................................................................................................... 5
2.2 Fundamental System Level Concepts ......................................................................... 5
2.3 Fundamental Circuit Level Concepts .......................................................................... 25

Chapter 3. Survey of Low Power Techniques for RF Transceivers

3.1 Overview of the State-of-the-Art ................................................................................ 37
3.2 System Level Power Saving Methods ....................................................................... 46

Chapter 4. System and Architectural Level Proposals

4.1 Introduction ................................................................................................................ 59
4.2 Proposed Energy-Aware Scheme .............................................................................. 59
4.3 Low Power Receiver Architectures ......................................................................... 75

Chapter 5. Circuit Design
5.1 Introduction ................................................................. 95
5.2 Available Technology ..................................................... 95
5.3 Impedance Matching ......................................................... 99
5.4 Biasing Point Optimization and Weak Inversion Biasing .......... 111
5.5 Passive Mixer Design and Analysis ................................. 128

Chapter 6. Full Front End Designs

6.1 Introduction ................................................................. 142
6.2 Energy-Aware Receiver ................................................... 142
6.3 Energy-Aware Receiver – Low-Power Architecture ............... 148
6.4 Summary ........................................................................ 166

Chapter 7. Conclusion

7.1 Conclusions ................................................................. 169
7.2 Future Direction ............................................................ 170

List of Publications .......................................................... 174

Full List of References ........................................................ 175
List of Figures

Fig. 2.1 Illustration of frequency translation and the image problem. The triangular signal is an undesired image while the other signal is desired. ................................................................. 7

Fig. 2.2 Illustration of the IEEE 802.15.4 standard’s system and channel bandwidths. ............................................................................................................. 8

Fig. 2.3 When the IF frequency (fIF) is less than half of the system bandwidth, the image frequency (fIM) falls inside the system bandwidth....................................................................................... 8

Fig. 2.4 The phase transfer function of a -90 degree phase shifting network .......................................................................................................................... 9

Fig. 2.5 A typical front-end for a high-IF receiver ......................................................... 11

Fig. 2.6 Low-IF front-ends. (a) The Hartley receiver and (b) the Weaver receiver........................................................................... 13

Fig. 2.7 Illustration of the effects of flicker noise on the IF for (a) 2 MHz IF (b) 4 MHz IF (c) 1 MHz IF (d) 2.5 MHz IF. .............................................. 14

Fig. 2.8 Illustration of LO leakage suppression when fLO ≠ fRF. ............. 15

Fig. 2.9 Illustration of BER ....................................................................................... 17

Fig. 2.10 Interference profile (translated to IF) and IIP3 requirements of the IEEE 802.15.4 standard..................................................... 22

Fig. 2.11 Illustration of the phase noise problem.............................................. 23

Fig. 2.12 Illustration of RF self-mixing................................................................. 24

Fig. 2.13 Model of an antenna connected to a BPF and a load. ................. 26

Fig. 3.1 Illustration of current-reuse. ................................................................. 39

Fig. 3.2 Illustration of effect of process variation and bandwidth on system gain for (a) 1 GHz RF bandwidth and (b) 0.5 GHz RF bandwidth. The center frequency of the RF bandwidth is shifted by approximately ±10 % from the nominal value. ............ 42

Fig. 3.3 (a) The simple TIA used in [5] and (b) its equivalent representation. ....................................................................................... 43
Fig. 3.4 Illustration of single-ended to differential conversion using a single-balanced mixer. ................................................................. 44
Fig. 3.5 A QPSK modulator.......................................................................................... 48
Fig. 3.6 A QPSK constellation diagram with four example vectors shown. ...................................................................................... 49
Fig. 3.7 Spreading and De-Spreading in CDMA .................................................. 52
Fig. 4.1 First and third derivatives of a MOSFET of unit aspect ratio with 0.6 V drain-source voltage..................................................... 62
Fig. 4.2 IIP_3 of the above device. ............................................................................ 62
Fig. 4.3 A generic differential amplifier.................................................................. 64
Fig. 4.4 NF (dB) and IIP_3 (dBm) requirements versus input signal power (dBm). ................................................................................. 65
Fig. 4.5 Illustration of how noise is added to a signal........................................ 67
Fig. 4.6 Block Diagram of a Typical RSSI............................................................ 69
Fig. 4.7 Output V_{RSSI} of the System in Fig. 4.6 for two different amplifier gains. The input power is referenced to a 50-Ω load. ............................................................................................................ 69
Fig. 4.8 Illustration of implemented NF and SNR_{out} for an EA receiver and a conventional receiver. ΔSNR_{out} and ΔNF are the uncertainties in the SNR_{out} and NF respectively.................................. 70
Fig. 4.9 A Direct-Conversion Receiver including RSSI. ...................................... 71
Fig. 4.10 Average power consumption breakdown for five recent works in literature and two proposed designs. The proposed designs have 4 power states represented by four levels.............. 73
Fig. 4.11 A simple active-RC low-pass filter............................................................ 77
Fig. 4.12 3D plot of a low-pass filter transfer function for a pole frequency of 0.3 +j0 rad/s. ................................................................. 78
Fig. 4.13 3D plot of a band-pass response formed using a complex pole frequency of 0.3 – 0.5j rad/s......................................................... 79
Fig. 4.14 Implementation of a complex band-pass filter....................................... 80
Fig. 4.15 Integration of the passive mixer and the CSF ..................................... 81
Fig. 4.16 Effect of addition of filtering capacitors on the op-amp input of an op-amp based CSF. (a) Without capacitors and (b) with capacitors. ................................................................................. 82
Fig. 4.17 Schematic of the op-amp used in [5] for the IF section TIA........ 83
Fig. 4.18 Comparison between (a) the active-RC based CSF-mixer
and (b) the Gm-C based CSF-mixer. ........................................... 84
Fig. 4.19 1st order filtering of the IEEE 802.15.4 interferers. ............... 85
Fig. 4.20 The simple biquadratic filter used in [21] to design a 3rd
order Butterworth filter.............................................................. 86
Fig. 4.21 Illustration showing the placement of replica passive mixers
and their connection to the LNA output node............................... 87
Fig. 4.22 Figure to illustrate the principle of operation of the tuning
loop. (a) The overall loop and (b) the digital amplitude
comparator.................................................................................. 87
Fig. 4.23 Illustration of a folded Gilbert Cell for use as a low frequency
multiplier circuit. ....................................................................... 88
Fig. 4.24 Response of the tuning loop at the loop filter output (V)
versus time (ms). ......................................................................... 89
Fig. 5.1 (a) Cross-sectional diagram of an n-type MOSFET and (b)
its small signal equivalent model................................................. 96
Fig. 5.2 Variation of fT with VGS in the 0.18 µm RF CMOS process........ 97
Fig. 5.3 Cross-sectional view of a MIM capacitor.............................. 98
Fig. 5.4 A cross-sectional view of a VNPN BJT................................. 98
Fig. 5.5 A simple system showing different impedance levels. ........... 101
Fig. 5.6 (a) inductive degeneration matching and (b) its equivalent
representation.............................................................................. 102
Fig. 5.7 resistive termination in (a) shunt and (b) series form. .......... 104
Fig. 5.8 (a) Common-Gate rmatch synthesis and (b) de-coupling
transconductance and rmatch in a differential design. ................. 106
Fig. 5.9 Inductive degeneration matching......................................... 107
Fig. 5.10 Different forms of the L-match including the (a), (c) low-pass
L-match, (b), (d) high-pass L-match, (a), (b) single-ended L-
match, and (c), (d) differential L-match. .................................... 109
Fig. 5.11 Cascaded matching networks. (a) A low-pass π network, (b)
a high-pass T network, and (c) a band-pass wideband
matching network...................................................................... 110
Fig. 5.12 Schematic of NFmin testbench. The op-amps are ideal. .......... 112
Fig. 5.13 $\text{NF}_{\text{min}}$ for different biasing currents and values of $R_{\text{max}}$ versus $V_{\text{GS}}$ at 2.4 GHz. .............................................................. 113
Fig. 5.14 $I_{\text{IV}}$ versus $V_{\text{GS}}$ for a MOSFET................................................... 114
Fig. 5.15 Small signal model for estimation of $G_{\text{TUmax}}$ of a MOSFET. ....... 116
Fig. 5.16 $G_{\text{TUmax}}$ versus frequency for different biasing voltages. $R_{\text{max}}$ is equal to 500 $\Omega$ .................................................... .................... 117
Fig. 5.17 Schematic of a weak-inversion biased LNA. ......................................... 118
Fig. 5.18 Micrograph of the proposed LNA showing area, inputs, and outputs. ...................................................................................... 119
Fig. 5.19 Voltage Gain and $S_{11}$ of the proposed LNA............................... 120
Fig. 5.20 Measured NF of the proposed LNA. ............................................. 121
Fig. 5.21 Schematic of the current-reuse LNA. ............................................. 123
Fig. 5.22 Voltage gain and $S_{11}$ of the proposed current reuse design........ 125
Fig. 5.23 Noise performance of the proposed current reuse LNA.............. 126
Fig. 5.24 Simulation of the settling time of the receiver. The receiver settles to the desired state within approximately 1 $\mu$s. ............... 127
Fig. 5.25 System model of the current-mode passive mixer. ....................... 129
Fig. 5.26 Decomposition of $g_{\tau}(t)$ in the time and frequency domain. (a) $g_{\tau}(t)$ (b) $G_{\tau}(f)$ (c) pulse train in time (d) pulse train in frequency (e) the sampling function in time (f) the sampling function in frequency. (e) and (f) show the sampling function for two different sampling function widths. ........................................ 130
Fig. 5.27 Comparison between theoretically calculated and simulated conversion gain ($CG$ in V/V), input conductance ($G_{\text{mix}}$) and output conductance ($G_{\text{out}}$). In simulation, the LO was 2.45 GHz, 250 mV$_{\text{pk}}$ per phase. ................................................. 135
Fig. 5.28 Switch width optimization including input-referred noise (IRN), op-amp loop-gain, and voltage gain. The LO was 2.45 GHz, 250 mV$_{\text{pk}}$ per phase. ............................................. 137
Fig. 6.1 Overall architecture of the energy-aware receiver in [1].............. 143
Fig. 6.2 Micrograph of the energy-aware receiver in [1]............................. 144
Fig. 6.3 Overall SSB NF of the full system............................................... 145
Fig. 6.4 Conversion gain of the full system in all gain modes. ................. 146
Fig. 6.5 Overall IIP$_3$ and LNA IIP$_3$ for all gain modes. ....................... 147
Fig. 6.6  Overall Receiver Architecture. ..................................................... 150
Fig. 6.7  Schematic of the LNA and Mixers................................................. 151
Fig. 6.8  Plot of S11 and S_{21} of the proposed LNA where port 2
reference impedance is 2 \, k\Omega, and port 1 reference
impedance is 50 \, \Omega. ........................................................................ 154
Fig. 6.9  Full receiver front-end gain at the CSF Input and Output............. 155
Fig. 6.10 Full receiver front-end NF. .......................................................... 155
Fig. 6.11 IF Section IIP_3 with (improved IIP_3) and without (poorer IIP_3)
the tuned passive mixer output pole. .................................................. 156
Fig. 6.12 Layout of the full RX................................................................. 157
Fig. 6.13 Schematic of the output buffers. .................................................. 158
Fig. 6.14 Measured S-parameters of the LNA........................................... 160
Fig. 6.15 Measured NF of the LNA. .......................................................... 161
Fig. 6.16 Receiver voltage gain up to the passive mixer output before
and after calibration.............................................................................. 162
Fig. 6.17 System IIP_3 results taken at all four gain modes at the mixer
output and the CSF output ................................................................. 163
Fig. 6.18 Overall system SSB NF for all four modes of operation.......... 164
Fig. 6.19 Micrograph of the proposed design............................................ 165
List of Tables

Table 2.1 Design Tradeoffs................................................................. 32
Table 3.1 Comparison of Prior Published Work............................... 37
Table 4.1 Active Versus Passive Down-Conversion............................. 76
Table 5.1 Comparison of Proposed LNA and Current Literature........... 122
Table 6.1 Comparison To Prior Published Work.................................. 147
Table 6.2 Simulated Performance of the Proposed Design.................... 153
Table 6.3 Measured Performance of the Proposed Design.................... 165
Chapter 1

Introduction

1.1 Motivation

The previous decade saw the near complete integration of the wireless transceiver and the rise of Complementary Metal-Oxide Semiconductor (CMOS) as the choice technology for digital intensive consumer based applications. While technologies such as Silicon Germanium (SiGe) Bipolar-CMOS (BiCMOS) provide better performance than CMOS for high speed analogue electronics, CMOS offers two important advantages: lower cost, and the ability to be easily integrated with digital electronics. The CMOS components in SiGe BiCMOS generally trail those of the CMOS technology by two generations [1] making SiGe BiCMOS less appealing for digital intensive applications.

Low power consumption is a phrase which has been the centre of attention for many technologies (the automobile industry for example) mainly for environmental reasons. In the context of mobile wireless applications we are mainly concerned with battery life both due to the problem of disposing of batteries and in the interests of extending the time over which a mobile device can be used without having to recharge. In particular, there are a slew of applications requiring low data-rate, short range wireless capabilities such as wireless sensor networks, wireless personal area networks (WPAN), wireless keyboards and mice, wireless headsets for mobile phones, cordless phones etc. The IEEE 802.15.4 standard [2] was specifically designed to cater to such applications, and has allowed designers to reach new low levels of power consumption.
Despite such advances, we believe there is ample room for further energy reduction through both novel circuit design and system planning.

1.2 Objectives and Major Contributions

The objectives of our research are to develop a thorough understanding of low-power RFIC design from system to circuit level and introduce new methods for low-power design.

We evaluate the tradeoffs between gain, bandwidth, noise figure, and linearity, and discuss their optimization with respect to power consumption. At the circuit level, this led us to study the use of subthreshold biasing for low power LNA design and active mixer design. This study resulted in an LNA requiring only 0.6 mA of DC current from a 1.8 V supply [3] and a current reuse active mixer which also consumed 0.6 mA of DC current from a 1.8 V supply [4].

A careful review of state-of-the-art designs revealed that receivers are always designed to cater to the worst case received signal. This led us to explore the concept of energy-aware design whereby a receiver’s performance dynamically caters to the in-situ requirement [5]. We designed an IEEE 802.15.4 standard receiver front-end based on this concept which demonstrated 72% power savings under good channel conditions [6]. Our final receiver design based on our proposed design methods consumes just 2.2 mW in the nominal state while saving up to 68 % of this power under good channel conditions [7].

1.3 Organization of the Thesis

Chapter 2: In order to gain a proper appreciation of the rest of the thesis, we first introduce some fundamental concepts in RFIC. We also introduce system level receiver specifications.

Chapter 3: The second chapter will be a literature review of the state-of-the-art in short-range low data-rate receivers at both the system and circuit levels. We also discuss different system level techniques designed to lower the receiver’s energy consumption.
Chapter 4: Chapter three discusses system level optimization under the proposed energy-aware scheme, and the proposed architecture is discussed.

Chapter 5: Here we delve into circuit level analysis and optimization of individual receiver blocks such as the low-noise amplifier (LNA) and down-conversion mixer. We also look at the technology available to us.

Chapter 6: Finally we present measurement results of our proposed designs demonstrating potential power savings.

Chapter 7: In the final chapter, we discuss the conclusions of this study as well as possible direction for future research.
References


2. IEEE 802.15.4 Standard For Local and Metropolitan Area Networks, 2003.


2.1 Introduction

This chapter introduces fundamental concepts in RFIC design starting with system level concepts and ending with circuit level concepts. We will first introduce some simple receiver architectures and explain their basic operation. Next we discuss signal impairments which will set up the discussion on receiver specifications. Following our system level discussion, we discuss the maximum power transfer theorem, the Bode-Fano criterion, and power efficiency. Finally we will take a look at the available technology.

2.2 Fundamental System Level Concepts

The starting point of most signals is analog in nature (voice, video, temperature, humidity etc). In a digital communications system, the first step is to quantize the signal. Harry Nyquist told us that a signal must be sampled at more than twice its bandwidth in order to retain the information[1]. In the GSM standard, for example, a 3.1 kHz voice signal can be sampled at 8 kSamples/s with 13 b/Sample to result in a 104 kb/s data rate. Although the bandwidth of the digital signal is substantially wider than the initial audio signal, it is significantly more robust to impairments. In order to minimize the data rate, the signal is usually compressed and coded. The 104 kb/s GSM signal is usually compressed to 6.5 kb/s or 13 kb/s [2].

The data pulses are shaped in order to minimize the transmit bandwidth, and then modulated onto an RF carrier. In the case of the IEEE 802.15.4
standard, the carrier frequency is in the 2.4 GHz band. There are several reasons why we transmit at RF as opposed to low frequencies.

a) The size of passive components such as antennae and filters is related to the wavelength of the signal. The wavelength of a 3 MHz signal is 100 m while that of a 3 GHz signal is only 10 cm in free space.

b) There is more bandwidth at higher carrier frequencies. For example, a 1 MHz signal bandwidth is only 0.1 % of 1 GHz, but 10 % of 10 MHz.

c) High frequency carriers cannot travel as far as low frequency carriers permitting their frequencies to be reused in different geographic zones. The received signal strength is inversely proportional to the square of the product of carrier frequency and distance [3].

The modulated RF carrier is then amplified up to the desired level and transmitted via an antenna. On the receiver side, the RF carrier is first amplified, filtered, and then demodulated. After demodulation, the digital signal can be processed and converted back to its original analog form. The process of amplifying, filtering and removing the RF carrier is the main focus of this work.

2.2.1 Frequency Translation and the Image Problem

When signals of two different frequencies are multiplied, the output will include both the sum and difference of the input frequencies. First let us note that the Fourier transform of a cosine, $w_{LO}(t)$, is,

$$W_{LO}(f) = \frac{1}{2} \delta(f - f_{LO}) + \frac{1}{2} \delta(f + f_{LO})$$

(2.1)

where $f_{LO}$ (LO stands for local oscillator) is the frequency of the sinusoid. Supposing that $w_{RF}(t)$, with fourier transform $W_{RF}(f)$, is a bandpass signal centred at $f_{LO} + f_{IF}$, multiplication in the time domain with the LO is equivalent to a convolution in the frequency domain. This is shown mathematically below.
\[ W_{RF}(f) * W_{LO}(f) \]
\[ = \frac{1}{2} \int [W_{RF}(\lambda + f_{LO} + f_{IF}) + W_{RF}(\lambda - f_{LO} - f_{IF})] \delta(f - \lambda + f_{LO}) + \delta(f - \lambda - f_{LO}) \, df \]
\[ = \frac{1}{2} \int [W_{RF}(f + f_{IF}) + W_{RF}(f - f_{IF}) + W_{RF}(f - 2f_{LO} - f_{IF}) + W_{RF}(f + f_{LO} + f_{IF})] \]
\[ \Rightarrow \frac{1}{2} [W_{RF}(f + f_{IF}) + W_{RF}(f - f_{IF})] \] (2.2)

The arrow indicates low-pass filtering of the high frequency components. Now suppose that a bandpass image signal, \( W_{IM}(f) \), resides at \( f_{LO} \). Repeating the convolution process above, we find that the output signal is,

\[ W_{IM}(f) * W_{LO}(f) \]
\[ = \frac{1}{2} \int [W_{IM}(\lambda + f_{LO} - f_{IF}) + W_{IM}(\lambda - f_{LO} + f_{IF})] \delta(f - \lambda + f_{LO}) + \delta(f - \lambda - f_{LO}) \, df \] (2.3)

\[ \Rightarrow \frac{1}{2} [W_{IM}(f + f_{IF}) + W_{IM}(f - f_{IF})] \]

The down-converted image signal overlaps with the desired signal. This is illustrated in Fig. 2.1. This image signal acts as interference to the desired signal and degrades the quality of the reception. The image problem described above can generally be solved in two different ways, but before discussing them, we need to introduce system and channel bandwidth.

![Diagram of frequency translation and the image problem](image.png)

Fig. 2.1 Illustration of frequency translation and the image problem. The triangular signal is an undesired image while the other signal is desired.
2.2.2 System Bandwidth and Channel Bandwidth
The IEEE 802.15.4 standard has a system bandwidth spanning 83.5 MHz. This 83.5 MHz bandwidth is divided into 16 channels each with a bandwidth of 2 MHz, and spaced by 5 MHz. This is illustrated in Fig. 2.2.

![Fig. 2.2 Illustration of the IEEE 802.15.4 standard’s system and channel bandwidths.](image)

2.2.3 Image Rejection
The first method to reject an image signal is to simply filter it before down-conversion. Such an image-rejection filter has a fixed frequency response. One requirement of the down-conversion is that the signal must be down-converted to a center frequency higher than half of the system bandwidth, called the intermediate frequency (IF). If the IF were less than half of the system bandwidth, then for certain channels, the image signal would be inside the system bandwidth, and therefore, in practice it would not be possible to filter it away. This is illustrated in Fig. 2.3.

![Fig. 2.3 When the IF frequency (f_{IF}) is less than half of the system bandwidth, the image frequency (f_{IM}) falls inside the system bandwidth.](image)

The second method of image rejection is to use an image reject mixer. The idea works on the principle of the Hilbert transform which corresponds to a -90 degree phase shift network [4]. The phase transfer function is shown in Fig. 2.4.
\[ H(f) = \begin{cases} -j, & f > 0 \\ +j, & f < 0 \end{cases} \] (2.4)

\[ \text{Fig. 2.4 The phase transfer function of a -90 degree phase shifting network} \]

Let us represent the signal as a desired signal \( W_{RF}(f) \) and an image signal \( W_{IM}(f) \),

\[
W(f) = W_{RF}(f + f_{LO} + f_{IF}) + W_{RF}(f - f_{LO} - f_{IF}) \\
+ W_{IM}(f + f_{LO} - f_{IF}) + W_{IM}(f - f_{LO} + f_{IF})
\]

(2.5)

Applying a -90 degree phase shift,

\[
W(f) = jW_{RF}(f + f_{LO} + f_{IF}) - jW_{RF}(f - f_{LO} - f_{IF}) \\
+ jW_{IM}(f + f_{LO} - f_{IF}) - jW_{IM}(f - f_{LO} + f_{IF})
\]

(2.6)

Next, both (2.5) and (2.6) are mixed with the LO signal. With \( W_1(f) \) as the down-converted and low-pass filtered (LPF) version of (2.5), and \( W_2(f) \) as the down-converted and LPF version of (2.6),

\[
W_1(f) = \frac{1}{2} \left[ W_{RF}(f + f_{IF}) + W_{RF}(f - f_{IF}) + W_{IM}(f + f_{IF}) + W_{IM}(f - f_{IF}) \right]
\]

(2.7)

\[
W_2(f) = \frac{j}{2} \left[ W_{RF}(f + f_{IF}) - W_{RF}(f - f_{IF}) + W_{IM}(f - f_{IF}) - W_{IM}(f + f_{IF}) \right]
\]

(2.8)

Notice in (2.8), how the phases of the positive and negative frequencies of the image signal have switched compared to the RF signal. Mixing with both in-phase and 90 degrees out-of-phase components is termed quadrature mixing and can be done with quadrature phases of either the LO signal or the RF
signal. Careful inspection of Fig. 2.1 and Fig. 2.4 reveals why. If we now apply a -90 degree phase shift to (2.7), we obtain,

$$W_i(f) = \frac{j}{2} [W_{RF}(f + f_{IF}) - W_{RF}(f - f_{IF}) + W_{IM}(f + f_{IF}) - W_{IM}(f - f_{IF})]$$  \hspace{1cm} (2.9)$$

Finally take (2.8) and add it to (2.9),

$$W(f) = j[W_{RF}(f + f_{IF}) - W_{RF}(f - f_{IF})]$$  \hspace{1cm} (2.10)$$

The choice of how image-rejection is performed is a key consideration which affects the receiver architecture. Receiver architectures will be discussed next.

### 2.2.4 Receiver Architectures

There are three general categories of heterodyne-based receiver architectures [5]: zero-IF or direct-conversion receivers (DCR), low-IF receivers, and high-IF or superheterodyne receivers. They are defined by the image problem discussed above. A zero-IF receiver mixes the RF signal directly to baseband and therefore does not have an image problem. However, quadrature mixing is still required due to the possible phase mismatch between the transmitter and receiver LO. In a low-IF receiver, the IF is less than half of the system bandwidth and therefore requires image-reject mixing. In a high-IF receiver, the IF is more than half of the system bandwidth. Therefore, image-reject filtering can be used.

#### 2.2.4.1 High-IF Receivers

A typical front-end for a high-IF receiver is shown in Fig. 2.5. The band-select filter (BSF) removes unwanted interference outside of the system bandwidth thereby relaxing the intermodulation requirements of the LNA and mixer. The image-reject filter (IRF) may be simply a notch filter centered at the image frequency. As image-rejection in this case is reliant on attenuation rather than cancellation, very high values (more than 60 dB) of image-rejection can be achieved [6]. Unfortunately, the high-IF topology cannot be integrated on-chip while still yielding the full benefits of the topology. The quality factor (Q) requirements of the filters are so high as to be impossible to be integrated [6].
Furthermore, off-chip filters are designed with typical input impedances of several hundred ohms. This may not be optimum for the mixers or LNA and may require impedance matching or even buffering at lower frequencies, increasing the required power consumption. The high-IF receiver also requires two LO frequencies. LO\text{tune} varies with the desired channel while LO\text{fixed} is fixed. Again, high power is needed to generate two tones. On the plus side, only a fixed LO frequency requires quadrature generation which allows for excellent IQ matching.

**Fig. 2.5 A typical front-end for a high-IF receiver**

### 2.2.4.2 Low-IF Receivers

As previously discussed, low-IF receivers make use of image cancellation. The level of cancellation that can be practically achieved depends on process variation, unless some form of calibration is used. It is shown in [7], that the image-rejection-ratio (IRR) can be calculated as,

\[
IRR = \frac{(\frac{\Delta A}{A})^2 + \theta^2}{4}
\]

where \(A\) is the LO amplitude, \(\Delta A\) is the LO amplitude mismatch, and \(\theta\) is the phase mismatch of the LO. Image rejection without tuning is usually limited to less than 40 dB, however Behbahani et. Al [8] have reported a circuit which can repeatedly achieve nearly 60 dB \(IRR\). It is of note that the circuit required double-quadrature mixing and a total of 7 stages of polyphase filtering. Such a circuit would present significant noise and power consumption issues as to be impractical for low power design.
Compared to high-IF receivers, low-IF receivers do not require off-chip components. Their IF is usually selected high enough to avoid flicker noise problems while allowing for simple DC-offset cancellation. Image rejection requirements are based on tolerable image frequency levels and are usually lower when the image is closer to the desired frequency. Therefore, although low-IF receivers cannot achieve as high \( IRR \) as high-IF, they usually do not need to.

**Basic Architectures**

Generic implementations of low-IF receivers are shown below. Fig. 2.6a shows a Hartley receiver while 2.6b shows a Weaver receiver. An efficient implementation of the Hartley Receiver is shown in [6]. It is also possible to apply the 90 degree RF phase shift to the desired signal as was done in [9]. Both receivers operate on the principles discussed in 2.2.3. The important thing to remember is that when the 90 degree phase shift is applied at RF, both the image and desired signal undergo the same phase shift while after down-conversion, the second 90 degree phase shift causes the image and desired signals to undergo opposite phase shifts allowing for cancellation of the image.

Note that in the Weaver receiver, there are two mixing stages each with it’s own image. Therefore, the first LO, \( LO_1 \), should select the correct channel and the bandpass filters (BPF) must filter the image of the second mixing stage.
IF Planning

If the IF in a low-IF receiver is planned carefully, the required level of image rejection can be relaxed significantly. The idea is simply to choose an IF such that no channel falls in the image channel of the desired signal. Unfortunately, due to other considerations, that may not always be possible. Fig. 2.7 illustrates the effects of flicker noise and IRR on the signal for different choices of the IF in the IEEE 802.15.4 standard. The signal is received along with an interferer in the adjacent channel. Filtering is applied at the IF so that the further away (in frequency) the interferer is from the desired signal, the more it is filtered. From the figure, it would seem that 4 MHz is the best choice of IF. However, with a 4 MHz IF, the IF blocks must be faster than they would be with a 2 MHz IF. Other considerations for choice of IF include DC offset filtering, and the frequency resolution of the LO. More discussion on IRR and DC offset is found in section 2.2.5.
2.2.4.3 Zero-IF Receivers

In a zero-IF receiver, the signal is down-converted directly to baseband. The biggest benefit of this topology is that channel filtering can be implemented with simple LPFs as opposed to BPFs required in low-IF or high-IF receivers. Furthermore, a zero-IF receiver has no image problem. However, a zero-IF receiver still requires quadrature mixing in order to account for possible phase mismatch between the transmitting LO and the receiving LO. This is shown in (2.12) and (2.13).

