AN ULTRA LOW POWER NB-IOT RF POLAR TRANSMITTER IN 65nm CMOS TECHNOLOGY

By
Ahmed Abdulraouf Awad Ahmed
Ahmed Magdy Abdelmoneim Elsayed
Amgad Osama Ahmed Shehab
Mohamed Abdelrahman Bakr Elbadry
Mohamed Hesham Mahrous Khattab

Under Supervision of: Dr. Hassan Mostafa

Technically Sponsored by: Si-Vision, LLC

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
Bachelor of Science in
Electronics and Electrical Communication Engineering

July 2018

Acknowledgments

We would like to thank Dr. Hassan Mostafa for his guidance and providing the necessary tools and climate for the completion of this work.

Additionally, we want to express our gratitude for Eng. Ahmed Salah, VP of Technology at Si-Vision for his continuous and tremendous efforts and support. Without his advices and follow-up, this work would not have been finished on time.

Moreover, we want to demonstrate how grateful we are to all Professors and TA's who taught us throughout our five years of higher education for their endless support.

We also say thanks to our parents whose patience and psychological support were critical to our success. We thank our families and friends, too, whose motivation helped us a lot during the many hard times we had. . .

Abstract

The *Narrowband Internet of Things* is a new radio interface firstly introduced by 3GPP in Release 13. NB-IoT is designed for narrowband, Low-Power Wide Area Networks (LPWAN). It provides good coverage for indoor low-data-rate and low-powered devices as well. This project focuses on the RF PHY layer of a NB-IoT transmitter. In this work, a polar transmitter based on the offset PLL architecture was implemented. The Transmitter works at 1.8G band which supports also several other standards like GSM, and E-UTRA. The designed transmitter proved to meet the required spectral mask with an rms EVM of **1%**. It supports a maximum output power level of **14** dBm consuming less than **90** mW from **1.2V** supply. The noise transmitted at an offset of 100 KHz is -89.9 dBc.

Table of Contents

List of Figures

List of Tables

Chapter 1 : Introduction

1.1 Introduction to Internet of Things

The Internet of Things (IoT) is the network of physical devices, vehicles, home appliances, and other items embedded with electronics, software, sensors, actuators, and connectivity which enables these things to connect and exchange data, creating opportunities for more direct integration of the physical world into computer-based systems, resulting in efficiency improvements, economic benefits, and reduced human exertions.

The number of IoT devices increased 31% year-over-year to 8.4 billion in 2017 and it is estimated that there will be 30 billion devices by 2020 [1]. The global market value of IoT is projected to reach $7.1 trillion by 2020. IoT involves extending internet connectivity beyond standard devices, such as desktops, laptops, smartphones and tablets, to any range of traditionally *dumb* or non-internet-enabled physical devices and everyday objects. Embedded with technology, these devices can communicate and interact over the internet, and they can be remotely monitored and controlled.

IoT targets to potentially create the integration of different wireless technologies, and subsequently will create market for new services. Some of the existing PHY layer protocols related to wireless IoT are IEEE 802.15.4, IEEE 802.15.6, Bluetooth Low Energy (BLE), EPCglobal, LTE-A, Z-Wave, 6LowPAN, and Near Field Communication (NFC).

Future 5G networks may need to ensure the rapidly emerging requirements of IoT applications. Some relevant Quality of Service (QoS) requirements include spectral efficiency, energy efficiency, connectivity and latency. To meet these diverse requirements, an efficient, scalable and flexible air-interface is required and, therefore, different modules of Physical (PHY) and Medium Access Control (MAC) layers should be optimized so that they can be configured flexibly according to the technical specifications of each use case. One of the important aspects in this regard is the design of PHY layer for IoT-based wireless systems considering the practical constraints of energy efficiency, spectral efficiency, cost-effectiveness, and quality of experience.

However, the design of IoT-enabled wireless networks which can deliver a variety of services with desirable quality of experience under energy/resource constrained practical wireless scenarios is crucial. In contrast to other wireless communication paradigms, IoT has its own unique features and diverse requirements, as shown in Fig.1.1, such as group-based communication, time-tolerant, small data transmission, secure connection, monitoring surrounding environment/parameters, low cost and low energy consumption. Besides, several

requirements such as bandwidth, reliability and latency of different existing services are highly diverse. In terms of connectivity, it's challenging to find out which devices need to be connected and which communication technology is suitable to connect them. Furthermore, several other issues such as dynamic resource allocation, harmful interference mitigation and interoperability of different technologies have to be investigated while devising communication technologies for IoT.

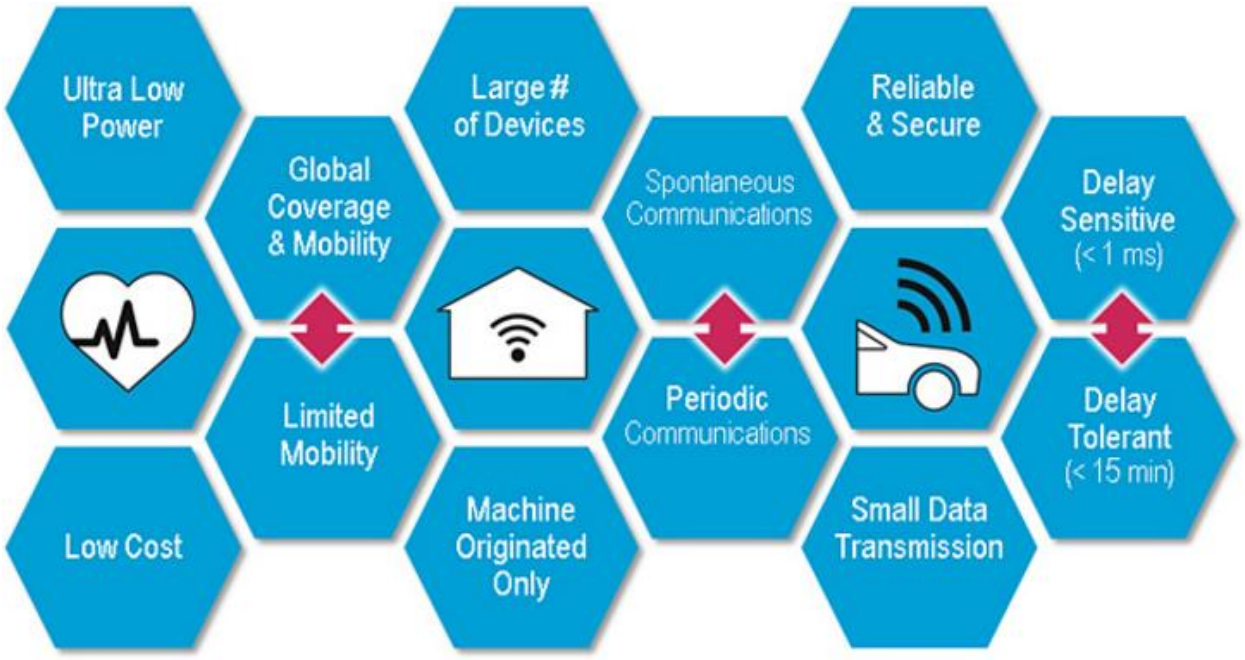


Figure 1.1: Main IoT PHY layer characteristics

1.2 Development of Narrow Band IoT Standard

One of the characteristics of Machine Type Communication (MTC) is the broad spectrum of capabilities. For example surveillance cameras have to deliver a huge amount of data while being almost stationary, whereas devices for fleet tracking have a small amount of data while performing a lot of handovers.

Yet another class of devices has neither of these capabilities. Examples are devices for meter reading like electricity, gas, or water consumption. They are often stationary, thus need not an optimized handover. Only a small amount of data is usually transferred, which is even not delay sensitive. However, the number of these MTC devices may become quite big, even up to several orders of magnitude compared to the traditional devices. Using existing LTE technology would lead to a network overload, because despite of their small amount of user data the amount of signaling is about the same. The first specification of NB-IoT focuses on this class of devices.

These devices are often installed at places without power supply. Consequently they run completely on battery and it may be very expensive to change the battery, because they may only be accessed by trained staff. Hence, in some cases the battery lifetime can even determine the lifetime of the whole device. Optimized power consumption is therefore essential for a proper

operation. In addition, the coverage at these places is often quite bad. Therefore, the indoor coverage has to be significantly improved, up to 23 dB are regarded as necessary. Finally, due to their sheer amount of required devices, they have to be in the low cost range. As a goal, each module shall be in the price range of less than 5 US$ [2].

In order to reach devices achieving all these requirements, the 3rd Generation Partnership Project (3GPP) organization, which unites seven telecommunications standard development organizations, worked on three different IoT standard solutions; LTE-M, NB-IoT, and EC-GSM.

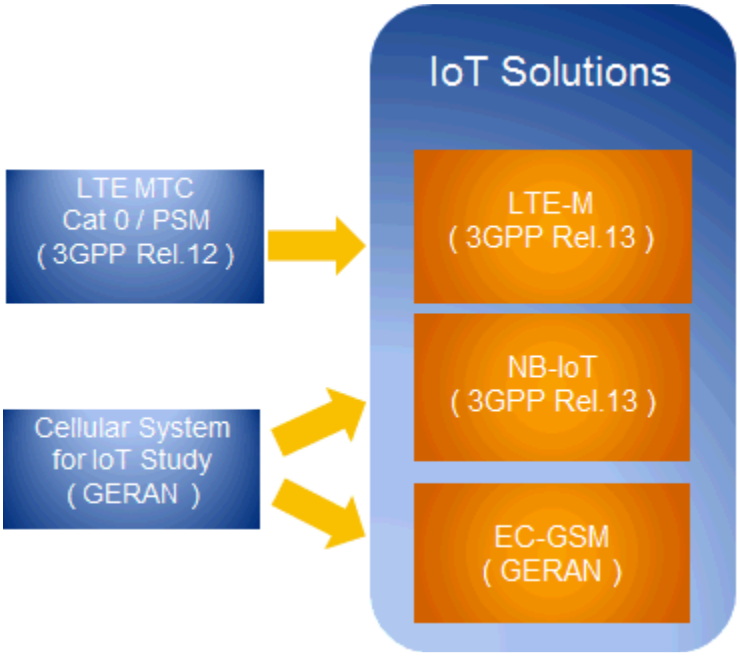


Figure 1.2: Three different solutions for specifying an optimized Internet of Things standard

The normative phase of NB-IoT work item in 3GPP started in September 2015 [2] and the core specifications were completed in June 2016 in Release 13. Commercial launch of NB-IoT products and services started around the end of 2016 and the beginning of 2017.

1.3 Narrow Band IoT (NB-IoT) Classification and Uses

The Narrow-Band Internet of Things (NB-IoT) is considered as a massive Low Power Wide Area (LPWA) technology proposed for data perception and acquisition intended for intelligent low-data-rate applications. It is considered also a new 3GPP radio-access technology in the sense that it is not fully backward compatible with any existing 3GPP devices; however it is designed to achieve excellent co- existence performance with legacy GSM, General Packet Radio Service (GPRS) and LTE technologies. Therefore, it is a promising technology as it doesn't need any new infrastructure to be operated; it can operate using the present base stations and equipment allowing fast low-cost deployment.

NB-IoT is intended to be used in applications that require low rate of signaling and long life time. Examples of such applications include:

- Smart metering for electricity and gas consumption.
- Earth global observation to monitor the rates of emissions, pollutants and different gases in air, to monitor also forests fire destruction and to quantify the purity levels of water.
- Personnel usage by individuals for applications regarding tracking children remotely.
- Fire Alarms in factories and industries as well as the optimization of the supply chain performance.
- Access monitoring and control of the smart homes.
- Smart lighting and camera systems in smart cities
- It can also be used in remote payment transactions.

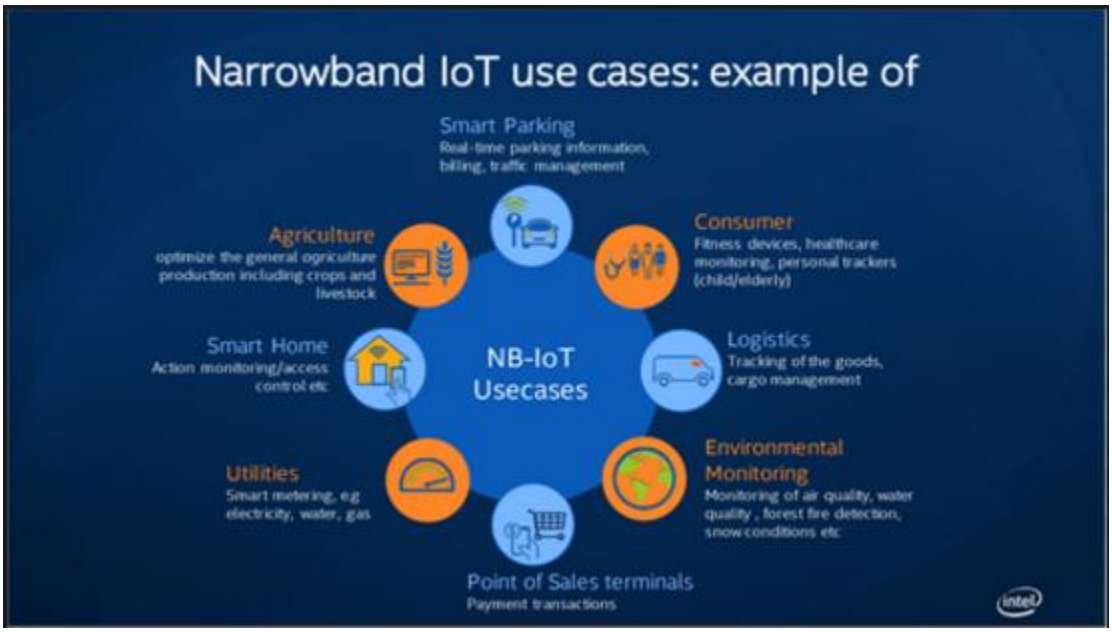


Figure 1.3: NB-IoT use cases

1.4 NB-IoT Spectral Issues

The RF bandwidth of NB-IoT physical layer is 200 KHz, only 180 KHz of them are allowed for transmission and the rest of the bandwidth acts as a guard as shown in Fig.1.4. So, a single NB-IoT carrier spans bandwidth of 180 kHz in uplink and 180 kHz in downlink with Frequency Division Duplexing (FDD). And to reduce RF front end costs, half duplex operation has been chosen in Release 13, so devices can only transmit or receive at any one time, unlike LTE where full duplex is supported

In downlink, NB-IoT adopts QPSK modem and OFDMA technology with sub-carrier spacing of 15 KHz, while in uplink, BPSK or QPSK modem and SC-FDMA technology including single sub-carrier and multiple subcarrier are adopted. A single sub-carrier technology with sub-carrier spacing of 3.75 kHz and 15 kHz is applicable to IoT terminal with ultra-low rate and ultra-low

power consumption. It can be noted that the coverage ability for 3.75-kHz spacing is higher than that for 15-kHz spacing because of the higher power spectral density. Also, the 3.75-kHz sub-carrier spacing provides larger system capacity than 15-kHz sub-carrier spacing. However, in the in-band operation mode scene, 15-kHz sub-carrier spacing has better LTE compatibility because for 15-kHz sub-carrier spacing, the NB-IoT uplink frame structure (frame size and time slot length) is the same as the LTE network [3].

The uplink supports both single sub-carrier and multiple sub-carrier transmissions. In single sub-carrier transmission, the sub-carrier spacing can be 3.75 kHz or 15 kHz; while in multiple sub-carrier transmission, the sub-carrier spacing of 15 kHz is adopted. The terminals need to support both single sub-carrier and multiple sub-carrier transmissions for easier selection of suitable mode by base station.

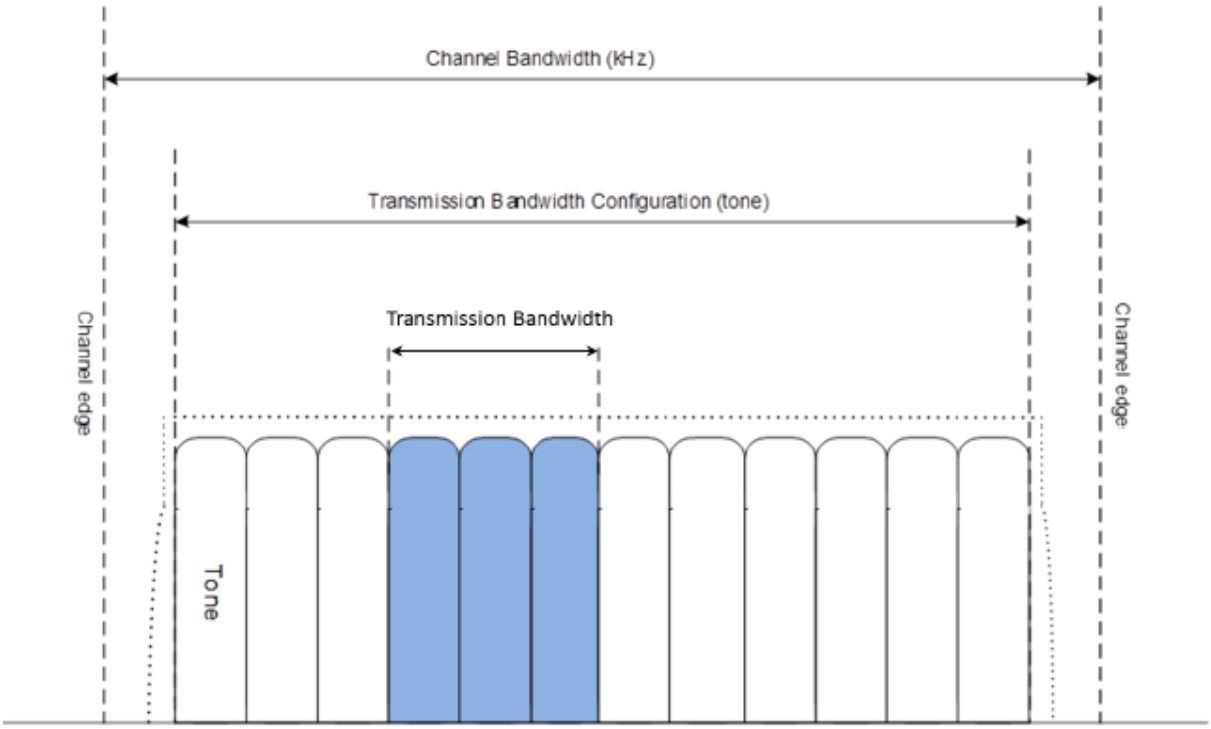


Figure 1.4: Channel Bandwidth and Transmission Bandwidth Configuration

LTE PHY layer standard specifies 19 bands for NB-IoT deployment. According to GSA's Evolution to LTE report – April 9, 2015, the mostly used LTE band among 393 commercially launched networks is the 1800 MHz band (B3) with 45% share as shown in Fig.1.5. Hence, this band (B3) is chosen for the NB IoT transceiver to be compatible with most of the deployed networks. This band specifies the range of frequencies from 1710 MHz to 1785 MHz as the uplink operating band and specifies the range of frequencies from 1805 MHz to 1880 MHz as the downlink operating band.

The NB-IoT system supports three deployment modes due to its extremely narrow bandwidth; independent deployment mode (standalone), guard-band deployment mode and in-band deployment mode, as shown in Fig.1.6.

- Independent deployment mode (standalone) utilizes frequency spectrum occupied by current GSM/EDGE wireless access network system to replace the existing single or multiple GSM carrier waves (Re-farming process). It is also called standalone deployment because it utilizes independent frequency band that does not overlap with the frequency band of LTE.
- Guard-band deployment mode utilizes the resource blocks which are not used for current LTE carrier like the wave guard band (edge band of the LTE).
- In-band deployment mode utilizes the resources blocks of LTE carrier wave. It takes 1 Physical resource block (PRB0 of LTE frequency band resources for deployment.

