NOTE TO USERS

Page(s) not included in the original manuscript are unavailable from the author or university. The manuscript was microfilmed as received

ii and iii

This reproduction is the best copy available.

TRANSCEIVER ARCHITECTURE FOR SHORT-RANGE WIRELESS APPLICATIONS

by

Gholamreza Yousefi Moghaddam

A Thesis Submitted to the
Faculty of Graduate Studies and Research
In Partial Fulfillment of the Requirements
For the Degree of

Doctor of Philosophy

In Electrical Engineering

Ottawa-Carleton Institute of Electrical and Computer Engineering
Department of Electronics

Carleton University

September, 2009

Your file Votre référence
ISBN: 978-0-494-60126-6
Our file Notre référence
ISBN: 978-0-494-60126-6

Abstract

A new concept in transceiver architectures is introduced in which the receiver front-end is used to do the job of the transmitter back-end as well, or vice versa. The re-use of major blocks saves silicon real-state and reduces the die cost. The modified high-IF dual-conversion receiver is intelligently re-configured into a dual-loop constant envelope modulator. The modulator is capable of transmitting high bit-rate data, such as video, or a short packet of slow-varying data, such as temperature measurements, offering prospect of achieving an energy-efficient transceiver for a time-multiplexed standard such as WPAN.

Furthermore, a novel mechanism for frequency and bandwidth tuning of RF-filters is presented. The idea is applied to the LNA in the receiver front-end; it is suggested that a narrow-band LNA centered at the desired channel alleviates the dynamic range requirements of the following stages in the receive chain. The proposed tuning mechanism adds a new dimension to the main architecture which turns it into a potential candidate for developing an adaptive, a multi-standard or a software defined radio.

A theoretical analysis on characteristics and limitations of low noise amplifiers with active load-loss compensation is presented. The study shows that the maximum achievable dynamic range is defined based on the power budget, bandwidth, and quality factor of the load. The analysis explores the impact of commonly practiced techniques on noise, linearity and dynamic range of the amplifier.

Acknowledgments

I would like to thank my supervisor, Professor Ralph Mason, for the possibility to work on this thesis and his support. His patience with the very slow last phase of work is also greatly appreciated. I also warmly thank Prof. Tad Kwasniewski for reviewing my thesis and for his valuable comments. I thank Prof. Calvin Plett and Prof. John Knight for their unconditional support, trust and encouragement.

I moved to industry full time in June 2006, I thought that the remaining work for the thesis would be a walk in the park. It wasn't. Finding motivation and energy to complete the work was almost impossible. It has taken a lot of support from colleagues, friends, and especially family to get this work done. I am grateful to you all for that.

I would like express my gratitude to Mrs. Blazenka Power, the department of Electronics administrator, for her compassionate support. Her fabulous smile, sparkling eyes with passion in finding the best possible solution and taking the action, immediately earned my everlasting respect. Thanks belong to Nagui Mikhail for his assistance and providing lab equipments.

Thanks to Irfan Ahmed for his unconditional support since my arrival to the Carleton University; to Gord Allan for his friendship, passionate technical discussions and brainstorming in the blue room; to Yasser Soliman for his encouragement and support; to Ryan Griffin, Bachir Tamer, Andrea Liao, James Chiu, Peter Noel, Farhad Zarkeshvari, Steve Knox, John Danson, Kimia Ansari, J. J., Salia Mirbaha, William Che Knisely, Siva Kanesapillai; to Mike Bailey, Suzzanne Alexander, Richard, little John, Myriam, Lynne and, of course, Julie.

Of course, none of this would have been possible without the unwavering support of my family. Thanks to my parents and siblings who kept me motivated with their warm and encouraging words from the other side of the world.

Reza Yousefi

Ottawa, Canada

September 21, 2009

Contents

LIST OF FIGURES

List of Tables

List of Acronyms

ADC	Analog to Digital Converter
AGC	Automatic Gain Control
BB	Baseband
BER	Bit Error Rate
BPF	Band Pass Filter
BW	Bandwidth
CMOS	Complementary Metal Oxide Semiconductor
CP	Charge Pump
DAC	Digital to Analog Converter
DECT	Digital European Cordless Telephone
DSP	Digital Signal Processor
FSK	Frequency Shift Keying
FTL	Frequency Translational Loop
GSM	Global System for Mobile Communications
ICP	Input Compression Point
IF	Intermediate Frequency
IIP3	Input referred 3rd order Inter-modulation Point
ISM	Industrial, Scientific, and Medical
LNA	Low Noise Amplifier
LNTA	Low Noise Transconductance Amplifier
LO	Local Oscillator
LPF	Lowpass filter
MPLL	Master PLL
NF	Noise Figure
OTA	Operational Transconductance Amplifier
PA	Power Amplifier
PLL	Phase Locked Loop
PFD	Phase-Frequency Detector
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
rms	Root mean square
RX	Receiver
SAW	Surface Acoustic Wave
SNR	Signal to Noise Ratio
TDMA	Time Domain Multiple Access
TSPC	True-Single-Phase-Clock
TX	Transmitter
VCO	Voltage Controlled Oscillator
WPAN	Wireless Personal Area Network

List of Symbols

Cgs	Gate-Source capacitance
Cgd	Gate-Drain capacitance
Cox	Thin oxide capacitance
Pin	Input power
Q	Quality factor
Ron	MOS On resistance
Vth	MOS threshold voltage
f_t	Transit frequency
g_m	MOS gate transconductance
g_{ds}	MOS drain transconductance
λ	Channel-length modulation factor
γ	MOS channel thermal noise factor
μ	Carrier mobility
ω_n	Natural frequency
ξ	Damping constant

1 Introduction

1.1 Wireless Micro-sensor Network

The idea of wireless microsensor networks has received a great deal of attention and interest. Distributed networks of thousands of collaborating microsensors promise a maintenance-free, fault-tolerant platform for gathering rich, multi-dimensional observations of the environment [1]. Wireless integrated network sensors combine sensing, signal processing, decision capability and wireless networking capability in a compact, low power system [2]. Such networks enable a reliable monitoring and control of a variety of applications that range from medical and home security to machine diagnosis, chemical/biological detection and other military applications.

The constraints of a wireless micro-sensor network are quite different from those of conventional wireless handheld devices. First of all, sensors have small packet size (hundreds of bits) and low average data rate (hundreds of bits per second) due to low event rates. Second, the transmission distance is very short, typically on the order of ten meters or less, and the communication link is asymmetric (i.e., traffic flow is mostly up-link from sensors to the base station). Third, sensors

have low mobility and form a quasi-static network. Last and most important, the battery life time of the sensor network is crucial and must be maximized [3, 4]. Due to such unique characteristics of the sensor network, design methodologies for conventional wireless devices would result in inefficient use of energy if they are applied to microsensor network.

From the sensor network's perspective, energy efficient transceiver implies the combination of the following parameters in a transmitter: fast start-up time, high data rate and low power consumption [5].

Fast start-up time is necessary in order to minimize the energy consumption when a sensor node is turned on from an off state. For short packet sizes, the transient energy during the start-up can be significantly higher than the energy required by the circuitry during the actual transmission. Several techniques have been suggested in order to reduce the start-up power loss reduction. The warm-up time can be dominated by the settling time of transients in the signal path, especially the (integrated active) channel select filters and frequency synthesizer. Wideband techniques, such as direct sequence spread spectrum (DSSS), have an advantage in that their wide channel filters have inherently short settling times. Also, with their greater channel spacing, DSSS frequency synthesizers may also employ higher frequency references, reducing lock time [5, 6]. A variable loop bandwidth technique is probably the most popular approach[1]; the idea is to start the PLL with a wide loop bandwidth and change it to a smaller loop bandwidth when the loop is closed to being locked [7-11].

High data rate reduces the transmit time of a packet and hence reduces the active time of the radio. The power consumption of the transmitter is dominated by the frequency synthesizer and is not affected by the data rate. Therefore, from a circuit perspective, a high data rate will allow lower

[1] And it is assumed that the proposed architecture in this thesis may adopt a similar approach in this regard.

transmitter energy consumption by reducing the transmit time of a fixed sized packet [4]. It is also shown that many types of batteries exhibit a recovery effect in which their lifetimes may be extended if current is drawn from them in bursts, rather than an equivalent average current. Therefore, the shorter the burst, the more battery life is extended [12, 13, 14].

1.2 Related Works

In industry, IEEE 802.15 Working Group has been formed for wireless personal area network (WPAN), which focuses on the development of consensus standards for short distance low rate wireless network [15]. The concept of Personal Area Networks (PANs) was first demonstrated by researchers at IBM in 1996 [17]. IBM engineers developed a method that utilized the human body to exchange digital information. Engineers used pico-amp currents through the body at very low frequencies, around 1MHz. The low power and frequencies prevented eavesdropping and interference to neighbouring PANs. The engineered system allowed two researchers to electronically exchange a business card with a handshake. IBM engineers met their goals of creating a way for communication between body-borne appliances by using the human body as the communications channel. The only problem was that this method required some form of human contact between devices. Human contact is not always wanted or possible and sometimes the devices doing the communicating are not body-borne but just in proximity to other PAN devices. To get around the limitation of human contact alternates to near-field communications such as infrared (IR) or far-field (radio) communications are being considered. Using wireless methods such at IR or radio frequency (RF) for PANs is known at Wireless Personal Area Networking (WPAN) [16-19].

In academia, there are several projects that involve wireless sensor network on various topics. One of the leading researches in this area is μAMPS (Micro Adaptive Multi-domain Power-aware

Sensors) projects that focus on developing a complete and flexible power-aware system for wireless sensor networks [20]. The goal is no longer simply the development of low-power techniques in hardware and software, but rather to create and develop a flexible platform that can adapt computation in order to trade-off quality and system lifetime [1]. Other research in the field of wireless sensor networks include *PicoRadio* [21, 22] and *WINS* [23] that focus on radio communication aspect of the sensor network and *smartDust* [24] that focuses on micro-electro-mechanical-system (MEMS) technology. Research has also been conducted in digital and RF circuits as well as the MEMS area that explore low power systems and circuit techniques as well as miniature implantable devices with energy harvesting techniques.

1.3 Contribution and Scope of Thesis

This thesis primarily introduces a new ***concept*** in transceiver architectures in which the receiver is used for the transmit purpose *as well*, or vice versa. This is more than just the "re-use" of certain blocks to merely save area or power. The transmitter formed by the re-configured receiver offers the opportunity to modulate *high bit-rate data*, such as video[2], or a short packet of slow-varying data, such as temperature, offering prospects of achieving an energy efficient transceiver. The modified high-IF dual conversion receiver uses two *correlated frequencies* as local oscillators (LO), but it allows *moving of the center frequency* of the LNA accordingly.

Although not limited to, the concept is presented as the architecture tailored to the characteristics of a wireless microsensor network to satisfy the expectations from such a link. It takes advantage of the fact that in most micro-sensor networks, the transmitter and the receiver (at each node) are not active at the same time. This creates an opportunity to maximize the "re-use" of building

[2] It is estimated that the system can support bit-rates up to 350-Kbps suitable for videoconferencing and videophone

blocks, which may result in a significant saving on the silicon real estate and a reduction in die cost.

Meanwhile, the architecture itself is designed such that it is potentially capable of receiving and transmitting phase or frequency modulated signal with minimum modifications. It allows the transmitter modulating a high bit rate data. This is achieved using a *dual loop frequency synthesizer* which also provides an opportunity to have more control on the start-up time and settling time. Power consumption is further minimized by using a few inexpensive off-chip components. The premise of such a combination is the manifestation of novel low cost, energy-efficient and configurable transceiver architecture.

Furthermore, in a parallel effort, the thesis challenges the idea of *partial channel-filtering* at early stages of the receive front-end and by incorporating it into the system, shows it may mitigate the dynamic range requirements of mixers and baseband channel-select filters which result in power savings at the cost of an increase in the complexity of the system. In this case, the frequency tuning is an indispensable part of the proposed architecture, while the bandwidth tuning provides an option to have more control on the passband (gain) of the LNA. The thesis proposes a *new tuning mechanism* which takes advantage of the existing architecture and *re-uses* the transmitter modulator and the receive chain in order to tune the LNA center frequency and bandwidth. This results in a considerable saving in hardware, which could only be used during the calibration; and just for a small fraction of the transceiver operating period - which tends to function more often in the transmit mode. The proposed tuning mechanism adds a new dimension to the architecture which turns it into a *potential candidate for developing a software defined radio.*

The prototype design is intended to demonstrate the functionality of the architecture in micro-sensor applications, but it can be adopted for any time-multiplexed standard, software defined radio and especially in hardware extensive structures such as phased-arrays.

1.4 Overview of Thesis

Major architectures for receivers and transmitters are reviewed in Chapter 2, which sets the stage for introducing the motivation and concept behind this thesis. In Chapter 3, the details of the proposed system, the architecture and its functionality, frequency plan, and link budgeting are explained followed by the circuit topologies used in major building blocks. Techniques are presented for generating the required local frequencies and low power building blocks. Characteristics of ultra-narrow band low noise amplifiers are studied and formulated in Chapter 4 to highlight trade-offs involved in the design of one the most demanding building blocks in the transceiver. The frequency and bandwidth tuning mechanism for the RF filter at the front end, in the receive mode, is presented in Chapter 5. The design of the test fixture, validation plan and measurement results are discussed in Chapter 6. The thesis is concluded in chapter 7 by suggesting modifications to the presented architecture, defining a road map towards achieving a fully wireless sensor node.

2 Transceiver Architectures: a brief review

2.1 Introduction

In this chapter different architectures commonly used in wireless data transmission systems are briefly reviewed and compared with respect to their performances. Although extensive, the review is far from complete and the sole intention is to create a language, as a common denominator, to facilitate the discussion on the proposed concepts in this thesis. The architectures are reviewed with no chronicle or purposely causal order and emphases are put on the issues less highlighted in the literature.

2.2 Receiver Architectures

The basic function of any receiver is to distinguish the desired signal from all other signals and to amplify the signal to a sufficient level for reliable detection. It has to tackle with many difficulties rising either from the environment where the receiver has to operate or from the fundamental

properties of various circuit functions and the non-idealities of their real life implementations. Complexity, cost, and power dissipation are some of the criteria used to evaluate the pros and cons of various architectures.

2.2.1 Conventional Super-heterodyne

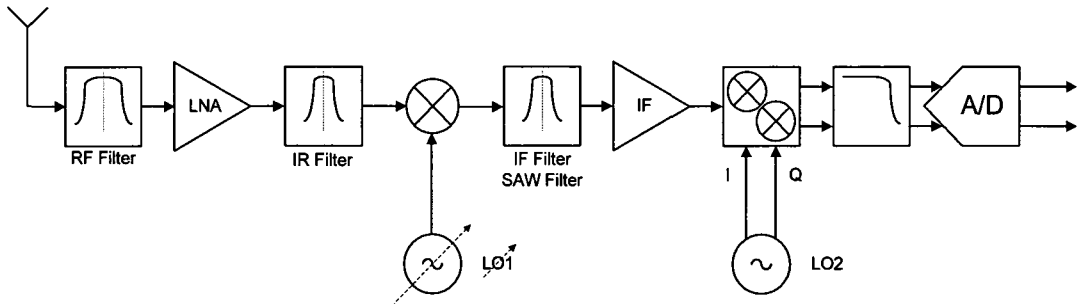


Figure 2.1 Schematic of a conventional Super-heterodyne receiver

Many RF Communication transceivers manufactured today utilize a conventional super-heterodyne approach shown in Fig.2.1. The off-chip front-end RF filter after the antenna serves to attenuate out-of-band energy and perform a partial rejection of the image band. The LNA and the following filter provide some amplification on the desire band and further attenuate the undesired signals present at the image frequency. An RF channel select frequency synthesizer (LO1) tunes the desired *band* to a *fixed* IF where a discrete component filter performs a first order attenuation of alternate channel energy. The IF filter in combination with a variable gain amplifier, reduces the distortion and dynamic range requirements of the subsequent receiver blocks.

A superior performance with respect to selectivity (a measure of a receiver's ability to separate the desired band about the carrier from signals received at other frequencies) and sensitivity (the minimum signal at the receiver input such that there is a sufficient signal-to-noise ratio at the

receiver output) can only be achieved with the use of high-Q discrete components found on a super-heterodyne receiver [25, 26, 27].

2.2.2 Direct Conversion

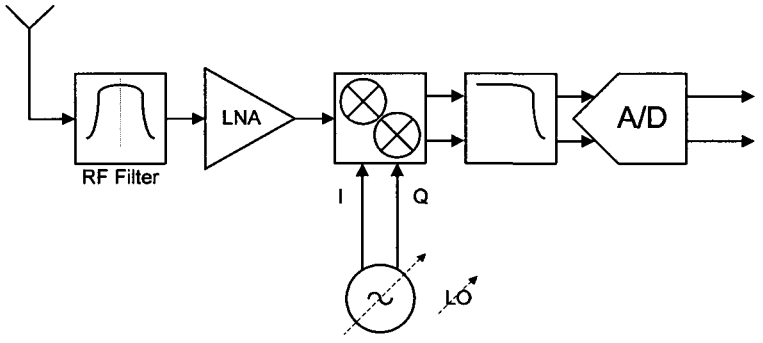


Figure 2.2 Schematic of a direct conversion receiver

One receiver architecture that eliminates many off-chip components in the receive signal path is the direct conversion architecture, also known as zero-IF or homodyne [28]. In this approach, shown in Fig.2.2, all of the in-band potential channels are frequency translated from the carrier directly to the baseband using a single mixer stage. Energy from undesired channels is easily removed with on-chip filtering at baseband. In a direct conversion receiver, the IF is eliminated as well as the need for image-rejection filtering.

Although the direct conversion receiver allows for higher levels of integration than a super-heterodyne system, problems are associated with this architecture. The adjacent channel blocker cannot be filtered out until the final baseband filter stage, which prevents the received in-band signal from having enough amplification to achieve a low noise figure comparable to the super-heterodyne. Also, because the local oscillator is at the same frequency as the RF carrier, the

potential exits for LO leakage to either the mixer input or to the antenna where radiation may occur. The unintentionally transmitted LO signal may reflect off the nearby objects and be "re-received", reducing the SNR at the output of the mixer and eventually the dynamic range of the receiver by pushing the baseband filter to saturation and its non-linear region [28, 29].

2.2.3 Low IF

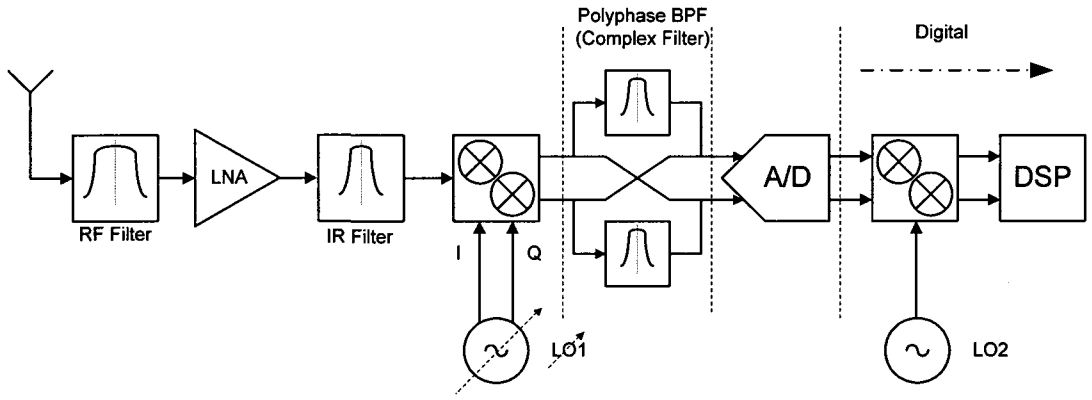


Figure 2.3 Schematic of a low IF receiver

A low-IF architecture combines the advantages of both super-heterodyne and direct conversion. It uses two down conversion paths in a dual IF receiver, and postpones the image rejection from high IF to the lower IF. A high quality RF filter at high IF would not be necessary. A broadband filter, similar to that of zero-IF would be satisfactory. The information at the lower IF, which is usually chosen to be at one or two times of the bandwidth of the desired signal, now can be extracted using a set of low Q bandpass filters. The two functions of suppressing the image and adjacent channels can be rolled into a higher order complex active channel-select Hilbert filter.

Imperfections in complex systems, such as mismatch between two paths, degrade the performance of the system. In digital domain matching can be perfect. In analog part, mismatch is unavoidable.

This mismatch might exist between I and Q local oscillators, similar to direct conversion, or the propagation delay between two parallel paths in the complex filter. Crosstalk between positive and negative frequency paths also reduces the SNR before ADC [30, 31].

2.2.4 Wide-band IF with Double Conversion

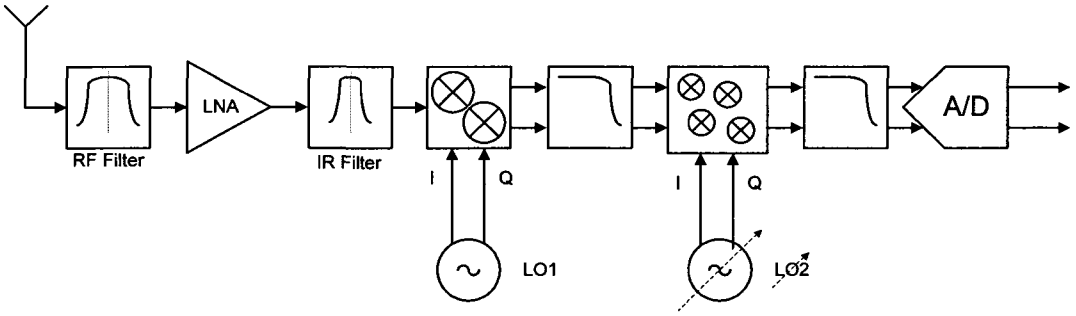


Figure 2.4 Schematic of a wide-band IF with double conversion

An alternative architecture well suited for integration of the entire receiver is wide-band IF with double conversion [32]. Shown in Fig.2.4, this receiver system takes all of the potential channels and frequency translates them from RF to IF using a mixer with a single frequency local oscillator. A simple low pass filter is used at IF to remove any up-converted frequency components, allowing all channels to pass to the second stage mixers. All channels at the IF are then frequency translated directly to baseband using a tuneable channel-select frequency synthesizer. Unlike a conventional super-heterodyne receiver, the first local oscillator frequency translates all of the received channels, maintaining a large bandwidth signal at IF. The channel-select filtering is then realized with the lower frequency tuneable second LO.

In the wide band IF receiver, the signal is mixed to a non-zero IF; therefore, image problem is reintroduced in the system, which can be handled using six mixer configuration similar to Weaver

technique [33]. Assuming that the up-converted terms are removed, the image rejection is very wide band, and the edge of the image attenuation band is set by the frequency of the first LO.

By removing the channel-select filter at the IF, strong adjacent channel interferers are now a concern for the second mixer stage in Weaver structure as well as baseband blocks. This implies a higher dynamic range requirement of these latter stages. In addition, the spurious tones generated by the IF LO can mix with undesired IF channels creating in-band interference at the output of the second mixer stage. Also as with conventional image reject mixer systems, the magnitude of the image attenuation in the wideband IF architecture is a function of the phase mismatch between I and Q of the first and second LO and matching between signal paths.

2.2.5 Sub-sampling Receiver

Bandpass sampling theorem proves that one can reconstruct the information content of the received RF signal by sampling it at an integer fraction of the carrier frequency which is greater than twice the bandwidth of the modulating signal [34]. Spectral images of the modulating signals are repeated, and down-conversion can be achieved by lowpass filtering the desired image at a lower frequency [35]. Sampling the carrier frequency f_c at a rate f_s results in spectral images located at $nf_s \pm f_c$ where n is an integer. A desired spectral image can be filtered using a discrete time analog filter.

Subsampling receiver eliminates the need for IF filters, image reject mixers, and analog I/Q branches thereby making these suited for integration. Further, the LO in such a scheme operates at much lower frequency than the carrier. However, this trades one set of problems with another. The

major issues in such receivers are noise-folding, sensitivity to clock jitter[3] and linearity requirement. Noise folding can be reduced by using a narrow band pre-select filter [36, 37, 38]. The sensitivity to timing jitter and the noise figure of the system can be improved by placing N sampling switches in parallel collecting successive samples that are dumped to an output buffer simultaneously every N clock cycles [39].

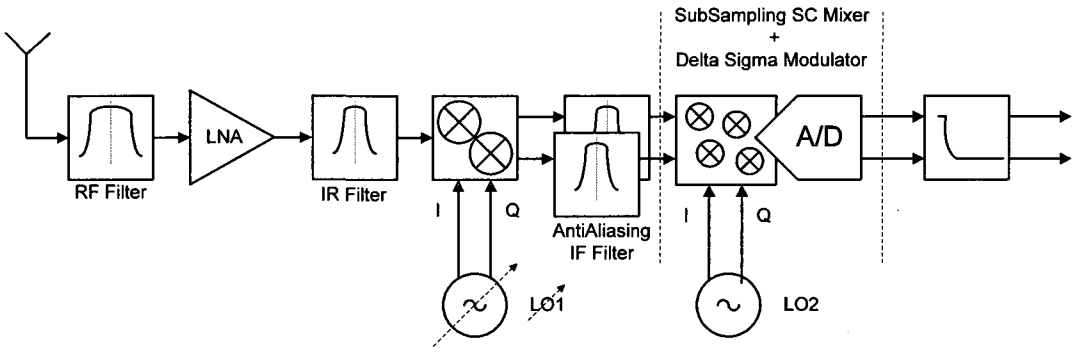


Figure 2.5 Schematic of a sub-sampling receiver employing IRSSADC

Azeredo-Leme [40], took a more ambitious approach to embed the subsampling switch-capacitor (SC) mixer in the ADC, and introduced the Image-Reject Sub-Sampling ADC (IRSSADC) technique in a wideband subsampling system. Shown in Fig.2.5, the required anti-aliasing filter before ADC is eliminated and its function is performed by IF filters. This approach results in a significant reduction of receiver complexity and a potential power dissipation reduction since fewer blocks are required. Two IRSSADCs are required, realizing a double-quadrature down-conversion from I and Q. The mixing and addition operations can be realized directly at the input SC branches of the ADC.

[3] The lower clock frequency does not reduce the jitter requirement; one might even argue that satisfying the jitter expectations at higher frequency is easier where LC oscillators can be used. They have shown a better performance with this regard with a comparable or even lower power consumption.

2.2.6 Bandpass Sigma-Delta Modulator

Digitizing signals early in a receiver, at a high intermediate frequency or even in the radio frequency stage, makes for flexibility, simple frequency plans, and low component count at the cost of demanding specifications on the ADC [41, 42]. Delta-Sigma modulators including both lowpass modulators and bandpass are preferred architectures for high resolution ADCs. Continuous time modulators are much faster and provide a certain amount anti-aliasing filtering [43- 48].

To reduce the power consumption of the digital signal processing stage following the modulator, the sampling rate must be made as low as possible. Gourgue [49] studied a bandpass sigma delta modulator architecture in which the RF or IF signal is quantized based on the subsampling theorem and fed back after an upconversion.

The most serious limitation with subsampled bandpass Sigma-Delta is the digital-to-analog converter virtually implemented in the feedback loop which behaves as a lowpass filter. The feedback signal is then multiplied by its transfer function, a zero order hold. The replica centered

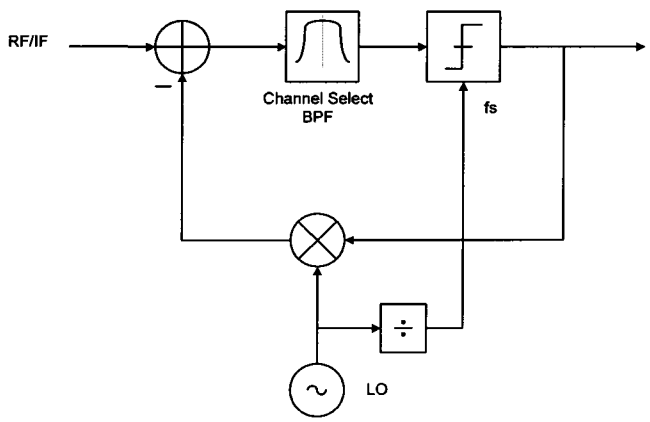


Figure 2.6 Schematic of a subsampling bandpass Sigma Delta modulator

on IF frequency is the only one, useful for the feedback which is attenuated by the DAC transfer function. To avoid the loss of the resolution encountered in the architecture, the modulator is modified by introducing a mixer in the feedback to transpose the main lobe of the feedback signal to the IF frequency [50]. The modulation frequency must be chosen such that the desired image, within [-fs/2 , fs/2], being upconverted to the IF/RF frequency. The study claims the possibility to develop a receiver offering the performances of a super-heterodyne scheme without requiring extra mixers to downconvert the IF signal to baseband.

2.2.7 Discrete-Time Receivers

(Charge Sampling and Direct RF Sampling)

The charge sampling technique was initially reported by Carley and Makherjee [51]. They showed the current mode sampling architecture with interleaving for sample and hold amplifier results in a substantial reduction in error due to sampling clock jitter and aperture time.

For a conventional voltage sampling circuit, the frequency response has a non-linear phase response which is a function of the product of sampling switch resistance and sampling capacitance (τ),

$$H(\omega) = \frac{1}{\sqrt{1+(\tau\omega)^2}} e^{-j\arctan(\tau\omega)} \qquad (2.1)$$

In order to make the phase shift as linear as possible, the value of $\tau\omega$ must be small enough. The -3dB bandwidth of the transfer function is also directly related to τ. Therefore the high frequency performance of the voltage sampling is directly limited by the time constant. Therefore the value of the sampling switch resistance and the tracking capacitance must be small enough to satisfy the constant group delay and small magnitude attenuation. Because of the required wide bandwidth,

after sampling, all the high frequency thermal noise in band will be folded into the Nyquist band [52].

Charge sampling performance, on the other hand, is directly determined by signal integrating window width which is defining the bandwidth; and the phase shift is linearly proportional to the frequency, resulting to a constant group delay. The sampling bandwidth can be narrowed to lower the folded noise. Even in order to reduce the effect of the device noise, such as flicker noise, charge sampling offers more space to choose larger capacitor (reducing thermal noise or KT/C) and switch area (reducing flicker noise). The capacitor can be large enough to minimize the clock feed-through which makes a high speed and high accuracy sampling system possible [52, 53, 54]. In [55], Yuan describes a charge sampling mixer with embedded bandpass filter. The bandpass charge sampling circuit is simultaneously a high-Q filter, a homodyne mixer and baseband signal sampler. The center frequency, the bandwidth, and adjacent selectivity can be programmed through the clock frequency and weighting function.

The concept was expanded to the quadrature bandpass charge sub-sampling of IF into baseband [56, 57]. The embedded complex FIR filter of the sampler can provide some of the required channel selection filtering and image band rejection in a receiver chain.

In a parallel effort, Poborzekiy [58, 59] proposed a different technique (in system level) to sample a narrowband channel with built-in anti-aliasing, after it is downconverted to zero IF. The wave-shaping built into mixing, followed by integration, realizes a programmable multi-rate filter. However, many functions merges into a mixer, making its realization impractical. A simplified version of this idea has been realised by Manku [60]. The report shows that the jitter performance of the wave-shaping clock is the main concern, although generating a high frequency non-uniform clock seems to be a challenge by itself.

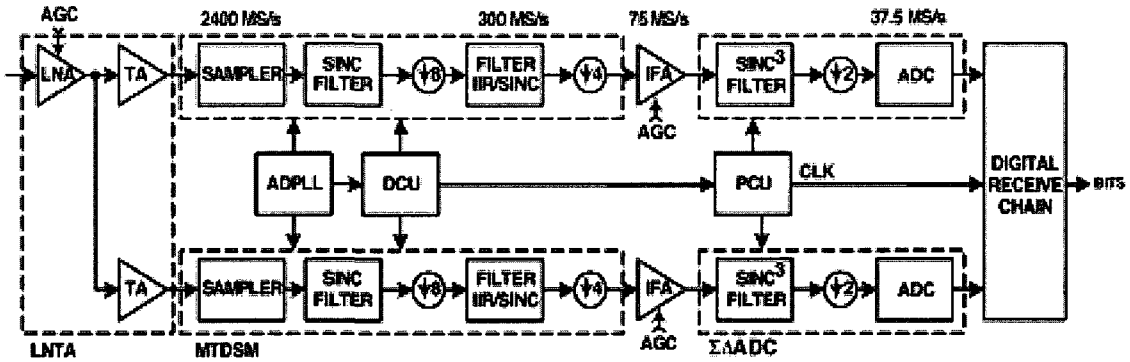


Figure 2.7 Direct RF sampling receiver by Texas Instrument [61]

Texas Instrument incorporated the charge sampling technique into a sequence of analog multi-rate signal processing, and introduced the Multi-Tap Direct Sampling Mixer (MTDSM) [61]. Shown in Fig.2.7, the input RF signal is first converted to a current, and then down-converted at the Nyquiste rate, integrated and dumped on a sampling capacitor

and filtered [62-64]. A large amount of programmable filtering and decimation gives them more flexibility than traditional receivers. The architecture uses a pre-select filter and tuned front-end for the respective band. A linear trans-conductance amplifier is a necessary part of the latter. Decimation in time is performed on the collective sampling and rotating capacitor while the incoming RF signal is accumulated on this total capacitance. This creates filtering which provides good anti-aliasing to the folding frequencies in narrowband RF standards such as Bluetooth. The IIR filtering relaxes the linearity requirements of subsequent analog blocks while a narrow selectivity could be achieved that is controlled by a capacitor ratio.