Supposing LO₁ and LO₂ are out-of-phase by 90 degrees and $W_{BB}(f)$ represents the original signal to be transmitted,

$$[W_{BB}(f) * W_{LO1}(f)] * W_{LO2}(f) = W_{BB}(f) * [W_{LO1}(f) * W_{LO2}(f)]$$  \hspace{1cm} (2.12)$$

Looking at the second term in (2.12) and noting that $f_{LO} = f_{LO1} = f_{LO2}$,

$$W_{LO1}(f) * W_{LO2}(f)$$

$$= \frac{1}{4} \int_{-\infty}^{\infty} [-j\delta(\lambda - f_{LO}) + j\delta(\lambda + f_{LO})] \delta(f - \lambda - f_{LO}) + \delta(f - \lambda - f_{LO}) d\lambda$$

$$= \frac{j}{4} \left[ -\delta(f) - \delta(f - 2f_{LO}) + \delta(f + 2f_{LO}) + \delta(f) \right]$$

$$= \frac{j}{4} \left[ \delta(f + 2f_{LO}) - \delta(f - 2f_{LO}) \right]$$  \hspace{1cm} (2.13)
As we can see, there is no baseband component of (2.13). If LO\textsubscript{1} and LO\textsubscript{2} were in-phase, a baseband component would be recovered. Since the phase mismatch of between LO\textsubscript{1} and LO\textsubscript{2} is unknown, quadrature mixing is required to ensure a baseband component is recovered. Generally speaking, modulation schemes based on quadrature-amplitude modulation (QAM) involve separately modulating parallel bit streams onto an in-phase and quadrature-phase component of the LO.

Unfortunately, two major problems exist in zero-IF receivers. Firstly, the DC-offset is difficult to filter off. In analog receivers, DC-offset removal is accomplished by periodic calibration or by using high-pass filters (HPF) [10]. Secondly, low frequency flicker noise falls in the same band as the desired signal. For higher data-rate applications with wide signal bandwidth, flicker noise has less of an impact on the overall signal to noise ratio (SNR), however, DC-offset removal is still an issue.

2.2.4.4 Sub-Harmonic and Super-harmonic LO Receivers

A critical issue in zero-IF and even low-IF design is LO and RF self-mixing which basically means either the LO or RF signal is mixed with itself. This is discussed in more detail in section 2.2.5.7. LO and RF self mixing result in DC-offsets which can saturate baseband amplifiers, and in zero-IF receivers, can severely degrade SNR. The problem of LO self-mixing can be mitigated by generating the LO at a frequency far removed from the RF signal. This is illustrated in Fig. 2.8. When the LO signal frequency is sufficiently far removed from the RF signal, it is filtered by the resonant networks of the LNA, and by the input band-select filter.

![Fig. 2.8 Illustration of LO leakage suppression when \(f_{LO} \neq f_{RF}\).]
One way in which to generate the LO at a different frequency from the RF signal is to generate it at double the nominal LO frequency and then divide it by 2 [11]. The divide by 2 circuits can conveniently be designed to output quadrature LO signals, but due to the high frequency of operation of the VCO and dividers, this technique generally consumes high power. Another method is to generate a sub-harmonic of the LO and use a high IF [12], [13]. In [13] it is identified that the image frequency of the first down-conversion is baseband. Due to the high flicker noise at baseband, this method requires image rejection. Either of the methods discussed above suffice to reject the image, however, filtering as applied in [13] obviously leads to higher image rejection.

2.2.5 System Requirements

Impairments to the desired signal occur in three main forms: interference, noise, and distortion. Interference is any unwanted signals unintentionally picked up by the receiver. Noise finds its way into a receiver through blackbody radiation picked up by the antenna as well as random electron motion within the receiver circuitry. Linear distortion occurs due to multi-path propagation as well as limited channel bandwidth while nonlinear distortion is caused by a nonlinear transfer characteristic [14]. We do not intend to further discuss linear distortion, but suffice it to say that linear distortion causes inter-symbol interference (ISI) and affects the minimum bandwidth which must be provided by the channel filtering (to prevent ISI), and the minimum SNR required for successful demodulation (due to ISI caused by the multipath propagation).

An important difference between noise and interference is that noise (as defined in this work) is broadband while interference is narrowband. This means that if the receiver is experiencing too much interference, it may be able to select a different channel for communication, or even wait until the interfering signal is gone. This is not possible if the noise level is too high as all channels experience approximately the same level of noise at all times.

As receivers employ filters to remove any power outside the desired frequency band, we are normally primarily concerned with the overall
unwanted signal power in the desired band. We will see that calculating the unwanted signal power in the desired band is a recurring theme when considering system requirements.

2.2.5.1 Bit Error Rate

All of a system’s impairments sum up to the most important performance metric, the bit error rate (BER) which is the rate at which bits are incorrectly interpreted. In other words, the rate at which a ‘1’ is judged to be a ‘0’ or a ‘0’ is judged to be a ‘1’ by the receiver. This is best understood with an illustration. In Fig. 2.9, a zero is ideally represented by $V_0$ and a one is ideally represented by $V_1$. For the actual signal, the value will take on one of the two distributions shown depending on its ideal value. The variance of the distribution is equal to the mean squared noise voltage. If the signal is higher than $V_A$, it will be interpreted as a one and if it is lower than $V_A$, it will be interpreted as a zero. Therefore, based on the ratio of the signal power to the noise power (SNR), we can calculate the probability of a bit error which is essentially the BER. The BER depends, among other things, on the type of modulation used, and the method of demodulation and is discussed in [15].

![Fig 2.9. Illustration of BER](image)

For the RF front-end design, the BER requirement is translated to an SNR requirement, and this is split up into the basic impairments introduced at the beginning of this section.

**IEEE 802.15.4 Requirement: BER**

In the IEEE 802.15.4 standard, the digital bit sequence is first multiplied with a spreading code which is a pseudo-random bit sequence [16]. The new bit
sequence is then modulated via offset quadrature phase shift keying (O-QPSK) which can encode two bits per symbol. The original 250 kb/s after a 16 times spreading and 2 b/symbol modulation therefore becomes 2 Mchips/s which results in a 9 dB processing gain and a transmit bandwidth of 2 MHz [9]. O-QPSK is used with half-sine pulse shaping which is effectively minimum shift keying (MSK) and can therefore be implemented by direct VCO modulation. System simulations in [17] show that for a packet-error rate (PER) of 1% which corresponds to a BER of 0.00625%, the $SNR_{out}$ should be at least 0.5 dB, however, these simulations may have neglected to include delay spread due to the channel. In [18], system simulations including delay spread showed that the $SNR_{out}$ should be at least 14 dB (accounting for the processing gain of 9 dB). It should be noted also that the IEEE 802.15.4 standard itself [16] mentions that a detector should be able to achieve the required PER with an input SNR of only 5-6 dB, however, it is unclear whether multipath effects were considered. An actual system study to determine the required input SNR is beyond the scope of this work. Therefore, we will derive all requirements based on the more stringent SNR (14 dB). The reader is referred to [3] for more discussion on multipath fading, and path loss models.

2.2.5.2 Noise
All matter radiates noise corresponding to its surface temperature, termed blackbody radiation. Up to around 80 THz, this noise can be considered white [19], and is calculated as,

$$ P_n = kT\Delta f $$

(2.14)

where $k$ is Boltzmann’s constant, $T$ is the temperature in Kelvin, and $\Delta f$ is the bandwidth. The system itself adds additional noise to the system and if the noise sources are uncorrelated, their powers add. The noise factor is defined as the ratio of the total output noise power to the output noise power due to the source alone and for reference, is defined at 290 K. The earth’s average temperature is close to 290 K, but the real reason this value is used is because $kT$ is then $4.00 \times 10^{-21}$ J which made calculations easier [15]. The noise figure, $NF$, is simply the noise factor in decibels. Noting that the noise
factor effectively expresses the ratio of the input SNR to the output SNR, we can write,

\[
SNR_{\text{in}}(dB) = SNR_{\text{out}}(dB) + NF
\]  

(2.15)

\[
NF_{\text{req}} = \text{Sensitivity} - 10\log(kT\Delta f) - SNR_{\text{out}}(dB)
\]  

(2.16)

Here we have expressed the required system NF, \(NF_{\text{req}}\), in terms of the required output SNR, \(SNR_{\text{out}}\), and the sensitivity of the receiver, which is the minimum signal strength the system can receive with a pre-defined BER.

**IEEE 802.15.4 Requirement: Noise Figure**

The sensitivity requirement of an IEEE 802.15.4 standard compliant receiver is -85 dBm. Using the aforementioned 2-MHz bandwidth and \(SNR_{\text{out}}\) of 14 dB, the required \(NF\) is -85 - (-174) - 10\(\log(2M)\) - 14 = 12 dB where \(kT\) is -174 dBm/Hz at 290 degrees Kelvin.

### 2.2.5.3 Interference

Interference is simply any unwanted signal picked up by the receiver. It can affect the receiver in several ways and these effects are quantified by performance parameters such as the input-referred \(n^{\text{th}}\) order intercept, \(IIP_n\), the 1-dB gain compression point, \(P_{1\text{dB}}\), image-rejection-ratio, \(IRR\), and the phase noise, \(PN\). The most straightforward way in which interference can affect the desired signal is if it falls in the desired channel. The power of the interference must then be less than the power of the desired signal by the required SNR. It is of note that the IEEE 802.15.4 standard specifies that adjacent and alternate channel rejection requirements should be met with the desired signal 3 dB above the sensitivity level, which is why the desired signal level is taken as -82 dBm instead of -85 dBm in the following cases.

### 2.2.5.4 Intermodulation

When more than one tone passes through a nonlinear transfer function, the tones intermodulate, which is to say that they mix with each other. Consider the nonlinear transfer function,

\[
y(t) = a_0 + a_1x(t) + a_2x^2(t) + a_3x^3(t) + ...
\]  

(2.17)
Now suppose the input is represented by a two-tone signal,

\[ x(t) = A \cos(\omega_1 t) + B \cos(\omega_2 t) \]  

(2.18)

Substituting (2.18) into (2.17) yields components at all possible combinations of the fundamental frequencies: \( f_{\text{out}} = n f_{\text{in}} + m f_{\text{out}} \), where \( n \) and \( m \) can take on any integer value up to the highest order of the polynomial transfer function.

Taking (2.16) up to the 3rd order, we can write the output as,

\[ y(t) = a_0 + a_2 (A^2 + B^2) + a_2 (AB \cos(\omega_1 t - \omega_2 t)) \]  

(2.19a)

\[ + a_1 [A \cos(\omega_1 t) + B \cos(\omega_2 t)] \]  

(2.19b)

\[ + \frac{3a_3}{2} \left[ \left( \frac{1}{2} A^3 + AB^2 \right) \cos(\omega_1 t) + \left( \frac{1}{2} B^3 + BA^2 \right) \cos(\omega_2 t) \right] \]  

(2.19c)

\[ + \frac{3a_3}{4} \left[ AB^2 \cos(\omega_1 t - 2\omega_2 t) + BA^2 \cos(\omega_2 t - 2\omega_1 t) \right] \]  

(2.19d)

\[ + \frac{a_2}{2} \left[ A^2 \cos(2\omega_1 t) + B^2 \cos(2\omega_2 t) + 2AB \cos(\omega_1 t + \omega_2 t) \right] \]  

(2.19e)

\[ + \frac{3a_3}{4} \left[ AB^2 \cos(\omega_1 t + 2\omega_2 t) + BA^2 \cos(\omega_2 t + 2\omega_1 t) \right] \]  

(2.19f)

\[ + \frac{a_3}{4} \left[ A^3 \cos(3\omega_1 t) + B^3 \cos(3\omega_2 t) \right] \]  

(2.19g)

Distortion components represented by (2.19e)-(2.19g) generally fall out of the desired band and can be readily filtered. Therefore they are not of much interest. Equation (2.19a) represents a static DC offset and potentially a slow time-varying DC offset, and can be used to calculate \( \text{IIP}_2 \). These are important in direct-conversion receiver design. (2.19c) is a nonlinear fundamental component and is not important for constant-envelope modulation schemes. (2.19d), however, is used in calculating the \( \text{IIP}_3 \). The \( \text{IIP}_3 \) is defined as the input power whereby the output in-band third-order intermodulation product, \( \text{IM}_3 \), is equal to the linear output. Since the receiver is never driven to a point where the \( \text{IM}_3 \) is higher than the desired output, the actual \( \text{IIP}_3 \) as a power
level is not of much significance. However, the \( IIP_3 \) allows us to calculate the input power level at which the IM\(_3\) is higher than the noise level thereby making the SNR insufficient. \( IIP_3 \) is measured with two equal amplitude tones and therefore is found when \( a_1 A = 3a_3 A^2 / 4 \), or \( A = \sqrt{(4a_1/3a_3)} \).

First let \( x_1(t) \) equal to \( A\cos(\omega_1 t) \) and \( x_2(t) \) equal to \( B\cos(\omega_2 t) + C\cos(\omega_3 t) \). The tones in \( x_2(t) \) will intermodulate to yield components at \( 2\omega_2 - \omega_3 \) and \( 2\omega_3 - \omega_2 \) which can be in the desired band. Taking the \( 2\omega_2 - \omega_3 \) term, we need,

\[
\frac{a_1 A}{(\text{SNR}_{\text{out}})} \geq \frac{3a_3 B^2 C/4}{\sqrt{(4a_1/3a_3)}} \geq \sqrt{(B^2 C \cdot \text{SNR}_{\text{out}} / A)}
\]

where all terms are in units of volts or amperes. We can express this in decibels with the correct reference impedance as,

\[
IIP_{3,\text{req}} (\text{dBm}) \geq \frac{2B + C - A + \text{SNR}_{\text{out}}}{2}
\]  \( (2.20) \)

where \( IIP_{3,\text{req}} \) is the required system \( IIP_3 \), and \( B \), \( C \), and \( A \) are now in dBm.

**IEEE 802.15.4 Requirement: \( IIP_3 \)**

The IEEE 802.15.4 standard defines a spectrum mask which tells the designer the level of interference and its frequency spectrum which an IEEE 802.15.4 compliant receiver must be able to tolerate. As previously mentioned, channels are spaced by 5 MHz while \( \text{SNR}_{\text{out}} \) is equal to 14 dB. The receiver should be able to tolerate 0 dBc in the adjacent channel and 30 dBc in all alternate channels relative to the desired channel when the receiver is operating at a signal level 3 dB higher than the sensitivity. This is illustrated in Fig. 2.10. The strictest \( IIP_3 \) requirement occurs when considering the intermodulation of two signals of 10 MHz and 20 MHz offset from the desired signal (each at -52 dBm) resulting in \( IIP_{3,\text{req}} \geq -30 \) dBm. For two signals of 5 MHz and 10 MHz offset from the desired signal (illustrated below), \( IIP_{3,\text{req}} \geq -60 \) dBm.
2.2.5.5 Image Rejection Ratio

As previously discussed, image rejection is required in both high-IF and low-IF receivers. The level of image rejection required is that which forces the image frequency to be less than the desired SNR. The power at the image frequency which must be tolerable by the receiver depends on the spectrum mask.

IEEE 802.15.4 Requirement: IRR

For the IEEE 802.15.4 standard, apart from the adjacent and alternate channel interference already mentioned, the receiver must also be able to tolerate a -30 dBm blocker outside of the system bandwidth [10]. Therefore, for a high-IF receiver the IRR must be more than -30 - -85 + 14 = 69 dB. For a low-IF receiver with an IF of 5 MHz or higher, the IRR must exceed 30 + 14 = 44 dB, while for an IF of less than 5 MHz, the IRR should exceed 14 dB. Therefore, for a low-IF receiver, there is considerable advantage to choosing an IF less than 5 MHz. In [20], the argument was made for a 6 MHz IF, but the SNR_{out} used was only 0 dB resulting in the IRR barely meeting requirements.

2.2.5.6 Phase Noise

Phase noise is a requirement of the frequency synthesizer and we will not spend much time on it. Phase noise is usually measured in the frequency domain as a power per unit bandwidth away from the desired LO tone [21]. For example, if we measure the power of the LO at 5 MHz offset from the carrier as -80 dBm with a resolution bandwidth of 1 kHz, while the carrier has
a power of 3 dBm, then the phase noise is -113 dBc/Hz at 5 MHz offset. The problem arises when this LO power at 5 MHz offset translates the unwanted adjacent channel to the IF. This is illustrated in Fig. 2.11. We can see from the graph on the right, that after down-conversion, the SNR of the desired signal at $f_{IF}$ is significantly degraded.

![Fig. 2.11. Illustration of the phase noise problem.](image)

**IEEE 802.15.4 Requirement: Phase Noise**

Suppose we wish to receive a -82 dBm signal at 2.402 GHz and the interferer in the alternate channel (2.412 GHz) is -52 dBm. Suppose also that the IF is 2 MHz. Assuming low-side injection, the LO power at 2.4 GHz would mix with the 2.402 GHz signal producing the desired 2-MHz IF signal while the LO power at 2.410 GHz would mix with the -52 dBm signal, producing unwanted IF noise. The total unwanted LO power is approximated as $P_{LO} + PN + 10\log(BW)$ at 2.410 GHz. Taking $P_{int}$ as the interferer power, with $SNR_{out}$ equal to 14 dB and a 2 MHz signal bandwidth, the required phase noise is approximately,

$$PN = Sensitivity - P_{int} - SNR_{out} - 10\log(BW)$$  \hspace{1cm} (2.21)

which equals to -107 dBc/Hz at 10 MHz offset, or -77 dBc/Hz at 5 MHz offset for the IEEE 802.15.4 standard.

**2.2.5.7 $IIP_2$ and Self-Mixing**

Another concern mainly in direct-conversion receivers and low-IF receivers is self-mixing. The problem arises because the LO frequency is the same as the RF frequency. We can split self-mixing up into two categories: LO self-mixing...
and RF self-mixing. When the LO leaks to the mixer input ports, it is down-converted to form a static DC-offset. This can potentially saturate the baseband amplifiers. In some cases, the LO-RF input port isolation is insufficient, and the LO leaks to the antenna and is transmitted. If the LO signal then reflects of a nearby object and is received, it can be down-converted to form a slow time-varying DC-offset [12]. For RF self-mixing, we are mainly concerned that a high power interfering signal will leak to the LO ports thereby causing unwanted down-conversion to baseband. This is shown conceptually in Fig. 2.12.

![Fig. 2.12 Illustration of RF self-mixing.](image)

Note from the figure how the baseband bandwidth of the self-mixed signal is double its original bandwidth. This is a result of convolving the signal with itself. The IIP\textsubscript{2} is defined as the input power which results in the second order intermodulation product power (IM\textsubscript{2}) to be equal to the desired output power. The analysis for the required IIP\textsubscript{2} is similar to that for the required IIP\textsubscript{3}. Firstly, assume an input signal with power $A$. The desired down-converted signal is proportional to the LO power, $P_{LO}$, and some factor $G_{\text{Conv}}$ resulting in an output power of $P_{LO}G_{\text{Conv}}A$. The undesired product is proportional to the RF-LO leakage, $G_{\text{leak}}$, the and the same factor, $G_{\text{conv}}$, with a resulting output power of $G_{\text{leak}}G_{\text{Conv}}A^2$. Setting these two output powers to be equal, we find the IIP\textsubscript{2} of the down-converter to equal to $P_{LO}/G_{\text{leak}}$. To calculate the required IIP\textsubscript{2}, we require that the unwanted output power is less than the wanted output power by a factor equal to the desired SNR. Taking the interferer
power as $B$, we require that $G_{\text{leak}}G_{\text{Conv}}B^2 < P_{\text{LO}}G_{\text{Conv}}A/\text{SNR}_{\text{out}}$ or $\text{IIP}_2 > \text{SNR}_{\text{out}}B^2/A$. Expressed in decibels, the required $\text{IIP}_2$ is,

$$\text{IIP}_{2,\text{req}} (\text{dBm}) \geq 2B - A + \text{SNR}_{\text{out}}$$

(2.22)

where $A$ and $B$ are now in decibels. The situation is complicated by the fact that all interferer’s must be accounted for since they can all self-mix down to baseband.

IEEE 802.15.4 Requirement: $\text{IIP}_2$

For the IEEE 802.15.4 standard, the maximum alternate-channel interference is taken to be -52 dBm. Considering only one alternate-channel interferer, the $\text{IIP}_2$ requirement is equal to $2^*(-52) - (-82) + 14 = -8$ dBm, where once again the desired signal is -82 dBm and the required SNR is 14 dB.

2.3 Fundamental Circuit Level Concepts

In this section, we discuss fundamental circuit level concepts such as power, gain, and noise. In low frequency analogue circuits it is often sufficient to represent a source as either an ideal voltage source (a circuit driven by an op-amp for example) or an ideal current source (a circuit driven by an operational transconductance amplifier (OTA) for example). However, in RFIC design, this is rarely the case. Care must be taken when talking about concepts such as input power and power gain.

2.3.1 Power

At the antenna, the desired signal is received along with unwanted interference. If the signal is impaired by an IEEE 802.11a standard signal, then we would expect interference in the 5.15 GHz to 5.35 GHz band [22]. Suppose we must be able to tolerate a -30 dBm 802.11a interferer. The antenna is usually connected to a band-select filter to remove unwanted out-of-band interference such as 802.11a signals (See Fig. 2.5 and Fig. 2.6). As a simple case, let us represent the LNA input as a 50 Ω resistor and assume that the antenna is represented by a voltage source in series with a 50 Ω
resistor. The filter is modelled as a lossless passive network. This is shown in Fig. 2.13.

![Fig. 2.13 Model of an antenna connected to a BPF and a load.](image)

As the BPF is a lossless network, we know that any power injected into the circuit at node \( \text{in} \) will be dissipated in \( R_{\text{load}} \) connected to node \( \text{out} \). If we inject our -30 dBm 802.11a interferer into \( \text{in} \), then we can expect that the same amount of power will reach \( \text{out} \). This may seem strange since the purpose of the BPF is to filter out unwanted interference, but our 802.11a signal has passed right through. The problem is easily understood when we define the -30 dBm 802.11a signal not as the input power, \( P_{\text{in}} \), but as the available power, \( P_A \), which is the maximum power available from the antenna and is achieved under matched conditions [23]. Since the BPF is not matched outside of the pass-band, \( P_{\text{in}} \neq P_A \) in the stop-band, i.e. the receiver is able to tolerate a blocker available power level of -30 dBm but not an input power level of -30 dBm. This leads to different definitions of the term gain.

2.3.2 Power Gain Definitions

The power gain or operating power gain of a circuit is defined as the ratio of the power delivered to the load to the power supplied to the circuit, \( P_{\text{in}} \) [23], which as we have discussed may or may not be a useful specification. Another commonly used power gain is the transducer gain, \( G_T \), which is the ratio of the power delivered to the load to the power available from the source. From Fig. 2.13, the available power from the source is equal to \(|E|^2/4R_{\text{source}}\) [24] (here \( E \) is the open circuit rms voltage of the generator). Note that \( P_A \) is frequency independent (as opposed to \( P_{\text{in}} \)) which allows \( G_T \) to include filtering effects. Unfortunately, for two cascaded circuits, we cannot simply multiply their transducer gains to obtain the overall transducer gain. Finally, another important power gain measure is the maximum unilateral transducer power
gain, \( G_{TU_{\text{max}}} \). \( G_{TU_{\text{max}}} \) is essentially the maximum transducer power gain of a unilateral device and can be used as a comparison basis for transistors.

### 2.3.3 Noise Figure

Repeating section 2.2.5.2, the \( NF \) of a system is defined as the ratio of the total output noise power (not including the load noise) to the output noise power resulting from the source only, and is taken at 290 K. However, we will see that the actual power of the noise is unimportant, and the voltage level of the noise alone can be used to define \( NF \). The \( NF \) is calculated as,

\[
NF = \frac{P_{n,\text{load}}}{P_{n,\text{load-source}}} \quad (2.23)
\]

where \( P_{n,\text{load}} \) is the total output noise power, and \( P_{n,\text{load-source}} \) is the output noise power delivered to the load from the source only. These powers can be equivalently represented as a voltage noise across the load resistance, where \( P_{n,\text{load}} \) equals to \( |V_{n,\text{load}}|^2/R_{\text{load}} \) and \( P_{n,\text{load-source}} \) equals to \( |V_{n,\text{load-source}}|^2/R_{\text{load}} \) (\( R_{\text{load}} \) is the load resistance).

\[
NF = \frac{|V_{n,\text{load}}|^2}{|V_{n,\text{load-source}}|^2} = \frac{|V_{n,\text{load}}|^2}{|V_{n,\text{load-source}}|^2} \quad (2.24)
\]

Clearly, \( NF \) is not restricted to measuring power ratios, but can be taken as voltage ratios or even current ratios.

### 2.3.4 Voltage Gain and Matching

As IC designers, we are often more concerned with voltage gain than power gain. This is exemplified by the situation described above (section 2.3.3). Much like power gain, voltage gain can be defined in different ways. In order to include filtering effects, we can define voltage gain as \( 2V_{\text{out}}/E \) otherwise it is commonly defined as \( V_{\text{out}}/V_{\text{in}} \). Going back to Fig. 2.13, we note that if the BPF includes impedance matching, then there will be a voltage gain (since we have assumed the input is matched, either voltage gain definition suffices) equal to \( \sqrt{(R_{\text{load}}/R_{\text{source}})} \). This is a direct consequence of the conservation of
energy. From our discussion on $NF$ above, it would seem that we should make $R_{\text{load}}$ as large as possible in order to maximize the voltage gain without affecting the $NF$. Forgetting about the effect on the linearity of the system, and whether or not a matching network could be designed for this situation, there is another important problem. The Bode-Fano Criterion tells us that there is a trade-off between the achievable matching bandwidth, and the voltage gain achievable through impedance matching [25]. For the readers interest, we repeat the pertinent formula which assumes a load consisting of a parallel resistor, $R$, and capacitor, $C$,

$$\int_0^\infty \ln \frac{1}{\Gamma(\omega)} d\omega \leq \frac{\pi}{RC}$$

(2.25)

where $\omega$ is the frequency in radians/second. This equation tells us that the better the matching, the smaller the bandwidth over which this bandwidth can be achieved. As such, if the designer is only aiming for -10 dB $S_{11}$, better matching should be avoided since it reduces the bandwidth over which -10 dB $S_{11}$ can be achieved. The trade-off between bandwidth and voltage gain is exploited throughout literature as we will see in our literature survey. While this can lead to better performance with a few measured chips, it can be a problem if the design is to be mass-produced, as it can affect the overall yield.

2.3.5 Maximum Power Transfer and Power Efficiency

We have mentioned before that under matched conditions the power delivered to a load is equal to the available power from the source (section 2.3.1). It is easy to show that the power delivered to a load is equal to,

$$\left| \frac{V_{\text{load}}}{R_{\text{load}}} \right|^2 = \frac{R_{\text{load}}|E|^2}{(R_{\text{load}} + R_{\text{source}})^2 + (X_{\text{load}} + X_{\text{source}})^2}$$

(2.26)

where $V_{\text{load}}$ is the voltage across the load, $X_{\text{load}}$ is the reactive part of the load impedance and $X_{\text{source}}$ is the reactive part of the source impedance. Setting $X_{\text{load}} = -X_{\text{source}}$ and $R_{\text{load}} = R_{\text{source}}$ yields the maximum output power. It is important to note that while (2.26) tells us the maximum power delivered to a load, maximum power transfer does not imply maximum efficiency. In fact
since half of the power is dissipated in the source, the maximum efficiency achievable under matched conditions is 50%. Neglecting the reactive parts of the impedances, we can see that the total power dissipated is equal to $|E|^2/(R_{\text{load}} + R_{\text{source}})$, and the efficiency is equal to $R_{\text{load}}/(R_{\text{load}} + R_{\text{source}})$. Keeping $R_{\text{source}}$ fixed, we can see that the maximum efficiency is achieved by maximizing $R_{\text{load}}$, however this condition results in zero power transferred to the load. Here we have used a Thévenin equivalent circuit for the source. If we had used a Norton equivalent representation, we would have found that the maximum power efficiency results from minimizing $R_{\text{load}}$.

A real amplifier dissipates DC power as well as power at unwanted harmonics of the signal (caused by the nonlinearity of the amplifier), neither of which are accounted for in the preceding analysis. Power efficiency is important in the design of amplifiers which must deliver significant power to the load, and the topic has its own extensive literature [26]. For small signal amplifiers such as the low-noise amplifier of a receiver, maximum power transfer is more important.

### 2.3.6 Cascaded IIP$_3$

The IIP$_3$ up to a certain point in a system depends on both the bias point of a circuit as well as the gain of the previous stages. A well known equation to calculate cascaded IIP$_3$ is shown below [27]-[29].

$$\frac{1}{IIV_{3,\text{total}}} = \frac{1}{IIV_{3,1}^2} + \frac{G_1^2}{IIV_{3,2}^2} + \frac{G_2^2G_3^2}{IIV_{3,3}^2} + ...$$  \hspace{1cm} (2.27)

Here the IIP$_3$ is represented as a voltage and is appropriately symbolized as $IIV_3$. $IIV_{3,n}$ represents the $IIV_3$ of the $n^{th}$ stage while $G_n$ represents the voltage gain from input to output of the $n^{th}$ stage. Unfortunately, this equation does not tell the full story. Referring back to (2.17), at low frequencies, we can see that $a_3$ can be either positive or negative. (2.17) is repeated below.

$$y(t) = a_0 + a_1x(t) + a_2x^2(t) + a_3x^3(t) + ...$$  \hspace{1cm} (2.28)

The expanded terms in (2.19) likewise depend on the phase of $a_n$. Hence (2.27) is actually a pessimistic result [27]. At RF, any phase is possible and
the addition of IM\textsubscript{3} components should be done using vector addition. For instance, assume the voltage output third-order intermodulation distortion product of each stage is designated the variable $OIM_{3,n} = \text{Re}\{OIM_{3,n}\} + j\text{Im}\{OIM_{3,n}\}$ where $n$ represents the stage number. The output of two cascaded stages results in,

\[
OIM_{3,\text{total}} = \text{Re}\{OIM_{3,1}G_1\} + \text{Re}\{OIM_{3,2}\} + j[\text{Im}\{OIM_{3,1}G_1\} + \text{Im}\{OIM_{3,2}\}]
\]

(2.29)

Note that each component of $OIM_{3,n}$ can be either positive or negative. Furthermore, (2.29) can easily be iteratively extended to $n$ stages.