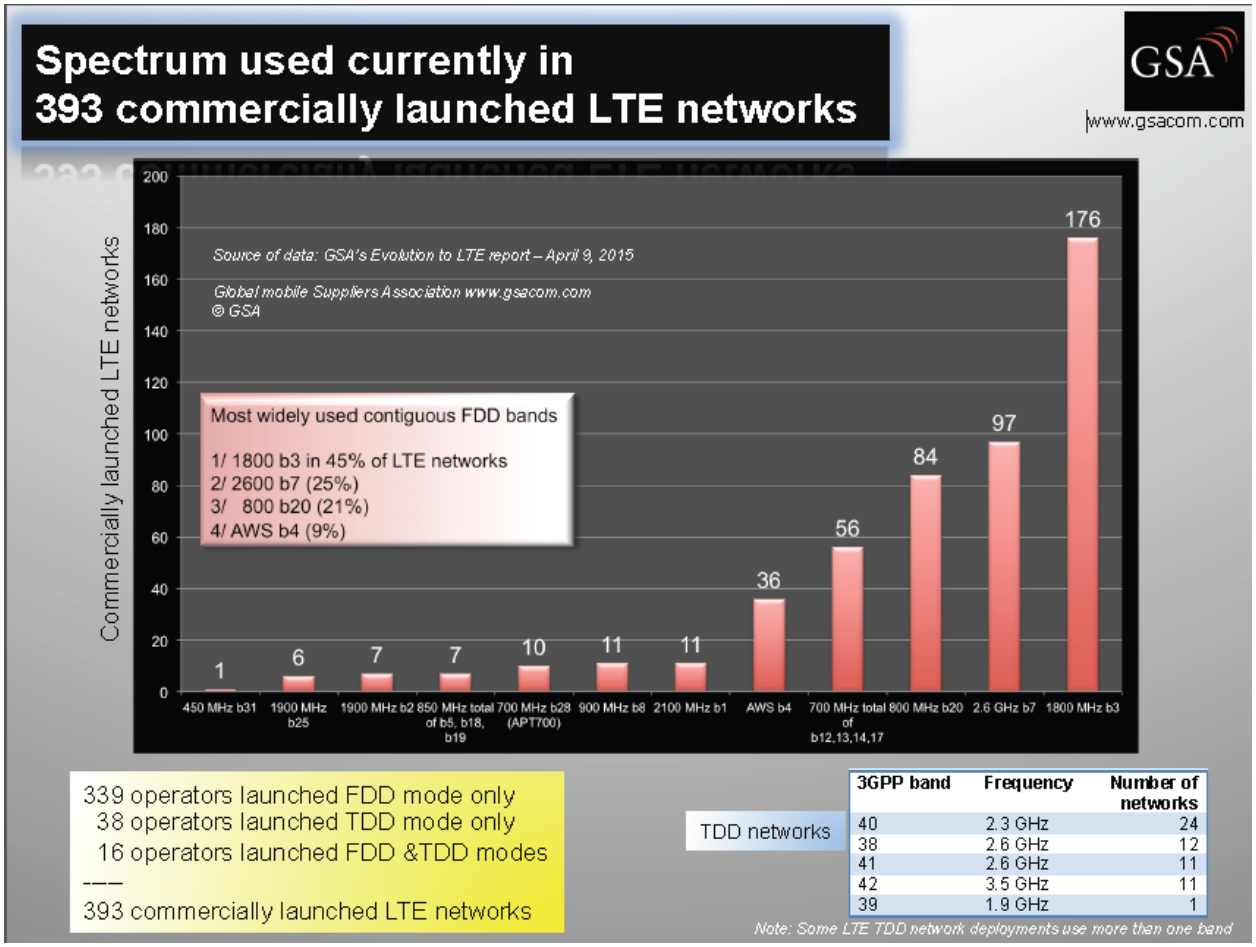


Figure 1.5: GSA report concerning the domination of networks operating at the 1.8G band.

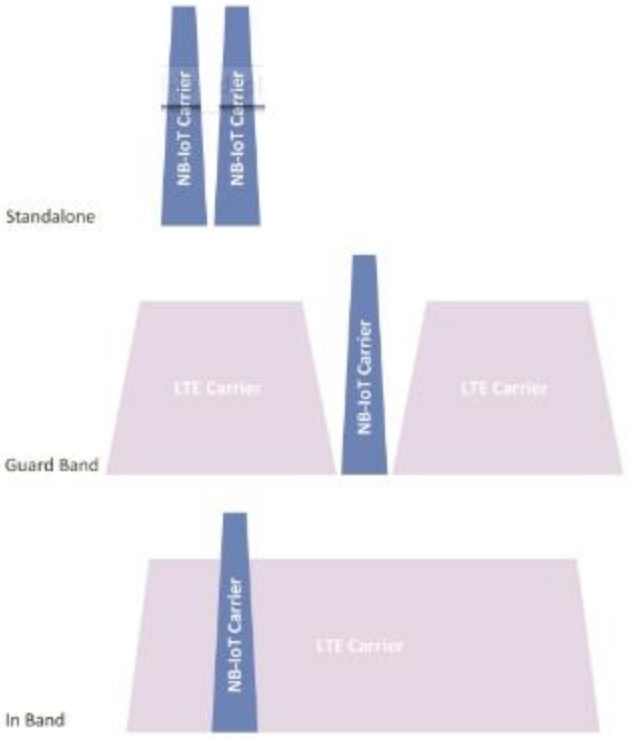


Figure 0.6: The three deployment schemes of NB IoT system

1.5 NB-IoT Transmitter Standard Specifications

The 3GPP organization issued the full technical requirements of the NB-IoT physical layer Transmitter in Release 14 [4]. These specifications concern on the quantification of the permitted level of out of band emissions, in band emissions, spurious emissions, Error Vector Magnitude (EVM), and maximum output power.

1.5.1 Out of Band Emissions

The out of band emissions specification in the standard was done on two levels; first of them is defining a certain power spectral mask. This mask represents the permitted power spectral density (PSD) profile that specifies the permitted amount of transmission in various frequency ranges around a given channel frequency. The goal of standardizing such a mask is to prevent transmitters of the standard from transmitting spectral components on other frequencies that could interfere with other transmissions and corrupt them. The spectral mask defined for NB-IoT transmitters is depicted in table 1.1.

Table 1.1: Power Spectral Mask

Δf_{ooB} (KHz)	Emission Limit (dBm)	Measurement Bandwidth
± 0	26	30 kHz
± 100	-5	30 kHz
± 150	-8	30 kHz
± 300	-29	30 kHz
$\pm 500 - 1700$	-35	30 kHz

The second level of defining the out of band emission level is defining the Adjacent channel leakage ratio (ACLR). ACLR is the ratio of the filtered mean power centered on the assigned channel frequency to the filtered mean power centered on an adjacent channel frequency. Since NB-IoT is deployed over existing GSM and UTRA bands, two ACLR specs are determined for both GSM and UTRA channels.

GSM's ACLR is 20 dB at a frequency offset ±200 KHz from the NB-IoT channel edge with 180 KHz adjacent channel measurement bandwidth. UTRA's ACLR is 37 dB at a frequency offset ±2.5MHz from the NB-IoT channel edge with 3.84 MHz adjacent channel measurement bandwidth. By comparing the ACLR requirements with the requirements of the power spectral mask; it can be noticed that achieving the mask is enough to achieve the ACLR requirements.

1.5.2 In Band Emissions

The in-band emission is measured, according to the standard, as the ratio of the UE output power in a non–allocated resource block to the UE output power in its allocated one. These in band emissions include the carrier leakage power and the image rejection criteria.

Carrier leakage is an additive sinusoid waveform that has the same frequency as the modulated waveform carrier frequency. The relative carrier leakage power is the power ratio of this additive sinusoid waveform to that of the modulated waveform and it's required to be as small as possible. Carrier leakage limits for different power levels is shown in table 1.2.

Table 1.2: NB-IoT Relative Carrier leakage power specifications

Power level	Relative Limit (dBc)
0 dBm < output power	-25
-30 dBm ≤ output power ≤ 0	-20
-40 dBm ≤ output power ≤ -30	-10

Regarding the image rejection criteria, the standard deals with the image as an in-band emission and puts specification regarding the image rejection to be greater than 25 dB rejection for both in-phase and quadrature signals.

1.5.3 Spurious Emissions

The standard defines the spurious emissions as the emissions which are caused by unwanted transmitter effects such as harmonics emission, parasitic emissions, intermodulation products and frequency conversion products. The specs for these emissions are applied in the spectrum parts which are not included by both spectral emission mask and ACLR requirements. These specified emission levels are given in table 1.3.

Table 0.3: NB-IoT Relative Spurious Emissions specifications

Frequency Range	Maximum Level	Measurement Bandwidth
9 KHz : 150 KHz	-36 dBm	1 KHz
150 KHz : 30 MHz	-36 dBm	10 KHz
30 MHz : 1 GHz	-36 dBm	100 KHz
1 GHz : 12.75 GHz	-30 dBm	1 MHz

1.5.4 Error Vector Magnitude (EVM)

The error vector magnitude (EVM), also known as constellation error, is a measure of the accuracy of the signal constellation of the transmitted signal and can be given in percentage or dB in wireless standards. This is effectively a measure of the receiver's ability to decode the signal, since errors in symbol positions on the constellation have a direct impact on the bit-error rate (BER) of the system.

The Error Vector Magnitude is a measure of the difference between the reference waveform and the measured waveform. For NB-IoT using QPSK modulation scheme, the EVM percentage allowed by the standard is $\leq$17.5%. EVM is affected and degraded by the effects of many factors such as Inter symbol Interference, LO phase noise, carrier leakage, I-Q mismatch [5]. Thus, a margin should be provided to overcome the mentioned non-idealities.

Fig.1.7 shows the deviation of the constellation points from their reference positions for QPSK modulation.

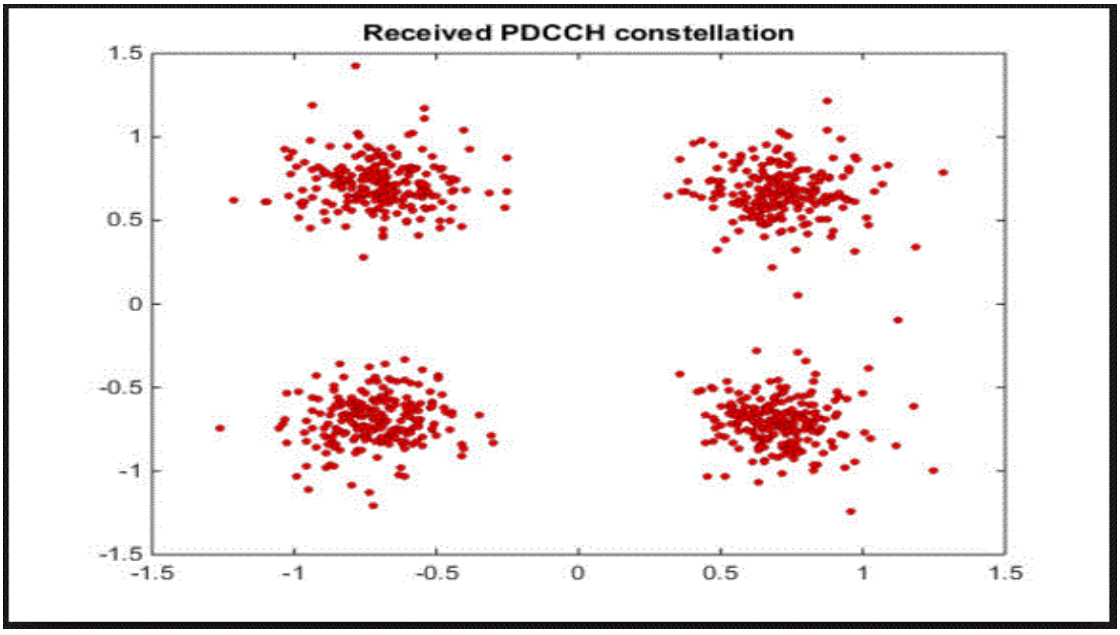


Figure 0.7: EVM for QPSK modulation

1.5.5 Transmitter Output Power Level

According to the NB-IoT standard, the permitted maximum power, when using the 1.8G band for uplink transmission, is classified into 3 classes according to the region of operation in order to obey the regional guidelines either concerning environmental codes or other technologies band requirements. These classes are:

- Class 3: 23dBm ± 2
- Class 5: 20dBm ± 2
- Class 6: 14dBm ± 2.5

Concerning the minimum power to be transmitted; it's specified to be at -40 dBm while the transmit-off power shouldn't exceed -50 dBm

References

[1] Hsu, Chin-Lung; Lin, Judy Chuan-Chuan (2016). "An empirical examination of consumer adoption of Internet of Things services: Network externalities and concern for information privacy perspectives". Computers in Human Behavior. **62**: 516–527.

[2] Qualcomm, Incorporated, "Narrowband IoT (NB-IoT)," RP-151621, 3GPP TSG RAN Meeting #69, Sept. 2015.

[3] Miao, Yiming, et al. "Narrow Band Internet of Things: Simulation and Modeling." IEEE Internet of things Journal (2017).

[4] 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (Release 14).

[5] Gu, Q. (2005). RF system design of transceivers for wireless communications. Springer Science & Business Media, 2005

Chapter 2 : Transmitter Architectures & System Design

2.1. Introduction

The choice of an architecture is determined by not only the RF performance that it can provide but other parameters such as complexity, cost, power dissipation, and the number of external components.

In the past ten years, it has become clear that high levels of integration improve the system performance along all of these axes. It has also become clear that architecture design and circuit design are inextricably linked, requiring iterations between the two [1].

2.2. Transmitter Architectures

In this section, a literature review is presented to reach the most suitable transmitter architecture for our applications.

2.2.1. Super-Heterodyne Transmitter

The super-heterodyne transmitter is a classical approach. In this architecture, as shown in Figure 2.1, the modulation is done at an intermediate frequency (IF) and subsequently upconverted to the transmit frequency by a mixer.

This architecture requires a significant amount of circuitry, typically including both an IF filter and an RF filter to overcome noise and spurious issues. In addition, this architecture requires a PA with good linearity, thus compromising the potential for high-power efficiency and complicating output power control [1].

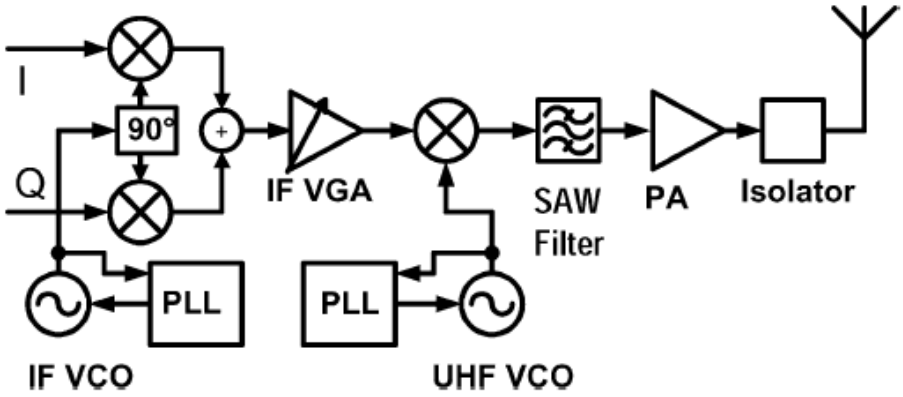


Figure 2.1: Super-heterodyne up conversion with IQ vector modulator

2.2.2. Direct Upconversion

The direct-conversion architecture (see Figure 2. 2) is used in many transmitter designs for different standards and modulation formats.

The amount of circuit blocks is reduced compared to the super-heterodyne transmitter, and, for several applications, this is the architecture that offers the best compromise of performance versus power consumption and circuit complexity.

Disadvantages

- The direct-conversion transmitter therefore often needs additional filtering after the PA.
- This solution requires a PA with good linearity, compromising the overall transmitter power efficiency and complicating output power control.
- A pre-driver to buffer the PA's heavy capacitive load.
- I/Q Mismatch Effect on constellation and EVM [2].
- DC offset comes from Digital baseband leads to carrier leakage [2].

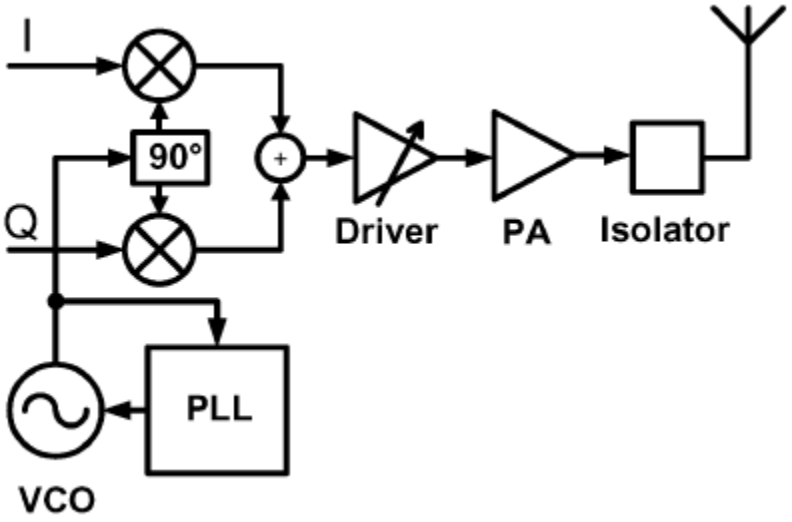


Figure 2.2: Direct Upconversion with IQ vector modulator

2.2.3. Phase Modulation Using a Fractional PLL

One approach in phase modulation is to modulate the phase using a fractional-N Sigma -Delta PLL where the PLL reference is a crystal oscillator depicted in Figure 2.3. The phase and envelope are extracted using a cordic processor. The phase is differentiated then applied to the sigma- delta modulator then to the PLL divider. It relies on that the PLL contains a VCO which acts as an integrator to restore the phase at the PLL output. This architecture is different from our choice to where the phase modulation is done on two steps using two PLLs in the system. The phase signal applied to the Sigma-Delta modulator is generated from a power hungry cordic processor [3]. The power reported in [4] shows that the cordic processor for a 12 bit word size will exceed 10 mW according to the implementation and the throughput.
The sigma delta modulator will require a high resolution sigma delta modulator to reduce the quantization noise. This would require more bits and more power consumption. This shows that adding a second PLL can save the power consumed by the cordic processor. Moreover, there are other issues that make it less attractive to use this architecture [5]. One issue is that separating the signal into phase and amplitude using the Cordic causes high out of band emissions. Hence, Finite Impulse Response (FIR) filters need to be designed. Another issue, is the need for high sampling rate to get a higher image rejection and spectrum purity. This increases the clock speed and the power consumption. It is shown that these filters, clock distribution circuits power consumption can reach 50 mW depending on the frequency of operation and number of bits [6]. Thus, this architecture would consume much higher power consumption than the proposed architecture.

Another transmitter architecture is the Digital Transmitter proposed in [6]. This transmitter architecture suffers from the same disadvantages related to power consumption in the fractional N Sigma- Delta PLL based transmitter.

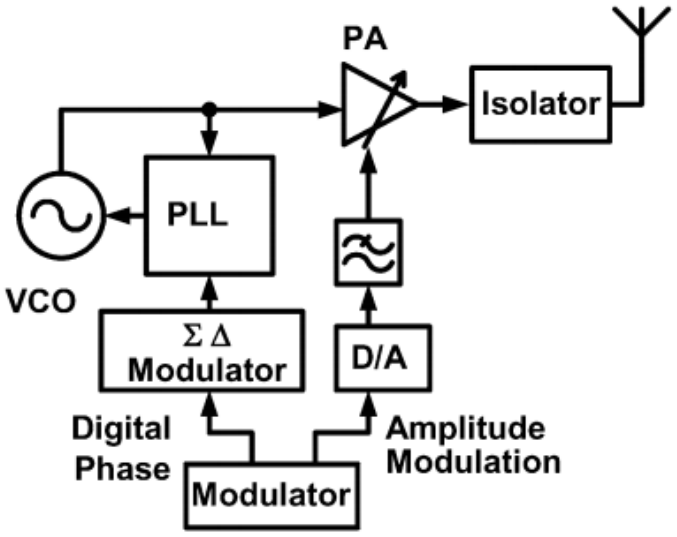


Figure 2.3: Fractional N Sigma-Delta PLL based Transmitter

2.2.4. Modulation By Offset PLLs

Another transmitter architecture is the offset PLL based modulation depicted in Fig. 2.4. The offset PLL transmitter proposed in [1] upconverts the baseband I and Q signals to an intermediate frequency (IF) then the phase of the upconverted signal is extracted using a limiter. This is done to avoid extracting the phase using the power hungry cordic processor. The IF frequency is chosen such that the limiter does not introduce AM/PM conversion [1]. The extracted phase is applied to the offset PLL depicted in Fig. 2.5. The offset PLL uses a mixer, filter and limiter in the feedback loop. This is done to translate the PLL output frequency the IF frequency without affecting the phase. Thus, this offset PLL is considered as a unity feedback loop. The phase at the PLL output tracks the phase at the PLL input. The envelope is extracted and applied to the PA supply.

2.2.5. Proposed Architecture

Our proposed architecture replaces the sigma delta modulator discussed in section 2.23 with the offset PLL and the IQ DACs with the upconversion mixers and the limiter for phase extraction. The envelope path consists of a DAC followed by a reconstruction filter and a DC/DC converter to control the PA supply. A divide by 2 is inserted after the VCO such that the VCO runs at twice the PA frequency and avoid VCO pulling. The divider by 2 circuit is inserted inside the loop in the forward path such that the PLL output phase tracks the extracted phase at the IF frequency. The PLL feedback mixer gets its LO frequency from another PLL. The PLL must be chosen such that its VCO runs at a different frequency from the PA and the offset PLL VCO by 20% [2]. Hence the IF frequency was chosen to be 250 MHz and the LO frequency at the feedback mixer is 1.5 GHz for an output frequency of 1.75 GHz. The PLL output is applied to a non-linear PA to drive the antenna. For the envelope path, the envelope is extracted in the digital domain without the need of the cordic processor. Then the envelope is applied to a DAC and an antialiasing Filter. To solve the issue of the delay mismatch, a delay will be inserted in the digital domain in the envelope path to compensate for the mismatch [2]. As will be shown in the next chapters, The phase path consumes 26.32 mW in the phase path and consumes only 6.78 mW in the PLL. Thus, adding another PLL means that the digital circuitry that consumes more than 50 mW [6] is replaced by analog circuitry that consumes only 35 mW. This means more than 15 mW reduction in the power consumption but on the expense of a larger area. In our approach, power was traded for area.