2.2.8 Super-regenerative Receiver

The super-regenerative receiver was first described in 1922 by Armstrong [65]. It was widely used in the Second World War, to develop Walkie-Talkie communication receivers, and as pulse responder for radar identification of ships and aircrafts (I.F.F.) [66]. Today, it is well suited for wireless short distance low-bit-rate data exchanges, like those in ISM band, because of its simplicity and low power consumption. It is mainly applied to demodulation of on-off keying (OOK) binary signals [69].

Super-regeneration is a process in which a circuit is made to oscillate intermittently, a repeated build-up and decay of self-oscillations in an oscillator whose free-running frequency is near or equal to the signal frequency. The start-up time of the self-oscillation is proportional to the power and frequency of the injected signal. In some topologies a periodic quench oscillator is used separately to allow the build-up and dampening of the self-oscillations. The incoming signal is sampled for a short period in each quench cycle.

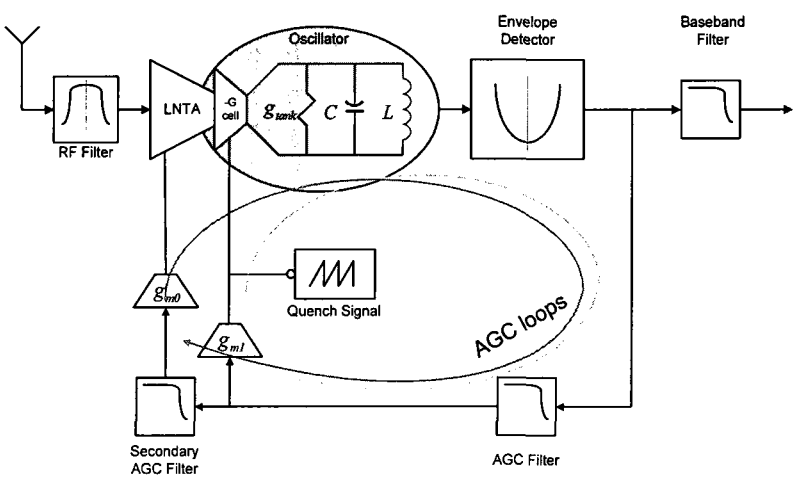


Figure 2.8 Schematic of a super-regenerative receiver

The quench frequency is much lower than the natural frequency of the super-regenerative oscillator, i.e. carrier frequency. Because of this internal sampling operation, the frequency of the modulation or carrier on the input signal must be less than half the sampling frequency in order for the information to be recovered without distortion [66]. Input noise and circuit noise may have a bandwidth much wider than half the sampling frequency; and therefore there is a situation in which the input signal does not experience folding in passing through the super-regenerative detector while the input noise and also noise generated within the parametric amplifier do experience folding. Such a system may have poor sensitivity and still have a low measured noise figure [67, 68].

Loughlin (1947) used an off-tune super-regenerative super-heterodyne receiver for frequency-modulated signals; and he called it "Fremodyne"[4]. Interestingly, it is observed that the distortion in the detector is less than that resulting from amplitude-modulated signal [66]. In a more recent research, a low power and low voltage super-regenerative receiver operating at 1GHz is published, in which several quenching strategies in different modes of operation are compared [69].

2.3 Transmitter Architectures

An RF transmitter performs modulation, upconversion, and power amplification. In the contrast to the variety of approaches invented for RF reception, transmitter architectures are found only in a few forms. The transmitter circuitry processes basically only one signal at a time, the strength and spectral contents of which is well-known. The challenges in transmitter design are thus more related to the spectral purity of the final modulated signal entering the antenna - particularly in

[4] Some might prefer to call this a FSK receiver based on the injection-locking mechanism [].

systems with many channels closely spaced in frequency, any kind of spectral deterioration may disturb adjacent channels. Thus, the attention in the following review is concentrated on the basic transmitter architectures used to realize the required vector modulation and not on the PA.

2.3.1 Direct Modulation

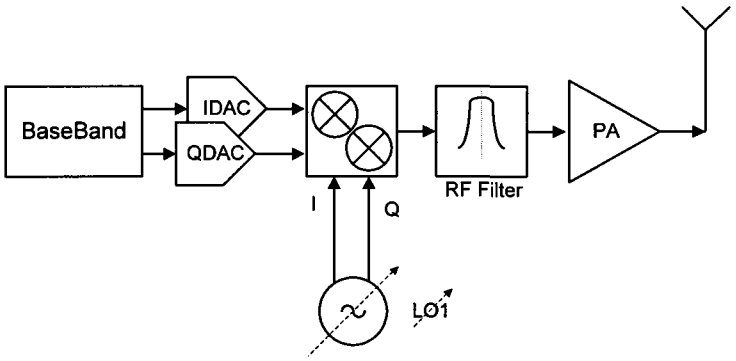


Figure 2.9 Schematic of a direct modulation transmitter

For linear modulations, the IQ or Cartesian upconversion transmitter is quite popular. Shown in Fig.2.9, two DACs driven by complex outputs from a digital baseband can synthesize an arbitrary modulation centered at 0 Hz, and two mixers driven by quadrature phases of an LO can upconvert this to the required carrier frequency without creating an image sideband. In reality non-linearity in the input DACs and LO drivers (such as hard limiting) creates inter-modulation components, spreading the spectrum around the ideal single tone SSB. Furthermore, the RF VCO generating the LOs operates at the output frequency. Therefore, depending on the quality of shielding, a re-modulation may occur by the high power modulated output of the PA to pull the frequency of local oscillator, which results in a phase error. The direct conversion requires a duplexer in the front-end because the noise floor generated by the modulator is quite high and can cause interference with the reception of a nearby receiver [70].

2.3.2 IF Modulation

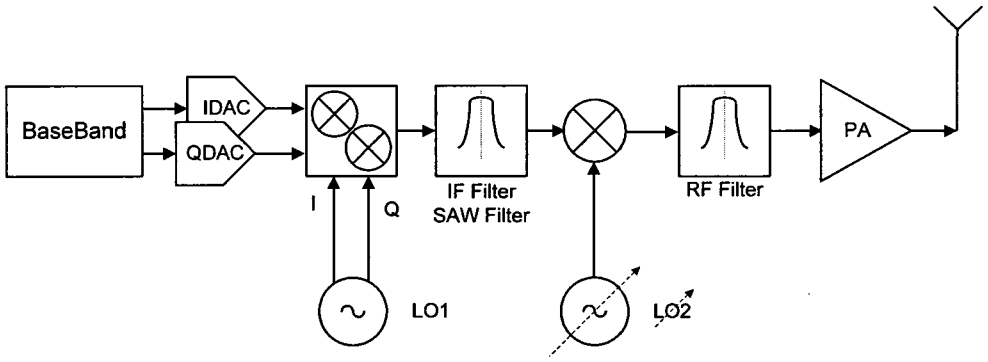


Figure 2.10 Schematic of an IF modulation transmitter

In order to avoid the pulling problem in the direct modulation architecture, almost all the commercially available implementations use an offset mixing scheme in order to generate the desired output. This will relax the isolation requirements. Moreover, a narrow IF SAW filter might be used to reduce the unwanted emissions. The IF modulation comes up with a lower noise floor but still high in meeting the specifications of some standards such as GSM; therefore it can not avoid the expensive duplexer. The modulation loops, explained next, are the system approach which avoids a duplexer.

2.3.3 Modulation Loops

Many wireless mobile products use nonlinear saturated power amplifiers (PAs) because they have higher efficiency than linear amplifiers. A nonlinear amplifier changes the signal amplitude by different amounts depending on the instantaneous amplitude of the signal. The more the amplitude of a signal varies, the more nonlinear amplification occurs, which will result in a distorted signal. Therefore, modulated signals with constant envelope are often preferred in wireless

communications. In constant envelope modulations, only phase or frequency information is employed to carry the user data, with the carrier amplitude being constant.

Architecture for low power FSK modulation based on a ΣΔ fractional-N synthesizer is shown in Fig.2.11. The idea is that fine frequency contents of the modulation can be generated by proper dithering of the divider value with a high resolution ΣΔ modulator [71]. The modulation is applied as a stream of digital words to the quantizer input, and controls the instantaneous VCO phase. The constant envelope phase modulated waveform suffers none of the adverse effects associated with the non-linear power amplifier, and yet benefits from its high efficiency. The main disadvantage is that the data rate is limited by the PLL loop bandwidth, since the PLL shows a lowpass transfer function from the ΣΔ. Although the loop bandwidth can be increased, this leads to higher power dissipation in the ΣΔ in order to maintain low quantization noise.

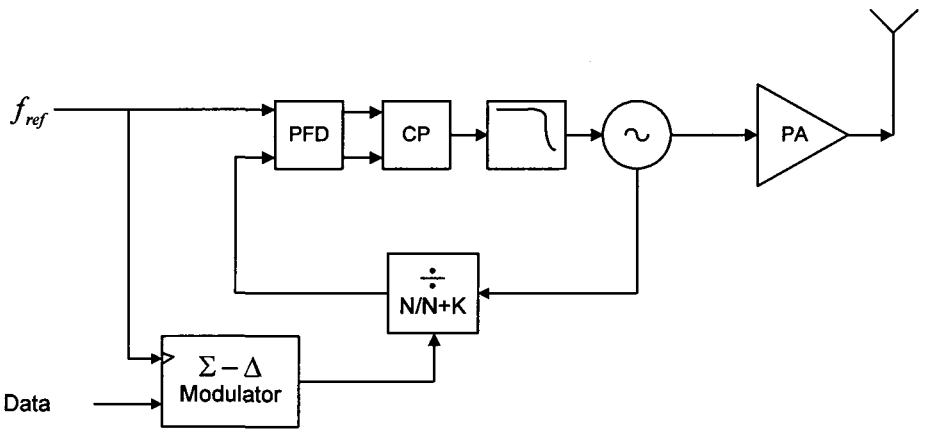


Figure 2.11 Schematic of a FSK modulator

To avoid the limitation of data rate from PLL loop bandwidth, a solution was suggested by Perrott [72], by adding a transmit pre-emphasis filter to the ΣΔ modulation path as shown in Fig.2.12.

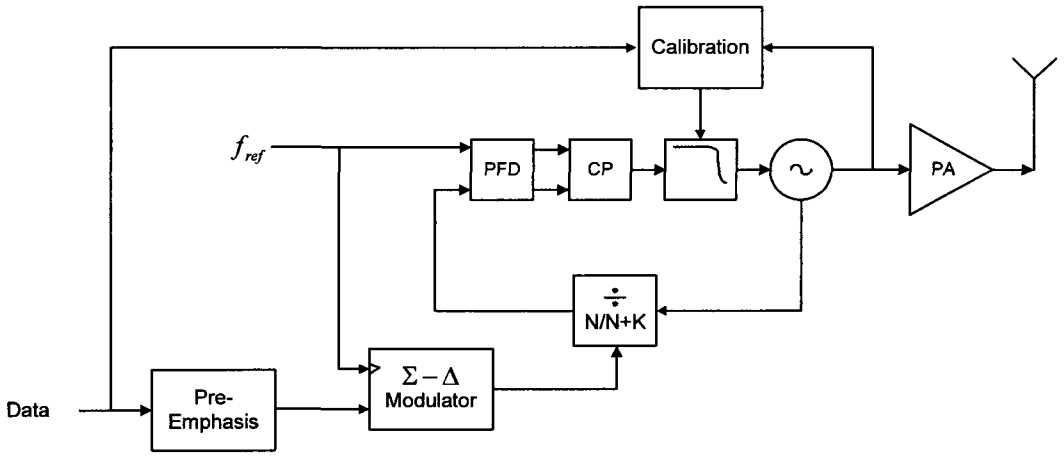


Figure 2.12 FSK modulator with pre-emphases filter in its data path

The modulated data and quantization noise each see different loop characteristics. While both the data and the quantization noise goes through the lowpass nature of the PLL, the data sees a much wider loop bandwidth with the help from the pre-emphasis filter. Hence, high data rate can be achieved even with a small loop bandwidth, which in turn enables low power consumption at the ΣΔ. A potential problem of this approach is the gain mismatch between the PLL loop and the pre-emphasis filter. If the data does not see a flat band over its spectrum, the modulated signal will be distorted. Since the PLL loop bandwidth depends on parameters that cannot be controlled precisely (e.g. VCO gain, charge pump current, etc), manual calibration of the PLL loop bandwidth or the pre-emphasis filter is necessary. The solution to this problem has been proposed by McMahill [73], by adding an automatic calibration circuitry. The basic idea is to compare the sampled output phase to the original data and correct the loop response by adjusting the loop gain.

Another popular method that achieves high data rate low power modulation is the open loop direct VCO modulation architecture. Shown in Fig.2.13, the underlying idea is to open the loop after the PLL settles to a desired frequency and directly modulate the VCO. Since the loop is open during modulation, the PLL does not affect the modulation in any way and hence the data rate is not limited by the PLL. Furthermore, power consumption is dramatically reduced since the VCO is the only component that needs to be turned on during transmission. A critical drawback of this approach however, is the instability of the VCO when the loop is open. The charge stored in the capacitor of the VCO control input will eventually leak away and cause the carrier frequency to drift. In addition, great care must be taken as to avoid any VCO pulling when the loop is opened [74].

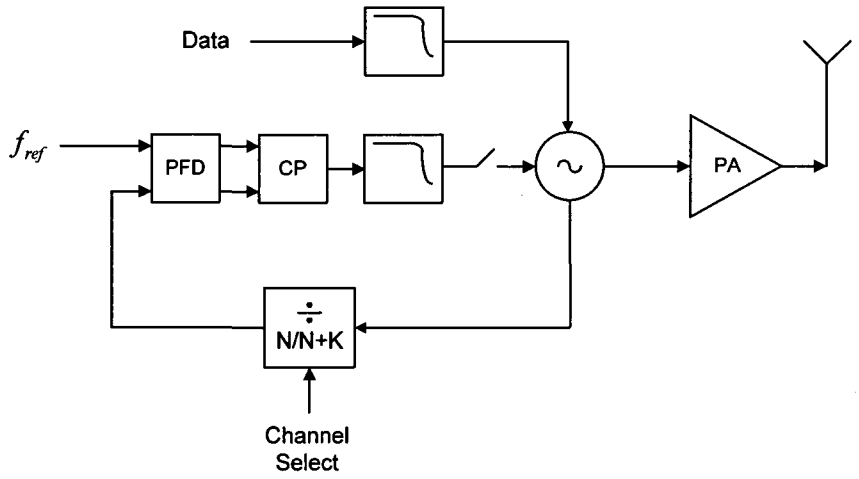


Figure 2.13 Open loop direct VCO modulation architecture

Closed loop direct VCO modulation, i.e. using the same structure as above without opening the loop [75], on the other hand is robust to these problems and still has the advantage that the upper bound on data rate is not affected by the PLL loop bandwidth. Since the data is modulated in closed loop, any drifts that are seen in the open loop architecture are no longer an issue. The disadvantage is that the modulated waveform will be distorted from the negative feedback loop of

the PLL. Since the PLL acts as a high pass filter when viewed from the VCO, low frequency components of the modulated data will be corrupted by the PLL. In other words, long strings of zeros or ones will not be modulated correctly.

The effect of PLL on modulation accuracy can be decreased by reducing the loop bandwidth, which is also prefered for low power frequency synthesizer. Unfortunately, even for an arbitrary small loop bandwidth, the error will eventually show up if the transmitted data has a sequence of zeros or ones that is longer than the time constant of the loop response. Closed loop modulation is more effective if the DC component of the data is removed. This can be done through Manchester encoding, which replaces 0's and 1's with transitions [76]. The drawback is that the effective data rate is halved.

2.3.4 TX modulating The Reference path

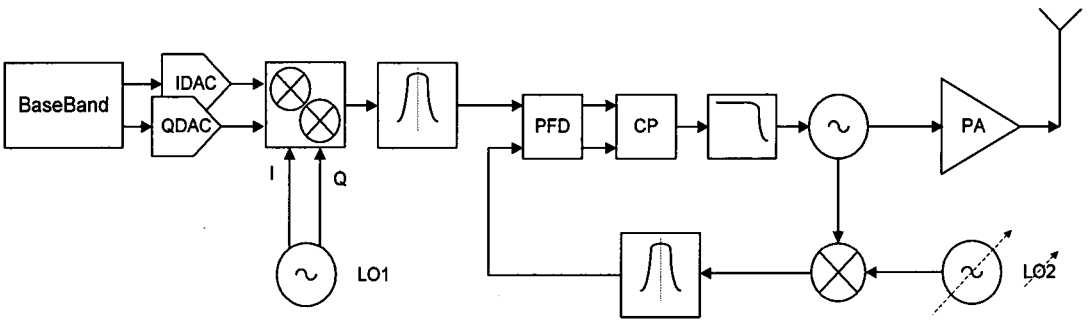


Figure 2.14 Schematic of a constant envelope modulator using frequency translational loop

A pure phase-modulating transmitter may be realized by embedding the VCO in an offset PLL, where a conventional I-Q Cartesian upconverter first creates an approximately constant envelope waveform at some IF, and then a PLL with a mixer in feedback translates this to RF; the phase-modulated RF VCO in the loop then drives the non-linear PA. Due to the high reference frequency, there are several hundred phase comparisons for one transmitted bit. Therefore the loop

is able to copy the phase modulation of the reference path onto the output of the power VCO, which is within the loop.

2.4 Conclusion

The brief review of the commonly practiced architectures, presented in this chapter, shows a wide spectrum of trade-offs within each and amongst different architectures. For example, while the super-heterodyne receiver provides the best performance in term of dynamic range, there are others promising different opportunities to realize a fully integrated radio. At the same time the latter might need to deal with image rejection issues, IQ mismatch, sampling jitter, higher power consumption, or other non-idealities for which the solution, if practically possible, doesn't come for free.

In reality, eventually, it is the *cost* of the commercial product which makes an idea survive its space-time or get buried inside ears; and cost is defined according to the nature of the product. Sometimes it is only referred to the Bill-Of-Material (BOM), and sometimes it has been interpreted as energy consumption, power dissipation, or even yield. Often it is related to the level of complexity of the idea and required skills/resources/man-hours to get the idea materialized, or its potential in being migrated and applied into different standards, protocols, or technologies within a certain period of time (time to market). It is also important to identify the subject of that cost; that is, whether it is the designer/producer/manufacturer or the end user/customer. For example, one might choose an architecture which needs a few off-chip components (an increase in the total cost of the product), but offers less energy consumption and as a result longer battery life time which favours consumers (and a delusive premise of capturing a bigger slice of the pie!). The proposed idea in this thesis promises a low-cost solution for wireless applications where the receiver and transmitter ought not to be operating at the same time. It offers a significant saving in

die cost (saves material for the manufacturer), a high bit-rate or short burst data transaction (saves energy or battery for consumers), and through the use of a few off-chip components (costs material for manufacturer) allows employing circuit topologies which dissipate less power (saves power for consumers).

3 Transceiver Architecture for Micro-Sensor Networks

A low power, energy efficient and configurable transceiver architecture is introduced. The circuitry employed in the high-IF dual conversion receiver is re-used in the dual loop frequency synthesizer to form a constant envelope modulator and function as the transmitter. Loops in the dual loop frequency synthesizer are coupled through an injection locking mechanism such that in addition to providing local oscillator frequencies for each mode, it allows a high bit rate data to be modulated in the transmit mode.

Although the prototype design is demonstrating the functionality of the architecture in micro-sensor applications, it can be easily adopted for any time-multiplexed system; it can also be scaled up in frequency and incorporated in phased-array antennas where hardware re-use results in a significant saving on the silicon real estate and a reduction in die cost.

3.1 Introduction

The constraints of a wireless microsensor network are different from those of conventional hand-held devices. They communicate small packet sizes, and the average data rate is low due to low event rate. Also, the traffic is mostly up-link from sensors to the base station with a very short transmission distance. Furthermore, the battery lifetime of the sensor network is crucial and must be maximized. Due to such unique characteristics of sensor networks, conventional design methodologies for wireless devices may result in an inefficient use of energy if they are applied to microsensor network.

The proposed architecture in this chapter is tailored to the aforementioned characteristics to satisfy the expectations from such a link. It takes advantage of the fact that in most micro-sensor networks, the transmitter and the receiver (at each node) are not active at the same time. This creates an opportunity to maximize the "re-use" of building blocks. The transmitter is activated more often than the receiver; and considering the facts that the packet size is small and transmit power required to communicate with a node at a short distance is low, special emphases is put on the transmitter capabilities to pass through the start-up time quickly and handle a high data rate while consuming very low power. The receiver is formed by reconfiguring the transmitter blocks, in particular the VCO and the mixer in the Frequency Translational Loop (FTL), to act as the LNA and the RF mixer, respectively, in a high-IF dual conversion scheme. The multipurpose circuitries put more stringent demands on the associated building blocks which results in emergence of new trade-offs.

3.2 SYSTEM ARCHITECTURE

The proposed transceiver architecture is shown in Fig.3.1. The receiver (Rx) is a high-IF dual conversion system. The captured signal is first filtered by the pre-select RF filter with a deep notch at the image band; A fraction of the desired band is then amplified by the tunable LNA and down-converted to the high intermediate frequency (IF). The bandpass filter at the IF further attenuates the blockers, facilitating the design of the quadrature IF mixers and baseband stages. The IQ mixers translate the channels at the IF

to the baseband where the channel-select filters extract the desired channel from around dc.

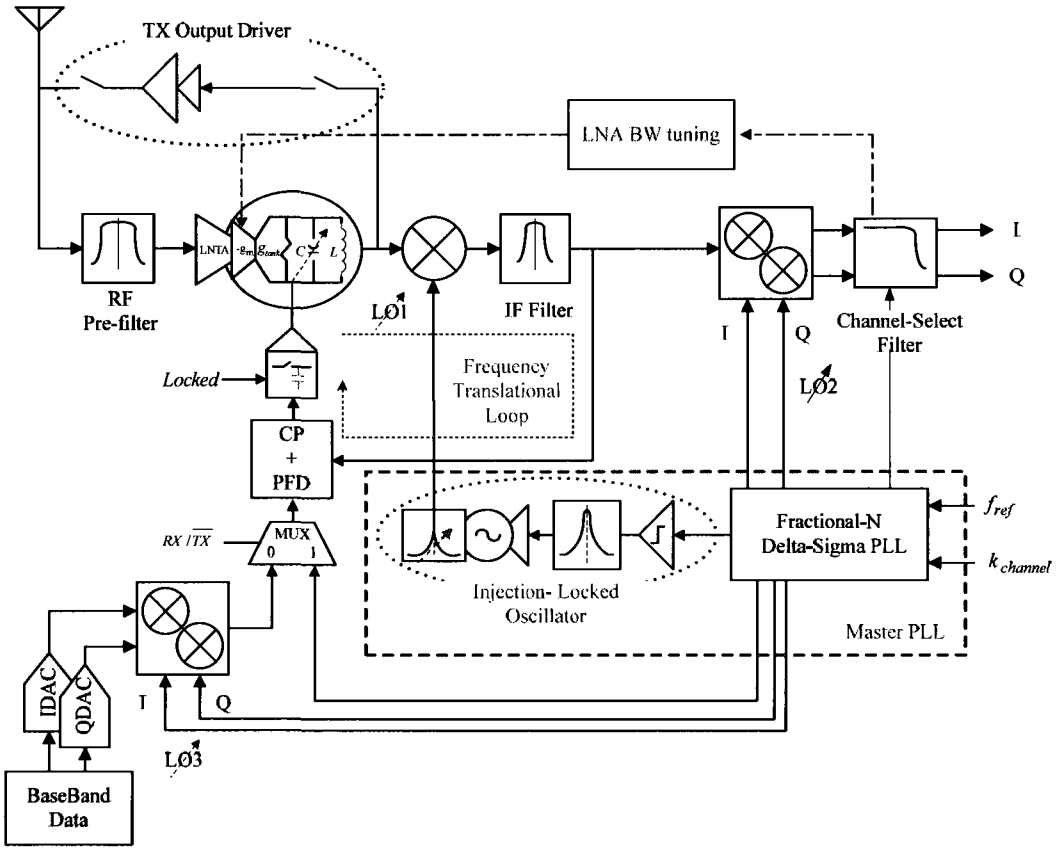


Figure 3.1 Schematic of the proposed transceiver architecture

In the receive mode, the LNA is first tuned to the desired channel. The Frequency Translational Loop (FTL) is used to simply switch its PFD reference frequency from the transmitter IQ-mixer to LO2. Once the loop is settled, i.e. LNA is locked to the band of interest, the loop is opened and the tuning voltage is stored in the sampling capacitor of the loop filter. The bandwidth of the LNA can be adjusted to a set of discrete values. The narrow band LNA further attenuates out-of-band blockers and amplifies the in-band blockers with less gain than that of the desired channel. This relieves the phase noise expected from LO1 at high offset frequencies. Due to the high frequency LO1, the RF mixer suffers from large 1/f noise at its output, but this does not corrupt the signal, which lies at the IF. At the output of the second mixer, the flicker noise is very important since the signal is now down-converted to zero. Since, the first IF is well bellow RF, the flicker noise at the second mixer output is proportionally lower. Besides, the second LO, is at lower frequency and can be a sharp square-wave signal generated by the ring oscillator and its CMOS buffers at the driving nodes. The sharp transition in the LO signal reduces the time period in which both LO switches are ON and contribute noise at the mixer output.

The transmitter (Tx), on the other hand, is a constant envelope modulator which exploits the same frequency translational loop (FTL) used in tuning the center frequency of the LNA in the receiver. In the transmit mode, the LNA, which has already been pushed into oscillation, functions as the VCO for the FTL. The QPSK data is first upconverted to the IF using LO3 and then fed into the Phase-Frequency Detector (PFD) of the FTL. Due to the high reference frequency, there are several hundred phase comparisons for one transmitted bit. Therefore, the loop is able to copy the phase modulation of the reference path onto the output of the VCO, which is within the loop. Furthermore, none of the local oscillator signals in the FTL is at the carrier frequency, and thus the VCO will not suffer from frequency pulling. Since the FTL reference frequency, i.e. LO3, is high,

a wider loop bandwidth is desirable in order to transfer a high bit rate up-converted data to the output of the modulator. In addition, the wider bandwidth reduces the settling time and as a result the energy consumption of the transceiver.

The receiver and the transmitter share the master PLL (MPLL) comprising the Fractional-N Delta-Sigma PLL and the frequency multiplier, which are providing the required local oscillator frequencies, LO1, LO2 and LO3, for IF and RF mixers. The dual-loop frequency synthesizer formed by the chained MPLL and FTL plays a major role in breaking many trade-offs that exists in regular architectures. The MPLL provides a high reference frequency for the second loop (FTL) allowing the modulator handling a high data rate through the wide loop-bandwidth as well as settling quickly. The same synthesizer is used during calibration of the receiver and tuning the LNA.

This is achieved using a dual loop frequency synthesizer which also provides an opportunity to have more control on the start-up time and settling time. A master PLL provides a high reference frequency for the modulating loop with a high loop bandwidth. The coupling is performed through the injection lock mechanism. The sequential combination of the loops allows a more elaborate control on the settling time of the synthesizer in different scenarios. The modulator loop locks quickly because of its high loop bandwidth.

3.3 Frequency Plan

The receiver adopts high IF dual-conversion scheme with variable correlated LOs and a tuneable LNA. In order to select any of the 16 channels located between 2400MHz and 2480MHz, the master PLL (MPLL) generates IF frequencies, LO2 (LO3 in Tx mode), between 300-MHz to 310-MHz; LO1 is the 7^{th} harmonic of the IF, 2100-MHz to 2170-MHz, produced by the frequency

multiplier through the injection locking mechanism. This combination results in the downconversion of the desired channel to the baseband and requires the IF-filter to be transparent only for 12MHz.

The dielectric pre-select RF filter, has a passband of 100MHz centered around 2.45GHz with a maximum 2.2dB insertion loss; and it attenuates the image band, around 1800MHz, by at least 60dB. The frequency characteristics of such a filter is shown in the Fig.3.2.The existence of the family of off-chip filters with such characteristics is one of the main factors defining the frequency plan and the choice of the receiver architecture in this project. This level of attenuation is crucial to the overall performance of the system considering the fact that several other standard such as DECT, DCS1800, and PCS1900 are operating at nearby frequencies as shown in Fig.17. The pre-select filter simplifies the receiver architecture by choosing a high-IF without being concerned about image issues.

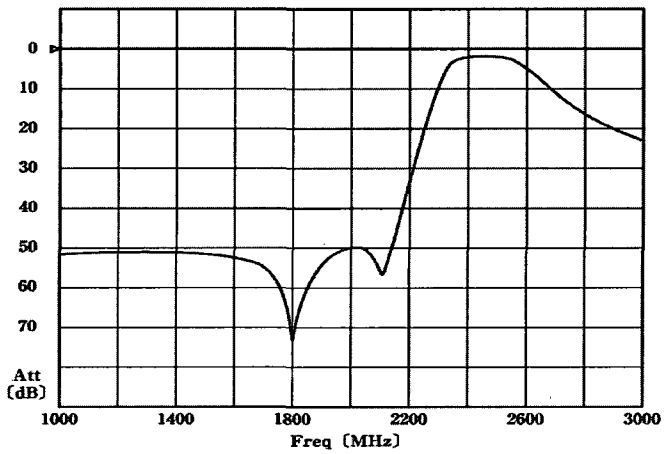


Figure 3.2 Frequency response of the pre-select filter[5]

[5] MDR747F by Soshin

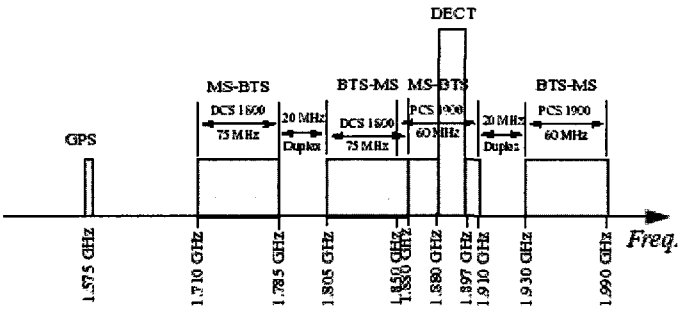


Figure 3.3 standards operating at potential image bands

The impact of the choice of the IF on a few other parameters in the system is listed in the Table.1, where it is assumed that the system needs to cover 100MHz bandwidth. The table shows how the choice of the multiplication factor in the dual-loop synthesizer affects the local oscillator frequencies and their relative offsets, location of the image band and phase noise penalty due to the injection-locking mechanism.

Several other factors influence the choice of the IF frequency;

1. sub-optimal budgeting of delay, noise, and power in the chain of the fractional-N PLL, frequency multiplier and FTL; for example, choosing a lower IF frequency demands a higher multiplication factor from the injection lock system, resulting in a longer settling time which is not desirable in an energy-efficient network. The noise penalty caused by the injection locking mechanism might be balanced out with the fact that the ring oscillator in the PLL is operating at lower frequencies and therefore a better phase noise can be achieved at the same offset. Injecting a stronger harmonic helps to speed up the transition at the cost of higher power consumption.
2. hardware power consumption and bandwidth limitations; CMOS topologies are employed throughout the system, where it is possible, to save static power consumption at the low

frequency part of the design such as the Fractional-N PLL. For example, the TSPC topologies used in dividers can only operate up to 400MHz with reasonable power consumption.

3. proximity of LO1 to the carrier frequency and frequency pulling in the transmit mode; that is, a lower IF frequency needs a higher LO1 which brings the free-running oscillation frequency of the injection-locked VCO closer to the carrier frequency in transmit mode which makes it prone to frequency pulling. This defines the lower limit for the IF.
4. locking range expected from injection locked oscillator; it implies that a lower IF frequency requires a wider tuning/locking range from the injection locked VCO.
5. A lower IF allows the use of a narrow-band IF filter which might be taken as an opportunity to further attenuate undesired channels as well as to save power in the RF mixer for a given conversion gain. It also reduces the dynamic range requirement of baseband stage.

In order to balance the aforementioned trade-offs the optimum choice is an IF in the neighbourhood of 300MHz, highlighted in the Table.1.

Setting the second IF to zero has the advantage that there will be no need for an image reject filter after the RF mixer. It also allows the re-use of the receive chain to extract the envelope required by the LNA bandwidth tuning circuitry. However, the IF filter - an off-chip LC load tuned at IF frequency - is placed after the RF mixer to reject the un-wanted products of the mixing, particularly $\omega_{RF} + n\omega_{IF}$ in order to prevent them from circulating in the second loop of the

Table 1 some of the impacts of the choice of the IF on other features in the system

Multiplication Factor M	IF/LO2/LO3[6] (MHz)	LO1 (MHz)	Phase Noise[7] (dB)	Image band (MHz)
3	600-625	1800-1875	9.5	1200-1250
5	400-417	2000-2085	14	1600-1668
7	300-313	2100-2191	17	1800-1878
9	240-250	2160-2250	19	1920-2000
11	220-228	2180-2272	21	1960-2044

synthesizer or FTL. Harmonic content as well as the non-linearity the first mixing operation has a significant impact on the functionality of the FTL and the spectral purity of the transmit signal.