### 2.3.7 Dynamic Range

It is worthwhile to mention the issue of dynamic range in the IEEE 802.15.4 standard. Dynamic range is a measure of the ratio of the largest signal to the smallest signal which can be received with sufficient BER. The smallest signal is specified by the sensitivity level of the receiver while the maximum signal strength is normally limited by the linearity of the receiver. The modulation scheme used in the IEEE 802.15.4 standard is a constant envelope type scheme and hence, the desired signal itself is not susceptible to nonlinear distortion or AM-PM conversion [31]. Most IEEE 802.15.4 standard receivers therefore use limiting amplifiers in the IF section in order to maintain a constant output amplitude signal even when the receiver signal strength is changing rapidly [9], [11]. The reader should bear in mind that for the IEEE 802.15.4 standard, the maximum interference level (and not the desired signal level) that a receiver should tolerate sets the output IM\textsubscript{3} levels. Therefore, the IM\textsubscript{3} level is independent of the received signal strength (of the desired signal) and does not affect the maximum signal strength which can be received.

### 2.3.8 Summary of Trade-Offs

In RF receiver design, numerous performance parameters trade with one another in various ways. Table 2.1 summarizes some of the more important trade-offs in RFIC design. In matching networks, gain and bandwidth trade off via the Bode-Fano criterion while in op-amps, they trade off via negative feedback. Both gain and bandwidth are related to power consumption since
increasing device sizes reduces output resistance thereby increasing driving capability. Gain trades with NF based on the cascaded NF equation [28],

\[ NF_j = 10 \log(F_j) \]  

\[ F_{total} = F_1 + \frac{F_2 - 1}{GP_1} + \ldots \]  

where \( F_n \) is the noise factor of the \( n^{th} \) stage, \( NF_n \) is the noise figure of the \( n^{th} \) stage, \( F_{total} \) is the overall noise factor and \( GP_n \) is the power gain of the \( n^{th} \) stage. We have already discussed how linearity and gain trade with each other through (2.27). Therefore, linearity also trades indirectly with bandwidth through gain. Attempting to increase gain by reducing bandwidth inevitably reduces a circuit's linearity. Noise performance also trades directly with power consumption through a device's input-referred noise [30]. Linearity also trades directly with power consumption through an op-amp's unity-gain bandwidth (UGB). Lastly, noise trades with linearity through (2.30) and (2.27). Increasing gain reduces linearity while improving noise performance.

Apart from the trade-offs mentioned above, one of the most practical trade-offs not yet mentioned is cost. In RFIC design, cost manifests itself in the design process in the choice of technology, the size of the die (influenced by the number of inductors and capacitors), the number of off-chip components etc. Through such numbers, it can also be linked indirectly to the parameters in Table 2.1.
<table>
<thead>
<tr>
<th>Bandwidth</th>
<th>Power</th>
<th>Noise</th>
<th>Linearity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gain</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Bode-Fano</td>
<td>$g_m \propto \sqrt{I_{DS}}$</td>
<td>$F_{\text{total}} = F_1 + \frac{F_2 - 1}{GP_1}$ (2.27)</td>
<td></td>
</tr>
<tr>
<td>2. Negative Feedback</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bandwidth</td>
<td>$x$</td>
<td>$R_{ds} \propto \frac{1}{I_{DS}}$</td>
<td>Via Gain</td>
</tr>
<tr>
<td>Power</td>
<td>$x$</td>
<td>$x$</td>
<td>$\text{IRN} = \frac{4kT\gamma}{ag_m^2}$</td>
</tr>
<tr>
<td>Noise</td>
<td>$x$</td>
<td>$x$</td>
<td>$x$</td>
</tr>
</tbody>
</table>

$^1$ IRN is the input-referred noise
References


3.1 Overview of the State-of-the-Art

In this section, we present the state-of-the-art in low-power receiver design. Most of the works presented are targeted for the IEEE 802.15.4 standard which as previously discussed is a recently ratified standard designed specifically for low-power, low-data-rate applications. Table 3.1 presents the performance of several recently published works.

3.1.1 Discussion of Results

The lowest power consumptions were achieved by [2] and [3]. [1]-[4] achieved similar NF, while [3], [5], and [6] achieved commendable $\text{IIP}_3$. With the exception of [2] and [3], 0.18 $\mu$m CMOS was the technology of choice, and the most widely adopted system architecture consisted of an LNA, IQ mixers, a complex channel-select filter, limiting amplifiers, and a demodulator.
### Table 3.1
**Comparison of Prior Published Work**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency (GHz)</td>
<td>2.4</td>
<td>2.5</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>Estimated BW (GHz)</td>
<td>0.5</td>
<td>0.35</td>
<td>0.35</td>
<td>3.0</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Current (mA)$^1$</td>
<td>$2.4^D$</td>
<td>$0.6^D$</td>
<td>$1.875^D$</td>
<td>$4^D$</td>
<td>$3.5^S$</td>
<td>$3^S$</td>
<td>$5^D$</td>
</tr>
<tr>
<td>Voltage Supply (V)</td>
<td>1.8</td>
<td>1.2</td>
<td>0.4</td>
<td>1.35</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
</tr>
<tr>
<td>Noise Figure (dB)</td>
<td>5.7</td>
<td>5</td>
<td>5.1</td>
<td>6</td>
<td>7.3</td>
<td>-</td>
<td>10</td>
</tr>
<tr>
<td>IIP$_3$ (dBm)</td>
<td>-16</td>
<td>-37</td>
<td>-7.5</td>
<td>-12</td>
<td>-8</td>
<td>-4</td>
<td>-</td>
</tr>
<tr>
<td>Voltage Gain (dB)</td>
<td>33</td>
<td>43</td>
<td>17</td>
<td>37</td>
<td>30</td>
<td>30</td>
<td>-</td>
</tr>
<tr>
<td>Technology ($\mu$m)</td>
<td>0.18</td>
<td>0.18</td>
<td>0.13</td>
<td>0.09</td>
<td>0.18</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>Blocks Designed$^2$</td>
<td>LMCAD</td>
<td>LMV</td>
<td>LMCA</td>
<td>LMC</td>
<td>LM</td>
<td>LMCAD</td>
<td>LMCAD</td>
</tr>
<tr>
<td>IEEE 802.15.4</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Mixer Type$^3$</td>
<td>PC</td>
<td>A</td>
<td>PV</td>
<td>PV</td>
<td>PC</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Sensitivity (dBm)$^4$</td>
<td>-91.3</td>
<td>-68</td>
<td>-91.9</td>
<td>-91</td>
<td>-89.7</td>
<td>-85</td>
<td>-87</td>
</tr>
</tbody>
</table>

$^1$ S: Single-Ended, D: Differential. Only the front-end is considered (LNA, mixer, filtering)
$^4$ Based on the IEEE 802.15.4 standard according to the poorer of either IIP$_3$ or NF.

#### 3.1.1.1 Op-Amp Linearity and Current Reuse [1]

This work implements a full IEEE 802.15.4 standard receiver. The current consumption is only 5.6 mA, with 2.4 mA going to the LNA, and a supply voltage of 1.8 V. The voltage gain of the LNA is high (33 dB) which relaxes the NF requirements of the following stages. This also allows the design to achieve excellent NF (5.7 dB). The penalty for the high gain is in the system IIP$_3$, but at -16 dBm, it is more than sufficient for the application (-30 dBm, see
Section 2.2.5.4). The design used op-amp based complex channel-select filtering which can provide high linearity resulting from the use of negative feedback [8], [9]. In particular, in [8] and [9] it is shown that,

\[
V_{\text{IP 3,after}} = V_{\text{IP 3,before}} (1 + LG)^{\frac{3}{2}}
\]

\[
V_{\text{IP 2,after}} = V_{\text{IP 2,before}} (1 + LG)^{3}
\]

where \(V_{\text{IP n,after}}\) is the IIP\(_n\) in volts after feedback is applied, \(V_{\text{IP n,before}}\) is the IIP\(_n\) in volts before feedback is applied, and \(LG\) is the small signal loop-gain [10].

Since the NF requirements of the IF stages is relaxed by the high LNA gain, and the linearity of the IF stages is high through use of feedback, the bottleneck in both the NF and the linearity is in the front-end LNA and mixer. In order to optimize the performance of the LNA, current-reuse was used. The basic technique of current-reuse is to stack one block on top of another, and separate the blocks by an AC ground. The justification is that the MOSFET’s small signal performance is relatively independent of its voltage headroom as long as the device operates in the saturation region [11]. An illustration of a simple method of current reuse is shown in Fig. 3.1. Large inductors and capacitors are used to AC-separate the two amplifying stages.

![Fig. 3.1 Illustration of current-reuse.](image)

As discussed in Chapter 2, section 2.2.5.5, the IRR requirement is very much dependent on the IF selection. In [1], the IF is 2 MHz which results in a very
relaxed IRR requirement (14 dB). However, given the signal bandwidth of 2 MHz, the flicker noise corner frequency should be less than 1 MHz for good overall noise performance. In [1], passive mixing was used which results in minimal flicker noise [12]. Specifically, current-output passive mixing was used, where the stage following the passive mixer was a transimpedance amplifier (TIA). The op-amp based channel-select filters (CSF) doubled as TIAs and can provide high linearity to the IF section. Passive mixers are discussed in more detail in chapter 4.

3.1.1.2 Linearity versus Noise [2]

This design operates with the lowest current consumption (0.6 mA for the LNA and 1.17 mA total). Furthermore, the design requires a voltage supply of only 1.2 V. We can immediately see that lowering the supply voltage is similar to current reuse (discussed above) except that all blocks will benefit from the reduced power consumption. The obvious drawback is that all blocks must be able to operate under the low supply voltage, including blocks which would normally benefit from high voltage supplies such as telescopic op-amps. This design achieves excellent NF for the power consumption, but there were several tradeoffs made. The design consists of an LNA, a mixer, and a variable-gain amplifier (VGA).

The paper quoted an output-referred IP$_3$ (OIP$_3$) of +6 dBm, and IIP$_3$ was estimated here by simply subtracting the gain to get -37 dBm. This level of IIP$_3$ may not be sufficient for the IEEE 802.15.4 standard. The poor IIP$_3$ is directly connected to the good NF since, effectively, three cascaded gain stages have been used without any channel filtering (the LNA, an active mixer, and the VGA).

The active mixer achieves better NF and gain than a passive mixer at the expense of a higher flicker noise corner frequency. In this work, the flicker noise corner frequency was not measured due to limitations in the measurement equipment, and the NF was measured at an IF above 10 MHz.

This work also used a very high $Q$ matching network to achieve a matching network gain of 14.8 dB. As discussed in Section 2.3.4, matching network voltage gain is a direct consequence of the conservation of energy, and
trades with matching bandwidth. The high matching network voltage gain helped in suppressing the later stage’s noise contribution.

In order to improve the transconductance of the individual devices, subthreshold biasing was used. Subthreshold biasing results in a reduction in device transit frequency, $f_T$ [11]. However, in practical low-power situations, subthreshold biasing will generally result in better performance than strongly inversion biasing. This will be discussed in further detail in Chapter 4.

### 3.1.1.3 Low Voltage Design [3]

This work achieves the lowest power consumption both for the front end as well as the entire design. In fact, the front end consists only of a step-up impedance transforming network, and a passive mixer. Therefore, the front end consumes no power. Despite the lack of a true LNA, this design achieves excellent NF (5.1 dB). Notice that we did not call the input LC network a matching network. In [3] and [13], it becomes apparent that the designers did not match the input of the RX to the source. This allowed for an NF of 1.1 dB in the input LC network and a voltage gain of 16.2 dB up to the quadrature mixer output. The gain of the input LC network alone was probably between 17 dB and 18 dB.

Note that as only a single RF resonant network was used, the gain bandwidth of this design could be large. As the source was not matched to the RX, it is irrelevant to talk about the matching bandwidth.

When comparing this design with others, it is important to realize that matching bandwidth is a much more stringent requirement than gain bandwidth (-10 dB versus -3 dB). If the designers had decided to include input matching, this design could still have achieved sufficient gain bandwidth for the IEEE 802.15.4 standard, but not necessarily sufficient matching bandwidth. This is because the designers used a very high $Q$ input LC network in order to boost the voltage gain.

With the reduced bandwidth comes a more distressing problem: susceptibility to process variation. This is because the bandwidth of the LNA is usually designed to be large enough not only to pass the entire system bandwidth, but to account for variation in both the bandwidth and the center frequencies.
of the resonant circuits. This is illustrated in Fig. 3.2 where $BW_{\text{sys}}$ is the required system bandwidth. The overlapping areas of the three shapes in each figure determine the achieved RF bandwidth. For the 1 GHz RF bandwidth, the effect of process variation is marginal while for a 500 MHz RF bandwidth, the effect is catastrophic. Note that the overall RF bandwidth is also affected by the number of resonant networks in the signal path. In [3], fortunately there is only one such network. The reader should note that it comes as no surprise that the two designs with the lowest power consumption [2], [3] also featured the smallest RF bandwidth.

![Fig. 3.2 Illustration of effect of process variation and bandwidth on system gain for (a) 1 GHz RF bandwidth and (b) 0.5 GHz RF bandwidth. The center frequency of the RF bandwidth is shifted by approximately ±10 % from the nominal value.](image)

The design featured voltage-output passive mixers. This supported the use of $G_m$-C type channel select filtering. High linearity was possible since the pre-CSF gain was only 17 dB. Apart from the bandwidth trade-off, the lack of a true LNA also results in poor LO-RF isolation, however, high LO-RF isolation may not be necessary for such relaxed design requirements. Overall this work presents an excellent view of what is possibly achievable in terms of performance and power consumption.

### 3.1.1.4 Wideband and Improved Technology [4]

This is one of the more recent works in literature on the IEEE 802.15.4 standard. At first glance, the performance is not amazing. Despite using a technology node two generations newer (0.09 µm) than most of the other works (0.18 µm), the current consumption is the highest. However, this design features a significantly wider (3 GHz) bandwidth than any other design. Furthermore, this design required no inductors in the LNA saving valuable die
area. It was the lack of inductors which contributed to both the higher power consumption, and the wide RF bandwidth.

In this design, the authors made a case for 6 MHz IF, using simulations to show that a 6 MHz IF results in the best rejection of in-band interference. Unfortunately, the simulations performed were very specific and far from exhaustive. Based on previous discussions in 2.2.4.2 and the fact that all other works chose a lower IF, we believe that the choice of a 6-MHz IF may have been a mistake.

### 3.1.1.5 Passive Mixer and TIA [5]

To our knowledge, this was one of the earliest works to use a current-output passive mixer. In this case, the TIA used was a current reuse amplifier configured as a simple CMOS inverter, with a feedback resistor from output to input and common-mode degeneration resistors (Fig. 3.3). Such a design can be fast since the number of capacitive elements in the circuit is small, but overall the TIA used here cannot compare with the op-amp based TIA used in [1]. Firstly, the simple TIA used in [5] does not allow for setting of the common-mode output voltage. Secondly, the forward gain is limited due to the resistive loads. Lastly, the low-frequency common-mode rejection afforded by the degeneration resistors could be improved using current sources.

![Fig. 3.3 (a) The simple TIA used in [5] and (b) its equivalent representation.](image)

Although the LNA in this design consumed only 1.8 mW, the overall power consumption of the design was 6.3 mW. Most of the power was consumed in
the TIA in order to support a good overall NF. The TIA power consumption could have been reduced by increasing the gain of the LNA at the cost of higher power consumption in the LNA, and poorer overall IIP$_3$.

This design used a single ended LNA, while the differential LO signal was used to split the IF signal into a differential signal. Such a strategy was possible due to the low IIP$_2$ requirements of the IEEE 802.15.4 standard. A single-ended LNA requires approximately half of the power consumption of a differential one. In section 2.2.5.7, we showed that the IIP$_2$ requirement for the IEEE 802.15.4 standard is only -8 dBm.

### 3.1.1.6 MGTR Linearity Enhancement [6]

This is one of the earliest publications to use the IEEE 802.15.4 standard, and was based on preliminary specifications for the standard. The work uses very standard design techniques such as active mixing, inductive source degeneration, and active-RC based CSFs. The LNA is single-ended, and the active mixers convert the single-ended inputs into differential outputs. This is possible because the LO is a differential signal. The output is therefore a differential signal multiplied by a single-ended signal resulting in a differential output. Fig. 3.4 shows how this would work.

![Fig. 3.4 Illustration of single-ended to differential conversion using a single-balanced mixer.](image)

As the input stages of the active mixers constitute a second nonlinear gain stage, they are required to have high linearity. This was accomplished by
using a technique known as the multiple-gate transistor technique (MGTR) [14]. The idea behind the technique was hinted at in section 2.3.6. Essentially, the input output relationship of a device can be expanded in a Taylor series around a biasing point,

\[
 f(V_{gs} + \nu_{gs}) = f(V_{gs}) + \frac{df(V)}{dV}_{V=V_{gs}} \nu_{gs} + \frac{1}{2} \frac{d^2f(V)}{dV^2}_{V=V_{gs}} \nu_{gs}^2 + \frac{1}{3!} \frac{d^3f(V)}{dV^3}_{V=V_{gs}} \nu_{gs}^3 + ...
\]

(3.3)

where \( f(V) \) is the drain-source current as a function of the gate-source voltage, \( V_{GS} \) is the DC gate-source voltage, \( \nu_{gs} \) is the small-signal gate source voltage. The second term on the right side of (3.3) is the small signal linear output component \( (g_m \nu_{gs}) \) while the fourth term is mainly responsible for IM3. If two biasing points, \( V_{GS1} \) and \( V_{GS2} \) can be found such that,

\[
 \frac{d^3f(V)}{dV^3}_{V=V_{gs1}} = -\frac{d^3f(V)}{dV^3}_{V=V_{gs2}}
\]

(3.4)

while the transconductances, \( g_m \), at the two biasing points are in-phase, then the two transistors could be added in parallel to cancel the IM3 component. While theoretically sound, this method is not practical, and has not found widespread use in RF applications. The main problem is process variation. In practice, variations in transistor threshold voltage result in deviations from the desired biasing point, and without near perfect cancellation of the IM3, little improvement in the IIP3 is generally observed.

Reference [6] describes a full transceiver including PLL, transmitter, receiver and baseband. The total power consumption in the receive chain is 9 mW while the PLL consumes 12 mW.

3.1.1.7 An Alternative Approach [7]

This work took an entirely different approach to the design. The design uses an LNA, and a mixer, followed by a 2 MHz continuous time bandpass ΣΔ ADC. As the receiver is highly digitized, it can expect to achieve excellent filtering, and re-configurability. The total power consumption is 9 mW from a 1.8 V supply, but the power consumed by the LNA and mixer are not quoted.
As the output is not passed through a limiting stage, the receiver can also potentially process amplitude modulated signals. Unfortunately, the 9 mW consumed by the front-end alone is not indicative of how much power could potentially be required in this design. The output of the ΣΔ ADC is a 1-bit 64-MHz signal which needs to be decimated before further signal processing can be done.

In 0.18 μm CMOS technology, highly digitized receivers for low-power applications are not likely to be able to compete with analog based designs in terms of performance. However, with improving technology, digital circuits are becoming faster, and inevitably, using such technology nodes as 65 nm CMOS and beyond will make digital implementations of low-power low-data-rate transceivers much more competitive.

3.1.2 Summary of Circuit Techniques

In an effort to reduce circuit power consumption, several general techniques have been used. First of all, it has been recognized that many circuits do not need to use the full supply voltage in order to achieve certain performance. This led to the use of the current reuse technique in [1] and [2]. Rather than reuse the current, one can simply reduce the supply voltage as was done in [2]-[4]. Most works rely on impedance transformation to achieve some measure of voltage gain in exchange for reduced bandwidth [2], [3], [5], [6]. The use of passive mixers has allowed for low-power, low-flicker noise, high-linearity frequency translation [1], [3]-[5]. For IF CSFs, active-RC based poly-phase filters [15] have gained wide usage [1], [4], [6]. For transistor biasing, subthreshold biasing has shown to be useful [2]. In Chapter 4 subthreshold biasing will be discussed in more detail. Also of note is that only two [5], [6] of the presented designs used inductive source degeneration in the input matching network. We will see in Chapter 4 why this is so.

3.2 System Level Power Saving Methods

In this section, we introduce several architectural and system level methods which have been employed in order to reduce overall energy consumption. The concept of energy-aware design is introduced which leads to the
proposed method of energy-aware design. Several other methods are then introduced which are found to be compatible with energy-aware design. The proposed energy-aware method will be presented in Chapter 4.

### 3.2.1 Basic Energy-Aware Concept

A normal receiver (non-energy-aware) is designed based on its sensitivity requirement, where the sensitivity is the weakest signal receivable by the receiver. An energy-aware receiver is aware of the power consumption it requires to meet its real-time needs. It is therefore able to reduce its power consumption when the performance requirements are relaxed, and increase its performance when the situation demands.

In a typical scenario, a receiver may be receiving a significantly higher input signal than the sensitivity level. For proof, simply look at the signal meter on your cell-phone. Clearly, the receiver’s performance is normally better than it needs to be. This results in unnecessarily high power consumption. Next we will discuss three different methods to make use of the energy-aware concept.

#### 3.2.1.1 Power Consumption Control based on EVM

This method was introduced by Senguttuvan et al. in 2007 [16], [17] and is named Virtually Zero Margin Adaptive RF or simply VIZOR. The idea can be split into two parts: the sensing mechanism and the power control mechanism.

According to the authors, the sensing mechanism aims to sense the quality of the reception. This is done in the DSP by calculating the error-vector-magnitude (EVM). EVM is a performance parameter which quantifies the deviation of a received symbol diagram form its ideal constellation. Essentially, any quadrature amplitude modulation (QAM) based modulation scheme can be represented by a constellation diagram. This includes quadrature phase shift keying (QPSK) which is used by the IEEE 802.15.4 standard. QPSK involves splitting a serial bit stream into two parallel bit streams. One of the parallel bit streams is modulated onto the in-phase carrier, \( I \), while the other bit stream is modulated onto the out-of-phase carrier, \( Q \). This is shown schematically in Fig. 3.5.
Fig. 3.5 A QPSK modulator.

The DSP usually includes pulse shaping in order to limit the bandwidth of the resulting signals. Based on the above figure, there are four possible output symbols. These four symbols are represented in a constellation diagram like that in Fig. 3.6. In Fig. 3.6, we also show four examples of received vectors. The difference between the received symbol constellation and the transmitted symbol constellation is due to the addition of impairments in the channel and the receiver itself. An error vector is essentially the difference between the received vector and the ideal vector. Taking the ratio of the average magnitude of this vector to the magnitude of the ideal vector approximately results in the EVM although the actual calculation is slightly more involved [18]. As a result,

\[
EVM_{RMS} \approx \sqrt{\frac{1}{\text{SNR}}} \quad (3.5)
\]

According to [16], EVM can easily be measured by the DSP in short time. Furthermore, EVM is clearly correlated with BER, and additional headroom can be given in the EVM to ensure adequate BER.
Having sensed the EVM, the DSP must then control the power consumption of the receiver. Senguttuvan et al. propose to control the power consumption via various control knobs, namely the power supply voltage and certain biasing voltages, in a feedback fashion where the EVM is continuously sensed and refined until it is just above the required value. Power consumption, being a product of the supply voltage and the average current consumption, can be controlled by adjusting the supply voltage through voltage regulators or current consumption through the biasing voltages.

Unfortunately, there are several problems with the above approach, both in the sensing mechanism and the power control mechanism.

a) On the sensing mechanism side, we note that EVM is a single metric to include all impairments. On the up side this ensures that all impairments are treated. However, on the down side, the impairments are not treated independently. The power consumption requirements of a receiver are based on many parameters which are in many cases uncorrelated (see Chapter 2, Section 2.2.5). Therefore, more power can be saved by treating the impairments independently.

b) The authors propose to continuously monitor EVM and control power consumption. Adjusting the receiver’s power consumption is not normally done during normal data transfer because it can disrupt certain key blocks such as the frequency synthesizer.
c) Controlling power consumption via the power supply voltage is not very practical. A single global power supply voltage could not be used since different circuit blocks have vastly different effects on the overall power consumption (LNA and CSF for example). Having more than one voltage regulator is costly. Furthermore, voltage regulators cannot normally maintain a high efficiency over a broad range of output voltages.

To our knowledge, the method proposed in [16] and [17] has not at this point gone past the simulation stage. A more practical method which is widely used is presented next.

3.2.1.2 Data-Rate Control
Channel estimation or estimating the quality of the received signal can be accomplished by transmitting a known signal and estimating the quality of the received signal by the BER. Rather than attempting to control the power consumption of the receiver, we may simply control the data rate. Assuming we have a fixed amount of data to send, 1 MB for example, the faster we send the data, the shorter the amount of time which the transceivers need to be on. A new emerging standard which is meant to replace wireless LAN, the IEEE 802.16 standard, or WiMax, allows for various channel bandwidths and modulation schemes. The mobile WiMax standard supports channel bandwidths from 5 MHz to 10 MHz and QPSK, 16-QAM and 64-QAM modulation schemes allowing for data rates between 1 and 5 Mbps [19].

Allowing for variable data rates not only reduces average receiver power consumption, but also reduces average transmitter power consumption since their required communication times go hand-in-hand. The same cannot be said about any of the other schemes we will talk about. In order to achieve variable data rates, the receiver must also be able to adjust its channel filter bandwidth. It should also be noted that much like the previously discussed energy-aware method, the variable data-rate method does not treat signal impairments independently. Therefore, the power saving is not optimal.
3.2.1.3 Transmit Power Control

An obvious method to design an energy-aware system would be to allow control of the transmit power. For example, if device A and device B are communicating with each other, device A only transmits enough power such that the power reaching device B is equal to device B’s sensitivity level. This allows device A to save power. This kind of arrangement is actually used in code-division multiple-access (CDMA) systems to solve the near-far problem [20].

Essentially, systems where multiple users must share the same communications channel can split this resource in a limited number of ways: frequency, time and code. Splitting in frequency involves dividing the system bandwidth into channels as is done in the IEEE 802.15.4 standard where the 83.5 MHz bandwidth is split into sixteen 2-MHz wide channels with 5 MHz spacing. Another multiple access technique is known as time-division multiple-access (TDMA) where each user is given a specific time slot so that users do not interfere with each other. Lastly, users can use the exact same bandwidth at the same time without interfering with each other by using CDMA. Each user is assigned one of a set of pseudorandom (PN) codes. The data to be transmitted is then multiplied by the PN code. The resulting signal is spread in frequency and needs to be de-spread on the receiver side. This is illustrated in Fig. 3.7. Note that if the codes do not match, then the resulting output signal is not de-spread and has a much larger bandwidth than the properly de-spread output. As a result, an interfering CDMA signal occupying the same channel as the desired signal will appear as noise.

Problems arise when two CDMA radios physically close to each other are both communicating with the same base station which is located far away. They each must receive a weak signal coming from the far away base station while dealing with a strong interferer from a nearby radio. In order to prevent such a situation from arising, CDMA systems require that each transmitter only transmit the minimum power required for proper communication.
Getting back to the energy-aware concept, allowing for variable transmit power has some major drawbacks.

a) First of all, in low data-rate systems, the amount of transmitted power may not be a significant portion of the total power consumption.

b) Secondly, backing off PA output power generally results in poorer efficiency.

c) A third issue is that for accurate information about how much power should be transmitted, a two way communication is necessary. i.e. radio 1 transmits information regarding how much power is being transmitted, radio 2 measures the received power and sends back information on how much the transmit power can be reduced.

d) Lastly, like the other methods introduced, different RF impairments are not treated independently making this method sub-optimal.

Resolution of these problems will be left until Chapter 4 when we discuss the proposed energy-aware scheme. We will now turn our attention to other energy saving methods.

### 3.2.2 Other Energy Saving Schemes

In this section, we discuss different energy saving schemes in literature which in general can be used in conjunction with energy-aware methods. The
methods we will discuss are the wake-up receiver (WuRX), energy harvesting, and fast start-up receivers.

3.2.2.1 The Wake-up Receiver

In general, a receiver may only actually be receiving data for a short duration in a given interval. However, if it cannot anticipate when it is going to receive a signal, it must be set to its best sensitivity state in case it does receive a weak signal. One way to get around this problem is to use a WuRX [21], [22]. The principle of operation is implied in its name. While waiting for a signal, the main receiver is set to sleep mode. During which time, an auxiliary receiver is turned on whose sole function is to wake up the main receiver when a desired signal is incoming. Such a WuRX can be designed with extremely low power consumption since the specifications it must meet can be relaxed. If desired, the data rate requirement can be significantly lower than the main receiver. Given that Sensitivity (dBm) = \(-174 + BW \text{ (dB)} + NF + SNR_{req}\) (See Chapter 2, Section 2.2.5.2), the channel bandwidth in dB must be reduced by the same amount that the WuRX NF is increased.