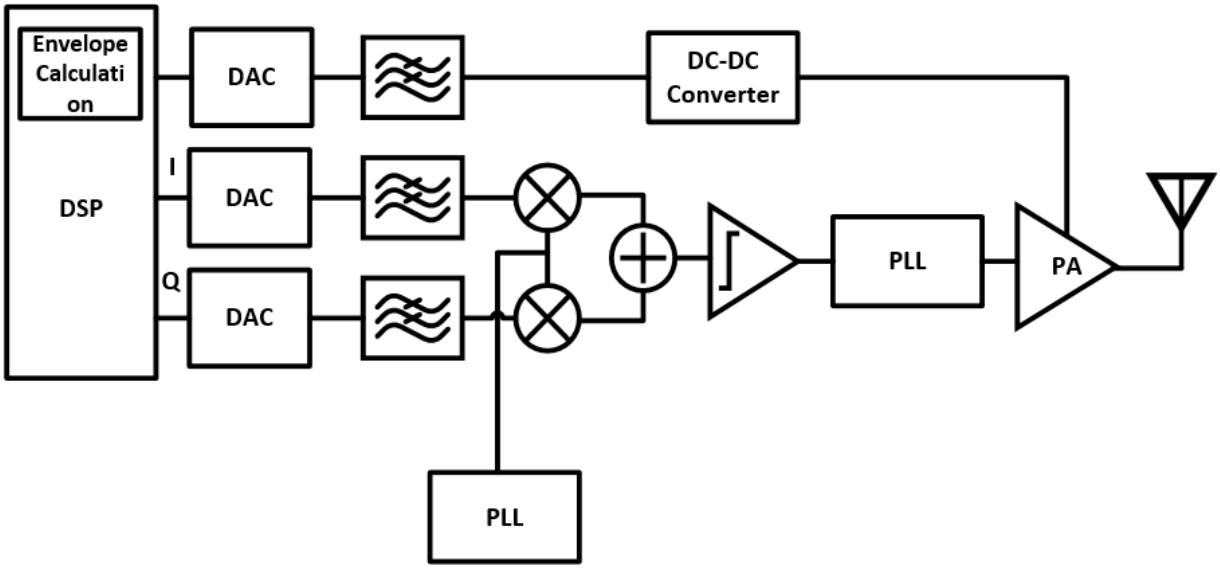


Figure 2.4: Offset PLL Transmitter

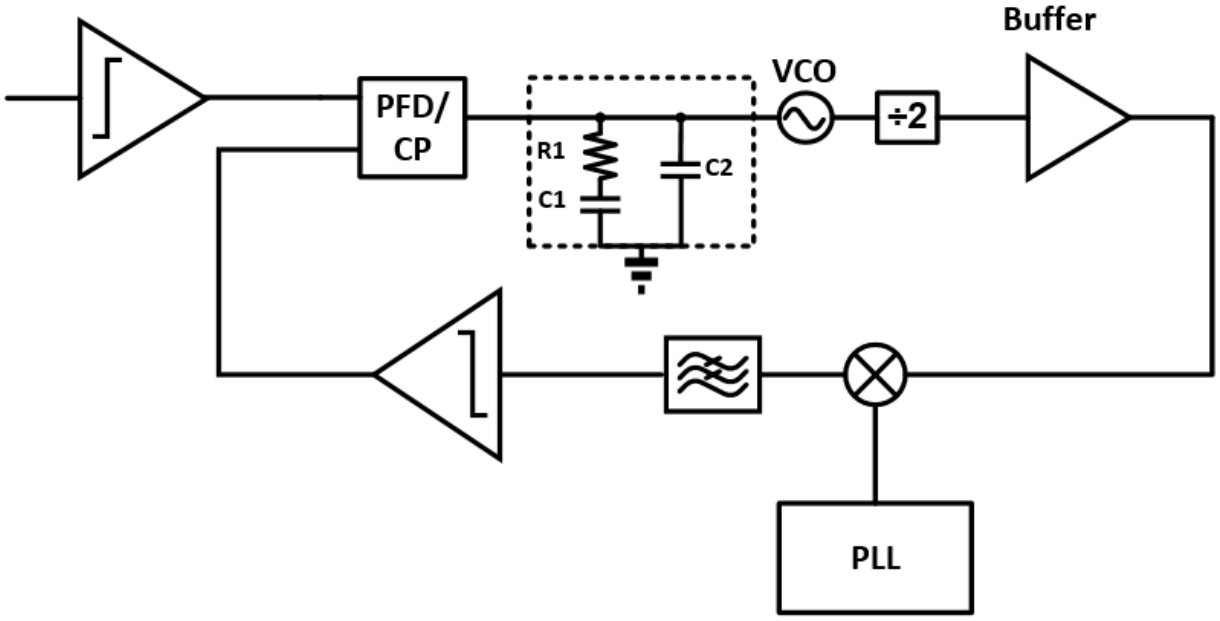


Figure 2.5: Offset PLL

2.3. System Design

The next step after choosing suitable architecture for the transmitter is performing system level design to reach the targeted gain, noise, swing, and linearity specs. Unlike receivers, RF transmitters do not have equations to relate the blocks' specs with the cascade specs. This requires extra efforts in performing the system design.

2.3.1. Emission Levels

The most critical issue in RF transmitter systems is the level of unwanted emissions to the neighboring channels [7]. Hence, a good starting point is the required level of unwanted emissions specified by the standard. The standard divides the unwanted emissions into two different categories. Firstly, the spectral emission mask is defining the emissions levels at a specific offset from the transmitting channel edge. This offset ranges from 0 to 1.7 MHz. The emissions in any part of the frequency spectrum other than the range specified by the spectral emission mask are defined by the spurious emission level in the standard. Spurious emissions are emissions, which are caused by unwanted transmitter effects such as harmonics emission, parasitic emissions, intermodulation products and frequency conversion products. Tables 2.1 and 2.2 illustrate the specified levels by the standard for both spectral emission mask and spurious emissions. An important note from the standard states that the spurious emissions requirements are applied for the frequency ranges that are more than 1.7 MHz. Thus, the focus will be on achieving the spectral emission mask specs.

Table 2.1: Spectral Emission Mask

Δf (KHz)	Emission Limit (dBm)	Measurement Bandwidth
± 0	26	30 KHz
± 100	-5	30 KHz
± 150	-8	30 KHz
± 300	-29	30 KHz
$\pm 500 - 1700$	-35	30 KHz

Table 2.2: Spurious Emissions

Frequency Range	Maximum Level	Measurement Bandwidth
$9\ KHz \leq f \leq 150\ KHz$	-36	1 KHz
$150\ KHz \leq f \leq 30\ MHz$	-36	10 KHz
$30\ MHz \leq f \leq 1000\ MHz$	-36	100 KHz
$1\ GHz \leq f \leq 12.75\ GHz$	-30	1 MHz

2.3.2. Design Procedure

From Table 2.1, the emission limit at offset 400 KHz from channel's center is -29 dBm. By subtracting the measurement bandwidth, the noise density is then -74 dBm/Hz. This value is required at the PA's output. For design simplicity, this spec is translated to the PLL's input by adding a 7-dB margin. Normalizing to the 14-dBm carrier power gives a normalized noise density of -95 dBc/Hz. The target now is to determine the amplitude noise level corresponding to this phase noise level. This was estimated by measuring the phase noise of ideal, noiseless limiter. The input of this limiter is the superposition of an ideal sinusoidal signal and a noise source, which is simply a resistance with noise power equals to 4KTR. Fig. 2.6 shows the setup illustrated. The resistance value is tuned until reaching the required phase noise value. Fig. 2.7 presents the output phase noise of the limiter at resistance value equals to 95 MΩ. The amplitude noise corresponding to this 150 KΩ resistance is 1.58 PV^2/Hz. This value will be considered as the sum of contributions of both limiter's input referred noise and the baseband chain output noise. A 0.1 fV^2/Hz input referred noise is assumed for the limiter which corresponds to 20 dB noise figure. The remaining noise for the baseband chain is 1.58 PV^2/Hz which is equivalent to -105 dBm/Hz. 3-dB is subtracted from this value to take into account both I and Q branches. Therefore, -108 dBm/Hz is the required output noise from the baseband chain.

The baseband chain includes: DAC, reconstruction filter, mixer and summer. A good point to start with is to assume several reasonable assumptions and then move from these assumptions to the definition of rest of the parameters. The assumptions used are:

- DAC full scale is 200 mV.
- Oversampling ratio (OSR) = 32 (Sampling frequency = 6.4 MS/sec.)
- Limiter Sensitivity (minimum accepted signal amplitude) is 50 mV.

As the full scale output from the DAC is 4 times the sensitivity of the limiter, then the signal must be exposed to attenuation by 1/4 of its strength along the path from the DAC to the limiter. This attenuation can be achieved across the mixer and summer by designing a mixer and summer stage with conversion gain of -10 dB (corresponds to attenuation of 1/4).

From literature, it was found that to implement a mixer and summer with conversion gain of -10 dB, the accompanied noise figure will be in the range of 45 dB. Therefore, when applying this fact with the mentioned assumptions, we get:

$$Noise\ level\ before\ mixer = Noise\ level\ after\ summer - Gain\ - NF$$

$$= -108 - (-10) - 45 = -143\ dBm/Hz \qquad (2.1)$$

This noise level must be achieved at offset of 500 KHz (PLL bandwidth). Moving to the filter, as the signal channel transmission bandwidth is limited to 90 KHz, the filter -3 dB point can be designed to be at 120 KHz in order to pass the signal without any attenuation. Now the specification on the attenuation of the filter at stop band can be created.

Filter approximation chosen is Butterworth approximation as it is characterized by its low group delay. According to this approximation, the stop band attenuation can be calculated as:

$$As = 10\log(1 + \epsilon^2\ \Omega s^{2n}) \qquad (2.2)$$

For ϵ =1, and Ωs = 500/120 = 4.2, it can be found that for a 4th order butter worth low pass filter, attenuation of 50 dB can be achieved at 500K. By assuming a noise figure factor of 45 dB for such filter, it can be found that the noise level before filter is calculated as:

$$Noise\ level\ before\ filter = \ Noise\ level\ after\ filter + Filter\ Attenuation - NF$$

$$= -143 + 50 - 45 = \ -138\ dBm/Hz \qquad (2.3)$$

This is the noise level required after the DAC, so in order to define the DAC SNR specification, this noise level needs to be integrated on the sampling bandwidth of the DAC (fs/2). So by integrating, it can be found that the noise level is -138 + 10 log (3.2M) = -72 dBm. With signal full scale amplitude of 200 mV (-4 dBm), we can easily find that the required SNR of the DAC is -4 – (-72) = 68 dB.

After defining the required SNR level of the DAC, the effective number of bits can be estimated using the following relation:

$$ENOB = \frac{SNR - 1.76 - 10\log(OSR)}{6.02} = \frac{68 - 1.76 - 10\log(32)}{6.02} = 8.5\ bits$$

Thus, the system requires 10 bit Digital to Analog Converter as a margin.

The PLL feedback path is characterized individually. Starting from the mixer, assuming that the mixer's input noise is above the thermal noise by 3 dB. Again, similarly to (2.1) and (2.3), gain and noise figure of the mixer and filter cascade are related by Gain + NF = 37 dB. The filter's gain is chosen to be 0 dB and the mixer's gain is chosen to be 3.5 dB. Thus the total noise figure of both blocks is 33.5 dB. A simple cascade analysis is made results into 25 dB noise figure for the mixer and 35 dB noise figure for the filter.

The frequency planning on the system level was applied in order to achieve reasonable design specifications on different components of the chain. The digital base-band channel has transmission bandwidth of 90 KHz, when a signal is transmitted; it is first up-converted to 250 MHz then introduced to the limiter which extracts only the phase information which is then introduced to the offset PLL to make it converted to its transmission 1.7 GHz band.

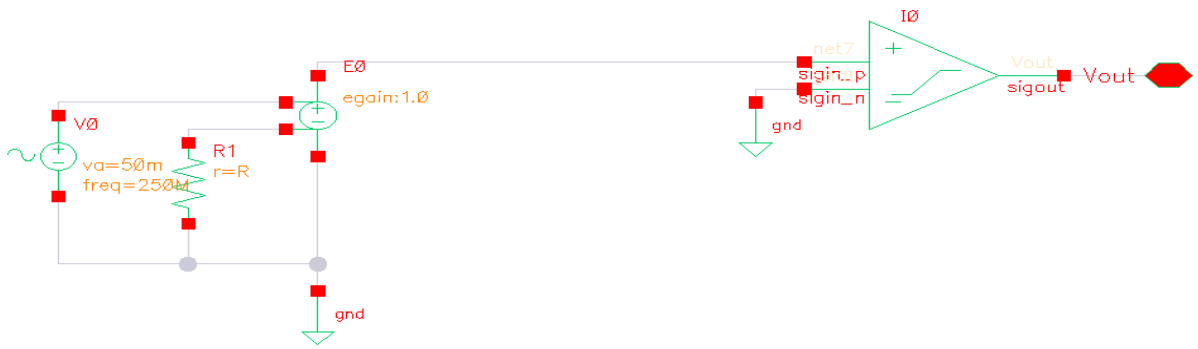


Figure 2.6: Noise Testbench

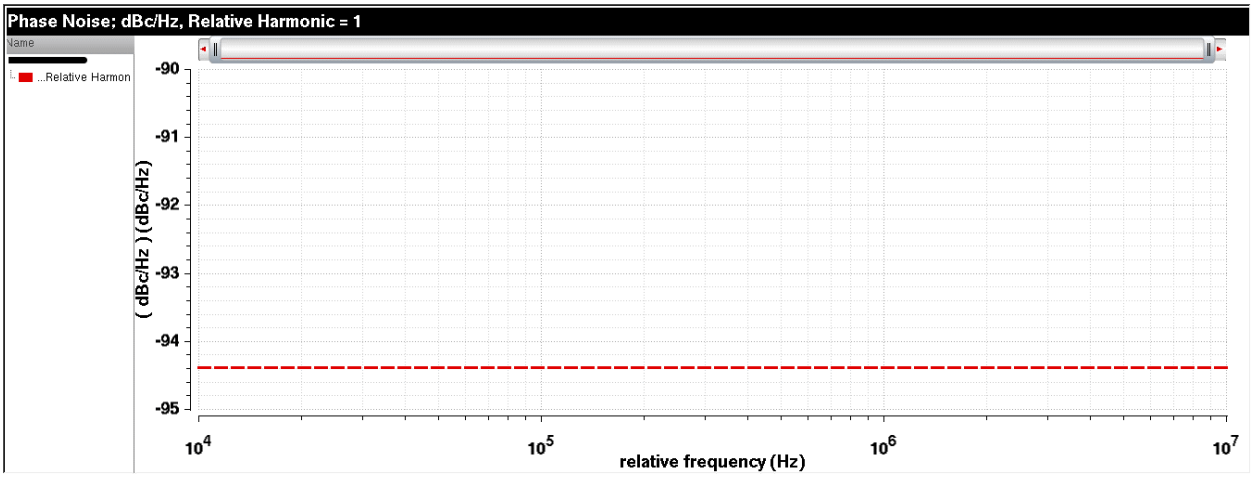


Figure 2.7: Output Phase Noise

References

[1] Tirdad Sowlati, Raja Pullela, Morten Damgaard, Evan McCarthy,“ Quad-Band GSM/GPRS/EDGE Polar Loop Transmitter” IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 39, NO. 12, DECEMBER 2004

[2] B. Razavi, “RF microelectronics”, Upper Saddle River, NJ: Prentice Hall, 2012

[3] Choi, D. Kang, D. Kim, J. Park, B. Jin and B. Kim, "Power Amplifiers and Transmitters for Next Generation Mobile Handsets", JSTS:Journal of Semiconductor Technology and Science, vol. 9, no. 4, pp. 249-256, 2009.

[4] J. Manica, R. Passerone and L. Rizzon, "Power-aware architectural exploration of the CORDIC algorithm", Proceedings of the 2013 9th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME), 2013.

[5] Z. Song, X. Liu, X. Zhao, Q. Liu, Z. Jin, and B. Chi, “A Low-Power NB-IoT Transceiver With Digital-Polar Transmitter in 180-nm CMOS,” IEEE Trans. Circuits Syst. I Regul. Pap., vol. 64, no. 9, pp. 2569–2581, 2017.

[6] D. Chowdhury, L. Ye, E. Alon, and A. M. Niknejad, “An efficient mixed-signal 2.4-GHz polar power amplifier in 65-nm CMOS technology,” IEEE J. Solid-State Circuits, vol. 46, no. 8, pp. 1796–1809, 2011.

[7] B. Razavi, "RF transmitter architectures and circuits", Proceedings of the IEEE 1999 Custom Integrated Circuits Conference (Cat. No.99CH36327), 1999.

Chapter 3 : Digital to Analog Converter (DAC)

3.1 DAC Overview

Digital-to-Analog converters (DAC) are key building blocks in many mixed-signal system-on-chip (SoC) designs. They are responsible for providing the interface between the digital and the analog domains within the chip as well as interfacing the system to the real world. Fig. 3.1 illustrates the basic fundamental operation of a DAC where a stream of digital samples is reconstructed to a piecewise constant output analog signal. The output of the DAC is usually smoothened with a reconstruction filter depending on the target application.

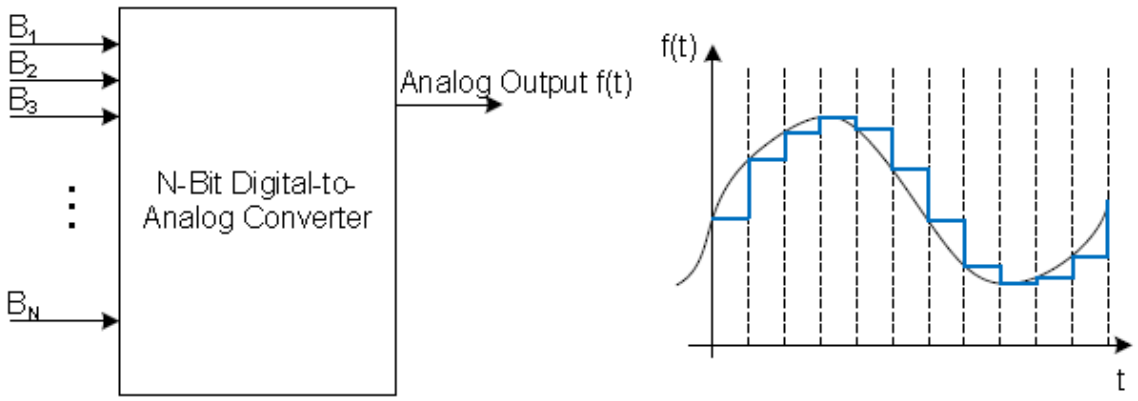


Figure 3.1: Ideal DAC output without reconstruction filter

The performance of DAC often proves to be the bottleneck in electronic systems. With the growing demand for high bandwidth communications, it is crucial to develop high speed converters. Great care is needed to design a data converter that is fast, accurate, consumes little power, and takes up minimum area. Historically high speed data converters were implemented using technologies such as SiGe or GaAs [1]. The downside to these technologies is that they are expensive to manufacture, consume large amounts of power, and cannot be integrated on the same chip as traditional CMOS technology. So, to counter these trade-offs, research has been done to implement high speed converters using standard CMOS technologies.

Main concept of operation of DAC is taking a certain binary word, and outputting a reference voltage corresponding to its decimal weight. These reference voltages create a pulse amplitude modulated signal that has a staircase-like pattern. The number of binary inputs the DAC receives is known as the *resolution.* For an N bit resolution DAC, there are 2^N possible output voltage levels. The binary word input D can be characterized as:

$$D = \sum_{m=0}^{N-1} 2^m (b\,m) \qquad (3.1)$$

where m is the index of the binary word, and b is the '1' or '0' value of the bit at that index.

Another important term to be highlighted is the *reference* voltage of the DAC which is the analog value from which all the outputs are scaled. With this full scale voltage V_{ref} , the analog output at any given digital input can be determined by:

$$vout = Vref * \frac{D}{2^N} \quad (3.2)$$

The least significant bit (LSB) is the bit b_0 in the digital word. In converters, the term *one LSB* is associated with the minimum step size in the analog output. The LSB can be defined as:

$$LSB = \frac{Vref}{2^N} \quad (3.3)$$

Since the minimum analog output voltage corresponds with the digital word consisting of all '0's, the maximum analog output voltage will be one LSB less than V_{ref} . The range from the minimum output voltage to the maximum is known as the *full scale range*. The full scale voltage V_{FS} can be calculated as:

$$V_{FS} = \frac{2^N - 1}{2^N} * Vref \quad (3.4)$$

The output of the DAC can be described as a series of rectangular functions. This is because the DAC acts as a sample hold, holding the output level for 1/fs seconds, where fs is the sampling frequency [2]. The rectangular pulse in time domain will form the function a sinc function in the frequency domain. The frequency response characteristic of the DAC output can be expressed as:

$$H(f) = \frac{1}{fs} * \frac{\sin\left(\frac{\pi\, fin}{fs}\right)}{\frac{\pi\, fin}{fs}} \quad (3.5)$$

This output characteristic is known as sinc weighting, which can be problematic for the converter. Fig.3.2 shows the frequency response of the DAC output. The problems that arise from such response can be categorized into two main problems:

- Magnitude attenuation across frequency increase. The magnitude of the signal is attenuated by 3.92 dB at the Nyquist frequency, and reaches zero at the sampling frequency.
- The spectrum replica each fs/2 which creates signal aliases that accordingly causes high distortion in the output signal.