3.4 General System Requirements

According to the specifications of WPAN standard, the receiver is required to detect signals as small as -85dBm. Using the definition of the noise figure, the sensitivity of the receiver can be calculated based on the noise floor power density at the input, the channel bandwidth and minimum SNR at the output of the receiver required by the baseband modulator to guarantee certain BER. That is

$$NF = Sensitivity - NoiseFloor - SNR_O \tag{3.1}$$

Or

$$NF = -80dBm - (-174dBm + 10 * \log(2MHz)) - 5dB = 26dB$$

This is the noise figure budget for the entire receiver. Moving along the receive chain starting from the Antenna, as we pass gain stages the contribution of the following stages to the noise figure is

[6] IF=RF/(M+1)

[7] This only counts on the noise degradation due to the injection locking mechanism when the injected harmonic is located perfectly at the free-running frequency of the slave oscillator and therefore can be estimated by 20log(M).

reduced by the total power gain preceding them. Therefore, the first few stages, such as Antenna, Duplexer, matching network, LNA and RF mixer have the dominant impact on defining the noise figure of the system. Leaving a margin of 5dB, the target noise figure is assumed to be 21dB.

The allowable gain in the receive chain is calculated based on the maximum differential swing that can be handled by the baseband circuits. Assuming a swing of 1 volt or a peak signal handling capability of P_{peak} equal to 13dBm, and 5dB margin in the case the receiver gain is larger than expected, and if the worst case blocker which is going to amplified by the gain is 40dB stronger than the minimum detectable level of the signal, that is -40dBm, then

$$\text{Gain} = P_{peak} - P_{Blocker} - 5\text{dB} \tag{3.2}$$

Which results is a maximum allowable gain of 48dB in the receiver. Note that here it is assumed that the worst case blocker is not attenuated in the receive chain; this is not the case in the proposed architecture where the preselect filter is limiting the bandwidth of the signal delivered to the LNA; besides the RF mixer has only 12MHz bandwidth and if the LNA is used in its ultra-narrowband mode, the worst case blocker power is attenuated before reaching to the baseband[8]. Taking this attenuation, $ATN_{blocker}$, into account results in a higher allowable gain, and as it is shown shortly, higher minimum detectable signal at the input of the baseband and lower required dynamic range. Therefore eq.3.2 can be modified by the amount of $ATN_{blocker}$, or

$$\text{Gain}_{ATN} = P_{peak} - (P_{Blocker} - ATN_{blocker}) - 5\text{dB} \tag{3.3}$$

Assuming an overall attenuation of 10dB, the maximum allowable gain is 58dB.

Assuming the noise contribution of the baseband circuitry to the noise figure is negligible, the available noise power at the input of the LNA (noise floor) is amplified by the gain, and appears as

[8] If the blockers are located at offsets well beyond the bandwidth of the LNA or RF mixer.

the larger noise floor at the input of the baseband circuits; therefore the minimum detectable signal at the input of the baseband stage is

$$P_{\min BBin} = (-174 + 10\log(B)) + Gain_{ATN} + NF - 10dB \quad (3.4)$$

in which 10dB margin is included; and for 2MHz channel bandwidth, results in -42dBm.

An additional 5dB of margin is required on the noise floor at the input of the baseband in the case the gain of the receiver is lower than expected. This yields a total minimum detectable level of signal of -47dBm or -60dBV at the input of the baseband[9]. Considering the 1 volt peak signal handling capability of the baseband (0dBV), the overall dynamic range is 60dB. This requires a 10-bit resolution ADC in the baseband[10]. Note that, if the blocker is not attenuated by the time it reaches the baseband the dynamic range would be 70dB and a 12-bit ADC would have been required. The two bits difference typically translates into sixteen times higher power consumption at the data converter [11][101].

The baseband circuitry needs to be linear to the level that their inter-modulation[12] products generated by two out-of-band blockers are still small enough relative to the minimum detectable level of the signal, to meet the required output SNR. That is

$$IM_3 = Sensitivity - SNR_O - 10dB \quad (3.5)$$

The 10dB deduction is the margin required to make sure that this is noise rather than distortion limits the overall receiver performance. Eq.3.4 results in a maximum input referred IM3 of

[9] dBV=dBm-13dB

[10] In reality one need to use 11-bit considering the ADC suffers from its own non-linearity and therefore the effective number of bits (ENOB) is lower than the nominal.

[11] In analog circuits with performance limited by the thermal noise, every extra bit costs four time extra power consumption. This is because each bit corresponds to 6dB of the dynamic range; therefore increasing the DR by a single bit, translates into 6dB or four times less noise power. This requires four times larger capacitor in the sampler; and in order to keep the time constant of the sample-and-hold constant, the transconductance of the sampler must be increased by four times. The power dissipation is proportional to this transconductance and thus increases by four times (for a given supply voltage and over-drive voltage).

[12] Here we only consider the third order inter-modulation products.

-95dBm.

Based on the fact that the input power (in this case the out-of-band blocker) and the IM3 vary with rates of 10dB/decade and 30dB/decade respectively, one can calculate the overall out-of-band third order intercept point of the receiver, $IIP3_{Receiver}$, using the following geometric relationship [26]

$$IIP3_{Receiver} = 1.5P_{Blocker} - 0.5IM_3 \tag{3.6}$$

Eq.3.5 results in an overall receiver out-of-band input third order intercept point of -12.5dBm.

Once again if the out-of-band blocker is attenuated in the receive chain before being applied to the baseband, eq.3.6 needs to be modified to

$$IIP3_{Receiver} = 1.5(P_{Blocker} - ATN_{blocker}) - 0.5IM_3 \tag{3.7}$$

The overall $IIP3_{Receiver}$ needs to be -27.5dBm.

If the linearity of the baseband circuitry is dominant (since it has to deal with the large amplified signal), and therefore assuming the rest of the receiver contributes only 1dB to the overall $IIP3_{Receiver}$, that is -11.5dBm, $IIP3_{rest}$,one can estimate the input referred $IIP3_{BB}$ of the baseband circuitry using the following relationship

$$\frac{1}{IIP3_{Receiver}} = \frac{1}{IIP3_{rest}} + \frac{Gain}{IIP3_{BB}} \tag{3.8}$$

which results in $IIP3_{BB}$ of about 42dBm. For the case where the blocker is attenuated and higher gain is allowed[13], one needs to be aware of the non-linearity of the filtering process in the rest of the chain (the above assumption of 1dB for the rest might not be valid anymore). However, assuming the filtering is conducted only using high quality passive components (possibly off-chip), the 1dB assumption can still be valid which in that case eq.3.8 is modified to

[13] Note that this is the maximum gain and therefore gain can be designed to be lower!

$$\frac{1}{IIP3_{Receiver}} = \frac{1}{IIP3_{rest}} + \frac{Gain_{ATN}}{IIP3_{BB}} \tag{3.9}$$

Which results in an input referred baseband intercept point of 78.6dBm.The larger number is due to the use of the maximum allowable gain.

If non-linearity of the rest of the receiver is significant (similar to the case where an active load is used in the ultra-narrow band LNA), and assuming the same amount of $ATN_{blocker}$, i.e 10dB, but using an equivalent amount of $IIP3_{BB}$ given by eq.3.8, the minimum linearity required of the rest of the chain would be at least -27.2dBm. This budget needs to be partitioned between the LNA and two mixers.

In order to define the requirements for the phase noise at the output of the synthesizer (RF LO), one method would be to assume that the receiver channel is noiseless and only interference produced within the signal band moving through the receive chain is due to the phase noise reciprocal mixing with out-of-band blockers[14]. Phase noise is assumed to be flat across the channel of interest at a certain offset from carrier (RF LO). The interference component that is then produced when the blocker mixes with the phase noise sideband is then compared to the desired signal which mixes with the LO energy. Based on minimum Carrier-to-Interference Ratio, CIR, the output of the mixer, the blocker level along with the position in frequency relative to the signal , Δf, and the desired signal, the required phase performance of the synthesizer in dBc/Hz may be estimated using,

$$PN(\Delta f) = P_{signal} - P_{blocker}(\Delta f) - CIR_{min} - 10\log(B) \tag{3.10}$$

[14] This is a rough assumption though. A better picture of the true CIR at the output of the receiver should include the white noise added to the desired signal band as well as the effect of blockers reciprocal mixing with phase noise and the effects of gain compression in the receiver signal path due to a large blocking signal [100].

Considering the channel bandwidth of 5MHz, for the minimum detectable desired signal of -80dBm, an adjacent channel blocker of -40dBm, and a minimum CIR of 5dB, the phase noise of the RF LO at 5MHz offset needs to be less than -112dBc/Hz. Attenuated blockers at larger offsets reduce the maximum phase noise requirements at those offsets. The requirements for phase noise at closer offsets are tighter as direct mixing becomes the dominant producer of the interference. At the limit, when the offset frequency falls well within the desired channel bandwidth, one may assume that the signal blocks itself (perfect overlap), through direct mixing, and therefore the first two terms in eq.3.10 cancel out each other and the remaining is an estimate for the phase noise at the close-in offsets. If minimum CIR stays the same, eq.3.10 results in the close-in phase noise[15] of -72dBc/Hz.

3.5 Circuit Implementations

The major building blocks designed for the transceiver are briefly introduced. The intention is to highlight the challenges and trade-offs involved in each design and to report their specifications extracted based on simulation. Some of the successful measurement results are reported in chapter 6.

[15] Note that this is the estimate for the extremely small offset frequency of zero; and it is calculated to give an idea about its difference from the phase noise requirements at larger offsets. A more accurate estimate needs to consider the type of the modulation as well.

3.5.1 RxLNA or TxVCO

A common-gate topology with impedance transformation at its input is chosen for the LNA (shown in Fig.3.4). In order to decrease the power dissipation while achieving a relatively low Q matching network at its input [79]. Cascode transistors are added in order to further improve the input-output isolation. The isolation and wide-band matching network are crucial since the tuning of the LNA load affects the input matching characteristics. The cross-coupled transistors are used for the tuning the bandwidth of the LNA. They are designed such that for certain settings the LNA is pushed into an oscillatory state and act as the VCO. The bandwidth tuning is performed by the digitally controlled source degenerating NMOS transistors.

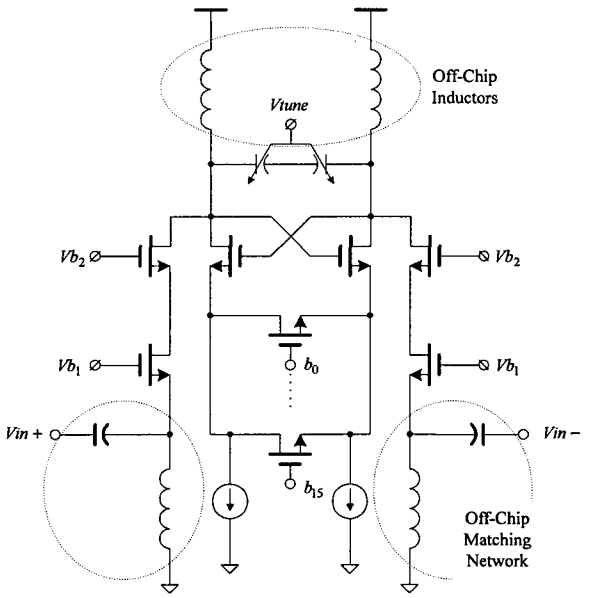


Figure 3. 4 Common-gate topology with input impedance transformation chosen for the LNA stage

The LC tank inductors are realized by bond-wires connecting the pads to the PCB, in series with high-Q off-chip inductors. The capacitor is made of the parallel combination of a set of switched MIM capacitors and a pair of MOS varactors. The varactors are designed using PMOS transistors to protect them from substrate noise coupling. A tuning range of 150MHz is achieved around the neighbourhood of 2.44GHz. The Kvco is kept relatively low in order to reduce the sensitivity of the tuning node in the receive mode when the FTL is opened. In the receive mode the achievable gain varies depending on the bandwidth setting. The LNA has a gain of 18dB when it is tuned to a

20MHz bandwidth. Such a high gain reduces the impact of the following stages on the sensitivity of the receiver.

The center frequency of the LNA is tuned first by switching to the VCO mode; then by turning off certain number of the degenerating transistors in the –gm-cell, it is switched into the amplification mode, and as a result, the Q of the tank is changed. During the tuning, when switching from VCO mode to the LNA mode with a limited but high Q , the center frequency of the LNA changes, however, the desired channel still falls in the extended bandwidth, which is what matters at the end of the tuning process. Also, the center frequency might change in the normal operation when the LNA is receiving small inputs due to voltage dependant parasitic capacitances. This sets a limit on the minimum tuneable bandwidth, introducing a new trade-off between the maximum swings allowed in VCO mode and the linearity of the LNA. The smaller the VCO swing is, the lower the LNA linearity gets, albeit the narrower the tuneable bandwidth is going to be. Therefore a co-design methodology has been developed to find the optimum setting.

The maximum achievable dynamic range is shown to be inversely proportional to the quality factor of a bandpass filter [80]. Tuning the LNA for a narrower bandwidth degrades its linearity. Fig.3.5 depicts the simulation results for a -1-dB compression point test when the bandwidth is tuned for 2MHz, or the channel bandwidth. Considering the fact that in WPAN the maximum signal power at the receiver is -20dBm, a compression point of -25dBm at the LNA already defies the possibility of the receiver being a standard compliant. This sets a limit on how narrow the filter could be in a given standard.

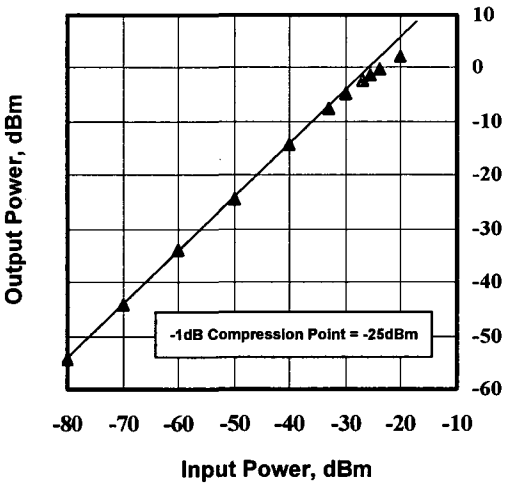


Figure 3.5 1-dB compression point simulation result for the narrow-band LNA

3.5.2 Mixers

The RF mixer is designed based on the doubly-balanced Gilbert cell topology with a parallel LC tank as the load, which also acts as the IF filter. A 470nH off-chip inductor with a relatively high Q of about 15 is used to achieve a 2dB conversion gain with a very small biasing current of 58μA. This choice allows the IF filter to be transparent for a 20-MHz of bandwidth at 305MHz. The IF bandpass filter rejects the image frequency of RF+LO1 which is around 2.75GHz. The IIP3 of the RF mixer is mainly determined by the overdrive voltage of its RF input transistors. A supply voltage of 1.6V allows choosing a Vdsat of 400mV which results in about 15dBm IIP3. The linearity of the RF mixer has a significant impact on the phase noise performance of the FTL.

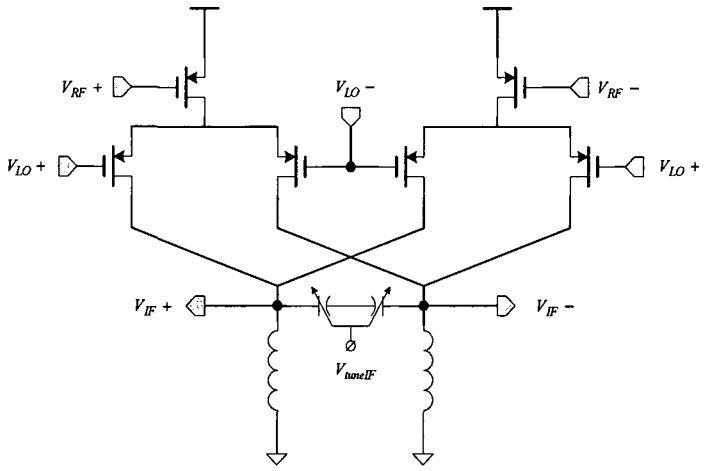


Figure 3.6 standard Gilbert cell used as the RF mixer; the narrow-band tuned load uses a relatively large off-chip inductor

Differential to Single-ended converter (shown in Fig.3.7) is used to translate the differential output of the RF mixer into a single ended CMOS level for the PFD in the FTL.

A pair of IQ mixers forms the IF mixer in the receive chain. PMOS transistors are used as the input transconductor and

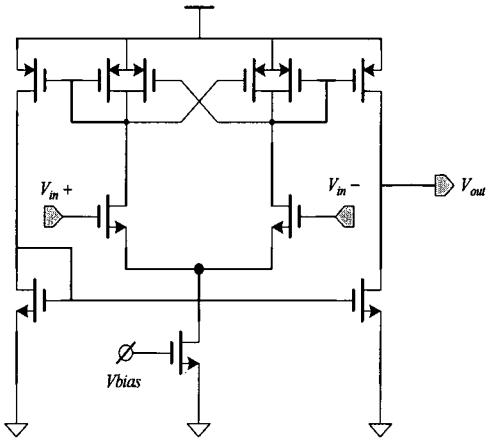


Figure 3.7 differential to single-ended stage used to convert the harmonic signal to the saturated CMOS level signal

switches for their lower flicker noise power. Each mixer adopts a Gilbert topology with diode connected NMOS as the load. This load also acts as the gm-cell for the first Biquad filter in the baseband channel select filter which saves an intermediate trans-conductance

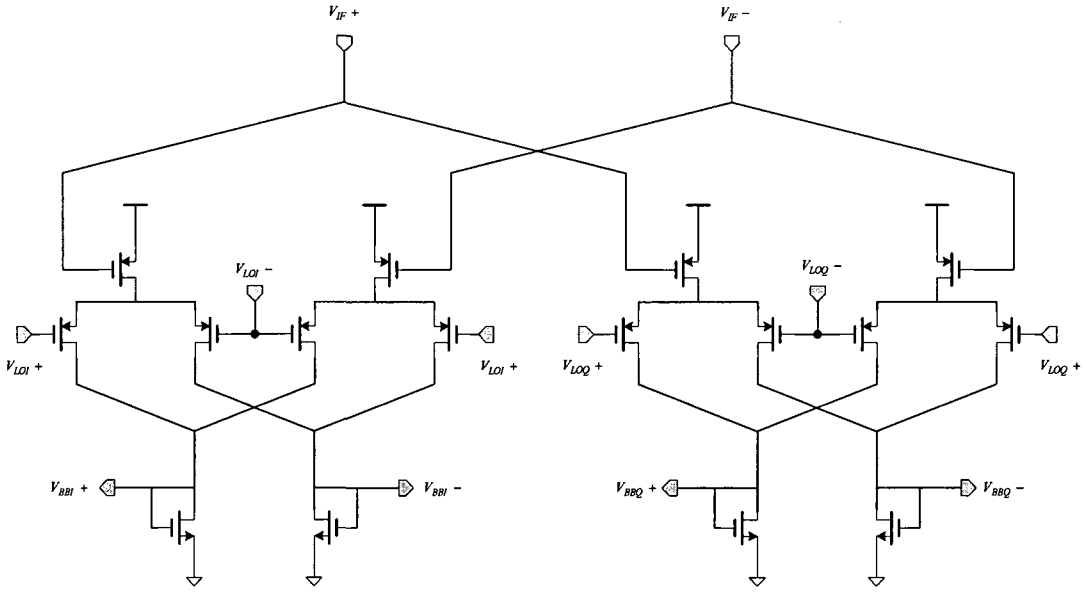


Figure 3.8 IF IQ mixer with gm cell as the load used as the interface to the channel select filter

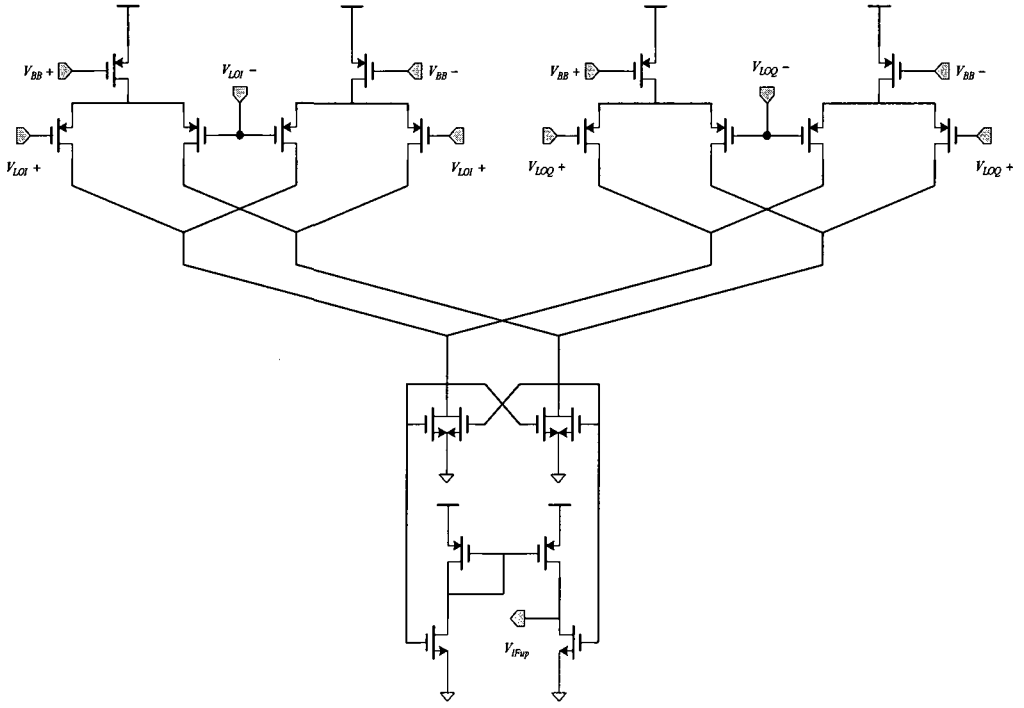


Figure 3.9 up-conversion mixer used in the IF stage of the transmitter; the differential to single-ended converter is merged into the load of the mixer to reused the current.

stage. Since the IQ mixers are driven by square waves generated by the ring oscillator, a conversionsion factor of $2/\pi$ needs to be taken into account in the gain formula.

$$A_{IQmixer} = \frac{g_{mRF}}{g_{mLoad}} = \frac{2}{\pi}\sqrt{\frac{\mu_p}{\mu_n}.\frac{W_{rf}}{W_{load}}.\frac{L_{load}}{L_{rf}}} \tag{3.11}$$

where μ are mobility, and W and L are the width and the length of the respective devices. The combination results in 7dB conversion gain.

A similar structure is employed for the up-conversion mixers in the transmitter, where the described topology is combined with a differential-to-single ended converter to provide a CMOS level signal to the PFD in FTL. The input transistors are fed by a 100mVpp binary data around 900mV common- mode produced by single-bit DACs.

The baseband modulator is basically an Offset Quadrature Phase Shift Keying (OQPSK) demultiplexer, which is coded in Verilog, according to Fig.3.10 [99]. In the transmit mode, the CMOS level IQ data drive input buffers to the upconversion mixer after the amplitude scaling. The level conversion is simply done using transmission gates in Fig.3.11.

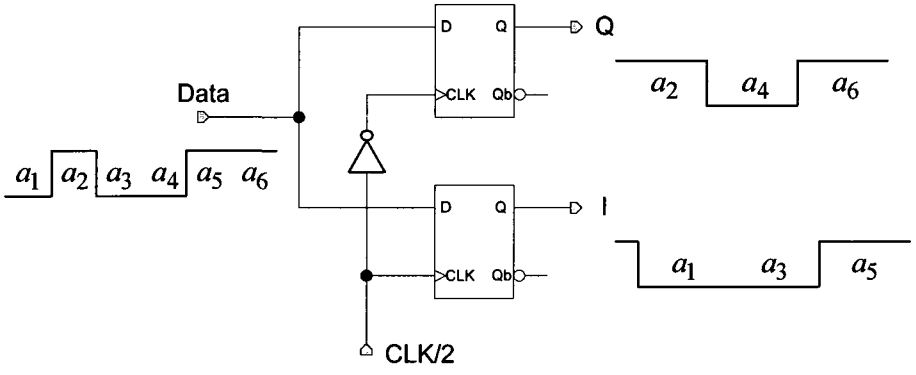


Figure 3.10 OQPSK demultiplexer (serial to parallel) used to modulate the baseband data

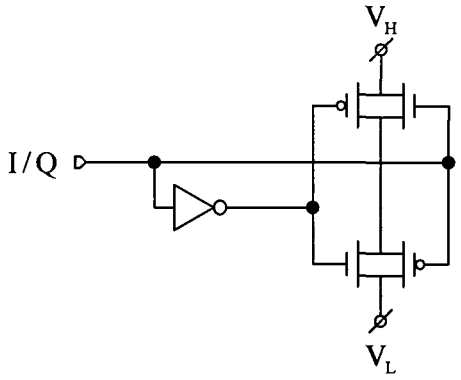


Figure 3.11 simple level conversion used to convert the CMOS level IQ data generated by the OOPSK modulator to levels VH and VL for which the upconversion mixer is optimized for.

3.5.3 Frequency Synthesizer

The body of the frequency synthesis system can be decomposed and discussed in three major building blocks, the master PLL (MPLL), the frequency multiplier, and the frequency translational loop (FTL). In order to select any of the 16 channels located in between 2400MHz to 2480MHz, the master PLL needs to generate IF frequencies, LO2 (Rx mode) and LO3 (Tx mode), between 300MHz to 310MHz with a resolution of 625kHz from a reference frequency of 10MHz. A frequency multiplier generates the 7^{th} harmonic of the IF, 2100MHz to 2170MHz, according to the LO2 (LO3), which is used as the LO1.

3.5.3.1 Master PLL

A Fractional-N Delta-Sigma PLL is used as the master PLL (MPLL) to generate the local oscillator frequencies primarily for the IF mixers. The block diagram of the master PLL is shown in the Fig.3.12.

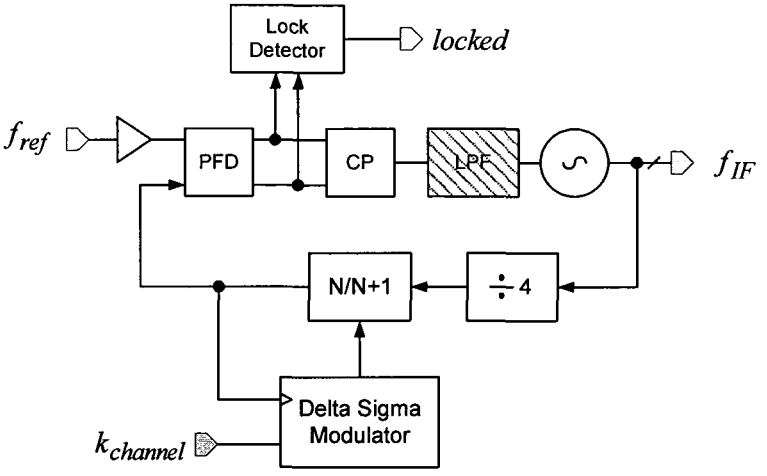
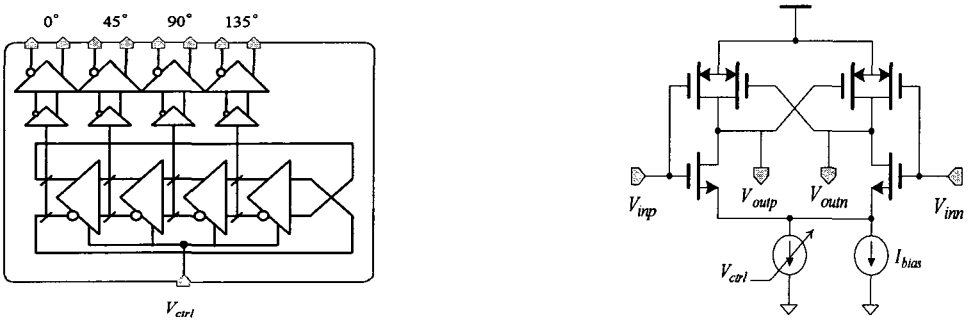


Figure 3.12 Block diagram of the master PLL

The 4-stage ring oscillator is used to produce eight phases of IF frequencies, capable of covering a tuning range of 280MHz to 320MHz. The delay cell used in the ring oscillator is shown in Fig.3.13. Experiments with various topologies show that this structure provides a better PSRR and higher loop gain for a given current while it can operate at supply voltages as low as 1.3V. The phase noise is better than -127dBc/Hz at 5MHz offset. Each unit delay cell drains 450μA from the supply at 305MHz.



(a) (b)

Figure 3.13 Four-stage differential ring oscillator (a) and the schematic of the unit delay cell employed in its design (b)

Two stages of divide-by-2 are followed by a divide-by-7/8 to cover the required output range of 280MHz to 320MHz from a 10-MHz reference frequency. The latches used in the dividers are all based on True-Single-Phase-Clock (TSPC) topology [81] which consumes very low power while it is fast enough for the frequency range of the MPLL. The divide-by-7/8 generates a non 50% output clock which is not an issue for the tri-state PFD sensitive to rising edge only.

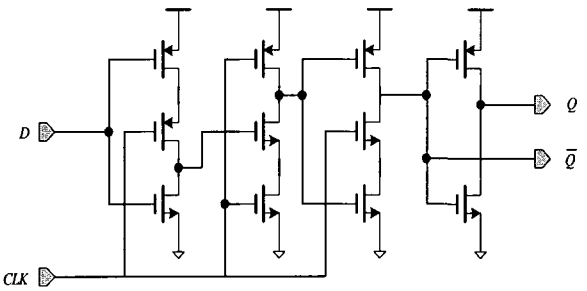


Figure 3.14 Positive edge-triggered register used in TSPC

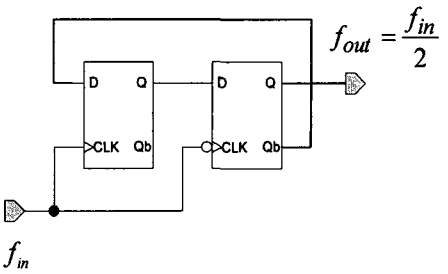


Figure 3.15 Schematic of the divide-by-2

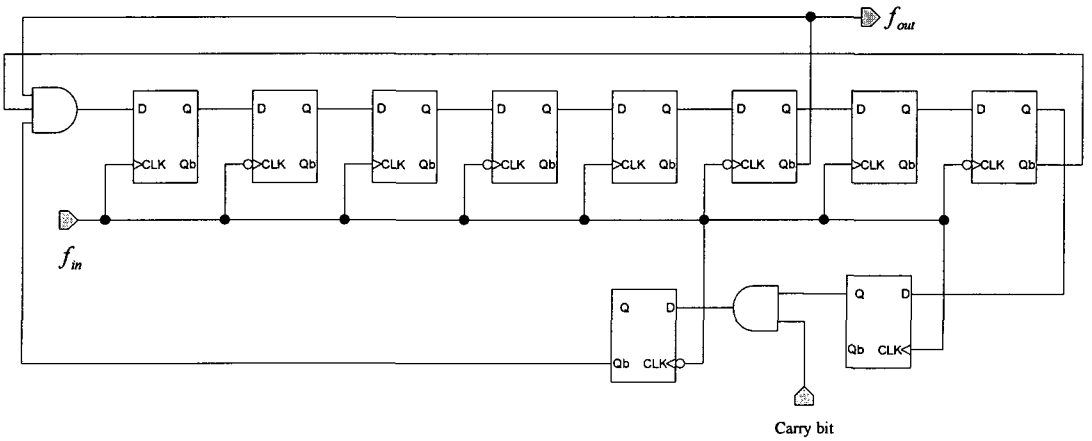


Figure 3.16 divide-by-7/8 is controlled by the carry bit generated by the Sigma-Delta modulator

A single-bit third-order Delta-Sigma Modulator (DSM), based on [71], controls the mode bit of the Divide-by-7/8 according to the 16-bit control word, K-channel. This requires the use of a loop filter of third order or higher to compensate for the 40dB/dec noise shaping caused by the DSM. The third order loop filter is made of off-chip components. A lock detector is implemented in order to turn on the FTL only after the MPLL is settled. The worst case settling time of 20μs is observed for the full scale frequency change; Note that the desired range, 300-310MHz, is only a part the

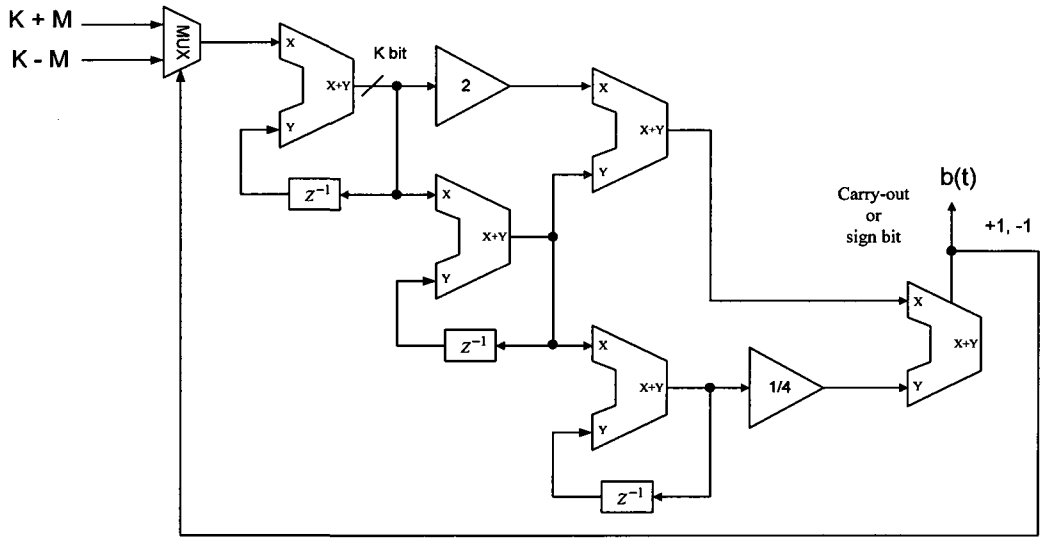


Figure 3.17 Implementation of a single-bit 3rd order Delta-Sigma Modulator

full range. The adder structure, shown in Fig.3.18, is based on conventional one-bit static mirror. Each unit draws an average current of about 10uA of the supply at 10MHz. The limit for speed of the circuit is set by the critical node *Cout* which needs to be optimized for the clock frequency. All other transistors in the Sum node are minimum size.