Unfortunately, there are certain power overheads which are for the most part unrelated to the NF requirement. The most obvious is in the frequency synthesizer. The PLL often requires a large portion of the overall power consumption, especially in low power designs [1], [6], [23]. In [22], a receiver architecture was proposed specifically to be used in a WuRX. Essentially, the channel selection was performed by a high Q front-end filter which obviated the need for a tunable PLL. This architecture allowed for only a single frequency channel since the filter’s center frequency could not be changed. Rather than attempt to demodulate the signal at RF, the authors proposed to down-convert the incoming signal using a free-running oscillator. This gave rise to the term “uncertain IF”. Given normal process variation, the down-converted signal was known to be within a certain frequency range. The modulation used was on-off keying (OOK) which allowed for simple detection. In [22], the 2-GHz WuRX used only 52 µW while providing a data rate of 100 kb/s at -72 dBm sensitivity.
Since the WuRX is separate from the main receiver, it can be used in conjunction with energy-awareness.

### 3.2.2.2 Energy Harvesting

Although energy harvesting does not actually reduce the power consumption of a device, the main goal of a low-power radio is to extend battery life or remove the battery entirely. This is possible with energy harvesting. Energy harvesting via solar cells is an obvious way to do this and was proposed for the use in wireless sensor networks in [3]. From a research point of view, harvesting electromagnetic energy is an interesting idea. This was used in [24] to power a demodulation circuit used in a wireless sensor node.

### 3.2.2.3 Fast Start-up

In [25], the authors proposed a receiver architecture which allowed for fast start-up. Like that in [22], the transmission was modulated using OOK, but in this case, the receiver architecture did not include an uncertain IF. In fact, no frequency synthesizer was used which greatly reduced the start-up time. The receiver achieved a 2.5-µs start-up time. Such a receiver is extremely useful in situations where the total amount of data to be transferred is small. In which case, a normal receiver’s start-up time could be a significant portion of the total running time. Much like the OOK receiver in [22], the receiver in [25] supports only a single frequency channel for communication.

### 3.2.3 Summary of System Level Techniques

We have grouped system level energy saving techniques into energy-aware techniques and other techniques. Among the energy-aware techniques, we find that control of the receiver’s performance via EVM is the easiest to implement since no special communications standard is needed. In the transmit power control, two way communication of the channel conditions is required, while a special standard supporting different data rates and types of modulations is required for data rate control. None of the current energy-aware techniques support independent control of interference and noise performance.
Several other energy-saving techniques have been discussed, and all of which can potentially be coupled with energy-awareness. OOK modulation seems to be a good scheme for designing ultra-low power receivers, however for the lowest power receivers, special high-Q front end filters are required which will most likely limit their use. Lastly, if the radio’s total power consumption is sufficiently low, the possibility arises of powering the whole system via energy harvesting circuitry. The next chapter will discuss the proposed system level energy-awareness scheme as well as an architectural refinement to typical low power receivers which will allow for significant power savings.
References


Chapter 4
System and Architectural Level Proposals

4.1 Introduction

In this chapter, we will discuss the system and architectural level contributions of this study. In particular, we will discuss the proposed energy-aware scheme, and a novel receiver architecture for low data-rate applications. The benefits of the proposed system novelties will be discussed alongside their limitations.

4.2 Proposed Energy-Aware Scheme

The proposed energy-aware scheme involves adjusting the RX front end’s power consumption based on its in-situ required NF. While the final design merit for an RX is its bit-error rate (BER), RFIC designers generally split the performance requirements up into nearly independent specifications. As discussed in Chapter 2, signal non-idealities arise due to linear distortion [1], interference, and random noise. In [2], and [3], the authors proposed to group the signal non-idealities into a single parameter, the EVM. EVM was chosen for its strong correlation with BER. The main drawback of this approach is that signal impairments are not treated independently. Therefore, the required power consumption as dictated by the EVM is always a worst case scenario rather than an optimum case (See Section 3.2.1.1). Since it is not possible to measure true BER in real-time, additional headroom on the measured EVM must also be given.
Our proposed energy-aware scheme [4], [5] is to treat interference and noise independently (linear distortion is not treated). This can lead to the optimum power consumption. However, we will see that interference performance is not easy to control. As a result, a sub-optimal design can be achieved by only controlling the noise performance. This method will be shown to be well suited to low-power, low-cost designs.

4.2.1 Noise and Interference

In Chapter 3, section 3.2.1.1, we discussed the sensing mechanism for an energy-aware control system. Essentially, four conditions make for a good sensing parameter to use in the sensing mechanism.

a) The parameter that is being controlled should have a strong correlation with power consumption.

b) Furthermore, it must be correlated in the correct direction. What this means is that if the required performance drops, the power consumption required to achieve the new performance also drops.

c) Thirdly, as much as possible, we do not want a change in one parameter to degrade another performance parameter. For example, if we increase (making it worse) the system NF, it should not degrade the system $IIP_3$

d) Lastly, a good sensing mechanism should be readily controllable in order to avoid high system complexity.

We will now look at several performance parameters, and see how well they fit this description.

4.2.1.1 Noise Figure

Noise figure ($NF$) correlates well with power consumption both at the system level and the circuit level. The input-referred noise of a MOSFET (only channel noise is considered) is approximately equal to,

\[
V_{n_{in}}^2 \approx \frac{4kT\gamma T\gamma}{ag_m}
\]  

(4.1)
where \( k \) is Boltzmann’s constant, \( T \) is temperature in Kelvin, \( \gamma \) is a parameter approximately equal to 2/3 in saturation for long-channel devices, \( \alpha \) is the ratio of \( g_m \) to the transconductance when the drain source voltage is zero, and \( g_m \) is the device transconductance [6]. Since \( g_m \) improves with current consumption,

\[
g_m = \sqrt{2\mu_0 C_{ox} \frac{W}{L}/I_{DS}}
\]  

(4.2)

where \( \mu_0 C_{ox} \) is process dependent, \( W/L \) is the aspect ratio, and \( I_{DS} \) is the drain-source current, current consumption can be directly linked to NF. At the system level, NF is also indirectly related to current consumption through the gain of a cascaded system. The cascaded NF can be calculated as,

\[
NF_i = 10 \log(F_i)
\]  

(4.3a)

\[
F_{total} = F_1 + \frac{F_2 - 1}{G_i}
\]  

(4.3b)

where \( F_{total} \) is the noise factor of the system, \( NF_n \) is the noise figure of the \( n^{th} \) stage, \( F_1 \) is the noise factor of the first stage, \( F_2 \) is the combined noise factor of all subsequent stages, and \( G_1 \) is the power gain (which is proportional to the square of the voltage gain) of the first stage. When \( F_1 \) is small compared to \( F_{total} \), \( F_{total} \) is inversely proportional to the first stage’s squared-gain. Assuming we have a common-source LNA representing \( G_1 \), its voltage gain is proportional to the square-root of the current consumption, and therefore, \( F_{total} \) is inversely proportional to the current consumption.

We can easily see that the noise figure is both correlated with power consumption, and correlated in the correct direction. Furthermore, from the discussion in chapter 2, section 2.3.5, we know that a reduction in gain generally leads to an improvement in linearity. Since LNA gain control is readily implemented, all four sensing mechanism requirements are met.

### 4.2.1.2 IIP_3

From a circuit level perspective, \( IIP_3 \) is dependent on the biasing conditions, and also on any linearization methods used such as negative feedback, or
MGTR (See Ch. 3, Section 3.1.1.6). Attempting to control the $IIP_3$ of a MOSFET can have varied results. For example, Fig. 4.1 shows the first and third derivatives of the current with respect to the gate source voltage of a MOSFET of unit aspect ratio with a 0.6 V drain-source voltage. The corresponding $IIP_3$ is shown in Fig. 4.2.

Fig. 4.1 First and third derivatives of a MOSFET of unit aspect ratio with 0.6 V drain-source voltage.

Fig. 4.2 $IIP_3$ of the above device.
Note that the $IIP_3$ was calculated simply by taking $\sqrt{|8g_m/g_m'|}$ (the factor of 8 is from $4/3 \times 3!$, see sections 2.2.5.4 and 3.1.1.6). Hence this $IIP_3$ does not include high frequency effects. Although there is a general trend towards reduction in $IIP_3$ with power consumption, there is a large zone around $V_{GS} = 0.6$ V where the $IIP_3$ is maximum. Furthermore, Fig. 4.2 fails to show the change in sign of the IM$_3$ product. Overall, although the device $IIP_3$ is correlated with power consumption, it is difficult to manage.

Another way to control a circuit’s $IIP_3$ would be to control the loop gain of a negative feedback loop. LNA topologies which employ negative feedback include inductive degeneration and resistive shunt feedback [7], [8]. However, adjusting an amplifiers loop gain can potentially de-stabilize the circuit if the phase margin is degraded [9], and can also affect loading conditions between the stages. This is particularly important in LNA design where the input impedance is usually designed to match a 50-$\Omega$ source.

At the system level, controlling $IIP_3$ faces a greater problem, namely, condition (b) in section 4.2.1 is not met. When input power is high, we can reduce the gain of the system. In general, this naturally has the effect of improving system $IIP_3$. However, the required $IIP_3$ is lower. Therefore, it is inconvenient to control a system’s $IIP_3$.

4.2.1.3 1-dB Compression Point
Gain compression occurs mainly due to one of two effects: current limiting or voltage limiting. Consider the generic differential amplifier in Fig. 4.3. The maximum current which can flow through either input transistor is limited to $I_1$. This is the current limit. Furthermore, the maximum voltage which can develop at node $V_{out}$ is $V_{DD}$ while the minimum is the voltage across the transistor and the current source, $V_{low}$. Therefore, the maximum output swing is $2(V_{DD}-V_{low})$. Clearly, the circuit in Fig. 4.3 is current limited if,

$$\frac{g_mV_{in}}{I_1} > \frac{g_mZ_{load}V_{in}}{2(V_{DD}-V_{low})}$$

or,

$$\frac{g_mV_{in}}{I_1} > \frac{g_mV_{in}}{2(V_{DD}-V_{low})}$$

or,

$$\frac{g_mV_{in}}{I_1} > \frac{V_{in}}{2(V_{DD}-V_{low})}$$

or,
The opposite is true for a voltage limited circuit. Note that $V_{\text{low}}$ tends to be a soft limit while $I_1$ is generally a hard limit. This leads to more gradual roll-off in gain compression curves for voltage limited devices.

![Fig. 4.3 A generic differential amplifier.](image)

The point we are trying to make is that whether a circuit is voltage limited or current limited, the 1-dB compression point is dependent on the biasing conditions of the device. If the device is current limited, reducing $I_1$ when the input signal is large would actually degrade the 1-dB compression due to an increased $g_m/I_1$. This is opposite to the requirement. If a device is voltage limited, reducing $I_1$ would reduce $g_mZ_{\text{load}}$ and thereby improve the 1-dB compression point.

At the system level, reducing gain generally improves 1-dB compression point. It is therefore correlated in the correct direction with power consumption. Although reducing gain also degrades NF, the NF requirement also relaxes. Overall it seems that 1-dB compression point is suitable for power control. Unfortunately, at low signal levels, the 1-dB compression point is generally much higher than necessary. As 1-dB compression point correlates well with gain and indirectly NF, an improved 1-dB compression point at low gains can simply be seen as a bonus.
4.2.1.4 Image Rejection Ratio
In a image-reject mixer architecture, the IRR of a system is mainly dependent on the matching between the I and Q paths of the receiver, while in a image-filter architecture, the IRR is dependent on the filter order and roll-off. Although architectural level decisions can be made which will affect the power consumption of the receiver, the IRR is generally not very easy to control by affecting the power consumption. Therefore, IRR will not be further discussed.

4.2.1.5 Phase Noise
Phase noise is a parameter of the local oscillator. Architectural decisions related to power consumption can lead to improved phase noise, but like IRR, it is not easy to control phase noise in situ in order to optimize power consumption.

4.2.1.6 Overall
We have discussed several parameters related to noise performance and interference performance, and overall it should be clear that $NF$ is the most suitable parameter to control. It is well correlated with a system’s power consumption, and is easily controlled at the system level through the gain. Fig. 4.4 summarizes the $NF$ and $IIP_3$ requirements are a function of input signal power. As the input signal power increases, the required $NF$ and $IIP_3$ become more relaxed.
In Fig. 4.4, we assumed that the input SNR increases with input signal power indefinitely. In practice, the transmitter can only transmit signals at a limited SNR. This is usually specified as an EVM requirement. For instance, the RMS EVM requirement for the IEEE 802.15.4 standard is 35% over 1000 measured chips [10]. As the data rate is relatively low, the actual EVM achieved is normally significantly better than this value. In [11], the EVM achieved was 5.3% which corresponds to an SNR of about 25.5 dB. Note that the output SNR of a system can be expressed as,

\[
SNR_{\text{out}} = \frac{P_{\text{in}}}{N_{\text{total,in}}} = \frac{1}{N_{\text{TX,in}}/P_{\text{in}} + N_{\text{env}}/P_{\text{in}} + \frac{\text{IRN}_{\text{system}}}{P_{\text{in}}}}
\]

(4.6)

where \(\text{IRN}_{\text{system}}\) is the input-referred-noise (IRN) power of the system, \(P_{\text{in}}\) is the input power, \(N_{\text{total,in}}\) is the total noise referred to the input of the system, \(N_{\text{TX,in}}\) is the noise received from the transmitter, \(N_{\text{env}}\) is the noise received from the environment. In (4.6), the term \(N_{\text{TX,in}}/P_{\text{in}}\) is constant and equal to the transmitter’s output SNR, \(SNR_{\text{TX}}\). When \(P_{\text{in}}\) is large, \(N_{\text{env}}/P_{\text{in}}\) is small and negligible. In order for the receiver to maintain a constant output SNR, the ratio \(\text{IRN}_{\text{system}}/P_{\text{in}}\) must be kept constant. This is easier to see in (4.7)
The approximation in (4.7) holds for high values of the NF which occur in low data-rate systems such as the ones under consideration. When \( SNR_{in} \) reaches the limit defined by the transmit EVM, we only need to keep the ratio \( IRN_{system} \) to \( P_{in} \) constant in order to maintain a constant output SNR. Therefore, the transmit EVM does not set a hard limit for the energy-awareness of a system.

### 4.2.2 Sensing the NF

The total noise (considered in the calculation of NF) includes two parts: the received noise and the noise added by the receiver (Fig. 4.5) [12]. We have established in (4.7) that the received noise is small compared to the noise added by the receiver when the NF is high. Therefore, in order to control the noise figure of the receiver, we only need to consider the noise added by the receiver. This is easily seen from (4.6) and (4.7). Remember that when the receiver is in its highest gain states, \( 1/SNR_{TX} \) is negligible. We can approximately say that the input noise is unimportant when the NF of the receiver exceeds 10 dB. Under this situation, the receiver’s noise accounts for more than 90% of the total output noise.

Noise picked up by the receiver comes from the black body radiation of the environment. The amount of noise picked up by the antenna depends on the temperature of the environment in view of the antenna. The surface of the earth is approximately 290 K, and because it makes for simple calculations,
this is the temperature normally taken in the calculation of the antenna noise power (and defined by $NF$). In certain environments, the background noise temperature can be significantly higher. 290 K corresponds to 17 °C. The input noise power is $kT\Delta f$, and $NF$ is a factor in decibels. Hence, the received noise in decibels is calculated as $10\log(\frac{k}{T}) + 10\log(\Delta f)$. Ignoring the bandwidth term, if the background noise temperature is increased from 17 °C to 100 °C, the received noise increases from -174 dBm/Hz to -172.9 dBm/Hz. Under such a situation, we only need to improve the receiver $NF$ by 1.1 dB in order to achieve the same output SNR. As we will not normally design for such an extreme situation, we can safely say that the received noise is around -174 dBm/Hz, and add a safety margin to the receiver $NF$ to ensure sufficient sensitivity.

In the above discussion, we have established two points: the received noise can be taken as -174 dBm/Hz, and the received noise is insignificant to the calculation of $SNR_{out}$ when the receiver $NF$ is more than 10 dB. Remember that since we are treating the individual signal impairments independently, we are not worried about other noise-like impairments such as interference, DC-offset, phase noise, etc. The receiver will only switch to a poorer $NF$ state when the required $NF$ is relaxed. For the IEEE 802.15.4 standard, we calculated the overall required $NF$ to be around 12 dB. The total required dynamic range of the IEEE 802.15.4 standard is 65 dB. If we allow for $NF$ steps of 6 dB, then the first step occurs when the required $NF$ is 18 dB. Under this condition, we can safely assume that the received noise (0.98% contribution to the total output noise) is negligible. Therefore, for the case of interest, we only need to measure the receiver’s output noise and the signal output power in order to know the output SNR.

### 4.2.2.1 Typical RSSI
Measuring the received signal strength is a standard feature in many modern radios, and the block which takes care of the operation is called the received signal strength indicator (RSSI) [13]. A block diagram of a standard RSSI is shown in Fig. 4.6 [14]. The multipliers have the effect of squaring their input. The resulting output is filtered by a low-pass filter. Assuming that each amplifier has a maximum output voltage of 0.125 V, the output, $V_{RSSI}$, will be
similar to the plot in Fig. 4.7 (two different small signal amplifier gains are shown: 2 V/V and 4 V/V). The output is shown assuming a hard limit to the output voltage of each limiter. If a soft limit is used (see section 4.2.1.3), the output curve will be smoother. The analog output can be converted to a digital one using a simple analog to digital converter (ADC).

![Block Diagram of a Typical RSSI](image)

*Fig. 4.6 Block Diagram of a Typical RSSI*

![Output V_RSSI of the System in Fig. 4.6 for two different amplifier gains. The input power is referenced to a 50-Ω load.](image)

*Fig. 4.7 Output V_{RSSI} of the System in Fig. 4.6 for two different amplifier gains. The input power is referenced to a 50-Ω load.*
4.2.2.2 The Need to Measure IRN

According to the arguments made so far, we only need to measure IRN and input power. In fact, while measuring IRN could help, it is not necessary. This is because unlike input signal power, IRN is a fixed value which is set by the designer. If the modeling of the receiver is adequate in terms of noise and gain, then the IRN will be known in advance. The inaccuracy of the modeling must be taken into account with additional headroom. This is true for any receiver, energy-aware or not, but it is more important for an energy-aware receiver since if the IRN is higher than expected in a non energy-aware design, the receiver's sensitivity will simply degrade., but in an energy-aware design, the receiver could fail for all input powers. This is because the receiver is designed to maintain a fixed output SNR. If this fixed output SNR is too low for the one input power, it could be too low for all input powers. This issue is illustrated in Fig. 4.8. The dotted line represents the required $NF$ and $SNR_{out}$ for successful detection, and $\Delta NF$ and $\Delta SNR_{out}$ represent the uncertainty in the receiver's implemented $NF$ and $SNR_{out}$ due to variation in the process. Clearly, without sufficient room for uncertainty, the EA system could fail entirely while a conventional system would merely exhibit a reduced sensitivity.

![Illustration of implemented NF and SNR for an EA receiver and a conventional receiver. ΔSNR out and ΔNF are the uncertainties in the SNR out and NF respectively.](image)

While we have argued against the need to measure a receiver's IRN, it should be noted that such a chore can possibly be accomplished by simple means. In theory, the output noise power of the receiver could be measured by simply
increasing the sensitivity of the RSSI and either recording the RSSI when the receiver is not receiving any signals, or blocking input signals via a switch. This is shown schematically for a direct conversion receiver in Fig. 4.9. When the receiver IRN is to be measured, the switch is connected to a dummy 50-Ω source to prevent the receiver from picking up signals. When the receiver input signal power is to be measured, the switch is closed. The switch could be integrated into the design of a T/R switch. In this thesis, we have focused on the design of the front end. Therefore, the full receiver implementation is not within the scope of this work.

4.2.3 Timing Issues

The method by which we have proposed to control the noise figure of the receiver involves sensing the signal power, and setting the receiver to the correct power/gain state. Here, power consumption is controlled in discrete steps in an energy-aware fashion where each step corresponds to a particular power/gain state. Changing the power/gain state of the receiver introduces a glitch into the reception. In the worst case this glitch could cause the PLL to “unlock”. It could also cause a bit error and require the information packet to be resent. Fortunately, the receiver state control does not need to be done during the reception of useful information. The IEEE 802.15.4 standard allocates 128 µs worth of preamble at the start of each data packet for the receiver PLL and AGC to lock [10]. Therefore, power/gain state control does not disturb the operation of the receiver when useful data is being transmitted. Nevertheless, it is important to minimize the time required for the front-end to switch power/gain state.
4.2.4 Receiver Architecture Optimization for Energy Awareness

Energy-aware receiver design relies on a receiver’s ability to regulate its power consumption based on the in-situ required specifications. However, any practical receiver design has overhead power requirements which can be considered fixed. For example, the power consumption of the frequency synthesizer has little correlation with the overall NF of the receiver. Although we can conceivably adjust the frequency synthesizer’s output power based on the required NF (as was attempted in [4]), there is still a minimum power consumption required by circuit blocks such as the frequency dividers and phase-frequency detector. The goal of the receiver designer therefore should be to minimize the overhead power consumptions, and try to compensate for the degraded noise performance using blocks whose noise performance depends heavily on power consumption. This will allow for the greatest control of the overall power consumption.

4.2.4.1 Overhead Power Consumption

The biggest limitation on the controllability of the receiver power consumption is in the power consumption required by the frequency synthesizer. In [13], and [15], the frequency synthesizer required 9.72 mW, and 12 mW, respectively, while in [16] it required just 2.4 mW. All frequency synthesizers were designed using CMOS for the IEEE 802.15.4 standard but [13] and [15] used 0.18 µm technology and [16] used 0.13 µm technology. Improving technology and frequency synthesizer architectures can therefore lead to very low power overhead for the frequency synthesizer.

Another required power overhead is due to the bandwidth requirements of the op-amps used in the channel filter. In order to provide proper filtering, the op-amp’s loop gain should be more than one over the entire system bandwidth which is 83.5 MHz in the case of the IEEE 802.15.4 standard [10]. This is because when the loop gain is less than one, the feedback paths become feed-forward paths which make the filtering transfer function non-ideal. Lastly there is some power overhead required by support circuit blocks such as bandgap references and calibration circuitry.
4.2.4.2 Proposed Design Methodology

This leads to our proposed design methodology. By pushing the requirements of the receiver to the front-end LNA, we can increase the amount of controllable power consumption in the receiver. In other words, we reduce power consumption required by circuit blocks whose power consumption cannot be readily controlled such as the channel-select filter, and frequency synthesizer, and increase the power consumption in blocks whose power consumption is easily controlled such as the LNA. For instance, the channel select filter can be designed with the lowest possible power consumption which allows it to meet bandwidth and linearity requirements. The high IRN can then be compensated with increased front-end gain. High front-end gain requires additional front-end power consumption. Overall, the front-end power consumption will have more room for controllability. Of course, this could lead to a suboptimal design at the sensitivity level, but for typical input power levels, significant amounts of energy could be saved. Fig. 4.10 shows the break down of the average power consumption for receivers designed in [13], and [16]-[20], and the author’s designs in [5] and [21]. [21] presents simulated results which have been submitted for publication and fabrication. Further discussion on the results of our works are presented throughout the thesis.

![Average power consumption breakdown for five recent works in literature and two proposed designs. The proposed designs have 4 power states represented by four levels.](image_url)

Fig. 4.10  Average power consumption breakdown for five recent works in literature and two proposed designs. The proposed designs have 4 power states represented by four levels.
In order to push the requirements to the LNA, the LNA must be able to provide a high voltage gain. Furthermore, in order to compensate for the high LNA gain, all subsequent circuit blocks up to and including the channel-select filter must exhibit high linearity. High linearity can be achieved in the down-conversion and channel filtering stages by using passive mixers and active-RC filtering [13]. An alternative to this approach will be discussed in section 4.3.

4.2.5 Summary of the Proposal
The arguments presented so far can be summarized in the following points.

a) A receiver’s power consumption is minimized by controlling the power consumption based on the in-situ required performance. This performance is optimized by splitting up receiver impairments into independent measures.

b) The most readily controllable performance parameter is the NF.

c) The IEEE 802.15.4 standard requires only a 12 dB NF. When the NF of the receiver is more than 10 dB, 90 % of the output noise is accounted for by the receiver’s internally generated noise.

d) In this situation, the receiver’s output SNR can be approximated by measuring the signal power and the receiver IRN. Measuring the receiver’s input signal power is a standard feature of modern radios and is accomplished by the RSSI.

e) Noise power can be measured by the RSSI, but it is unnecessary with adequate modeling.

f) It is important to minimize the time required by the receiver to switch power state. However, power state control does not occur during reception of useful information.

g) Optimizing an energy-aware design involves maximizing the controllable power consumption. This can be done by pushing performance requirements to blocks whose power consumption and NF can be readily controlled.
The proposed energy-aware scheme has been published in two works [4], [5]. In the next section, we will discuss two different power saving receiver architectures.

4.3 Low Power Receiver Architectures

As discussed, the proposed energy-aware system design requires for a low power high linearity IF section. This will allow us to push the noise, gain and linearity requirements to the front end. In this section we will discuss two receiver architectures which are able to do this. The first architecture was originally employed in [13] and allowed the authors to use a very high gain LNA (33 dB), while the second architecture is a novel architecture and achieves ultra-low power consumption. Both architectures employ passive mixers in the down-conversion. Therefore, before discussing the two architectures, we will make the case for the use of passive down-conversion.

4.3.1 The Case for Passive Mixing

Among recent low-power research works, architectures using passive mixers have generally out-performed those using active mixers in terms of overall sensitivity (for the IEEE 802.15.4 standard) [13], [15], [17]-[20], [26]. This is mainly attributable to the fact that passive mixers distort the input signal less (due to the passive operation), and do not add flicker noise to the system. However, a standard Gilbert-Cell mixer does both. Given the IEEE 802.15.4 standard receiver blocking profile [26], we can calculate the sensitivity based on $IIP_3$ as,

$$\text{Sen}_{IP3} = 3P_{blk} + SNR_{req} - 2IIP_3 \quad (4.8)$$

where $P_{blk}$ is the interfering power, $SNR_{req}$ is the required output signal to noise ratio (SNR), and $IIP_3$ is the receiver input-referred third order intercept power. From chapter 2 (Section 2.2.5.1), we estimated the $SNR_{req}$ to be approximately 14 dB while $P_{blk}$ is -52 dBm in the worst case when the input power is 3-dB higher than the required sensitivity (-85 dBm). The sensitivity based on $NF$ can be easily calculated as,
$Sen_{NF} = NF + 10 \log(kT\Delta f) + SNR_{req}$ \hspace{1cm} (4.9)

where $10 \log (kT\Delta f)$ is -111 for a 2 MHz signal bandwidth. Table 4.1 shows the sensitivity of designs [13], [17]-[20], [26], and [5] and the mixer type used (this information is also found in Chapter 2). We have included the overall power consumption as was published, but it is important to note that different works presented more or less complete systems. Furthermore, certain designs [18], [19] were not specifically designed for the IEEE 802.15.4 standard. Table 4.1 clearly shows the advantage of using passive mixers in IEEE 802.15.4 systems. Of the two designs using active mixers, [17] fails to meet sensitivity requirements based on $IIP_3$, and [20] requires more power consumption than other works. In [5] we used passive down-conversion. We will not further discuss active mixers, but refer the readers to [27], [28] for excellent discussions on active mixer operation.

**TABLE 4.1**

<table>
<thead>
<tr>
<th>Reference</th>
<th>[13]</th>
<th>[17]</th>
<th>[18]</th>
<th>[19]</th>
<th>[20]</th>
<th>[26]</th>
<th>[5]$^A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sen$_{NF}$ (dBm)</td>
<td>-91.3</td>
<td>-92</td>
<td>-91.9</td>
<td>-91</td>
<td>-89.7</td>
<td>-87</td>
<td>-90.7</td>
</tr>
<tr>
<td>Sen$_{IIP_3}$ (dBm)</td>
<td>-110</td>
<td>-68</td>
<td>-127</td>
<td>-118</td>
<td>-126</td>
<td>-112</td>
<td>-88</td>
</tr>
<tr>
<td>Mixer Type$^A$</td>
<td>PC</td>
<td>A</td>
<td>PV</td>
<td>PV</td>
<td>PC</td>
<td>A</td>
<td>PC</td>
</tr>
<tr>
<td>Tech. (nm)</td>
<td>180</td>
<td>180</td>
<td>130</td>
<td>90</td>
<td>180</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td>Power (mW)</td>
<td>10</td>
<td>1.4</td>
<td>0.75</td>
<td>4.05</td>
<td>6.3</td>
<td>10.8</td>
<td>5.4</td>
</tr>
<tr>
<td>IF Power (mW)</td>
<td>5.76$^C$</td>
<td>0.5</td>
<td>0.75</td>
<td>1.15</td>
<td>4.5</td>
<td>-</td>
<td>0.36</td>
</tr>
</tbody>
</table>

$^A$ A: Active, PV: Passive Voltage-mode, PC: Passive Current-mode

$^B$ Second gain mode

$^C$ Estimated only.

