UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL AND APPLIED SCIENCE

SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

NANO RESEARCH GROUP

SPECTRUM MONITOR FOR COGNITIVE RADIO

By

Siwen Liang

Thesis for the degree of Doctor of Philosophy

December 2010

Supervisor: Professor William Redman-White

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL AND APPLIED SCIENCE
SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

SPECTRUM MONITOR FOR COGNITIVE RADIO

by Siwen Liang

The concept of a cognitive radio assumes that the receiver is able to determine the activity level across a large range of spectrum in order to assign a channel for its use. Hence a key function is a spectrum monitor to detect the spectrum availability as a first step. This thesis explores the requirements and design issues for the spectrum monitor receiver. The main challenge of this receiver design is to draw a spectrum map covering a wide range of frequency that is fast and accurate enough while consuming low power compared with the main transceiver circuits. The history and applications of the concept of cognitive radio (CR) are overviewed, followed by a wideband receiver architecture review, giving a wide range of scheme options for the proposed spectrum monitor. The concept of figure of merit (FoM) is then introduced. This concept helps to predict the performances versus power consumptions for active components over the next few years. By exploring the trend and relationship among FoMs, performances and time scales, a design approach is obtained to be used as a guide for system level receiver budget design. Then the spectrum monitor architecture is explored depending on the application and the figures of merits. For a representative cognitive radio application it is shown that the dual-down conversion architecture is suitable for the spectrum monitor, and the system specifications are given. Using these system specifications, the circuit level design of two of the key blocks is explored, where the requirements are significantly different from conventional designs reported in literature. A wide tuning range ring oscillator based PLL that is suitable for the frequency conversion and tuning functions is designed, fabricated and tested. A design method for high frequency on-chip bandpass filters is presented and experimentally tested. Comments on this research and future works are finally discussed.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Siwen Liang, declare that the thesis entitled

……Spectrum Monitor for Cognitive Radio ……

and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- Parts of this work have been published before submission as

Siwen Liang; Redman-White, W.; , "Lumped element band pass filter design on 130nm CMOS using delta-star transformation," *Research in Microelectronics and Electronics, 2009. PRIME 2009. Ph.D.* , vol., no., pp.32-35, 12-17 July 2009

Signed: ……

Date: 19th Jan, 2011

Acknowledgements

I'd like to thank Professor William Redman-White for his wonderful supervision, kind guidance and patience. I not only learnt knowledge in this area but also learnt the way of thinking. His inspiration and passion has been and will continuously be encouraging me all through my life. Particularly thank for the supports on the chip fabrication and measurements.

Really appreciate the important advice from Professor Lajos Hanzo and Dr Lie-liang Yang for academic issues. Also, I have got lots of help from Bill's colleagues in NXP. Olivier Tesson in Eindhoven helped on inductor modelling and simulation in ADS. Dr Domine Lenaert helped with the RF on-chip measurement of passive band pass filters. Dr Martin Bugbee of NXP Southampton helped with the PLL Cadence library, DRC and tape-out the PLL.

Thank also to Professor Peter Ashburn for financial support for chip fabrication.

I would like to express gratitude to my colleagues for help, advices and supports during the past years, they are Dr. Harold Chong, Dr. Yoshishiqe Tsuchiva, Dr. Kian Kiang, Dr. Xiaoli Li, Dr. Du Yang, Kai Sun, Yudong Wang, Ricky Chen and Ke Li. It has been a great and memorable time to work with them, and it's an honour to be their colleague.

Finally, I am deeply grateful to my parents and my fiancée Jing Liu for encouraging and supporting me, and for staying with me all the time along.

Definitions and Abbreviations Used

ADC	Analogue to Digital Converter
ADS	Advanced Design System
ASITIC	Analysis and Simulation of Spiral Inductors and Transformers for ICs
BPF	Band Pass Filter
CML	Current Mode Logic
CP	Charge Pump
CMOS	Complementary Metal-Oxide-Semiconductor
CR	Cognitive Radio
DAC	Digital to Analogue Converter
DECT	Digital Enhanced Cordless Telecommunications
DFS	Dynamic Frequency Selection
EM	Electro-Magnetic
ENOB	Effective Number of Bits
EvDO	Evolution Data Optimized
FCC	Federal Communications Commission
FOM	Figures of Merit
GSGSG	Ground-Signal-Ground-Signal-Ground
GSM	Global System for Mobile Communications
HSDPA	High Speed Downlink Packet Access
IF	Intermediate Frequency
IIP3	Input-referred 3rd order Intercept Point
IRR	Image Rejection Ratio
ITRS	International Technology Roadmap for Semiconductors
JFET	Junction gate Field Effect transistor
LNA	Low Noise Amplifier
LO	Local Oscillator
LPF	Low Pass Filter
MAC	Media Access Control
OFDM	Orthogonal Frequency Division Multiplexing
PAPR	Peak to Average Power Ratio
PCS	Personal Communications Service
PFD	Phase Frequency Detector

PGS	Patterned Ground Shield
PLL	Phase Locked Loop
RF	Radio Frequency
RKRL	Radio Knowledge Representation Language
RMS	Root Mean Square
SDR	Software Defined Radio
SFDR	Spurious Free Dynamic Range
SNDR	Signal to Noise and Distortion Ratio
SNR	Signal to Noise Ratio
TPC	Transmit Power Control
UMTS	Universal Mobile Telecommunications System
UWB	Ultra Wide Band
VCO	Voltage Controlled Oscillator
VGA	Variable Gain Amplifier
WCDMA	Wideband Code Division Multiple Access

Chapter 1 Introduction

1.1 Overview of Research

In this PhD project, the research interest is cognitive radio (CR) transceiver architecture design. This concept was presented by Joseph Mitola [1] in 2000 in his dissertation of Doctor of Technology in Royal Institute of Technology (KTH) in Sweden. In 2003, the FCC (Federal Communications Commission) gave the definition of cognitive radio as "A radio that can change its transmitter parameters based on interaction with the environment in which it operates" [2]. Generally speaking, cognitive radio is an emerging approach for using the existed precious radio spectrum resources more effectively. This concept is considered as the extension of the Software Defined Radio (SDR), which is already in use to some extent in some modern communication systems such as Wi-Fi (IEEE 802.11a/b/g) and Bluetooth etc.

A cognitive radio must have the following properties:

1. Sensing: RF technology that "listens" the huge swaths of spectrum
2. Cognition: Ability to identify primary users
3. Adaptation: Ability to configure the transmit power, frequency and modulation intelligently and flexibly to best use white spaces and minimize interference to primary users.

And in addition, for mobile applications, it also need:

4. Low cost: Not much extra cost to the entire radio function.
5. Low power: Power consumption must be low enough for portable application.

The main research work in this project is focused on the first step of a cognitive radio, namely, spectrum sensing. A high quality RF receiver monitoring occupation of the widely used spectrum is the goal of this research work.

Figure 1-1 illustrates the position of the spectrum monitor in a transceiver chip and Table 1-1 lists the comparison of the general different roles between the spectrum monitor and the main radio. Note that the 'narrow band' in the 'main radio' column could mean up to a few MHz, and 'wide band' in the 'spectrum monitor' column could mean hundreds of MHz or even GHz, depending on the application.

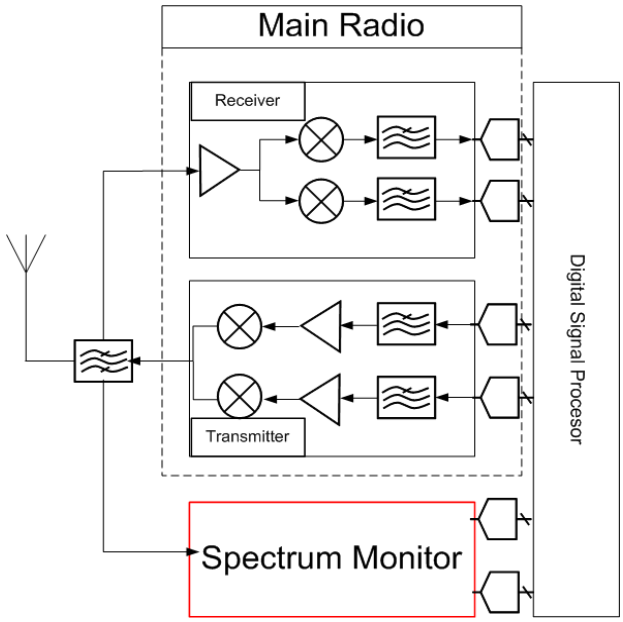


Figure 1-1 Spectrum monitor position in a cognitive radio handset

	Spectrum Monitor	Main Radio
Purpose	Detect the channel occupation	Demodulate the signals
RF Band	Wide band	Narrow band
Base Band	Multi-channels co-exist	Single channel

Table 1-1 Differences between spectrum monitor and main radio

1.2 Ideal Spectrum Monitor Requirement

An ideal spectrum monitor is able to scan the entire band of interest, e.g. from DC to 6GHz where most modern communication systems operate, with narrow enough frequency resolution and large enough dynamic range. Figure 1-2 shows an example of the real time spectrum map (0~6GHz) collected within 50μs at 20GS/s sampling rate in Berkeley downtown [3]. The x-axis is frequency with the unit of Hz. The unit of y-axis is not provided in the reference, but is believed to be in dBm. The actual values depend on the FFT bin bandwidth. From this figure, the spectrum resource usage situation can be roughly observed.

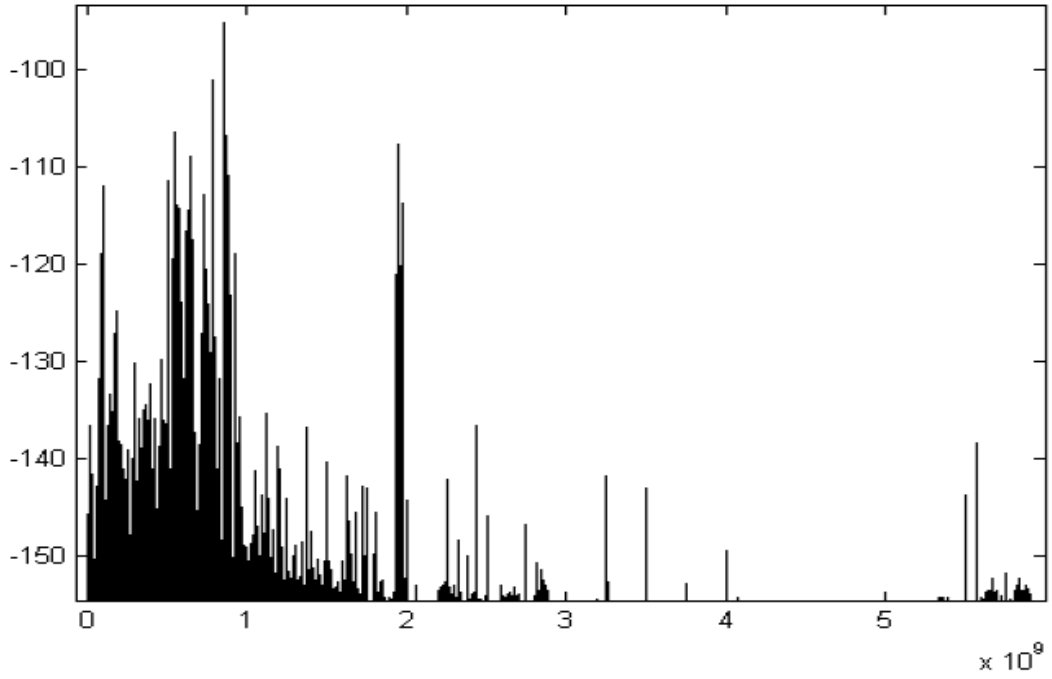


Figure 1-2 An example of the spectrum map [3]

The GSM system has nearly the narrowest bandwidth among most of the modern communication systems, hence if the spectrum monitor can recognize a single GSM channel, it should be able to detect the channel occupation for most other communication channels. Hence, a bandwidth of 200 kHz is initially chosen in this project.

A minimum sensitivity of -102dBm/200kHz and maximum acceptable signal strength of -15dBm/200kHz as in GSM system among most modern communication standards can also be considered as suitable specifications of the proposed spectrum monitor, which means a dynamic range of at least 90dB is required, with some reasonable margins.

While achieving the above specifications, the desired monitor needs to be fast enough in order to detect the variation of the spectral occupation in real time for the main receiver to adapt rapidly to changing channel occupancy. Furthermore, the power consumption must be acceptable for portable applications, i.e. tens of milli-watts for an RF front-end in a mobile handset, and the chip area also needs to be small.

In this PhD project, the requirements mentioned above are discussed and aimed to be addressed at system level and/or circuit level.

1.3 Document Structure

This PhD project describes the research into the architecture of a spectrum monitor. The following is the structure of the thesis.

Chapter two presents an overview of the research and the applications of software defined radio (SDR) and cognitive radio (CR). As discussed in the next chapter, the spectrum monitor for the cognitive radio is essentially a flexible wideband receiver, so some of the wideband transceiver architectures in modern communication systems are introduced and explored to provide some references for the spectrum monitor.

Chapter three discusses the Figures of Merits (FoM) of the active circuit blocks in an RF receiver chain, consisting of LNA, mixer, active low pass filter, baseband amplifier and ADC. The figures of merit of the VCO and divider are also included because of the non-negligible power consumption contributed to the whole RF front-end system. The principle of the relationships between performances and figure of merits are discussed. Then the improvements likely to be achieved in figures of merits of each function are predicted for the next few years. A system approach is developed based on the trends and relationships among FoMs, specifications and time scales, according to comprehensive statistical results. This approach can be used as a design guide for the spectrum monitor.

Chapter four describes the system level research for the spectrum monitor, showing that the dual-down conversion architecture is probably the most suitable option. The full system level analysis for this architecture is performed and the power consumption is then predicted using the strategy developed in chapter three.

Chapter five explains the design of the band pass filter. The conventional synthesis method is shown to be complicated and not suitable for on-chip designs. To address this, a method for the design of coupled resonator band pass filters design is introduced. Two filters are designed for different types of receiver architectures. Both types require some unusual design methods, including delta-star transformation and the introduction of additional transfer function zeros. An experimental chip is implemented for these filters.

Chapter six explains the design of the special PLL needed for the local oscillator in the spectrum monitor receiver, including system level design and circuit level design of the ring oscillator, the high frequency divider, the phase-frequency detector and the loop filter. The PLL block is integrated in the same testing chip as the filters.

Chapter seven summarizes the system level modelling (in chapter three, chapter four) and circuit level design (in chapter five, chapter six), showing the feasibility of a spectrum monitor for the cognitive radio function in mobile devices in the near future. Important comments on the research and potential future work are then discussed.

1.4 Declaration

This thesis describes the research undertaken by the author. All of the work has been done by the author alone, except assistance stated in the acknowledgements, such as testing, modelling, etc. The original contributions by the author include the (1) FoM derivations and predictions, (2) spectrum monitor receiver architecture analysis, (3) star-delta transformation technique in the integrated band pass filter, (4) the PLL tuning methods and (5) high speed PLL integer divider architecture, while all the other ideas from published works are given in cited references.

Chapter 2 Cognitive Radio and Wideband Receiver Review

2.1 Introduction

This chapter gives an overview of some of the important background topics for the research. First of all, the history of software defined radio (SDR) is reviewed, including the research and applications, and this is followed by the overview of cognitive radio, of which the motivation, evolution and key techniques are introduced. In the next section, wideband RF receiver architectures are explored and discussed as useful candidates or references for spectrum monitor architecture design.

2.2 Software Defined Radio and Cognitive Radio

2.2.1 Software Defined Radio History

The software defined radio architecture was firstly envisaged by Mitola in 1995 [4], as seen in Figure 2-1. The signal from DC to radio frequency is digitized by an ADC directly, and all the signal processing is done in the DSP. (DDC means digital down conversion).

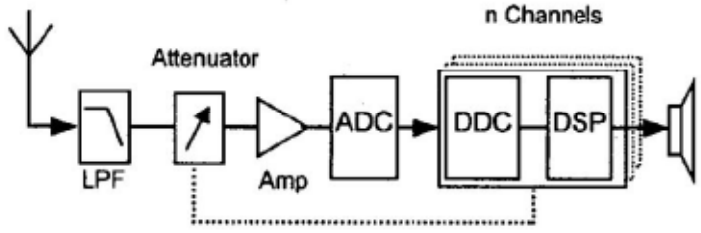


Figure 2-1 Ideal software defined radio architecture [5]

The only realization of this kind of SDR is the UK DERA which dealt with frequency from 3MHz to 30MHz in 2000 [6]. This Mitola type ideal software radio is limited by the ADC's technology when the frequency is increasing. After the antenna, RF anti-aliasing pre-filtering is adopted in Toshiba's SDR receiver [7] for PDC (Personal Digital Cellular) at 1.5GHz and DCS (Digital Cellular System) at 1.9GHz applications. The bandwidth of this receiver is 10MHz covering over 50 channels. With the principle of sub-sampling, a GPS (Global Positioning System)/GLONASS (GLObal Navigation Satellite System) receiver is published in [8], which also used a pair of RF pre-filters after the LNA to attenuate the wideband LNA noise. The sub-sampling combined with analog decimation technology was applied [9] for the purpose of minimising power consumption. To avoid the use of an anti-aliasing pre-filter, which limits the flexibility of the receiver, quadrature charge-domain sampling circuits was introduced [10] at an IF of 100MHz. In industry, this technology was exploited by Texas Instruments in its Bluetooth and GSM receivers. However, relying on the RF preselect filter makes it limited to narrowband applications. For example, the preselect filters have 100MHz bandwidth around 900MHz for GSM band and 83.5MHz bandwidth around 2400MHz for Bluetooth receiver, but not the whole commonly used wireless band, say several GHz.

The detailed review of the advantages and limitations of the above schemes is presented in [5], where Abidi also introduced an SDR architecture being able to tune to any channel from 800MHz to 6GHz (Figure 2-2). In this receiver, a zero-IF architecture ensures the high flexibility and low image rejection requirement. A second-order RC filter is driven by the mixer to eliminate the RF preselect filter, which is to achieve the full anti-aliasing function. The sampler is placed immediately after the RC filter and leaves the rest of the filter in the discrete-time domain.

The above architectures aim to detect signals within different channel bandwidths, which is the usual function of software defined radio. Nowadays, the most popular SDR technology can be found in IEEE 802.11a/b/g and Bluetooth. However, the detected bandwidths are typically 20MHz with the carrier frequency at hundreds of MHz or several GHz. Hence, the advantage of software-defined radio is in fact the flexibility but not the wide bandwidths, as these are essentially narrow band receivers.

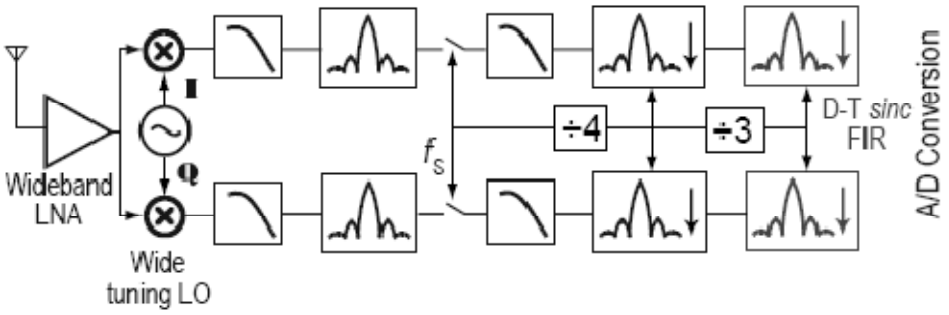


Figure 2-2 Abidi's SDR receiver architecture [5]

2.2.2 Evolution to Cognitive Radio

Motivation

The cognitive radio is a much wider concept than software defined radio. The motivation of this important concept is the scarcity of frequency resources with the increasing applications of wireless communication nowadays, while in the meantime, the licensed spectrum is wasted seriously. Usually, unlicensed bands are often very crowded, e.g. 2.4GHz ISM band, whereas some licensed bands, e.g. TV band, are often left unused. A report measured that the frequency usage efficiency is less than 5.2% below 3GHz on average [11]. This leads to the idea of how to use the frequency much more effectively to solve the conflict. One of the main functions of cognitive radio will focus on a radio detecting these unused bands and using them as long as the primary users are not affected.

Evolution

Although the cognitive radio is a relatively new concept, the essential idea has been applied in a few communication systems. A sort of automatic channel selection scheme is applied in the cordless phone working at 45MHz to avoid using the occupied channels. The unlicensed PCS (Personal Communication Service, provided in United States and Canada) devices listen to the spectral occupation before transmission. The DFS (Dynamic Frequency Selection) and TPC (Transmit Power Control) technology are adopted in the IEEE 802.11a network to avoid interference with radar signals. Besides, the cognitive modulations are also used in the HSDPA and CDMA1x EvDO transmission, by configuring the optimum modulation scheme, data rate and transmit power according to the environment and the users' demands.

There are several definitions of the cognitive ratio. One is from the Royal Institute of Technology (KTH) in Sweden as presented by Mitola [4], suggesting the

SDR as a proposed plant for cognitive radio (CR) based on the RKRL (Radio Knowledge Representation Language) realized at the application level. Another one is supported by Virginia Institute, presented by Riese [12], pointing that the SDR is not the necessary plant of CR, instead, the modelling at the MAC (Media Access Control) level of communications. Nowadays, a more widely acceptable and simplified definition of CR is given by FCC (Federal Communications Commission), suggesting that any radio with the function of adaptive spectral cognition can be considered as cognitive radio [2]. The legal licensed users who are called primary users have higher priority for certain spectrum bands while the unlicensed users with CR function are allowed to access to the spectrum as long as they don't interfere with the primary users.

Key Techniques

There are several key techniques required in cognitive radio. The first one is spectrum monitoring. The spectrum monitor needs to be able to detect independently the unoccupied band and the emergence of primary users. This requires successive listening and some acceptable accuracy (depends on the actual environment and system configurations) to avoid or minimize mistakes.

For the unoccupied channel detection, the challenges exist both in RF front-end design and the digital signal processing stage. In RF front-end design, widely separated signals with different power means that the detection of weak signals in the presence of strong signals is a frequent requirement. Besides, the dynamic range might need to be controlled to keep the input signal of the ADC (Analog to Digital Converter) at a reasonable level with respect to its Full-Scale specification, by means of adaptive tunable notch filters. As for the DSP stage, Cabric summarized spectrum sensing techniques [3] in the signal processing stage, including matched filtering, energy detector and cyclostationary feature detection methods. However, the reduction in the signal strength caused by multipath and fading may limit detection ability and accuracy significantly [13]. Hence the cooperative spectrum sensing technique is studied [3] to improve the sensing detection and recognize modulations, numbers and types of the signals.

Besides, there are also some other techniques such as the detection of the position of a primary user, which was studied by Wild [14], via exploiting the Local Oscillator leakage power emitted by super-heterodyne receivers.

2.3 Wideband Receiver Architectures

As mentioned before, a spectrum monitor is essentially a wideband receiver. Therefore, some existing wideband receivers are worth being investigated to provide some references for the design of the potential spectrum monitor. There are mainly two types of communication systems need that need to deal with wideband signals. One is the UHF TV band from 400MHz to 800MHz. This band is also considered as a good candidate for the realization of the cognitive radio because there are always unoccupied channels. The other one is the UWB band from 3.1GHz to 10.6GHz. This is a relatively new standard, but there are already many realizations for this high frequency wideband system.

2.3.1 TV Tuner

A wide frequency range spanning from 48 MHz to 860 MHz is covered by various analog and digital TV standards all around the world. The design of the wideband TV tuner involves several key technologies to deal with the problem of harmonic mixing, image, linearity and dynamic range (Figure 2-3). Problems exist for a zero IF tuner in the lower bands when the 3rd order of the local oscillator falls in the wanted band, resulting in the higher unwanted channel being mixed to the baseband together with the wanted channel. The image problem can be solved by using a tunable RF band pass filter. It is easier to achieve this compared with in the narrowband case, where the Q factor needs to be very high. The linearity is also an important issue because the in-band interference appears due to the wideband amplifications.

There are three main receiver architectures used in these applications: (1) conventional super-heterodyne architecture, (2) up/down dual-conversion architecture and (3) low-IF fully integrated tuner with poly-phase filters.

The conventional super-heterodyne (Figure 2-4) architecture is the simplest approach for a TV tuner. The tunable band pass filter in the RF stage and the SAW filter in IF stage filter out the unwanted channels, minimizing the interference and achieving a good performance of linearity. Then the commonly used 36/44MHz IF channels will be demodulated in digital domain. The problems of this architecture are the difficulty of the integration of a tunable high Q band pass filter, the more cost and less integration due to the extra SAW filter needed and the image and harmonic interferences.

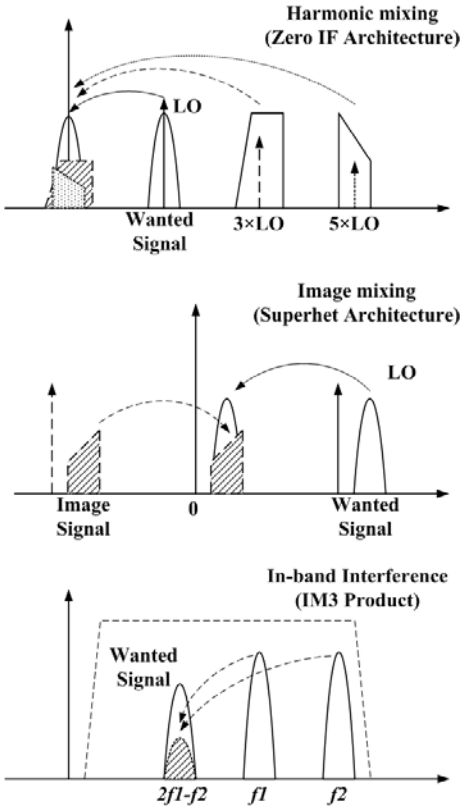


Figure 2-3 Problems encountered in TV tuner design

The up/down dual conversion architecture (Figure 2-5) eliminates the requirement of the tunable RF band pass filter. Instead, it solves the harmonic problem by means of converting the wanted bands to a higher frequency. An external SAW is still needed at the higher IF, which is followed by the down mixer to move the signal to a standard IF of 36/44MHz for the demodulation. With the given Q factor of the band pass filter, the image rejection ability of the SAW filter at the higher first-IF is limited (about 30~40dB) compared with the situation in a conventional super-heterodyne architecture. Due to the SAW filter's limited image rejection capability, the second mixer is usually an image rejection mixer (IRM). There are different ways to implement an image rejection mixer. Usually, it consists of two mixers mixing the quadrature input RF signals with a single local oscillator, or mixing the input RF signal with quadrature local oscillators (Figure 2-6), both are followed by a Hilbert filter or a polyphase filter, which respond to the complex representation of the input signal instead of the magnitude only. Hence they can recognise the negative frequency (for example) and remove it. These approaches are also referred as single quadrature mixing. It achieves the image rejection by nulling the unwanted image frequencies and passing all the other frequencies. The image rejection level, which is usually about 40dB, is mainly limited by the amplitude and phase mismatch of RF and LO inputs to

the mixer and the gain mismatch of the mixer itself. However, the cost of the SAW filter reduces overall integration level, and the power consumption due to the circuits at higher frequencies still forms the bottlenecks of this system.

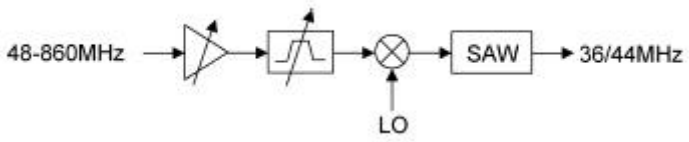


Figure 2-4 Conventional architecture [15]

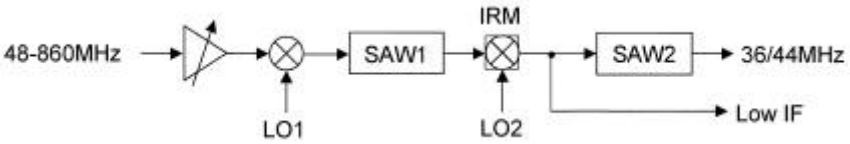


Figure 2-5 Up/Down dual conversion architecture [15]

The fully integrated tuner (Figure 2-7) replaces the tunable RF bandpass filter with selectable on-chip RF bandpass filters. The RF polyphase filter generates differential quadrature phases from a differential RF input. This signal is then fed into a double quadrature mixer (DQM) [15] consisting of four mixers to convert the signal down to a low-IF frequency. A following IF polyphase filter selects the channel for demodulation. This low-IF architecture avoids the drawbacks of the zero-IF architecture, such as incompatibility with existing channel decoders, the matched ADC pair, and the most important is the DC offset which is difficult to remove without the loss of useful signal information around DC. However the image problem worsens compared with the zero-IF architecture, due to some unrelated channels which might be larger than wanted channels. Hence the typical image attenuation of over 50 dB is required compared with that of about 15 dB for zero-IF.

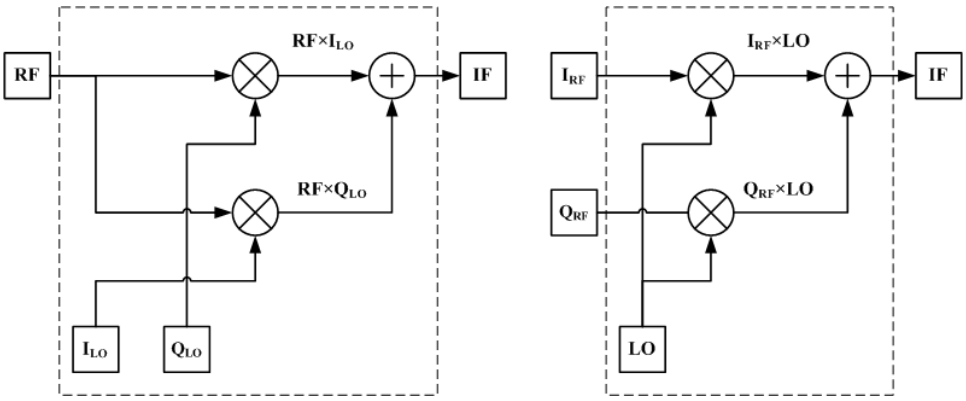


Figure 2-6 Single quadrature mixing architectures

The low-IF architecture is more popular in recent technology, because the DQM technique can improve the image rejection level significantly. The DQM scheme is essentially the combination of two single quadrature mixing circuits (Figure 2-8). The differential quadrature RF inputs and LOs mix through the four real mixers, as shown in Figure 2-8. The negative frequency is removed in the output differential quadrature signals. It is much less susceptible to the inputs (LO and RF) gain and phase mismatches and the image attenuation is ultimately limited by the mismatch of the mixers and IF polyphase filter, and it is shown that an image rejection of over 50dB can be obtained from the DQM structure [16]. Furthermore, a complex one-tap LMS (Least Mean Square) decorrelation algorithm [17] could be adopted in the digital domain to improve the cancellation of the image.

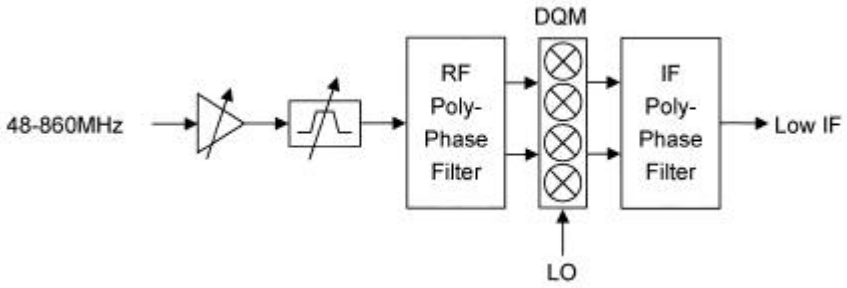


Figure 2-7 Low-IF scheme using double quadrature mixers [15]

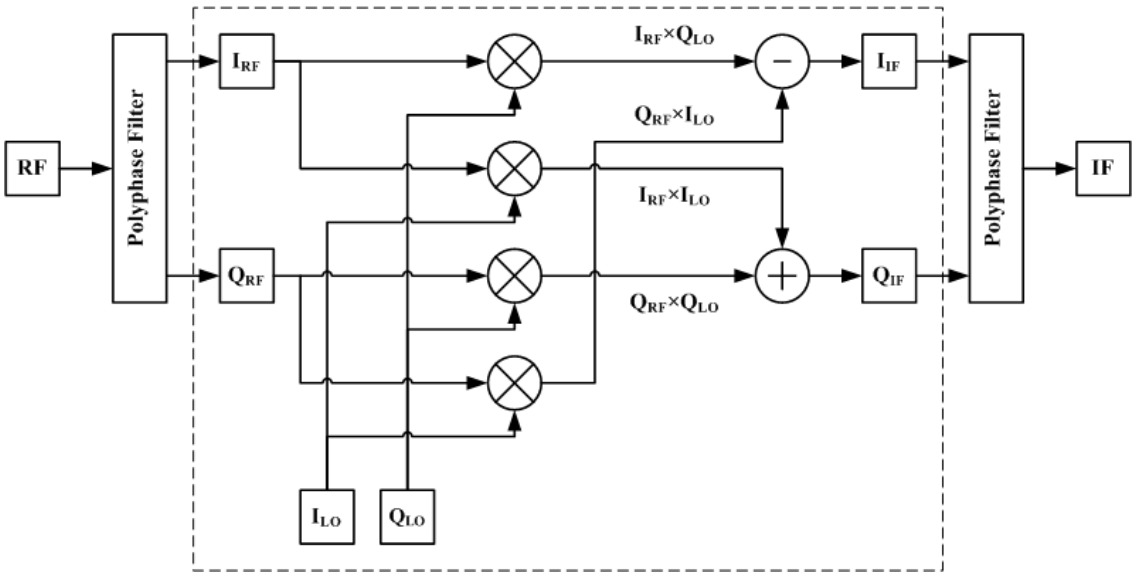


Figure 2-8 Double quadrature mixing architecture

2.3.2 UWB Receiver

Any wireless transmission scheme occupying a fractional bandwidth (BW/f_c, the ratio of transmission bandwidth over centre frequency) of over 20% or an absolute

bandwidth of more than 500MHz can be considered as an Ultra Wide Band (UWB) technology [18]. The FCC allocation for the UWB frequency is 7.5GHz unlicensed band from 3.1GHz to 10.6GHz [18]. A direct method is to develop an impulse-radio technique so that the signal bandwidth could cover the entire 7.5GHz frequency range.

For the impulse-radio UWB scheme, the receiver architecture (Figure 2-9) is similar to the ideal software defined radio receiver. The bottleneck is the high speed ADC, which is required to be at least the signal Nyquist rate of 15GHz, as well as a reasonable dynamic range. Basically, there are two problems of the ADC design. The first one is whether this high performance ADC is achievable with present day technology. Second is the high power consumption even if the ADCs can be designed and fabricated. Heydari [19] discussed briefly that, to implement a 4-bit 15GHz full flash ADC, one single comparator needs a preamplifier with the unity gain-bandwidth of roughly 330GHz, which is very difficult for present day CMOS technology. Also, such an ADC could consume hundreds of milliwatts of power. Alternatively, some other approaches has been involved in terms of the ADC design, such as a time-interleaved architecture [19]. A number of parallel ADCs are needed in this architecture. Each ADC performs at a sampling period of integer that is an integer multiple of the original sampling period, and these ADCs are clocked by equally delayed clock signals. The sum of the converted signals is equivalent to the digitized signal with the original sampling rate. In this way, the design of the ADC is feasible for present day technology. However, the power consumption is almost the same as the flash architecture given the same bandwidth and resolution.

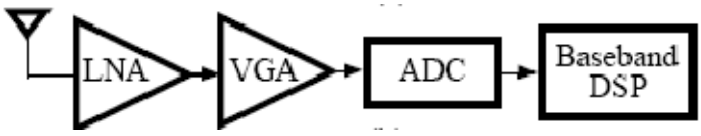


Figure 2-9 Impulse-radio UWB architecture [19]

Consequently, multiband UWB transceiver architecture is more attractive because of the lower difficulty in practical realizations. The Multiband OFDM (MB-OFDM) Alliance (MBOA) formed in 2003 was started in order to support the UWB specifications based on OFDM. The whole band from 3.1GHz to 10.6GHz is divided into several sub-bands, each of 528MHz, and each set of three sub-bands is called a band group. Figure 2-10 illustrates the band plan. The multiband OFDM system results in a satisfying trade-off between different design criteria and a low-power multi-band UWB transceiver. This scheme relaxes the impractical requirements in terms of the

ADC to a more achievable sampling rate of 1.1GHz, which is twice the bandwidth of 528MHz to satisfy the Nyquist criteria. In spite of this, the requirements of the gain, noise and linearity are also challenging.

A bank of LNAs and mixers could be used to cover the whole band of 7.5GHz but suffers from the high frequency switches and the power consumption. Therefore, one wideband LNA/mixer front-end is usually designed and a high quality wide tuning range frequency synthesizer is needed to accomplish the frequency generation.

The following two UWB receiver examples represent two popular architectures: direct conversion and dual conversion. These two receivers are integrated in a single chip and use one wideband LNA/mixer front-end, while generating the LO frequencies in different ways, which also represents distinct frequency generation plans.

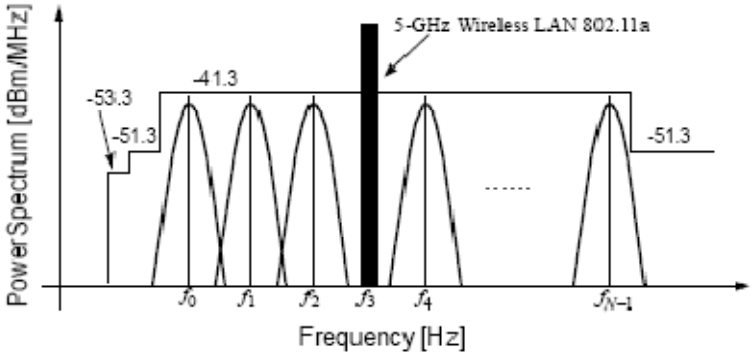


Figure 2-10 Multiband OFDM UWB band plan [19]

A Zero-IF direct-conversion UWB is reported [20] from NXP semiconductors as shown in Figure 2-11. This transceiver is designed for band group one (3168MHz~4752MHz) and band group three (6336 MHz~7920MHz) using MB-OFDM. The main interferers are from the 2.4GHz and 5GHz ISM bands, which are normally used in WLAN and Bluetooth. In this design, however, they fall out of the band of interest. A wideband LNA covering about 5GHz is needed in the first stage. An integrated transformer is used to achieve the passive phase splitting to transform the single-ended RF input into a differential signal. This avoids the external wideband balun, which usually causes loss and incurs more cost. After the down conversion, the 264MHz (300MHz measured) baseband filter is implemented by the low noise transimpedance amplifier (TIA), which consists of an operational amplifier and a bridged-T RF feedback network. The local oscillator is a set of three RF-ring-oscillators with four differential amplifiers in cascade for each of them.

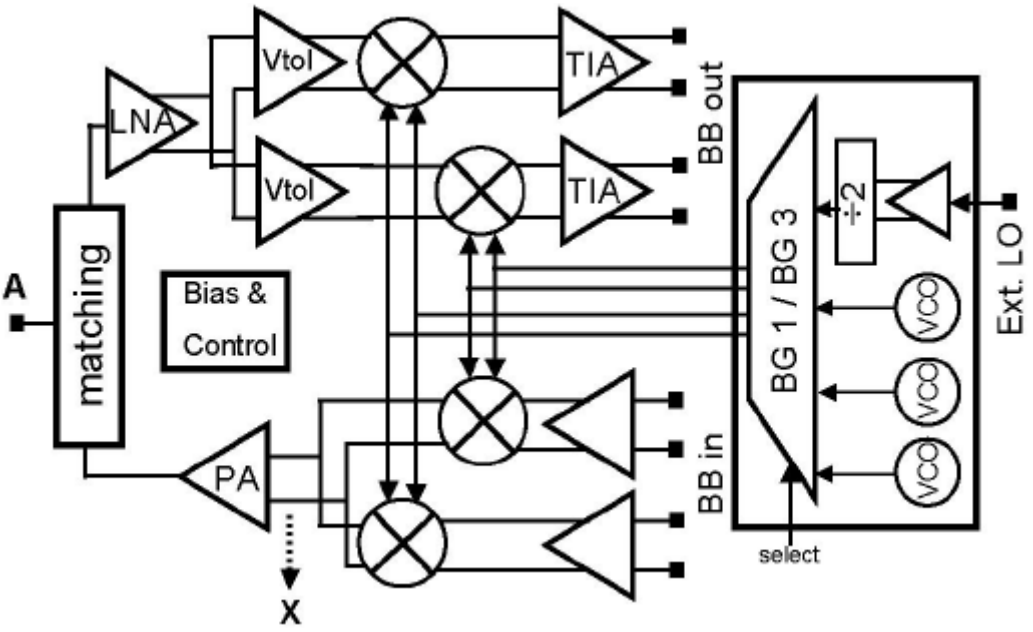


Figure 2-11 UWB transceiver (Bergervoet, ISSCC 2007) [20]

Another dual conversion UWB receiver [21] first converts an 9 bands to a fixed IF frequency using a first LO, which uses a single LC oscillator and generates multiple frequencies via multiple frequency dividers, wideband SSB mixers and multiplexers. The selected frequency bands are then converted to the baseband by the second LO, which is also generated from the same LC oscillator. Figure 2-12 and Figure 2-13 show the receiver architecture and frequency plan.

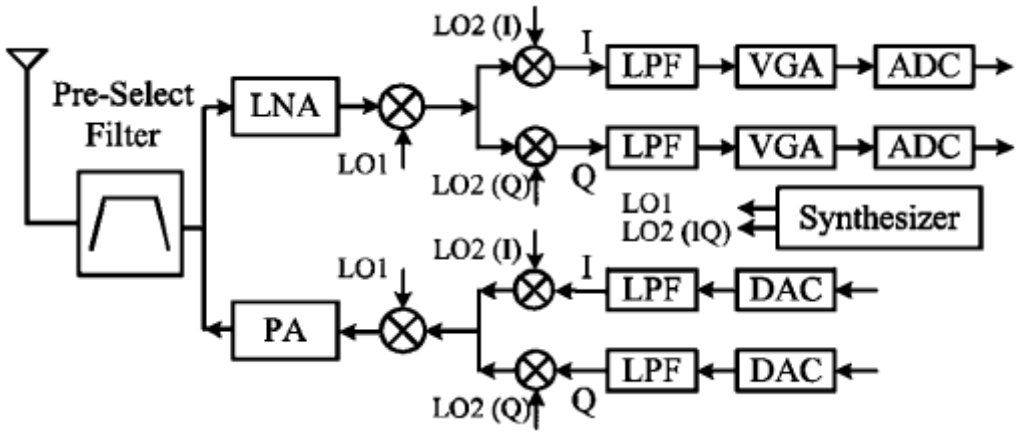


Figure 2-12 UWB transceiver (Hui, JSSC 2009) [21]

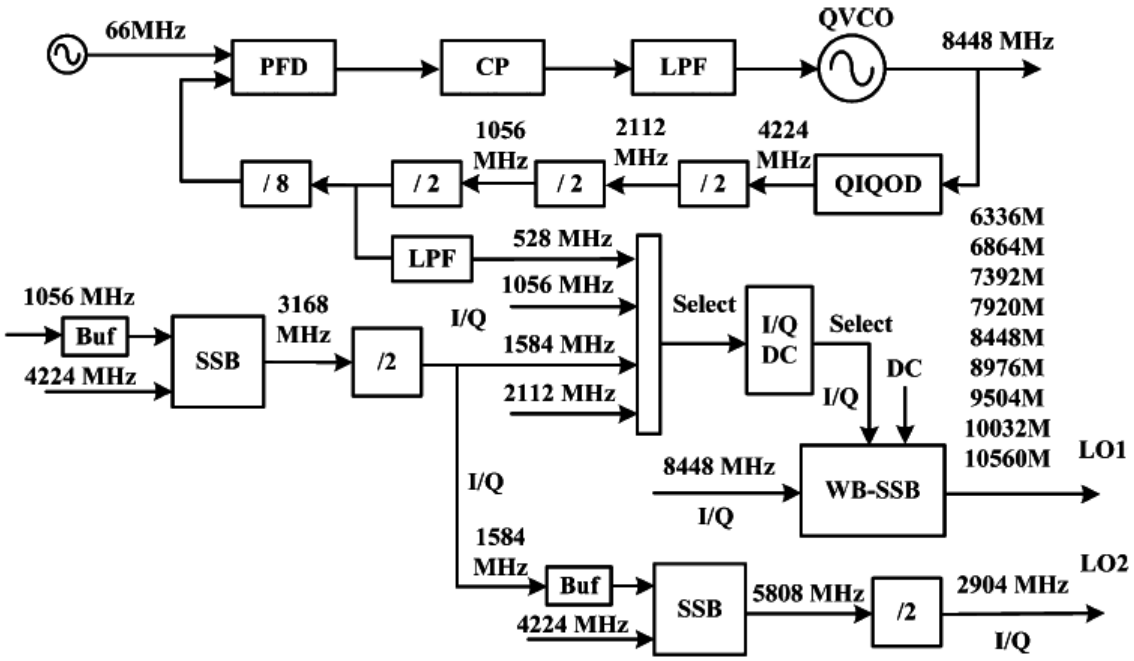


Figure 2-13 UWB transceiver frequency plan (Hui, JSSC 2009) [21]

2.3.3 Receiver Architectures Comparison

Because of the limited performance and high power consumption, the ideal structure where the ADCs are placed directly after the antenna and LNA is not feasible in the near future in CMOS technology. Therefore, TV tuners and UWB receivers adopt alternative frequency plans.

Although all the three TV tuner architectures mentioned above can achieve satisfactory performance, complicated high performance passive components in the first two architectures are very difficult to integrate. Traditionally, TV tuners are not designed for portable applications, so there are not strong demands in terms of size and low power. Therefore, these structures cannot be transformed entirely to the proposed spectrum monitors. Nevertheless, the second architecture (up/down conversion) could be useful, because it has only one major off-chip component, namely the high IF filter. If this filter can be replaced by a moderate complexity integrated filter, and if the performances of the active circuits are also improved, there might be a compromise solution to meet the spectrum monitor specifications. The third option removes the necessity of complicated on-chip passive components, at the cost of introducing multiple mixers, which means increasing the power consumption.

Author	Chun-Huat Heng [15] 2005	J. R. Bergervoet [20] 2007	Hui Zheng [21] 2009
Application	TV Tuner	UWB	UWB
Architecture	Low-IF	Direct Conversion	Dual Conversion
Frequency Band	48MHz ~ 860 MHz	3168~4752MHz 6336~7920MHz	3168~7920MHz
Receiver chain			
Voltage Gain	63 dB	24 dB	25~84 dB
NF (dB)	14 dB	5~5.5 dB	4.5~5.8 dB
S_{11} (dB)	N/A	-7 dB	-13 dB
In-Band IIP3	-5 dBm	N/A	-13 dBm
OutBand IIP3	N/A	+5dBm	-3.5 dBm
Frequency Synthesizer			
Architecture	3×LC VCOs (PLL1) 1×LC VCO (PLL2) Multiplexer	3×Ring oscillators Multiplexer	1×LC oscillator 2×WB-SSB mixers Multiplexer
Reference frequency	5.4 MHz (PLL1) 27 MHz (PLL2)	N/A	66MHz
Spur	-102dBc @ 5.4MHz (PLL1)	N/A	-42dBc @ 10MHz
Phase Noise	-100dBc @ 300kHz (PLL1)	-88dBc/Hz @ 1MHz	-126dBc/Hz @ 10MHz
Power Consumption	125 mW (PLL1 + PLL2)	62.4 mW (w/o PLL)	102.6 mW
Chip			
Technology	0.25 μm CMOS	65nm CMOS	0.18μm CMOS
Supply	2.5V	1.2V	1.8V
Power	763 mW (RX + PLLs)	114 mW (RX + VCOs)	285 mW (RX + PLLs)
Area	36 mm^2	0.4 mm^2	15.6 mm^2

Table 2-1 TV tuner and UWB receivers comparison

Similar to the TV tuner, the multiband OFDM UWB transceivers also deal with wideband signals at the front-end stages. The difference is that, due to the frequency

bands used, the UWB receiver usually only needs to deal with 5.2 GHz 5.8 GHz interferers in the foreseeable future, while the strong interferers existing in the whole band must be considered in TV tuners. This explains the reason for the up-conversion architecture, the extensive usage of off-chip passive components to attenuate in-band interferers (It can be seen that the IIP3 of the UWB receiver is lower than that of the TV tuner), and multiple quadrature mixers and polyphase filters to suppress images for TV while, on the other hand, the UWB receivers' architecture can choose a direct-conversion or dual-conversion architecture just like other narrow band applications, and is also easy to integrate. The comparisons of different wideband receiver examples are listed in Table 2-1. In general, TV tuners need to meet more stringent specifications of gain, linearity, phase noise and spurs, and hence have greater power consumption and larger areas. These requirements on UWB receivers are relaxed to some extent, but wide range high frequency PLLs are one of the most difficult blocks to design to meet the specification and reduce area and power consumption. A ring oscillator-based PLL has the advantages of low power and low cost, with the worse phase noise. A combination of an LC oscillator and SSB mixers achieves excellent phase noise but suffers from high power, high cost and higher spurs.

2.4 Summary

In this chapter, the concepts, history and evolution of software defined radio/ cognitive radio are reviewed, and the key techniques of the realization of cognitive radio are discussed. As explained, a high performance, low power, low cost spectrum sensing technique is the crucial first step. This spectrum monitor is essentially a wideband receiver.

Therefore some existing wideband receiver architectures are analyzed, including TV tuner and UWB receiver. Comparisons of complexity, performances and power consumptions are made among different receiver examples. The further design of the spectrum monitor can be based on the reviewed architecture and techniques, while making necessary modifications according to its own specifications.

Chapter 3 Figures of Merits

3.1 Introduction

A Figure of Merit (FoM) is useful as a method for comparison, typically reflecting the relationship between performances and power consumption of a component. People usually use a figure of merit to plan the products design for the next several years. A well-known example of this strategy can be seen in the digital IC world, where 'Moore's Law' is commonly used. Moore's Law predicts that the scaling trends of transistors yield a doubling of the number of gates per unit area every 18~24 month as shown in Figure 3-1. Hence, when planning the architecture and design of a large digital application, e.g. microprocessor and DSP, it is sensible to consider what the most appropriate technology is to use so that the cost and performance trade-off is optimum at the time the product comes to production, rather than at the start of the conceptual design.

In this chapter, this strategy is used in RF and analogue design, which typically involves more complicated functions and analysis. The circuits include RF blocks such as a narrowband/wideband low noise amplifier (LNA) and mixer, baseband blocks such as active low pass filter (LPF) and variable gain amplifier (VGA), frequency synthesising blocks such as voltage controlled oscillator (VCO) and frequency divider, and Nyquist and Sigma-Delta ADCs as the interface between the analogue and digital world. The performance of a component usually includes gain, noise, linearity, speed (bandwidth and frequency) for general RF and analogue circuits, and, additionally, phase noise for the oscillator, and digitizing resolution for the ADC. These specifications and power consumption are always related to device parameters to some extent. Generally, with a fixed technology, more power dissipation is needed to achieve higher performances, such as high operating frequency, wide bandwidth, high linearity,

high gain in some cases (e.g., Mixer with resistor load), low noise and high resolution. With the development of the technology, e.g. shrinking of the size of the transistor, the device parameters vary through the years, such as maximum oscillation frequency, the input flicker noise and MOSFET internal gain, as well as density of capacitor, quality of inductor and temperature linearity of resistors, etc. As will be discussed, these physical parameters actually improve the achievable performance with certain power consumption, or in other words, improve the figure of merit.

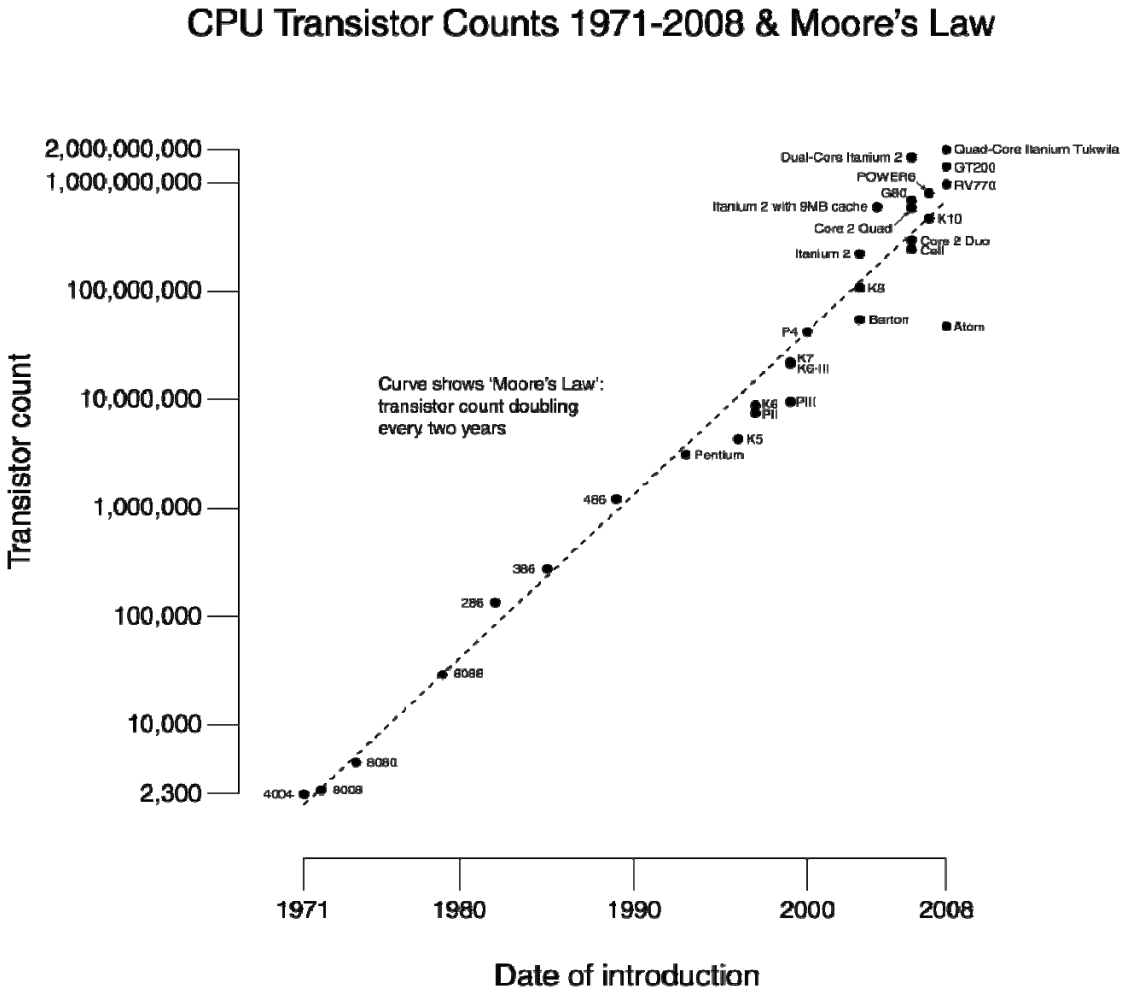


Figure 3-1 Moore's Law in the digital world [22]

Therefore, the investigation of FoM is meaningful at the starting stage of an RF transceiver design because it provides a general guide, sometimes called a roadmap, to determine the trade-offs between performance and power consumption of transceiver design in the future. In this chapter, this strategy is to be introduced as follows.

First, the general analysis approaches are explained, including the theoretical and practical methods. Then, the various FoMs are defined and analysed for the low noise amplifier (LNA), mixer, low pass filter (LPF), variable gain amplifier (VGA), voltage controlled oscillator (VCO), frequency divider and analogue-to-digital converter

(ADC). It is important to realize that the definitions of the figures of merit for these functions are not unique. Instead, there may be several useful figures of merit for a given function. For some of the functions, there has been a consensus among researchers over a particular FoM definition, while for others there is less clear agreement over the 'best' FoM definition. In this project, different FoM definitions are to be mentioned and the most popular ones and/or the most suitable ones for cascaded receiver system level design are chosen. When collecting the FoM values among published works, they are either provided directly, or could be calculated from the published performance figures. Although RF BiCMOS technology generally achieves better performance than CMOS, all analysis in this chapter is based on CMOS technology because this is the prime technological driving force in the consumer wireless communication market, especially for portable devices.

3.2 Theoretical (ITRS) and Practical Analysis Approaches

Given any circuit architecture, circuit performances are directly related to the device parameters. Therefore, theoretically, the FoM can be calculated according to the provided device parameters. So the FoM improvement through years can be predicted by changing these parameters, which are published by ITRS annually.

The International Technology Roadmap for Semiconductors (ITRS) is a group of documents published annually by expert representatives from the semiconductor industry, who aim to give a technology assessment through the years. The most relevant part of the ITRS for this PhD project is the section entitled 'RF and Analog/Mixed-signal Technologies for Wireless Communications'. In these articles, physical trends of active and passive device parameters are produced. For example, for high speed RF/analog transistors, the supply voltage, gate length, internal gain, flicker noise, matching variance, current density, peak transition frequency and minimum noise figure parameters are included. And the on-chip passive devices parameters include the inductor's Q factor, MOS varactor' tuning range, resistors parasitic and temperature linearity, etc. These physical improvement trends set some fundamental limits for the circuits built from corresponding devices. The actual achievable performance is related to these physical parameters either directly or indirectly. Figure 3-2 and Figure 3-3 illustrate some of the CMOS device and on-chip passive components technology trends from 2003 to 2014, according to the ITRS relevant articles [23-29].

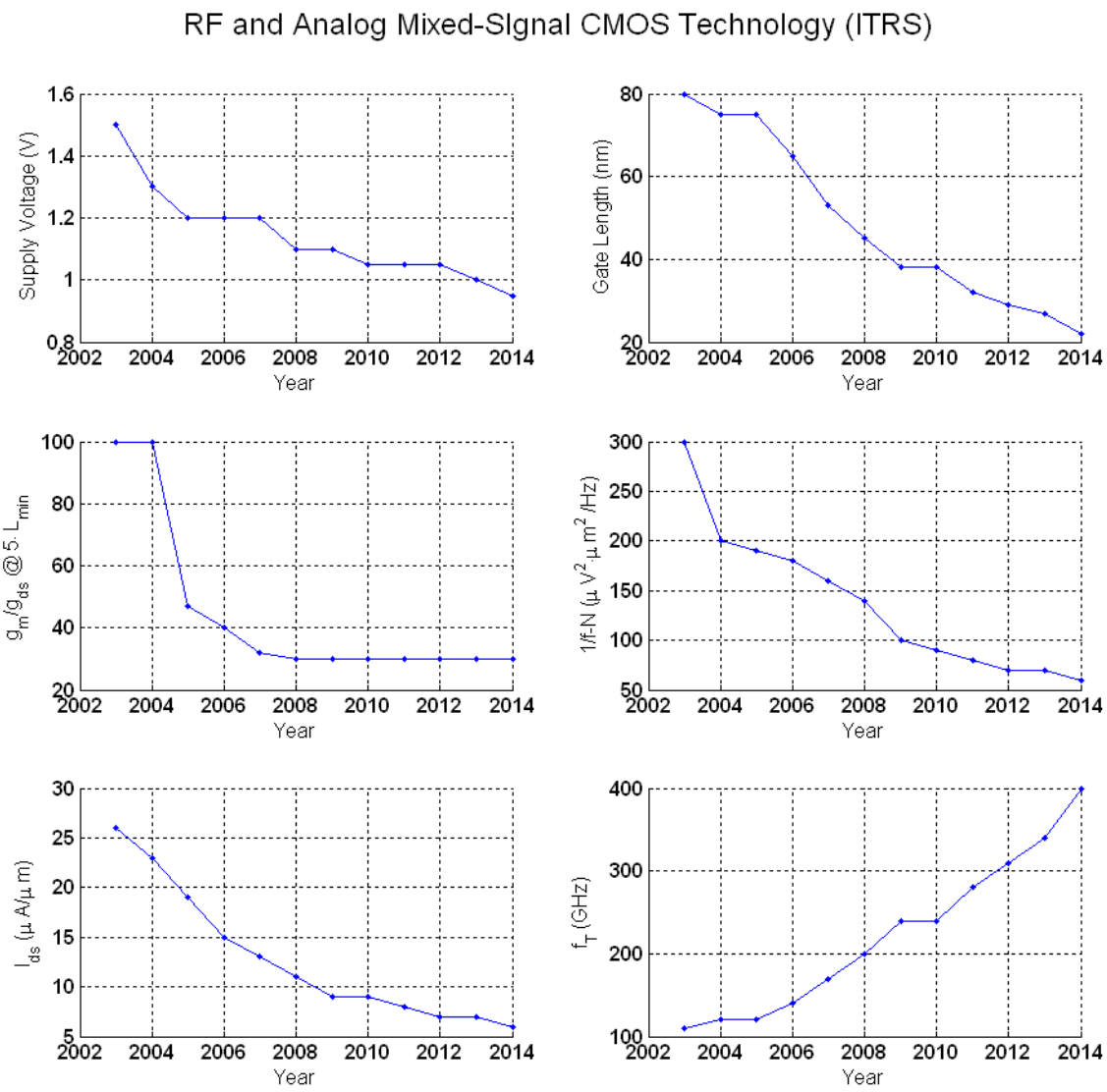


Figure 3-2 RF/Analog Mixed-Signal CMOS Technology (ITRS)

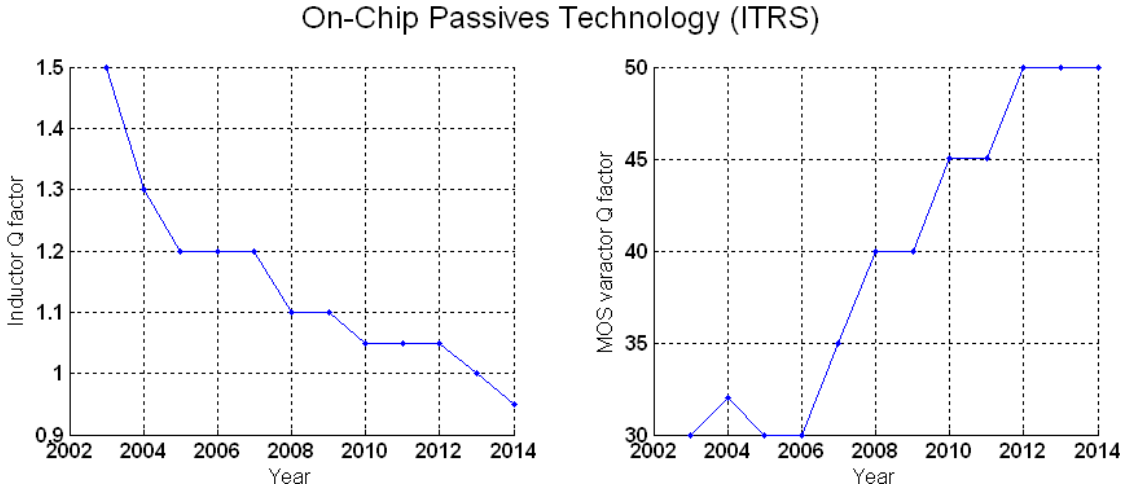


Figure 3-3 On-chip Passive Technology (ITRS)

From the engineering point of view, however, the theoretical approach is not very convenient to use, not only because of the difficulty of FoM equation derivations, but also due to the inaccuracy caused by simplifications and assumptions during

calculations. Hence, an alternative way of prediction by obtaining the trendline of a FoM after collecting published circuit measurement data within past decades is chosen in this project. The measured data are mainly from published results in well recognized journals and international conferences. By analysing the relationships among collected performances and FoM through past years, the power consumption can be predicted given certain specifications in a certain year. Note that this is a fast estimation method, instead of an accurate calculation, for system level design at the beginning of the transceiver development. For each block, four diagrams could be obtained, as illustrated in Figure 3-4 ~ Figure 3-7. All the specifications and FoM values are plotted on a log-scale, so that the trendline and relationships are linear. The reason is that the improvements in physical device parameters generally have an exponential relationship with respect to the time scale. The subscripts i and j are the indices of different specifications.

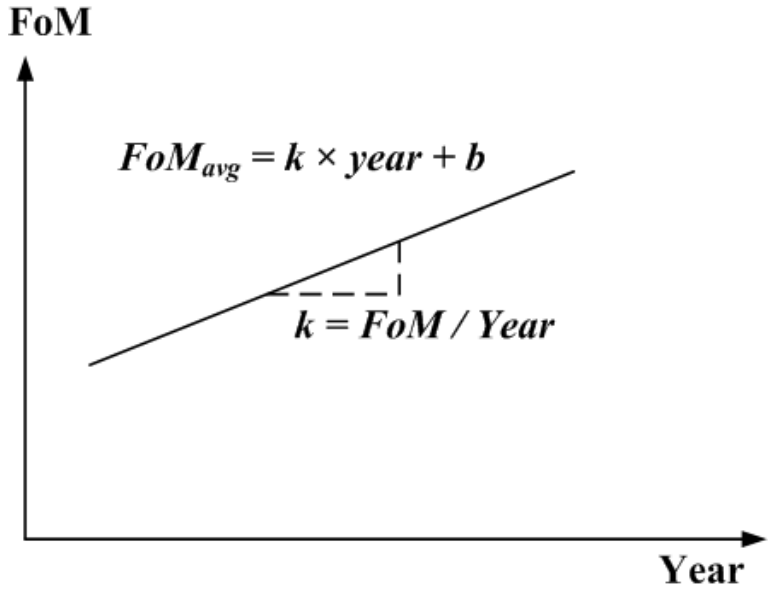


Figure 3-4 Form of a typical average FoM improvement through years

Trendline: FoM versus years (Figure 3-4)

The fundamental diagram of a FoM strategy is the average FoM versus year, which will be derived by a linear fitting technique according to the collected data points, and shows the improvement with technology development. The linear fitting equations are provided in Appendix A-1.

$$FoM_{avg} = k \times year + b. \quad (3\text{-}1)$$

The variable *year* is typically from 1995 to 2015. The slope k is the FoM improvement speed, which will be calculated and summarized in the unit of

months/3dB, for the convenience when comparing with the digital CMOS transistor density (18~24 months/3dB).

As can be seen later in the chapter, the FoM of one type of circuit from published results could have a very large variation, even with the same circuit architecture, using the same technology and in the same year. The FoM/year trendline therefore only reflects average value of FoM.

To reflect the FoM improvement as accurately as possible, an adequate number of data points from publications is needed to prevent too much influence of individual data point on the slope (k) and intercept (b) of the fitted line. In this chapter, about 100~150 data points are collected for most blocks such as the LNA, mixer, VCO and ADC. This number of samples can reduce the influence of any individual data point on the fitted line and hence usually provide acceptable accuracy. However, due to the limited number of published works, fewer data points have been collected for frequency dividers, baseband low pass filters (LPF) and variable gain amplifiers (VGA). To guarantee the greatest possible accuracy under all conditions, automatic selections are done by MATLAB programs, to eliminate data points that fail to meet certain criteria. These criteria include the influence of individual data point on slope and interception of the trendline, for example:

- Influence on FoM trendline slope (k) is less than 10%
- Influence on FoM trendline intercept (b) is less than 10%

The diagram of FoM versus year alone is useful when a block has already been designed to achieve certain specifications and consume certain power, and the FoM, hence power consumption, can be predicted for the next few years, with the same circuit architecture and specifications. Note that the important factor is the FoM improvement slope k. However, when none of the specifications or power consumption is known yet, the FoM relationship with specifications, specification variation versus years and the relationships between specifications are worth investigating in order to predict the power consumption accordingly.

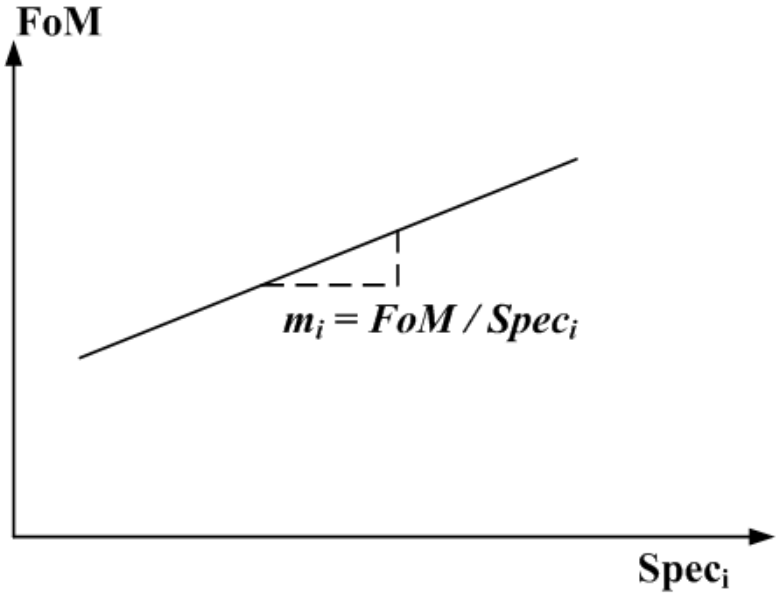


Figure 3-5 Relationship between FoM and specifications

Relationship: FoM and individual specification (Figure 3-5)

Although the FoM is defined as performance normalized by power consumption, it doesn't necessarily mean that power consumption is in proportion to each specification. Consequently, it is unavoidable to investigate how much correlation exists between each specification and the FoM, and hence power consumption. Instead of pure theoretical analysis, a more useful and practical way is to explore the slope m_i in Figure 3-5. It could be any value between -1dB/dB and +1dB/dB, for parameters such as gain, noise, linearity, dynamic range, SNR, and could be between -10dB/dec and +10dB/dec for frequency and bandwidth, which are in magnitude. If the slope m_i equals zero, this means that the FoM doesn't change with specification variation. For example, if 6dB higher voltage gain needs a doubled power consumption, and 6dB lower voltage gain saves half the power consumption, then the FoM value remains unchanged and FoM vs. Gain slope equals to zero. This means that the gain and power consumption are fully correlated. On the other hand, if doubling or halving the absolute value of one specification causes no change in power consumption, the FoM will also increase or decrease by the same amount, which means that the FoM is uncorrelated with this specification and the FoM vs. specification slope m_i could be ±1 (or ±10 for frequency, BW etc.). In reality, any specification depends on all the active and passive devices within the circuit, and cannot be fully correlated or uncorrelated with power consumption and FoM.

As mentioned above, the average FoM could be obtained with very large variations. One important reason for this variation is the different specification

combinations, which can influence the actual FoM in different ways. The average FoM is usually corresponding to average specifications. As a result, an offset from the average FoM value, denoted ΔFoM, should be taken into account when actual specifications are chosen as opposed to average specification values:

$$\Delta FoM_i = m_i \times \left(Spec_i - Spec_{i,avg}\right). \tag{3-2}$$

For the first order FoM prediction, assuming that FoM offset from different specifications are independent, the total offsets is the sum of each specifications' on top of average FoM:

$$\begin{aligned} FoM &= FoM_{avg} + \sum_i \Delta FoM_i \\ &= (k \times year + b) + \sum_i \left[m_i \times \left(Spec_i - Spec_{i,avg}\right)\right]. \end{aligned} \tag{3-3}$$

This equation doesn't take into account the correlation between different specifications. This correlation could be very complicated, depending on the relationships among specifications and different biasing adjustment methods to improve or reduce performance. Consequently, for first order estimation, the method in Equation (3-3) is usually a good compromise between prediction accuracy and complexity. The most accurate prediction can be obtained with moderate specifications.

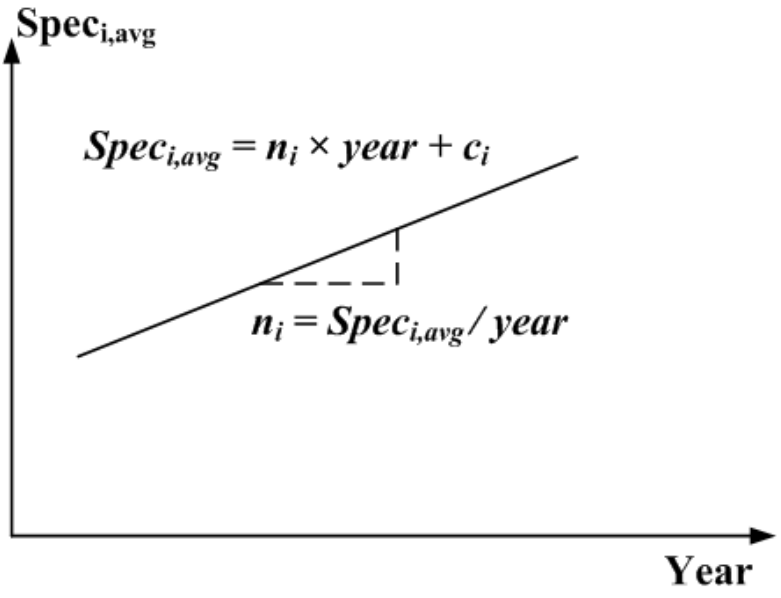


Figure 3-6 Average specification variation through years

Trendline: average specification versus years (Figure 3-6)

To calculate the FoM, average values of specifications $Spec_{i,avg}$ are needed as a reference. One simple solution for determing average values is to select the mean value of each specification from all the collected data points (as will be shown soon).

However, this average value is the calculation result from the past 10~15 years, and cannot be guaranteed to be still an average value in the future. In fact, according to ITRS, device physical parameters change with time, as does the achievable performance. Apart from this, emerging new communication systems also force designers to vary the performance of their circuit blocks. Hence, the actual average specification through the years is a complicated parameter that depend on both transistor level and system level factors. In spite of this, the average specification through the years is obtained and a linear fitted trendline is derived, assuming the trend is to be maintained as for past years. Actually, this assumption is often valid to some extent. Take the ADC, for example; the industry is always pursuing higher bandwidth from the communication system level point of view, and digitizing resolution tends to be reduced due to the falling voltage supply, from a transistor level point of view. These trends can be observed clearly and it is reasonable to assume that this will be maintained in the foreseeable future. With this assumption, the linear fitted trendline for certain specification through years, which is involved in equation (3-3), can be written as:

$$Spec_{i,avg} = n_i \times year + c_i \,. \tag{3-4}$$

By substituting equation (3-4) into equation (3-3), the final FoM prediction with calibration can be obtained as:

$$FoM = (k \times year + b) + \sum_i \{m_i \times [Spec_i - (n_i \times year + c_i)]\}. \tag{3-5}$$

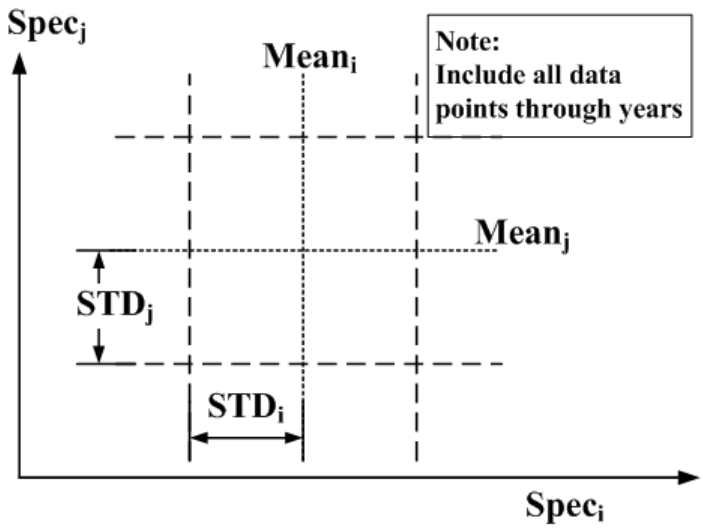


Figure 3-7 Specification relationships and boundaries

Specification limitations and combinations (Figure 3-7)

The limitations of specifications are given with mean values and standard variations. In addition to this, the boundaries are plotted in pairs, so that the effective combinations and trade-offs among specifications can be observed. Some obvious relationships between specifications can then be observed through the figure, such as high frequency/high noise figure, high gain/low linearity, high frequency/ high phase noise, high resolution/low bandwidth, etc. With these figures as guides, the engineer can select reasonable combinations, according to the data density on these specification figures. Generally, higher data density means more actual design has been done to achieve the combination, and hence more it is more likely to be feasible to achieve performance consistent with the FoM trends for a new design.

In the following sections, the trendlines of FoM vs. year, the FoM vs. specifications relationship, the trendlines of specifications vs. year and the relationships among specifications are obtained respectively for each receiver block, followed by the analysis and finally the power consumption predictions are demonstrated.

3.3 FoMs of Receiver Blocks

3.3.1 Low Noise Amplifier

The low Noise Amplifier (LNA) is usually the first active on-chip stage of a receiver chain. Generally, the LNA can be classified as either a narrowband LNA or a wideband LNA. In this section, because the design styles and trade-offs for these tend to be different, the FoMs of narrowband and wideband LNAs are defined and discussed separately.

Narrowband LNA

A typical narrowband LNA architecture is shown in Figure 3-8. The degeneration inductor L_s and series-connected inductor L_g set the real part of the input impedance of the LNA (typically to 50Ω) as well as tune out the gate capacitance at certain frequencies, therefore achieve the narrowband power matching [30].

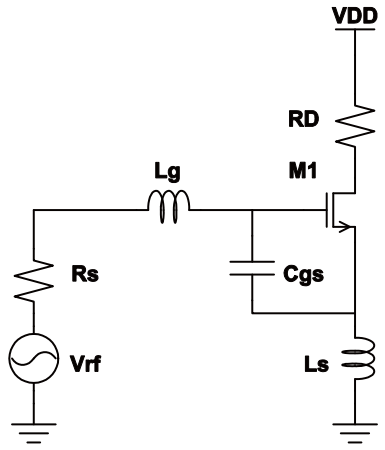


Figure 3-8 L-degenerated Common-Source LNA

The factors affecting the FoM of the LNA include gain, noise figure, linearity, operating frequency and power consumption. There are mainly two kinds of FoM definitions in publications. For narrow band applications, when the linearity is not an important specification, the LNA performance can be compared according to frequency, gain and noise figure. However, in modern communication systems, there might be strong in-band interferers, so linearity should be taken into account in such receivers. Equations (3-6) and (3-7) give the FoM definitions with and without linearity specifications, respectively [31].

$$FoM1_{NBLNA}\left[\frac{GHz}{mW}\right] = \frac{Gain[abs] \times Freq[GHz]}{(F[abs]-1)} \times \frac{1}{P[mW]}. \tag{3-6}$$

$$FoM2_{NBLNA}[GHz] = \frac{Gain[abs] \times Freq[GHz] \times IIP3[mW]}{(F[abs]-1)} \times \frac{1}{P[mW]}. \tag{3-7}$$

All the parameters are measured as absolute values. These equations are essentially *the performances normalized by power consumption*. They indicate that more power consumption is needed to achieve more gain, better linearity, higher frequency and lower noise figure. The reason for these trade-offs will be explained separately in the following paragraphs.

3.3.1..1 Operating Frequency

The fundamental maximum bandwidth for an amplifier depends on the transition frequency f_T of the MOS device itself.

$$Freq \propto f_T \propto f_{max}. \tag{3-8}$$

For CMOS devices, the definition of f_T is [30]

$$f_T = \frac{g_m}{2\pi(C_{gs}+C_{gd})} \approx \frac{g_m}{2\pi C_{gs}}. \tag{3-9}$$

According to Equation (3-9), for a certain technology, a higher operating frequency requires a higher transconductance, which means a higher bias current I_{DS} if the overdrive voltage is fixed ($g_m=2I_{DS}/(V_{GS}-V_{TH})$), and hence higher power consumption in the case of a fixed supply voltage. Note that the gate-source capacitance usually dominates the denominator item, which is equal to $(2/3)WLC_{ox}$ in the strong inversion case.

For different technologies, assume V_{gs} is set so that the device is at the onset of strong inversion, and the transistor is biased in saturation. Note that if V_{gs} is very large, vertical field mobility begins to influence the transconductance, and the operating range will be reduced at the drain terminal. The f_T of a long-channel device is then given by

$$\begin{aligned} f_{T_{long}} &\approx \frac{g_m}{2\pi C_{gs}} \\ &\approx \frac{\mu C_{ox}(W/L)\left(V_{gs}-V_t\right)}{2\pi(2/3)WLC_{ox}} \\ &= \frac{3\mu\left(V_{gs}-V_t\right)}{4\pi L^2} \propto \frac{V_{DS,sat}}{L^2}. \end{aligned} \tag{3-10}$$

The f_T in a short channel CMOS device is different because of the velocity saturation effect [32]:

$$\begin{aligned} f_{T\_short} &\approx \frac{g_m}{2\pi C_{gs}} \\ &\approx \frac{\mu C_{ox} W E_{sat}}{2\pi(2/3)WLC_{ox}} \\ &= \frac{3\mu E_{sat}}{4\pi L} \propto \frac{v_{sat}}{L}. \end{aligned} \tag{3-11}$$

Here, E_{sat} is the field strength where the carrier velocity drops to half of the value extrapolated from the low-field mobility. Note that with transistor feature size scaling down, the maximum transition frequency increases, and so does the possible operation frequency of LNA.

3.3.1..2 Gain

For a long channel CMOS transistor, the voltage gain parameter (Gain) in Equation (3-6) is approximately proportional to the load resistance and the input transistor's transconductance at low frequency:

$$Gain[abs] \propto g_m \times R_{load}. \tag{3-12}$$

In the case of impedance matched LNA circuits, however, it can be shown that the effective transconductance is not directly dependent on the device transconductance [33]. Instead, it is determined by

$$G_m = \omega_T / 2\omega_0 R_s. \quad (3\text{-}13)$$

A detailed derivation of this expression can be found in Appendix A-2.

From this equation, it can be observed that with the same bias condition and the same operating frequency, the effective transconductance increases with increased ω_T/ω_0 which means that the maximum achievable gain increases using new technologies with the same power consumption for the same application.

In LNA design, input impedance matching is needed to transfer maximum possible power from the antenna. Output matching is sometimes needed as well if the output of the LNA needs to terminate to an off-chip load, for example, a band-pass-filter, etc. In many published works, the measurement of the LNA is made with a vector network analyzer, which provides S-parameter measurement results with ideal matching conditions. However, when attempting to reflect the technology factor in the LNA performance improvement, especially for integrated CMOS circuits, the voltage gain is more closely related to technology parameters. The output of the LNA to a 50Ω probe are normally either connected to a source follower buffer or matching network.

In the process of data collection of this literature review, If the measured power gain is provided, the voltage gain is simply determined as 6dB added to the power gain for both the source follower buffer and the output matching network cases, while the measured IIP3 is adopted directly.

Some published results also provide voltage gain with a capacitive load instead of power gain. In this case, the provided voltage gain is adopted. There are also some designs using a common source amplifier as the output buffer, in which case the load is usually a 50Ω probe. For these designs, the common source amplifier gain should be given, and the voltage gain is the measured power gain minus the buffer's gain.

3.3.1..3 Noise Figure

The *(F-1)* term in the denominator of Equation (3-6) is used instead of merely *F* based on consideration of the amplifier's noise contribution to the total system. At the system level, the noise figure of a single stage common source LNA with inductive degeneration is defined as

$$F = \frac{N_{source} + N_{LNA}}{N_{source}} = 1 + \frac{N_{LNA}}{N_{source}}. \quad (3\text{-}14)$$

The N_{source} is the noise power generated by the source, which is typically a 50Ω resistor, and N_{LNA} is the input referred noise generated by the LNA itself, and therefore *(F-1)* separates the noise already presents at the input and the noise generated by the LNA itself.

$$F - 1 \propto \frac{N_{LNA}}{N_{source}}. \quad (3\text{-}15)$$

The minimum achievable noise factor for a narrowband LNA is analysed in [33]. Assume a linear two-port noise model, and the noise in the LNA is dominated by the thermal noise of the channel current, it shows that with an optimized device width and constrained by the fixed power consumption, the minimum noise factor can be approximately obtained and is given by

$$F_{min,P} \approx 1 + 2.4\frac{\gamma}{\alpha} \cdot \frac{\omega_0}{\omega_T}. \quad (3\text{-}16)$$

The drain current noise coefficient γ (Note, this is not the body effect coefficient) is typically 2/3 for long channel devices and typically 2~3 for short channel devices. The parameter α equals g_m/g_{d0}, and is unity for long-channel devices, and decreases as the channel length shrinks. The parameter g_{d0} is the drain-source conductance at zero V_{DS}. This equation implies that the noise performance improves with increasing ω_T, which is in turn improving with the scaling down of the feature size.

3.3.1..4 IIP3

In a wireless communication receiver, signals are usually treated as small signals before the baseband amplifier, and the most important non-linearity effects for small signals are the 2nd and 3rd order intermodulation products, of which the former should be minimized in a direct-conversion receiver architecture, while the latter is to be suppressed enough for any architecture to avoid large interference from adjacent channels.

For a single transistor amplifier, assume the IIP3 is caused mainly by the transconductance of the transistor; then the IIP3 can be expressed as:

$$V_{IIP3} = \propto \sqrt{V_{ov}} \propto \sqrt{I_D/g_m}. \quad (3\text{-}17)$$

The derivation of this equation can be found in Appendix A-3, in which the theoretical IIP3 vs. overdrive voltage is calculated for some technologies, according to the Spice model provided on the MOSIS website [34].

This conclusion is also supported in [35], where a double-balanced mixer is analyzed for its linearity performance. As the experiment results illustrated in [36] show, the CMOS device IIP3 is generally getting worse with the improvement of technology, given the same bias current density.

Despite this, in most LNA designs, the linearity of the amplification transistor is partially de-coupled from the overall linearity by the feedback of inductive degeneration. In fact, by properly increasing the overdrive voltage, the linearity is also improved. So, it's difficult to determine theoretically the linearity trends. However, it is expected that the FoM improvement with IIP3 should be no better than that without IIP3 performance.

3.3.1..5 FoM Prediction

The power consumption of the LNA is simply

$$P_{diss} = V_{DD} \times I_D \,. \tag{3-18}$$

As described in all the discussion above, the FoM definition in Equation (3-7) can be expressed in the form of the device parameters by replacing the Freq, Gain, NF, IIP3 and P_{diss} with Equation(3-8), (3-12), (3-16), (3-17) and (3-18) respectively, resulting in the expression below [37] :

$$FoM2_{NBLNA} \propto \frac{g_m \times R_{load}^2 \times f}{\frac{N_{LNA}}{N_{source}} \times V_{DD}}. \tag{3-19}$$

However, this equation is only partially related to the device parameters, which can be used as a guide when deciding the trade-offs in the design of a practical LNA. A better comparison method is to express the FoM by the device parameter improvement or degradation.

By substituting Equation (3-13) and (3-16) into Equation (3-6), the FoM without the IIP3 parameter is obtained:

$$\begin{aligned} FoM1_{NBLNA} &= \frac{Gain \times f_0}{P_{diss}} \times \frac{1}{(F-1)} \\ &\propto \frac{(\omega_T/\omega_0) \times (R_L \times \omega_0)}{P_{diss}} \times \frac{1}{\omega_0/\omega_T} \\ &\propto \omega_T^2 \,. \end{aligned} \tag{3-20}$$

The values of R_L, f_0 and P_{diss} are fixed. With Equation (3-20), the theoretical figure of merit improvement rate over the years can be obtained according to the values of ω_T in the published ITRS documents from year 2003 to year 2009.

The effects of MOS bias and width tradeoffs on the LNA for a certain technology (fixed power supply voltage and fixed channel length) and fixed power consumption (fixed power supply voltage and fixed bias current) are worth investigating to get a good guide for the design issues. The transistor is usually biased at around the onset of strong inversion. When the width of the transistor reduces, the overdrive voltage must increase by the square root of the reduction in width. The transconductance (and hence gain) and transition frequency also reduce and increase with the same rate of overdrive voltage, respectively. The relationships of IIP3 vs. NF performance and device biasing/size are also discussed in Equations (3-16) and (3-17). These trade-offs are summarized in Table 3-1, as explained in [38, 39].

MOS Bias Tradeoffs			MOSFET Performance				
Bias	W/L	V_{gs}-V_{th}	g_m/I_D	Gain	f_T	IIP3	NF
Strong Inversion	↓	↑	↓	↓	↑	↑	↑
Desired			↑	↑	↑	↑	↓

Table 3-1 MOS bias/sizing trade-offs

Figure 3-9 and Figure 3-10 illustrate the FoMs collected from the literature and predicted for narrowband LNAs with and without the IIP3 parameter. Note that the FoMs are defined on a linear scale in equation (3-6) and (3-7), but in the figures, the FoM values are expressed in units of dB, in order to present the general expected exponential improvement discussed in section 1.2 more effectively. For all the following receiver blocks through this chapter, the FoM values are also expressed with the units of dB in the figures. It can be observed that the theoretical predictions obtained in Equation (3-20) are quite similar to the statistical predictions. Take the example of Figure 3-9, the theoretically predicted $FOM1_{NBLNA}$ improves 3dB every 34.1 months as a result of the CMOS technology scaling progress, while the statistical predicted value improves 3dB every 34.8 months. For the FoM with IIP3 parameter, the statistical value of $FOM2_{NBLNA}$ improves 3dB every 38.4 month, which is slightly slower than $FOM1_{NBLNA}$ and supports the arguments regarding scaling in the IIP3

section. Note that it is valuable to compare this FoM improvement speed with that of digital transistor density, which is doubled, or 3dB higher every18~24 month.

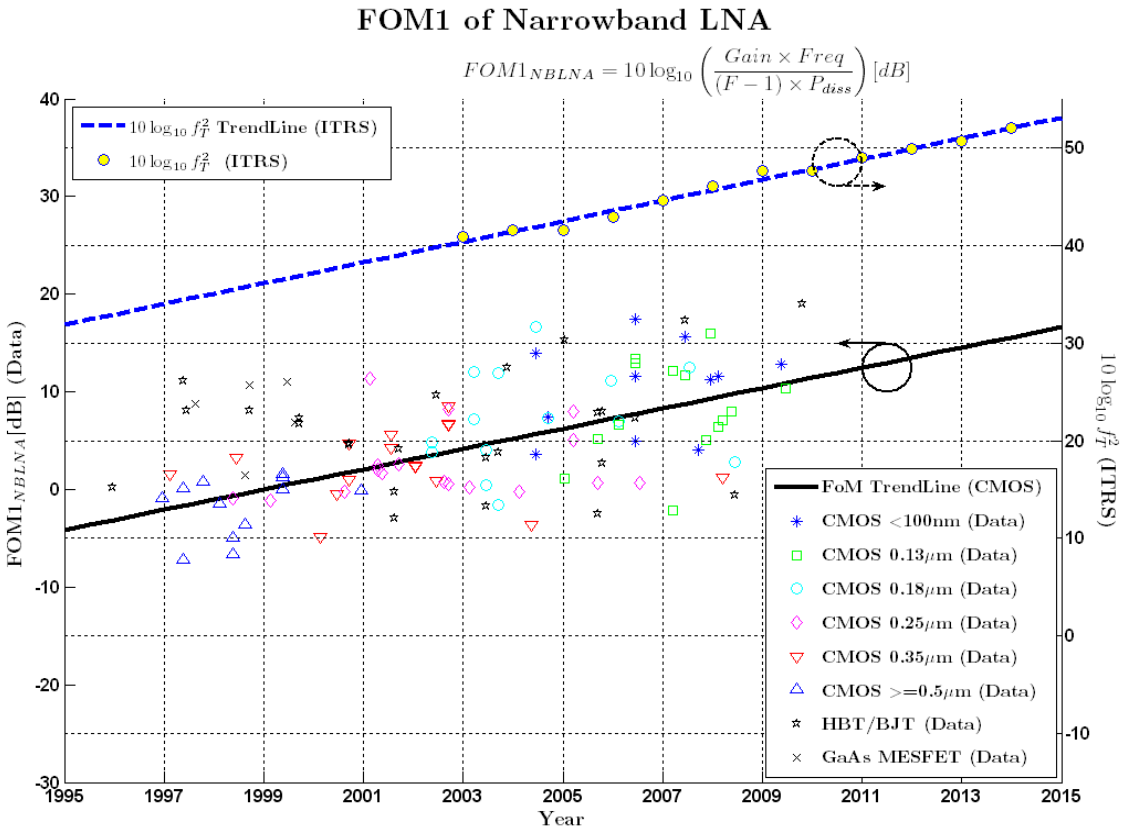


Figure 3-9 FoM tendency line narrowband LNA (without IIP3). The Y-axis on the right is 10log (f_T^2), representing the theoretical improvement rate of the FoM due to the f_T improvement

Also, it can be seen clearly that the adopted technologies are changing gradually, resulting in the improved FoM. Take the example of Figure 3-10, in the year of 1999, when most LNAs are fabricated in 0.35μm CMOS or older technology, the average FoM is about -0.7 dB, corresponding to 0.85 GHz on a linear scale, while in the year of 2009, the FoM improved to 8.7 dB, or 7.41 GHz, with 0.13μm CMOS or later technologies. This figure is therefore predicted to reach 14.3 dB in the year of 2015.

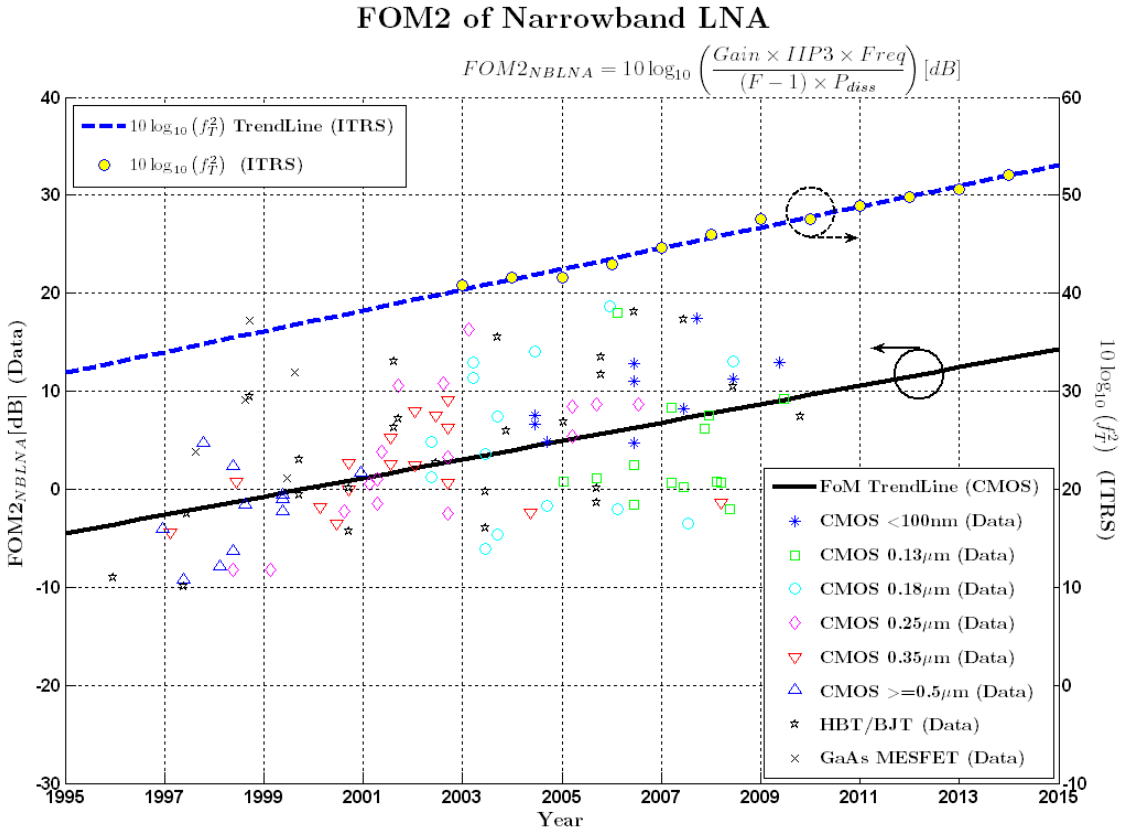


Figure 3-10 FoM tendency line of narrowband LNA (with IIP3)

The relationships of FoM vs. gain, FoM vs. noise, FoM vs. linearity and FoM vs. frequency are plotted in Figure 3-11 to Figure 3-14. As discussed before, FoM is usually some performance metric normalized to power consumption. If power consumption increases or decreases with the same rate of performance, the power vs. performance slope will be nearly one, resulting in the slope of FoM vs. performance being almost zero (constant value). For example, if the power consumption doubles when the voltage gain is raised by 6dB, then the resulting FoM remains unchanged, implying a strong correlation between voltage gain and power consumption. On the other hand, if correlation between power consumption and performance is relatively low, the FoM vs. performance will have a non-zero slope. In Figure 3-11 and Figure 3-12, the slopes of FoM versus gain and FoM versus noise are almost zero, which means that power consumption is nearly proportional to performance. For most conditions, the slope of FoM vs. gain and noise can be treated as zeros. Note that it is *F-1*, instead of noise figure that is investigated according to the FoM definitions. In Figure 3-13, however, the slope of FoM versus IIP3 is about +0.6dB/dB. This means that if IIP3 in dBm is increased by 3dB, the FoM will increase by 1.8dB, resulting in 32% more power consumption, instead of 100% more power consumption as in the

case of gain and noise. Similarly, the power consumption can only be saved by about 25% if IIP3 is reduced by 3dB. As for the frequency, in Figure 3-14, the FoM increases by about 5dB for a 10 times' higher frequency. These figures imply that the gain and noise are strongly correlated with power consumption, while the correlation between IIP3 and frequency with power consumption is less strong.

It can also be observed in the figures that the FoM vs. performance relationships are maintained for different technologies. Take the example of Figure 3-13, where the FoM of 0.5µm CMOS LNAs are mostly below the fitted line, the FoM of CMOS less than 100nm are mostly located above the fitted line, and their trends can be roughly recognized as similar to the fitted curve. This supports the proposition that the slope of a fitted linear relationship is valid for different technologies.

The specification trends over several years are illustrated in Figure 3-15 ~ Figure 3-18. The average gain and IIP3 are nearly unchanged over the past ten years. The noise figure increases slowly, and the average frequency increases by 24% every year.

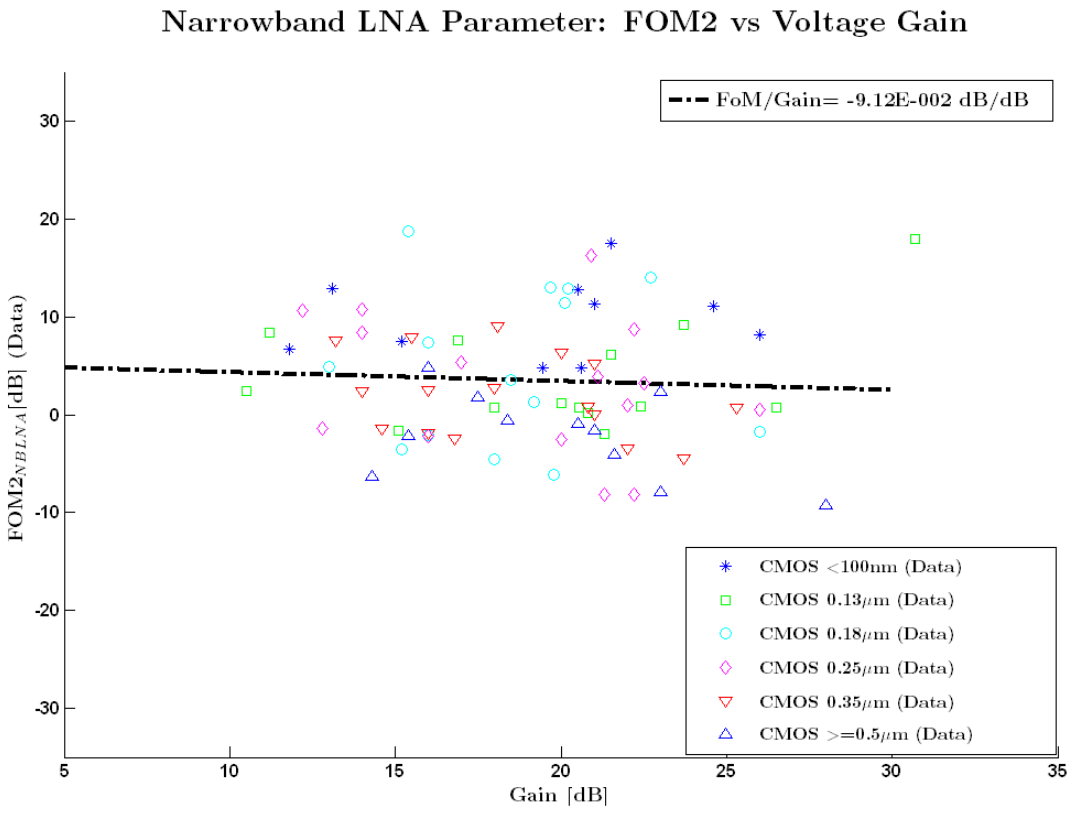


Figure 3-11 Narrowband LNA parameters: FoM2 versus voltage gain

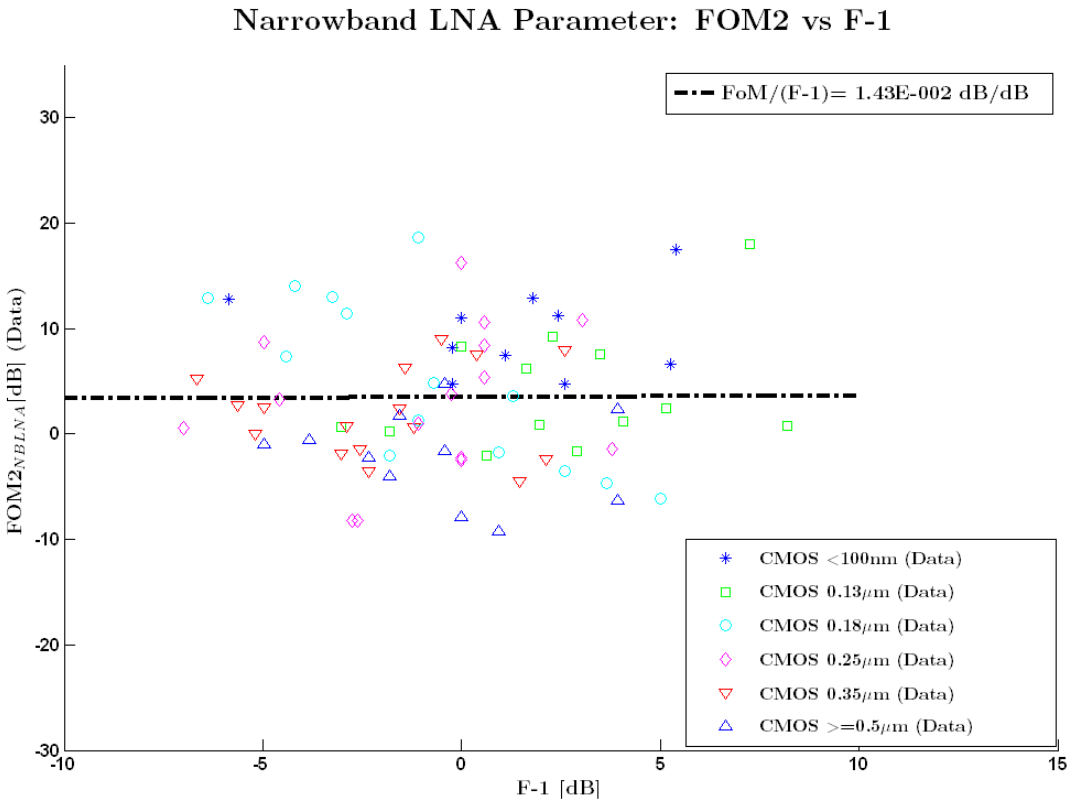


Figure 3-12 Narrowband LNA parameters: FoM2 versus noise (F-1)

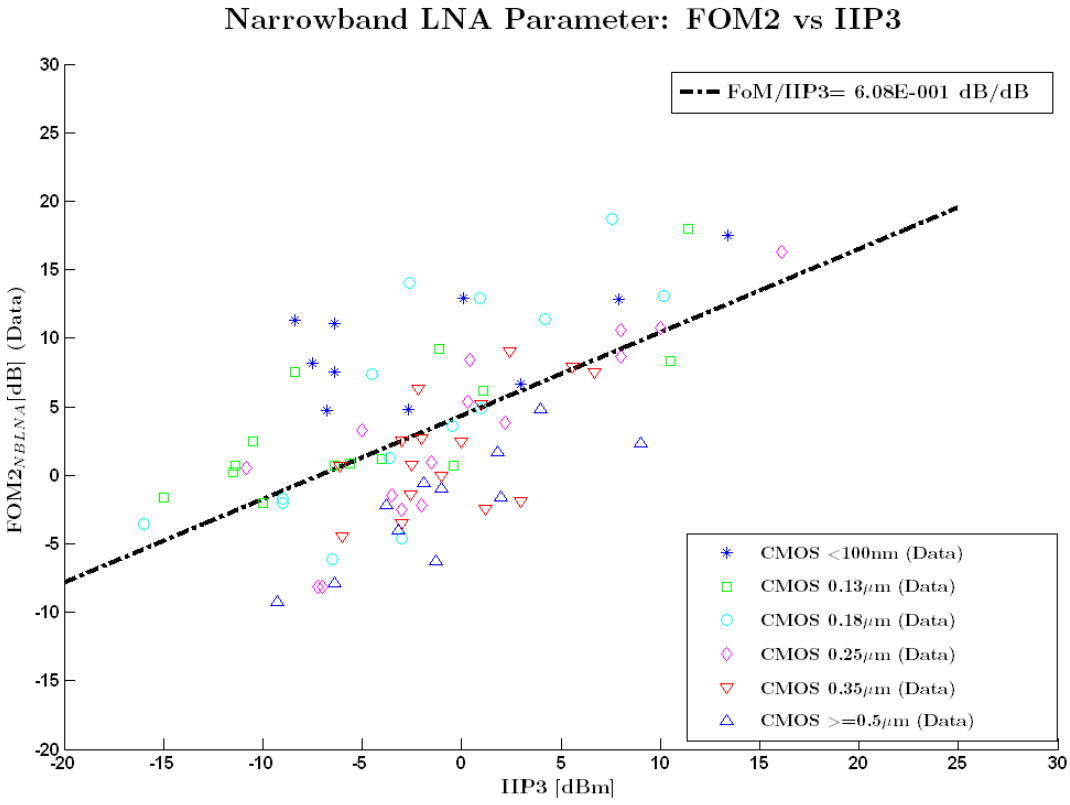


Figure 3-13 Narrowband LNA parameters: FoM2 versus IIP3

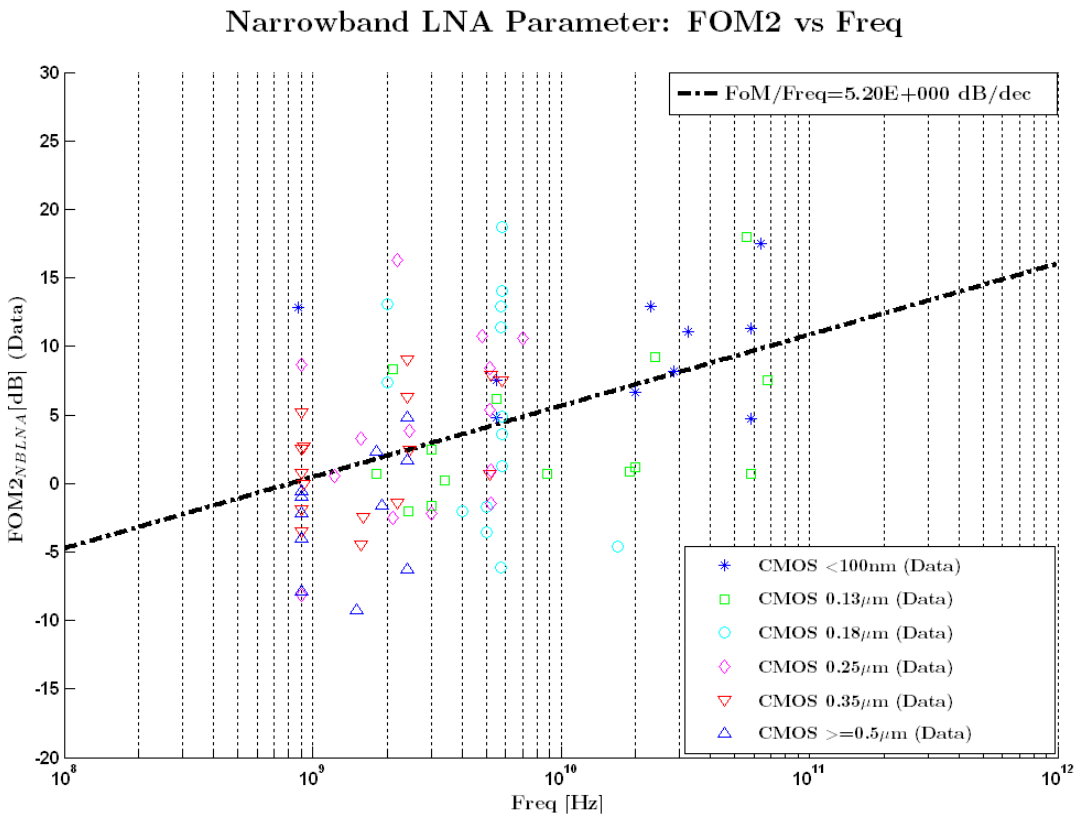


Figure 3-14 Narrowband LNA parameters: FoM2 versus frequency

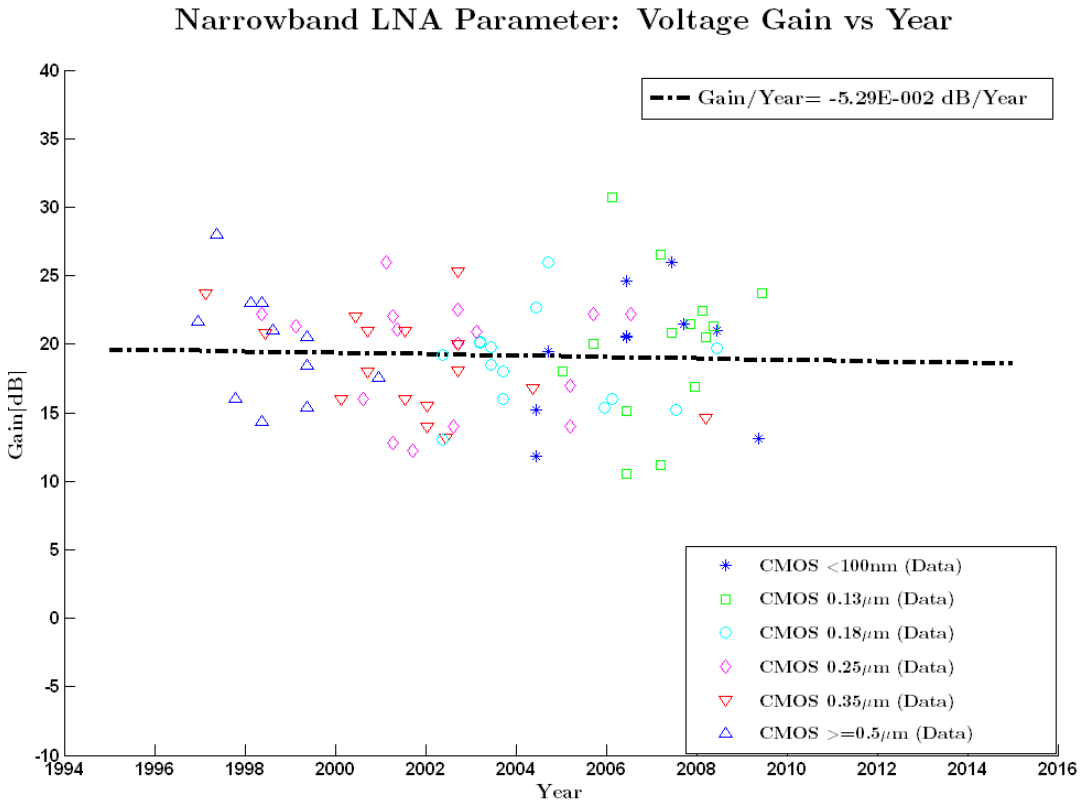


Figure 3-15 Narrowband LNA parameters: voltage gain versus year

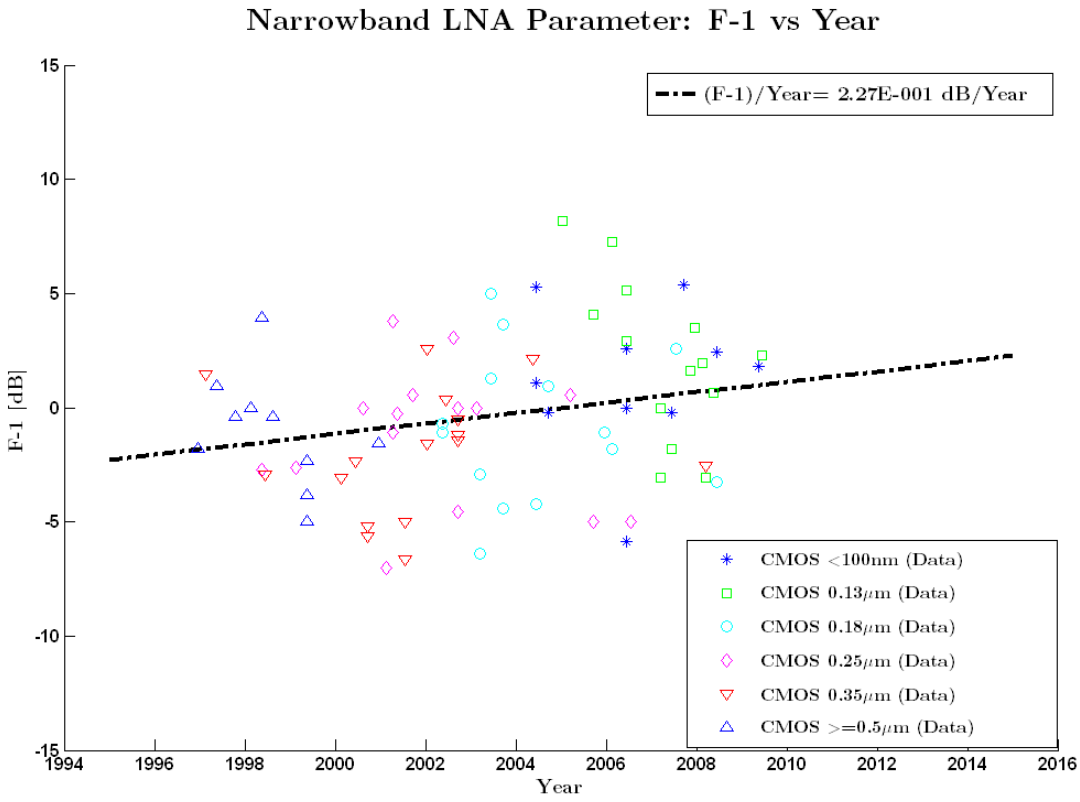


Figure 3-16 Narrowband LNA parameters: F-1 versus year

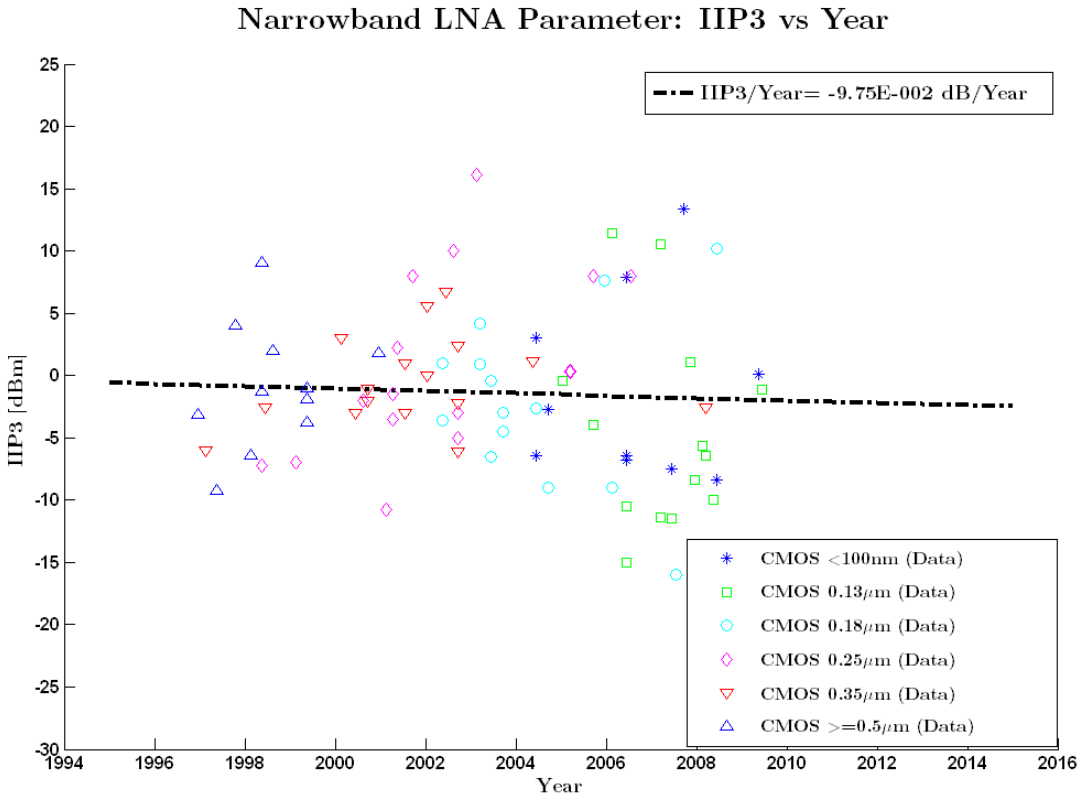


Figure 3-17 Narrowband LNA parameters: IIP3 versus year

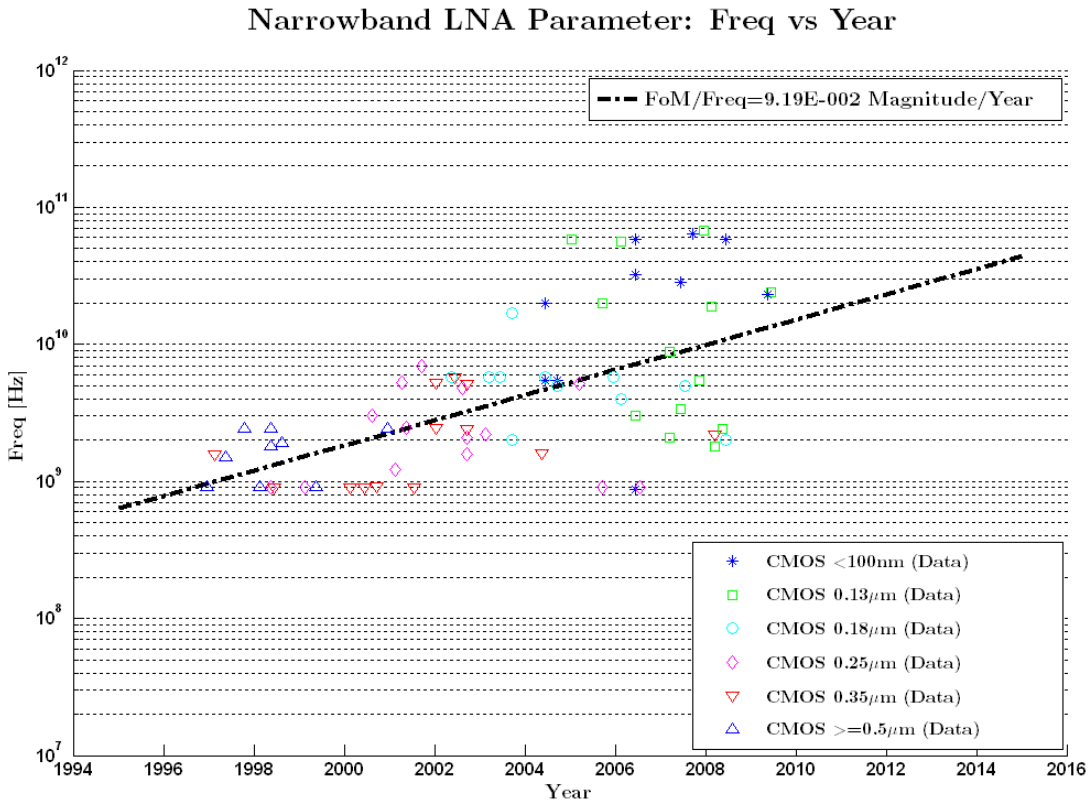


Figure 3-18 Narrowband LNA parameters: frequency versus year

Theoretically, there are also design trade-offs between performance requirements. Generally speaking, lower noise can be achieved with higher gain, at the cost of frequency. On the other hand, there is a general inversely proportional relationship between IIP3 and noise figure, as well as between IIP3 and gain. The performance relationships are investigated as illustrated in Figure 3-19 to Figure 3-22. The discussed relationships can be observed roughly in the figures, especially the gain-noise, gain-IIP3 and frequency-noise relationships. Because of the different design and measurement methods in publications, these trade-offs can have very large variations as demonstrated, instead of a generally linear relationship when plotted on a log scale, as indicated by the theoretical conclusions. Therefore, it is more meaningful to treat these figures as references to verify the effectiveness the combinations between these performances, as well as the performance limitations. When choosing specifications, areas with higher data density are more convincing and easier to achieve because more practical designs are published. Generally, circuits are easier to implement by selecting performance near the mean value and within standard variation range. According to the figures, this criterion includes the voltage gain between 15.0dB and 23.3dB, a noise figure between 1.6dB to 4.8dB, IIP3 between -7.9dBm to 5.2dBm, and frequency

between 1.1GHz to 13.1GHz (and of course, there are many 900MHz GSM band LNAs). If some extreme specifications are to be achieved, these figures provide estimations of the difficulty. For example, if a high linearity LNA with IIP3 over 10dBm is required, according to Figure 3-20, it is difficult to achieve a voltage gain of over 20dB simultaneously. Similarly, when the signal frequency is over 10GHz, it is difficult to achieve a noise figure below 3dB, as observed in Figure 3-22.