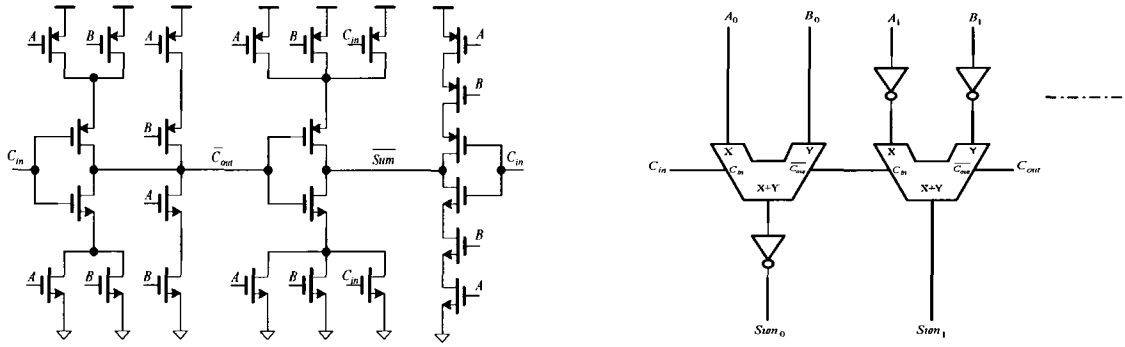


Figure 3.18 Schematics of (a) the one-bit static mirror adder (b) the m-bit adder based on carry look ahead

Figure 3.19 shows the calculated noise spectra referred to different sources; the spectra are shown for the midrange output frequency of 305MHz and based on a reference frequency of 10MHz. The loop bandwidth of 160kHz is found to be an optimum choice in the balancing the noise contribution from the ΣΔ modulator and that of the VCO. The rms jitter is estimated to be 15psec.

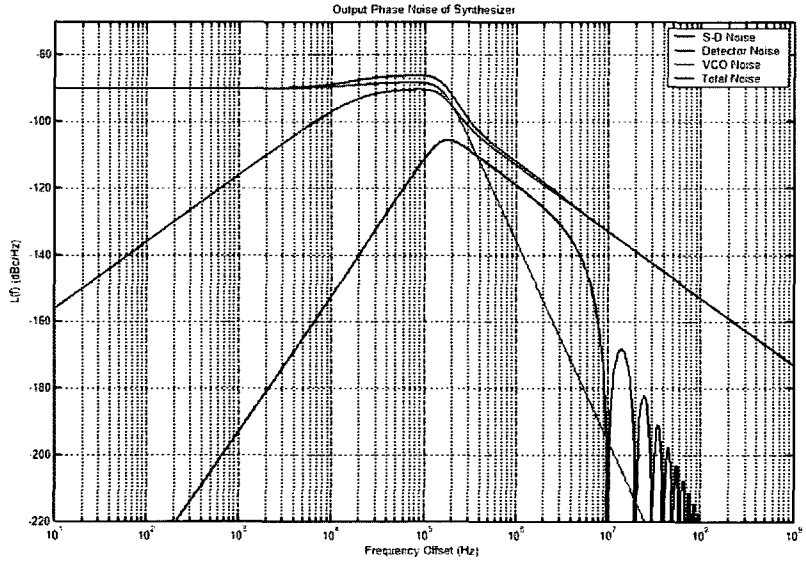


Figure 3.19 calculated noise spectral densities referred to different sources in the MPLL

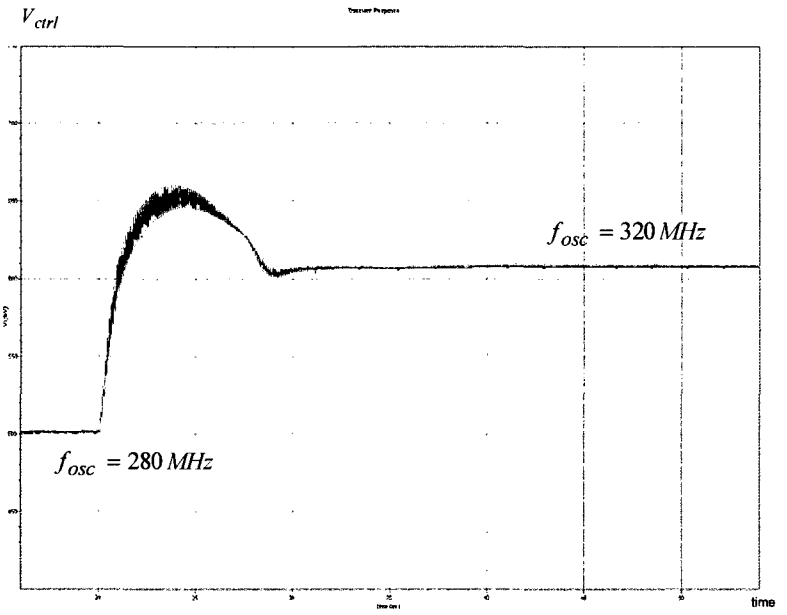


Figure 3.20 Transistor level simulation results for the transient response of the VCO control voltage (MPLL) for a full-range frequency step

3.5.3.2 Frequency Multiplier

The frequency multiplier, shown in Fig.3.21, is designed based on the sub-harmonic injection locking of a system oscillating with free running frequency close to an odd harmonic of the original injected signal. The efficiency of a pure sub-harmonic multiplier system strongly depends on the non-linearity of the slave oscillator and the offset between its free-running frequency and the closest harmonic of injected signal. The farther the injected frequency (harmonic) to the free running frequency of the slave VCO , a larger portion of the injected energy has to overcome the oscillator inclination to oscillate at its inherent natural frequency. The wasted energy is being translated as increase in noise and excess phase noise in the oscillator. The oscillation amplitude is inversely proportional to the degree of non- linearity of the oscillator, with supply voltage as the upper limit. On the other hand, driving the mixer with a larger swing LO is desired since it results in a higher conversion gain and a less flicker noise power. The injected power, bandwidth of the filter and the slave VCO are optimised to minimize the amplitude modulation of the output frequency and variations of the peak output amplitude with the frequency.

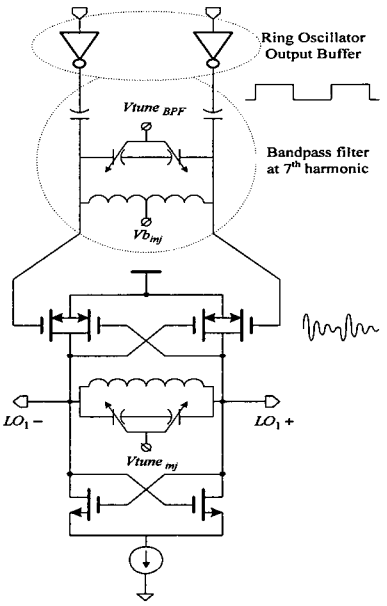


Figure 3.21 frequency multiplier based on injection of the fundamental frequency chosen from the harmonics of the low frequency signal generated by MPLL

A complementary pair of phases of the ring oscillator is applied to CMOS inverters in order to enhance its harmonic contents. The pulse-shaped signal is then ac coupled to an oscillator through an on-chip bandpass filter with its pass-band tuned around the 7^{th} harmonic. Theoretically, it is expected that a realisable inductor with a Q of about 6 can boost the voltage level of the 7^{th} harmonic by the Q factor. In order to minimize the risk of the frequency pulling by the coupling of strong interferers produced by the rest of the transceiver, the locking range of the slave VCO is reduced. To cover the entire band of frequency of about 10Mz, the free-running frequency of the slave oscillator is tuned according to the injected pulse, to be brought closer to this harmonic since excess noise is higher in injection at higher offsets [96, 97]. This makes the phase noise degradation within a minimum 20*log(7) or 17dB.

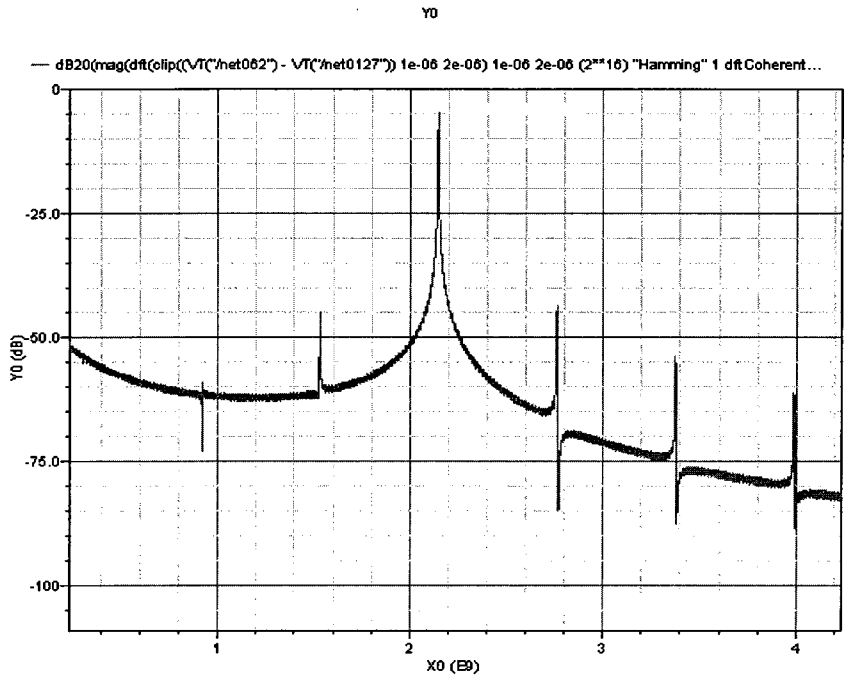


Figure 3.22 Spectrum of the injection-locked VCO at 2.135GHz

The CMOS drivers inject –40dBm into the bandpass filter. This is calculated by integrating the product of the voltage swing at the output of the driver and the current leaving the ac coupling capacitor in the bandpass filter, over one hundred cycles of the LO2. The bandpass filter is made using a 9.4nH symmetric on-chip inductor with a center tap used to provide biasing voltage for the PMOS injector. A set of binary weighted MIM capacitors in parallel with the inductor make the center frequency tuneable. A similar LC tank is used as the resonator for the exploited complementary differential oscillator with a different decoder for the tuning. The average dc current drawn by the oscillator is 800μA.

3.5.3.3 Frequency Translational Loop

The FTL is highlighted in Fig.3.1. The differential IF signal is converted to single-ended before being applied to the PFD. The integrator capacitor in the loop filter is connected to the charge pump through the sampling switch used during the receive mode. The minimum size NMOS switch is accompanied by dummy switches to minimize the sampling clock feed-through and charge re-distribution. The sampling capacitor is large enough to filter out the even-harmonics coupled to the common node of the varactors, but it is mediated by a buffer stage before being connected to the varactor in the LNA tank. This is necessary in order to prevent accumulation of the dc charge produced by the even-order non-linearity of the varactors, which eventually changes the voltage stored on the capacitor and as a result, the LNA gets out of tune.

The bandwidth of the FTL is determined by several factors. Ideally, since the FTL reference frequency is high, a wider loop bandwidth seems possible and therefore desirable in order to transfer a high bit rate up-converted data to the output of the modulator in transmit mode. A wider bandwidth reduces the settling time and as a result the energy consumption of the transceiver.

Several other factors define a limit for such a choice; given the phase noise performance of the MPLL and the channel spacing, the loop bandwidth is set to 2MHz. A settling time of about 7μs is observed.

Due to the large loop bandwidth, the close-in phase noise of the FTL is dominated by that of reference frequency from MPLL and LO1 rather than VCO. The LO1 is correlated to the reference frequency via the frequency multiplier and there are reference spurs present on the spectrum of the injection locked oscillator at about –52dBc.

3.5.4 Baseband Channel-Select Filters

A pair of fourth order gm-C filters is designed for channel selection in the I and Q paths at the baseband. Filters are based on the cascade of two Biquad filters. The single-ended equivalent of each Biquad stage representing an RLC filter has been shown in the Fig.3.23.

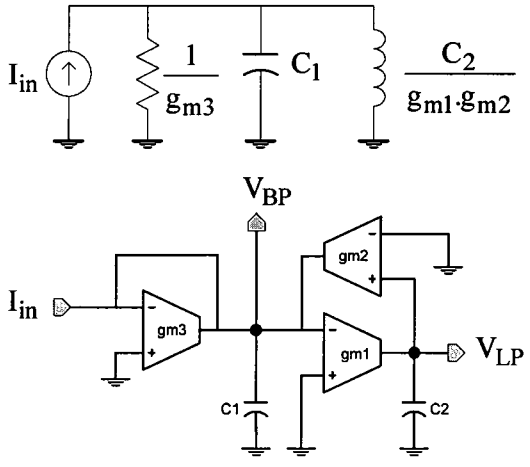


Figure 3.23 half circuit equivalent of each Biquad stage in the filter

The transfer functions of the filter respect to both bandpass and lowpass outputs are simply expressed as

$$\frac{V_{BP}}{I_{in}} = \frac{c_2 S}{c_1 c_2 S^2 + g_{m3} c_2 S + g_{m1} g_{m2}} \tag{3.12}$$

$$\frac{V_{LP}}{I_{in}} = \frac{g_{m1}}{c_1 c_2 S^2 + g_{m3} c_2 S + g_{m1} g_{m2}} \tag{3.13}$$

It is insightful to recognise that at $\omega_o = \sqrt{\frac{g_{m1} \cdot g_{m2}}{c_1 . c_2}}$, the imaginary part of the denominator is zero, and therefore maximum output voltages occur and both transimpedances are equal to $1/g_{m3}$; that is

$$\left|\frac{V_{LP}(j\omega_o)}{I_{in}(j\omega_o)}\right| = \left|\frac{V_{BP}(j\omega_o)}{I_{in}(j\omega_o)}\right| = \frac{1}{g_{m3}} \tag{3.14}$$

Therefore, the maximum internal voltages seen by each OTA is given by I_{in}, which can be controlled easily without scaling; Thus all available OTA designs suffer from a limited linear signal range, which is desirable. Based on eq.3.2 and eq.3.3, the quality factor can be defined by $Q = \omega_o \frac{c_1}{g_{m3}}$.

In Fig.3.23, all OTA's parasitic capacitors are in parallel with the tuning capacitors C1 and C2. Thus, they do not increase the order of the transfer function and can be absorbed by pre-distortion. Besides, the output conductance of the OTA2 and the OTA3 is in parallel with the conductance gm3 and can be absorbed by pre-distortion. The only separate parasitic is the output conductance of the OTA1; this needs to be considered in the design of the intermediate gain stage. The input capacitance of the gain stage combined with output conductance of the OTA1 should not limit the bandwidth of the filter.

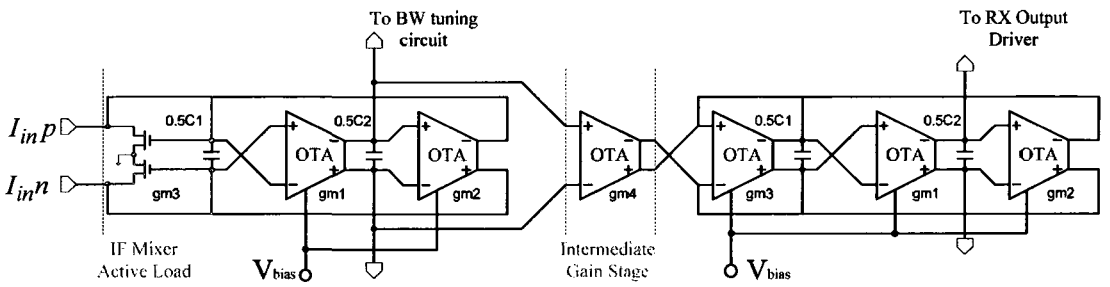


Figure 3.24 Baseband channel select filter designed based on cascading two Biquad filter

The structure of the baseband filter has been shown in Fig.3.24; two differential Biquad filters are chained together using the intermediate transconductance stage. The intermediate stage is used to balance out the magnitude of the gain of the first Biquad to keep the second Biquad well within its linear region. The input transconductance of the first Biquad is actually the diode connected active load of the I (Q) mixer; thus the input to the baseband filter is the downconverted signal in its original current format. The bandwidth of the filter is tuned by changing the bias voltage of the OTAs. When the filter is tuned for 1MHz bandwidth, it attenuates the edge of the adjacent channel at 4MHz by about 48dB (Fig.4.25). Combining this with the selectivity of the partial channel select filter built on the LNA the total adjacent channel attenuation is better than 60dB. Should the LNA bandwidth be set wider, the channel select filter had to have higher order to provide this level of adjacent channel rejection; this would result in a slower filter, i.e. a longer settling time.

Filter nonlinearity sets the upper bound on the dynamic range. The filter uses standard two-stage OTA to drive the capacitive load. Each OTA drains 64μA and gives 82dB dc gain with a unity gain frequency of 11MHz and 60 degree phase margin. Large gate area is used for the input NMOS transistors to lower the flicker noise corner. The lower f_t of the transistors is not an issue because the maximum input frequency of the filter is 10MHz. The combined gain of the IF mixer and channel-select filter is 38dB. The baseband filters consume a total current of about 750μA.

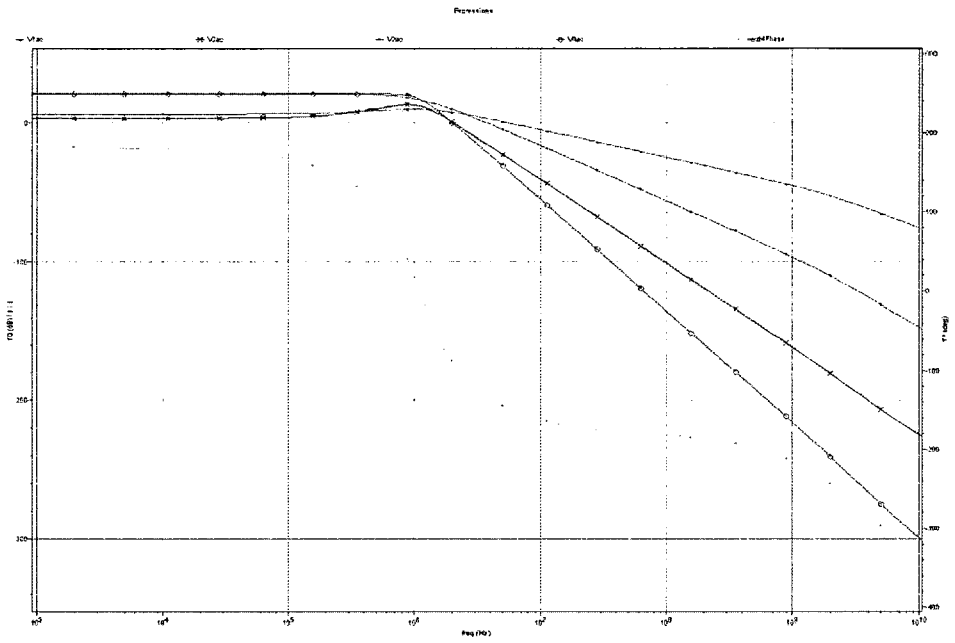


Figure 3.25 The incremental frequency response of the channel-select lowpass filter; the red and green are the magnitude responses of the first and second Biquads; the yellow is the phase response of the entire filter.

The capacitors C1 and C2 are fixed and designed using the MIM capacitance. It is necessary to maintain the symmetry throughout the entire filter in order to maintain the differential signal and to minimize harmonics produced by the even order non-linearity. One of the major contributors to the asymmetry is floating capacitor between differential pairs. Consider the capacitor in Fig.3.26a, where one end of the capacitor is grounded. The bottom plate parasitic capacitance with the substrate is grounded, or if the capacitor is flipped up-side down, the parasitic can be taken into account by pre-distortion. However, as shown in Fig3.26b using this capacitor between the differential pair results in the residue capacitance on one side of the pair. One solution is to replace the capacitor, C, with two larger capacitors, 2C. Figure 3.26c shows that the bottom plate parasitics

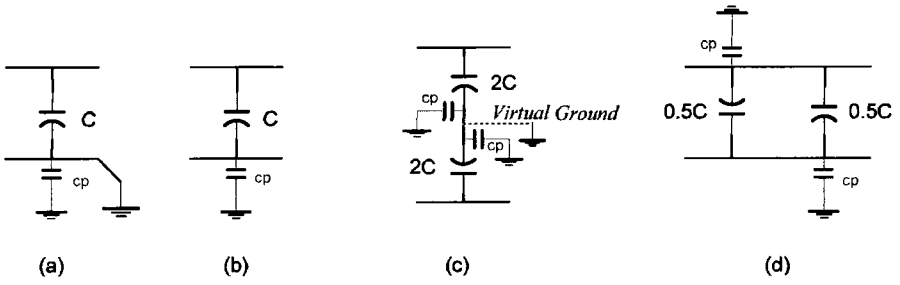


Figure 3.26 bottom plate parasitic capacitance makes a floating capacitor assymetric

are now either virtually grounded or cab be absorbed by pre-distortion as a part of the resonator capacitance; but the required area is quadrupled. The alternative approach, which is used in this design, is what is shown in the Fig.3.26d where two smaller capacitors, 0.5C, are used in parallel such that their opposite polarities are connected together. Therefore, a proper common-centroid layout, guaranties bottom plate parasitics are loading the differential pair equally and can be taken as a part of the time constant of the filter; while they occupy almost the same area as the single capacitor.

3.6 Conclusion

Major contributions of the thesis have been reviewed throughout this chapter and can be summarized as follows.

- A new concept in transceiver architectures is introduced in which the receiver is used for the transmit purpose as well, or vice versa. The dual-personality structure can be accommodated in a smaller area which reduces the die cost.
- A true energy-efficient transceiver for a time-multiplexed system is presented in which the transmitter back-end is formed by re-configuring the front-end of the receiver such that it allows modulating high bit-rate data or short bursts, and managing the power dissipation.
- A Dual-loop injection-lock-coupled fractional-N frequency synthesizer is proposed not only as the core of the constant envelope modulator in the transmit mode or frequency tuning mechanism in the receive mode but also as a technique to cope with losses in clock distribution network. This suggests distributing a low frequency clock across the chip and generating the high frequency clock locally at the destination or the end user.
- A modified high-IF dual conversion receiver is proposed. Two correlated frequencies are used as local oscillators, while the center frequency of the LNA can follow them accordingly.
- New circuit topologies are presented for LNA and IQ mixers in the receiver and transmitter chain.

As mentioned earlier in this chapter, when the bandwidth of the LNA in not wide enough to cover the entire desired band, there is a need for a tuning mechanism in order to move the passband of the LNA to the frequency range transparent to the targeted channel. The proposed tuning mechanism is explained in chapter5.

4 On Characteristics of Ultra-Narrowband LNAs

General design considerations for the Low-Noise Amplifier (LNA) were briefly reviewed in the chapter 3, as one of the building blocks in the transceiver. In this chapter, major characteristics of the block are discussed in details. While there may not be a single electronic component that is ever required to "do it all", the circuit that comes the closest might be the LNA in portable RF receivers. It is required to operate at high frequencies while still providing a significant level of isolation, to deliver a low noise signal with minimal power consumption, to provide gain and a reasonable degree of linearity, and maintain its stability across a band wider than its passband [92]. A narrower filter as the first stage receiver reduces the complexity and the cost of following stages, by reducing the dynamic range of the signals they are to process[16]. The ultra-narrow-band terminology is used to differentiate the case from LNAs regularly used in radios which are categorized as narrow-band compare to wide-band or ultra-wide-band ones. An ultra-narrow-band is supposed to have a bandwidth narrower than the target frequency band, or provides partial band

[16] Note that this requires a more complex filter.

coverage, and therefore needs a frequency tuning mechanism to cover the entire band. Using active devices to narrow the bandwidth of the LNA introduces additional challenges which need to be considered. Characteristics are often benchmarked against the standard Common-Source or common-gate topologies for comparison.

4.1 Input-Output Isolation

The ultra narrowband nature of the load in the LNA makes the matching properties at the input of a single transistor amplifier varying with center frequency of the tuneable load. The undesirable situation can get worse if the quality factor of the input matching network is high. In a common-source (CS) topology this coupling is due to gate-drain capacitive feedback. Considering the effect of the Cgd, the input impedance of an inductively source degenerated CS LNA can be expressed as

$$Z_{in}^{cgd}(j\omega) = \frac{Z_{in}(j\omega)}{1 + \frac{\omega_T}{\omega} \cdot \frac{C_{gd}.\omega.Z_{load}(j\omega)}{jC_{gd}.\omega.Z_{in}(j\omega)+1}} \tag{4.1}$$

where $Z_{in}^{cgd}(j\omega)$ is the input impedance taking into account the Cgd, and $Z_{in}(j\omega) = 1/jC_{gs}.\omega + jL.\omega + L.\omega_T$ is the familiar input impedance of the stage without considering Cgd. This impedance needs to satisfy the matching criteria required by a 50 Ω source impedance. For simplicity, the unity current gain bandwidth is approximated by $\omega_T \cong g_m/C_{gs}$, neglecting Cgd[17]. It can be seen that a larger transistor and/or a larger load impedance translates into a larger deviation of the input impedance with the load center frequency. This requires a tuneable matching network at the input to follow the nonlinear coupling. An immediate solution would be to minimize the miller effect by using a cascode amplifier and at the same time using smaller

[17] The more accurate expression for the unity gain bandwidth is $\omega_T = g_m/(C_{gs} + C_{gd})$; and in submicron technologies Cgd values are comparable to that of Cgs; it is about half of Cgs in the 65nm technology. Although, here, we are investigating the impact of Cgd, logically we are not supposed to drop it out the equation. We used the no longer valid approximation only to simplify the expression into a form decomposable into more familiar one at limits. The designer can simply neglect this simplification and derive the complete form following a standard circuit analysis method.

transistors to minimize Cgd. The former necessitates scaling of the biasing current proportionally to maintain the ω_T at its peak in order to avoid the noise figure degradation.

Compared to the CS topology in a standard single transistor common-gate amplifier, this is not the gate-drain capacitance which creates a feedback between the input (source) and the output (drain), but the transconductance of the input transistor provides a direct channel between the two. In general the impedance seen at the source of a transistor, loaded by the impedance of the resonator $Z_{load}(\omega)$, can be expressed by

$$Z_s(j\omega) = \frac{r_o + Z_{load}(j\omega)}{1 + g_m r_o} \tag{4.2}$$

When the impedance of the load is much smaller than the output impedance of the transistor, $|Z_{load}(j\omega_o)| << r_o$, the above expression is simplified to the familiar $\frac{1}{g_m}$. In an ultra-narrowband resonator this criterion might not be satisfied since the equivalent parallel resistance of the tank is in the same order as r_o. Thus, once again the input impedance varies depending on the load impedance for every channel setting. Adopting the cascode topology helps to reduce this dependency. Fig.4.1 shows an single-ended LNA using the cascode CG topology exploiting the negative transconductance of $-G_q$ to compensate for the loss of the tank, G_p. It can be shown that the impedance seen looking into the source of M1, i.e. Z_{s1}, can be written as

$$Z_{s1}(j\omega) = \frac{r_{o1} + Z_{s2}(j\omega)}{1 + g_{m1} r_{o1}} \tag{4.3}$$

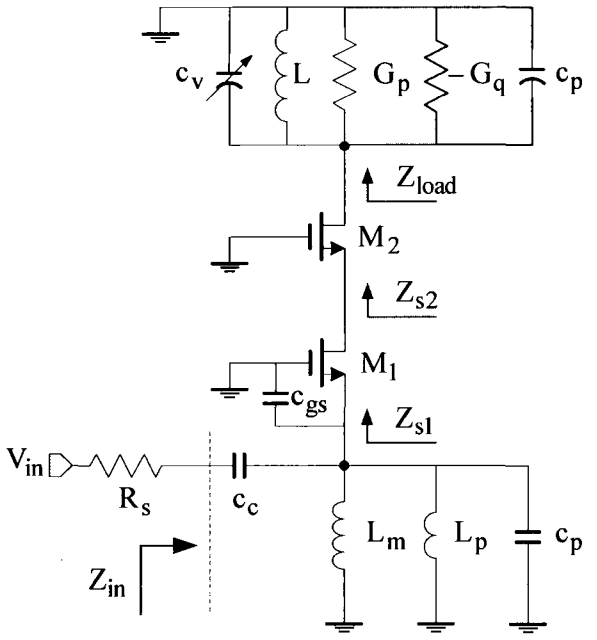


Figure 4.1 Cascode Common-Gate with Q-enhanced load

where $Z_{s2}(j\omega) = \dfrac{r_{o2} + Z_{load}(j\omega)}{1 + g_{m2} r_{o2}}$.

The assumption $g_m r_o >> 1$ leads to the further simplification of the above expression to

$$Z_{s1}(j\omega) \approx \frac{1}{g_{m1}} + \frac{Z_{load}(j\omega)}{(g_m r_o)^2} \tag{4.4}$$

Here it is assumed that C_{gs1} is absorbed into the input matching network. And the effect of C_{gs2} is ignored without reducing the argument. It is clear that the load and the input impedance inter-dependency is now reduced by the higher order of $(g_m r_o)^2$ compare to that of non-cascode case, $g_m r_o$.

We can therefore conclude that a cascode CG with relatively high intrinsic voltage gain is necessary in order to improve input-output isolation, especially when load impedance is high

(ultra-narrowband). Since maintaining a high ω_T requires the use transistor with minimum length, a higher transcoductance then is preferred to meet the above expectations.

A higher tranconductance means higher current consumption which is not desirable. In fact, in order to reduce the power consumption, it is preferred to design for even lower g_m and using an impedance transformer matching the LNA to the source impedance or 50 Ω. We will come back to this once again when discussing the input matching network issues.

Before leaving this section, let's have another look at the expression derived for Z_{sl}. Substituting the load impedance at its resonance frequency with the equivalent resistance $Q_L L\omega_o$ we will get

$$Z_{sl} = \frac{1}{g_m} + \frac{1}{\Delta\omega_{3dB}.(g_m r_o)^2}.\frac{1}{C} \tag{4.5}$$

where $\Delta\omega_{3dB} = \frac{\omega_o}{Q_L}$, $Q_L = \frac{G_p}{G_p - G_q}Q_o$ and Q_o is the inherent quality factor of the load's inductance. This means for a given bandwidth, lower inductance will enhance the isolation; and this is clearly the practical option available to improve the isolation while maintaining other performance specifications, such as gain and noise figure, intact.

4.2 Input Matching

The effective transconductance of the CS LNA with inductive degeneration is equal to

$$G_{m,eff} = g_m.Q_{cs} = \frac{g_m}{2R_s C_{gs}\omega_o} \tag{4.6}$$

$$Q_{cs} = \frac{1}{2R_s C_{gs}\omega_o} \tag{4.7}$$

As the bias current goes down, the device transconductance gm decreases while according to eq.4.6, the effective transconductance remains constant if the width of the transistor is also reduced respectively. Keeping the current density constant will guaranty that ω_T is independent of the absolute bias current [91]. Therefore, proper sizing of the device makes it possible for it to operate at it peak unity gain frequency. This is desired since the noise figure of the amplifier is reciprocally proportional to ω_T as it is shown by eq.4.8

$$F_{cs} = 1 + \gamma . g_m R_s \left(\frac{\omega_o}{\omega_T} \right)^2 \tag{4.8}$$

where γ is the noise factor of the channel thermal noise of the transistor [95]. Therefore, any effort to reduce the power consumption of the CS LNA, while maintaining its NF, requires reducing the width of the transistor respectively with its current. This results in a smaller C_{gs}, which in turn, according to eq.4.7, increases the Q_{cs} of the input matching network. Typically the quality factor of the input matching network of the CS LNA is about 8 to 10 at moderate current consumption; a further increase in its value would bring it the neighbourhood of the desired bandwidth for the target application.

Therefore, although reducing the size of the transistor works favourably towards achieving a higher isolation and lower noise figure, this comes at the price of the higher quality factor of the input matching network, making the input return loss sensitive to component mismatches or variations in the input matching network.

A common-gate (CG) topology with impedance transformation would be a potential candidate to reduce the above trade-off to a more manageable level. The input matching network for the

cascode CG is shown in the Fig.4.1. Assume that a portion of the total input matching inductance, L_p, is dedicated to resonate out the parasitic capacitance of M1. This dictates

$$L_s = \frac{L_m L_p}{L_m + L_p} \tag{4.9}$$

$$L_p(C_p + C_{gs}) = \frac{1}{\omega_o^{\,2}} \tag{4.10}$$

Therefore the input impedance of the LNA can be written as

$$Z_{in}(j\omega) = \frac{1}{jC_c\omega} + \left(\frac{1}{g_m} \| jL_m\omega\right) \tag{4.11}$$

The conjugate impedance matching requires $Z_{in}^{\,*} = R_S$, which results in the following resonance frequency and values for the remaining elements of the network,

$$\omega_o = \frac{1}{\sqrt{L_m C_c - (g_m L_m)^2}} \tag{4.12}$$

$$C_c = \frac{1}{R_s\omega_o}\sqrt{\frac{g_m R_s}{1 - g_m R_s}} \tag{4.13}$$

$$L_m = \frac{1}{g_m\omega_o}\sqrt{\frac{g_m R_s}{1 - g_m R_s}} \tag{4.14}$$

where the identity $R_s = g_m.L_m/C_c$ needs to be held to satisfy the conjugate matching requirements. On the other hand, the voltage gain from the input of the matching network (Cc) to the source of the transistor has a quality factor which is expressed by

$$Q_{match} = \frac{1}{g_m}\sqrt{\frac{C_c}{L_m}} = \sqrt{\frac{1}{g_m R_s}} \tag{4.15}$$

It can be seen that for $R_s = 50\Omega$, where the required transconductance is equal to 20mS, the quality factor is as low as unity which is a matching network as wide as the center frequency of the LNA. In this case, there is no need for the impedance transformation and the matching could be done directly by proper sizing of the transistor. To achieve such gm, an overdrive voltage of 200mV requires 2mA current (single-ended). This is at least four times the budget allocated for the LNA stage. In order to reduce the power consumption a lower transconductance is preferred. A target transconductance of 5mS requires only 0.5mA for a differential amplifier with an equivalent first order linearity range. Eq.4.16 shows that, in this case, the quality factor is increased only by the square-root of the inverse of the transconductor which is only two times. Thus the matching network is still wide enough to stay robust against mismatches, parasitics and variation of off-chip matching elements and their interface with the die. This saving in power and robustness of the matching network comes at the cost of a higher noise figure which is discussed in detail later in this chapter.