**4.3.2 Op-Amp Based CSF**

The first architecture which we will discuss uses an op-amp based channel-select filter (CSF) [13]. Op-amp based active-RC filters use negative feedback to achieve very high linearity. In Chapter 3 we mentioned that negative feedback improves both $IIP_3$ and $IIP_2$ according to [22], [23],
The large boost in linearity allows us to achieve more than sufficient linearity for the IF section. This in turn lets us increase the gain of the front-end thereby pushing the performance requirements to the front. Before showing the full receiver architecture, it is useful to show how the op-amp based CSF is implemented as a complex band-pass filter. This is useful when down-converting to a low IF. In a direct conversion receiver, simple low-pass filters could be used.

4.3.2.1 Active-RC Complex BPF

A topology for a differential active-RC low pass filter is shown in Fig. 4.11. In [24], it was shown how this active RC low pass filter can be transformed into an active-RC complex band-pass filter. We can illustrate this by first supposing that we have a low-pass transfer function,

\[ H_{lp}(j\omega) = \frac{1}{1 + j\omega/\omega_0} \]  

\( \omega \) is a complex value, however only real values are physically possible. We can plot a 3D graph of the magnitude of the transfer function, \(|H_{lp}(j\omega)|\), versus the imaginary and real parts of \( \omega \). The actual magnitude of the transfer function observed in the filter is found by extracting the points on the graph where \( \text{Im}[\omega] = 0 \). The two graphs are plotted in Fig. 4.12 with \( \omega_0 = -0.3 \) rad/s.
In order to transform the low-pass response into a band-pass response, we need to shift the pole from a purely imaginary value to a complex value. In [24], this is done by transforming the low-pass response according to,

\[
H_{bp} (j\omega) = H_{lp} (j\omega - j\omega_c) = \frac{1}{1 - j\omega_c/\omega_0 + j\omega/\omega_0} = \frac{1}{1 - 2jQ + j\omega/\omega_0} \tag{4.13}
\]

In Fig. 4.13, the transformation is applied with \(\omega_c\) set to -0.5 rad/s. The new pole frequency occurs when \(\omega = 0.5 - j0.3\) rad/s. Under normal circumstances, it is not possible to synthesize a single complex pole as they normally occur as complex conjugate pairs. However, in a low-IF design, the down-converted signal is a quadrature signal. The basic unit which realizes a single complex pole is shown in Fig. 4.14, where \(A\) is the center frequency gain, and \(Q\) is the quality factor. In [13], a cascade of three such units was used to realize a 3rd order Butterworth response.

The op-amps’ input nodes at sufficiently low frequencies are AC ground. Therefore, the input resistors (and hence the passive mixer core) determine
the input resistance to the filter. For the in-phase half, we use Kirchoff’s current law to say,

\[
\frac{(V_{\text{OUT}} - V_{\text{OUT}P})(1 + RsC)}{2R} = \frac{(V_{\text{INP}} - V_{\text{INM}})A}{2R} + \frac{(V_{\text{OUT}Q} - V_{\text{OUT}P})Q}{R} \tag{4.14a}
\]

\[
\frac{(V_{\text{OUT}} - V_{\text{OUT}P})(1 + RsC)}{2R} = \frac{(V_{\text{INP}} - V_{\text{INM}})A}{2R} + j(V_{\text{OUT}Q} - V_{\text{OUT}P})Q \tag{4.14b}
\]

\[
\frac{(V_{\text{OUT}} - V_{\text{OUT}P})(1 + RsC - 2jQ)}{2R} = \frac{(V_{\text{INP}} - V_{\text{INM}})A}{2R} \tag{4.14c}
\]

\[
\frac{V_{\text{OUTP}}}{V_{\text{INP}}} = \frac{A}{(1 - 2jQ + RsC)} \tag{4.14d}
\]

It is easy to see the correspondence between (4.14d) and (4.13). For both the LPF and the BPF, the total bandwidth (two sided) is \(2\omega_0\). For the BPF in Fig. 4.13, the IRR is found by taking the ratio of the magnitude of the transfer function at \(\text{Re}[\omega] = 0.5\) rad/s and -0.5 rad/s (equal to 10.8 dB).

Fig. 4.13 3D plot of a band-pass response formed using a complex pole frequency of \(\omega = 0.5 - 0.3j\) rad/s.
4.3.2.2 Integrated Passive Mixer and CSF

In [13], a simple way in which the passive mixer can be integrated into the CSF is shown. We have reproduced this in Fig. 4.15. Here it is recognized that the passive switching transistors are nothing more than time-varying resistors. Design equations for this type of passive mixer are presented in the next chapter. Essentially, the operation is very similar to the ideal operation of the complex band-pass filter in Fig. 4.14. Here the input resistance is mainly determined by the passive mixer core. The passive mixer core converts the RF input voltage into an IF input current. This IF input current forms a voltage drop across the resistors and capacitors connected across the op-amps. Generally speaking, smaller mixer core on-resistance leads to higher gain, and lower noise, but higher capacitive loading to the frequency synthesizer and LNA, and reduced loop-gain (refer to the next chapter for a detailed analysis).
4.3.2.3 Effect of Filtering on $IIP_3$

In order to optimize the filter design, it is important to understand the effect of filtering on the $IIP_3$ of the filter. Suppose we have two interfering signals, $B$ and $C$, and a desired signal, $A$. At the output of the filter, $B$ and $C$ are attenuated by the filtering function, while $A$ is amplified by the filter's pass-band gain. Therefore, at the input of the op-amp, the three tones must be equal to their output level divided by the op-amp gain. If the three tones are within the op-amp’s 3-dB bandwidth, then they will have the same frequency profile at the input and the output. The effect is illustrated in Fig. 4.16. At the frequencies of tones $B$, and $C$, the difference in input signal level to the input of the op-amp with and without filtering capacitors is denoted as $GD_B$ and $GD_C$. Following the arguments presented in section 2.2.5.4, it is easy to see that the $IIP_3$ is improved by approximately $\frac{1}{2}(2GD_B + GD_C)$. This is equivalent to filtering the signal with a unity pass-band gain filter and then amplifying the signal by the desired amount. Which one comes first (filtering or amplification) is an important distinction since if amplification had come first the $IIP_3$
achieved would be significantly lower. Both pre-filtering and negative feedback can contribute to a very high IF section $IIP_3$.

We decided in [5] not to make any assumptions on the IF frequency (no filtering was included). This turned out to be a mistake since a significant improvement in IF section $IIP_3$ could have been achieved if we had provided filtering and specified the receiver interference profile.

![Diagram](image)

*Fig. 4.16 Effect of addition of filtering capacitors on the op-amp input of an op-amp based CSF. (a) Without capacitors and (b) with capacitors.*

### 4.3.2.4 Full Front-End Architecture

Based on the above discussion, the full front-end architecture is formed with Fig. 4.15 preceded by an LNA. We used this architecture in [5] without the filtering capacitors. Hence the op-amp was configured as a trans-impedance amplifier (TIA). The principle as has been discussed is that the IF section’s $IIP_3$ is high enough to be neglected, and its power consumption is low compared to the LNA’s power consumption. As a result, there is more room for energy-aware control. As the receiver is expected to be used in a low-IF configuration, some method to keep the IF section flicker noise corner frequency low was necessary.

The fully-differential op-amp is shown in Fig. 4.17. The input differential pair uses parasitic NPN transistors which provide better matching, DC-offset and flicker noise performance than MOS devices [25]. In a CMOS process, NPN bipolar junction transistors (BJT) are formed using the deep n-well, p-well and
n-well layers. The current consumption of the op-amp is defined by PMOS current sources, and common-mode feedback (CMFB) is used in the output stage to set the input and output common-mode voltages to 1 V. This common-mode voltage propagates back to the input of the passive mixer. Miller compensation was used to set the phase margin to 60 degrees. The TIAs were designed to consume 100 µA each from the 1.8 V supply. Using this op-amp, the overall flicker noise corner frequency was simulated to be around 100 kHz.

![Fig. 4.17 Schematic of the op-amp used in [5] for the IF section TIA.](image)

### 4.3.3 $G_m$-C Based CSF with Tuned Passive Mixer Output Pole

As shown in 4.2.3.3, an op-amp based CSF has the property that the interferers are filtered before being amplified by the op-amps’ nonlinear gains. This results in a great improvement in $I_{IP3}$. In [21], we showed that this is also possible in $G_m$-C filter implementations. An illustration of the op-amp CSF based method and the $G_m$-C based CSF method is shown in Fig. 4.18. It was earlier recognized that the switching transistors of the passive mixer can be represented as a variable resistor. When a shunt capacitor is connected to the output end of the variable resistor, a first-order low-pass filter is formed. In order not to disturb the corner frequency of this filter, the stage following the
passive mixer must have high input impedance. This naturally leads to the use of a $G_m$-C based CSF.

![Diagrams of mixers](image)

**Fig. 4.18** Comparison between (a) the active-RC based CSF-mixer and (b) the $G_m$-C based CSF-mixer.

### 4.3.3.1 Advantages

Compared to the active-RC approach, this method only works properly with a low-pass filter which implies that it works best with direct-conversion receivers. Like the active-RX approach, the $G_m$-C approach involves filtering before amplification. 1\textsuperscript{st}-order filtering of the IEEE 802.15.4 standard interferers is shown in Fig. 4.19. The result is a significant improvement in IF section $IIP_3$. For example, if we assume the two interferers are at 5-MHz offset and 10-MHz offset from the desired signal, they will be filtered by 14 dB and 20 dB respectively. Therefore, the improvement in $IIP_3$ is $14 + \frac{1}{2} \times 20 = 24$ dB. For 10-MHz and 20-MHz offset interferers, the improvement is 33 dB.
Another advantage of using the $G_m$-C approach which does not arise in the active-RC approach is that the first stage filtering requires no power consumption since it is passive. For a third order filtering function, only a single biquad is needed. In [21], the simple biquad shown in Fig. 4.20 was used. The overall filter was a 3rd-order low-pass Butterworth filter. It is easy to show that the DC gain of the filter is equal to $g_{m1}/g_{m2}$ while the corner frequency is equal to $C^{-1}\sqrt{(g_{m2}g_{m3})}$ and the $Q$ is equal to $\sqrt{(g_{m3}/g_{m2})}$. With four variables and three equations, we have one degree of freedom. This was used to select $g_{m1}$ to provide the desired overall noise performance of the receiver system. The individual transconductors are configured as simple differential pairs.

The use of $G_m$-C type filters also potentially results in significantly lower power consumption. The reason is that in active-RC filters, the filter function is only ideal as long as the loop gain is high. As the loop gain falls below zero, the feed-forward path starts to dominate the overall transfer function, and the gain becomes non-ideal. In order to overcome this, high unity-gain bandwidth op-
amps are essential and this puts a lower limit on the power consumption (in a cascaded filter, each stage must meet this requirement). In Fig. 4.18, the parasitic feed-forward path is evident in the active-RC filter implementation, but in the \( G_m \)-C implementation it is ideally absent. In a \( G_m \)-C implementation, gain-bandwidth is not a limitation for the proposed application.

![ Circuit Diagram ]

Fig. 4.20 The simple biquadratic filter used in [21] to design a 3\(^{rd}\) order Butterworth filter.

### 4.3.3.2 Disadvantages

The principle drawback in the proposed method is that the pole formed by the passive mixer and the output capacitor must be tuned. In general, the real pole formed by the switch resistance and the output capacitance is not used for filtering because of the considerable variation in the switch resistance. The switch resistance can vary due to variations in the LO voltage (\( V_{LO} \)), the switch threshold voltage, the switch size, and even the output impedance of the previous stage (the LNA output resistance affects the passive mixer output resistance [5]). In [21], a tuning loop was proposed which requires a replica of the passive mixer to be included in the design (Fig. 4.21). This causes a reduction in the impedance at the LNA output node potentially degrading the LNA gain. The replica passive mixers also cause additional loading to the frequency synthesizer.
4.3.3.3 The Tuning Loop

The principle of the tuning loop is illustrated in Fig. 4.22. The passive mixers are represented by variable resistors controlled by the LO signal. The output pole of the passive mixer consists of the resistance of the passive mixers and a bank of digitally controllable metal-insulator-metal (MIM) capacitors. A replica of the passive mixers without the output capacitors is also implemented. At the desired pole frequency, the real passive mixer will have a 3-dB lower output impedance than its replica.

Fig. 4.22  Figure to illustrate the principle of operation of the tuning loop. (a) The overall loop and (b) the digital amplitude comparator.
A signal at the desired pole frequency (1 MHz in this case) is fed into the passive mixers’ outputs via high output impedance transconductors which do not affect the passive mixers’ output impedances. The effect of the 3-dB lower output impedance of the real passive mixer is imitated on the replica side by a 3-dB attenuation of the 1 MHz tuning signal. The output amplitudes of the transconductors are then detected and compared. This signal is filtered and fed into a digital comparator. The output of the comparator drives a 6-bit counter which is connected back to the passive mixers output capacitor to close the loop. An effective design for a multiplier is a Gilbert Cell [14], however, this was modified to form a folded Gilbert Cell which is more suited to low-voltage operation. The proposed folded Gilbert Cell is shown in Fig. 4.23.

Referring to Fig. 4.22a, if the transconductor output on the real side is lower than that on the replica side, then the counter will count down in order to lower the capacitance, and vice-versa. The tuning scheme implemented in this work is rather primitive and is only designed to illustrate the potential of tuning the output pole of the passive mixer. In a more advanced implementation, a successive approximation architecture [17] for the loop would reduce the required tuning time significantly. Fig. 4.24 shows the response of the loop at the output of the loop filter. This is a differential signal and is therefore DC zero at steady-state.

In the proposed tuning loop, the final digital output will oscillate around the desired steady-state voltage with an amplitude of 1 least-significant bit (LSB).
Therefore the maximum error is less than 1 LSB. The loop-filter corner frequency is set so that the maximum ripple on the loop-filter output is less than ½ LSB. The loop-filter corner frequency establishes the settling time of each step, and therefore the maximum rate at which the loop can be clocked.

The ripple on the loop-filter output arises due to the squaring of the signals. A sinusoid, \( \sin(\omega t) \), when squared produces a DC component and an AC component at twice the input frequency, \( \sin^2(\omega t) = \frac{1}{2}(1-\cos(2\omega t)) \). As the signal at the real mixer side is 45 degrees out-of-phase with that at the replica mixer side, the input to the loop-filter is approximately equal to \( \sin^2(\omega t) - \sin^2(\omega t + \pi/4) = \frac{1}{2}(1-\cos(2\omega t)) - \frac{1}{2}(1-\cos(2\omega t + \pi/2)) = \cos(2\omega t) - \sin(2\omega t) \) when calibrated. Therefore, we could potentially have reduced the ripple by feeding in the real side reference signal at a phase shift of \( 3\pi/4 \). In that way, the input to the loop filter would be proportional approximately \( \sin^2(\omega t) - \sin^2(\omega t + 3\pi/4 + \pi/4) = 0 \). With less ripple, we could have used a higher loop-filter corner frequency and could potentially have sped up the loop. In this work we decided not to overcomplicate the design with the generation of octet phases.

![Fig. 4.24](image_url)  
*Fig. 4.24  Response of the tuning loop at the loop filter output (V) versus time (ms).*
4.3.4 **Summary of Architectures**

We have presented two different receiver architectures which involve integrating the passive mixer with the CSF: the active-RC filter approach and the $G_m$-$C$ filter approach. The active-RC approach is well suited to both low-IF and direct-conversion receivers, while the proposed $G_m$-$C$ approach is best suited for direct-conversion approaches. It was shown that both have the effect of filtering the desired signal after down-conversion, but before subsequent amplification. This results in significant improvement in $IIP_3$. For low power consumption, the $G_m$-$C$ approach was shown to have the advantages of requiring one less filter stage (in a cascaded implementation), and no gain-bandwidth limitation to the amplifiers. Unfortunately, the proposed $G_m$-$C$ approach requires a special tuning scheme.
References


10. IEEE 802.15.4 Standard For Local and Metropolitan Area Networks, 2003.


18. B. W. Cook, A. Berny, A. Molnar, S. Lanzisera, K. S. J. Pister, “Low-Power 2.4-GHz Transceiver With Passive RX Front-End and 400-mV


Chapter 5

Circuit Design

5.1 Introduction

In this chapter, we will look at circuit design methods which can be employed to minimize power consumption of front-end circuits. We will start our discussion by addressing the technology which is available to us. We will then look at matching networks and discuss their optimization. Next we will look at biasing point optimization. We will also look at different circuit arrangements for current reuse. Finally, we analyze the passive mixer to see how it interfaces with the low-noise amplifier.

5.2 Available Technology

In this section, we will explore the technology available to us. We will be working with the Global Foundries 0.18-µm RF CMOS technology. The technology features a triple-well process, with one poly layer, five standard thickness (0.55 µm) metal layers, and one thick top metal layer (2.52 µm). The technology allows for metal-insulator-metal (MIM) capacitors.

5.2.1 MOSFETs

A cross-sectional diagram of an n-type MOSFET is shown in Fig. 5.1(a). The gate is separated from the drain, source and channel by a dielectric which results in gate-source, gate-drain, and gate-channel capacitances. The gate-channel capacitance is the intrinsic capacitance and is normally split up into intrinsic gate-source and gate-drain capacitances in a ratio depending on the region of operation of the device [1]. In addition to the intrinsic capacitance,
there is an extrinsic capacitance due to the unavoidable overlap of the gate over the drain and source regions. This capacitance is fixed and independent of biasing conditions. Additionally, a depletion region forms between the P-type bulk and the N-type channel/source/drain. This results in source-bulk, drain-bulk and channel-bulk capacitances. Another important parasitic is the gate resistance. This resistance arises both due to the physical resistance of the gate material, and due to non quasi-static (NQS) effects [1]. A possible small signal model valid in the triode and saturation regions of a MOSFET is shown in Fig. 5.1(b) [2]. The intrinsic and extrinsic gate-source and gate-drain capacitances are grouped into $C_{GS}$ and $C_{GD}$. $C_{SB}$ and $C_{DB}$ represent the depletion region capacitances between the source/drain regions and the bulk. $r_{ds}$ represents the drain-source resistance which arises mainly due to channel-length modulation [2]. $r_g$ represents the gate resistance discussed above, and $g_m$ and $g_{mb}$ represent the transconductance and the body-effect transconductance [2]. More complicated models and systems of equations are normally used in circuit simulators [3], but simple ones such as that in Fig. 5.1 readily offer insight into the device parasitics and operation.

Some important parameters of the MOSFETs are the transit frequency, $f_T$, and $g_m/V_{DS}$. At RF, the intrinsic gain, $g_m r_{ds}$, is less important as RF amplifiers cannot be designed to reach their intrinsic gain with typical passive component $Q$ values. For matching purposes, it is also helpful to know the quality factor of the input impedance of the device. The threshold voltage of

![Cross-sectional diagram of an n-type MOSFET and its small signal equivalent model.](image)
the technology is 0.48 V. This limits the minimum supply voltage of the circuit designs. The transit frequency is useful as a measure of how fast a transistor can operate. $f_T$ is shown versus gate-source voltage, $V_{GS}$, in Fig. 5.2. In the subthreshold region, $f_T$ drops exponentially with $V_{GS}$. This is explained by the fact that the main device capacitance in the subthreshold region is the fixed extrinsic overlap capacitance. Since the device current varies exponentially with $V_{GS}$ in the subthreshold region, $f_T$ also naturally drops exponentially. This is one of the main reasons that subthreshold biasing is avoided in RF design.

![Graph showing $f_T$ versus $V_{GS}$](image)

*Fig. 5.2 Variation of $f_T$ with $V_{GS}$ in the 0.18 µm RF CMOS process.*

Another useful parameter is $g_m/I_{DS}$. This parameter gives us an idea of how efficiently a transistor converts an input voltage into an output current. $g_m/I_{DS}$ is shown for the 0.18-µm RF CMOS technology in Fig. 5.2. Clearly $g_m/I_{DS}$ improves as the device moves from strong inversion to weak inversion to subthreshold operation. From Fig. 5.2 there is a tradeoff between how efficiently a signal can be amplified, and the maximum frequency at which a signal can be amplified.
5.2.2 Inductors

Inductors are critical components in RFIC design for two main reasons. Firstly, they require a very large amount of die area. Not only are they physically large, but other components should not be placed within their vicinity to prevent unwanted magnetic coupling. Secondly, on-chip inductors typically exhibit low quality factor (about 5 – 10) at 2.4-GHz. This makes them the limiting factor in designs which use series or parallel LC tanks requiring very small series resistance or very large parallel resistance respectively. The Global Foundries 0.18 μm RFCMOS process design kit (PDK) which we are using offers spiral inductors with polysilicon patterned ground shields. The patterned ground shield prevents coupling to the lossy substrate while not allowing eddy currents to build up in the shield itself [4]. The top metal layer is 2.52 μm thick resulting in a low sheet resistance. Lastly, the PDK offers center tapped inductors which are useful in differential design. Center-tapped inductors behave like two inductors which are magnetically coupled to each other. The mutual coupling effectively results in a boost in the total inductance [5] which can be calculated as,

\[ L = 2L_0(1 \pm k) \]  

(5.1)

where \( L_0 \) is the inductance of each individual coil, and \( k \) is the coupling coefficient. The plus sign holds for differential signals while the minus sign holds for common-mode signals. An interesting use of this effect is that if the LC tank is designed to resonate at the operating frequency for differential signals, then it will resonate at a much higher frequency for common-mode signals. This gives the LC tank a measure of common-mode rejection.

5.2.3 Other Components

The MIM capacitors are formed between the metal 5 layer and a fuse top layer. The cross-sectional view is shown in Fig. 5.3. With the thin dielectric layer, the MIM capacitors have a high capacitance density of 1 fF/µm². The metal 5 layer forms an additional capacitance to ground which is approximately 1.5 % of the desired capacitance. This is called the bottom plate capacitance and is important when designing high-pass AC coupling
circuits. The PDK offers various resistor options including high sheet resistance N+ Poly resistors (1320 Ω/sq) and good performance (tolerance, matching, temperature variability) P+ Poly resistors.

Lastly, the PDK offers vertical NPN bipolar junction transistors (BJT) formed between the n-well layer, the p-well layer and the deep n-well layer. NPN transistors provide better matching, DC-offset and flicker noise performance than MOS devices, but in this process, have significantly lower $f_T$, potentially making them unsuitable for RF design. Nevertheless, these BJTs have found use in IF circuitry as well as RF mixers [6], [7]. It is worth mentioning that BJTs also have significantly higher $g_m/I_{DS}$ than MOSFETs ($1/V_t \approx 40 \text{ V}^{-1}$). A cross-sectional view of a VNPN BJT is shown in Fig. 5.4.

5.3 Impedance Matching

RFIC designers are generally interested in boosting voltage levels (as opposed to power levels) in order to reduce a signals sensitivity to noise. This can be done in two basic ways: either the power of the signal can be amplified, or the impedance level can be increased. In integrated RFICs, the
impedance level of different nodes may vary drastically. Generally speaking, the reference impedance at the input of the receiver is equal to 50 Ω. In order to save power, the impedance level is normally stepped up in the RF front end. However, the amount by which it can be stepped up usually depends on the quality factor of the parasitic components, and the bandwidth requirements (see Section 2.3.4). In the IF section, the reference impedance level can be considerably higher as the bandwidth requirements are reduced.

5.3.1 Where (and why) is matching done?

Despite the different impedance levels, in monolithic designs, typically only a single impedance transform network is used and is located at the interface between the off-chip source (antenna and filtering), and the on-chip LNA. A circuit diagram illustrating the different impedance levels and matching is shown in Fig. 5.5. Starting from the left hand side, we note that $r_{\text{source}}$ is not matched to $Z_{\text{in}1}$. Rather, a matching resistance is deliberately introduced specifically for matching (In Fig. 5.5, we have assumed $r_{\text{match}}$ is a simple resistor, but later we will show various ways of producing a virtual $r_{\text{match}}$). It is not possible to match directly to $Z_{\text{in}1}$ because the real part of $Z_{\text{in}1}$ is too small requiring unachievable quality factor values in the passive components. We can consider the input network of the input transistor as a capacitor with a quality factor of approximately [8],

$$Q_{\text{GS}} = \frac{5g_{d0}}{\omega_0 C_{\text{GS}}} = \frac{5\alpha r}{\alpha \omega_0}$$

(5.2)
where $g_{d0}$ is the channel conductance when the drain-source voltage is zero, $\omega_0$ is the operating frequency in radians per second, $\omega_T$ is the transit frequency in radians per second, and $\alpha$ is equal to the ratio of $g_m$ to $g_{d0}$ (approximately equal to 1). At a frequency of 2.5 GHz with an $f_T$ of just 20 GHz, the quality factor would be close to 40. We have already mentioned that on-chip inductors can only achieve $Q$ values of around 5-10.

Likewise, we note that $Z_{01}$ is not matched to $Z_{in2}$. In this case, even if it were possible to match the impedances, we would not do so as having $Z_{in2} << Z_{01}$ serves to improve the stability of the design (It also reduces the miller multiplication of $C_{GD}$ of the input transistor, and makes matching easier). Again, $Z_{02}$ cannot be matched to $Z_{inmix}$ as the $Q$ of $Z_{02}$ is significantly higher than that of $Z_{load}$ (therefore, $Z_{load}$ would modify the matching condition). Instead, the output node of the LNA is normally resonated. $Z_{0mix}$ cannot be matched to $Z_{inIF}$ because of the low frequency of operation. In practice, matching is not done at the IF in integrated designs even when the IF frequency is in the range of hundreds of MHz.

In summary, we normally only need to be concerned with the impedance matching at the input of the receiver, and we need to find a way to create a virtual $r_{match}$. This leads to an interesting question. The purpose of impedance matching is to maximize the power delivered from a source to a load.
However, since we are not really matching to $Z_{in}$, we are not maximizing the power to the input device and therefore, why do we need to do impedance matching? It is entirely possible that the network which leads to the optimum power transfer to the input device may not use impedance matching at all. Let us look, for instance, at the equivalent input impedance representation of an inductively degenerated LNA (Fig. 5.6) [8].

From the design values in Fig. 5.6, at resonance we can calculate the power which is dissipated in the input transistor’s input network as,

$$P_{in,ms} = \frac{2V_s^2 r_g}{(r_{source} + r_{match} + r_g)^2}$$

(5.3a)

Clearly, $P_{in,ms}$ is maximized by letting $r_{match}$ tend towards zero. This is equivalent to removing $L_s$ and resonating $L_g$ with $C_{gs}$. Obviously we would incur a bandwidth penalty (see section 2.3.4), and the gate inductor, $L_g$, would have some parasitic resistance anyway, but certainly we would still achieve more power transfer to the input. In fact, avoiding input matching can even lead to better $NF$ [9]. The truth is that impedance matching at the input is not always necessary [9], [10], [11]. However, one concern is that off-chip components (in particular filters) are designed to be impedance matched, and without impedance matching, could exhibit unpredictable performance. If input impedance matching is not used, the designer should carefully model the input interface to ensure proper operation. In this work, we will always match the input to $50 \ \Omega$. 

Fig. 5.6 (a) inductive degeneration matching and (b) its equivalent representation.
In terms of gain maximization, we are more concerned with the AC current supplied by the device than the power delivered to the device. Therefore, (5.3a) can be modified as,

\[
I_{\text{out}} = \frac{2g_m V_s \left( r_g + \frac{1}{sC_{\text{GS}}} \right)}{r_{\text{source}} + r_{\text{match}} + r_g}
\]

(5.3b)

where \( I_{\text{out}} \) is the current leaving the device. Clearly the overall transconductance is also maximized when \( r_{\text{match}} \) is minimized.

We have shown that since we are not matching to the input device itself (the MOSFET), we do not achieve maximum power transfer to the device. A side-effect of this is that we may end up with an unwanted frequency response. We will see that a typical L-match can be designed as a low-pass, a high-pass or even a band-pass filter (in a two-stage L-match). However, this filtering function only applies to the load, \( r_{\text{match}} \). The power transferred to the device may have an entirely different frequency response. In [12] for example, the power transferred to \( r_{\text{match}} \) was broadband and band-pass, however, the power transferred to the device was low-pass. Although this frequency response was compensated for at the load, the \( NF \) was distinctly higher at the higher end of the desired band.