In order to overcome these problematic effects of the sinc weighting response of the DAC; oversampling technique is used, where the signal is sampled in the digital domain to frequency much greater than the Nyquist frequency. This technique makes the attenuation that accompanies the roll-off of the DAC frequency response be shifted away from the signal bandwidth. So, the signal can pass through the DAC without attenuation. The oversampling ratio (OSR) can be obtained simply as the ratio between the sampling frequency and the Nyquist one. Fig. 3.3 gives better insight on the effect of oversampling on the spectrum.

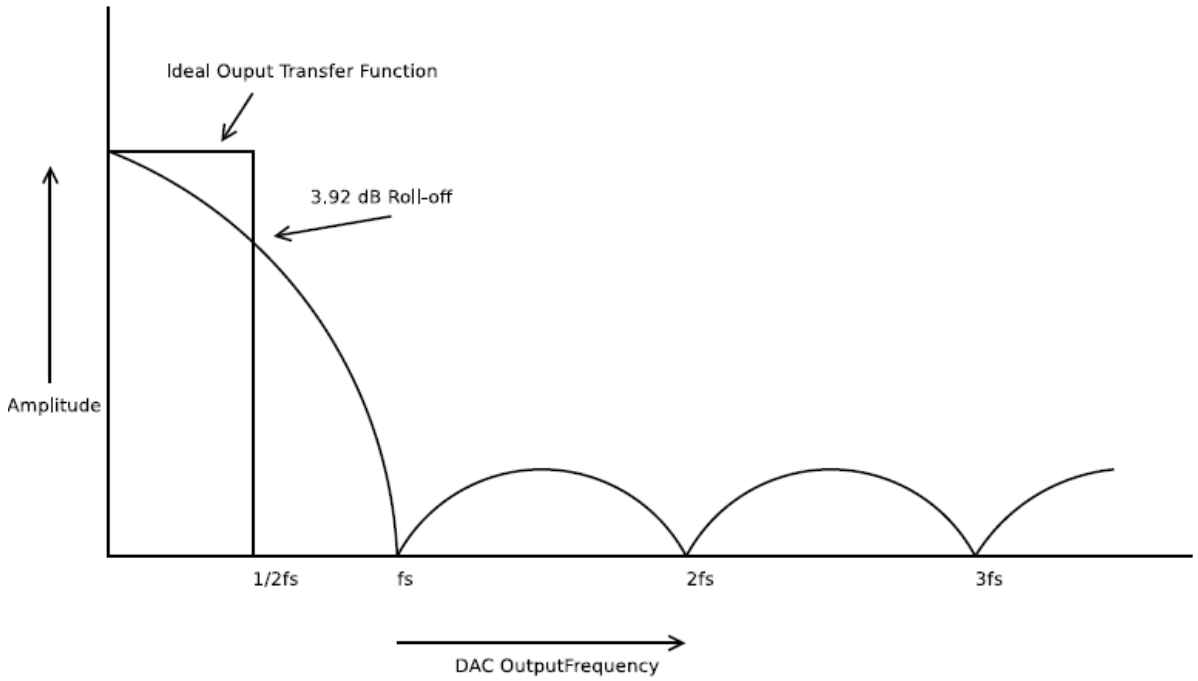


Figure 3.2: Sinc Weighted output transfer function

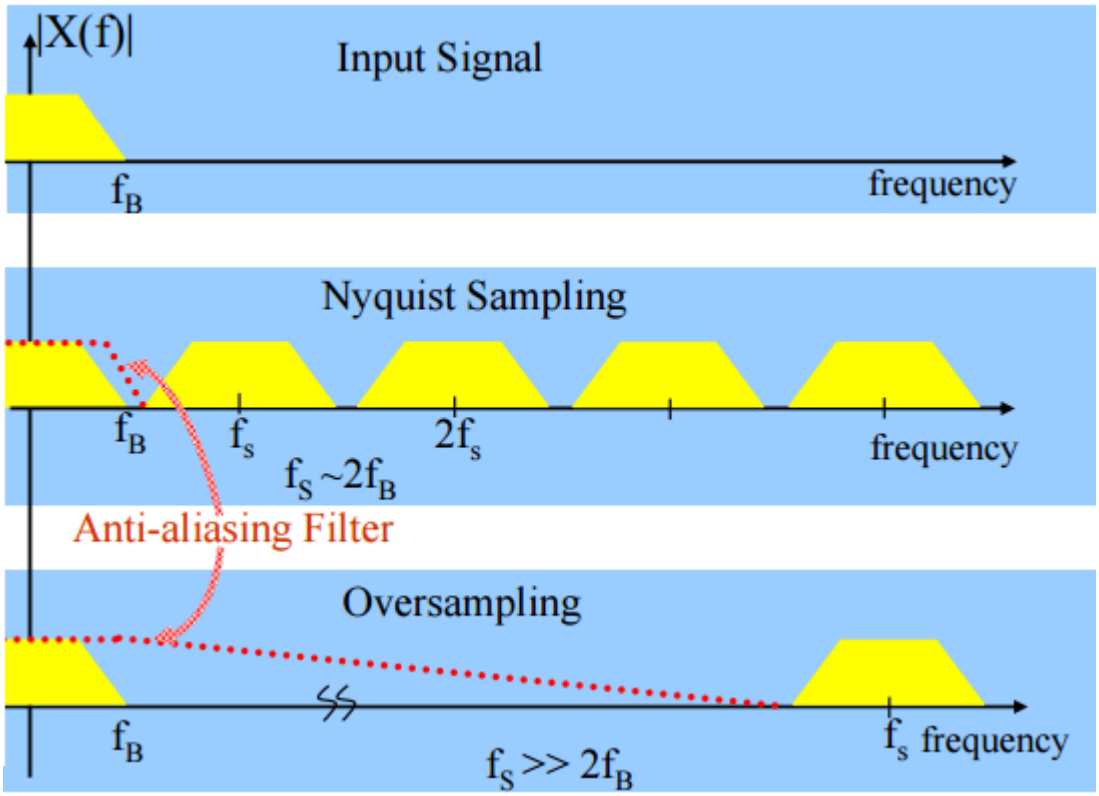


Figure 3.3: Effect of oversampling on the DAC output spectrum

Oversampling also has a great contribution in the anti-aliasing process; as due to oversampling the aliases are much separated from the main signal giving opportunity for designing a reasonable reconstruction filter to get rid of these aliases without affecting the signal itself. The reconstruction filter smooths the rectangular output of the DAC into a smooth sine-like output.
An ideal reconstruction filter would have a brick wall response [3], attenuating everything passed the Nyquist frequency. In practice, a brick wall filter cannot be designed. Designing a filter with a large amount of attenuation in a short transition band requires an analog filter consisting of many orders. High order filters add complexity and increase the area of the design. Increasing the transition band of the filter reduces the filter complexity, but the bandwidth of the DAC will be sacrificed.

In this work, full procedure for designing a digital to analog converter with a reconstruction filter will be presented through the following sections.

3.2 DAC Performance metrics

The performance of a DAC can be expressed in terms of its static performance and dynamic performance. Static performance metrics analyze the accuracy of the converters output versus the expected output at a given code. Typically the input is swept to produce the full range of binary codes, so the DAC output appears as the characteristic staircase plot. The voltage level at each code is measured to determine the accuracy. Measuring the performance this way provides good insight on how accurate the DAC is once the output settles.

However, the dynamic performance is more important. The dynamic accuracy of the DAC is measured by applying a sinusoidal input to an ideal analog to digital converter, which then supplies the DAC with the digital corresponding inputs. The DAC then reconstructs the sinusoid, and then FFT analysis can be done to see how the DAC performs.

3.2.1 Static Performance Measures

Static performance measures include: Differential Non-Linearity (DNL), Integral Non-Linearity (INL), Gain error, Offset error, and monotonicity.

3.2.1.1 Static Monotonicity

If a DAC is monotonic, then the output always increases as the input increases. It is normally not necessary to state such a condition for a DAC explicitly to have this behavior. In order to ensure this property in the DAC, specs regarding DNL and INL need to be defined as stated in sec. 3.2.1.2. Fig. 3.4 shows obviously the difference between a monotonic DAC and non-monotonic one. For the ideal DAC, each code has an analog corresponding voltage greater than the previous code. While for the non-monotonic one, the corresponding voltage decreases when moving from code 010 to code 011.

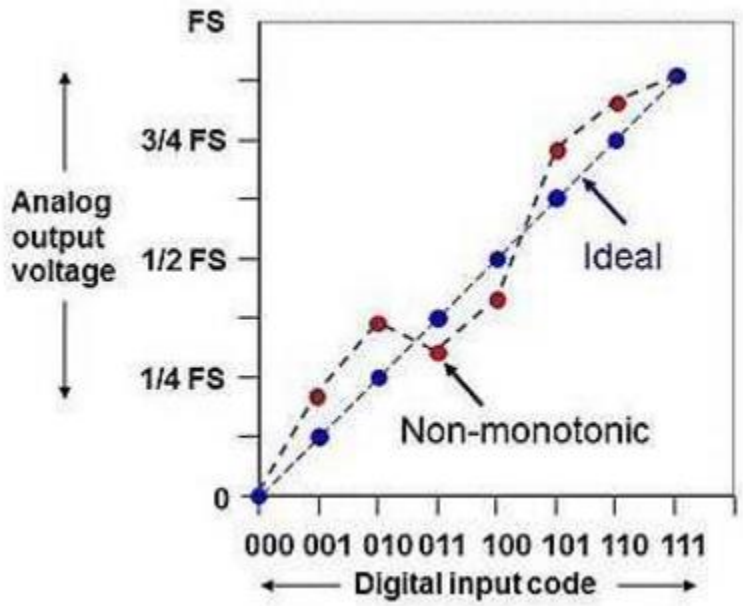


Figure 3.4: Comparison between ideal and non-monotonic 3 bit DAC

3.2.1.2 Static Linearity

The two main static linearity performance metrics are the differential nonlinearity (DNL), and the integral nonlinearity (INL).

Differential nonlinearity is the difference of the output level between two adjacent codes. Often times DNL is measured in terms of LSB, in an ideal case the difference between two codes is 1 LSB. If the step between two adjacent codes is 1.25 LSB instead of 1 LSB, then the DNL is said to be 1/4 LSB. In order to maintain sufficient accuracy, the DNL must be between -1/2 to +1/2 LSB. In order to guarantee the output is monotonically increasing, the DNL should never be greater than 1 LSB [4]. Non-monotonic behavior can produce large unacceptable errors in the output.

Integral nonlinearity is the measure of the actual output voltage level minus the ideal level. One way to measure INL is to sweep the digital input and plot the analog output. A line is then drawn from zero to full scale; the output deviation from this line is the INL. The INL per code can also be defined as the summation of the DNL from all previous codes. As with DNL, the acceptable range of values for the INL is within $\pm$ ½ LSB. In order to maintain monotonicity in the output, the INL must be less than 1 LSB for every code.

Fig.3.5 and Fig.3.6 clarifies the definition of DNL and INL respectively on a 3 bit Digital to Analog Converter with maximum DNL of 0.75 LSB and maximum INL of 0.75 LSB as well.

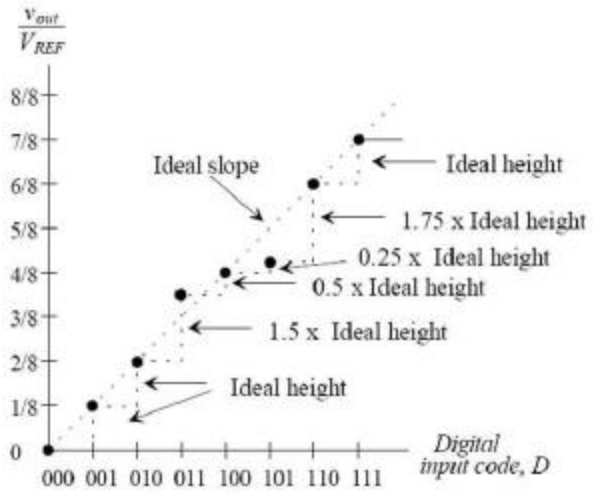


Figure 3.5: DNLof a 3 bit DAC

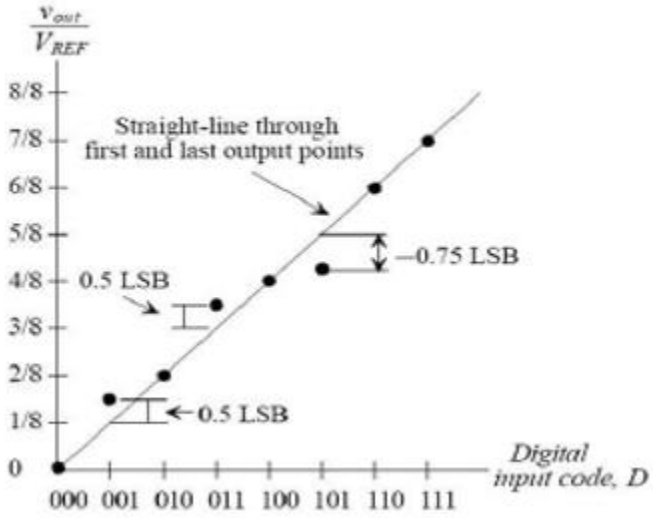


Figure 0.6: INL of a 3 bit DAC

Other linearity issues associated with digital to analog converters are offset and gain errors. Offset error is when all output codes are uniformly increased by a DC voltage. This error usually does not impact the performance of the converter, but must be compensated when measuring the DNL and INL. Offset error factor is measured to represent the ratio between voltage generated from the DAC when all zeros code is input to the LSB value of the DAC.

Gain error is when the output varies from the ideal best-fit line in a linear or non-linear fashion. A linear gain error will not affect the performance of the converter; however a non-linear gain will cause distortion. Gain error can simply be calculated as the difference between the ideal gain to the actual one.

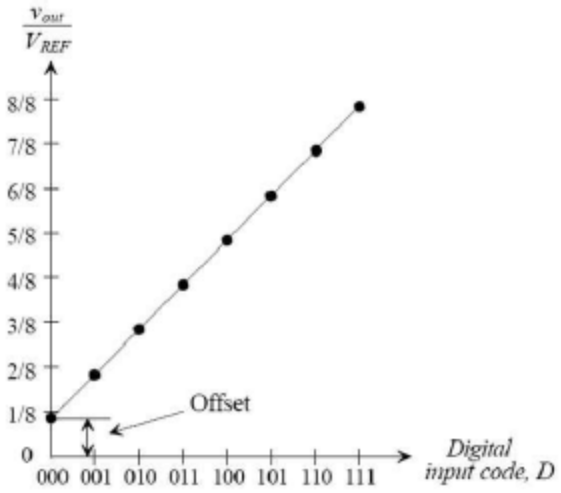


Figure 3.7: Offset error of a 3 bit DAC (0.125LSB)

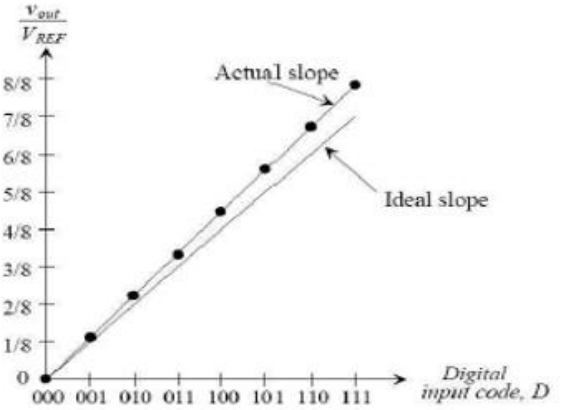


Figure 3.8: Gain error of a 3 bit DAC

3.2.2 Dynamic Performance Measures

The dynamic performance describes the behavior of the DAC when the input word makes transitions between different values. The major dynamic measures can be described through frequency domain measurements. Dynamic performance measures include: settling time, glitch peak, clock feed through, sampling time uncertainty, SNDR, SFDR, and effective number of bits.

3.2.2.1 Settling time and glitch peak.

The output signal of an actual DAC cannot change its value instantly. The time it takes for the output to settle within a certain accuracy of the final value, for example 0.1%, is called the settling time [5]. This value of time determines the highest possible speed of the circuit. To properly test the settling time of the converter; the output can either move from zero to full scale, or in a binary weighted architecture it may be important to know the mid-scale settling time. When the input changes, there will be a delay before the output begins to rise. Once the output begins to rise, the speed it rises with depends on the slew rate of the converter. As the output reaches the final value, there may be overshoot and ripple that will increase the settling time as depicted in Fig.3.9.
The overshoot, or undershoot in the rising output causes glitches.
Glitches can occur from charge feed through errors or from timing skew from components within the DAC. Charge feed through happens when charge builds up between the drain and gate of a CMOS switch. When the switch changes state, the charge of this coupling capacitance is discharged to the output of the switch. Switches in converters have to be carefully designed to ensure charge feed through does not reduce the quality of the output [6]. The glitch energy can be estimated by approximating the glitch as a triangle, then taking the area under it. Special design techniques can be used to reduce the glitches in the DAC.

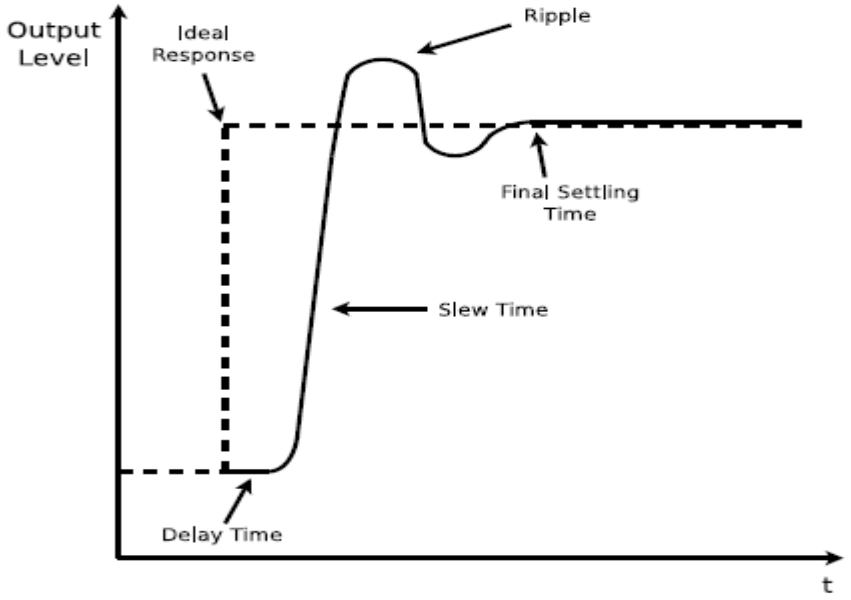


Figure 3.9: Settling time and glitch power of an actual DAC

3.2.2.2 Clock Feed-Through (CFT) and Sampling time Uncertainty in DAC

Due to capacitive coupling in switches; the clock (or digital switching signals) affects the analog output signal [7]. The clock feed-through (CFT) can introduce both harmonic distortion and distortion tones at multiples of the clock signal. The CFT is reduced when reducing the capacitive coupling and therefore the switch transistor sizes should be small to decrease the size of the parasitic capacitances. However, with a smaller transistor, the on-resistance increases which may degrade the performance due to an increased settling time.

Due to noise and other non-ideal effects in the circuit, the time between two samples will change. This sampling time variation gives an error in the output that is determined by the size of the output step and the time variation. The average power of this error (Pe) can be calculated as:

$$Pe = X.\ \frac{Te}{Ts} \quad (3.6)$$

where X is the step size, Te is the error in the sampling time, and Ts is the sampling time of the converter. This power error increases as the signal frequency increases since the step size gets larger.

3.2.2.3 Dynamic Frequency Domain Measures

In real world applications, the dynamic frequency domain measurements provide a better idea of how the converter performs. This is done by applying a full scale sinusoid with a certain frequency to the converter. For some communications applications, the converter may have to handle multiple-tone input. Special metrics such as intermodulation distortion can be used for such applications.