The collected published LNA performance figures are also plotted as a histogram of frequency in Figure 3-23. Most of the LNAs are within the popular band from about 1GHz to 10GHz, including GSM, 3G, Bluetooth and WLAN etc. Therefore, the FoM statistical data can be used more convincingly in these communication systems. By contrast, mm-wave LNAs only occupy a small portion of the data, so it is not well fitted in the mm-wave band such as radar and 60GHz applications.

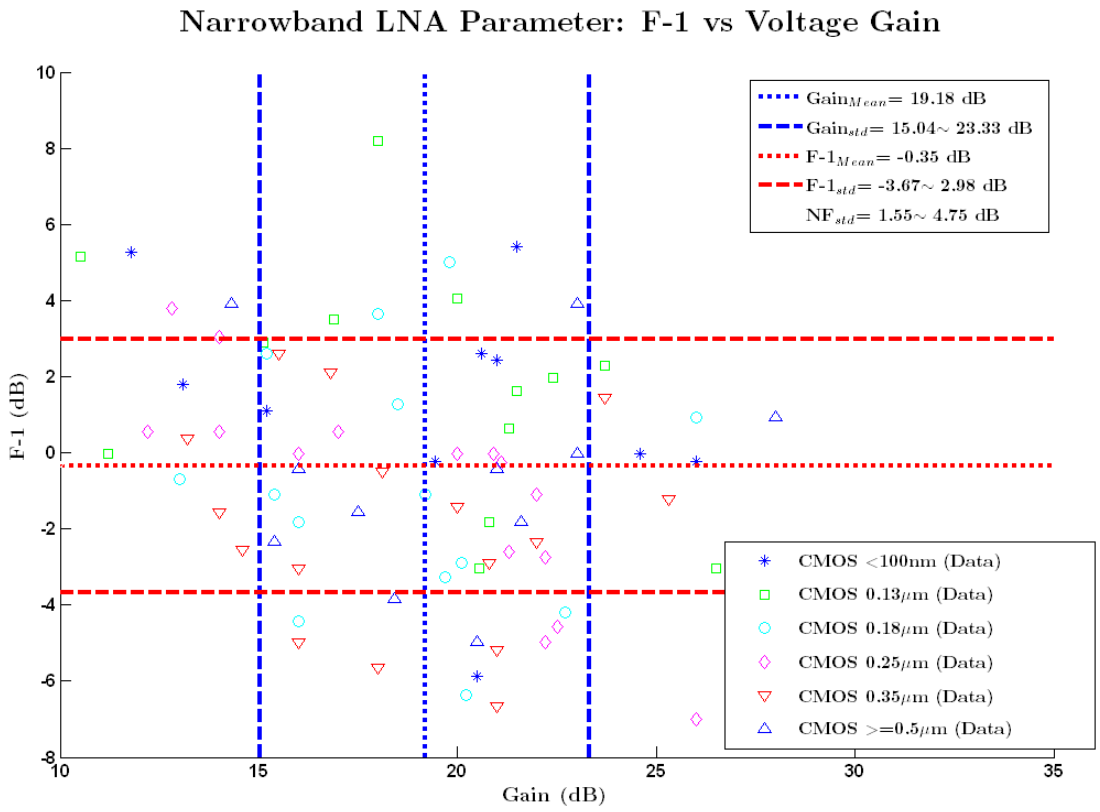


Figure 3-19 Narrowband LNA parameters: noise figure versus gain

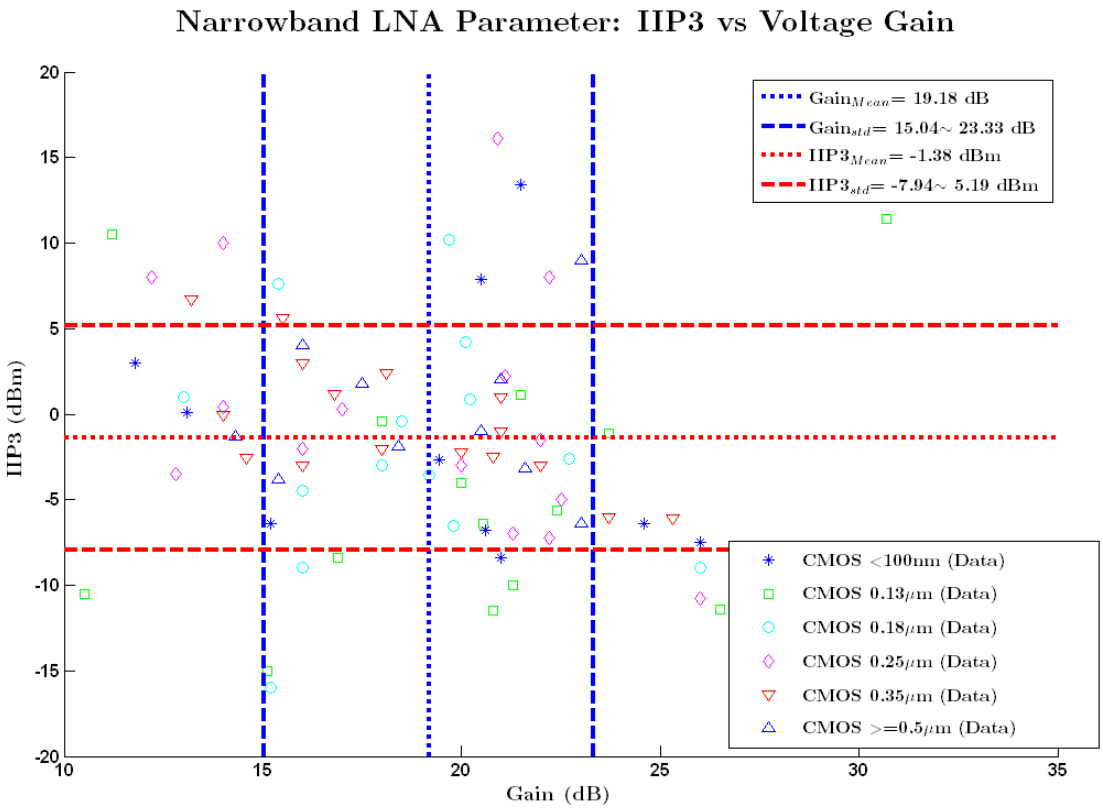


Figure 3-20 Narrowband LNA parameters: IIP3 versus gain

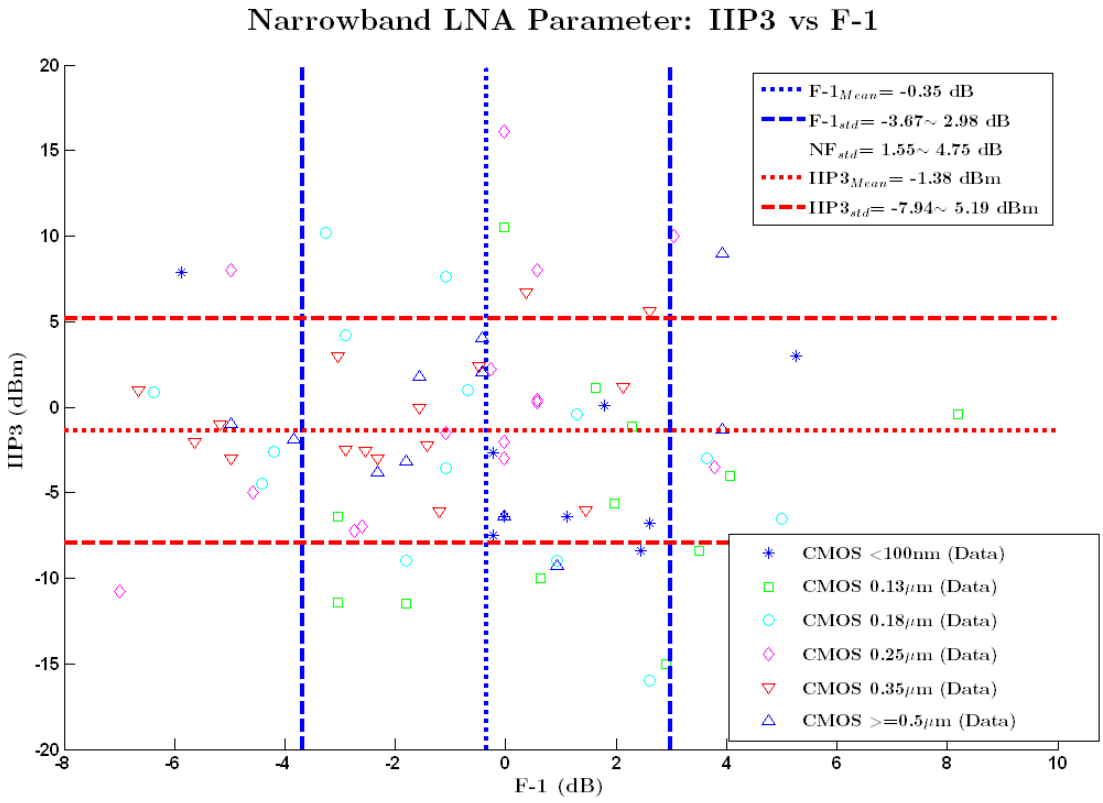


Figure 3-21 Narrowband LNA parameters: IIP3 versus noise figure

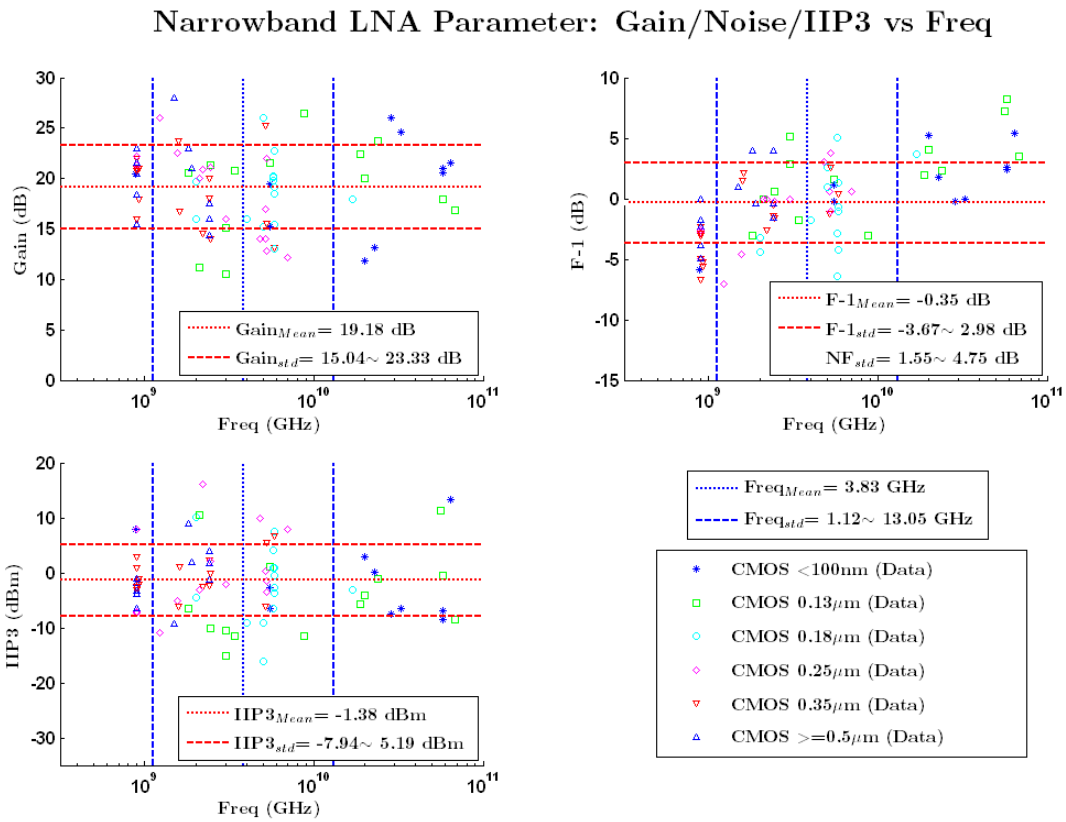


Figure 3-22 Narrowband LNA parameters: gain, noise, IIP3 versus frequency

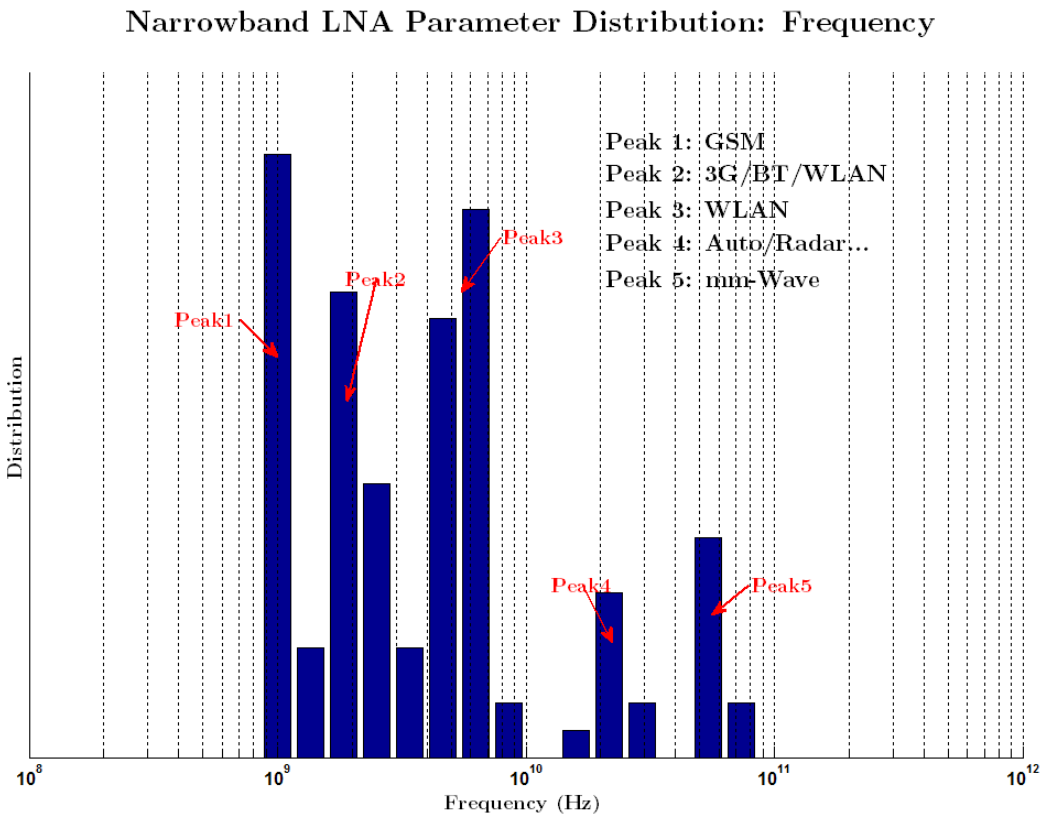


Figure 3-23 Narrowband LNA parameters: frequency distribution

The full information of average FoM vs. year trendline, FoM vs. specification relationships and specification vs. year trendline are summarized in Table 3-2. With this information, the power consumption for certain specifications in certain years can be predicted using Equation (3-5). To verify the effectiveness of this method, examples of the prediction LNA power consumptions are made. In Figure 3-24, the power consumptions of LNAs with moderate, high and low performances are calculated and predicted and compared with the power consumption of all the collected data. Ideally, the power consumption of collected data points should be grouped with respect to specifications. However, due to the large variation of the performance combination among the designs, it is impossible to do that. Nevertheless, it can be observed that most published LNAs' power consumption are within the predicted range and have very similar distributions and trends. This supports the effectiveness of the FoM strategy for the prediction of narrowband LNA power consumption. According to the prediction, the power consumption of moderate specifications of 19dB gain, 2.77dB noise figure, -1.5dBm IIP3 and 4GHz frequency could be lowered to 3.2mW in the year 2015. Higher performance of 22dB voltage gain, 2.12dB noise figure, 3dBm IIP3 and 8GHz frequency might consume 14.4mW power. With relaxed specifications, the power consumptions are expected to achieve sub-mW level.

$FoM_{avg}[dB] = k_i \times year + b_i = 9.4 \times 10^{-1} \times year - 1882.5$	
$Spec_{i,avg} = n_i \times year + c_i$	$FoM/Spec_i = m_i$
$Gain_{avg}[dB] = 5.3 \times 10^{-2} \times year + 125.1$	$\frac{FoM}{Gain} = -9.1 \times 10^{-2}\ dB/dB$
$(F-1)_{avg}[dB] = 2.3 \times 10^{-1} \times year - 456.0$	$\frac{FoM}{F-1} = 1.4 \times 10^{-2}\ dB/dB$
$IIP3_{avg}[dBm] = 9.8 \times 10^{-2} \times year + 194.0$	$\frac{FoM}{IIP3} = 6.1 \times 10^{-1}\ dB/dB$
$Freq_{avg}[Hz] = 10^{9.2\times10^{-2}\times year - 174.6}$	$\frac{FoM}{Freq} = 5.2\ dB/dec$

Table 3-2 Narrowband LNA FoM statistics summary

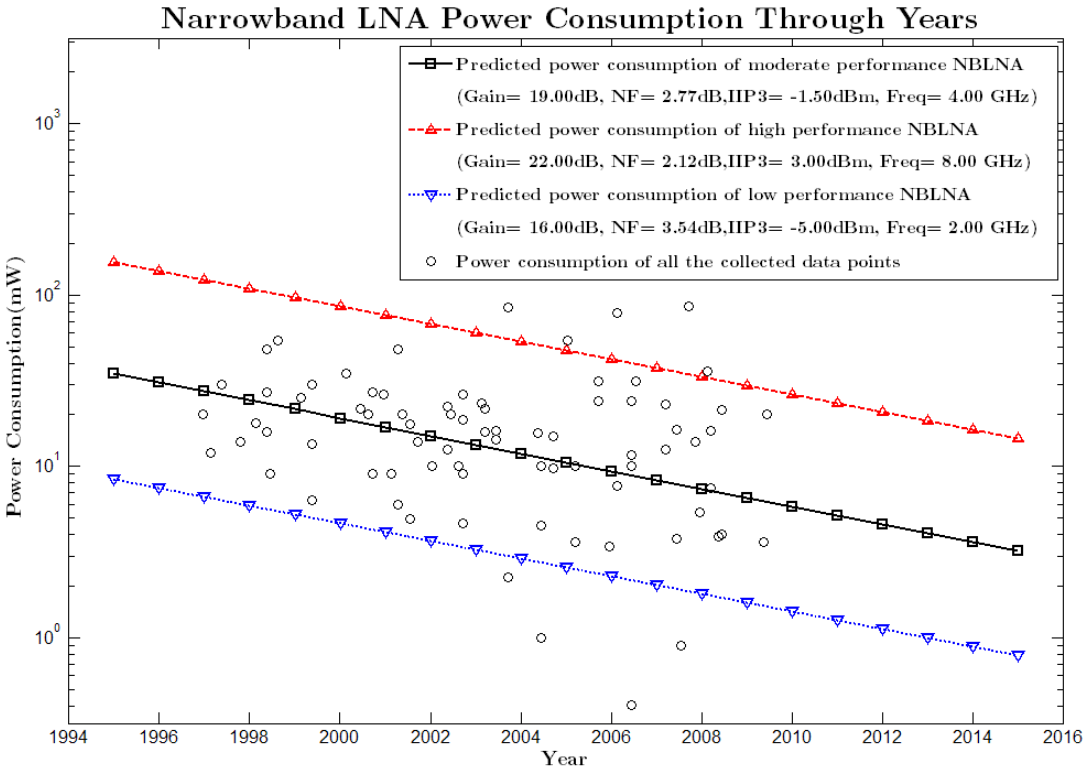


Figure 3-24 Prediction of narrowband LNA power consumption through years

For the following receiver blocks from the next section, including wideband LNA, mixer, VCO, frequency divider, baseband circuits and ADC, in order to keep the chapter more compact and convenient for readers, only the figures of FoM versus year (as in Figure 3-9 and Figure 3-10) and the power consumption prediction (as in Figure 3-24) are shown explicitly in the chapter, while most of the other figures, including the relationships of FoM versus performance, trendlines of specifications versus year and the relationships among specifications are given in Appendix A. Some of these figures may be shown in the chapter for analysis purposes when necessary. All these trendlines and relationships will then be summarized in the form of tables similar to Table 3-2. It is worth emphasising that all the trendlines and relationships are calculated and derived automatically from the collected data points by MATLAB programs. Although in some of the figures, data points have very large variations, the intrinsic trends and relationships could be revealed by the linear fitting method and sometimes supported by theoretical analysis.

Wideband LNA

Unlike narrowband LNAs, in which the inductive degeneration architecture is usually adopted, there are different types of matching structure for wideband LNAs as

shown in Figure 3-25, including resistive feedback [40], common-gate [41], distributed amplifier [42] and bandpass filter matching [43] techniques. Despite differing architectures, the fundamental transistor performance improvement trends are the same as for narrowband LNAs. Therefore, wideband LNAs are considered as being in the same category.

As analyzed before, the maximum operation frequency, instead of bandwidth, is directly related to a transistor's f_T. In spite of this, the fractional bandwidth is usually higher than 50%, as in TV tuners and UWB transceivers, where the wideband LNAs are frequently used. Therefore, the bandwidth is highly correlated with the operating frequency and hence the transistor's f_T. So it is suitable to replace the operation frequency with bandwidth in the case of wideband LNAs.

Another difference between wideband and narrowband LNAs is the performance parameters. In wideband LNAs, the gain, noise figure and IIP3 are not guaranteed to be flat over the entire bandwidth. So the average values are calculated from the literature.

Integrated wideband LNA have started to become popular in publications only in the last 10 years, due to improvements in CMOS technology. The motivation has come from UWB standards, and has been followed by on-chip TV tuners. Most early publications focus on the wideband performance, as opposed to noise figure and IIP3 as is usual for narrowband LNAs. David Barras [44] defined the figure of merit normalized to technology parameter f_{max} (maximum frequency of oscillation as a technology benchmark), which can be used to compare designs using the same technology:

$$FoM_{WBLNA}[mW^{-1}] = \frac{Gain[abs] \times BW[GHz]}{(F[abs]-1)} \times \frac{1}{P[mW]} \times \frac{1}{f_{max}[GHz]}. \qquad (3\text{-}21)$$

However, when attempting to predict the figure of merit with technology improvement, this normalization should be removed, as in Equation (3-22). This definition is very popular in wideband LNA publications.

$$FoM1_{WBLNA}\left[\frac{GHz}{mW}\right] = \frac{Gain[abs] \times BW[GHz]}{(F[abs]-1)} \times \frac{1}{P[mW]}. \qquad (3\text{-}22)$$

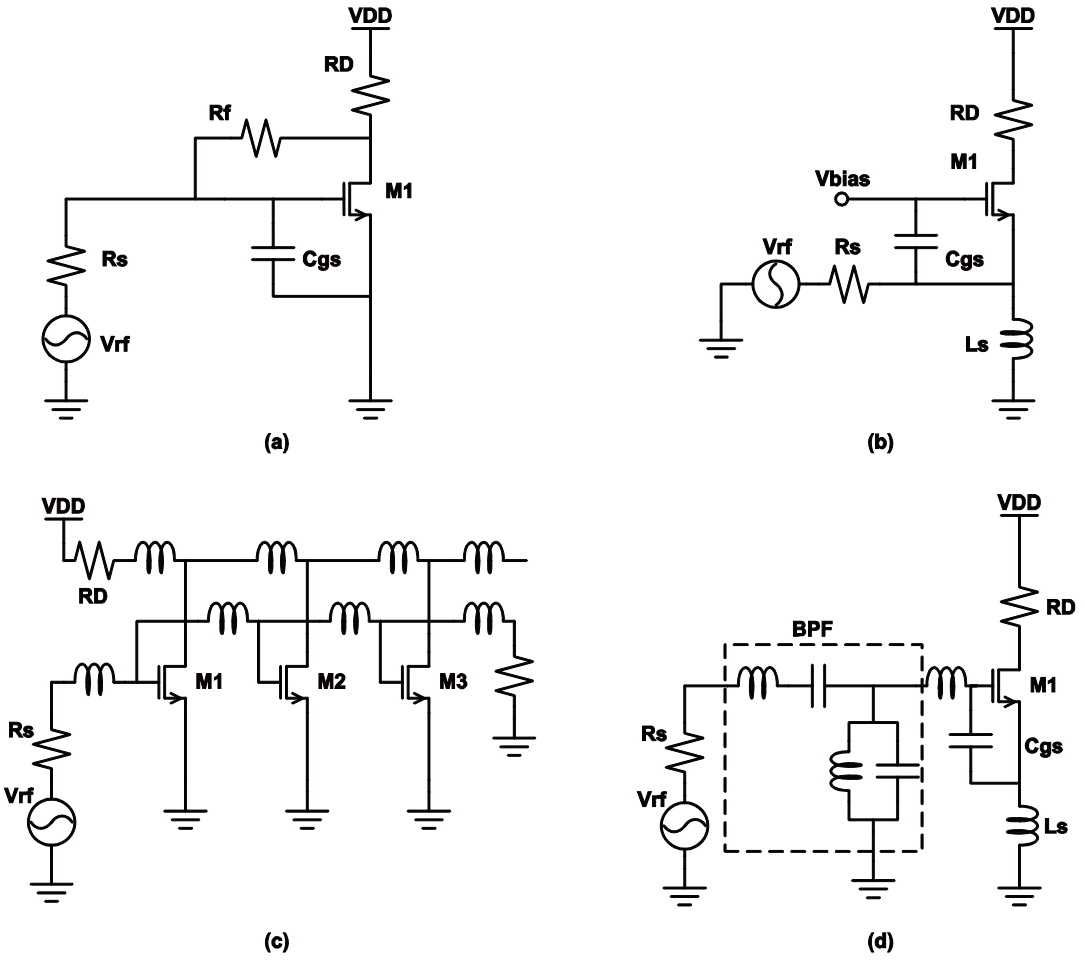


Figure 3-25 Wideband LNA architectures: (a) Feedback, (b) Common-gate, (c) Distributed amplifier, (d) Equivalent input bandpass filter.

In fact, one of the reasons why the linearity is normally not included is that there are few communication standards that cover the UWB band, except 802.11a WLAN. Therefore, strong interferers are seldom present. The 5.2GHz and 5.8GHz WLAN signals are suppressed by notch filters, which is a relatively simple approach compared with high a linearity wideband LNA. However, both the licensed and unlicensed parts of the spectrum are getting crowded rapidly, leading to large number of interferers in the foreseeable future. Therefore, the linearity specifications should be taken into account for future applications, as defined by Amer [45] in Equation (3-23).

$$FoM2_{WBLNA}[GHz] = \frac{Gain[abs] \times BW[GHz] \times IIP3[mW]}{(F[abs]-1)} \times \frac{1}{P[mW]}. \qquad (3\text{-}23)$$

The wideband FoM with and without linearity parameters are illustrated according to different technologies and different architectures in Figure 3-26 ~ Figure 3-29, respectively. It can be observed that the FoM improvements for wideband LNAs are quite similar to those for narrowband LNAs. The predicted FoM without IIP3 improves by 3dB every 31.3 months, and it is 39.4 months if IIP3 is included. Although

the respective trends are quite similar, the FoM of a wideband LNA is about 3~4dB lower than a narrowband LNA all through the years, which means that power consumption for a wideband LNA is normally higher than a narrowband one, given the same gain, NF, IIP3 and maximum signal frequency.

It can also be observed that the variations are very large for the FoM of a wideband LNA where the IIP3 parameter is included; the reason for this is mainly due to the different linearization techniques adopted. The influence of a source follower buffer (which is very popular because of its wideband character) also varies a lot.

Integrated wideband LNAs have only been published over the last 10 years, so CMOS technology older than 0.35µm is seldom involved. Published wideband LNAs in the year of 2004 mainly adopted 0.18µm technology, achieving an average FoM with IIP3 included of about 0.06 dB; this improves in 90nm nowadays. This figure is predicted to reach 14.32 dB in the year of 2015.

FoM values not including the IIP3 parameter are similar for different architecture, while, when IIP3 are included, it can be observed that, somehow, distributed LNAs have the best FoM, followed by the common-gate LNAs and the resistive feedback LNAs fall behind the other architectures.

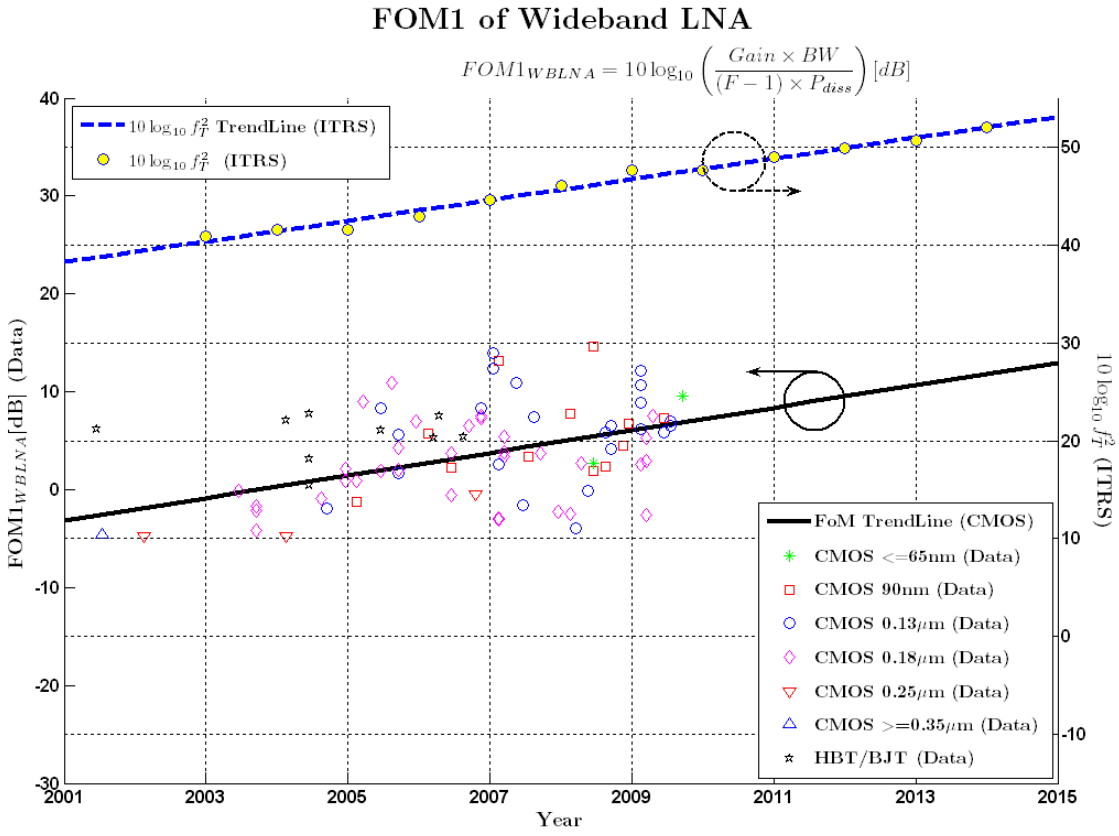


Figure 3-26 FoM tendency of wideband LNA (without IIP3)

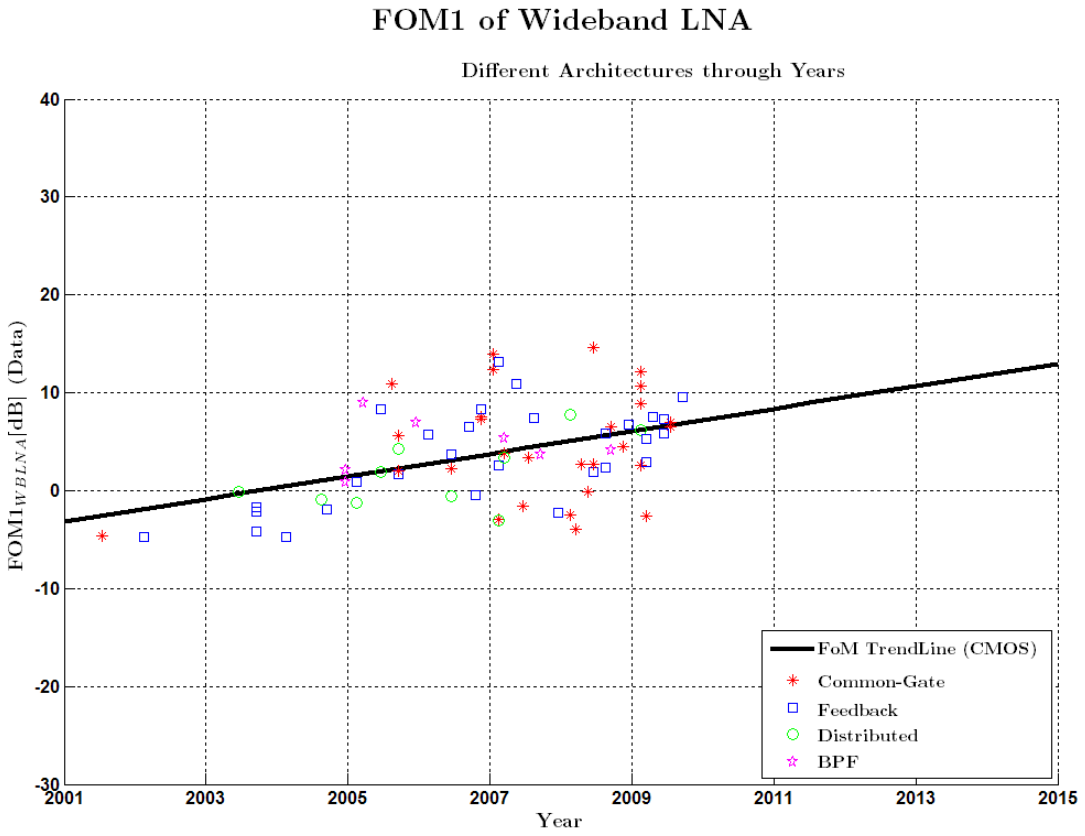


Figure 3-27 FoM tendency of wideband LNA (without IIP3) for different structures

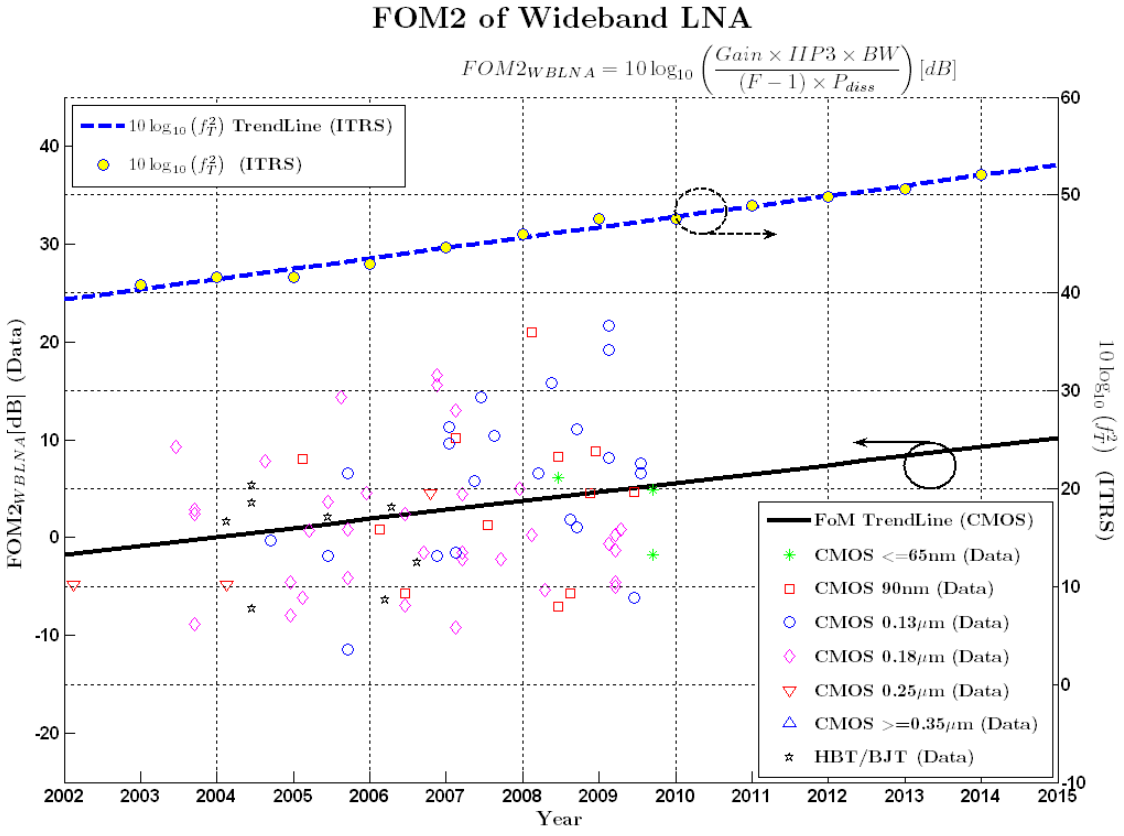


Figure 3-28 FoM tendency of wideband LNA (with IIP3)

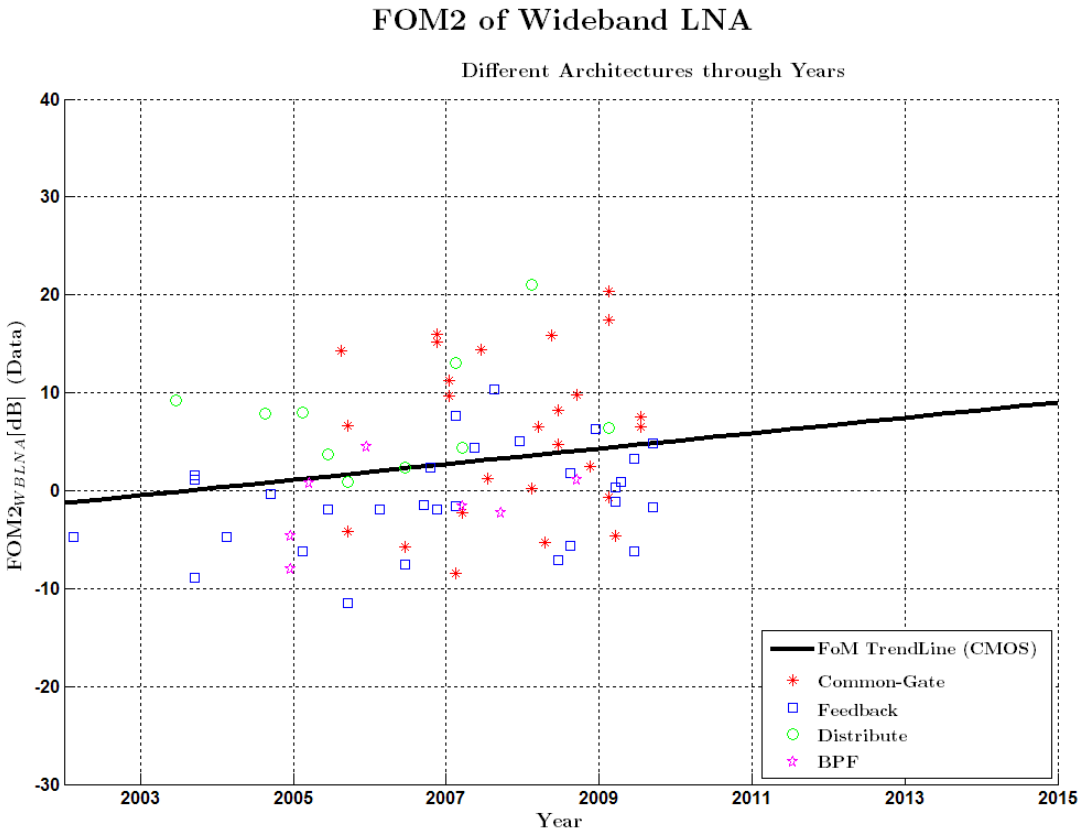


Figure 3-29 FoM tendency of wideband LNA (with IIP3) of different structures

As mentioned in the narrowband LNA section, the other trendlines and relationships among FoM, specifications and time scale can be referred to Appendix A-4, including the relationships between FoM vs. specifications (gain, noise, linearity and bandwidth), the specifications variation through years, as well as relationships and distributions among specifications. The power consumption is proportional to voltage gain and noise figure as for narrowband LNAs, resulting in almost zero-slope of FoM vs. gain and FoM vs. noise. The FoM vs. IIP3 slope is 0.7dB/dB and the FoM vs. BW slope is 4.3dB/dec. These two trends are comparable to those of narrowband LNAs, which are 0.6dB/dB and 5.2dB/dec, respectively.

The gain improves by about 3dB over 10 years, while noise figure and IIP3 remain almost the same. The bandwidth doesn't change a lot, mainly because the wideband LNA application is almost limited to TV and UWB receivers, which have fixed spectrum allocations. It is important to realize that these trendlines are determined not only by the technology improvement, but are also affected by the applications to some extend.

Again, the noise-gain, gain-IIP3 and frequency (BW) -noise trade-offs can also be observed roughly. It is worth comparing the parameters between narrowband and

wideband LNAs. The average voltage gains are similar for both LNAs, from about 12 to 22dB. The mean value of the narrowband LNA noise figure is 3.2dB while it is 4.1 dB for wideband LNAs, about 1dB higher than narrowband LNAs, as is expected theoretically. Both LNAs' IIP3 figures are similar as well. Most wideband LNAs' bandwidths are around 800MHz and 7GHz, corresponding to the TV tuner and UWB, for wideband LNAs, while for narrowband LNA, the frequencies are mostly found in cellular band, WLAN and Bluetooth ISM bands.

$FoM_{avg}[dB] = k_i \times year + b_i = 9.2 \times 10^{-1} \times year - 1839.5$	
$Spec_{i,avg} = n_i \times year + c_i$	$FoM/Spec_i = m_i$
$Gain_{avg}[dB] = 3.0 \times 10^{-1} \times year - 574.3$	$\frac{FoM}{Gain} = -6.6 \times 10^{-3}\ dB/dB$
$(F-1)_{avg}[dB] = 9.2 \times 10^{-2} \times year - 187.6$	$\frac{FoM}{F-1} = 3.9 \times 10^{-2}\ dB/dB$
$IIP3_{avg}[dBm] = 4.5 \times 10^{-2} \times year - 91.5$	$\frac{FoM}{IIP3} = 7.2 \times 10^{-1}\ dB/dB$
$BW_{avg}[Hz] = 10^{1.4\times10^{-2}\times year-18.3}$	$\frac{FoM}{BW} = 4.3\ dB/dec$

Table 3-3 Wideband LNA FoM statistics summary

The average FoM versus year, average specifications versus year and the relationship between FoM and specifications are summarized in Table 3-3. The predicted power consumptions are demonstrated in Figure 3-30, together with the collected data. Again, moderate, better and worse performances are combined to provide a general boundary for power consumption. According to the prediction, the average power consumption could be reduced to less than 3mW in the year 2015 with a moderate performance of 18dB gain, 4.1dB noise figure, -0.5dBm IIP3 and 5GHz bandwidth. This might become as low as 10mW even with higher performances of 21dB gain, 3dB noise figure, 3dBm IIP3 and 10GHz bandwidth. Similar to narrowband LNAs, sub-mW power consumption is expected with relaxed specifications.

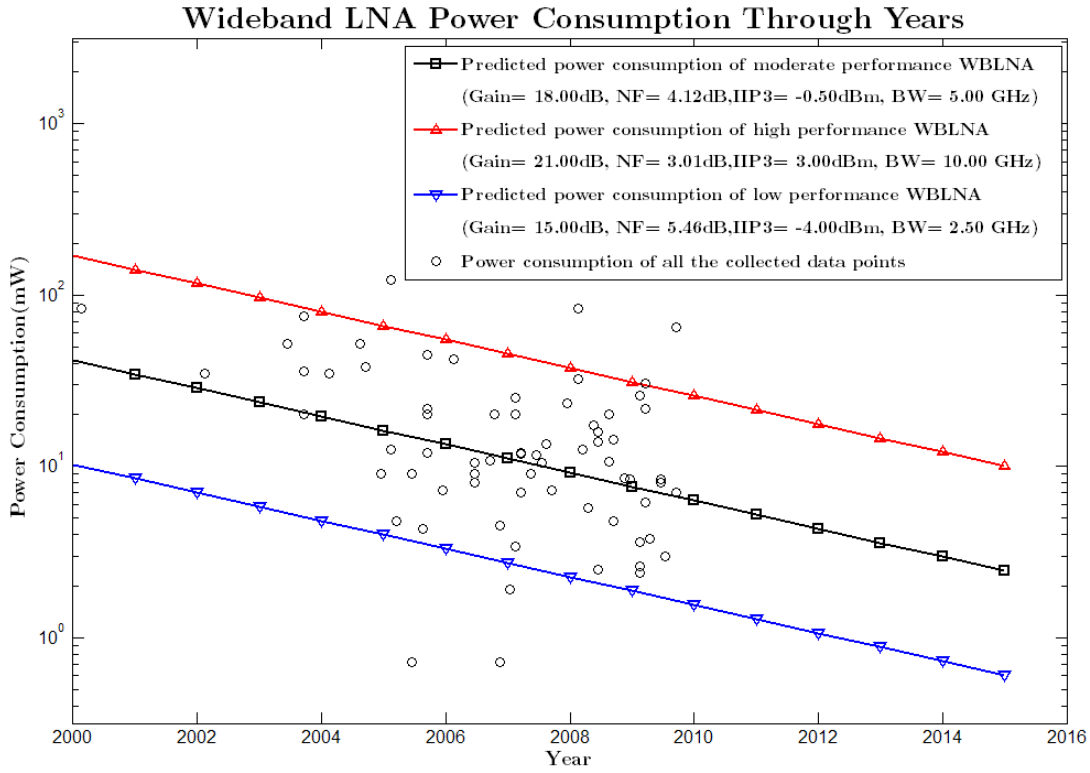


Figure 3-30 Prediction of wideband LNA power consumption through years

3.3.2 Mixer

The other very important block in the RF front-end is the mixer, which converts the RF signal to an intermediate frequency (IF) for further signal processing. This IF frequency could be baseband (in a zero-IF architecture), comparable with the signal bandwidth (in a low-IF architecture), just lower than the RF signal (in a superhet architecture), or two or three times higher than the RF signal (in a wideband up-conversion architecture).

A mixer is essentially a combination of an amplifier and current switches driven by the local oscillator. It can be classified as a passive or active mixer. A differential double balanced passive mixer consists of a transconductor, four cross-coupled connected MOS switches and an op-amp gain/filter stage. This type of mixer has been becoming popular in recent years because of the popularity of zero-IF and low-IF receiver architectures [46], but isn't suitable for other receiver architectures where the following stage is still an RF signal instead of baseband. Besides, the noise and linearity are highly dependent on the op-amp stage and the large LO driving level. What's more, the limited conversion gain provided by the op-amp reduces the

flexibility of using this kind of mixer in the receiver chain. On the other hand, the conventional current steering double balanced mixer shown in Figure 3-31 is more versatile in different architectures, and hence is investigated in this section. Besides, there are enough samples in publications for the purpose of predicting the FoM.

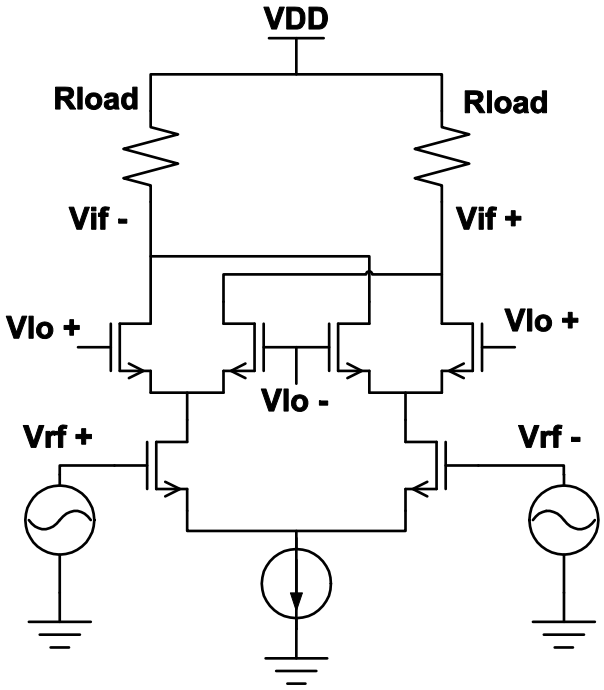


Figure 3-31 Gilbert Mixer

The FoM of a mixer can be defined in many different ways. The most popular one takes the dynamic range over power consumption [47]:

$$FoM_{mixer}[GHz] = \frac{IIP3[mW]}{F[abs]-1} \times \frac{1}{P[mW]}. \quad (3\text{-}24)$$

Other publications adopt the 1dB gain compression point as the linearity specification, which is usually about 10dB less than IIP3, and at the same time take the conversion gain into account [48]:

$$FoM_{mixer}[GHz] = \frac{Gain[abs] \times P1dB[mW]}{F[abs]-1} \times \frac{1}{P[mW]}. \quad (3\text{-}25)$$

In this thesis, however, in order to consider the receiver cascade analysis at system level, the same parameters as the LNA are involved in the FoM definition, which means that the operating frequency is added. Depending on whether the IIP3 specification is included or not, the FoM equations are given as follows, where conversion gain, DSB (double-sideband) noise figure and RF input frequency replace the voltage gain, noise figure and operating frequency in the LNA FoM definitions, respectively. Unlike the SSB (single sideband noise) figure, where the input noise sources could come from the image frequency as well, all the input noises are from the

converted channel frequency, as in an amplifier (ignoring harmonic mixing), hence F_{dsb}-*1* can more accurately reveal the actual noise generated by the mixer itself.

$$FoM1_{mixer}\left[\frac{GHz}{mW}\right] = \frac{CG[abs] \times F_{RF}[GHz]}{(F_{dsb}[abs] - 1)} \times \frac{1}{P[mW]}. \quad (3\text{-}26)$$

$$FoM2_{mixer}[GHz] = \frac{CG[abs] \times F_{RF}[GHz] \times IIP3[mW]}{(F_{dsb}[abs] - 1)} \times \frac{1}{P[mW]}. \quad (3\text{-}27)$$

Published data points giving the relevant figure of merit with and without the IIP3 specification are collected and presented along with predictions in Figure 3-32 and Figure 3-33. Although the mixer circuit structure is a little more complicated than an LNA, the FoM improvement trend is almost the same as for LNAs: to get a 3dB improvement, it takes 32.4 months for FOM1 and 35.9 months for FOM2. Note that the average FoMs of mixers are generally 15dB less than narrowband LNAs and about 11dB less than wideband LNAs.

The relationships between FoM and specifications (gain, noise, linearity and bandwidth), the variation of specifications through years, as well as relationships and distributions among specifications are presented in Appendix A-5. The FoM hardly changes with the conversion gain, implying again, the proportionality between gain and power consumption. Unlike in LNAs, there is an obvious FoM change of noise in terms of *F-1* in the mixer, which is about -0.3dB/dB. This means that the power consumption will increase by 59%, instead of 100%, with a 3dB lower *F-1*. And relaxing *F-1* by 3dB will only save 37% of the power consumption. This relatively low correlation between noise and power consumption is due mainly to the mixing process, which has a different noise mechanism compared with that of the transconductor differential pair stage. Similar to an LNA, a mixer's FoM improves with a higher IIP3, which is 0.27dB/dB, and so leads to 65% more power consumption with 3dB higher IIP3, and 58% less power consumption with 3dB less IIP3. Note that the IIP3 of the mixer is correlated with the power consumption more than that of the narrowband and wideband LNAs. The FoM vs. RF input frequency is about 7dB/dec, which is comparable with that of LNAs.

The conversion gain of active mixers increases by about 0.48dB per year on average. The noise *F-1* increases slowly and the IIP3 has been almost unchanged through years, as in a narrowband LNA, while the RF input frequency increases 14% every year, on average.

The conversion gain of mixers is normally distributed between 0dB and 13dB, which is much lower than LNAs. The noise, in terms of *F-1*, ranges from about 7dB to

18dB. Note that the noise figure and *F-1* are nearly the same for this high noise level. The largest noise sources are mainly due to white noise and flicker noise during the switching operation [49]. The IIP3 of a mixer is generally between -6.7dBm and +8.6dBm, comparable with that of an LNA. The input frequency ranges from 1GHz to 9GHz, which is pretty similar to LNAs, depending strongly on the application. The proportional gain vs. noise relationship and inverse proportional gain vs. IIP3 relationships can be observed roughly. However, within the standard variation ranges, any combination of these performances could be possible, according to the figures.

$FoM_{avg}[dB] = k_i \times year + b_i = 1.0 \times year - 2028.5$	
$Spec_{i,avg} = n_i \times year + c_i$	$FoM/Spec_i = m_i$
$Gain_{avg}[dB] = 4.7 \times 10^{-1} \times year - 951.9$	$\frac{FoM}{Gain} = -4.8 \times 10^{-2}\ dB/dB$
$(F-1)_{avg}[dB] = 3.4 \times 10^{-2} \times year - 56.1$	$\frac{FoM}{F-1} = 3.3 \times 10^{-1}\ dB/dB$
$IIP3_{avg}[dBm] = 6.6 \times 10^{-2} \times year + 133.4$	$\frac{FoM}{IIP3} = 2.7 \times 10^{-1}\ dB/dB$
$Freq_{avg}[Hz] = 10^{5.9\times10^{-2}\times year-109.2}$	$\frac{FoM}{BW} = 7.0\ dB/dec$

Table 3-4 Active mixer FoM statistics summary

According to FoM vs. year trendlines, FoM vs. performance, and specification vs. year trendlines, which are summarized in Table 3-4, together with the relationships and distributions among specifications as references, the predictions of mixer power consumptions through the years are calculated and demonstrated in Figure 3-34. The moderate, high and low performance of conversion gain, double sideband NF, IIP3 and RF input frequency are combined to provide the power consumption boundaries. It is clearly shown that most published mixers' power consumptions are within the boundaries and have the very similar trend through years. This supports the effectiveness of the FoM methods applied to predicting the active mixer power consumption.

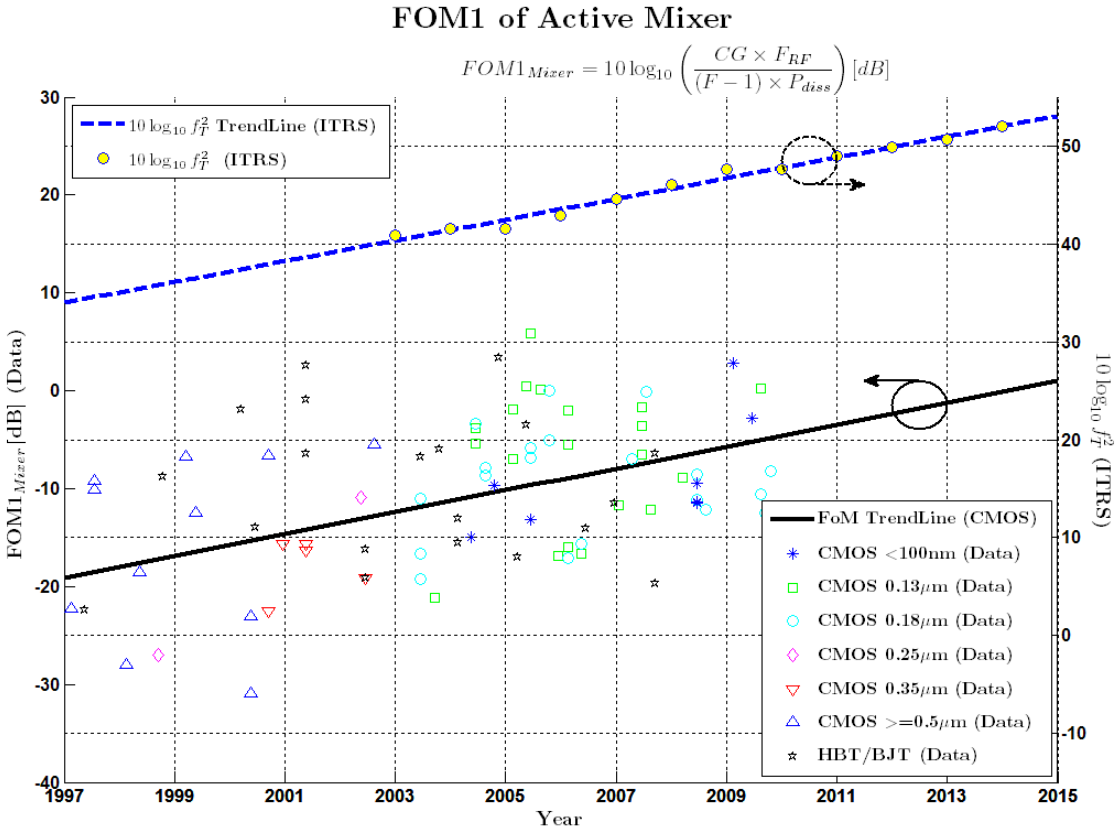


Figure 3-32 FoM tendency of active mixer (without IIP3)

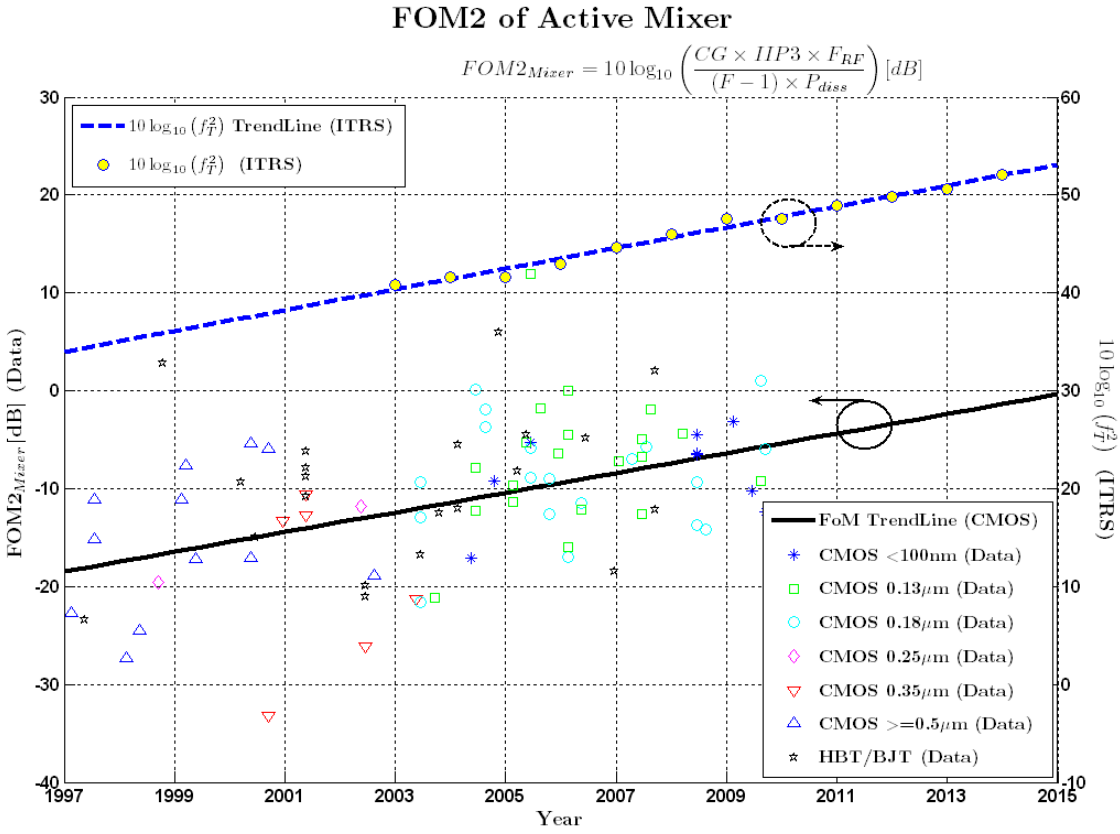


Figure 3-33 FoM tendency of active mixer (with IIP3)

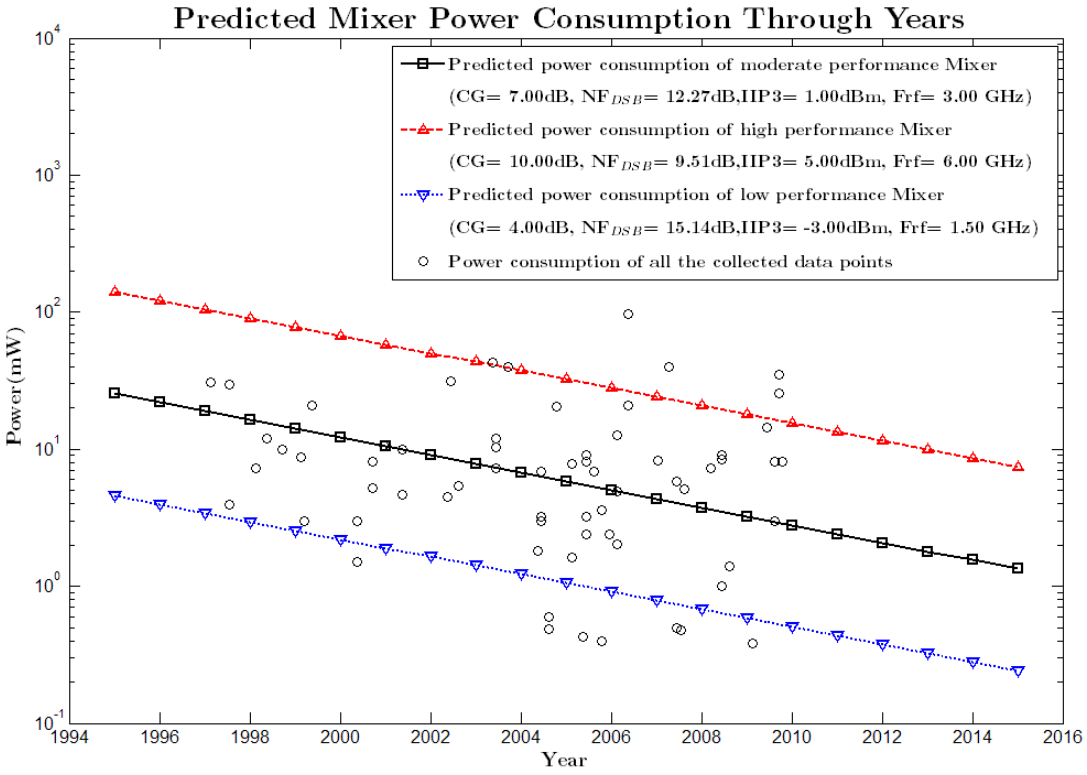


Figure 3-34 Prediction of active mixer power consumption through years

3.3.3 Voltage Controlled Oscillator

A stable and local oscillator is one of the essential factors in a receiver architecture, where the nominal tuning frequency is set by means of a phase locked loop synthesiser. A voltage controlled oscillator (VCO) is the most important block in the local oscillator subsystem. Parameters involved in VCO include operating frequency, phase noise at specified frequency offsets, tuning range and power consumption. Among these specifications, the most important one for the VCO for RF applications is the phase noise, which may degrade the receiver SNR performance through reciprocal mixing [30]. Therefore, the design challenge of the VCO is to minimize the phase noise while minimizing the power consumption.

There are two main distinct VCO architectures available: LC oscillator and ring oscillator. In this section, the FoMs of LC and ring oscillators are collected and predicted separately. However, a widely used definition of FoM is applied for both architectures [50].

$$FoM_{VCO}\left[\frac{1}{J}\right] = \left(\frac{f_0[Hz]}{f_m[Hz]}\right)^2 \times \frac{1}{L\{f_m\}\,[1/Hz]} \times \frac{1}{P_{diss}[mW]}. \qquad (3\text{-}28)$$

Parameters f_0, f_m and $L\{f_m\}$ are the centre frequency, the offset frequency and the phase noise at offset frequency. The centre frequency is in the numerator for the same reason as in the LNA and Mixer definitions, namely that more power consumption is needed to achieve higher bandwidth for active devices. Ignoring flicker noise near the centre frequency and the noise floor far away from centre frequency, the VCO feedback system converts the white noise from the active device to phase noise, which decreases by 20dB/dec with offset frequency [51]. As will be shown in the following analysis, for both the LC-VCO and ring VCO, trade-off exists between phase noise and power consumption; therefore, the product of these parameters is in denominator in both of the definitions.

Apart from the FoM definition in Equation (3-28), the tuning range of the oscillator is sometimes included in the FoM definition[52] as expressed in Equation (3-29).

$$FoM2_{VCO}\left[\frac{1}{J}\right] = \left(\frac{f_{0,max}[Hz] - f_{0,\min}[Hz]}{f_m[Hz]}\right)^2 \times \frac{1}{L\{f_m\}[1/Hz]} \times \frac{1}{P_{diss}[mW]}. \quad (3\text{-}29)$$

For LC oscillators, a large tuning range usually depends on the availability of high quality MOS varactors or capacitor arrays with digital switches, which do not necessarily reflect the LC circuit's own performance. And wide tuning range is an inherent feature for ring oscillators, as will be explained in detail later. Therefore, in this project, tuning range is not considered as a FoM parameter.

LC Oscillator

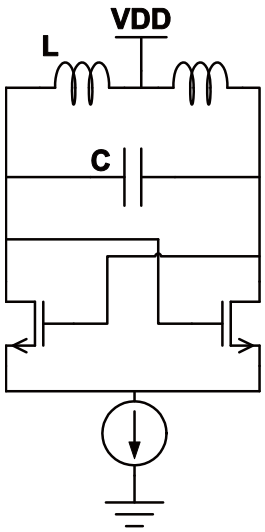


Figure 3-35 LC oscillator

An LC oscillator as shown in Figure 3-35 can be treated as the combination of a lossy LC resonator and an amplifier providing negative resistance to compensate for

the resistive loss in the LC tank. Essentially, this kind of oscillator amplifies broadband white noise and performs noise shaping by the filtering of the LC resonator. Therefore the amplifier's gain must be large enough to start-up the oscillator. The output sideband phase noise power level at a certain offset from the centre frequency is determined by two factors: the filtering of the LC resonator and the noise sources from the amplifier and resonator.

Phase noise has been shown to be inversely proportional to the Q factor of the LC tank in [53], [54]. At circuit level in this section, a fully differential LC VCO biased by a tail current and loaded with on-chip components is selected as the default structure. In this circuit, the main noise sources come from white noise of the differential pair and the tail current source. All of these noise sources are proportional to the oscillating frequency and inversely proportional to the Q factor of the resonator and the output voltage swing[55], which can be expressed as:

$$L\{f_m\}\,[1/Hz] = \left(\frac{f_0[Hz]}{f_m[Hz]}\right)^2 \times \frac{kT \times (1+\gamma)}{Q_{LC}^2 \times P_{sig}}. \tag{3-30}$$

The parameter γ is the noise coefficient, which is about 2/3 for long channel devices and 2~3 for short channel devices. The Q factor of the LC tank is determined by the Q factor of the inductor and the capacitor, where the inductor's loss dominates within the frequency range of most common communication systems. Note that the inductor's quality factor, Q_L, is a frequency dependent-variable itself ($Q_L=\omega L$). This might be substituted into equation (3-30) as well, which is true for investigating an inductor at different frequencies. However, from the design and technology improvement point of view, it is more meaningful to investigate the *achievable* phase noise (and hence the *achievable* Q_L according to the above equation) at a certain centre frequency and offset frequency. Therefore, Q_L should be seen as a function of year, instead of a function of frequency in this FoM investigation. In fact, ITRS also predicts the Q_L of a 1nH inductor at a fixed frequency of 5GHz through the years. The output signal voltage swing is directly related to the supply voltage, which means that the signal power, P_{sig}, is proportional to the oscillator's power consumption. Therefore, by combining Equation (3-28) and Equation (3-30), the FoM as a function of technology parameters can be obtained as follows:

$$FoM_{LCVCO}\left[\frac{1}{J}\right] \propto \frac{Q_L^2}{kT \times (1+\gamma)}. \tag{3-31}$$

The figure of merit data points from published designs and the associated trendline are illustrated in Figure 3-36. In 1998, most VCOs were implemented using

0.35μm, the figure of merit was 0.21×10^{21} /J, and it improves to about 4.12×10^{21}/J in the year of 2010, when 0.13μm CMOS and later technologies dominated the published results. This is predicted to reach 8.26×10^{21}/J by the year of 2015. According to the fitted trendline, it takes roughly 60 months for the FoM of an LC VCO to improve by 3dB, which is quite similar to the rate of improvement in the reports of the inductor's quality factor.

When comparing the phase noise of oscillators, one could simply use the phase noise at a fixed frequency offset, e.g. 1MHz, which is a conventional way for engineers, and the phase noise is indeed measured at 1MHz offset from the centre frequency in most publications to make the comparison. However, the corresponding judgements are only valid when the centre frequencies of the oscillators are also the same. As shown in equation (3-30), with a fixed offset frequency, Δf, the phase noise is bound to increase with higher centre frequency even with all the other performances the same, including the same Q factor. Therefore, a better comparison method is introduced here, that is, using the *normalized offset frequency*, Δf_{norm}, which is defined as the ratio of the absolute offset frequency to the centre frequency. Assume there are two oscillators running at 1GHz and 5GHz respectively, and that the phase noise is measured at offset frequencies of 200kHz and 1MHz, respectively. If their phase noises are at the same level, we could say that their performances are the same. The *normalized offset frequency* is arbitrarily set to $\Delta f_{norm}=2\times10^{-4}$ (1MHz offset from 5GHz centre frequency) in this chapter. All the phase noise measurement of the collected data points are modified accordingly with a -20dB/dec rate (only suitable for white noise induced phase noise). For example, if a 1GHz oscillator's phase noise is measured as -120dBc/Hz at 1MHz offset, then the phase noise at Δf_{norm} (the absolute value is 200kHz in this case) could be estimated to be -120dBc/Hz-$20\times\log_{10}$(1GHz/5GHz) = -106dBc/Hz, so that a comparison can be made with any other oscillator with the same normalized offset frequency. Note this doesn't mean that the oscillator actually achieves -106dBc/Hz at 200kHz, where the flicker noise could have already taken effect. Instead, this is a useful comparison method between different oscillators running at different frequencies.

The figures of FoM vs. performances, specifications vs. years and relationships between specifications can be seen to Appendix A-6. In particular, the FoM vs. phase noise at Δf_{norm} is repeated in Figure 3-37, and the linear fitted trendline shows that the FoM decreases by 0.9dB with every 1dB increase in phase noise. This result reveals the

fact that the correlation between phase noise at Δf_{norm} and power consumption is very weak. This observation can be supported by a theoretical explanation, that is, the phase noise is determined mainly by the inductor's quality factor, which has little relationship with power consumption. Further observation shows that a very similar relationship exists for different technologies. The distinguishable trends can be noticed for 0.35µm, 0.25µm, 0.18µm and 0.13µm, respectively, while data points using older technologies generally have a lower FoM compared to newer technologies.

On average, the phase noise decreases by 0.4dB and the frequency increases by 24% every year, according to the linear fitted trendline. In these figures, the changes in the adopted technologies can be observed clearly.

In terms of the specifications, VCOs can achieve a normalized phase noise from -125dBc/Hz to -110dBc/Hz. The frequencies of the data points, which are highly dependent on existing communication systems, are mostly found between 900MHz and 25GHz. Note that the VCO sometimes needs to reach a frequency of twice the signal frequency in order to generate I/Q signals. As the frequency goes higher, older technologies are gradually replaced by newer technologies. In addition, it can be roughly found that, for a fixed frequency, the phase noise is generally lower for a newer technology.

$FoM_{avg}[dB] = k_i \times year + b_i = 0.6 \times year - 998.8$	
$Spec_{i,avg} = n_i \times year + c_i$	$FoM/Spec_i = m_i$
$PN@\Delta f_{norm_{avg}}[dB] = -3.7 \times 10^{-1} \times year - 628.2$	$\frac{FoM}{PN@\Delta f_{norm}} = -0.92\ dB/dB$
$Freq_{avg}[Hz] = 10^{9.5\times10^{-2}\times year - 179.8}$	$\frac{FoM}{Freq} = -1.31\ dB/dec$

Table 3-5 LC-VCO FoM statistics summary

The relationships between FoM, performances and time scales of LC-VCOs are summarized in Table 3-5. The LC-VCO power consumption is predicted through the years, with moderate, high and low performances, as demonstrated in Figure 3-38. Note that the centre frequencies are selected as 6GHz for all of the three performances, and hence the offset frequency could also be selected as the same, which is 1MHz in this demonstration figure. The corresponding phase noises are -117dBc/Hz, -125dBc/Hz and -109dBc/Hz, respectively. It can be seen that the power consumptions of most of the collected data points are within the predicted boundaries and have a similar trend, validating the FoM method of estimating power consumption for an LC-VCO.

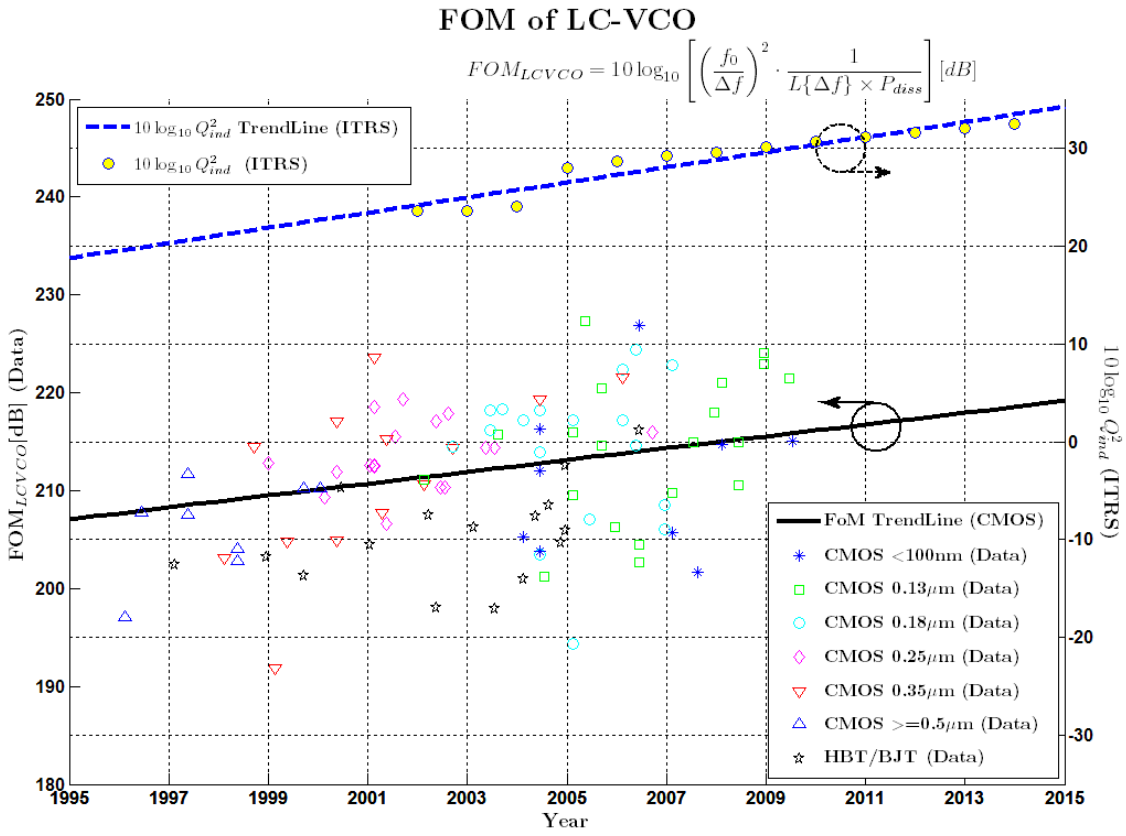


Figure 3-36 FoM tendency of LC oscillator

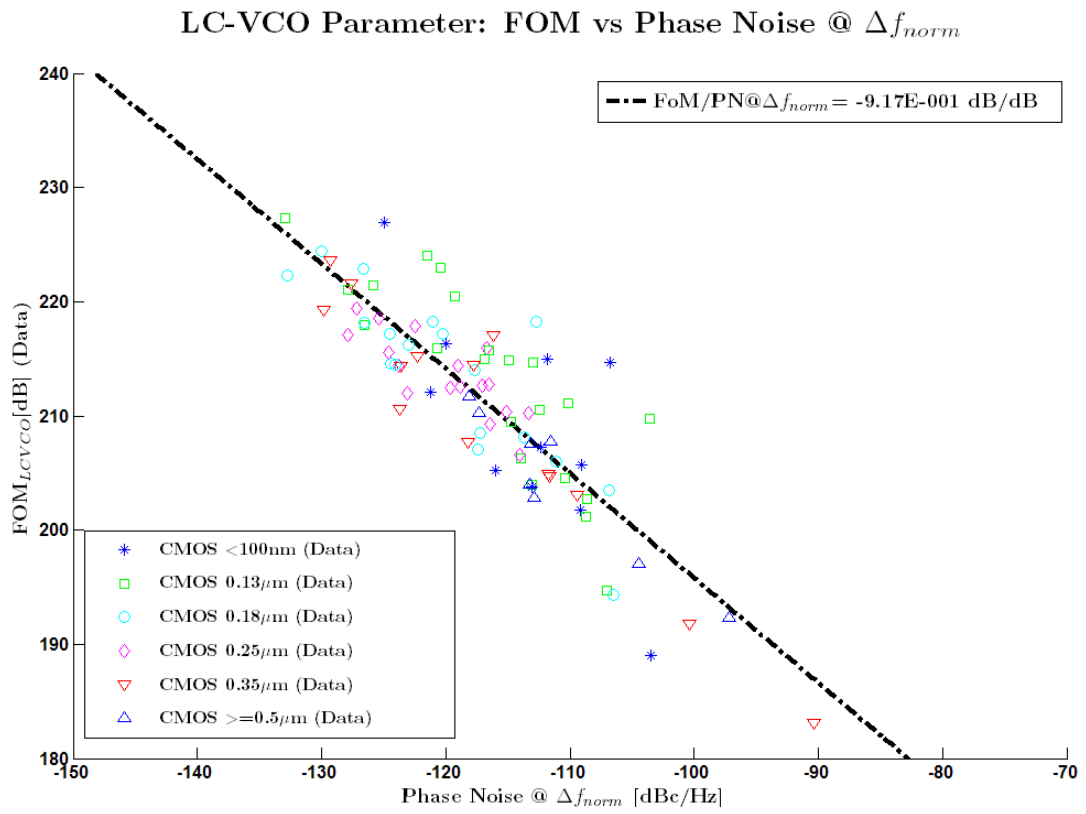


Figure 3-37 LC oscillator parameters: FoM versus phase noise

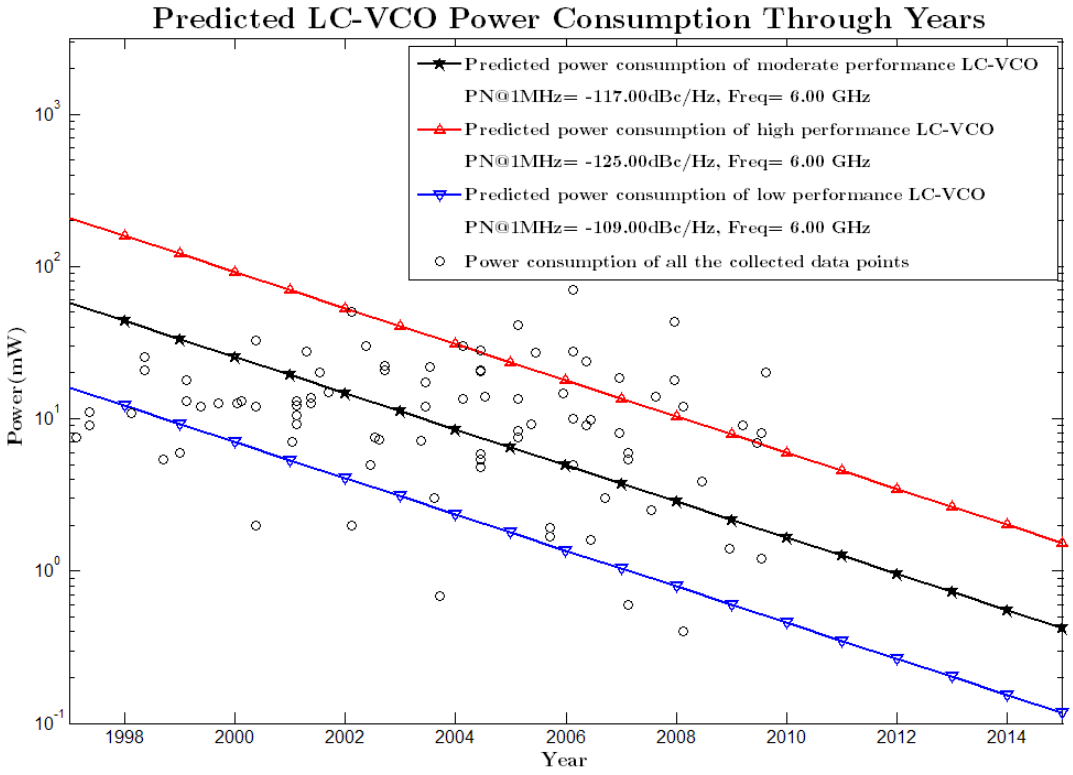


Figure 3-38 Prediction of LC-VCO power consumption through years

Ring Oscillator

Besides the LC VCO, another important type of VCO is the ring oscillator. Unlike the LC VCO, the ring oscillator consists of a series of connected inverters or amplifiers to provide positive feedback. The operating frequency is determined by the delay per stage and the number of stages. Figure 3-39 shows the basic ring oscillator architecture and two different delay cell circuits.

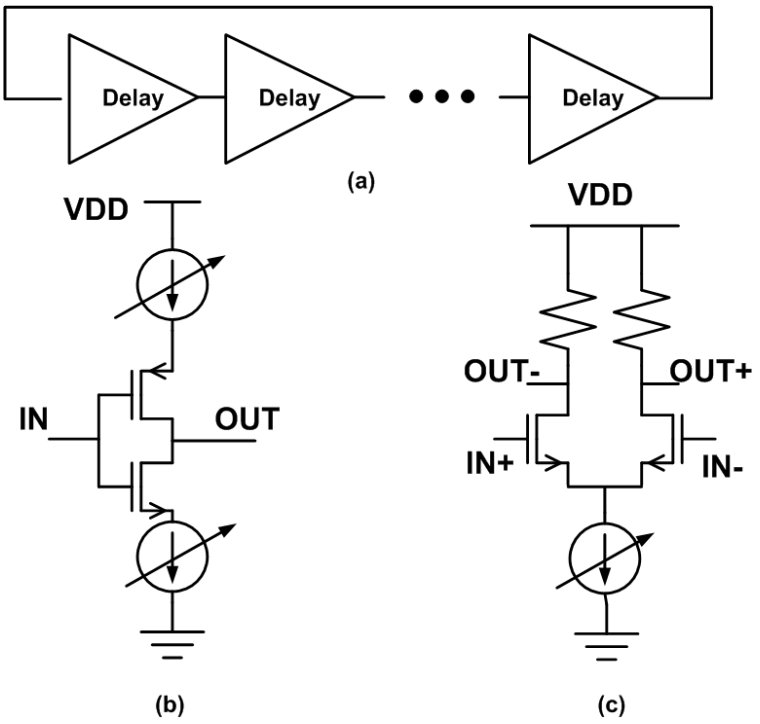


Figure 3-39 Ring Oscillator (a) Structure, (b) current starved CMOS inverter delay cell, (c) Differential amplifier delay cell

With a fixed number of delay stages, the tuning range of a ring VCO is roughly governed by the ratio of the transconductance over the load capacitance of each stage. The transconductance can be increased by increasing the current (for example, increase the VDD for an inverter type delay cell, or increase the tail current for a differential amplifier delay cell). Hence the tuning range could be much wider than that of an LC-VCO, which is mainly tuned by the load capacitance. Even a ring-oscillator tuned by an MOS varactors load (as in an LC-VCO) is able to obtain a wider tuning range than an LC VCO because:

$$f_{tune\_ring} \propto \frac{g_m}{C_{load}}. \tag{3-32}$$

$$f_{tune\_LC} \propto \frac{1}{\sqrt{L \times C}}. \tag{3-33}$$

Hence, if very wide tuning range is required, the ring oscillator could be a simpler choice rather than an LC-VCO. However, the phase noise of a ring oscillator is much higher than that of an LC oscillator. As analyzed in [56], if we only consider the phase noise due to white noise, the SSB phase noise can be obtained for CMOS inverter and differential types in Equation (3-34) and (3-35), respectively.

$$\mathcal{L}(f_m) = \frac{2kT}{I}\left(\frac{1}{V_{DD} - V_t}(\gamma_N + \gamma_P) + \frac{1}{V_{DD}}\right) \cdot \left(\frac{f_0}{\Delta f}\right)^2. \tag{3-34}$$

$$\mathcal{L}(f_m) = \frac{2kT}{If_0 \ln 2}\left(\gamma\left(\frac{3}{4V_{effd}} + \frac{1}{V_{efft}}\right) + \frac{1}{V_{op}}\right) \cdot \left(\frac{f_0}{\Delta f}\right)^2 . \quad (3\text{-}35)$$

Note that V_{effd} and V_{efft} are overdrive voltage of the differential pair and tail current device, which are constant to some extent, and that V_{op} is the signal swing range, which usually depends on the tail current and load resistance for a differential pair amplifier. The parameter γ is the MOSFET's noise coefficient, as discussed in earlier sections.

It can be found that the phase noise of a ring oscillator is highly dependent on the current, and hence the power consumption, for a fixed supply voltage. More current flow reduces the phase noise. Furthermore, it can also be observed from these equations that the phase noise is strongly correlated with the power consumption ($I \times V_{DD}$) for inverter-based ring oscillators. For differential amplifier type ring oscillators, the correlation between phase noise and signal output power ($I \times V_{op}$) can also be observed to some extent, which is in turn related to the power consumption. Therefore, the phase noise of ring oscillators can be estimated very roughly to be proportional to the product of the power consumption and the term $(f_0/\Delta f)^2$. Hence, by substituting the phase noise into Equation (3-28), an almost constant FoM can be obtained.

The data collection and fitted trendline of the FoM for ring oscillators is illustrated in Figure 3-40. As can be seen in the figure, the FoM in different technology keeps almost the same level through the years at around 183.5dB. Figure 3-41 illustrates the phase noise at a normalized offset frequency ($\Delta f_{norm}=2\times10^{-4}$) versus oscillation frequency, of which the average value is about -93dBc/Hz, generally over 20dB higher than the LC counterpart. The normal operating frequency is also lower than that of an LC-VCO. A ring oscillator is not good enough for most radio system local oscillator applications due to its relatively higher phase noise than that of an LC-VCO. However, it could be a candidate for fast, low cost, low power spectrum scanning if the phase noise requirement is not very tough. This will addressed in later chapters.

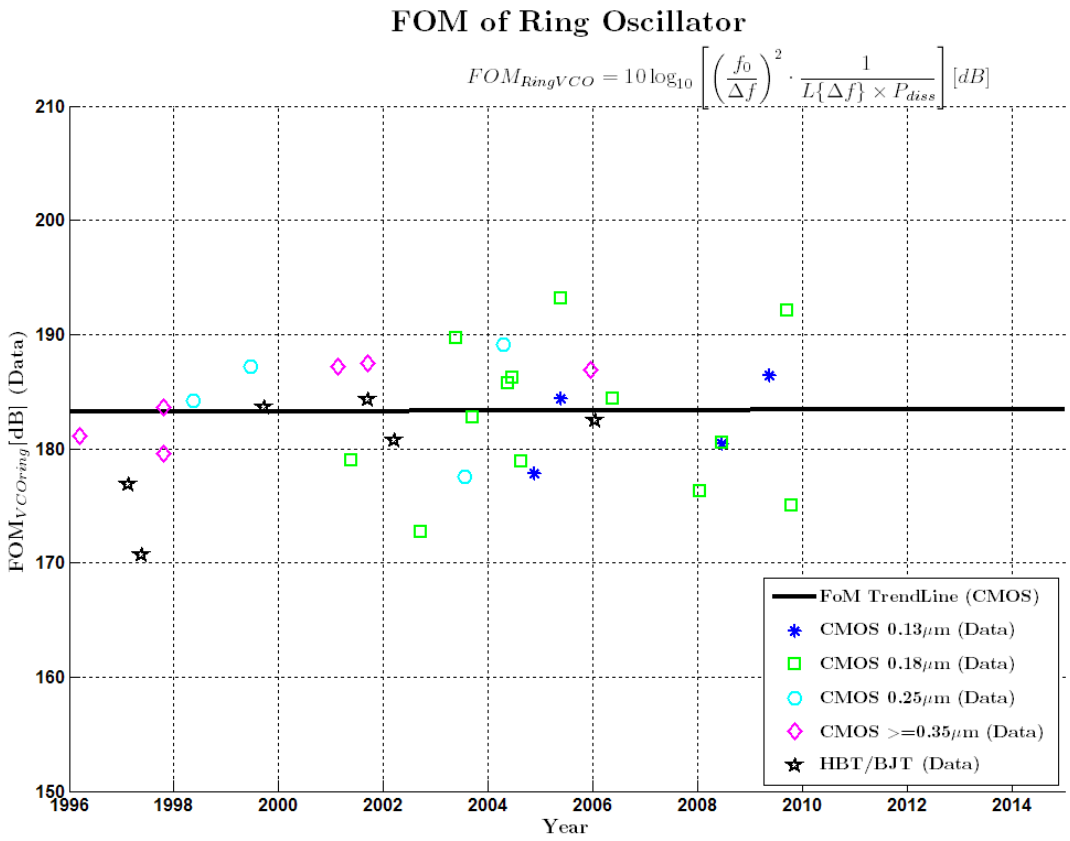


Figure 3-40 FoM tendency of ring oscillator

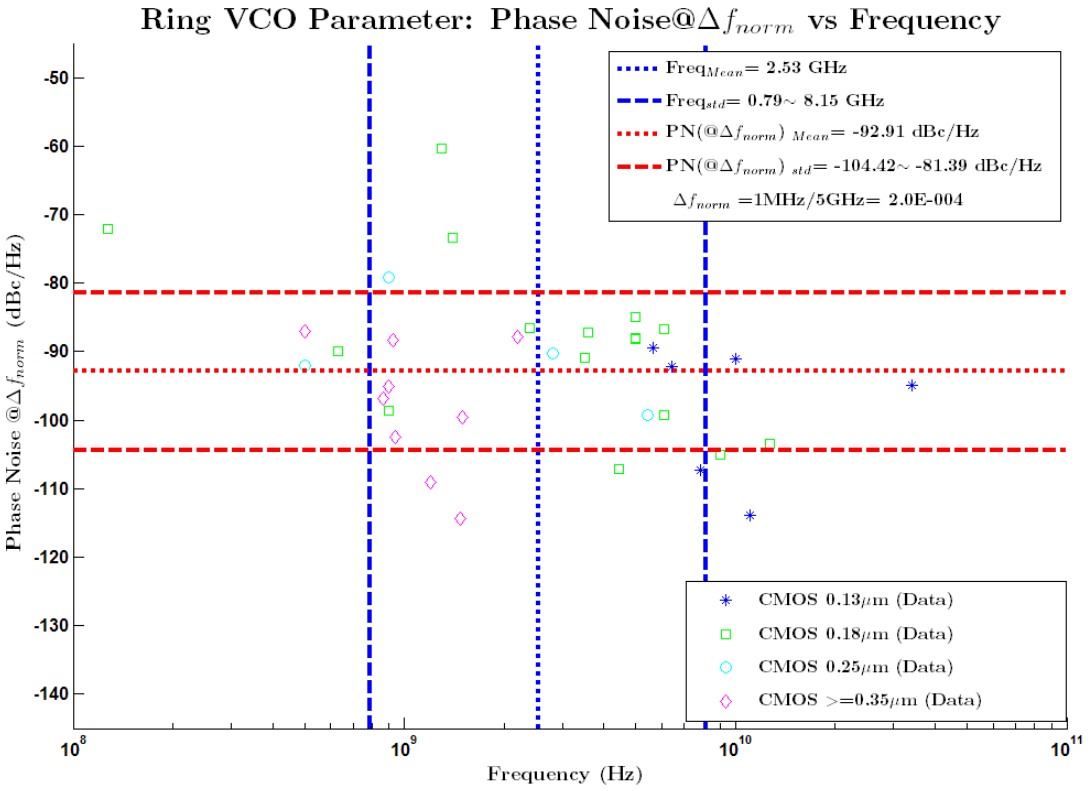


Figure 3-41 Ring oscillator parameters: phase noise versus frequency

3.3.4 Frequency Divider

The frequency divider is the other RF block in a frequency synthesizer apart from the VCO. High speed dividers are usually placed following the VCO as frequency pre-scalars before further lower speed logic implementing integer-N or fractional-N division blocks. Two of the most popular frequency dividers are D-flip-flop based and injection-locked, as shown in Figure 3-42. In the frequency synthesizer design for radio communication applications, D-flip-flop based frequency dividers are usually implemented by differential current mode logic (CML) circuits at RF frequencies instead of conventional rail-to-rail digital CMOS latches [57]. In an injection-locked divider, a free-running LC or ring oscillator, where the resonant frequency is approximately at a multiple or a sub-multiple of the incoming signal, locks to the injected signal in phase and frequency [58]. Generally, CML based static frequency dividers achieve relatively lower frequency and higher power consumption compared to injection-locked dividers due to the necessary charging and discharging or the load capacitance, but benefit from a wider locking range. On the other hand, injection-locking dividers can achieve higher frequency with lower power consumption than CML dividers, since the operation is essentially an oscillator and, for the same reason, the locking range is limited around the free-running frequency and could be very narrow for a high Q inductor.

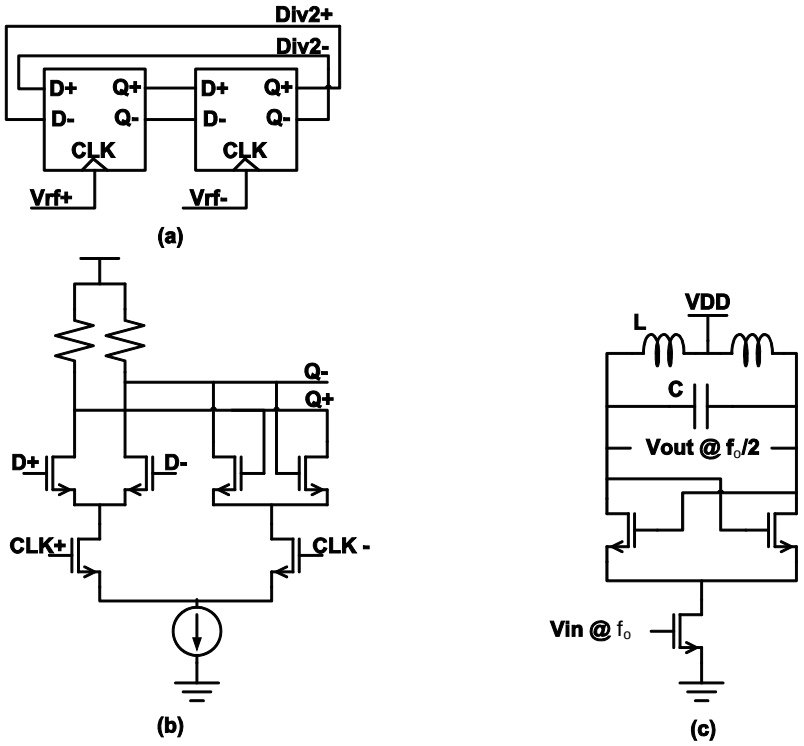


Figure 3-42 Divider (a) Static DFF based (b) CML D-latch (c) Injection-locked divider

There are different FoM definitions for frequency dividers in the literature. For CML dividers, power gain and speed are usually concerned [59], and for injection-locked dividers, the highest frequency and locking range are often compared [60]. The locking range is also a specification that is used to compare the dividers frequently. In fact, the most important performance requirements are the highest achievable frequency and the power consumption. Therefore, the FoM of a frequency divider for both architectures is defined as:

$$FoM_{Div}\left[\frac{GHz}{mW}\right] = \frac{f_{inmax}[GHz]}{P_{diss}[mW]}. \tag{3-36}$$

Note that the parameter f_{inmax} refers to the maximum operating frequency of the divider. In the case of CML dividers, the maximum achievable frequency is directly proportional to the transition frequency f_T in the technology roadmap. For injection-locked dividers, the speed of the active devices also benefits from a higher f_T.

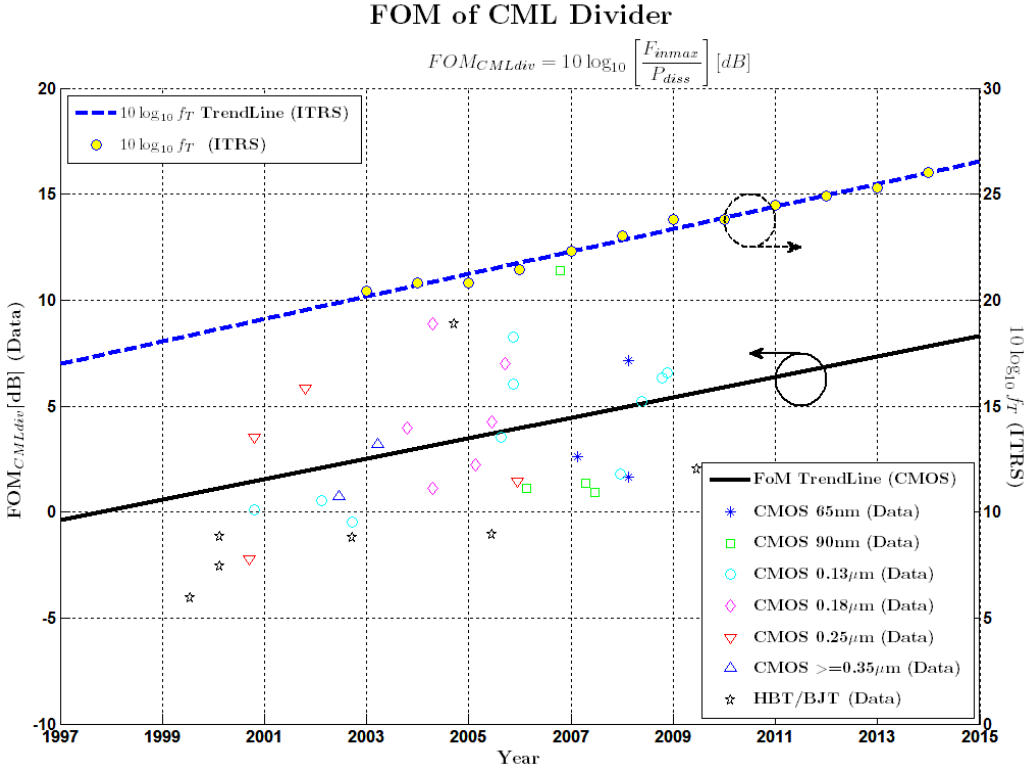


Figure 3-43 FoM tendency of CML frequency divider

Figure 3-43 and Figure 3-44 illustrate the FoM data points from the literature with the predictions. It can be observed in the figures that the CML divider's FoM improves by 3dB within about 75 months, while it takes injection-locked dividers only 45 months to achieve a 3dB improvement. The overall FoM of injection-locked dividers is higher than CML dividers, and generally can achieve maximum frequencies

that are about 50% higher. The maximum frequency distribution of CML and injection-locked frequency dividers can be found in Appendix A-7.

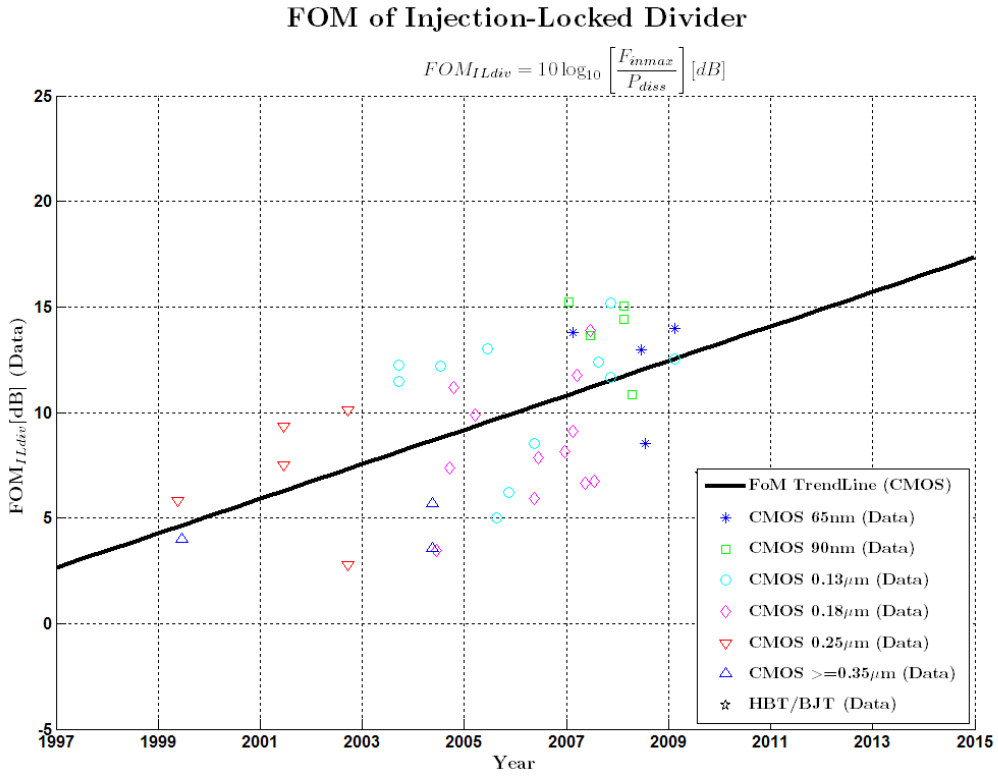


Figure 3-44 FoM tendency of injection-locked frequency divider

3.3.5 Baseband Blocks (LPF and VGA)

The baseband low pass filter (LPF) and variable gain amplifier (VGA) are usually the last stages in the receiver chain before the signal goes into the ADC. The low pass filter selects the channel and the variable gain amplifier provides gain or attenuation in order to scale the signal to the ADC's dynamic range.

The most commonly adopted type of low pass filter is a continuous-time filter, which can achieve high bandwidth and also act as an anti-alias filter for the ADC. There are mainly three different architectures to realize continuous-time filters: transconductance-C filter [61], MOSFET-C filter [62] and active-RC filter [63]. All of these architectures involve op-amp or transconductor circuits and capacitor (array) for bandwidth tuning. The baseband gain is provided either by the low pass filter itself, or by a following baseband amplifier, which usually has a similar structure to the low pass filter. Therefore, sometimes, they are designed together. Take an example of the baseband filter and gain stage shown in Figure 3-45.

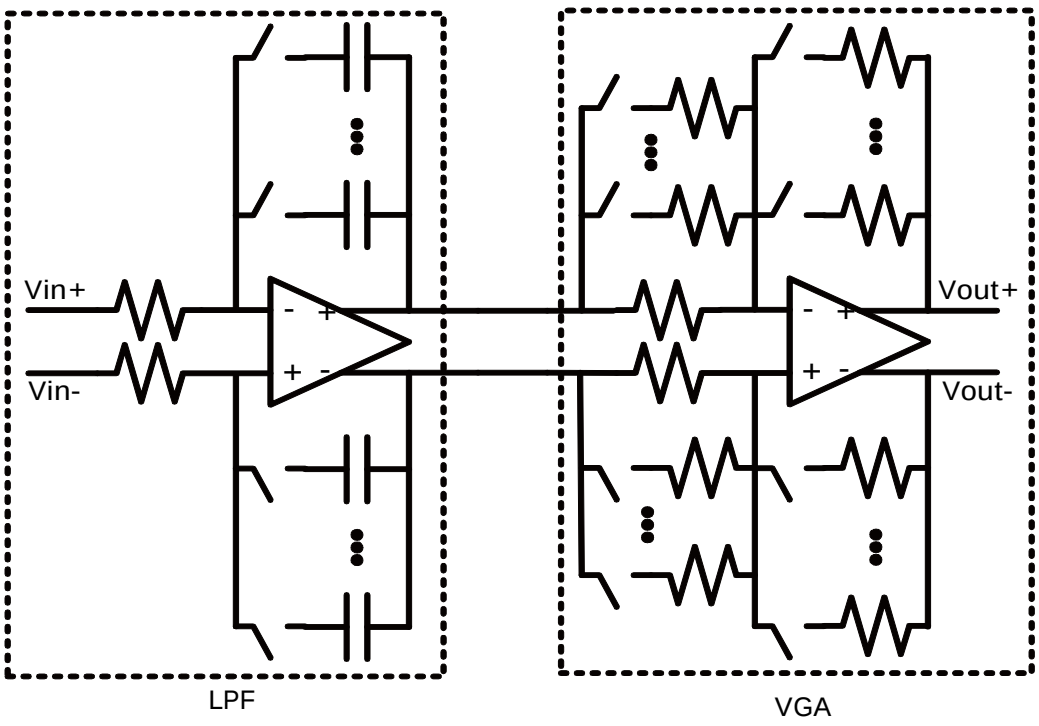


Figure 3-45 Example: Active RC filter + VGA

In this project, the low pass filter and variable gain amplifier are investigated separately. Because of their similar circuit architectures, the main limitations on the bandwidth, noise and linearity are imposed by the op-amp in both cases. Therefore, a simplified strategy is adopted for these two blocks: the low pass filter provides the bandwidth, non-linearity and noise, while the VGA provides voltage gain. Because the bandwidth of the VGA is usually dominated by a single pole, it is better to ensure that this pole is high enough with respect to the filter's cut-off frequency to avoid any signal attenuation within the band. In this way, the bandwidth and gain of the combination can be determined by the filter and the amplifier, respectively. Usually, the noise of the VGA is not as important as that in the other blocks, because its noise contribution to the entire front-end chain is usually suppressed enough by the LNA and other gain blocks. After the received signal goes through the baseband lowpass filtering stage, the interferers' level is usually attenuated sufficiently so that the linearity requirement is also not very high. Besides, the lack of full measurement results for published VGA designs makes it very difficult to collect enough data points to draw an accurate FoM trendline with noise and linearity specifications involved. Therefore, for a rough estimation on the performance of the receiver chain, only the gain and bandwidth specifications are considered in the FoM calculations.

Baseband LPF

The most popular FoM definition for a low pass filter is expressed as [62]:

$$FoM_{LPF}\left[\frac{1}{J}\right] = \frac{Order \times Fc[Hz] \times SFDR[abs]}{P_{diss}[mW]}. \quad (3\text{-}37)$$

The filter order is sometimes proportional to the number of op-amps (e.g. full differential biquad filters or leapfrog type filters), and hence in proportional to the total power consumption although, in other cases, this is not entirely true (for example, a two pole low pass transfer function could be obtained from a basic Sallen-Key structure, which consists of only one op-amp). The *Fc* is the cut-off frequency, which is normally the 3dB bandwidth for Butterworth response for example. The SFDR (spurious free dynamic range) is related to the linearity and integrated input-referred noise of the circuit [64]:

$$SFDR[dB] = \frac{2}{3} \cdot (IIP3[dBm] - Noise[dBm]). \quad (3\text{-}38)$$

IIP3 is often given using a scale of dBm or dBV. If the total output harmonic distortion (THD) is given, the method in [65] can be used to obtain approximately the corresponding IIP3, which can be found in Appendix A-8 of this thesis.

In a baseband LPF, the integrated noise is directly related to the bandwidth. According to the preceding discussion, when adopting FoM vs. performance relationships for estimating the power consumption, lower correlations among the specifications are preferred. Therefore, noise and frequency should ideally be de-coupled. In the data collected from the literature presented in this chapter, the bandwidth should be de-embedded from the integrated noise, leaving the noise power density as the reference specification. In the case that a published design quotes the noise power spectral density (*nV/sqrt(Hz)* or V^2/Hz), it is adopted directly. In the case that the integrated noise power over the pass-band is given, the noise density is calculated as below (note that this is only for low pass filtering):

$$Noise\left[\frac{dBV}{Hz}\right] = 10\log\left(\frac{Noise[V^2]}{Fc[Hz]}\right). \quad (3\text{-}39)$$

This noise density is an average value that includes both white noise and flicker noise. The filter is usually at least 3rd or 4th order in a practical design, so that the noise bandwidth can be assumed to be the same as the 3dB cut-off bandwidth.

Now, by substituting equation (3-38) and (3-39) into (3-37), the FoM of a low pass filter can be expressed as:

$$FoM_{LPF}\left[\frac{1}{J}\right] = Order \times \left(\frac{IIP3[mW]}{Noise\ Density\ \left[\frac{mW}{Hz}\right]}\right)^{\frac{2}{3}} \times \frac{(Fc[Hz])^{\frac{1}{3}}}{P_{diss}[mW]}. \quad (3\text{-}40)$$

Note that this equation is equivalent to the FoM definition in equation (3-37), and the label in the figure showing the FoM trend will still quote the original FoM definition. In this expression, however, the filter order, IIP3, noise density and frequency are relatively independent of each other and hence it is more suitable for developing a power consumption prediction strategy.

The FoM of the baseband filter over the years is summarized as shown in Figure 3-46 with different technologies and in Figure 3-47 with different architectures. It can be observed that the FoM of a low pass filter improves by 3dB within about 23 month according to the figures. This is over 50% faster than RF blocks such as LNA and mixer, for which the 3dB improvement takes about 30~40 months. It also reveals that there is no obvious preferable architecture from the FoM point of view, although active-RC filters are becoming popular in recent years because of their generally higher linearity than other architectures, which is more demanded by recent communication systems.

The figures of FoM vs. performances, performances vs. years and relationships among performances are shown in more detail in Appendix A-9. The linear fitting technique is applied to obtain the trendline for filters implemented in CMOS technology. The FoM increases about by 0.6dB for every 1dB IIP3 improvement and by about 0.6dB for 1dB noise density reduction. This implies a trade-off between linearity and noise. For example, from the point of view of the input signal with an opamp-RC architecture, by increasing the input resistors so that there is less input current flowing, the capacitors can be reduced accordingly. This effectively improves the filter's linearity because less output current is needed (and therefore less distortion), while the noise performance gets worse (i.e. the noise is proportional to kT/C). The FoM improves slightly, by 1.2dB, for every 10 times higher cut-off frequency. This value is much lower than that of LNA and mixer, implying that the bandwidth is more related to power consumption. The linear fitted trendline also suggests that the FoM changes with a rate of -1.3dB per order, which means the efficiency of power consumption being reduced with higher order. The variation is expected to be due to the number of op amps per pole is not the same for different architectures such as the Sallen-Key type blocks as well as the Biquad and leapfrog filter.

According to the figures, the IIP3 and noise density vary by +7.9dB and -7.2dB over 10 years, respectively. Because the low pass filter is a baseband circuit, flicker noise is a significant parameter in the whole noise specification. According to the ITRS, the flicker noise reduces through the years, and this is one of the reasons for the lower noise level. The cut-off frequency, on the other hand, doesn't change a lot through the years. This is because the signal bandwidth of commonly used communication systems hasn't changed a lot and is normally below 20MHz.

In terms of the relationships among specifications, the IIP3 and Noise density distribution and relationship for all the collected data points are particularly shown in Figure 3-48. Note that these two parameters are converted to power (dBm) instead of voltage (dBV), referred to a 50Ω resistor. This makes it convenient in receiver chain budget analysis. The average IIP3 is about 18dBm, which is much higher than a general RF blocks such as LNA and mixer. The average noise is -135dBm/Hz, corresponding to a 39dB noise figure, which is much worse than RF blocks. In fact, good IIP3 and relaxed noise level is suitable for baseband circuits from the system cascade budget analysis point of view. The bandwidth of most low pass filters reported is normally lower than 100MHz, which is generally enough bandwidth for existing communication systems. A 5MHz mean value of bandwidth normally corresponds to WLAN systems. Some designs have design bandwidth of around 200MHz~300MHz, corresponding to an MB-OFDM UWB baseband circuit, where the signal bandwidth is 528MHz. It also reveals some distinguishable trends for IIP3 and noise. For each technology node (0.5μm, 0.35μm and 0.13μm), the linearity improves with higher noise (of course with some exceptions due to the design variation in published works). These results support the trade-off analysis in the FoM vs. IIP3 and FoM vs. noise sections.

$FoM_{avg}[dB] = k_i \times year + b_i = 1.6 \times year - 3038.2$	
$Spec_{i,avg} = n_i \times year + c_i$	$FoM/Spec_i = m_i$
$IIP3_{avg}[dBm] = 7.8 \times 10^{-1} \times year - 1557.1$	$\frac{FoM}{IIP3} = 5.9 \times 10^{-1}\ dB/dB$
$NOISE_{avg}[dBm/Hz] = -7.2 \times 10^{-1} \times year + 1304.1$	$\frac{FoM}{NOISE} = -5.6 \times 10^{-1}\ dB/dB$
$Fc_{avg}[Hz] = 10^{-4.5\times10^{-3}\times year+15.8}$	$\frac{FoM}{Fc} = 1.2\ dB/dec$
	$\frac{FoM}{Order} = -1.3\ dB/dec$

Table 3-6 Baseband LPF FoM statistics summary

The FoM improvement vs. year, FoM relationship with performances and specifications' variation vs. years are summarized in Table 3-6.

The power consumptions for low pass filters are predicted through the years, given moderate, high and low performances, as shown in Figure 3-49. Compared with the RF blocks, this baseband analogue block consumes much more power, but the power consumption reduces more rapidly through the years, also ending up with sub-mW power consumption for moderate specifications such as 5th order, 10MHz bandwidth and 55dB SFDR (18dBm IIP3 and -135dBm/Hz noise density) by the year of 2015. The power consumption of the collected published data has a very similar trend and distribution to the predicted curves, supporting the power estimation method for the low pass filter.