5.3.2 Creating \( r_{\text{match}} \)

From the discussion above, a MOSFET’s gate resistance is too small to be matched to at low GHz range frequencies. Therefore, \( r_{\text{match}} \) must be synthesized. Going back to section 2.3.4, we note that for narrowband amplifiers, we can create a much larger value of \( r_{\text{match}} \) than in broadband amplifiers. This is a simple result of the Bode-Fano criterion which we repeat here. Given a load consisting of a either a parallel or series resistor and capacitor/inductor,

\[
\int_0^\infty \ln \left| \frac{1}{\Gamma(\omega)} \right| d\omega \leq \begin{cases} 
\frac{\pi}{RC} & \text{parallel RC} \\
\frac{\pi R}{L} & \text{series RL}
\end{cases}
\]

(5.4a)
where $\omega$ is the frequency in radians/second. This equation tells us that the better the matching, the smaller the bandwidth over which this level of matching can be achieved. Since the load $Q$ is equal to $1/(\omega_0 RC)$ (series RC case), (5.4) also tells us that the achievable matching bandwidth is inversely proportional to load $Q$. It is also of note that since the device input $Q$ drops with increasing operating frequency, at millimeter-wave frequencies $r_{\text{match}}$ does not need to be synthesized. In ultra-wideband systems, $r_{\text{match}}$ is generally created to be equal to the source resistance. As mentioned in Section 2.3.4, the tradeoff is voltage gain. Larger bandwidth results in smaller voltage gain. A more rigorous investigation would look at the power gain of the circuit. We will look into this when we discuss biasing point optimization. For now, let us look at several common techniques which can be used to synthesize $r_{\text{match}}$. Naturally, other ways to synthesize $r_{\text{match}}$ are conceivable. For the following derivations, we will assume that the $r_g$ is absorbed into $r_{\text{match}}$.

### 5.3.2.1 Resistive termination

Resistive termination is the simplest way to create $r_{\text{match}}$. A resistor is simply added either in shunt or in series with the input transistor as shown in Fig. 5.7. If $r_{\text{match}}$ is designed to be equal to $r_{\text{source}}$, then matching can be achieved in the shunt case by simply adding parallel resonance of $C_{GS}$, and in the series case by series resonance of $C_{GS}$. Despite how simple this matching network appears, it is actually quite useful in low-power narrowband designs [10], [13]-[16]. In such designs, the parasitic resistance of the on-chip inductors is normally absorbed into $r_{\text{match}}$. For shunt termination, $r_{\text{match}}$ is selected to be large compared to $r_{\text{source}}$ while in series termination, $r_{\text{match}}$ is normally selected to be 50 Ω. There are two main drawbacks with this method. Firstly, the tolerance of $r_{\text{match}}$ can only be made to be around 20 %, and secondly, the minimum possible $NF$ with this matching network is 3 dB. It is the 3-dB minimum $NF$ that normally draws designers away from resistive termination, but in cases where the overall system $NF$ is high, a 3-dB LNA $NF$ may be
acceptable. In narrowband designs, this design requires only a single on-chip inductor.

![Diagram of resistive termination in (a) shunt and (b) series form.](image)

**Fig. 5.7** Resistive termination in (a) shunt and (b) series form.

### 5.3.2.2 Common-Gate Matching

An $r_{\text{match}}$ synthesis method which is very popular in wideband LNA design is common-gate matching. A diagram of the method is shown in Fig. 5.8(a). It is easy to show that $r_{\text{match}}$ is equal to $1/(g_m + g_{mb})$. $I_1$ is normally replaced by an inductor at RF [5]. It can be shown that the minimum $NF$ of this matching scheme is $1 + \gamma/\alpha$ [8], which is approximately equal to 2.2 dB for long-channel devices ($\gamma$ is approximately equal to 2/3 for long-channel devices and $\alpha$ is approximately equal to 1). $r_{\text{match}}$ can be well controlled by matching the device transconductance to an off-chip resistor. The biggest problem with this architecture is that under matched conditions, the overall transconductance is fixed. This is easily seen since under matched conditions, the source will see a load resistance equal to $r_{\text{source}}$. Therefore, the input current will equal to $V_{\text{source}}/r_{\text{source}}$. The only path for the current to flow is directly through the common-gate transistor and it manifests as the output current, $i_{\text{out}}$. Hence $i_{\text{out}}/V_{\text{source}}$ is limited to $1/r_{\text{source}}$ under matched conditions.

In differential designs, this can easily be overcome as shown in Fig. 5.8(b). This type of input network also has some interesting noise cancellation properties, although the actual $NF$ achieved in noise cancellation designs are rarely spectacular [17], [18].
The common-gate architecture enjoys the additional benefit of being more linear than a common-source architecture. Assuming $r_{\text{match}}$ is designed to be 50 $\Omega$, it is easy to see that,

$$\frac{i_{\text{out}}}{V_{\text{source}}} \approx \frac{2(g_m + g_{\text{mb}})}{1 + r_{\text{source}}(g_m + g_{\text{mb}})}$$ (5.5)

Fig. 5.8 (a) Common-Gate $r_{\text{match}}$ synthesis and (b) de-coupling transconductance and $r_{\text{match}}$ in a differential design.

### 5.3.2.3 Inductive Degeneration

Inductive degeneration is one of the most popular techniques [19]-[21] used to create $r_{\text{match}}$ owing to its ability to achieve near concurrent impedance and noise matching [21]. A figure depicting inductive degeneration matching is shown in Fig. 5.9. It is relatively easy to show that the input impedance can be written as [8],

$$Z_{\text{in}} = sL_s + \frac{1}{sC_{GS}} + \frac{g_m}{C_{GS}}L_s \approx sL_s + \frac{1}{sC_{GS}} + \omega r L_s$$ (5.6)

This formula led to the small signal representation in Fig. 5.6(b). $r_{\text{match}}$ is approximately equal to $\omega r L_s$. The simplest way to implement the matching network in narrowband designs is to let $r_{\text{match}}$ equal to 50 $\Omega$ and add an inductor in series with the network for series resonance. It should be noted that the quality factor of the series resonant network is limited to $1/(\omega_0 C_{GS} r_{\text{match}})$. Therefore, this may not be the optimum matching network at
millimeter-wave frequencies. As this matching method uses negative feedback, the transconductance is more linear than in the resistive termination case. It is easy to see that at the operating frequency, the overall transconductance is equal to,

\[
\frac{i_{\text{out}}}{V_{\text{source}}} \approx \frac{2g_m}{\omega_0 C_{GS} \left( r_{\text{source}} + \frac{g_m L_s}{C_{GS}} \right)}
\]  

Equation (5.7) clearly shows that \(g_m\) is linearized. Inductive degeneration also offers the possibility of nearly concurrent impedance and noise matching. The idea of noise matching is to find the source impedance which minimizes the input-referred noise of the design. It is easily seen from [8] that if gate-induced noise is neglected, the optimum source impedance is \(-j\omega C_{GS}\), where \(\omega\) is the frequency in radians per second. At the operating frequency, this simply means that the source should be transformed from 50 \(\Omega\) to as high a resistance as possible, while resonating with \(C_{GS}\). In low power designs, the LNA’s gate-induced noise is rather insignificant compared to the rest of the noise contributed by the system and we can safely ignore noise matching. In designs where achieving the minimum possible \(NF\) is critical, noise matching may be a useful concept to apply.

A last point which should be made is that for low-power designs, it is imperative to maximize the voltage gain of the matching network (as long as bandwidth requirements are met) in order to minimize the noise contribution of the following stages [10], [13]. As a result, the gate inductor, \(L_g\) (Fig. 5.6), normally needs to be quite large. With typical low \(Q\) on-chip inductors, the series resistance of \(L_g\) can often be so large that \(L_s\) is not even required for
proper matching. For instance, for a voltage gain of 12 dB, we need a load $Q$ of 4, which means that a 50-Ω series resistance would require a series reactance of $j200$ Ω. With a $Q$ of 6, the inductor would already contribute 33.3 Ω of series resistance which is equivalent to a reflection coefficient, $\Gamma_{11} = -14$ dB. This value of $\Gamma_{11}$ is normally sufficient for impedance matching, and therefore, $L_s$ is not necessary.

5.3.3 LC Impedance Matching

Having created $r_{\text{match}}$, we need to design an LC network to transform the source impedance to $r_{\text{match}}$. We could also use a transformer, however this requires additional modeling work as transformers are not included in the Global Foundries 0.18 µm RFCMOS PDK. Inductors and capacitors do not dissipate energy, they only store energy. Therefore, ideal LC matching networks are lossless. As discussed in section 2.3.4, a direct result of this is that all energy must be dissipated in either the source, or the load resistance. Hence, in a matched case, when transforming from $r_{\text{source}}$ to $r_{\text{match}}$, we will have a voltage gain equal to,

$$\frac{V_{\text{match}}}{V_{\text{source}}} = \sqrt{\frac{r_{\text{match}}}{r_{\text{source}}}}$$

(5.8)

where $V_{\text{match}}$ is the voltage across $r_{\text{match}}$. The simplest matching network is a simple L-match. An L-match can be designed as either a high-pass or a low-pass network. An L-match is shown in Fig. 5.10 where $r_{\text{large}}$ is larger than $r_{\text{small}}$. The easiest way to understand the L-match is to transform the parallel networks into series ones. For instance, in 5.10 (a), $Q = \frac{r_{\text{large}}\omega_0}{C}$. We need to have $\frac{r_{\text{large}}}{Q^2} = \frac{r_{\text{small}}}{}\text{under matched conditions. However, we should also remember that matching bandwidth trades with the degree of matching (see 5.3.2). Once } r_{\text{large}} \text{ and } r_{\text{match}} \text{ are decided, } Q, L \text{ and } C \text{ and hence matching bandwidth are fixed.}
We can reduce the matching bandwidth by using a T match or a π match [22]. Such matching networks essentially place step-up and step-down matching networks back to back. Since a step-up is always used with a step-down, the virtual impedance in between the networks is always smaller than \( r_{\text{small}} \) or \( r_{\text{large}} \) resulting in higher \( Q \) requirements for the matching networks. T matches and π matches are shown in Fig. 5.11. Alternatively, we can increase the bandwidth of a matching network by cascading step-up with step-up or step-down with step-down matching networks [23], however more than one inductor is generally required. In wideband LC matching networks, the virtual resistance, \( r_{\text{virtual}} \), in between the L-matches is a value in between \( r_{\text{large}} \) and \( r_{\text{small}} \). For a two-stage matching network, the value of \( r_{\text{virtual}} \) which results in the widest matching bandwidth is equal to \( \sqrt{(r_{\text{large}}r_{\text{small}})} \) [23]. As die area is a concern, we avoid the use of high-order matching networks.
In the differential implementation of the matching network (Fig. 5.10(d)), we note that the high-pass network has a distinct advantage: only one inductor is required. This is important due to the large die area requirement of monolithic inductors. Another important benefit is that the parallel parasitic resistance of the inductor is easily absorbed into $r_{\text{large}}$. Lastly there is a boost in the effective parallel inductance due to the mutual magnetic coupling between the coils (section 5.2.2). This allows us to use a physically smaller inductor.

5.3.4 Impedance Matching Summary

There are a few points worth reiterating when it comes to impedance matching. Firstly, it is generally not possible to match directly to a MOSFET’s input impedance at low frequencies. Therefore, we must create a resistance to match to, and as a result, impedance matching does not result in maximum power transfer to the input device. Of the different methods to create $r_{\text{match}}$, resistive termination allows for high voltage gain while requiring a minimum number of passive components. Inductive degeneration results in good linearity, noise performance and voltage gain, but requires an additional inductor as compared to resistive termination, and is generally not suitable in
designs which require high voltage gain matching networks (low-power designs). LC matching networks can be designed as high-pass, low-pass and band-pass (second order or higher matching networks) networks. The high-pass L-match is particularly useful in differential designs.

5.4 Biasing Point Optimization and Weak Inversion Biasing

At the circuit level, we have paid particular attention to weak inversion biasing. We have already demonstrated in Section 5.2.1 that weak inversion biasing results in better $g_m/I_{DS}$ than strong-inversion biasing, although it comes at the price of poorer $f_T$. In this section, we will also look at the noise performance and linearity of a device at different biasing levels. We will also look at the maximum unilateral power gain of a composite device, and show why this is a useful figure of merit.

5.4.1 Noise Performance

As power consumption is the critical limitation in our designs, we are interested in comparing the noise performance of a design at different biasing points for the same power consumption. We can compare the biasing points using the parameter $NF_{\text{min}}$, which is the minimum possible noise figure achievable by the device. The value $NF_{\text{min}}$ is achieved at an optimum source impedance. Fig. 5.12 shows the circuit setup for how $NF_{\text{min}}$ was simulated. Essentially, at DC we have a wide-swing current mirror which allow good reproduction of the current, $I$, through $M_1$. All capacitors and inductors are large and are used for noise isolation. The drain-source voltage of $M_1$ is held constant as $M_3$ and $M_4$ are essentially source followers for $V_b$. As the device sizes are changed, the gate-source biasing voltage of $M_1$ will change. Also of note is the resistors in parallel with the ports, $R_{\text{max}}$. These resistors were added to represent the maximum impedance which an impedance matching network can step up to. $R_{\text{max}}$ of 250 $\Omega$, 500 $\Omega$, and 1 k$\Omega$ correspond to voltage gains of 7dB, 10 dB, and 13 dB respectively, for the impedance matching network. In order to sweep $V_{GS}$ while keeping the drain current $I_{DS}$ and the
drain-source voltage, $V_{DS}$, constant, the number of fingers of the transistors must be swept.

Fig. 5.12 Schematic of $NF_{\text{min}}$ testbench. The op-amps are ideal.

Fig. 5.13 shows the variation of $NF_{\text{min}}$ with $V_{GS}$ for different biasing currents and values of $R_{\text{max}}$ at a frequency of 2.4 GHz. There are several things to note from the curves. Firstly, the curves are discontinuous at a very low value of $V_{GS}$. This is a limitation of the device models. Better $NF_{\text{min}}$ is achievable for higher values of current and $R_{\text{max}}$. Lastly, the optimum value of $V_{GS}$ appears to be around 0.52 V. This is quite close to the threshold voltage of the device which is 0.48 V. $NF_{\text{min}}$ is directly related to the input-referred noise of the device. The input-referred noise due to the channel thermal noise can be written in terms of an input referred voltage noise as,

$$
\overline{V_n^2} = \begin{cases} 
\frac{2kT\gamma(V_{gs} - V_T)}{I_{ds}} \frac{4kT\gamma}{g_m} & : \text{strong inversion} \\
\frac{2kTn^2V_t}{I_{ds}} \frac{2kTn}{g_m} & : \text{subthreshold}
\end{cases}
$$

(5.9)

where $V_t$ is the thermal voltage (approximately 25 mV), and $n$ is the subthreshold slope (equal to 1.5 in this process). From (5.9), we can observe that in the strong inversion region, with $I_{ds}$ fixed, the input referred noise drops with the overdrive voltage, $V_{GS} - V_T$, while in the subthreshold region, the
input referred noise is constant. In simulation, $NF_{\text{min}}$ reaches a minimum value and then starts to increase in the subthreshold region. This is not due to an increasing amount of noise, but due to the reduced $Q$ of the input impedance.

We note that the series gate resistance of the device is approximately $1/(5g_{d0})$ (it is normally larger due to the physical layout of the device). Representing this as a resistance in parallel with $C_{GS}$, it is approximately equal to $5g_{d0}/\omega^2C_{GS}^2$. Noting that the ratio $g_{d0}/C_{GS}$ is roughly proportional to the $f_T$ of the device, we can conclude that the equivalent parallel resistance of the device decreases as the device goes from strong inversion to subthreshold operation. In the strong inversion region, the effective resistance in parallel with the device (from Fig. 5.12) is limited by $R_{\text{max}}$. However in the weak inversion region, it is limited by $r_g$. As a result, the maximum voltage gain achievable by the input matching network is reduced in the weak inversion region leading to poorer $NF_{\text{min}}$. This effect can be alleviated by increasing $R_{\text{max}}$.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figures.png}
\caption{\textit{NF}_{\text{min}} for different biasing currents and values of $R_{\text{max}}$ versus $V_{GS}$ at 2.4 GHz.}
\end{figure}

### 5.4.2 Linearity Performance

In Section 4.2.1.2, we discussed the possibility of controlling a circuit’s $IIP_3$, and we also showed how the $IIP_3$ varies with $V_{GS}$. In general, it is true that the $IIP_3$ of a device is poorer in the subthreshold region than in the strong-inversion region. However, we must remember a few things. Firstly, the $IIP_3$ of the front-end LNA is generally less critical than the $IIP_3$ of amplifying blocks.
further down the receiver chain. This is because the input to the receiver is at its smallest level at the LNA input. Secondly, in the subthreshold region, the IM₃ signal is opposite in sign as compared to the IM₃ signal in the strong-inversion region (see Section 2.3.6). Therefore, if a subthreshold biased device is cascaded with a strong-inversion biased device, IM₃ cancellation is possible. Lastly, the $IIP₃$ requirements of the IEEE 802.15.4 standard are extremely relaxed compared to the noise requirements. This is evident from Table 4.1 in Section 4.3.1 which shows that 5 out of 7 recently published low power low data-rate receivers achieved significantly higher $IIP₃$ limited sensitivity than noise limited sensitivity. Therefore, high voltage gain and low power consumption are a higher priority than high linearity in IEEE 802.15.4 standard designs.

Fig. 5.14 shows the $IIIV₃$ (input-referred third-order intercept voltage) versus $V_{GS}$. We have used the same setup as in Fig. 5.12, except that the output port has been replaced by a short to ground, and the output current was measured. The $OIC₃$ (output third-order intercept current) was referred back to the gate voltage to find the $IIIV₃$. We have used 1 mA of drain current while sweeping the device width.

![Graph showing $IIIV₃$ versus $V_{GS}$ for a MOSFET.]

*Fig. 5.14 $IIIV₃$ versus $V_{GS}$ for a MOSFET.*
There are a couple of points to remember when analyzing this graph. Firstly, the y-axis is in terms of dBV which when referred to a 50-Ω resistance, is equivalent to $IIP_3 - 10$ in dBm (e.g. 1 V = 0 dBV = 10 dBm into 50 Ω). At the threshold voltage, the $IIP_3$ would be approximately -5 dBm. In order to achieve an overall -15 dBm $IIP_3$, we could use an impedance matching voltage gain of no more than 10 dB. Given that the $IIP_3$ requirement for the IEEE 802.15.4 standard is around -30 dBm (Chapter 2, Section 2.5.4.2), it is fair to say that weak-inversion biasing is a viable option for LNA design in terms of linearity performance.

### 5.4.3 Power Gain

In Section 5.2.1, we discussed how $g_m/I_{DS}$ trades with $f_T$ as a device goes from strong-inversion to weak-inversion biasing. Although it was evident that an LNA design could benefit from weak-inversion biasing, a more rigorous proof is needed to show which biasing point is truly optimal for LNA design. This has led us to explore the use of $G_{TU\text{max}}$ as an appropriate figure of merit. $G_{TU\text{max}}$ is the maximum transducer gain under the assumption that the device is unilateral, which is to say that no power is fed back from the output to the input. This is equivalent to saying that $y_{12} = 0$ [24]. $G_{TU\text{max}}$ can be calculated as,

$$
G_{TU\text{max}} = \frac{|y_{21}|^2}{4 \text{Re}\{y_{11}\}\text{Re}\{y_{22}\}}
$$

(5.10)

Finding $G_{TU\text{max}}$ is equivalent to finding the transducer gain, setting $y_{12}$ to zero, and applying complex conjugate matching at the input and output of the device. In practice, making a device unilateral is almost always done as it ensures stability, reduces the miller multiplication of $C_{GD}$, and simplifies input matching. Perhaps the most common technique used to make a circuit unilateral is to simply add a cascode transistor [25]. However, other techniques exist which serve to neutralize the gate-drain capacitance of the input MOSFET [25]. In order to estimate the $G_{TU\text{max}}$ of a MOSFET, let us use the small signal model in Fig. 5.1(b). Once again, let us add $R_{max}$ to represent the maximum possible $Q$ factor of the input and output resonant networks.
Furthermore, we have replaced $r_{ds}$ at the drain by $R_{\text{max}}$ for simplicity. Since we are using a common-source routing, $V_{BS} = 0$. This leads to Fig. 5.15. It is easy to show that,

$$\text{Re}\{y_{11}\} = \frac{1}{R_{\text{max}}} + \omega_0^2 (C_{GS} + C_{GD})^2 r_g \approx \frac{1 + g_m^2 R_{\text{max}} r_g}{R_{\text{max}}} \frac{\omega_0^2}{\omega_T^2}$$  (5.11)

![Fig. 5.15 Small signal model for estimation of $G_{T\text{Umax}}$ of a MOSFET.](image)

At low frequencies, $y_{11} = y_{22} = 1/R_{\text{max}}$, $y_{21} = g_m$, and $G_{T\text{Umax}}$ is equal to $g_m^2 R_{\text{max}}^2 / 4$, but at a certain frequency, which we will call the first pole frequency, the power gain will drop by 3 dB. From (5.11), this frequency is approximately equal to $(r_g R_{\text{max}} (C_{GD} + C_{GS})^2)^{-1/2}$ or $\omega_T [g_m \sqrt{(R_{\text{max}} r_g)]}$. This leads to the approximation in equation (5.12), which is good for low frequencies.

$$G_{T\text{Umax}} = \frac{g_m^2 R_{\text{max}}^2}{4 \left(1 + g_m^2 R_{\text{max}} r_g \frac{\omega_0^2}{\omega_T^2}\right)}$$  (5.12)

Also, it is important to remember that $r_g$ is not a fixed value, but is approximated as $1/5g_{d0} = \alpha/5g_m$ (this value can be larger due to the physical layout of the device) [8]. Therefore, at the pole frequency we can expect a $G_{T\text{Umax}}$ of no more than $5ag_m R_{\text{max}} \omega_T^2 / 8 \omega_0^2$. Looking back at Fig. 5.2, we see that $g_m$ and $\omega_T$ vary inversely with each other. Therefore, there is an optimum biasing point for the highest possible $G_{T\text{Umax}}$, which changes with frequency. $R_{\text{max}}$ should also be optimized bearing in mind that it trades with bandwidth. Fig. 5.16 plots $G_{T\text{Umax}}$ versus frequency for different values of the gate-source (and drain-source) voltage. At a frequency of 2.4 GHz, the optimum setting for $V_{GS}$ appears to be slightly less than 515 mV which is quite close to the
threshold voltage (480 mV). Looking back at (5.15), we can see that without $R_{\text{max}}$, the pole and maximum DC power gain created by $y_{11}$ would not exist. This is illustrated by the solid curves in Fig. 5.16. Therefore, plotting $G_{\text{TUmax}}$ without $R_{\text{max}}$, we would have wrongly assumed that $\omega_T$ should be maximized for maximum $G_{\text{TUmax}}$.

![Fig. 5.16 Scatter plots represent $G_{\text{TUmax}}$ versus frequency for different biasing voltages. $R_{\text{max}}$ is equal to 500 $\Omega$. Solid curves represent $G_{\text{TUmax}}$ when $R_{\text{max}}$ is removed.](image)

Although $G_{\text{TUmax}}$ is a measure of power gain, it is easily related to voltage gain through the impedance level.

### 5.4.4 Summary

We have shown that the optimum biasing point for the input transistor is in the weak-inversion region where $G_{\text{TUmax}}$ rather than $\omega_T$ is maximized. Weak-inversion biasing apparently leads to optimum gain and from section 5.4.1, optimum noise figure. The linearity is generally good enough to meet the demands of the IEEE 802.15.4 standard.
5.4.5 A Weak-Inversion Biased LNA [15]

One of our early works was to establish the usefulness of weak-inversion biasing in LNA design for the IEEE 802.15.4 standard [15]. Fig. 5.17 shows a schematic diagram of the fabricated LNA, and the micrograph is shown in Fig. 5.18. There are several points to note in the proposed design. Firstly, it was recognized that for the frequency of operation (2.4 GHz), the output pole introduced at the load could be made high enough that we could avoid a load LC tank. This was done by biasing the cascode transistor in the strong-inversion region (high $f_T$). Therefore, only a single inductor was required in this design. As previously argued, inductive degeneration is not the most suitable design method for low-power designs. We used a series resonant matching network where $r_{\text{match}}$ was created by the parasitic resistance of the gate inductor. Given the 11.3 nH gate inductance, we can see that the input $Q$ is approximately 3.5 which results in a 10.9 dB voltage gain due to the input matching network.

![Passive Mixer](attachment:image.png)

**Fig. 5.17 Schematic of a weak-inversion biased LNA.**
5.4.5.1 Measurement Setup

The LNA was loaded by a pair of single-balanced passive mixers biased in an average LO condition in order to accurately reflect the load conditions. The NF was measured from Output 2, and this is possible since the output port does not contribute to the measured NF. NF is a measure of the ratio of the total output noise to the output noise contributed by the input port. Since the output port does not affect this ratio, we do not need to worry about the output port loading down the node Output 2. In order to measure gain accurately, Output 2 is left open, and Output 1 is loaded by the output port. An on-chip resistor divider (which can provide excellent matching) was used to provide good output reflection coefficient. This obviates the need for an output buffer to drive the 50-Ω output port. Although the ratio of the 900-Ω and 50-Ω resistances is accurately known, the actual value of the two resistances needs to be calculated from,

\[ R_{\text{out}} = 50 \cdot \frac{1 + S_{22}}{1 - S_{22}} \]  

(5.13)
where $R_{out}$ is approximately equal to the true resistance of the 50-Ω resistor. This is true because looking back from the Output 1, the 900-Ω resistor is in series with the output impedance of the cascode transistor (which is large).

5.4.5.2 Voltage Gain and $S_{11}$

Fig. 5.19 shows the voltage gain and $S_{11}$ of the LNA. In simulation, we expected an increase in the parasitic capacitance and this proved true resulting in excellent matching at 2.4 GHz (-19 dB from 2.4 to 2.5 GHz). We can see that the -10 dB $S_{11}$ matching bandwidth is very broad making it robust to process variation (2.2 GHz to 2.8 GHz). The same is true for the voltage gain.

**Fig. 5.19 Voltage Gain and $S_{11}$ of the proposed LNA.**

5.4.5.3 Noise Figure

Fig. 5.20 shows the measured $NF$ of the proposed LNA. The measured $NF$ is optimum around 2.4 GHz, at around 5 dB. This is sufficient for the proposed application. Unfortunately, there was a somewhat large increase in the $NF$ from simulation to measurement. Given the accuracy of the rest of the performance, we suspect that the noise discrepancy is due to unreliable noise modeling in the weak-inversion biasing region. The rapid increase in $NF$ at
high frequencies is due to the low-pass filtering effect of the input matching network.

![Graph showing noise figure (NF) vs. frequency for the proposed LNA.](image)

Fig. 5.20 Measured NF of the proposed LNA.

### 5.4.5.4 Overall Discussion

Table 5.1 shows a comparison between the proposed LNA and some state-of-the-art LNA designs at the time of publication. As we can see, the proposed LNA design offered one of the lowest power consumptions among competing designs. Furthermore, with only a single inductor, the design can save considerable die area. Most importantly, the design was able to achieve a very wide matching bandwidth, with high gain while consuming only 630 µA from the 1.8 V supply. Our final LNA design [30] is very similar to this one, except that the load resistor was replaced by a load inductor, and the supply voltage was reduced to only 1 V. The design in [30] also featured variable power/gain control and the ability to manually tune the frequency selection networks. This proved necessary due to the questionable modeling of the on-chip inductors.
Table 5.1

Comparison of Proposed LNA and Current Literature

<table>
<thead>
<tr>
<th>Specification</th>
<th>This Work [15]</th>
<th>[5]$^a$</th>
<th>[26]</th>
<th>[27]</th>
<th>[28]</th>
<th>[28]</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC Current (µA)</td>
<td>630</td>
<td>1200</td>
<td>4400</td>
<td>230</td>
<td>670</td>
<td>260</td>
</tr>
<tr>
<td>Voltage Gain (dB)</td>
<td>21.4</td>
<td>33</td>
<td>18</td>
<td>13.6</td>
<td>9.1</td>
<td>4.5</td>
</tr>
<tr>
<td>$S_{11}$ (dB)</td>
<td>-19</td>
<td>-</td>
<td>-</td>
<td>-5</td>
<td>-13</td>
<td>-13</td>
</tr>
<tr>
<td>$P_{1\text{dB}}$ (dBm)</td>
<td>-15</td>
<td>-</td>
<td>-13</td>
<td>-0.2</td>
<td>-25</td>
<td>-19.5</td>
</tr>
<tr>
<td>$I_{\text{IP}_3}$ (dBm)</td>
<td>-11</td>
<td>-8.7</td>
<td>-3</td>
<td>7.2</td>
<td>-11</td>
<td>-10.5</td>
</tr>
<tr>
<td>NF (dB)</td>
<td>5.2</td>
<td>5.7</td>
<td>3.5</td>
<td>4.6</td>
<td>4.7</td>
<td>6.3</td>
</tr>
<tr>
<td>Frequency (GHz)</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Inductor Count</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Technology</td>
<td>0.18µm</td>
<td>0.18µm</td>
<td>0.35µm</td>
<td>0.18µm</td>
<td>0.13µm</td>
<td>0.13µm</td>
</tr>
</tbody>
</table>

$^a$NF is for entire receiver. The theoretical minimum NF for a common gate amplifier with matched input (used in [5]) is 2.2 dB [29].