The first dynamic frequency domain specification is the signal to noise ratio (SNR) of the DAC. The SNR is defined as the spectral power of the input compared to the noise floor. The input signal must be at full scale in order to measure the SNR. Smaller amplitude will reduce the SNR, which is intuitive since the signal to noise ratio is directly proportional to the input signal power. In the ideal case, the noise floor would only consist of the quantization noise produced by the converter. In practical converters, errors from linearity, glitches, clock skew, and output settling time will increase the noise floor. The RMS quantization error can be expressed as:

$$Qrms = \frac{VLSB}{\sqrt{12}} \quad (3.7)$$

By dividing the signal power by Qrms, we can get SNR as in equations (3.8), and (3.9)

$$SNR = \frac{2^N * VLSB \Big/ 2\sqrt{2}}{Qrms} \quad (3.8)$$

$$SNR\ in\ dB = 6.02N + 1.76(dB) \quad (3.9)$$

From (3.9), it is obvious that increasing the resolution of DAC by one bit will increase the SNR by about 6 Db. Equation (3.9) can be applied for Nyquist rate DACs but in case of presence of oversampling, SNR will be calculated as:

$$SNR\ in\ dB = 6.02N + 1.76(dB) + 10\log\left(\frac{fs}{2.BW}\right) \quad (3.10)$$

The second important frequency domain dynamic specification is the Spurious Free Dynamic Range (SFDR). It represents the measure of the fundamental signal versus the highest distortion spur in the spectrum. Typically in a DAC, the largest spur will either be the second or third order harmonics, or their aliases. Fig. 3.10 shows a typical FFT representation of an actual DAC showing its achievement of more than 70 dB SFDR.

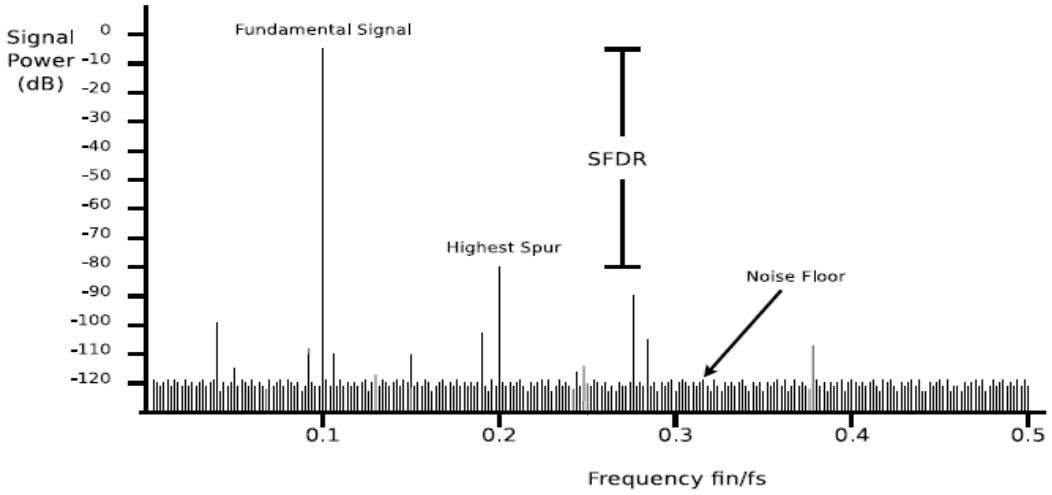


Figure 3.10: FFT representation of an actual DAC

SFDR like SNR is also affected by the static linearity of the DAC. Equation (3.11) gives a better insight on the relation between SFDR and maximum INL.

$$SFDR \approx 20\log\left(\frac{2^N}{INLmax}\right) \quad (3.11)$$

From the work in [8], it can be seen that the shape of the INL curve will determine which harmonic distortion order will dominate the SFDR. An arch shaped INL curve will cause the second order harmonic to be the dominant spur, while an 'S' shaped INL curve will cause the third order harmonic to be the highest spur.

Due to the mentioned spurs, harmonics, noise and distortion; the practical resolution of the converter usually degrades at high frequencies as the DAC is not ideal. The effective number of bits (ENOB) is the measure of this reduced resolution. From the SNR equation, the ideal signal to noise ratio is determined from a given resolution. Working in reverse, the resolution is calculated by factoring in the distortion; replacing the SNR variable with the SNDR. So, by measuring the SNDR of the converter, the real resolution of the output can be determined using equation (3.12). OSR is the oversampling ratio which is equal to 10 log (fs/ 2BW).

$$ENOB = \frac{SNDR - 1.76 - OSR}{6.02} \quad (3.12)$$

Other frequency domain dynamic metrics are briefly discussed in the following points:

- Total Harmonic Distortion (THD): It is defined as the ratio between the signal power to the sum of the powers of all harmonics.
- Signal to Noise and Distortion Ratio (SNDR): is the ratio of the power of the fundamental and the total noise and distortion power within a certain frequency band.
- Dynamic Range (DR): The range from full scale (FS) to the smallest detectable signal usually is called the dynamic range (DR) of the converter. It can be calculated as:

$$DR = 10\log\left(\frac{Maximum\ full\ scale\ signal\ power}{Smallest\ Detectable\ signal\ power}\right) \quad (3.13)$$

3.3 DAC Architectures

There are many different architectures to convert digital signals to an analog output. Choosing the appropriate architecture depends on the application of the converter. Requirements such as area, power, bandwidth, accuracy must be considered. As the operating speeds increase, it becomes harder to design a high resolution DAC.

The main building block in every DAC is a component that creates an appropriate analog output level by dividing a reference voltage. Elements such as resistors split the reference voltage into smaller voltage or current levels. Transistor current sources can also create output currents that are converted to the analog output voltage by the load resistors. Capacitors are also used in some architectures to store charge from the reference voltage and discharge it to the output.

3.3.1 Resistor String DAC

It is considered the simplest DAC architecture as seen in Fig. 3.11. It is made up of 2^N resistors in series, each resistor corresponding to one LSB. A reference voltage is connected at one end of the resistor string, and ground at the other. Switches are connected after each resistor, the output of each switch is tied together to form the DAC output.

The digital input must go through an N: 2^N decoder, which then enables or disables the switches. When the switch is enabled, a voltage division of the reference occurs which becomes the analog output.

This structure results in good accuracy and is inherently monotonic. It is also fast for 6-8 bits DAC and compatible with purely digital technologies [9]. The disadvantage of this topology is that the output is always connected to (2^N-1) switches in “off” state and one switch in “on” state. When resolution becomes a big value, large parasitic capacitance appears at the output and conversion speed becomes much slower. Also this architecture occupies large area and needs large settling time especially for DAC with 8 bit or higher resolution. Matching this large number of resistors properly is also a very difficult task.

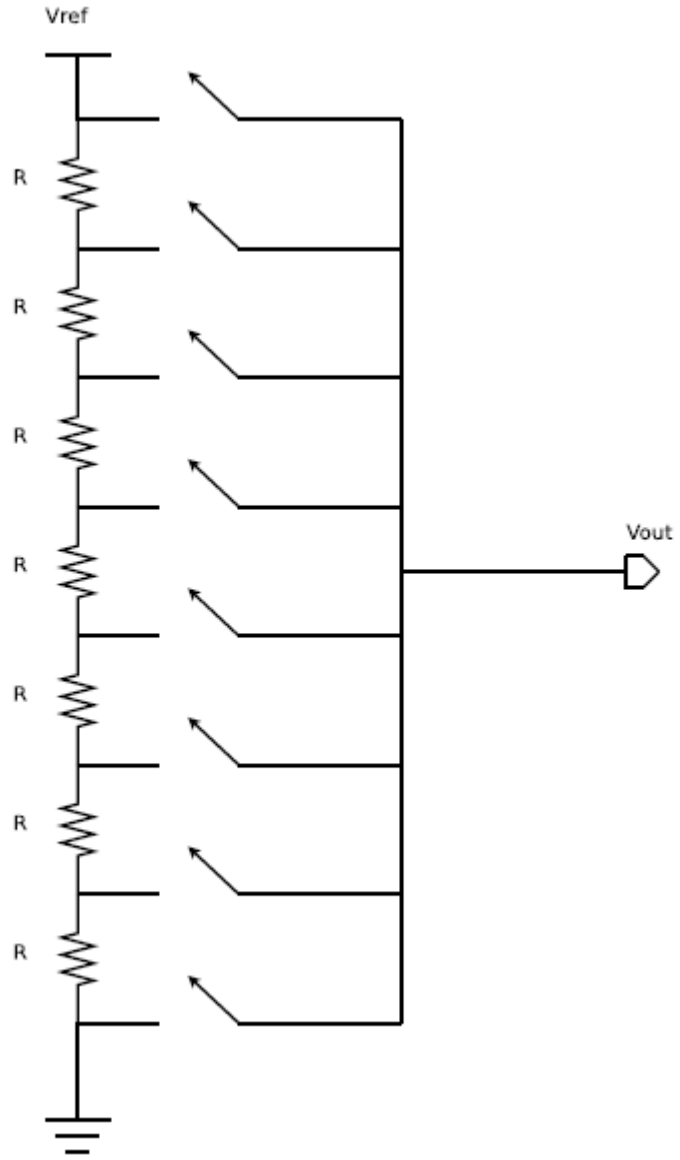


Figure 3.11: Resistor String 3 bit DAC. The control signals to the switches are generated from 3:8 decoder

3.3.2 Binary Weighted Resistor Ladder DAC

The binary weighted resistor DAC shown in Fig. 3.12 is built using a resistor ladder network consisting of N resistors. The MSB resistor has a resistance of R, the resistor value will increase by two for each descending bit. Each resistor is connected to a switch that is controlled by the digital input representing the bits. The outputs of the switches are all connected to an op amp which is acting as a buffer. A reference voltage connected to the resistor ladder allows current to flow through the resistors to ground, or the virtual ground in the op amp to form the output.

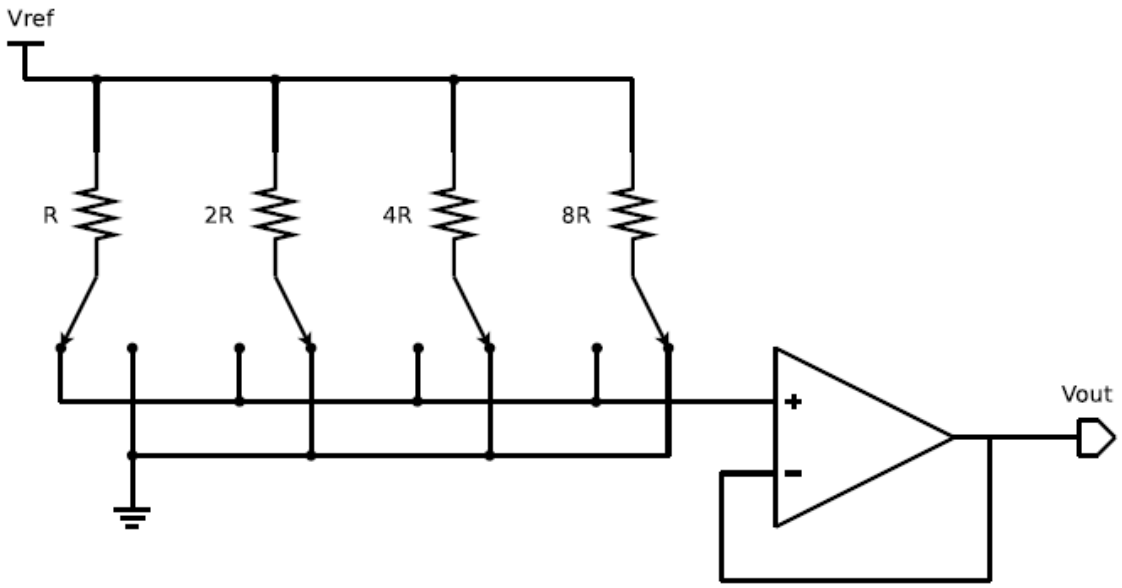


Figure 3.12: Binary weighted resistor ladder 4 bit DAC

The problem with this architecture is that in higher resolution designs, the difference in resistor values becomes large. Process variation in fabrication makes matching of resistors difficult. If the MSB resistance is off by 1% in an 8 bit DAC of this design, an error of more than one LSB will occur. Also, this architecture depends on an op amp which accordingly will limit the possible speed of the converter.

3.3.3 R-2R Ladder DAC

The main drawback of the binary weighted ladder DAC is the difficulty in matching resistors having large difference between their values, so in order to overcome this drawback; R-2R ladder DAC is implemented as shown in Fig.3.13.

A resistor ladder is formed by having a series of resistors of value R, then rungs of resistors of value 2R. Depending on how the reference voltage is arranged, the R-2R can be described as being in voltage mode or current mode. In voltage mode, the switches switch between V_{ref} and ground. The output is taken at the end of the resistor network. While the current works by having the switches switch between the output and ground, while the reference voltage is connected to the resistor ladder.

The advantage of this type of implementation is that it uses a small number of components (2N resistors only), and only two different sizes of resistors. This improves the precision due to the relative ease of producing equal valued matched resistors. Also, the output impedance for this design is always constant independent of the applied code. This is because the equivalent resistance on each side of the ladder rung is 2R. The disadvantage is that if the number of bits is high; there will be a time delay between the LSB and the MSB.

The current mode R-2R DAC must use an operational amplifier to switch the reference voltage between ground and virtual ground. The op amp output buffer can affect the performance of the DAC by limiting the bandwidth, and thus causing linearity errors.

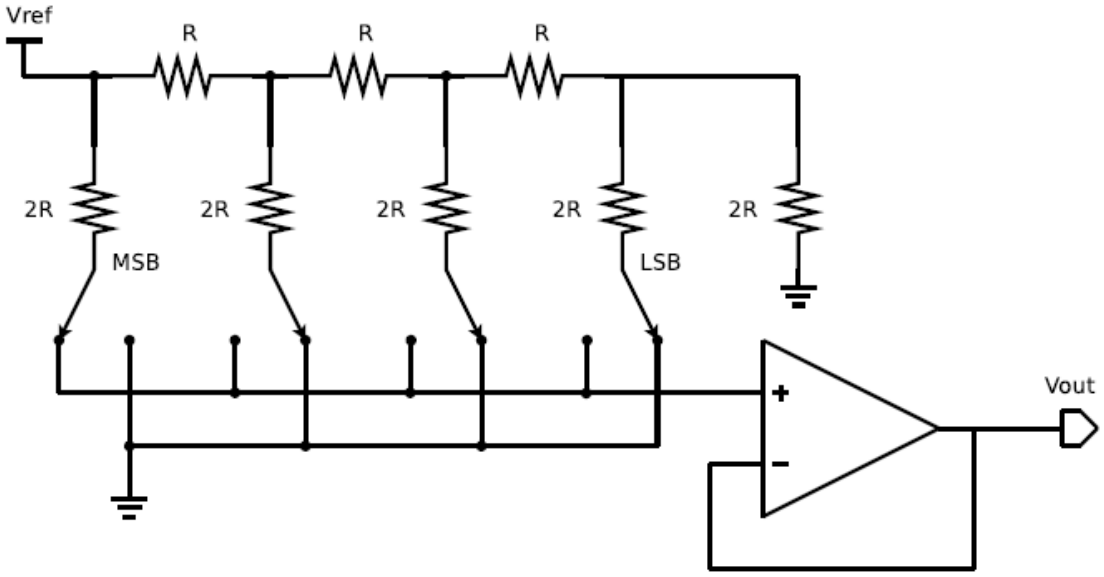


Figure 3.13: R-2R ladder current mode 4 bit DAC

So in case of choosing the current mode R-2R DAC, the linearity of the implemented resistors should be as high as possible.

3.3.4 Charge Scaling DAC

Charge scaling DAC is considered a switched-capacitor (SC) DAC, where the charge stored on a number of scaled capacitors is used to perform the conversion. The converter can be implemented in a binary weighted fashion, or using a C-2C network like in resistor networks. One difference with the charge scaling architecture is the need for a reset switch that discharges all the capacitors between each conversion.

The charge scaling converter is fast, accurate, and easily implemented in CMOS after the emergence of sub-micron CMOS. The disadvantage of the architecture is that it requires an op amp which can limit its performance. Parasitic loading of the capacitors from the op amp can limit the resolution of the converter; this can be solved by using a switched capacitor integrator. This architecture also suffers from great charge feed-through errors in the output.

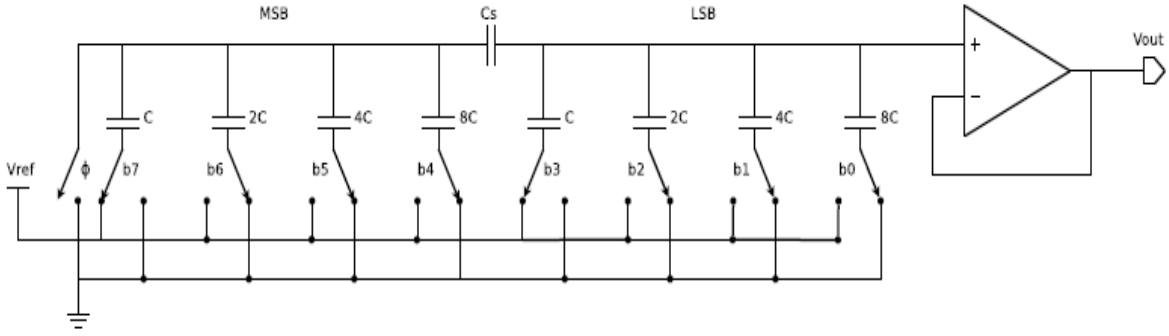


Figure 3.14: Charge Scaling DAC Architecture

3.3.5 Current Steering DAC

The current steering DAC is the preferred architecture for high speed converters. This architecture is shown in Fig.3.15. It works by summing the current produced by an array of current sources. A load resistor is connected to the output of the current sources on one end, and the supply voltage at the other. When the current source is enabled, current flows through the load resistor. This current is dumped by the current source, thereby converting the current into a voltage eon the resistor. The output voltage is formed from the voltage division between the power supply connected to the resistor, and the voltage drop produced by the current draw from the current sources. Since an op amp is not required to form the output voltage, higher performance can be achieved.

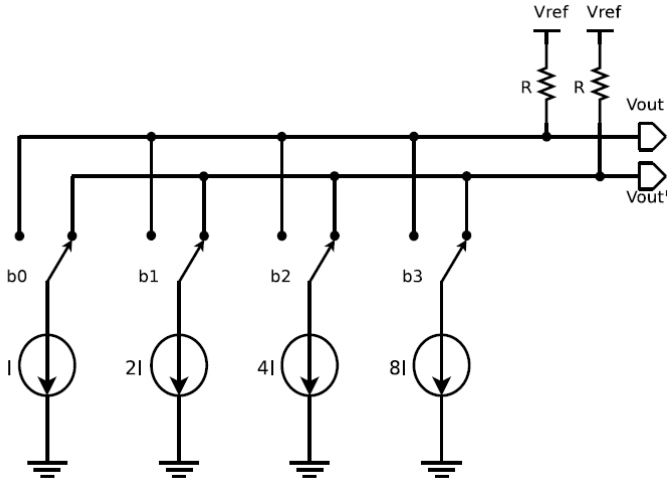


Figure 3.15: Current Steering 4 bit DAC

A binary weighted current steering DAC can easily be implemented as in Fig. 3.15, only N current sources and switches are needed. The problem with the binary weighted design is that a large glitch can occur at the output when several current sources are being switched at once. For example, when the MSB current source switches on (mid code), all the other current sources will be turning off. If the MSB current source does not reach its final value before the other sources switch off, then a glitch will occur.

Also as the current sources are binary weighted, the sizes of the transistors representing the current sources will be binary weighted leading to a great ratio between LSB current source transistor and MSB transistor. The ratio between them is equal to 1: 2^{N-1}, this ratio becomes unreasonable for high resolution DAC as achieving matching between transistors with such difference in their aspect ratios is an extremely complicated task.

A solution to these drawbacks of the binary weighted implementation is to move to the unary weighted current steering DAC as obvious in Fig.3.16. In this design, every current cell produces a unit current of I. The LSB will consist of a single unary current cell, while the MSB will be comprised of 2^{N-1} current sources. In order to send the correct control signals to the unary current cells, a binary to thermometer decoder is used. Thermometer coding is preferred over binary because only one bit switches at a time, eliminating the MSB current source glitch.