4.3 Voltage Gain

At perfect input impedance match, the voltage gain from a voltage source at the input to the source of the device M1 is

$$A_{v1}(\omega_o) = \left|\frac{V_1}{V_{in}}\right| = \frac{1}{2}\sqrt{\frac{1}{g_m R_s}} \tag{4.16}$$

and therefore the total gain the LNA is expressed as

$$A_{vLNA}(\omega_o) = \frac{R_{load}}{2}\sqrt{\frac{g_m}{R_s}} \tag{4.17}$$

where the impedance of the load at the center frequency of the tank is given by $R_{load} = (G_p - G_q)^{-1}$, and g_m is the transconductance of the M1. Equation 4.17 is valid up to frequencies where the loading effect of the gate-source capacitance of the cascode device, M2, and the gate-drain capacitance of M1 can be neglected. Considering these factors, the current division ratio manifests itself as a reduced current at the load which depends on the ratio of unity current-gain frequency to the operating frequency of the circuit. It is also assumed that the output impedance of the cascode transistors is much higher than that of the load[18]. Note that the effect of the Cgs1 is resonated out by the partial inductance, L_p, in the matching network. The gain expression is used in calculating the noise figure later on in this section. One may prefer to reframe the gain as

$$A_{vLNA} = \frac{1}{2} \cdot Q_{match} \cdot g_m R_{load} \tag{4.18}$$

where $Q_{match} = \frac{1}{g_m}\sqrt{\frac{C_c}{L_m}} = \sqrt{\frac{1}{g_m R_s}}$. Equation 4.19 decomposes the original expression into the gain of cascaded blocks. As discussed earlier, a low Q of around 2 to 3 is desired. This translates to a g_m in the neighbourhood of 2mS to 5mS. It is obvious that as the load bandwidth gets narrower, the voltage gain increases for a given input transconductance. One may try to balance this gain out by reducing g_m of the M1, which jeopardizes the linearity of the input transconductance stage before delivering the signal to the load.

Eq.4.18 can also be written as

---

[18] The output impedance can be written as $r_{out} = r_{o2}\left[1 + g_{m2}\left(r_{o1}(1 + g_{m1}Z_s) + r_{o1}\right)\right] + r_{o2}$ where Z_s is the total impedance seen at the source of M1 looking into the matching network (and not the source of M1). If transistors are sized equally and assuming $r_{o1} = r_{o2}$, at the perfect match, i.e. $g_{m1}Z_s = 1$, the output resistance of the cascode stage can be simplified as $r_{out} = 3g_m r_o^2 + 2r_o \cong 3g_m r_o^2$.

$$A_v = \frac{1}{2}.Q_{match}.g_{m1}.\frac{1}{G_p}.\frac{Q_L}{Q_o} \tag{4.19}$$

Considering $G_pQ_o = C\omega_o = \frac{1}{L\omega_o}$ and $Q_L = \frac{\omega_o}{\Delta\omega_{3dB}}$ then

$$A_{vLNA} = \frac{1}{2}.Q_{match}.g_{m1}.\frac{1}{C.\Delta\omega_{3dB}} \tag{4.20}$$

This reveals that the gain of the stage is defined by the quality factor of its input matching network and the bandwidth of the load at the given center frequency. The voltage gain increases at the rate of 20dB/dec of the decrease in the bandwidth. The only design parameter is the tank capacitance (inductance) which is defined according to the desired gain. For example, a $Q_{match} = 2$, $\Delta\omega_{3dB} = 2\pi.(20MHz)$ and C=1pF, results in 32dB voltage gain (single-ended).

4.4 Noise

The noise figure of an ultra-narrowband amplifier is another distinguishable performance metric which need to be considered; focus of works carried out by most researchers has been on the general characteristics of the noise figure in tuned amplifiers and the common-source topology has been the case of their studies [80, 87]. Here, the goal is to use the standard approach to highlight the major differences in the noise figure behaviour of the ultra-narrowband common-gate stage with input impedance transformation.

Before presenting the noise figure analysis, a general circuit analysis lemma is introduced. This facilitates the estimate of the contribution of the channel thermal noise of a transistor at any possible configuration or terminations at its source and its drain. The technique is frequently referred to throughout the noise analysis.

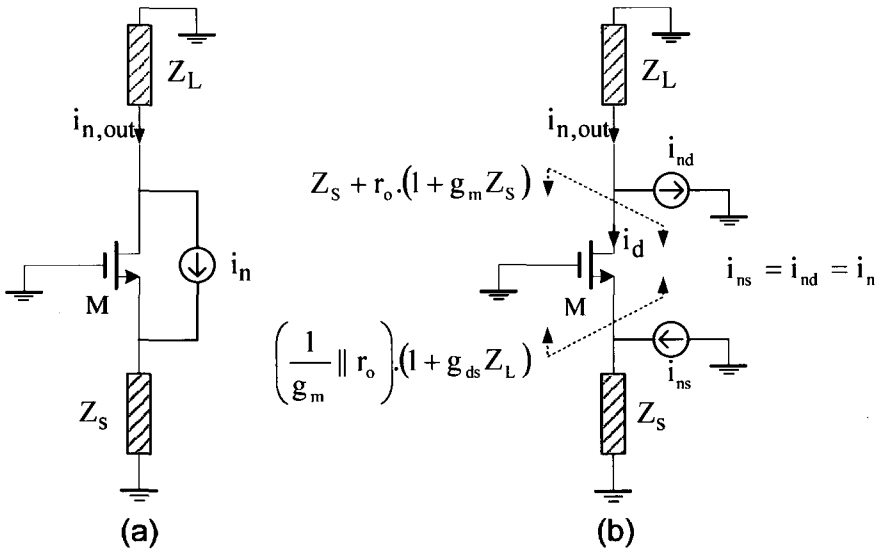


Figure 4.2 A general technique to calculate the contribution of channel thermal noise of a transistor to the output considering the effects of source and drain termination impedances

Lemma: Fig.4.2a shows the transistor M with its source degenerated with the impedance Z_s and its drain driving the load Z_L. The output response, $i_{n,out}$, to the current excitation, i_n, can be calculated by the superposition of the response to each of the two current sources, i_{ns} and i_{nd}, each carrying the same amount of current as i_n, as shown in Fig.4.2b.

A straightforward current division shows the contribution of i_{ns} at the source of the transistor to the drain current, $i_{n,out}^{ns}$ is

$$i_{n,out}^{ns} = \frac{-Z_s}{Z_s + (1 + g_{ds}.Z_L)/(g_m + g_{ds})} i_{ns} \tag{4.21}$$

On the other hand, the portion of i_{nd} flowing to the load Z_L is

$$i_{n,out}^{nd} = \frac{Z_S + r_o.(1 + g_m.Z_S)}{Z_L + Z_S + r_o.(1 + g_m.Z_S)} i_{nd} \tag{4.22}$$

Considering the fact that $i_{ns} = i_{nd} = i_n$, and applying superposition principle results in

$$i_{n,out} = i_{n,out}^{ns} + i_{n,out}^{nd}$$

or

$$i_{n,out} = i_n \left(\frac{-Z_s}{Z_s + (1 + g_{ds}.Z_L)/(g_m + g_{ds})} + \frac{Z_S + r_o.(1 + g_m.Z_S)}{Z_L + Z_S + r_o.(1 + g_m.Z_S)} \right) \quad (4.23)$$

The drain-source transconductance $g_{ds} = 1/r_o$ is used to highlight the symmetry in the impedance transformation factors when looking into the source of the transistor, $(1 + g_{ds}Z_L)$ compared to that seen looking into the drain, i.e. $(1 + g_m Z_s)$.

The parasitic effects of Cgs and Cgd are assumed to be included in Z_S or Z_L where applicable. Also, body effects is neglected for simplicity.∴

The above rather monstrous result might discourage pursuit of any manual analysis in calculating the noise figure. Two special cases have been found to be accurate enough, easy to interpret and handle in noise analyses. The first case is when the load impedance is much lower that the impedance seen looking into the drain of M; and when the source impedance Z_S is large. At the extreme, i_{nd} drains its current solely from the load, reducing the second part of the above expression to unity. Therefore

$$i_{n,out} = \frac{i_n}{Z_s (g_m + g_{ds})/(1 + g_{ds}.Z_L) + 1} \quad (4.24)$$

This is the expression which will be used to extract the noise contribution of both transistors in the cascode configuration (even with a high impedance load).

When dealing with wide-band amplifiers, one can further simplify the original expression by neglecting the term $g_{ds}.Z_L$, and reduce the noise expression to

$$i_{n,out} = \frac{i_n}{Z_s.(g_m + g_{ds}) + 1} \tag{4.25}$$

Figure 4.3 shows the major noise contributors in the CG LNA with input impedance transformation[19].

The lemma can be used to calculate the noise voltage at the output generated by M1. Since C_{gs1} is cancelled out by the extra inductance L_p at the source as a part of matching network, one can drop the respective term from the equation. It is also sensible to assume that the impedance seen looking into the source of M2 is considerably lower than that of C_{gs2}, and therefore $i_{n1,out}$ is perfectly passed to the output[20]. This results is the following expression:

$$\overline{v_{n1,out}^2} = \frac{4kT\gamma.g_{m1}}{[Z_s\,(g_{m1} + g_{ds1})/(1 + g_{ds1}Z_{s2}) + 1]^2} \cdot \frac{1}{(G_p - G_q)^2} \tag{4.26}$$

Assuming $g_m >> g_{ds}$[21], and knowing that at the perfect match $g_{m1}Z_s = 1$, and considering eq.4.3, the effect of the frequency dependant load impedance is further reduced by the second power of r_o, and eq.4.26 can be simplified to

$$\overline{v_{n1,out}^2} = \frac{kT\gamma.g_{m1}}{(G_p - G_q)^2} \tag{4.27}$$

---

[19] For simplicity, only the channel thermal noises of the active devices are considered; correlated excess noises such as gate-induced noise are neglected.

[20] This is reasonable for the target frequencies (2.4GHz) which is much lower than the cut-off frequency of the device. The reactance of Cgs at this frequency is about 5KΩ which is much higher than 1/gm2=200Ω.

[21] In submicron technologies, as the channel becomes shorter, making such assumption is certainly questionable; here this simplification is only made to simplify the equation without distorting the intended analysis. For those inclined, the original form can be used towards the final extraction.

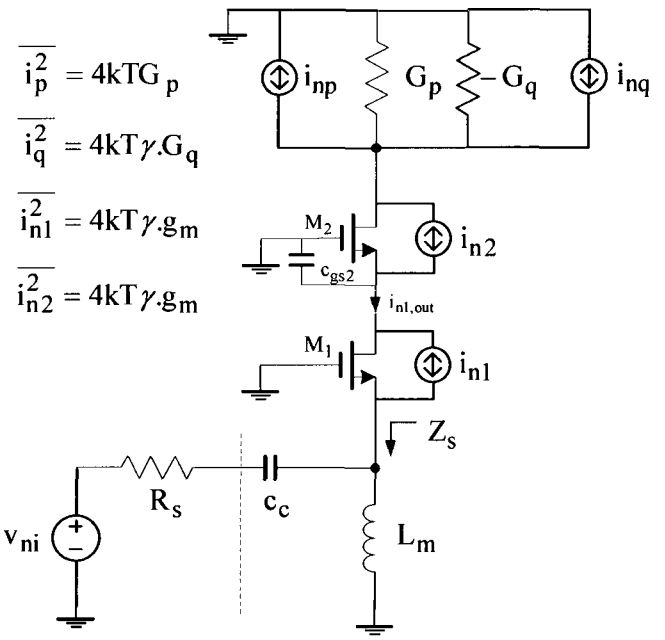


Figure 4.3 a single-ended equivalent of the Common-Gate LNA with input impedance transformation, used to calculate the noise figure

Next we focus on the role of channel noise of M2 in the noise figure. Following the lemma in its original form, one needs to substitute the load impedance with $(G_p - G_q)^{-1} = (C.\Delta\omega_{3dB})^{-1}$ and the source impedance with the impedance seen looking into the drain of M1 which is approximately $2r_o$[22] , in order to get the noise contribution of M2 to the output current[23]. One could picture Fig.4.2b and simply observe that since the impedance seen looking into the source of M2 is much smaller than that seen looking into the drain of M1, the noise current i_{ns} is almost completely drawn into the source of

[22]Equivalent impedance hanging on the source of M2 can be approximated as $r_{outM1} = r_o(1 + g_m Z_s) + Z_s$ which at the perfect input match is reduced to $r_{outM1} \approx 2r_o$.

[23] It is assumed that the reactance of Cgs and Cgd of both transistors are very small at the target frequency and can be neglected relative to their rivaling impedances in parallel.

M2 and passed to the output. On the other hand, i_{nd} finds the load impedance much smaller that the impedance seen looking into the drain of the M2, and flows towards it. These two equal noise currents are out of phase, resulting in a net noise current of zero. Figure 4.4 shows that within the range of the validity of the above approximation, one can look at the channel noise source and the cascade device transconductance as a degenerated loop, without any current projected at the output. As an example, for a load tuned for the bandwidth of about 2MHz, using 1pF capacitance in the tank, the load impedance is approximately 80KΩ compared to the cascode output resistance of 6MΩ for a transconductance of 5mS and a bias current of 500μA[24]. On the source side, the ratio of 1/5mS=200Ω to the impedance of 40KΩ seen at the drain of the M1 is very small. Therefore, although in the ultra-narrow band LNA the load impedance is high, it is still a valid approximation to neglect the noise contribution of the channel thermal noise of the cascode transistor.

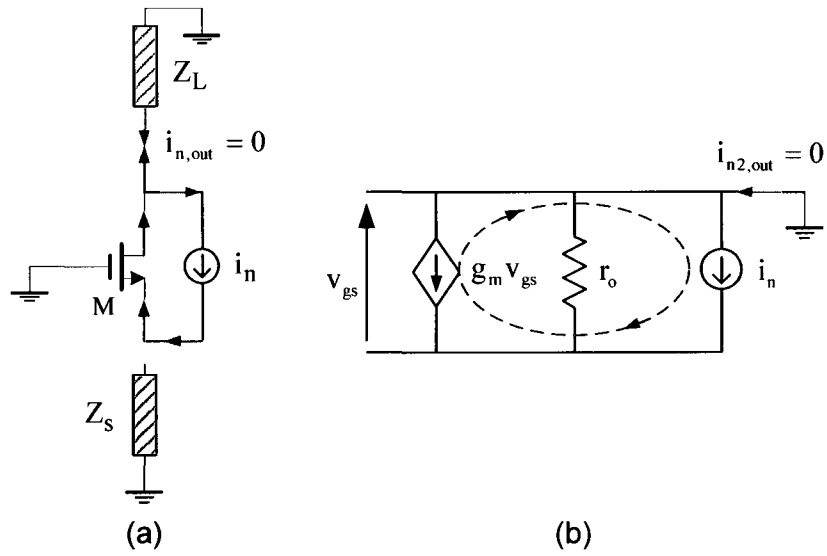


Figure 4.4 channel thermal noise contribution of the cascode transistor acts as the degenerated loop

[24] The channel length modulation factor, λ, is assumed to be 0.1 v^{-1}.

And finally, the noise power generated by the load can simply be expressed by

$$\overline{v_{n,load}^2} = 4kT\frac{G_p + \gamma.G_q}{(G_p - G_q)^2} \tag{4.28}$$

Assuming equally sized transistors and reframing the gain expression of the amplifier as

$$A_{vLNA} = \frac{1}{2}\sqrt{\frac{g_m}{R_s}}.(G_p - G_q)^{-1} \tag{4.29}$$

all the ingredients are ready to calculate the noise figure based on

$$F = \frac{N_{out}}{A_{vLNA}^2.(4kTRs)} \tag{4.30}$$

where N_{out} is the total voltage noise power at the output of the amplifier. Therefore the noise figure of the CG LNA with impedance transformation is

$$F = 1 + \gamma + \frac{4}{g_m}\left(G_p + \gamma.G_q\right) \tag{4.31}$$

where each term represents the partial contribution of its respective element, that is source resistor, input transistor M1, and enhanced load[25]. The immediate conclusion is that the common-gate topology, regardless of its gain or bandwidth, has a minimum noise figure of as high as 1+γ which for submicron technologies can easily exceed 4dB considering the fact that the device noise factor γ is much higher than 2/3 and values higher than 2 have been reported [90]. Therefore the CG topology is not recommended for applications or standards demanding high sensitivity such as GPS or GSM.

[25] Noise contribution of the gate resistance is neglected; in practice, in order to minimize this resistance, the gate is broken into parallel fingers; and fingers are connected from both side to further reduce the distributed resistance by four times.

If the negative transconductor is turned off, $G_q = 0$, the LNA is operating with its intrinsic load G_p and the noise figure is given by

$$F\big|_{Gq=0} = 1 + \gamma + \frac{4}{g_m}.G_p \tag{4.32}$$

As expected, increasing the intrinsic gain of the LNA, g_m / G_p, reduces the noise figure.

On the other extreme, when G_q is very close to the inherent conductance of the tuned load, $G_q \approx G_p$, the noise figure is increased to

$$F\big|_{Gq \to Gp} \cong 1 + \gamma + \frac{4}{g_m}(1+\gamma).G_p \tag{4.33}$$

This sets the maximum noise figure of the LNA for the given bandwidth. Also, it can be seen that the negative transconductor, deteriorates the noise figure by $1+\gamma$. Equation 4.33 shows that the higher the intrinsic gain of the amplifier the smaller the negative impact of Q-enhancement on the noise figure. This suggest using high quality elements constructing the tuned load (inductor and varactor), or lower G_p, in order to minimize the noise figure[26].

Using an on-chip inductor with an inductance of 2nH and quality factor of 6, results in the tank effective conductance of G_p=5.5mS at 2.4GHz; and if the LNA devices are sized to provide a transconductance of g_m =5mS, the noise figure of the LNA without any Q-enhancement at its load is 8.7dB of which approximately 6.4dB is the contribution of the loss in the tank which is about

[26] The noise figure expression seems to have no dependency to ω_T, compare to that of the CS topology. This is, first, because the limiting impact of the Cgs1 is resonated out within the input matching network, and secondly, it is assumed that the operating frequency of the circuit is much lower that ω_T, and therefore the attenuating impact of the Cdg2 and Cgs1 on the drain current of M1 is neglected. It is assumed that i_{d1} is passed to the output intact; however if operating frequency is close to ω_T one needs to also consider the current division in the voltage gain expression as well.

60%. On the other hand, if the loss in the tank is entirely compensated using the G_q, the noise figure increases to 12.1dB which 11.2dB of it is because of the q-enhanced load (about 80%)[27]. Repeating the same experiment using an off-chip inductor or bondwire with a Q of 15, results in corresponding noise figures of 6.8dB and 9.2dB. Therefore, using a high quality inductor in the tank of the LNA helps to improve the sensitivity of the receiver by a few dB. The second experiment shows that as the Q of the tank increases the noise contribution of the input stage of the LNA become dominant. In this case, if the sensitivity requirements are still a concern one can apply the capacitive-cross-coupled input differential pair to further reduce the noise contribution of the M1 by 50% [94].

4.5 Linearity

The main purpose of using an LNA at the front-end of a receiver is to facilitate matching with the antenna, to amplify the incident RF signal before it is processed by high noise receiver circuitry and to help isolate the antenna from local oscillator tones and unwanted mixing products generated in the receiver chain. Usually, the LNA passes the entire band of interest to the next stage and therefore it is expected to maintain its linearity across the entire band; i.e. it needs to handle incoming signals from a few tens of micro volts minimum desired signal up to levels of in-band and out-of-band blockers which might be 40dB or more stronger than the minimum detectable level. In most cases, the tuned load of the LNA is made of passive components, and therefore the input transconductor is the main source of the non-linearity and therefore the focus of analyses.

[27] The device noise factor, γ, is assumed to be as bad as 2.

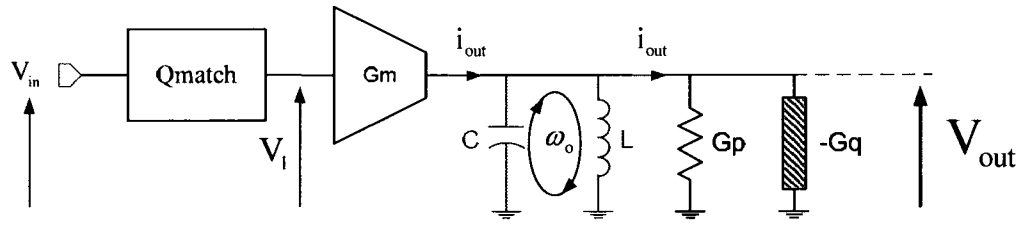


Figure 4.5 LNA stage decomposed in its building blocks. The negative transconductor which needs to handle the amplified signal is pushed into its non-linear region before preceding stages.

The non-linearity of stages following the LNA is of more concerns in the sense that they have to deal with an amplified and often wide-band signal.

The above statements are applied to ultra-narrow-band amplifiers too. Although the ultra-narrow band LNA has a very limited pass-band, the process of filtering itself is still dealing with a wideband incoming signal and therefore has to be linear across the entire band. The situation gets worse where the tuned load has to use active devices which impose their own range of linearity on the LNA. As shown in the Fig.4.5, the negative transconductor, Gq, responsible for narrowing the pass-band of the LNA, still needs to be able to manage a wideband amplified signal in order to fulfill its duty. The input matching network, the input transcondcutor (Gm), and the tuned load gradually see a larger voltage across their output. It is most likely that the Gq is pushed into its non-linear region before preceding blocks, and therefore the tuned load often defines the linearity of the LNA. Hence, the study of the non-linearity of the load is the focus of the rest of this section. A simplified but general analysis of the non-linearity of transconductors is presented mainly to highlight related issues and methodologies in characterization of a given topology.

One of the most commonly used negative transconductors is the cross-coupled differential pair. Fig.4.6 shows a hypothetical test-bench used to study the linearity of the

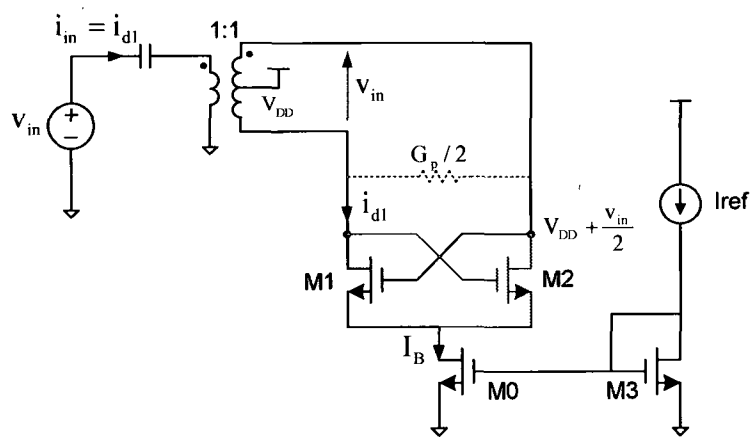


Figure 4.6 a test-bench for linearity analysis of the transconductor

transconductor. The use of a Balun allows biasing the transconductor to the equivalent condition to that of an actual LNA. In the following analysis, the LC tank is substituted by its equivalent loss. Similar to the analysis of the standard differential pair one can write

$$v_{in} = V_{gs1} - V_{gs2}$$

or

$$v_{in} = \sqrt{\frac{2.(I_D + i_{in})}{\mu.C_{ox}W/L}} + V_{th1} - \sqrt{\frac{2.(I_D - i_{in})}{\mu.C_{ox}W/L}} - V_{th2} \qquad (4.34)$$

where I_D is DC bias current flowing in each branch of the source coupled transistors. In short channel or narrow channel devices, the threshold voltage V_{Th} tends to rise with V_{DS} [84]. In our case study the drain-source voltage of the two devices in the differential pair are varying in two different direction, that is, for example, when input voltage v_{in} increases, V_{DS1} goes up while V_{DS2} falls down and vice-versa. Therefore $V_{Th1} > V_{Th2}$ and in Eq.4.34 there would be a voltage dependant component left. Here, for simplicity and without distorting the generality of the

discussion, we assume that $V_{Th1} = V_{Th2}$[28]. Therefore the input voltage can be expressed in terms of the input current as:

$$v_{in} = \sqrt{\frac{2I_D}{K}}\left(\sqrt{1+\frac{i_{in}}{I_D}} - \sqrt{1-\frac{i_{in}}{I_D}}\right) \tag{4.35}$$

where $K = \frac{1}{2}\mu C_{ox} W/L$.

After a simple manipulation, the current can be expressed versus voltage as:

$$\frac{i_{in}}{I_D} = -v_{in}\sqrt{\frac{K}{2I_D}}\sqrt{\left(1-\frac{K}{8I_D}.v_{in}^2\right)} \tag{4.36}$$

Substituting $I_B = 2I_D$ and $\frac{K}{4I_B} = \frac{1}{2}\left(\frac{g_m}{I_B}\right)^2$ results in:

$$i_{in} = -\frac{g_m}{2}v_{in}\sqrt{1-\frac{1}{2}\left(\frac{g_m}{I_B}\right)^2 v_{in}^2} \tag{4.37}$$

Equation 4.37 provides the foundation for the rest our analysis on the non-linearity. If $v_{in} << \sqrt{2}(V_{GS} - V_{Th})$ then the square root term in eq.4.37 can be replaced with its Taylor Series approximation[29] which results in

$$i_{in} = -\frac{g_m}{2}v_{in}\left(1-\frac{1}{4}\left(\frac{g_m}{I_B}\right)^2 v_{in}^2\right) \tag{4.38}$$

---

[28] Body effects and its effect on the threshold voltage and the effective transconductance of the devices are neglected. The channel length modulation is also neglected to simplify the argument. The mobility reduction as a result of normal field effect and velocity saturation in the short channel devices is also neglected; although taking the later into consideration may result in a more linear system.

[29] $\sqrt{1-x^2} \approx 1-x^2/2 \quad x<<1$

And therefore the small signal differential transconductance[30] is expressed as

$$G_q^{ss} \triangleq \frac{\partial i_{in}}{\partial v_{in}} = -\frac{g_m}{2} + \frac{3}{8} I_B \left(\frac{g_m}{I_B} \right)^3 v_{in}^2 \tag{4.39}$$

Let's define the linear part of G_q^{ss} as $G_{ql} = g_m / 2$ and the non-linear part as $\Delta G_q = \frac{3}{8} I_B \left(\frac{g_m}{I_B} \right)^3$.

On the other hand, based on Fig.4.5, the output voltage can be written in its constructing terms, that is

$$v_{out} = G_{in} \frac{1}{G_p^{diff} - G_q^{ss}} v_{in} \tag{4.40}$$

where $G_{in} = \frac{1}{2} Q_{match} g_{ml}$ and $G_p^{diff} = G_p / 2$. The voltage gain is compressed when $G_q^{ss} < G_{ql}$, that is, when the voltage dependant part of the G_q^{ss} (i.e. ΔG_q) is such that for a given output voltage the effective negative transconductance is lower than its linear part.

The gain compression point is defined as the ratio of the actual gain to the linear gain which is

$$\frac{G_p^{diff} - G_{ql}}{G_p^{diff} + G_q^{ss}} \triangleq \frac{1}{x} \tag{4.41}$$

where x is representing the compression ratio. Equation 4.41 can be further manipulated and written as

$$\frac{G_p^{diff} + G_q^{ss}}{G_p^{diff} - G_{ql}} = 1 + \frac{\Delta G_q}{G_p^{diff} - G_{ql}} \tag{4.42}$$

or

$$\frac{\Delta G_q}{G_p^{diff} - G_{ql}} = x - 1 \tag{4.43}$$

---

[30] The large signal transconductance can also be defined as $G_m^{LS} \triangleq \frac{i_{in}}{v_{in}} = -\frac{g_m}{2} \sqrt{1 - \frac{1}{2} \left(\frac{g_m}{I_B} \right)^2 v_{in}^2}$.

For example, for the 1dB compression point $x = 10^{1dB/20} = 1.122$.

Now all the ingredients are ready to calculate the compression point for the negative transconductance cell in the Fig.4.6. Combining eq.4.39 and eq.4.43 results in

$$v_{out}^{(x)} = \frac{2I_B}{G_{q1}} \cdot \sqrt{\frac{1}{3}(x-1) \cdot \frac{G_p^{diff}}{G_{q1}} \cdot \frac{Q_o}{Q_L}} \tag{4.44}$$

It is clear that the compression point power is proportional the inverse of the Q-enhancement ratio. Note that this is actually the output referred compression point and in order to reference it to input it has to be divided by the gain.

At the extreme, when G_{q1} approaches G_p^{diff}, the channel bandwidth is very narrow and the voltage gain approaches infinity and eq.4.44 becomes:

$$v_{out}^{(x)}\Big|_{Gq \to Gp} \cong \frac{2I_B}{G_p^{diff}} \cdot \sqrt{\frac{1}{3}(x-1) \cdot \frac{Q_o}{Q_L}} \tag{4.45}$$

where $Q_L >> Q_o$, which in turn results in a very small compression point. At the same time, the noise figure NF is at its maximum. So there is a need for a figure of merit to provide a more inclusive number considering the linearity and the noise in the system; this measure is known as the dynamic range and is discussed in the next section.

Before warping up our analysis on the non-linearity, it is insightful to take a look at a few particular cases. For a 1dB compression point x=1.122, eq.4.45 can be expressed as

$$v_{out}^{1dB} \cong 0.4 V_{dsat}^{(q)} \sqrt{\frac{Q_o}{Q_L}} \tag{4.46}$$

where $V_{dsat}^{(q)} = \frac{I_B}{G_p^{diff}}$. As expected, when there is no loss compensation applied, that is $G_q \to \infty$, or in other word $Q_L = Q_o$, the compression point is as large as $0.4V_{dsat}^{q}$. This would be the situation where the non-linearity of the input stage of the LNA might become the limiting factor.

4.6 Dynamic Range

From the noise analysis in section 4.4, it was learned that the load is a major contributor to the noise figure of the LNA; It is also shown that since the active part of the load or negative impedance generator is dealing with an amplified signal, its non-linearity is the dominant factor in defining the linearity of the LNA, particularly when the load is tuned to an ultra-narrow bandwidth. Although the noise figure by itself is a major factor in defining the sensitivity of the system, the non-linearity of each unit in the receive chain indicates the amount of distortion and un-wanted harmonics produced. Moving along the receive chain from the front-end to the baseband, one can monitor the amount of noise floor at every node and also the maximum level of the signal allowed to prevent the following stage causing a certain level of distortion in the system. The ratio of these to two measures defines the dynamic range of the system at that node. Depending on the type of distortions there are several definitions for the dynamic range such as Spurious Free Dynamic Range (SFDR), or 1dB compression point dynamic range.

The 1dB compression point dynamic range is defined as:

$$DR = \frac{P_{1dB}}{P_{min}} \tag{4.47}$$

where P_{min} is minimum detectable signal representing the noise floor at the node, and P_{1dB} is the power of the signal causing the gain to compress by 1dB[31]. Both quantities need to be powers referred to either input or output of the stage under study. Substituting the output referred powers in the above definition for the dynamic range results in

$$DR = \frac{(V_{out}^{(1dB)})^2}{\overline{v_{n,load}^2}\,B} \tag{4.48}$$

where the denominator is the output noise power generated by the load which is integrated across the channel bandwidth[32].

In our case study, the ultra-narrowband CG LNA, this represents the noise floor at the output. Substituting eq.4.28 and eq.4.46, the dynamic range when the load is tuned to a very narrow band ($G_{q1} = G_p^{diff}$) can be expressed by[33]

$$DR = \frac{0.2V_{dsat}^{q} I_B}{kT.(1+\gamma).B}\left(\frac{Q_o}{Q_L}\right)^3 \tag{4.49}$$

where $I_B = G_p.V_{dsat}^{(q)}$. By a rough approximation, one may Replace the transconductance overdrive voltage by half of the power supply[34] and come up with the following expression relating the dynamic range to the DC power consumption of the block and the Q-enhancement ratio,

---

[31] To be more accurate, based on this definition for DR, the denominator is total noise referred to either ports and can be calculated, knowing the noise figure of the stage and the bandwidth of the channel, that is kT(NF)B. As mentioned, since the first two terms in the NF expression is fixed, we put out focus on the contribution of the active load on the DR. To be precise, this will introduce an error in the DR prediction; one could define a rather unconventional metric of the inverse of DR, that is 1/DR, and easily stay loyal to the DR definition and conduct the argument.

[32] Not the bandwidth of the LNA.

[33] Since the compression point is calculated for the differential case, the single-ended noise power is doubled.