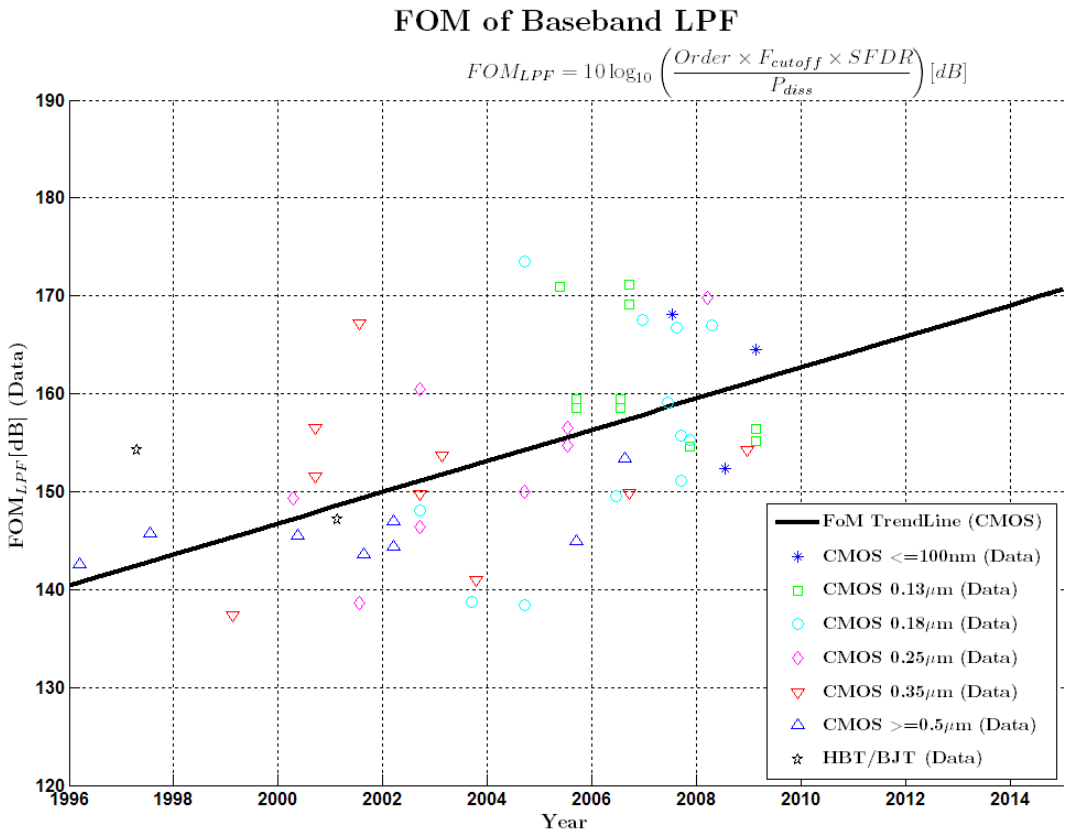


Figure 3-46 FoM trend of low pass filter

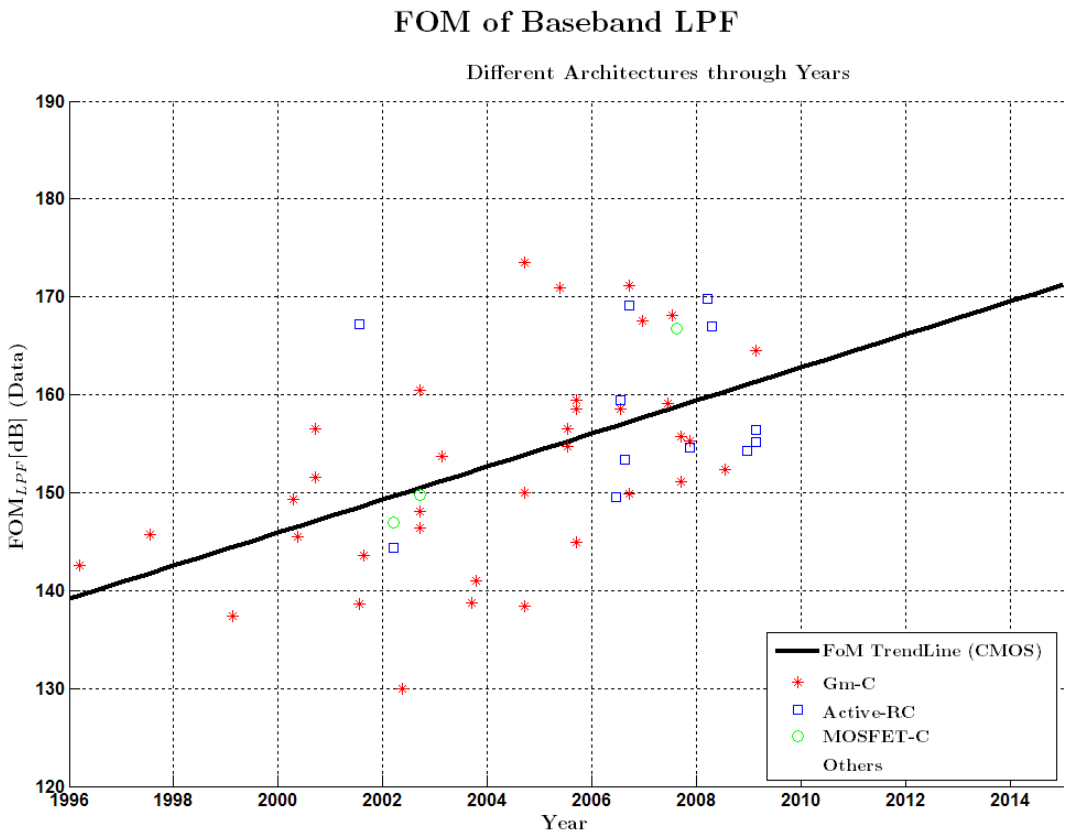


Figure 3-47 FoM tendency of low pass filter of different structures

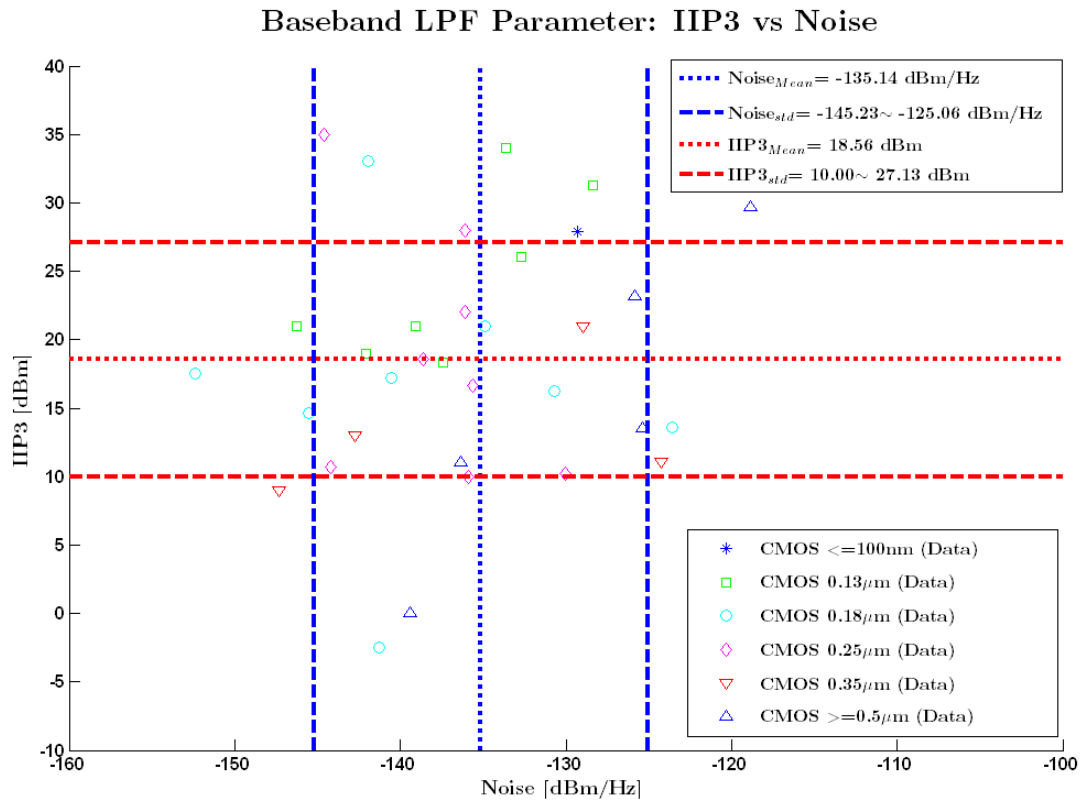


Figure 3-48 Low pass filter parameters: IIP3 versus Noise

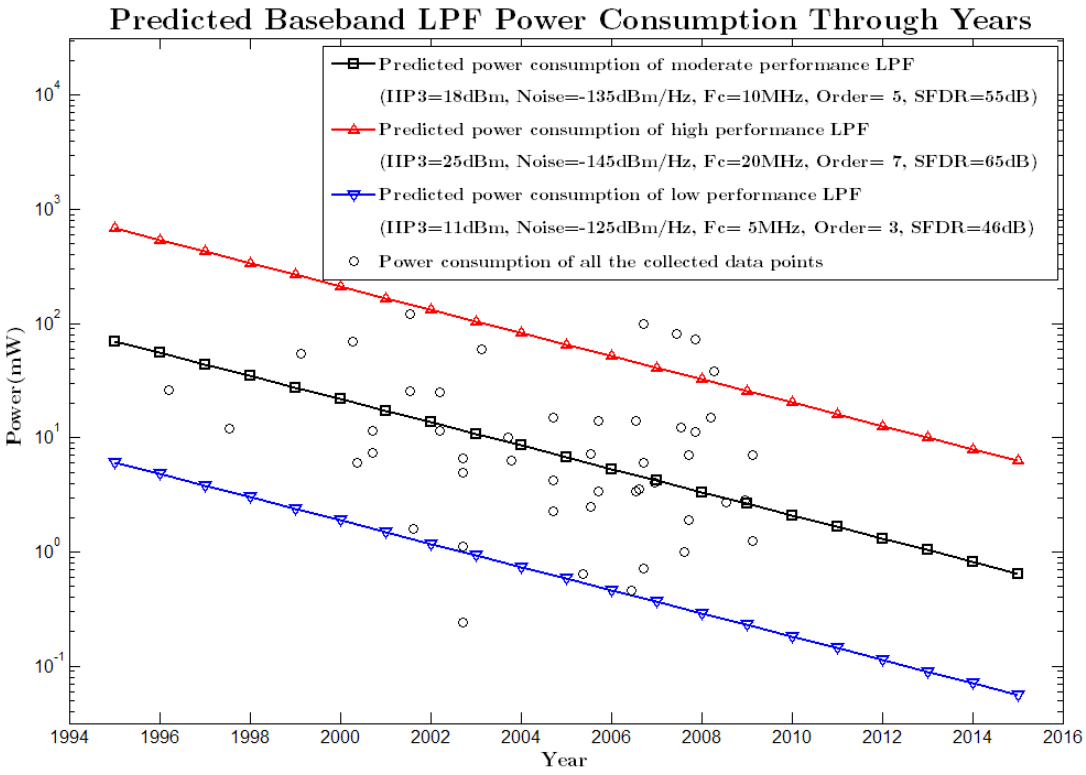


Figure 3-49 Prediction of baseband LPF power consumption through years

Baseband VGA

As discussed before, only the voltage gain and bandwidth of the VGA are involved in the definition of FoM:

$$FoM_{VGA}\left[\frac{1}{J}\right] = \frac{Gain[abs] \times BW[Hz]}{P_{diss}[mW]}. \tag{3-41}$$

The FoM values for the collection of published results and the associated linear fitted trendline are shown in Figure 3-50. The result shows that for the same power consumption, the gain-bandwidth product improves at a rate of 25.7 months per 3dB, which has a similar rate of improvement as for baseband low pass filters. The f_T of the ITRS data and the corresponding trendline are also plotted as comparison.

The figures with collected data points and fitted trendlines for FoM vs. performances, performance variation through the years and the relationships among specifications can be found in Appendix A-10. FoM improves with a moderate rate of 0.4dB for every 1dB more gain. The bandwidth has low correlation with power consumption, and FoM improves 7.5 dB for every 10 times higher bandwidth. On average, gain is getting 0.7dB higher and the bandwidth is improving by 10% every year.

The gain of a VGA used here is the maximum achievable gain, which is generally between 16~60 dB. The bandwidth is usually within the range of about 20MHz ~ 500MHz. The trade-off between gain and bandwidth can be roughly observed. For example, the gain can be as high as 80dB when the bandwidth is less than 100MHz, while the gain reduced below 60dB when bandwidth approaches 1GHz. It can also be observed clearly that the gain and bandwidth are getting better with technology improvements, supporting the theoretical constant gain-bandwidth production for a certain technology.

The FoM improvement through the years, FoM vs. performance, together with the specification variations through the years are summarized in Table 3-7. The power consumption of VGA through the years is predicted with moderate, better and worse specifications. Similar to the baseband LPF, the power consumption reduces rapidly with technology improvement, and is expected to achieve sub-mW power consumption with moderate specifications of 38dB gain and 100MHz bandwidth. Even if the gain and bandwidth are required to be 60dB and 500MHz, the power consumption can still be lower than10mW by the year of 2015. The power consumptions of data points from practical designs have a similar trend and distribution to the predicted values, supporting the effectiveness of power consumption prediction using FoM methods.

$FoM_{avg}[dB] = k_i \times year + b_i = 1.4 \times year - 2697.2$	
$Spec_{i,avg} = n_i \times year + c_i$	$FoM/Spec_i = m_i$
$Gain_{avg}[dBm] = 6.9 \times 10^{-1} \times year - 1342.4$	$\frac{FoM}{Gain} = 4.3 \times 10^{-1}\ dB/dB$
$BW_{avg}[Hz] = 10^{4.2\times10^{-2}\times year-76.0}$	$\frac{FoM}{BW} = 7.5\ dB/dec$

Table 3-7 Baseband VGA FoM statistics summary

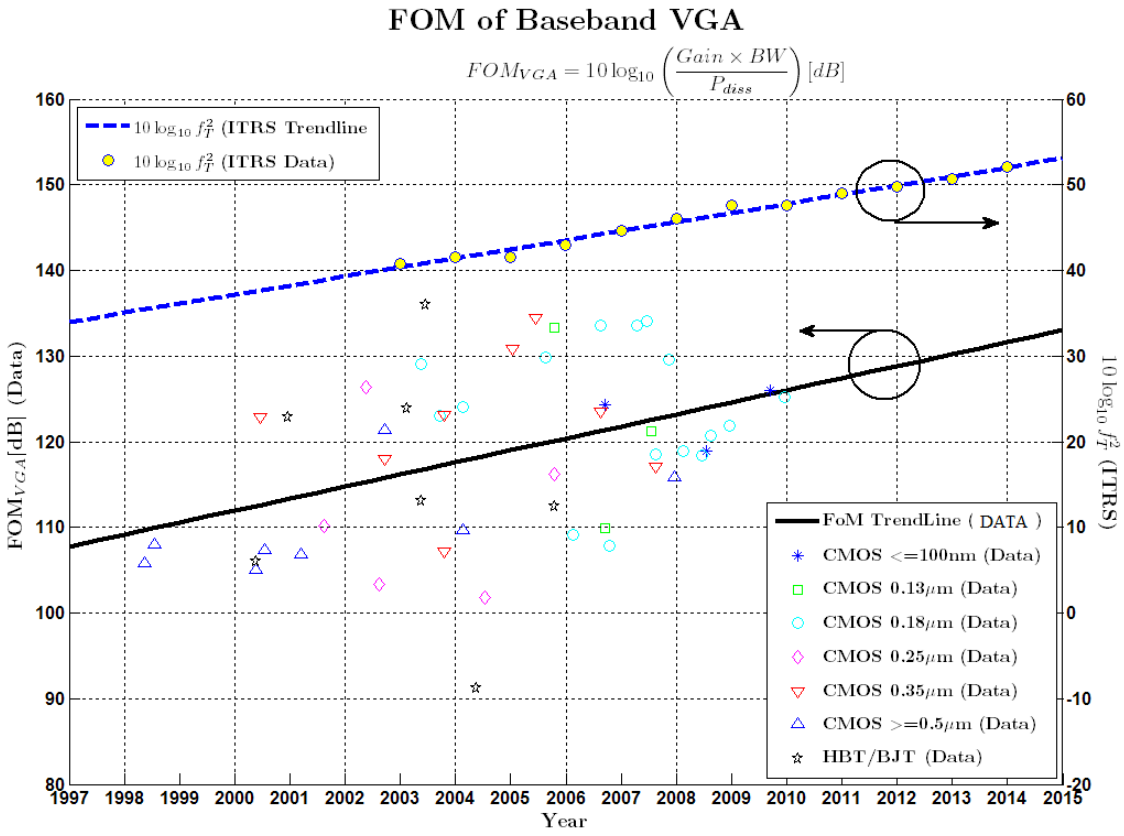


Figure 3-50 FoM tendency of baseband VGA

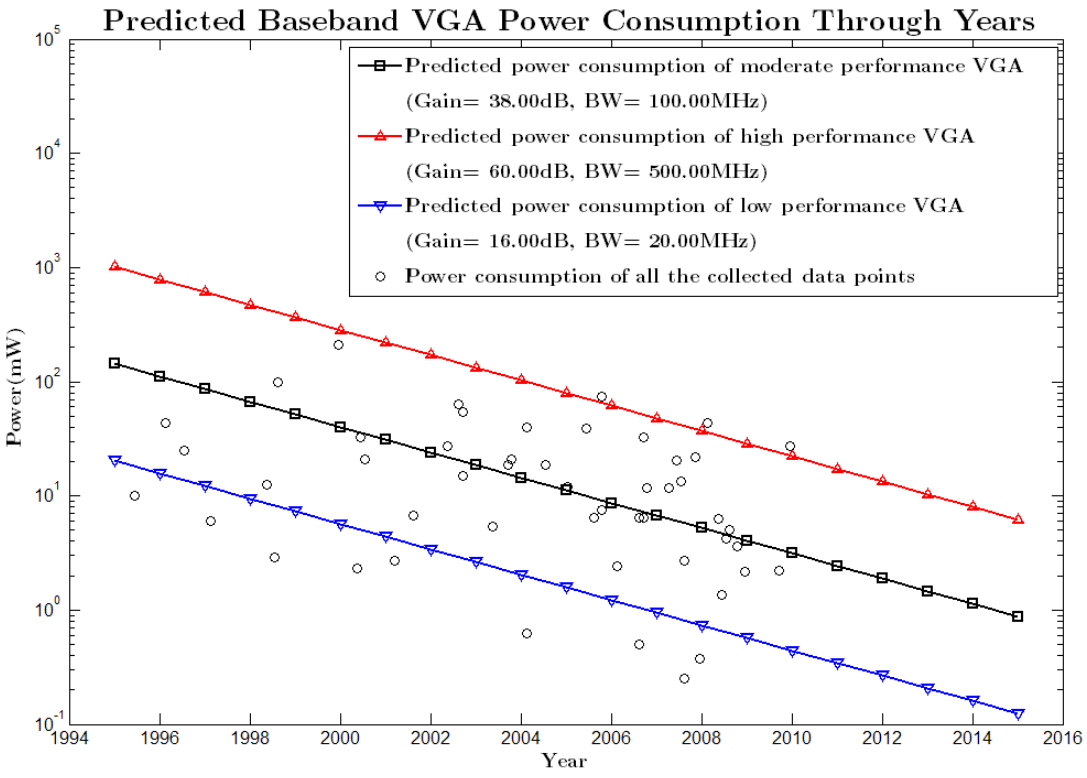


Figure 3-51 Prediction of Baseband VGA power consumption through years

3.3.6 Analogue to Digital Converter

As an interface between analogue and digital domains, analogue to digital converters (ADC) play a critical role in radio systems and are often a bottleneck in the whole system. There are two main kinds of ADC used in radio systems: Nyquist ADC and Sigma-Delta ADC. These two distinct ADC categories are to be discussed separately. Nevertheless, from the system point of view, the most important performance parameters of ADCs are similar to the other analog blocks, which are dynamic range and bandwidth.

The dynamic range is proportional to the ratio of linearity and noise. As discussed for the other front-end blocks, higher linearity and lower noise usually lead to higher power consumption for general analogue/RF circuits. For general purpose applications, the input signal of an ADC is often near full scale level for best resolution. Therefore, the total in-band noise power plus distortion power is usually provided as the minimum effective digitizing level, with a single frequency input signal with some margin below full scale level. This is usually specified as signal to noise and distortion ratio (SNDR).

The sum of these two parameters can be specified by the effective-number-of-bits (ENOB), which has the relationship with SNDR as shown in [66]:

$$SNDR = 6.02 \times ENOB + 1.76\,. \tag{3-42}$$

Ideally, the bandwidth of a Nyquist ADC is usually half of the sampling frequency. However, many Nyquist ADCs do not have a bandwidth at full resolution that gets to half the sample rate. Instead, many designs can only resolve a reduced number of bits at higher frequencies, due to the internal low-pass filtering from the signal source internal resistance into the capacitance of the sample-hold circuits or comparator inputs. Therefore, during the data collection, the effective resolution bandwidth (ERBW) is used, which is the actually input signal frequency achieving the corresponding targeted ENOB. For Sigma-Delta ADCs, the bandwidth is equal to the sampling frequency divided by the over-sampling-rate (OSR). The loop filter in the Sigma-Delta ADC usually has the analogue bandwidth of at least 10 times higher than the signal frequency, and hence the signal can normally get to the quantising operation without significant loss due to the internal filtering effect. For both types of ADCs, higher bandwidth means higher transconductance, thus higher current, and power consumption:

$$BW \propto \frac{g_m}{C} \propto I_D. \tag{3-43}$$

Integrated white noise is proportional to temperature and bandwidth. Therefore, a trade-off exists between noise and bandwidth.

$$P_{white\ noise} \propto Temp \times BW\ . \tag{3-44}$$

According to the above discussions, a commonly adopted definition of FoM is therefore defined [67] for both Nyquist ADCs and Sigma-Delta ADCs.

$$FoM_{ADC}\left[\frac{step}{pJ}\right] = \frac{2^{ENOB} \times (2 \times ERBW[Hz])}{P_{diss}[mW]}\ . \tag{3-45}$$

Nyquist ADC

There are many different types of Nyquist ADC. For modern communication systems, the most popular Nyquist ADC architectures include flash ADC [68], pipeline ADC [69] and successive-approximation (SAR) ADC [70], as shown in Figure 3-52.

Flash ADC is one of the most basic architectures. In order to convert the analog signal to an N-bit digital signal, the number of comparators needed is equal to 2^N-1. This architecture provides the fastest sampling speed because all the digits are converted simultaneously. However, for high resolution, power consumption will be exponentially increased, and matching among a large number of comparators increases the design difficulty. Therefore, flash converters are usually used in high speed and low resolution applications. In SAR converters, a single comparator outputs one bit at a time by comparing the analog input and a DAC output which is updated by previously decided bits, which are stored in a register. Each successive DAC value is set to half of the previous uncertainty range, and so the final digital output is thus successively approximated to the analog input signal. This type of ADC can provide higher resolutions at the cost of speed. A pipelined ADC generates digital bits from a series connected stages, from the most significant bits (MSB) to the least significant bits (LSB). This method is slower than a pure flash ADC and generally faster than an SAR ADC, and with a moderate resolution. Various other techniques may be also involved in ADC design, including folding and interpolation [71], time-interleaved [72], etc, aiming to reduce the number of comparators (to same the power and area) or increase the amplitude of LSB voltage (to achieve higher resolution).

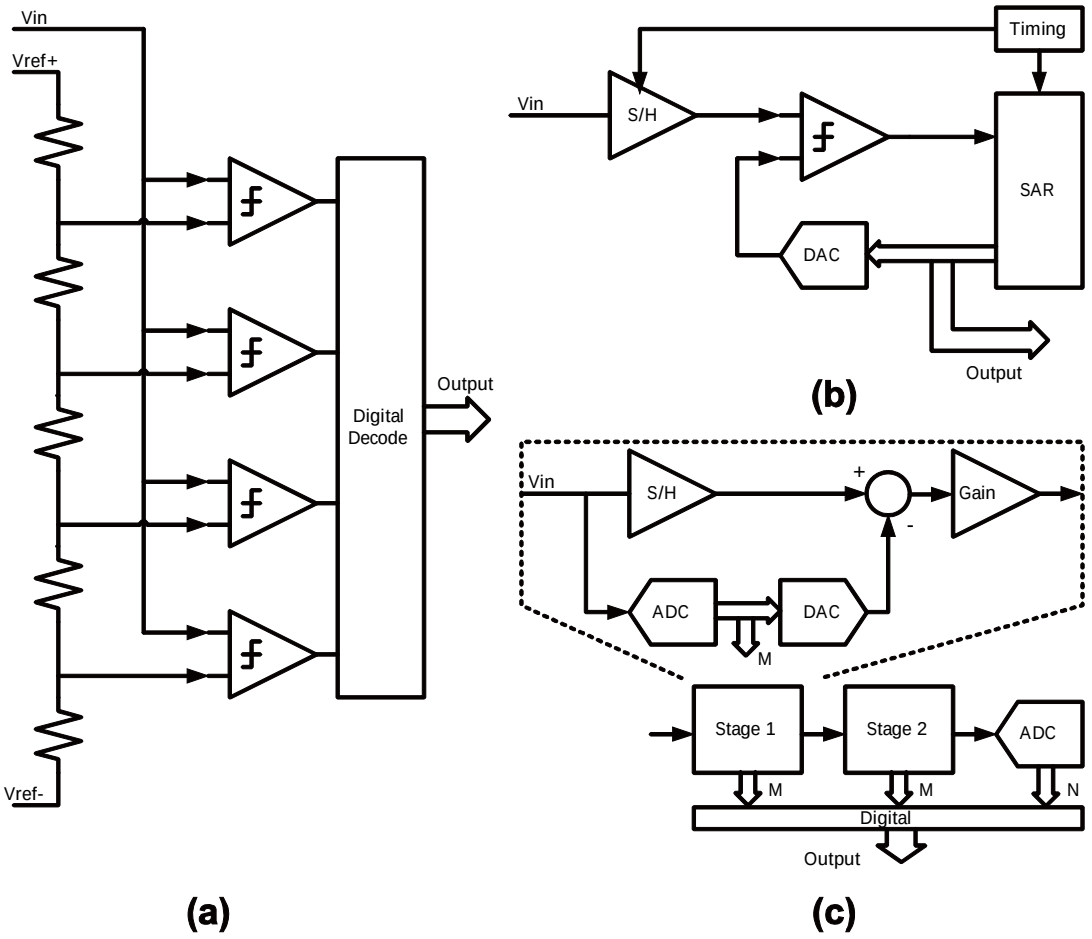


Figure 3-52 Nyquist ADC architectures: (a) Flash (b) SAR (c) Pipelined

In spite of the different architectures, the fundamental block for all types of Nyquist ADCs is a single comparator, which is essentially a 1-bit ADC. The implementation of a comparator could be similar to an op-amp, using positive feedback, or using switched capacitors combined with a CMOS inverter.

The FoMs of published Nyquist ADCs, together with associated fitted trendlines, are illustrated in Figure 3-53 and Figure 3-54 according to technology and architectures. As observed in the figures, the FoM of a Nyquist ADC improves by 3dB every 22.6 months, which is very similar to Moore's law. This means that for the same performance, the power consumption can be halved within less than two years for CMOS technology. Intuitively, the comparator circuit behaviour is ideally more like digital rather than analog circuits because the output is either VDD or ground. In spite of this, one of the most important issues in comparators is the matching of transistors such that the smallest input step size can be resolved reliably, which is essentially an analogue issue. To maintain the resolution, the size of the transistors sometimes needs to be large to reduce the mismatches, instead of continuing scaling down as digital circuits. Therefore, the scaling with Moore's law of the FoM improvement is believed

not to help from the comparators. Instead, the architectural improvement is probably one of the main reasons that the FoM is improving with Moore's law. With technology improvements, more digital correction circuits *can* be put on the Nyquist ADC, and hence these digital circuits finally influence the FoM improvement significantly. The FoM is expected to achieve 20.8 step/pJ (or 0.048pJ/step) by the year of 2015. The figures also reveals that SAR ADCs can normally achieve better FoM and flash ADCs' FoM are usually lower, while pipelined ADCs provide moderate FoM values.

The figures of the data points and fitted trendlines of FoM vs. performance, and performance vs. year can be seen in more detail in Appendix A-11. Measurements reported with one bit ENOB difference show that this only causes 0.17dB difference in FoM, where FoM in dB is calculated as 10×log(FoM[step/pJ]), indicating that ENOB are strongly correlated with power consumption. The FoM varies by 2.58dB with 10 times bandwidth scaling. The figures also show that the ENOB figures are decreasing slightly with time (more likely to be due to the lowered supply voltage so that the amplitude of LSB is also reduced) and bandwidth are increasing over 30% every year.

The ENOB vs. BW relationship of all the collected data points is repeated in the chapter and shown in Figure 3-55. The effective resolution of published circuits is generally from 5 bits to 11 bits, while the bandwidth is usually from 2MHz to 600MHz. A strong correlation between ENOB and bandwidth can be observed in the figure. For higher speeds (BW>30MHz), ENOB reduces by 0.96 bits for every doubling of bandwidth, which is very close to theoretical analysis of -1dB/octave [66], or the SNDR of -6dB/octive. The average product of performance for higher bandwidth, which is defined as $P=2^{ENOB}\times(2\times ERBW[Hz])$ by Walden [66], maintains a constant value of 4.37×10^{10} with various resolution and bandwidth combinations. Note that the product of performance is the numerator of the FoM definition in equation (3-45), with the units of Hz. At lower speeds, the resolution stops increasing. The limit in the ENOB is probably determined by the limits of analogue device matching for practical sizes in conventional circuit architectures. Above a certain frequency, this matching limit of ENOB is reduced. The corner bandwidth where this is observed is about 20~40MHz, according to Figure 3-55.

The FoM improvement vs. year, FoM variation with performances and specification variation through the years are summarized in Table 3-8. The power consumptions are predicted for ADCs with moderate (9 bits, 60MHz), higher (11 bits, 200MHz) and lower (6 bits, 30MHz) performances, respectively. To achieve the

moderate requirement, the power consumption can be reduced to about 5mW by the year of 2015. However, high specification requirements still demand consumptions of more than 40mW by then, which will continue to be a bottleneck for the whole receiver, compared with other receiver blocks which are generally achieving less than 10mW power consumption. Most of the collected published results have a similar distribution and trend with respect to predicted values. This implies the effectiveness of the developed power consumption prediction method using the FoM strategy.

$FoM_{avg}[dB] = k_i \times year + b_i = 1.6 \times year - 3203.8$	
$Spec_{i,avg} = n_i \times year + c_i$	$FoM/Spec_i = m_i$
$ENOB_{avg}[dBm] = 6.2 \times 10^{-2} \times year + 131.2$	$\frac{FoM}{ENOB} = 1.7 \times 10^{-1}\ dB/bit$
$BW_{avg}[Hz] = 10^{1.2\times10^{-1}\times year - 236.0}$	$\frac{FoM}{BW} = 2.6\ dB/dec$

Table 3-8 Nyquist ADC FoM statistics summary

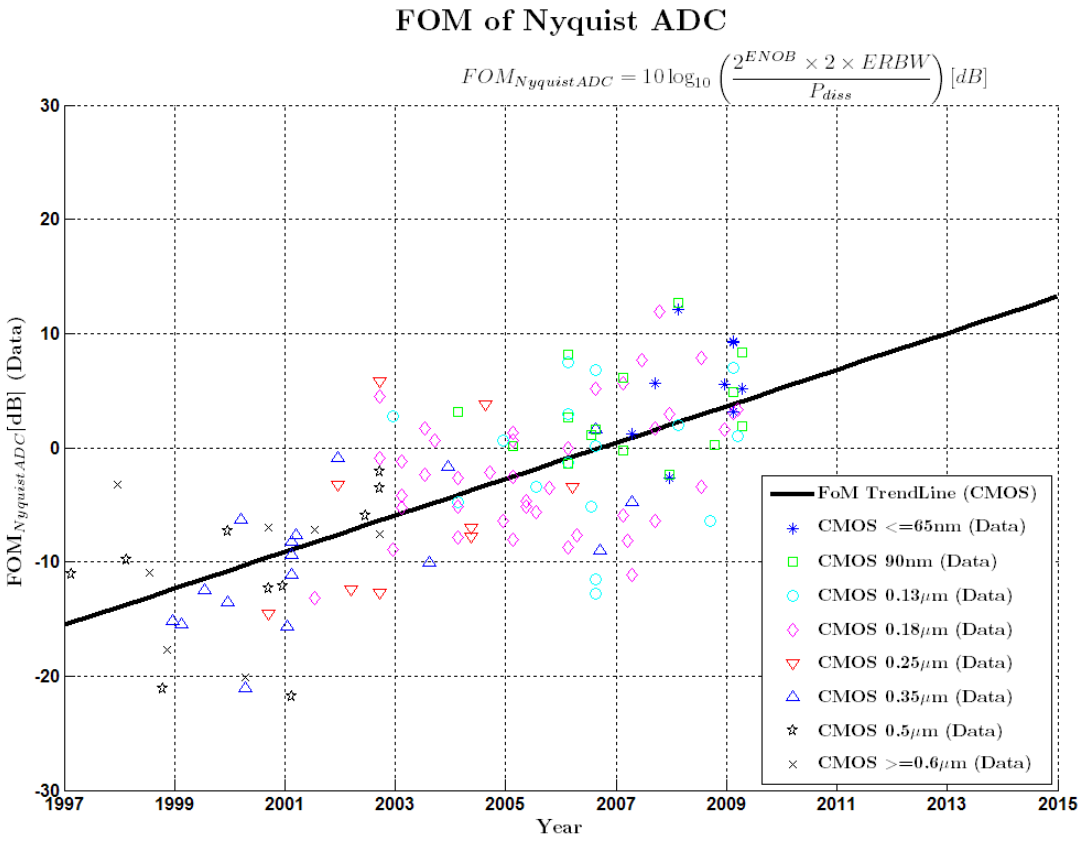


Figure 3-53 FoM tendency of Nyquist ADC

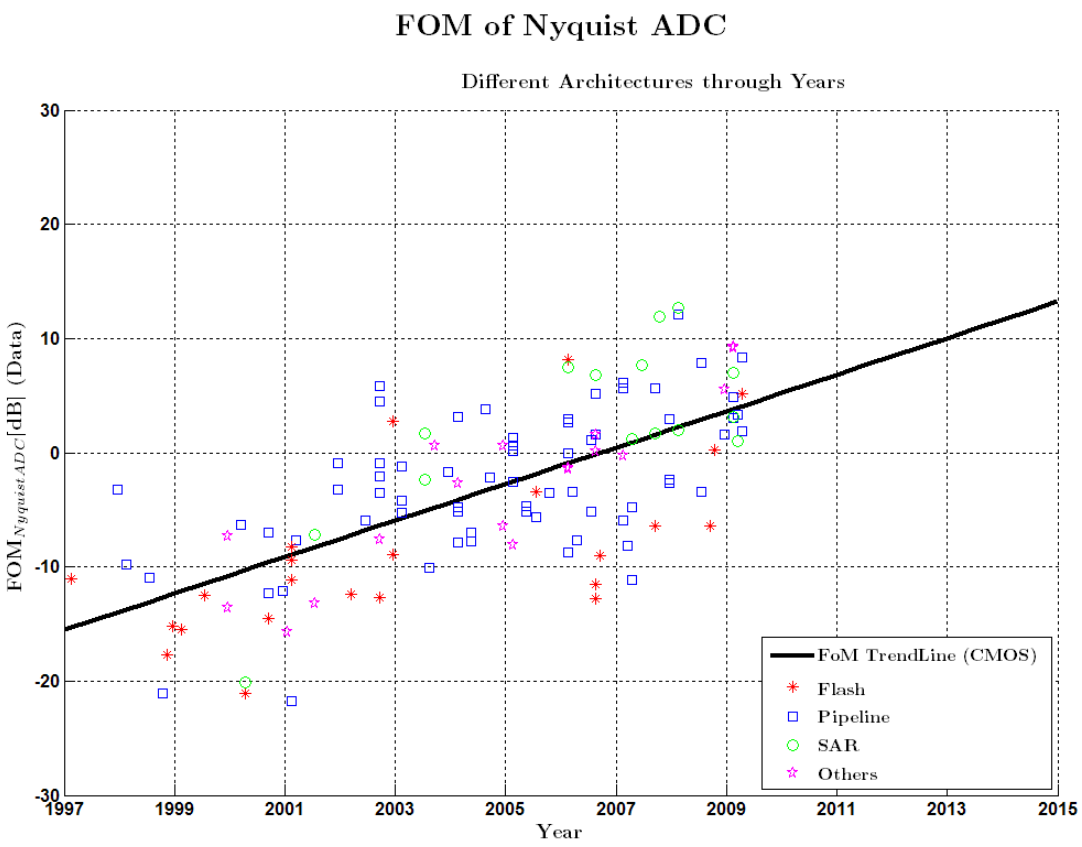


Figure 3-54 FoM tendency of Nyquist ADC with different architectures

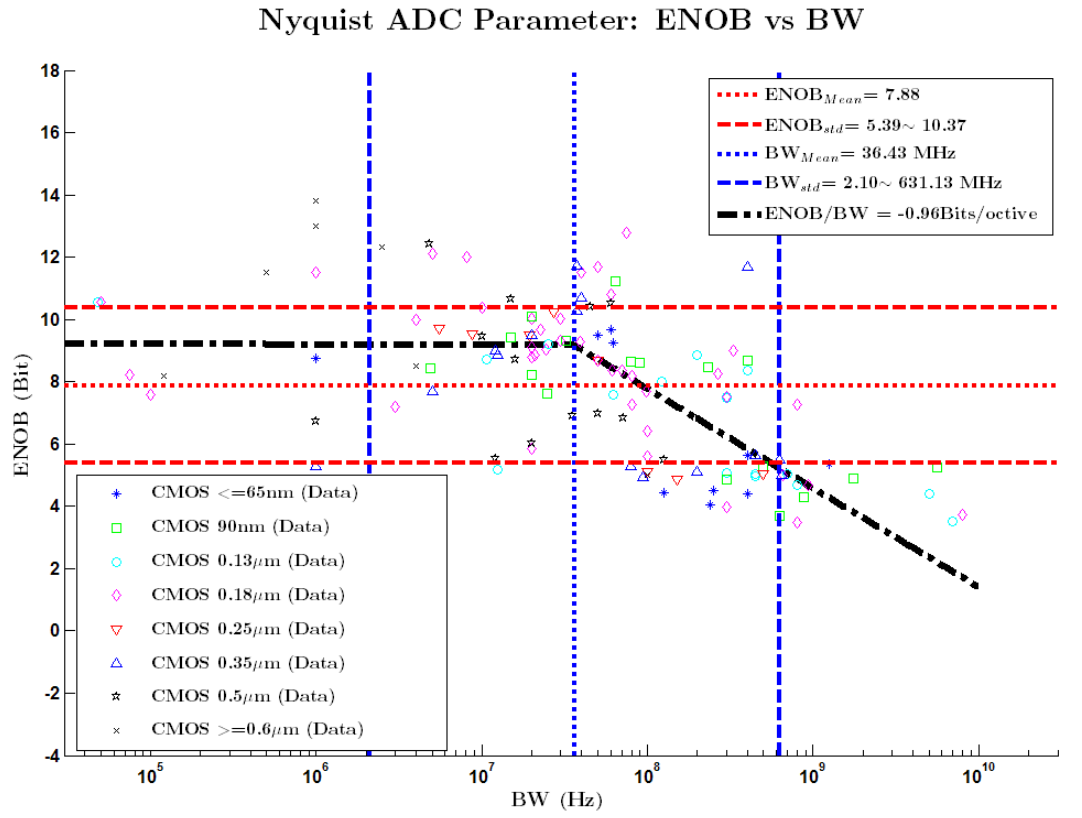


Figure 3-55 Nyquist ADC parameters: ENOB versus bandwidth (break point of ~30MHz explained in pp. 85)

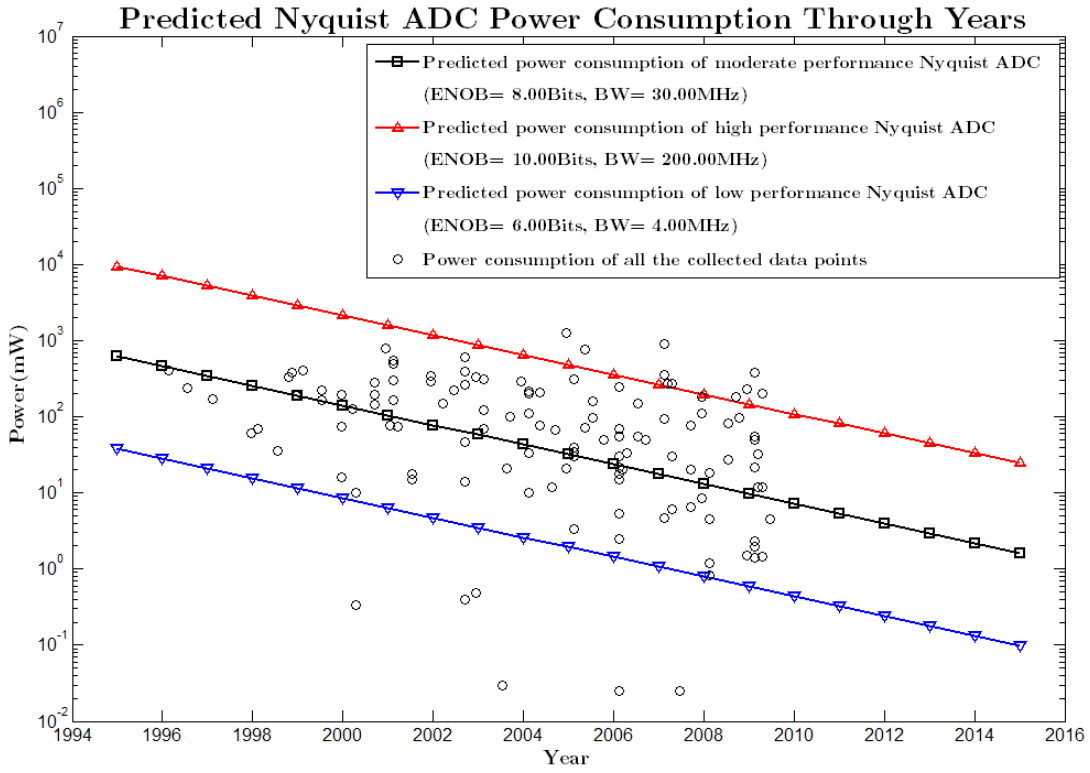


Figure 3-56 Prediction of Nyquist ADC power consumption through years

Sigma-Delta ADC

Another category of ADC is the Sigma-Delta ADC, sometimes called an oversampling ADC, in which a sigma-delta modulator, as a core block, performs quantization noise shaping and hence reduces the in-band quantization noise at the cost of much lower effective bandwidth, typically less than the Nyquist frequency by a factor of the over-sampling-rate (OSR) [73]. A typical second-order Sigma-Delta ADC architecture is shown in Figure 3-57, containing a sigma-delta modulator and followed by a digital low pass filter. The N-bit ADC is typically a 1-bit comparator.

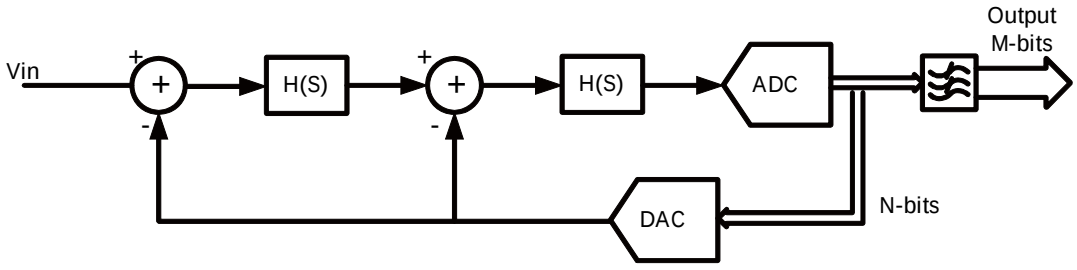


Figure 3-57 Second order Sigma-Delta ADC architecture

Figure 3-58 shows the FoM of collected data points in the literature, together with the linear fitted trendline through years. It shows that the FoM of Sigma-Delta ADC improves 3dB every 33.1 months, compared to 22.6 months for the Nyquist ADC. The

improvement rate is more similar to RF blocks such as the LNA or Mixer. The main reason is that, the limiting parts in an over sampling ADC tend to be analogue parts of the sigma-delta demodulator, particularly the feedback DAC, the signal and feedback summation in the loop filter, and to a lesser extent the comparator speed (Also note that for the 1-bit comparator, the accuracy is not important). According to the trendline, the FoM is expected to fall to 6.7 step/pJ (or 0.15pJ/step) in the year of 2015, which is about three times higher than that of a Nyquist ADC.

The figures of the data point and fitted trendlines of the FoM vs. performances, performances variation through years can be found in Appendix A-12. The FoM is hardly changed with ENOB, implying that the ENOB is strongly correlated with the power consumption, while the FoM increases by 1.1dB for every 10 times higher bandwidth. In a similar way to Nyquist ADCs, the ENOB also reduces slowly over the years, and the bandwidth increases by about 30% every year on average.

Figure 3-59 shows the dynamic range and bandwidth trade-off for collected data points for Sigma-Delta ADCs. The ENOB are generally between 9 and 15 bits, which is about 4 bits or 24dB SNDR higher than Nyquist ADCs. However, the bandwidths are typically below 10MHz. The ENOB decreases by less than 0.4bits for every doubling of the bandwidth, according to the figure. The average product of performance, $P=2^{ENOB}\times(2\times ERBW[Hz])$, keeps a constant value of 4.58×10^9 Hz, which is almost one-tenth of the Nyquist ADCs' value. This comparison implies that the oversampling ADC has the advantage of high dynamic range, for certain application like audio devices, but has lower overall performance and FoM than its Nyquist counterparts.

$FoM_{avg}[dB] = k_i \times year + b_i = 1.6 \times year - 3203.8$	
$Spec_{i,avg} = n_i \times year + c_i$	$FoM/Spec_i = m_i$
$ENOB_{avg}[dBm] = 6.2 \times 10^{-2} \times year + 131.2$	$\frac{FoM}{ENOB} = 1.7 \times 10^{-1}\ dB/bit$
$BW_{avg}[Hz] = 10^{1.2\times10^{-1}\times year-236.0}$	$\frac{FoM}{BW} = 2.6\ dB/dec$

Table 3-9 Sigma-Delta ADC FoM statistics summary

The relevant FoM statistics for Sigma-Delta ADCs are summarized in Table 3-9, and the power consumptions are predicted for different performances in Figure 3-60. With moderate specifications of 12 bits ENOB and 2MHz bandwidth, a Sigma-Delta ADC is expected to consume about 3mW power in the year of 2015, while 14 bits ENOB with 5MHz bandwidth still consumes nearly 30mW power. Again, the collected

data points have power consumptions with very similar trend and distribution to the predicted values, implying the effectiveness of FoM method in predicting power consumption.

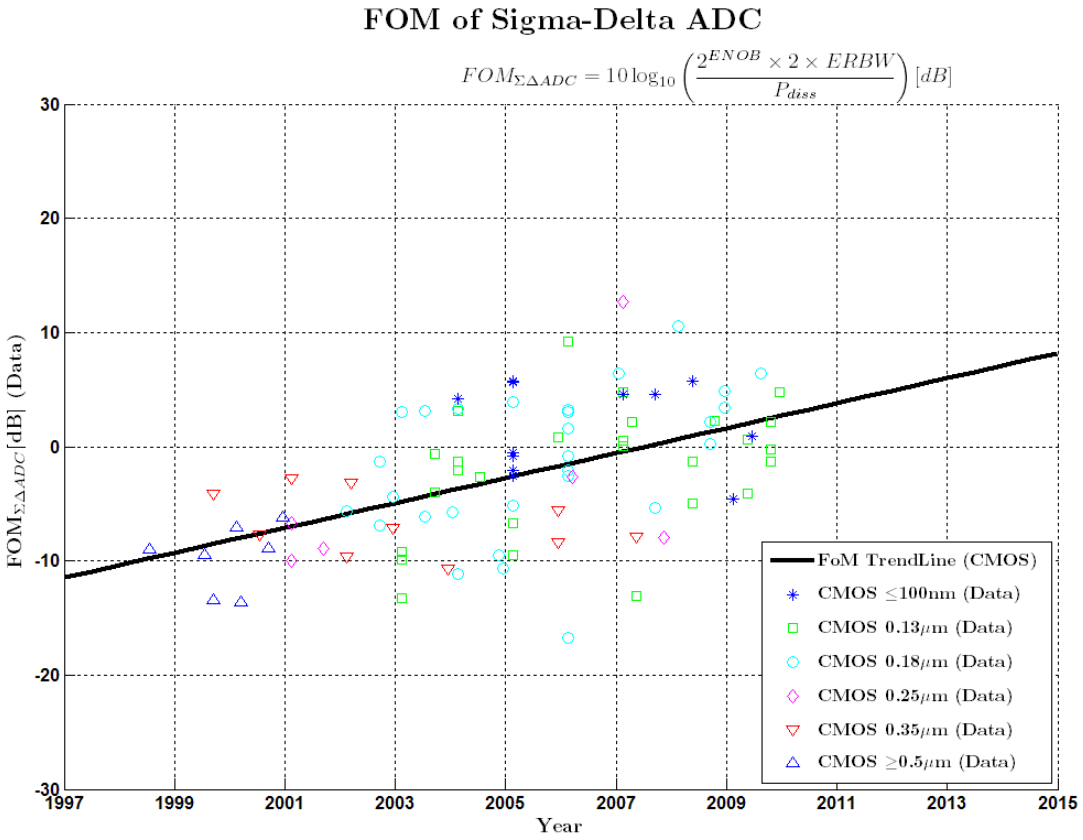


Figure 3-58 FoM tendency of Sigma-Delta ADC

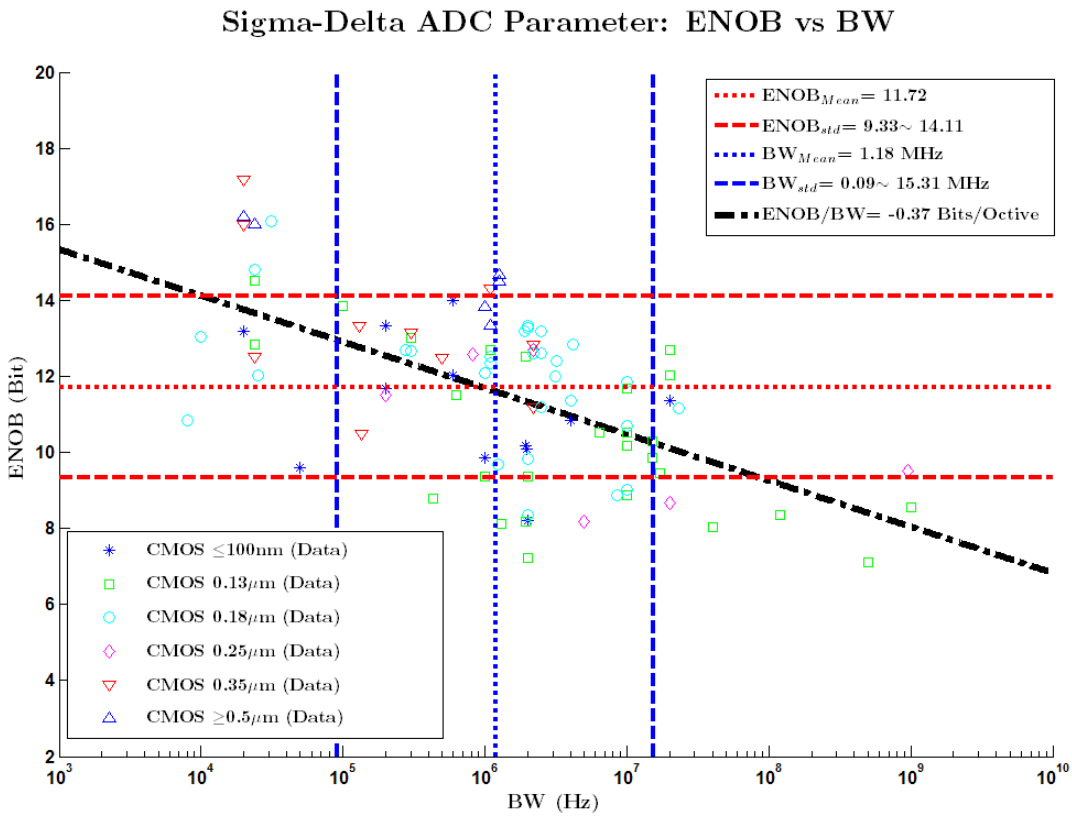


Figure 3-59 Sigma-Delta ADC parameters: ENOB versus bandwidth

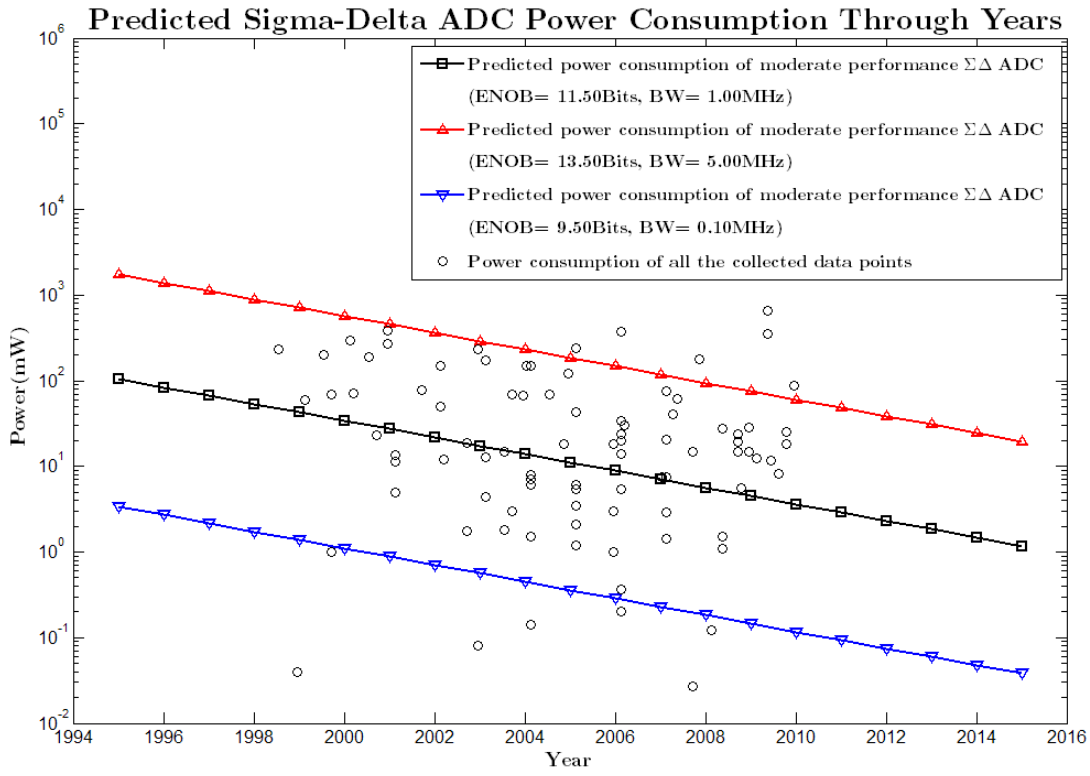


Figure 3-60 Delta Prediction of Sigma-Delta ADC power consumption through years

3.3.7 FoM Prediction Summary

The above sections investigated and explained the figures of merits of the main receiver blocks, including wideband/narrowband LNA, mixer, LC/ring oscillator, CML /injection-locked frequency divider, baseband LPF/VGA, and Nyquist/Sigma-Delta ADC. The FoM in 2010, the predicted FoM in 2015 and the improvement rates are summarized in Table 3-10.

RF front-end blocks, including the LNA and mixer generally improve their FoMs by 3dB within 30~40 months. The FoMs of baseband blocks (LPF and VGA), improve over 50% faster than RF blocks, typically within 22~25 months. Note that not all the specifications are included in baseband blocks, for example, the gain of the LPF and the IIP3/noise of the VGA are ignored. The frequency synthesizer blocks such as VCO and dividers have different FoM mechanisms, and hence their FoMs vary with respect to each other, but they principally improve slower than other blocks. The FoMs of Nyquist and Sigma-Delta ADCs are improving more like digital and RF blocks respectively, due to their different principles.

		Unit	FOM 2010 [abs]	FOM 2010 [dB]	FOM 2015 [abs]	FOM 2015 [dB]	MM Per 3dB	Eq.
LNA	$FOM1_{NBLNA}$	$\frac{GHz}{mW}$	13.84	11.42	45.84	16.61	34.8	(3-6)
	$FOM2_{NBLNA}$	GHz	9.14	9.61	27.05	14.32	38.4	(3-7)
	$FOM1_{WBLNA}$	$\frac{GHz}{mW}$	5.26	7.21	19.85	12.98	31.3	(3-22)
	$FOM2_{WBLNA}$	GHz	3.61	5.57	10.38	10.16	39.4	(3-23)
Mixer	$FOM1_{Mixer}$	$\frac{GHz}{mW}$	0.35	-4.59	1.26	0.99	32.4	(3-26)
	$FOM2_{Mixer}$	GHz	0.29	-5.39	0.92	-0.35	35.9	(3-27)
VCO	FOM_{LCVCO}	$\frac{10^{21}}{J}$	4.12	216.15	8.26	219.17	59.8	(3-28)
	$FOM_{RingVCO}$	$\frac{10^{18}}{J}$	2.18	183.39	2.24	183.47	-	(3-28)
Div	FOM_{ILdiv}	$\frac{GHz}{mW}$	31.18	13.26	54.19	17.34	44.3	(3-36)
	FOM_{CMLdiv}	$\frac{GHz}{mW}$	3.88	5.89	6.78	8.31	74.7	(3-36)
LPF	FOM_{LPF}	$\frac{10^{16}}{J}$	1.86	162.70	11.64	170.66	22.7	(3-37)
VGA	FOM_{VGA}	$\frac{10^{12}}{J}$	3.98	126.00	20.09	133.03	25.7	(3-41)
ADC	$FOM_{NyquistADC}$	$\frac{step}{pJ}$	3.34	5.24	20.98	13.22	22.6	(3-45)
	$FOM_{\Sigma\Delta ADC}$	$\frac{step}{pJ}$	1.88	2.73	6.60	8.19	33.1	(3-45)

Table 3-10 Receiver Blocks FoM Prediction Summary

3.4 Summary

In this chapter, the FoMs of all the main blocks in a typical receiver chain have been defined and investigated. A large number of published designs have been reviewed to collect FoM data and to allow quite confident predictions of the future FoM values for these cells. A systematic approach is analysed and applied to power consumption estimation, according to the relationships among FoM, specification and time scales. The derived power consumption prediction curves are expected to be accurate enough for the first order system level estimation at the starting stage of receiver design. The FoM data and the approach of predicting the power consumption are used in chapter four as an example of its application.

Chapter 4 Integrated CMOS Spectrum Monitor Architectures

4.1 Introduction

The spectrum monitor aims to obtain a map of the spectrum occupation that is fast and accurate enough for the management of the cognitive radio function, while maintaining relatively low power consumption and with a small cost overhead. This is a very challenging topic for state of the art technology.

The spectrum monitor function can be considered in two parts: the receiver which acquires the target frequency range in some way, and the sensing function which determines if there is activity in a potentially usable channel that must be avoided. There are many methods for sensing the activity in a part of the spectrum. Basic spectrum sensing methods include matched filter, energy detector and feature detection [3]. Matched filter detection requires advance knowledge of modulation related information, and hence cannot be adopted in the spectrum monitor, where the nature of the spectrum occupation must be assumed to be unknown. The energy detection method evaluates the average energy in the spot channel to decide the occupancy. Just a short time could be used to get the instantaneous energy present, so this method is very fast. The main problem is the sensitivity requirement. The threshold of the decision level is difficult to specify due to the variation of modulated signals and communication channels. Besides, this method cannot handle a negative SNR such as in spread-spectrum system. The feature detection method, on the other hand, evaluates the 'cyclostationary feature' of the modulated signals over a long period. Modulated signals usually exhibit periodicity because of intentionally introduced signals which assist the receiver in detecting the pulse timing, carrier phase, etc. Hence the evaluation

time depends on the modulation scheme, and is expected generally to be longer than the energy detection method. Despite the slower evaluation speed, even signals with negative SNR can be distinguished from white noise [74]. In a practical wideband spectrum sensing scheme, a two stage strategy is recommended to achieve fast speed and accurate decision [75], as shown in Figure 4-1. In this strategy, the receiver first performs the fast energy detection over a wide frequency range. An acceptable decision level is specified to identify the channels occupied by strong signals, and these channels are marked as unusable. The second step is to apply feature detection to the rest of the channels, which could be occupied just by white noise or by modulated signals below the noise level (including the input noise floor and the noise generated by the receiver). A combination of the two steps can be expected to be optimised to achieve a good balance between speed and accuracy.

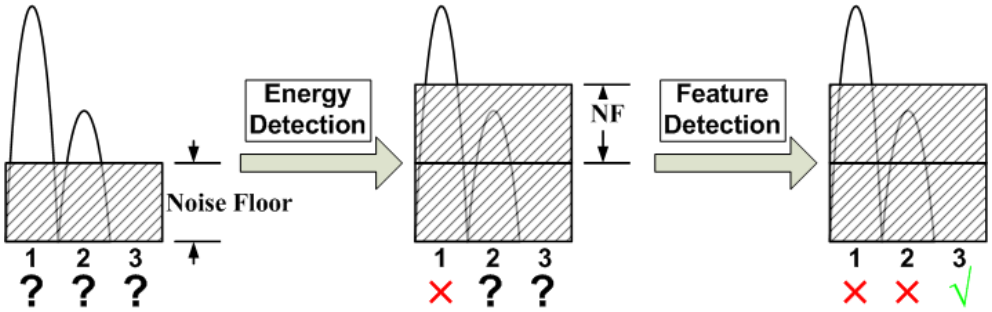


Figure 4-1 Two-stage detection example

This chapter will focus on the energy detection part, which involves the receiver architecture investigation based on CMOS technology. A direct way of detecting the received channel's energy is in the analogue domain, including a baseband channel selection filter, a squarer and an integrating low pass filter [76]. Although the burden on the ADC and DSP functions is relaxed significantly, the channel selection filter should clearly be tunable. This involves quite complicated circuitry, typically with a large silicon area, and is difficult to integrate. On the other hand, with the development of CMOS technology, ADC and DSP performance figures are improving rapidly, as discussed in chapter three. Therefore, it is reasonable to move more functions from analogue circuits to the ADC and DSP. In the architecture proposed for this study, an ADC digitizes the wideband signal received by the RF/analogue front-end. A DSP function performs an FFT and then calculates the energy falling in each FFT bin (channel) to finish the spectrum energy mapping function. This is followed by a further feature detection algorithm to make the decision.

In this chapter, a simplified spectrum occupation model is first given. Then the specifications and architectures of a practical spectrum monitor are discussed. Finally, the practical system level design is developed and power consumptions are predicted.

4.2 Spectrum Monitor Specifications

The specifications of the spectrum monitor for the front-end, ADC and the frequency synthesizer are listed in Table 4-1, and will be discussed in this section.

System	Frequency	2GHz~5GHz
	Channel BW	200kHz
	Sensitivity	-85dBm/200kHz
	Decision Margin	3dB
	Sub-band BW	100MHz
ADC	Type	Nyquist
	Full Scale	$1V_{pp}$
	BW	≥50MHz
	ENOB	≥8.2 bit
Front-End	Gain	18dB
	Noise Figure	≤20dB
	Linearity	OIM3≤-83dBm
		IIP3 depends on architecture
	Image Rejection	Ideally -65dBc, typically -55dBc
Frequency Synthesizer	Phase Noise and Spur	-80dBc/Hz through 100MHz band
	Tuning Range	3GHz
	Tuning Step	100MHz

Table 4-1 Spectrum monitor specification for 100MHz sub-band

4.2.1 Receiving Chain (Front-End and ADC)

In order to determine the specifications of the spectrum monitors, the spectrum occupation situation should be investigated first. Table 4-2 lists some popular communication systems, including TV band, cellular band, ISM band or even higher frequency bands for WLAN and UWB, etc. Among these frequencies, the cellular bands are always too crowded for cognitive radio. Therefore, generally speaking, the

lower frequency TV bands and higher frequency bands are potential candidates for future cognitive radio applications. In this chapter, the 2GHz~5GHz band is selected as the region of interest, which is relatively quieter than the TV band. Clearly an extended study could include the TV bands as these become more available.

	DVH-H	GSM 900	DECT	PCS 1900	UMTS FDD	Blue-tooth	802.11a	MB-UWB
Frequency (MHz)	470~750	935 ~ 960 (DL)	1880~1897	1930~1990 (DL)	2110~2170 (DL)	2400 ~ 2483	5150 ~ 5350	3100~10600
Band (MHz)	280	45	16	60	60	83	200	7500
BW (MHz)	8	0.2	1.76	0.2	5	1	20	512
Max. Pwr (dBm)	-28	-15	-33	-23	-	-20	-30	-40
Min. Pwr (dBm)	-80	-102	-86	-102	-107 (384kb/s)	-70	-72	-80
Max. Pwr dBm /200kHz	-44	-15	-42	-23	-	-26	-50	-74
Min. Pwr dBm /200kHz	-96	-102	-95	-102	-	-77	-92	-114

Table 4-2 Some popular communication systems ('DL' stands for 'Down Link').

As can be seen in the table, the channel bandwidth varies among different communication systems. Note that the modulation scheme is WCDMA for the UMTS standard; therefore the input signal is below the white noise level due to the spread-spectrum algorithm. In the following discussions, the FFT bin is assumed to have 200kHz bandwidth, which is the narrowest bandwidth in commonly used modern communication systems. The maximum and minimum received signals power is therefore normalized to 200kHz bandwidth. Note that the simplified normalization procedure assumes that there is a flat power spectrum density within the band, which is often not the truth, and that the signal power near the centre frequency is usually higher, so this is only a rough estimation. The required dynamic range of the receiver could be defined to be from -102dBm to -15dBm as in the GSM system, over the 2GHz~5GHz frequency band. The actual achievable dynamic range is determined by the combinations of ADC performances, as well as the gain, noise and linearity

specifications of the front-end. In most situations, it is better to analyze the required dynamic range requirement from the ADC part, and work back to the antenna.

Because of the necessity of phase information in the feature detection, the final baseband should be complex signals; therefore, a pair of I/Q ADCs is needed. The 3GHz band signal requires a bandwidth of 1.5GHz after conversion to zero IF. Assume that the 87dB dynamic range corresponds to the SNDR requirement of ADC, then the effective number of bits (ENOB) is equal to 14.16 bit. According to the discussions in chapter three, neither a Nyquist nor a Sigma-delta ADC is currently able to achieve this high bandwidth and high resolution simultaneously. Even if this could be done, according to the power consumption estimation method in chapter three, the power consumption could reach several Watts, which would be unacceptable in mobile devices. As a consequence, a lower dynamic range and/or lower bandwidth are needed for practical ADCs. For this reason, instead of obtaining the channel occupancy situation of the whole band in one go, it is more practical for the proposed spectrum monitor to scan a fraction of the band of interest at one time to reduce the required ADC's bandwidth, and then to sweep the scanning segment within the entire band, as demonstrated in Figure 4-2. This strategy is similar to multi-band OFDM systems as reviewed in chapter two. Essentially, this method shares the performance burden between both RF/analogue and ADC functions, so that both parts have practical performance requirements and consume reasonable power.

With a defined overall target band of 3GHz and a channel bandwidth of 200kHz, the selection of the bandwidth of each sub-band, and thus the number of channels, is determined by whether the ADC's speed and resolution are able to deal with the in-band channels. In this chapter, this sub-band is selected as a moderate bandwidth of 100MHz and there will be 30 sub-bands within the entire band of interest. By defining the number of channels in each sub-band as M, then M=500 in this case. These configurations are shown in Figure 4-2.

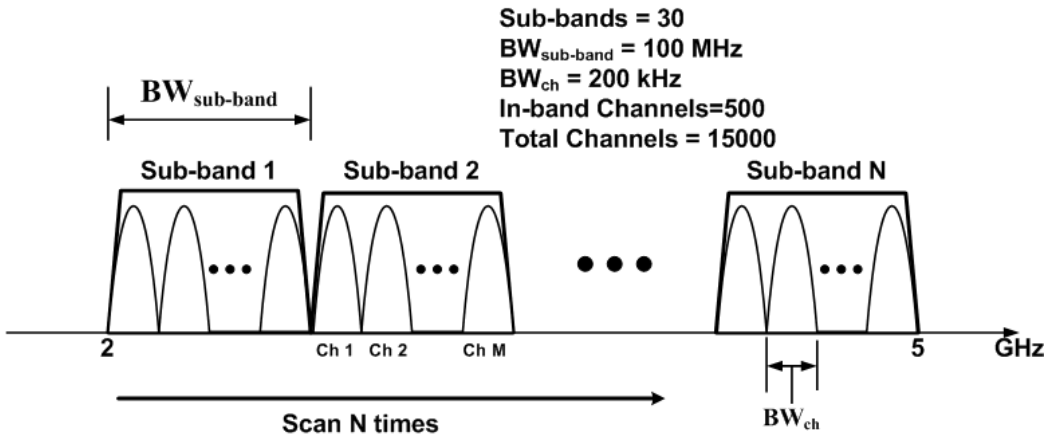


Figure 4-2 Spectrum monitoring plan

To determine the front-end and ADC's specifications to handle the 500 channels each with 200kHz bandwidth, a simplified signal power distribution model is first developed. Firstly, for the sake of argument, assume any 200kHz busy channel's signal power is within -100dBm to -20dBm, which is reasonable according to Table 4-2. The input signal power at the antenna of the spectrum monitor is decided by the instantaneous transmitting power of other mobile/base stations and path loss. The total pass loss between a transmitter and a receiver is related to the distance (which determines the average path loss) the large-scale fading margin (which could be 6~10dB, according to the actual environments) and small-scale fading margin (20~30dB) [77]. The average path loss L_p is a function of the distance from the transmitter, d, and can be expressed in decibels as:

$$\overline{L_p(d)}\ dB = L_s(d_0)\ (dB) + 10n\log(d/d_0). \quad (4\text{-}1)$$

The value of n is a fitting parameter that depends on different factors, such as the environment, antenna and frequency. The parameter d_0 is a reference distance where the reference attenuation of the transmitted RF signal can be ideally obtained as below, where λ is the wavelength.

$$L_s(d_0) = (4\pi d_0/\lambda)^2\ . \quad (4\text{-}2)$$

The actual path loss at a single point from the transmitter usually can be seen as Gaussian distributed (X_σ) with respect to the average path loss. Hence the pass loss can be finally expressed as:

$$L_p(d)\ dB = L_s(d_0)\ (dB) + 10n\log(d/d_0) + X_\sigma\ (\text{dB}). \quad (4\text{-}3)$$

The parameter X_σ depends on the actual environment (such as urban or suburban areas). For simplicity, the large-scale fading margin is assumed to have a fixed value of 8dB.

In addition to path loss, there is also a small-scale fading phenomenon, which can usually be described as Rayleigh fading [77]. The actual attenuation of the signal could be 10dB less or 20dB higher than the path loss.

Assume that the overall signal dynamic range of -100dBm~-20dBm is actually obtained with the above loss and fading present; the corresponding signal power diagrams are demonstrated in Figure 4-3 and Figure 4-4. The *average signal power* is assumed to be normally distributed within the range of *maximum and minimum average power* calculated in these figures (-72dBm ~ -38dBm); hence the final signal power level distribution can be obtained.

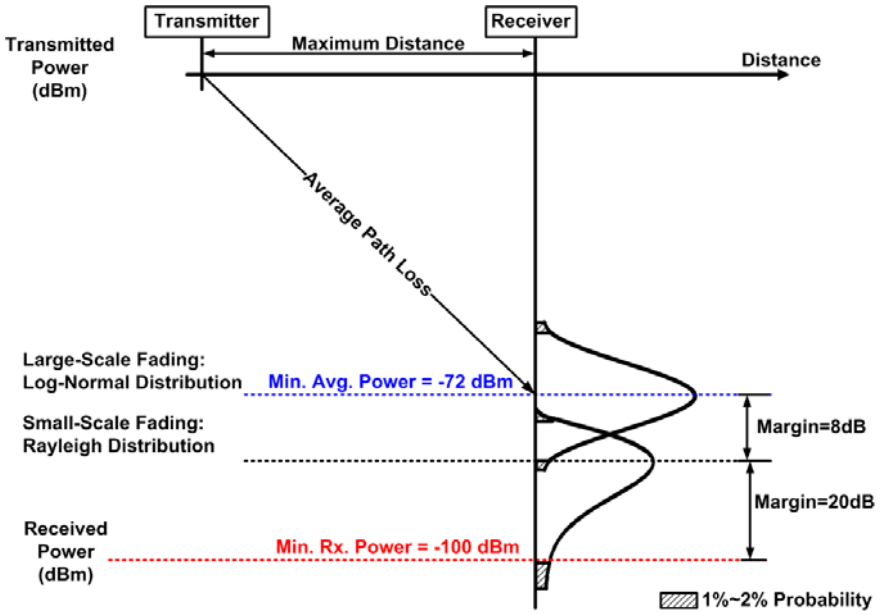


Figure 4-3 minimum signal power modelling

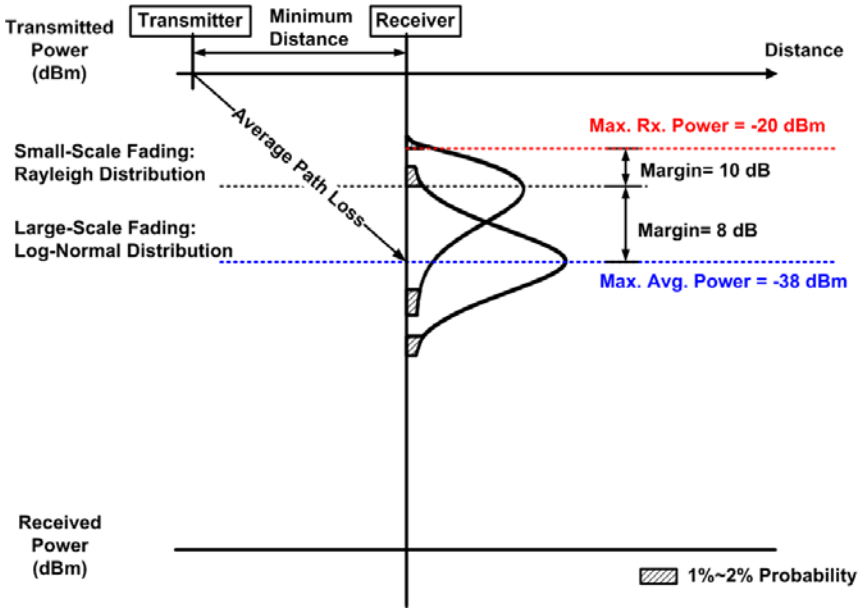


Figure 4-4 maximum signal power modelling

Given this information, the input signal power spectrum can then be randomly generated. In fact, by modelling each channel as a sine-wave with average distributed phase, simulation shows that the average total signal power of 500 channels is about -23dBm, referred to 50Ω resistance, and is equivalent to an RMS voltage of -36dBV_{rms}. An example of the signal spectrum is provided in section 4.4. Note that this is due to the fact that very high power signals are seldom present. Very weak signals also have a low possibility to be present. Therefore, the sensitivity of the proposed spectrum monitor is set to -85dBm/200kHz. In fact, since feature detection algorithms have the potential ability to detect signals with -20dB SNR [3], it is expected to be able to recognize signals as low as -105dBm, which covers most communication systems' sensitivity. Here, the concept of the 'sensitivity' of the spectrum monitor means that if the input power level is equal to or higher than -85dBm/200kHz, then the spectrum monitor will recognize the channel as occupied.

The gain requirement of the spectrum monitor is different from other radios. In most communication systems, variable gain is applied according to the received signal power. Strong signals are only amplified a little in total, while the receiver will switch to high gain modes for weak signals. In this way, the dynamic range of the ADC's input signal is reduced from the original signal range, and thus the ADC's resolution can be reduced as well. In the spectrum monitor, however, large variations of signal levels are present at the same time at the input, so the gain of the RF blocks must be limited to avoid saturating the baseband circuits. At the same time, it must also be high enough at each stage to suppress the noise from later blocks. In fact, only a small range of moderate gains is expected be needed to be able to satisfy these requirements; hence a variable gain amplifier, which is usually required in baseband and sometimes in RF stages, is no longer an essential block.

For wideband signals, an important issue is the peak-to-average-power-ratio (PAPR). For simplicity (it is difficult to generate different modulated signals over wideband frequency), by modelling the channels as ideal sine-waves, with a certain amplitude distribution (which is the same as signal power distribution), and an average distribution for the phase, the PAPR versus number of channels is modelled by MATLAB simulations, as shown in Figure 4-5. For 500 channels, the fitted curve of PAPR is 9.5dB, and is set to 10dB in the following discussions.

Assume that an ADC with a full scale of $1V_{pp}$ is adopted in the system, which is equivalent to -6dBV_p peak voltage. Also assume that the sub-band of interest is a busy

band, which means that all the channels have some signal present. Then the peak voltage of 500 busy channels in one sub-band is equal to -36dBV$_{rms}$+PAPR=-26dBV$_p$. The front-end gain is then expected to have -6dBV$_p$-(-26dBV$_p$) =20 dB to amplify the in-band signal to ADC's full scale. Allowing 2dB margin from full scale, the front-end gain could be set to 18dB.

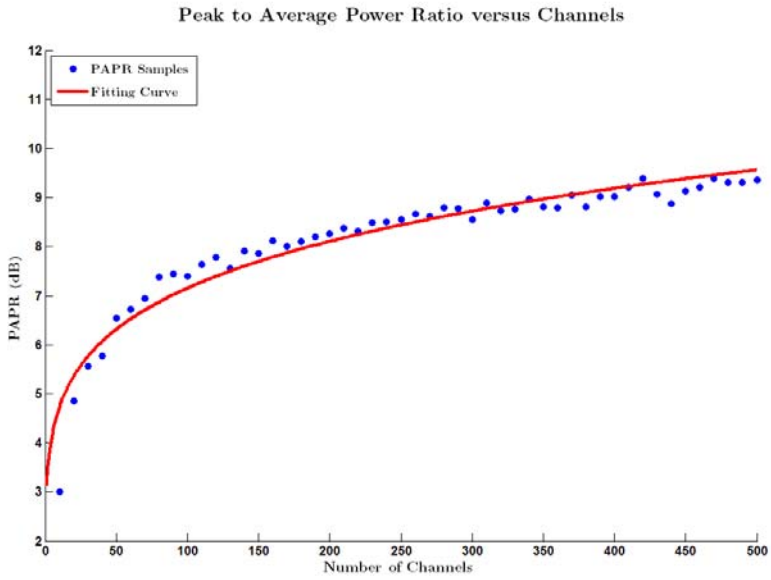


Figure 4-5 signal PAPR (Peak to Average Power Ratio) versus number of channels

In narrowband systems, the signal should be amplified by the front-end gain before the input stage of ADC. However, in the case of the spectrum monitor, desensitizations and cross modulations are expected to be introduced because of the large number of in-band strong blocking signals. This reduces the apparent gain for weak signals. Hence a 3dB gain-reduction for a -85dBm/200kHz input signal is a reasonable assumption. Therefore, signal power level (-85dBm/200kHz) of the weakest target could be amplified by only 15dB by the front-end section, and at the input stage of the ADC reaches -70dBm/200kHz. Assume the ADC gives 0dB gain and arbitrarily allow another 3dB as decision margin. Then the noise level at the output of ADC, which is the sum of the noise from the front-end and the ADC's quantization noise, should be lower than -73dBm/200kHz. The white noise level in each 200kHz channel at the input of the front-end will be -174dBm/Hz+10×log(200kHz) =-121dBm, which will be amplified to the level of -103dBm/200kHz by the front-end gain. Assume the ADC's quantization noise level is set to -74dBm/200kHz (which is 1dB lower than the required total noise level at the output of the ADC), and the total quantization noise over the 100MHz bandwidth would be -74dBm+10×log$_{10}$(500)=-47dBm. Because the full-scale (1V$_{pp}$) of the ADC is +4dBm with respect to a 50Ω resistance, the SNDR of

the ADC is then determined as +4dBm-(-47dBm)=51dB, and the equivalent ENOB is calculated as (51-1.76)/6.02=8.2dB. The noise figure of the front-end could be configured as 20dB, which is a much relaxed specification compared to that of most other wireless receivers. The noise level at the output of the front-end (input of the ADC) would then be -121+20+18=-83dBm/200kHz, which is much lower than the ADC's quantization noise level.

High linearity is required to make sure that interferers do not give rise to false measurements of significant signal power in channels that are actually empty, and lead to false 'occupied' indications. The wideband spectrum sensing approach makes this requirement even more important, because there could be so many in-band interferences that might accumulate false inputs at the wanted channel frequency. The main requirement is that the intermodulation products should be limited to be below the noise level, and the blocking effect on weak signals should also be negligible. Note that interferers could be from anywhere within the 3GHz band at the input stage of the receiver. The concept of 'in-band' changes as one moves along the signal path of the receiver chain, and is reduced due to filtering, until finally reduced to 100MHz at the end of the front-end. Therefore, the IIP3 specification is highly dependent on the receiver architecture and cannot be easily specified without reference to the architecture. Hence this issue will be discussed after the receiver architecture's introduction. Assume that the intermodulation products at the output of the front-end are allowed to be at the same as the noise level, that is, -83dBm in a 200kHz channel. Therefore, the total noise and distortion level in a 200kHz channel at the output of the ADC is the sum of -74dBm (ADC's quantization noise), -83dBm (front-end's output noise floor) and -83dBm (front-end's output intermodulation product level), which is equal to the targeted level of -73dBm/200kHz.

All the above analysis of the gain, noise, linearity and dynamic range is summarized in Figure 4-6.

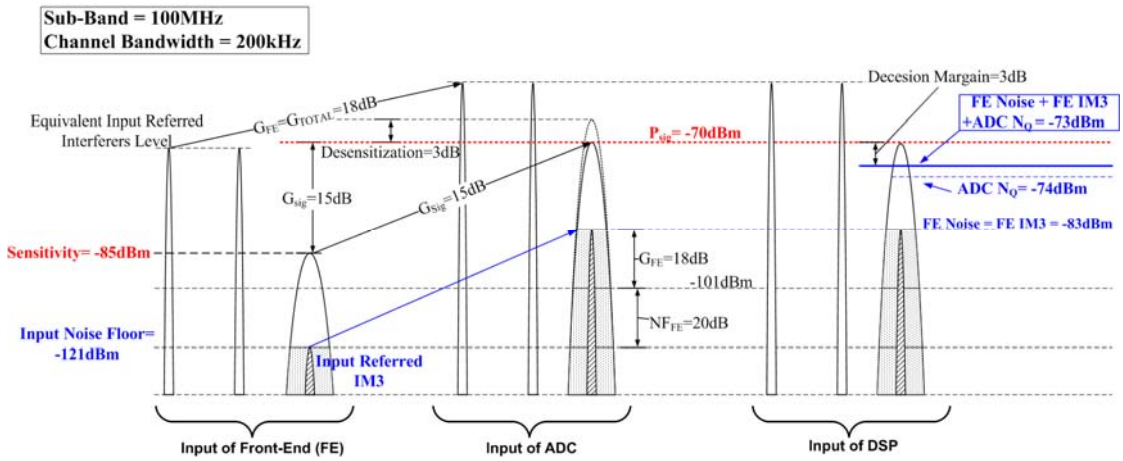


Figure 4-6 Cascade performance (100MHz Sub-band)

Another important issue is image rejection. An image response could arise either from a single sideband mixing (real mixing) stage, or the mismatch between quadrature complex branches. In the spectrum monitor, it is more likely to have image signals on top of in-band wanted signals because of its wideband configuration, no matter what architecture is adopted. Because the sensitivity of the receiver is configured (as mentioned before) to -85dBm/200kHz, the image rejection ratio (IRR) is ideally over 65dBc when an in-band strong signal of -20dBm/200kHz is present, which is a tough requirement. In practice however, because very strong signals such as -20dBm/200kHz are seldom present (In fact, as mentioned before, the average total power in a 100MHz band is only about -23dBm), an IRR of 55dBc could be acceptable in most cases, which means that the image response of a -30dBm/200kHz signal can be suppressed below the receiver's sensitivity level. For real mixing, achieving this IRR is related to the degree of filtering of the image signal prior to the mixer signal, or by selecting a good frequency plan that avoids the overlap of the image and wanted signal. For complex mixing, achieving good IRR depends on minimizing the imbalance between the I/Q branches; in this case a fixed, lower IF makes it easier to achieve the specification.

Harmonic mixing in a wideband architecture is another systematic problem that needs some careful attention. Harmonic mixing appears when the mixer operates in switching mode, i.e., with an apparently non-sine-wave LO signal, which is normally the case for the reason of reducing the noise. The mixing converts the input signals close to the harmonic frequencies of the LO signal to the converted wanted signal IF region, and hence those signals overlap with wanted ones. Two common approaches can be used to address this. One can effectively remove the LO's harmonic

components, using harmonic rejection mixers, for example [78], or devise a frequency plan for the receiver that avoids input signals at harmonics of the LO.

4.2.2 Frequency Synthesizer

As well as the receiver chain, the frequency synthesizer's specification is also of considerable importance. There could be many candidate architectures, integer-N, fractional-N, or even mixing solutions. From the system level, no matter what frequency synthesizer architecture is finally selected, the essential requirement is to be able to tune over the 3GHz band with frequency steps of 100MHz. This is generally a tough requirement for LC oscillators. The phase noise requirement is also different from many common communication systems. Because there is no requirement for demodulation of the signals, it is not important to be concerned with low phase noise to avoid the reduction in SNR due to phase distortion caused by the phase noise to quadrature signals. Rather, the issue of most concern is the reciprocal mixing [30] effect. The phase noise has to be low enough to avoid the reciprocal mixing power of a strong in-band signal overwhelming an adjacent weak signal. To guarantee that the reciprocal mixing product is lower than -85dBm/200kHz (the receiver's sensitivity level), the phase noise of -65dBc/Hz at 200kHz offset frequency is needed for the strongest signal of -20dBm/200kHz. Note that this is the sum of all the LO phase noise contributions if more than one LO is adopted. As will be discussed later, the necessary number of LOs is expected to be at least two or three to complete the spectrum monitoring. Therefore, for a single LO, a phase noise much lower than -65dBc/Hz @ 200kHz offset is needed. In this chapter, this requirement is roughly set to a reasonable value of -80dBc/Hz over the entire 100MHz band of interest. This requirement is still much relaxed compared with most of the other common communication systems. Remember this is also the requirement for reference spurs.

4.3 Architectures Selection

4.3.1 Direct Conversion

According to the specifications, two candidate receiver architectures are to be introduced in this section: the up-down-down conversion architecture and dual-down conversion architecture. Before discussing these two options, a conventional direct-

conversion receiver architecture shown in Figure 4-7 is first analyzed to show the difficulty in meeting the system requirements.

In this architecture, the baseband variable gain amplifier is eliminated for the reason discussed in the last section. The baseband filter is assumed to have zero gain. Therefore, all the front-end gain is provided by the LNA and mixer, of which the input signals are wideband from 2GHz to 5GHz. With such a wide input frequency range, it is likely that there could be too many 'in-band' interferers in this range to achieve 18dB gain in total. Hence, some band selection must be done before the LNA, and this bandpass filter must be tunable over the entire band, which involves complicated and bulky passive circuits that are difficult to integrate, increasing the cost significantly. Next, the quadrature mixers must have good matching in order to reject the image signal to an acceptable level. However, it is very difficult to guarantee matching over the entire 2GHz~5GHz band. Lastly, the local oscillator needs to have a tuning range from 2GHz~5GHz; this is a very large fractional bandwidth, and is also difficult to implement. Furthermore, as is usual for zero-IF architectures, the DC offset is a problem for down-converted signals near DC. Consequently, this architecture is not considered suitable for the spectrum monitor application, and is not pursued further in this study.

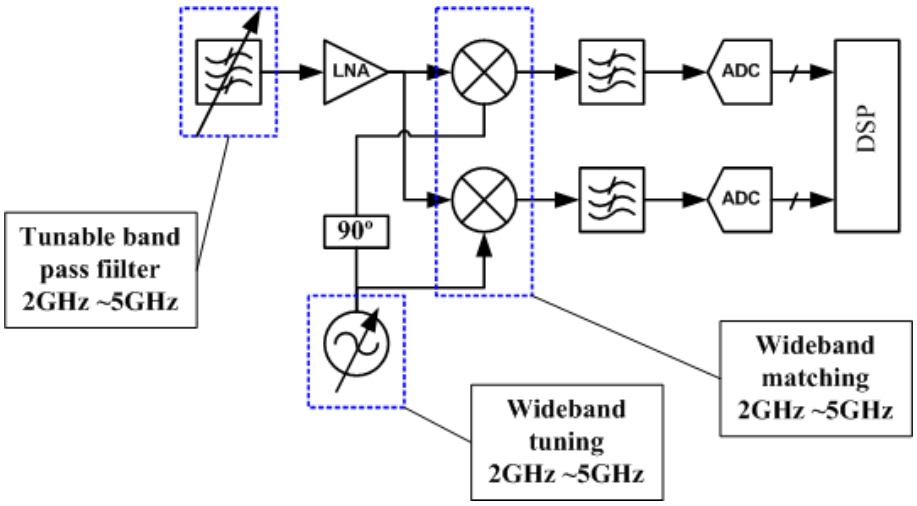


Figure 4-7 Direct conversion architecture

4.3.2 Up-Down-Down Conversion

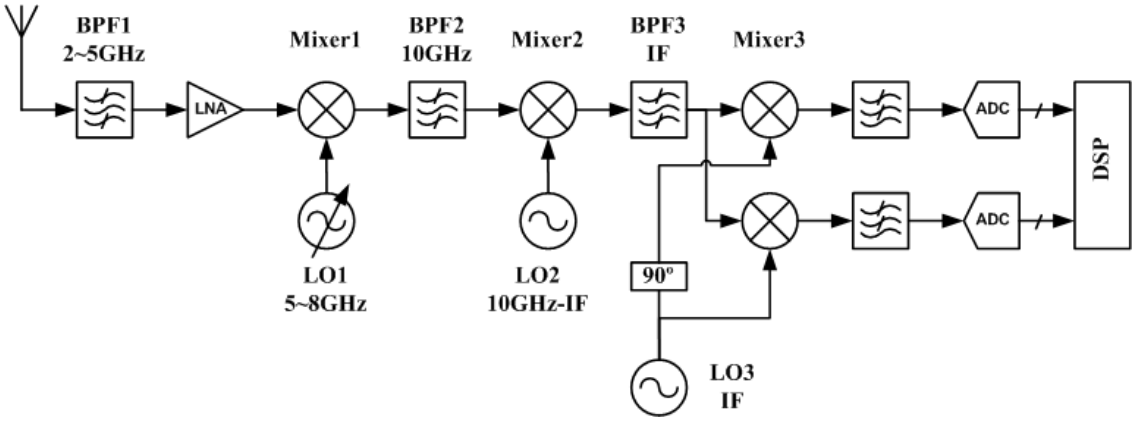


Figure 4-8 Up-down-down conversion architecture

Inspired by TV tuner designs, an up-conversion based architecture is proposed in Figure 4-8. To examine the operation, taking the example of receiving the sub-band from 4.9GHz~5GHz, the frequency conversion is illustrated in Figure 4-9. The frequency band of interest from 2~5GHz is first selected by BPF1, followed by a wideband LNA. The real mixer, Mixer1, then up-converts the wanted 100MHz sub-band to 10GHz through the first local oscillator, LO1, with a 5~8GHz tuning range. In this example, LO1 is 5GHz, so that the wanted signal is converted to 9.9GHz~10GHz. It can be observed that the 3rd order harmonic product is over 10GHz, and doesn't overlap with any wanted signal. BPF2 does some filtering at 10GHz. Note that the fractional bandwidth of an on-chip LC bandpass filter usually cannot be lower than 10% because of the losses of an integrated planar inductor. This will be discussed in detail in chapter five. In this case, the actual bandwidth of BPF2 is expected to be about 1GHz. In TV tuners, the entire band is typically up-converted to around 1~2GHz, and this is usually followed by complex down-conversion to DC or low-IF for demodulation. Various techniques [79], including image rejection mixers and RF/IF polyphase filters, are involved to guarantee the image rejection ratio. This involves lots of extra active and passive components, increasing the cost and power consumption. For signals around 10GHz, it is even more difficult to achieve the matching and image rejection requirements. The effectiveness of an image rejection mixer depends on very low mismatch between corresponding components in the I and Q paths. These mismatches are getting worse at frequencies as high as 10GHz, partly due to the components themselves, and also due to the parasitics at such high frequencies, making

it difficult to realise reliable matched paths. Because of these demands, an alternative architecture was investigated.

Instead of complex down-conversion, Mixer2 is still a real mixer, down converting the wanted signal to an IF frequency. There are several considerations in the selection of the IF frequency. First of all, this IF frequency should be selected as low as possible to make it easier to minimise and compensate for the mismatch between the I and Q paths at the next complex down-conversion stage. This requires that the frequency of LO2 should be as close to 10GHz as possible. However, making LO2 too high would lead to image rejection problems due to real mixing. This is similar to the situation for a traditional Superhet architecture. The problem is worsened due to the fact that the filtering bandwidth of 10GHz for BPF2 is relatively wide (about 1GHz as mentioned above and will be further discussed in chapter 5). Higher order filters could be designed to have more out-of-band attenuation, and the small area inductors/capacitors at this high frequency won't increase the chip size significantly. However, more passive components in a filter bring greater insertion loss, and this is worsened when on-chip inductors are involved. As a reasonable assumption, the second IF frequency after the Mixer2 could be chosen around 2GHz. This means that the frequency of LO2 is about 8GHz, and that BPF2 must provide enough filtering at the image frequency of around 6GHz; these requirements will be discussed in detail in chapter five. With the IF set to 2GHz, the mismatch requirement, and hence the image rejection ratio of the following complex mixer, Mixer3, can be achieved with less difficulty. Note that when expressing the signal in complex form of $I(\omega)+jQ(\omega)$ after the real mixer, Mixer2, the frequency components due to positive *or* negative LO will be cancelled as shown in Figure 4-9. The wanted signal could also be further filtered by BPF3 to some extent at this intermediate frequency. However, at this low frequency of around 2GHz, the filter components' size could be increased significantly, compared with the 10GHz bandpass filter. Therefore high-order complicated filter structures are not very attractive. Instead, simple solutions, such as a resonator load at the output of Mixer2, might be a better choice. The final complex down conversion moves the signal to DC and is followed by two 50MHz low pass filters in I/Q paths.

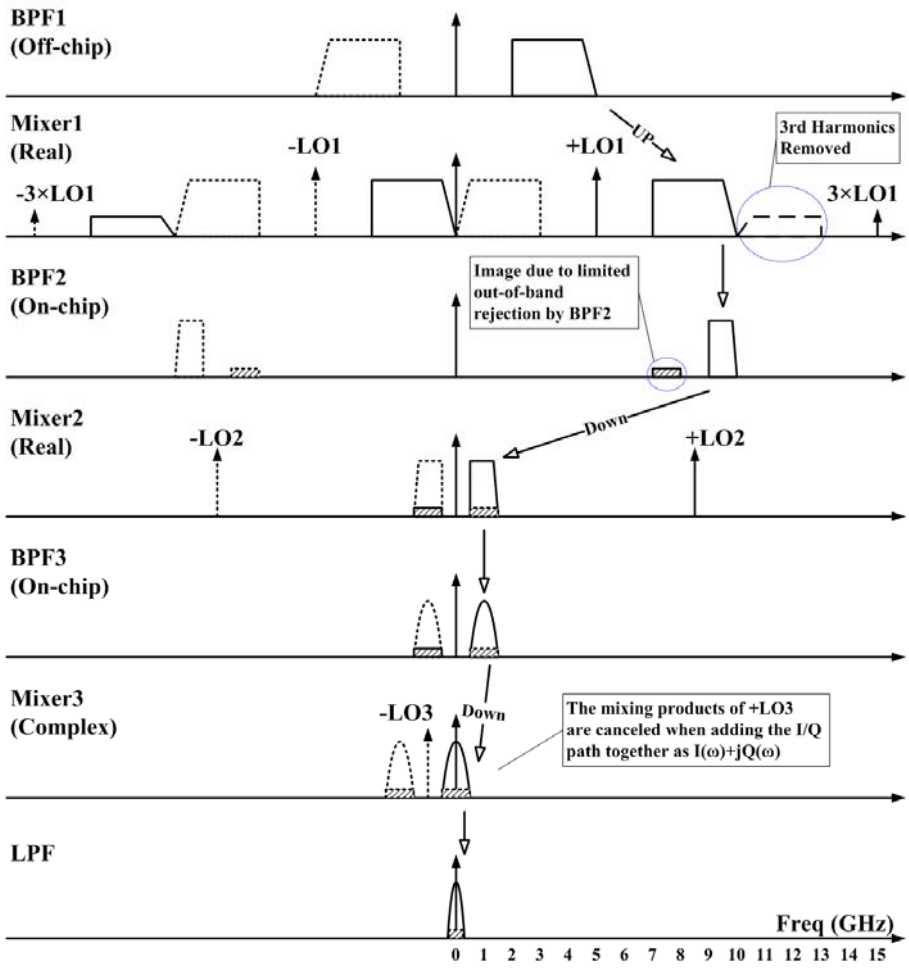


Figure 4-9 Up-down-down conversion frequency plan

The most important advantage of this architecture is to remove the odd order harmonics problem in the first wideband mixing stage. For example, if the entire 0~6GHz band is of interest, the tuning range of LO1 could be set to 6~12GHz, followed by a 12GHz BPF2. Also, the total gain of the receiving chain can be distributed among four stages, including WBLNA and three mixers, making it very flexible to configure the cascade NF and IIP3. For example, although the two on-chip filters are expected to bring losses and thereby raise the noise figure, the active stages in front of the filters can be set to have enough gain to suppress the loss. Also because of the losses, the total gain of the receiver chain in turn is not likely to exceed the limited gain requirement. For the IIP3, the inter-stage filters narrow the 'in-band' frequency range step by step while, the signal level alone stages gets higher by the applifications, hence the total intermodulation products could be limited.

In this architecture, apart from the first filter, BPF1, which is usually implemented off-chip, there are three mixer stages and two further integrated bandpass filtering stages. Hence one of the disadvantages of this system is that the area and

power consumption could be increased accordingly. Also, the 10GHz BPF2 is a design bottleneck, as a trade-off must be made between selectivity and possibly quite significant insertion loss.

In summary, this is a versatile architecture, and it is expected to have more flexibility to handle different situations, leading to potentially higher performance, while the high frequency selectivity is one of the main design challenges for on-chip filtering. Greater power consumption is also expected due to more stages, and many operating at high frequencies.

4.3.3 Dual-Down Conversion

Another possible architecture saves one mixing stage in the up-down-down architecture, and is hence called 'dual down conversion', as shown in Figure 4-10, its frequency plan is illustrated in Figure 4-11.

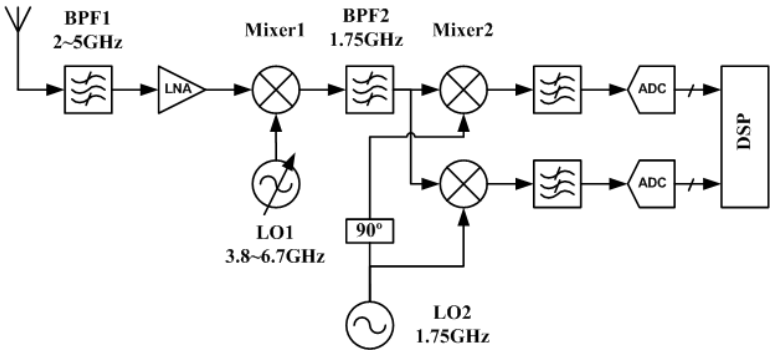


Figure 4-10 Dual down conversion architecture

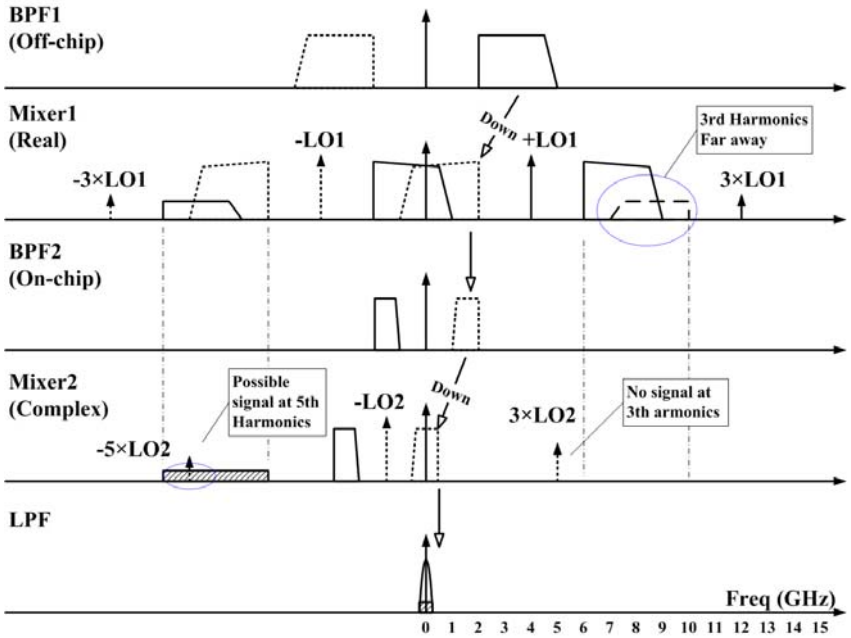


Figure 4-11 Dual down conversion frequency plan

Taking the example of detecting the sub-band of 2GHz~2.1GHz, the 2GHz~5GHz signal is firstly selected and amplified entirely as in the up-down-down conversion architecture. Mixer1 then down-converts the whole band to a lower frequency intermediate frequency. As can be seen in the frequency plan diagram, the down-conversion mixing introduces images that overlap with the entire band itself. However, by choosing the IF frequency carefully, it is possible to prevent the image from overlapping with the wanted 100MHz sub-band. For the 2GHz~5GHz band, the minimum IF frequency is 1.5GHz, corresponding to LO1 of 3.55GHz. In this case, the original 2~2.1GHz sub-band at negative frequencies is converted to 1.55~1.45GHz. Note that the channels' order is symmetrically reversed in the frequency domain, because this is converted from the negative frequency sub-band. The positive frequency sub-band of 4.9~5GHz is converted to 1.35~1.45GHz, thereby avoiding overlap with the wanted sub-band. The 3rd order harmonic of LO1 is also far away from the sub-band of interest. When higher frequency sub-bands are to be detected, the LO1 frequency is swept from 3.55GHz to 6.45GHz in 100MHz steps, and the resulting image signal and 3rd order harmonics are always far from wanted signal. After the first down conversion mixing, a bandpass filter is needed at the IF frequency. Because the IF frequency can be lower than 2GHz, the bandpass filter could have a bandwidth as low as 200MHz, providing better selectivity than the 10GHz bandpass filter. The area of the BPF2 is expected to be relatively large due to the requirement for a high linearity low power passive filter at the relatively lower frequency.

The complex mixer, Mixer2, then converts the IF signal to DC through the fixed local oscillator at the IF frequency, LO2. For complex mixing, the 5th order harmonics instead of 3rd order could be a problem. As shown in Figure 4-11, there is no signal present at the 3rd order harmonic frequency of LO2. However, if the 2~2.1GHz sub-band is down-converted to 1.55~1.45GHz, the 5th order harmonics of LO2 would be located at 7.5GHz , just within the upper side band frequency of 5.55~8.55GHz. Strong signals at this frequency would be down-converted to the DC in the Mixer2 stage, and corrupt the in-band signal. As LO1 increases from 3.55GHz in 100MHz steps, the upper sideband signals are also pushed to higher frequencies. The 5th order harmonic of LO2 stays within the upper sideband until LO1 reaches 5.55GHz (to the sub-band of 3~3.1GHz), when the upper sideband is moved to 7.55~10.55GHz. The 5th order harmonic level is ideally 13.98dB lower than the fundamental components, and therefore a -20dBm signal must be attenuated by at least 51dB by BPF2 at 7.75GHz

frequency in order to be guaranteed to be lower than the -85dBm receiver sensitivity; hence this corresponds to the higher frequency edge attenuation of the filter's specification. Furthermore, the poles at the Mixer1's output and Mixer2's input also give additional attenuation of the harmonics. Apart from this attenuation, the higher IF frequency is used to push the 5^{th} order harmonic higher, ideally out of band. By setting the IF frequency at 1.75GHz, the upper sidebands of the sub-bands within the 2.1~5GHz range are pushed away from the 5^{th} order harmonic of 8.75GHz, and thus the BPF2's higher frequency edge attenuation specification is relaxed. The nearest image is to the IF is located at 1.1~1.2GHz (original 4.9~5GHz sub-band), which is 500MHz away from wanted signal.

Another significant issue is the handling of interferers. At the output of Mixer1, the entire signal band and its image are overlapped from DC to 1.2GHz. According to classical intermodulation analysis, a 3dB higher signal power results in 9dB higher IM3 power. Hence, any intermodulation involving the frequency from DC to 1.2GHz produced by the next stage could be very strong, and this could corrupt the in-band signals. For example, when detecting the 2.0~2.1GHz sub-band, the frequency of LO1 is 3.8GHz. After the Mixer1 stage, the interferers originally at 3.1GHz and 2.6GHz are converted to 0.7GHz and 1.2GHz, respectively. These two signals could generate a 3^{rd} order intermodulation (IM3) product at the frequency of 1.7GHz, where the originally targeted 2.1GHz signal is present. In addition, another other two interferers originally at 4.5GHz and 5GHz are also converted to 0.7GHz and 1.2GHz respectively. Assume these four interferers' powers are all the same, then the actual IM3 product at 1.7GHz caused by these two pairs of interferers after the frequency conversion of Mixer1 is expected to be 9dB higher than that from only one pair of them present. For a fixed IF frequency, two solutions can be considered. The first is to make sure that the attenuation of the lower frequency edge of BPF2 is high enough to suppress the interferers. The second is to specify that the IIP3s of the active blocks are sufficiently high to reduce the potential intermodulation products. If these factors are limited, a higher intermediate frequency is a better choice, because the overlapped band could be narrower and it will also be attenuated more by BPF2 because they are further away from the wanted signal. Consequently, fewer intermodulation products would be expected.

This dual down conversion architecture is more compact than the up-down-down conversion architecture, and hence is expected to consume less power and to introduce

less insertion loss. The harmonic problems can be solved and the image problem of real mixing is avoided. Also, the difficulty of compensating mismatch is reduced due to the fixed lower IF. Another advantage is that the IF filter, BPF2, can achieve good selectivity because of the lower centre frequency, although it could occupy larger area for the same reason. The challenge is the lower edge frequency attenuation of BPF2 and higher linearity requirement of stages following Mixer1.

According to all the analysis above, the dual-down conversion architecture is believed to be the best candidate for an integrated 2~5GHz spectrum monitor.

In order to realise this design, there are two key techniques functions having specifications that are significantly different from the corresponding functions in other common radio architectures, specifically:

(1) An integrated narrowband filter, BPF2, is required at 1.75GHz.

(2) The first local oscillator, LO1, has a very wide tuning range from 3.8GHz~6.7GHz with 100MHz steps and a phase noise of lower than -80dBc/Hz for all the frequencies.

These detailed designs of these two blocks are discussed in chapters 5 and 6, respectively.

4.4 System Level Design

Before the budget analysis of the receiver chain, the spectrum occupancy needs to be generated randomly. When generating these signals, the 100MHz sub-band of interest is assumed to be a busy band, which means that all the 200kHz channels are occupied by signal. The 2.9GHz region outside the normal receiver band is unlikely to be absolutely quiet or busy. Hence, it is reasonable to assume that only a part of the channels with signal present, and the power in the other channels is purely white noise of -121dBm/200kHz. In this section, we assume that the targeted sub-band is 2GHz~2.1GHz, and that 30% of channels within the 2.1~5GHz region are assumed to be busy (An arbitrary but reasonable assumption). One example of the randomly generated signal spectrum is shown in Figure 4-12. Note that the ‘visual bandwidth’ of 10MHz means that the bandwidth of the displayed frequency bin is 10MHz, and the signal power is the sum of the signals falling into each 10MHz frequency bin. After generating the signals (according to Figure 4-3 and Figure 4-4 together with the associated discussions) within 2~5GHz band, the spectrum at the input stage of each

block is then determined by the gain (or loss), mixing and filtering effects of the previous stages, and could be obtained by simulation as well.

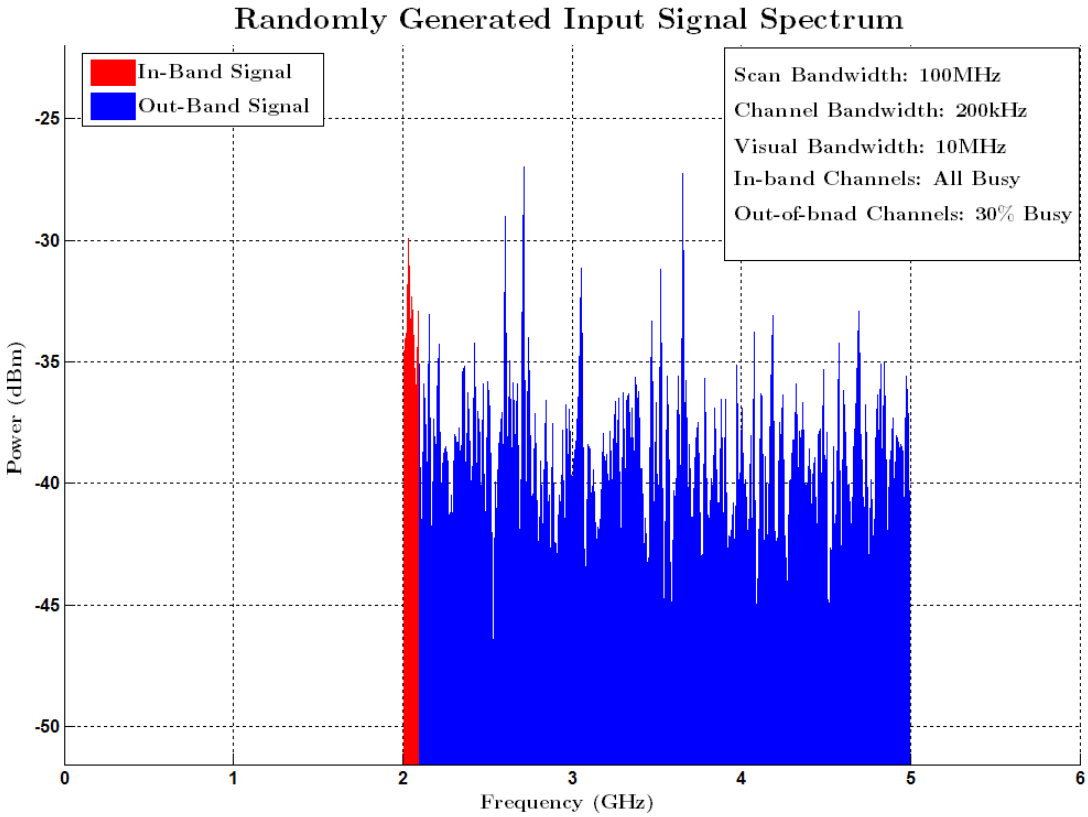


Figure 4-12 Example of input signal spectrum

With the spectrum monitor specifications in Table 4-1 and the dual down conversion architecture in Figure 4-10, the receiver chain's system level design is discussed in this section. The analysis generally includes gain, noise figure and linearity to achieve the proposed receiving dynamic range. Using the above mentioned configurations, the power consumption is to be predicted according to the FoM discussions presented in chapter three. The system level specification diagram is shown in Figure 4-13.

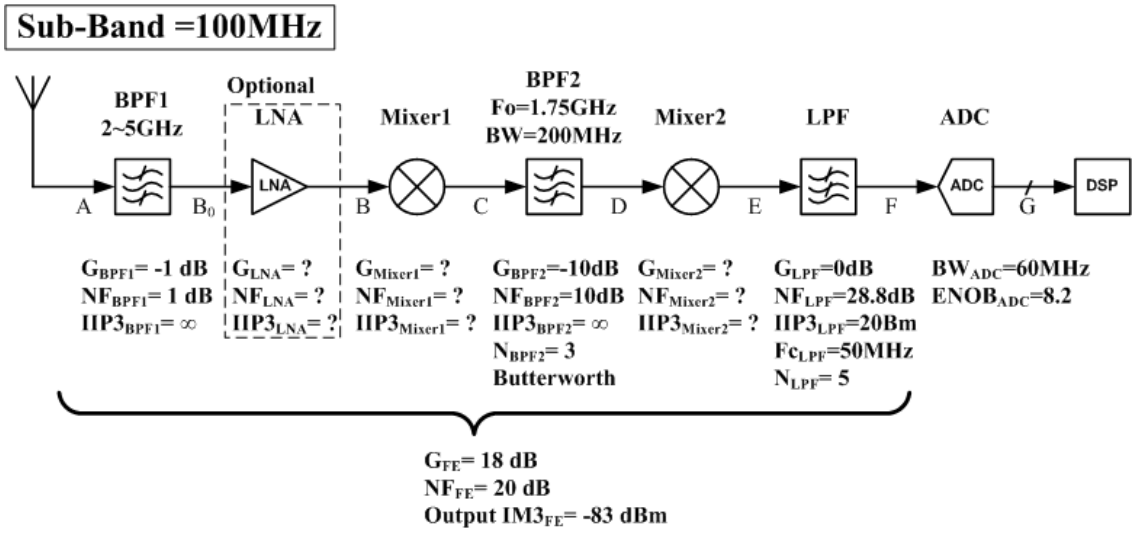


Figure 4-13 Dual down conversion system level configurations

To simplify the discussion, some assumptions are made before beginning the analysis. The filter BPF1 is usually implemented off-chip, and a realistic insertion loss of -1dB is assumed. The inter-stage 1.75GHz filter BPF2 should be on-chip, and -10dB insertion loss is expected because of the likely low Q factor of on-chip inductors. Good out-of-band attenuation is needed for the reasons discussed above. Therefore, the filter should have enough poles in the response, and hence an all-pole response such as Butterworth is preferred, so that the stop band is monotonic. At this frequency, the passive components, especially the inductors, are expected to occupy a large area; hence, the number of poles (hence usually the number of inductors), should be just enough to save area. The ADC and LPF's specifications are pre-determined, as demonstrated in the figure. The baseband lowpass filter doesn't provide gain, so that all the gain is provided in the RF and IF stages. Moderate noise and IIP3 are applied to the filter. According to the specification statistics obtained in chapter 3, typically, an input-referred noise voltage of $25\ nV/\sqrt{Hz}$ can be expected. By assuming the filter has infinite input resistance, the effective noise figure can be calculated as

$$NF_{LPF} = 1 + \frac{\left(25\ nV/\sqrt{Hz}\right)^2}{4kTR} = 28.8\ dB. \tag{4-4}$$

The Boltzmann constant k equals to 1.38×10^{-23}JK^{-1}, the temperature is assumed to be T=300K, and the filter's noise power is referred to a resistor of R=50Ω.

The IIP3 of the LPF is selected as a moderate value of 20dBm referring to the voltage across a 50Ω resistor. By making these assumptions, the system cascaded performances are determined by the configurations of LNA, Mixer1 and Mixer2.

There are two main reasons for using an LNA in a receiver: (1) to provide enough gain, and (2) to suppress the noise generated in the following stages. As the input signals of the LNA and Mixer1 are essentially wideband signal, a huge number of interferers could be present to generate intermodulation products on top of a single in-band 200kHz channel. The IM3 products generated by Mixer1 are much higher after the signals are amplified by the LNA. So intuitively, the LNA should have moderate gain, just enough to compensate for the lack of the mixer's gain and enough to control the overall noise figure, and low enough to prevent Mixer1 from generating too much and too high IM3 products. Consequently, instead of being a necessity in the signal chain, the LNA could be only adding some design freedom to optimize the overall performance and design, for example, removing the difficulty of designing high gain mixers. For the 100MHz sub-band case, the required noise figure of the front-end is relaxed to 20dB, as opposed to being less than 10dB for lots of other radio receivers. The total gain of the active blocks is 29dB. Therefore, for simplicity and for demonstration purposes, the LNA block is omitted from the following discussion.

The noise figure can be calculated through classical Friijs equation:

$$F_{total} = 1 + (F_1 - 1) + \frac{(F_2 - 1)}{G_1} + \frac{(F_3 - 1)}{G_1 G_2} + \cdots + \frac{(F_n - 1)}{G_1 G_2 \cdots G_{n-1}}. \qquad (4\text{-}5)$$

Note that, in the spectrum monitor, both real and complex mixing procedures generate single-side band noise figure, because the detected sub-band (100MHz) itself is not symmetrical.

The direct usage of the Friijs cascade IIP3 equation is not appropriate in the spectrum monitor receiver. The classical IIP3 calculation assumes two large in-band tones are always present before the precise channel selection by the baseband low pass filter. In this receiver, however, the concept of 'in-band' changes over the various stages. For LNA and Mixer1, 'in-band' means within the 3GHz bandwidth, while for Mixer2, 'in-band' is equal to the bandwidth of BPF2. Therefore, the calculation of IIP3 should be performed in a different way. In the following discussion, assume that all the non-linearity products are generated at the input stage of each block, which generally means at the input trans-conductance of the MOS devices.

The total output IM3 power of the front-end, $OIM3_{FE}$, is the sum of the IM3 power produced by each non-linear block, including Mixer1, Mixer2 and LPF. The design goal is to ensure that the value of $OIM3_{FE}$ is lower than -83dBm/200kHz.

$$OIM3_{FE}[dBm] = IIM3_{Mixer1} + (G_{Mixer1} + G_{BPF2} + G_{Mixer2} + G_{LPF}) + IIM3_{Mixer2} + (G_{Mixer2} + G_{LPF}) + IIM3_{LPF} + G_{LPF}. \quad (4\text{-}6)$$

The unit of IIM3 is dBm and the unit of gain is dB. In this equation, the total gain from Mixer1 to LPF is 19dB, given the gain of 18dB from the front-end chain and 1dB loss from the BPF1. According to the assumptions of G_{BPF2}=-10dB and G_{LPF}=0dB, the two mixers should provide a total gain of 29dB, and hence G_{Mixer2}=29-G_{Mixer1}. Therefore, equation (4-6) is reduced to

$$OIM3_{FE}[dBm] = IIM3_{Mixer1} + IIM3_{Mixer2} + IIM3_{LPF} + 48 - G_{Mixer1}. \quad (4\text{-}7)$$

The IIM3 power is a function of the block's linearity specification, IIP3, given a pair to interferers, when the power of each interferer is assumed to be equal to P_{INT}[dBm].

$$IIM3[dBm] = f(IIP3[dBm]) = -2 \times IIP3[dBm] + 3 \times P_{INT}[dBm]. \quad (4\text{-}8)$$

Due to the potentially huge number of interferers within the wide bandwidth, it's not appropriate to assume only a pair of interferers being present. Instead, it is more meaningful to obtain the *equivalent* two tone interferers, according to the actual spectrum occupancy. For each block, by randomly generating the spectrum at the input stage, the sum of the IM3 product power falling into a certain in-band channel can be calculated, given a certain IIP3 specification. This procedure can be repeated many times so that an average IIM3 can be obtained. With this average IIM3 product power and the given IIP3, the equivalent interferer power, P_{INT} can then be calculated in Equation (4-9), which will then be used in further IM3 product calculation using Equation (4-8).

$$P_{INT}[dBm] = \frac{IIM3[dBm] + 2 \times IIP3[dBm]}{3}. \quad (4\text{-}9)$$

Starting from the first block, Mixer1, the input signal is simply 1dB lower than the original signal because of the insertion loss of BPF1. Hence Mixer1 faces very large numbers of weak interferers. Note that the selectivity of BPF1 is assumed to be ideal, which is equivalent to the infinite out-of-band attenuation. The equivalent interferers' power at the input of Mixer1 can be obtained according to the simulation as:

$$P_{INT\_Mixer1}[dBm] = -31.5\ dBm\,. \quad (4\text{-}10)$$

After Mixer1, the interferers' power is expected to be amplified to some moderate level. Further, the non-ideal on-chip filter BPF2 can only achieve moderate filtering and only attenuates the out-band interferers to a limited extent. Hence the main

interferers are expected to be at moderate power levels within several hundred MHz around the IF of 1.75GHz. If BPF2 is assumed to have a 3rd order Butterworth response with a fixed 10dB insertion loss, the equivalent interferer power at the input of Mixer2 can be obtained as a function of Mixer1's gain:

$$P_{INT\_Mixer2}[dBm] = G_{Mixer1} - 46.6\,. \tag{4-11}$$

A small number of strong interferers are present at the baseband filter stage, after the Mixing and amplification of the previous stages. The simulated equivalent interferers' power at the input of the filter, LPF, is obtained by simulation as:

$$P_{INT\_LPF}[dBm] = -15.62\ dBm. \tag{4-12}$$

Therefore, given the IIP3 specifications and the gain of each stage, the input referred 3rd order intermodulation product power falling into a wanted signal channel at each stage can be calculated using Equation (4-8), and the total OIM3 contribution at the output of front-end can then be obtained using Equation (4-7). The design goal is to limit the total $OIM3_{FE}$ to be lower than -83dBm/200kHz.