5.4.6 A Current-Reuse LNA [14]

In [14], we proposed a current-reuse design in order to use the 1.8 V supply more efficiently. A schematic of the design is shown in Fig. 5.21. The process included symmetric spiral center-tapped inductors with poly-silicon ground shields. Three such inductors were used in the design. The first stage of the LNA looks like a current-source to the second stage, and this improves the common-mode rejection. It is also of note that deep n-wells were used to isolate the bulk connections of the transistors. This allowed us to tie the bulks of the transistors to their respective sources. This was necessary to prevent an increase in the threshold voltage of the cascade transistors due to the
body effect [2]. By keeping a low threshold voltage, the transit frequency \(f_T\) of the devices is maintained at the required value. All three inductors are 16.9 nH with a quality factor \(Q\) of 8.2 at the operating frequency. Additional resistors were added in parallel to the inductors (not shown in Fig. 5.21) in order to broaden the matching-bandwidth for the matching inductor, and the gain-bandwidth for the load inductors. The LNA was biased using a constant-
\(G_m\) biasing circuit with 2-bit gain control of the \(G_m\).

![Diagram](image)

Fig. 5.21 Schematic of (a) the gain control biasing circuit and (b) the current-reuse LNA.

The input of the LNA was matched to a 50-Ω source using a high-pass LC matching network. Compared to a low-pass matching network [13], a high-pass matching network requires only a single inductor (versus two) which can make use of mutual coupling between the coils to boost the effective inductance resulting in considerably smaller die area usage. A low-pass LC matching network requires differential inductors in order to achieve the same effect (not included in our process design kit). An additional 1-kΩ resistor (not shown in Fig. 5.21) was added in parallel with the input inductor in order to broaden the matching bandwidth. The effective \(Q\) of the inductor was therefore approximately 2.6. The overall impedance matching network gain
was equal to 11.3 dB which is slightly higher than that used in our previous design, but still smaller than that in [13]. Note that including input pad parasitics, and the device capacitance, the input matching network used does not easily fall into the L-match category. Impedance matching was done using Smith Chart concepts [31].

This LNA design was used for an energy-aware design, and as the power consumption of the LNA is changed with the gain state, the device capacitances of all transistors and most importantly, M1 and M2, are also changed. These changing device capacitances could potentially alter the frequency at which the LNA is matched to the 50-Ω source. We can reduce this effect by ensuring that the resonant frequency between the matching inductor and the device capacitances is significantly higher than the operating frequency (2.4 GHz). The same holds true for the two load inductors. Obviously this puts a restraint on the minimum $f_T$ of the devices.

**5.4.6.1 Gain Performance**

The voltage gain of the impedance match can be calculated as,

$$G_1 = \sqrt{\frac{\omega_0 L Q_{L1}}{R_s}}$$  \hspace{1cm} (5.14)

where $L$ is the inductance (16.9 nH), $Q_{L1}$ is the quality factor of the inductor including the additional parallel resistor (2.6), and $R_s$ is the source resistance (50 Ω). The LNA actually consists of three isolated gain stages with the last stage being a transconductance stage loaded by a finite $Q$ inductor and the passive mixer. The first stage is due to the matching network described above. The second gain stage consists of a V-I conversion by M1 and M2, and an I-V conversion by the first load inductor. The output impedance of the cascode V-I converter consisting of M1-M4 is significantly higher than the parallel parasitic resistance of the first load inductor. As a result, the gain of the second stage can be closely approximated as,

$$G_2 = g_m \omega_0 L Q_{L2}$$  \hspace{1cm} (5.15)
where $g_m$ is the transconductance of $M_1$ and $M_2$, and $Q_{L2}$ is the quality factor of the load inductor. The final stage of the LNA is loaded by a quadrature passive mixer and an inductor of the same inductance and $Q$ as the previous stage. The biasing and device sizes are the same as the second stage resulting in the same $g_m$. Therefore, with $G_{\text{mix}}$ as the input conductance of the passive mixer, the overall voltage gain is,

$$ G_{\text{total-LNA}} = \frac{(g_m \omega_0 Q_{L2})^2}{\omega_0 L Q_{L2} G_{\text{mix}}} + 1 \sqrt{\frac{\omega_0 L Q_{L1}}{R_s}} \quad (5.16) $$

In order to achieve sufficient gain-bandwidth, $Q_{L2}$ was reduced from 8.2 to approximately 3.4 using additional resistors parallel to the load inductors. Our expression, (5.16), shows that the LNA gain is proportional to $g_m^2$. Fig. 5.22 shows the measured voltage gain and $S_{11}$ of the design.

![Fig. 5.22 Voltage gain and $S_{11}$ of the proposed current reuse design.](image)

Unfortunately, due to a poor choice in on-chip biasing scheme, the current consumption in all modes was approximately 20% higher than expected. This also affected the gain control step, in particular the 5 mA to 3 mA mode step. In fact the 3 mA mode shows comparable performance to the 5 mA mode while consuming 40% less power. Another anomaly is the shift in center
frequency from the design value (2.4 GHz) to 2.25 GHz. This is attributed to poor modeling of the on-chip inductors.

5.4.6.2 Noise Performance

In this LNA, there is not much to say about the noise performance. The overall $NF$ is dominated by the 3 dB matching network $NF$, and the drain noise of the input device. From Fig. 5.23, the LNA $NF$ is not much affected by the change in gain/power states. However, as we will see in chapter 6, the overall system $NF$ is greatly affected. The minimum $NF$ achieved is around 5 dB which again is sufficient for the intended application.

![Graph](image)

Fig. 5.23 Noise performance of the proposed current reuse LNA.

5.4.6.3 Timing

As changing the gain state of the receiver involves a change in the DC operating point, the receiver must be able to change state fast enough to meet requirements. The IEEE 802.15.4 standard specifies a 128 μs preamble [32] at the head of each data packet which can be used for the PLL and AGC to lock. An advantage of designing the gain control in the RF section is that RF circuitry is designed with short time constants. Therefore, the circuits can reach steady-state quickly. Fig. 5.25 illustrates the settling time of the receiver
power consumption as the receiver goes from the highest gain state to the lowest gain state. The receiver requires approximately 1 µs for the current consumption to be within 1% of the steady-state value leaving ample time for the PLL to lock.

Fig. 5.24 Simulation of the settling time of the receiver. The receiver settles to the desired state within approximately 1 µs.

5.4.6.4 Overall Performance
The LNA achieved an $IIP_3$ of approximately -11 dBm in all gain states. The LNA was designed for a gain step of 6 dB. However, the biasing network was designed using on-chip resistors in order to meet limitations on the number of probes. In a more robust biasing scheme, at least one off-chip resistor should be used to set the desired current consumption. Unfortunately, the measured bias current deviated significantly from the nominal value resulting in a change in the gain step. Future iterations of this work will use a more accurate gain-step. In Chapter 6, we will do a comparison between the full receiver which used this LNA, and state-of-the-art designs in literature. For now we will turn our attention to passive mixer design and analysis.
5.5 Passive Mixer Design and Analysis

In Section 4.3.1, we looked at the case for passive mixing over active mixing. It was shown that in recent literature, designs using passive mixer exhibited better overall sensitivity for the power consumption than designs using active mixers. This was mainly attributed to three facts: passive mixers are highly linear, they don’t consume power, and they don’t contribute flicker noise. We also introduced two different types of passive mixers: the voltage-mode, and the current-mode passive mixers. We used current-mode passive mixers in [14] for their high linearity, and a novel technique was introduced using voltage-mode passive mixers in [30] which also allows for high linearity. In this section, we will analyze the current-mode passive mixer in three parts; the LNA-mixer interface, the mixer core, and the mixer-TIA interface.

For the LNA-mixer interface, we are mainly concerned with the passive mixer’s input impedance since (5.16) shows that it will affect the LNA voltage gain. For the passive mixer core, we will concentrate on the conversion gain from the switching transistors to the IF. For the mixer-TIA interface, we are mainly concerned with the output impedance which as mentioned, affects loop stability and linearity. We can model the current-mode passive mixer using the system in Fig. 5.25. The LNA is represented by a current source with an LC tank and a parallel resistance. The op-amp with feedback resistors forms a TIA which can easily be changed into a filter (Section 4.3.2). The double-balanced quadrature mixer core provides in-phase and out-of-phase components at the output while ensuring good isolation between the LO, RF and IF ports. We will set up our analyses by briefly discussing convolution matrices [33].
5.5.1 Convolution Matrices

A simple model for the time-varying conductance of a single switch (Fig. 5.25) in the ON state is,

\[ g_T(t) = K(V_{LO} \cos(\omega_{LO} t) + V_{DC} - V_T) \]  \hspace{1cm} (5.17)

where \( K \) is a constant which depends on the switch sizes and the technology, \( V_{LO} \) is the LO signal swing, \( V_{DC} \) is the bias voltage across the gate and source of the switches, and \( V_T \) is the threshold voltage of the switches. In the OFF state, \( g_T(t) = 0 \). As the LO is available in quadrature phases, we can define \( LO_Ip \) by (5.17). For the switches driven by \( LO_{Im} \), \( LO_{Qp} \) and \( LO_{Qm} \), the cosine in (5.17) is replaced by negative cosine, positive sine and negative sine respectively. The conductance of these switches are \( g_{T2}(t) \), \( g_{T3}(t) \) and \( g_{T4}(t) \). It should be noted that (5.17) assumes that the LO signal appearing at the sources of the switching transistors is negligible, which is true for typical biasing conditions. In practice, LO leakage to the mixer input is dependent on the output impedance of the LNA, and it can in turn change the conductance of the switching transistors. However, since we have assumed no leakage, the mixers operation is independent of the LNA output impedance.
Fig. 5.26 Decomposition of $g_T(t)$ in the time and frequency domain. (a) $g_T(t)$ (b) $G_T(f)$ (c) pulse train in time (d) pulse train in frequency (e) the sampling function in time (f) the sampling function in frequency. (e) and (f) show the sampling function for two different sampling function widths.

From Fig. 5.26, we can see how $g_{T1}(t)$ to $g_{T4}(t)$ can be mapped into the frequency domain. $g_{T1}(t)$ is a convolution between an impulse train and a sampling function which in the frequency domain is represented by a multiplication between a frequency domain impulse train and a frequency domain sampling function. A mixer multiplies in the time domain, and hence the output in the frequency domain is a convolution of the input and $G_{T1}(f)$. $G_{T1}(f)$ only has values at discrete frequencies because we assumed that the LO is periodic. We can therefore write convolution matrices for $G_{T1}(f)$ to $G_{T4}(f)$ [27]. If only the first two harmonics are considered, then the result is,

$$G_{T1}(f) = \begin{bmatrix} G_0 & G_{-1} & G_{-2} \\ G_1 & G_0 & G_{-1} \\ G_2 & G_1 & G_0 \end{bmatrix}$$

(5.18)
where we have limited \( \mathbf{G}_{\text{TN}}(\mathbf{f}) \) to a three-by-three matrix for simplicity. Note how the \( G_1 \) in (5.18) and (5.19) is in-phase while it is \( 90^\circ \) out-of-phase in (5.20) and (5.21). This is a simplification since in a real MOSFET, the internal capacitances of the device result in both in-phase and quadrature components for each term in (5.18)-(5.21). The subscripts, \( n \), for each entry correspond to \( f_{RF} + nf_{LO} \). The convolution matrix components for \( \mathbf{G}_{\text{T2}}(\mathbf{f}) \) to \( \mathbf{G}_{\text{T4}}(\mathbf{f}) \) are given in terms of those calculated for \( \mathbf{G}_{\text{T1}}(\mathbf{f}) \). As an example of how to use the convolution matrices, assume we apply a small voltage, \( V_A \) which has a spectral component at \( f_{RF} \), across a switch governed by (5.18). We can calculate the output components at the zero, positive and negative sidebands as,

\[
\begin{bmatrix}
I_{A,f_{RF}+f_{LO}} \\
I_{A,f_{RF}} \\
I_{A,f_{RF}+f_{LO}}
\end{bmatrix} =
\begin{bmatrix}
G_0 & G_1 & G_2 \\
G_1 & G_0 & G_1 \\
G_2 & G_1 & G_0
\end{bmatrix}
\begin{bmatrix}
V_A \\
0 \\
0
\end{bmatrix}
\]

(5.22)

This is obviously just a simple extension of Ohm’s law. We can then use Kirchhoff’s laws to analyze the entire mixer. The TIA’s op-amp is assumed to be ideal at IF frequencies and hence the IF bandwidth is not apparent from our derivations. Let \( Y_{\text{TIA}} \) be the TIA differential input admittance, and \( V_{RF} \) the voltage across the mixer input terminals. Therefore, we can write,

\[
V_X = (\mathbf{G}_{\text{T1}} + \mathbf{G}_{\text{T2}} + 2Y_{\text{TIA}})\mathbf{G}_{\text{T1}}^{-1}(\mathbf{G}_{\text{T1}} - \mathbf{G}_{\text{T2}})V_{RF}
\]

(5.23)
\[ V_Y = (G_{T3} + G_{T4} + 2Y_{TIA})^{-1}(G_{T3} - G_{T4})V_{RF} \]  

(5.24)

where \( V_X \) and \( V_Y \) are defined in Fig. 5.25. At high frequencies, the op-amp gain tends to zero, and we can approximate the TIA input admittance as \( R_f \) in parallel with some node capacitance, \( C_X \). With \( A \) (the final answer takes the limit as \( A \) tends towards infinity) as the op-amp low frequency open-loop gain, \( Y_{TIA} \) can be written as,

\[
Y_{TIA} = \begin{bmatrix}
  sC_X + \frac{1}{2R_f}(1 + A) & 0 & 0 \\
  0 & sC_X + \frac{1}{2R_f} & 0 \\
  0 & 0 & sC_X + \frac{1}{2R_f}
\end{bmatrix}
\]  

(5.25)

### 5.5.2 The LNA Mixer Interface

To find out how the mixer loads down the LNA, we first apply the same voltage, \( V_{RF} \), as above, and measure the current going into the mixer core.

We find that the input current is calculated as,

\[
2I_{RF} = (G_{T1} + G_{T2} + G_{T3} + G_{T4})V_{RF} + (G_{T2} - G_{T1})V_X + (G_{T4} - G_{T3})V_Y
\]

(5.26)

This equation was solved using MATLAB to obtain \( I_{RF} \) as a function of \( V_{RF} \).

The solution we are looking for is \( I_{RF,0}/V_{RF,0} \) which allows us to write the differential input conductance as,

\[
G_{mix} = 2G_0 - \frac{4G_0G_{RF}}{1 + 2R_f(sC_X + G_0)}
\]

(5.27)

where \( s_1 \) equals to \( 2\pi(f_{RF} + f_{LO}) \). The first term, \( 2G_0 \), can be seen by inspection. However, an additional term arises following our assumption that the TIA input impedance tends towards \( 2R_f||C_X \) at high frequencies. From Fig. 5.25, the input signal is up-converted due to \( G_{T1} \) and forms a voltage at \( V_X \). This high frequency signal then gets down-converted through \( G_{T2} \) which is out-of-phase with \( G_{T1} \) resulting in an overall negative input admittance term. Similar paths exist through \( G_{T3} \) and \( G_{T4} \). If the op-amp bandwidth were infinite, this additional term would not arise.
The additional term in (5.27) increases the input impedance of the passive mixer, and it would seem that if $R_I$ is large and $s_1C_X$ is minimized, and $G_1$ can be made equal to $G_0$, the input impedance would be infinite. While infinite input impedance may seem like a benefit (no loading to the LNA), we will see in the analysis on the mixer core, that it goes hand-in-hand with conversion gain. The higher the input impedance, the lower the conversion gain.

As a side note, the ratio $G_1/G_0$ is dependent on the peak to average conductance of the switches, and tends towards a value of one as the duty cycle of the switches is reduced [34]. The ratio of $2G_0$ to $s_1C_X$ will depend on the technology used and the frequency of operation. Clearly for $s_1C_X$ to be considered negligible, the technology’s $f_T$ would have to be at least an order of magnitude higher than the operation frequency. Taking into account the op-amp’s input capacitance, $s_1C_X$ was found to be significantly greater than $2G_0$ in our design.

5.5.3 The Mixer Core

Going back to (5.23), we can calculate the conversion gain by simply multiplying $V_X$ or $V_Y$ by the op-amp gain, $A$. In this case, we are interested in the terms $V_{X,-1}/V_{RF,0}$ and $V_{Y,-1}/V_{RF,0}$ multiplied by $A$. We can calculate the conversion gain of the passive mixer as,

$$
\frac{V_{EI}}{V_{RF}} = -2G_1R_I + \frac{4G_2G_1R_I^2}{1 + 2R_I(s_1C_X + G_0)} \quad (5.28)
$$

$$
\frac{V_{IFQ}}{V_{RF}} = 2jG_1R_I - \frac{4jG_2G_1R_I^2}{1 + 2R_I(s_1C_X + G_0)} \quad (5.29)
$$

The term -$2G_1R_I$ can be seen on inspection due to the shunt-shunt feedback configuration. Needless to say, solving the problem using higher order matrices will lead to more complex solutions. Once again we note that $s_1C_X$ is large and it therefore limits the influence of the term involving $G_2$. Otherwise, if $C_X$ were equal to zero, we would find that the conversion gain would be reduced by the second terms in (5.28) and (5.29).
5.5.4 The Mixer TIA Interface

The op-amp is conveniently designed as a two-stage amplifier where the first stage provides DC gain and the second stage is used to drive the output impedance. Assuming the second stage is a transconductance, $G_{m2}$, and the first stage provides DC gain, $A_1$, it is easy to see that the DC loop-gain of the TIA is $A_1 G_{m2}/G_{out}$, where $G_{out}$ is the output conductance of the passive mixer. Reducing $G_{out}$ improves the loop-gain thereby improving the linearity of the TIA while also degrading its phase margin. The resonator at the output of the LNA can be approximated as having conductance $G_{LNA}$ (equal to $(\omega_0 L Q_2)^{-1}$) at $f_{RF}$ and infinity at other frequencies. $G_{out}$ is calculated by applying a voltage across the output of the mixer core, $V_{test}$ and measuring the resulting current, $I_{test}$. The result is,

$$I_{test} = \frac{1}{2} \left[ (G_{T1} + G_{T2}) V_{test} + (G_{T2} - G_{T1}) V_{RF} \right]$$  \hspace{1cm} (5.30)

where $V_{RF}$ is the matrix representing the resulting voltage which develops across the input terminals of the mixer core. With some effort, it can be shown that,

$$V_{RF} = \left[ 2 G_{LNA} + G_{T1} + G_{T2} + G_{T3} + G_{T4} + (G_{T4} - G_{T3}) V_{V} V_{RF}^{-1} \right] (G_{T1} - G_{T2}) V_{test}$$  \hspace{1cm} (5.31)

The value of $G_{out}$ which we are interested in is equal to $I_{test, -1}/V_{test, -1}$ which is calculated as,

$$G_{out} \approx G_0 - \frac{G_1 G_{-1}}{G_0 + G_{LNA} + G_{mix}^2}$$  \hspace{1cm} (5.32)

We can see from (5.32) that the output impedance of the passive mixer depends not only on the conductance of the switches, but on the output impedance of the LNA. Note that $C_X$ was assumed to be part of the TIA. The second term in (5.32) results from mixing up and then back down in frequency. The output impedance was calculated using three-by-three matrices rather than five-by-five due to the computational difficulty.
5.5.5 Accuracy of the Analysis

Equations (5.27), (5.28), and (5.32) and their simulated counterparts are plotted versus the switch width in microns in Fig. 5.27 \((G_n\) were extracted from the simulation of a single MOSFET). The trend derived in the equations holds true for the most part in simulation although \(G_{out}\) appears to deviate from theory as the device width is made large. As mentioned earlier, a real MOSFET includes a distribution of capacitances and resistances which were not modeled by our simple model, and this is the biggest factor contributing to the equations inaccuracy.

\[
\begin{align*}
\text{Conversion Gain (V/V)} & \\
\text{Width (microns)} & \\
\end{align*}
\]

\[
\begin{align*}
\text{G}_{\text{mix}} & = \left( \frac{V_{T} - V_{DC}}{V_{LO}} \right) + \left( \frac{V_{DC} - V_{T}}{\pi} \right) \sin^{-1} \left( \frac{V_{T} - V_{DC}}{V_{LO}} \right) \\
& + \frac{V_{LO}}{4\pi} \sin 2 \cos^{-1} \left( \frac{V_{T} - V_{DC}}{V_{LO}} \right) \\
\text{G}_{0} & = \frac{V_{DC} - V_{T}}{\pi} \cos^{-1} \left( \frac{V_{T} - V_{DC}}{V_{LO}} \right) + \frac{V_{LO}}{\pi} \sin \cos^{-1} \left( \frac{V_{T} - V_{DC}}{V_{LO}} \right).
\end{align*}
\]

Fig. 5.27 Comparison between theoretically calculated and simulated conversion gain (CG in V/V), input conductance (G_{mix}) and output conductance (G_{out}). In simulation, the LO was 2.45 GHz, 250 mV_{pk} per phase.

For a sinusoidal LO, \(G_0\) and \(G_f\) can be shown to be equal to

\[
G_f = \frac{V_{DC} - V_{T}}{2\pi} \cos^{-1} \left( \frac{V_{T} - V_{DC}}{V_{LO}} \right) + \frac{V_{DC} - V_{T}}{\pi} \sin \cos^{-1} \left( \frac{V_{T} - V_{DC}}{V_{LO}} \right)
\]
From (5.33) and (5.34), if we were to bias the voltage-output passive mixer at the threshold voltage of the transistor (i.e. $V_{DC} = V_T$), the ratio $G_1/G_0$ would equal to $\pi/4$ which is -2.1 dB. This agrees with the analysis in [34].

5.5.6 Overall Implementation

A current-mode passive mixer was used in [30] to design a highly linear mixer/IF stage. The current-mode passive mixer was connected to the current-reuse LNA presented in Section 5.4.6. Based on the preceding analyses, we can optimize the switch size, LO strength, and $R_f$. Increasing the LO voltage improves the conductance of the switches without greatly affecting the switches capacitance. Therefore, for minimum capacitive loading to the frequency synthesizer and LNA, we should maximize the LO voltage. We chose a 250 mV peak per LO phase as this value does not require excessive driving capability of the LO. For $R_f$, we note from the section above that $R_f$, to first order, does not affect the op-amp loop-gain. However, if the non-dominant pole is at the output of the op-amp, then a smaller $R_f$ leads to higher op-amp unity-gain bandwidth. As a compromise between overall voltage gain ((5.28) and (5.29)), and bandwidth, we selected $R_f$ as 4 k$\Omega$. The simulation data in Fig. 5.28 illustrates the optimization of the switch width. When using Fig. 5.28, we must take into account the increasing $G_{mix}$ (Fig. 5.27) loads down the LNA thereby reducing the LNA voltage gain (this is evident from (5.16)). Therefore, there is an optimum width for minimum overall NF. For the LNA output impedance, of 880 $\Omega$, this was found to be 4 $\mu$m, but because larger switch size is more forgiving in terms of process variation, we chose a switch width of 5 $\mu$m. Details of the op-amp were given in Section 4.3.2.4. In Chapter 6, we will discuss the overall measurement results of the full receiver front-ends including the energy-aware design presented here.
Fig. 5.28 Switch width optimization including input-referred noise (IRN), op-amp loop-gain, and voltage gain. The LO was 2.45 GHz, 250 mV\text{pk} per phase.
References


10. B. W. Cook, A. Berny, A. Molnar, S. Lanzisera, K. S. J. Pister, “Low-Power 2.4-GHz Transceiver With Passive RX Front-End and 400-mV


32. IEEE 802.15.4 Standard For Local and Metropolitan Area Networks, 2003.


6.1 Introduction

In this chapter, we will present measured results of two systems which we have published [1], [2]. The first system [1] is our energy-aware system. We have presented the current reuse LNA and current-mode mixer designs for this system in Chapter 5, and in this chapter, we present the overall system measurements and include some discussion. Then we look at a very low power system [2] which we have designed that incorporates many of the power savings techniques which we have already discussed. We will show how the proposed system is among the least power hungry designs in literature while it retains excellent performance and robustness. Finally we will compare all of the proposed systems with state-of-the-art designs in literature. All of our measurements were done using direct on-wafer probing.

6.2 Energy-Aware Receiver

The focus of this receiver was simply to illustrate the principle of energy-aware design. We proposed a design methodology for maximizing the energy-awareness which involved pushing the noise and linearity requirements of the receiver to the LNA. The implementations of the current-reuse LNA and the current-mode passive mixer were discussed in Chapter 5. The overall receiver architecture is shown in Fig. 6.1. The input is fed into a high-pass L-match which offers 11.3 dB of voltage gain. The LNA is modeled as a transconductor with finite output impedance $R_{\text{out}}$. The passive mixer core feeds a pair of op-amp based TIAs to provide IQ voltage signals to the output. Using the
strategy of pushing the noise and linearity requirements to the LNA, we then
designed the LNA for high nominal gain and a wide gain/power control range.
Details of the LNA are found in Section 5.4.6. The overall power consumption
was controllable from 9 mW down to 2.5 mW. Therefore, in its lowest
performance state, the receiver saves 72 % of its nominal power consumption. The key principles used in this design which have been
discussed in the thesis include:

a) Energy-aware design

b) Current reuse for optimal use of supply voltage

c) Pushing the linearity and noise requirements to the LNA

d) Current-mode passive mixer for high linearity down-conversion and IF
section.

Fig. 6.1 Overall architecture of the energy-aware receiver in [1].

6.2.1 Overall Measurement Setup

A micrograph of the chip is shown in Fig. 6.2. Four sets of outputs are
available for measurement. From the left, the first output ground-signal-signal-
ground (GSSG) pad is connected as G-RXI-RXQ-G where G represents
ground, RXI represents the in-phase receiver output, and RXQ represents the
90 degrees out-of-phase receiver output. The second output pad from the left
is connected as G-RXI-G-RXQ-G and allows us to measure the RXI and RXQ outputs one at a time. The third set of output pads from the left is a GSSG pad used for measuring the LNA’s $NF$, and the final GSSG pad is used for measuring the LNA’s voltage gain (see Fig. 6.1). The LNA was tested separately from the overall system using the same strategy described in Section 5.4.5.1.

The LO polyphase splitter was implemented on-chip as a two-stage RC polyphase splitter. This was done in order to reduce the pad count. Due to the limitation on the number of RF probes which could be used, the biasing circuitry was implemented using on-chip resistors. The drawback is that current consumption of the chip can deviate significantly from the designed value. We used a constant-$g_m$ biasing circuit [4] for the LNA with a resistor which could be varied in three steps. This is an extremely simplistic method for gain tuning and in retrospect, a more robust method involving power detection should have been used. Such circuits are readily found for gain control in automatic gain-control (AGC) loops [5] and often involve decision making by the digital signal processor (DSP).

Fig. 6.2 Micrograph of the energy-aware receiver in [1].
As no channel-filtering was included in this design, it was important to provide linear output buffering in order to drive the 50-Ω measurement instruments. Standard three op-amp instrumentation amplifiers were used for this purpose which provide differential to single-ended conversion with high common-mode rejection ratio (CMRR) [6]. A ten-pin-probe was used for DC biasing and control signals. Despite its name, the ten-pin-probe actually only allows for up to seven signal inputs. The last three probes are allocated to ground pads.

### 6.2.2 System Performance

The front end was characterized for noise, gain, linearity and power consumption performance. The noise figure and conversion gain of the front end were measured using the Agilent E4407B spectrum analyzer which has a built-in noise figure personality. Unfortunately, neither the spectrum analyzer nor our noise source were designed to be used below 10 MHz (we used them down to 5 MHz). The current consumption in the highest to lowest power modes are 5.01 mA, 2.97 mA, 1.88 mA, and 1.39 mA respectively with a 1.8 V supply.

From Fig. 6.3, the front-end single-sideband (SSB) NF is around 9 dB (approximately 6 dB double sideband (DSB) NF) in the highest gain mode and
increases with the reduced LNA gain. The front end gain, as seen in Fig. 6.4, agrees with the LNA gain. The $IIP_3$ for the front end is -31 dBm in the highest gain mode and improves with lower LNA gain. This is shown in Fig. 6.5. This was sufficient for our application but can be improved by simply turning the TIA into a filter (as discussed in Section 4.3.2.3). This is because the bottleneck on the $IIP_3$ turned out to be in the IF section. Based on our discussion, in Section 4.3.2.3, we could reasonably expect a significant improvement in IF section $IIP_3$ by turning the TIA into a band-pass filter. The overall $IIP_3$ would then have been limited by the LNA. The LNA $IIP_3$ is also shown in Fig. 6.5. The front end gain of 35 dB in the highest gain mode is sufficient such that the noise performance of the subsequent blocks can be made insignificant without requiring high power consumption.

![Conversion gain of the full system in all gain modes.](image)

*Fig. 6.4 Conversion gain of the full system in all gain modes.*
6.2.3 Discussion

Table 6.1 shows a comparison between the proposed design and current literature. The NF quoted in this work is SSB NF while that in [7] is DSB NF (which is in principle up to 3 dB lower than SSB NF). [5] and [8] use image-reject mixers which are able to suppress the noise in the image band, however, the work in [8] uses two IFs and it is not clear how well the first image noise is suppressed. The authors of [7] used high $Q$ input matching and active mixing to achieve excellent NF for its current consumption. This came at the cost of a low IIP$_3$ and possibly a high flicker noise corner frequency. It should be noted that the key point in [8] was the innovative use of a digital demodulator which allowed the authors to achieve a low overall power consumption and good performance.

Certain concepts which could have been employed to reduce the overall power consumption include implementing lower supply voltage, subthreshold biasing, and use of a single-ended LNA. Although linearity requirements were
met, we should also have designed a proper filter rather than simply a TIA. All of these issues were addressed in our final design.