[34] $V_{dsat}^{(q)} = V_{GS} - V_{Th} \approx V_{DD} - V_{dsat0} - V_{Th} \approx V_{DD}/2$.

$$DR \cong \frac{(0.1)P_{DC}}{kT.(1+\gamma).B}.\left(\frac{Q_o}{Q_L}\right)^3 \tag{4.50}$$

This shows a very strong relationship between the dynamic range and the enhancement ratio; For a quality factor ration of 10 (i.e. $Q_L = 10Q_o$), the cost is a drop of the dynamic range by 30dB. This suggests the use of a higher quality inductor (or a tank) where there is less enhancement required to achieve to the desired bandwidth. This can be highlighted further by incorporating $\frac{1}{G_p}.\frac{Q_L}{Q_o} = \frac{1}{C.\Delta\omega_{3dB}}$ into the eq.4.50 and expressing the dynamic range in terms of the target bandwidth, power consumption and the intrinsic conductance of the tank (inductor) as:

$$DR \cong \frac{0.2P_{DC}}{kT(1+\gamma)B}.(\Delta\omega_{3dB})^3.\left(\frac{C}{G_p^{diff}}\right)^3 \tag{4.51}$$

For a desired bandwidth $\Delta\omega_{3dB}$ at each center frequency ω_o (defined by C)[35], and for a given power dissipation budget, using a tank with a higher quality factor (or lower G_p) results in a higher achievable dynamic range[36]. The strong relationship of the dynamic range with the tank capacitance in Eq.4.51 suggests the active loss compensation is more beneficial at lower frequencies (larger C) for a given bandwidth. This conclusion is trivial since loaded quality factor is higher at lower frequencies[37]. *Euqation.4.51 shows that for a given bandwidth, center frequency, DC power budget, and intrinsic load conductance, the maximum achievable dynamic range using the active loss compensation mechanism is fixed.* As it is shown shortly, *any attempt to increase*

[35] This defines the gain according to eq.4.20.

[36] Note that the dynamic range is derived for the extreme case where the loss in the tank is completely compensated and therefore the bandwidth would be zero.

[37] $Q_L = \frac{(G_p - G_q)^{-1}}{L.\omega_o}$

this dynamic range by improve linearity will cost noise performance of the circuit and therefore the dynamic range stays the same. However, if the dynamic range exceeds requirements, but a narrower bandwidth is desired, one can rely on active Q-enhancement techniques.

On a side note, it is interesting to see that "only" across a limited bandwidth, tuning to a higher ω_o requires reducing the tank capacitance and the effective load transconductance, $G_p - G_q$, by the same rate, in order to keep the bandwidth constant[38]. This is shown as follows

$$\Delta\omega_{3dB} = \frac{\omega_o}{Q_L} = \frac{L.\omega_o^2}{G_p - G_q} = \frac{1}{C.(G_p - G_q)} \tag{4.52}$$

Let's assume that the capacitance is changed by the small amount of ∂C relative to the C; According to eq.4.52, It is expected that the effective transconductance to be adjusted by $-\delta G_q$, that is

$$(C + \delta C)(G_p - (G_q - \delta G_q)) = \frac{1}{\Delta\omega_{3dB}} \tag{4.51}$$

or

$$C.\delta G_q = \frac{1}{\Delta\omega_{3dB}}.(\frac{1}{1 + \delta C/C} - 1) \tag{4.52}$$

Assuming the relative variation in capacitance is small, one can use the Taylor series approximation to further simplify the above expression to

$$C.\delta G_q \cong \frac{1}{\Delta\omega_{3dB}}\left(\frac{-\delta C}{C}\right)$$

Or

$$\frac{\delta G_q}{G_p - G_q} = \frac{-\delta C}{C} \tag{4.53}$$

[38] Constant load quality factor.

4.7 Linearization, Noise and Dynamic Range vs. Power Trade-off

So far our study shows that the non-linearity of the active load puts a tight constraint on the dynamic range and how narrow the bandwidth of the LNA can be made. The literature of analog circuit design contains a huge list of techniques used to improve the linearity of different circuits but they are mostly applicable in baseband circuitry. Every extra device used to improve the linearity demands its own linearity requirements, contributes to the noise and, more importantly, introduces extra poles to the system which limits the bandwidth and raise stability concerns. In addition, extra devices may require extra power. Therefore, at high frequencies the choices are limited.

Negative feedback is probably the first linearization technique that comes to mind. Fortunately, it can be easily incorporated into the negative transconductor used in Fig4.6. The result looks like what is shown in the Fig.4.7. The resistive feedback, $2R_{fb}$, used between two sources on M1 and M2 allows the input voltage to partially drop across the resistor in addition of two gate-source voltages and therefore reduces the stress on the devices.

It can be shown that applying the negative feedback desensitizes the effective transconductance to the non-linearity of the cross-coupled transistors. The effective transconductance is expressed as:

$$\frac{i_{in}}{v_{in}} = \frac{-g_{m,eff}}{2} = \frac{1}{2}\cdot\frac{-g_m^*}{1+g_m^* R_{fb}} \tag{4.54}$$

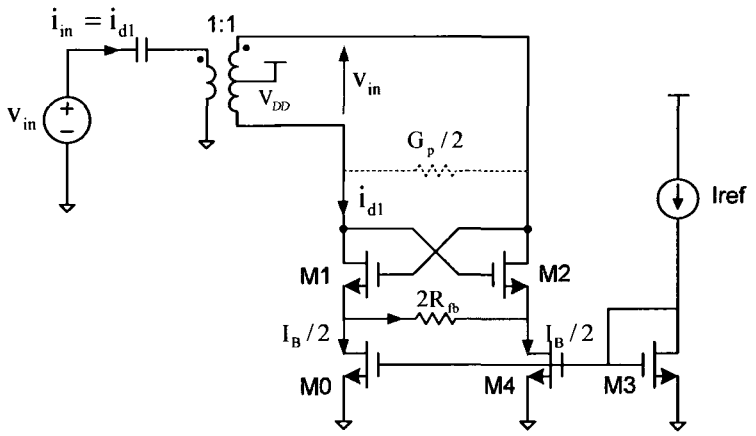


Figure 4.7 Linearization of the transconductance using resistive degeneration feedback

where g_m^* is the transcondcutance of each device. To preserve the consistency of the notation, eq.4.55 may be reframed as:

$$G_q^{fb} = \frac{G_q^{ss}}{1 - 2G_q^{ss}R_{fb}} \tag{4.55}$$

where G_q^{ss} is derived in eq.4.39 and its small signal or linear part is negative.

Taking the derivative of both sides with respect to the input voltage results in:

$$\frac{\partial G_q^{fb} / \partial v_{in}}{G_q^{fb}} = \frac{\partial G_q^{ss} / \partial v_{in}}{G_q^{ss}} \cdot \frac{1}{1 - 2G_q^{ss}R_{fb}} \tag{4.56}$$

That is the fractional variation of the G_q^{fb} is equal to the fractional variation in G_q^{ss} attenuated by a factor of $1 - 2G_q^{ss}R_{fb}$. At the limit, where the loop gain is much larger than unity, $g_m^*R_{fb} >> 1$, the effective transconductance is reduced to

$$G_q^{fb} \cong \frac{-1}{2R_{fb}} \tag{4.57}$$

Satisfying the loop gain condition results in the proportionally increase of the power consumption of the circuit. If the tank uses a 2nH inductor with the quality factor of 10 at 2.4GHz; that is the equivalent conductance of the tank is G_p^{diff}=1/300 S, the perfect loss compensation needs a 300Ω resistor between the sources of the devices. In this case, the loop gain condition demands $g_m^* >> 1/R_{fb}$ which ideally would be in the neighbourhood of $10G_p^{diff}$. Comparing this to the standard cross-coupled cell, where the required transconductance under the same condition is only G_p^{diff}, shows ten times more current consumption is the price paid for the linearity[39]. The immediate conclusion is that to save power one needs to employ an inductor with the highest possible quality factor (lower G_p^{diff}). On the other hand, following the large loop gain condition makes the transconductance of the input stage of the LNA have more influence in defining the linearity of the amplifier.

Although negative feedback helps to desensitize the load effective transconductance to voltage swing, the dynamic range depends also to the noise which the negative feedback cannot reduce. In fact, as it is shown shortly, at bests, the negative feedback may preserve the original noise characteristic of the stage.

Consider Fig.4.8 where the major noise contributors in the linearized cross coupled are shown for the half-circuit equivalent[40]. Using the cross-coupled nature of the circuit and the Lemma introduced in section 4.4, it is straightforward to calculate the effective noise at the input (output) ports of the circuit. Assuming the frequency of operation is much lower than ω_T, the

[39] Note that this is for the case where we prefer to ideally completely eliminate the non-linearity of the load. It does not necessarily have to be like this. One can optimize the value of the resistor to meet certain compression point, which in that case the power consumption increase will be more moderate.

[40] The half circuit noise is calculated in order to have a fair comparison with that of calculated for the half circuit LNA and to be able to re-use the results derived thus far. The noise contributions of bias current sources are neglected.

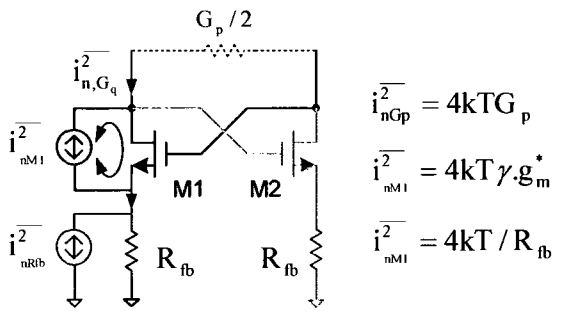


Figure 4.8 half circuit noise equivalent circuit; showing the noise current of M1 being dissipated in a degenerated loop, leaving the feedback resistor as the sole contributor to the output noise under the large loop gain condition

noise current of the feedback resistor is passed to the output (similar to a common-gate stage), and therefore $\overline{i^2_{Out,Rfb}} = 4kT/R_{fb}$. Following the Lemma, M1 channel thermal noise needs to consider the impedances seen looking into its source and its drain. Based on eq.4.2 and assuming the large loop gain criteria is satisfied, the impedance seen at the source of M1 can be approximated by $1/g^*_m$. This means almost the entire noise current is directed towards the source of the device rather than the feedback resistor.

On the drain side, the impedance is seen as $R_{fb} + (1 + g^*_m R_{fb}).r_o$, which again respecting the large loop gain results in an impedance much larger than the impedance of the Q-enhanced load (partially loss compensated). Thus the noise current prefers the low impedance path of the load. Interestingly, the current noise of the transistor has no net contribution to the noise at the port of the circuit if the large loop gain condition is observed. This is quite similar to the degenerated noise loop of the cascode device at the input stage of the LNA, described in Fig.4.4. The total noise power density at the output of the cell can be summed up as:

$$\overline{i^2_{n,G_q^{fb}}} \cong \frac{4kT}{R_{fb}} \tag{4.58}$$

The total noise power density of the transconductor "before" linearization which is basically the channel thermal noise of the devices constructing the cell is $\overline{i^2_{n,G_q}} \cong 4kT\gamma.G_q$ [41]; thus, under equal conditions and large loop gain it is expected that:

$$G_q = \frac{1}{R_{fb}} << g_m^* \tag{4.59}$$

$$g_m^* >> g_m \tag{4.60}$$

where the g_m is the required transconductance to provide $G_q = 1/g_m$ for the target loss compensation ratio. Equation 4.60 is translated into the fact that applying, so to speak, the linearizing feedback causes a much higher noise level at the output of the LNA. Once again the results emphasize using a tank with a high intrinsic quality factor to alleviate the noise performance degradation.

One might virtuously argue that applying the negative feedback had no impact on the dynamic range of the filter, as far as the output stage of the LNA concerns. The system is desensitized against the non-linearity of the load, however the noise power have gotten worse with the same ratio. It seems the high price of the increase in power consumption is paid for nothing! These are all true facts within the window provided so far. Let's look at the issue from a higher level.

By trading the noise for the linearity or vice-versa for a given (achievable) dynamic range, we are basically moving the fixed ratio window to the level that meets the system requirements. In a

---

[41] using a simple low-frequency small signal model and superposition one can show that the total noise current at the output of the standard cross-coupled transconductor can be expressed as $\overline{i^2_{nout}} = \frac{2*4kT\gamma.G_q}{(1-G_q/G_p)^2}$ where G_q/G_p is the loop gain. And therefore the voltage noise is $\overline{v^2_{nout}} = \overline{i^2_{nout}}/G_p^2 = \frac{2*4kT\gamma.G_q}{(G_p-G_q)^2}$. The similar result can be achieved by considering the effect of the positive feedback on the load separately and assuming the channel thermal noise is injected to the effective load.

receiver demanding a high sensitivity, any excess noise introduced by the components needs to be avoided and therefore, for example in our case study, the original standard cross-coupled pair are the topology of the choice. But in applications such as short-range wireless sensor networks, while the required sensitivity is moderate or low, the linearity of the receiver is the main challenge since the sensor node needs to handle strong in-band blockers. The following stages in the receiver are not able to compensate for the excess non-linearity introduced by the LNA, as the first stage in the chain. In this case a more linear transconductor is needed and therefore the linearized negative transconductance is the right candidate. Systems with the latter nature demand a more energy-efficient design and therefore the power consumption needs to be optimized. The bottom line is, in order to gain the highest dynamic range achievable in a given technology, it is a must to improve the noise figure and the linearity by using "passive" elements with the highest possible (and desirable) quality factors as the load while reducing the number of active elements in the signal path, as long as other performance metrics are satisfied.

4.8 Stability

Traditional microwave definitions of stability parameters and circles can be completely established from a set of small-signal (e.g., s-parameter) measurements on a two port element. Stability circles are collection of impedances which, when presented as a termination on one port, results in a unity reflection at the other port which can lead to instability. There are seemingly many measures of stability, often scalar metrics, encountered in microwave circles. Although in the course of a design we usually rely on such figures as phase margin, K1, or Bf, this section intended to highlight some of the stability issues from a circuit and topology point of view.

From the simplified representation given for the load so far, we can also deduce some stability criteria concerning relative large input voltages. The effective load conductance may never become zero or negative for any input. This means that the function describing the load, $G_{load} = G_p - G_q$, must have a positive minimum larger than zero. That requires first calculating the derivative of the function with respect to the voltage, and ensuring that the function is positive to the voltage value. To ensure that the found extreme is a minimum, the second derivative has to be larger than zero also.

Applying the above criteria to the large signal and the small signal expressions derived for load conductance, $G_p^{diff} - G_q^{LS}$ and $G_p^{diff} - G_q^{SS}$, shows that functions have minimums at the origin and they are stable as long as $G_p > G_{q1}$. This result is used in the proposed transceiver in this work to push the LNA into its unstable region and make it act as the oscillator in the transmitter.

The fact that the stability criteria resulted in no voltage dependant measure is the result of analysing the system in differential mode in which, ideally, the even order terms expressing the current versus voltage are null. In reality, non-idealities such as mismatch or substrate coupling introduce even order voltage dependency which might change stability conditions.

The above criteria are exercised only on the active load as if the preceding or following stages are perfectly isolated and have no impact on its behaviour. In fact couplings due to parasitics can cause dramatic changes on the state of the circuit. Such couplings may establish a negative feedback which can raise stability concerns. For example, in the case of the (ultra-narrow band) cascode LNA, where the voltage gain is high, the strong signal at the output can easily find its way to the gate of the cascode device through the gate-drain parasitic feedback capacitance. On the other hand, since the source of the cascode device, M2, as shown in Fig.4.9, is dominantly capacitive loaded, the impedance seen looking into the gate of M2 possess a negative real part.

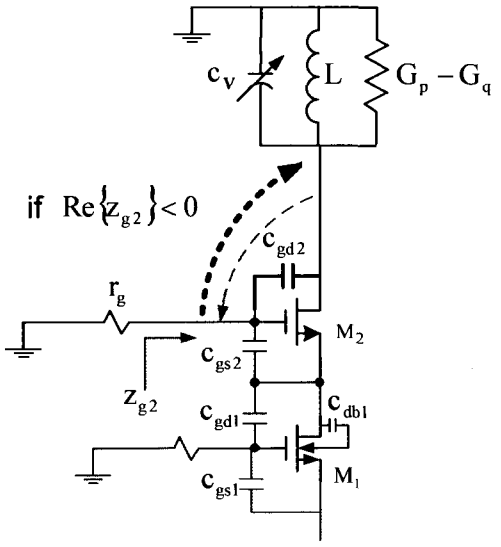


Figure 4.9 if the real part of the impedance seen looking into the gate of the cascode device becomes negative in conjunction with the high quality load can cause instability and eventually oscillation.

This impedance can be expressed as

$$z_{g2}\Big|_{C_{gd2}=0} \cong \frac{-g_{m2}}{c_{gs2}(c_{gd1}+c_{db1})\omega^2} + \frac{1}{j\omega c_{gs2}} + \frac{1}{j\omega(c_{gd1}+c_{db1})} \tag{4.61}$$

$$c_{d1} \overset{\Delta}{=} c_{gd1} + c_{db1} \tag{4.62}$$

If the real part of the effective impedance at the gate, i.e. $z_{g,eff} = r_g \,||\, z_{g2}$, is still negative, it can serve power to the LC load and potentially destabilize the stage[42]. Consider Fig4.10, where the feedback loop is broken at the gate of M2 while the loading effect of the input impedance of the cascode device, z_{g2}, is preserved at the coupling node. Now when an incremental voltage change at the output of the LNA is coupled to the gate of M2, it should normally create a current in $z_{g,eff}$ flowing to the ground. However, if the real part of the impedance is negative, this current needs to

[42] Ignoring the gate resistance of the device hides the negative impact of this coupling, as it is like the leaked power has been led to the ground.

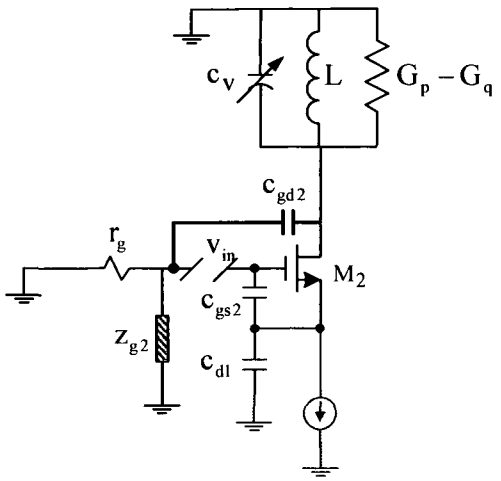
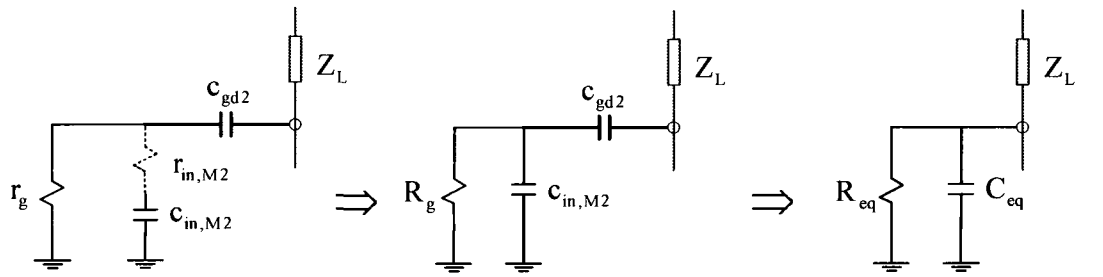


Figure 4.10 breaking the feedback loop while maintaining the loading effect of the capacitive impedance at the source of M2, on its gate

flow to the opposite direction, towards the coupling node and through C_{gd2}, the load, where it is stored in the magnetic field of the inductor, and increasing the gain of the LNA. The energy leaked to the gate is partly bounced back and restored in the inductor, helping to partially compensate the loss of the LC load, rather than being completely dissipated at the gate. If the magnitude of the equivalent negative impedance is large enough to perfectly or overly compensate for the loss in the tank, the resonator will oscillate. The effect of the feedback in transferring the negative impedance to the load has been shown in Fig.4.11 to help to quantify the impact. The series RC in Fig.11a is representing z_{g2} with its real and imaginary elements defined in the figure. The resistively loaded capacitive multiplier (divider) in Fig.4.11b translates the effective resistance at the gate, R_g, to its equivalent value, R_{eq}, seen by the load in Fig.4.11c. Note that the destabilizing influence of the



$$z_{in,M2} = r_{in,M2} + \frac{1}{j\omega.c_{in,M2}} \qquad R_g = r_g \parallel R_{in,M2} \qquad R_{eq} = R_g.\left(1 + \frac{c_{in,M2}}{c_{gd2}}\right)^2$$

$$r_{in,M2} = -\frac{g_{m2}}{c_{d1}.c_{gs2}}.\frac{1}{\omega^2} \qquad R_{in,M2} \cong (Q^2+1)r_{in,M2} \qquad C_{eq} \cong \frac{c_{in,M2}}{Q^2_{in,M2}} \parallel c_{gd2}$$

$$c_{in,M2} = \frac{c_{d1}.c_{gs2}}{c_{d1}+c_{gs2}} \overset{\Delta}{=} c_{d1} \parallel c_{gs2} \qquad Q_{in,M2} = \frac{c_{d1}+c_{gs2}}{g_{m2}}.\omega \cong \frac{\omega}{\omega_T} < 1$$

(a) (b) (c)

Figure 4.11 procedure of translating the effect of the impedance at the gate to the load

gate impedance is "boosted" via the capacitive impedance multiplier[43]. Therefore the total conductance of the load is

$$G_{load} = G_p - G_q - |G_{eq}| \tag{4.63}$$

This means the resonator is pushed further towards instability by $G_{eq} = 1/R_{eq}$. Figure 4.11c shows that the effect of the gate-drain capacitance on the center frequency of the bandpass filter (or the resonance frequency of the oscillator) is "reduced" by the effective capacitance seen at the gate of the cascode device[44].

[43] This is similar to the transformer impedance transformation. Note that this is an approximation; the exact admittance seen at the load is expressed as $Y_{eq} = \frac{j\omega C_{gd2} - R_g C_{gd2} C_{in,M2} \omega^2}{j\omega R_g (C_{gd2} + C_{in,M2}) + 1}$ which at sufficiently high frequencies, its real and imaginary parts is simplified to those used in Fig.4.11c.

[44] Approximately by one half.

The effective distributed gate resistance is proportional to the width of the device, that is

$$r_g = \frac{1}{3} R_{sh} \frac{W}{N^2 L} \tag{4.64}$$

where W is the total width of the device, R_{sh} is the sheet resistance of the gate poly-silicon which is about 4.5Ω/□, L is the length of the device and N is the number of fingers in parallel. If the fingers are connected at both ends, the distributed resistance is four times smaller than that of expressed with eq.4.64.

On the other hand, the quality factor of z_g can be expressed as

$$Q_{in,M2} = \frac{c_{dl} + c_{gs2}}{g_{m2}} \omega \cong \frac{\omega}{\omega_T} \tag{4.65}$$

In most cases, the operating frequency of the circuit is much less that the unity current-gain frequency and therefore the quality factor is much less than one. This translate into

$$R_{in,M2} \cong r_{in,M2} \tag{4.66}$$

$$C_{eq} = \frac{c_{in,M2}}{Q_{in,M2}^2} \tag{4.67}$$

According to eq.4.66, it is clear that $R_{in,M2}$ is proportional to the inverse of the width of the device for a given current density and therefore r_g and $R_{in,M2}$ are moving towards the same direction with width of transistors[45]. Equation 4.67 shows one can neglect the impact of the equivalent capacitance on the center frequency of the tank. One only needs to take the gate-drain capacitance of M2 into account.

[45] Another case of negative numbers; they are smaller as they get larger! The absolute values of two resistors though, move towards opposite directions.

It is concluded that in order to for the impedance at the gate to have a positive real value, or to stabilize the gate of the cascode device, the following condition must be negative if

$$r_g < \left|R_{in,M2}\right| \quad \rightarrow \quad R_g > 0 \tag{4.68}$$

Based on eq.4.64 and eq.4.66, for a given transconductance, increasing the number of fingers helps to reduce r_g and improve the stability. Combining the two equations shows the maximum allowed width for each finger for a given current density in order to stabilize the gate.

As showed earlier in this chapter, the common-gate topology with impedance transformation helps to save power by demanding a smaller transconductance compare to that of required by the standard common-gate topology; and since the smaller g_m is achieved using smaller width the stability condition is closer to be satisfied.

A numerical example may help to clarify the statement. The LNA with the width of 13um in 0.18um CMOS process provides a transconductance of about 5mS. Gross estimates for Cgs and Cgd are 13fF and 8fF. Neglecting the contribution of the drain-bulk capacitance results in $R_{in,M2}$ of -210kΩ at 2.4GHz; If the transistor is build using a single finger gate (worst case), it has a resistance of about 100Ω. Therefore effective parallel resistance on the gate is positive. However, if the standard common-gate topology requires 20mS transconductance and thus devices need to be four times wider for the same current density; this results in a four times bigger $R_{in,M2}$ of -50kΩ; and a single finger transistor will have a four times larger gate resistance or 400Ω which means the gate has a smaller positive real part. The stability criteria are harder to satisfy as operating frequency goes higher or the required transconductance is larger. Both of these situations are readily present in high power high frequency circuitries such as power amplifiers.

A similar approach can be taken to calculate the effect of the stage immediately following the LNA. For example, if a source follower is used as the buffer to drive the capacitive load (input

impedance) of the mixer or a pad, the negative real part of the input impedance of the buffer have the similar loss compensation effect on the inductively loaded LNA and may have significant impact on the stability considering the relatively larger size of the buffer.

4.9 Conclusion

Some of the main characteristics of the LNA and in particularly the ultra-narrow band LNA were presented. Parameters playing a major role in the input-output isolation were highlighted. It is also shown that for given power budget the common-gate topology with input impedance transformation facilitates the design of the input matching network by reducing its quality factor and therefore its sensitivity to the component variations and mismatch. The price paid for this gain is the degradation of the noise performance of the LNA, which is justifiable where the receiver specifications do not demand a high sensitivity. The linearity of the LNA is studied with emphasis on the impact of the active load on the compression point of the stage.

The dynamic range of the amplifier is shown to be determined by the bandwidth, center frequency, DC power budget and intrinsic quality factor of the resonator. Given these parameters, any attempt to improve the dynamic range is baseless. This is highlighted by studying a common practice linearization technique.

The stability of the cascode topology is discussed from a circuit point of view with emphasis on the impact of the preceding and following stages in destabilizing the amplifier.

In conclusion, as far as dynamic range and complexity concern, it is recommended to use a "passive" resonator with the best possible quality factor meeting the target bandwidth, while optimising the transconductance of the input stage to the cascode for the best linearity for the given power budget and avoid dealing with the active loss compensation.

5 RF Filter Tuning Mechanism

A novel frequency and bandwidth tuning mechanism is introduced in order to control and adjust those features of RF filters, in general, and front-end of a receiver, in particular. Although not fundamentally required for proper functionality of the proposed architecture in the previous chapter, in a parallel effort, this chapter challenges the idea of partial channel-filtering at early stages of the receiver front-end and shows its potential in mitigating the dynamic range requirements of mixers and baseband channel-select filters which might result in a considerable power savings.

5.1 Introduction

The narrow-band LNA further attenuates the image frequency and out-of-band blockers, and amplifies in-band blockers with less gain than that of the desired channel. This allows extracting data right at the IF or, through a simple down-conversion, at dc, without a complex and/or costly

image-reject mixer stage. It also reduces the linearity required of the following stages, i.e. mixers and baseband filters. Furthermore, the lower gain at the adjacent channels and blockers translates into a commensurately lower phase noise requirements from LO1 at respective offsets.

On the other hand, the narrowband LNA might provide a relatively high gain at the desired channel which in turn reduces the linearity of the stage. This means the receiver pays the costs of all the aforementioned benefits. A system level analysis is required in order to estimate the optimum gain and bandwidth settings for a given architecture.

The proposed methodology takes advantage of the existing architecture and re-uses the transmitter modulator and the receive chain in order to tune the LNA frequency and bandwidth. This results in a considerable saving in hardware, which could only be used during calibration; and just for a small fraction of the transceiver operating period- which tends to function more often in the transmit mode.

Phase locking mechanism, which is commonly used for the tuning of the center frequency of filters [82-85], is chosen to fulfill the similar purpose, without introducing any extra circuitry. The dual loop frequency synthesizer, i.e. the frequency translational loop in conjunction with the master PLL, which is used as the modulator in the transmitter, is exploited to operate simply as a phase locked system for the frequency tuning of the LNA. The proposed approach is based on a digitally controlled feedback system seeking a setting for which the envelope feature of the decaying oscillation converges to the target value.

5.2 *Proposed frequency and bandwidth Tuning Mechanisms*

The principle idea behind the proposed mechanism is the fact that the bandwidth of the desired filter (system) stays invariable in the receive chain as long as it acquires the narrowest bandwidth in the chain, while its quality factor is lower as one moves away from the antenna. That triggers this notion that one could monitor the bandwidth of the RF filter at the baseband which could be translated into a quality factor of the original RF-filter.

Prior of starting the receive operation, the LNA center frequency and the bandwidth need to be tuned. The self-tuning process is conducted through an iterative sequence of frequency tuning followed by a bandwidth tuning mechanism (BTM). The schematic of the tuning scheme is shown in Fig.5.1. In order to tune the center frequency of the LNA, the $-g_m$ cell incorporated in the LNA is used to over-compensate the loss of the tank, making the LNA oscillate and act as the VCO for the FTL. The coupled MPLL and FTL form a dual loop phase locked system. Once locked, a locked-signal is generated and opens the FTL at its loop filter, holding the steady state control voltage of the varactor on the loop filter integrating capacitor.

As soon as the center frequency of the LNA is tuned, reducing the $-g_m$ value results in a decaying oscillation whose envelope is extracted and used as a feature for monitoring and tuning the bandwidth of the LNA.

Assuming the target Q>>1 at RF, the envelope of the decaying oscillation at the LNA (former VCO) can be approximated by

$$v_{env}(t) = v_O.\exp^{-\frac{BW}{2}t} \tag{5.1}$$

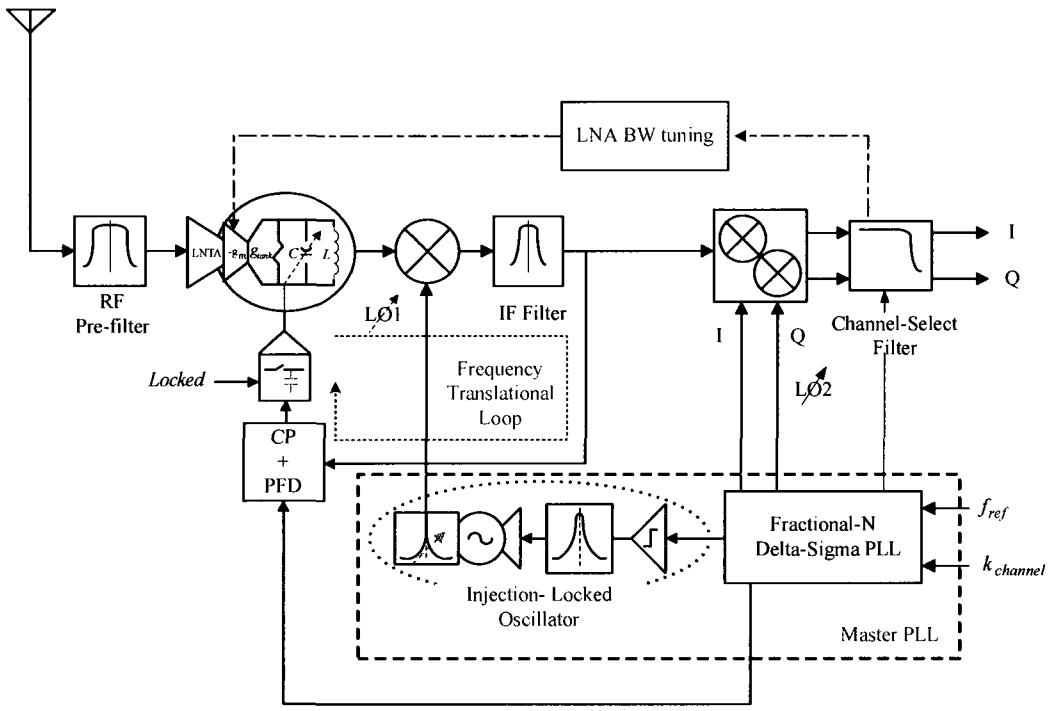


Figure 5.1 Receiver in the tuning mode

where $BW = (g_p - g_m)/C$ and g_p is the equivalent tank parallel conductance. At the steady state, for $g_m = g_p$ the envelope is constant, that is the LNA is functioning as an oscillator.