So far, with the gain, noise figure and IIP3 specifications of each block, the cascaded performance can be obtained. The optimum configuration with respect to power consumption can then be found.

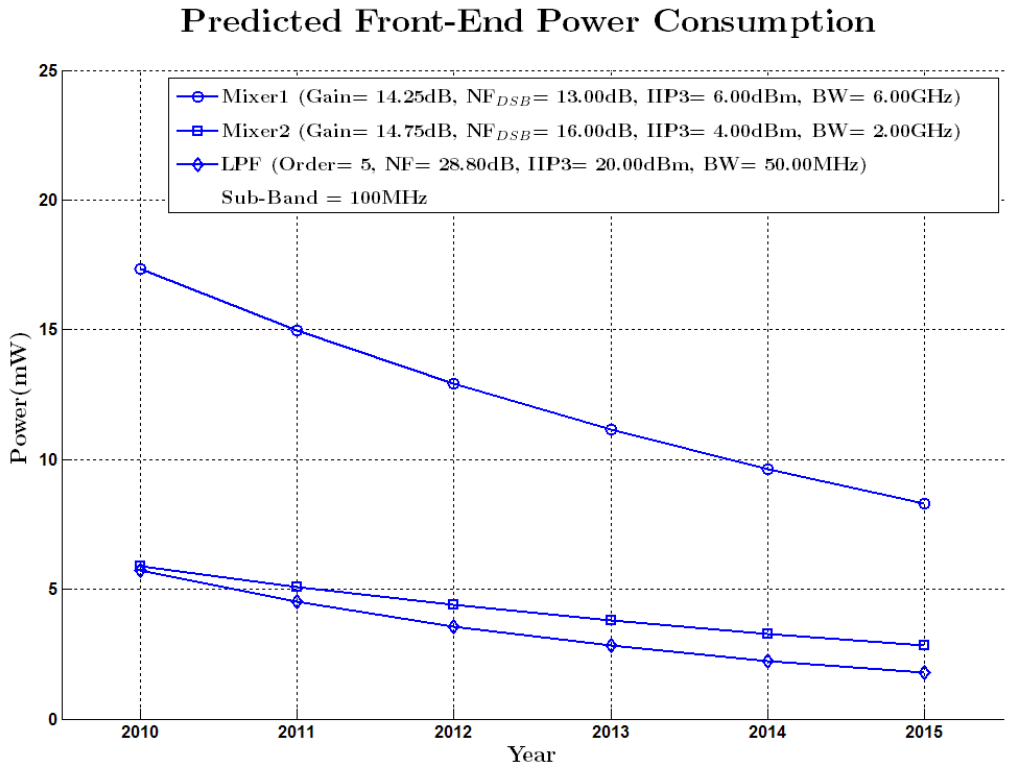


Figure 4-14 Predicted power consumption of front-end blocks (100MHz sub-band)

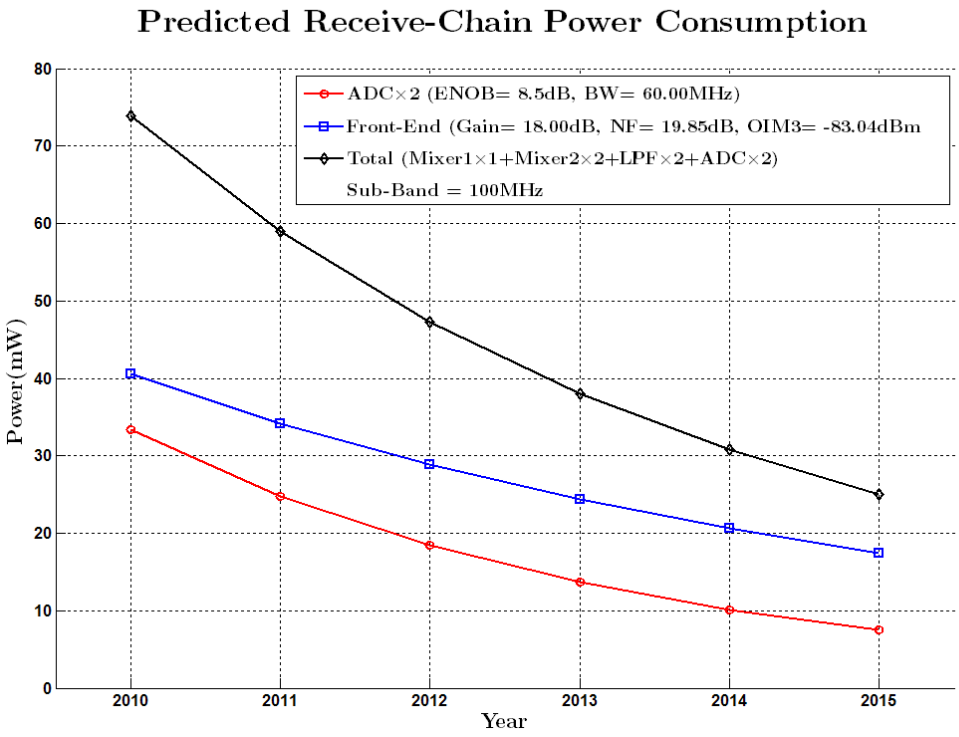


Figure 4-15 Predicted power consumption of FE+ADC (100MHz sub-band)

The power consumption of the front-end is predicted according to the methods described in chapter three. The minimum power consumption is then selected among the avaliable configurations and the predicted achievable power for each of the main blocks is shown in Figure 4-14. The total power consumption of the receiver chain is thereby calculated and shown in Figure 4-15. The gain, NF_{DSB} and IIP3 of Mixer1 are 14.25dB, 11dB and 6dBm, respectively. These specifications of Mixer2 is correspond to 14.75dB, 16dB and 4dBm, respectively. A pair of 5^{th} order lowpass filters has the input referred noise of 25 $nV/\sqrt{Hz}$ (equivalent to 28.8dB noise figure referred to 50Ω resistor), an IIP3 of +7dBV (equivalent to +20dBm IIP3 referred to 50Ω resistor) and a cutoff frequency of 50MHz. The resolution of the ADC is selected as 8.5bits, and the bandwidth is selected as 60MHz, which are slightly better performances than the minimum requirements.

In the year 2010 (2010 will be over when this thesis is finally hard bounded), the power consumption of Mixer1, Mixer2, LPF and ADC are expected to be 17.3mW, 5.9mW, 5.7mW and 13.7mW, respectively. Hence the total power consumption is approximately 68mW. By the year of 2015, the power consumption of Mixer1, Mixer2, the LPF and ADC are estimated to be reduced to 8.3mW, 2.8mW, 1.8mW and 3.1mW, respectively, resulting a total power consumption of about 23.7mW. The system block's specifications and power consumptions are listed in Table 4-3 for reference.

The above system level simulation results show that designing a low power spectrum monitor receiver is possible using present day technology, and that the power consumption is expected to be reduced significantly in the near future.

Spectrum Monitor Sub-band Resolution		100MHz
Mixer1	Gain	14.25 dB
	NF_{DSB}	11 dB
	IIP3	6 dBm
	Power	8.3 mW
Mixer2	Gain	14.75 dB
	NF_{DSB}	12 dB
	IIP3	4 dBm
	Power	5.9 mW × 2
LPF	Order	5
	BW	50 MHz
	NF	28.8 dB (25 $nV/\sqrt{Hz}$)
	IIP3	+20dBm (+7 dBV)
	Power	1.8 mW × 2
ADC	ENOB	8.5
	BW	60MHz
	Power	3.8 mW × 2
Front-End	Power	17.5 mW
Total	Power	25.1 mW

Table 4-3 Spectrum monitor configurations and power consumptions by the 2015

4.5 Summary

In this chapter, the spectrum monitor for cognitive radio application is investigated. The specifications of the receiver are first analysed, including the front-end, ADC and frequency synthesizer. Then the candidate receiver architectures are discussed, drawing the conclusion that the dual-down conversion architecture is the most suitable architecture for low cost and low power design. The system level design is then performed and simulated to achieve the proposed specifications. The power consumption of the receiver chain is finally predicted using the FoM strategy discussed

in chapter three. As a conclusion, this chapter shows the feasibility of designing a low cost, low power spectrum monitor receiver for portable cognitive radio application in the near future.

Chapter 5 High Frequency Integrated Passive Band Pass Filters

5.1 Introduction

As discussed in chapter four concerning the spectrum monitor architecture, there are two possible approaches for the spectrum monitor: up-down-down conversion and dual-down-conversion architectures. Both of these require passive filters after the mixing stages. For the target input signal range used, the up-down-down-conversion architecture requires a bandpass filter with a centre frequency of 10GHz while that for the dual-down-conversion architecture is 1.75GHz. The ideal bandwidths of the filters are 100MHz according to the system level requirements. There are several solutions for filtering in a receiver design. Most communication systems adopt off-chip high frequency filters, in order to achieve low insertion loss so that the overall front-end noise figure requirement is met for the demodulation purpose. However, off-chip bandpass filters occupy large board area and increase the cost. In this project, the spectrum monitor needs to be small and cheap enough compared with the main transceiver design to avoid too much extra cost for the whole cognitive radio. Therefore, an on-chip bandpass filter solution is required. For implementation using standard CMOS technology, the filter could consist of either a transmission line or lumped components. Because a transmission line's size is usually comparable to the wavelength, a filter implemented using this method for frequencies below 10GHz will have an area of the order of square centimetres, which is too large for a chip design. On the other hand, a lumped component filter has a much smaller area and lower cost,

while the main drawback is the large expected loss. However, in this project, the spectrum monitor can tolerate much more noise than other communication systems, and hence the lumped component solution is worthy of investigation.

Because of the inconvenient design of conventional bandpass filter solutions, as will be explained in the next section, a series-coupled-resonator topology has been chosen. As the most important element in a filter, inductor modelling on silicon is first described. After that, both the bandpass filters at 10GHz and 1.75GHz frequencies are designed. Some novel modifications are then applied to these two bandpass filters to improve the bandpass filter's performance or reduce the difficulty of implementation.

The filters are simulated with the ADS Momentum CAD suite, and then implemented on 130nm standard CMOS technology. On chip measurements are made using ground-signal-ground-signal-ground (GSGSG) probe pairs and a vector network analyser.

5.2 Bandpass Filters Topology

The straightforward way of realizing a bandpass filter is based on network element transformation techniques. A normalized lowpass filter with a certain type of response can be de-normalized by frequency and then converted to a bandpass filter by means of an element impedance transformation. However, some drawbacks exist in this transformation, particularly with unrealistic element values, making it impractical for the required high frequency bandpass filter design. Therefore some other topology must be chosen to overcome these problems. In this section, the conventional bandpass filter is firstly discussed, followed by an introduction to the topology to be used, the 'coupled resonator bandpass filter'. The advantages and disadvantages of both synthesis methods are analyzed and compared in the discussion.

5.2.1 Conventional Bandpass Filter Topology

A bandpass filter is conventionally synthesized from a normalized lowpass filter, which usually involves two transformation techniques: frequency transformation and impedance transformation. The frequency transformation converts the normalized frequency Ω to the actual centre frequency ω_0 and the bandwidth. This transformation equivalently converts an inductor to a series combination of inductor and capacitor, while converting a capacitor to a parallel combination of capacitor and inductor. The

bandpass filter obtained has a normalized source and load impedances of 1Ω, and hence the impedance transformation converts all the components values with respect to actual source/load impedances, which are often 50Ω for RF applications. An example of the low-pass to band-pass transformation is shown in Figure 5-1, where a 3rd order Butterworth filter with a centre frequency of 10GHz and 3dB bandwidth of 1GHz is synthesised.

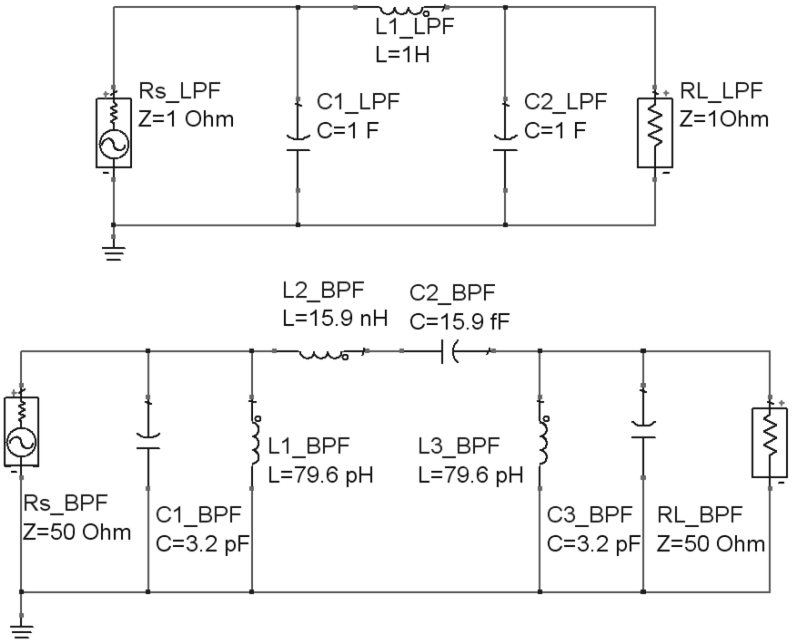


Figure 5-1 Convention bandpass filter synthesis

One of the most important disadvantages of this standard synthesis method is the large variation of the element values. In the above example, the capacitor values of the shunt and series resonator are 3.2pF and 15.9fF, respectively, while the corresponding inductor values range from 80pH to 16nH. As will be discussed later, the inductance is directly related to the inductor size. In practice, a spiral inductor of 1nH could occupy 200μm×200μm with a reasonable quality factor (10~20) on modern CMOS technology. Hence an inductor on the order of 10nH is expected to occupy too much silicon area from the cost point of view. Such a large inductor could also introduce lots of parasitic capacitance and resistance loss. On the other hand, the smallest capacitor is usually very difficult to implement accurately. Besides, the variation of element values is three orders of magnitude, which leads to a lot of difficulty of tuning because of the sensitivity of the values.

5.2.2 Coupled Resonator Bandpass Filter Topology

A coupled bandpass filter is an approximate narrow bandpass filter technique [80]. Essentially, the approximations are accurate when the fractional bandwidth is less than 5% and remains almost accurate within 20% of the fractional bandwidth. Since the bandwidth is very narrow in our application, this bandpass filter realization method can be considered as accurate. There are two types of coupled resonator bandpass filter: series coupled shunt resonators and shunt coupled series resonators, as shown in Figure 5-2. Both filters are the duals of each other. The coupling elements can be either capacitors or inductors. However, due to the much larger losses of inductors and the requirement to reduce the number of inductors, the coupling elements are chosen as capacitors, as shown in Figure 5-2.

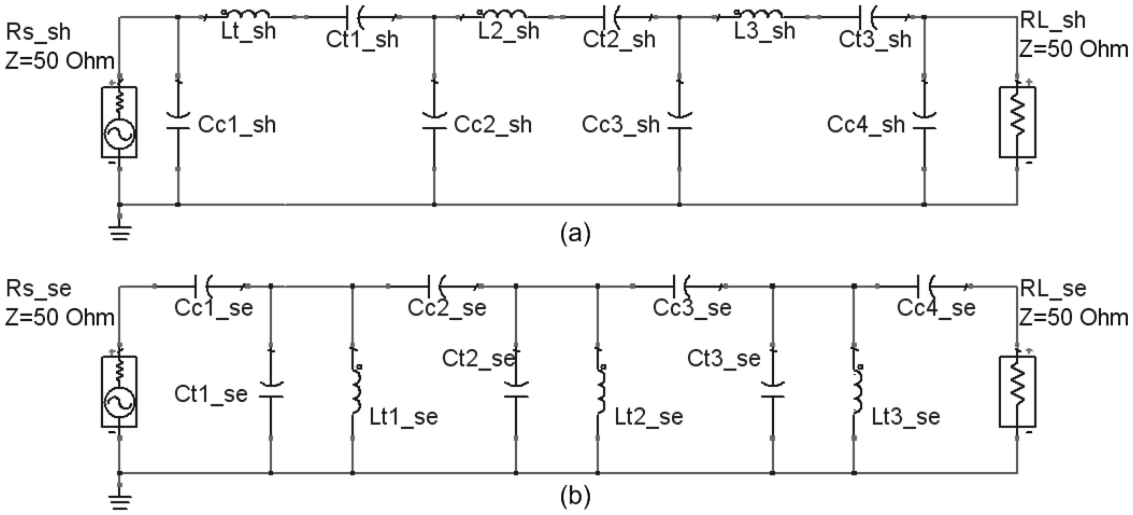


Figure 5-2 Filter types: (a) shunt-coupled series resonator, (b) series-coupled shunt resonator

Rather than performing the selectivity by means of both series and shunt resonators in the classical way, the selectivity can be only performed by means of series or shunt resonators coupled together. The coupling elements are called *Impedance (K)* inverters for a shunt coupled filter and *Admittance (J)* for a series coupled filter, which reduce themselves to a single series or shunt reactance. In principle, each coupling capacitor (Cc) is determined by the frequency response form of the prototype low pass filter, while each resonator is tuned at the same frequency. The tuned capacitor in the resonator value (Ct) equals the original tuning capacitance, excluding the effect of the coupling capacitance. The inductor of each resonator can be chosen as the same value to simplify the design.

Note that only all-pole filter responses are suitable for this prototype. Therefore, Butterworth, Chebyshev type I, Bessel and Gaussian responses could be selected, but

responses having stop band zeros such as Elliptic filters cannot be used for this type of filter.

The synthesising of coupled-resonator filters are discussed as below.

- Step 1: Define the initial requirements for the filter

Set the specifications by choosing the centre frequency (F_0), bandwidth (BW), filter order (N), filter response (Butterworth, certain ripple Chebyshev, Bessel etc), source and load impedance (R_S and R_L) or conductance (G_S and G_L), inductance of resonators (L), and the unloaded Q factor of the inductors. An example of these specifications and configurations is listed in Table 5-1. Note that the normalized filter components values g_0~g_4 correspond to the source resistance (R_S_LPF), the first shunt capacitor (C1_LPF), the series inductor (L1_LPF), the second capacitor (C2_LPF) and the load resistance (RL_LPF) of 1Ω, 1F, 1H, 1F and 1Ω in the standard low pass filter structure in Figure 5-1.

F_0	10 GHz	R_S	50 Ω
BW	1000 MHz	R_L	50 Ω
N	3	L	1nH
Type	Butterworth	Q_{ind}	15

g_0	g_1	g_2	g_3	g_4
1.000	1.000	2.000	1.000	1.000

Table 5-1 Initial filter specifications and configurations

- Step 2: Derive the coupling capacitors' values

First, calculate the coupling factors, K_i (or J_i), for $i=1...n+1$,

Note that G_S, G_L are the source and load conductance, and Q_{bp} is the quality factor of the bandpass filter, or the fractional bandwidth of the filter.

$$K_i = \begin{cases} \sqrt{\dfrac{R_S \cdot \omega_o \cdot L \cdot Q_{bp}^{-1}}{g_{i-1} \cdot g_i}} & (i = 1) \\ \dfrac{\omega_o \cdot L \cdot Q_{bp}^{-1}}{\sqrt{g_{i-1} \cdot g_i}} & (2 \le i \le n) \\ \sqrt{\dfrac{R_L \cdot \omega_o \cdot L \cdot Q_{bp}^{-1}}{g_{i-1} \cdot g_i}} & (i = n+1) \end{cases} \quad \text{, for a shunt coupled filter.} \qquad (5\text{-}1)$$

$$J_i = \begin{cases} \sqrt{\dfrac{R_S^{-1} \cdot \omega_o \cdot L \cdot Q_{bp}^{-1}}{g_{i-1} \cdot g_i}} & (i = 1) \\ \dfrac{\omega_o \cdot L \cdot Q_{bp}^{-1}}{\sqrt{g_{i-1} \cdot g_i}} & (2 \le i \le n) \\ \sqrt{\dfrac{R_L^{-1} \cdot \omega_o \cdot L \cdot Q_{bp}^{-1}}{g_{i-1} \cdot g_i}} & (i = n+1) \end{cases} \quad \text{, for a series coupled filter.} \qquad (5\text{-}2)$$

According to the coupling factors obtained, the values of the coupling capacitors, Cc_i, can be calculated for $i=1...n+1$,

$$Cc_i = \begin{cases} \dfrac{R_S}{\omega_o \cdot \left[K_i \cdot \sqrt{R_S^2 + K_i^2}\right]} & i = 1 \\ \dfrac{1}{\omega_o \cdot K_i} & 2 \le i \le n \\ \dfrac{R_L}{\omega_o \cdot \left[K_i \cdot \sqrt{R_L^2 + K_i^2}\right]} & i = n+1 \end{cases} \quad \text{, for a shunt coupled filter.} \qquad (5\text{-}3)$$

$$Cc_i = \begin{cases} \dfrac{J_i}{\omega_o \sqrt{1 - \left(\dfrac{J_i}{G_s}\right)^2}} & i = 1 \\ \dfrac{1}{\omega_o \cdot J_i} & 2 \le i \le n \\ \dfrac{J_i}{\omega_o \sqrt{1 - \left(\dfrac{J_i}{G_L}\right)^2}} & i = n+1 \end{cases} \quad \text{, for a series coupled filter.} \qquad (5\text{-}4)$$

- Step 3: Derive the tuning capacitors in the resonators

First calculate the proposed capacitance tuning to the centre frequency with the inductor value of L:

$$Ct = 1/(\omega_0^2 \cdot L)\ . \qquad (5\text{-}5)$$

Then, the actual tuning capacitance can be adjusted by removing the effect of a pair of coupling capacitances at each node for $i=1...n$,

$$Ct_i = \begin{cases} \dfrac{C_t \cdot Ce_i \cdot Cc_{i+1}}{Cce_i \cdot Cc_{i+1} - C_t \cdot Cce_i - C_t \cdot Cc_{i+1}} & i = 1 \\ \dfrac{-C_t \cdot Cc_i \cdot Cc_{i+1}}{Ct \cdot (Cc_i + Cc_{i+1}) - Cc_i \cdot Cc_{i+1}} & 2 \le i \le n-1 \\ \dfrac{C_t \cdot Ce_{i+1} \cdot Cc_i}{Cce_{n+1} \cdot Cc_i - C_t \cdot Cce_{i+1} - C_t \cdot Cc_i} & i = n \end{cases}, \text{ for a shunt coupled filter.} \quad (5\text{-}6)$$

$$Ct_i = \begin{cases} Ct - Ce_i - Cc_{i+1} & i = 1, i = n \\ Ct - Cc_i - Cc_{i+1} & 2 \le i \le n-1 \end{cases}, \text{ for a series coupled filter.} \quad (5\text{-}7)$$

So far the shunt and series coupled resonator bandpass filter's element values have been decided. Note that this is the lossless situation considering that the Q factor of the inductor is infinite. In fact, the finite inductor Q factor will cause significant insertion loss as explained later.

5.2.3 Comparison of the Two Topologies

The frequency responses of the conventional bandpass filter, series and shunt coupled filters are illustrated in Figure 5-3. All the filters are centred at 10GHz with 1GHz bandwidth, and lossless components. By comparing the frequency response, it can be observed that the conventional filter has a moderate attenuation rate on both edges: a slightly sharper edge at the lower frequency and a slightly gentle edge at the high frequency. This variation is insignificant, and could be seen as symmetric, particularly near the centre frequency. The coupled filter has a much faster attenuation rate on one edge, and has a rather slower rate on the other edge. The dual characteristic of the coupled filters can also be examined. They can be seen as approximately symmetric to each other with respect to the conventional filter response.

The element values of the filters are compared in Table 5-2. The comparison reveals the big advantage of coupled filters with respect to the conventional filter because of the smaller range of element values. The table summarizes the ranges of element values of the three topologies. It is observed that the ratio of values of the coupled resonator filter (~10) is much smaller than that of a conventional filter (~1000); hence it is expected to be easier to implement.

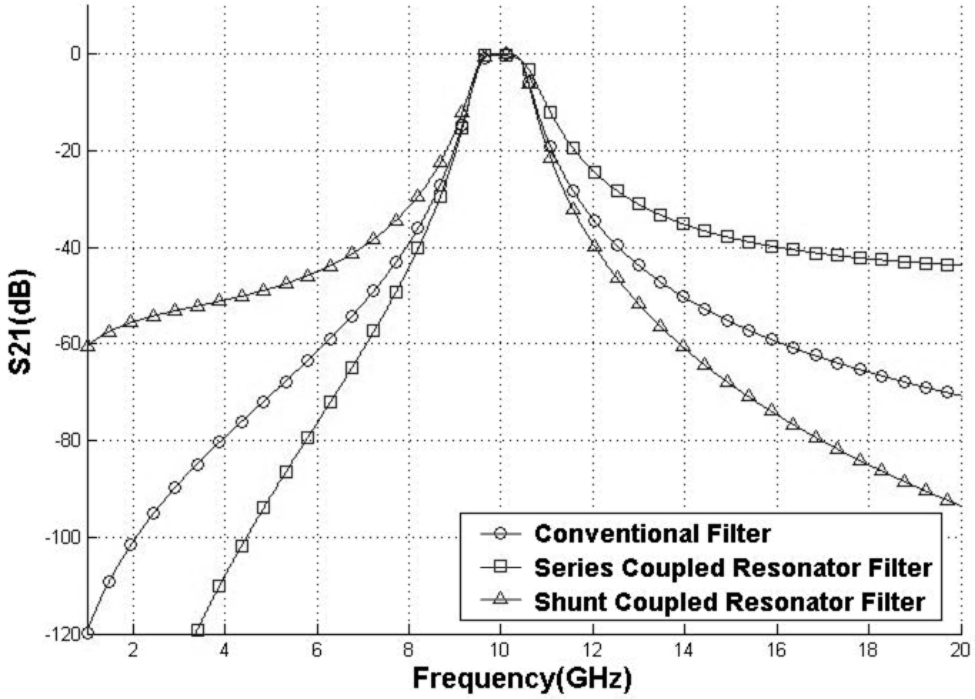


Figure 5-3 Frequency response of conventional and coupled bandpass filter

	Element Value Range	
	L	C
Conventional	pH~nH	fF~pF
Series Coupled	Same (nH)	10s fF~100s fF
Shunt Coupled	Same (nH)	100s fF~ pF

Table 5-2 Element value range of conventional and coupled bandpass filter

Additionally, practical design issues are to be noticed. The conventional and shunt coupled filters both have inductors with floating terminals, while inductors in the series coupled filter all have one terminal grounded for single-ended configuration, and terminals that could be connected to plus/minus signals in a differential topology. This reduces the parasitics and resistive losses, increases the Q factor, and makes the filter much easier to design. Also, the sharper lower edge of the series coupled filter gives more image attenuation when the real down-mixer operates at lower edge frequencies, which is often desirable since this reduces the power consumption. The drawback of the series coupled filter is that the coupling capacitors have very small values when the centre frequency is very high and the bandwidth is very low, increasing the difficulty of accurate implementation. This will be discussed later in the context of a practical design. Last, but not least, one of the obvious advantages of a series-coupled-resonator

filter is that the inductors' values are all the same, which reduces the design difficulty significantly, since only one inductor needs to be designed.

5.3 Inductor Design

The inductor design is the key issue for an on-chip bandpass filter because the insertion loss of the filter is dominated mainly by inductor losses. In this section, the on-chip inductor design is to be introduced in detail. Firstly, the layout parameters and integrated inductor modelling are demonstrated and explained, followed by a simplified frequency dependent model, which is used for the initial fast inductor dimension estimation. To increase the Q factor of an inductor on a silicon substrate, a patterned ground shield (PGS) is then introduced. Finally, the two inductor structures (at 10GHz and 1.75GHz) are designed in the ST 130nm standard CMOS substrate and metal layers. These two inductor structures are modelled in the 2.5D electromagnetic field simulation software ADS Momentum to obtain the full S-parameter data, and the inductor parameters are then extracted accordingly.

Modelling

For on-chip bandpass filter design, integrated inductors are needed, and the planar inductor construction is really the only choice of implementation. Figure 5-4 shows a classical square spiral inductor layout. The layout parameters are tabulated in Table 5-3 and explained. Note that by implementing the inductor tracks at the top metal layer (M6), the parasitic capacitance to ground is minimized because they are far away from the substrate.

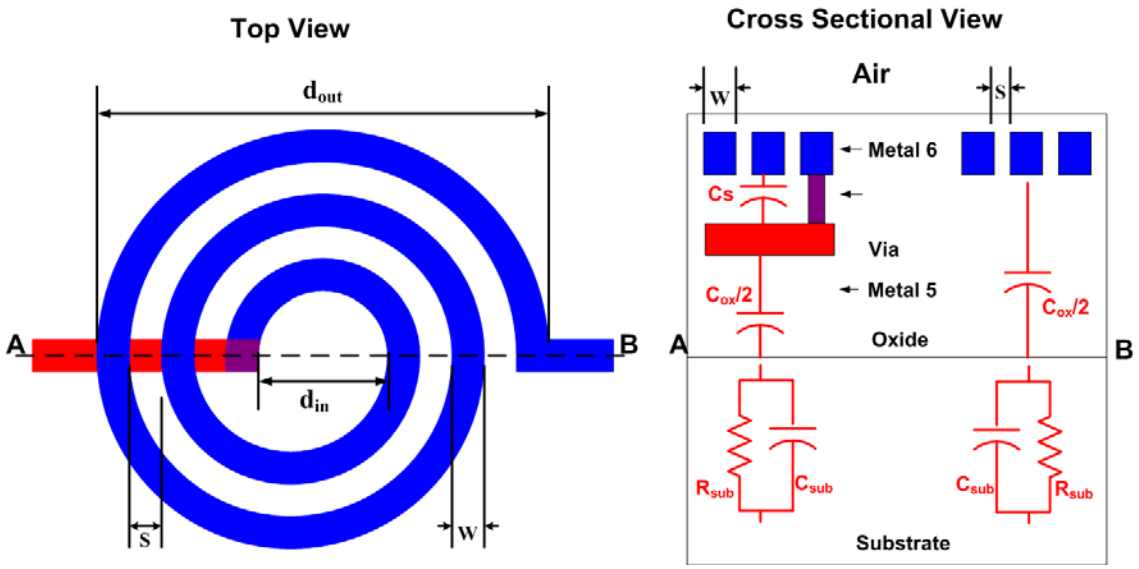


Figure 5-4 Top view and cross sectional view of square spiral inductor

Parameter	Explanation
d_{out}	Outer diameter
d_{in}	Inner diameter
W	Metal trace width
S	Spacing between metal edge to metal edge
N	Number of turns

Table 5-3 Layout parameters of a circle spiral inductor

An equivalent circuit corresponding to the inductor layout is shown in Figure 5-5, including losses and parasitic elements due to the tracks and substrate. In this model, the inductor is seen as a passive symmetric two-port Pi-network. The physical meaning and estimated values of the elements in this network are tabulated in Table 5-4. All these elements are explained as below. The final complex conductance of the series and shunt branches can be obtained by measuring the y-parameters, as shown in Figure 5-5.

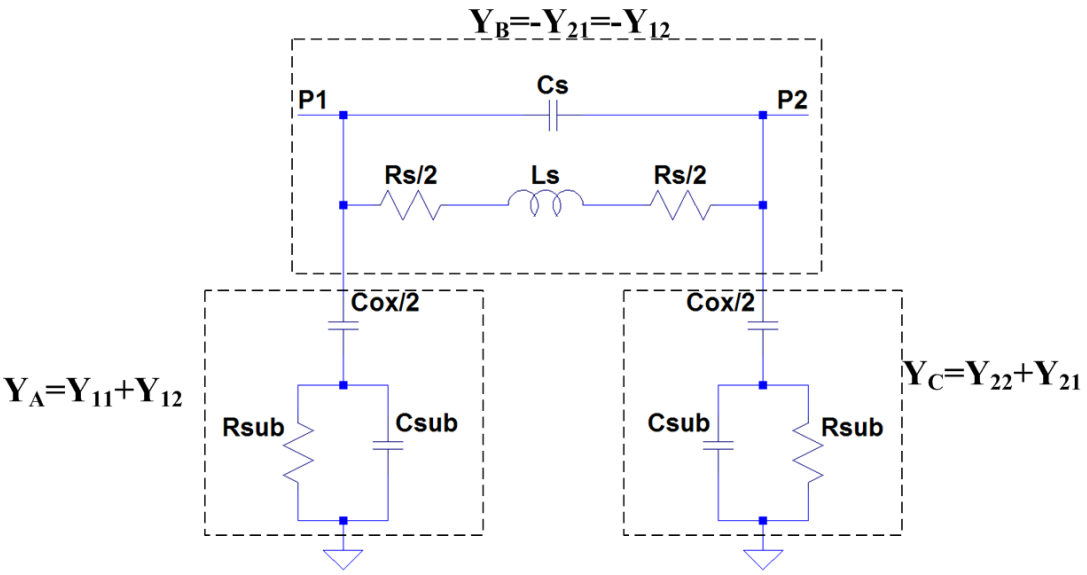


Figure 5-5 Planar inductor model: physical equivalent model [81]

Element	Physical Meaning	Equation	Note
Ls	Inductance	$Ls = \mu_0 \frac{N^2 d_{avg}}{1+\rho}$	$\mu_0 = 1.257 \times 10^{-6}$ $d_{avg} = (d_{out} + d_{in})/2$ $\rho = (d_{out} + d_{in})/(d_{out} - d_{in})$
R_S	Series Resistance	$Rs \approx \frac{l \cdot \rho}{w \cdot \delta \cdot \left(1 - e^{-t/\delta}\right)}$	Skin effect is included Metal thickness: *t (μm)* Total length: *l (μm)* Metal resistivity: $\rho = R_{sheet} \times t(\Omega - cm)$ Skin depth: $\delta = \sqrt{\frac{2\rho}{2\pi f \cdot \mu_0}}(\mu m)$
C_{ox}	Parallel plate capacitance	$C_{ox} \approx wl \cdot \frac{\varepsilon_{ox}}{t_{ox}}$	Oxide permittivity: $\varepsilon_{ox} = \varepsilon_r \cdot \varepsilon_0$ $\varepsilon_0 = 8.854 \times 10^{-12}$ $t_{ox}(\mu m)$: Distance between spiral inductor and substrate
C_S	Shunt capacitance	$C_p \approx nw^2 \cdot \frac{\varepsilon_{ox}}{t'_{ox}}$	$t'_{ox}(\mu m)$: Distance between main spiral layer and the cross under layer
R_{sub}	Substrate resistance	$R_{sub} \approx \frac{2}{wlG_0}$	$G_0 \propto 10^{-7}\left(S/\mu m^2\right)$: Substrate conductance per unit area
C_{sub}	Substrate capacitance	$Csub \approx \frac{wlC_0}{2}$	$C_0 \propto 10^{-3} \sim 10^{-2}\left(fF/\mu m^2\right)$: Substrate capacitance per unit area

Table 5-4 Planar inductor physical model elements [82]

▪ Ls: The given estimation formula is usually accurate within 5%. In practice, the inductance of the inner turns is actually degraded due to cancellation of the magnetic field for very close edges. Hence the inner windings will decrease the Q factor because of the normal contribution to resistance and weakening contribution to inductance. The inductance can be seen as a constant value over a large frequency range because it is determined mainly by the external magnetic flux of the conductor, which doesn't change significantly with varying frequency.

▪ Rs: For high frequency applications using modern technology, skin depth must be taken into account. The skin depth is due to eddy current induced by the alternating current. This factor effectively reduces the thickness of the actual cross-sectional area, and hence the sheet resistance, which is inversely proportional to the series resistance. The skin depth finally reduces to less than the metal thickness in CMOS technology operating in the GHz range. In order to obtain a high quality factor, Rs should be as small as possible. The reasonable value range of Rs is roughly 1Ω~5Ω for inductors less than 10nH with Q of 10~20, for GHz frequency applications.

▪ Cs: The shunt capacitance is a small constant value representing the capacitance of the cross-over area between the main metal and the under-cross metal (overlap area between blue and red metal tracks in Figure 5-4). Because of such a small area, this capacitance is expected to be very small. In our application, this small amount of parallel capacitance reduces the inductance Ls and leads to a slightly decreased Q factor of the unloaded inductor.

▪ C_{ox}: This capacitance is also a constant value representing the capacitance of the area between the inductor and substrate. The area is relatively large compared with Cs. Therefore, the windings should have as large a distance to substrate as possible to reduce this parasitic element.

▪ C_{sub} and R_{sub}: The substrate capacitance and resistance introduce losses and degrade the inductor's Q.

Although almost all of the element values are independent of frequency, except Rs which depends on frequency because of the skin effect, their actual effects on the whole inductor network vary with frequency. There are basically two situations to be considered. The first one is the substrate capacitance and loss effects. Since the electric field terminates at the oxide-Si interface, at low frequencies, C_{ox} dominates the shunt capacitance to ground. At high frequencies, the penetration of electric fields into the substrate connects C_{sub} and C_S in series, which reduces the total shunt capacitance to ground, while the current flow in the substrate makes the effect of R_{sub} more important. Finally, C_{ox} will effectively be shorted by C_{sub} and R_{sub}. The second factor is the self-resonant effect due to the parasitic capacitances, including C_{ox}, Cs and C_{sub}.

For this reason, it is not a good choice to extract and calculate the exact element values in the physical model at the initial design stage. Alternatively, a frequency dependent model, which is specifically focused at the design frequency, is usually extracted to have a rough estimate of the inductor's value and layout parameters, as

shown in Figure 5-6 (Assume one terminal is grounded). Note that this inductor model is single-ended, with one terminal grounded, as discussed in the filter topology section.

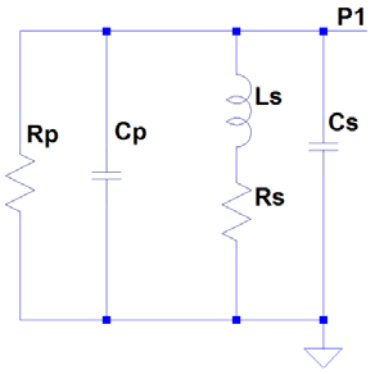


Figure 5-6 Frequency dependent inductor model

The series branches, including the series inductance and resistance, together with the overlap capacitance, are kept in this model, while the substrate parasitics and silicon dioxide capacitances are combined into a parallel connection of a resistor and capacitor, the values of which are given by [83]

$$R_P = \frac{1}{\omega^2 C_{ox}^2 R_{sub}} + \frac{R_{sub}(C_{ox} + C_{sub})^2}{C_{ox}^2}. \tag{5-8}$$

$$C_P = C_{ox} \cdot \frac{1 + \omega^2 (C_{ox} + C_{sub}) C_{sub} R_{sub}^2}{1 + \omega^2 (C_{ox} + C_{sub})^2 R_{sub}^2}. \tag{5-9}$$

The quality factor of the inductor can then be expressed by Equation (5-10). Note that the first term reflects the magnetic energy and resistive losses related only to the tracks. The substrate loss is described by the second term. The self-resonant frequency phenomenon also alters the quality factor, as expressed by the third term.

$$Q = \frac{\omega L_s}{R_s} \cdot \frac{R_p}{R_p + \left[1 + \left(\frac{\omega L_s}{R_s}\right)^2\right] R_s} \cdot \left[1 - \frac{R_s^2 (C_p + C_s)}{L_s} - \omega^2 L_s (C_{ox} + C_{sub})\right]. \tag{5-10}$$

The substrate resistive losses can be reduced by inserting a conductive grounded shield, usually in the silicided polysilicon layer or the lowest metal layer, between the inductor and substrate, which stops the electric field from penetrating the substrate. This still leaves the problem associated with ‘eddy currents’ [84], which can be generated in the substrate from the magnetic field of the inductor spiral. This current exists in conductors, including the resistive substrate and the grounded shield, near the structure and has the opposite direction to the current along the inductor’s metal tracks. Therefore, a negative inductance is formed and could reduce the actual inductance significantly as well as increasing the losses and degrading the Q factor. The eddy current in the grounded shield (as opposed to that in the substrate) has the potential to

be dominant, because of the high conductance of this layer compared to that of the substrate. For this reason, the grounded shield should be patterned with narrow fingers and slots to prevent an induced current flowing through it. The fingers should be narrow enough to minimize current loss and the slots should be narrow enough to stop the electric field reaching the substrate. The structure of the inductor with a patterned ground shield is shown in Figure 5-7. It can be observed that the substrate parasitic capacitance and resistive losses due to the electric field penetration are eliminated by the shield. However, the induced magnetic eddy current losses in the substrate are still present.

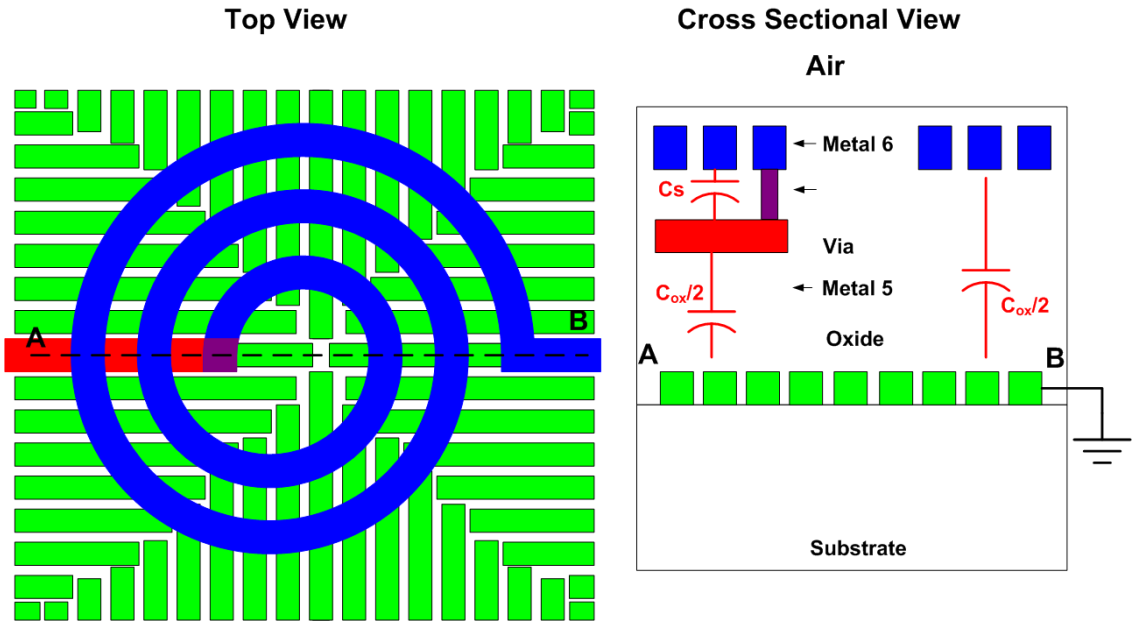


Figure 5-7 Planar inductor model: with patterned ground shield

Implementation

The inductor design procedure involved in this project starts width a coarse design using the ASITIC program [84] and then a full EM simulation design using ADS Momentum.

The ASITIC is adopted for the fast initial inductor dimension estimation. Generally speaking, the inductor design is essentially the best combination of length, width, space and number of turns to achieve the best compromise between Q factor and layout area.

For narrowband modelling, there are some basic guidelines for designing an inductor:

1. For a certain diameter, *2r*, and number of turns *N*, there is a peak Q factor as *W* is varied. If *W* is made very narrow, the increased resistance dominates and degrades

the Q while if *W* is very wide, the reduced average radius dominates and degrades the Q. So the peak Q appears with a moderate *W*, which has achieved a good balance between reducing the resistance and increasing the average radius.

2. For a given *W* and *N*, there is a peak Q factor for different values of *2r*. If *2r* is very small, the reduced average radius dominates and degrades Q, while if *2r* is large, the increased resistance dominates and degrades the Q. So the peak Q appears with a moderate value of *2r*, where there is a good balance between increasing the average radius and reducing the resistance.

3. For a given *N*, the optimum choices of 'moderate' *2r* and *W* referred to above will vary with inductor size. So a tail of peak Q factors exists. It approximately follows a line from small *2r* and W to large *2r* and *W*.

4. When *N* is the design variable, for bigger *N*, the peak Q factor appears for small values of *2r* and *W*, and Q decreases very quickly when *2r* increases. On the other hand, for smaller *N*, the peak Q factor appears for larger *2r* and *W*, and Q decreases more slowly as *2r* increase.

5. When the inductor size is increased, the inductance and the parasitic capacitance increase as well, decreasing the self-resonant frequency.

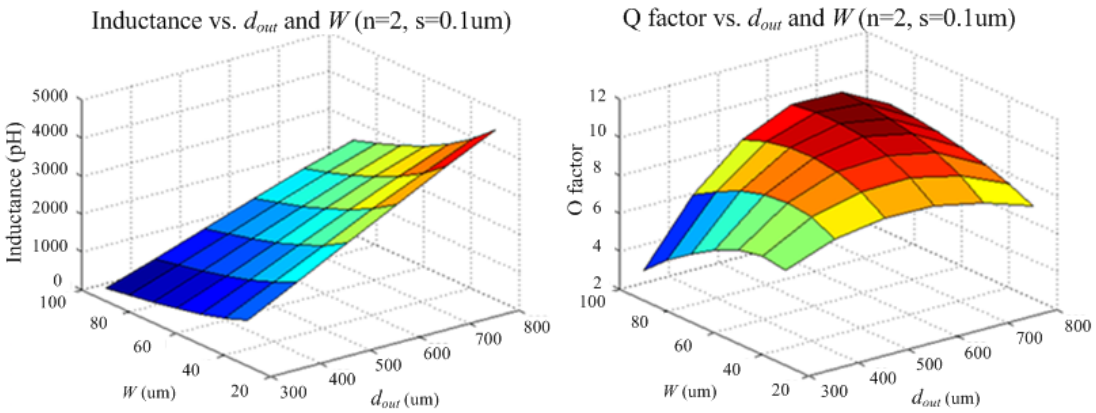


Figure 5-8 Example of Q factor and inductance versus *W* and d_{out}

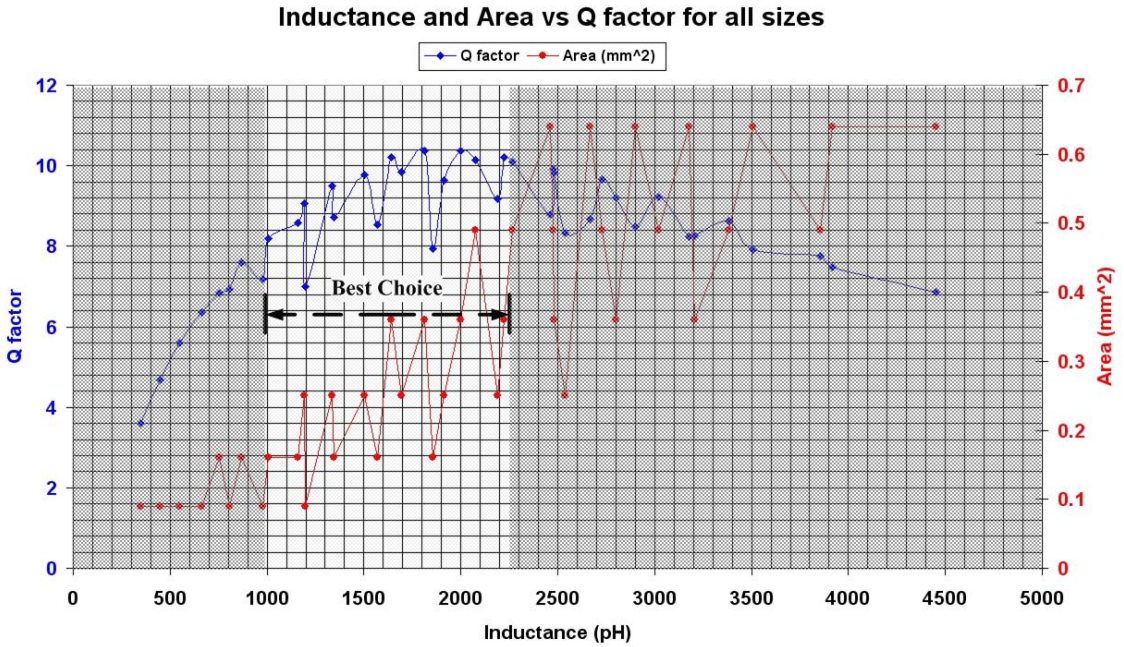


Figure 5-9 Example of Q factor and area vs. inductance

Some inductance, area, quality factor trade-off examples are shown in Figure 5-8 and Figure 5-9. By doing the fast simulation using ASITIC, a best combination of the outer diameter (d_{out}), number of turns (N), track width (W) and track spacing (S) can be found, to achieve the best combination of quality factor and area combinations, within the acceptable inductance range.

After performing an approximate design using ASITIC, ADS Momentum is used to simulate the S-parameters of the inductor structures. For the practical design, the inductor layout is selected as a symmetrical octagonal spiral structure. This structure generally leads to minimized common-mode current in the shield and the substrate and hence a higher Q factor [85], and better area efficiency, especially in differential circuits. The 45 degree angles in the structure are compatible with most CMOS technology layout design rules. The 3D-view layouts of 10GHz and 1.75GHz inductors are shown in Figure 5-10 and Figure 5-11, respectively.

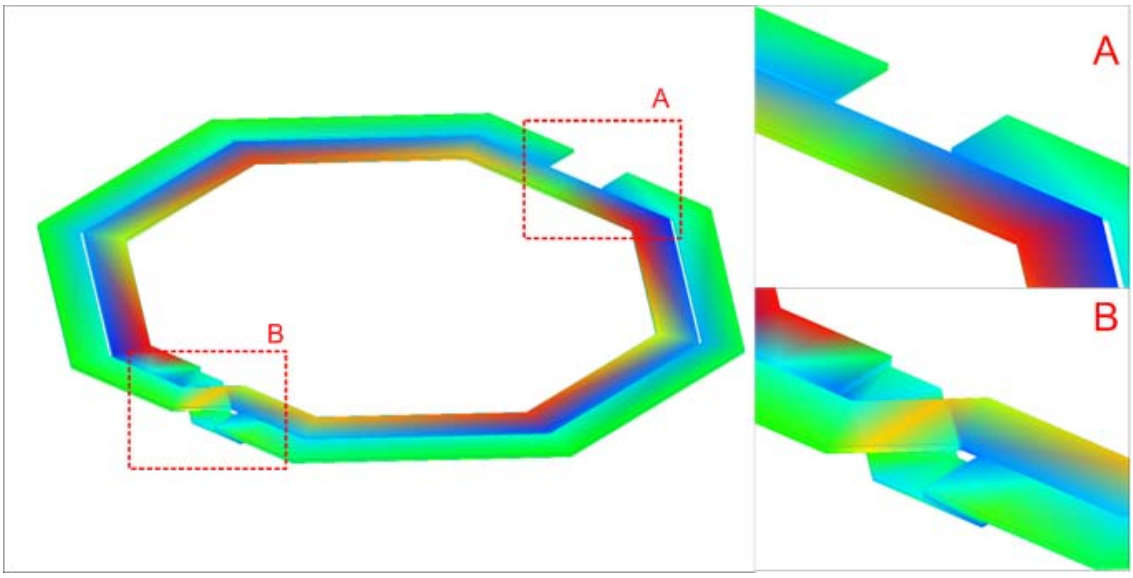


Figure 5-10 10GHz inductor 3D view (Octagonal in practical design)

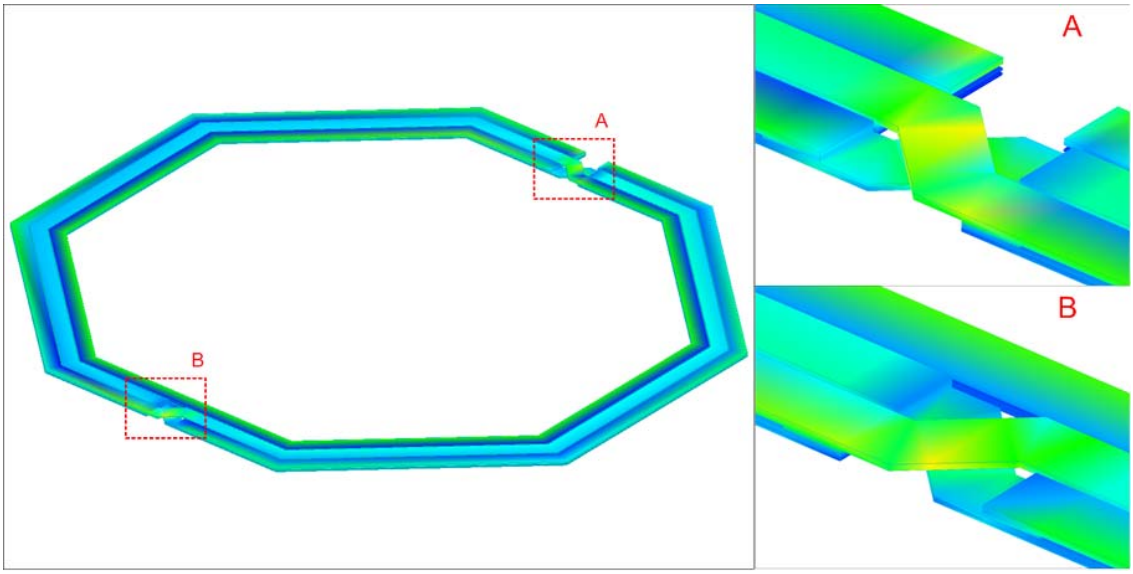


Figure 5-11 1.75GHz inductor 3D view (Octagonal in practical design)

While a relatively higher Q factor can be achieved for the 10GHz inductor because of the higher frequency, it is difficult for the 1.75GHz inductor to achieve a higher Q factor within a relatively small area. Therefore, the stack layout structure [86] is adopted. In this design, identical turns are built in the top three metal layers respectively, with parallel connection to each other, as illustrated in Figure 5-11. This structure increases the inductance due to the stacked turns, while reducing the series resistance because of the parallel connections between stacked turns. The drawback of this structure is the increase in the parasitic capacitance to ground, although, as will be shown later in this chapter, this can be absorbed by the tuning capacitance of the filter to some extent. The higher parasitic capacitance also results in a lower self-resonating

frequency, which is taken into account during the design to make sure that it is well above the inductor operating frequency of 1.75GHz.

The differential and single-ended inductor parameters can be extracted from the Y-parameters and Z-parameters, which can be obtained directly from the simulated S-parameters. The expressions are given below.

$$L_{DD} = \frac{Imag(Z_{11} - Z_{12} - Z_{21} + Z_{22})}{2\pi f}. \tag{5-11}$$

$$R_{DD} = Real(Z_{11} - Z_{12} - Z_{21} + Z_{22}). \tag{5-12}$$

$$Q_{DD} = \frac{Imag(Z_{11} - Z_{12} - Z_{21} + Z_{22})}{Real(Z_{11} - Z_{12} - Z_{21} + Z_{22})}. \tag{5-13}$$

$$L_{SE} = \frac{Imag(1/Y_{11})}{2\pi f}. \tag{5-14}$$

$$R_{SE} = Real(1/Y_{11}). \tag{5-15}$$

$$Q_{SE} = -\frac{Imag(Y_{11})}{Real(Y_{11})}. \tag{5-16}$$

$$C_{CM} = \frac{Imag(Y_{11} - Y_{12} - Y_{21} + Y_{22})}{2\pi f}. \tag{5-17}$$

The terms L_{DD}, R_{DD}, Q_{DD} are the inductance, series resistance and Q factor of the differential mode impedance, while the terms L_{SE}, R_{SE}, Q_{SE} are for the single-ended mode. The common mode impedance reflects the parasitic coupling to the shield or substrate in a differential situation, and is usually in the form of a parasitic capacitance.

Simulation Results

The layout-parameters, simulation results and extracted values are listed in Table 5-5 and Table 5-6, for the 10GHz and 1.75GHz inductors, respectively.

The extracted results of the differentially driven 10GHz inductor are shown in Figure 5-12. The original S-parameter simulation results can be found in Appendix B-1. The two-turn inductor occupies an area of 200μm×200μm implemented in the top metal (M6), with a track width of 11.9μm and a track spacing of 1μm. The differential inductance is about 1.34nH, with the peak Q factor of 14.22, almost at 10GHz. The common-mode capacitance is 57fF. The self-resonant frequency is located above 30GHz, which is well beyond the operating frequency.

The extracted results of the differentially driven 1.75GHz filter inductor are shown in Figure 5-13. The 1.75GHz filter inductor has three turns, with an area of 500μm×500μm, and implemented in the top three layers (M6, M5 and M4). The peak

Q factor achieves 11.84 at 1.737GHz frequency, with an inductance of 10.89nH, and a self-resonating frequency of 4GHz.

10 GHz Inductor			
Inductor Layout Parameters			
Diameter	200 μm	Track Width	11.9 μm
Track Space	1 μm	No. Turns	2
Metal Layers	M6	No. Stacks	1
Extracted Inductor Parameters @ 10GHz			
L_{DD}	1.338 nH	R_{DD}	5.913 Ω
$F_{Qmax-DD}$	10 GHz	Q_{DD}	14.22
$Fres_{DD}$	30.66 GHz	C_{CM}	57.03 pF
L_{SE}	1.427 nH	R_{SE}	11.89 Ω
$F_{Qmax-SE}$	5.179 GHz	Q_{SE}	7.513
$Fres_{SE}$	25.52 GHz		

Table 5-5 10GHz inductor layout parameter and simulation results

1.75 GHz Inductor			
Inductor Layout Parameters			
Diameter	500 μm	Track Width	11.9 μm
Track Space	1 μm	No. Turns	3
Metal Layers	M6,M5,M4	No. Stacks	3
Extracted Inductor Parameters @ 1.75GHz			
L_{DD}	10.89 nH	R_{DD}	10.14 Ω
$F_{Qmax-DD}$	1.737 GHz	Q_{DD}	11.84
$Fres_{DD}$	4.01 GHz	C_{CM}	0.55 pF
L_{SE}	12.67 nH	R_{SE}	32.09
$F_{Qmax-SE}$	1.01 GHz	Q_{SE}	4.293
$Fres_{SE}$	3.275 GHz		

Table 5-6 1.75GHz inductor layout parameter and simulation results

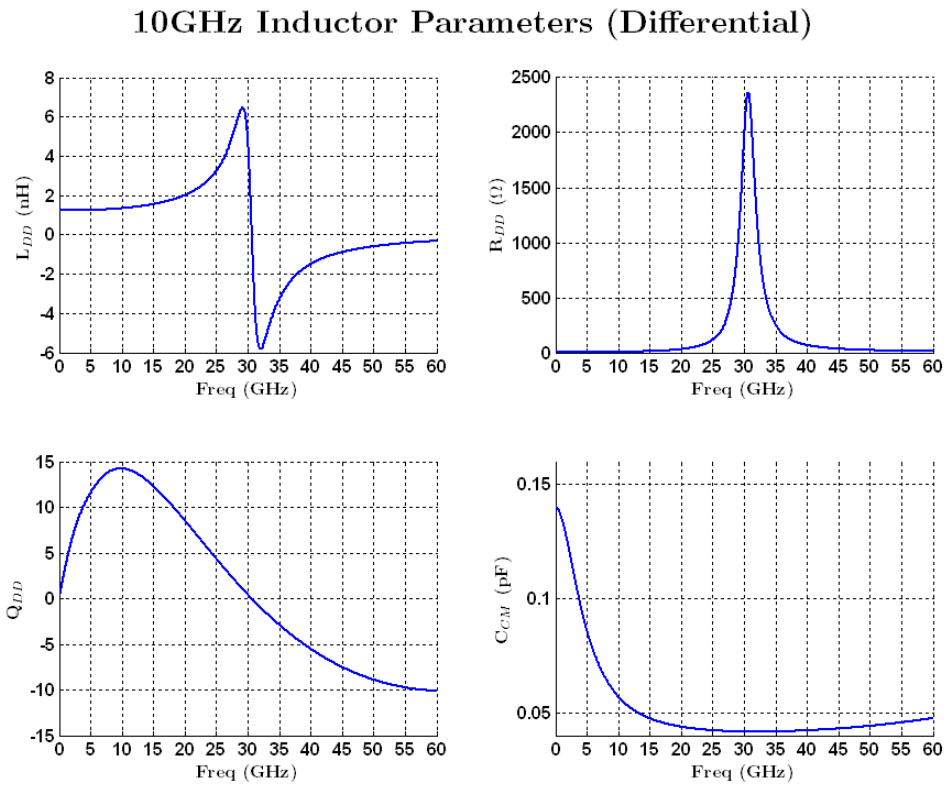


Figure 5-12 10GHz inductor extracted results (differential)

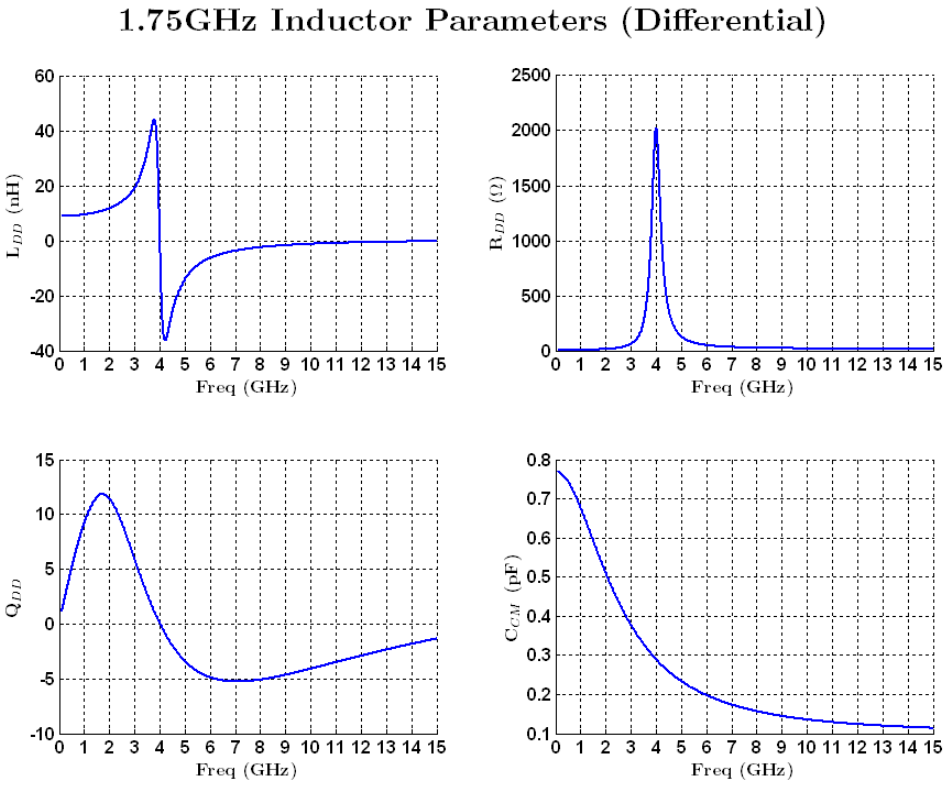


Figure 5-13 1.75GHz inductor extracted results (differential)

5.4 10GHz BPF Designs for Up-Conversion Architecture

5.4.1 Specifications

After deciding the topology and inductor design for the bandpass filter, the next design issues are the specification decisions, including the filter response type, filter order and bandwidth. Before selecting these parameters, the insertion loss of the filter should be discussed first.

The insertion loss is introduced as a result of power dissipation in the capacitors and inductors, of which the loss in the inductors is usually the dominant factor. It can be proven that [87] a reciprocal relationship exists between actual insertion loss and bandwidth, assuming that the inductor quality factors are fixed. This can be validated by simulation.

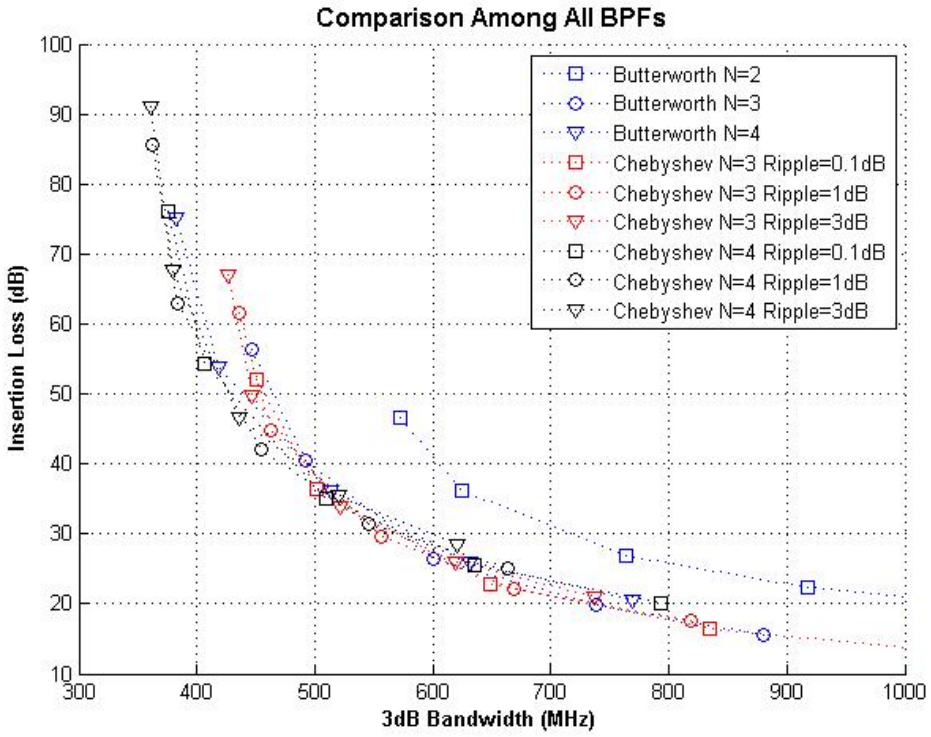


Figure 5-14 Insertion loss versus BW (using series coupled resonator topology with the same inductor Q factor and iterated design)

The simulated insertion losses versus bandwidth are shown in Figure 5-14, assuming that the inductors of the series coupled filters have 1nH inductance and a Q factor of 15. Different filter responses (Butterworth, Chebyshev 0.1dB/1dB/3dB ripple) and different orders (2^{nd}, 3^{rd} and 4^{th}) are examined. These simulation results reveal several trends to assist in filter selection. The second order filter has the worst

bandwidth vs. insertion loss relationship. The insertion losses of the 3rd order filters are higher than for the 4th order filters when the bandwidth is below about 500MHz. When the bandwidth is more than about 500MHz, the insertion losses of the 3rd order filters are reduced faster than those of the 4th order filters to below 20dB. Despite the fact that a higher noise figure can be tolerated by the spectrum monitor receiver, an insertion loss of less than 20dB is still expected to be necessary. For this reason, according to Figure 5-14, a third order filter with about 800MHz~1GHz bandwidth could be a reasonable choice to achieve 10~20dB insertion loss. Also, the 3rd order filters occupy less area than 4th order filters. The filter with a Butterworth response has a better phase response (more like a linear phase), while a filter with a Chebyshev response has greater out-of-band attenuation. The choice of response depends on the actual application. As a compromise for this application, a standard 3rd order Butterworth filter is selected. Table 5-7 summarizes the relevant specifications.

Response	Butterworth
Order	3
Actual bandwidth	800MHz~1000MHz
Insertion loss	10dB~20dB

Table 5-7 10GHz BPF design specifications

5.4.2 Delta-Star Transformation Techniques

In spite of the foregoing advantages, the series coupled resonator bandpass filter has one major problem: excessively small coupling capacitors for a high frequency bandpass filter. For the filter in this work, the two coupling capacitors' values are as small as 23fF, which makes it too hard to implement them accurately and reliably on chip. Fortunately, the delta-star transformation technique [88] can be used to solve this problem effectively.

The delta-Star transformation is used to establish the equivalence of networks with three terminals. The three elements terminate at three common nodes. This transformation guarantees that the impedance between any pair of terminals is the same for both networks. Note that none of the three terminals is defined as the source. Usually, the Delta network can be seen as a Pi network if one node is grounded, and a Star network can be seen as a T network if the same node is grounded. As illustrated in

Figure 5-15, the delta network can be transformed to a star network using equation (5-18), while maintaining the input and output impedances.

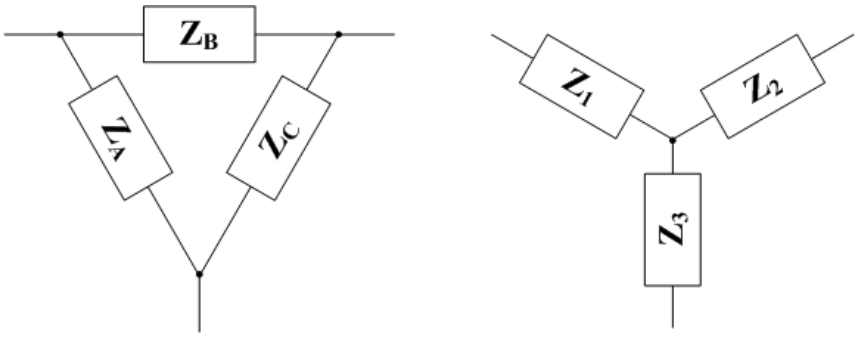


Figure 5-15 Delta-Star transformation

The general equation is to compute the impedance Z_i (i=1,2,3), in Star network with impedances corresponding Z_A, Z_B and Z_C in Delta network.

$$\begin{cases} Z_1 = \dfrac{Z_A Z_B}{Z_A + Z_B + Z_C} \\ Z_2 = \dfrac{Z_B Z_C}{Z_A + Z_B + Z_C} \\ Z_3 = \dfrac{Z_A Z_C}{Z_A + Z_B + Z_C} \end{cases} \tag{5-18}$$

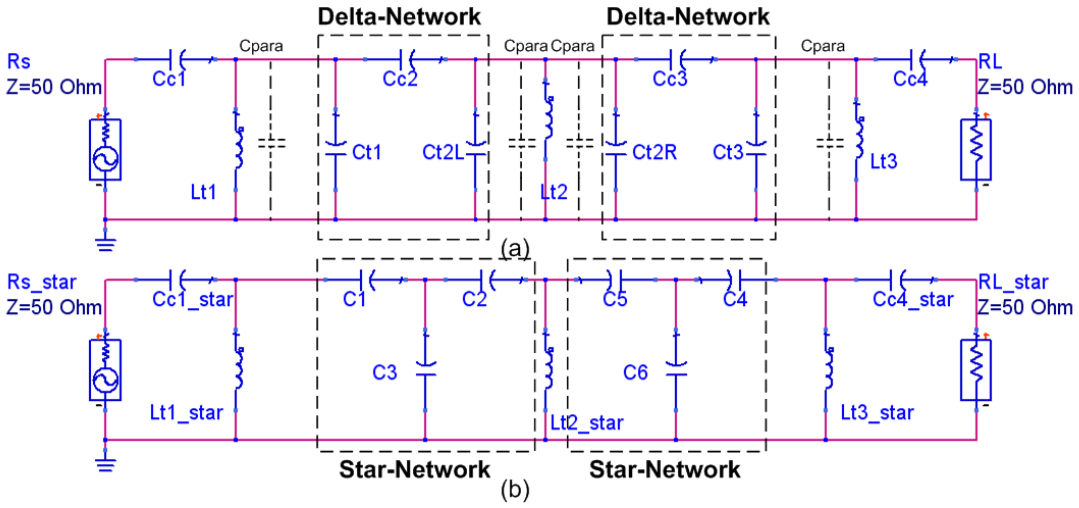


Figure 5-16 Delta-Star transformation in filter design

Because the filter in this work is third-order, the tuning capacitor in the central resonator can be split into two parallel capacitors, so that two symmetric delta-networks can be formed, as shown in Figure 5-16. Note that Cc_1=Cc_4, Cc_2=Cc_3, Ct_1=Ct_3, Ct_{2L}=Ct_{2R}. The delta-star transformation can be applied to both of these networks. The resulting transformed capacitor values are more reasonable and can be fabricated easily. Take the example of the left hand delta-network before transformation as illustrated Figure 5-16. Here Ct_1=80fF, Cc_2=14fF, Ct_{2L}=73fF, while the transformed

capacitor values are C_1=109fF, C_2=100fF and C_3=568fF. Hence the excessively small capacitor values are transformed to larger values while maintaining the same frequency response.

A similar situation can always be obtained for transforming coupled resonator bandpass filters. In general, two moderate value tuning capacitors Ct_1, Ct_{2L} are changed into C_1 and C_2 within the same order, while the smallest coupling capacitor Cc_2 will be transformed to C_3, which is at least 10 times larger, reaching the same order of magnitude as the other two capacitors. The increased area can be ignored compared with the inductor sizes. This technique can be adopted widely in high frequency coupled resonator bandpass filters.

Note that the inductors have parasitic capacitances that appear in parallel with the tuning capacitors. These parasitic capacitors can actually be 'absorbed' by the tuning capacitors by replacing the synthesized values of the tuning capacitors with values where the parasitic capacitances obtained by simulation have been subtracted.

5.4.3 Filter Implementation

As discussed above, the attenuation at the upper band edge of a series-coupled-resonator filter is not as great as in a conventional bandpass filter. To compensate for this, an additional zero is necessary at a higher frequency in shunt with the source or load resistance. There are two main benefits from this additional zero. First, for the 10GHz filter, the bandwidth achieved with acceptable insertion loss could be as wide as 1GHz, which means that a lot of the out-of-band signals will not be filtered out, and so some additional attenuation near the centre frequency will help to reduce the total intermodulation product levels in the following mixer. Secondly, as discussed in the previous chapter, the mixer after the 10GHz filter is a real down conversion mixer at about 8GHz, which means that the 3rd order harmonics at 24GHz should have sufficient attenuation. Ideally, the attenuation should be at least the same as the dynamic range of the signal, e.g. 65dB. However, at this high frequency, the internal bandwidth of the mixers at both sides of the 10GHz BPF also provides significant attenuation. Therefore, the design goal is to add moderate attenuation to the filter above the upper band edge.

The zero-branch is a series connection of an inductor and a capacitor. To minimize the influence of the zero-branch on the input impedance, the reactance at the 10GHz centre frequency needs to be much higher than the source or load resistance. This requires large inductor values and a higher frequency for the zero. However, too

large an inductor value results in too small a capacitor (and hence tare that is hard to fabricate) for a certain zero frequency, as well as occupying too much area. Furthermore, the zero frequency shouldn't be too far way from the centre frequency because the proposed extra attenuation is expected to be several GHz away from the centre frequency. Hence some trade-off must be made. The small area inductor can have many turns with a lower quality factor.

The final filter has a differential structure, which can be transformed straightforwardly from the single-ended version. The shunt admittances remain the same while the series impedances should be half of those in the single-ended filter, which means that the values of the series connected capacitors' should be doubled. This transformation is demonstrated in Figure 5-17.

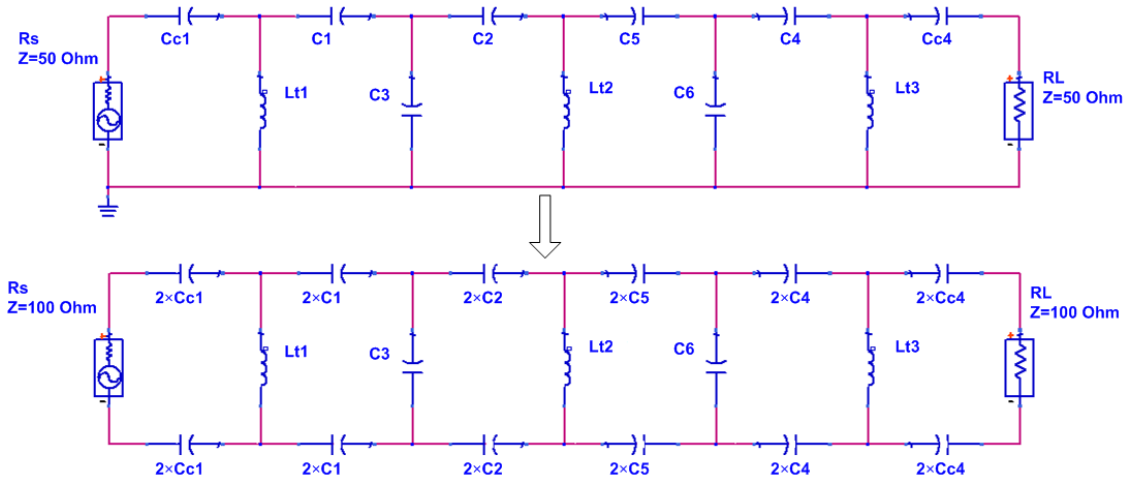


Figure 5-17 Differential to single-ended transformation

Since the filter operates at a very high frequency, the input and output terminals should be placed as far as possible from each other to minimize coupling effects due to parasitic capacitance. To avoid negative mutual inductance between adjacent inductors, the signal paths are twisted and the signal flows are in the form of a figure-of-eight. The capacitors are implemented by fringe capacitors from metal one to metal five.

Furthermore, patterned ground shields are placed in the metal one layer under the inductors to reduce the substrate losses. The shields are implemented as rectangular structures, with the edge length the same as the inductor diameter (200µm). The finger width and pitch are 0.13um and 0.18um respectively, which are the minimum allowed values in this ST 130nm technology.

The 10GHz filter schematic is shown in Figure 5-18 and the layout in Figure 5-19. The three 200µm×200µm tuning inductors dominate the chip area. The fringe capacitors are placed within the 50µm gaps between inductors. The zero-branch

inductor is only 40μm×40μm. The total size of the filter is 780μm×200μm. Figure 5-20 shows the filter die photo.

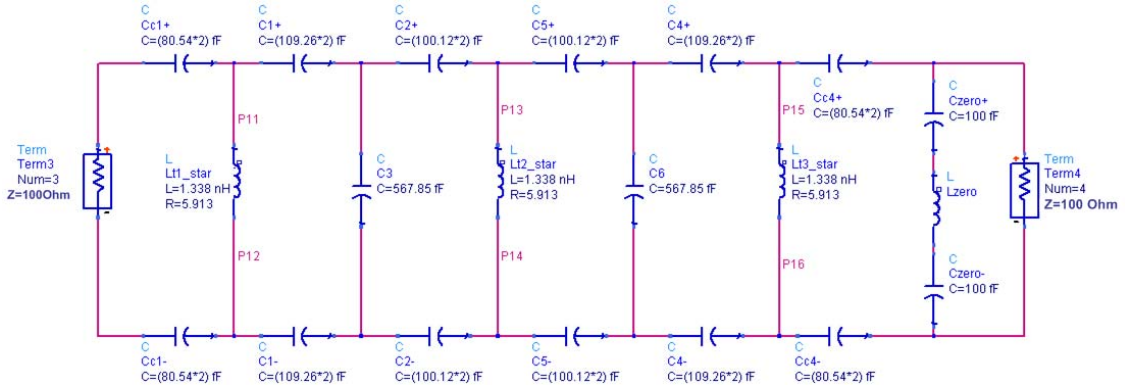


Figure 5-18 10GHz BPF schematic

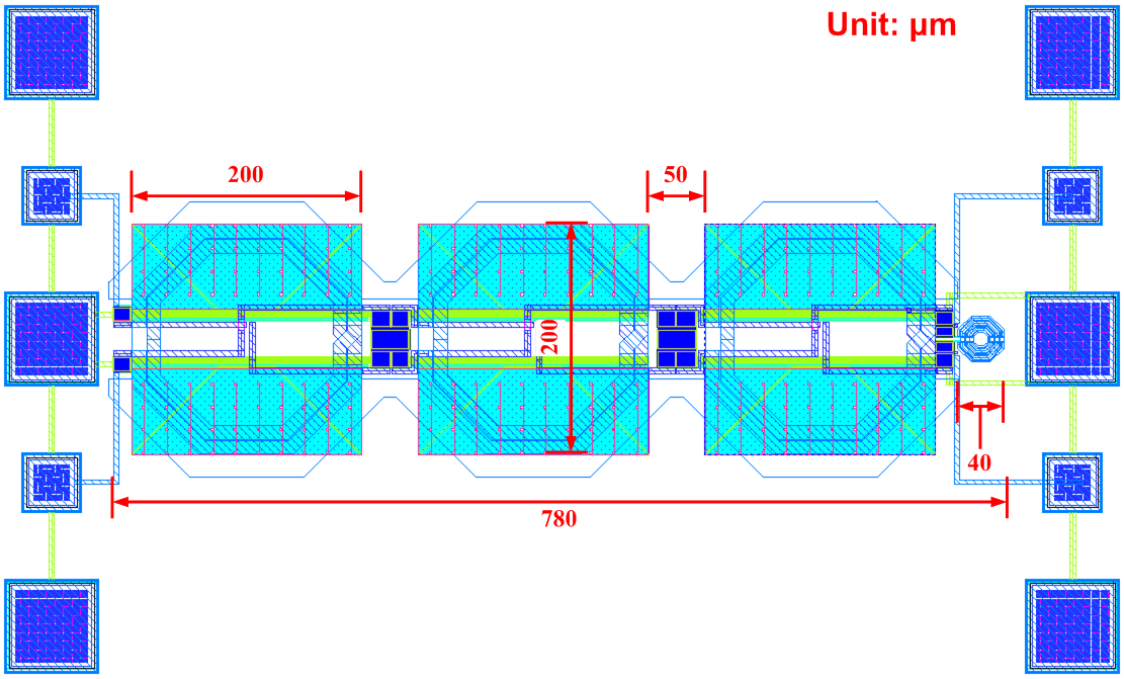


Figure 5-19 10GHz BPF layout (Cadence) and dimensions

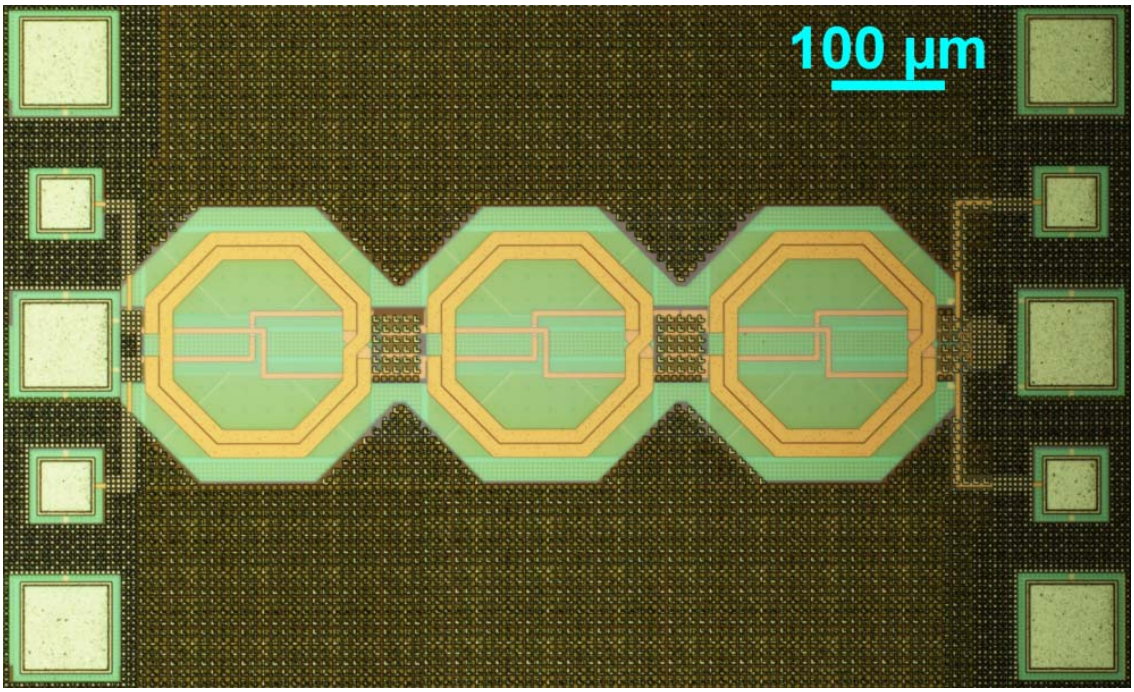


Figure 5-20 10GHz BPF photo with GSGSG Pads

5.4.4 Simulation and Measurement Results

Figure 5-21 to Figure 5-23 show the electromagnetic simulation of the 10GHz filter's frequency response, and input and output matching, compared with a conventional filter with the same centre frequency and bandwidth. The conventional 3rd order filter's topology is the same as that shown in Figure 5-1. The simulated insertion loss at 10GHz is 10.92dB and is less than 11dB within the monitored band from 9.9GHz to 10GHz.The actual 3dB bandwidth is 1148MHz from 9.43GHz to 10.58GHz. The lower frequency stop band attenuates by 73dBc at 6GHz, compared with 46dBc for conventional filter. Note the unit here is dBc, instead of a conventional expression in dB. This expression gives the attenuation with respect to the signal level in the passband, and so removes the effect of the passband attenuation. Because of the additional zero, the higher frequency stop band response maintains the same attenuation as a conventional filter up to 16GHz, where 42dBc attenuation is observed. At frequencies higher than this, the effects of the zero are diminished and the original series-coupled-resonator filter frequency response dominates, while the attenuation flattens out and remains at about 40dBc up to 40GHz. The phase of the filter remains nearly linear around the centre frequency. The input matching return loss (S_{11}) is below -15dB, while the output matching (S_{22}) is below -10dB. Because the zero-branch is

located at the load resistance, the output matching is not as good as the input matching, but is still good enough for this application.

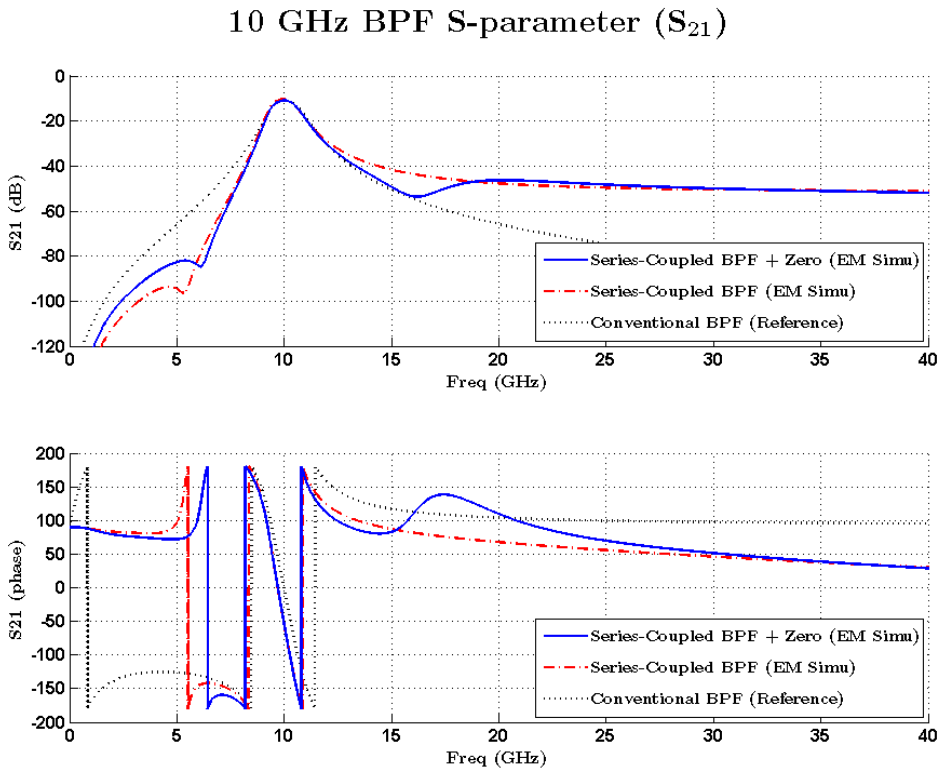


Figure 5-21 10GHz BPF simulation results: transfer function (S_{21})

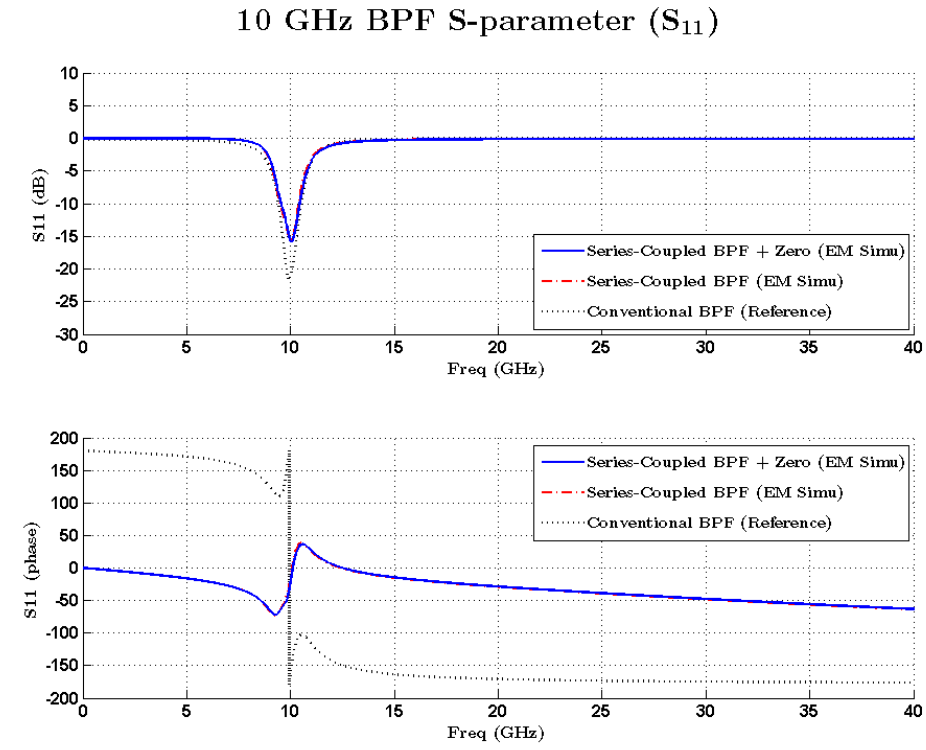


Figure 5-22 10GHz BPF simulation results: input matching return loss (S_{11})

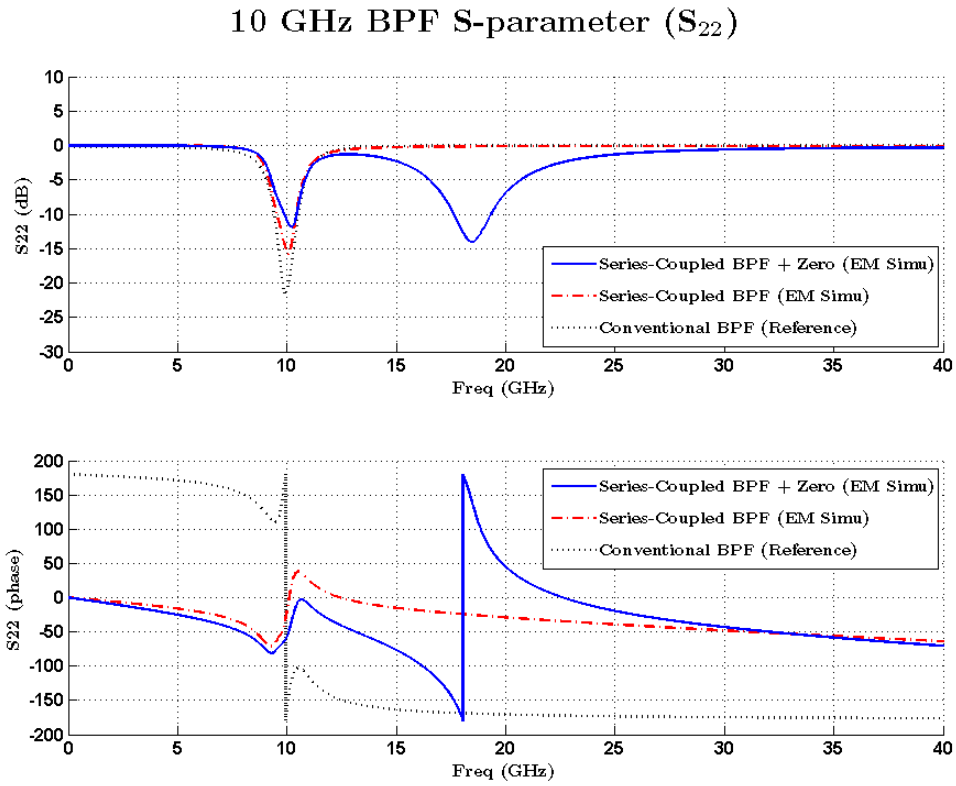


Figure 5-23 10GHz BPF simulation results: output matching return loss (S_{22})

The bandpass filter is measured using a pair of GSGSG RF probes and an Agilent E8361A PNA Network Analyzer (10MHz~67GHz). The S-parameters of the filter are measured directly with on-chip probing, as shown in Figure 5-24. The cable losses from the output connector of the network analyzer to the tips of RF probe are also measured. The actual filter S-parameters of the filter are obtained via a calibration procedure by subtracting the test cable loss (S_{21}) at the input and output ports from the measured filter S-parameters. The raw data of S_{21}, S_{11} and S_{22} data of the filter without calibration and cable loss can be found in Appendix B-2. The calibrated filter response is then smoothed using the MATLAB program to get the final data. Note that the calibration procedure doesn't remove the parasitic capacitance and resistance of the GSGSG pad pairs.

The final calibrated measurement results are illustrated in Figure 5-25 to Figure 5-27. The insertion loss is 15.5dB at the centre frequency of 9.47GHz. The 3dB bandwidth is still about 1GHz. The stopband attenuation at the lower edge matches the simulation results down to 8GHz. At the lower frequency of 6GHz, the attenuation is 45dBc, and it is more than 55dBc at the frequency of 4.7GHz. At the upper band edge,

the stop band attenuation is more like a pure series-coupled-resonator filter without an additional zero-branch. There are some disagreements between the simulation and measurement results. The measured centre frequency is lower than the simulated results. The reason is believed to be due to the pair of GSGSG pads, which can be essentially modelled as parasitic capacitance to ground, lowering the centre frequency. The effect of the extra zero is not observed in the measurement result, and this is mainly due to the lower Q factor of the capacitors at this high frequency, because the fringe capacitor model is adopted directly from the ST 130nm design kits and its high frequency model is not provided. The measured input and output matching return losses are both below 10dB.

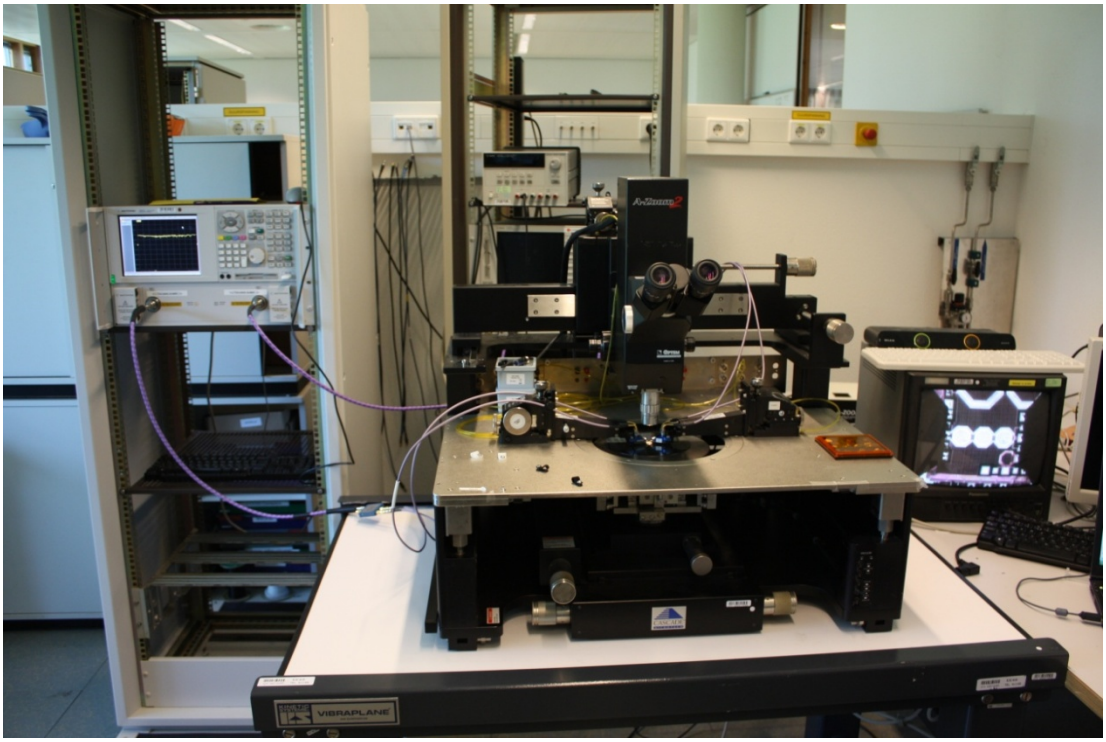


Figure 5-24 BPF measurement environment

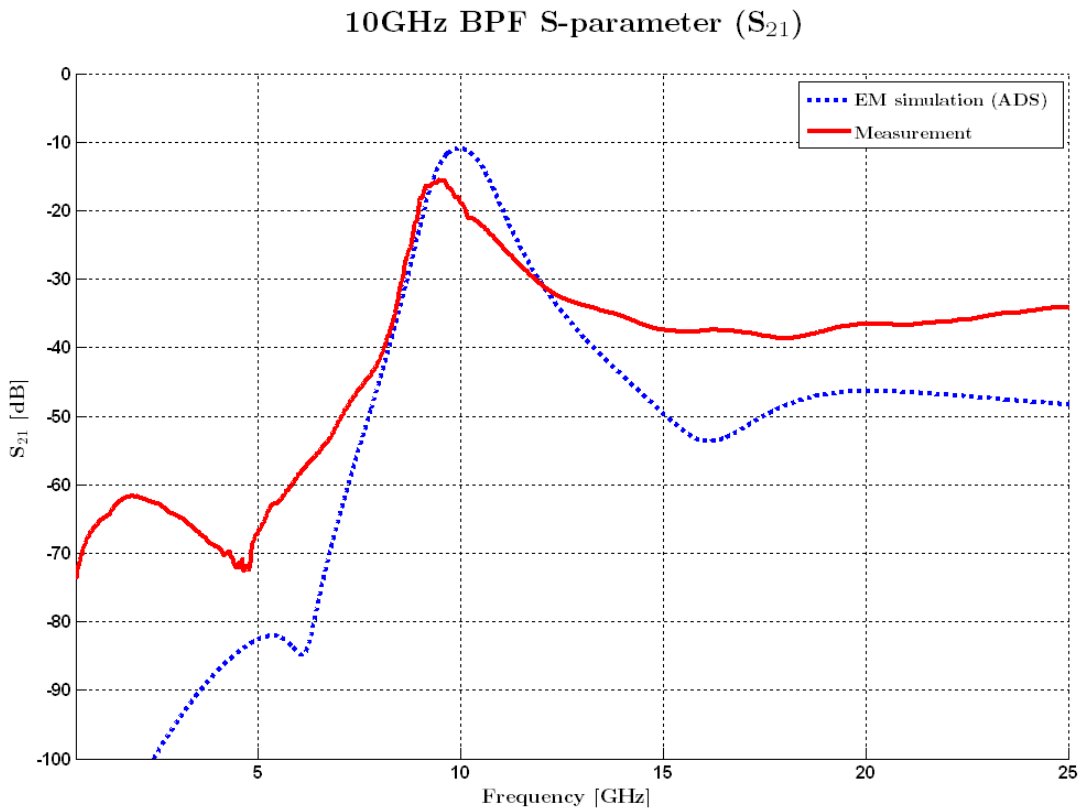


Figure 5-25 10GHz BPF calibrated measurement results (S_{21})

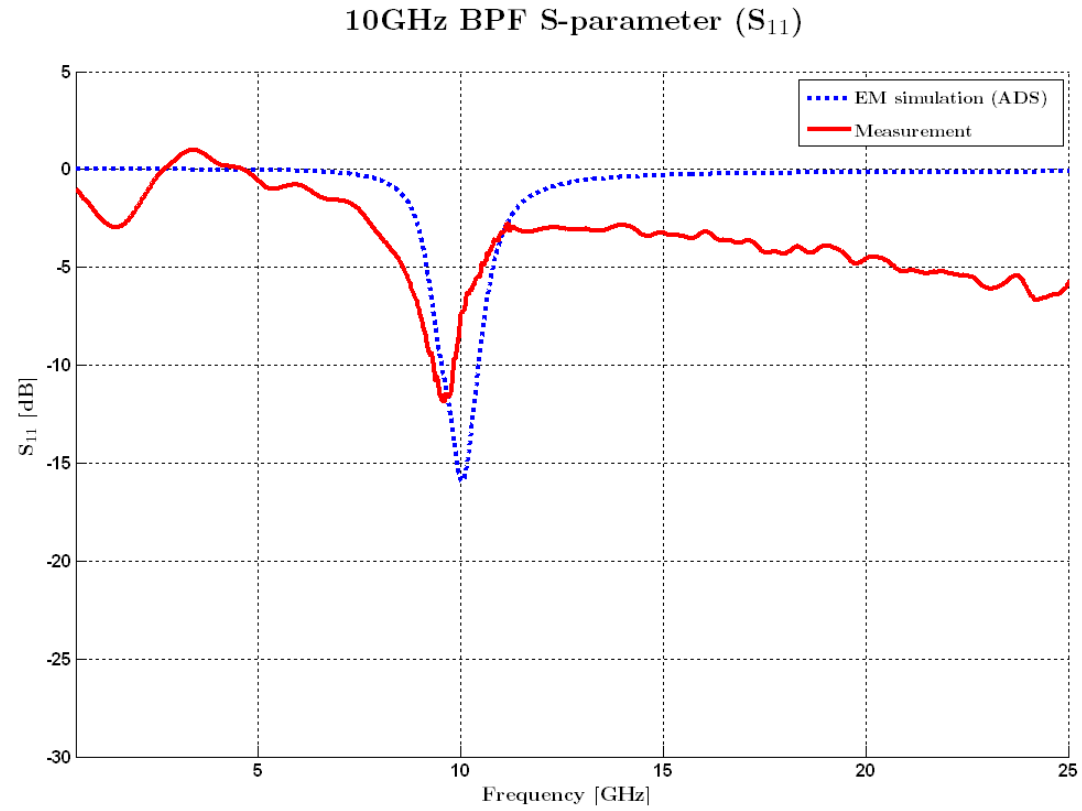


Figure 5-26 10GHz BPF calibrated measurement results (S_{11})

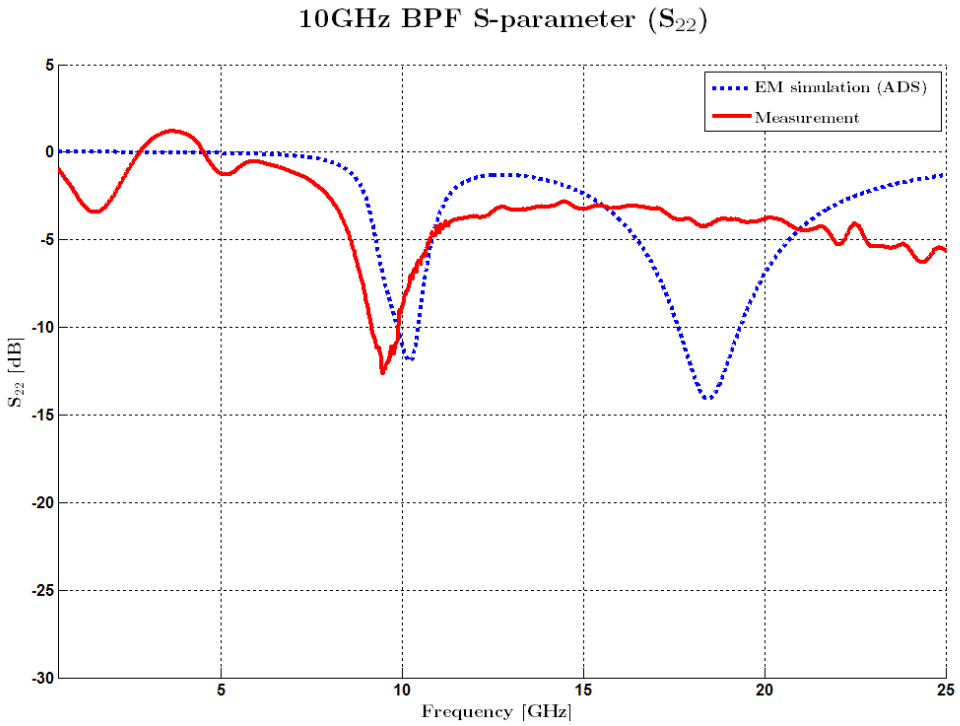


Figure 5-27 10GHz BPF calibrated measurement results (S_{22})

5.5 1.75GHz BPF Design for Down-Conversion Architecture

5.5.1 Specifications and Implementations

The ideal integrated bandpass filter in down-conversion is centred at 1.75GHz with a bandwidth of 100MHz, as discussed in chapter four. The initial design of the 1.75GHz bandpass filter design is the same as for the 10GHz bandpass filter, with a 3rd order Butterworth response. However, there are some different specifications for the 1.7GHz filter.

Compared with the 10GHz filter, the 1.75GHz filter's centre frequency is lower by about 80%, and therefore with the same Q factor of inductor, it is reasonable to estimate that the 3dB bandwidth of the filter can be also narrowed by 80%, down to about 200~300MHz. This means that no more passive filters should be needed before the base-band low pass filtering.

The 1.75GHz bandpass filter is placed after the first real down-conversion mixer and is followed by a pair of quadrature down-conversion mixers. Therefore, the image rejection ratio requirement is only related to the I/Q mismatch of the complex mixers. As explained in chapter four, the filter must be able to attenuate the signal at the

frequency of the 5th order harmonics of the complex down-conversion mixer, ideally by 56dBc to meet the dynamic range of 70dB.

In the 1.75GHz filter, the coupling capacitors are of the order of several hundred femto-farads, which can be implemented reliably in the ST 130nm CMOS technology. Therefore, the delta-star transformation used in the 10GHz filter is not necessary in this case.

Because the inductor is implemented in three layers, the lowest layer is metal four, which is much closer to the metal one and poly-silicon layers that could be used as the ground shield. This introduces more parasitic capacitance between the inductor and the ground shield. By removing the shield, the parasitic capacitance can be reduced, at the cost of greater insertion loss due to electrostatic coupling to a lossy substrate. Therefore, some trade-off must be made between the parasitic capacitance and the resistive loss. In this design, the grounded shield is removed, so that the parasitic capacitance can easily be 'absorbed' by the tuning capacitance, as discussed before. The schematic of the 1.75GHz BPF is shown in Figure 5-28.

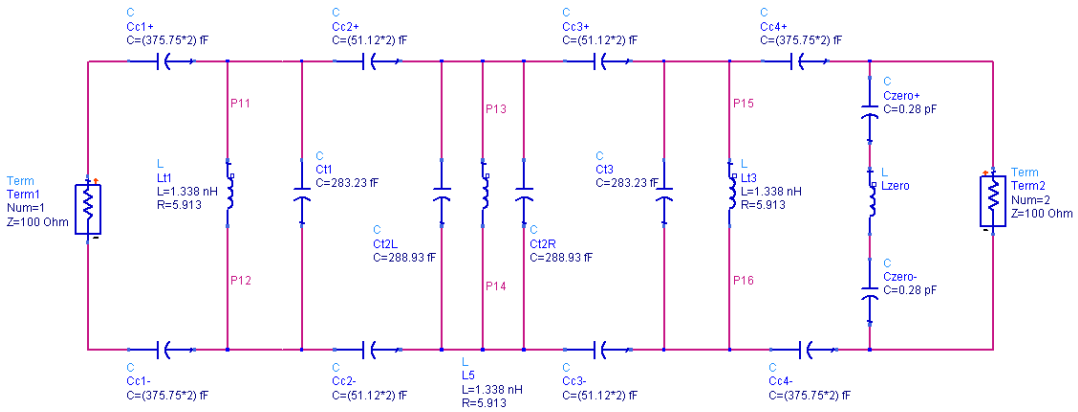


Figure 5-28 1.75GHz BPF schematic

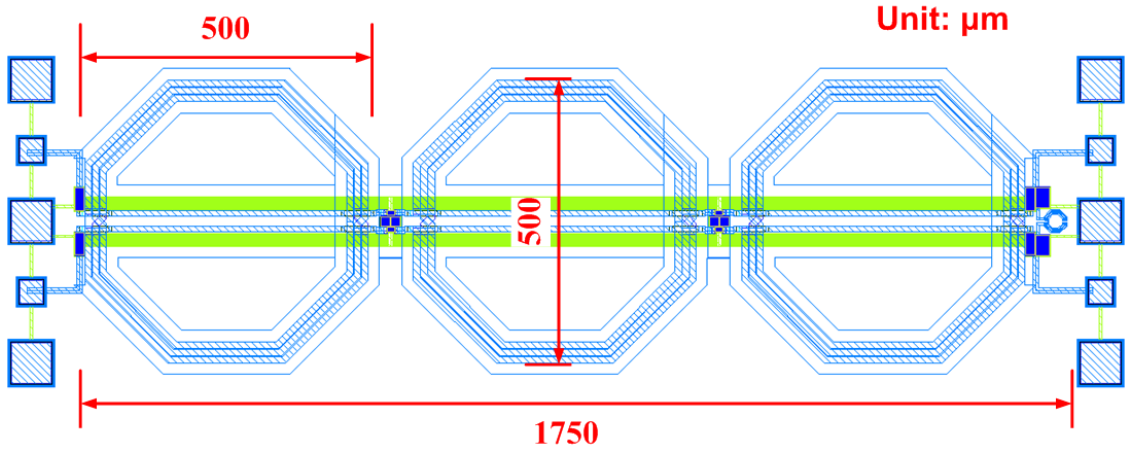


Figure 5-29 1.75 GHz BPF layout (Cadence) and dimensions

In the 10GHz bandpass filter discussed in section 5.4, the bandwidth is over 1GHz, while the targeted signal bandwidth is 100MHz, so some offset in the centre frequency is tolerable. However, in the 1.75GHz filter, the bandwidth is of the same order as the signal bandwidth, so the frequency accuracy is much more important. For this reason, the filter is designed at a little higher frequency of 1.85GHz than the specified frequency to have some margin for unpredicted parasitics.

The 1.75 GHz filter layout is shown in Figure 5-29. The three 500μm×500μm tuning inductors dominate the chip area. The fringe capacitors are placed in the 40μm gaps between the inductors. The total size of the filter is 1750μm×500μm, excluding the GSGSG RF probe pads. Figure 5-30 shows the filter die photo.

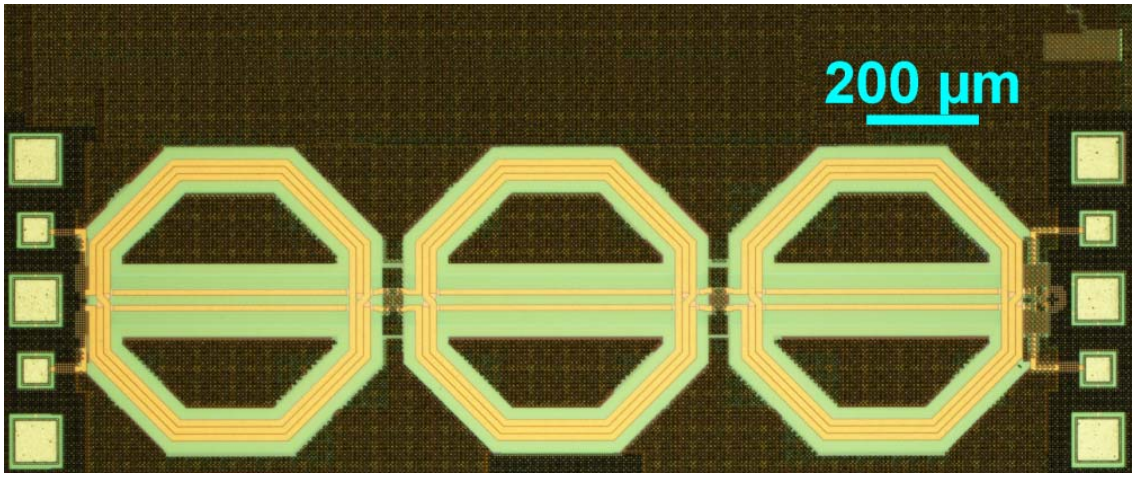


Figure 5-30 1.75 GHz BPF die photo

5.5.2 Simulation and Measurement Results

The electromagnetic simulation results are illustrated in Figure 5-31 to Figure 5-33, compared with a reference bandpass filter design implemented using a conventional topology. The simulated insertion loss is 9.9dB, and the 3dB bandwidth is about 250MHz. At 400MHz and 800MHz frequency offsets from the 1.85GHz centre frequency at the lower edge, the stop band attenuation achieves 32dBc and 75dBc respectively. At higher frequency edge, the additional zero is placed at about 8GHz, where the attenuation achieves 62dBc, and remains at more than 50dBc within the frequency range from 6GHz to 11GHz. At very high frequencies, the attenuation maintains better than 44dBc. The phase remains linear near the centre frequency. The input and output matching return loss figures, S_{11} and S_{22}, are more than 15dB and 20dB, respectively.

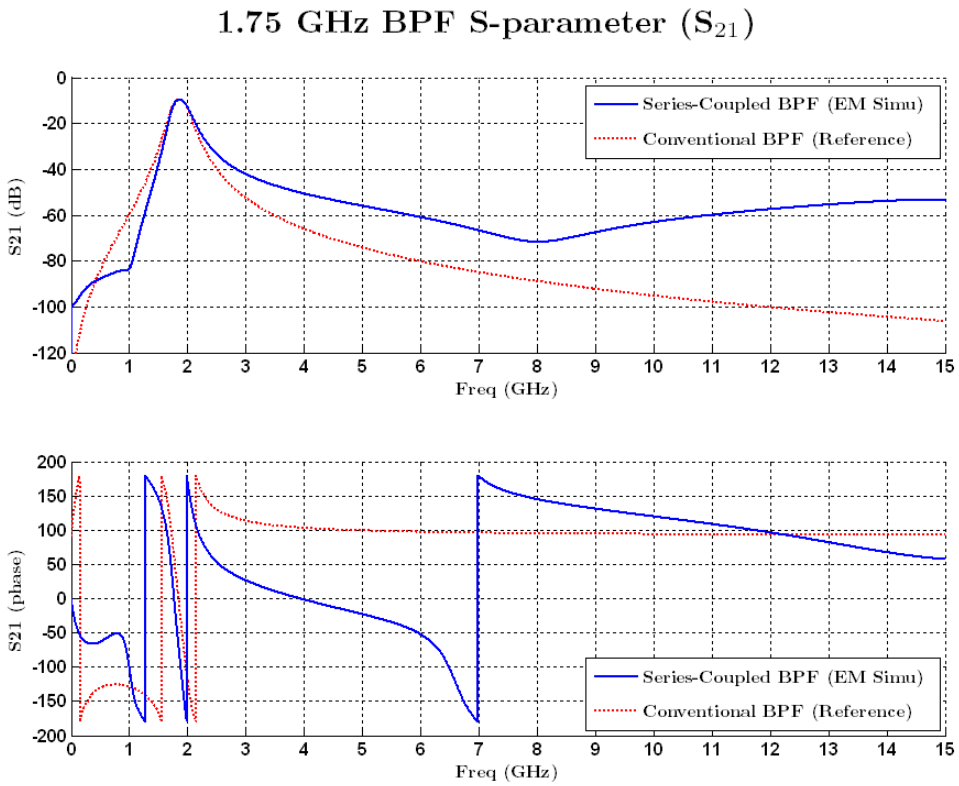


Figure 5-31 1.75 GHz BPF Simulation: transfer function (S_{21})

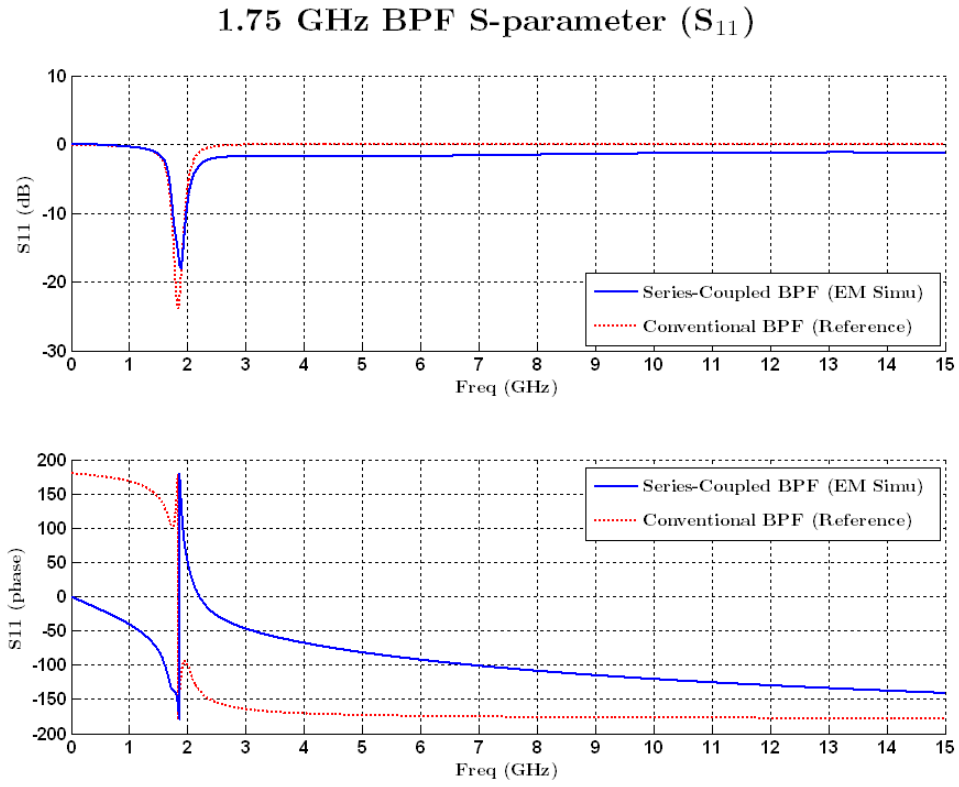


Figure 5-32 1.75 GHz BPF Simulation: input matching return loss (S_{11})

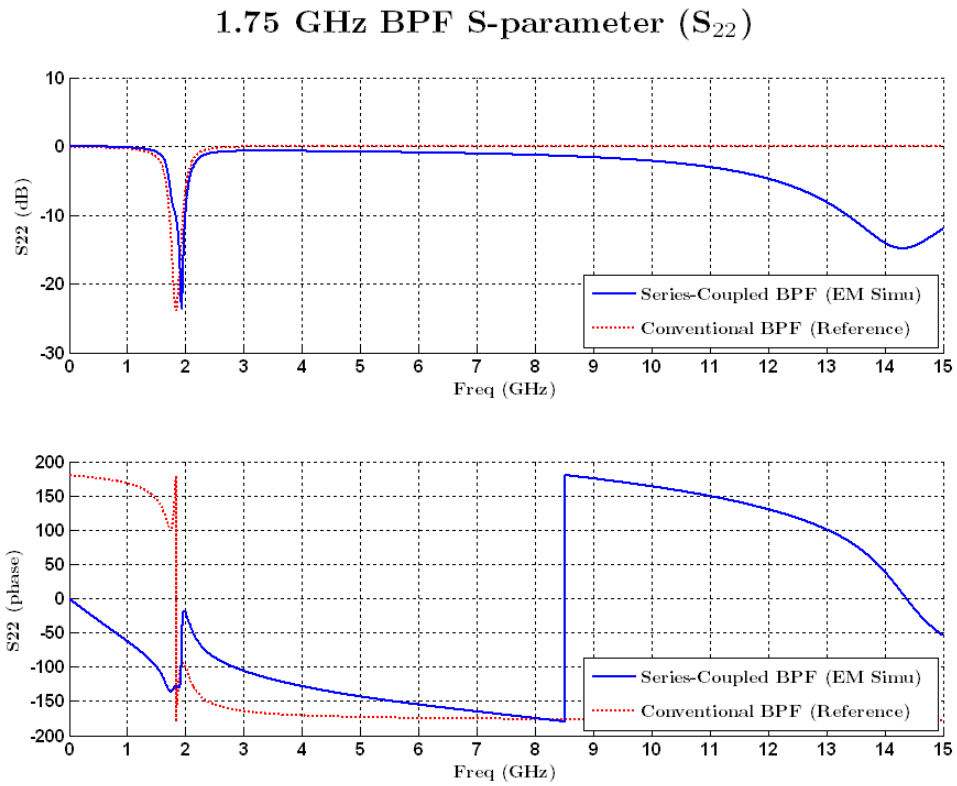


Figure 5-33 1.75 GHz BPF Simulation (S_{22}) output matching return loss

The calibration method is the same as for the 10GHz BPF. The raw measurement data of the transfer function, input and output return loss, together with cable loss can be found in Appendix B-3. The calibrated measurement results are shown in Figure 5-34 to Figure 5-36, compared with the EM simulation results. The centre frequency as measured has moved from the initial design value of 1.85GHz to exactly 1.75GHz, due to the effect of the GSGSG pads. The measured insertion loss is 8.6dB, and the 3dB bandwidth is 210MHz. The lower edge stop band frequency response matches the simulation results very well down to 1.2GHz, where 50dBc attenuation is achieved. At the higher frequency edges, the stop band attenuation is less than 3dB higher than simulation predictions up to 5GHz. The zero position is offset to 8.3GHz, where the attenuation is 53dBc, which is 9dB less than simulation predictions. At the complex mixers' 5^{th} harmonic frequency of 8.75GHz, the attenuation is 51.5dBc. Therefore, some additional internal bandwidth restriction is needed to achieve the specification of 56dBc. This pole can be either from the output buffer of the first real down-conversion mixer, or the input stage of the second complex down-conversion mixer, or both. The input and output matching return losses, S_{11} and S_{22}, are both over 15dB near the centre frequency of 1.75GHz. The main reason of the disagreement between the measured and

simulated S_{22} results at higher frequency is that the cable loss was not measured very accurately at high frequency.