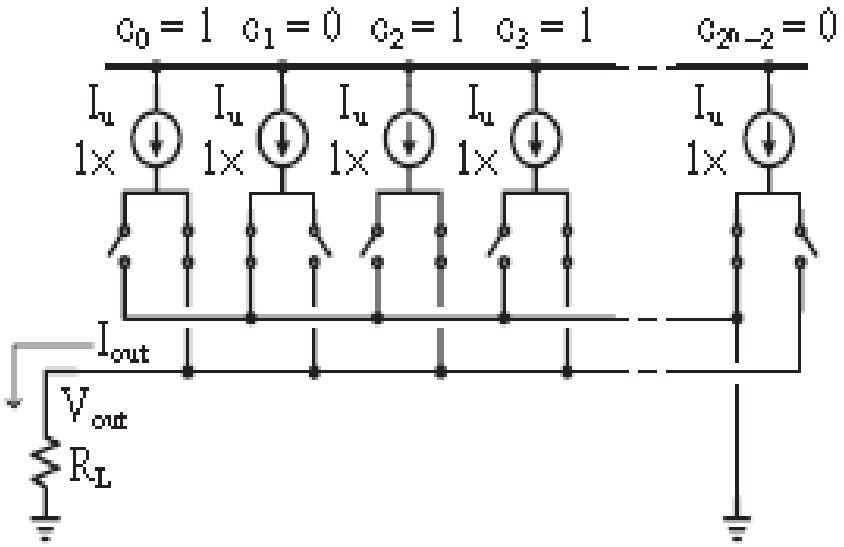


Figure 3.16: Unary weighted current steering n bit DAC. Control signals c (0) to c(2^n-2) are generated from the binary to thermometer decoder

The difference between binary and thermometer coding can be clarified obviously in Table 3.1

Table 3.1: Binary vs. Thermometer coding

Decimal	Binary	Thermometer
0	000	0000000
1	001	0000001
2	010	0000011
3	011	0000111
4	100	0001111
5	101	0011111
6	110	0111111
7	111	1111111

Unary weighted current steering DAC, although solves most of the drawbacks accompanied with the binary weighted corresponding DAC, but is limited to certain DAC resolution, above which its usage would be not the best choice. As resolution surpasses 10 bits, the amount of current sources required becomes very large. This is an issue since all of these cells will consume large area. The routing required to control thousands of current cells introduces a large amount of parasitic capacitance that will limit the speed of the converter. The thermometer coded design also requires the thermometer decoder which adds complexity, and increases power consumption from the logic gates. Therefore another enhanced technique for current steering DAC implementation was used, i.e. Segmented Current Steering DAC.

The segmented current steering DAC shown in Fig.3.17 represents a trade-off between high speed, and reduced area and complexity. In the segmented technique the DAC is divided into 2 sub DACs named after the segment of bits they represent as: Most Significant Segment DAC and Least Significant Segment DAC. The most significant segment is implemented in a unary fashion while the least significant segment is implemented in binary fashion. The segmentation factor expresses the ratio of the most significant segment (unary part) from the total number of bits. So for a 10 bit DAC if 4 bits are implemented in a unary weighted implementation (Most Significant Segment), then it's said to be 40% segmented.

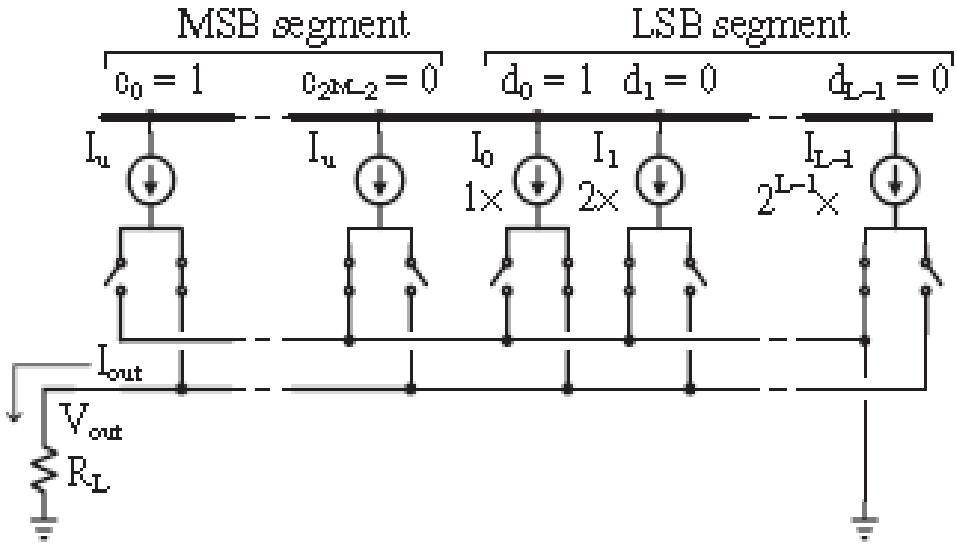


Figure 3.17: Segmented current steering DAC

The optimal ratio between binary weighted bits and thermometer coded bits must be determined before designing the converter. From the work in [10], the amount of segmentation for a 10 bit DAC should be approximately 30-70% for proper operation. Fig. 3.18 shows the summary of the work done.

When designing a current steering DAC, the finite output impedance of the current cells must be taken into consideration to maintain accuracy. The current cells are typically made up of transistors acting as current mirrors, with the drain of the current mirror connected to the DAC output. As the output voltage changes, the drain voltage on the current mirror will also change. Since the current mirror does not have infinite output impedance, the changing drain voltage will

cause the reference current to vary. To get an accurate reference, we need to add an output buffer and designing the current cells with increased output impedance.

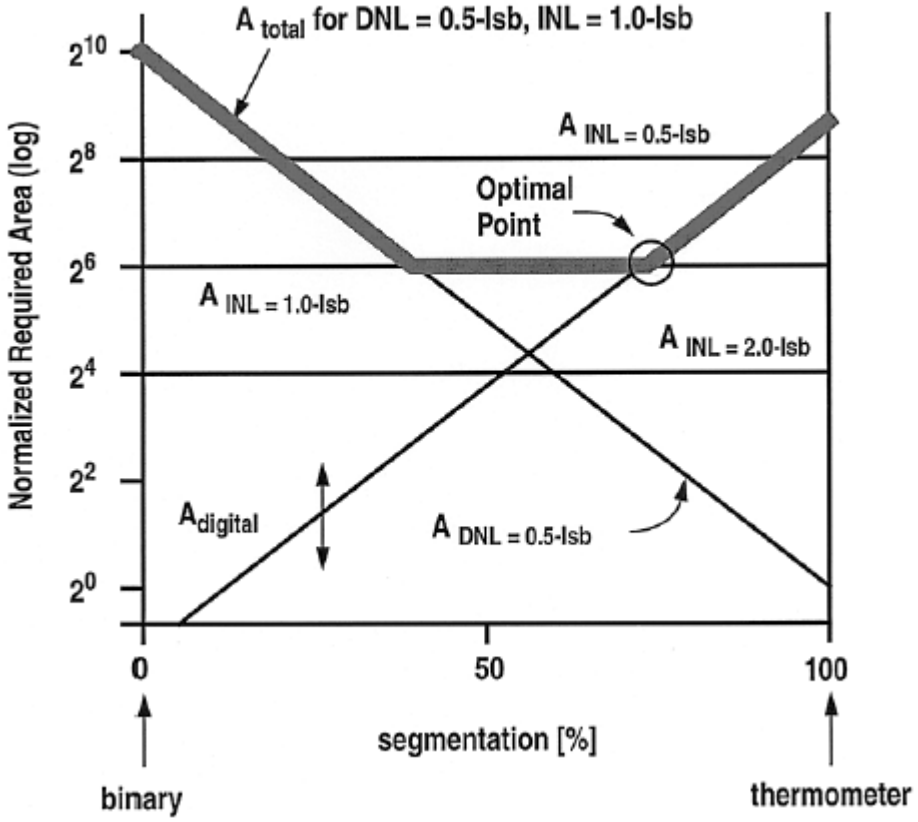


Figure 3.18: Normalized required area versus percentage of segmentation

Fig. 3.18 shows the normalized required area versus percentage of segmentation. Based on DNL performance only, the minimum analog area for 100% segmentation is 1/1024 from the minimum analog area for 0% segmentation. On a logarithmic scale, the minimum analog area requirement as a function of segmentation will form a straight line connecting the above mentioned points, as shown in Fig. 3.18. INL depends mainly on the area, so it has no variation across segmentation ratio. For the digital area in the converter, it increases as the segmentation increases due to the need for larger and more complicated decoders. This area has nothing to do with the linearity specs of the converter which are controlled by the ratio of segmentation and analog area only.

As we increase the percentage of segmentation, the required total area is first dominated by the DNL requirement, then by the INL requirement, and finally by the decoding logic. So, if the system requirement is DNL= 0.5 LSB and INL = 1 LSB. For minimum area, the flat part of this curve would be optimum and the total area would be determined by the INL requirement.

In the next section, the system level requirements of the DAC for the NB-IoT transmitter will be discussed and according to these requirements, the proper DAC architecture will be chosen.

3.4 NB-IoT DAC system level requirements

The NB-IoT DAC system level requirements are summarized in table 3.2:

Table 3.2: DAC System level Requirements

Quantity	**Value**
Number of Bits	10 bits
Sampling Rate	6.4 MS/sec
SNR	69 dB
SFDR	65 dB
DNL	< 0.5 LSB
INL	< 1 LSB

According to the required specifications, it was found that the best architecture to be used is the segmented current steering DAC with segmentation ratio of 40% (4 bits thermometer coded, 6 bits binary coded).

3.5 Proposed DAC Overview

The DAC system starts with a binary to thermometer decoder which converts the 4 most significant (MS) bits to 16 control lines. These control lines along with the least significant (LS) bits enter a master-slave latch (register) in order to be synchronized. The output of the latch is input to the switch drivers of the DAC core in order to finally generate the control signals that control the switches responsible for current steering. The output current is dumped into 200 ohm resistors in order to be converted to corresponding voltage. This voltage is buffered through an 80 MHz operational amplifier in order to increase the overall output impedance of the current sources. Finally, the signal passes through a 4^{th} order differential low pass reconstruction filter in order to get rid of the aliases in the spectrum and make the signal smoother by eliminating the glitches to avoid any undesired specs. Fig.3.19 clarifies the DAC overall system.

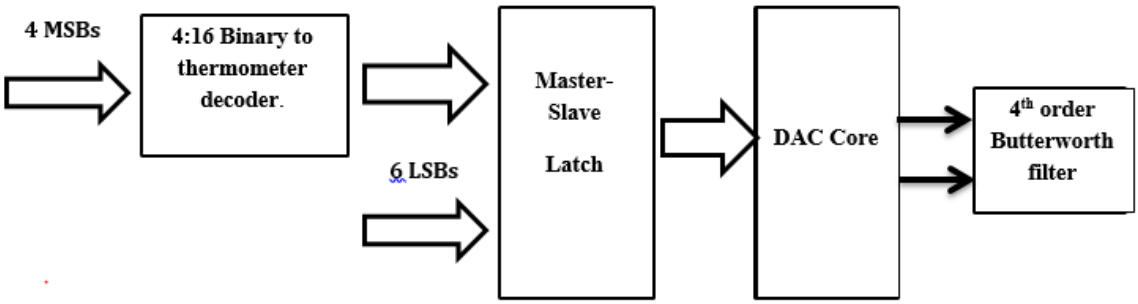


Figure 3.19: Proposed 10-bit DAC overall system

3.6 Digital Sub-Circuits

Digital Sub-Circuits include binary to thermometer decoder and master-slave latch.

3.6.1 Binary to Thermometer Decoder

The thermometer decoder can be implemented in a 1-D, or 2-D way [11]. The one dimensional design uses a single decoder to convert the binary bits and address each current cell. The two dimensional decoder allows the current cells to be arranged in a matrix. Two decoders are required to address the rows and columns in the current cell matrix. The advantage of using the 2-D design is that the complexity of the thermometer decoder can be reduced, which is why this method was chosen.

The four MSBs (B9, B8, B7, and B6) are divided into 2 groups; one group is input to the column sub-decoder (B7, and B6) while the other group (B9, and B8) is input to the row sub-decoder.

The logic for the 2:4 sub-decoders is easily obtained from table 3.3, where the 2 input bits are swept across all their possible values, and the output 4 signals are defined from the definition of thermometer code.

Table 3.3: Sub-decoder input and output signals

MS bit (Y)	LS bit (X)	T3	T2	T1	T0
0	0	0	0	0	1
0	1	0	0	1	1
1	0	0	1	1	1
1	1	1	1	1	1
		X AND Y	Y	X OR Y	‘1’

By rearranging the equations using inverted inputs, it can be seen that 2 level logic using NAND, NOR, and NOT gates can implement the required function. The logic gates were sized large

enough in order to ensure high speed and drive strength. Fig.3.20 shows the output of the implemented sub- decoder for all possible cases.

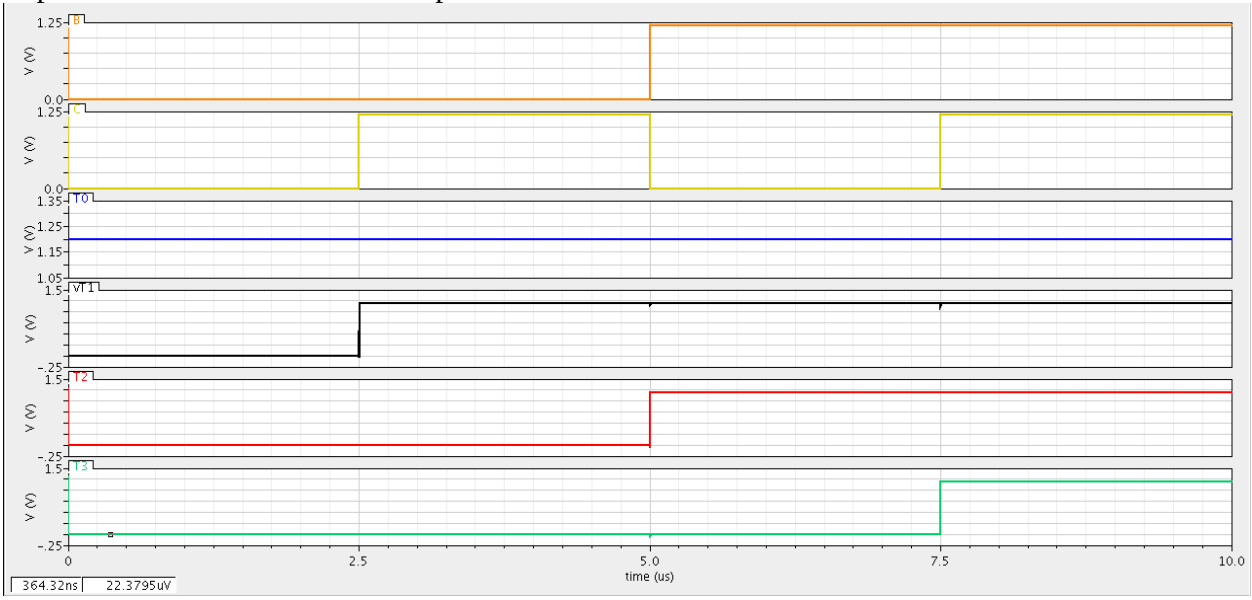


Figure 3.20: DAC Sub-decoder input and output signals

Now by having the row and column sub-decoders, the 15 unary weighted thermometer coded current cells corresponding to the MS segment of bits can be implemented in a 4*4 matrix 2D structure as shown In Fig.3.21.

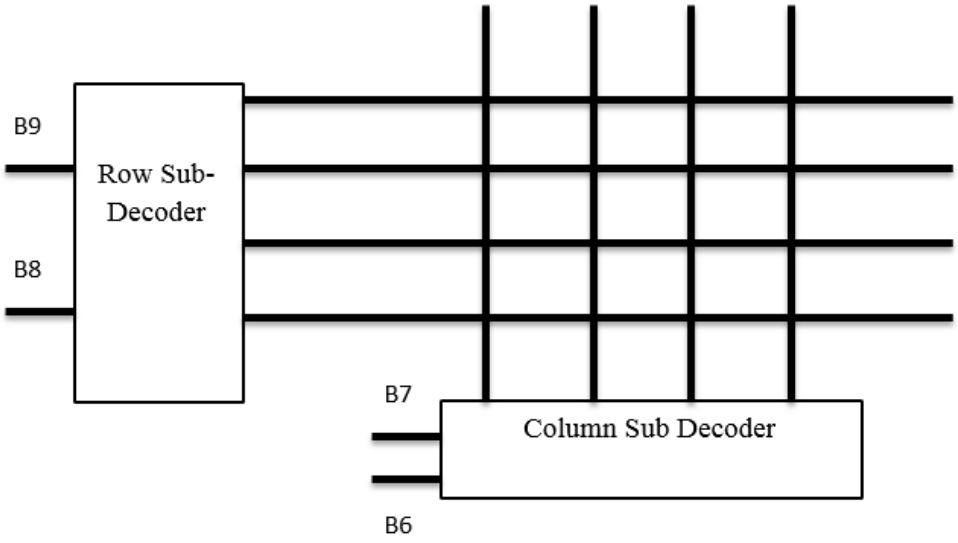


Figure 3.21: MS segment implemented in a 4*4 matrix

For each cell to be activated, it contains a local decoder unit which simply performs an ANDing operation between the row line and the column line of the cell. And for thermometer coding action, the local decoder also contains the next row line so that if the next row becomes

activated, then all the cells in the current row should be activated as well. Fig. 3.22 shows the full decoding blocks implementation.

3.6.2 Master Slave Latch

The MS segment passes through the digital decoder circuits unlike the LS segment of bits. Therefore, the delay of the signal path through the decoders can lead to unwanted deviation and skewing between the control signals connected to the core. In order to overcome these effects, dummy decoders can be placed in the path of the LS segment. Instead of using these dummy decoders, the 15 output control lines from the local decoders are input to a master slave latch in order to be synchronized with the LS segment 6 bits. This type of latch, shown in Fig.3.23, consists of two cascaded D latches and an inverter placed in the clock path.
Each D latch is simply implemented as cross coupling inverters in order to store the signal until the clock edge arrives. Fig.3.24 shows the implementation of the D-latch used. The signal is input with its inverted version as a differential input to the latch at the negative clock edge when clk' moves from low to high enabling the switches of the Master Latch. After that the signal is held by the feedback loop of the cross coupled inverters until the next positive edge of the clock, which when occurs will allow the signal to move from Master Latch to Slave Latch before going to the output. The inverters used at the end of each latch stage are put to boost the signal strength. The functionality of the master slave latch has been tested with a clock speed of 10 MHz and it proved high accuracy as depicted in Fig.3.25.