The system can be characterized by

$$\frac{\partial v_{env}(t)/\partial t}{v_{env}(t)} = -BW/2 \tag{5.2}$$

Form which the equivalent difference equation can simply be derived as

$$v_{env}[n + \Delta T] = (1 - \frac{BW}{2}.\Delta T).v_{env}[n] \tag{5.3}$$

where ΔT is the sampling period and must be smaller than half of the time constant of the system to satisfy Shannon criteria. Equation (3) implies that if $v_{env}[n]$ is known, the

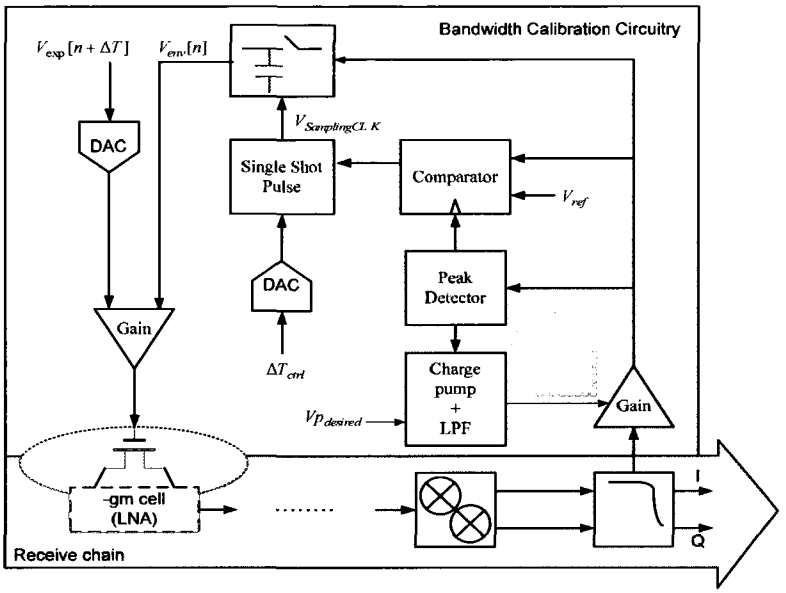


Figure 5.2 Block diagram of the LNA bandwidth tuning mechanism

expected amplitude after ΔT must be $v_{env}[n+\Delta T]$ providing the system time constant is $2/BW$. A system level realization of such interpretation is shown in Fig.5.2.

The RF mixer, half of the IF mixers, and first Biquad stage in the respective channel-select filter are exploited to extract the envelope. During the bandwidth tuning process, the bandwidth of the Biquad is changed to be wider than that of its normal operation in order to be transparent to the dynamics of the envelope. This is done through changing the transconductance of the gm-C filter. A comparator detects whether the decaying envelope crossed the pre-defined reference voltage ($v_{env}[n]=V_{ref}$); this will trigger the One-Shot Pulse generator which its duration is digitally controlled according to ΔT. The envelope amplitude, $v_{env}[n+\Delta T]$, is stored in the sampling capacitor. This value is compared to the expected amplitude for the desired bandwidth, $V_{\exp}[n+\Delta T]$. The result of this comparison, zero or one, is stored in a shift-register controlling the effective conductance of the $-g_m$ cell. This procedure continues iteratively until the first time the sampled value falls below the expected value.

The AGC loop is used to make sure that all envelopes are larger than certain level ($Vp_{desired}$). The peak detector is needed to assure that the sample is taken when the envelope is decaying.

Let's assume the V_{env1} shown in Fig.5.3 is extracted by the BTM. There are two points where the envelope crosses the V_{ref}. The first one which happens while the envelope is rising needs to be detected and dismissed. The peak detector is incorporated in the design to fulfill this mandate and used as an "enable" signal for the envelope sampling circuit.

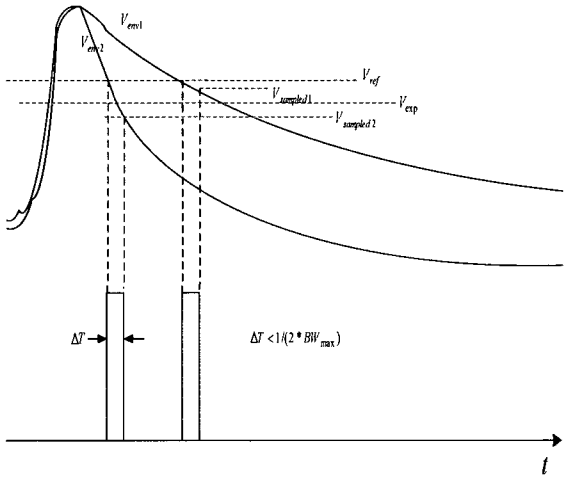


Figure 5.3 Envelope and single-shot sampling trigger pulse generated by bandwidth tuning mechanism

As soon as the envelope falls below the V_{ref} the Single-Shot Pulse (SSP) is generated; the envelope is sampled at the falling edge of this pulse and its value is compared with $V_{\exp}$ determining whether the bandwidth of the envelope is approaching the expected value. The width of the SSP is determined by the Shannon theorem and implementation limitations at extremes. A very short pulse demands a higher bandwidth and larger sampling capacitor of the sampling switch, as well as a high-resolution comparator in the final stage. A pulse wider than what Shannon theorem defines, results in a sampled value which is not truly representing the envelope.

During the tuning, when switching from VCO mode to the LNA mode with a limited but high Q, the center frequency of the LNA drifts; however, the desired channel still falls in the extended bandwidth, which is what matters at the end of the tuning process. Also, the center frequency might change in the normal operation when the LNA is receiving small inputs due to varactors

non-linearity, and voltage dependant parasitic capacitances of the varactor and $-g_m$ cell. This sets a limit on the minimum tuneable bandwidth, introducing a new trade-off between the maximum swing allowed in VCO mode and the linearity of the LNA. The smaller the VCO swing is, the lower the LNA linearity gets, albeit the narrower the tuneable bandwidth is going to be. A co-design methodology needs to be defined to meets the requirements of both modes of operation.

5.3 Circuit Implementation

The bandwidth tuning system introduced in system level in Fig.5.2 might be realized using the circuitry shown in Fig.5.4. The IF-mixers, at the top of the circuit, are shown along with their respective channel-select filters. The low-pass node of first Biquad filter in the Quadrature branch is fed into the gain stage in the bandwidth tuning circuit. The amplified envelope is sampled on the capacitor Cs using the sampling switch M1; the *comp2* compares the sampled value to the expected value. The SSB block is a programmable single-shot pulse generator, its rising and falling edges are used to trigger sampling and comparison. The peak detector built using the OTA, M3 and Cp. The charge-pump mechanism is employed in the quasi-gain control loop to adjust the gain of the envelope amplifier. The high dc gain of the charge pump reduces the steady state error to zero.

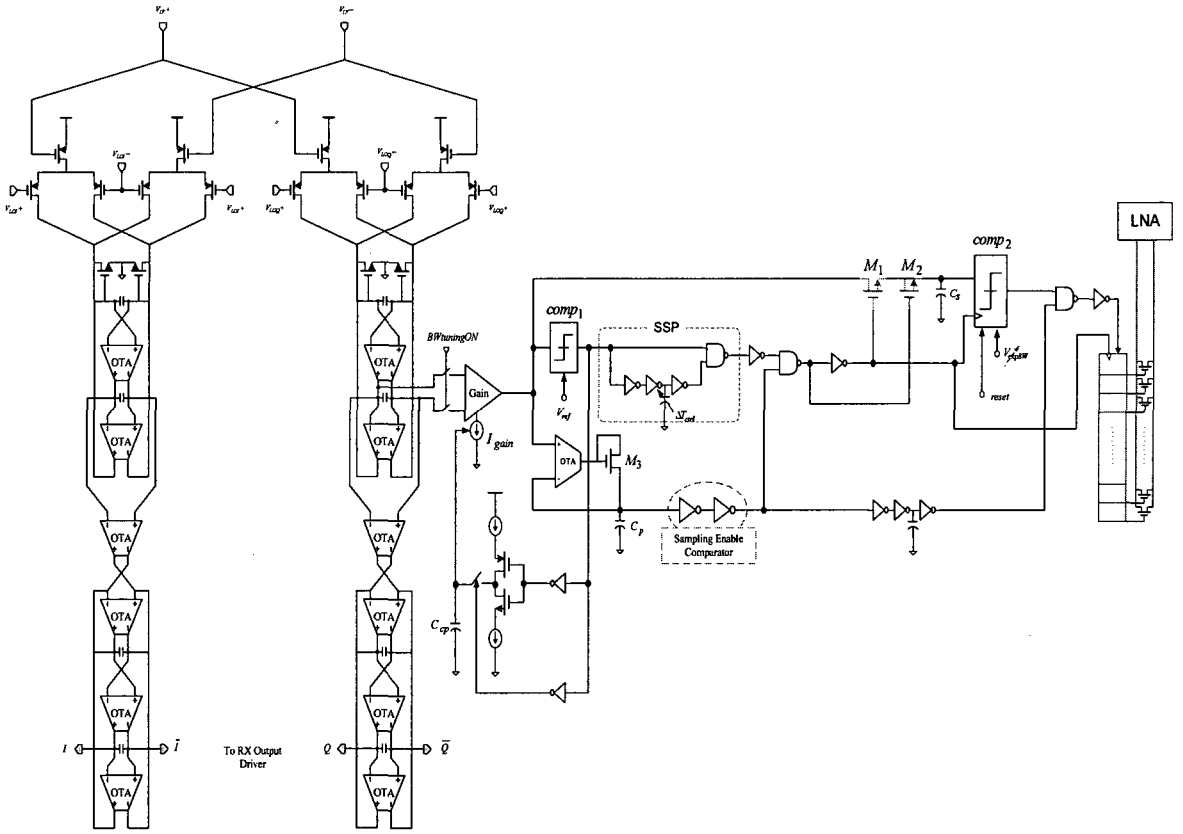


Figure 5.4 Schematic of the bandwidth tuning circuit

Fig.5.5 shows the simulation results for the bandwidth tuning circuitry. It is assumed that the center frequency has already been tuned and locked, the FTL loop is open, and then the bandwidth tuning circuitry is connected to the baseband filter with its bandwidth widened through changing its bias (gm). The single shot pulse is falling as soon as the decaying envelope passes the defined threshold of 0.75V. After about 100nsec the pulse rises and turns the sampling switch on. The envelope level of 0.68V is sampled on the sampling capacitor. This is equivalent of a bandwidth of about 290kHz.

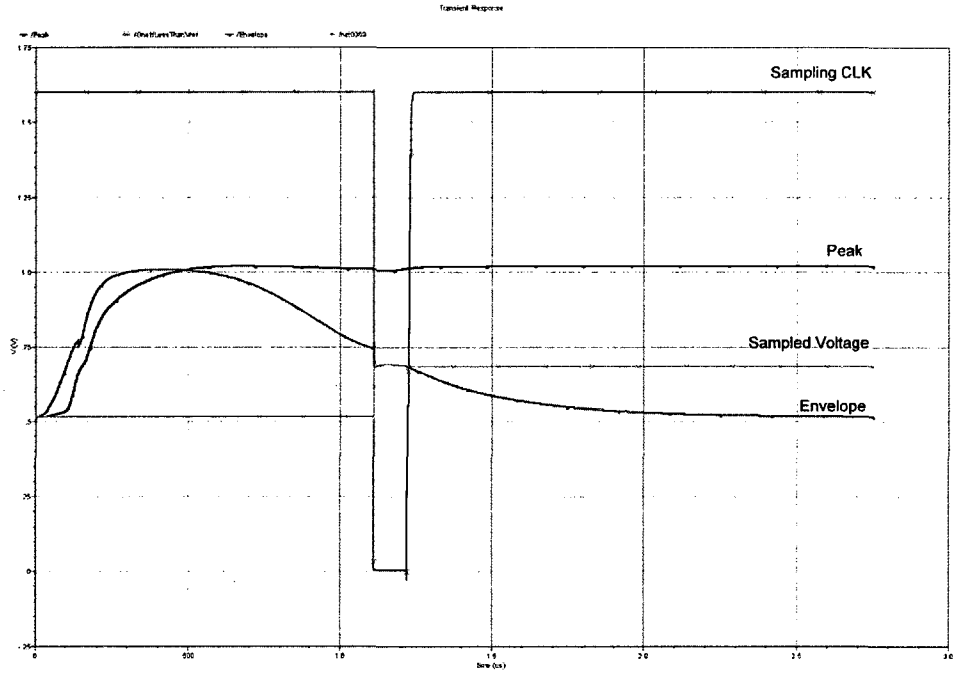


Figure 5.5 Simulation results showing sampling operation in the bandwidth tuning mechanism

5.4 Conclusion

A new methodology is proposed for the tuning of the frequency and bandwidth of an RF filter. The mechanism is employed in order to adjust the center frequency and bandwidth of the LNA in the receiver. The dual-loop PLL which is normally used for the transmit purpose is re-used for the frequency calibration. During the bandwidth tuning, the receive chain is re-used to extract the envelope of the decaying oscillation. Therefore the amount of hardware needed for the single purpose of tuning is minimized.

There are three major drawbacks in the proposed approach. First, after the FTL is settled, that is the center frequency of the filter is tuned, the loop is opened and that jeopardises the stability of the control voltage of the varactor. Second, the voltage-dependant parasitic capacitances, on both

amplifier transistors and varactors, have different values when the LNA is functioning as the tuned amplifier or as the VCO, which makes the LNA gain and center frequency voltage-dependant and prone to deviate when switching from one mode to another. And finally, the bandwidth tuning mechanism fundamentally drifts the LNA away from its tuned position; and because of parasitics, the actual position of the passband filter becomes unpredictable.

Provided the received package is short, and the varactors control voltage is kept well isolated from any source of interference such as substrate coupling, supply and ground fluctuations, it is observed that using a proper buffer one can assume that the control voltage stays stable after opening the loop.

The low power design technique proposed for the LNA results in small transistor considering the fact that they need to meet the constant current density required to optimize the noise figure. Therefore, their contribution to the parasitic capacitance would be minimal. On the other hand, the –gm cell is relatively big and its contribution to the resonator voltage-dependant capacitance is significant for its larger drain junction area. Proper layout strategies such as inter-digitizing would help to reduce these parasitics. However, the main culprit is the non-linear MOS varactor. Since the oscillation frequency of the VCO (LNA) is defined by the effective or rms capacitance over the period of the large swing output, a change in the amplitude of the output may shift the center frequency to a different position where the LNA gain in not maximum. The remedy for this problem would be to use linear varactors. The highest possible linearity might be achieved using switched-capacitors which require a DAC being introduced into the FTL.

Also, as long as the target channel falls well within the passband of the LNA, the frequency variation in operation mode switching seems to be tolerable.

6 Measurement Results

6.1 Introduction

The transceiver architecture presented in previous chapters has been implemented in a 0.18-μm CMOS process provided by TSMC through CMC. Fig.6.1 shows the die microphotograph of the fabricated chip. The chip takes an area of about 7.8 mm^2, mostly dictated by the number of pads required to provide the biasing of building blocks. An on-chip bandgap reference current generator could significantly reduce the size of the die.

The signal path in receive and transmit chains are differential. Careful attention is paid to symmetry throughout critical signal routings; such as LNA/VCO, RF and IF mixers, I and Q channel select filter.

Figure 6.1 Transceiver die microphotograph

Two different types of pads are used. For DC biasing or low frequency nodes, all the six metal layers are shorted together. To reduce the loss of the substrate, RF pads are realized with all the metal layers from metal2 to metal6 stacked on top of each other, shielded by the first metal layer at the bottom acting as RF ground (the bottom layer is connected to a ring surrounding the chip and connected to ground via multiple ground pads); RF pads are designed to have smaller area in order to lower pads capacitive loading on the sensitive building blocks. An alternative pad shielding, which results in lower capacitance, is done by covering the substrate under the pad by an active layer, p+ or n+, connected to ground. This is not exploited in this design, because the active layer is not allowed to be used under pads in this process. Also, immediate pads next to the RF pads are either grounded or connected to the power supply to reduce any coupling between high frequency pads and other pads.

Multiple supply pads are used to reduce supply bounce caused by the bond-wire inductance. Assuming a bond-wire inductance of 2-nH, and assuming the total current of 20-mA drained by

circuits (including biasing and buffers) switches on and off at 2.4-GHz, the total voltage drop across a single bond-wire will be 96-mV which is considerable. The current drawn by the actual receiver or transmitter is less than the above estimate but because the chip is tested with all the extra blocks starving for current, one needs to consider their impact on the performance before-hand and compensate for it. Also, in all the blocks, multiple substrate contacts are placed to reduce the effect of noise pick-up from the substrate. Guard rings, consisting of n-well contacts connected to the supply, have been also placed around the VCO, LNA, and dividers. Unfortunately, the design kit does not support a deep n-well option which could provide more isolation between NMOS devices and the substrate.

To reduce the coupling through supply between sensitive RF blocks and large signal drivers in the MPLL, two separate power supply routings are used throughout the chip which are eventually star connected to a common off-chip supply of 1.6-V.

All the non-RF pads are supported by ESD protection. ESD clamps increase the return loss if they are not compensated for their contributions on parasitic capacitance at the RF frequencies; and therefore are avoided in this design at RF ports.

6.2 Test Board Design

The transceiver loose die comprises 66 pads allowing biasing on-die circuitries, providing necessary off-chip tuning circuits such as matching networks or IF filter, and communicating the internal signals with the outside world. Handling this number of pads is not possible on the probe station a test board is designed to provide a medium between the chip, power supply and

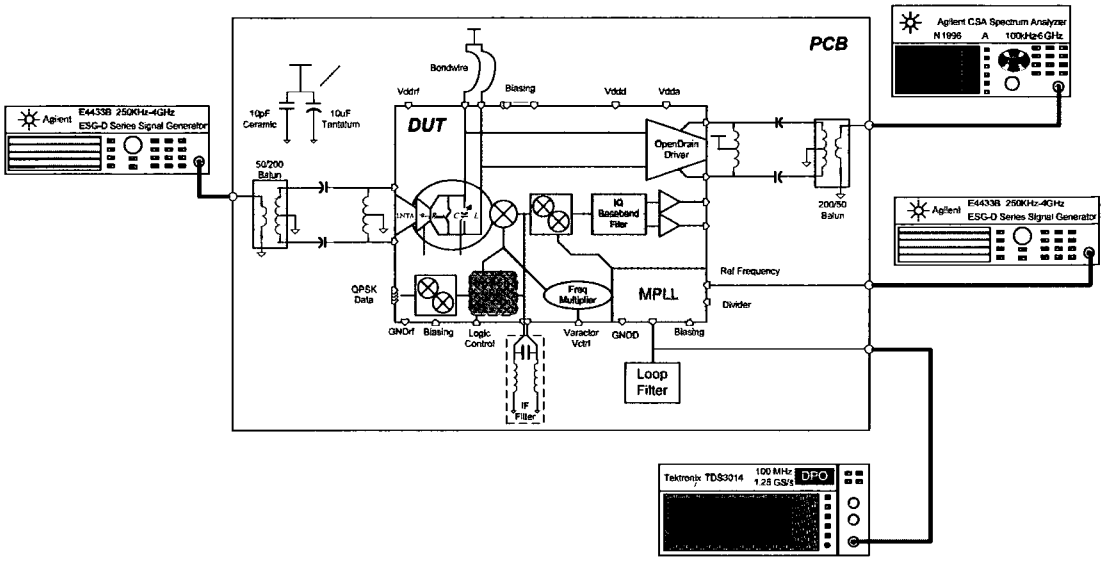


Figure 6.2 simplified diagram of the test bench, PCB and DUT connections

measurement instruments. The schematic of the test board is shown in the Appendix. Fig.6.2 shows the basic concept on which the test bench has been designed.

The PCB is built using a 4-layer 62-mil thick glass epoxy FR-4 material ($\varepsilon_r = 4.7$). The layout is done in PCAD. The layers are from top to bottom signal-ground-power-signal. The signal layers contain 1-oz copper. All the pads on the PCB are gold plated allowing wedge (or ball) bonding to the die. The impedance control is done on the routings connected to the LNA input and the output buffer. The characteristic impedance of the differential pair connected to the die is 200-Ohms differential and those connected to SMAs are 50-Ohms (single-ended). The die is attached to the PCB using conductive silver epoxy or regular super glue, on the allocated space covered by ground plane (which is connected to the global ground plane through large via holes). The chip is wire-bonded using a wedge bonding technique[46]. Ground pads are directly connected to the ground

[46] The first chip is bonded at NRC by Sylvain Laframboise; the rest are bonded at Carleton University.

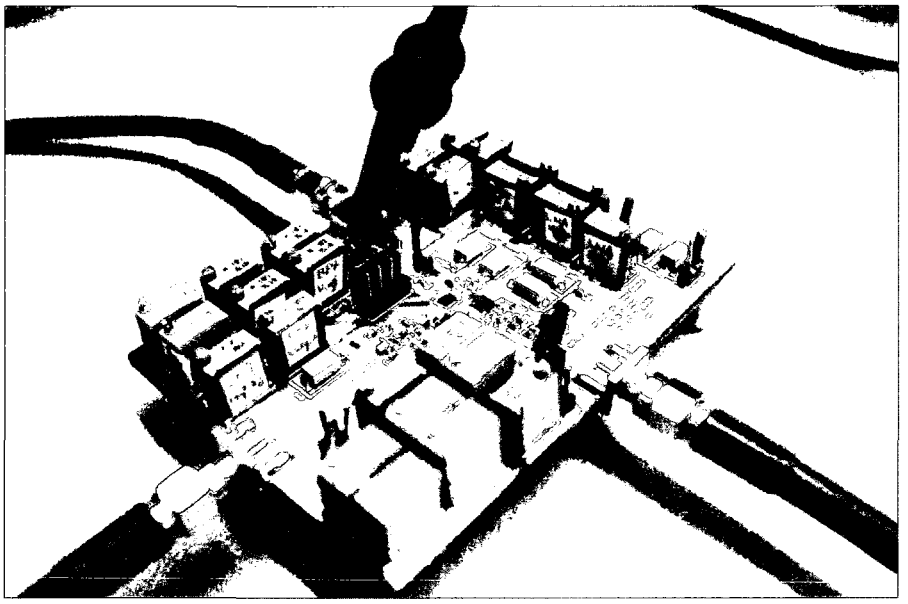


Figure 6.3 validation test fixture with DUT wire-bonded at the center

plane. Many via holes are used to produce low inductance grounds everywhere by connecting the top and bottom sides of the board (signal layers) to its ground layer. Fig.6.3 shows a photo of the test board with all components mounted.

6.3 Validation plan

Validating the basic functionality of the each major building block is chosen as the strategy to initiate the measurement process. The target circuits are master PLL (MPLL), frequency multiplier, and constant-envelope modulator. The next step is to test the functionality of the transmitter which is the combination of the above functional building blocks and an IQ up-converting mixer fed by OQPSK data. The test of the receiver is left for the last for its level complexity. Basically, once the functionality of the TX is proved, the remaining test is to open the OPLL loop and de-Q the LNA.

6.3.1 Master PLL

The focus of this section is mainly on the functionality of the frequency synthesizer. There are a few nodes (pads) which can be used to monitor and evaluate the dynamics of the MPLL; the following tests are mainly based on measuring the loop filter voltage or control voltage of the VCO versus the reference frequency. The output of the PLL (VCO) is not accessible, and therefore it is not possible to measure the phase noise of the synthesizer. The output frequency can also be changed by varying the biasing of the ring oscillator. The charge pump current is another variable which can be controlled off-chip; in the following tests the charge pump current is set to 200-μA, and the ring biasing voltage is fixed at 800-mV. Fig.6.4 shows a close-up micrograph of MPLL core on the die.

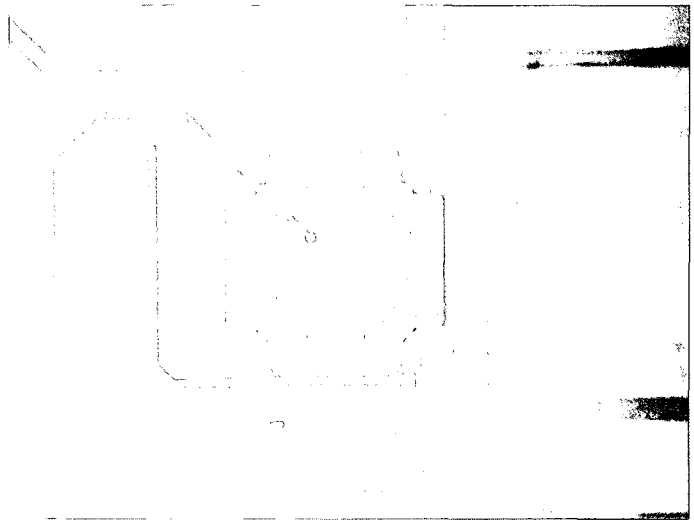


Figure 6.4 a close up look at the die microphotograph of the MPLL

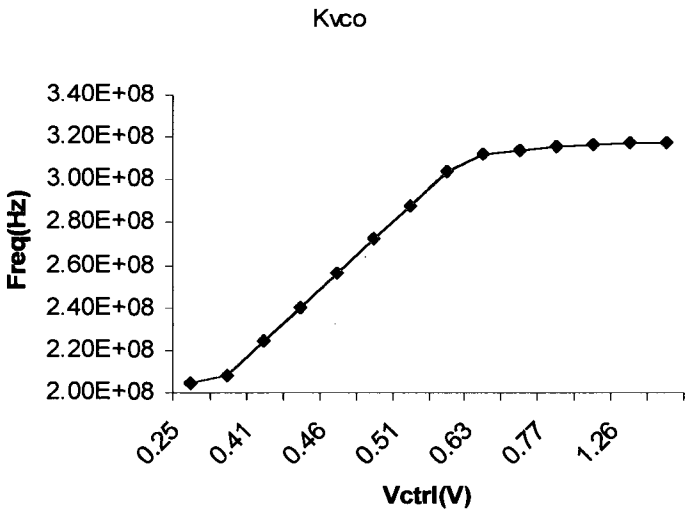


Figure 6.5 Measured VCO gain shows wide tuning range capability of the ring oscillator in MPLL[47]

The Kvco may be estimated by measuring the loop filter voltage for different values of reference frequencies. Given the divider value of 32, the reference frequency can be mapped to the output frequency of the VCO. The measurement results are shown in Fig.6.5. The figure shows a monotonic and wide locking/acquisition range of about 100-MHz, from 220-MHz to 320-MHz. This is clearly more than what is required by the proposed architecture. The curve can be moved to different offsets by changing the bias voltage on the delay cells forming the VCO.

By measuring the slope of the curve in the Fig.48 in the neighborhood of 305-MHz, the Kvco is estimated to be 186-MHz/V.

---

[47] Using CurveExpert™, it is turned out that the MMF model $f(x)=\frac{ab+cx^d}{b+x^d}$ with a=2.04e8, b=1.7e-4, c=3.17e8 and d=11.37 is the optimal description for Kvco in MMSE sense. The function is drawn with solid line in the Fig.48.

Next, to gain some insight on the dynamic behaviour of the loop, the reference frequency is modulated with different data rate and frequency deviations and transient response of VCO's control voltage is monitored. In the first experiment, a reference frequency of 9.88-MHz, corresponding an output frequency of 316-MHz is FM modulated with a fixed FM deviation of 20-kHz[48]. Fig.6.6 shows the captured photos of the output of the loop filter in the MPLL. From the under-damped transient response shown in the Fig.47a one can estimate the quality factor of the system to be approximately 2 (Q=2). This is corresponding to a damping rate of $\zeta = 0.25$. Measuring the period of the oscillation, T=80-usec, results in a damped natural frequency of

$$\omega_d = 2\pi / T = 2\pi * 12.5 \text{ krad/sec} \qquad 6.1$$

Then the natural frequency of the system can be calculated using

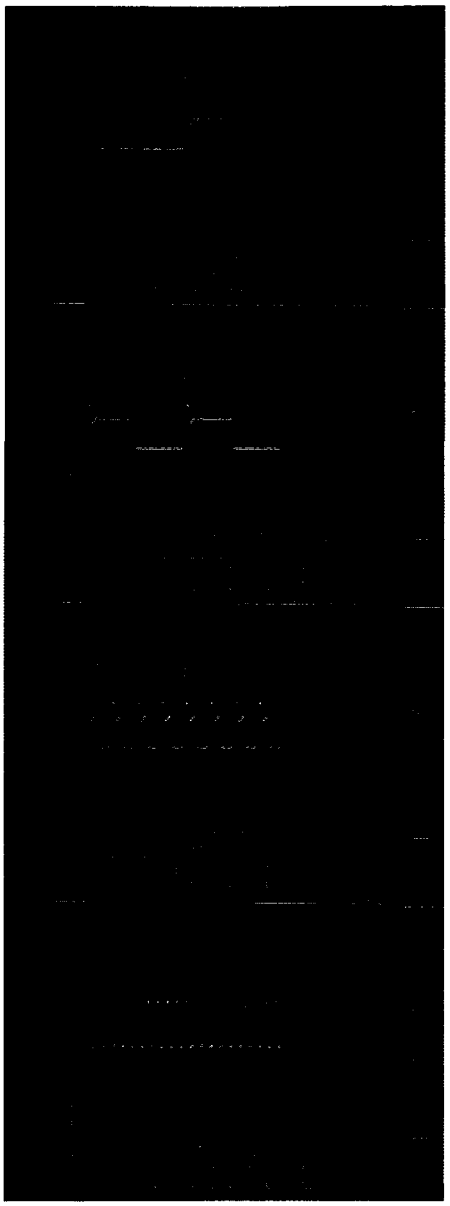


Figure 6.6 Loop filter voltage with an FM modulated reference frequency for different data rate settings

[48] This is not a full scale frequency step.

$$\omega_n = \frac{\omega_d}{\sqrt{1-\zeta^2}} = \frac{2\pi * 12.5\text{krad/sec}}{\sqrt{1-(0.25)^2}} = 2\pi * 13\text{krad/sec} \qquad 6.2$$

Therefore the loop bandwidth is stimated as

$$\omega_c = (1+\zeta\sqrt{2})\omega_n = 2\pi * 17.6\text{krad/sec} \qquad 6.3$$

The loop bandwidth dictates the maximum data rate the loop can handle; that is the shortest pulse width in which the loop can settle. Therefore an estimate for the highest bit rate would be about a quarter of the loop bandwidth or 4kbps. Fig.6.6 b to d confirm this estimate; as the bit rate approaches 4kbps the dynamic of the loop does not have enough time to settle, that is the output frequency of the VCO is not stable.

Measuring the current drained from the power supply also shows that the power consumption of the system is relatively independent of the data rate. Fig.6.7 shows the measured power = Idc*Vdd for different FM rate. This is power dissipated in the entire system, the actual chip and the biasing potentiometers between Vdd and ground. A sum of about 4-mW must be deducted from the measured values to remove the power drawn

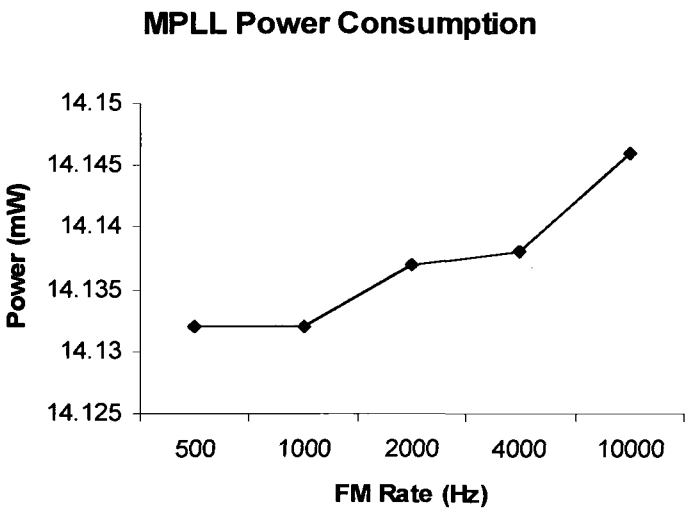


Figure 6.7 power consumption of the MPLL does not increase at higher modulation rate

in the biasing resistors. It can be seen that as the data rate increases from 500bps to 10kbps the power consumed by the system varies in the orders of a few tens of micro watts.

Although the architecture proposed in this thesis is based on a different type of modulation or method of up-conversion of the data, this test is a proof for the claim made in the previous chapters; as stated earlier power consumption of a PLL (or OPLL) does not increase by faster modulating signals, and therefore it is more energy conservative to have a wider loop bandwidth and to transmit the signals as quick as possible so that the system can be kept in the stand-by or off mode for a longer period of time.

In another experiment on the MPLL, the reference frequency is set to 9.88 MHz and the FM rate is kept at 1-Kbps. The FM deviation is then varied from 20-kHz to 100-kHz; and the voltage of the loop filter is monitored. Fig.6.8 shows the results of this test.

It can be seen that at steady state, the amplitude of the control voltage linearly increases with the frequency deviation (up to about 80-kHz). As the deviation approaches to the vicinity of 80-kHz the loop response enters into its nonlinear region (PFD) by showing cycle slipping behavior[49].