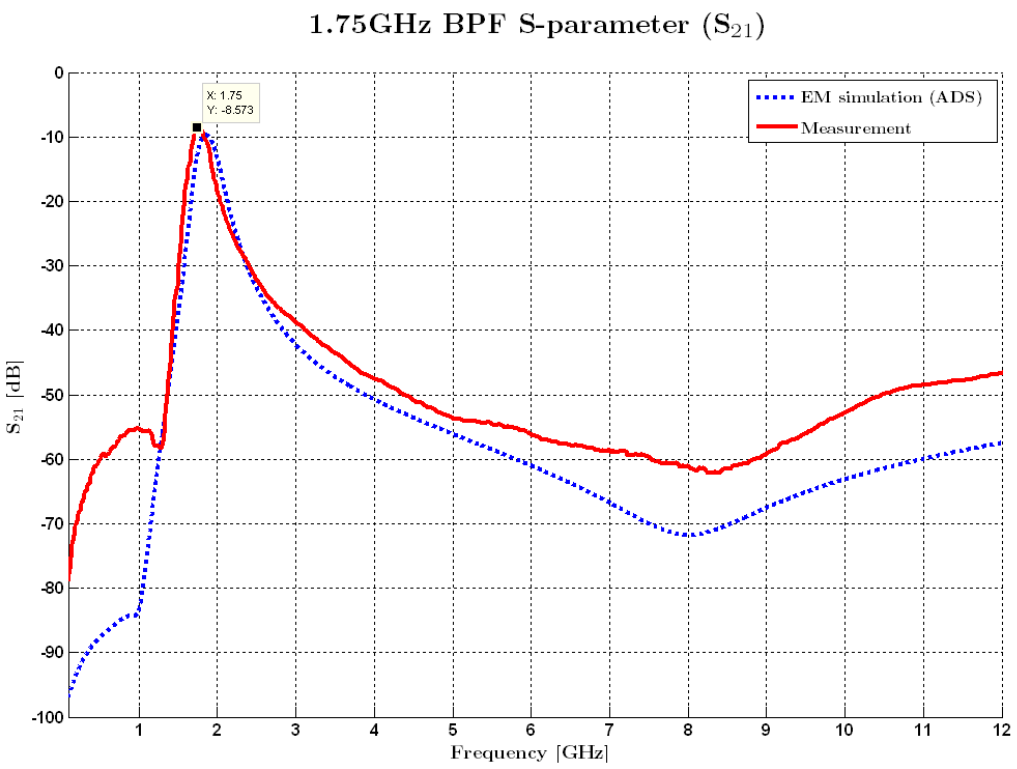


Figure 5-34 1.75GHz BPF calibrated measurement results (S_{21})

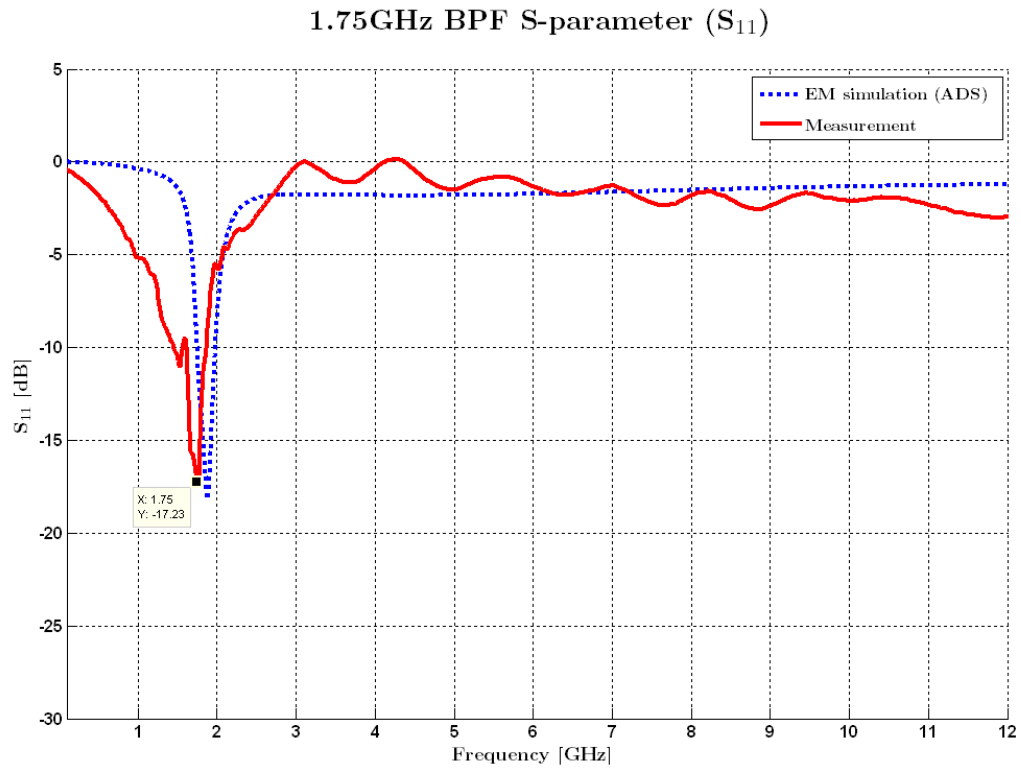


Figure 5-35 1.75GHz BPF calibrated measurement results (S_{11})

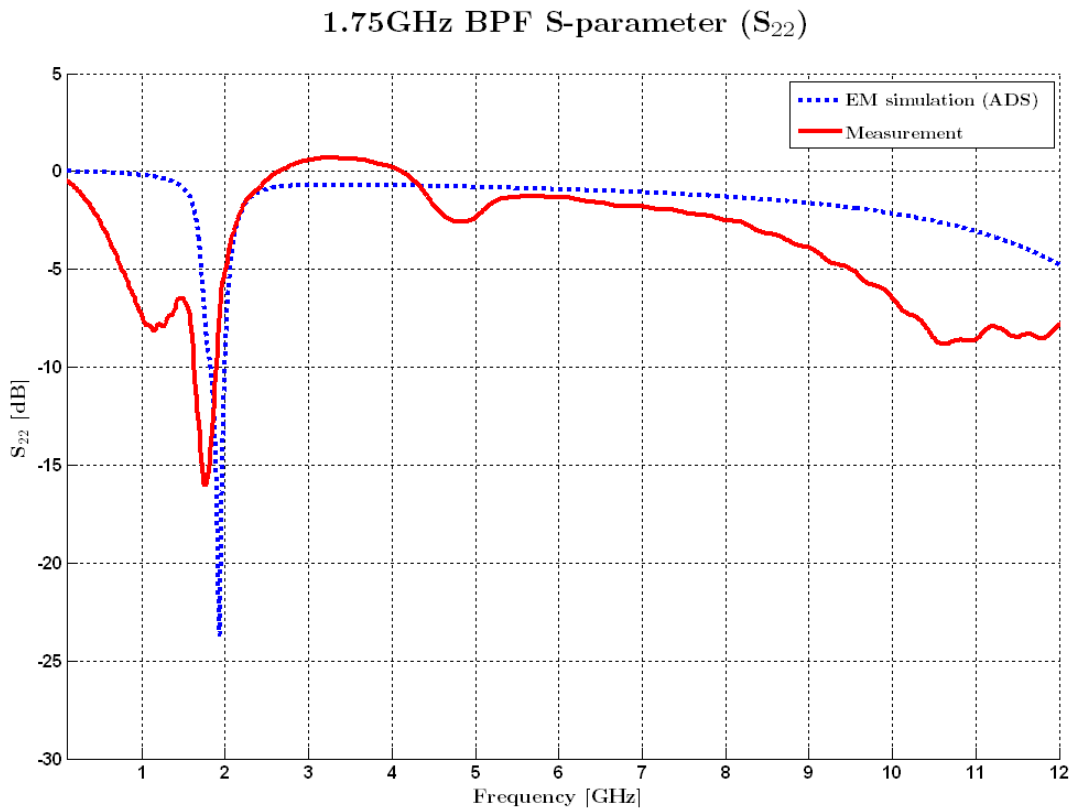


Figure 5-36 1.75GHz BPF calibrated measurement results (S_{22})

5.6 Summary

In this chapter, two integrated lumped element bandpass filters have been designed and implemented using ST 130nm standard CMOS technology, according to the requirement of the spectrum monitor architecture specifications in chapter four. Inductor modelling and designs are performed to determine the achievable Q factors when implemented in a filter. Special techniques have been used to improve the performance, such as the delta-star transformation, the addition of an out-of-band zero, and the use of figure-of-eight layout and stacked inductor structures.

For the up-conversion architecture, the 10GHz bandpass filter achieves an insertion loss of 15.5dB at the centre frequency of 9.47GHz, with a bandwidth of 1GHz. For the down-conversion architecture, the insertion loss of the 1.75GHz bandpass filter is 8.5dB, with a 3dB bandwidth of 210MHz. The two filters occupy 780μm×200μm and 1750μm×500μm die areas, respectively.

Chapter 6 Wide Tuning Range Frequency Synthesiser

6.1 Introduction

The local oscillator in a receiver chain is usually controlled by a frequency synthesis system. The key specifications of the local oscillator are the tuning range, frequency resolution and phase noise. In this project, as explained in chapter four, the tuning range is a more important issue because the monitor needs to sweep over 3GHz band, while the phase noise requirement is relaxed for this application compared with many other receiver specifications.

In this chapter, a 3rd order integer-N frequency synthesizer is designed. The theoretical system analysis is explained first. After that, the design of the key blocks is described, these being the voltage controlled oscillator (VCO), RF and digital frequency divider, phase frequency detector (PFD), charge pump (CP) and loop filter. Because of the relaxed phase noise requirement, a ring oscillator is selected as the VCO to provide a wide tuning range and low die area. Because the basic control gain could be as high as 3GHz/V, a novel tuning circuit is designed to guarantee the linear tuning and robustness to the loop filter's variation in the synthesiser. The commonly used dual modulus divider architecture arrangement has a relatively narrow range of division ratio which is only suitable for receivers with moderate tuning range. In this project, the divider is principally a counter divider with set/reset control using digital comparators. An optimized algorithm is designed to simplify the divider scheme, so that the speed can be guaranteed. With a 25MHz on-board reference, the division ratio range is set as 152-268, with increments of 4, corresponding to 3.8GHz to 6.7GHz in 100MHz steps. PFD/CP blocks are implemented with conventional architectures, followed by a 2nd

order loop filter. The phase noise and spur levels are designed to be below -80dBm/Hz at all frequency offsets to meet the spectrum monitor specification.

6.2 System Level Design

A frequency synthesiser is a feedback control system based on a phase lock loop (PLL). A PLL compares the phase of a reference oscillator and feedback signal with the phase of the feedback signal, whose frequency is a fraction of that of the VCO output. The low-pass-filtered output signal from this phase comparison controls the frequency of the VCO. The PLL can be also seen as controlling the frequency of the divided output signal so that it is the same as the reference frequency. If the division ratio following the VCO output frequency can vary, this system can generate different output frequencies, and hence is called a 'frequency synthesiser'.

The 3rd order integer-N frequency synthesiser architecture is shown in Figure 6-1. The frequency of the VCO output is divided by N times, and a feedback signal is generated. When the loop is in lock, the frequency of the feedback signal should be the same as the reference signal, which is usually generated by a crystal oscillator (either on or off the IC), while a very small constant phase different between the two should be observed. Two signals are generated from the phase/frequency detector, indicating the sign and magnitude of the comparison result between the reference and the feedback signals. The charge pump converts this result to current, which is integrated by the loop filter, and then converted to a control voltage. The PFD/CP combination has proven to be able to lock for any realistic frequency difference [89] and hence is widely adopted in PLL designs nowadays. The VCO's output frequency is then controlled by the output voltage of the loop filter.

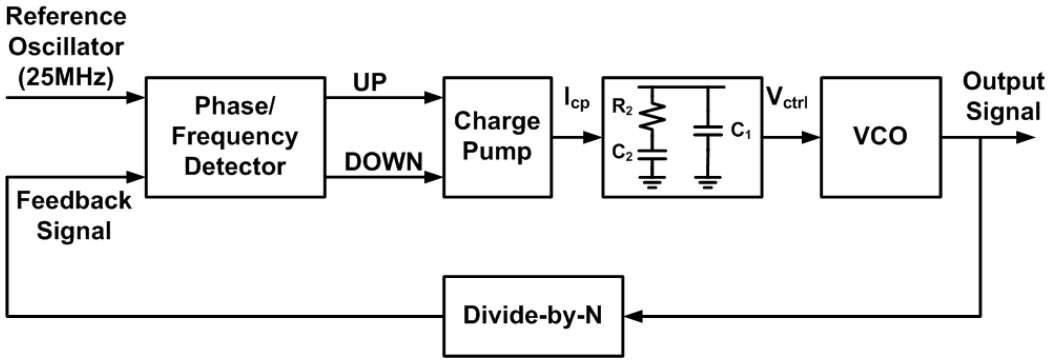


Figure 6-1 Frequency synthesiser structure (charge pump current is 30uA, ring VCO gain is 3GHz/V, feedback division ratio is from 142 to 248 with step of 4)

There are three main aspects in the design of the frequency synthesiser to meet the receiver system requirements: the loop stability, to ensure stable operation; the phase noise and spur level; and the settling time. To evaluate these performance issues in the PLL feedback system, the open loop and closed loop transfer functions should be obtained so that the gain and phase response can be analysed.

When the phases of the reference and feedback signals are the same or nearly the same, it is called the locked condition. In this situation, the PLL can be modelled as a linear time-invariant system in the phase domain. The phase of the reference oscillator is $\Theta_{REF}(t)$, the phase of the feedback signal is $\Theta_{FB}(t)$, and the difference between them is $\Theta_{ERR}(t)$, and the output signal phase is $\Theta_{OUT}(t)$. To analyse the stability, they can be transferred to the s-domain. This system architecture is shown in Figure 6-2.

By defining the forward transfer function as G(s) and the feedback transfer function as H(s), the open loop and closed loop transfer functions can be expressed as:

$$A_{openloop}(s) = G(s) \cdot H(s). \tag{6-1}$$

$$A_{closeloop}(s) = \frac{\Theta_{\mathrm{FB}}}{\Theta_{\mathrm{REF}}} = \frac{A_{openloop}(s)}{1 + A_{openloop}(s)} = \frac{G(s) \cdot H(s)}{1 + G(s) \cdot H(s)} . \tag{6-2}$$

The criterion for a stable system is that the phase margin (φ) of the open loop transfer function is no less than about 45 degrees at the frequency of the unity open loop gain. This frequency is called the loop bandwidth (f_u), and can be expressed as:

$$\left|A_{closedloop}(s)\right|_{f=f_u} = 1 . \tag{6-3}$$

The phase margin at the loop bandwidth is defined as:

$$\varphi = 180 + phase\left(A_{closedloop}(s)\right)\Big|_{f=f_u} . \tag{6-4}$$

The open loop transfer function is the product of the transfer functions of the PFD/CP, the loop filter, the VCO and the frequency divider, which are defined as shown below.

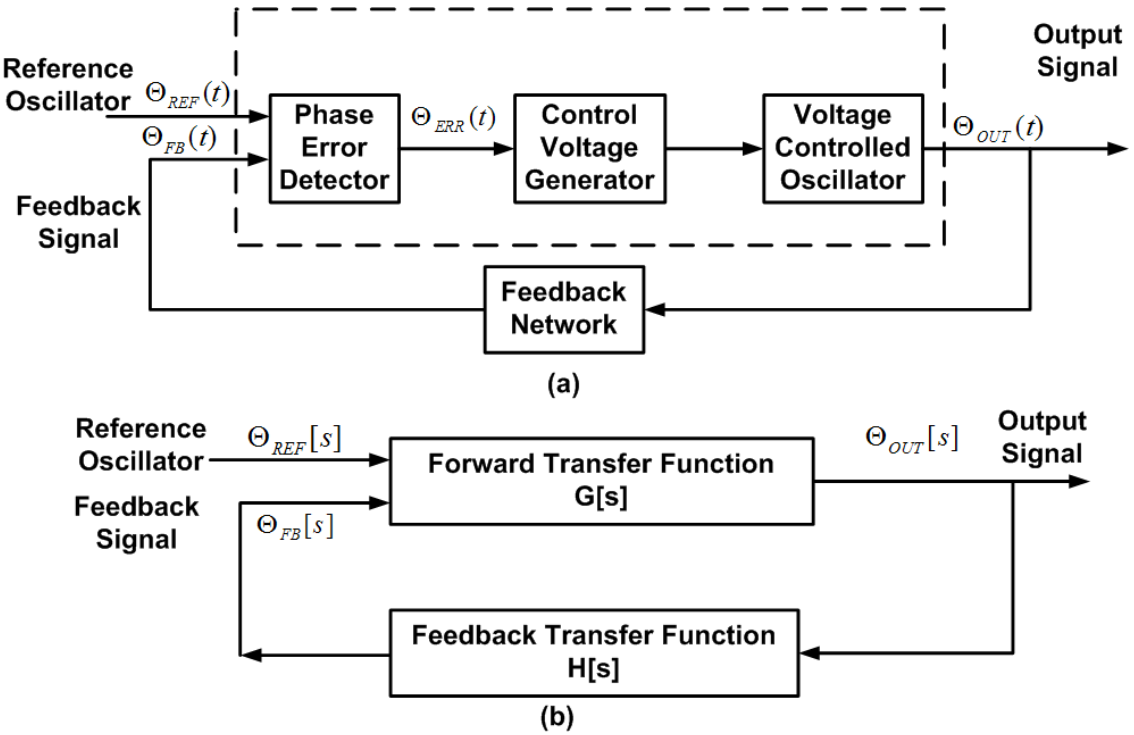


Figure 6-2 Frequency synthesiser modelling (a) time domain (b) S-domain

1. PFD-CP

The phase/frequency detector with the charge pump essentially operates in discrete time. However, when the reference frequency is much higher (8~10 times) than the loop bandwidth, the PFD-CP transfer function can be modelled as a continuous time process in the phase domain, because the sampling effects can be largely ignored if the phase variation is not very fast. The output of the PFD-CP is modelled as an instantaneous current that is proportional to the phase difference. The constant of proportionality is defined as the detector's gain K_{PFDCP} with the unit of Amps per radian. In the time domain, this relationship can be expressed as:

$$I_{PFDCP}(t) = K_{PFDCP} \times \left[\Theta_{REF}(t) - \Theta_{FB}(t)\right] = K_{PFDCP} \times \Theta_{ERR}(t) \quad (6\text{-}5)$$

The corresponding s-domain expression is

$$H_{PFDCP}[s] = \frac{I_{PFDCP}[s]}{\Theta_{\mathrm{ERR}}[s]} = K_{PFDCP}. \quad (6\text{-}6)$$

2. Loop Filter

The second order loop filter shown in Figure 6-1 is actually a trans-impedance because the input is the current from the charge pump and the output is the control voltage of the VCO. Hence, the low pass filter's transfer function is the impedance of

the passive network. This transfer function includes one zero and two poles, with one of the poles at zero frequency.

$$Z(s) = \frac{1 + sC_2R_2}{s(C_1 + C_2 + sR_2C_1C_2)} = \frac{1 + s \cdot T_2}{s \cdot (C_1 + C_2) \cdot (1 + s \cdot T_1)}. \quad (6\text{-}7)$$

In this equation, the times constants of the zero and the non-DC frequency pole are expressed as $T_2=C_2R_2$, and $T_1=C_1C_2R_2/(C_1+C_2)$, respectively.

3. VCO

The output of the VCO is a periodic angular frequency which is proportional to the control voltage from the loop filter and can be expressed as:

$$\omega_{VCO}(t) = K_{VCO} \times V_{LPF}(t). \quad (6\text{-}8)$$

Because the angular frequency is the derivative of the phase:

$$\omega_{VCO}(t) = \Delta\Theta / \Delta t. \quad (6\text{-}9)$$

The output phase of the VCO is the integral of the angular frequency over time:

$$\Theta_{VCO}(t) = \int_0^t \omega_{VCO}(t)dt = K_{VCO} \times \int_0^t V_{LPF}(t)dt. \quad (6\text{-}10)$$

And hence the s-domain transfer function expression is

$$H_{VCO}[s] = \frac{\Theta_{VCO}[s]}{V_{LPF}[s]} = \frac{K_{VCO}}{s}. \quad (6\text{-}11)$$

4. Frequency Divider

When the output signal is divided by N in frequency, the phase argument is also divided by N. Hence the transfer function of the frequency divider is simply a constant.

$$H_{FB}[s] = \frac{\Theta_{FB}[s]}{\Theta_{VCO}[s]} = \frac{1}{N}. \quad (6\text{-}12)$$

As the PLL is a feedback system, the stability issue is the first matter of concern. According to feedback theory, the open loop transfer function must have enough phase margin at the unity gain frequency, which is normally greater than about 45 degrees, to prevent the system from oscillating.

From the discussions above, the forward path transfer function of the PLL is the product of the transfer functions of the PFD-CP, the loop filter and the VCO, and is a third order transfer function.

$$G(s) = K_{PDFCP} \cdot \frac{1 + s \cdot T_2}{s \cdot (C_1 + C_2) \cdot (1 + s \cdot T_1)} \cdot \frac{K_{VCO}}{s}. \quad (6\text{-}13)$$

The transfer function of the feedback path of the PLL is the constant of the division ratio:

$$H(s) = \frac{1}{N}. \tag{6-14}$$

By substituting (6-13) and (6-14) into (6-1), the complete system open loop transfer functions can be obtained as:

$$A_{openloop}(s) = \frac{K_{PFDCP} \cdot K_{VCO}}{N} \cdot \frac{1 + s \cdot T_2}{s^2 \cdot (C_1 + C_2) \cdot (1 + s \cdot T_1)}. \tag{6-15}$$

In this equation, the VCO gain and the division ratio are normally pre-determined by the specifications and frequency planning. By making an assumption about the PFD-CP scaling values for the initial design, the loop filter component values are left as variables. Hence, the stability related parameters, loop bandwidth and phase margin, are determined by the loop filter design. As can be seen in Equation (6-15), the open-loop gain has two poles at DC, corresponding to a phase of -180 degrees. To ensure the requirement of a 45~60 degrees phase margin at unity loop gain, the zero and the non-DC pole should be placed below and above the unity gain loop bandwidth, respectively. The resulting phase margin at the unity gain loop bandwidth is hence:

$$\varphi = 180 + \arctan(\omega_u \cdot T_2) - \arctan(\omega_u \cdot T_1). \tag{6-16}$$

Given the fixed phase margin (system stability requirement) and loop bandwidth (phase noise/spur considerations), the two time constants can be calculated by taking the derivative of the phase margin and setting it to zero [90].

$$\begin{cases} T_1 = \dfrac{\sec(\varphi) - \tan(\varphi)}{\omega_u} \\ T_2 = \dfrac{1}{\omega_u^2 \cdot T_1} \end{cases} \tag{6-17}$$

The loop filter components can then be calculated by substituting equation (6-17) into (6-15) and setting equation (6-15) as one. The values of C_1, C_2 and R_2 can then be obtained accordingly. The results are directly given by:

$$\begin{cases} C_1 = \dfrac{K_{PDFCP} \cdot K_{VCO}}{N} \cdot \dfrac{T_1}{\omega_u^2 \cdot T_2} \sqrt{\dfrac{1 + \omega_u^2 \cdot T_2^2}{1 + \omega_u^2 \cdot T_1^2}} \\ C_2 = C_1 \left(\dfrac{T_2}{T_1} - 1 \right) \\ R_2 = \dfrac{T_2}{C_1} \end{cases} \tag{6-18}$$

The settling time, or locking time, corresponds to the transient response of the system. Instead of phase margin and loop bandwidth being given in the frequency

domain, the damping factor ζ and natural frequency ω_n are normally specified in the time domain, which are given as below [90].

$$\omega_n = \sqrt{\frac{K_{PFDCP} \cdot K_{VCO}}{N \cdot (C_1 + C_2)}}. \tag{6-19}$$

$$\xi = \frac{\omega_n}{2} \cdot R_2 C_2. \tag{6-20}$$

The transient response can then be calculated, given the start and stop frequencies (e.g. for a moderate change of divider setting), the damping factor and the natural frequency. By specifying the settling tolerance, *tol*, the locking time can then be obtained, and is often approximated by [90]:

$$T_{lock} = \frac{-\ln\left(\frac{tol}{f_2 - f_1}\sqrt{1-\xi^2}\right)}{\xi \omega_n}. \tag{6-21}$$

If the difference between the actual frequency and the targeted frequency is lower than the settling tolerance, *tol* in Hz, the PLL can be seen as locked.

In the spectrum monitor, the locking time is not necessarily required to be very short (as would be the case for say a frequency hopping system), because it takes some time for the receiver to evaluate the average energy falling into the selected channels anyway.

For the initial system level design, assume that the VCO gain is 3GHz/V, the geometric average of the division ratio is 201.83 (corresponding to 152~268), and a reasonable charge pump current of 30μA/rad. The phase margin is selected as 55 degrees, and then the loop bandwidth could be chosen as 2.5MHz, which is one-tenth of the 25MHz reference frequency. The loop filter components values can then be obtained from equations (6-18) to (6-20) as:

$$\begin{cases} C_1 = 0.57\ pF \\ C_2 = 5.16\ pF \\ R_2 = 39.12\ k\Omega \end{cases} \tag{6-22}$$

If the VCO output frequency needs to be changed from 3800MHz to 6700MHz, which is an extreme condition, and the tolerance is 1kHz, the transient response parameters are calculated as:

$$\begin{cases} \omega_n = 1.40\ MHz \\ \xi = 0.89 \\ T_{lock} = 1.99\ \mu s \end{cases} \tag{6-23}$$

The phase noise of the local oscillator and frequency synthesiser subsystem involves the combination of the system transfer functions and the noise generated by

each component. The overall phase noise should be lower than -80dBc/Hz for all the offset frequencies of up to 100MHz in order to meet the system requirement, according to the spectrum monitor receiver specifications discussed in chapter four. The spurious tone level, which is mainly due to reference frequency leakage through the PFD/CP and up/down mismatches in the charge pump, should also be suppressed to the same level, because of the wideband application of the PLL. All the blocks are effectively phase noise sources, including the reference oscillator, the phase/frequency detector, the loop filter, the voltage controlled oscillator, and the feedback frequency divider. From the system level point of view, these noise sources can be seen simply as additive noise at each node of the signal flow, as shown in Figure 6-3. The transfer functions from the phase noise sources are listed in Table 6-1.

It can be observed that the phase noise of the reference oscillator, the PFD-CP and the frequency divider are lowpass filtered by the PLL system, while the phase noise of the VCO and the loop filter are highpass filtered by the PLL system [91]. This observation implies that a compromise value of loop bandwidth should be chosen to maintain a good balance between the noise contributions of the different blocks. For example, a wider loop bandwidth can suppress the VCO's noise significantly, while the noise from the reference oscillator, the PFD-CP and the frequency divider might not be suppressed enough. A good rule of thumb is to select a loop bandwidth so as to keep all the noise sources at nearly the same level in the frequency offset range of most importance. In a practical design, the phase noise of the VCO usually dominates. Therefore, most applications are more concerned with getting a reasonably wide loop bandwidth so that the VCO noise can be filtered to a wide offset frequency. However, to avoid significant sampling effects, the loop bandwidth should always be selected to be much lower than the reference frequency. Therefore a reference oscillator with a high frequency is needed in this case.

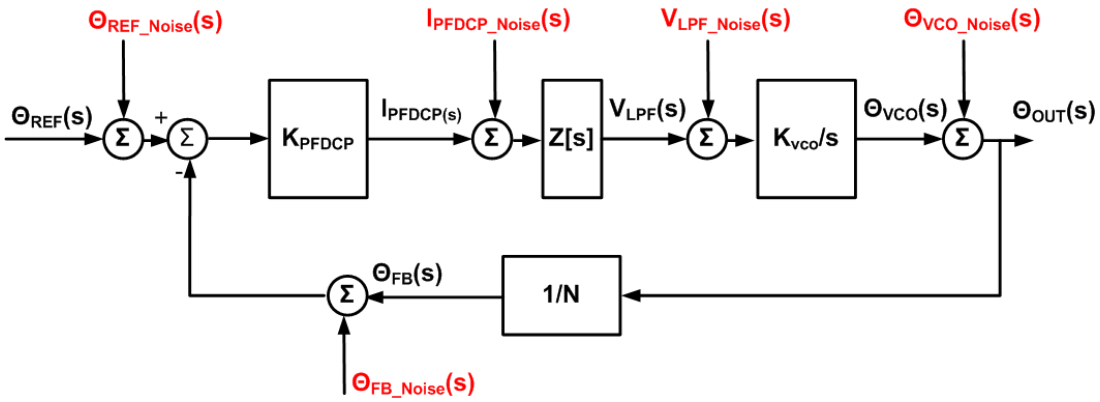


Figure 6-3 PLL phase noise model

From the above system analysis, the results of the overall PLL system level simulation using MATLAB program are shown below. The open and closed loop frequency responses are illustrated in Figure 6-4, verifying the phase margin and loop bandwidth. The transient response is plotted in Figure 6-5. Note that the overshoot of 7150MHz is under the extreme condition of a frequency change, and therefore this is the largest frequency tuning range that the VCO would be required to cover. The phase noise transfer functions are shown in Figure 6-6.

Reference Oscillator	$\frac{\Theta_{\mathrm{OUT}}}{\Theta_{\mathrm{REF\_Noise}}} = \frac{\frac{I_{PFDCP\_Noise} \cdot K_{VCO} \cdot Z(s)}{s}}{1 + \frac{I_{PFDCP\_Noise} \cdot K_{VCO} \cdot Z(s)}{s \cdot N}}$
Phase/Frequency Detector	$\frac{\Theta_{\mathrm{OUT}}}{\mathrm{I_{PFDCP\_Noise}}} = \frac{\frac{K_{VCO} \cdot Z(s)}{s}}{1 + \frac{I_{PFDCP\_Noise} \cdot K_{VCO} \cdot Z(s)}{s \cdot N}}$
Loop Filter	$\frac{\Theta_{\mathrm{OUT}}}{\mathrm{V_{LPF\_Noise}}} = \frac{\frac{K_{VCO} \cdot Z(s)}{s}}{1 + \frac{I_{PFDCP\_Noise} \cdot K_{VCO} \cdot Z(s)}{s \cdot N}}$
VCO	$\frac{\Theta_{\mathrm{OUT}}}{\Theta_{\mathrm{VCO\_Noise}}} = \frac{1}{1 + \frac{I_{PFDCP\_Noise} \cdot K_{VCO} \cdot Z(s)}{s \cdot N}}$
Frequency Divider	$\frac{\Theta_{\mathrm{OUT}}}{\Theta_{\mathrm{VCO\_Noise}}} = \frac{-\frac{I_{PFDCP\_Noise} \cdot K_{VCO} \cdot Z(s)}{s}}{1 + \frac{I_{PFDCP\_Noise} \cdot K_{VCO} \cdot Z(s)}{s \cdot N}}$

Table 6-1 Phase noise transfer functions in the PLL

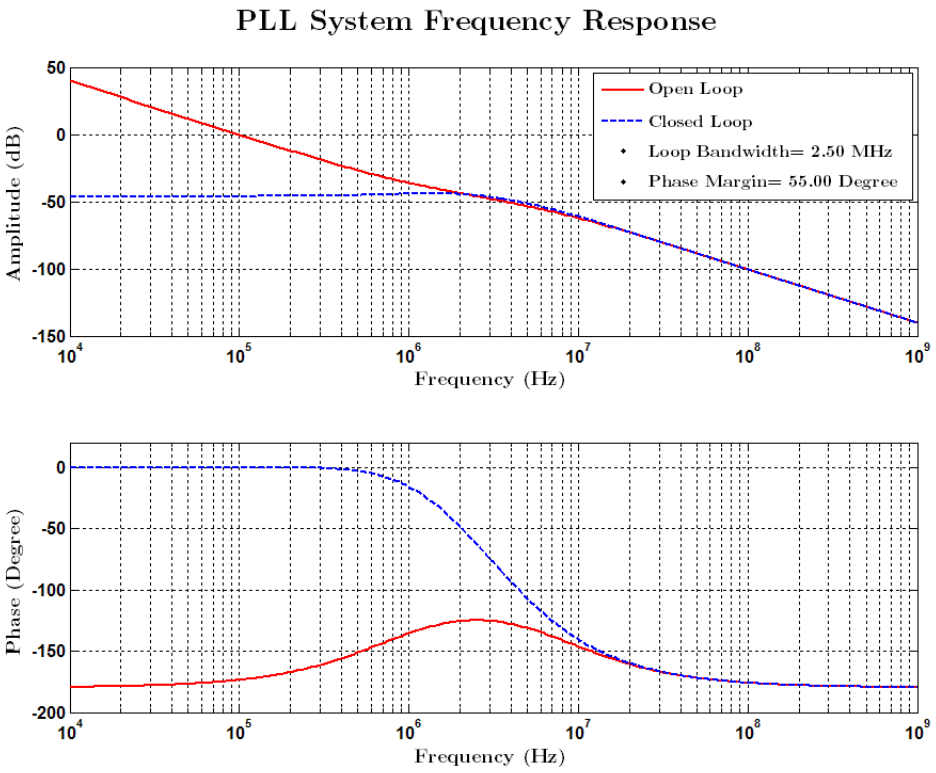


Figure 6-4 PLL frequency response

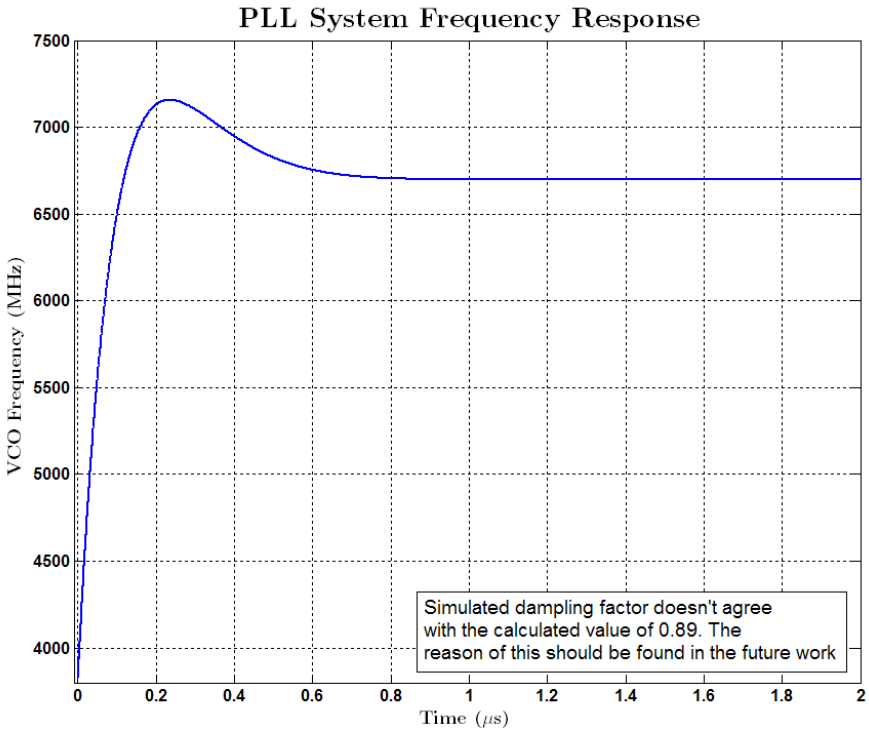


Figure 6-5 PLL transient response (Frequency change from 3.8GHz to 6.7GHz)

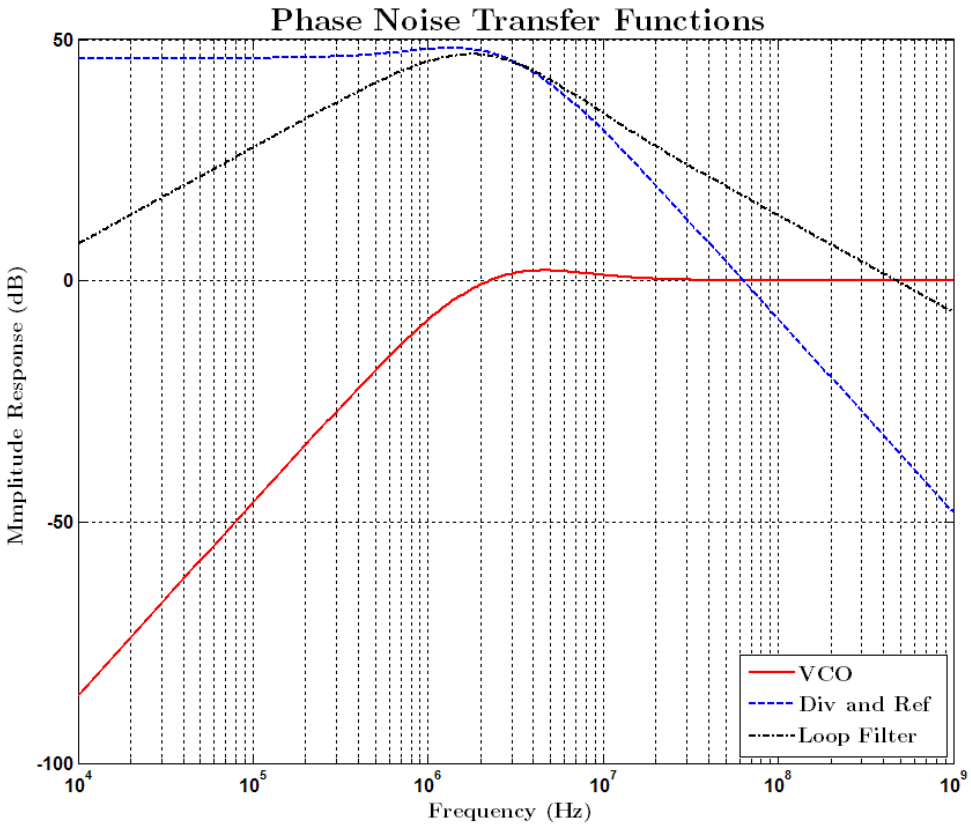


Figure 6-6 PLL phase noise transfer functions

6.3 Voltage Controlled Oscillator Design

The voltage controlled oscillator is the core block in a frequency synthesiser. As explained in chapter four, it is tuning range rather than the phase noise that needs more consideration in the spectrum monitor application. Furthermore, the chip area needs to be kept as small as possible in order to reduce the chip cost. Therefore, an LC VCO is not so attractive in this project because of its narrow tuning range and the large die area of the required inductor(s). On the other hand, despite the larger phase noise in a ring oscillator, by setting a proper loop bandwidth, it is expected that the system specifications can be met. In this project, a three stage, fully differential ring oscillator is used. Note that using three stage is to minimize the delay so that higher frequency can be achieved.

The ring oscillator is a feedback loop consisting of cascaded inverter stages, as shown in Figure 6-7. At a certain frequency, the cascaded delay cells will shift the phase of a signal by 180 degrees, plus another 180 degrees from the overall negative feedback, and hence the total phase shift around the loop is 360 degrees. Thus the

output oscillates at this frequency. To ensure that the circuit always oscillates, two conditions called the 'Barkhausen criteria' [92] must be satisfied:

$$\begin{cases} |H(j\omega_u)| \geq 1. \\ \angle H(j\omega_u) = 180°. \end{cases} \tag{6-24}$$

The minimum gain of each stage is usually not unity. Take the example of a three stage oscillator, if each stage's transfer function is $-A_0/(1+s/\omega_0)$, where ω_0 is the dominant pole of the stage, then the loop gain is

$$H(s) = -\frac{A_0^3}{\left(1+\frac{s}{\omega_0}\right)^3}. \tag{6-25}$$

If each stage contributes a 60 degree phase shift, the oscillation frequency is

$$\omega_{OSC} = \sqrt{3}\omega_0. \tag{6-26}$$

Hence, at this frequency, the loop gain equals unity:

$$\frac{A_0^3}{\left[\sqrt{1+\left(\frac{\omega_{osc}}{\omega_0}\right)^2}\right]} = 1, \tag{6-27}$$

and the required gain of each stage is then A_0=2.

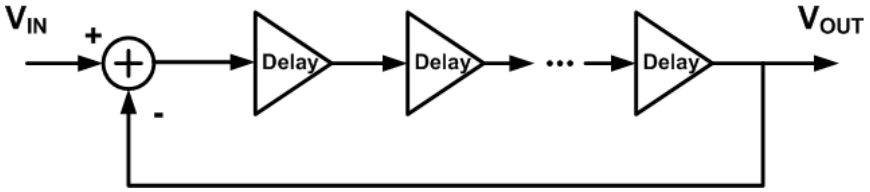


Figure 6-7 Ring oscillator structure

Note that the above oscillation conditions are for small signal analysis only. When the oscillator is oscillating steadily, the amplitude of the output voltage is nearly saturated with respect to the supply voltage. In this case, the circuits become nonlinear and the oscillation behaviour is essentially determined by the large signal, nonlinear current drive and the capacitances of each delay cell. The tuning frequency of the ring oscillator in this case is finally proportional to the stage delay, which is a large signal time domain parameter, and the number of stages [93].

$$f_{osc} = \frac{1}{2N \cdot T_D}. \tag{6-28}$$

In a practical design, the number of stages is usually fixed. Hence the stage delay is usually the variable that is used when tuning a ring oscillator, and so the key design

issue of the ring oscillator is the delay cell. There are two main types of ring oscillator delay cell: CMOS inverter based and differential pair amplifier based.

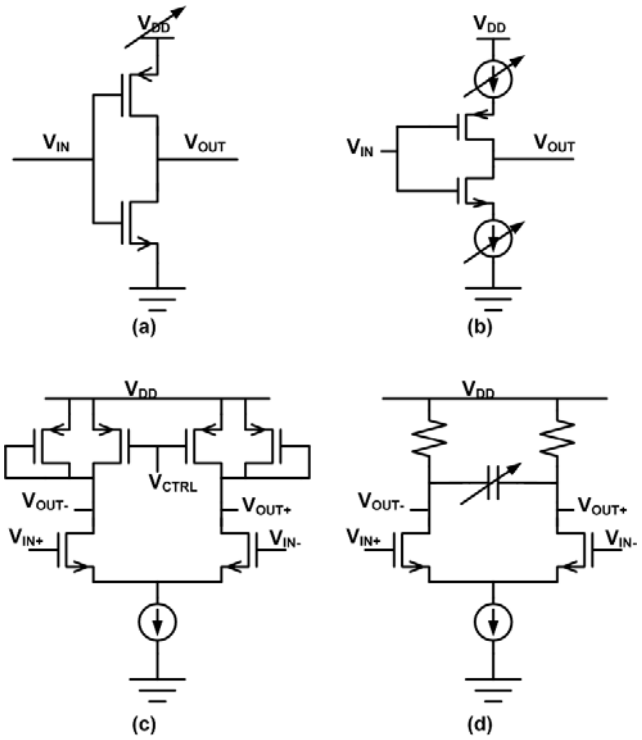


Figure 6-8 Ring oscillator delay cells (a) CMOS inverter tuned by VDD, (b) current starved inverter, (c) differential amplifier tuned by load resistance, (d) differential amplifier tuned by load capacitance.

The transmission delay is essentially controlled by the charging and discharging speed of the load capacitor. The tuning could be implemented by varying load resistance, load capacitance, or the bias current, and these methods are generally effective for both CMOS inverter and differential pair amplifier based delay cells. One of the most common examples of the delay cell is shown in Figure 6-8a, where a simple CMOS inverter is tuned by a variable power supply voltage. Note that it is essentially the current, and hence the g_m of the transistor, that is changed by varying the power supply voltage. Therefore the actual delay is approximately related to the g_m and load capacitance, and the capacitance is usually dominated by the input capacitance of the next stage. Increasing the device width doesn't alter the charging speed, because the load capacitance increases by the same rate as well. Another type of inverter based delay cell is called 'current starved' inverters [94, 95], as shown in Figure 6-8b. By varying the current sources, the g_m of the inverter MOS devices are also changed. For the differential amplifier delay cell, a common structure is shown in Figure 6-8c, where the PMOS loads are biased in the deep triode region, playing the role of variable resistors [96]. The delay cells tuned by resistance usually suffer from a non-linear

tuning voltage with respect to frequency, which could result in too large VCO gain in the centre of the tuning range. In Figure 6-8d, the load resistor is fixed, and the load capacitance can be designed as tunable so as to tune the oscillating frequency. This architecture is adopted in the proposed design and will be discussed in detail.

In this project, a novel differential amplifier based delay cell is designed, that is tuned by an array of MOS variable capacitances (MOS varactors) acting as the load, as shown in Figure 6-9. The differential amplifier delay cell in the three-stage ring oscillator has constant current consumption and a fixed load resistance. The capacitor array is comprised of five pairs of inversion mode MOS varactors. Five separate tuning voltages are needed to switch the MOS varactors on and to change the capacitances individually.

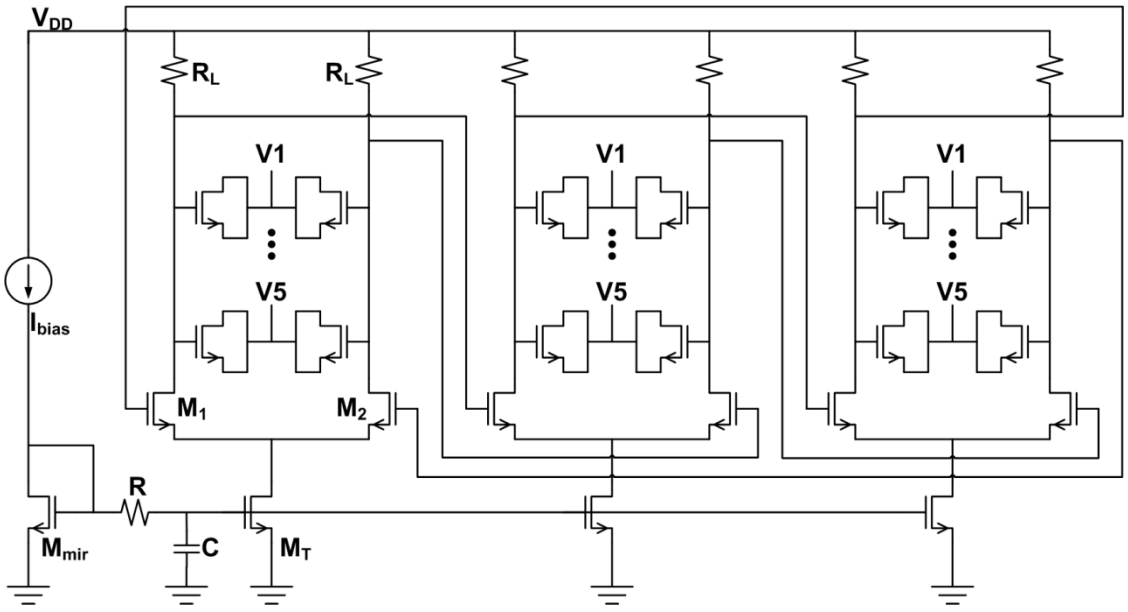


Figure 6-9 Proposed ring oscillator schematic

The tuning range of this scheme is not as wide as in a ring oscillator where the delay is varied by varying the load resistance or the bias current because of the inherent character of the MOS varactor capacitance variation range, which is normally less than 3:1 [97]. Despite this, it is still wider than for LC oscillators because the delay time, and hence the frequency, is proportional to the capacitance in a ring oscillator, while it is proportional to the square root of the capacitance in an LC oscillator.

$$f_{tune\_ring} \propto \frac{g_m}{C_{load}}.$$
$$f_{tune\_LC} \propto \frac{1}{\sqrt{L \times C}}. \quad (6\text{-}29)$$

The spectrum monitor requires a tuning ratio of 67:38, which is approximately 1.8:1, so it is expected that there will be enough margin to account for the effects of parasitics and tolerances from the differential pair NMOS devices. The configuration of the NMOS varactor is seen in Figure 6-9. The gate terminal is connected to the common mode of the VCO output at the middle point of the supply (1.2/2=0.6V). The connected source-drain terminals are connected to the tuning voltage. Note that the bulk terminal of the NMOS is connected to ground so that the device never enters the accumulation region (V_G is much lower than the bulk voltage). This configuration is referred as 'Inversion Mode' in [97]. When the gate-source voltage is lower than the threshold voltage, there are few mobile charge carriers (electrons) in the gate-oxide interface area, and the MOS capacitance is very small. As the gate voltage increases, an inversion layer channel is formed below the gate until the gate voltage is above the threshold voltage and the MOS device enters strong inversion, where the MOS capacitance reaches the largest value of C_{ox}.

The capacitance versus gate voltage is shown in Figure 6-10. Note that the relative capacitance variation is over a ratio of 3:1, and is steeply linear for a gate-source voltage from 200mV to 400mV, while capacitance is almost constant when gate-source voltage is lower than 100mV and higher than 500mV. Hence this could be the potential area of operation. Furthermore, the almost linear relationship between capacitance and gate-source voltage is one of the most important advantages of the MOS varactor tuning scheme.

A further advantage of this circuit compared with the complicated PMOS load arrangement is that the polysilicon resistor loads minimize the parasitic capacitance and load noise simultaneously. In addition, the fixed value resistors and constant power consumption eliminate the need for replica bias circuits, reducing the power consumption and circuit complexity significantly.

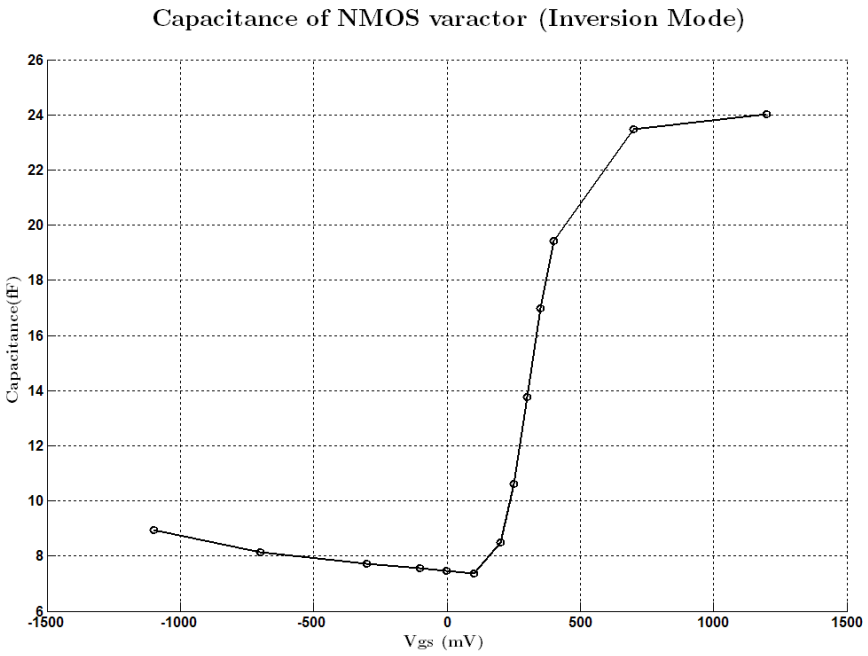


Figure 6-10 NMOS varactor capacitance of ST 130nm CMOS

The loop filter's output voltage can be within almost the whole range between VSS and VDD, and this is generally desirable in order to minimise corruption of the control voltage by external noise sources. However, this large range needs to be converted to five separate voltages to control the MOS varactor pairs separately. This conversion scheme is shown in Figure 6-11. As mentioned above, the VCO common-mode level is about 600mV, which is the middle point of the power supply, and is connected to the gate terminal of the MOS varactor. When the filter's output voltage, Vtune, sweeps from 0V to 1.2V, the initial source terminal voltages remain at constant values and the gate-source voltages are less than 100mV, so that the capacitances are almost constant as well. With Vtune increasing continuously, the source terminal voltage of the first pair of varactors starts to decrease and the capacitance starts to increase when the gate-source voltage exceeds 100mV, until the source terminal voltage is more than 500mV below the gate voltage, where the capacitance stops increasing and the source voltage stops decreasing. At this point, the source terminal voltage of the second pair of varactors starts to change in the same way as the first one, and hence the capacitance also changes accordingly. In this way, the MOS varactors' capacitance values increase and saturate one by one with increasing Vtune, so that the total load capacitance of the differential amplifier keeps accumulating until all the varactors reach their maximum capacitance value. The inversion cell's delay is then

increasing proportionally and the oscillating frequency is reducing in inverse proportion.

By using this configuration, the unwanted small offset of the filter output voltage can only influence one or two the capacitance of pairs of varactors, while the other capacitances are maintained either at their maximum or minimum values. Therefore the total capacitance variation due to the control voltage offset is significantly reduced.

In the practical design, the effective control voltage is set from 200mV to 1V, and each of the varactors' source terminal voltages varies from 500mV to 100mV. The control voltage requires some margin with respect to the power supply voltages, and this is mainly related to the voltage required to maintain the charge pump MOS devices in their saturation regions.

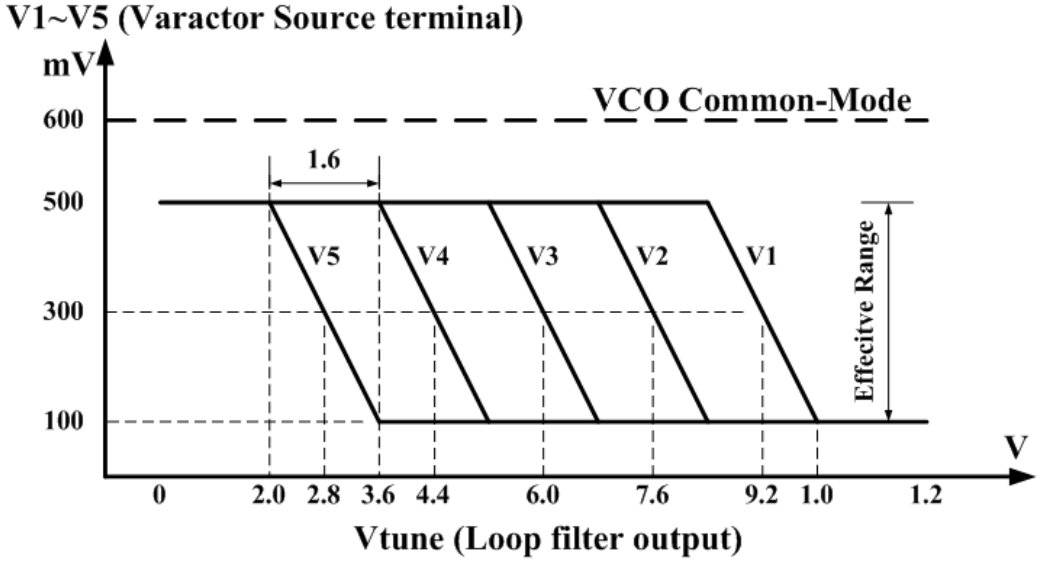


Figure 6-11 Ideal behaviour of varactor tuning scheme

To realize this tuning scheme function, a special tuning circuit has been designed, as shown in Figure 6-12. Five differential circuits generate the control voltages V1~V5, respectively. A resistor voltage divider branch sets the reference voltages as 280mV, 440mV, 600mV, 760mV and 920mV. When Vtune is low, Mx_L is on and the majority of the tail current, Ix_{tail}, flows through Mx_L and Rx_B. The voltage Vx approximately equals to $Ix_{tail}\times(Rx_A+Rx_B)$. Note here x=1~5, representing the five differential stages. As Vtune increases to 200mV, the current in $M5_L$ is gradually steered to $M5_R$ and finally most current flows through $M5_R$ when Vtune reaches to 360mV, and the voltage V5 is approximately equal to $I5_{tail}\times R5_A$. By adjusting the values of $I5_{tail}$, $R5_A$ and $R5_B$, the middle point of the current steering happens when Vtune reaches 280mV and V5 is reduced to 300mV. As Vtune keeps increasing, other differential pairs experience the

same process, and all the control voltages are generated. Note that each differential pair and the resistors should be calculated and configured separately because the common mode levels are different.

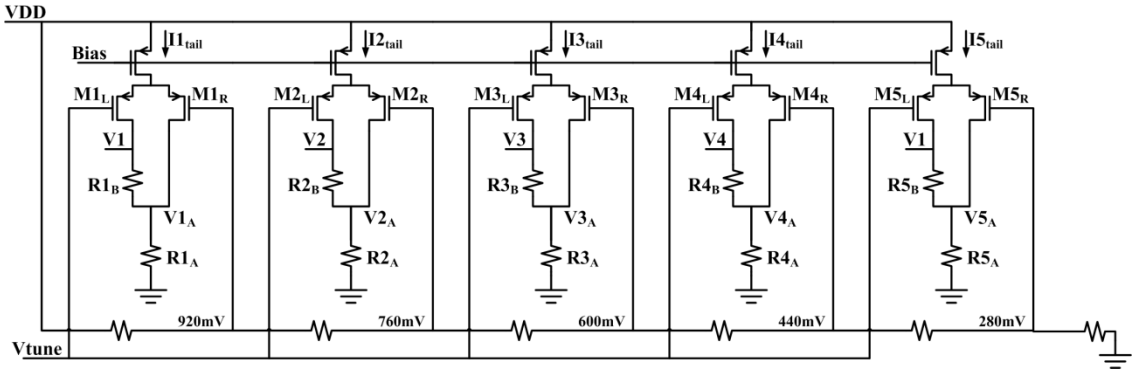


Figure 6-12 Tuning circuits

The phase noise sources in the ring oscillator include the differential pairs, the tail current sources, the current mirror, the load resistors (and the loss terms in the capacitors) and the control voltages. The control voltage noise is normally minimized by lowering the loop bandwidth and minimising the current spikes from the charge pump. The MOS varactor noise comes from channel resistance, which can be largely reduced using short channel length and wide devices. The main white noise sources in the ring oscillator are generated by the differential pairs, the tail current source and the load resistors. The sum of these white noise sources is given by [56].

$$\mathcal{L}(f_m) = \frac{2kT}{If_0 \ln 2}\left(\gamma\left(\frac{3}{4V_{effd}} + \frac{1}{V_{efft}}\right) + \frac{1}{V_{op}}\right)\cdot\left(\frac{f_0}{f_m}\right)^2 \tag{6-30}$$

Here, V_{effd} is the effective gate voltage of the differential pair, which represents the differential pair noise, and the V_{efft} term represents the tail current noise. The term V_{op} is the output voltage swing. To minimize the phase noise contribution due to white noise, a general design consideration is to set the overdrive voltage of the differential pair and the tail current source at a relatively high level, as long as not too much vertical field mobility degradation is introduced to reduce the transconductance. At the same time, the load resistances should be increased, and thereby also the output swing, so that the transistors are biased just at the onset of the strong inversion region, as long as the speed is still guaranteed. Note that V_{effd}, V_{efft} and V_{op} are fractions of the supply voltage, and these considerations in turn increase the current, and hence the power consumption. As a result, a trade-off must be made between phase noise, speed and power consumption.

Because of the modulating effect of the oscillator, the flicker noise is up-converted to appear as sidebands of the oscillator carrier frequency [53], and sometimes could be the dominant noise sources. The flicker noise sources include the differential pair devices, the tail current devices, and the current mirror device. The flicker noise can be seen as a slow fluctuation in the time domain, these fluctuations generally becoming much longer than the delays of the cells. As analysed in [56], the flicker noise originating in the differential pair appears at the second harmonic of the oscillating frequency and can be ignored. The flicker noises from the tail current sources are uncorrelated, so the total noise from these is proportional to the number of stages. On the other hand, flicker noise due to the current mirror devices drives the common gate voltages of all of the stages and is therefore correlated to each other, and is proportional to the square of the number of stages. Further, the current mirror device size is normally a fraction of the main tail current source devices, and hence the flicker noise frequency corner is much higher than that of tail current source. Consequently, the current mirror contributions dominate the flicker noise. By placing a large capacitance at the gate of the current mirror, the current mirror flicker noise and the delay cells can be decoupled [98]. Note that for differential ring oscillators, whose control voltage is also the gate voltage of the current mirror, this large capacitance results in an excessively long settling time. In the proposed ring oscillator, this isn't a problem because the tuning voltage is connected to MOS varactors, which is another advantage of the structure. To minimize the current mirror induced flicker noise, the number of delay cells should be reduced as far as possible; hence the three-stage oscillator in this design is a good choice. In addition, increasing the current mirror width by a reasonable amount to lower the flicker noise corner frequency also improves the phase noise performance.

6.4 Frequency Divider Design

Instead of the popular fractional-N dividers [99], the need for a very wide continuous range of division ratios suggests the use of an integer-N divider. In this section, an integer-N divider is designed to tune the local oscillator frequency continuously from 3800MHz to 6700MHz with steps of 100MHz. The most common integer-N divider is the dual modulus divider [100]. This type of divider involves a dual-modulus pre-scalar (P/P+1), a main counter (M) and a swallow counter (S), as well as the control circuits for these. The division ratio can be expressed as $N=P\times M+S$,

where $M=2^m$ and $0 \leq S \leq M$. If the reference is set to 100MHz or a fraction η of 100MHz, for example 25MHz when η=4, the division ratio should be an integer in the range from 38×η~67×η, with steps equal to η. However, calculations show that it is very difficult when using the normal dual modulus architecture to find an effective combination of P, M and S to achieve the required division ratio. In fact, the dual modulus frequency divider is really only suitable for narrow to medium tuning range applications. Consequently, a novel frequency divider architecture is developed based on a high frequency digital counter and reset logic circuits.

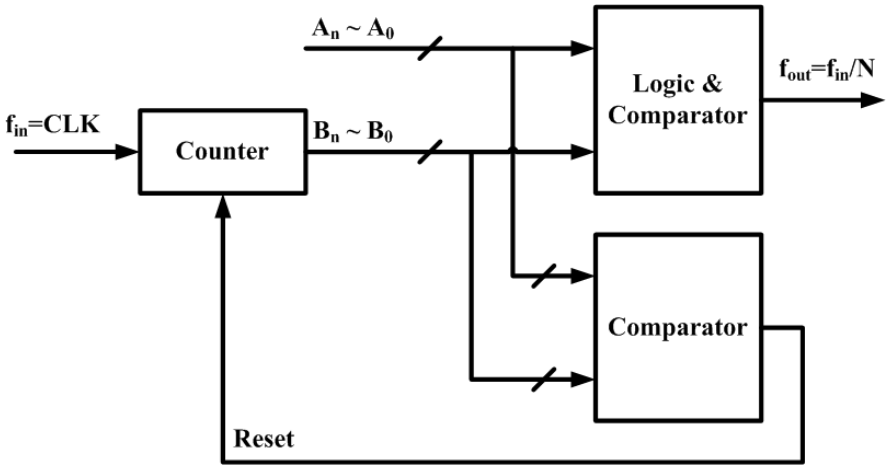


Figure 6-13 Integer-N frequency divider principle

Essentially, as shown in Figure 6-13, a frequency divider keeps comparing the output of an n-bit counter with pre-set values, and resets the counter when some defined criteria are met. The divided output signal is obtained by comparing and decoding the actual counter's output. Assume that the reference frequency is 25MHz, and the maximum division ratio would be 6700MHz÷25MHz=268, and hence a 9-bit counter is needed. The counter, comparator and logic blocks could be implemented either in customised high speed gates or using conventional rail-to-rail CMOS digital circuits. A customised high frequency divider would use differential current mode logic (CML) D-flip-flops and other logic cells, leading to high power consumption, while a pure rail-to-rail CMOS logic frequency divider generally can't reach speeds of several GHz. Consequently, a combination of logic styles is be needed to complete the function. The design goal is to minimize the usage of CML circuits while guaranteeing the speed.

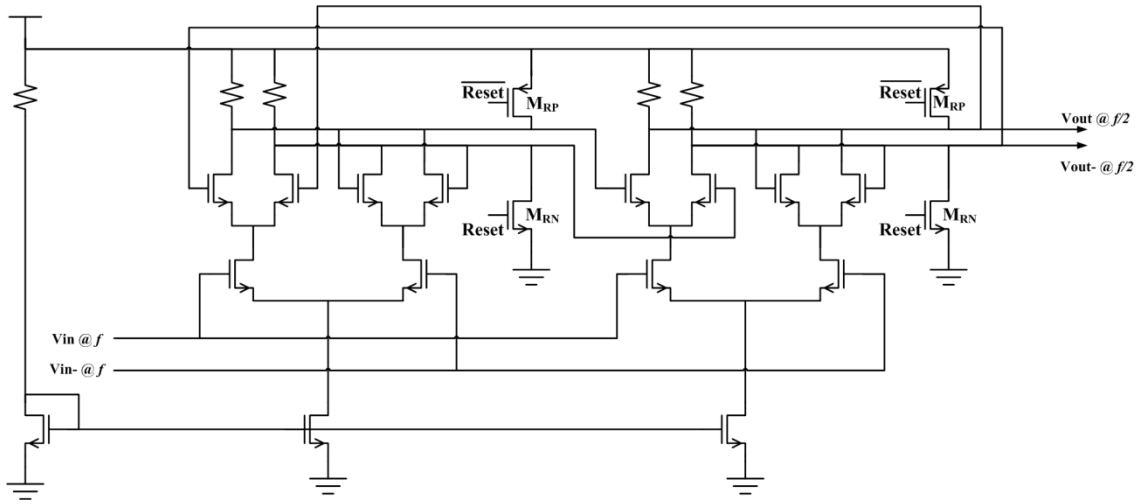


Figure 6-14 CML D-latch schematic (with Reset)

The diagram of the CML divide-by-2 block is a master-slave D-flip-flop circuit consisting of two CML D-latches, as shown in Figure 6-14. Note that the reset function is implemented by a pair of NMOS and PMOS with drain terminals connected to the differential outputs. When the reset is disabled, both MOS transistors are switched off, contributing as part of the load capacitances. When the reset is enabled, the gate voltage of NMOS M_{RN} is at the logic high level of V_{DD}, and hence it conducts all the current to ground, so that the voltage of Q is grounded. In the mean time, the PMOS M_{RP} is also switched on, supplying current from the power supply to the output node of $\overline{Q}$, and hence finally holding the voltage of $\overline{Q}$ at V_{DD}.

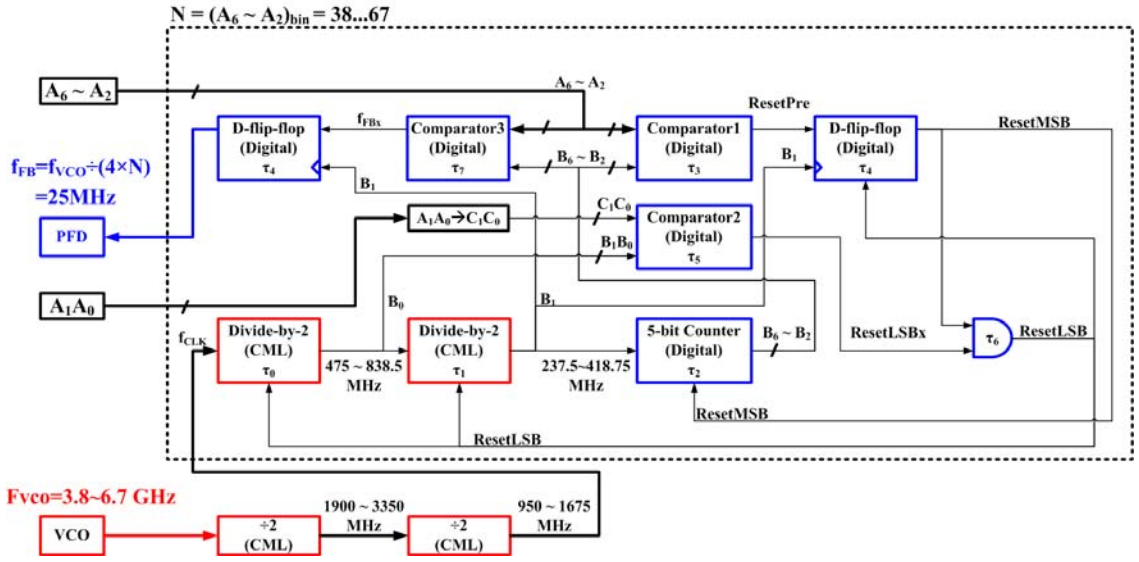


Figure 6-15 Integer-N frequency divider architecture

The proposed frequency divider architecture is illustrated in Figure 6-15. The red blocks are in the CML digital domain and blue blocks are in the CMOS digital domain. The PLL uses a 25MHz crystal oscillator as the reference signal, and therefore the

division ratio is set from 38×4 to 67×4 with steps of 4. The VCO output frequency, which is from 3.8GHz to 6.7GHz, is first fed into two cascaded divide-by-2 prescalers, which are standard high speed CML master-slave D-flip-flops. The resulting frequency f_{CLK} is set as the clock signal for the following integer-N divider, which has a continuous division ratio of 38~67, corresponding to a 7-bit counter. The clock frequency is from 950MHz~1675MHz, and hence the required following frequency division procedure must be completed within 597pS. Simulation shows that the delay of a 7-bit CMOS counter plus CMOS digital comparators is still not fast enough for this time scale. Therefore, the 7-bit division ratio is split into CML and CMOS digital circuits.

In Figure 6-15, the bus A_6~A_0 is the desired division ratio, and the bus B_6~B_0 is the actual output code, which should be equal to A_6~A_0 to generate the reset signal. The two least-significant-bits (LSB), B_0 and B_1, are obtained from two CML D-flip-flop based divide-by-2 circuits, respectively. The frequency of B_1 is from 237.5MHz to 418.75MHz, falling into the CMOS operation range of standard CMOS logic, and hence is configured as the *clock of the remainder of the digital circuits*. A 5-bit CMOS synchronous counter generates the five most-significant-bits (MSB), B_6~B_2. The Comparator1 compares the B_1B_0 with C_1C_0, which are mapped from A_1A_0, while Comparator2 compares the B_6~B_2 with A_6~A_2, and the comparison results are used to reset the CML dividers and digital counters. Finally, the divided signal is generated from output of Comparator3.

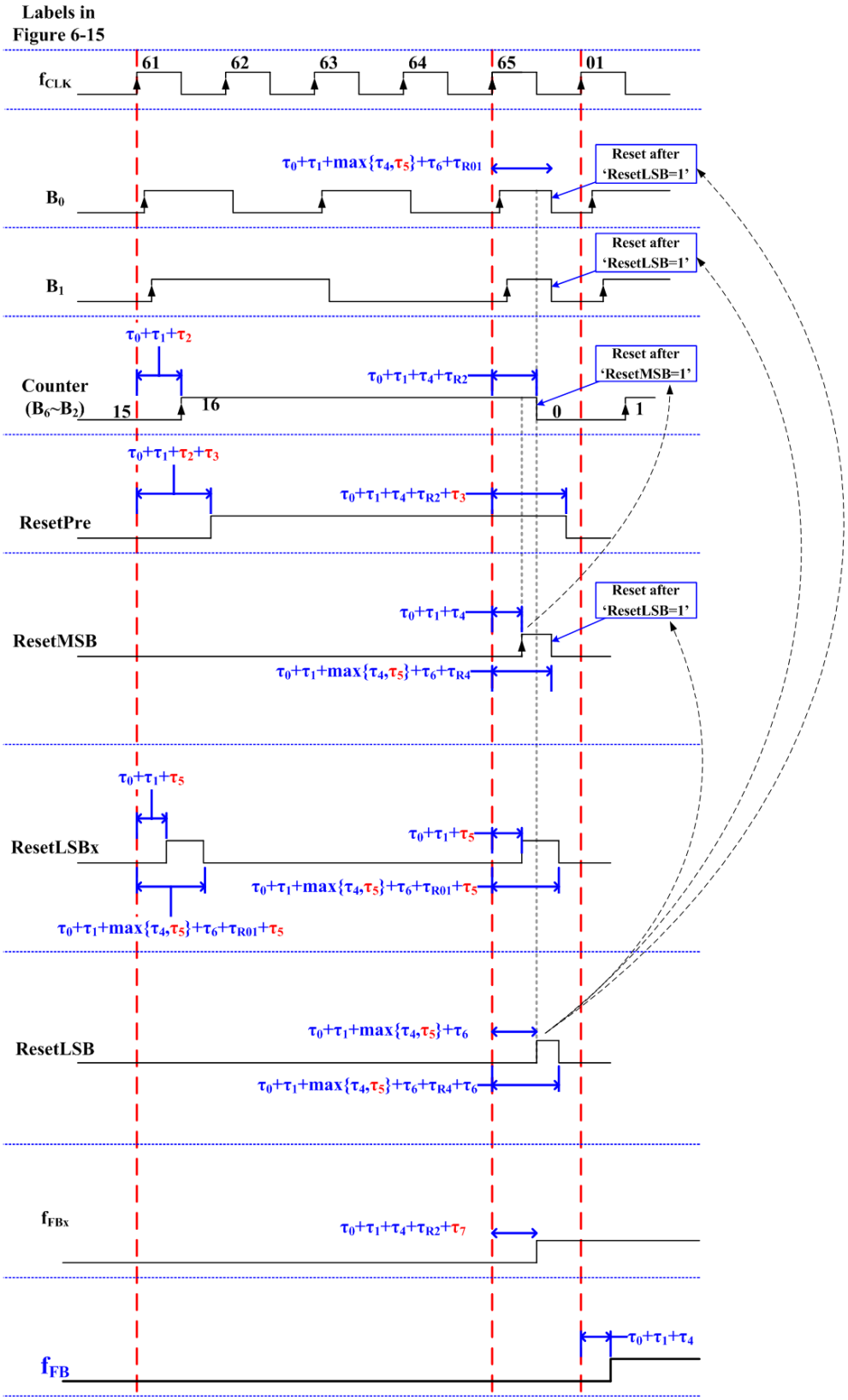


Figure 6-16 Integer-N frequency divider timing diagram

In the divider architecture, Figure 6-15, the label τ_x in each block represents its transition delay. Note that rather than specifically referring to a NAND gate, the expression of 'gate' below could be any of the NAND, NOR, AND, OR and XOR for

convenience. τ_0 and τ_1 are the time delays of the CML dividers, which could be much faster than their CMOS counterparts at the cost of higher current consumption. The 5-bit CMOS synchronous counter's time delay, τ_2, is generally the sum of three gates plus a CMOS D-flip-flop. The 5-bit comparator, Comparator1, has the time delay of τ_3, consisting of four cascaded gates, while the time delay of the 2-bit comparator, Comparator2, is τ_5, corresponding to three gates. The transmission gate based CMOS D-flip-flop has the time delay of τ_4. Note that τ_4 is a fraction of τ_2 because the counter includes several D-flip-flops. The delay of the final reset decision gate is τ_6. The Comparator3, which cascades three gates and is used to generate the final frequency divided signal, has the delay of τ_7. The resetting time delays of the CML dividers, 5-bit counter, and CMOS D-flip-flops are $\tau_{R01,}$ $\tau_{R2,}$ and τ_{R4}, respectively, and they are much shorter than the other delays. Among these delays, the longest ones are τ_2 and τ_3. As will be analyzed below, the proposed divider manages to avoid the cascade of long delay blocks in one f_{CLK} period, and hence guarantees the operating speed.

The division ratio can be expressed as N=m×16+n. Taking the example of F_{VCO}=6.5GHz, the required division ratio of 65 can be expressed by N=4×16+1, with m=4 and n=1, and the corresponding binary bits A_6~A_0 are $(1000001)_{bin}$. This means one period of f_{FB} contains 16 periods of B_1 and 1 period of f_{CLK}. The timing diagram is shown in Figure 6-16. In the following expressions, all the 'time delay' terms are defined with respect to the rising edge of the current f_{CLK} period unless specially mentioned. Also, all the reset flags are enabled with logic high for illustration purposes, although in the real design this is normally logic low. The 5-bit digital counter starts to count B_6~B_2 from 1 to 16, corresponding to the f_{CLK} periods from 1 to 64, with the delay of $\tau_0+\tau_1+\tau_2$ each time. Recall that the digital counter's clock input is B_1 rather than f_{CLK}. In the 61st f_{CLK} period, after counting to 16, Comparator1 sets the flag 'ResetPre' to high to prepare the reset in the next 17th B_1 (65th f_{CLK}) period. This is done one period before the actual reset cycle to avoid attempting the counting and comparing functions within one f_{CLK} cycle. The total time delay of this step is $\tau_0+\tau_1+\tau_2+\tau_3$, which can be designed to be lower than 600ps. Within the 65th f_{CLK} period, all the digits B_6~B_0 should be reset to zero to finish the divide-by-65 operation. Instead of continuing to count to 17, the rising edge of B_1 triggers the digital D-flip-flop and sets the flag 'ResetMSB' to a high state with a delay of $\tau_0+\tau_1+\tau_4$. This flag is to reset the 5-bit digital counter's value B_6~B_2 to $(00000)_{bin}$, within the delay of $\tau_0+\tau_1+\tau_4+\tau_{R2}$. This logic high 'ResetMSB' is also one of the two inputs of the 'ResetLSB' decision gate,

which is used to reset the two LSB digits B_1B_0 when both inputs are logic high. The CML outputs B_1B_0 are always compared by Comparator2 with C_1C_0 and the output is the flag 'ResetLSBx'.

The one-to-one mapping from A_1A_0 to C_1C_0 is $(01)_{bin}\rightarrow(11)_{bin}$ for N=65. This means that when $(B_1B_0)_{bin}$ equals $(11)_{bin}$, 'ResetLSBx' is set to high, indicating n=1 and $(A_1A_0)_{bin}=(01)_{bin}$. Note that this also happens periodically with f_{CLK} periods numbered 61^{st}, 57^{th}, 53^{rd}, etc. For this reason, the 'ResetLSB' decision gate needs another input 'ResetMSB' to determine when m=16 as well. The time delay of 'ResetLSBx' is equal to $\tau_0+\tau_1+\tau_5$. Note that both of the flags 'ResetMSB' and 'ResetLSBx' are set to logic high at the rising edge of B_1, and therefore the flag 'ResetLSB' is also set to high within the time delay of $\tau_0+\tau_1+\max(\tau_4, \tau_5)+\tau_6$. The logic high 'ResetLSB' is then used to reset the two CML dividers and the D-flip-flop simultaneously. The output digits from the CML divider, B_1B_0, are set to $(00)_{bin}$ with the delay of $\tau_0+\tau_1+\max(\tau_4, \tau_5)+\tau_6+\tau_{R01}$. This is followed by 'ResetLBSx' being set to low with a further delay of τ_5. Meanwhile, the output of the D-flip-flop, 'ResetMSB', is reset low by 'ResetLSB' with the delay of $\tau_0+\tau_1+\max(\tau_4, \tau_5)+\tau_6+\tau_{R4}$, followed by 'ResetLBS' itself being set low with a further delay of τ_6. Note that the operations 'ResetMSB' and 'ResetLSB' form a loop, so the delay of τ_6 is required to be long enough to avoid unpredicted values. In the 65^{th} f_{CLK} period, all the digits B_6~B_0 are reset to zero in a time less than $\tau_0+\tau_1+\tau_6+\max(\tau_4, \tau_5)+\tau_{R01}$ delay, and it is observed that the only block with a long delay is τ_5, which is present only once, so that this total time delay can be limited to one f_{CLK} period without much difficulty. All the flags, including 'ResetPre', 'ResetMSB', sResetLSBx' and 'ResetLSB', are set to high and then to low at different time slots during the 61^{st} to 65^{th} f_{CLK} periods. The long block delay τ_3 is involved once in 'ResetPre' and τ_5 is involved twice in 'ResetLSBx'. The total delays can be also limited to one f_{CLK} period with carefully design and layout.

The final output frequency of the divided signal is generated by Comparator3 and a D-flip-flop. If $(B6\sim B2)_{bin}$ is less than m/2, the output of Comparator3, f_{FBx}, is set to high. Otherwise it is set to low. For N=65 in this example, the signed f_{FBx} goes high with a further delay of τ_7, after $(B6\sim B2)_{bin}$ is set to zero during the 65^{th} f_{CLK} period. This delay also lies within the same f_{CLK} period. The final value of f_{FB} tracks the value of f_{FBx} with the trigger of the next f_{CLK} rising edge, completing the frequency division with time delay of $\tau_0+\tau_1+\tau_4$,

6.5 Phase/Frequency Detector and Charge Pump design

The structures of the PFD/CP and loop filter are shown in Figure 6-17. The phase/frequency detector compares the phase error between the two inputs: reference oscillator, f_{REF}, and feedback signal from the frequency divider, f_{FB}. This phase error is a voltage pulse, which is then converted to a current for charging or discharging the loop filter. The voltage developed at the filter output as a result of this current flow represents the VCO control voltage. The Q outputs of the D-flip-flops go high at the rising edges of their respective clock signal, and the D-flip-flops are reset when both Q outputs are high.

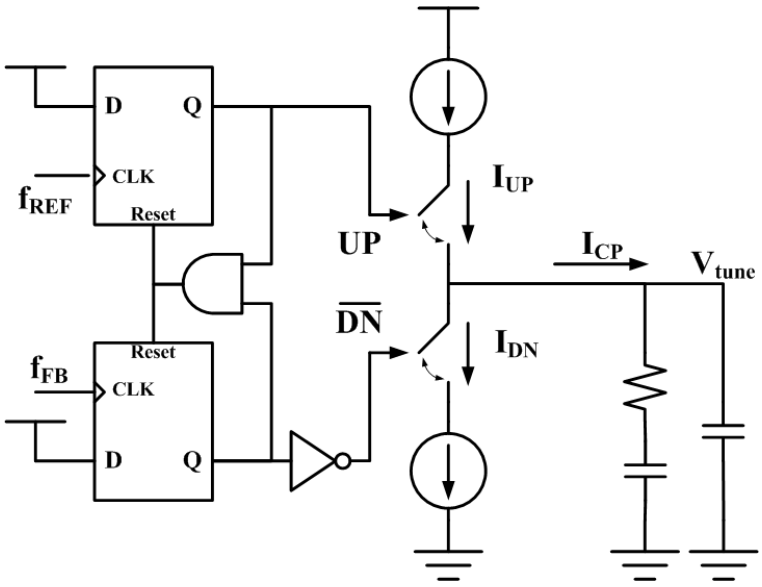


Figure 6-17 PFD-CP-Loop filter structure

An example signal flow diagram is shown in Figure 6-18. Assume that the loop is closed at time t_0, when the phase of f_{FB} lags with respect to f_{REF} and has a lower frequency than f_{REF}. Hence a voltage pulse appears at the 'UP' node with a width equal to the phase difference, while the voltage at the '$\overline{DN}$' node remains unchanged. During this pulse, the PMOS is off and the NMOS is still on, so the charge pump sinks current from the filter's capacitors, and thus the voltage Vtune is lowered, increasing the VCO frequency. This procedure repeats until the phases of f_{REF} and f_{FB} are aligned, reaching the 'lock' condition. The UP and $\overline{DN}$ branches of PFD-CP should be designed symmetrically so that for a defined active time interval equal current is sourced or sunk, minimizing any mismatch induced offset in the value of Vtune as well as the ripple. Note that during the 'lock' condition, enough delay should be introduced deliberately into the reset path (AND gate) so that coincident narrow pulses are generated on UP

and $\overline{DN}$. This is to ensure that there is enough time for the charge pump switches to be fully turned on in case a very small phase difference occurs, and thus to avoid the well know 'dead-zone' problem [99].

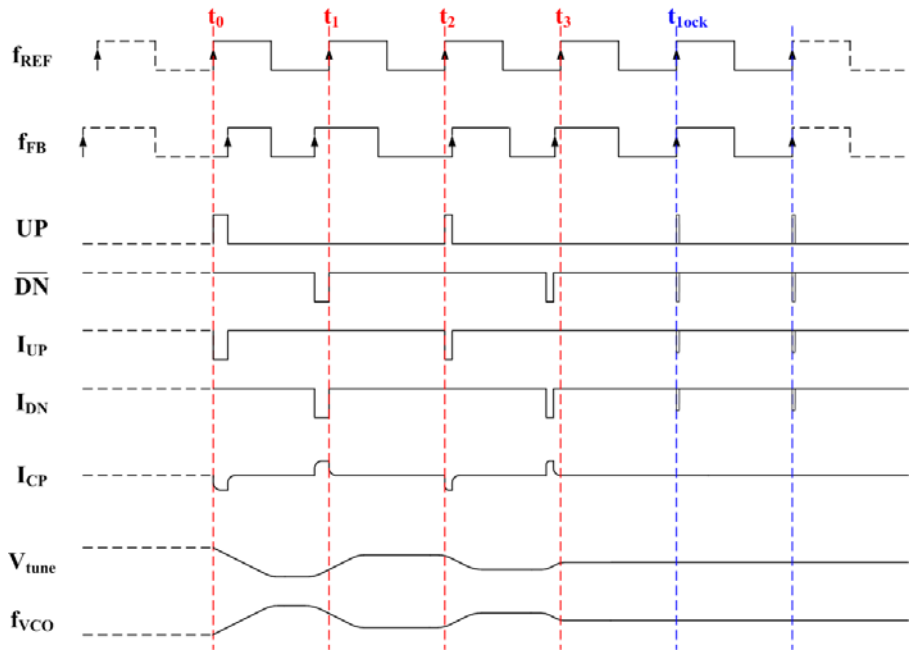


Figure 6-18 PFD-CP-Loop filter waveform diagram

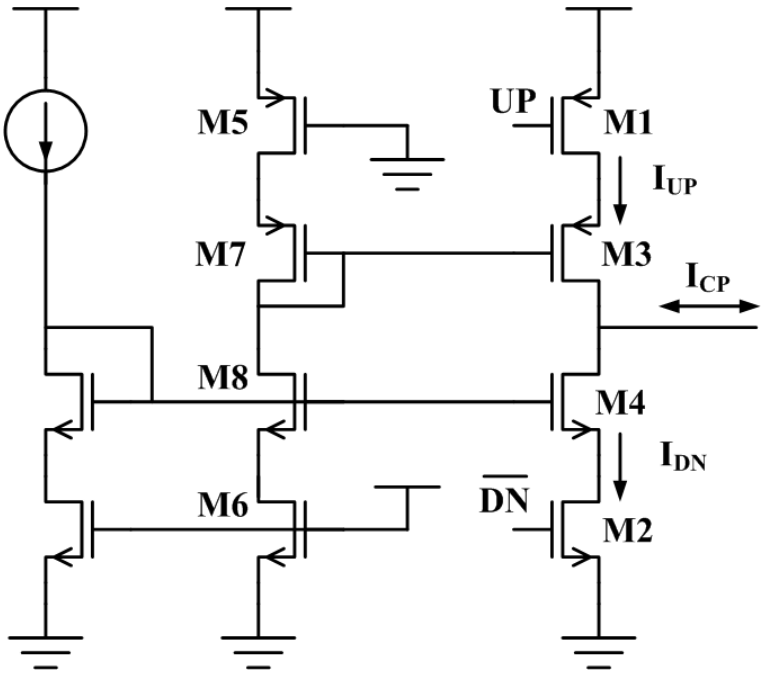


Figure 6-19 Charge pump circuit

The PFD is implemented in standard CMOS logic cells provided in the library of the technology supplier's design-kits (ST 130nm). The circuit of a single-ended charge pump is shown in Figure 6-19. The PMOS M1 and NMOS M2 are switches sensing the up and down signals from the PFD, diode connected M7 and M8 form current mirrors

for the current source devices M3 and M4 in the switches branch. The MOS devices M5 and M6 are dummy branches replicating the ON resistances of M1 and M2, respectively, and hence the gate of M5 is grounded and the gate of M6 is connected to the power supply.

6.6 Implementation, Simulation and Measurement

The completed frequency synthesizer's layout is shown in Figure 6-20. The digital and analogue circuits are separated and use different power supplies to reduce the coupling of noise from the digital circuits to the analogue functions. Each block is surrounded by power ring and ground ring structures. The output buffer, which includes a differential to single-ended amplifier and a source follower to drive the off-chip 50Ω load, takes the signal from the first divide-by-two output so that the main VCO's frequency won't be influenced during testing. The analogue circuits including the VCO, tuning voltage generator, loop filter (implemented with fringe capacitors and polysilicon resistor) as well as the strictly digital CML occupy a die area of 140μm×150μm, while the CMOS digital block's area is 50μm×50μm.

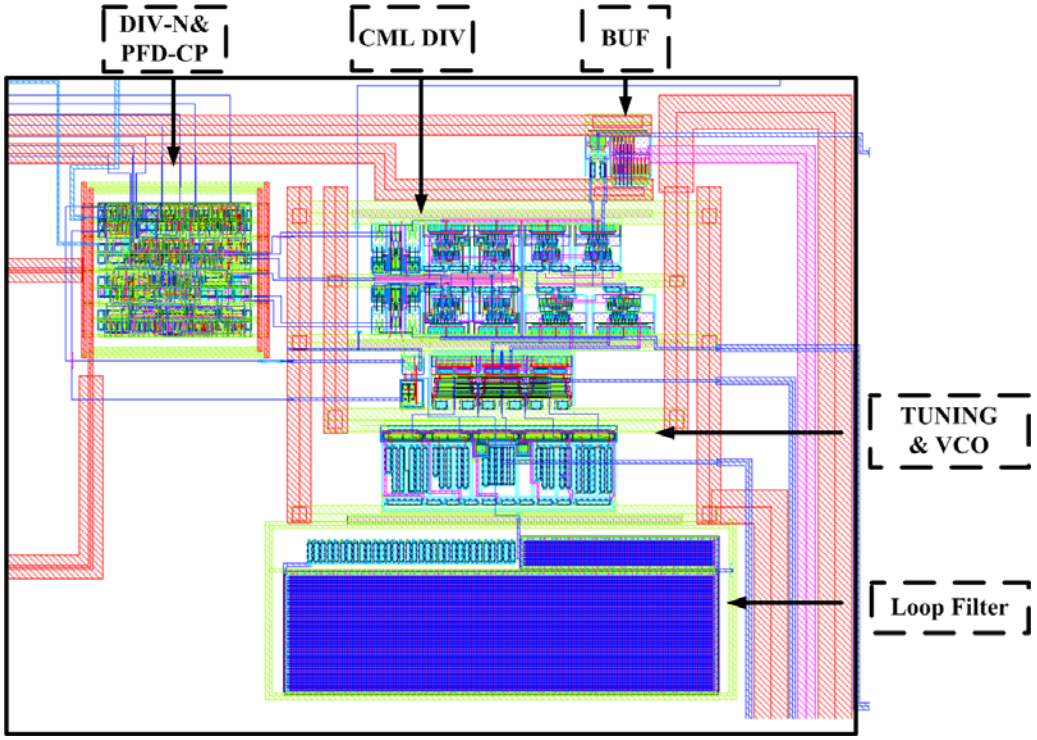


Figure 6-20 PLL layout

The test board picture is shown in Figure 6-21. Digital and analogue power supplies are provided by two tunable voltage regulators. The 25MHz reference frequency is generated by an on-board JFET Colpitts crystal oscillator. The 5V

reference signal is then fed into a Bipolar amplifier and ESD protection circuits before going into the chip. The 7-bit frequency divider control inputs are switchable between V_{DD} and ground. Because this is a test chip, the bias currents for the VCO, CML divider, control voltage generator, charge pump and output buffer are made adjustable through off-chip variable resistors. The output signal from the chip is connected to an Agilent E4443A 3Hz~6.7GHz spectrum analyzer via a 50Ω RF SMA connector on the board. Since the measured signal is the divided-by-2 output from the VCO, the output frequency is less than 4GHz, and hence can be handled by this equipment. A replica of the tuning voltage generation circuit is also measurable through on-chip analogue pads and on-board testing points.

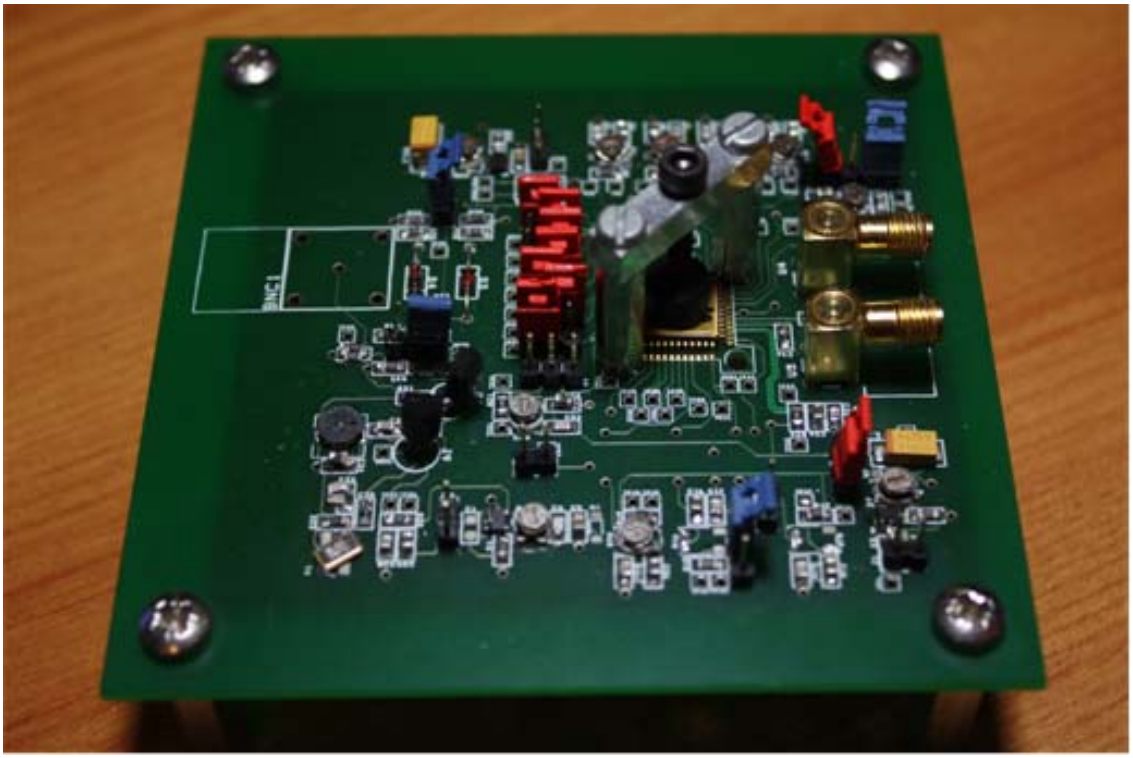


Figure 6-21 PLL testing board

The nominal power supply voltage is 1.2V. Simulation results show that each stage of the ring VCO draws 1.35mA current from the 1.2V power supply, the fastest CML frequency divider consumes 1.3mA current, and the currents are 0.83mA each for the following three CML dividers. The power consumptions of the tuning block, the charge-pump and the analogue to digital interface are not significant (less than 0.5mA all together). The measured total power consumption of the analogue blocks is 12mW, including the output buffer. The buffer is only used for testing and consumes 2.1mA current according to the simulation result. The total measured digital power consumption is 0.5mW, when the power supply of the digital block is set to 1.32V to

increase the speed of the digital blocks. The power consumption of the frequency synthesizer core is the sum of the measured power consumptions of analogue and digital blocks, subtracting the simulated output buffer power consumption, resulting in the total power consumption of 9.98mW.

6.6.1 Simulation Results

The VCO was simulated using a circuit netlist extracted post-layout. With the tuning frequency sweeping from 0V to 1.2V, the VCO's output frequency varies from about 7.35GHz to 4.1GHz, more than covering the required 3GHz band as shown in Figure 6-22. It can be observed that the relationship between control voltage and frequency is almost linear as expected. The effective frequency tuning happens for control voltages between 200mV and 1V as designed. Figure 6-23 and Figure 6-24 show the phase noise simulation results. The phase noise at 1MHz offset for a 4100MHz frequency is -91.84dBc/Hz, while increasing to -79.47dBc/Hz when frequency is 7095GHz. The phase noise versus frequency plots for 200kHz, 1MHz and 10MHz offsets are shown in Figure 6-25. Note that despite of the higher phase noise at higher frequency, the 200kHz offset close-in phase noise is less than -60dBc/Hz even for the highest frequency.

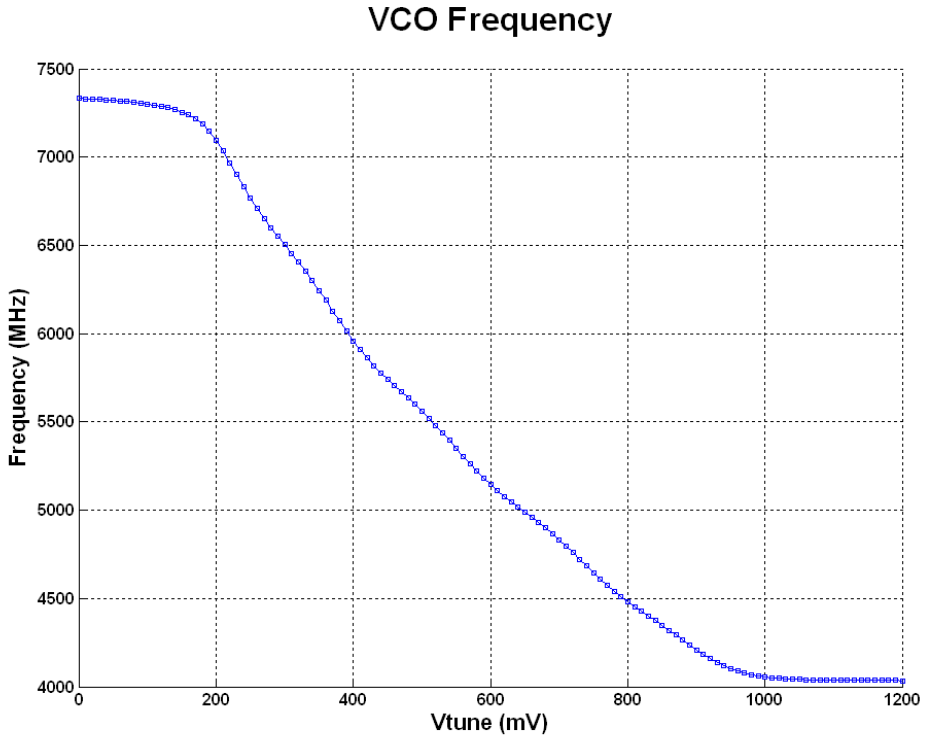


Figure 6-22 Ring VCO tuning range (post-layout simulation using spectreRF)

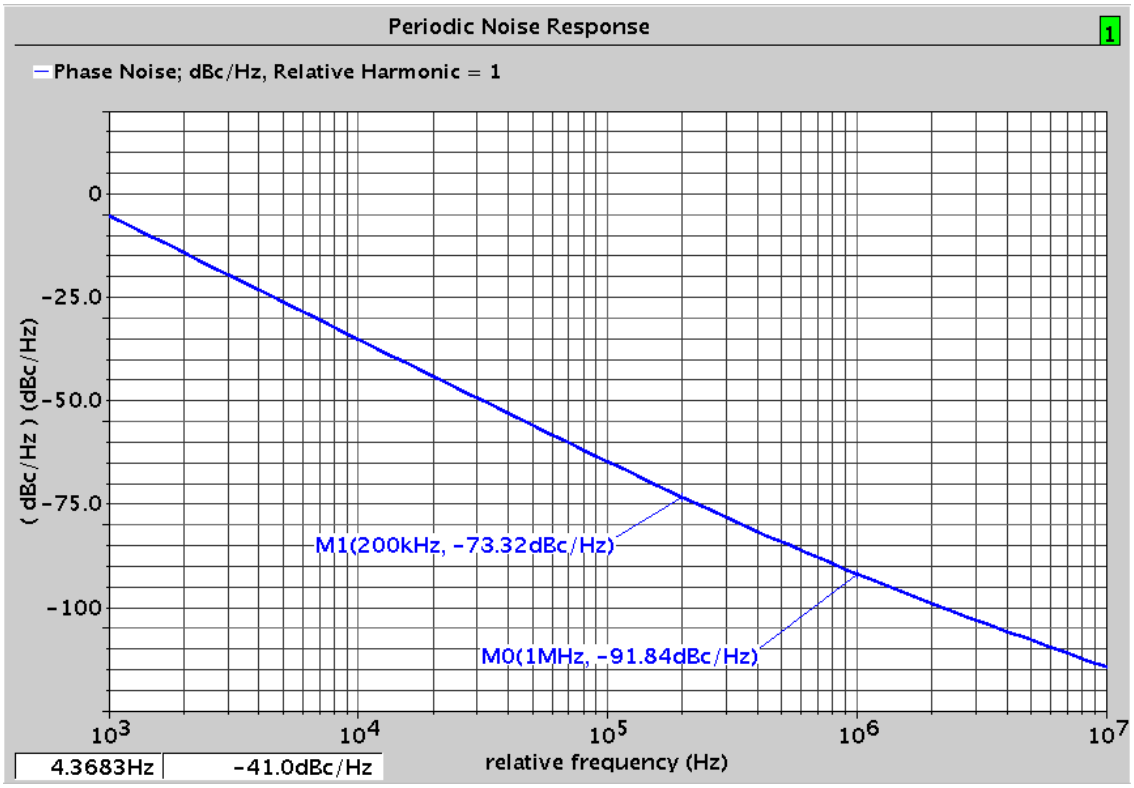


Figure 6-23 Ring VCO phase noise at 4100MHz (post-layout simulation using spectreRF)

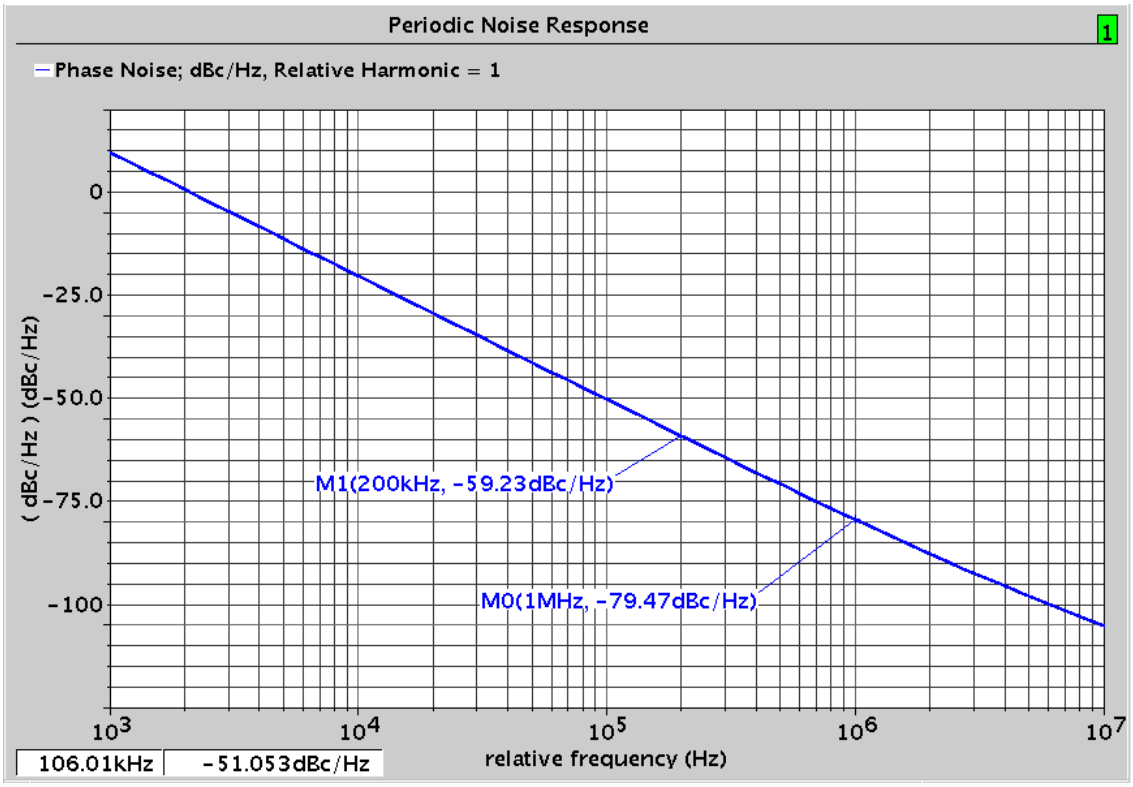


Figure 6-24 Ring VCO phase noise at 7095MHz (post-layout simulation using spectreRF)

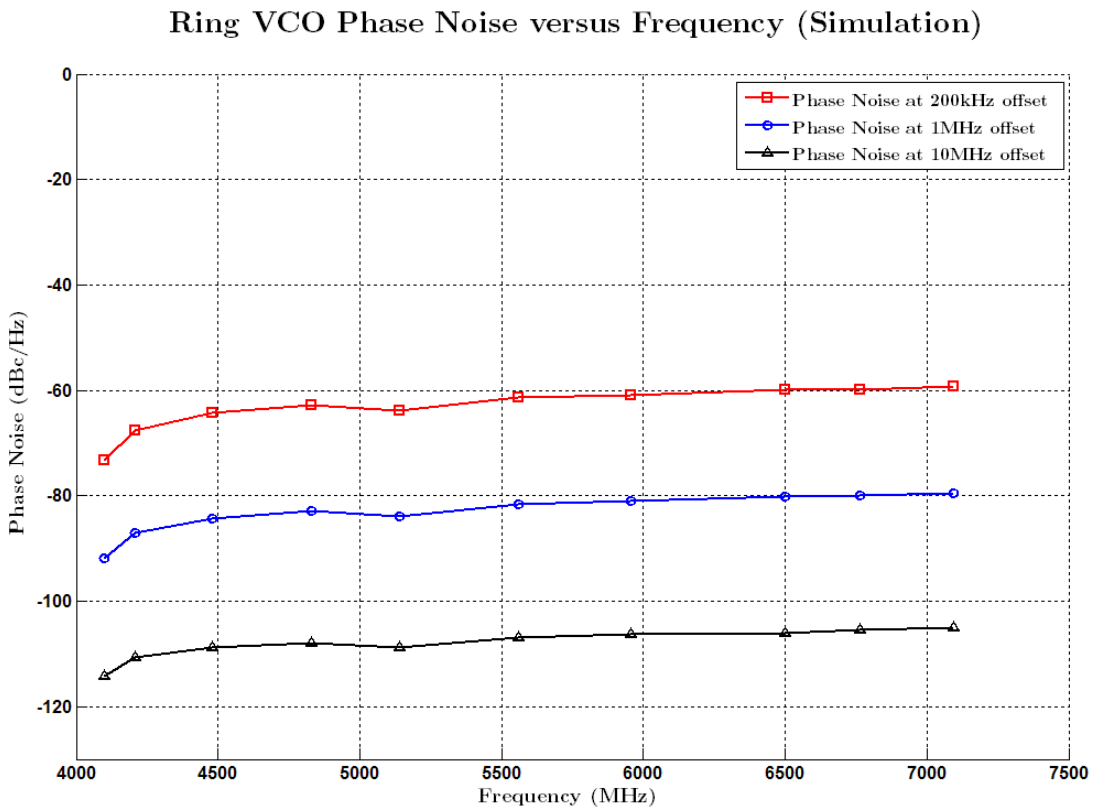


Figure 6-25 Ring VCO phase noise versus frequency (post-layout simulation using spectreRF)

6.6.2 Measurement Results

The comparison between simulation and measurement results of the tuning voltage generation circuit are compared in Figure 6-26. It can be observed that the tuning voltages are decreasing one by one within the effective MOS varactor range from 500mV to 100mV. The measurement results match the simulation results quite well, although some small offsets are observed as might be expected from device tolerances. However, these offsets are tolerable by the frequency synthesizer.

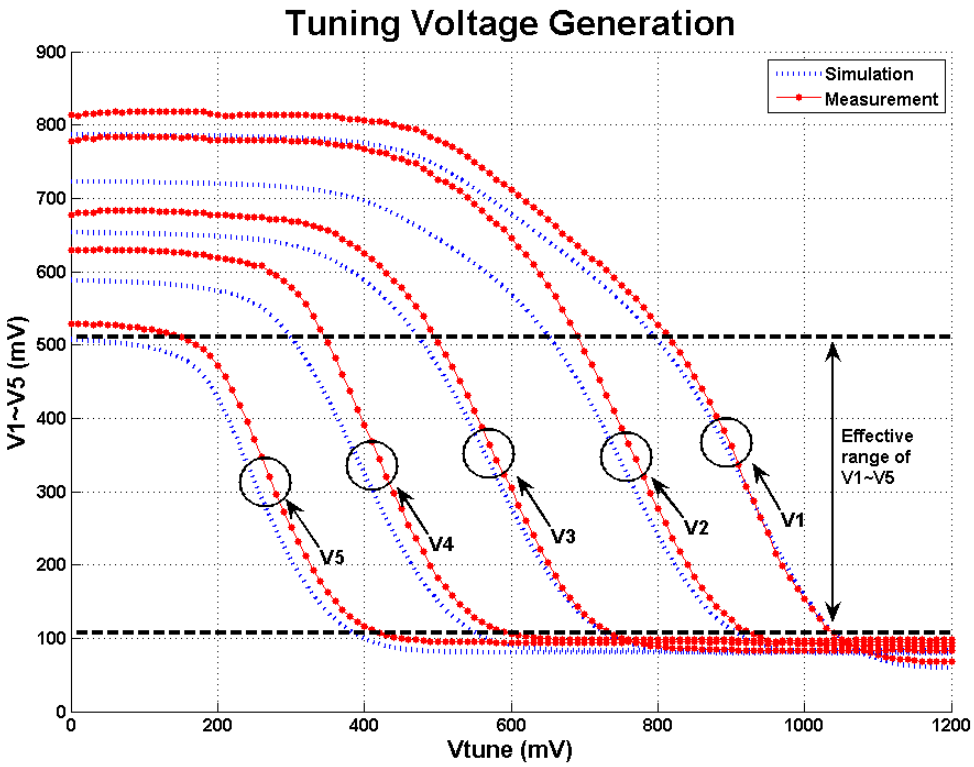


Figure 6-26 Measurement of tuning voltage generation

The measured phase noise of the PLL is shown below. During the design of this PLL, the loop bandwidth is made tunable for research purpose. According to Equation (6-18), different system parameters are related to the loop bandwidth. However, the VCO gain and the division ratio are designed as specifications, and the loop filter components are also fixed once the PLL is implemented on chip. Therefore, the only parameter left that can be used to adjust the loop bandwidth is the charge pump current. Hence, on this test chip the charge pump is made tunable by using a variable resistor to change the current of the current mirror in the charge pump (the left branch in Figure 6-19). According to Equation (6-18), the relationship between the charge pump and the loop bandwidth is plotted in Figure 6-27. It can be seen that more current is consumed to achieve higher loop bandwidth.

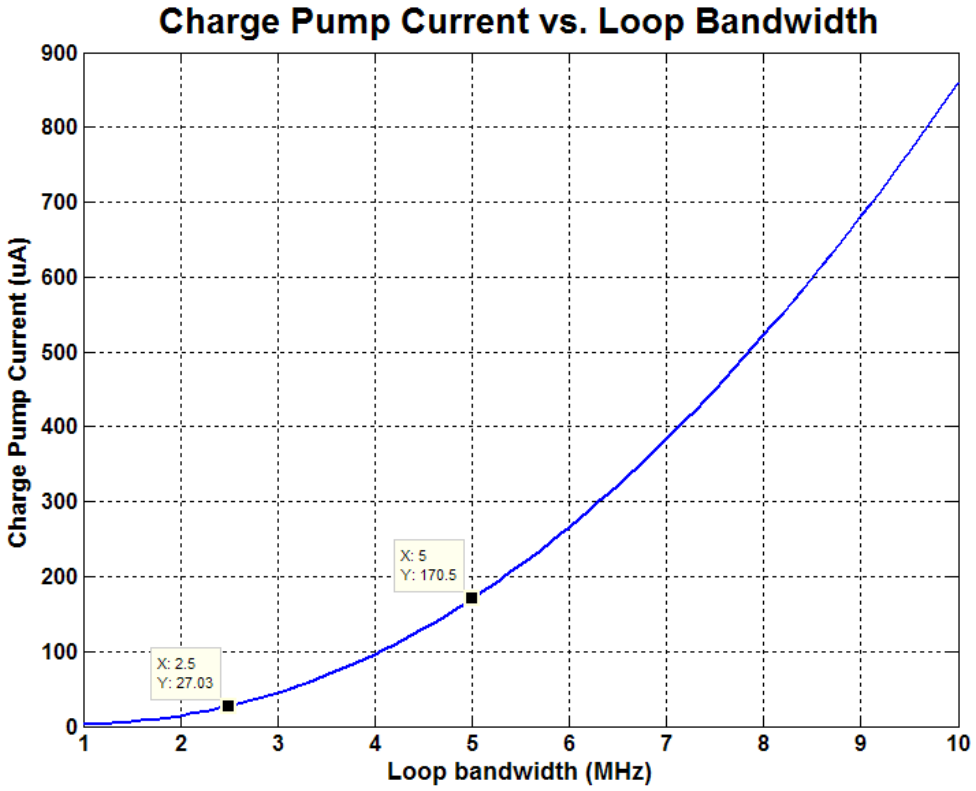


Figure 6-27 Charge pump current vs. loop bandwidth

Figure 6-28 and Figure 6-29 show the different phase noise results with the variation of the loop bandwidth. For a 2.5GHz divide-by-2 output, when the loop bandwidth is as narrow as 100kHz, the phase noise at 100kHz offset is about -65dBc/Hz, while the reference spur at 25MHz offset is suppressed to less than -94dBc/Hz. If the bandwidth is increased to 5MHz, at most offset frequencies within 5MHz the phase noise is lowered to -95dBc/Hz, while the reference spur increases to about -88dBc/Hz. In this case, the loop bandwidth is not far less than the reference voltage, which is often desirable as discussed in section 6.1. However, the measurement results show that the reference spur level is still lower than the specification. Hence this configuration (5MHz loop bandwidth) is an acceptable compromise solution. Note that a ripple in the phase noise appears around the loop bandwidth, but still lower than -90dBc/Hz.

This comparison illustrates the practical trade-off between in-band phase noise and out-of-band spur level. In most receivers, a lower in-band phase noise is more important for the demodulation. In the proposed spectrum monitor, however, the phase noise within 100MHz offset is of concern. According to chapter four, the -80dBc/Hz specification must be satisfied over the entire 100MHz range. Therefore according to the measurement results, the 5MHz loop bandwidth is adopted for this frequency

synthesizer. As will be shown later, the target of -80dBc/Hz from 200KHz out to 100MHz offset frequency has been met with the loop bandwidth of 5MHz for all the synthesised frequencies.