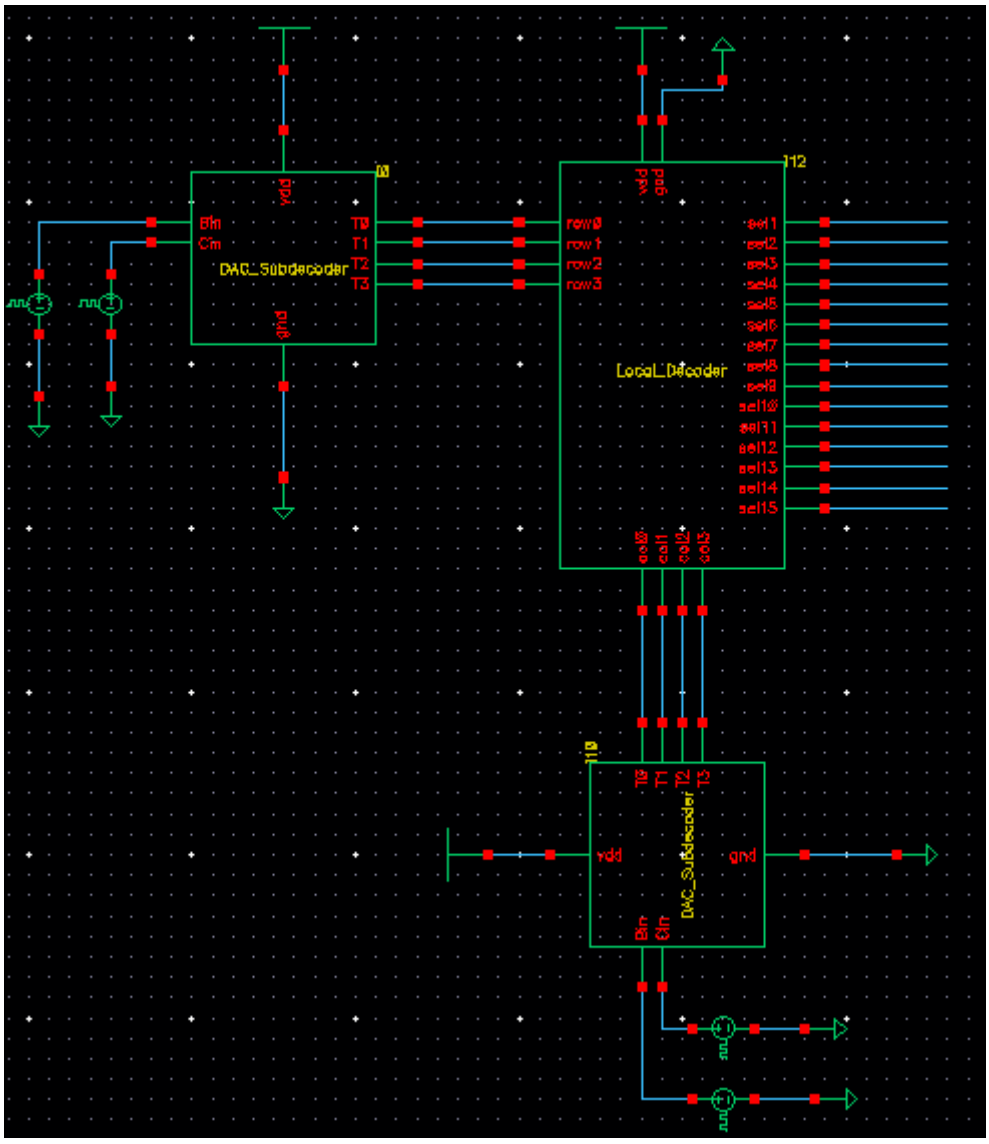


Figure 3.22: Row and Column Sub-decoders connected to the local decoders' matrix

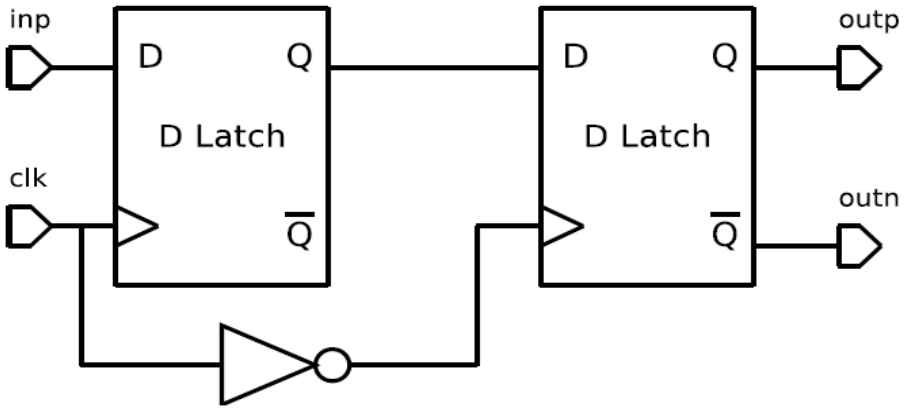


Figure 3.23: Master-Slave Latch block diagram

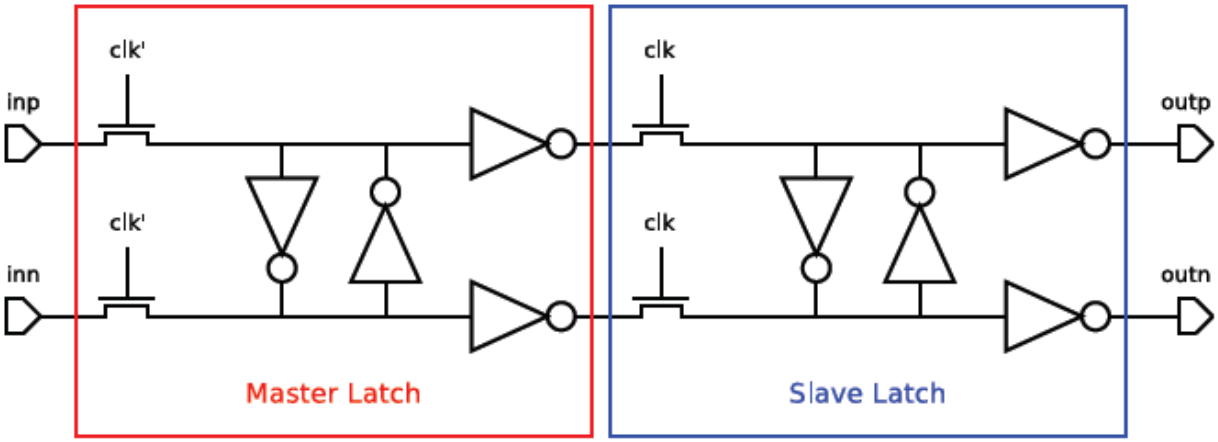


Figure 3.24: Master Slave Latch Gate level implementation

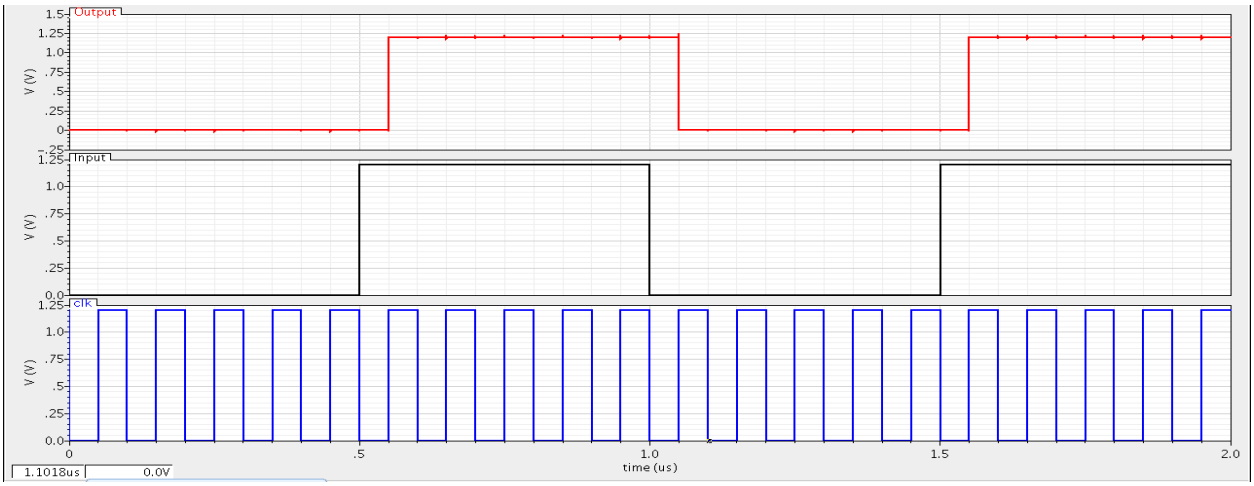


Figure 3.25: Master Slave performance with a clock of speed 10 MS/s

3.7 DAC Current Cell Design

The current cell is the main sub-circuit to the current steering DAC. It generates the required current to be provided at the output of the DAC. Based on the input word, the currents generated by number of current sources are properly switched and summed at the output node. Design of the current cell must take into account several factors like: transistor mismatch, output impedance, area, and power consumption. Inaccuracies in the current cell will greatly reduce the static and dynamic performance of the converter. A basic current cell consists of differential switches and a current source. The differential switches receive control signals from the local decoder and latch sub-circuits and then determine which output the current will go through.

3.7.1 DAC Current Source Design Issues

Design of the DAC current source is greatly affected by the fabrication non idealities which cause undesired mismatch between similar components. These variations impact the accuracy of the reference current being produced. So, in order to overcome these variations and improve the matching of the current source, its transistors need to be sized larger to decrease the fabrication possible errors but this would lead to greater area. So to compromise between these two effects; Pelgrom's model was developed to characterize the mismatch factors between transistors based on their area and distance between them. From Pelgrom's model, the minimum transistor area can be calculated according to (3.14):

$$(WL)\min = \frac{A\beta^2 + (\frac{2Avt}{Vgs - Vt})^2}{2\sigma^2} \qquad (3.14)$$

where Avt is the threshold voltage mismatch parameter, Aβ is the current factor mismatch parameter; both of these being process dependent variables. The relative standard deviation of the current source is denoted as σ. From Pelgrom's model, it can be noticed that in order to optimize the area needed with certain mismatch factor, the overdrive voltage can be increased. But this solution would lead to a limitation on the headroom for the current cells; this problem

was solved by using a current mirror with level shifter technique that will be discussed in the next section.

Another important parameter to be cared for during the design of the current cells is the output impedance of the current source. Ideally a current source achieves infinite output impedance. Output impedance of a current steering DAC is highly dependent on the applied code as the code defines the number of switched-on current branches and accordingly defines the output impedance. This code dependent nature of the output impedance causes the DNL and INL to vary. This variation in the INL will impact the dynamic performance of the converter leading to the degradation in the achieved SNDR and SFDR.

In order to achieve relatively high output impedance for all codes, 2 design techniques are used. First, the current cell is implemented using a cascode current source architecture with long-channel transistors to increase the r_o of the transistors used. This technique increases the output impedance and also decreases the parasitic capacitances accompanied with the current branches leading to a decrease in the clock feed through effect. The second technique is to add an op-amp buffer at the output of the DAC, as the input impedance of the buffer is usually huge. In this work, both the techniques were used together to obtain the highest possible output impedance.

Now we reach to the third important design parameter of the current steering DAC current source namely, the LSB current value (I_{LSB}). This parameter is greatly affected by the variations of the power supply. For high resolution DACs, designing for extremely low I_{LSB} will lead to overall low power consumption especially for modern low supply technologies, but this requires a very stable and sensitive power supply circuit. So there is a need to compromise between these 2 effects.

High I_{LSB} can achieve good linearity for the converter, i.e better DNL. But on the other hand increasing I_{LSB} leads to the reduction of the output impedance of the DAC current source leading to loss of linearity; therefore the design must choose a point in between achieving enough low DNL without losing the output impedance specification.

3.7.2 DAC Current Source Implementation

In order to achieve high output impedance, a cascode current mirror can appear as a good candidate as shown in Fig.3.26. The output impedance can be calculated through the well-known relation indicated in (3.15).

$$Rout = ro4 + ro2 + gm4.ro4.ro2 \quad (3.15)$$

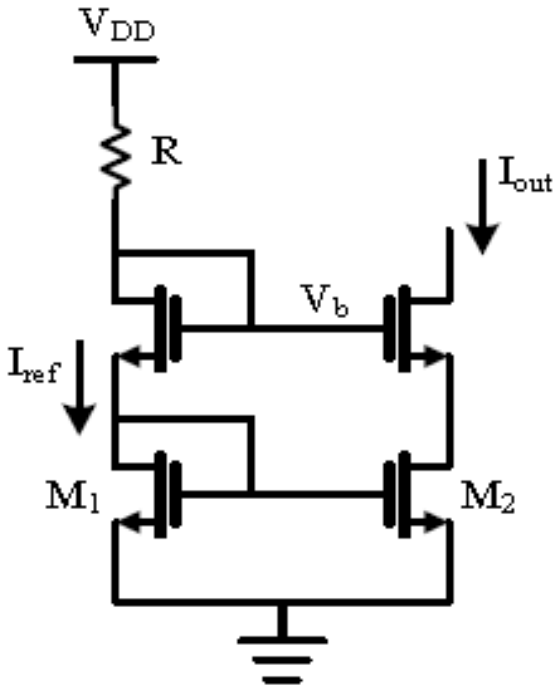


Figure 3.26: Cascode Current Mirror Architecture

For low supply technologies, the traditional cascode architecture would be not the best choice because the compliance voltage of this architecture is equal to the sum of the overdrive voltage of Q4 and the gate to source voltage of Q2, which leaves only a relatively small headroom for the connected switches and for the DC level of the next stages in the chain. Therefore, the need for low compliance high output impedance current mirror architecture is crucial. Fig. 3.27 suggests a better implementation for a cascode current mirror with overall compliance voltage of twice overdrive voltage.

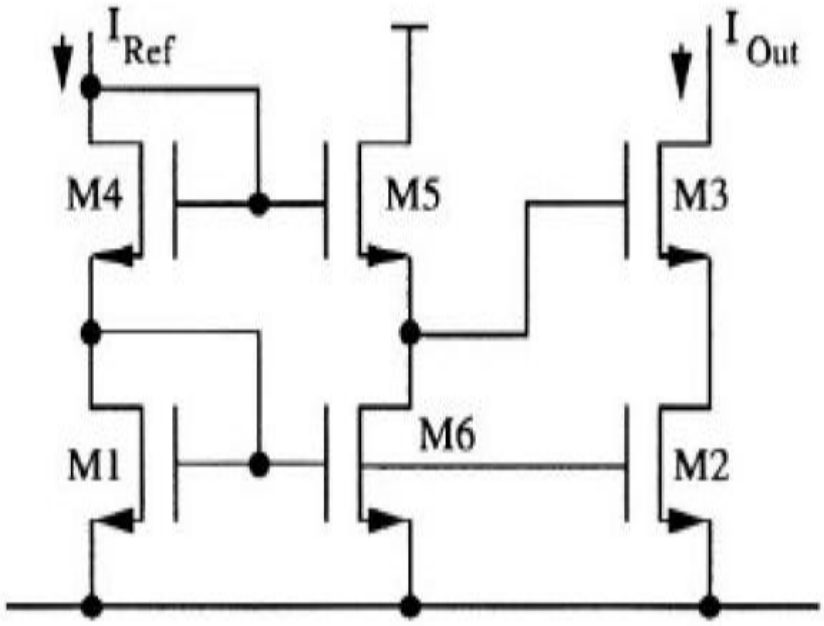


Figure 3.27: Cascode Current Mirror with level shifter

M5 behaves as the source follower and is biased by the output of the simple current mirror M6 and M1 which defines Vgs of M5 to be larger than its threshold voltage by the overdrive. However, Vds of M2 would be zero with equal thresholds and overdrives on all transistors. So, to bias M2 at the boundary between the active and triode regions, Vds of M2 is required to be equal to overdrive voltage. Thus, the overdrive voltage of M5 is doubled by reducing its W/L by a factor of four to satisfy the requirement of Vds of M2. This makes the overall compliance voltage of the mirror equivalent to the summation of the overdrives of M2 and M3 which is the lowest possible value. Thus this type of level shifting is considered as high-swing cascode current mirror. Source follower is used to implement level shift and the current mirror behaves as an active load.

Now, we have obtained a well-designed high output impedance with high output swing, the next important factor is the accuracy of the mirrored current. Lack of accuracy in the current mirror can lead to unequal steps which will affect the DAC linearity. So for a high accurate current mirror based on the previous discussed cascode structure, work in [12] was done. Fig.3.28 shows the implemented high accuracy current mirror.

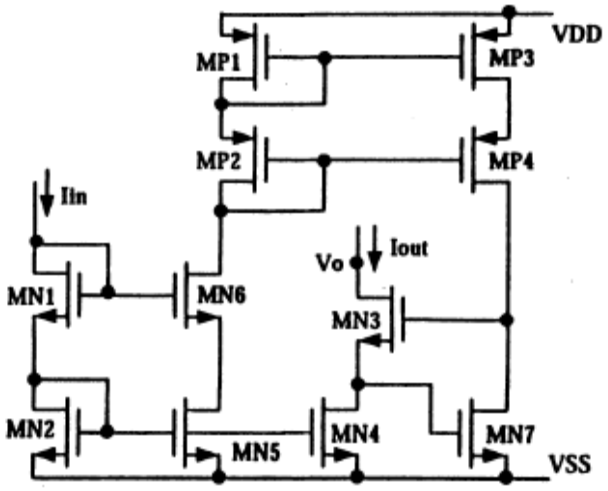


Figure 3.28: High accuracy cascode current mirror

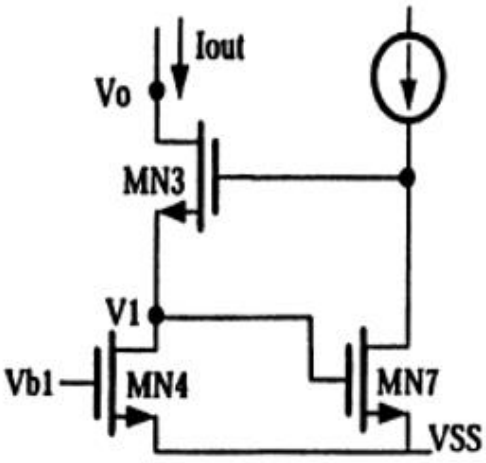


Figure 3.29: Feedback loop of the proposed design

The MOS transistors MN1 - MN4 are used as a two-stage cascode current mirror. The MOS transistors MN5 -MN6 and MP1 - MP4 are used to improve the match accuracy of the cascode current mirror.

The MOS transistors MP1 -MP4 are used to match the current IM6 and IM7. Putting the aspect ratio of MPI - MP4 the same, then Iin = IM6 and IM6 = IM7. Making the aspect ratio of M7 equal to that of M2, makes VGS7 equal to VDS2 because VGS7 is equal to VDS4 and it's easy to find out that VDS2 is equal to VDS4. With this situation, we can find that this circuit creates a negative feedback loop which ensures that Iout is equal to Iin.

Fig.3.29 gives a better insight on the created feedback loop. If Iout decreases, this will result in an increase in the source voltage of MN3. MN3 source node is also the gate node of MN7, so when this node increases, VGS of MN7 increases, so this transistor wants to sink more current but as the current sunk in MN7 is constant, the effect of increase of VGS7 will lead to an increase in VDS7 increasing the drain voltage of MN7 which is also the gate voltage of MN3, so it increases back to its biased value pushing back Iout to be equal to Iin.

Fig.3.30 shows the implemented current mirror which is implemented in a PMOS topology rather than NMOS because in the chosen DAC architecture, a current source is needed rather than a current sink. To meet the minimum area requirements along with all other specifications discussed through this section, the length of the transistor of the cascode current source was chosen to be 1µm while its width was chosen to be 2.4µm. I_{LSB} was pushed to 1µA and accordingly all the 6 branches representing the LS segment of bits carried currents of 1, 2, 4, 8, 16, and 32 µA respectively, while the 15 current cells representing the unary thermometer coded MS segment of bits carried 64 µA each.

The achieved mirroring accuracy was great, as the mirroring error for all branches had an average value of 0.18%. Fig.3.31 shows the variation of the mirroring error for all the 7 different current values.

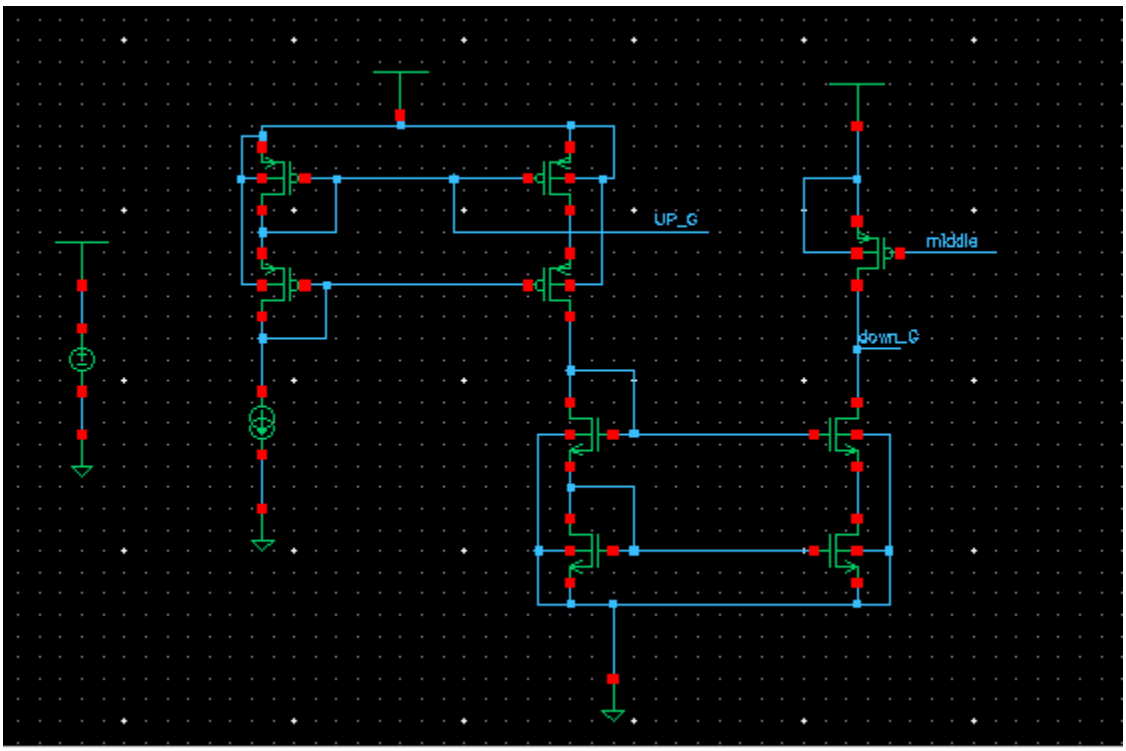


Figure 3.30: Implemented high accuracy cascode current mirror

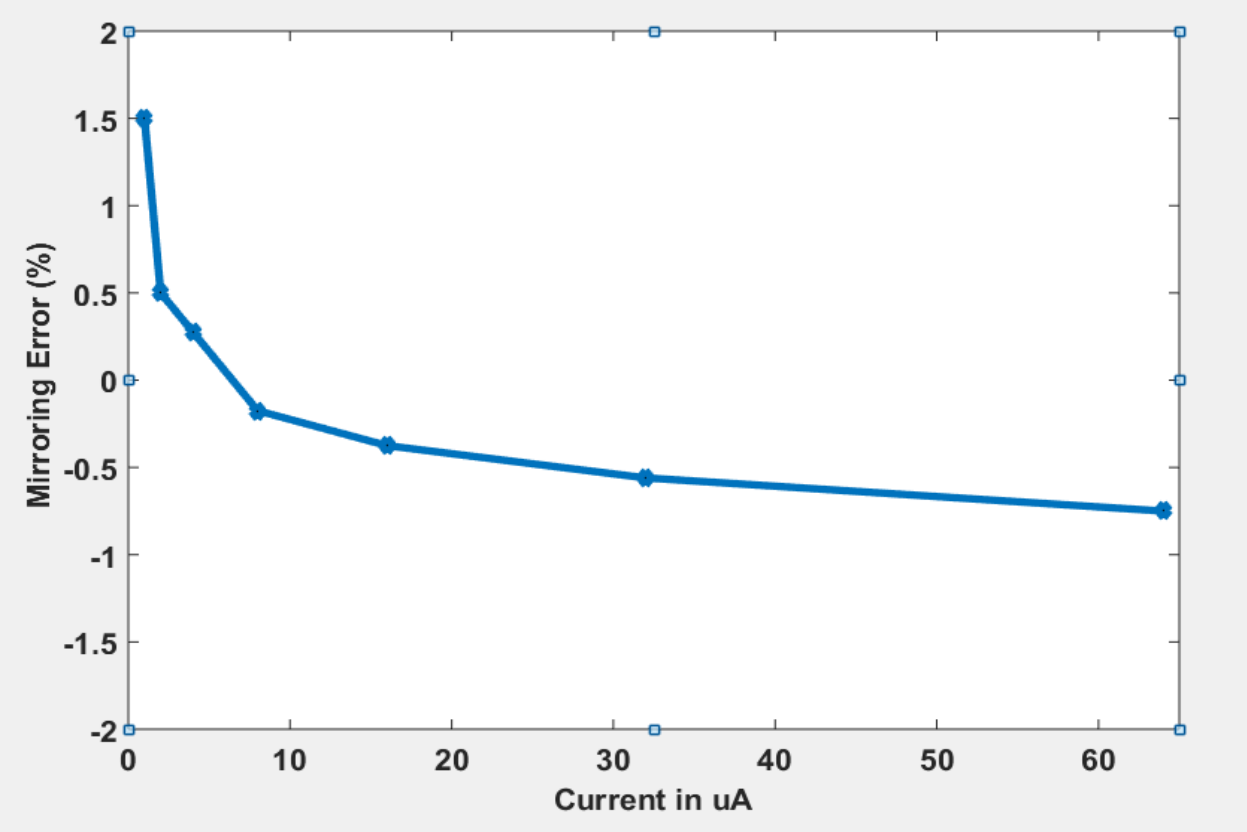


Figure 3.31: Mirroring Error % across different current values

When designing the current cells to achieve high output impedance, it should be noted that the LSB current cell has the most stringent impedance requirements. The output impedance of the overall converter can be represented as:

$$Zout = \frac{Runit}{n} \qquad (3.16)$$

where Runit is the LSB current cell impedance, and n is the decimal code that the DAC is converting [13].