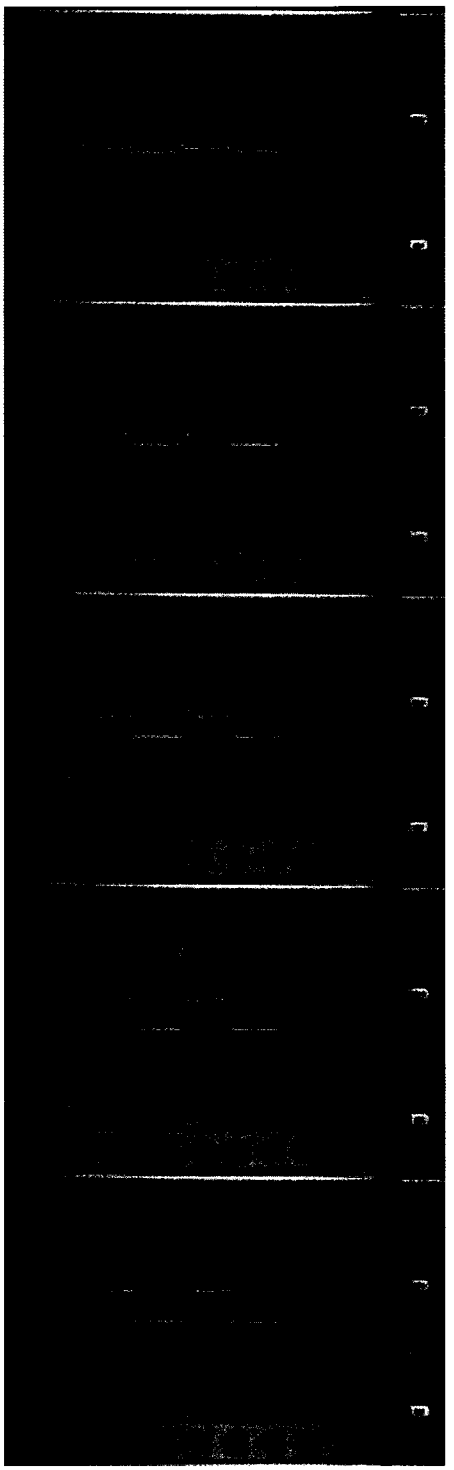
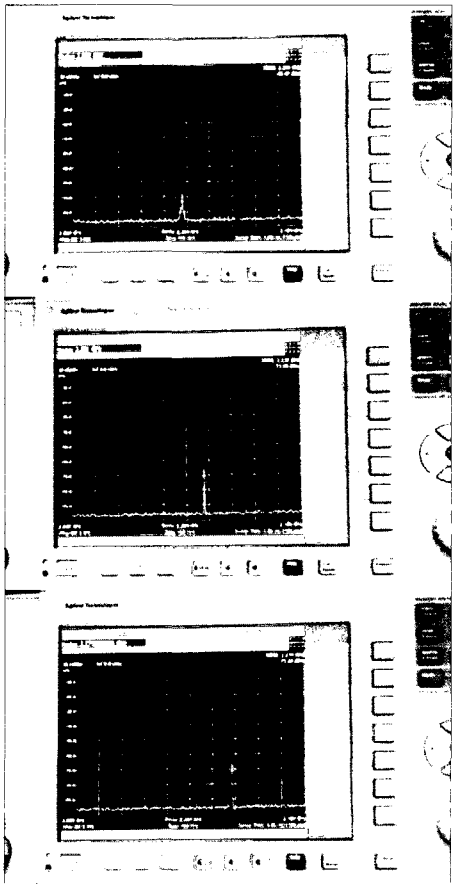


Figure 6.8 The loop filter voltage with FM modulated reference frequency for different FM deviation settings

[49] Non-linearity of the PFD is referred to its discontinuity every 2π once the phase difference goes beyond $\pm 2\pi$.

6.3.2 Frequency Multiplier

The output of the frequency multiplier is driving the RF mixer at its LO port and therefore there is no direct access to the output of the injection locked VCO. However when there is no RF input applied to the mixer, the leakage of the LO to the RF port through Cgd of the switching transistor can be observed at the output of the transmitter driver. Fig.6.9 shows the spectrum for three different values of the injection, proving that the RF LO is following the injection from MPLL. The power increases at higher frequency as the impedance of the parasitic capacitance reduces proportionally. At lower LO, the power of the leaked signal is low and buried in the noise floor and therefore it is not shown here.



a. 2.19GHz

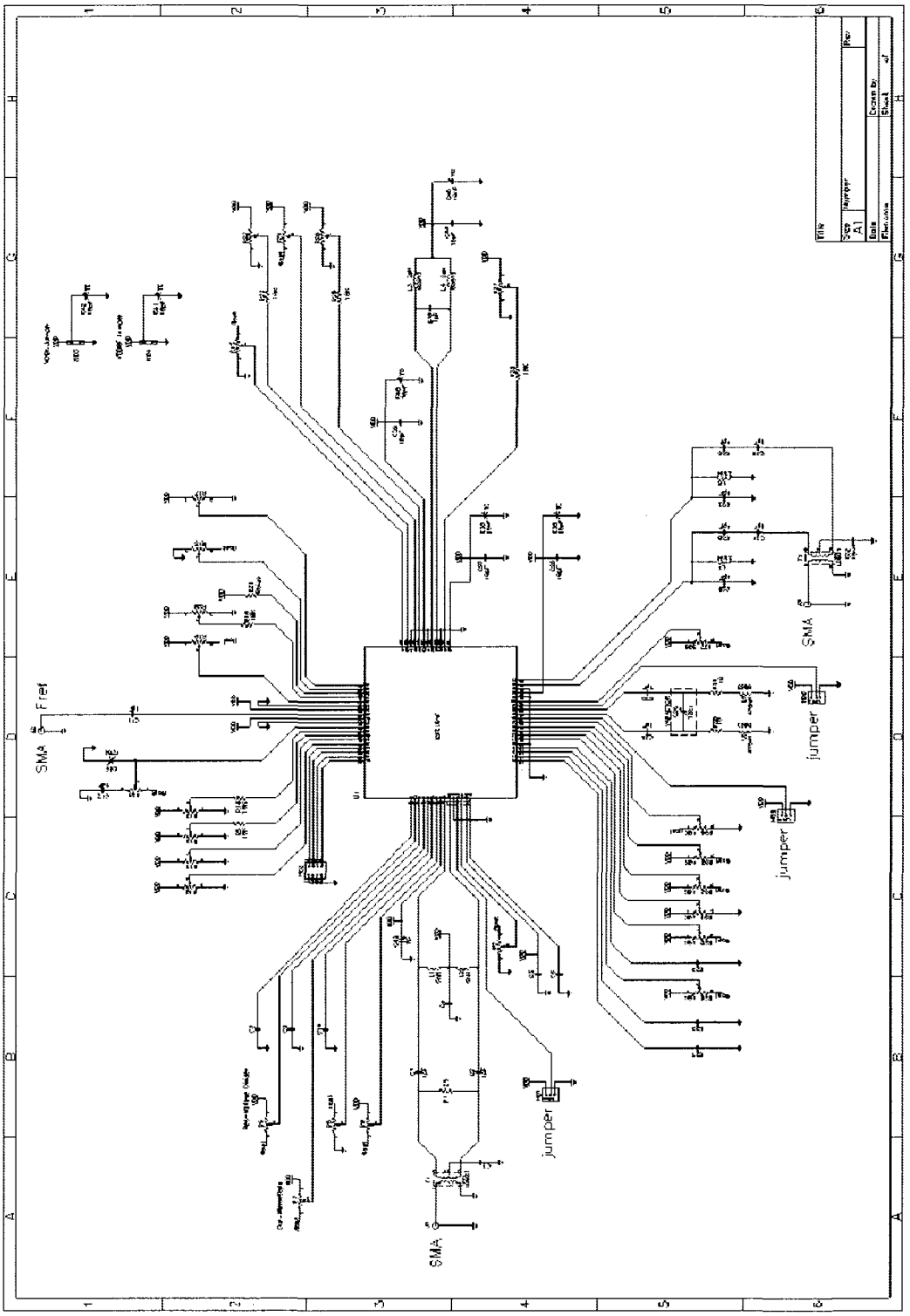
b. 2.23GHz

c. 2.27GHz

Figure 6.9. power spectrum of the output of the injection locked oscillator leaked to the output driver

Appendix A

Schematic of the test board.



6.4 *Summary*

Master PLL, frequency multiplier and output driver are proven to be functional and meet some of the design performance metrics. Although not precisely on the intended point of the architecture, the conducted tests proves one of the major arguments in this thesis that is invariance of the frequency synthesizer power consumption with modulating rate; and therefore, proposing the idea that it is more energy efficient to compress the data package and transmit it as fast as possible though a wide band modulator. In the proposed architecture though, this modulating loop is constructed by the offset PLL or constant envelope modulator which is different from MPLL; but the statement is general and fundamentally valid.

After testing a few chips, it turned out that two crucial building blocks, LNA and RF mixer are not functioning and as a result, it was not possible to close the OPLL loop and proceed with the rest of the validation plan to test the transmitter and receiver. Inspection of the layout does not show any clear fault in the design; surprisingly, varying the DC biasing of the RF mixer has no impact on its output common mode voltage which can be monitored through the off-chip IF filter. The voltage stays at the ground level; and since basic dc supply-ground test shows no short circuit between Vdd and ground, the only option would be an open circuit between the drain of the RF mixer and ground. A similar characteristic is observed on the LNA where the load is made of the bond-wires. The similarity of these two blocks is on their pad structures. These two circuits are the only ones using small RF pads WITHOUT ESD PROTECTION. A potentially valid hypothesis states that unprotected small devices on these two circuits are damaged by either HBM[50] or CDM[51] discharge phenomena in course of fabrication, wire-bonding or soldering the chip. Although a counter

[50] Human body model
[51] Charge coupled model

argument would be the probability of having all three tested chips damaged through the same cause.

7 Conclusions and recommendations

Seven years ago, at the beginning of this project, the author of these lines was well convinced that the use of active loss compensation did not improve the dynamic range of the LNA (or any circuits) subject to a given set of specifications or budget. However, one could move the lower boundary of the window defining the dynamic range to a higher level as long as the sensitivity requirement of the system allowed. This justified the use of linearization techniques to improve the linearity and pay the cost in the form of noise. At the same time, there was a demand for reducing the bandwidth of the first stage as much as possible in order to save power at the baseband.

The challenge of finding an affordable tuning mechanism for the center frequency and the bandwidth of the LNA or RF bandpass filter, ended up in the design of the entire receiver since some sort of lower frequency information was needed for the tuning of the center frequency. The bandwidth tuning was required some baseband features such as envelope to be extracted too. To avoid high sampling frequency (or sub-sampling version of that) a nested phase lock system was naturally developed (Frequency Translational Loop, FTL) for which the design of a master PLL was unavoidable.

Once the dual-loop fractional-N frequency synthesizer emerged, it revealed a new opportunity. What if we could take advantage of the established synthesizer and build the constant-envelope modulator for the transmit purpose; and furthermore, what if we could modulate the reference frequency of the FTL with data and build the full blown transmitter (without a PA). Further study of the characteristics of the dual loop transmitter proved potentials to handle higher data rates.

Now we had a single circuitry, which could work as both receiver and the transmitter. This idea was exciting and interesting enough in that it allowed us to drop the necessity of the bandwidth tuning mechanism or even adjustment of the center frequency for certain applications.

7.1 Conclusions

A new concept in transceiver architectures is introduced in which the receiver is used for the transmit purpose as well, or vice versa [86]. The re-use of major blocks to perform radically different tasks not only saves area but also, through an intelligent combination, offers an energy efficient transceiver. It is shown that the transmitter formed by the re-configured receiver offers the opportunity to modulate high bit-rate data, such as video, or a short packet of slow-varying data, such as temperature measurements. The modified high-IF dual conversion receiver uses two correlated frequencies as local oscillators, and it allows shifting the center frequency of the LNA accordingly, if it is necessary.

Although not limited to, the concept is presented as an architecture best fit for a wireless micro-sensor network to satisfy the expectations from such a link. It takes advantage of the fact that in most micro-sensor networks, the transmitter and the receiver (at each node) are not active at the same time. This creates an opportunity to maximize the "re-use" of building blocks, which may result in a significant saving on the silicon real estate and a reduction in die cost.

Meanwhile, the architecture itself is designed such that it is capable of receiving and transmitting phase or frequency modulated signal. It allows the transmitter modulating high bit rate data. This is achieved using a dual loop frequency synthesizer which also provides an opportunity to have more control on the start-up time and settling time. Power consumption is further minimized by using a few inexpensive off-chip components. The premise of such a combination is the manifestation of new low cost, energy-efficient and configurable transceiver architecture.

Furthermore, the thesis introduces a new mechanism for frequency and bandwidth tuning of RF-filters [98]. The original motivation for this challenging task springs upon the notion speculating a lower dynamic range requirements from the stages following the LNA provided the LNA bandwidth is kept narrow and centered around the desired channel. This could potentially reduce the power consumption significantly. The proposed tuning mechanism adds a new dimension to the architecture which turns it into a potential candidate for developing a software defined radio.

The prototype design is intended to demonstrate the functionality of the architecture in micro-sensor applications, but it can be adopted for any time-multiplexed system or standard and especially in hardware extensive structures such as phased-arrays.

A thorough analysis of the ultra-narrow band LNA shows that the best strategy to efficiently benefit from the new architecture is to use a high quality passive resonator in the LNA to gain the highest dynamic range achievable with a minimum power consumption. This significantly simplifies the architecture by avoiding the bandwidth tuning circuitry and the latency introduced to the protocol when it is active.

7.2 Future work and recommendations

The dual personality transceiver can be employed (and is preferred to) without any complex bandwidth tuning circuitry. The next generation of radio, based on the introduced architecture, should use a fully passive and high-quality resonator as the load of the LNA. The bandwidth must cover the entire band of interest. Therefore, the receiver is reduced to the standard super-heterodyne and can enjoy of the highest dynamic range achievable; it would now be capable of being optimized for any standard or application. In the transmit mode, the FTL loop still exploits the LNA as its VCO in the constant-envelope modulator. In a multi-standard application, the FTL can be reused in its original format to tune the LNA to a different "band"; this requires widening the tuning range of the LNA (VCO), while maintaining its sensitivity, dynamic range and stability. As emphasized in this thesis, the choice of transceiver architecture and type of modulation has direct impact on the performance of an energy efficient wireless sensor. The use of a hybrid charge-pump/loop-filter in the frequency synthesizer in order to reduce the start-up time is crucially important and needs further research. The author sees no significant advantage in using injection locking mechanism to generate the RF LO at the targeted frequencies. Strong AM to PM conversion creates spurs at neighbouring channels which reduces the SNR of the down or up converted signal. In fact, a more robust solution would be to have the slave VCO inside the MPLL replacing the ring oscillator. This significantly simplifies the design as there is no need to calibrate or adjust the free running frequency of a secondary VCO according to injected signal. However, at higher frequencies, injection locking or synchronization mechanism may be a solution to cope with clock distribution issues. It may also help to save power by avoiding high frequency divider in the MPLL.

Finally, the full utility of wireless devices can only be realized if both data transmission and supply of power to a node is wireless. The next generation of wireless sensor networks requires the energy production and storage mechanism to be integrated on-chip. The author is currently leading a research project in which a micro-scale battery can be deposited in the backside of the wireless sensor. The battery can continuously be charged by energy scavenging devices such as RF antennas, MEMS piezo-benders, mechanical torques or heat engines.

References

[1] A. Chandrakasan, R. Min, M. Bhardwaj, S.H. Cho, A. Wong, "Power aware wireless microsensor systems," *IEEE European Solid-State Circuit Conf. (ESSCIRC),* pp. 37-44, September 2002.

[2] G. Asada et al., "Wireless Integrated Network Sensors: low power system on a chip," *IEEE European Solid-State Circuit Conf. (ESSCIRC)*, pp.9-16, September 1998.

[3] A. Wang, S. Cho, C. Sodini, and A. Chandrakasan"Energy efficient modulation and MAC for asymmetric RF microsensor systems," in Proc. *Int. Symp. Low Power Design (ISLPED)*, 2001, pp. 106–111.

[4] A. Chandrakasan et al., "Design Considerations for Distributed Microsensor Systems," in *IEEE Proceeding of Custom Integrated Circuit Conference (CICC),*1999, pp. 279-286.

[5]A.Y. Wang et al., Energy efficient modulation and MAC for asymmetric RF microsensorsystems, *IEEE Intl. Symp. Low Power Electronics and Design,* 2001, pp. 106–111.

[6] Eugene Shih et al., Physical layer driven protocol and algorithm design for energy efficient wireless sensor networks, *Proc. MOBICOM,* 2001, pp. 272–287.

[7] J. Lee, B. Kim, "A low-noise fast-lock phase-locked loop with adaptive bandwidth control," *IEEE Journal o Solid State Circuits*, vol. 35, no. 8, pp. 490-502, Aug 2000.

[8] C.S. Vaucher, "An adaptive PLL tuning system architecture combining high spectral purity and fast settling time," *IEEE Journal of solid state circuits*, vol. 35, no. 4, pp. 490-502, Apr 2000.

[9] C-Y Yang and S-I Liu, "Fast-switching frequency synthesizer with a discriminator-aided phase detector," *IEEE Journal of Solid State Circuit*, vol.35, no. 10, pp. 1445-1452, Oct 2000.

[10] W.F. Egan, Frequency Synthesis by Phase Lock, Wiley-Interscience, New York, 1999.

[11] S. Willingham, M. Perrott, B. Setterberg, A. Grzegorek, and B. McFarland, "An integrated 2.4GHz frequency synthesizer with 5μs settling and 2Mbps closed loop modulation," in *IEEE ISSCC Digest of Technical papers,* 2000, pp. 200-201.

[12] S. Okazaki, S. Takahashi, and S. Higuchi, "Influence of rest-time in an intermittent discharge capacity test on the resulting performance of manganese-zinc and alkaline manganese dry batteries," *Progress in Batteries & Solar Cells,* 1987, v. 6, pp. 106–109.C.

[13] C.F. Chiasserini, R.R. Rao, "A model for battery pulsed discharge with recovery effect," *Proc. Wireless Communication Networking Conf.,* 1999, v. 2, pp. 636–639.

[14] C.F. Chiasserini, R.R. Rao, "Routing protocols to maximize battery efficiency," *Proc. MILCOM,* 2000, v. 1, pp. 496–500.

[15] IEEE, "IEEE 802.15 Working Group for WPAN," http://www.ieee802.org/15/.

[16] J. Lansford, P. Bahl, "The Design and Implementation of Home RF: A Radio-Frequency Wireless Networking Standard for the Connected Home," *Proceedings of the IEEE,* 1999.

[17] T.G Zimmerman, "Personal Area Networks: Near-field Intrabody Communication," *IBM Systems Journal,* vol. 35, no. 3, Jan. 1996, pp. 609-617.

[18] A. Kaleem, "Wireless Personal Area Networking Systems: A Comparison of Bluetooth, IrDA Data and HomeRF," *EPAC 2000*, pp. 1-15.

[19] B. Bougard et al., "Energy Efficiency of the IEEE 802.15.4 Standard in Dense Wireless Microsensor Networks: Modeling and Improvement Perspectives," *IEEE proceedings of the Design, Automation and Test in Europe Conference and Exhibition (DATE)*, pp. 1530-1591, 2005.

[20] A.P. Chandrakasan et al., "μAMPS: Micro Adaptive Multi-domain Power-aware Sensors," http://www-myl.mit.edu/research/icsystems/uamps.

[21] J.M. Rabaey et al., "PicoRadio support ad-hoc ultra-low power wireless networking" *IEEE Computer Magazine*, pp.42-48, Jul. 2000.

[22] J. Rabaey et al., "PicoRadios for wireless sensors networks: the next challenge in ultra-low-power design," *IEEE Int. Solid-State-Circuits Conf (ISSCC)*, pp. 200-201, February 2002.

[23] W.J. Kaiser et al., "WINS: Wireless Integrated Network Sensors," http://www.janet.ucla.edu/WINS.

[24] J. Kahn, R. Katz, and K. Pister, "Mobile networking for smart-dust," in *IEEE/ACM Conference on Mobile Computing and Networking (MobiCom)*, 1999.

[25] E.H. Armstrong, "Method of Receiving High Frequency Oscillations," US patent 1.342.885, 1920.

[26] U. L. Rohde and T. T. Bucher, *Communications Receivers Principles &Design*. New York: McGraw-Hill, pp. 63–74, 1988.

[27] W. Sabin and E. Schoenike, *Single-Sideband Systems & Circuits*. New York: McGraw-Hill, pp. 88–134, 1987.

[28] A. Rofougaran et al., "A 1 GHz CMOS RF front-end IC for a direct conversion wireless receiver," *IEEE J. Solid-State Circuits*, vol. 31, pp. 880–889, July 1996.

[29] B. Razavi, "Design considerations for direct-conversion receivers," *IEEE Transactions on Circuits and Systems II*, 44(6): 428-435, 1997.

[30] J. Crols and M. Steyaert, "A single-chip 900 MHz CMOS receiver front-end with a high performance low-IF topology," *IEEE J. Solid-State Circuits*, vol. 30, pp. 1483–1492, Dec. 1995.

[31] P. Choi et al., "An experimental coin-size radio for extremely low power WPAN (802.15.4) application at 2.4 GHz," in *IEEE Int. Solid-State-Circuits Conf. (ISSCC)* , Feb. 2003, pp. 92-93.

[32] J. Rudell *et al.*, "A 1.9 GHz wide-band IF double conversion CMOS integrated receiver for cordless telephone applications," in *IEEE Int. Solid-State-Circuits Conf. (ISSCC)*, Feb. 1997, pp. 304–305.

[33] D. Weaver, "A third method of generation and detection of single-sideband signals," *Proc. IRE*, pp. 1703–1705, Dec. 1956.

[34] Vaughan, R. G. Scott, N. L. White, and D. Rod, "The theory of bandpass sampling," *IEEE Transactions on Signal Processing*, vol. 39, pp. 1973–1984, Sept. 1991.

[35] D.H. Shen et al., "A 900 MHz integrated discrete-time filtering RF front-end," in *IEEE Int. Solid-State-Circuits Conf. (ISSCC).*, Feb. 1996, pp. 54-55.

[36] R. Mason, C. De Vries, E. Ivanov, "An RF Sub-sampling mixer, PGA and ΣΔ ADC for conversion at 900 MHz," ISCAS 2004, pp. 477-480.

[37] C. De Vries, R. Mason, "A 0.18μm CMOS 900 MHz receiver front-end using RF Q-enhanced filters," ISCAS 2004, pp. 325-328.

[38] H. Ahmed, C. De Vries, R. Mason, "RF Q-enhanced bandpass filters in standard 0.18μm CMOS with direct digital tuning," ISCAS 2003, pp. 577-580.

[39] S. Lindfors, A. Parssinen, and K. A. I. Halonen, "A 3-V 230-MHz CMOS decimation subsampler," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 50, no. 3, pp. 105–117, Mar. 2003.

[40] M. Helfenstein, G. S. Moschytz, *Circuits and systems for wireless Communications*, Kluwer Academic Publishers, 2000.

[41] R. Schreier, M. Snelgrove, "Bandpass Sigma-Delta Modulation," *IEEE Electronics Letters*, vol. 25, no. 23, Nov 1989.

[42] S. Jantzi, M. Snelgrove, P. Ferguson, "A fourth-order bandpass sigma-delta modulator," *IEEE Journal of Solid State Circuits*, vol. 28, Mar 1993.

[43] W. Gao, M. Snelgrove, "A 950-MHz IF Second-Order Integrated LC Bandpass Delta-Sigma Modulator," *IEEE Journal of Solid State Circuits*, vol. 33, May 1998.

[44] H. Tao, J. Khoury, "A 400-MS/s Frequency Translating Bandpass ΣΔ Modulator," *IEEE Journal of Solid State Circuits*, vol. 34, Dec 1999.

[45] J. van Engelen, J van de Plassche, E. Stikvoort, A. Venes, "A sixth-order continuous-time bandpass sigma-delta modulator fordigital radio IF," IEEE Journal of Solid State Circuits, vol. 34, Dec 1999.

[46] A. Namdar, B. Leung, "A 400-MHz, 12-bit, 18-mW, IF digitizer with mixer inside a sigma-delta modulator loop," *IEEE Journal of Solid State Circuits*, vol. 34, Dec 1999.

[47] J. Cherry, M. Snelgrove, W. Gao, "On the design of a fourth-order continuous-time LC delta-sigmamodulator for UHF A/D conversion," *IEEE Trans. on Circuits and Systems II*, vol. 47, no. 6, Jun 2000.

[48] R. Schreier, "A 10-300-MHz IF-digitizing IC with 90-105-dB dynamic range and 15-333-kHz bandwidth," *IEEE Journal of Solid State Circuits*, vol. 37, Dec 2002

[49] F. Gourgue, M. Bellanger, S. Azrouf, V. Bruneau, "A bandpass subsampled delta-sigma modulator for narrowband cellular mobile communications," in *Proc. IEEE Int. Symp. Circuits and Systems,* vol. 5, 1994, pp. 353-356.

[50] A.I. Hussein, W.B. Kuhn, "Bandpass ΣΔ modulator employing undersampling of RF signals for wireless communication," *IEEE Transactions on Circuits and Systems II, vol. 47, no.7, July 2000, pp. 614-620.*

[51] L.R. Carley,T. Mukherjee, "High-speed Low-Power Integrating CMOSSample-and-Hold Amplifier Architecture," in *Proc. Custom IC Conf.*, Santa Clara, CA, 1995, pp. 543–546.

[52] G. Xu, J. Yuan, "Performance analysis of general charge sampling," *IEEE Transactions on Circuits and Systems II: Brief papers*, 52(2): 107-111.

[53] G. Xu, J. Yuan, "A Low-voltage High-speed Sampling Technique," *Proceeding. 4th International conference on ASIC, 2001,* pp. 228 – 231.

[54] G. Xu, J. Yuan, "A CMOS 500 MS/s Charge Sampler," *Proceedings the IASTED International Conference on Circuits, Signals, and Systems*, Dec. 2004.

[55] J. Yuan, "A Charge sampling mixer with embedded filter function for wireless applications," *2nd International Conference on Microwave and Milimeter wave Technology*, Beijing, China, 2000, pp. 315-318.

[56] S. Karvonen, T. Riley, J. Kostamovaara, "A CMOS quadrature charge domain sampling circuit with 66-dB SFDR up to 10MHz," *IEEE TCAS-I*, vol. 52, no. 2, 2005, pp. 292-304.

[57] S. Karvonen, T. Riley, and J. Kostamovaara, "A low noise quadrature subsampling mixer," in *IEEE Int. Symp. Circuits and Systems (ISCAS 2001)*, Sydney, Australia, May 2001, pp. 790–793.

[58] Y. Poberezhskiy and G. Poberezhskiy, "Sampling technique allowing exclusion of antialiasing filter," *Electron. Lett.*, vol. 36, no. 4, pp.297–298, Feb. 17, 2000.

[59] Y. S. Poberezhskiy and G. Y. Poberezhskiy, "Sampling and signal reconstruction circuits performing internal anti-aliasing filtering and their influence on the design of digital receivers and transmitters," *IEEE Trans. Circuits Syst. I*, vol. 51, no. 1, pp. 118–129, Jan. 2004.

[60] T. Manku, C. Snyder, M. Ting, Y. Ling, J. Khajehpour, B. Kung, L. Wong, "Dual mixer down conversion architecture using complex mixing signals: enabling solution for software defined radio," in *Proc. IEEE Custom Integrated Circuit Conf., May 2002, pp. 227-234.*

[61] R.B. Staszewski, K. Muhammad, K.J. Maggio, D. Leipold, Direct Radio Frequency (RF) sampling with recursive filtering method, US 2003/0035499 A1, Feb003.

[62] R.B. Staszewski et al., "All-digital TX frequency synthesizer and discrete-time receiver for Bluetooth radio in 130nm CMOS," *IEEE J. Solid-State-Circuits*, vol. 39, no. 12, 2004, pp. 2269-2277.

[63] K. Muhammad et al., "A discrete time quad-band GSM/GPRS receiver in a 90nm digital CMOS process," in *Proc. IEEE Custom Integrated Circuit Conf.*, San Jose, CA, Sep. 2005, pp. 809–812.

[64] K. Muhammad et al., "A discrete-time Bluetooth receiver in a 0.13μm digital CMOS process," in *IEEE Int. Solid-State-Circuits Conf (ISSCC).*, Feb. 2004, pp. 268-269.

[65] E.H. Armstrong, "Some recent developments in the Audion receiver," *Proceeding of the IRE*, vol. 3, no. 9, pp.215–247, Sept. 1915.

[66] J. R. Whitehead, *Super-Regenerative Receivers*. Cambridge, U.K.:Cambridge University Press, 1950.

[67] W.F. Chow, "Transistor superregenerative Detection," *IRE Transactions on Circuit Theory*, March 1956, pp. 58-61.

[68] C.P. Wang, G. Wade, "Noise figure and sensitivity of a super-regenerative parametric amplifier," *IRE Transactions on Circuit Theory*, Dec. 1962, pp 321-327.

[69] A. Vouilloz, M. Declercq, and C. Dehollain, "A Low-Power CMOS Super-regenerative Receiver at 1 GHz," *IEEE Journal of Solid State Circuits*, vol. 36, no. 3, March 2001, pp. 440-451.

[70] A. Abidi, "RF CMOS Comes of Age," *IEEE Journal of Solid State Circuits*, vol. 39, no. 4, March 2004, pp. 449-461.

[71] T. A. D. Riley, M. A. Copeland, T. A. Kwasniewski, "Delta-Sigma Modulation in Fractional-N Frequency Synthesis," *IEEE Journal of Solid State Circuits*, vol. 28, pp. 553-559, May 1993.

[72] M. Perrott, T. Tewksbury, and C.G. Sodini, "27 mW CMOS fractional-N synthesizer/modulator IC," in *IEEE ISSCC Digest of Technical Papers*, Feb. 1997, pp. 366–367.

[73] D. McMahill, *Automatic Calibration of modulated fractional-N frequency synthesizers*, Ph.D. Thesis, Massachusetts Institute of Technology, 2001.

[74] S. Heinen et al., "A 2.7V 2.5GHz bipolar chipset for digital communication," in *ISSCC Digest of Technical Papers*, Feb. 1997, pp. 306–307.

[75] T. Cho et al., "A single-chip CMOS direct-conversion transceiver for 900 MHz spreadspectrum digital cordless phones ," in *ISSCC Digest of Technical Papers*, 1999, pp. 228–229.

[76] J.G. Proakis, *Digital Communications*, McGraw-Hill, New York, 1995.

[77] F.M. Gardner, *Phase lock techniques*, Wiley, New York, 1979.

[78] W. F. Egan, *Frequency Synthesis by Phase Lock*, Wiley-Interscience, New York, 1999.

[79] H. Darabi, Abidi A. A., "A 4.5-mW 900-MHz CMOS receiver for wireless paging," IEEE Journal of Solid State Circuits, vol. 35, pp. 1085--1096, August 2000.

[80] G. Groenewold, "The design of high-dynamic range contineous-time integratable bandpass filters," *IEEE Transaction of Circuit and System*, vol.38, no.8, pp. 838-852, August 1991.

[81] J. Rabaey, A. Chandrakasan and B. Nikolic, "Digital Integrated Circuits: A Design Perspective", 2/e, Prentice Hall, 2003.

[82] S. Pipilos, Y. Tsividis, J. Fenk, Y. Papananos, "A Si 1.8-GHz RLC filter with tunable center frequency and quality factor," *IEEE Journal of Solid-State Circuits,* vol.31, no. 10, pp. 1517-1524, October 1996.

[83] P. Katzin, B. Bedard, and Y. Ayasli, "Narrow-band MMIC filters with automatic tuning and Q-factor control," *IEEE Microwave and Millimeter-Wave Monolithic Circuits Symposium*, pp. 229-231, 1994.

[84] Y. Tsividis, "Integrated Contineous-time filter design – An Overview," *IEEE Journal of Solid-State Circuits,* vol. 29, no. 3, March 1994.

[85] R. Schaumann, "The Problem of automatic tuning in continuous-time integrated filters," *Proc. 1989 ISCAS*, pp. 106-109.

[86] R. Yousefi, R. Mason, "A 14-mW 2.4-GHz transceiver for short-range wireless sensor applications," *IEEE European Solid State Circuit Conference (ESSCIRC),* September 2008.

[87] W.B. Kuhn, N.K. Yanduru, S. Wyszynski, "Q-Enhanced LC Bandpass Filters for Integrated Wireless Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, no. 12, Dec. 1998.

[88] Y.P. Tsividis, *Operation and Modeling of the MOS Transistors*, McGraw-Hill, New York, 1987.

[89] B. Razavi, *Design of Analog CMOS Integrated Circuits*, McGraw-Hill, Boston, 2001.

[90] B. Razavi, *RF Microelectronics*, Prentice Hall PTR, Upper Saddle River, 1998.

[91] T.O. Dickson, et. al., "The invariance of characteristic current densities in nanoscale MOSFETs and its impact on algorithmic design methodologies and design porting of Si(Ge) (Bi)CMOS high-speed building blocks," *IEEE J. Solid-State Circuits*, vol. 41, no. 8, p. 1830, Aug.2006.

[92] ITRS-2005. [online]. Available: http://public.itrs.net.

[93] T. Soorapanth, S.S. Wong, "A 0-dB IL 2140±30 MHz Bandpass Filter Utilizing Q-Enhanced Spiral Inductors in Standard CMOS," *IEEE Journal of Solid-State Circuits*, vol. 37, No. 5, May 2002.

[94] W. Zhuo, S. Embabi, J. Pineda de Gyves, E. Sanchez-Sinencio, "Using Capacitive Cross-Coupling Technique in RF Low Noise Amplifiers and Down-Conversion Mixer Design," *IEEE European Solid State Circuit Conference (ESSCIRC),* September 2000.

[95] T.H. Lee, *Design of CMOS Radio Frequency Integrated Circuits*, Cambridge University Press, 1998.

[96] R. Adler, "A Study of Locking Phenomena in Oscillators," *Proc. IEEE*, vol. 61, pp. 1380-1385, Oct. 1973.

[97] E. Shumakher, G. Eisenestien, "On the Noise Properties of Injection-Locked Oscillators," *IEEE Transaction on Microwave Theory and Techniques*, vol. 52, no. 5, May 2004.

[98] R. Yousefi, R. Mason, "An Energy-Efficient Transceiver Architecture for Short-Range Wireless Sensor Applications," *IEEE Radio and Wireless Symposium,* San Diego, Jan. 2009.

[99] L.E. Larson, *RF and Microwave Circuit Design for Wireless Communications*, Artech house, Boston, 1997.

[100] U. Rohde, T. Bucher, "*Communications receivers Principles and Design*," McGraw-Hill, New York, 1988.

[101] R. van de Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters,* 2nd edition, Kluwer, 2003.