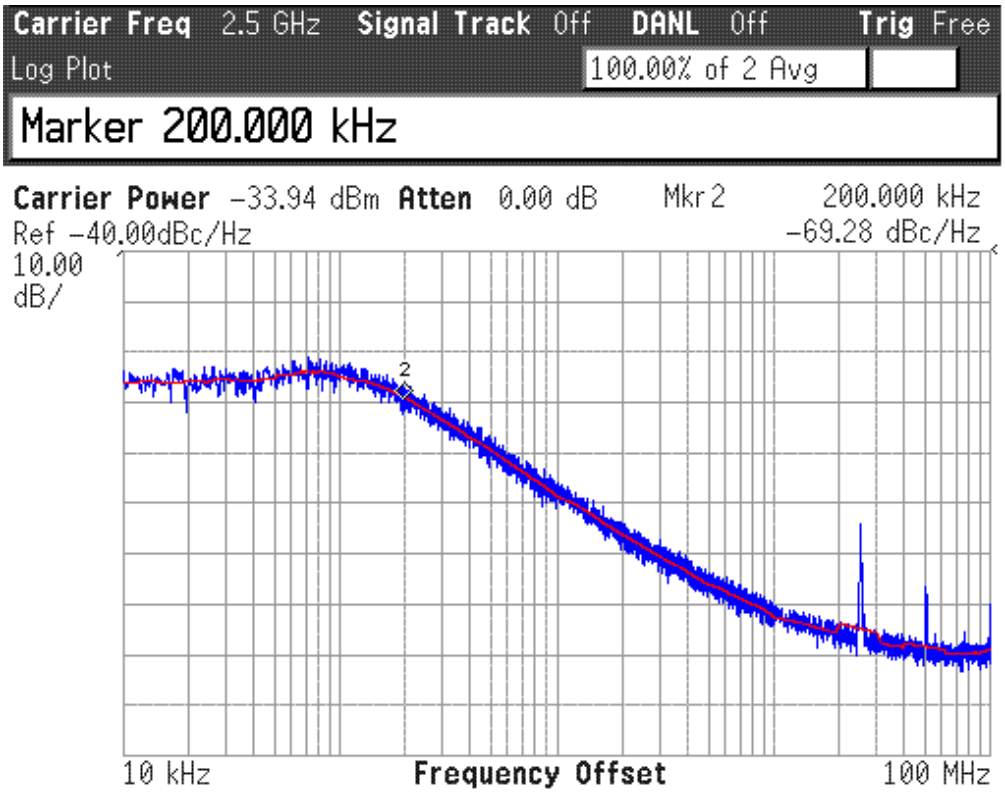


Figure 6-28 PLL Measurement: 5GHz phase noise (loop bandwidth=200kHz)

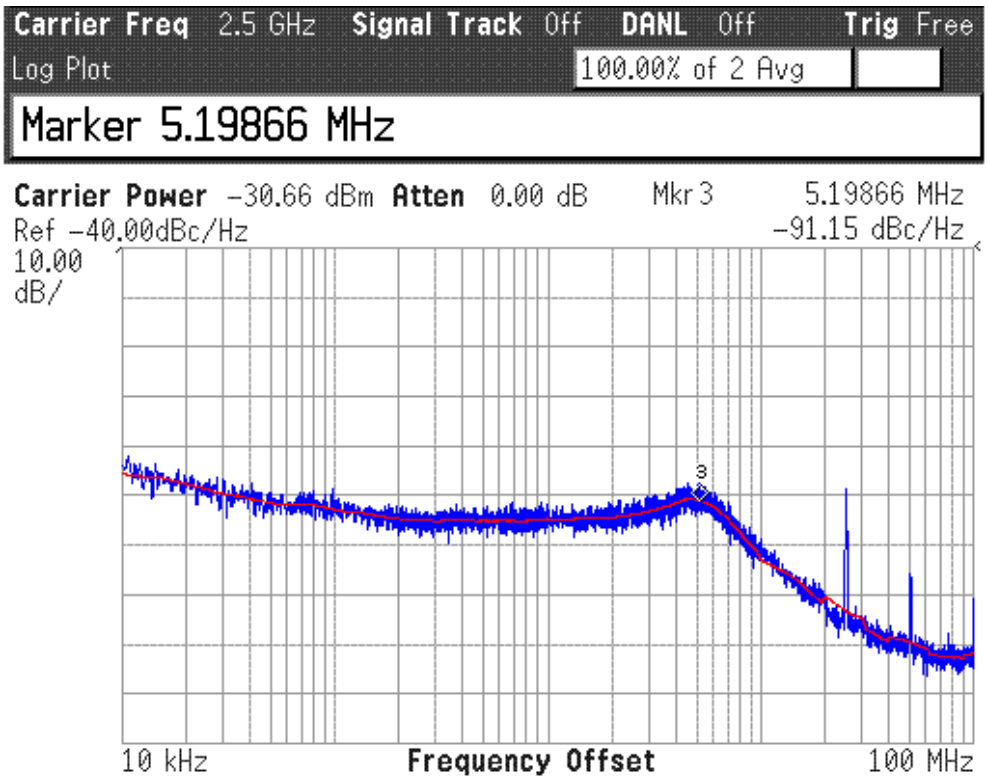


Figure 6-29 PLL Measurement: 5GHz phase noise (loop bandwidth=5MHz)

From the measurement result in Figure 6-30, the frequency synthesizer can be seen to lock the VCO frequency up to 7.3GHz, where the 3.65GHz divide-by-2 signal is observed. With the loop bandwidth set to 5MHz, the in-band phase noise is generally lower than -86dBc/Hz, and the out of band reference spur is about -89dBc/Hz at 25MHz offset. The phase noise after the divide-by-2 circuit decreases by 6dB. Therefore, a 6dB higher phase noise is expected at the VCO output node. The scaled phase noise versus VCO output frequency is given in Figure 6-31 to Figure 6-35, with offset frequencies of 200kHz, 1MHz, and loop bandwidths of 5MHz and 10MHz, as well as at the reference spur frequency of 25MHz. It can be observed that the phase noise between 200kHz and 1MHz offset is generally lower than -85dBc/Hz, while the reference spur and phase noise at the loop bandwidth are between -80dBc/Hz and -85dBc/Hz, satisfying the system level specification of the proposed spectrum monitor. The phase noise measured at other frequencies can be found in Appendix C-2.

Note that in the Figure 6-30, the close-in phase noise with less than about 100kHz offset is dominated by flicker noise and is higher than the specification. For the highest locked frequency (7.3GHz), it is increased to about -80dBc/Hz at 20kHz offset. This can be improved by using a 3rd-order loop filter to form a 4th-order frequency synthesizer in the future work, so that the close-in phase noise is flat.

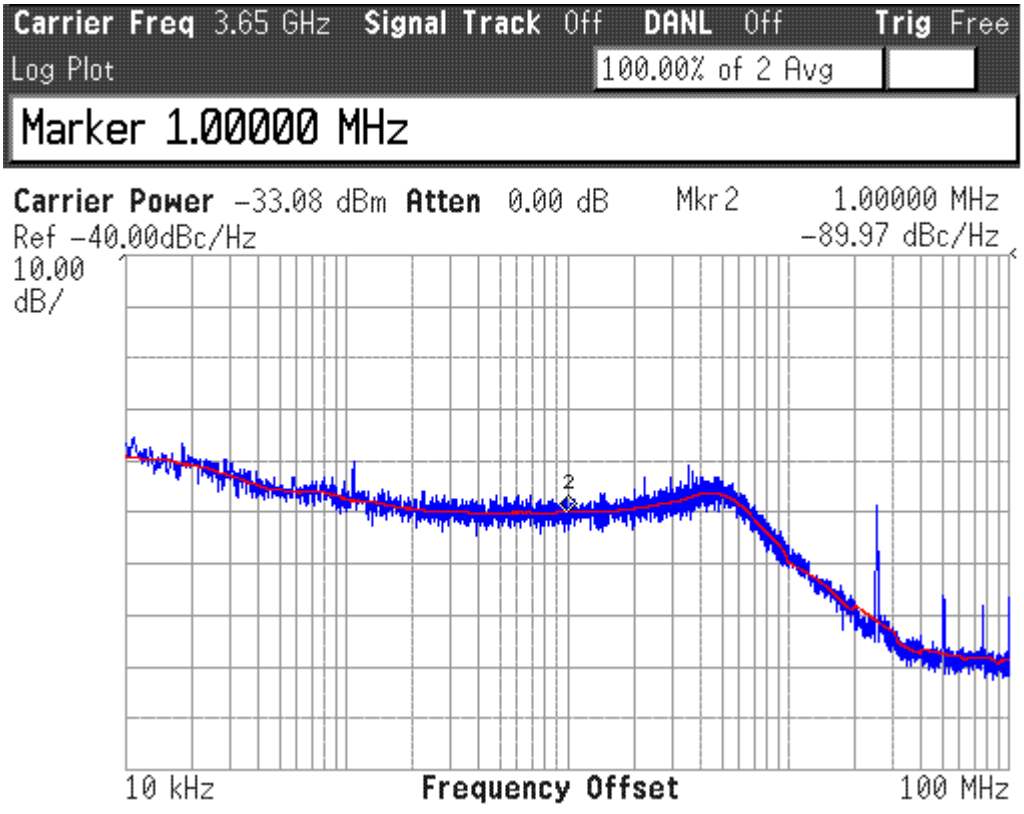


Figure 6-30 PLL Measurement: 7.3GHz phase noise (loop bandwidth=5MHz)

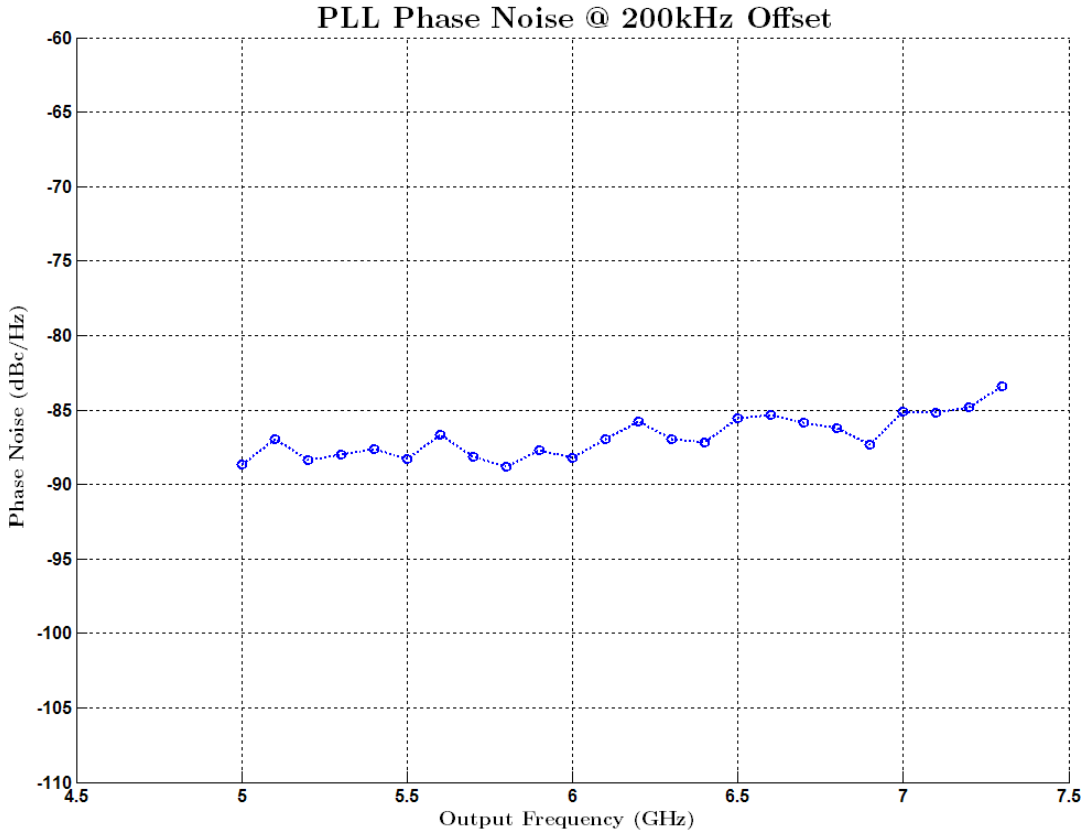


Figure 6-31 PLL Measurement: phase noise versus frequency @ 200kHz offset

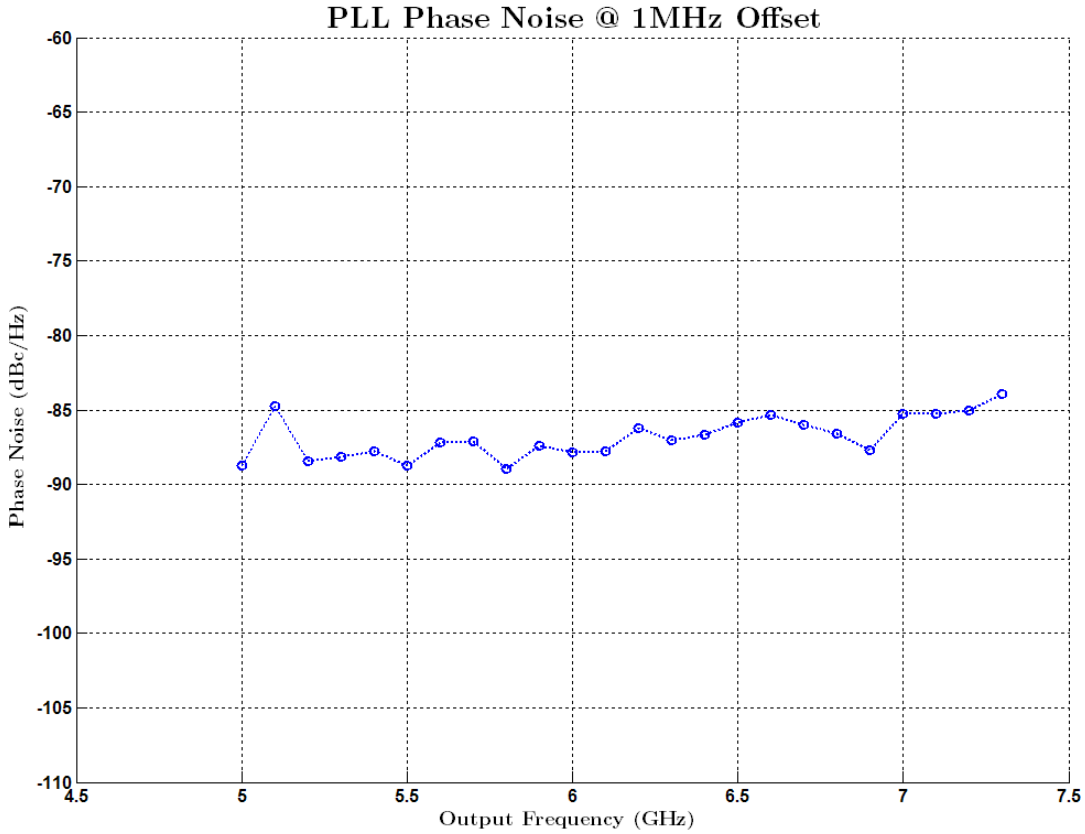


Figure 6-32 PLL Measurement: phase noise versus frequency @ 1MHz offset

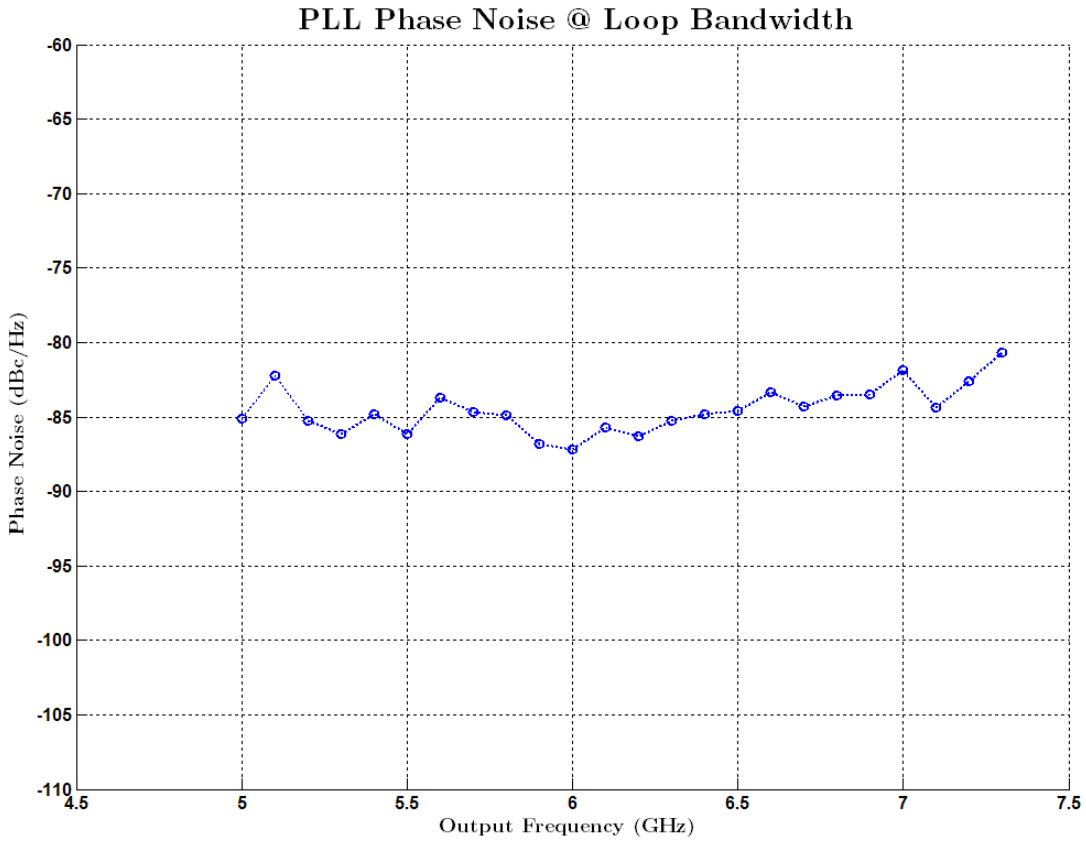


Figure 6-33 PLL Measurement: phase noise versus frequency @ loop bandwidth of 5MHz

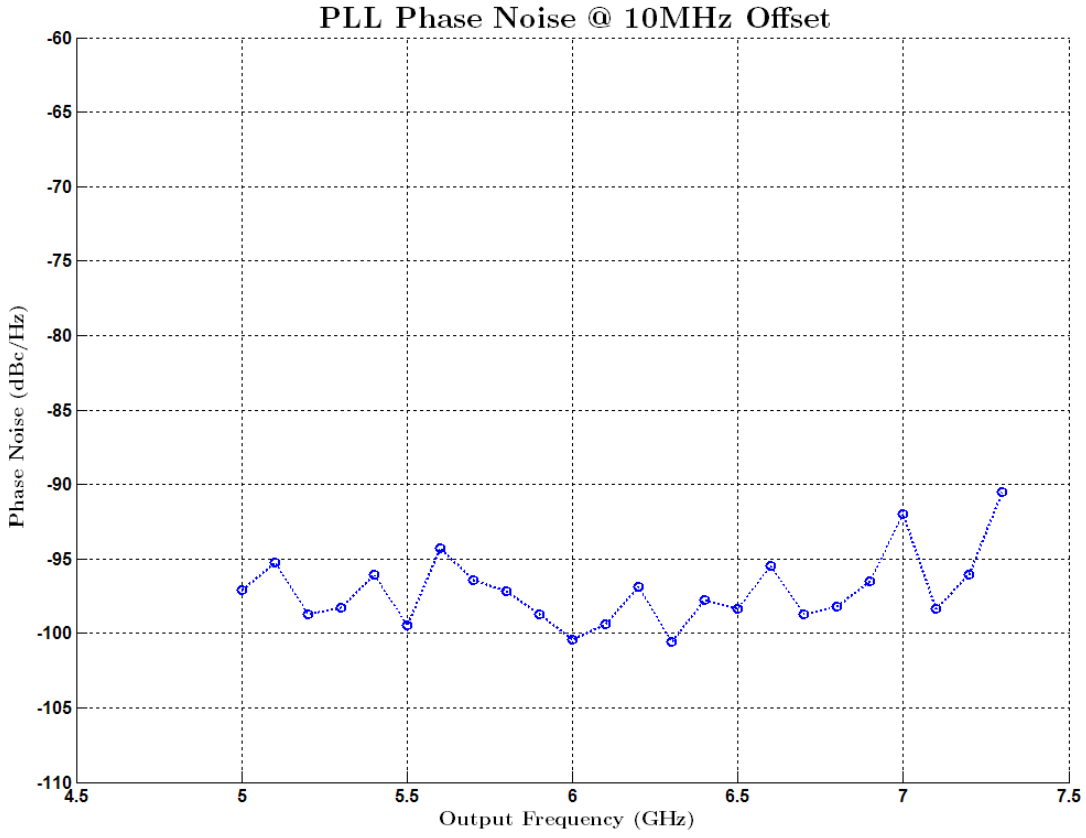


Figure 6-34 PLL Measurement: phase noise versus frequency @ 10MHz offset

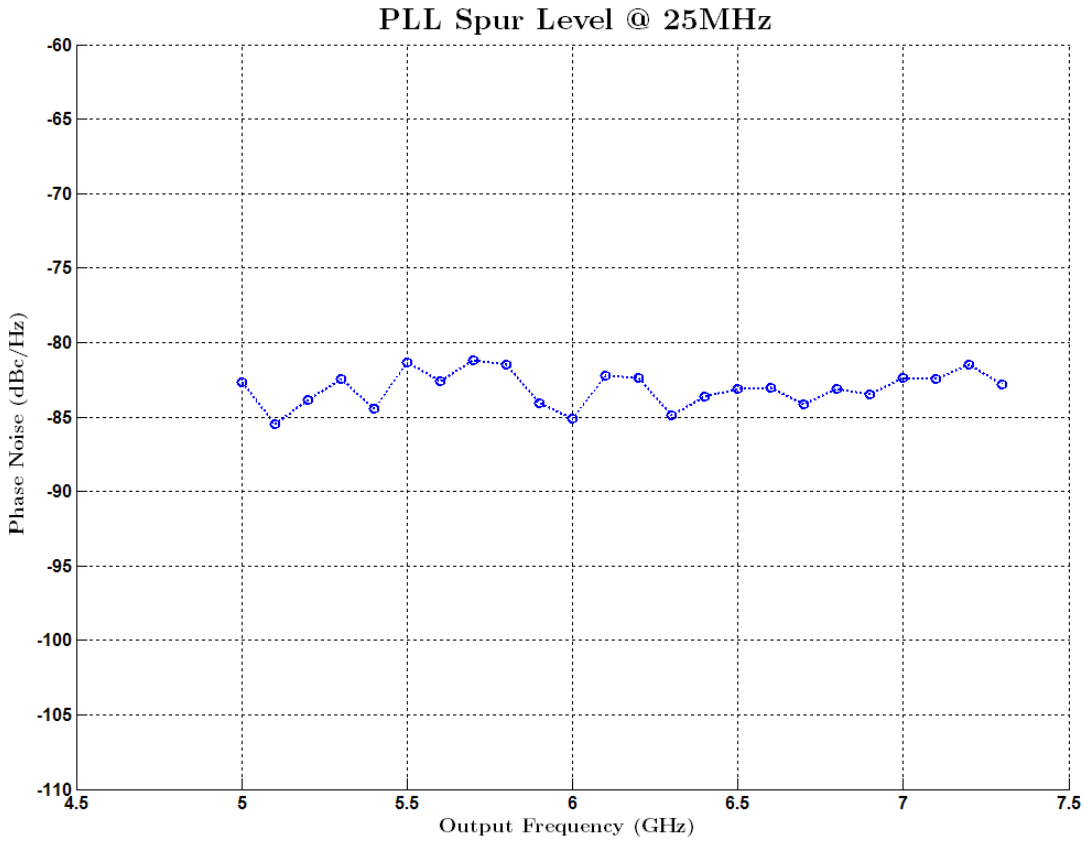


Figure 6-35 PLL Measurement: spur lever versus frequency @ 25MHz offset

6.7 Summary

In this chapter, a wide tuning range frequency synthesizer is designed, fabricated and measured. The system level design and simulation are described first, followed by the design of the key blocks of the system. Two blocks in particular using novel architectures are developed for this project, which are the ring oscillator with a staggered control voltage tuning scheme, and the fast reset-counter integer-N frequency divider. The measured staggered tuning voltage generation matches the simulation result quite well. The final phase noise and spur level are measured to be less than the required -80dBc/Hz within a very wide frequency offset, and over several gigahertz tuning frequency, satisfying the system specification for the spectrum monitor.

To compare this work with others', it is desirable to calculate the figure of merit as discussed in chapter three. However, the FoM of PLL is not investigated in chapter three due to too many variations in the design techniques. In spite of this, it is worthy calculating the FoM of the VCO according to the simulation results. With the phase noise of -91.84dBc/Hz at 1MHz offset when the VCO is tuned to 4100MHz (Figure 6-23), the VCO core consumes 4.86mW (1.2V×3×1.35mA, see section 6.6), the FoM

of the VCO alone can be obtained as 187dB, which is better than the average FoM of the published VCOs through the years.

Chapter 7 Conclusion

In this PhD project, the design of a spectrum monitor receiver for a future cognitive radio system is investigated. This spectrum monitor needs to be able to detect spectrum occupancy within popular communication bands, and be fast and accurate enough for further applications, as well as having low power and low cost that are essential for mobile devices.

To this end, new techniques are investigated based on the concept of the Figure of Merit for a particular function to assist in the system architecture design, and where likely improvements in technology are included to allow the design decision to be relevant to future developments. This theoretical work is followed by a complementary experimental study into the design, fabrication and measurement of two of the critical blocks that are essential for the spectrum monitor architecture.

7.1 Summary of the Chapters

Chapter 2 reviews the concept of cognitive radio and develops some of the requirements. The history of cognitive radio is first introduced, pointing out that among the several key techniques involved, an accurate, fast, low power and low cost spectrum monitoring solution is the first step to realize successful cognitive radio application. Because of the wideband requirement of the potential spectrum monitor receiver, some existing wideband receiver architectures (TV and UWB) are explored. Their advantages and disadvantages are analyzed as necessary references for the design of the spectrum monitor receiver.

In chapter 3, a thorough investigation is made into the concept of the Figure of Merit (FoM) for common circuit blocks used in integrated radio receivers. Of particular interest are the performance expectations for a particular block to be designed at the

current state of the art, and also, drawing on the inspiration of the well-known Moore's Law used in the digital world, predicting what state of the art performance can be expected for future designs as silicon technology advances and circuit innovation continues. The FoM are defined for various circuit functions, specifically narrowband and wideband LNAs, current steering mixers, LC and ring type VCOs, injection locked and CML frequency dividers, baseband LPF and VGA functions, as well as Nyquist and Sigma-Delta ADCs. The FoM origins are explained and a survey of the recent literature is undertaken to collect a large number of data points for the FoM of recently published designs. These data are used to identify the improvement trends in the FoM for each type of function, driven technology and design innovation.

In general, the trend in the published data for the FoM of RF stages, including the LNA and mixer, indicates that there is likely to be a 3dB improvement in FoM over about 30~40 months. This implies that designs done that far into the future are likely to have around 50% power consumption reduction for the same specifications. The FoM of LC VCOs mainly improves in line with the improvements in the Q factor of on-chip inductors, and the trend indicates a 3dB improvement over a period of about 60 months. Conversely, the trends shown for the published FoM of ring oscillators do not show significant improvement over time. Injection-locked and static CML dividers mainly benefit from higher f_T of silicon technology that is a direct consequence of the technology scaling in CMOS as driven by digital applications. A 3dB improvement in FoM for CML dividers and injection-locked dividers is expected over about 45 months and 75 months respectively, according to the trendlines. The baseband blocks including the channel selection LPF and VGA show an improvement trend of 3dB over about 25 months, which is much faster than the RF blocks. The trends for ADCs are somewhat dependent on architectures. The FoM trend of published Nyquist ADCs shows an improvement of around 3dB in about 22 months, which is quite similar to the rate of increase in the density of digital circuits predicted by Moore's law. By contrast, the trend in FoM in the published Sigma-Delta ADCs shows a rate of improvement very similar to the RF/analogue blocks, achieving 3dB in about 33 months.

Using these FoM data and the derived trends, a simple and effective strategy is developed to assist the designer in the system level architecture for designs that will be undertaken immediately, and also for future design projects. This latter factor is to help in achieving design choices that will make the finished receiver competitive with the state of the art at the time it is realised, and not be based on what will become historic

data. This strategy allows the estimation of the power consumption of a potential design of the circuit blocks given certain specifications, using the predicted FoM trends for the next few years. Taking the block specifications required and the expected FoM values, the system architecture can be investigated to achieve the optimum for the likely future FoM values.

In chapter 4, possible architectures for an integrated CMOS spectrum monitor are explored. The band of interest is selected to be 2GHz~5GHz based on a realistic use model concept, with the 200kHz frequency resolution to be provided in an associated DSP block. (This latter function is not dealt with in detail as it is outside the scope of the project.) Various modes of operation for scanning the band of interest are considered. Analysis shows that scanning multiple sub-bands (e.g. 100MHz) is the most feasible and effective way for current technologies, and a two stage scanning strategy is proposed. The system level specifications are analyzed using simplified signal modelling, and then the potential receiver architectures are discussed at length. For the 2GHz~5GHz input range, the dual-down conversion architecture is shown to be the most compact and effective candidate.

The system level design is then completed for the case where the sub-band resolution is set to 100MHz, together with the expected power consumption derived from the FoM studies. It is proposed that an acceptable spectrum monitor receiver can be developed in about 5 years time for mobile devices, according to the FoM based power consumption prediction discussed earlier. For the integrated parts of the receiver, the challenging design tasks include high gain/high linearity mixers, the integrated narrowband on-chip bandpass filter and the wide-tuning range frequency synthesizer.

In chapter 5 the design of the narrowband on chip filters required by the receiver is described. The topology of the on-chip bandpass filter is selected as a direct series-coupled-resonator architecture, instead of the conventional bandpass filter synthesis method. This topology is shown to have almost the same response as a conventional filter near the centre frequency but has much sharper stopband attenuation below the lower frequency edge. The 10GHz and 1.75GHz filters are designed for both the integrated up-down-down conversion and dual-down conversion spectrum monitor architectures. A novel application of the delta-star transformation is applied to the filter synthesis process for the higher frequency designs to produce a topology with component values that are realisable on silicon.

Both of these filters are designed in differential form, and are implemented on a standard 130nm CMOS technology using integrated planar inductors. Measurements are made using on-chip GSGSG RF probes pads. The 10GHz filter achieves 1GHz bandwidth with 15.5dB insertion loss, while the bandwidth of the 1.75GHz filter is as narrow as 210MHz, with 8.5dB insertion loss. The high insertion loss is mainly due to the limited quality factor of the fabricated inductors, although these losses are expected to be reduced in future technologies and designs. In spite of the high insertion loss, these filters are expected to be suitable for the spectrum monitor application, according to the specifications derived in chapter four.

Chapter 6 describes the local oscillator subsystem required for the spectrum monitor which is designed and implemented in a standard 130nm CMOS technology. Because of the moderate phase noise specifications and the very wide tuning range requirements for the spectrum monitor, a ring oscillator is selected as the core of the frequency synthesizer. To achieve the wide tuning range with a wide loop bandwidth an integer-N synthesiser architecture is used, with a mixture of CML and standard CMOS digital frequency dividers. The ring oscillator consists of three cascaded resistively load differential inverters, and is tuned by a MOS varactor array. To achieve a lower sensitivity to the loop filter control voltage and also better linearity in the tuning loop, internal staggered tuning voltages are generated for each varactor using a novel level translation circuit.

Using a discrete component 25MHz XTAL reference oscillator on the test board, the measurements of the frequency synthesizer show that it achieves a phase noise of lower than -80dBc/Hz within the entire 100MHz frequency offset required for the spectrum monitor, and over the frequency band of interest up to the highest frequency of 7.3GHz. The power consumption of the synthesiser without the test buffer is as low as 9.98mW.

7.2 Comments

The Figure of Merit strategy in this project involves a large amount of data collection, calculation and analysis for the main blocks in the entire receiver front-end. An effective and efficient systematic approach has been developed based on these research results to provide a quite confident prediction for the power consumption of the receiver into the future. This is believed to be able to provide very useful

information as reference and guidance in the initial system level design stage for product development.

Frequency synthesizer design is one of the most important blocks in the potential spectrum monitor. Novel tuning methods are involved for the VCO and frequency dividers. The measurement results show that this solution is suitable for the proposed spectrum monitor architecture.

The 1.75GHz bandpass filter achieves quite good performance and is believed to be competent for the spectrum monitor using the dual-down conversion architecture.

In spite of the above achievements, there are also some shortcomings in this project. The effects of GSGSG pads was not de-embedded, leading to lower centre frequencies compared to simulation results. In the design of the 10GHz bandpass filter, the degradation in Q factor of the capacitors at higher frequencies was not fully considered during design and simulation. As a result, the measured insertion loss and stop band attenuation performances do not fully match the simulation results very well. This could be improved by more accurate modelling of all of the components.

7.3 Future work

The future research would be first focused on further exploration of the spectrum monitor architecture for cognitive radio application. More realistic spectrum occupancy could be obtained by further research so that more meaningful specifications can be derived accordingly, including the frond-end's gain, noise, linearity and ADC's resolution and bandwidth, etc.

It is worth keeping on collecting the figure of merits of the investigated receiver functions in the future, especially for CMOS technologies with feature sizes of less than 100nm. The accuracy of the trendlines is expected to be improved by more data samples. Clear roadmaps of the technology improvement can be revealed gradually, which is expected to be providing very valuable information in many relevant areas.

As for the circuit design, the two key blocks could be improved to achieve better performance. The frequency synthesizer can be designed to have much wider tuning range, e.g. from 100MHz to 6GHz, as well as lower phase noise. More accurate inductor and capacitor modelling could be done to improve the filter's insertion loss and selectivity. The effects of the pads should also be fully de-embedded. The component value variations due to corner variations could also be taken into account for potential commercial realizations in the future.

Furthermore, the other receiver blocks, including wideband LNA, high gain and high linearity mixers, baseband LPF and high resolution ADC, could be designed to complete the entire spectrum monitor receiver.

It will be the key step in the evolution of the cognitive radio when this spectrum monitor receiver is finally designed and integrated using standard CMOS technology. Because of the foreseeable improvement of the DSP ability, it is reasonable to expect that suitable algorithms and solutions will be developed for the entire cognitive radio application. These low power, low cost solutions could be commercialized and this will lead to major changes in the spectrum resource usage in the wireless communication industry.

Appendix A

This appendix provides detailed explanations and derivations of the equations related to the FoM analysis in chapter three.

A-1 Linear fitting technique

The linear fitting technique, sometimes called linear regression, is a statistical approach to help modelling the intrinsic relationship between variables and could also be used to make predictions by extrapolating the fitted trendline. This method is used substantially in the FoM analysis. The following equations provide the detailed calculations. For a given group of data, assume that the value of each data, y_i, is a function of a variable, x_i, and the also assume that x_i and y_i could theoretically be related by linear relationship, then a fitted linear function $y=f(x)$ can be obtained from the data set to reflect this relationship.

$$y = f(s) = bx + a, \tag{A-1}$$

$$where\ b = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \tag{A-2}$$

$$and\ a = \bar{y} - b\bar{x}. \tag{A-3}$$

The values $\bar{x}$ and $\bar{y}$ are the mean values of the variables x and y, and the parameter n is the total number of data points.

Note this linear fitting is only valid if the variable x and y are expected to have linear relationships, which is true to a large extent for FoM analysis as discussed in chapter three.

A-2 Derivation of voltage gain of narrowband LNA

Assume that the CMOS device consists only of an input capacitance C_{gs} and a transconductance g_m, then the input impedance looking into the matching network of a narrowband LNA from the inductor L_g can be obtained by KCL equations and is given by

$$\begin{aligned} Z_{in} &= s(L_g + L_s) + \frac{1}{sC_{gs}} + \frac{g_m}{C_{gs}} \cdot L_s \\ &= s(L_g + L_s) + \frac{1}{sC_{gs}} + \omega_T \cdot L_s\,. \end{aligned} \tag{A-4}$$

At the LNA operating frequency, power matching is usually required. Hence the inductance and capacitance in Z_{in} should be cancelled, leaving the real part of $\omega_T L_s$ equal to the signal source resistance, R_s, which is usually 50Ω.

The overall effective transconductance, G_m (the ratio of the drain current to the signal source voltage, V_{rf}), of the LNA with a matching network can be derived to be Q_{match} times the MOSFET's transconductance, g_m, and is given by

$$G_m = g_m Q_{match} = g_m \frac{1}{\omega_0 C_{gs}(R_s + \omega_T L_s)}. \tag{A-5}$$

With the matching condition $\omega_T L_s = R_s$, and because $\omega_T = g_m/C_{gs}$, the effective transconductance can be determined as

$$G_m = \omega_T / 2\omega_0 R_s . \tag{A-6}$$

A-3 Derivation of IIP3 of the narrowband LNA

Usually, the transistor's non-linearity is a dominant factor. For the sake of argument, therefore, a single stage, resistive load (R_L), common-source amplifier is chosen to analyse the non-linearity.

The drain current of a MOSFET can be expressed in the form of a power series as:

$$I_D = g_0 + g_1 V_{gs} + g_2 {V_{gs}}^2 + g_3 {V_{gs}}^3 + \cdots, \tag{A-7}$$

where g_0 is the DC component, g_1 is the transconductance of the transistor, g_2 and g_3 are the non-linear coefficient of the second, third harmonics, V_{gs} is the small signal input voltage.

Therefore the output voltage is:

$$\begin{aligned} V_{out} &= I_D R_L \\ &= g_0 R_L + g_1 R_L V_{gs} + g_2 R_L {V_{gs}}^2 + g_3 R_L {V_{gs}}^3 + \cdots. \end{aligned} \tag{A-8}$$

The 3rd order input-referred intercept point is given as:

$$V_{IIP3} = \sqrt{\frac{4}{3}\left|\frac{g_1}{g_3}\right|}. \tag{A-9}$$

The drain current in an ideal long-channel MOSFET amplifier biased at the onset of the strong inversion region consists only of a second order harmonic, so the IIP3 is theoretically infinite. However, short-channel effects must be taken into account in the drain current equation in modern technology. Neglecting the channel length modulation, the drain current is:

$$I_D = \frac{1}{2}\mu_0 C_{ox}\frac{W}{L}\frac{\left(V_{GS}+V_{gs}-V_{TH}\right)^2}{1+\left(\frac{\mu_0}{2v_{sat}L}+\theta\right)\left(V_{GS}+V_{gs}-V_{TH}\right)}. \tag{A-10}$$

The term $\mu_0/2v_{sat}L$ represents the velocity saturation effect. Parameter θ is the vertical field effect fitting parameter with the units of V^{-1}, and can be roughly estimated as $2.3/t_{ox}$, where the thickness of the gate oxide, t_{ox}, has the units of *nm*.

Taking the example of a TSMC 0.35µm technology from MOSIS, the device parameters are: $\mu_0=358\times10^{-4}m^2/V\cdot s$, $v_{sat}=1.37\times10^5 m/s$ and $t_{ox}=7.8nm$. When the transistor is biased around the onset of the strong inversion region, which is the common situation in RF amplifier design, the overdrive voltage, $V_{ov}=(V_{GS}-V_{TH})$, is about 200mV. Then the item $(\mu_0 2v_{sat}L+\theta)\cdot(V_{GS}-V_{TH})$ equals 0.14, which is much less than unity. In practice, this condition can be satisfied in most situations.

Letting $\rho=\mu_0 2v_{sat}L+\theta$, and $K=(1/2)\mu_0 2C_{ox}(W/L)R_L$, the output voltage is then given by:

$$\begin{aligned} V_{out} = &\left[KV_{ov}^2-\frac{K\rho}{\rho+1}V_{ov}^3\right]+\left[2KV_{ov}-\frac{3K\rho}{\rho+1}V_{ov}^2\right]V_{gs} \\ &+\left[K-\frac{3K\rho}{\rho+1}V_{ov}\right]V_{gs}^2+\left[-\frac{K\rho}{\rho+1}\right]V_{gs}^3. \end{aligned} \tag{A-11}$$

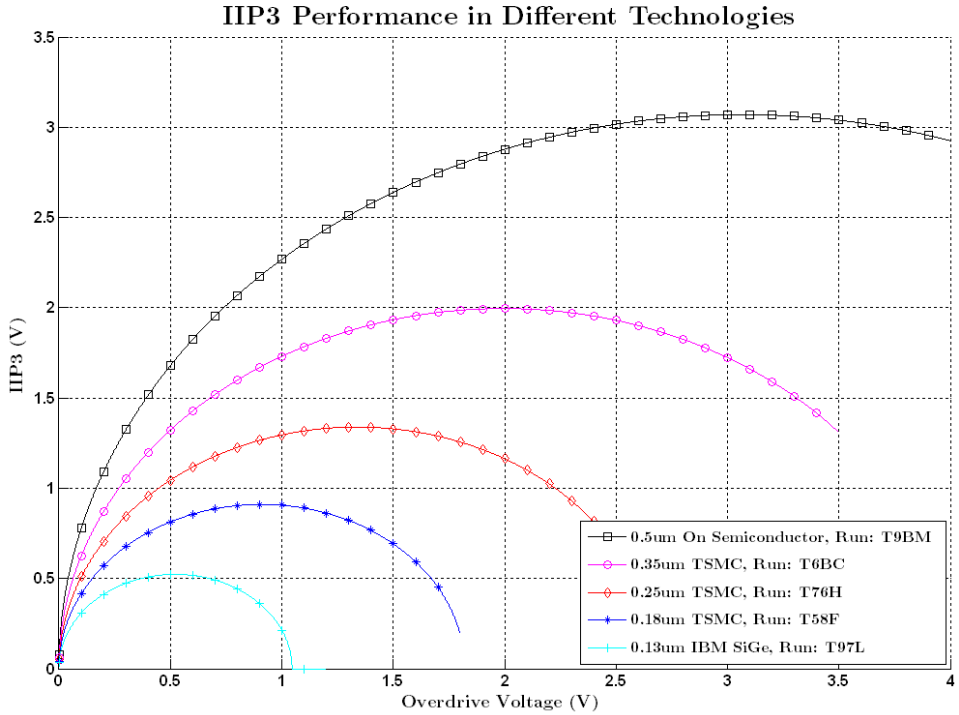
Comparing Equation (A-7) and (A-11), the coefficients g_1 and g_3 can be obtained and by substituting them into Equation (A-9), the 3rd order input referred intercept point is derived as:

$$V_{IIP3} = \sqrt{\frac{8}{3}\frac{V_{ov}}{\frac{\mu_0}{2v_{sat}L}+\theta}-V_{ov}^2}. \tag{A-12}$$

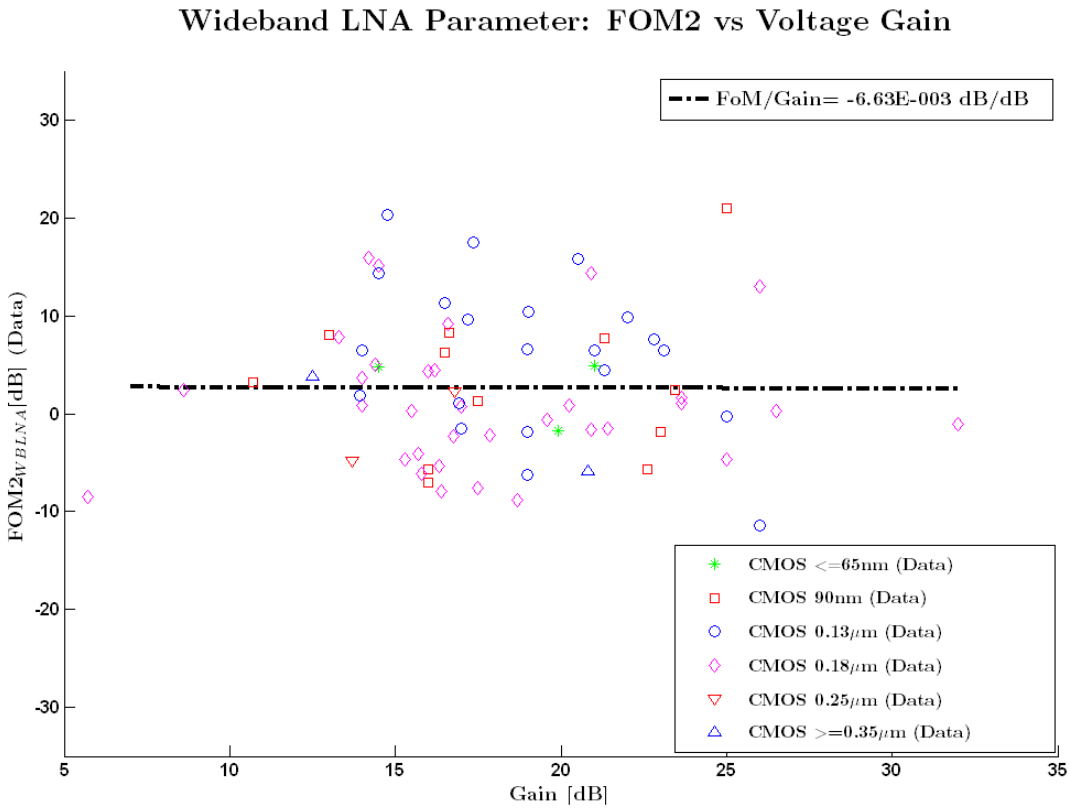
For a certain technology, Equation (A-12) shows that linearity in power is a quadratic function of the overdrive voltage. A diagram is plotted in the figure below, by taking examples of 0.5µm, 0.35µm, 0.25µm, 0.18µm and 0.13µm technologies from the MOSIS website. The parameters are taken from the Wafer Electrical Test Data and SPICE Model Parameters sections.

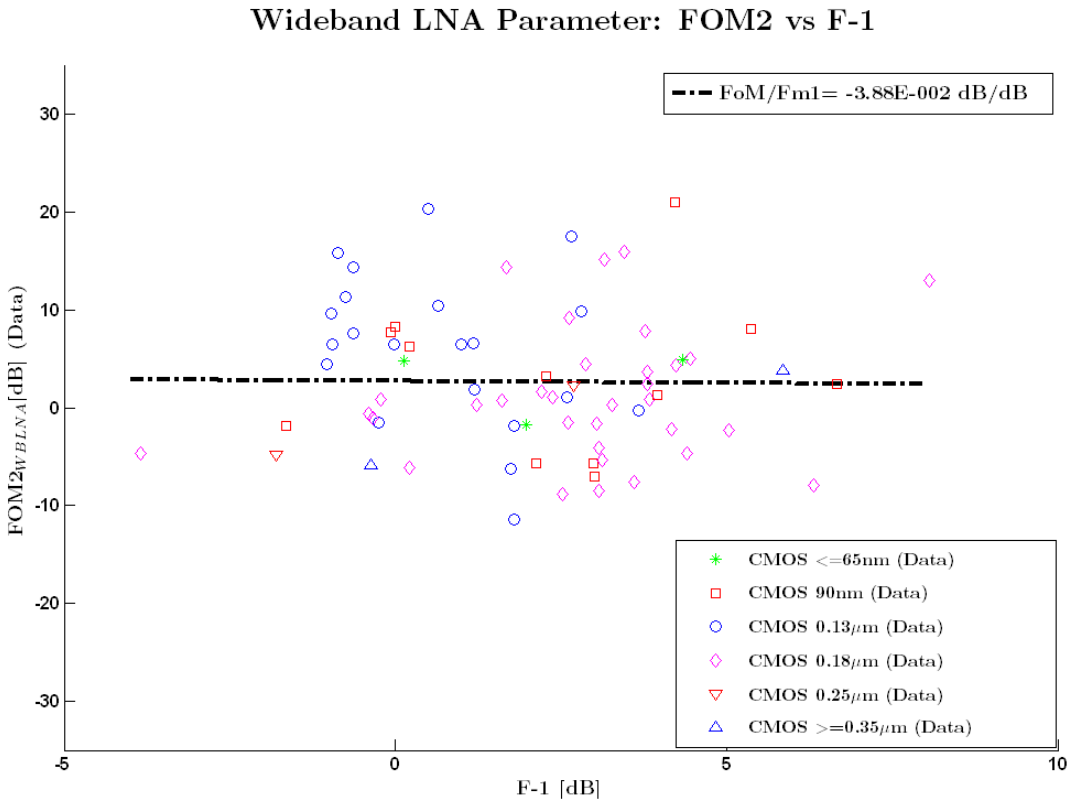
When the overdrive voltage is not very high, that is, from about 200mV above V_{TH}, the overdrive voltage is in general proportional to the ratio of current and transconductance, and then Equation (A-12) can be approximated as:

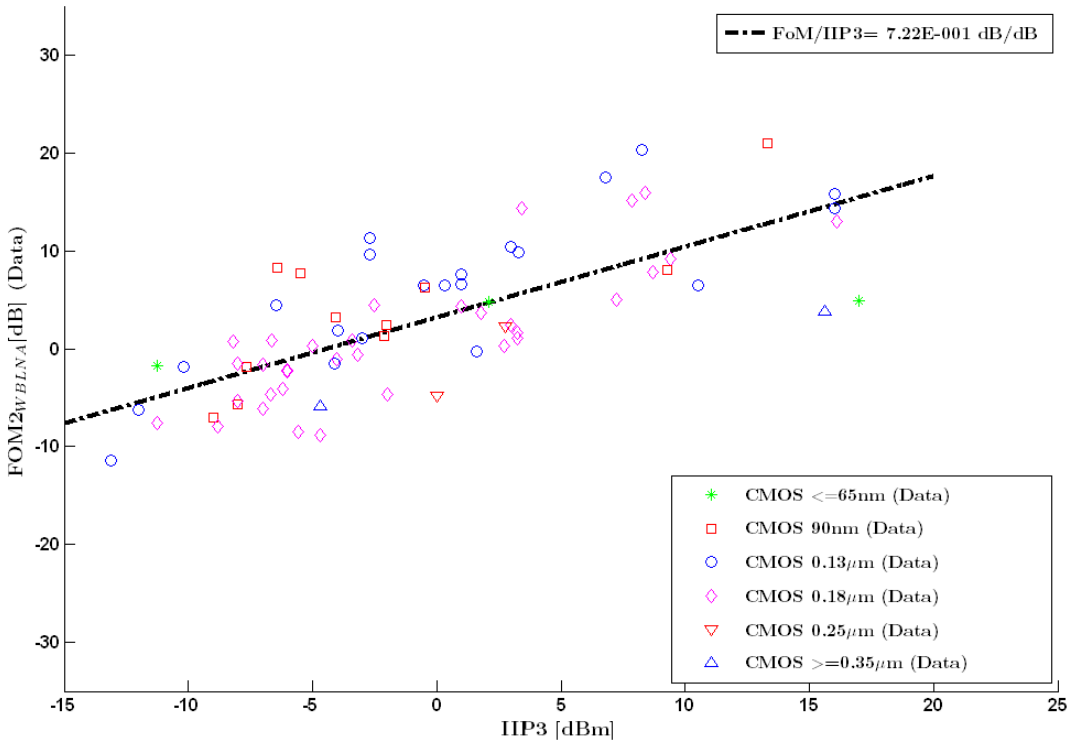
$$V_{IIP3} = \propto \sqrt{V_{ov}} \propto \sqrt{I_D/g_m}. \tag{A-13}$$



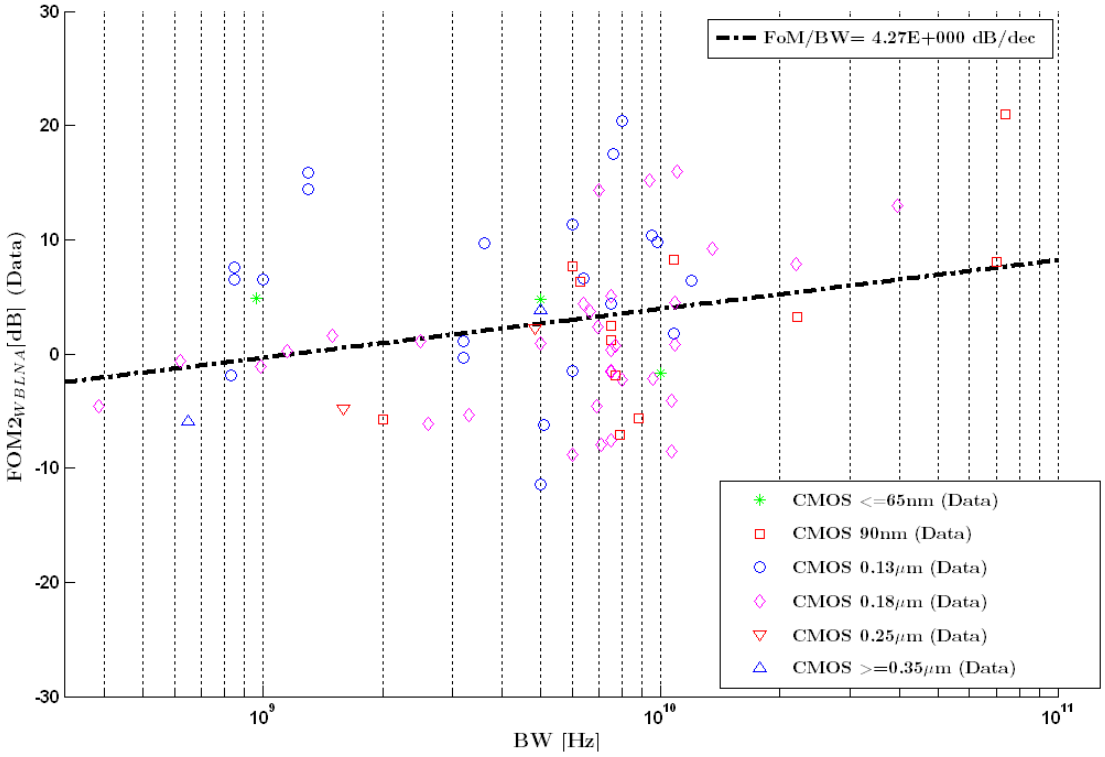
A-4 Wideband FoM Statistics



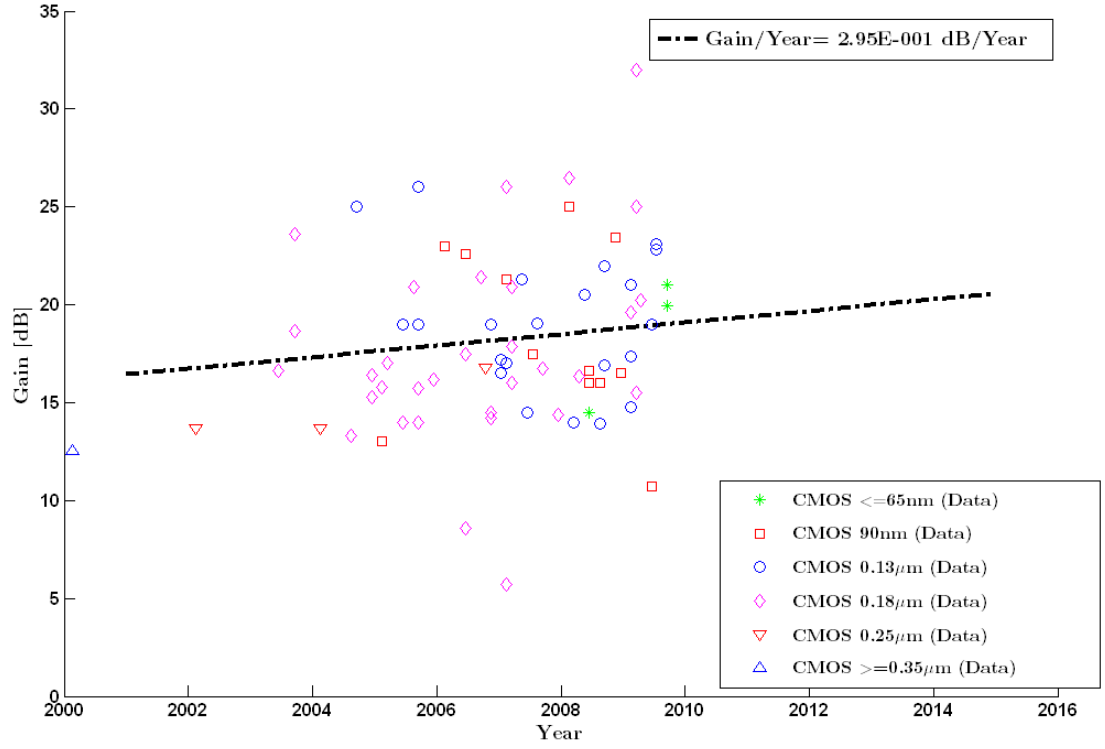

Wideband LNA Parameter: FOM2 vs F-1
FoM/Fm1= -3.88E-002 dB/dB
FOM2$_{WBLNA}$[dB] (Data)
F-1 [dB]
CMOS <=65nm (Data)
CMOS 90nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS >=0.35μm (Data)

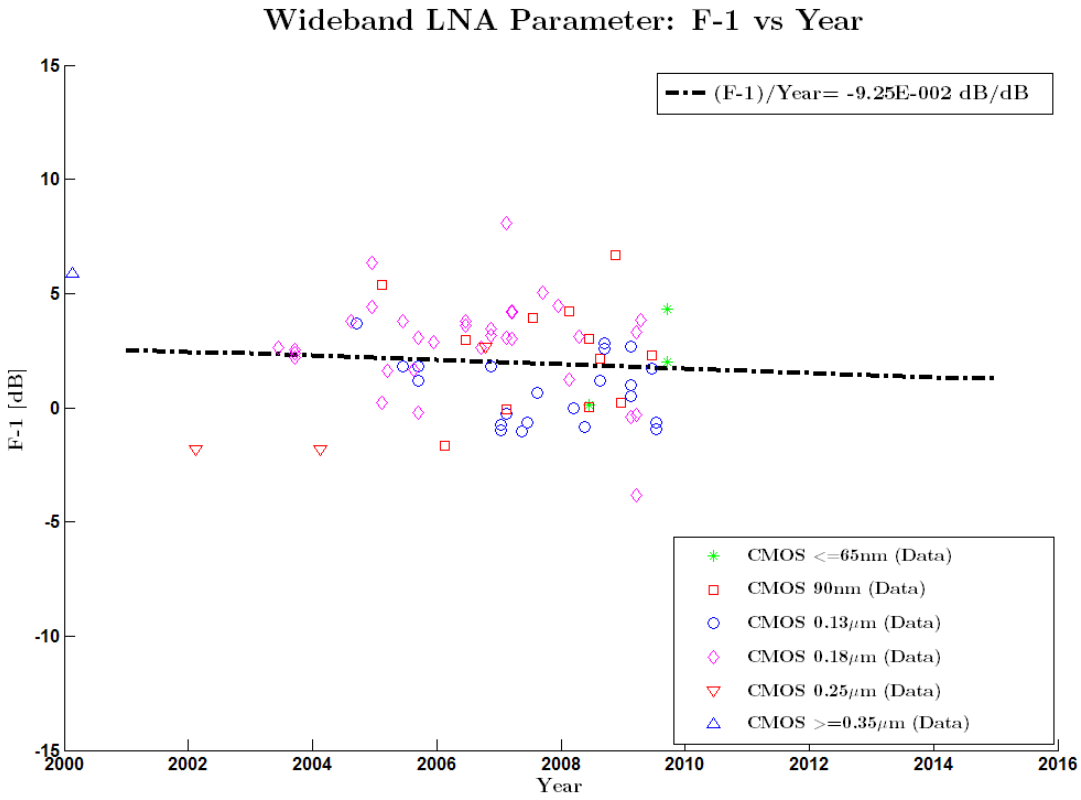

Wideband LNA Parameter: FOM2 vs IIP3
FoM/IIP3= 7.22E-001 dB/dB
FOM2$_{WBLNA}$[dB] (Data)
IIP3 [dBm]
CMOS <=65nm (Data)
CMOS 90nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS >=0.35μm (Data)

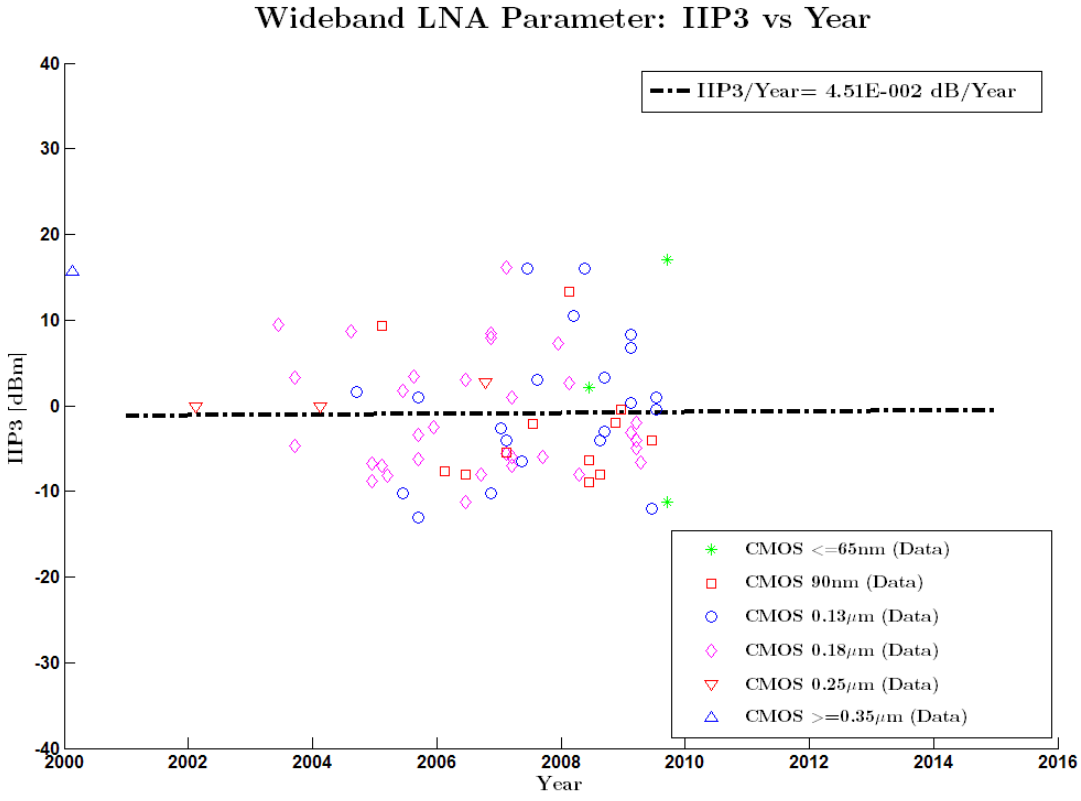
Wideband LNA Parameter: FOM2 vs BW



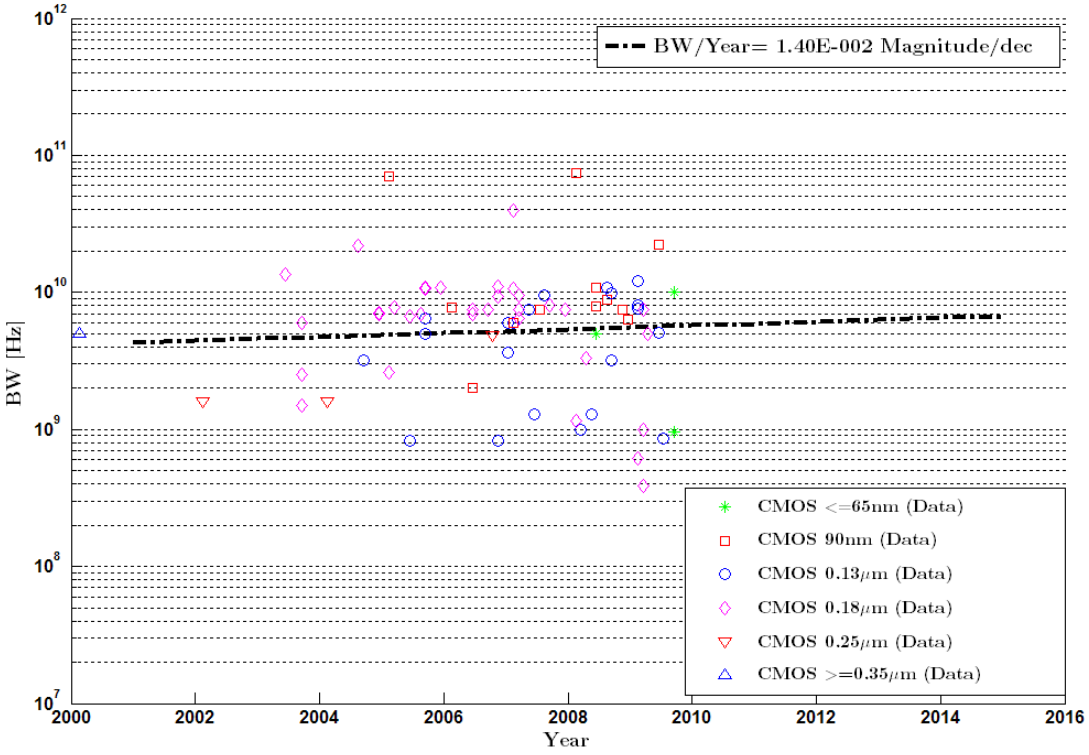
Wideband LNA Parameter: Voltage Gain vs Year



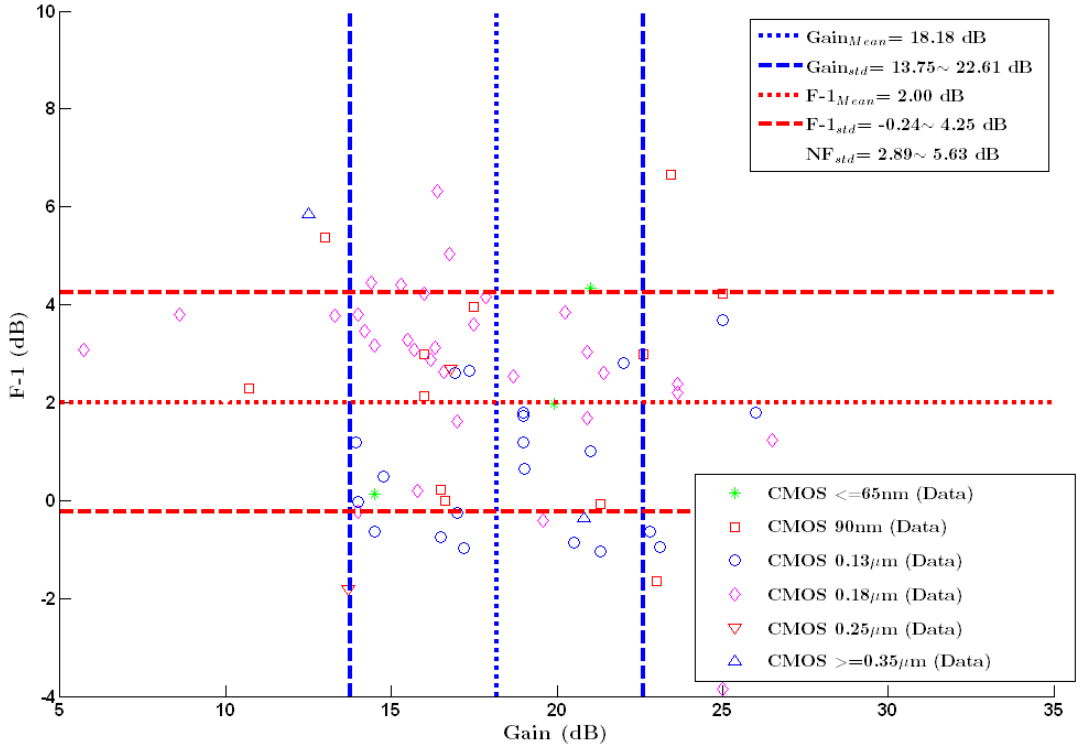

Wideband LNA Parameter: F-1 vs Year
(F-1)/Year= -9.25E-002 dB/dB
CMOS <=65nm (Data)
CMOS 90nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS >=0.35μm (Data)
F-1 [dB]
Year

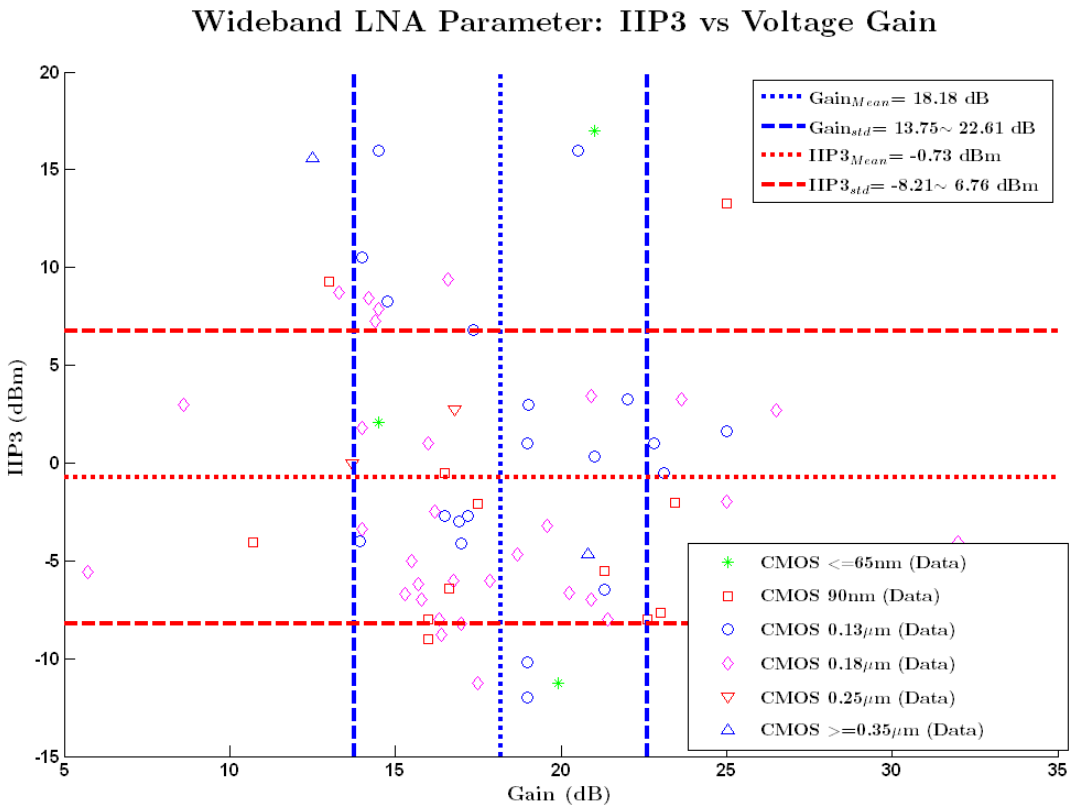

Wideband LNA Parameter: IIP3 vs Year
IIP3/Year= 4.51E-002 dB/Year
CMOS <=65nm (Data)
CMOS 90nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS >=0.35μm (Data)
IIP3 [dBm]
Year

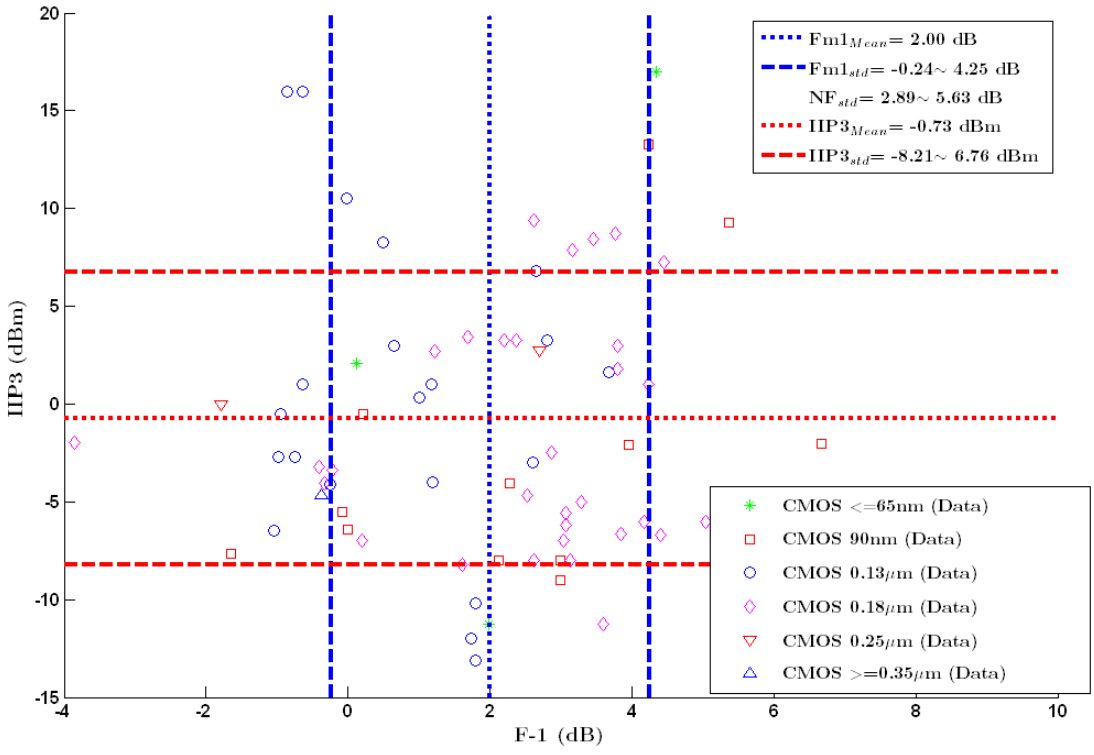
Wideband LNA Parameter: BW vs Year



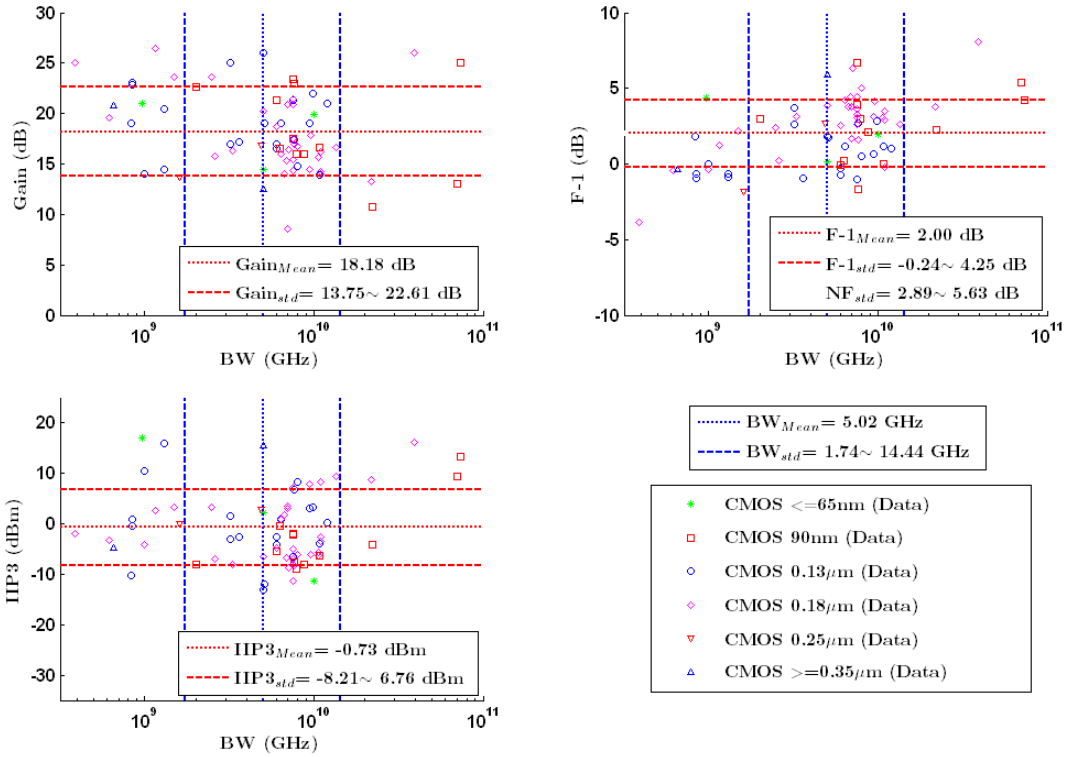
Wideband LNA Parameter: F-1 vs Voltage Gain



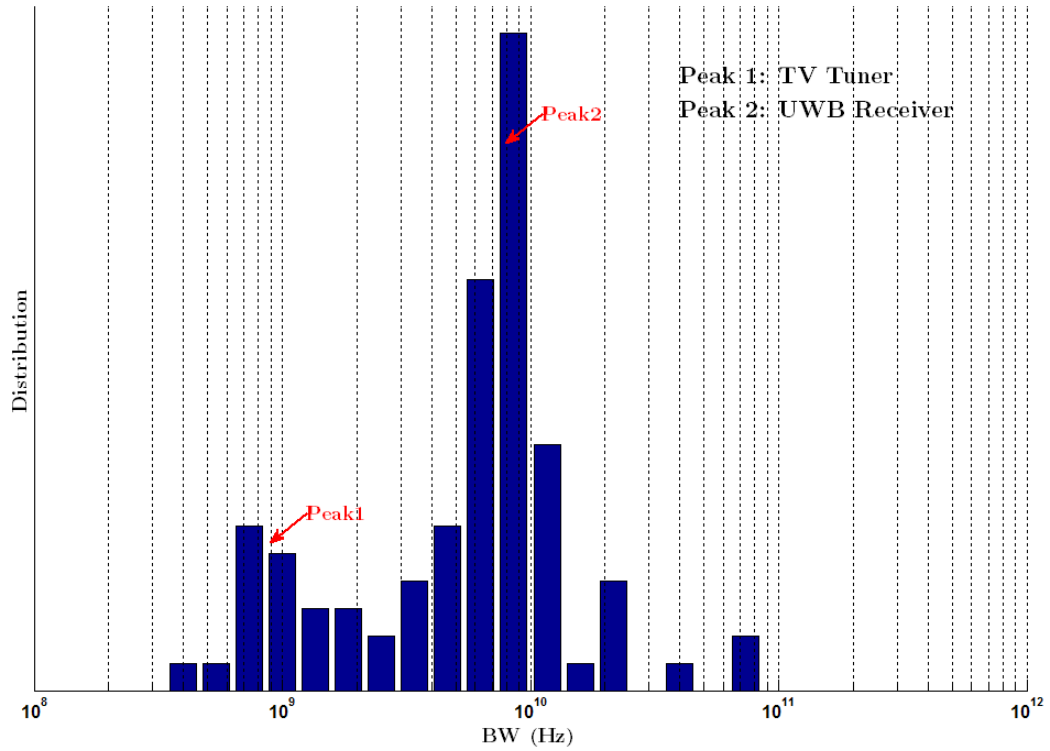

Wideband LNA Parameter: IIP3 vs Voltage Gain
Gain$_{Mean}$= 18.18 dB
Gain$_{std}$= 13.75~ 22.61 dB
IIP3$_{Mean}$= -0.73 dBm
IIP3$_{std}$= -8.21~ 6.76 dBm
CMOS <=65nm (Data)
CMOS 90nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS >=0.35μm (Data)
IIP3 (dBm)
Gain (dB)


Wideband LNA Parameter: IIP3 vs F-1
Fm1$_{Mean}$= 2.00 dB
Fm1$_{std}$= -0.24~ 4.25 dB
NF$_{std}$= 2.89~ 5.63 dB
IIP3$_{Mean}$= -0.73 dBm
IIP3$_{std}$= -8.21~ 6.76 dBm
CMOS <=65nm (Data)
CMOS 90nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS >=0.35μm (Data)
IIP3 (dBm)
F-1 (dB)

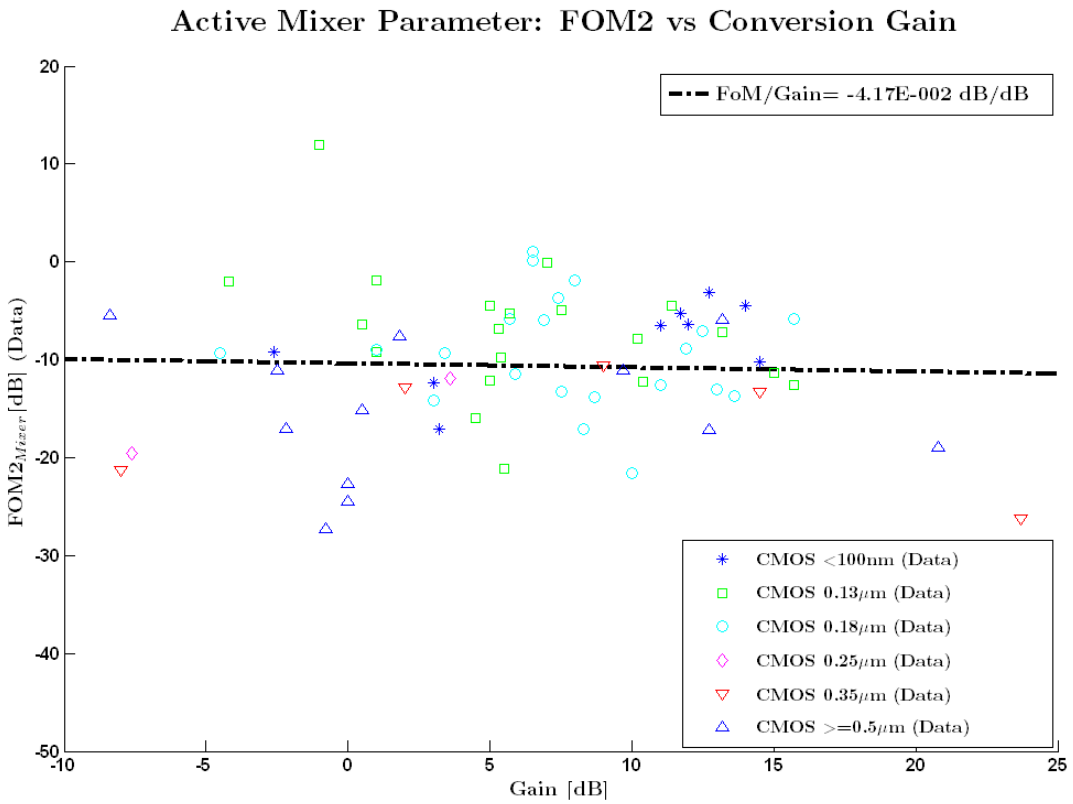
Wideband LNA Parameter: Gain/Noise/IIP3 vs BW

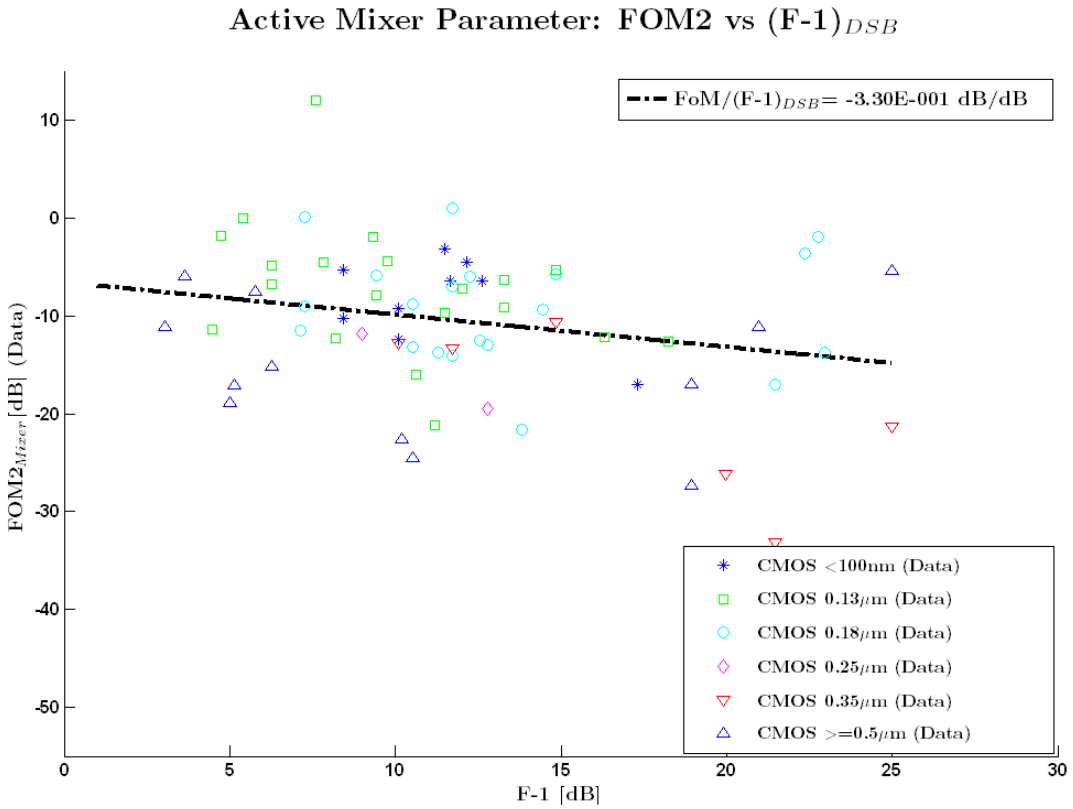


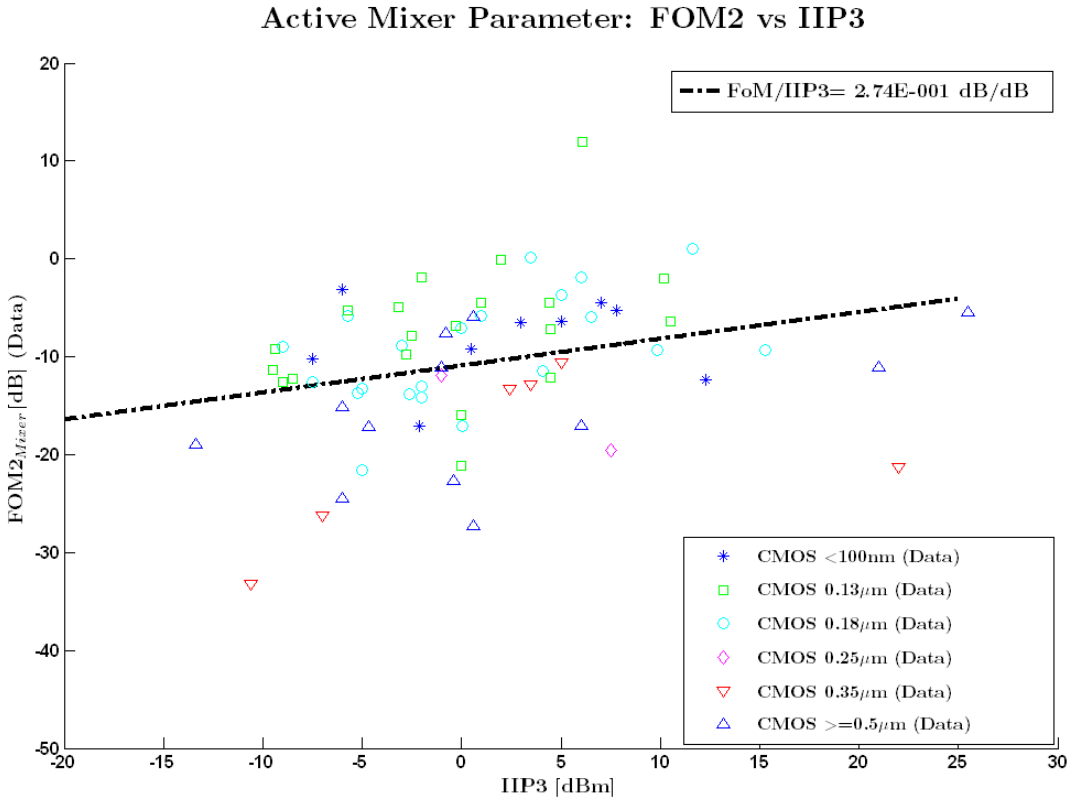
Wideband LNA Parameter Distribution: BW

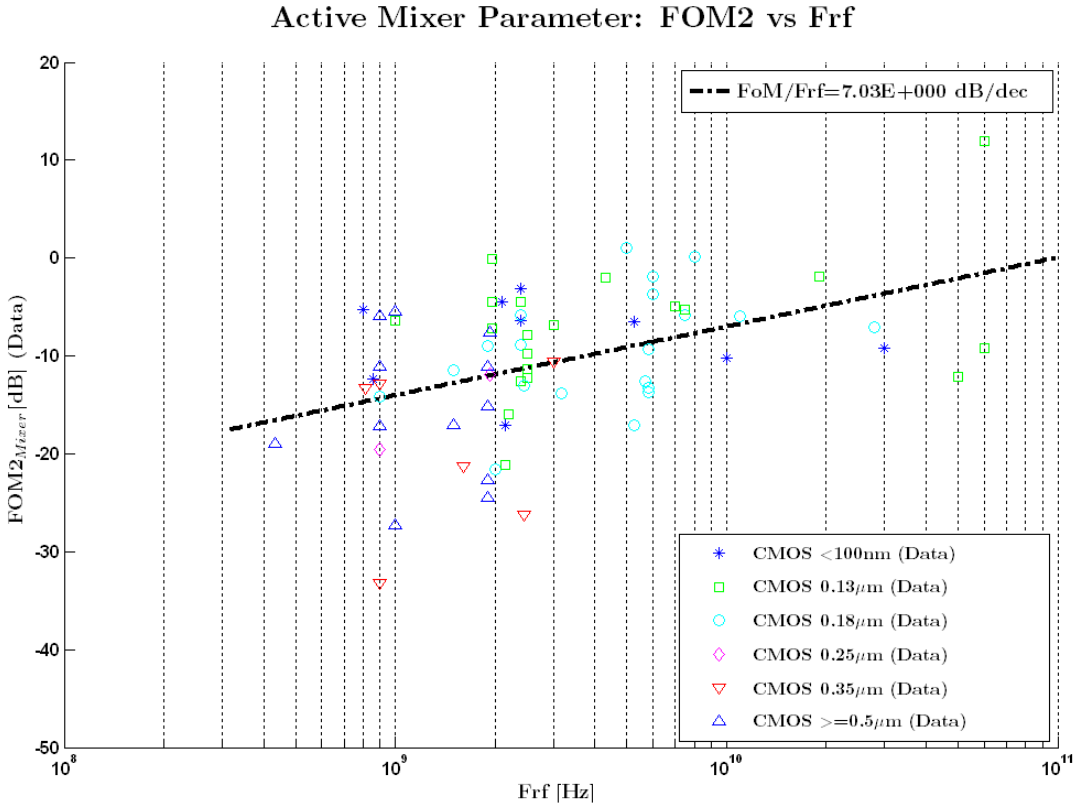


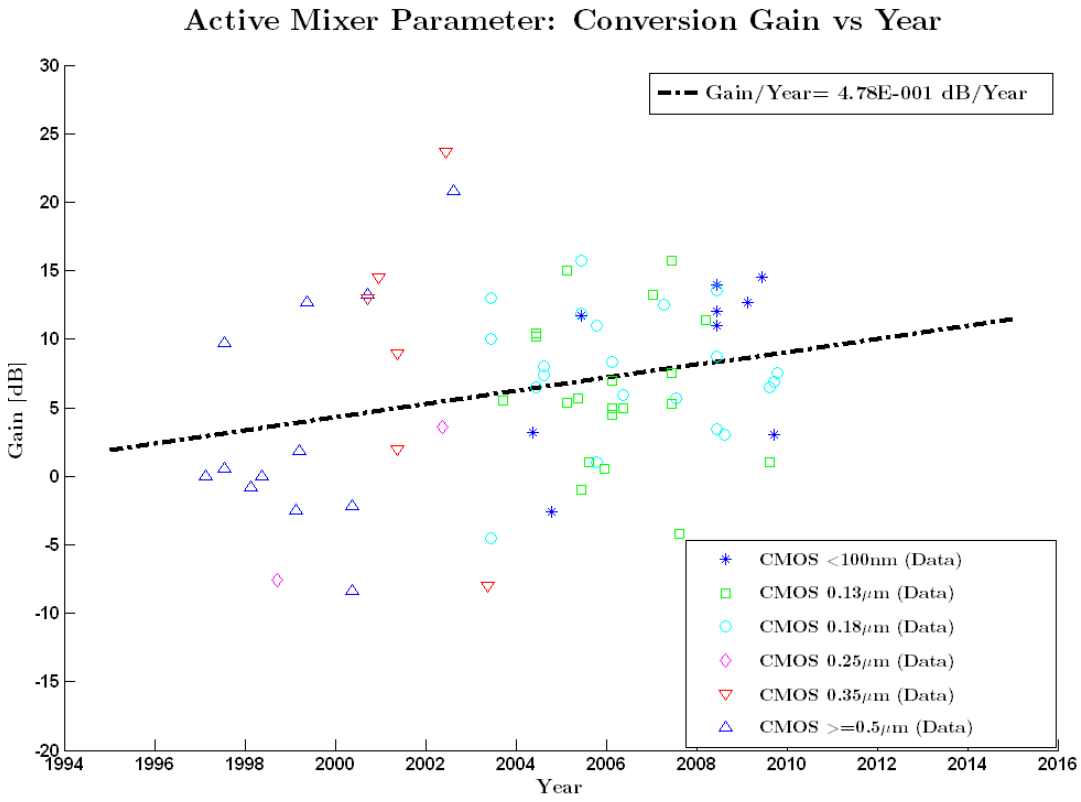
A-5 Mixer FoM Statistic

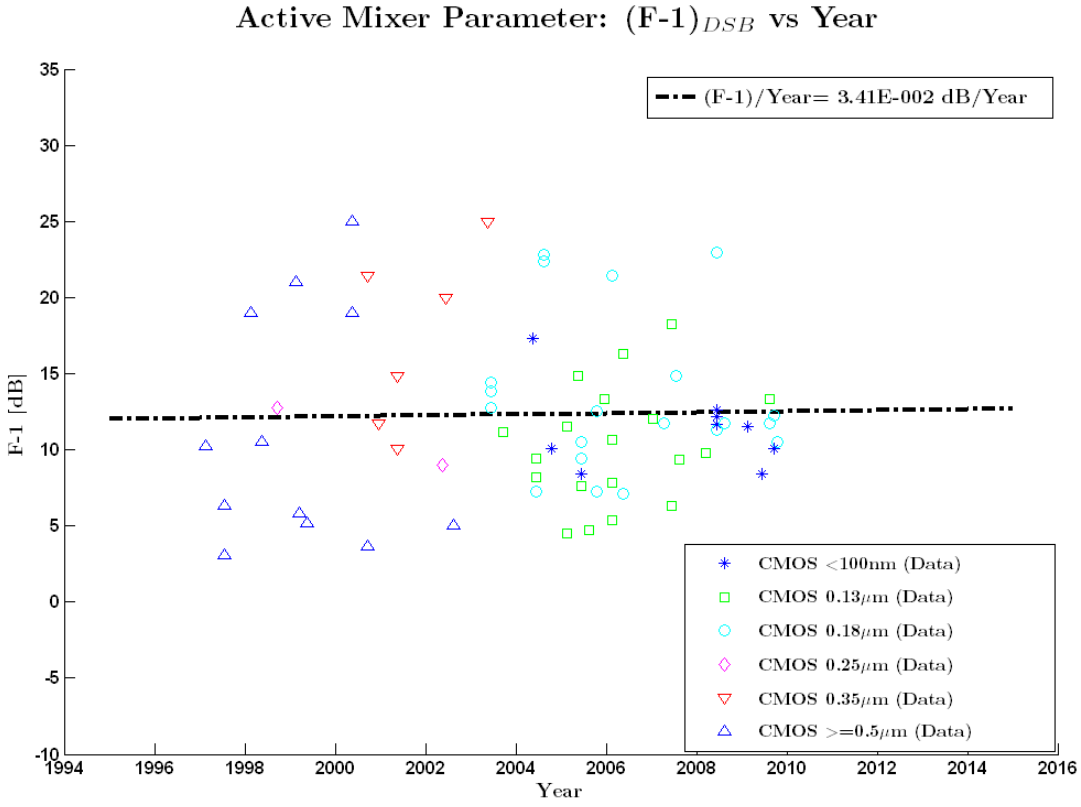


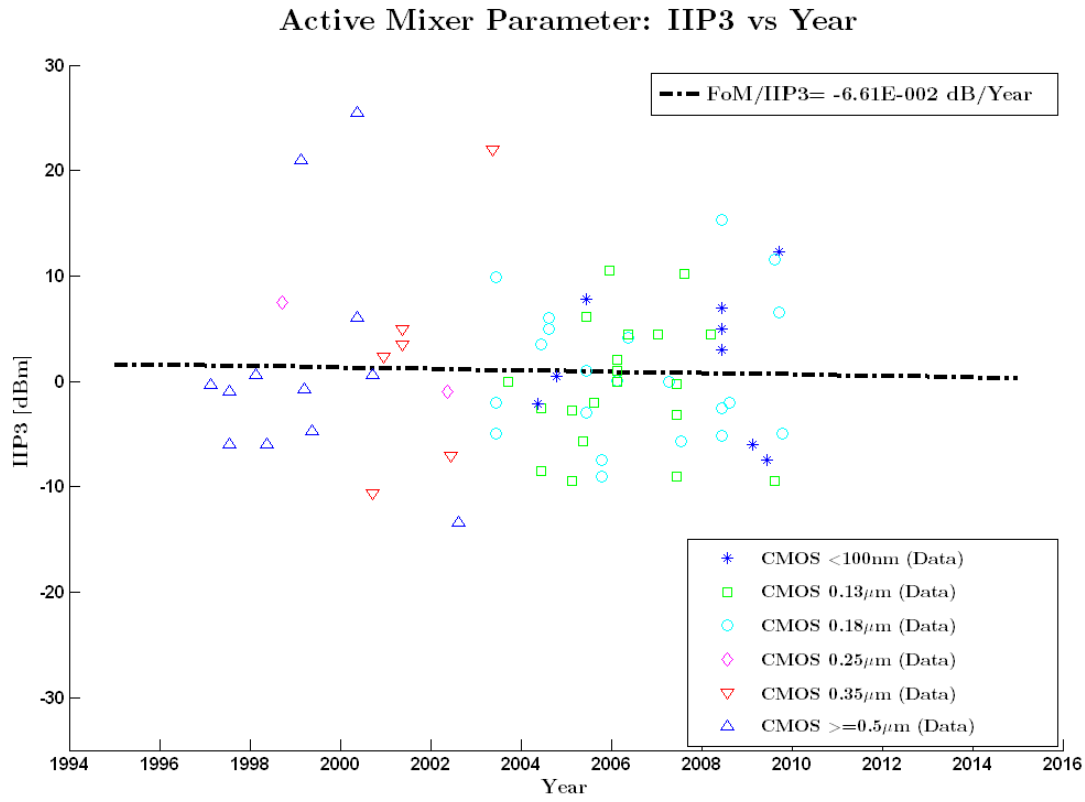


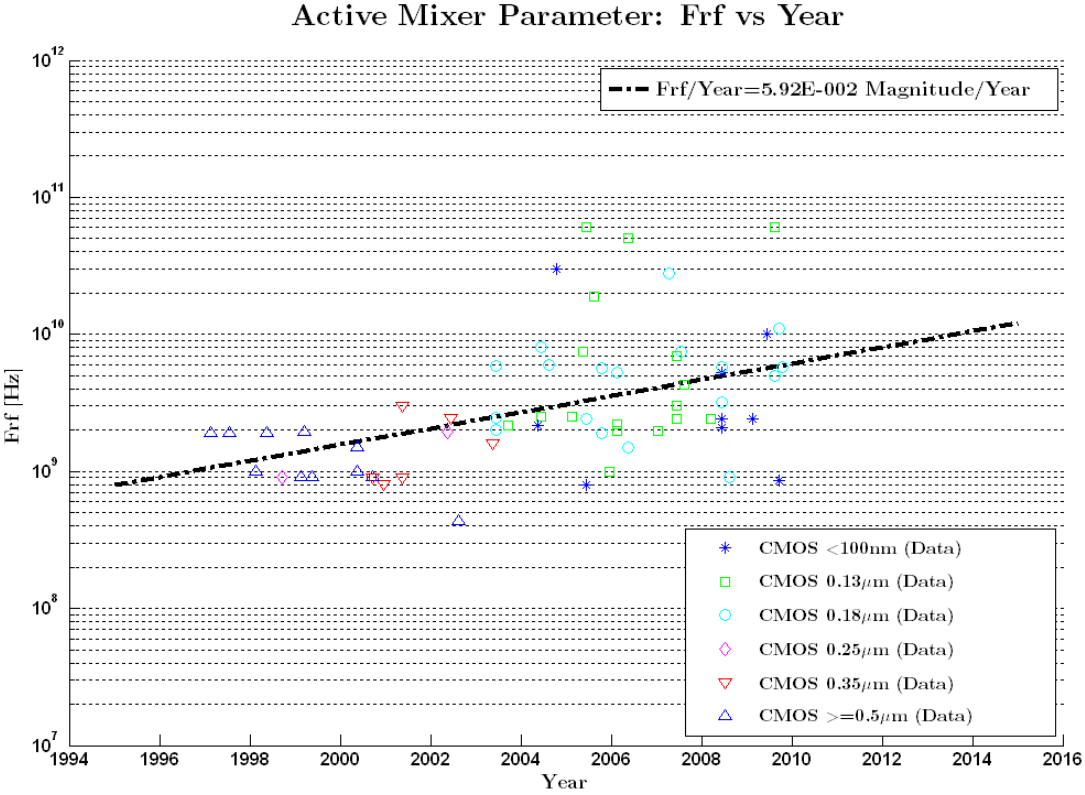

Active Mixer Parameter: FOM2 vs IIP3
FoM/IIP3= 2.74E-001 dB/dB
CMOS <100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
FOM2Mixer [dB] (Data)
IIP3 [dBm]


Active Mixer Parameter: FOM2 vs Frf
FoM/Frf=7.03E+000 dB/dec
CMOS <100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
FOM2Mixer [dB] (Data)
Frf [Hz]

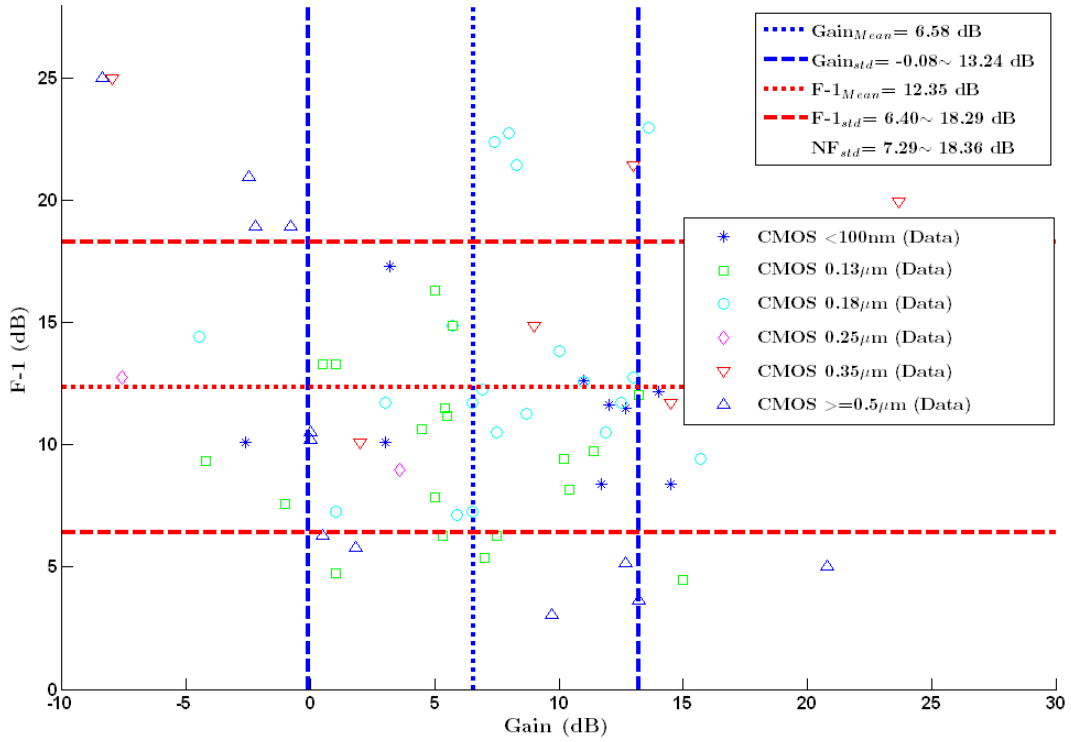

Active Mixer Parameter: Conversion Gain vs Year
Gain/Year= 4.78E-001 dB/Year
Gain [dB]
Year
CMOS <100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)


Active Mixer Parameter: (F-1)$_{DSB}$ vs Year
(F-1)/Year= 3.41E-002 dB/Year
F-1 [dB]
Year
CMOS <100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)

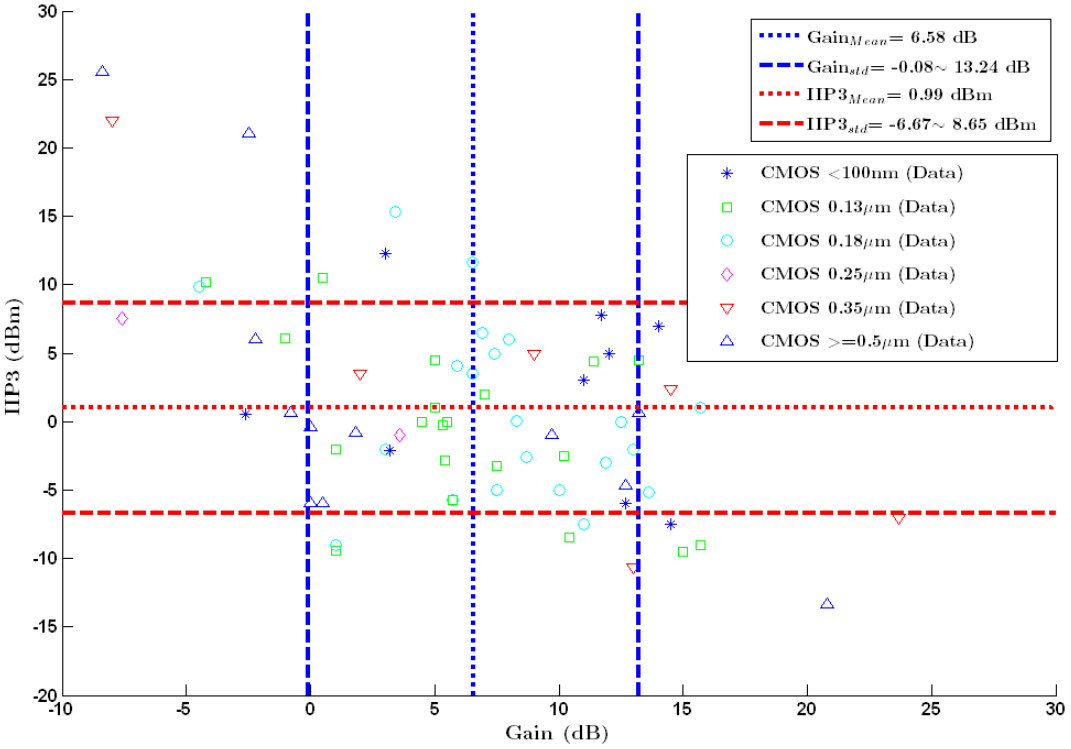

Active Mixer Parameter: IIP3 vs Year
FoM/IIP3= -6.61E-002 dB/Year
IIP3 [dBm]
Year
CMOS <100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)


Active Mixer Parameter: Frf vs Year
Frf/Year=5.92E-002 Magnitude/Year
Frf [Hz]
Year
CMOS <100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)

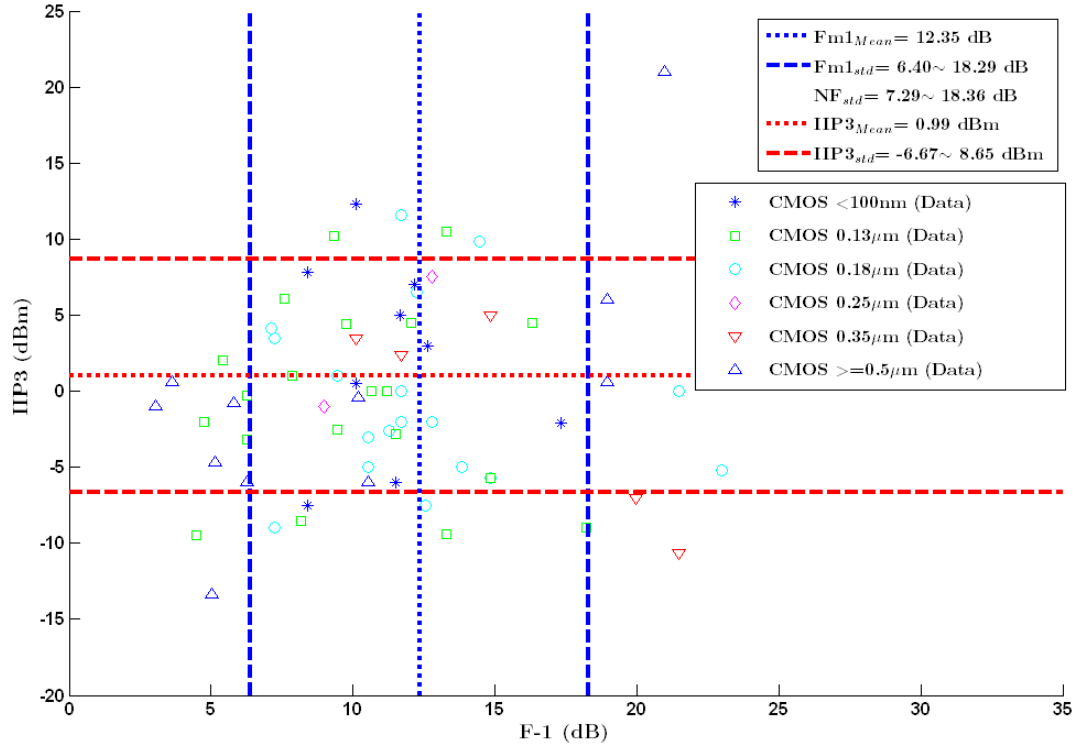
Active Mixer Parameter: $(\text{F-1})_{DSB}$ vs Conversion Gain



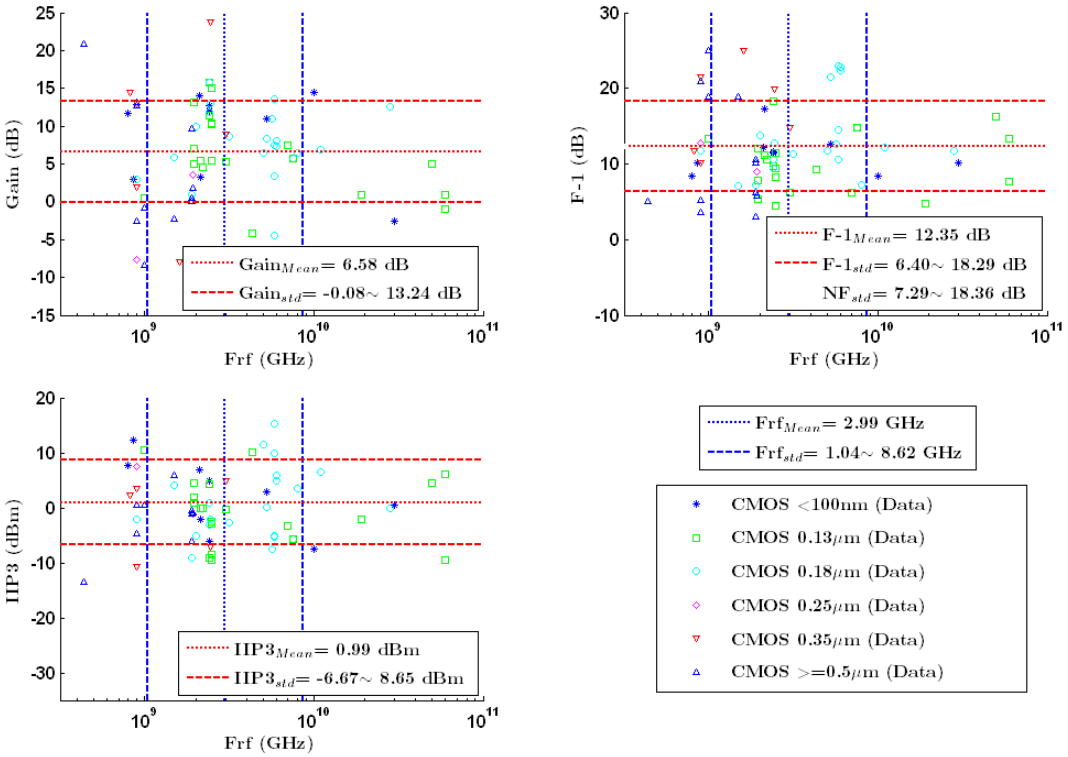
Active Mixer Parameter: IIP3 vs Conversion Gain

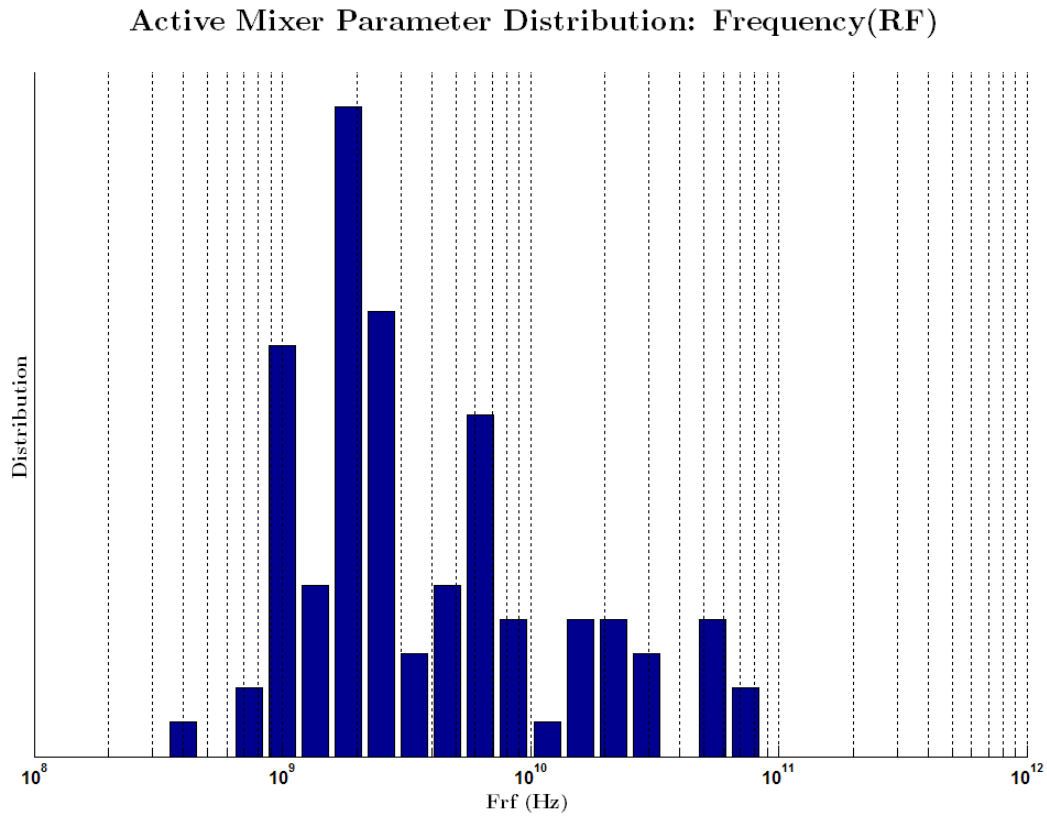


Active Mixer Parameter: IIP3 vs (F-1)$_{DSB}$

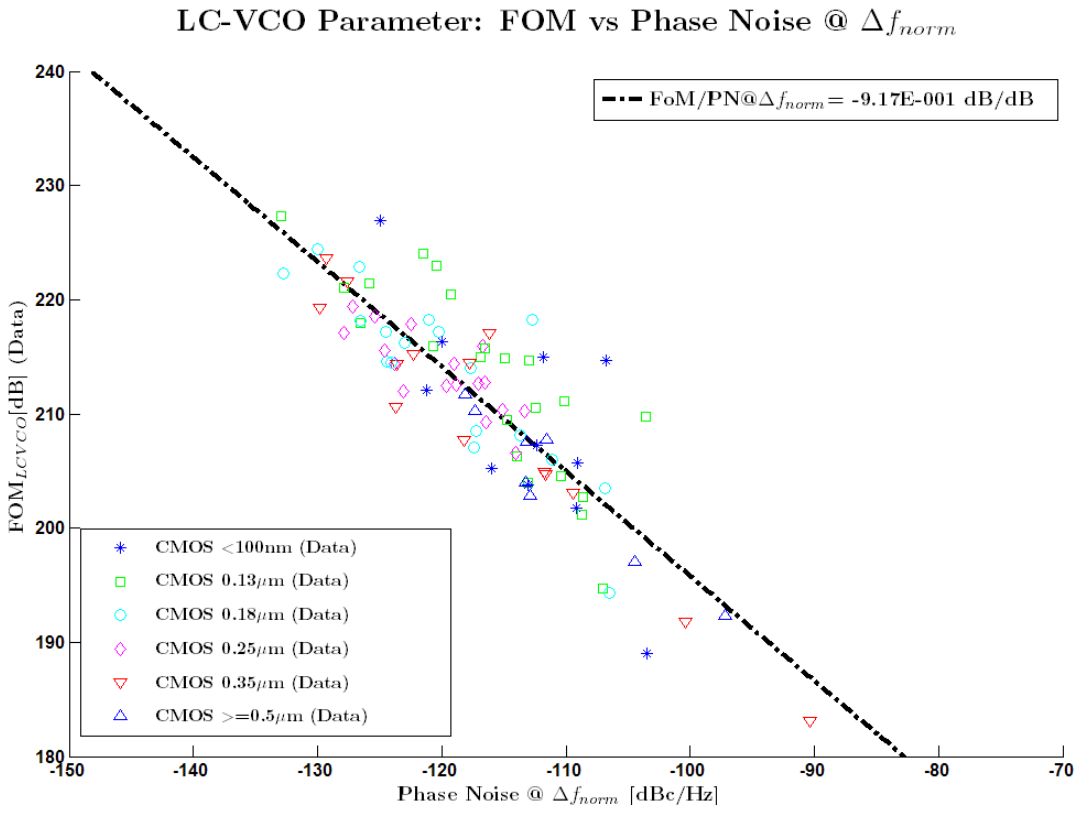


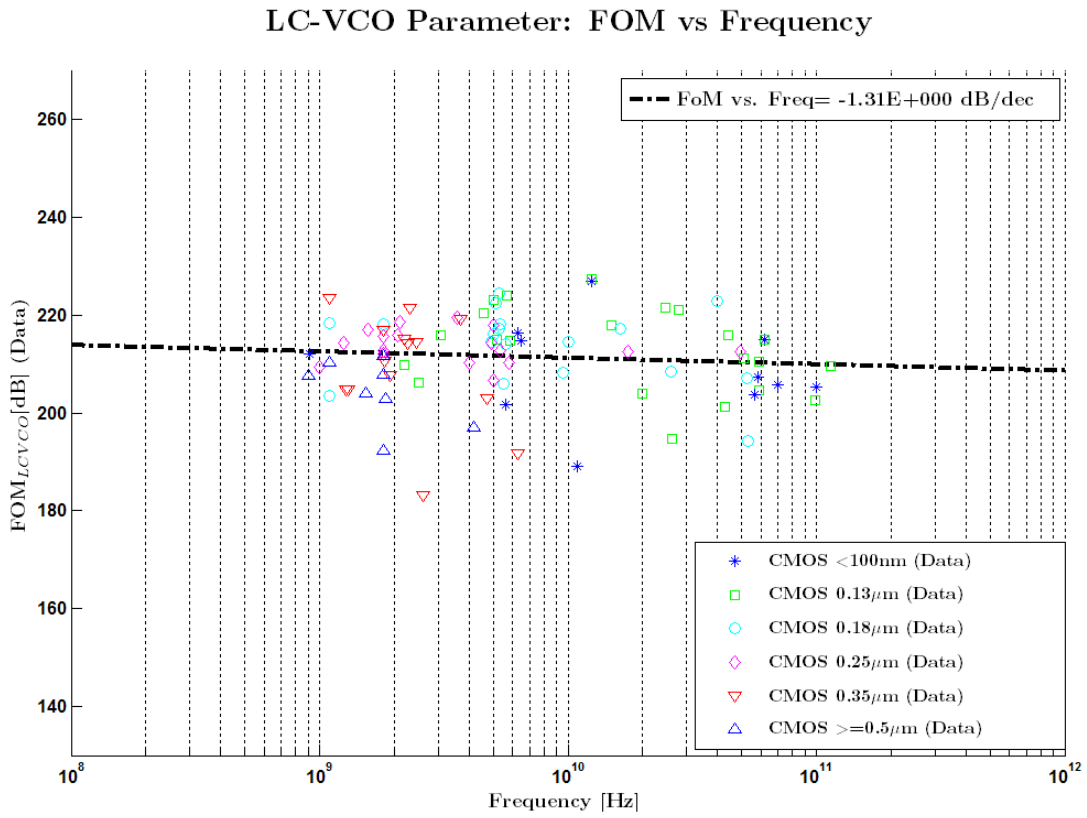
Active mixer Parameter: Gain/Noise/IIP3 vs Frf

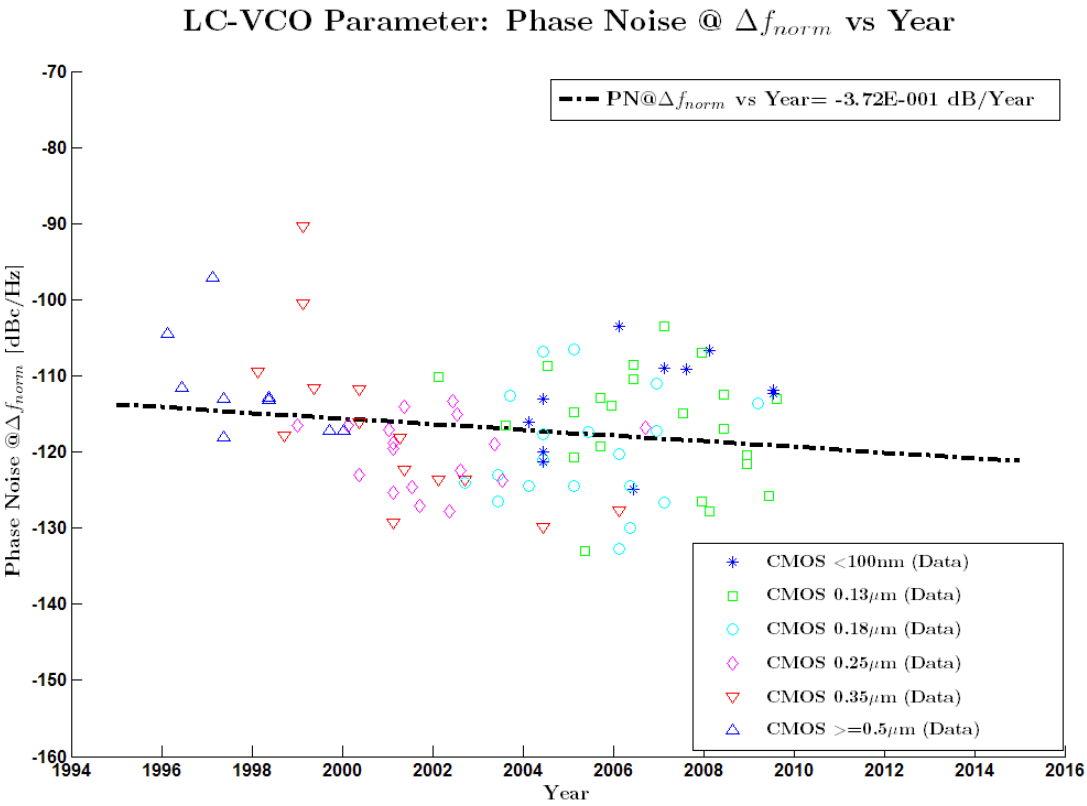


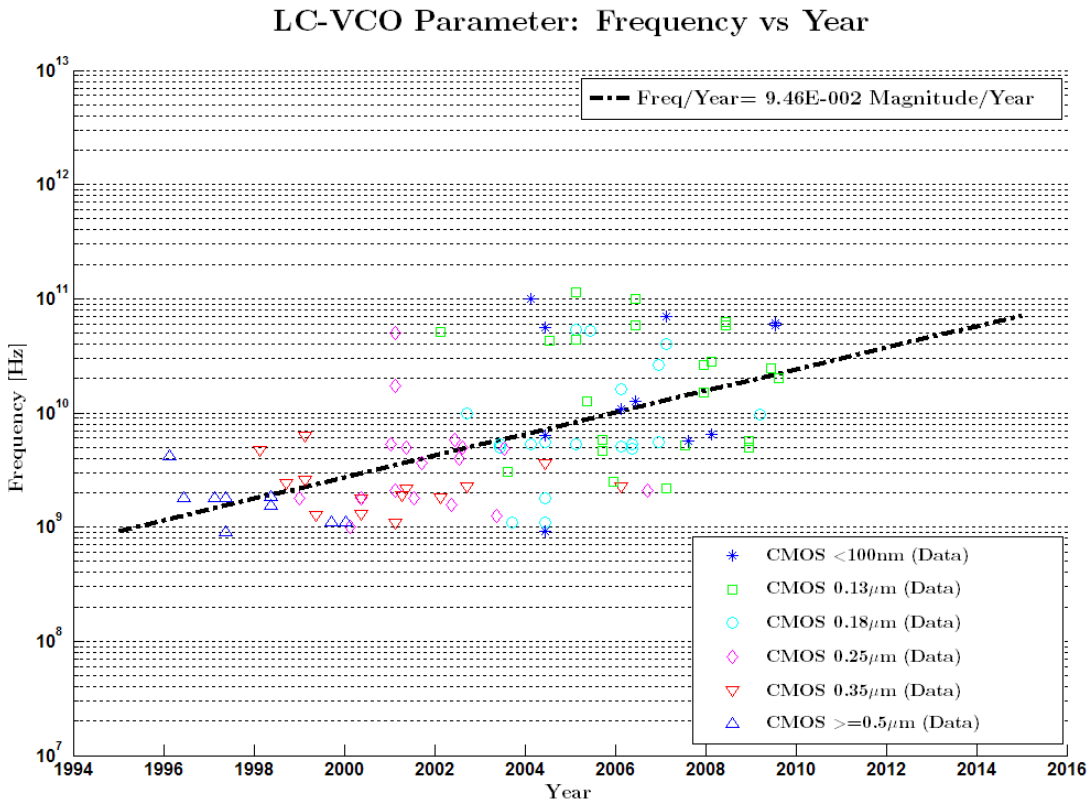


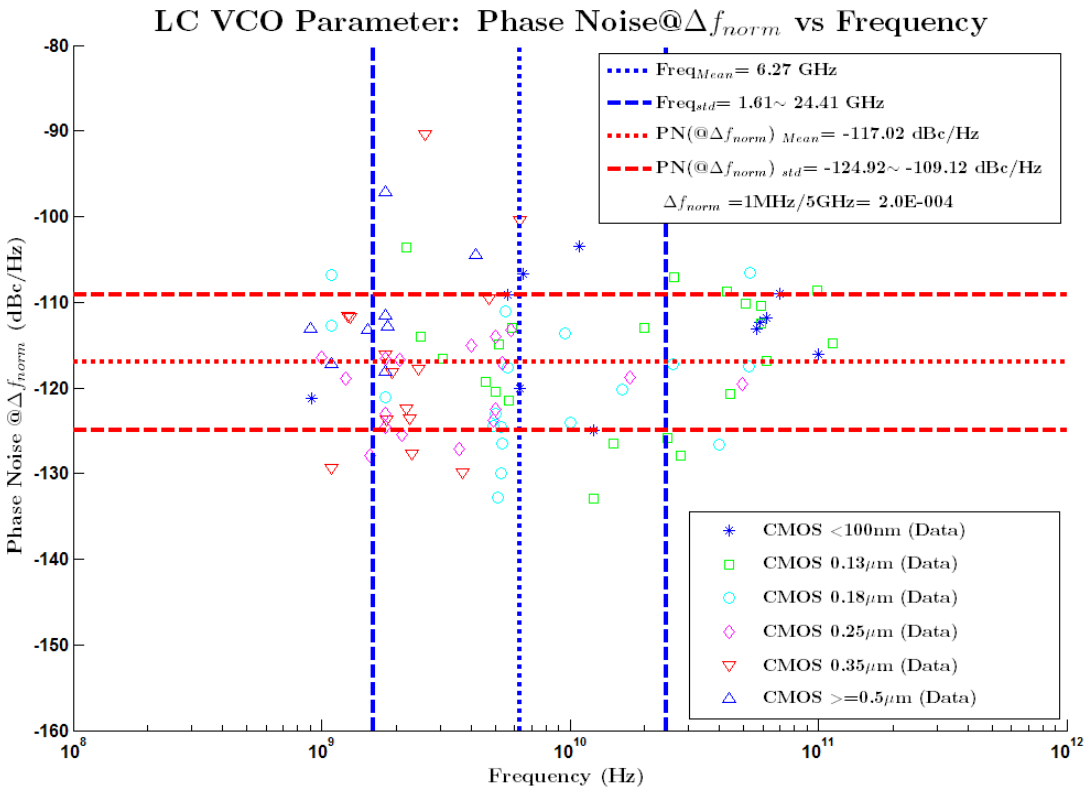
A-6 LC-VCO FoM Statistic



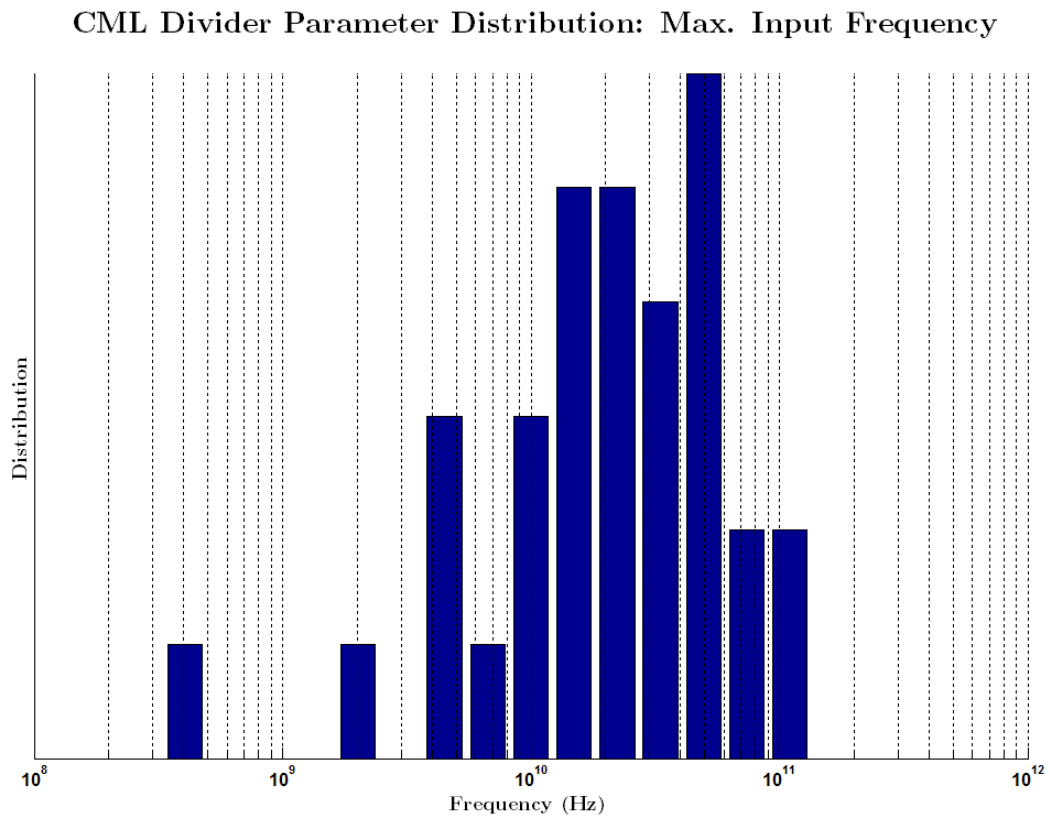

LC-VCO Parameter: FOM vs Frequency
FoM vs. Freq= -1.31E+000 dB/dec
CMOS <100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
FOM_LCVCO[dB] (Data)
Frequency [Hz]

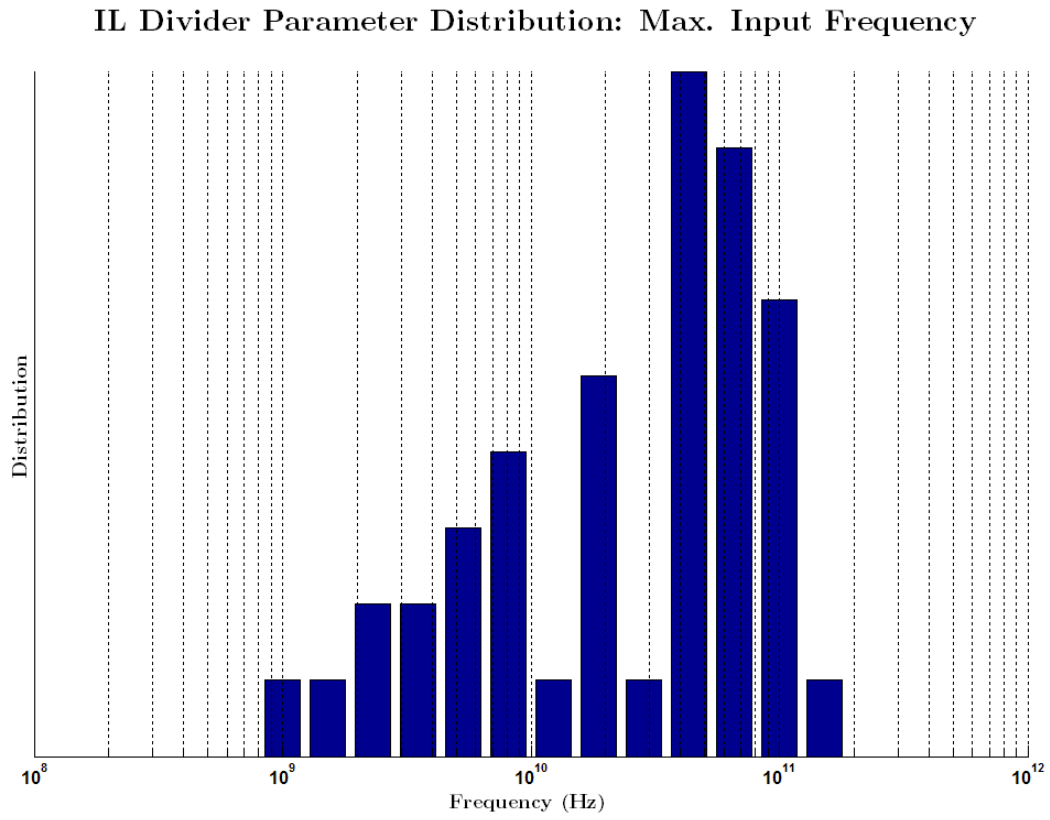

LC-VCO Parameter: Phase Noise @ Δf_norm vs Year
PN@Δf_norm vs Year= -3.72E-001 dB/Year
CMOS <100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
Phase Noise @Δf_norm [dBc/Hz]
Year


LC-VCO Parameter: Frequency vs Year
Freq/Year= 9.46E-002 Magnitude/Year
CMOS <100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
Frequency [Hz]
Year


LC VCO Parameter: Phase Noise@Δf_norm vs Frequency
Freq_Mean= 6.27 GHz
Freq_std= 1.61~ 24.41 GHz
PN(@Δf_norm) Mean= -117.02 dBc/Hz
PN(@Δf_norm) std= -124.92~ -109.12 dBc/Hz
Δf_norm =1MHz/5GHz= 2.0E-004
CMOS <100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
Phase Noise @Δf_norm (dBc/Hz)
Frequency (Hz)

A-7 Frequency Divider statistics





A-8 Derivation of Linearity of LPF

The reported linearity measurement results of LPF are given by many different ways. If IIP3 is not directly presented, one can also estimate the IIP3 value when the total harmonic distortion (THD) is reported instead, and is briefly introduced as below.