**Table 6.1**

**Comparison to Prior Published Work**

<table>
<thead>
<tr>
<th>Reference</th>
<th>This Work</th>
<th>[5]</th>
<th>[7]</th>
<th>[8]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency (GHz)</td>
<td>2.3</td>
<td>2.4</td>
<td>2.5</td>
<td>2.4</td>
</tr>
<tr>
<td>Current (mA)</td>
<td>5</td>
<td>3</td>
<td>1.9</td>
<td>1.4</td>
</tr>
<tr>
<td>Noise Figure (dB)</td>
<td>8.8(^B)</td>
<td>9.3(^B)</td>
<td>12.7(^B)</td>
<td>16.5(^B)</td>
</tr>
<tr>
<td>IIP(_3) (dBm)</td>
<td>-31</td>
<td>-27</td>
<td>-23</td>
<td>-19</td>
</tr>
<tr>
<td>Voltage Gain (dB)</td>
<td>35.6</td>
<td>34.7</td>
<td>28.7</td>
<td>24.5</td>
</tr>
<tr>
<td>Technology (µm)</td>
<td>0.18</td>
<td>0.18</td>
<td>0.18</td>
<td>0.18</td>
</tr>
</tbody>
</table>

\(^A\) Entire analog front-end included  
\(^B\) SSB NF which is approximately 3 dB higher than DSB NF  
\(^C\) Only LNA gain included

### 6.3 Energy-Aware Receiver – Low-Power Architecture

Having established the principle of energy-aware design, this receiver seeks to shore up all the low-power principles already introduced, to achieve the lowest possible power consumption. This work has been accepted for a full length presentation at VSLI-SoC 2010 [2]. The system architecture as well as the design of the IF blocks was shown in Section 4.3.3. The chip has been fabricated, and in this section, we will briefly look at the LNA and mixer designs, and present the overall simulation and measurement results. Fig. 6.6
shows the overall receiver architecture in detail. Some of the important concepts which went into the design of this receiver are:

a) Low voltage design. A supply voltage of only 1 V was used so that power is not wasted.

b) Single-ended LNA. We will show that second-order distortion requirements can be met using a single-ended LNA. Therefore, half of the LNA power can be saved.

c) Passive mixer output pole. Tuning the passive mixer output pole to 1 MHz allowed for significant improvement in the IF section’s $IIP_3$.

d) $G_m$-C filtering. $G_m$-C type filters do not have the high gain and wide bandwidth requirements of op-amp based filters.

e) Weak-inversion biasing. This was used in the LNA in order to achieve a good power to performance tradeoff.

f) Direct conversion. This architecture allowed the use of simple low-pass filtering at the IF. The flicker noise was reduced by using PMOS based transconductors in the $G_m$-C filters.

In the top left hand corner of Fig. 6.6, we show the replica bias circuit. Here, a current is injected externally to the chip and is replicated for all of the transconductors used. It is important to note that a true constant current source is difficult to generate on-chip due to resistor process variation. The tuning circuitry is shown in the upper right quadrant of Fig. 6.6 (details were discussed in Section 4.3.3.3). The LNA includes tuning for the frequency dependent LC networks, and gain/power control. The 2.4 GHz LO signal is generated externally to the chip and is split into quadrature signals via an on-chip polyphase filter. As filtering is included in this design, the outputs were buffered (to drive the 50-Ω instrumentation) using simple differential amplifiers without any linearization (shown in the bottom right hand corner). The common-mode voltage of the chip was derived on-chip using resistor division.
from the 1-V supply. The comparator and counter are clocked at 10 kHz (the clock is also generated off-chip).

![Figure 6.6 Overall Receiver Architecture.](image)

### 6.3.1 LNA and Mixer Designs

A schematic of the LNA and mixers is shown in Fig. 6.7. Input matching is achieved using a series resonant network with a resistor in series. Under matched conditions, the noise figure of the matching network is 3 dB, and the voltage gain of the matching network is equal to the quality factor \( Q \) of the network. In this work, an 11.5 nH inductor was used resulting in a \( Q \) of 3.54 and a voltage gain of 11 dB. An LC tank was used to tune the output node of the LNA. The load inductor was 6.5 nH with a \( Q \) of 8.6 resulting in an effective output resistance for the LNA of 860 Ω. Both LC networks were made tunable to avoid the modeling issues encountered in our previous design which
resulted in an offset in the frequency response (see Section 5.4.6.1). The LNA includes three gain control steps of 6 dB to ease the gain compression requirements of the IF section, and to allow for energy-aware design. The input transistors are biased in the weak-inversion region, and because their gates are always connected to the input node (whether on or not), their contribution to the input capacitance is almost fixed. This is true because in the weak-inversion and the off-state, the gate capacitance is dominated by the gate-drain and gate-source overlap capacitances, not the intrinsic device capacitances.

![Fig. 6.7 Schematic of the LNA and Mixers.](image)

Single-balanced passive mixers were used to convert the single-ended RF signal into a differential IF signal. This allowed the use of a single-ended LNA thereby saving half of the LNA power consumption. The justification for this strategy is the relaxed $IIP_2$ requirements of the IEEE 802.15.4 standard. The main concern is unwanted DC-offset related to the self-mixing of either LO signals or strong interfering signals. Self mixing of LO signals results in a static DC offset which must be filtered before introducing any high gain stages to the signal. We can estimate the require $IIP_2$ based on self mixing of interfering signals as,

$$IIP_2,\text{req} (dBm) \geq 2P_{\text{blk}} - Sen + SNR_{\text{req}}$$

(6.1)
where $Sen$ is the required sensitivity. Given alternate channel interferers equal to -52 dBm, the required $IIP_2$ is $2(-52) - (-82) + 14 = -8$ dBm. The achieved $IIP_2$ of the down-converter can be estimated as $P_{LO}/G_{\text{leak}}$ where $G_{\text{leak}}$ is the ratio of the differential RF signal at the gates of the switching stages to the single-ended RF signal at the source of the switching stages (note that $G_{\text{leak}}$ does not include common-mode leakage) and $P_{LO}$ is the LO power. $G_{\text{leak}}$ is effectively a single-ended to differential leakage gain. This formula is easily derived. If the mixer gain is proportional to $P_{LO}$, then the IM$_2$ product is proportional to the input power multiplied by $G_{\text{leak}}$. If the $IIP_2$ is the input power at which the output IM$_2$ level is equal to the level of the desired output, then it occurs when $P_{in}P_{LO} = P_{in}^2G_{\text{leak}}$ or $P_{in} = P_{LO}/G_{\text{leak}}$.

In [9], it is shown that the $IIP_2$ of active mixers has similar dependency, while $IIP_2$ on the order of +40 dB seems typical.

### 6.3.2 Simulation Results

The design was implemented in a 0.18 µm RFCMOS technology and sent for fabrication. The overall simulated performance of the receiver front end is shown in Table 6.2 and compared with recent literature. At this phase of the design, the proposed design compares favorably to recent literature, although the raw performance attained by [10] is still superior. It should be noted that in [10], several techniques were used which may or may not be allowable in a robust design, such as the lack of input matching [11] (instead, a step-up LC network was included using an inductor with a $Q$ of 20 – well beyond the achievable $Q$ in the process available to us), and the lack of an LNA (required for reverse isolation between the frequency synthesizer and the RF input. Furthermore, the receiver in [10] was not designed to meet a specific standard leaving some uncertainty in the performance. For our own design, Table 6.2 only shows simulated performance. The measured performance will be shown in Section 6.4 and shows degradation from simulations. The current consumption required in [7] was also less than that required in the proposed design, however, the $IIP_3$ was poor.
### Table 6.2

**Simulated Performance of the Proposed Design**

<table>
<thead>
<tr>
<th>Reference</th>
<th>[5]</th>
<th>[7]</th>
<th>[10]</th>
<th>[12]</th>
<th>This Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency Band (GHz)</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>$^A$RF $S_{11}$ Bandwidth, -10 dB (GHz)</td>
<td>$^B\infty$</td>
<td>0.4</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
</tr>
<tr>
<td>$^A$RF Gain Bandwidth (GHz)</td>
<td>0.5</td>
<td>0.4</td>
<td>-</td>
<td>3.0</td>
<td>0.4</td>
</tr>
<tr>
<td>Power Consumption (mW)</td>
<td>10</td>
<td>1.4</td>
<td>0.75</td>
<td>4.05</td>
<td>2.14</td>
</tr>
<tr>
<td>IF Power (mW)</td>
<td>5.76$^C$</td>
<td>0.5</td>
<td>0.75</td>
<td>1.15</td>
<td>0.31</td>
</tr>
<tr>
<td>Noise Figure (dB)</td>
<td>5.7</td>
<td>5.0</td>
<td>5.1</td>
<td>6.0</td>
<td>6.2</td>
</tr>
<tr>
<td>$II/P_3$ (dBm)</td>
<td>-16</td>
<td>-37</td>
<td>-7.5</td>
<td>-12</td>
<td>-12</td>
</tr>
<tr>
<td>$^D$Technology Node (µm)</td>
<td>0.18</td>
<td>0.18</td>
<td>0.13</td>
<td>0.09</td>
<td>0.18</td>
</tr>
</tbody>
</table>

$^A$ Estimated where necessary

$^B$ Very wide

$^C$ Estimated only

$^D$ CMOS only

### 6.3.2.1 $S_{21}$, $S_{11}$, and Voltage Gain

Fig. 6.8 shows a plot of $S_{21}$ and $S_{11}$ where the reference impedance of port 1 is 50 Ω and that of port 2 is 2kΩ. Output pads were added to the output node of the LNA which can be left open when testing the full receiver. When testing the LNA, the mixer was turned off. Rather than attempt to buffer the LNA output, for this fabrication, we simply renormalized the S-parameters for port 1 as 50 Ω and port 2 as 2kΩ [13]. As $S_{21}$ essentially shows the power gain of the circuit, in order to properly read off the voltage gain from Fig. 6.8, we must take the reference impedance ratio into account. The result is that we should
add a value of $10\log(2k/50) = 16$ dB to the plot of $S_{21}$. From Fig. 6.8, we do not expect much deviation in $S_{11}$ over the different power states of the LNA.

![Plot of $S_{11}$ and $S_{21}$ of the proposed LNA](image)

*Fig. 6.8 Plot of $S_{11}$ and $S_{21}$ of the proposed LNA where port 2 reference impedance is $2 \, k\Omega$, and port 1 reference impedance is $50 \, \Omega$."

### 6.3.2.2 Front end Gain and NF

The full receiver front-end gain is shown in Fig. 6.9, while the $NF$ is shown in Fig. 6.10. We show the gain at both the channel-select filter (CSF) input (the passive mixer's output), and the CSF output. The gain up to the CSF output follows a $3^{rd}$ order Butterworth response. The overall gain of approximately 40 dB in the pass-band is sufficient to make any noise contributed by the following stages negligible. The overall $NF$ shown in Fig. 6.10 is around 6.2 dB in the pass-band. The flicker noise corner frequency is around 10 kHz. Due to the $1/f$ nature of flicker noise, the total integrated noise contribution of flicker noise is the same over any decade. For example, the total noise contribution from 1 kHz to 10 kHz is the same as that from 10 kHz to 100 kHz or 100 kHz to 1 MHz. From Fig. 6.10, it is obvious that the flicker noise within the band from 100 kHz to 1 MHz is negligible compared to the thermal noise level in the same band. By extension, the flicker noise contribution from the 1 kHz to 10 kHz and 10 kHz to 100 kHz bands must be negligible compared to
the thermal noise contribution from 100 kHz to 1 MHz. This tells us that the flicker noise corner frequency is sufficiently low.

**Fig. 6.9** Full receiver front-end gain at the CSF Input and Output.

**Fig. 6.10** Full receiver front-end NF.
6.3.2.3 IF Section $IIP_3$

Due to limitations in the device models used, $IIP_3$ results of devices operating at zero drain-source voltage are highly inaccurate [14]. Regardless, passive mixers have been shown to demonstrate high $IIP_3$ [15]. Therefore, in this section, we only demonstrate the effect of the use of the passive mixer output pole on the $IIP_3$ of the CSF. The $IIP_3$ of the CSF can be simulated by taking interfering tones at either 5 MHz and 10 MHz, or at 10 MHz and 20 MHz. The requirement for the former condition is looser than the latter (Section 2.2.5.2), but conversely, the receiver $IIP_3$ under such conditions is worse (due to differing amount of filtering, section 4.3.3.1). The IF section $IIP_3$ is shown in Fig. 6.11 for two cases: with and without the tuned passive mixer output pole. This $IIP_3$ is shown for interferers at 5 MHz and 10 MHz offset from the desired signal. As expected, the improvement is 24 dB. The $IIP_3$ is +8.3 dBV which is equivalent to 18.3 dBm into a 50-$\Omega$ resistor. Therefore, with 30 dB front-end gain, the overall $IIP_3$ is expected to be approximately -12 dBm (we have assumed that the overall $IIP_3$ is dominated by the $IIP_3$ of the IF section. Simulations of the LNA $IIP_3$ support this assumption).

![Graph showing IF Section $IIP_3$ with (improved $IIP_3$) and without (poorer $IIP_3$) the tuned passive mixer output pole.](image_url)

*Fig. 6.11 IF Section $IIP_3$ with (improved $IIP_3$) and without (poorer $IIP_3$) the tuned passive mixer output pole.*
6.3.3 Chip Layout

The layout is shown in Fig. 6.12. The chip area is 1.5 mm by 1.5 mm in total, however much of this is taken up by output buffers, and metal fill. Of note is that the input is applied via a GSSG pad, despite the fact that only a single-ended LNA was used. For the input GSSG pad, one signal pad was used for the receiver input, and the other pad was used for a DC input. This was necessary due to the large number of external control signals used.

![Fig. 6.12 Layout of the full RX.](image)

6.3.4 Measured Performance

The chip was fabricated and has been characterized using on-wafer probing. In this section, we discuss the measured results and their discrepancy with the simulated ones. There are three major discrepancies: an improved
measured overall $NF$, degraded overall $IIP_3$, and poorer front end gain. These three problems appear to contradict each other, and although we can offer some explanations, in future, we will further investigate them.

### 6.3.4.1 Discrepancies in the Results

The first major discrepancy was a reduced gain front end gain. We observed approximately 5 dB reduction in the front end gain. This is surprising considering that the LNA gain was not very much different from simulation results. Since the mixer is passive, its gain should be very accurate. Therefore, the discrepancy has most likely arisen due to either poor output buffer gain, or loss between the DUT and the instrumentation. The first explanation is possible since the buffer was a very simple differential pair whose gain could not be set during the measurement. The second possibility is less likely, since for an output frequency of around 1 MHz, there is not much loss to speak of. Fig. 6.13 shows a schematic of the output buffers.

![Fig. 6.13 Schematic of the output buffers.](image)

From the schematic we can see that the output is matched to 50 Ω. A measure of common-mode rejection is provided by the common-mode 100-Ω at the virtual ground node. However, the transconductance of the devices is entirely process dependent. Although we could set the gain more accurately using feedback, we decided to keep the buffer design simple to ensure functionality.
Another major discrepancy in the measured results is the improvement in the measured $NF$ over the simulated $NF$. The measured $NF$ was found to be roughly 0.5 dB better than the simulated $NF$. Although it is possible that this improvement was a process variation, it is more likely that the loss compensation was not measured accurately enough. The pre-DUT loss was de-embedded from the $NF$ by simply assuming some attenuation in front of the LNA. In practice, the loss could have a more complex relationship with the measured $NF$ depending on the effect on the reflection coefficient. The pre-DUT loss was measured up to the probe tips by raising the probe tips from the wafer in order to create an open-circuit load, and then measuring the reflection coefficient. In a lossless line, all of the power should be reflected, but in a real transmission line, the signal is attenuated as it travels down the transmission line due to the finite conductance between the forward and return paths, and the finite resistance in series with the paths.

The last major discrepancy is in the overall $IIP_3$. The measured $IIP_3$ is around -17 dBm in the highest gain mode. Although we were not able to simulate the overall $IIP_3$, we estimated that it would be around -12 dBm. As the front end gain actually showed a reduction from the simulated value, our hypothesis that the reduced gain was simply due to underperforming output buffers becomes more plausible. The full reason behind degraded $IIP_3$ performance currently eludes us, and will be left for future work.

The mismatches in the measurements could have been possibly avoided by packaging the design. This would have allowed for more probing points, and would also have brought the external components closer to the DUT, reducing the loss and minimizing reflections. De-embedding and calibration would also be easier without having to work with wafer probes. On the downside, proper modeling of the chip-package-PCB interface would need to be done to ensure proper input matching of the receiver. Electro-static discharge (ESD) protection would also become a major issue. As time and money do not permit packaging at the moment, we will leave this for future work.

Apart from the above-mentioned discrepancies, the measured results agreed quite well with the simulated ones and are well within requirements. In
the next sections, we present the LNA, and full front-end measurement results.

6.3.4.2 LNA performance

The LNA was characterized for its $S_{11}$, $S_{21}$, and $NF$. It was mentioned in Section 6.3.3 that the input was probed using a GSSG pad. However, the output of the LNA was probed from a GSG pad. This mismatch prevented us from calibrating the system up to the probe tips. Instead, the system was calibrated up to the cable ends (which connect to the wafer probes). This caused a rather large error when attempting to convert the S-parameters from a two-port 50-Ω system to a two-port system with port one at 50-Ω, and port 2 at 2-kΩ. Therefore, the additional transmission lines were de-embedded from the S-parameter measurements. The measured S-parameters are plotted in Fig. 6.14.

![Fig. 6.14 Measured S-parameters of the LNA.](image)

We should once again note that there is a conversion factor of 16 dB when converting $S_{21}$ to voltage gain. This is because S-parameters are voltage waves normalized to the reference impedance (in this case, 2 kΩ). From Fig. 6.14, the $S_{11}$ was well modeled in simulation. Furthermore, the reflection coefficient does not change significantly between the different gain modes,
allowing for good reflection coefficient over a wide range of power consumption. The power consumption shown in Fig. 6.14 is the power consumption of the entire receiver. There is a slight reduction in the measured gain as compared to the simulated gain (around 1 dB) which could have been due to loss in the probes. Lastly, we point out that the gain difference between the gain modes is very close to the designed value of 6 dB. This and the constant input reflection coefficient are major improvements over our previous design (see Section 5.4.6.1). The measured LNA \( NF \) is shown in Fig. 6.15.

![Fig. 6.15 Measured NF of the LNA.](image)

The measured \( NF \) of the LNA is slightly lower than expected. This has been explained in section 6.3.4.1. As expected, the \( NF \) increases with the reduced power consumption.

### 6.3.4.3 The Calibration Loop

As discussed, a calibration loop was used to force the output pole of the passive mixer to 1 MHz. The gain of the receiver up to the mixer output was measured before and after calibration and is shown in Fig. 6.16. We also show the gain of the overall system. After calibration, the output pole frequency of the passive mixer settled slightly higher than 1 MHz (around 1.1
MHz). The CSF was not precisely tuned either which resulted in only 54 dB attenuation (instead of ideally 60 dB) at 10 MHz IF. However, the overall out-of-band rejection and $IIP_3$ are well within acceptable limits.

![Fig. 6.16. Receiver voltage gain up to the passive mixer output before and after calibration.](image)

### 6.3.4.4 System $IIP_3$

The system $IIP_3$ was measured up to the CSF output using a two-tone test with the first tone set to fall at 5.5 MHz IF and the second tone set to fall at 10.5 MHz IF. The third-order intermodulation tone therefore fell at 500 kHz, the center of the desired band. The $IIP_3$ results were taken for all four gain modes and are shown in Fig. 6.17. Given that it improves with the reducing LNA gain, we must conclude that the $IIP_3$ of the IF section was poorer than expected. However, it is still well within the acceptable $IIP_3$ range for our design. Nonetheless, it is something we should look into in our future studies.
6.3.4.5 Overall NF

The overall \( NF \) of the entire front end was measured using the gain method. In this method, the total output noise and the conversion gain are measured, allowing the overall \( NF \) to be calculated using the following formula,

\[
NF = N_{out} - 10 \log(kT\Delta f) + G_{conv} \quad (6.2)
\]

where \( N_{out} \) (in dB) is the total output noise, \( G_{conv} \) (in dB) is the conversion gain, and \( \Delta f \) is the resolution bandwidth. For accuracy, the noise level of the spectrum analyzer itself should be taken into account, however, due to the high gain of the device under test (DUT), it is insignificant in this case. The overall SSB \( NF \) for all gain modes is shown below. The DSB \( NF \) is approximately 3 dB lower when the \( NF \) is dominated by the received and LNA noise. Given that the \( NF \) follows the LNA gain mode closely, the overall \( NF \) is dominated by the IF section. The opposite was true in simulation. The use of a PMOS based IF section has made flicker noise insignificant.
6.3.5 Design Conclusions

We have presented both simulated and measured results of the proposed design. Apart from the lower than expected $IIP_3$, the low front-end gain, and the improved overall $NF$, the measured results are in good agreement with simulations. Furthermore, the results are well within requirements. The overall SSB $NF$ is around 9 dB in the maximum gain mode, while the power consumption is just 2.2 mW. This compares favorably with designs in literature where the conclusions drawn in section 6.3.2 still hold. In its lowest power state, the proposed design is able to conserve nearly 70 % of its nominal power consumption. Compared to the previous design, the proposed design offers better $IIP_3$, power consumption, and matching, while maintaining a respectable $NF$.

A micrograph of the design is shown in Fig. 6.19. The performance of the final design is summarized in Table 6.3.
Fig. 6.19 Micrograph of the proposed design.

### Table 6.3

**Measured Performance of the Proposed Design**

<table>
<thead>
<tr>
<th></th>
<th>IEEE 802.15.4 Requirement</th>
<th>This Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency Band (GHz)</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>RF $S_{11}$ Bandwidth, -10 dB (GHz)</td>
<td>N.A</td>
<td>0.6</td>
</tr>
<tr>
<td>RF Gain Bandwidth (GHz)</td>
<td>0.1</td>
<td>0.4</td>
</tr>
<tr>
<td>Power Consumption (mW)</td>
<td>N.A</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.7</td>
</tr>
<tr>
<td>$^A$Minimum Noise Figure (dB)</td>
<td>12</td>
<td>5.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21.2</td>
</tr>
<tr>
<td>$IIP_3$ (dBm)</td>
<td>-30</td>
<td>-17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-4</td>
</tr>
<tr>
<td>Technology</td>
<td>N.A</td>
<td>0.18 µm CMOS</td>
</tr>
</tbody>
</table>

$^A$ Estimated DSB $NF$ by subtracting 3 dB from SSB $NF$. 

165
For future iterations of this design, special care should be taken in the design of the supply bypassing circuitry. We should also consider a lower LNA gain since this ended up working out for the better in terms of the overall system $\text{IIP}_3$ (although the overall $\text{NF}$ was degraded somewhat).

6.4 Summary

In this chapter, we have presented measurement results of two published works based on the concepts proposed in this thesis. The first design was used to introduce the concept of energy-aware design, however in its nominal state, the proposed design cannot compete (in terms of raw performance) with other designs in literature. The second design uses many more of the proposed low-power concepts such as passive mixer output pole tuning, low-voltage design, weak-inversion biasing, single-ended LNA design, and most importantly, energy-aware design. Measured results of the proposed design compare favorably to the lowest power designs in literature.
References


Chapter 7

Conclusion and Future Direction

7.1 Conclusions

This thesis has presented a study of power consumption reduction techniques for low data-rate receivers based on the IEEE 802.15.4 standard. We have studied aspects of low-power design from the circuit level to the system level. We started by looking at the performance requirements of an IEEE 802.15.4 standard compliant receiver. This led us to propose several design techniques which take advantage of the relaxed performance requirements of the IEEE 802.15.4 standard. All work was done using a standard 0.18 µm RFCMOS technology.

At the circuit level, we have proposed the use of weak-inversion biasing as a means to achieve the maximum possible power gain LNA designs under the unilateral assumption. Our analyses on unilateral power gain under conditions of maximum possible impedance matching voltage gain and loading resistance have revealed that weak-inversion biasing can actually result in improved power gain, contrary to ideal simulations of unilateral power gain.

At the architectural level, we have proposed a novel voltage-mode passive mixer with a tuned output pole in order to relax the linearity requirements of the IF section. This allows the use of $G_m$-C filtering using transconductors without addition linearization. The proposed transconductors do not have the excessive bandwidth requirements of op-amps used in active-RC filtering. Furthermore, the proposed technique requires only a single biquadratic filter as compared to a third-order filter required for active-RC based filtering.
At the system level, we have proposed energy-aware design which seeks to optimize the energy consumption based on the *in-situ* performance requirements of the receiver. Our studies have shown that the proposed energy-aware method can be used alongside several other system level energy-savings methods such as fast receiver start-up, and wake-up receivers. The proposed system level design method was shown to save up to 70% of an RF front-end’s nominal power consumption when a strong desired signal is received.

### 7.2 Future Direction

Perhaps the biggest advantage of passive mixers is that they do not contribute flicker noise. In [1], the authors took a different approach to minimizing flicker noise. The authors replaced the NMOS type switching core in an active mixer with vertical NPN type transistors which are available in triple-well CMOS processes. Such bipolar transistors exhibit several useful advantages over NMOS devices including higher $g_m/I_{DS}$, better matching and most importantly, significantly lower flicker noise. The authors were therefore able to achieve a flicker noise corner frequency of less than 100 kHz. The overall power consumption of the LNA and mixer was 5 mW, but we feel that the design could have benefitted from some of the proposed design techniques employed in this work.

In particular, the use of an active mixer implies that the output impedance of the mixer is very large. We could therefore set the output pole frequency of the active mixer to whatever value necessary using a simple parallel resistor and capacitor. The tolerance of the pole would be significantly better than that in a passive mixer (whose output pole is affected by LO voltage variation, switch size variation, LNA output impedance variation, threshold voltage variation, etc) allowing output pole tuning to possibly be avoided. The biggest challenge in the design would be designing for a low supply voltage. This is because the turn-on voltage of bipolar transistors is around 0.7 V compared to only 0.5 V in the NMOS transistors in the Global Foundries 0.18-µm technology. The $f_T$ of the devices is also low (compared to NMOS devices)
and this would be a problem if higher frequency mixing were required. Nevertheless, the use of VNPN transistors in mixing presents possible future direction for research into low-power receiver design.

In Sections 2.3.4 and 3.1.1.3, we discussed the voltage gain, matching bandwidth tradeoff, as well as the effect of process variation on matching. If process variation were non-existent, we could potentially save a lot of power by pushing the voltage-gain, matching bandwidth tradeoff to the limit. This would imply designing very high $Q$ resonators. One way to get around the process variation problem is through calibration. To our knowledge not much research has been done on calibration at RF frequencies [2] (although the most obvious calibrating circuit at RF frequencies is the phase-locked-loop). On the bad side, calibration circuitry adds to design complexity, while taking up valuable die area. However, the lack of research in this area coupled with the potential gains is an open invitation to research into innovative RF calibration methods. Energy-aware design and calibration both involve the receiver adjusting its performance for optimization and we feel that when they are used together, they create the potential for ultra-low power receiver designs.

CMOS technology does not perform as well as bipolar technologies such as SiGe in the analog domain due to issues such as lower $f_T$ and transconductance efficiency, poorer device matching and higher flicker noise. However, the use of CMOS technology in the analog domain allows the designer to integrate analog and digital circuits onto the same die. While SiGe BiCMOS also offers CMOS devices, the advanced SiGe nodes currently support 0.18 $\mu$m CMOS devices compared to 28 nm CMOS devices in leading edge CMOS technologies (see for examples Global Foundries, TSMC, or IBM websites). It was observed in [3] that the primary advantage of CMOS over high performance analog technologies such as SiGe is the ability to design high performance switches. Leveraging such an advantage involves shifting the bulk of the signal processing in a receiver design to the digital domain. The author in [3] notes that an analog filter circuit is larger and more power hungry than an equivalent digital filter using the same CMOS process. This
suggests that in future we are likely to see the sampling operation in a receiver occurring much closer to the antenna. Indeed, research into direct-sampling architectures [4] is of great interest at the moment for receivers covering multiple applications and software-defined radios. Such architectures will only benefit from on-going technology scaling. Nevertheless, use of deep-submicron technologies brings about its own problems as the supply voltage scales faster than the threshold voltage [5]. This puts significant importance into the research of low-voltage circuit architectures. For instance, the design of an op-amp at a low supply voltage is non-trivial [6], and the problem is aggravated by the reduced intrinsic gain of deep-submicron devices.

The future is bright for RFIC design with more and more applications for wireless technology, each bringing its own new challenges. Scientists can choose to focus in a multitude of different directions to impact the research community and the wireless world.
References


List of Publications


Full List of References

2. IEEE 802.15.4 Standard For Local and Metropolitan Area Networks, 2003.


48. R. Senguttuvan, S. Sen, A. Chatterjee, “VIZOR: Virtually zero margin adaptive RF for ultra low power wireless communication”, *Proceedings*


78. C. Bowick, “Impedance Matching”, *RF Circuit Design*, Howard W. Sams & Co., Inc., 4300 West 62nd St, Indianapolis, Indiana 46268, USA, Chapter 4, pp. 72-75.