A single frequency sine wave is used as the input. With a given input signal voltage applied, the measured the total output harmonic distortion (THD) in dB should be reported. With the assumption that all the nonlinear distortion products are caused by the 3rd order harmonics, the equivalent IIP3 can be calculated by the following method.

Assume the large signal transfer function is (ignoring any filtering effect behaviour) $V_{out} = a_1 V_{in} + a_3 V_{in}^3$, and input testing signal is $V_{in} = A\cos\omega t$ $V_{in} = A\cos\omega t$, the output signal can be expressed as

$$V_{out} = \left(a_1 A + \frac{3a_3 A^3}{4}\right)\cos\omega t + \frac{a_3 A^3}{4}\cos 3\omega t. \tag{A-14}$$

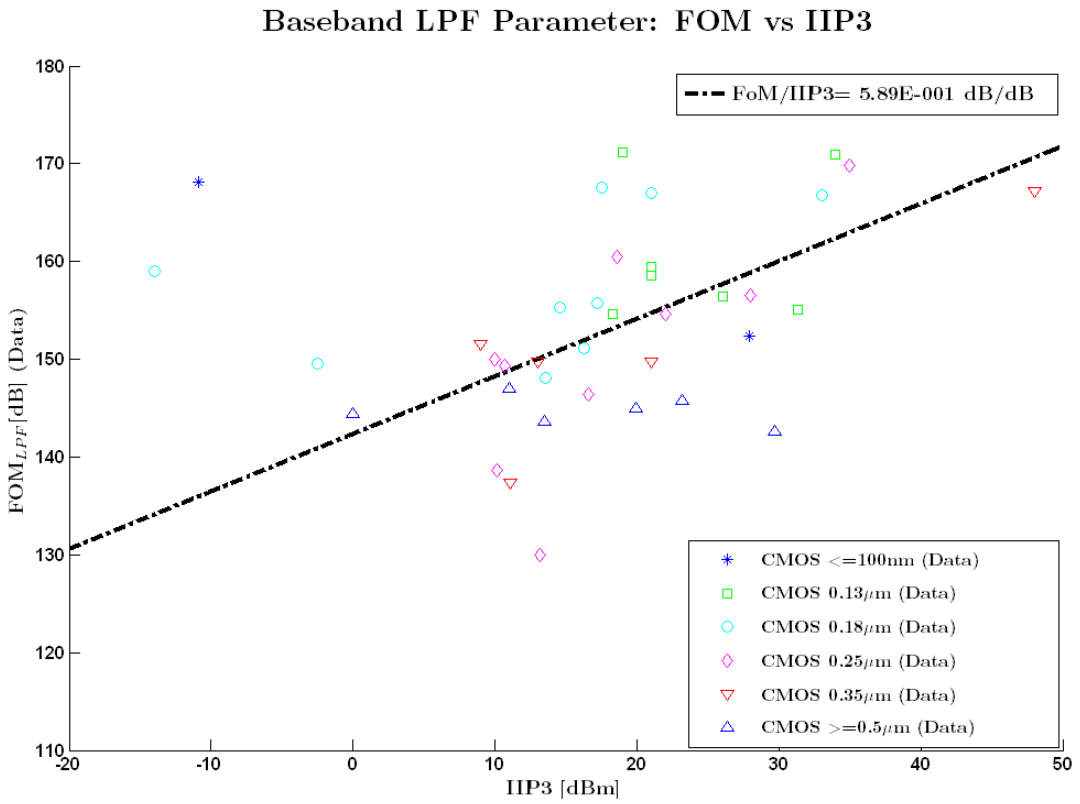
Equal the ratio of these two terms with the THD, and get the ratio of the DC gain and third order nonlinearity factor

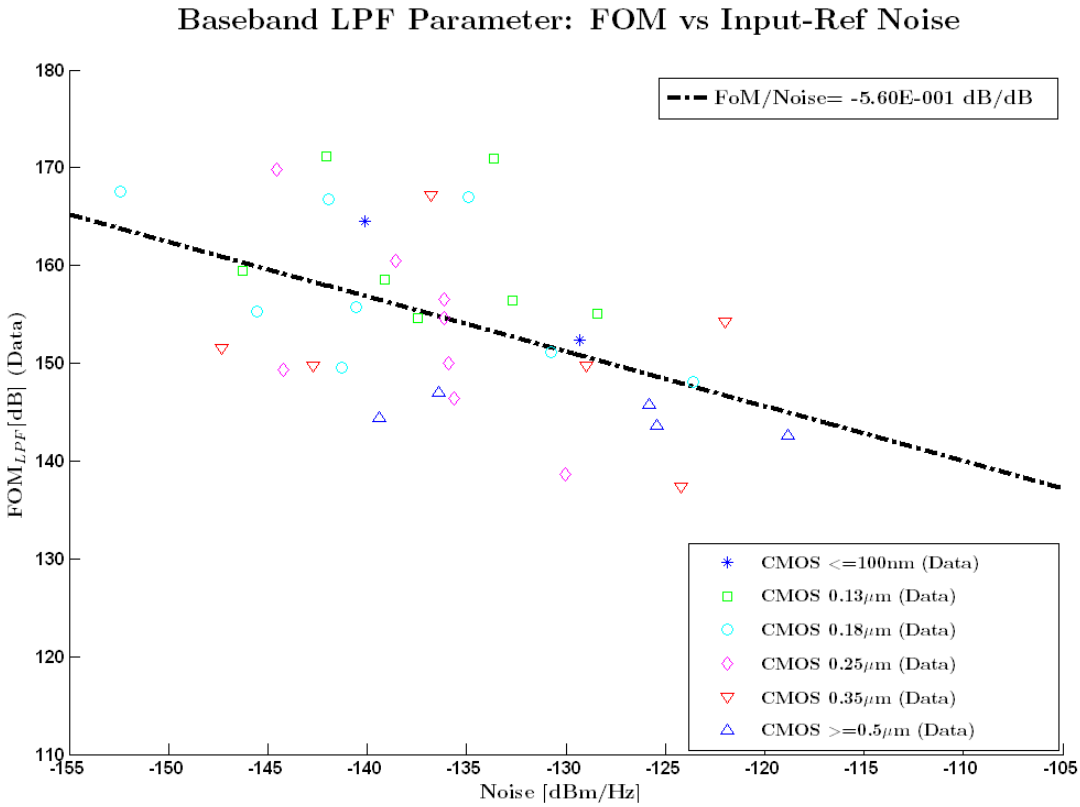
$$\begin{aligned} &\frac{\left(a_1 A + \frac{3a_3 A^3}{4}\right)}{\frac{a_3 A^3}{4}} = 10^{\frac{THD}{20}} \\ &\rightarrow \left|\frac{a_1}{a_3}\right| = \left|\frac{10^{\frac{THD}{20}} - 3}{4}\right| \cdot A^2. \end{aligned} \tag{A-15}$$

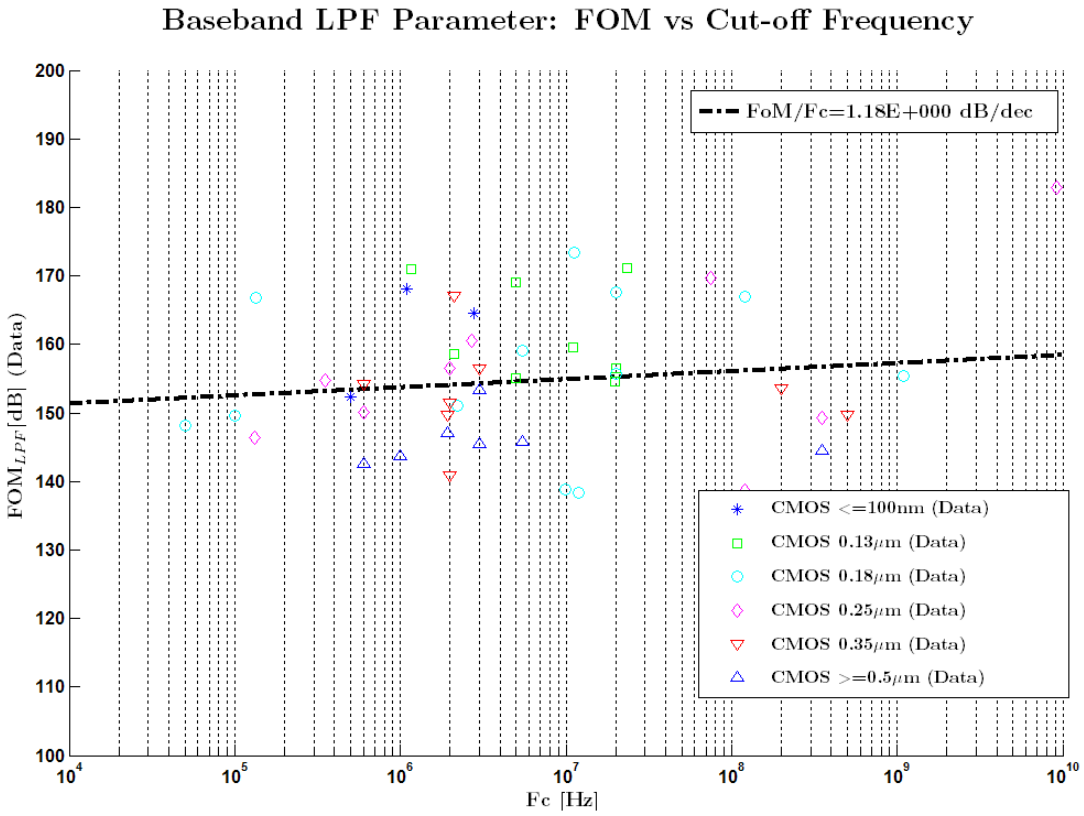
Hence the equivalent IIP3 in dBV can be calculated by

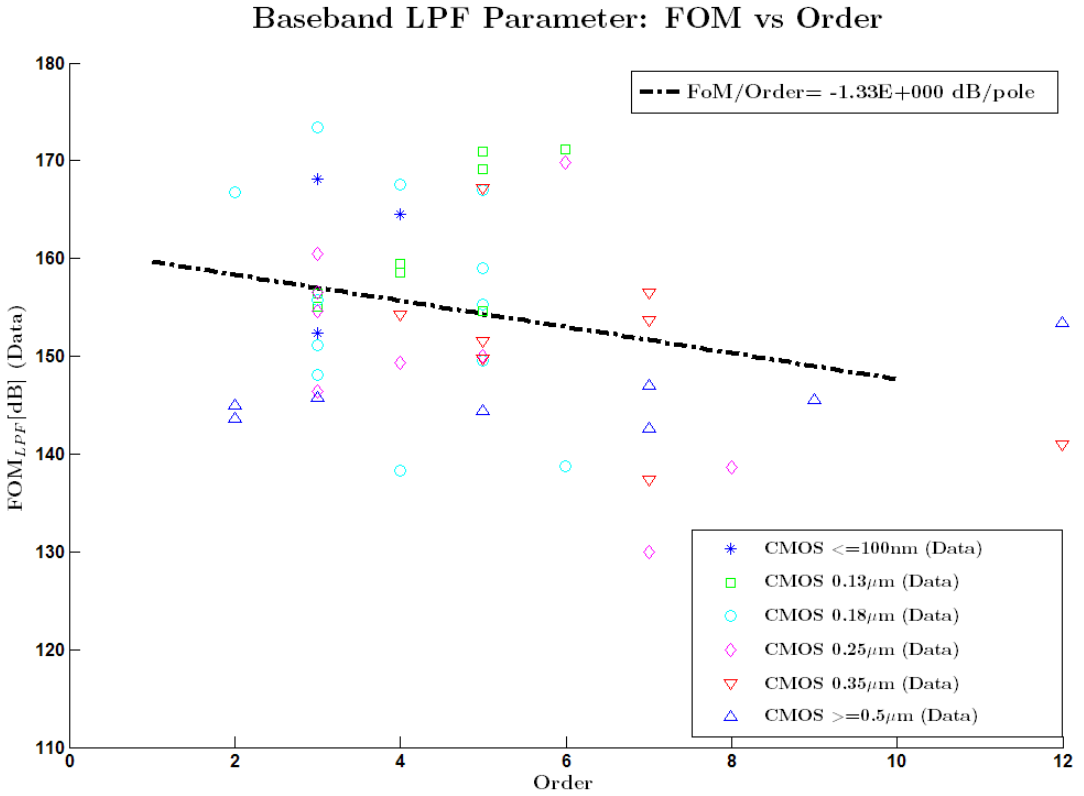
$$\begin{aligned} IIP3[dBV] &= 20\log\left(\frac{\left|\frac{4}{3}\cdot\left|\frac{a_1}{a_3}\right|\right|}{\sqrt{2}}\right) \\ &= 20\log\left(\sqrt{\left|\frac{10^{\frac{THD}{20}} - 3}{6}\right|}\cdot A\right). \end{aligned} \tag{A-16}$$

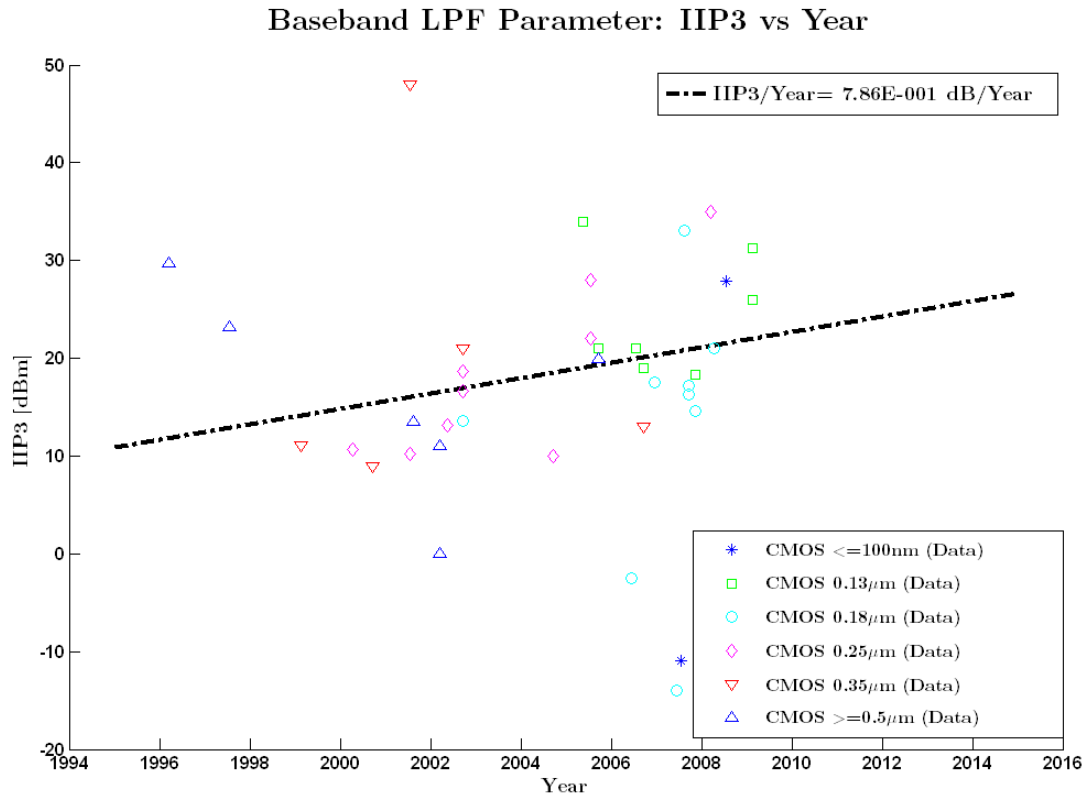
A-9 LPF FoM Statistics

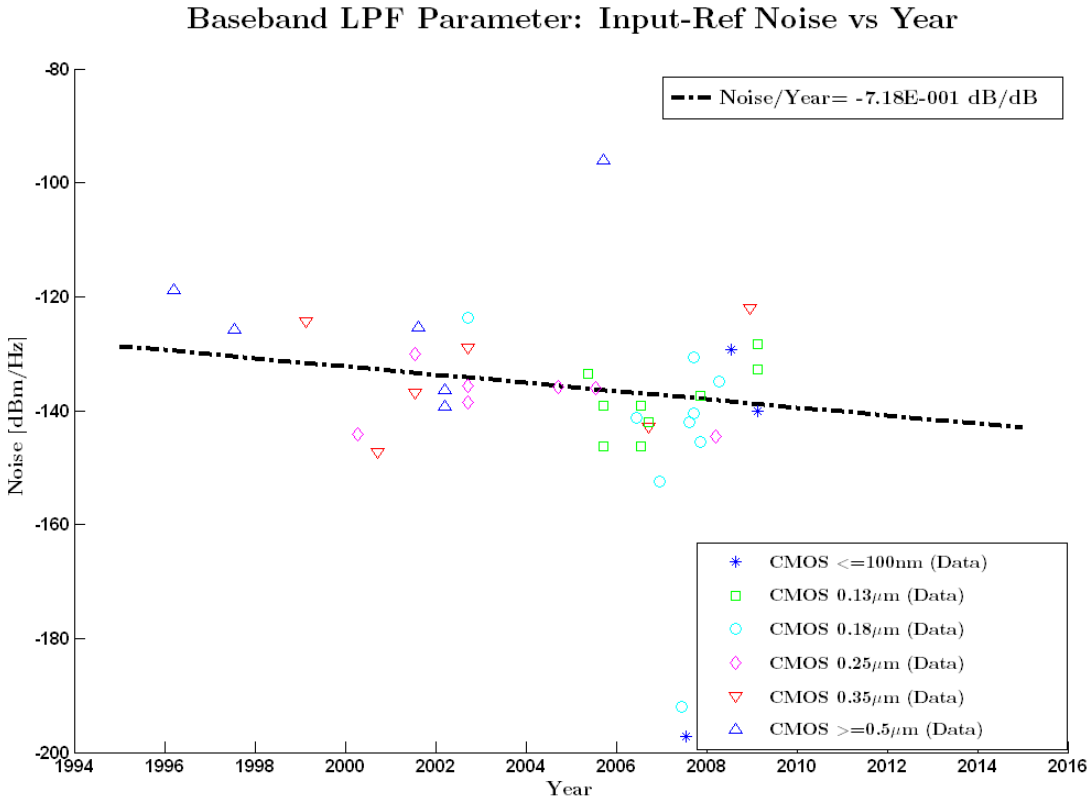


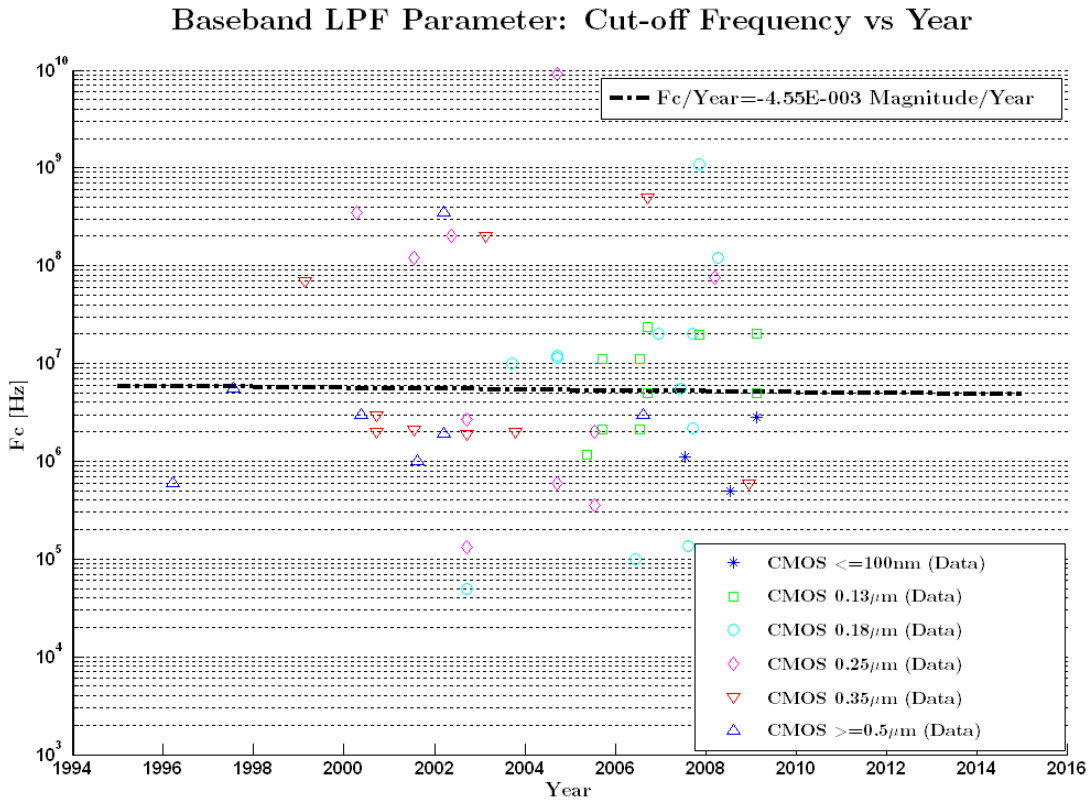


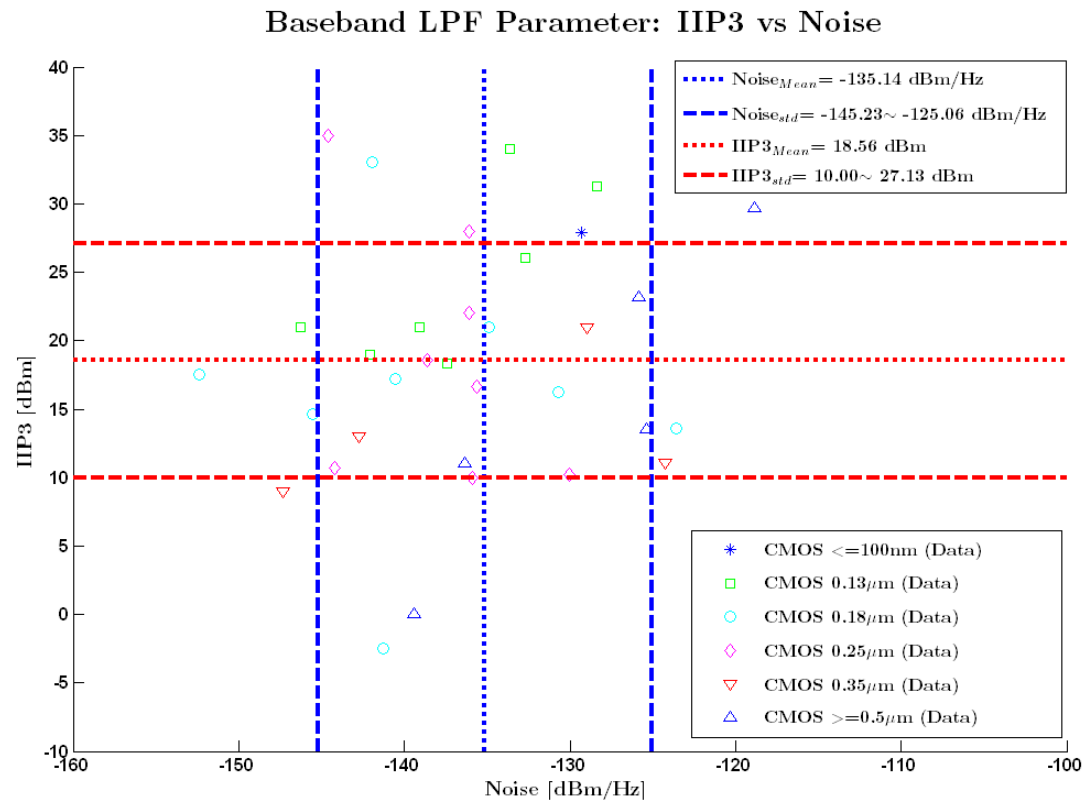

Baseband LPF Parameter: FOM vs Cut-off Frequency
FoM/Fc=1.18E+000 dB/dec
CMOS <=100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
FOM_LPF[dB] (Data)
Fc [Hz]

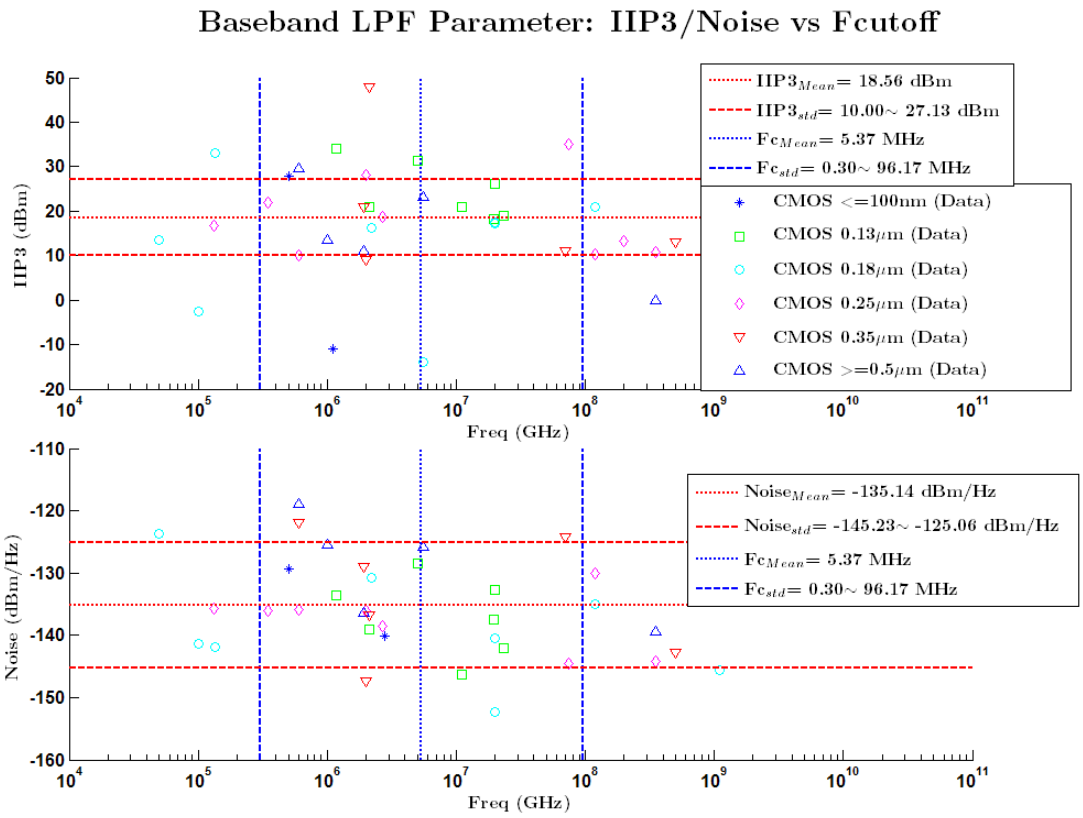

Baseband LPF Parameter: FOM vs Order
FoM/Order= -1.33E+000 dB/pole
CMOS <=100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
FOM_LPF[dB] (Data)
Order


Baseband LPF Parameter: IIP3 vs Year
IIP3/Year= 7.86E-001 dB/Year
CMOS <=100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
IIP3 [dBm]
Year

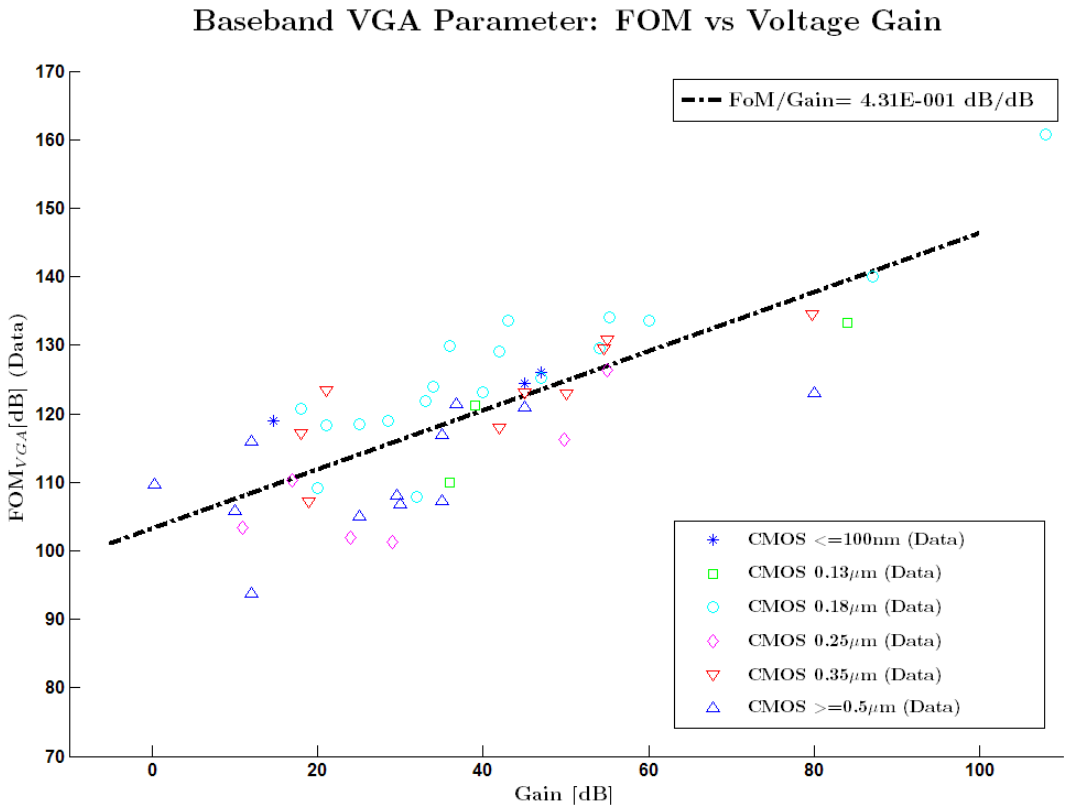

Baseband LPF Parameter: Input-Ref Noise vs Year
Noise/Year= -7.18E-001 dB/dB
CMOS <=100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
Noise [dBm/Hz]
Year

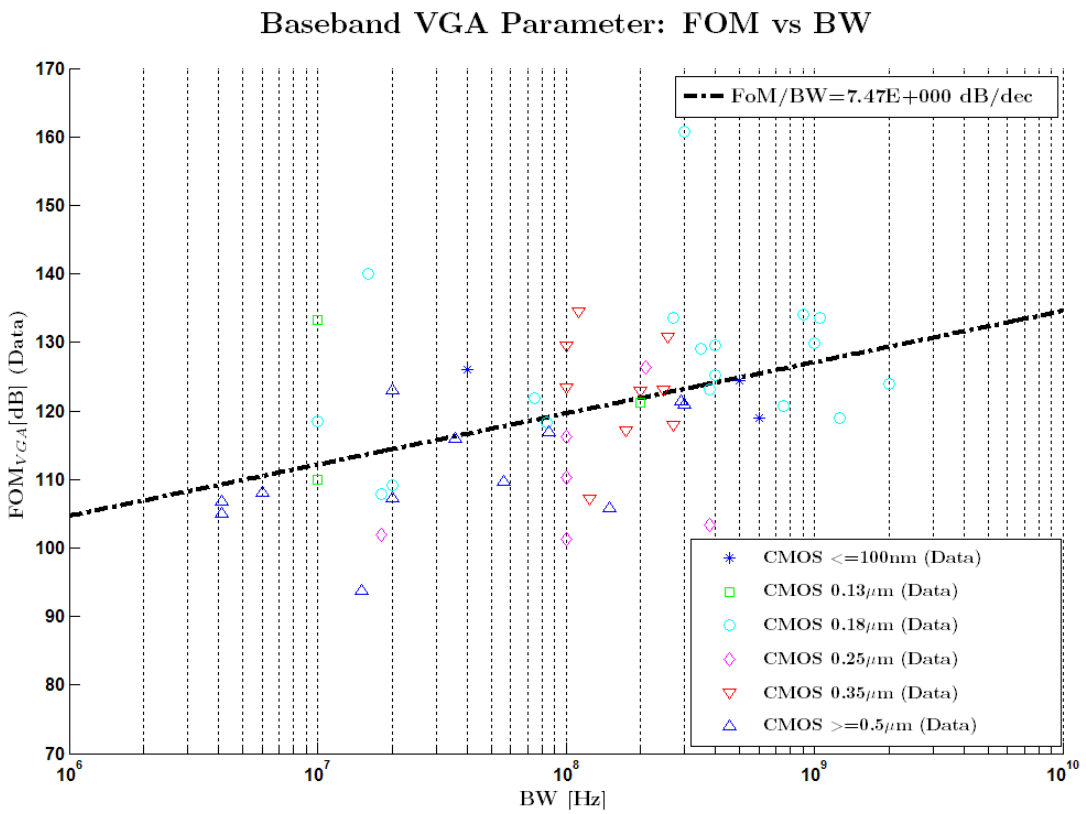

Baseband LPF Parameter: Cut-off Frequency vs Year
Fc/Year=-4.55E-003 Magnitude/Year
CMOS <=100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
Fc [Hz]
Year

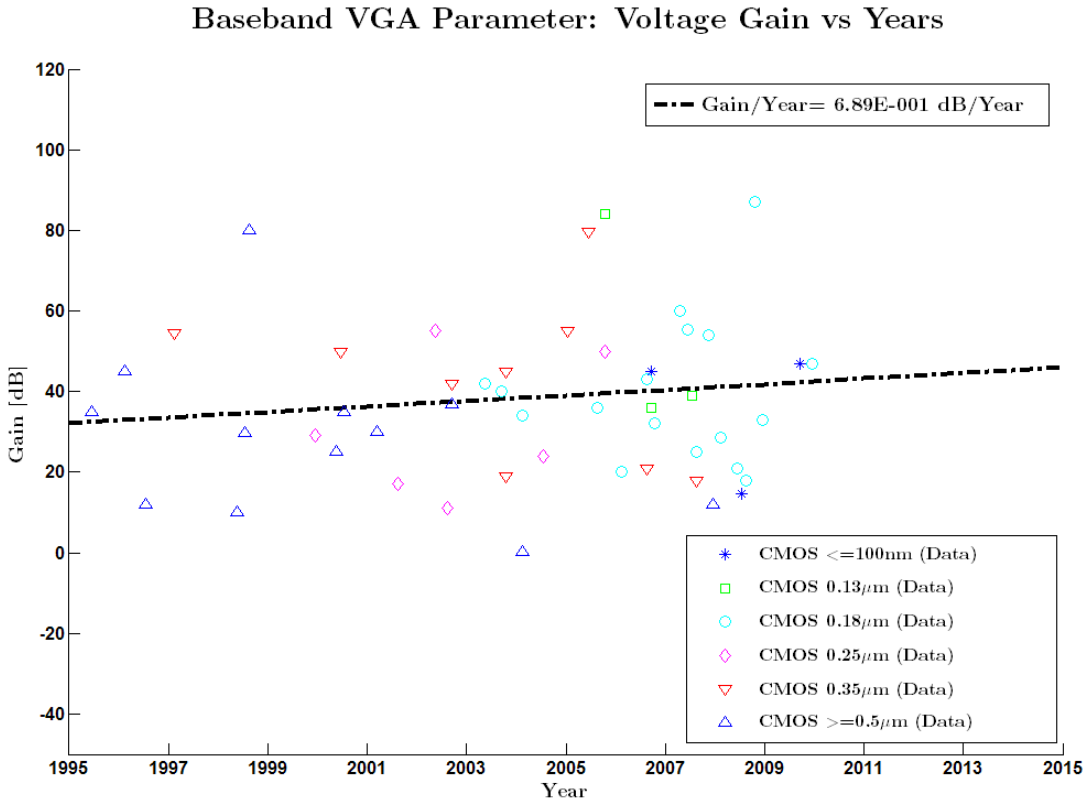

Baseband LPF Parameter: IIP3 vs Noise
Noise_Mean= -135.14 dBm/Hz
Noise_std= -145.23~ -125.06 dBm/Hz
IIP3_Mean= 18.56 dBm
IIP3_std= 10.00~ 27.13 dBm
CMOS <=100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
IIP3 [dBm]
Noise [dBm/Hz]

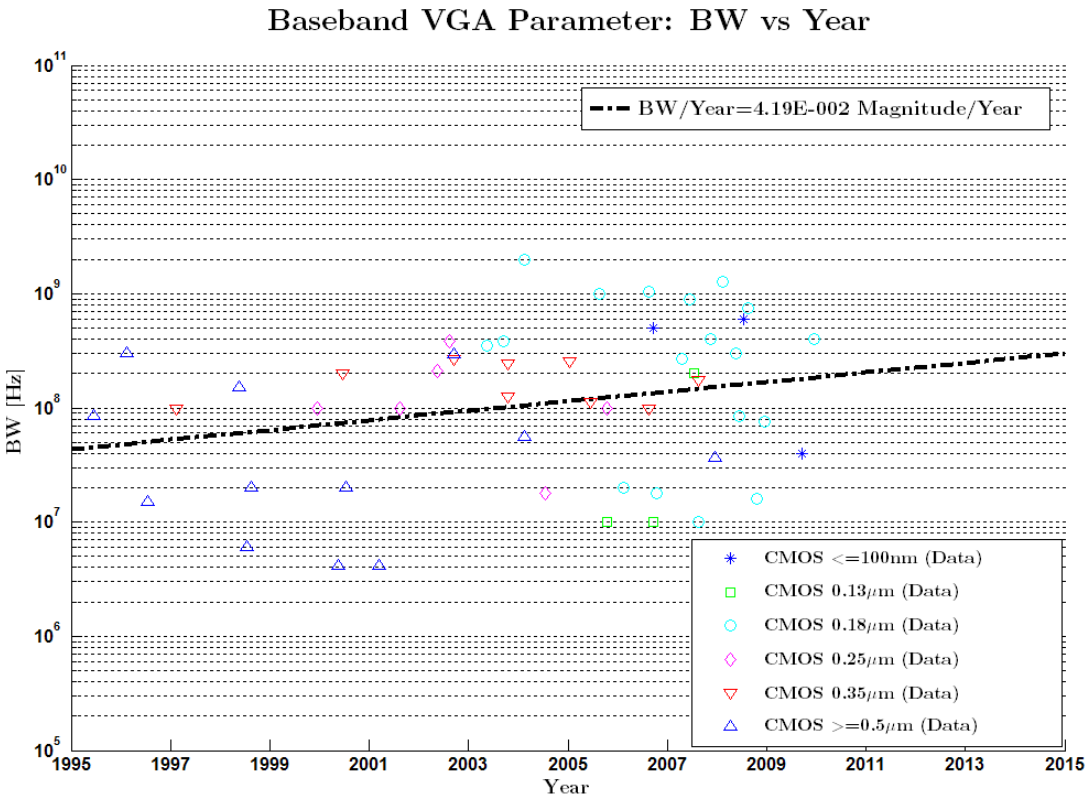


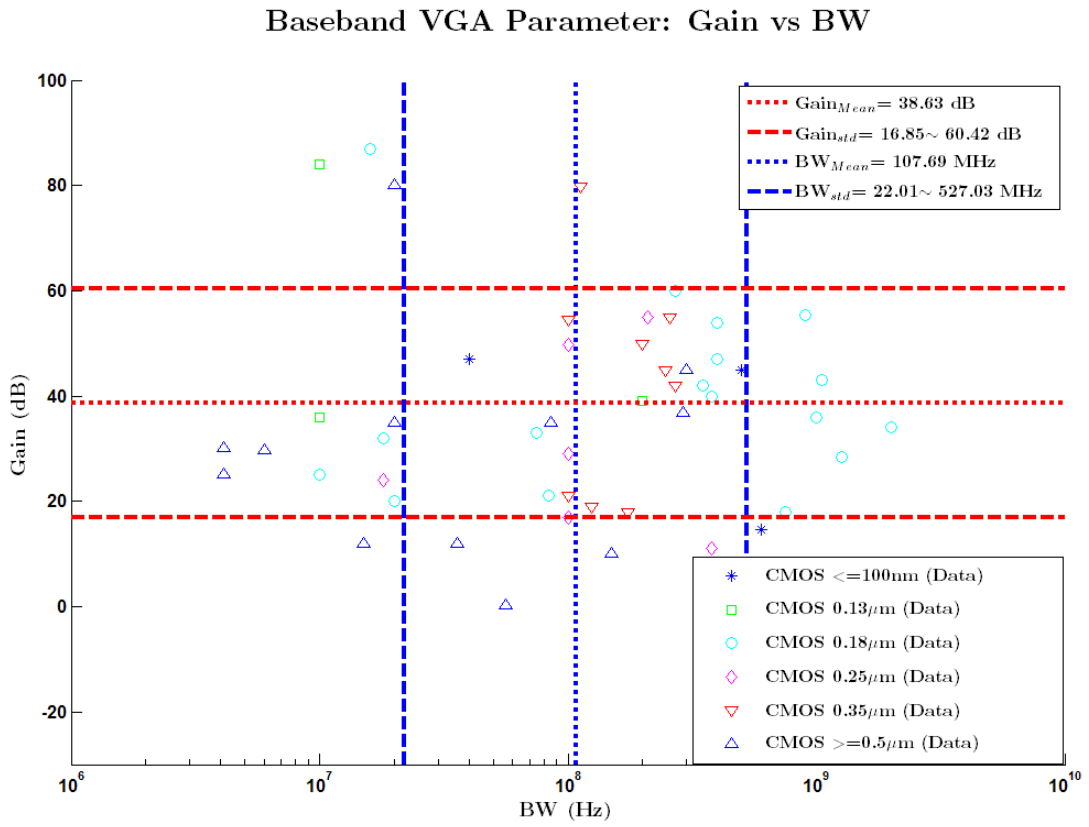
A-10 VGA FoM statistics



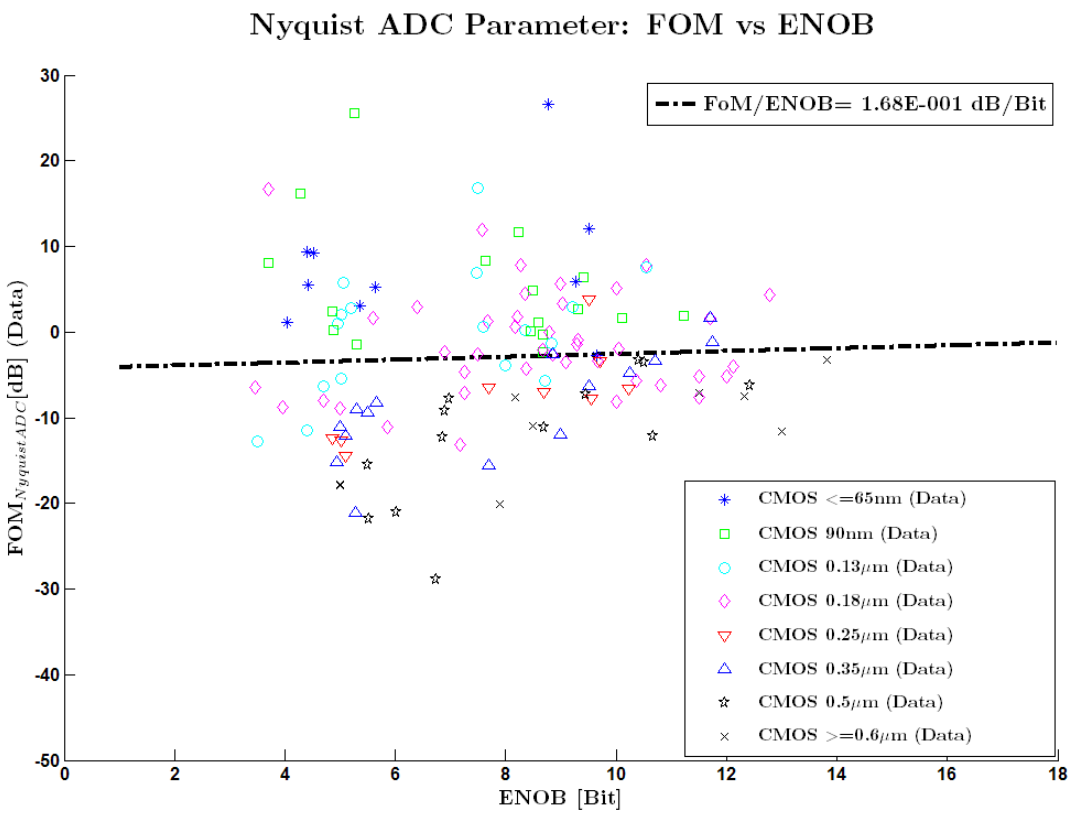

Baseband VGA Parameter: FOM vs BW
FoM/BW=7.47E+000 dB/dec
CMOS <=100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
FOM_VGA[dB] (Data)
BW [Hz]

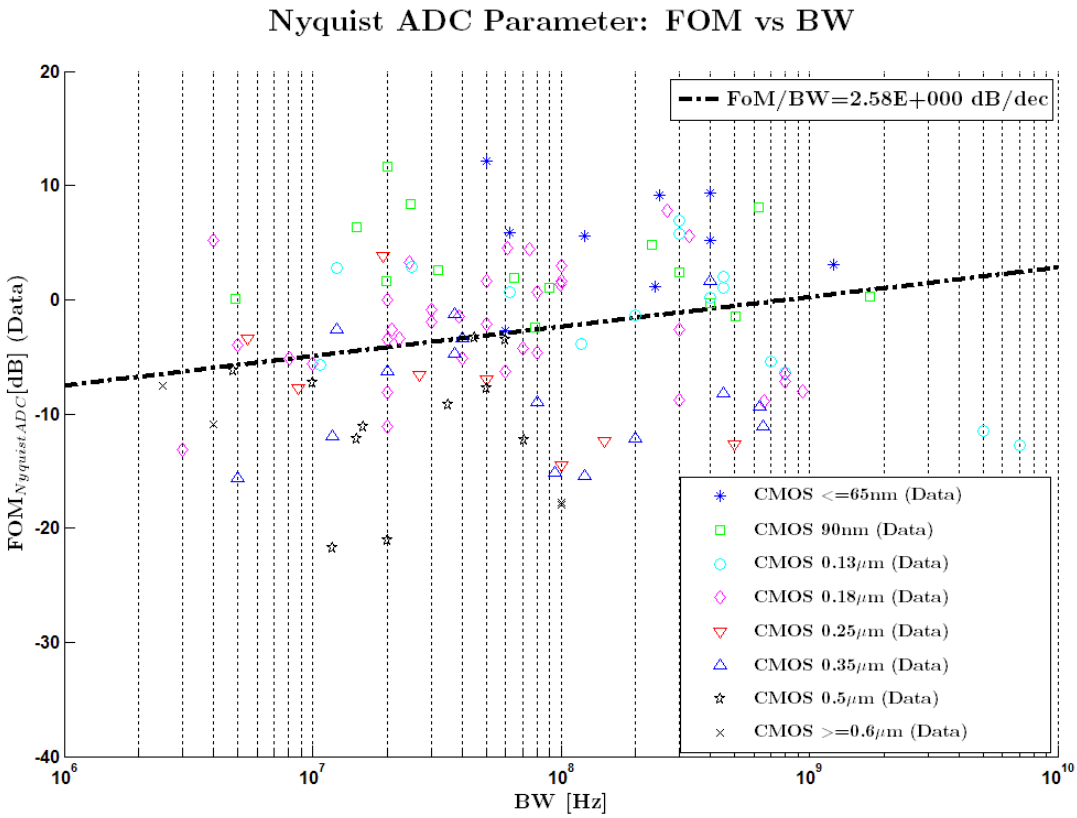

Baseband VGA Parameter: Voltage Gain vs Years
Gain/Year= 6.89E-001 dB/Year
CMOS <=100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
Gain [dB]
Year

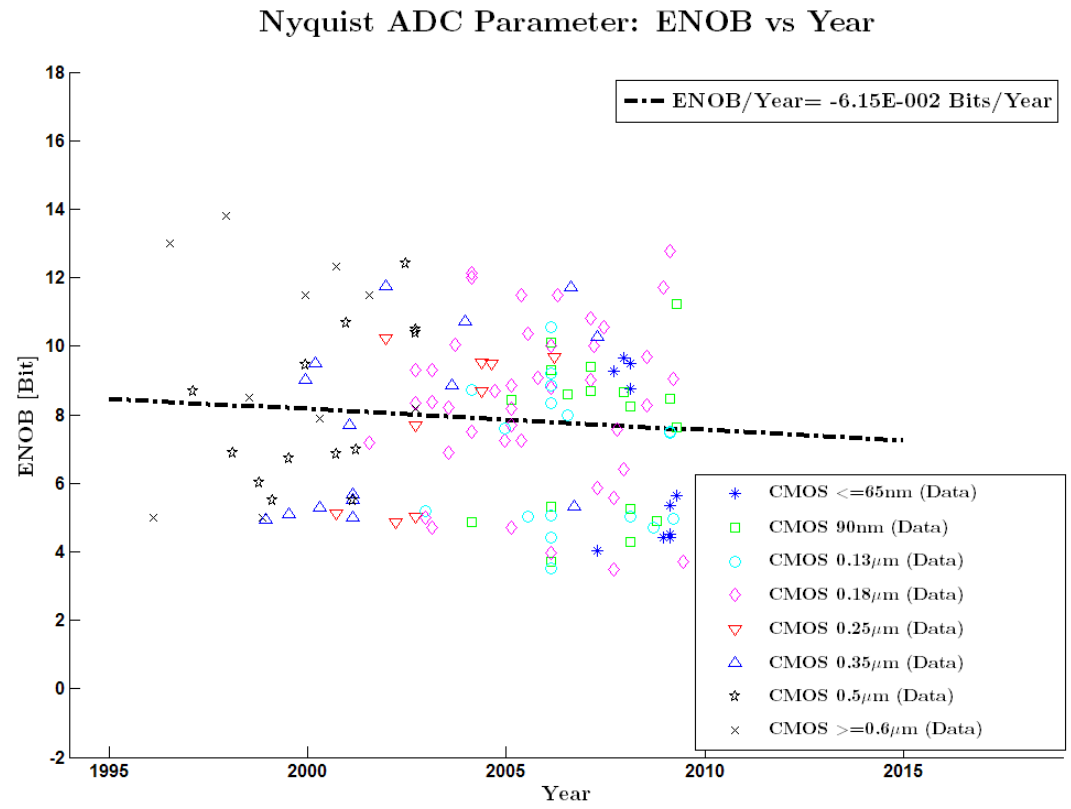

Baseband VGA Parameter: BW vs Year
BW/Year=4.19E-002 Magnitude/Year
CMOS <=100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
BW [Hz]
Year

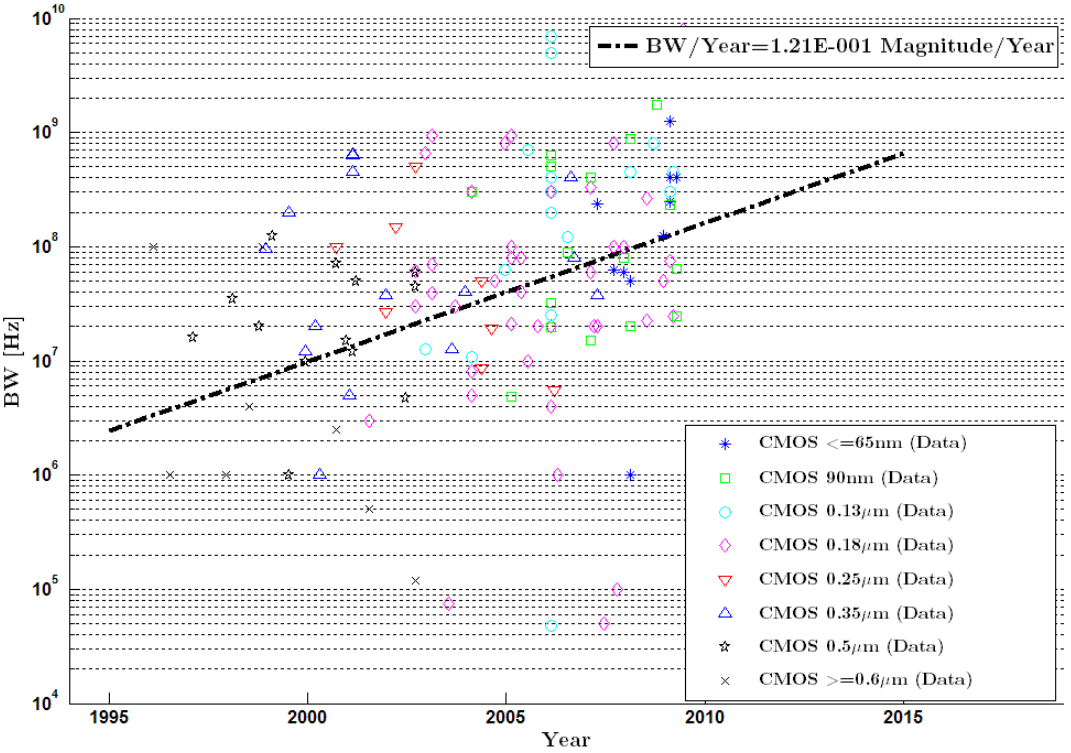

Baseband VGA Parameter: Gain vs BW
Gain_Mean= 38.63 dB
Gain_std= 16.85~ 60.42 dB
BW_Mean= 107.69 MHz
BW_std= 22.01~ 527.03 MHz
CMOS <=100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS >=0.5μm (Data)
Gain (dB)
BW (Hz)

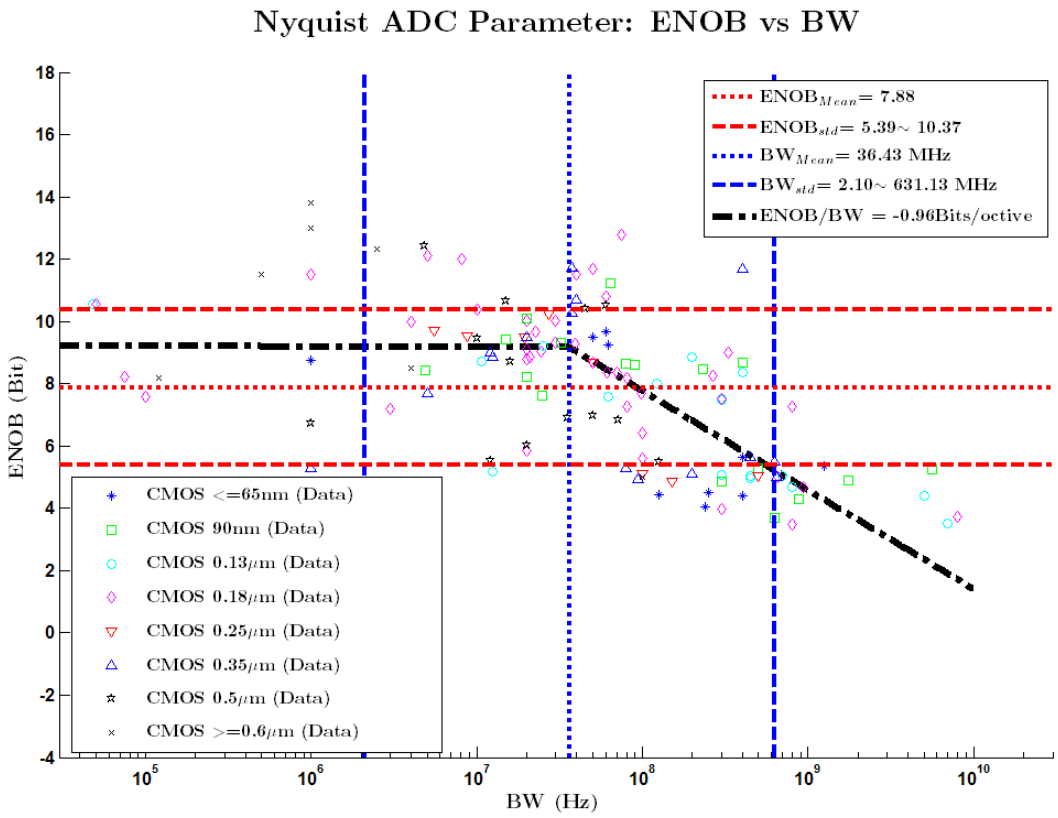
A-11 Nyquist ADC FoM Statistics



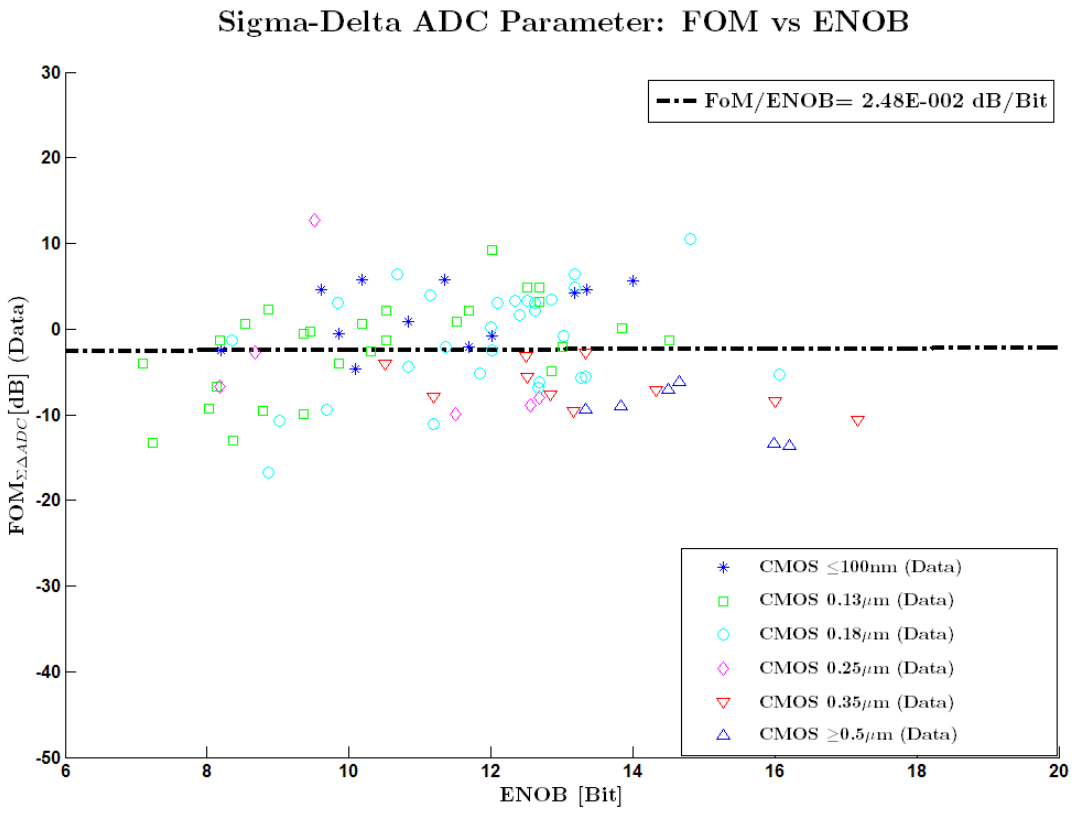


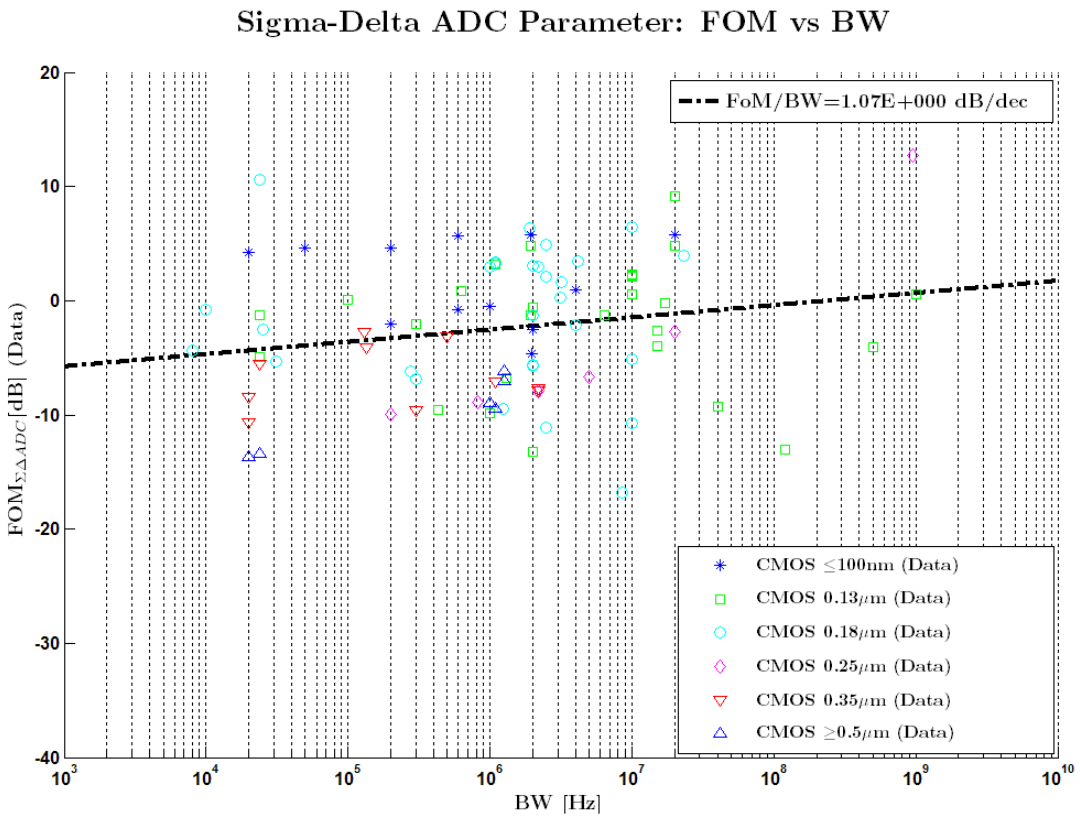

Nyquist ADC Parameter: ENOB vs Year
ENOB/Year= -6.15E-002 Bits/Year
ENOB [Bit]
Year
CMOS <=65nm (Data)
CMOS 90nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS 0.5μm (Data)
CMOS >=0.6μm (Data)

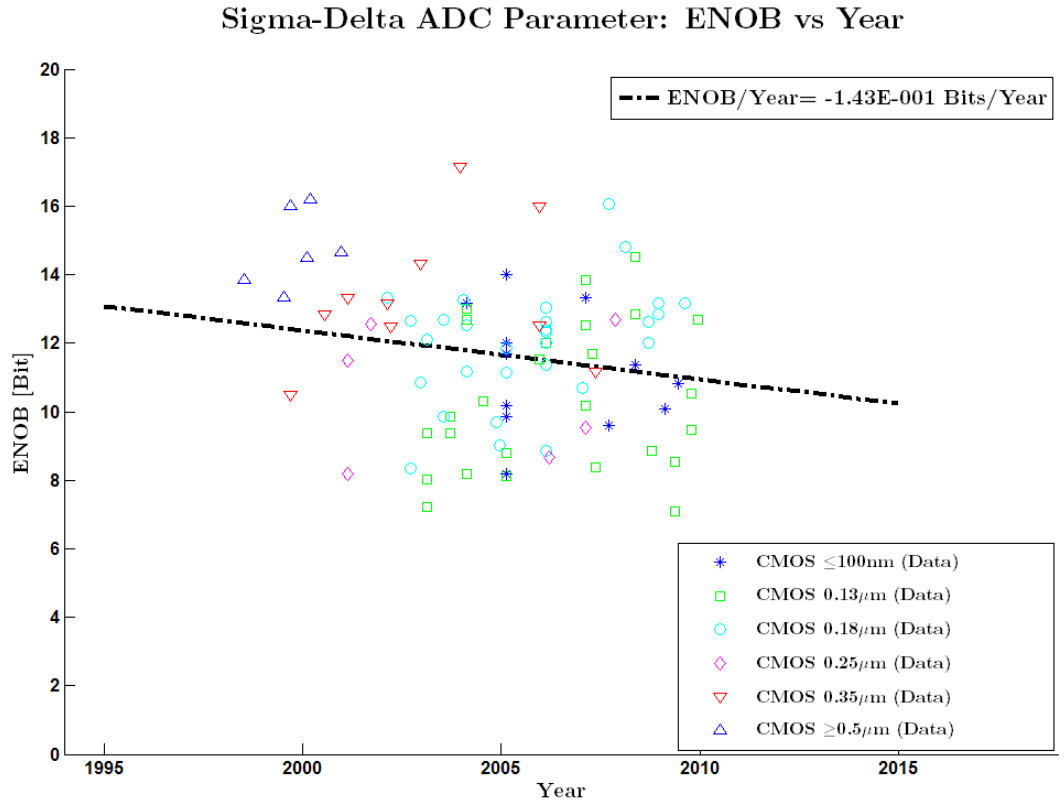

Nyquist ADC Paraameter: BW vs Year
BW/Year=1.21E-001 Magnitude/Year
BW [Hz]
Year
CMOS <=65nm (Data)
CMOS 90nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS 0.5μm (Data)
CMOS >=0.6μm (Data)

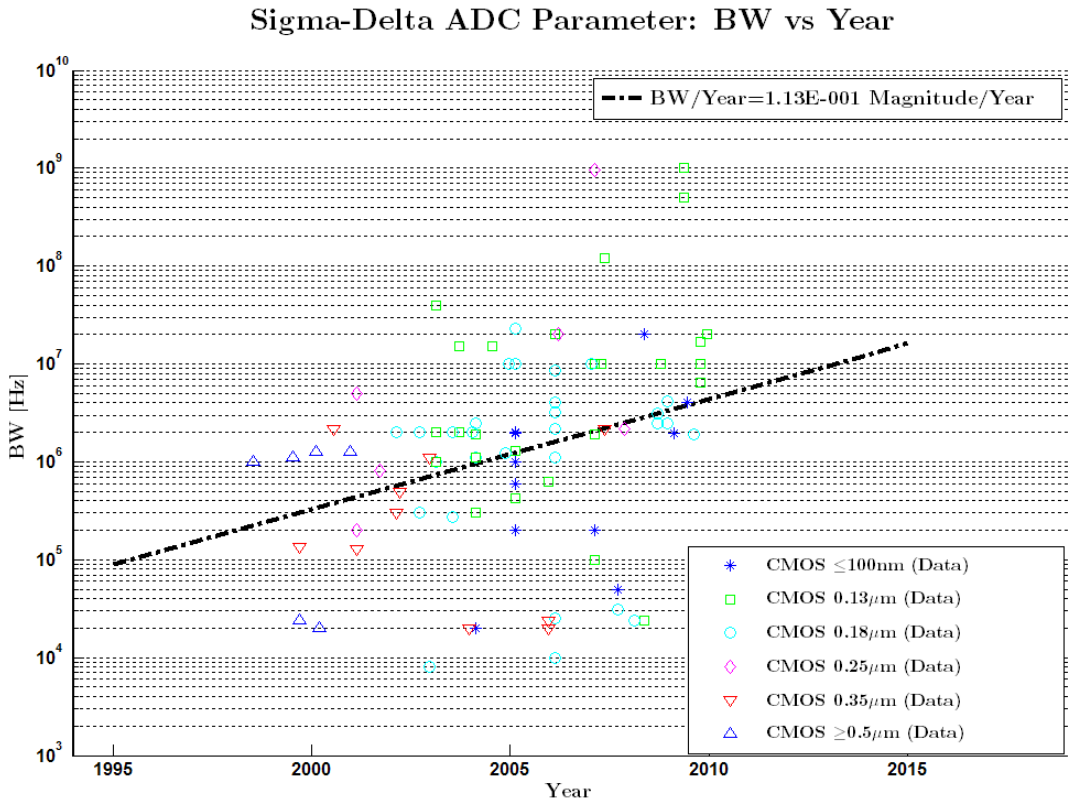


A-12 Sigma-Delta ADC FoM Statistics




Sigma-Delta ADC Parameter: FOM vs BW
FoM/BW=1.07E+000 dB/dec
CMOS ≤100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS ≥0.5μm (Data)
FOM_SΔADC [dB] (Data)
BW [Hz]

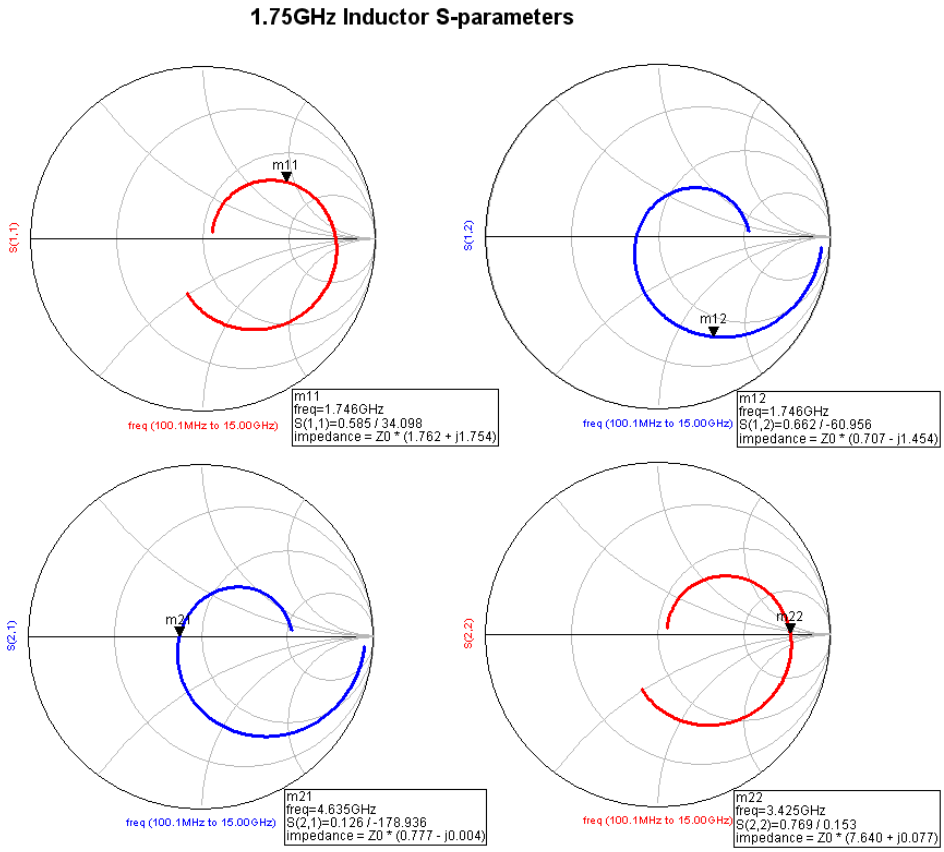

Sigma-Delta ADC Parameter: ENOB vs Year
ENOB/Year= -1.43E-001 Bits/Year
CMOS ≤100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS ≥0.5μm (Data)
ENOB [Bit]
Year


Sigma-Delta ADC Parameter: BW vs Year
BW/Year=1.13E-001 Magnitude/Year
CMOS ≤100nm (Data)
CMOS 0.13μm (Data)
CMOS 0.18μm (Data)
CMOS 0.25μm (Data)
CMOS 0.35μm (Data)
CMOS ≥0.5μm (Data)
BW [Hz]
Year
10^10
10^9
10^8
10^7
10^6
10^5
10^4
10^3
1995
2000
2005
2010
2015

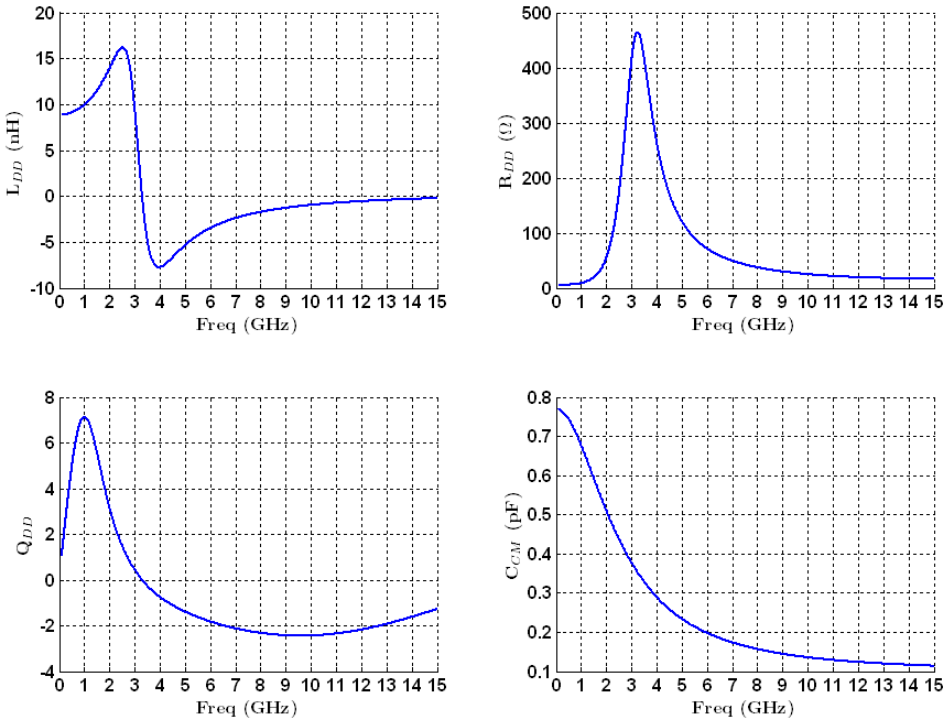
Appendix B

This appendix includes the simulation and measurement results of the BPF in chapter five.

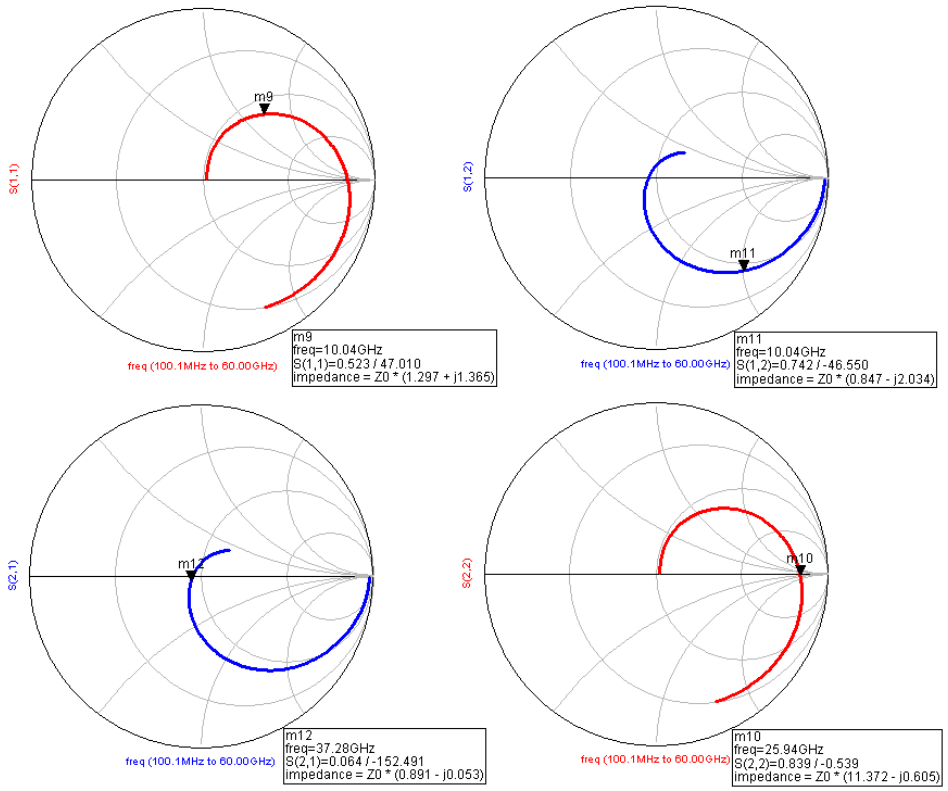
B-1 S-parameter Simulation Results of Single-ended Inductors

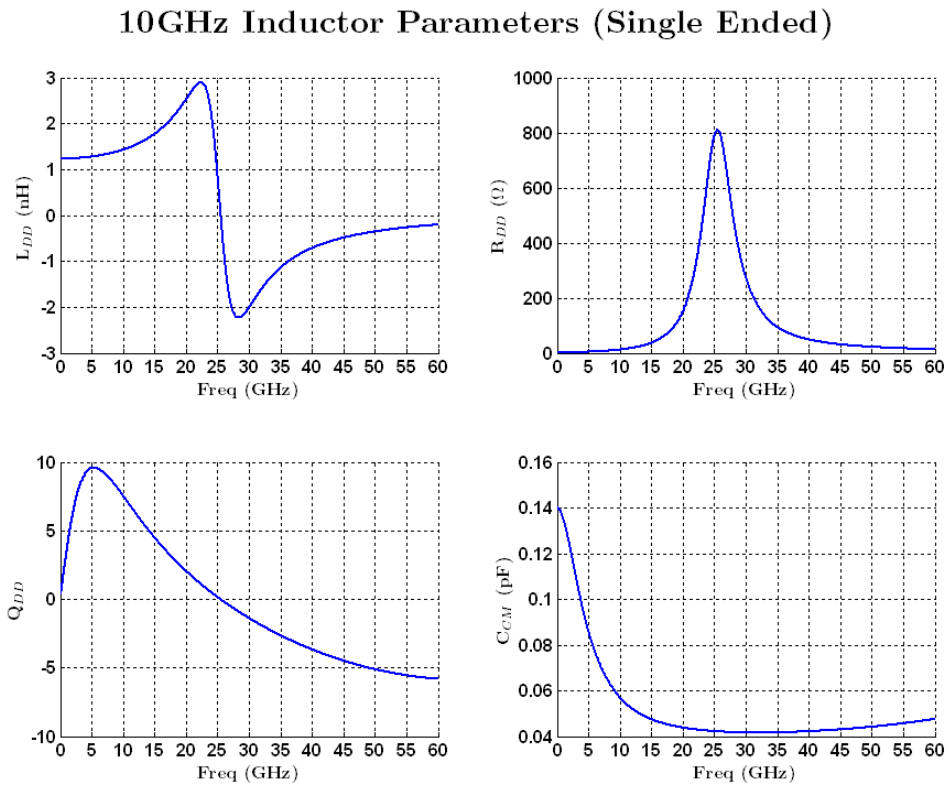


1.75GHz Inductor Parameters (Single Ended)



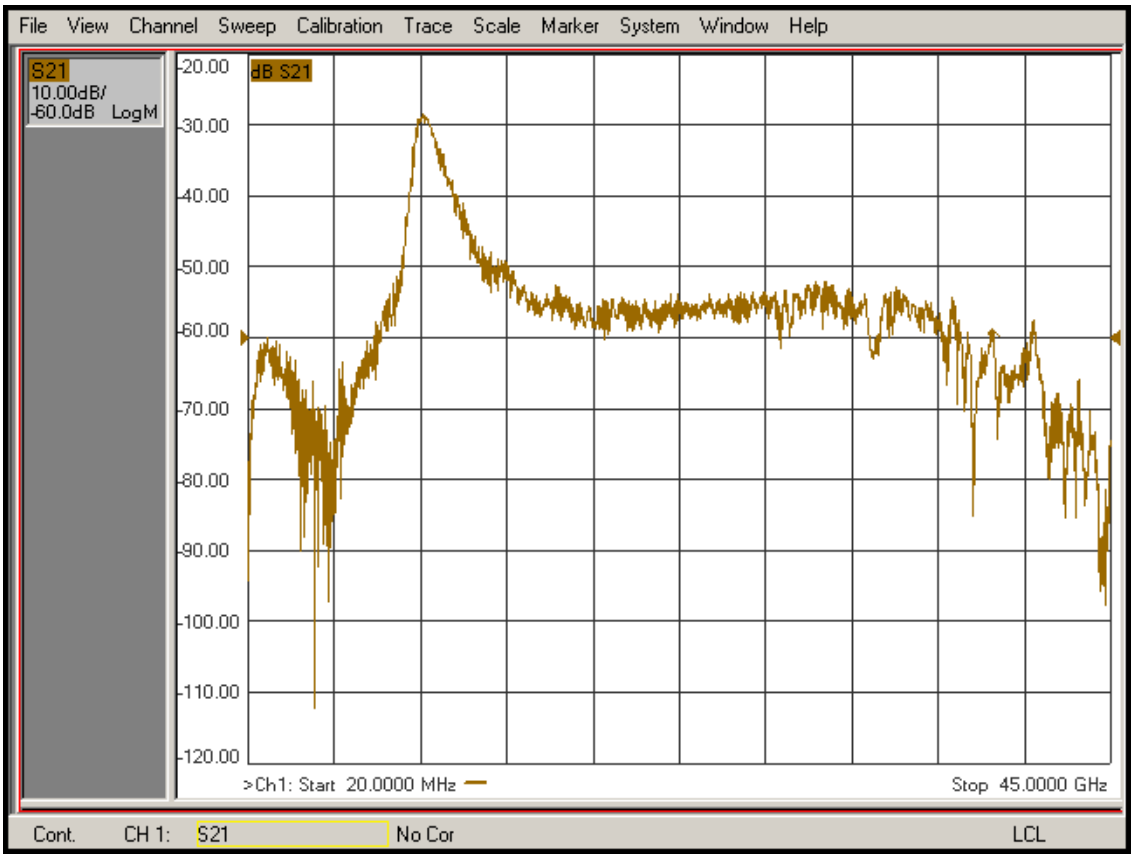
10GHz Inductor S-parameters



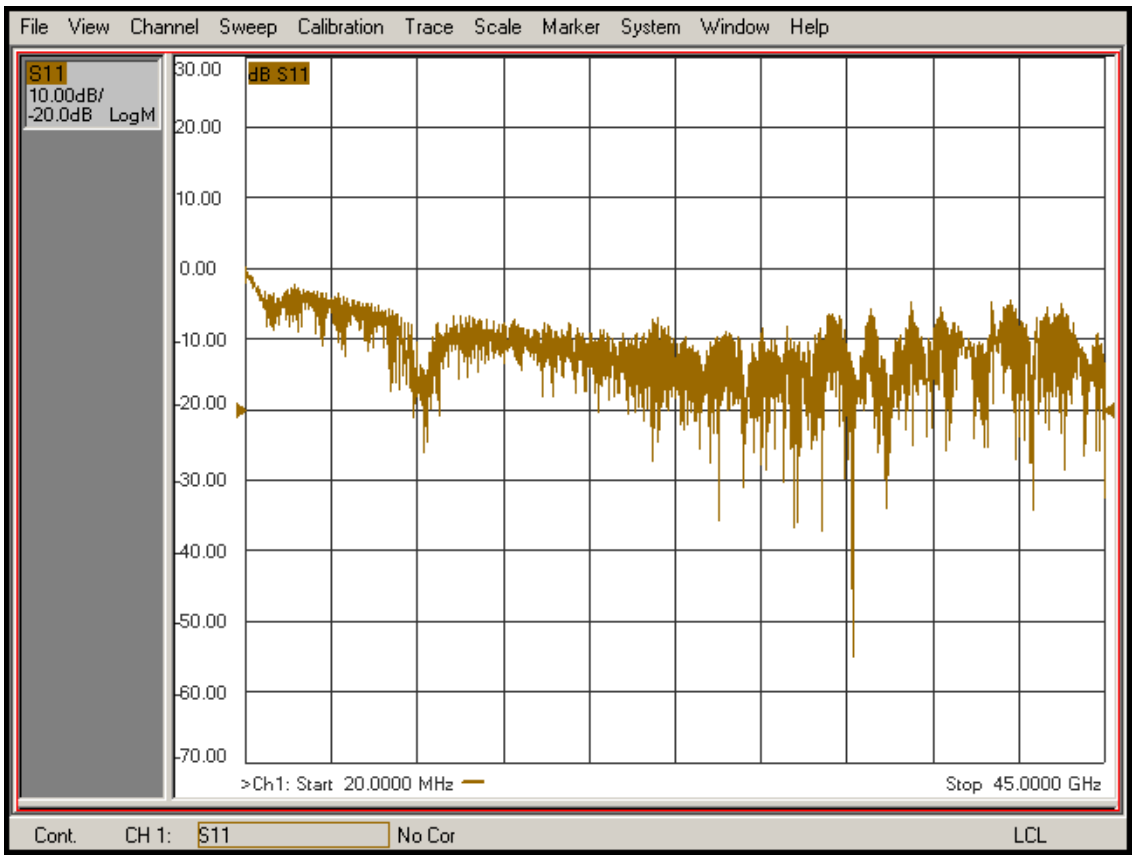


B-2 Measurement Raw Data of 10GGHz BPF

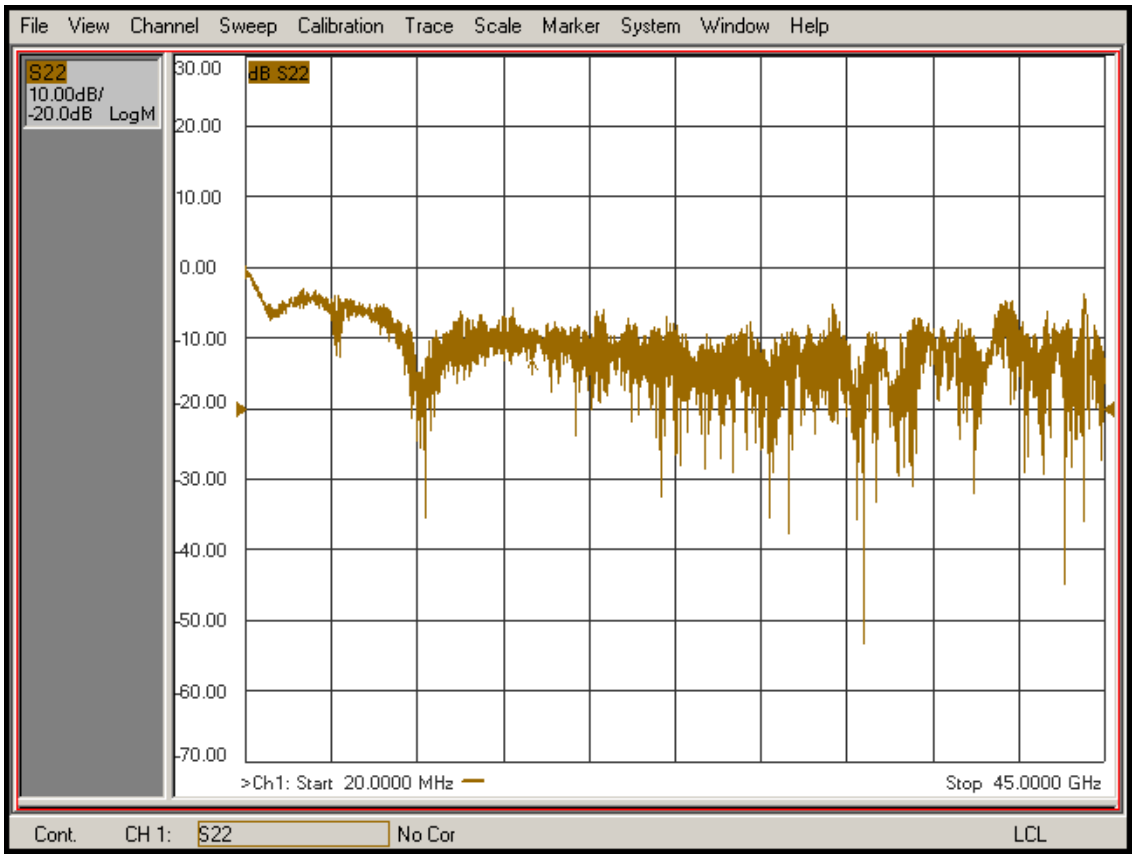
10GHz BPF measured raw data (S_{21})



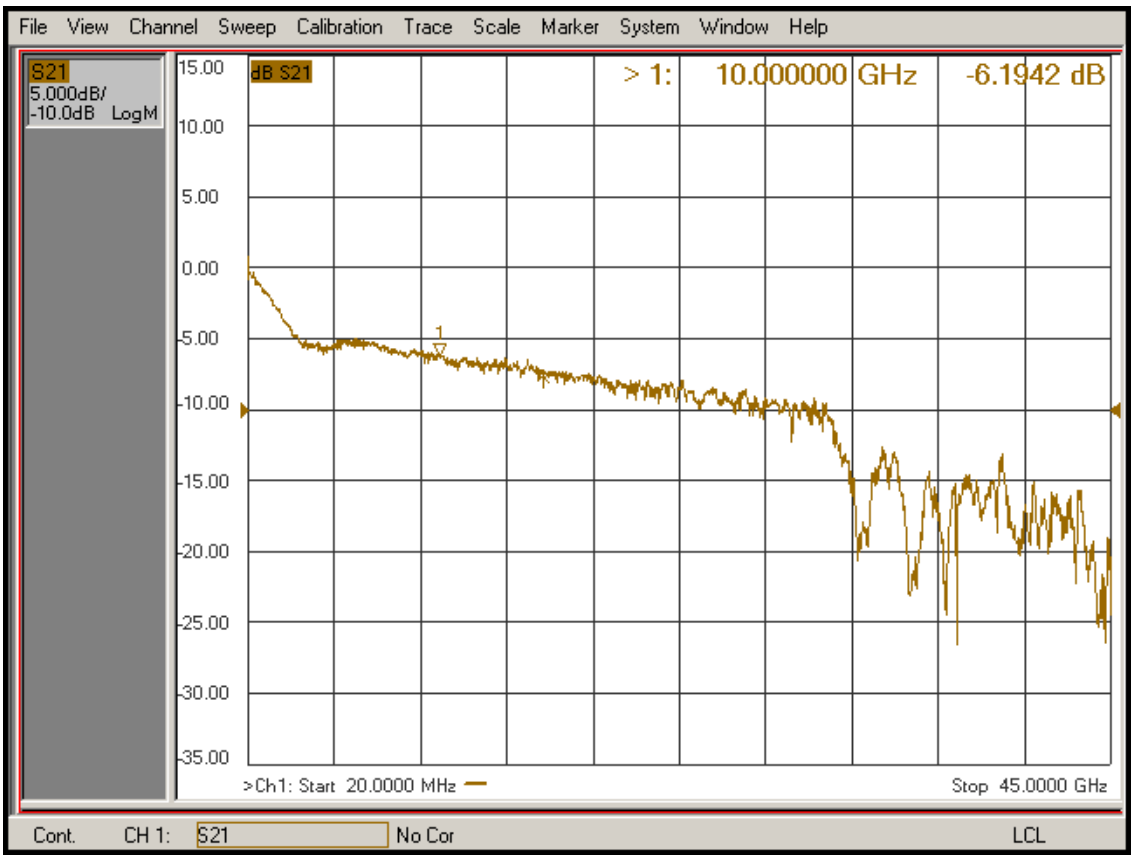
10GHz BPF measured raw data (S_{11})



10GHz BPF measured raw data (S_{22})

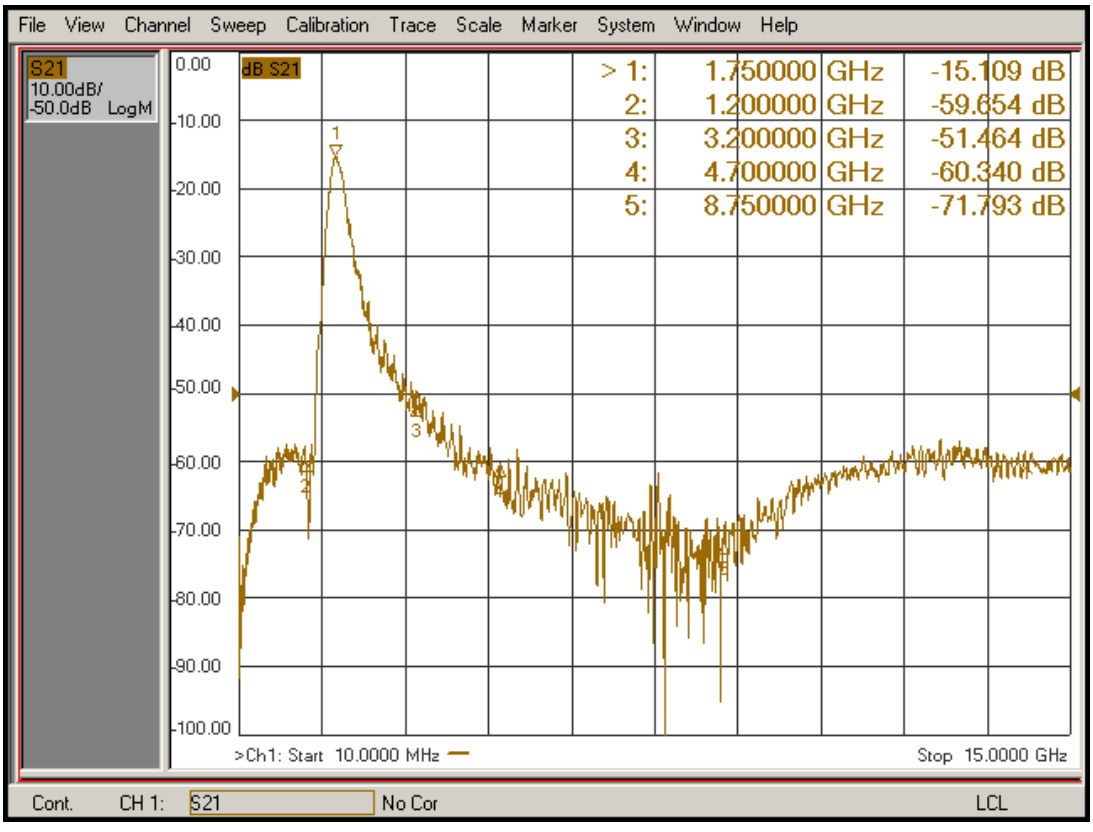


Cable loss of 20MHz ~ 45GHz (S_{22})

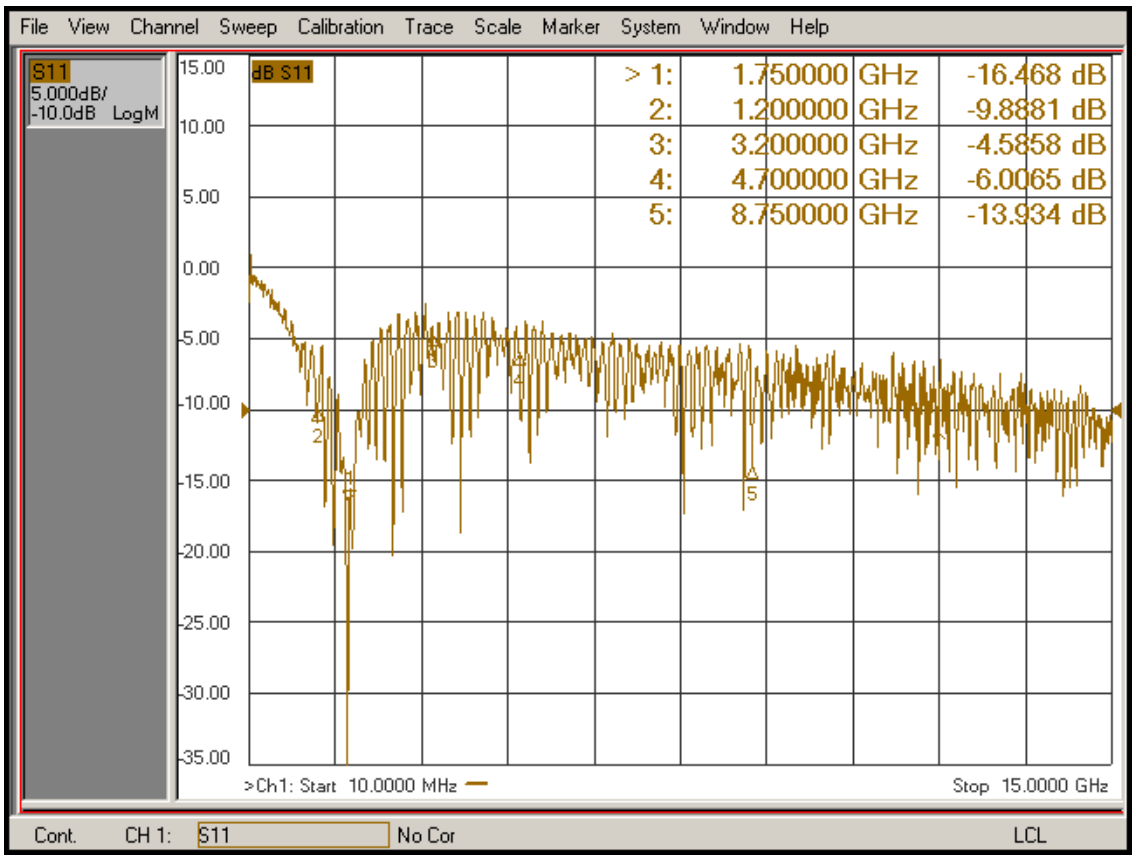


B-3 Measurement Raw Data of 1.75GHz BPF

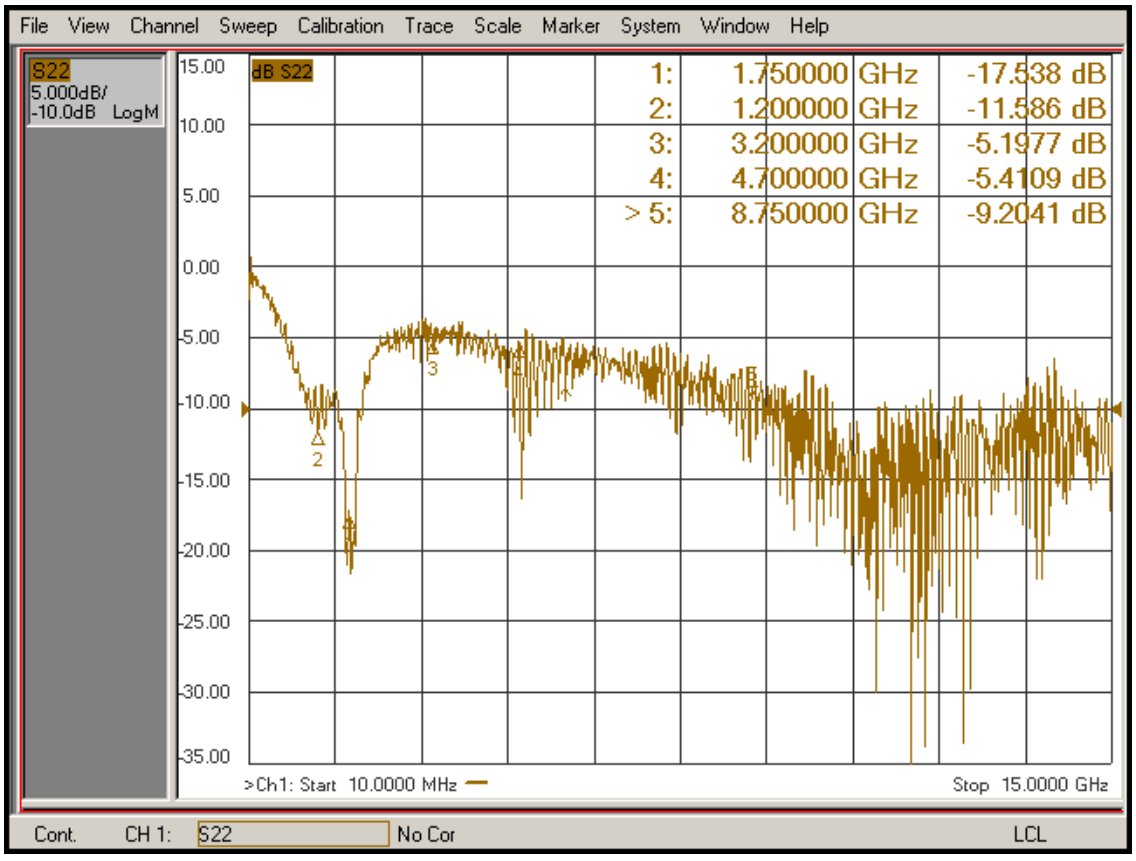
1.75GHz BPF measured raw data (S_{21})



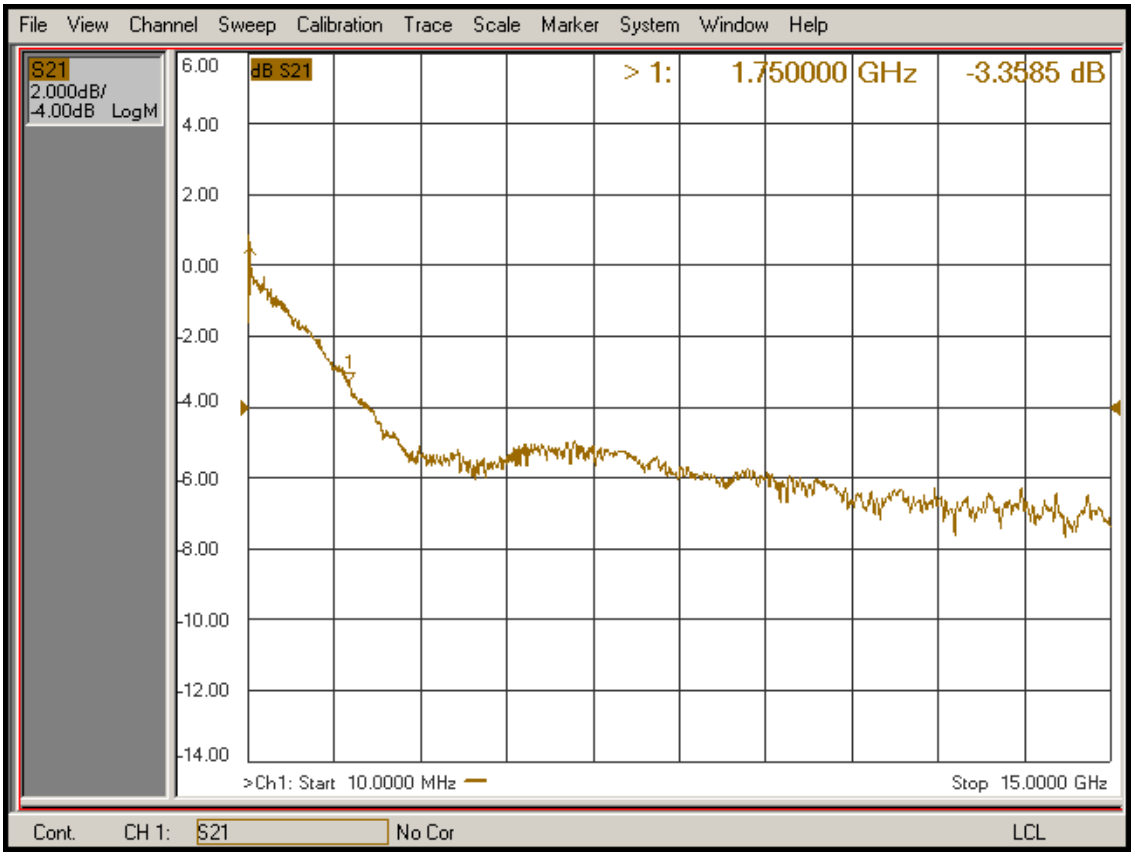
1.75GHz BPF measured raw data (S_{11})



1.75GHz BPF measured raw data (S_{22})



Cable loss 10MHz ~ 15GHz (S_{21})

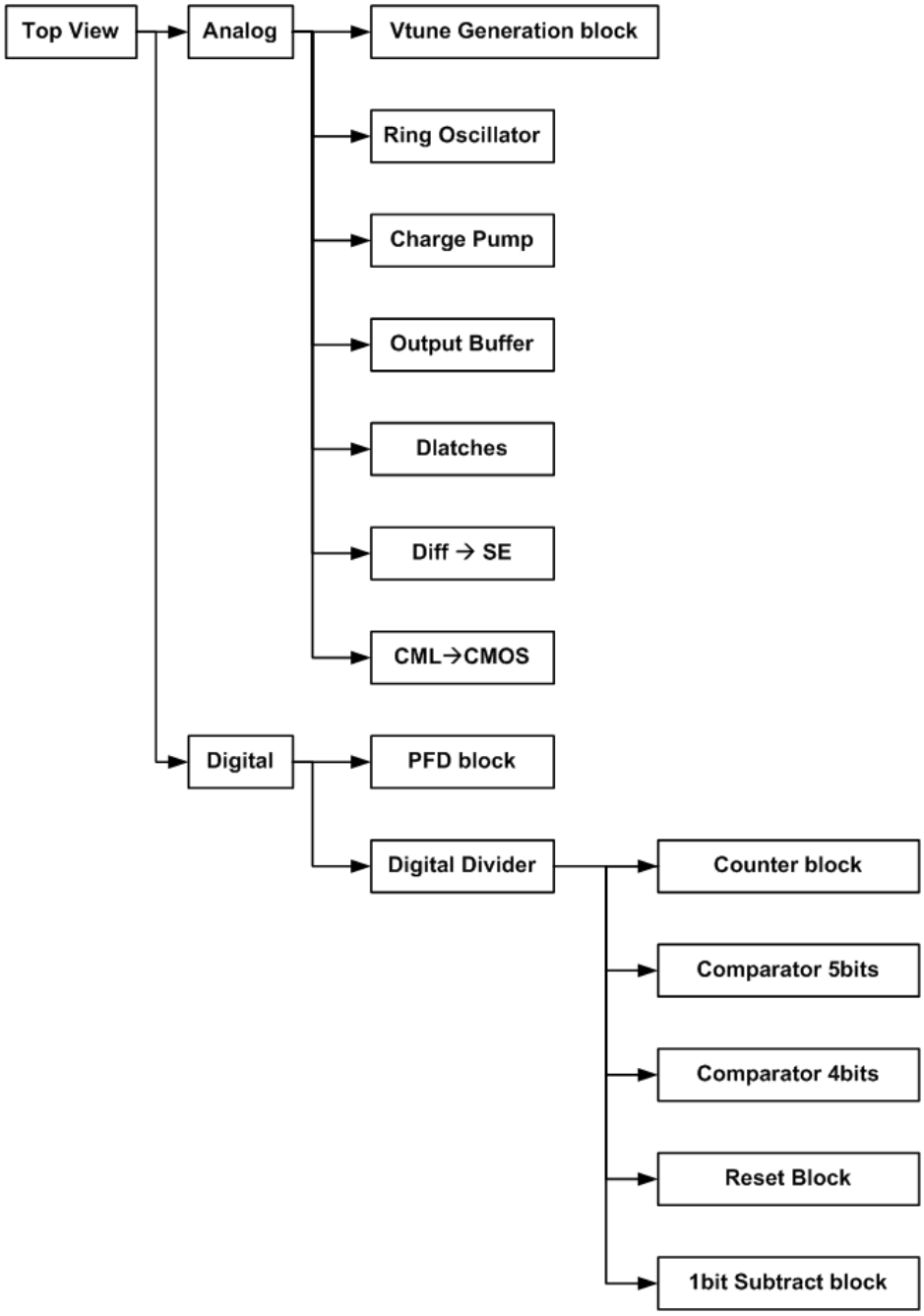


Appendix C

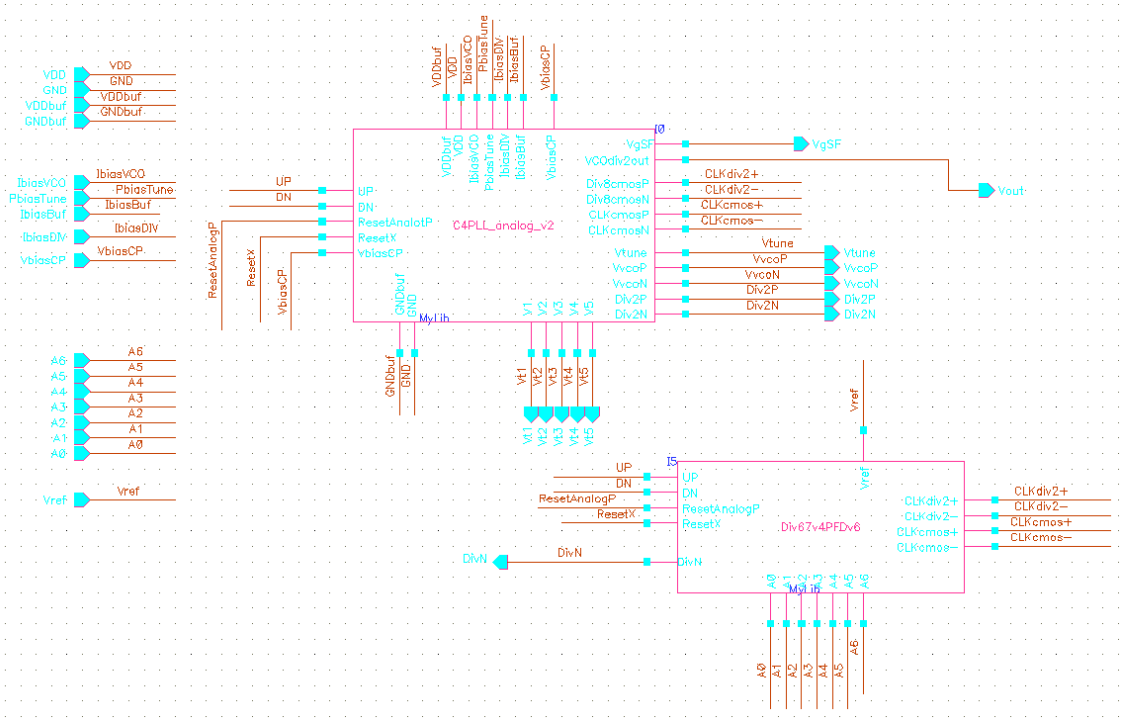
This appendix includes the figures related to frequency synthesizer in chapter six.

C-1 Schematics

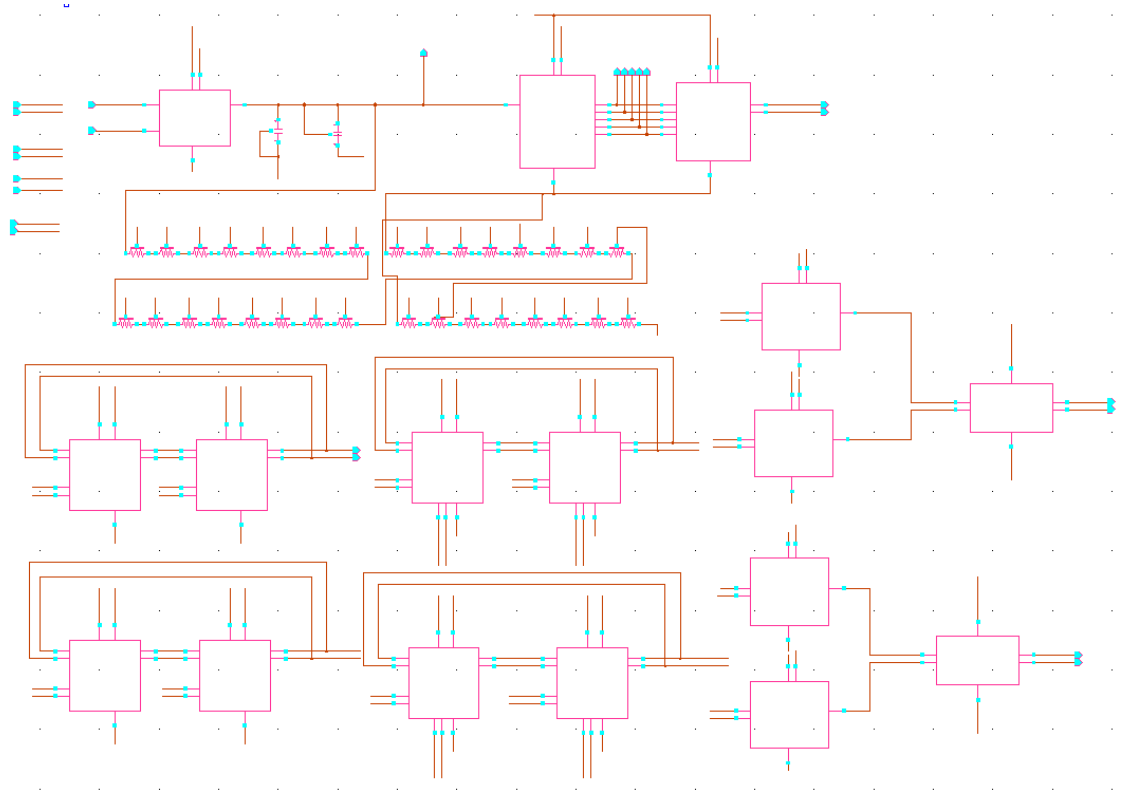
Schematic structure (Tree)



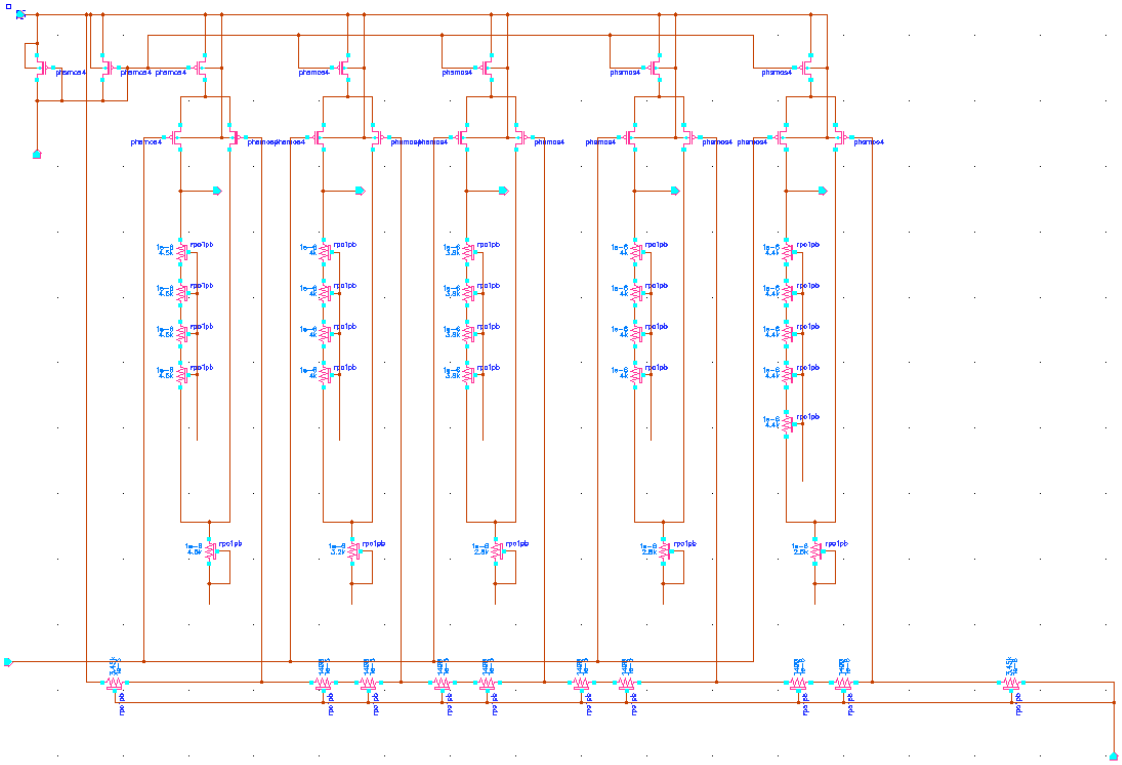
PLL Top Cell



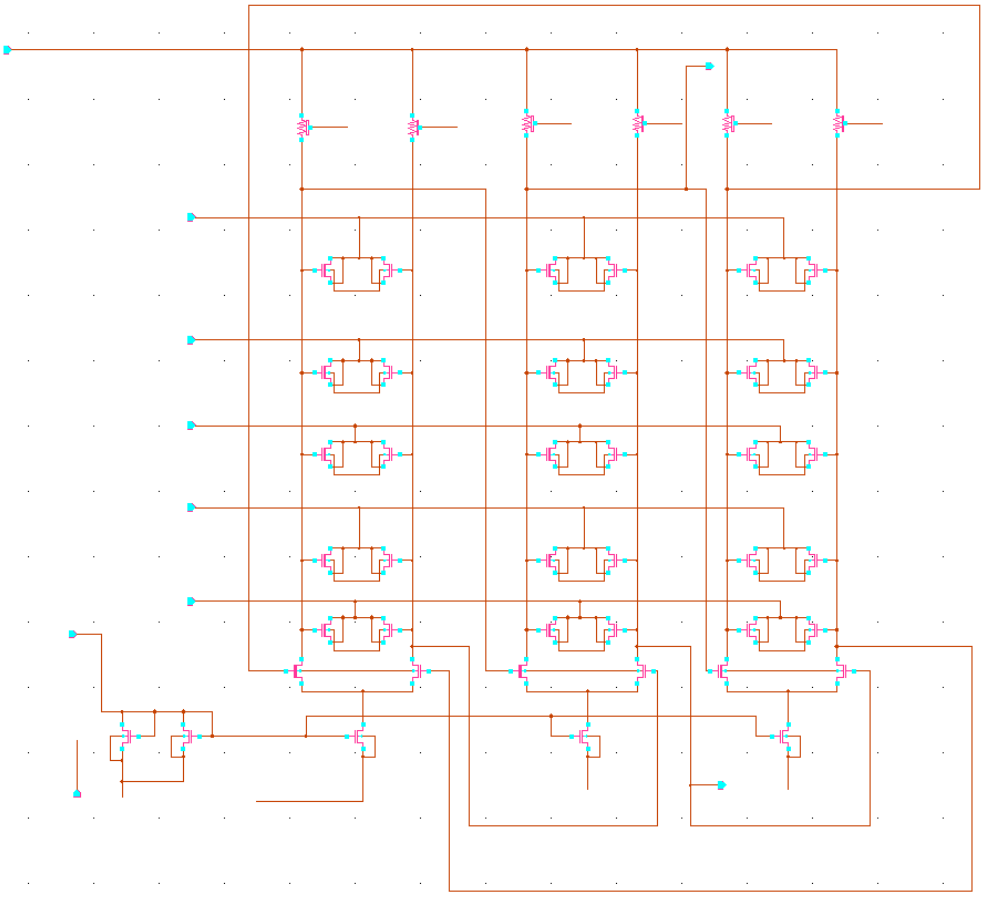
Analog block top view



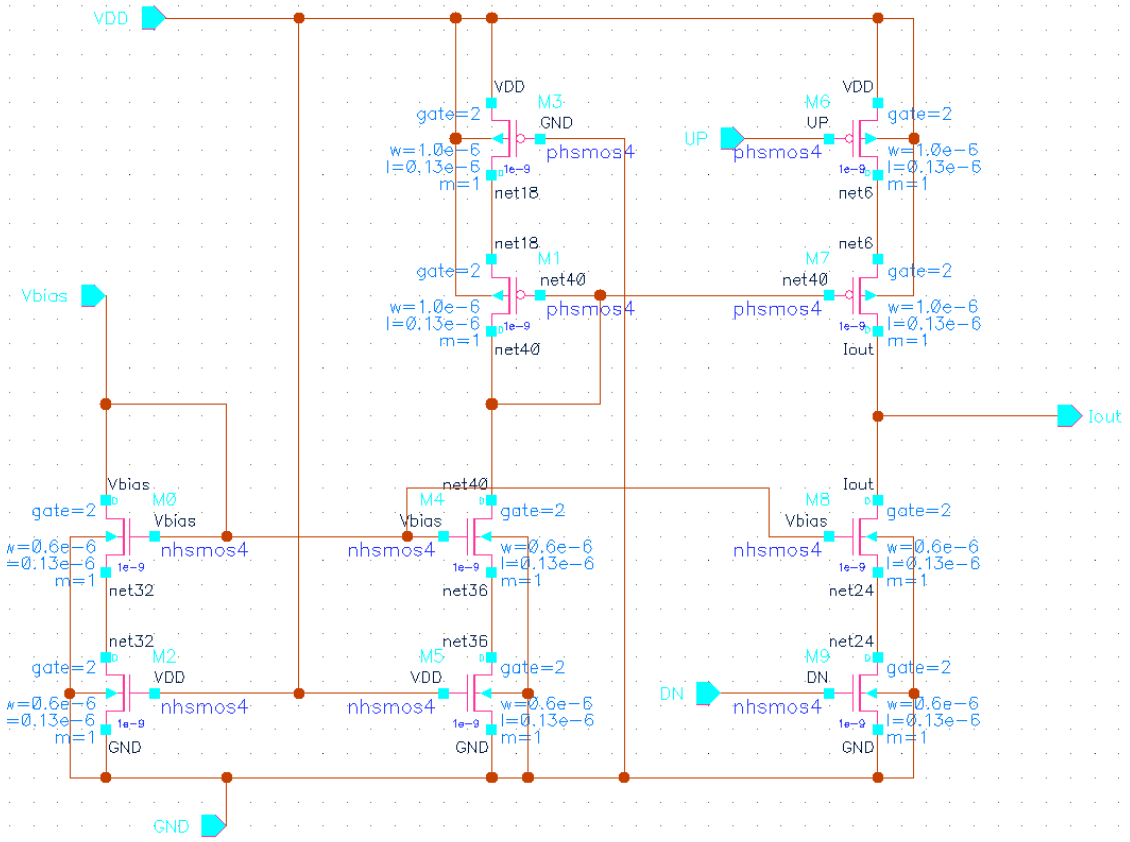
PLL tuning voltage generation block



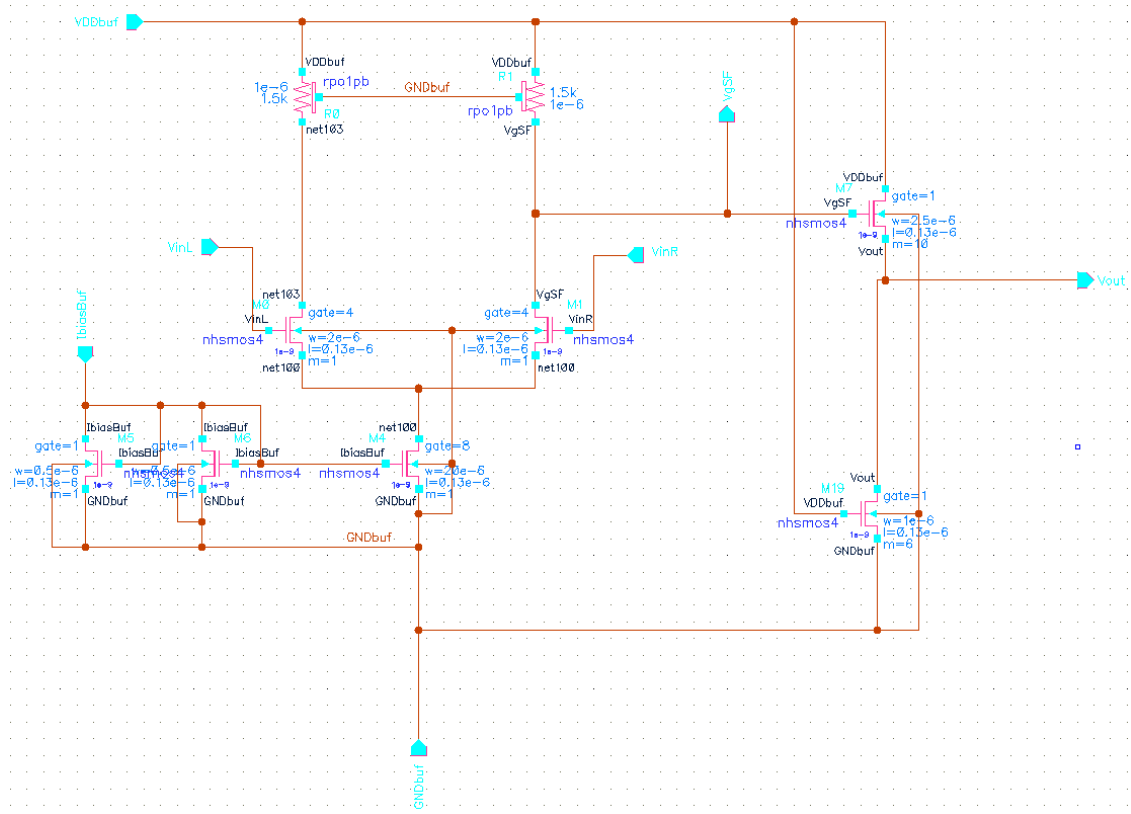
Ring oscillator block



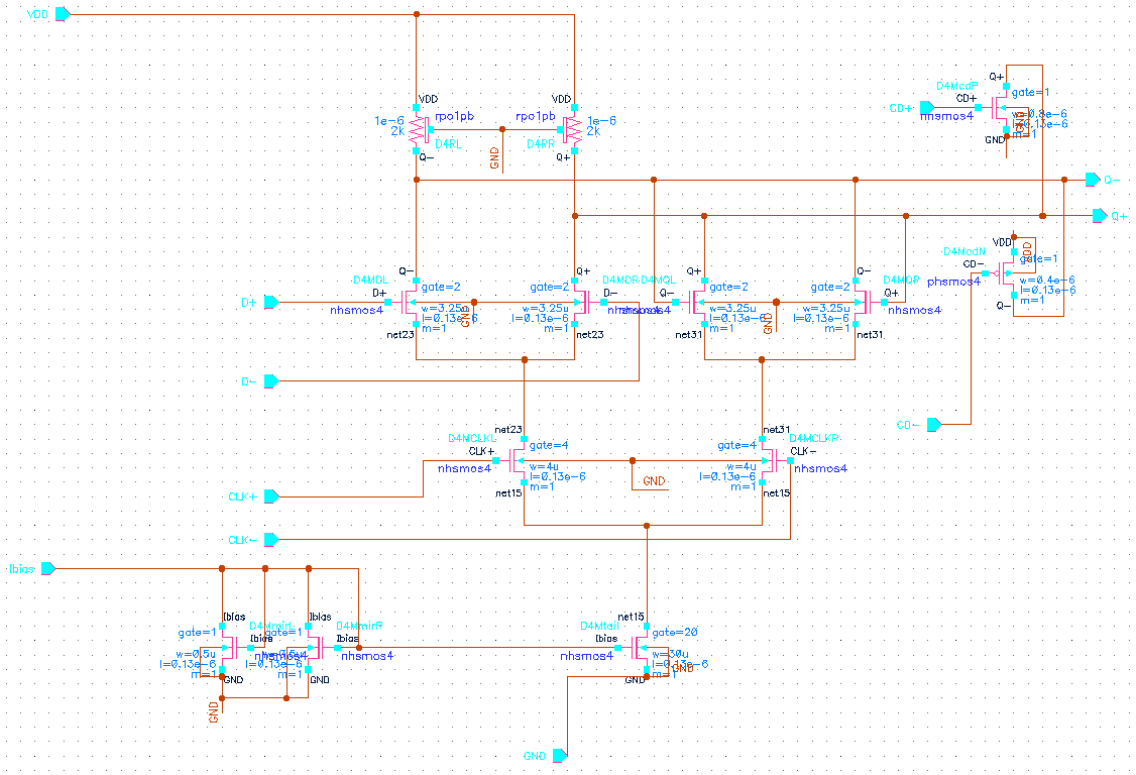
Charge pump block



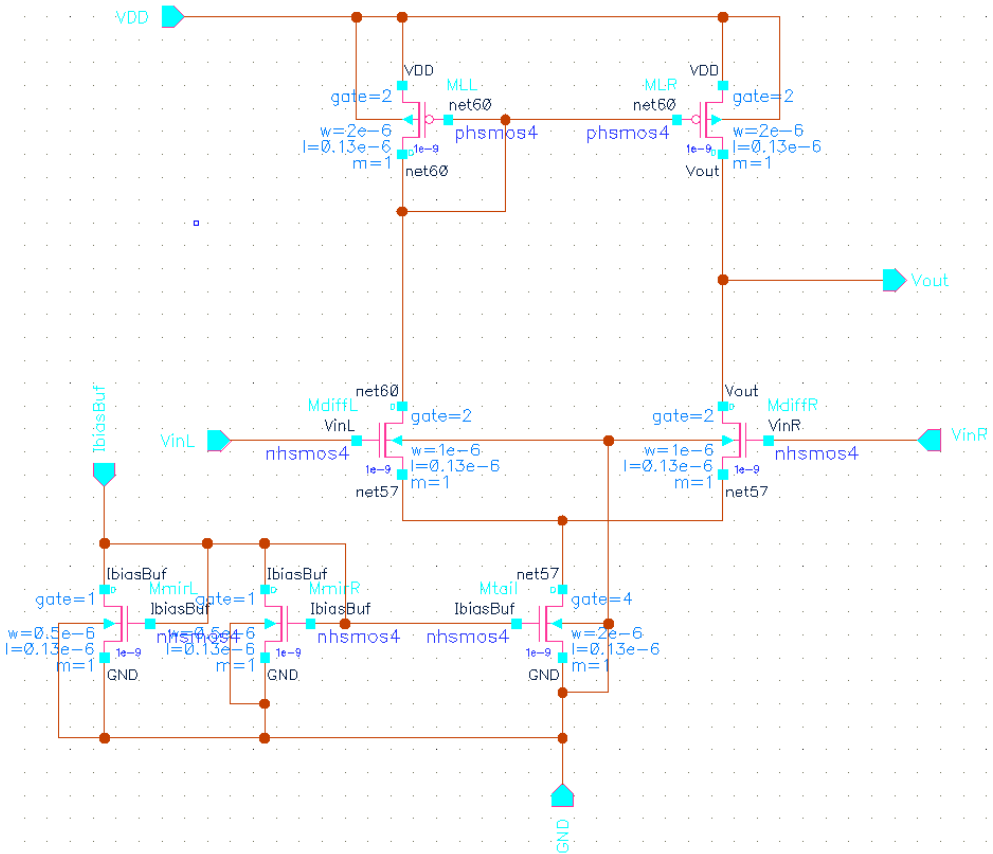
Output buffer



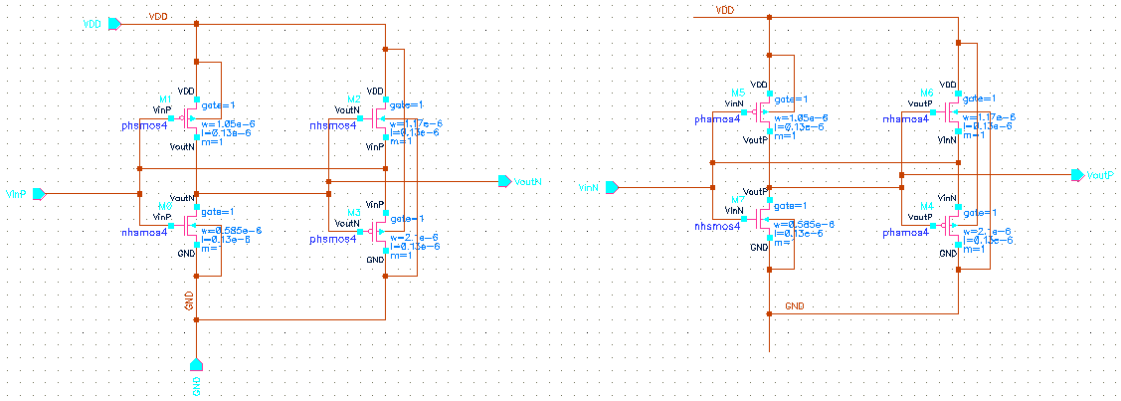
D-latches block (with reset)



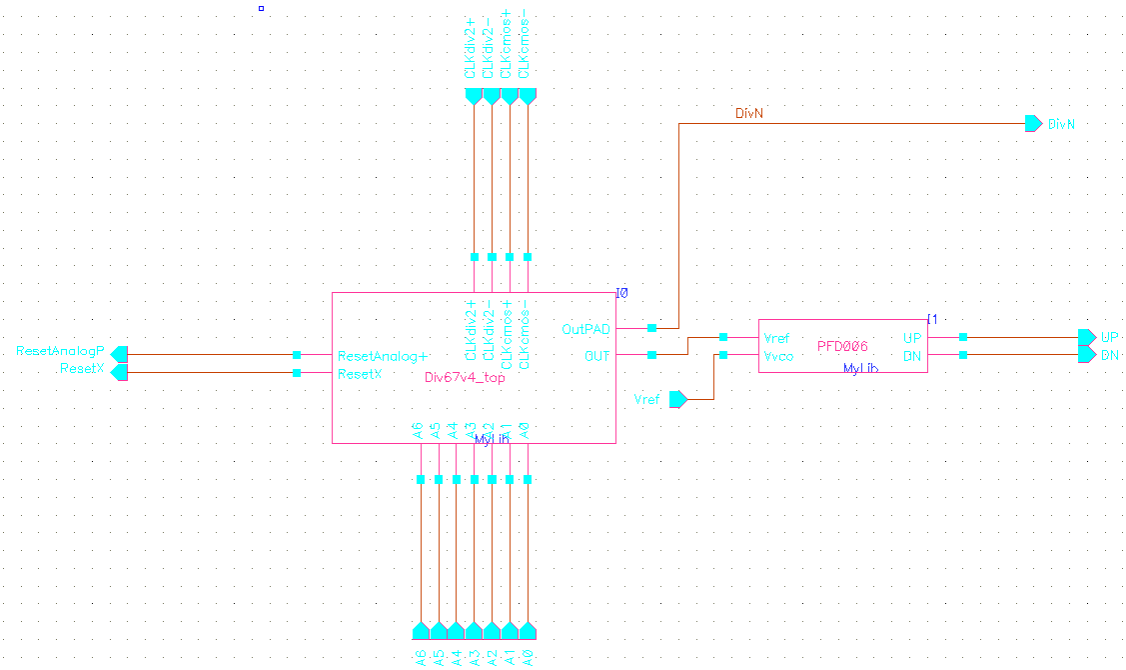
Differential to single block



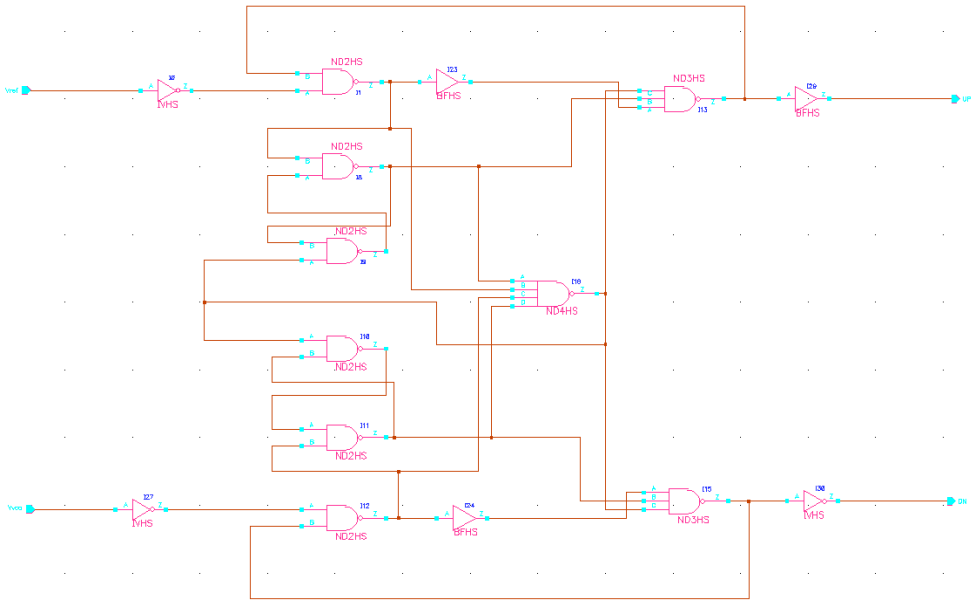
CML to CMOS conversion block



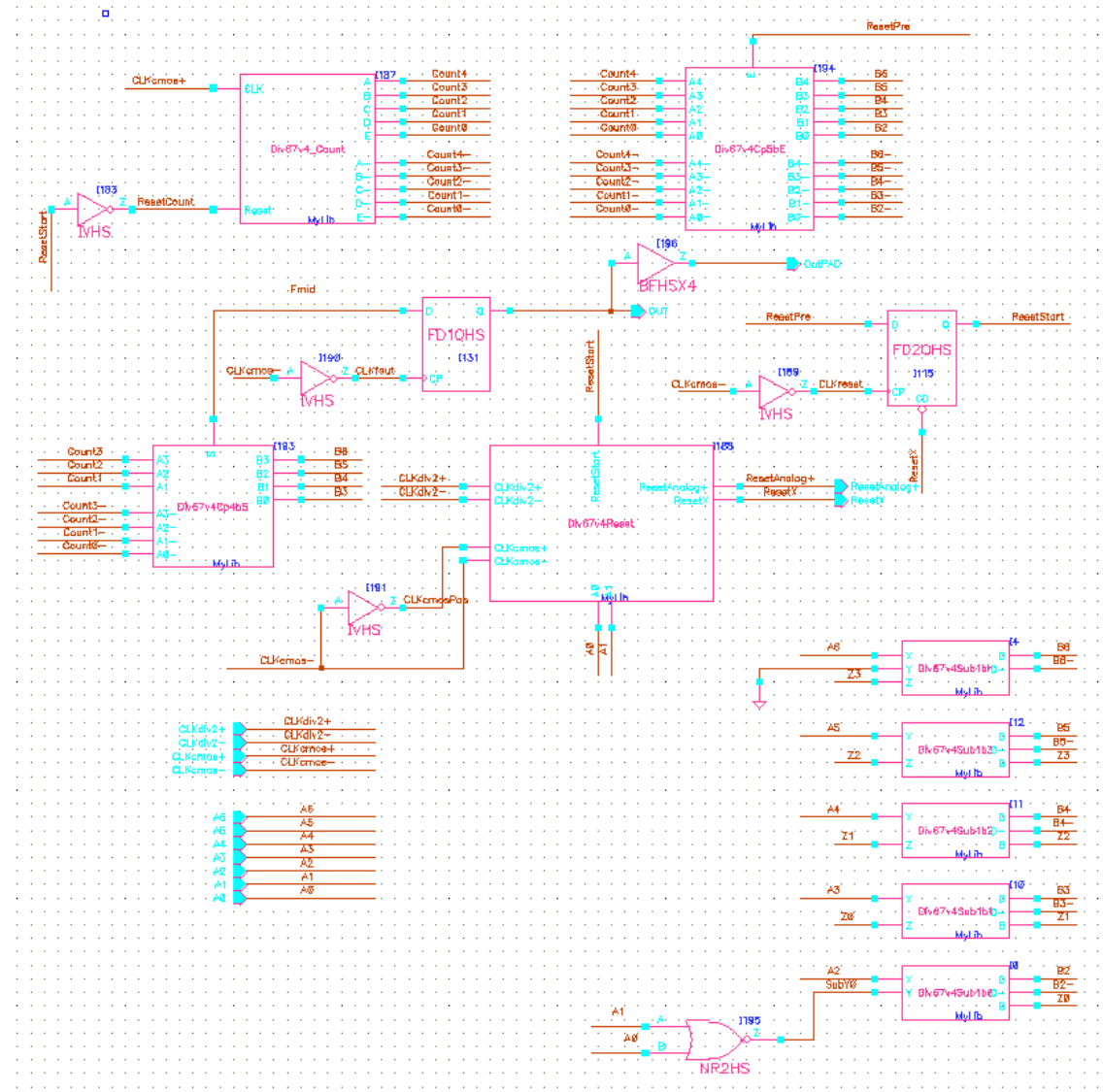
Digital block top view



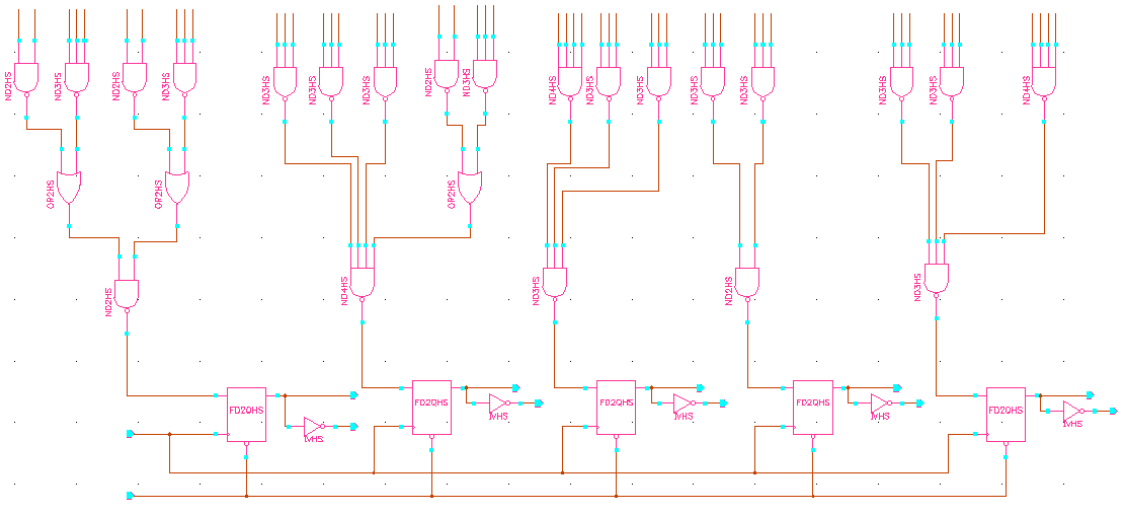
PFD block



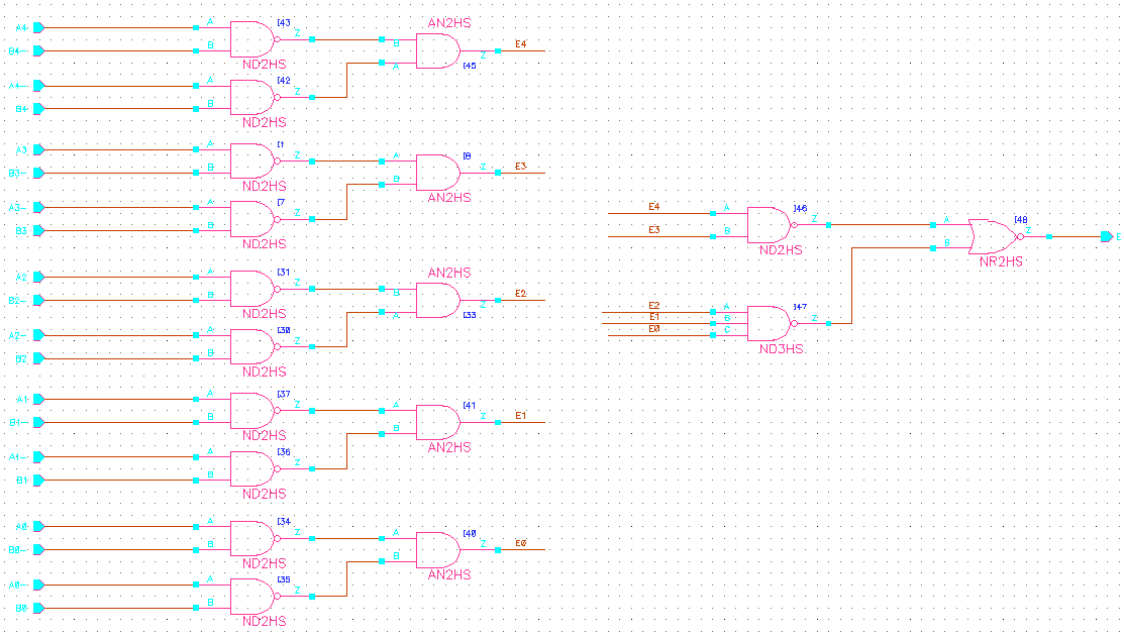
Digital divider top cell



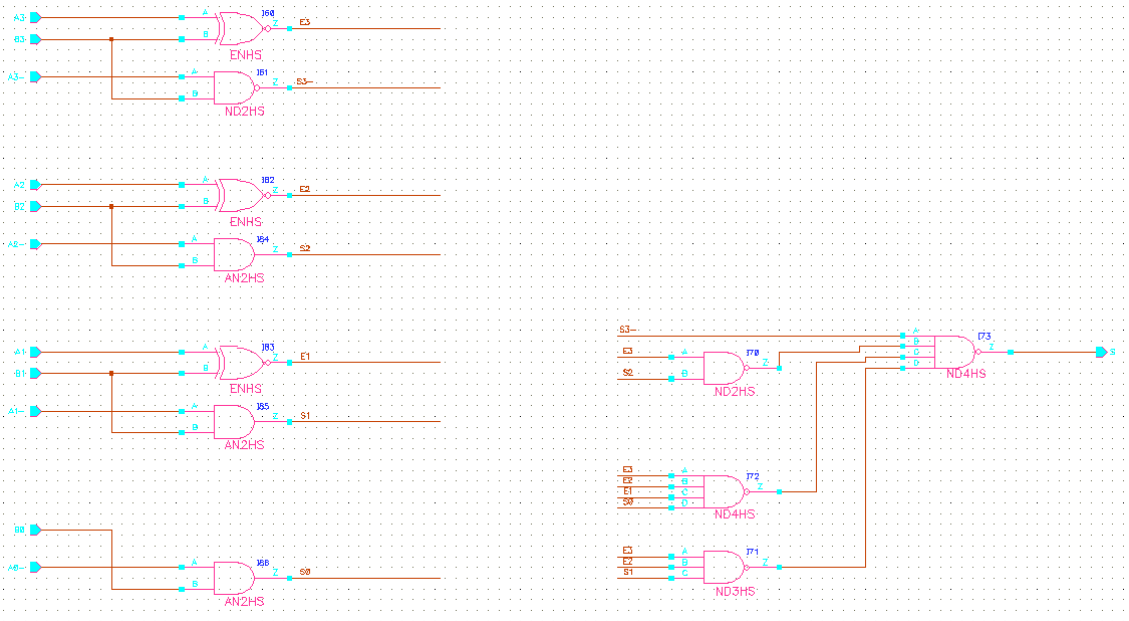
Digital counter block



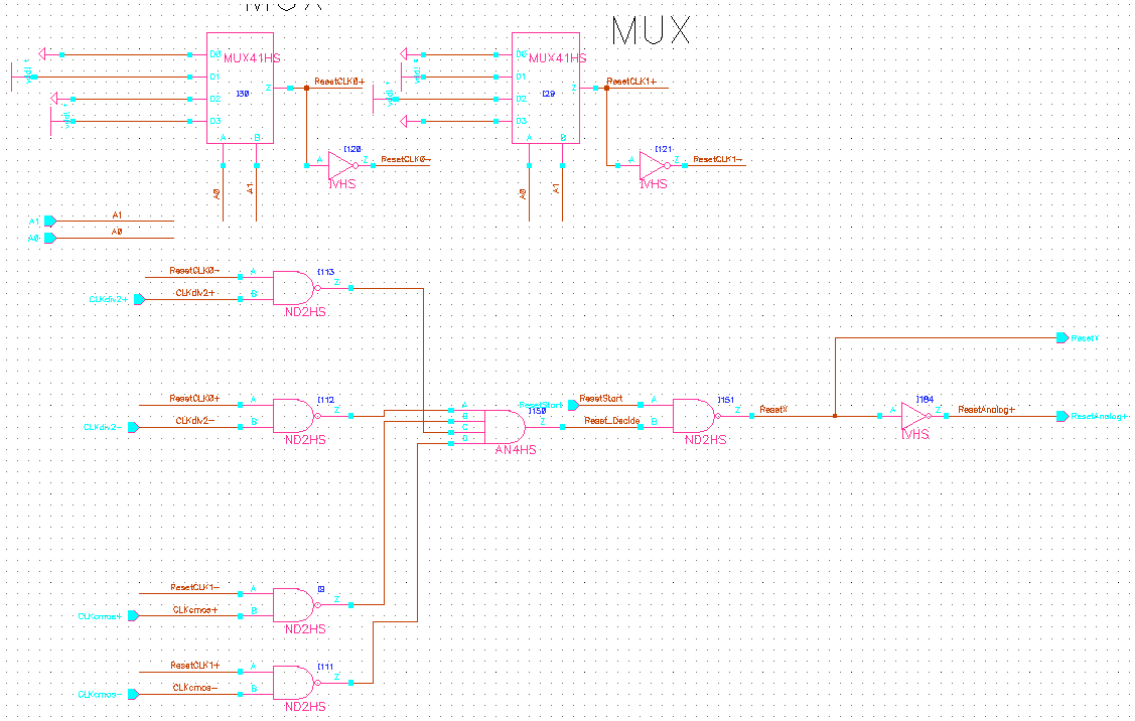
Digital comparator1



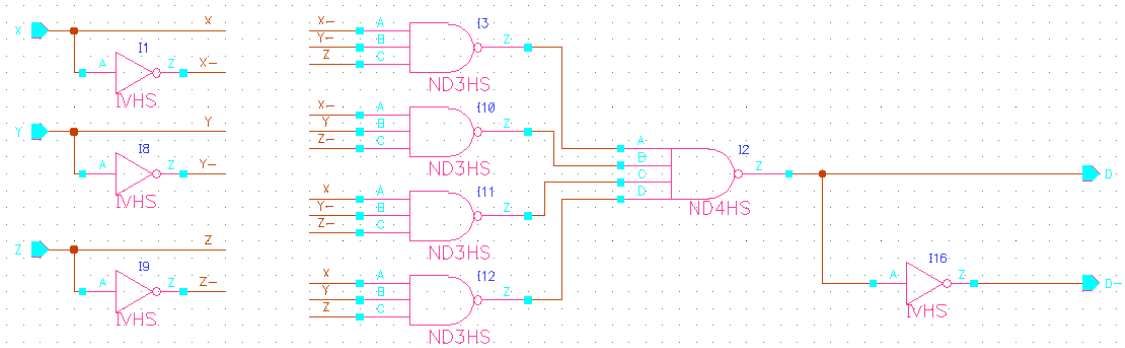
Digital comparator2



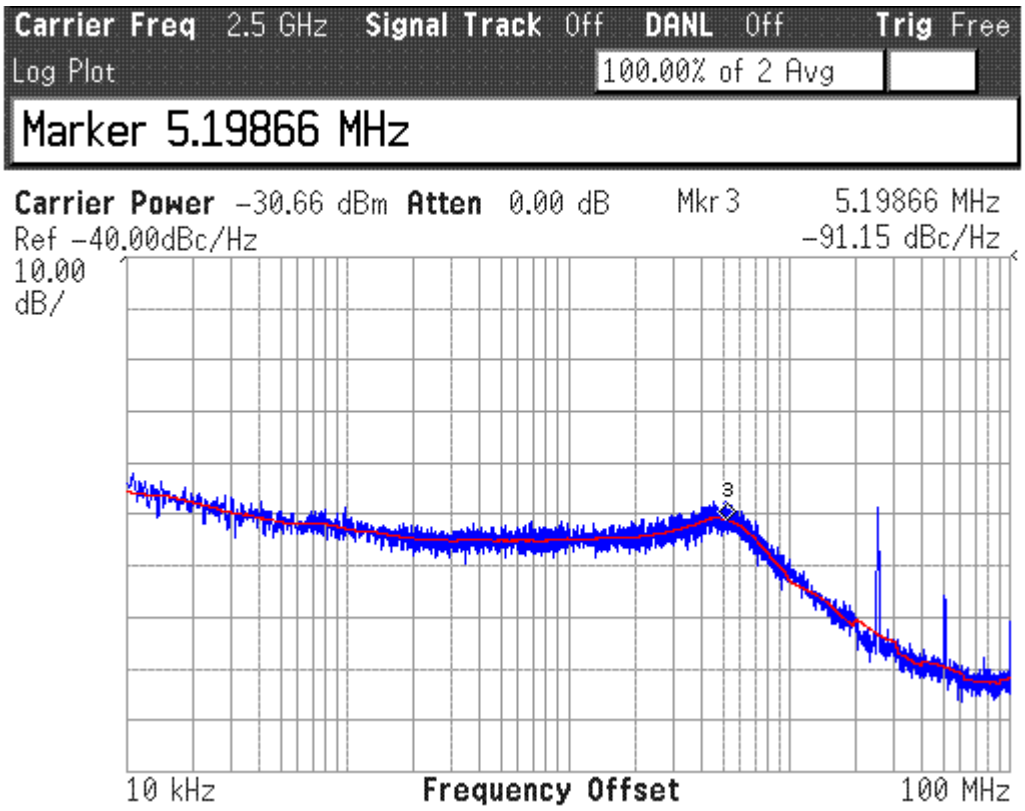
Digital reset block

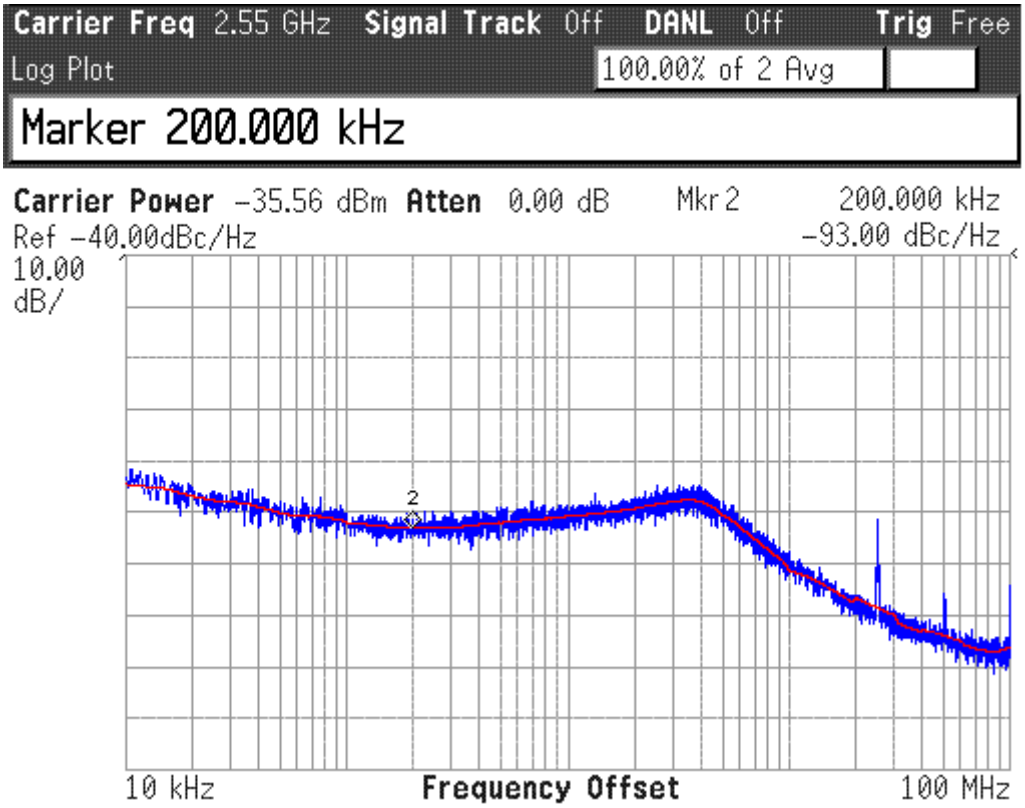


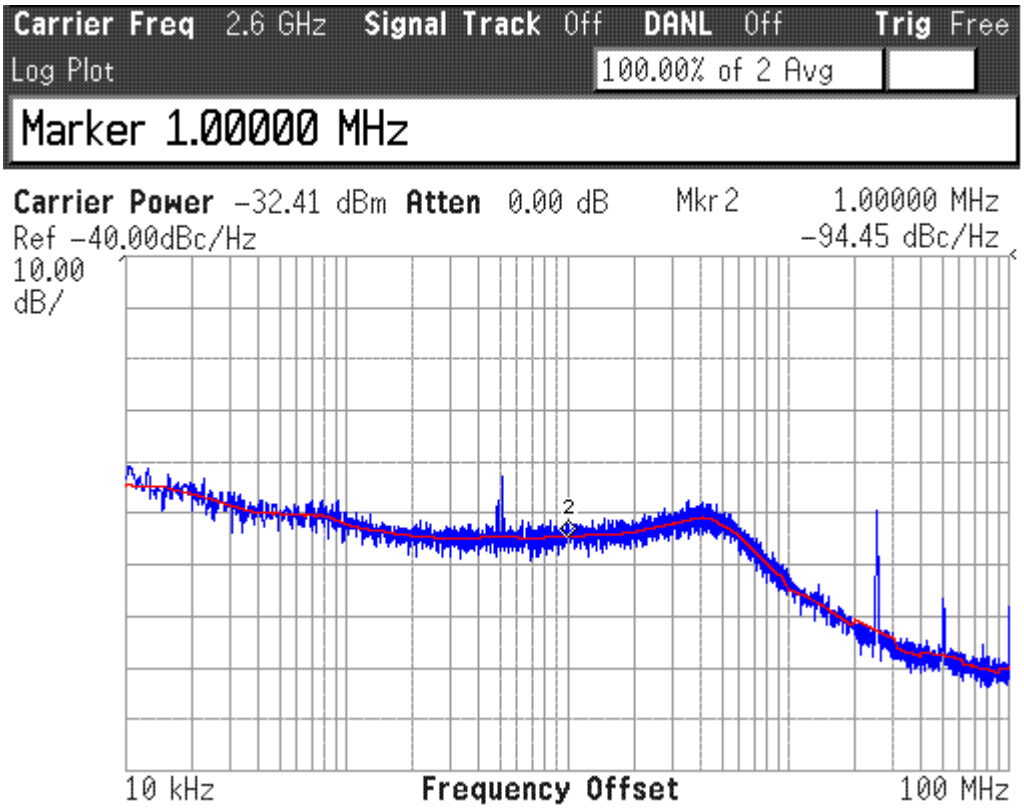
Digital 1bit subtracting block

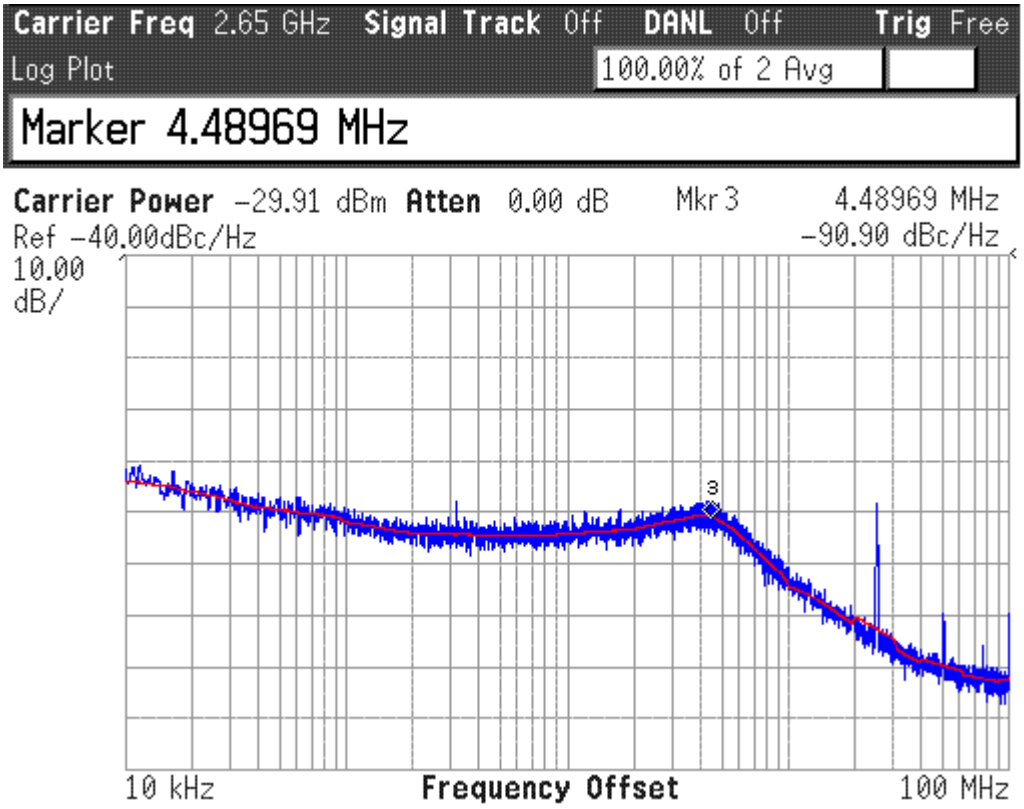


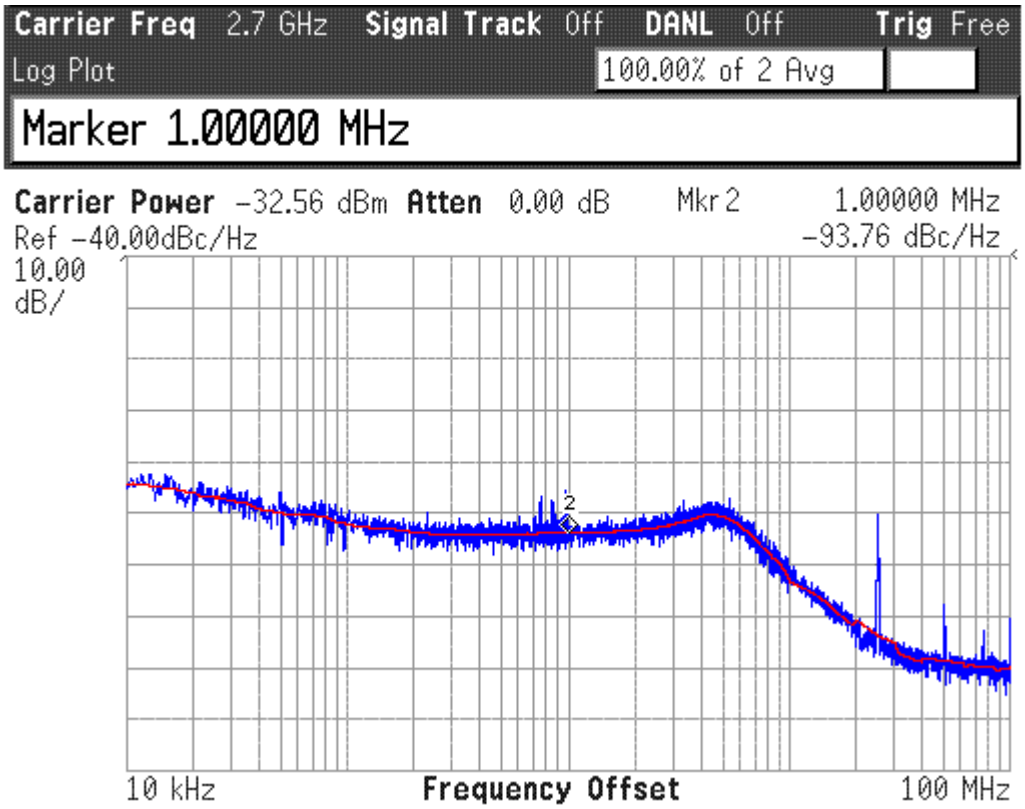
C-2 PLL Measurement Raw Data

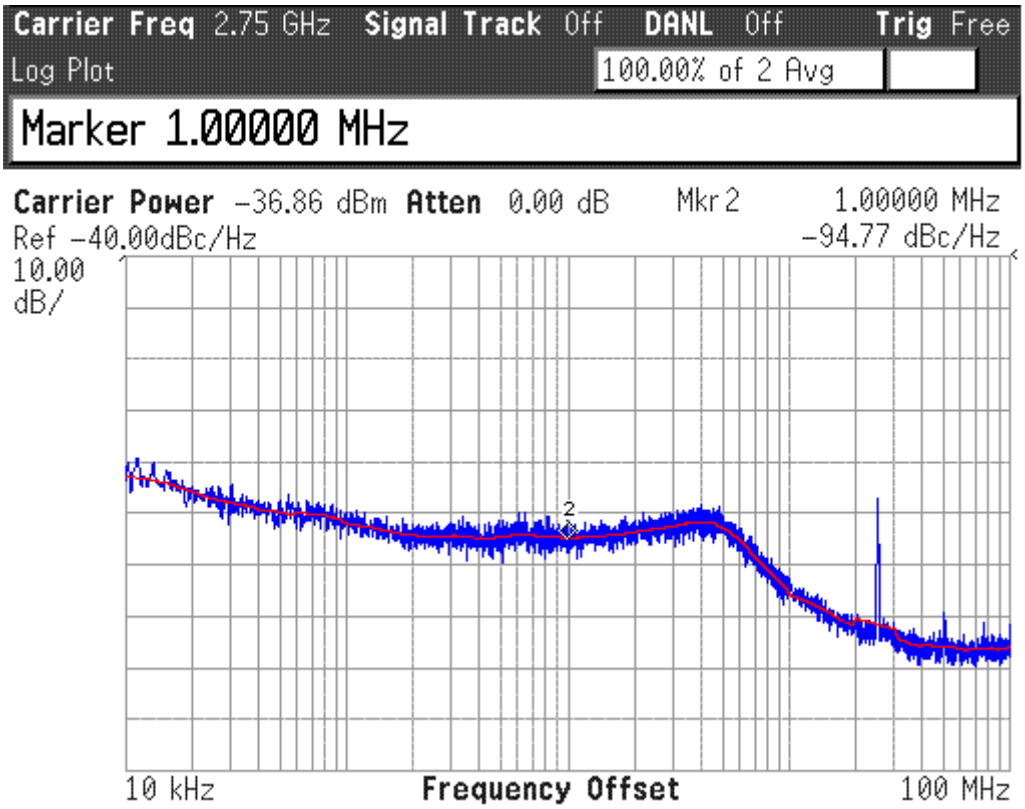


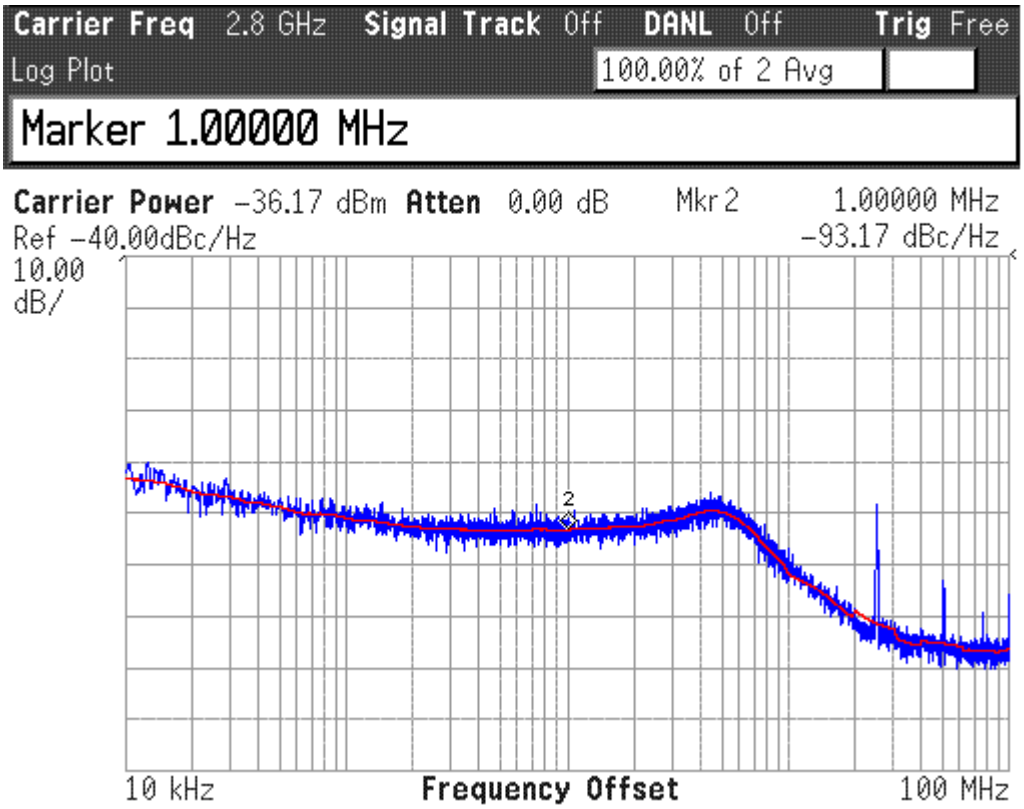


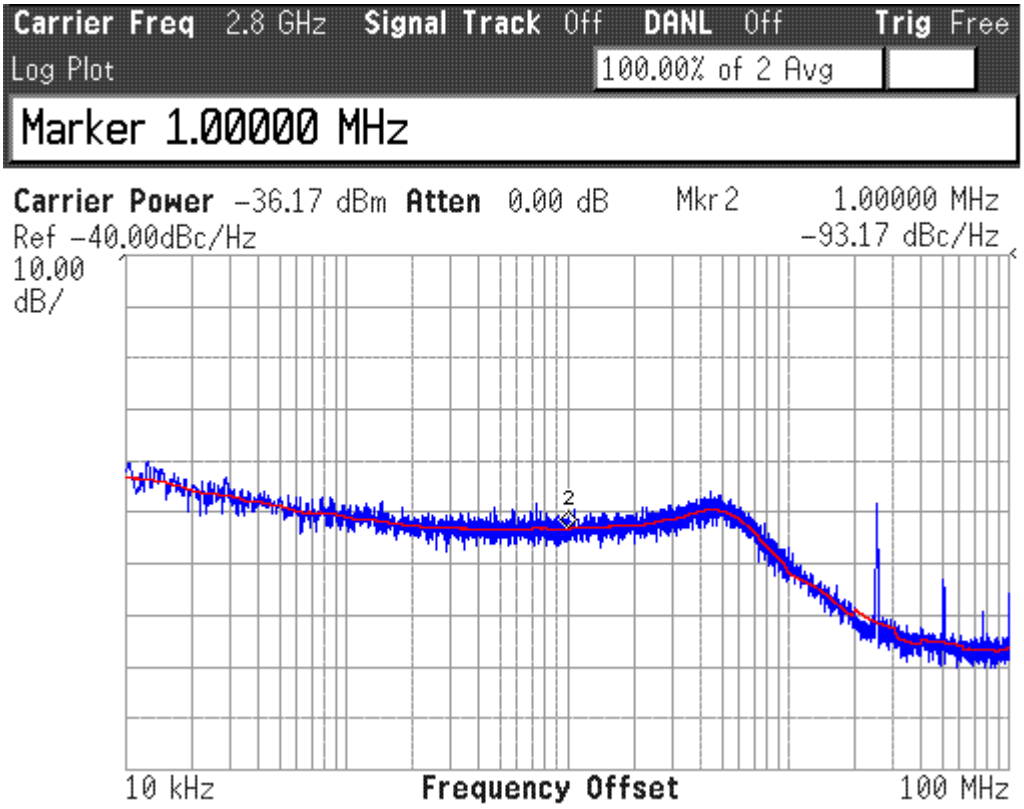

Carrier Freq 2.6 GHz Signal Track Off DANL Off Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -32.41 dBm Atten 0.00 dB Mkr2 1.00000 MHz
Ref -40.00dBc/Hz -94.45 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz

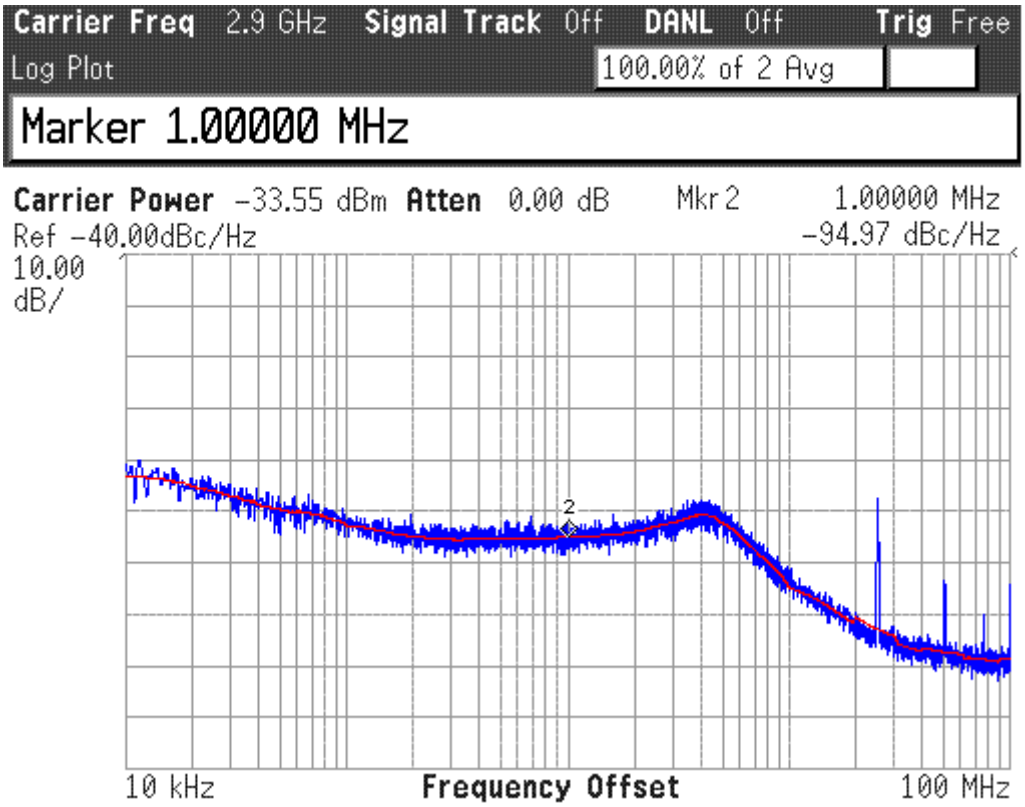

Carrier Freq 2.65 GHz Signal Track Off DANL Off Trig Free
Log Plot
100.00% of 2 Avg
Marker 4.48969 MHz
Carrier Power -29.91 dBm Atten 0.00 dB Mkr3 4.48969 MHz
Ref -40.00dBc/Hz -90.90 dBc/Hz
10.00
dB/
3
10 kHz
Frequency Offset
100 MHz

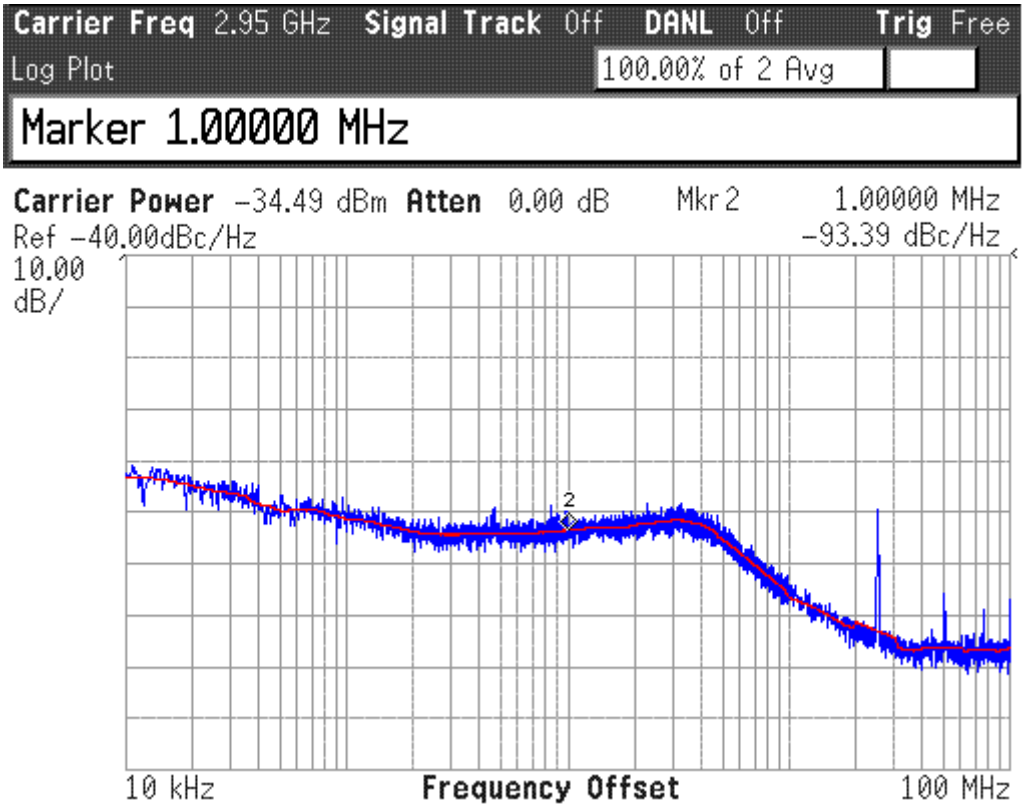

Carrier Freq 2.7 GHz
Signal Track Off
DANL Off
Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -32.56 dBm
Atten 0.00 dB
Mkr 2
1.00000 MHz
Ref -40.00dBc/Hz
-93.76 dBc/Hz
10.00
dB/
10 kHz
Frequency Offset
100 MHz

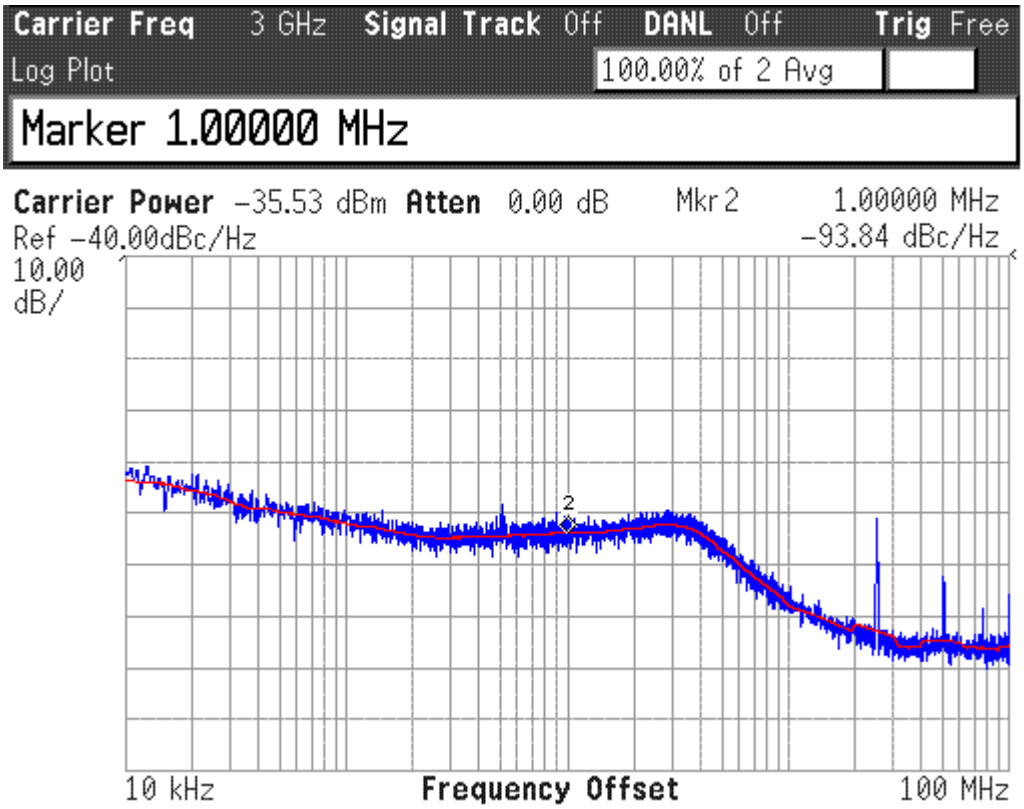

Carrier Freq 2.75 GHz
Signal Track Off
DANL Off
Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -36.86 dBm
Atten 0.00 dB
Mkr 2
1.00000 MHz
Ref -40.00dBc/Hz
-94.77 dBc/Hz
10.00
dB/
10 kHz
Frequency Offset
100 MHz

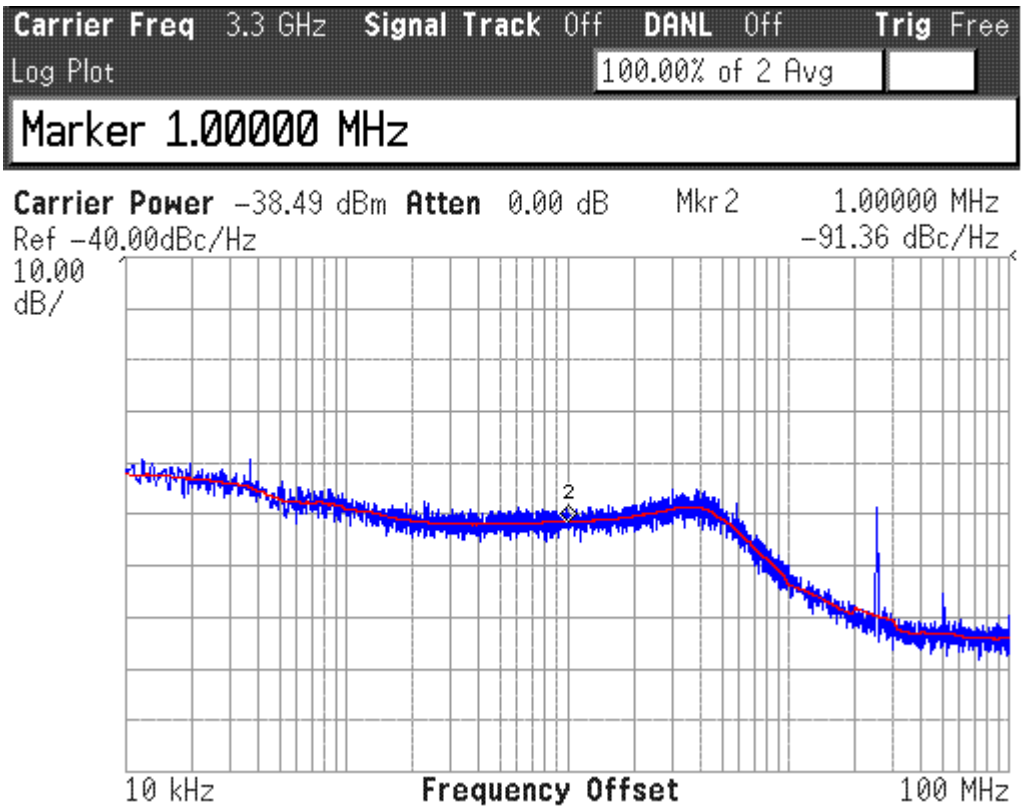

Carrier Freq 2.8 GHz Signal Track Off DANL Off Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -36.17 dBm Atten 0.00 dB Mkr2 1.00000 MHz
Ref -40.00dBc/Hz -93.17 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz

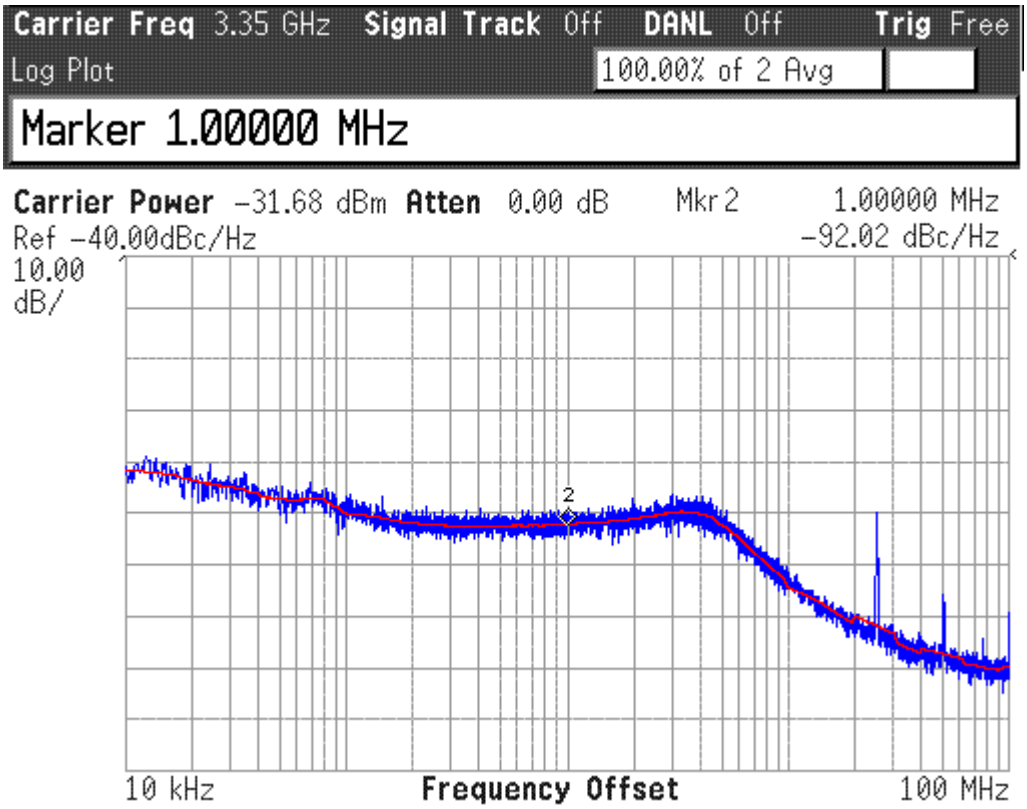

Carrier Freq 2.8 GHz Signal Track Off DANL Off Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -36.17 dBm Atten 0.00 dB Mkr2 1.00000 MHz
Ref -40.00dBc/Hz -93.17 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz

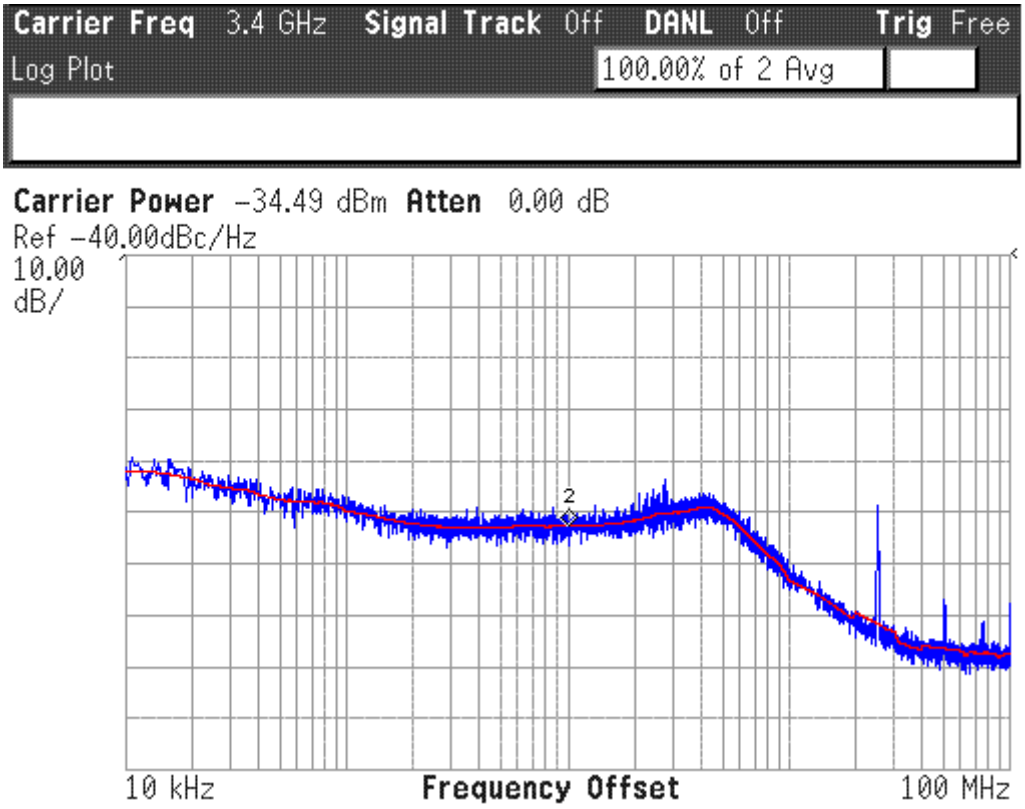

Carrier Freq 2.9 GHz
Signal Track Off
DANL Off
Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -33.55 dBm
Atten 0.00 dB
Mkr2 1.00000 MHz
Ref -40.00dBc/Hz
-94.97 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz

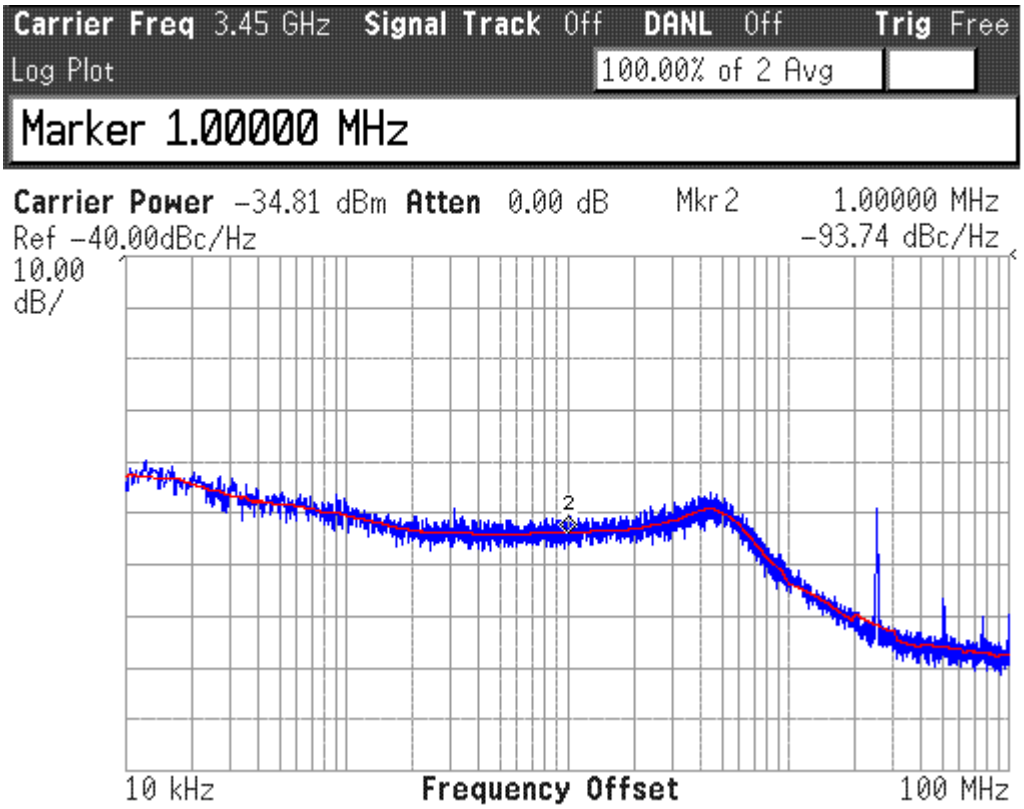

Carrier Freq 2.95 GHz
Signal Track Off
DANL Off
Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -34.49 dBm
Atten 0.00 dB
Mkr2 1.00000 MHz
Ref -40.00dBc/Hz
-93.39 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz

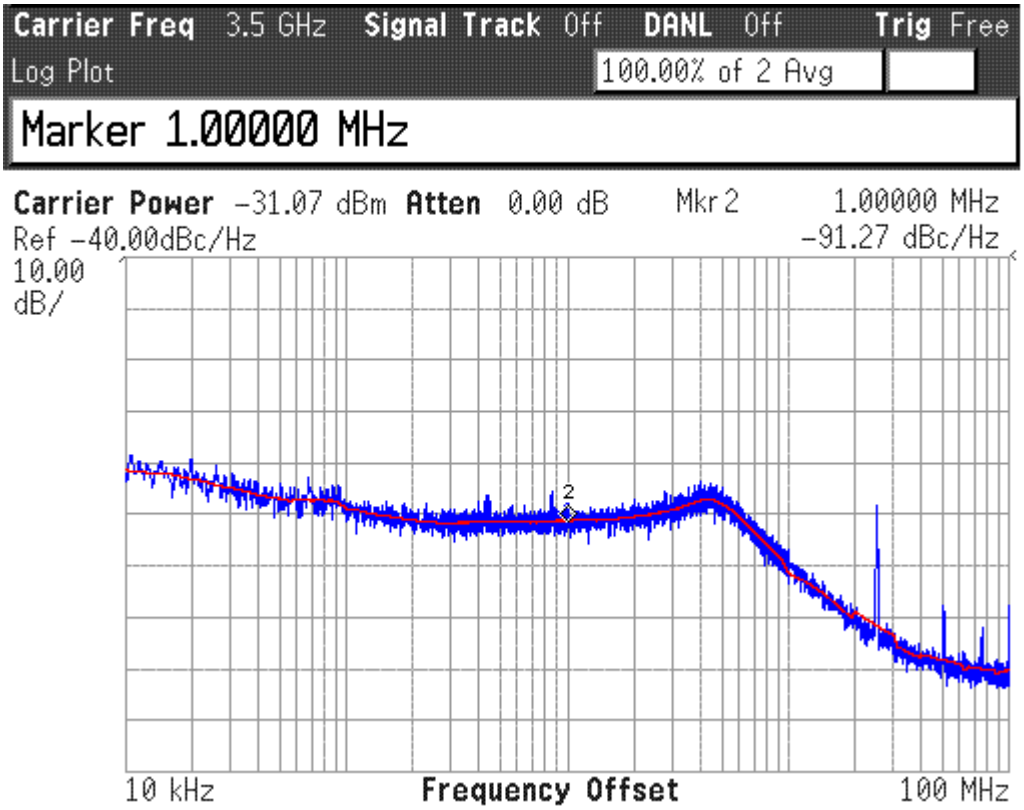

Carrier Freq 3 GHz Signal Track Off DANL Off Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -35.53 dBm Atten 0.00 dB Mkr2 1.00000 MHz
Ref -40.00dBc/Hz -93.84 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz

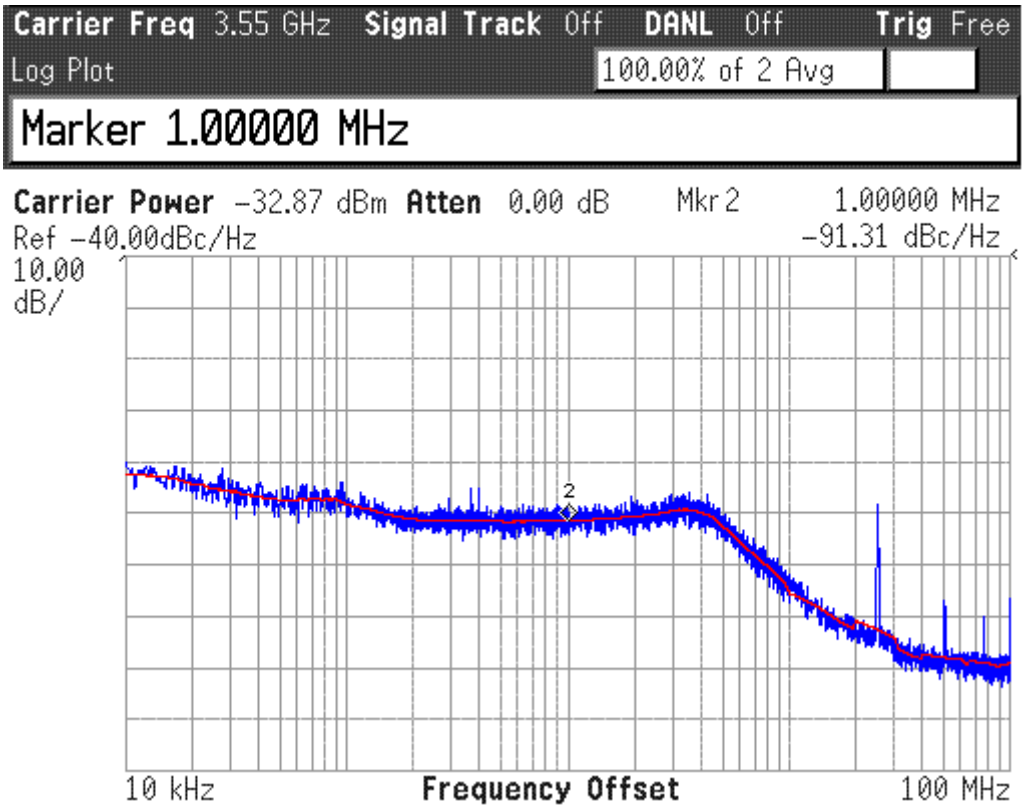

Carrier Freq 3.3 GHz Signal Track Off DANL Off Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -38.49 dBm Atten 0.00 dB Mkr2 1.00000 MHz
Ref -40.00dBc/Hz -91.36 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz

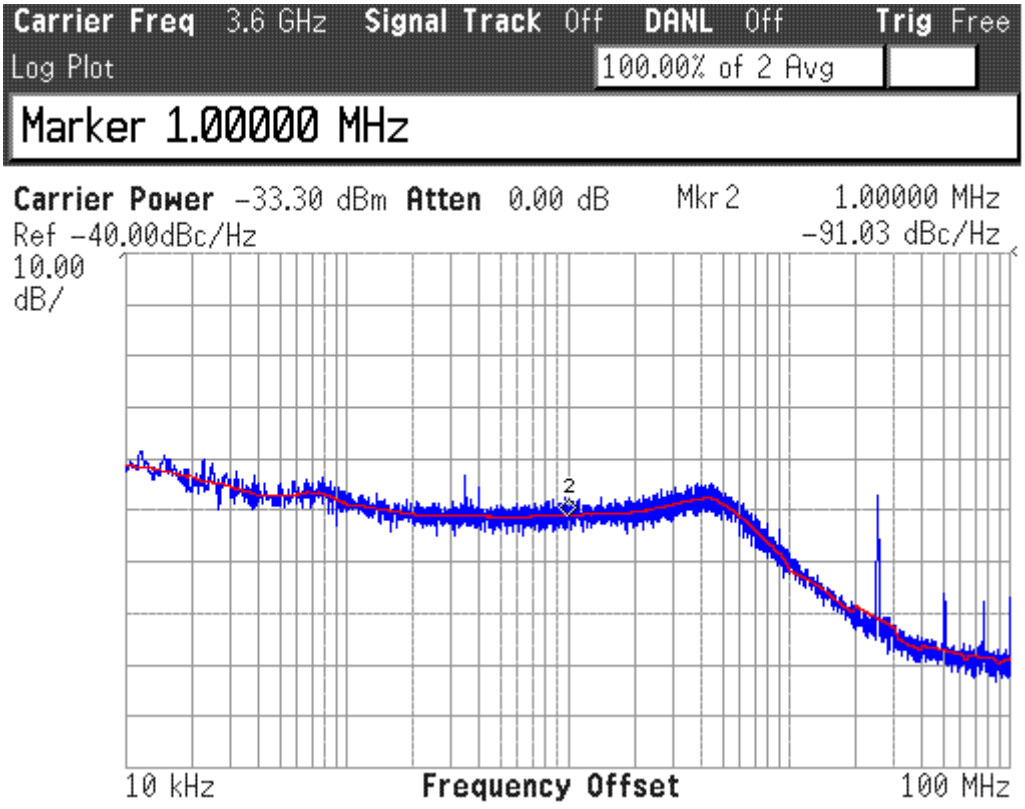

Carrier Freq 3.35 GHz
Signal Track Off
DANL Off
Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -31.68 dBm
Atten 0.00 dB
Mkr2 1.00000 MHz
Ref -40.00dBc/Hz
-92.02 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz

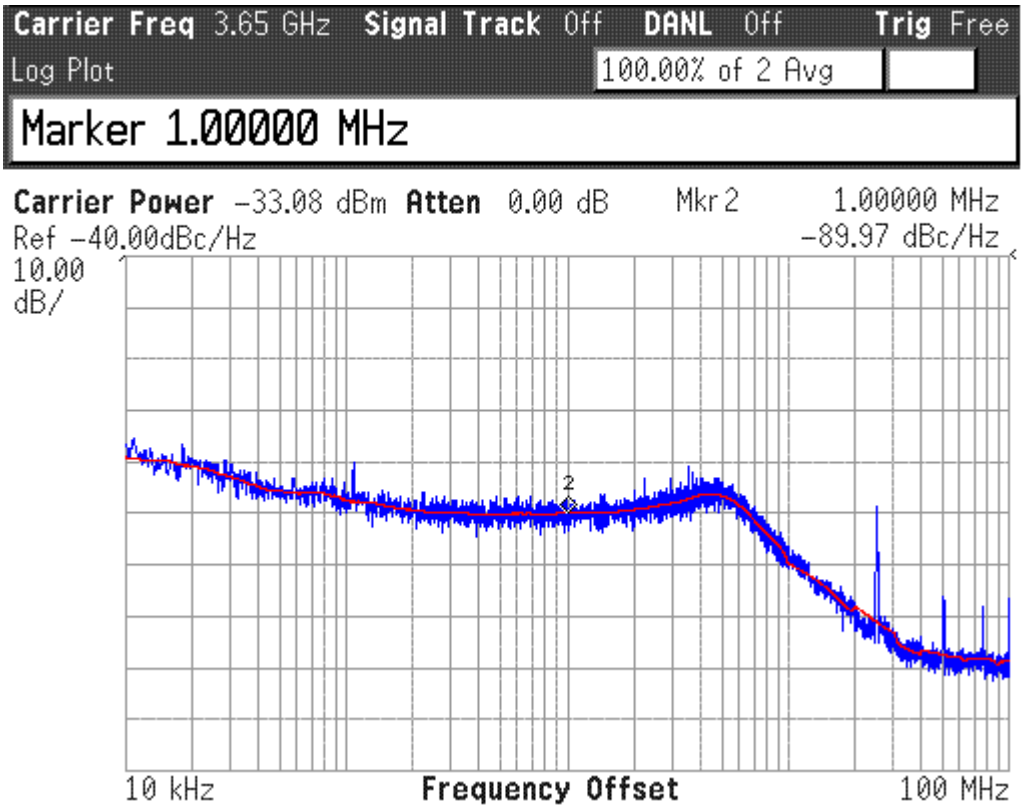

Carrier Freq 3.4 GHz
Signal Track Off
DANL Off
Trig Free
Log Plot
100.00% of 2 Avg
Carrier Power -34.49 dBm
Atten 0.00 dB
Ref -40.00dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz


Carrier Freq 3.45 GHz Signal Track Off DANL Off Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -34.81 dBm Atten 0.00 dB Mkr2 1.00000 MHz
Ref -40.00dBc/Hz -93.74 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz


Carrier Freq 3.5 GHz Signal Track Off DANL Off Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -31.07 dBm Atten 0.00 dB Mkr2 1.00000 MHz
Ref -40.00dBc/Hz -91.27 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz


Carrier Freq 3.55 GHz
Signal Track Off
DANL Off
Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -32.87 dBm
Atten 0.00 dB
Mkr 2
1.00000 MHz
Ref -40.00dBc/Hz
-91.31 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz


Carrier Freq 3.6 GHz
Signal Track Off
DANL Off
Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -33.30 dBm
Atten 0.00 dB
Mkr 2
1.00000 MHz
Ref -40.00dBc/Hz
-91.03 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz


Carrier Freq 3.65 GHz
Signal Track Off
DANL Off
Trig Free
Log Plot
100.00% of 2 Avg
Marker 1.00000 MHz
Carrier Power -33.08 dBm
Atten 0.00 dB
Mkr2 1.00000 MHz
Ref -40.00dBc/Hz
-89.97 dBc/Hz
10.00
dB/
2
10 kHz
Frequency Offset
100 MHz

References

[1] J. Mitola, "Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio," Doctor of Technology, Royal Institute of Technology(KTH), 2000.

[2] FCC, "NPRM-03-322. Notice of Proposed Rule Making and Order," 30 Dec, 2003 2003.

[3] D. Cabric*, et al.*, "Implementation issues in spectrum sensing for cognitive radios," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on*, 2004, pp. 772-776 Vol.1.

[4] J. Mitola, "The software radio architecture," *Communications Magazine, IEEE,* vol. 33, pp. 26-38, 1995.

[5] A. A. Abidi, "Evolution of a Software-Defined Radio Receiver's RF Front-End," in *Radio Frequency Integrated Circuits (RFIC) Symposium, 2006 IEEE*, 2006, pp. 17-20.

[6] N. C. Davies, "A high performance HF software radio," in *HF Radio Systems and Techniques, 2000. Eighth International Conference on (IEE Conf. Publ. No. 474)*, 2000, pp. 249-256.

[7] H. Yoshida*, et al.*, "A software defined radio receiver using the direct conversion principle: implementation and evaluation," in *Personal, Indoor and Mobile Radio Communications, 2000. PIMRC 2000. The 11th IEEE International Symposium on*, 2000, pp. 1044-1048 vol.2.

[8] D. M. Akos*, et al.*, "Direct bandpass sampling of multiple distinct RF signals," *Communications, IEEE Transactions on,* vol. 47, pp. 983-988, 1999.

[9] S. Lindfors*, et al.*, "A 3-V 230-MHz CMOS decimation subsampler," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on [see also Circuits and Systems II: Express Briefs, IEEE Transactions on],* vol. 50, pp. 105-117, 2003.

[10] S. Karvonen*, et al.*, "A quadrature charge-domain sampler with embedded FIR and IIR filtering functions," *Solid-State Circuits, IEEE Journal of,* vol. 41, pp. 507-515, 2006.

[11] M. McHenry. (2005, *Report on Spectrum Occupancy Measurments*. Available: http://www.sharedspectrum.com/measurements/

[12] C. J. Rieser, "Biologically Inspired Cognitive Radio Engine Model Utilizing Distributed Genetic Algorithms for Secure and Robust Wireless COmmunications and Networks," Doctor of Philosophy, Virginia Polytechnic Institute and State University, 2004.

[13] A. Sahai, *et al.*, "Some Fundamental Limits in Cognitive Radio," in *Allerton Conf. on Communication, Control and Computing*, 2004.
[14] B. Wild and K. Ramchandran, "Detecting primary receivers for cognitive radio applications," in *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, 2005, pp. 124-130.
[15] H. Chun-Huat, *et al.*, "A CMOS TV tuner/demodulator IC with digital image rejection," *Solid-State Circuits, IEEE Journal of,* vol. 40, pp. 2525-2535, 2005.
[16] J. van Sinderen, *et al.*, "A 48-860MHz digital cable tuner IC with integrated RF and IF selectivity," in *Solid-State Circuits Conference, 2003. Digest of Technical Papers. ISSCC. 2003 IEEE International*, 2003, pp. 444-506 vol.1.
[17] B. Widrow, *et al.*, "The complex LMS algorithm," *Proceedings of the IEEE,* vol. 63, pp. 719-720, 1975.
[18] S. Roy, *et al.*, "Ultrawideband radio design: the promise of high-speed, short-range wireless connectivity," *Proceedings of the IEEE,* vol. 92, pp. 295-311, 2004.
[19] P. Heydari, "A study of low-power ultra wideband radio transceiver architectures," 2005, pp. 758-763 Vol. 2.
[20] J. R. Bergervoet, *et al.*, "A WiMedia-Compliant UWB Transceiver in 65nm CMOS," in *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International*, 2007, pp. 112-590.
[21] Z. Hui, *et al.*, "A 3.1 GHz - 8.0 GHz Single-Chip Transceiver for MB-OFDM UWB in 0.18-um CMOS Process," *Solid-State Circuits, IEEE Journal of,* vol. 44, pp. 414-426, 2009.
[22] Wgsimon. (2008, *CPU Transistor Counts 1971-2008 & Moore's Law*. Available: http://en.wikipedia.org/wiki/Moore's_law
[23] (2003, *RADIO FREQUENCY AND ANALOG/MIXED-SIGNAL TECHNOLOGIES FOR WIRELESS COMMUNICATIONS*. Available: http://www.itrs.net/Links/2003ITRS/Home2003.htm
[24] (2004, *RADIO FREQUENCY AND ANALOG/MIXED-SIGNAL TECHNOLOGIES FOR WIRELESS COMMUNICATIONS*. Available: http://www.itrs.net/Links/2004Update/2004Update.htm
[25] (2005, *RADIO FREQUENCY AND ANALOG/MIXED-SIGNAL TECHNOLOGIES FOR WIRELESS COMMUNICATIONS*. Available: http://www.itrs.net/Links/2005ITRS/Home2005.htm
[26] (2006, *RADIO FREQUENCY AND ANALOG/MIXED-SIGNAL TECHNOLOGIES FOR WIRELESS COMMUNICATIONS*. Available: http://www.itrs.net/Links/2006Update/2006UpdateFinal.htm
[27] (2007, *RADIO FREQUENCY AND ANALOG/MIXED-SIGNAL TECHNOLOGIES FOR WIRELESS COMMUNICATIONS*. Available: http://www.itrs.net/Links/2007ITRS/Home2007.htm
[28] (2008, *2008 FOCUS ITWG TABLES: System Drivers, Design, Test & Test Equipment, RF and AMS for Wireless, and Process Integration, Devices, & Structures (PIDS)*. Available: http://www.itrs.net/Links/2008ITRS/Home2008.htm
[29] (2009, *RF and Analog Mixed-Signal CMOS Technology Requirements*. Available: http://www.itrs.net/Links/2009ITRS/Home2009.htm
[30] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, 1st ed. New York: Cambridge University Press, 1998.
[31] D. Linten, *et al.*, "A 5 GHz fully integrated ESD-protected low-noise amplifier in 90 nm RF CMOS," 2004, pp. 291-294.

[32] R. J. Baker, *CMOS Circuit Design, Layout , and Simulation*, 2nd ed. New Jersey: Jown Wiley & Sons, Inc., 2008.
[33] D. K. Shaeffer and T. H. Lee, "A 1.5-V, 1.5-GHz CMOS low noise amplifier," *Solid-State Circuits, IEEE Journal of,* vol. 32, pp. 745-759, 1997.
[34] Available: www.mosis.com
[35] Q. Huang, *et al.*, "The impact of scaling down to deep submicron on CMOS RF circuits," *Solid-State Circuits, IEEE Journal of,* vol. 33, pp. 1023-1036, 1998.
[36] P. H. Woerlee, *et al.*, "RF-CMOS performance trends," *Electron Devices, IEEE Transactions on,* vol. 48, pp. 1776-1782, 2001.
[37] R. Brederlow, *et al.*, "A mixed-signal design roadmap," *Design & Test of Computers, IEEE,* vol. 18, pp. 34-46, 2001.
[38] D. M. Binkley, *et al.*, "A CAD methodology for optimizing transistor current and sizing in analog CMOS design," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on,* vol. 22, pp. 225-237, 2003.
[39] F. Zhang and P. R. Kinget, "Low-power programmable gain CMOS distributed LNA," *Solid-State Circuits, IEEE Journal of,* vol. 41, pp. 1333-1343, 2006.
[40] F. Bruccoleri, *et al.*, "Wide-band CMOS low-noise amplifier exploiting thermal noise canceling," *Solid-State Circuits, IEEE Journal of,* vol. 39, pp. 275-282, 2004.
[41] Z. Heng, *et al.*, "A Low-Power, Linearized, Ultra-Wideband LNA Design Technique," *Solid-State Circuits, IEEE Journal of,* vol. 44, pp. 320-330, 2009.
[42] P. Heydari and D. Lin, "A performance optimized CMOS distributed LNA for UWB receivers," in *Custom Integrated Circuits Conference, 2005. Proceedings of the IEEE 2005*, 2005, pp. 337-340.
[43] A. Bevilacqua and A. M. Niknejad, "An ultra-wideband CMOS LNA for 3.1 to 10.6 GHz wireless receivers," in *Solid-State Circuits Conference, 2004. Digest of Technical Papers. ISSCC. 2004 IEEE International*, 2004, pp. 382-533 Vol.1.
[44] D. Barras, *et al.*, "A low supply voltage SiGe LNA for ultra-wideband frontends," *Microwave and Wireless Components Letters, IEEE,* vol. 14, pp. 469-471, 2004.
[45] A. Amer, *et al.*, "A Low-Power Wideband CMOS LNA for WiMAX," *Circuits and Systems II: Express Briefs, IEEE Transactions on [see also Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on],* vol. 54, pp. 4-8, 2007.
[46] Z. Sining and M. C. F. Chang, "A CMOS passive mixer with low flicker noise for low-power direct-conversion receiver," *Solid-State Circuits, IEEE Journal of,* vol. 40, pp. 1084-1093, 2005.
[47] A. Karimi-Sanjaani, *et al.*, "A 2 GHz merged CMOS LNA and mixer for WCDMA," in *VLSI Circuits, 2001. Digest of Technical Papers. 2001 Symposium on*, 2001, pp. 19-22.
[48] V. Vidojkovic, *et al.*, "A low-voltage folded-switching mixer in 0.18-um CMOS," *Solid-State Circuits, IEEE Journal of,* vol. 40, pp. 1259-1264, 2005.
[49] M. T. Terrovitis and R. G. Meyer, "Noise in current-commutating CMOS mixers," *Solid-State Circuits, IEEE Journal of,* vol. 34, pp. 772-783, 1999.
[50] N. H. W. Fong, *et al.*, "Design of wide-band CMOS VCO for multiband wireless LAN applications," *Solid-State Circuits, IEEE Journal of,* vol. 38, pp. 1333-1342, 2003.
[51] A. Hajimiri and T. H. Lee, "A general theory of phase noise in electrical oscillators," *Solid-State Circuits, IEEE Journal of,* vol. 33, pp. 179-194, 1998.

[52] K. Manetakis, *et al.*, "A CMOS VCO with 48% tuning range for modern broadband systems," in *Custom Integrated Circuits Conference, 2004. Proceedings of the IEEE 2004*, 2004, pp. 265-268.
[53] D. B. Leeson, "A simple model of feedback oscillator noise spectrum," *Proceedings of the IEEE,* vol. 54, pp. 329-330, 1966.
[54] B. Razavi, "A study of phase noise in CMOS oscillators," *Solid-State Circuits, IEEE Journal of,* vol. 31, pp. 331-343, 1996.
[55] J. J. Rael and A. A. Abidi, "Physical processes of phase noise in differential LC oscillators," in *Custom Integrated Circuits Conference, 2000. CICC. Proceedings of the IEEE 2000*, 2000, pp. 569-572.
[56] A. A. Abidi, "Phase Noise and Jitter in CMOS Ring Oscillators," *Solid-State Circuits, IEEE Journal of,* vol. 41, pp. 1803-1816, 2006.
[57] U. Singh and M. M. Green, "High-frequency CML clock dividers in 0.13-um CMOS operating up to 38 GHz," *Solid-State Circuits, IEEE Journal of,* vol. 40, pp. 1658-1661, 2005.
[58] S. Verma, *et al.*, "A unified model for injection-locked frequency dividers," *Solid-State Circuits, IEEE Journal of,* vol. 38, pp. 1015-1027, 2003.
[59] D. Huang and M. C. F. Chang, "Phase-coherent frequency divider with high speed-gain/power figure-of-merit," *Electronics Letters,* vol. 42, pp. 1152-1153, 2006.
[60] L. Tang-Nian, *et al.*, "A V-band wide locking range CMOS frequency divider," in *Microwave Symposium Digest, 2008 IEEE MTT-S International*, 2008, pp. 563-566.
[61] T. Itakura, *et al.*, "A 2.7-V, 200-kHz, 49-dBm, stopband-IIP3, low-noise, fully balanced gm-C filter IC," *Solid-State Circuits, IEEE Journal of,* vol. 34, pp. 1155-1159, 1999.
[62] A. Yoshizawa and Y. P. Tsividis, "Anti-blocker design techniques for MOSFET-C filters for direct conversion receivers," *Solid-State Circuits, IEEE Journal of,* vol. 37, pp. 357-364, 2002.
[63] J. Jussila, *et al.*, "A channel selection filter for a WCDMA direct conversion receiver," in *Solid-State Circuits Conference, 2000. ESSCIRC '00. Proceedings of the 26th European*, 2000, pp. 264-267.
[64] B. Razavi, *RF Microelectronics*: Prentice Hall, 1998.
[65] D. Chamla, *et al.*, "A G/sub m/-C low-pass filter for zero-IF mobile applications with a very wide tuning range," *Solid-State Circuits, IEEE Journal of,* vol. 40, pp. 1443-1450, 2005.
[66] R. H. Walden, "Analog-to-digital converter survey and analysis," *Selected Areas in Communications, IEEE Journal on,* vol. 17, pp. 539-550, 1999.
[67] R. H. Walden, "Analog-to-digital converter technology comparison," in *Gallium Arsenide Integrated Circuit (GaAs IC) Symposium, 1994. Technical Digest 1994., 16th Annual*, 1994, pp. 217-219.
[68] A. Ismail and M. Elmasry, "A 6-Bit 1.6-GSs Low-Power Wideband Flash ADC Converter in 0.13um CMOS Technology," *Solid-State Circuits, IEEE Journal of,* vol. 43, pp. 1982-1990, 2008.
[69] S. Junhua and P. R. Kinget, "A 0.5-V 8-bit 10-Ms/s Pipelined ADC in 90-nm CMOS," *Solid-State Circuits, IEEE Journal of,* vol. 43, pp. 787-795, 2008.
[70] N. Verma and A. P. Chandrakasan, "An Ultra Low Energy 12-bit Rate-Resolution Scalable SAR ADC for Wireless Sensor Nodes," *Solid-State Circuits, IEEE Journal of,* vol. 42, pp. 1196-1205, 2007.

[71] R. C. Taft, *et al.*, "A 1.8 V 1.0 GS/s 10b Self-Calibrating Unified-Folding-Interpolating ADC With 9.1 ENOB at Nyquist Frequency," *Solid-State Circuits, IEEE Journal of,* vol. 44, pp. 3294-3304, 2009.
[72] J. Xicheng and M. C. F. Chang, "A 1-GHz signal bandwidth 6-bit CMOS ADC with power-efficient averaging," *Solid-State Circuits, IEEE Journal of,* vol. 40, pp. 532-535, 2005.
[73] I. Galton, "Delta-sigma data conversion in wireless transceivers," *Microwave Theory and Techniques, IEEE Transactions on,* vol. 50, pp. 302-315, 2002.
[74] W. A. Gardner, "Signal interception: a unifying theoretical framework for feature detection," *Communications, IEEE Transactions on,* vol. 36, pp. 897-906, 1988.
[75] P. Jongmin, *et al.*, "A Fully Integrated UHF-Band CMOS Receiver With Multi-Resolution Spectrum Sensing (MRSS) Functionality for IEEE 802.22 Cognitive Radio Applications," *Solid-State Circuits, IEEE Journal of,* vol. 44, pp. 258-268, 2009.
[76] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proceedings of the IEEE,* vol. 55, pp. 523-531, 1967.
[77] B. Sklar, *Digital Communications Fundamentals and Applications*, 2nd Edition ed.: Prentice Hall, 2001.
[78] J. A. Weldon, *et al.*, "A 1.75-GHz highly integrated narrow-band CMOS transmitter with harmonic-rejection mixers," *Solid-State Circuits, IEEE Journal of,* vol. 36, pp. 2003-2015, 2001.
[79] F. Behbahani, *et al.*, "CMOS mixers and polyphase filters for large image rejection," *Solid-State Circuits, IEEE Journal of,* vol. 36, pp. 873-887, 2001.
[80] S. B. Cohn, "Direct-Coupled-Resonator Filters," *Proceedings of the IRE,* vol. 45, pp. 187-196, 1957.
[81] C. P. Yue and S. S. Wong, "Physical modeling of spiral inductors on silicon," *Electron Devices, IEEE Transactions on,* vol. 47, pp. 560-568, 2000.
[82] C. P. Yue, *et al.*, "A physical model for planar spiral inductors on silicon," in *Electron Devices Meeting, 1996., International*, 1996, pp. 155-158.
[83] C. P. Yue and S. S. Wong, "On-chip spiral inductors with patterned ground shields for Si-based," *Solid-State Circuits, IEEE Journal of,* vol. 33, pp. 743-752, 1998.
[84] A. M. Niknejad and R. G. Meyer, "Analysis, design, and optimization of spiral inductors and transformers for Si RF ICs," *Solid-State Circuits, IEEE Journal of,* vol. 33, pp. 1470-1481, 1998.
[85] O. Ban-Leong, *et al.*, "A comprehensive explanation on the high quality characteristics of symmetrical octagonal spiral inductor," in *Radio Frequency Integrated Circuits (RFIC) Symposium, 2003 IEEE*, 2003, pp. 259-262.
[86] O. H. Murphy, *et al.*, "Design of multiple-metal stacked inductors incorporating an extended physical model," *Microwave Theory and Techniques, IEEE Transactions on,* vol. 53, pp. 2063-2072, 2005.
[87] A. B. Williams, *Electronic filter design handbook*. New York: McGraw-Hill, 1981.
[88] A. E. Kennelly, "Equivalence of triangles and stars in conducting networks," *Electrical World and Engineer,* vol. 34, pp. 413-414, 1899.
[89] F. Gardner, "Charge-Pump Phase-Lock Loops," *Communications, IEEE Transactions on,* vol. 28, pp. 1849-1858, 1980.
[90] D. Banerjee, Ed., *PLL Performance, Simulation, and Design*. National Semiconductor, 2006, p.^pp. Pages.

[91] P. V. Brennan and I. Thompson, "Phase/frequency detector phase noise contribution in PLL frequency synthesiser," *Electronics Letters,* vol. 37, pp. 939-940, 2001.
[92] B. Razavi, *Design of Analog CMOS Integrated Circuits*, 1st ed. New York: McGraw-Hill, 2001.
[93] A. Hajimiri*, et al.*, "Jitter and phase noise in ring oscillators," *Solid-State Circuits, IEEE Journal of,* vol. 34, pp. 790-804, 1999.
[94] T. Rui and M. Berroth, "5 GHz voltage controlled ring oscillator using source capacitively coupled current amplifier," in *Silicon Monolithic Integrated Circuits in RF Systems, 2003. Digest of Papers. 2003 Topical Meeting on*, 2003, pp. 45-48.
[95] Z. Xuan and A. B. Apsel, "A low variation GHz ring oscillator with addition-based current source," in *ESSCIRC, 2009. ESSCIRC '09. Proceedings of*, 2009, pp. 216-219.
[96] L. Hai Qi*, et al.*, "A Low-Noise Multi-GHz CMOS Multiloop Ring Oscillator With Coarse and Fine Frequency Tuning," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on,* vol. 17, pp. 571-577, 2009.
[97] P. Andreani and S. Mattisson, "On the use of MOS varactors in RF VCOs," *Solid-State Circuits, IEEE Journal of,* vol. 35, pp. 905-910, 2000.
[98] D. Liang and R. Harjani, "Design of low-phase-noise CMOS ring oscillators," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on,* vol. 49, pp. 328-338, 2002.
[99] T. Bourdi, *CMOS Single Chip Fast Freqeuncy Hopping Synthesizers For Wireless Multi-Gigahertz Applications*: Springer, 2007.
[100] P. Larsson, "High-speed architecture for a programmable frequency divider and a dual-modulus prescaler," *Solid-State Circuits, IEEE Journal of,* vol. 31, pp. 744-748, 1996.