In order to calculate the output impedance of the implemented design, an AC source of amplitude 1V was placed at the output. AC analysis was applied to get the current value generated from the AC source at the required frequency of operation of the DAC at 6.4 MHz. Fig.3.32 shows that the generated current was around 730 nA indicating an output impedance of 1.36 MΩ for the LSB branch. For next branches, the output impedance will be simply equal to that of the LSB branch divided by multiples of 2 according to the binary weight of the branch.

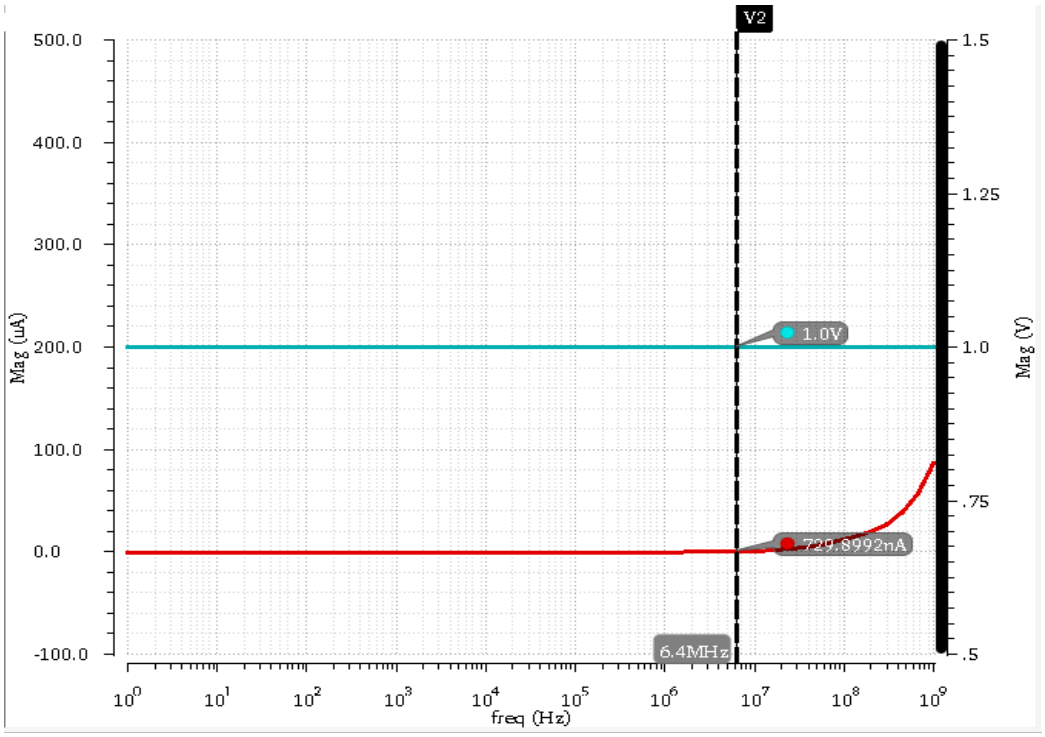


Figure 3.32: Calculation of the output impedance of the DAC current source

Fig.3.33 shows the stability of the current mirror feedback loop. It achieves phase margin of 66^o.

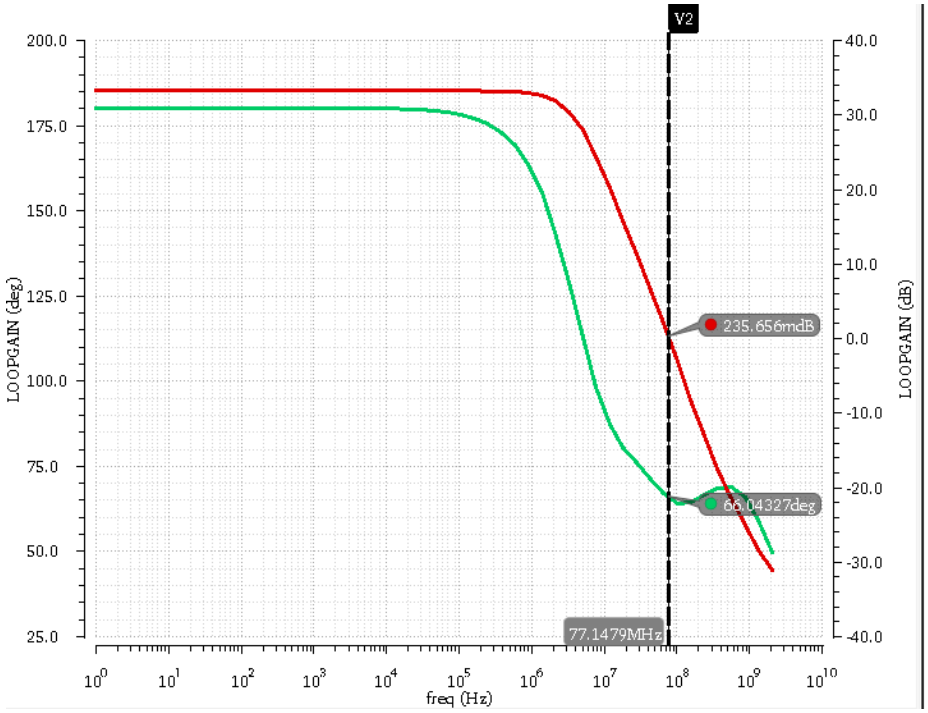


Figure 3.33: Stability of the current mirror feedback loop

3.7.3 DAC Current Cell Switch Design

To reduce the impact of glitches in the current cells, the crossing points of the differential current switch control signals can be modified. Instead of defining the high voltage input to a switch as V_{dd} and the low input voltage as Gnd, the swing of the control signals can be reduced by raising the logic low voltage. This method is used in [14] and achieved better linearity specifications.

A switch driver circuit is designed as shown in Fig. 3.34 to decrease the control signals swing. The driver works by having two NMOS transistors that are controlled by the digital output from the master-slave latch. When the digital input is zero, the driver output will be V_{dd} since the transistor will be turned off. When the digital input is high, the transistor will turn on. This causes some of the current from the resistor to sink to the current source, causing the output voltage level to decrease. The resistor and current source are set up so that the driver will output 1.2 V for a logic one, and 400 mV for a logic zero.

In the case of input high, the corresponding output level of the switch driver circuit will be 400 mV. This voltage will be input to the gate of one of the 2 differential switches of the cell. In order to ensure that this switch will be in cut-off region, the switch is made up of thick oxide high threshold voltage transistor. This type of transistors is characterized by its high threshold voltage; therefore 400 mV on its gate will not be enough to create a channel. Also in order to further ensure keeping the switch transistor in cut-off region, its aspect ratio is designed as large as possible in order to decrease its gate to source voltage and thus ensure that it will not exceed the threshold voltage. Increasing the aspect ratio also helps in decreasing the on resistance of the

switch, as the on resistance, as indicated in (3.17), of an NMOS transistor is inversely proportional to its aspect ratio.

$$Ron = \frac{1}{\mu n.Cox.\frac{W}{L}\ (VGS - Vth)} \quad (3.17)$$

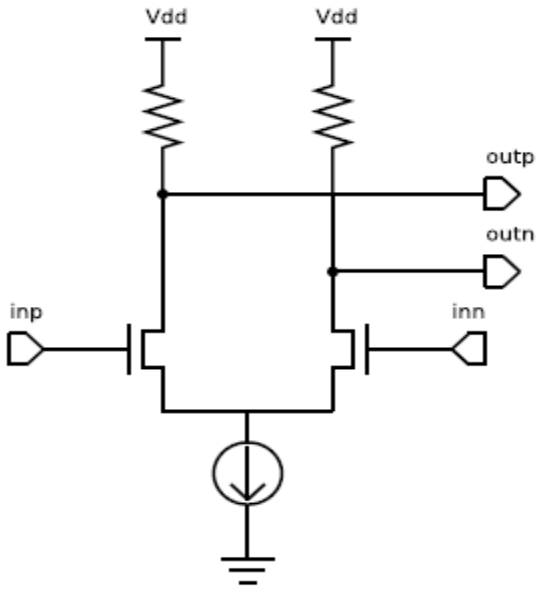


Figure 3.34: Current Cell Switch Driver Circuit

But the aspect ratio cannot be increased except for certain limit; this limit is related to the speed of switching required. Since the larger width transistor has more gate-drain capacitance, so it takes time for the switch to completely steer the current to either of the node. Hence, to decrease the internal capacitance of the transistor and improve the speed of switching, transistors are needed to have smaller width.

Since the current scales up by a factor of 2 for the six branches representing the 6 least significant bits (LS segment), the Ron resistances will need to be scaled down by the same factor, therefore, the aspect ratio of the switch transistors will be scaled up by a factor of 2 for each higher order bit. The other 15 current cell switches (representing the MS segment) will all have the same switch size as they all carry the same current.

The main drawback of using the implemented switch drivers is the increase in the leakage current passing from the switch. This is because the switch that is required to be off is working at a boundary region between the cut-off region and sub threshold. Therefore, this represents a drawback. But as this technique helps in decreasing the effect of glitches, this is considered a more important criterion for a transmitter than leakage current, because decreasing the glitches helps in minimizing the power of spurs that appear in the spectrum of the transmitter which is an essential characteristic for the transmitter.

Fig.3.35 shows the full switch block containing the differential switches and their switch driver circuit.

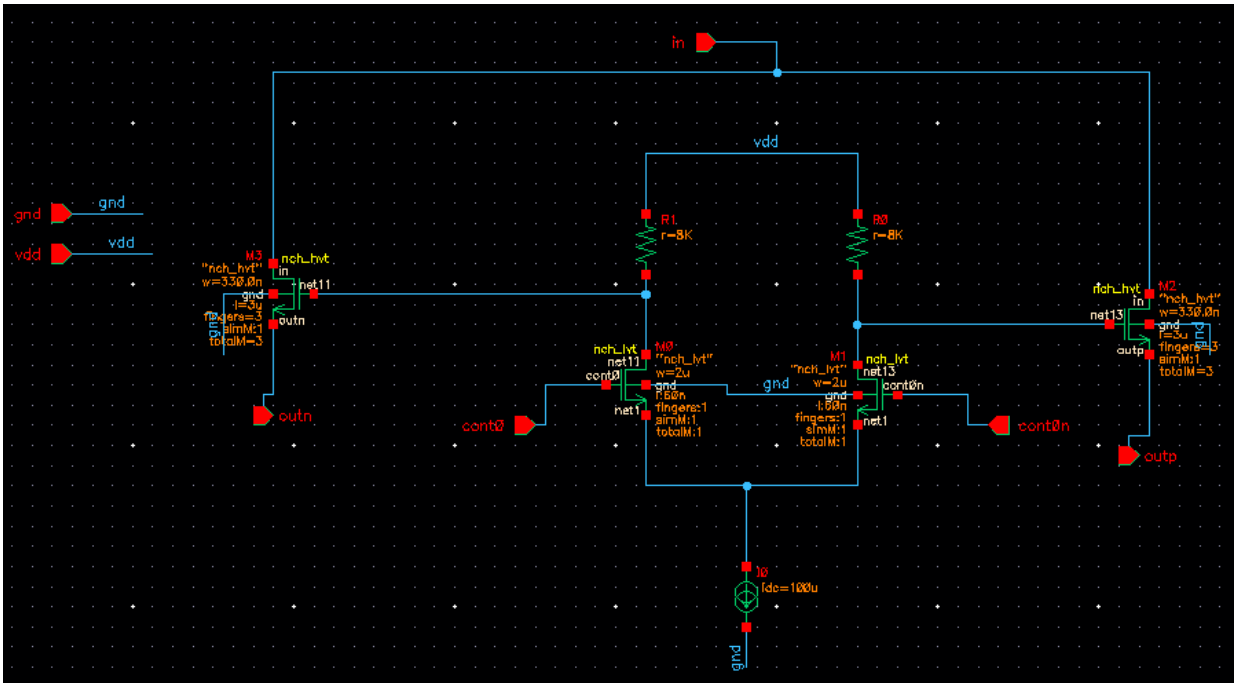


Figure 3.35: Full implemented switch block

Fig.3.36 and Fig.3.37 show the performance of the differential switch block.

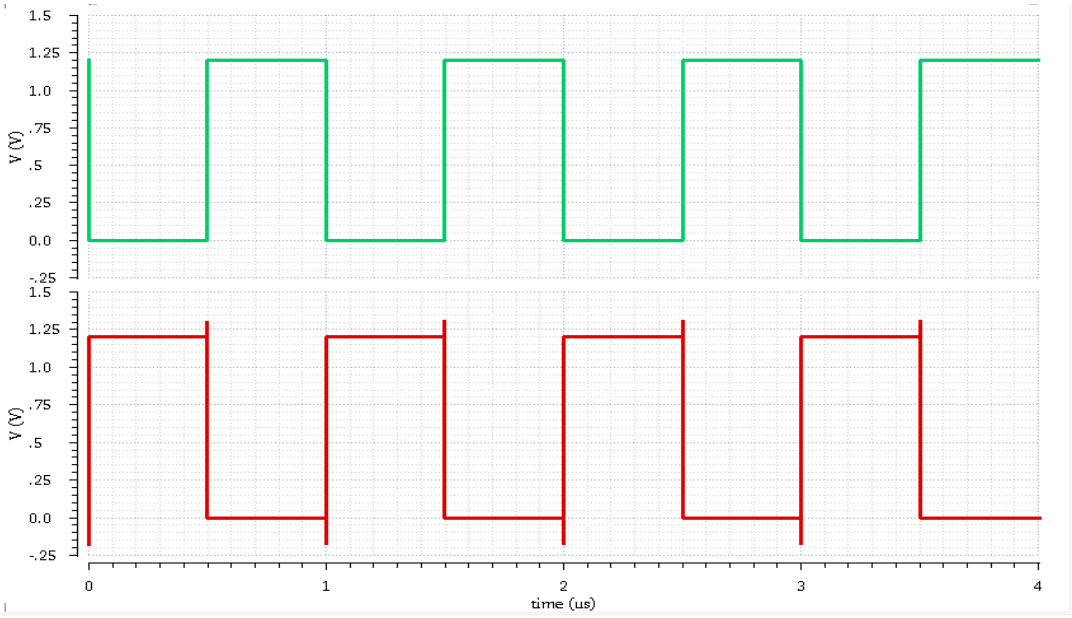


Figure 3.36: The output voltage signals from the 2 differential switches. Glitches that appear on the signals are due to the inverter CMOS placed between the 2 differential control signals entering the switch block

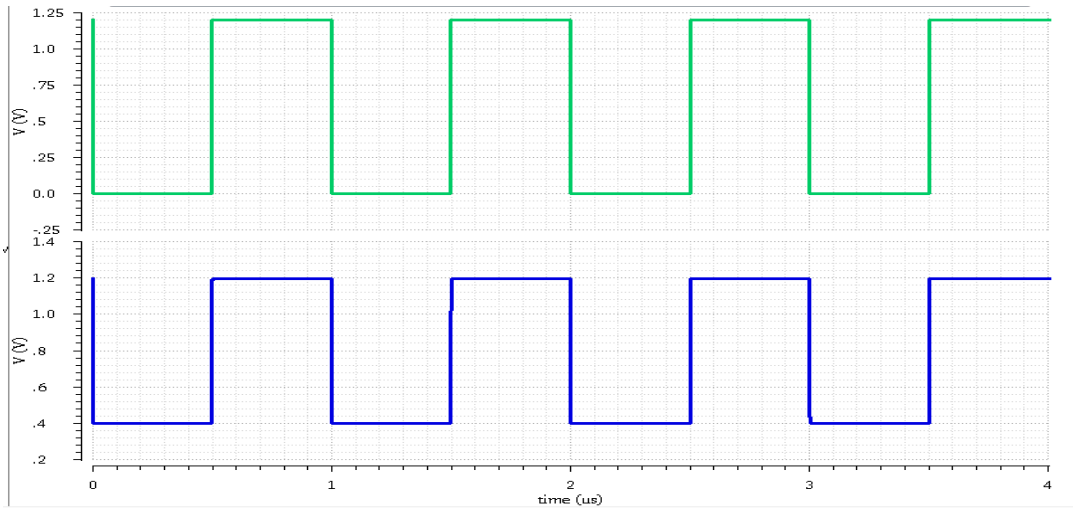


Figure 3.37: Input Control Signal to the Switch block (in green), and output signal from the Switch block (in blue). The output signal has rail-to rail voltage of 0.8 V unlike the input which has rail-to rail of 1.2V

Fig.3.38 shows the full implemented current cell containing one of the current source branches, the current mirror, and the switch block (Differential switches and their switch driver).

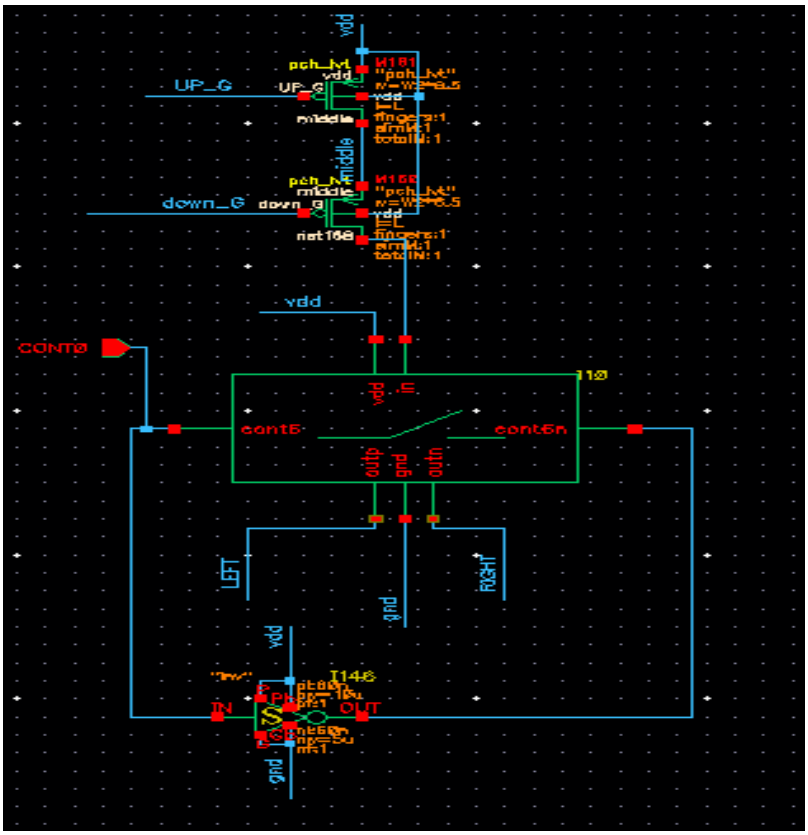


Figure 3.38: Full Detailed current cell

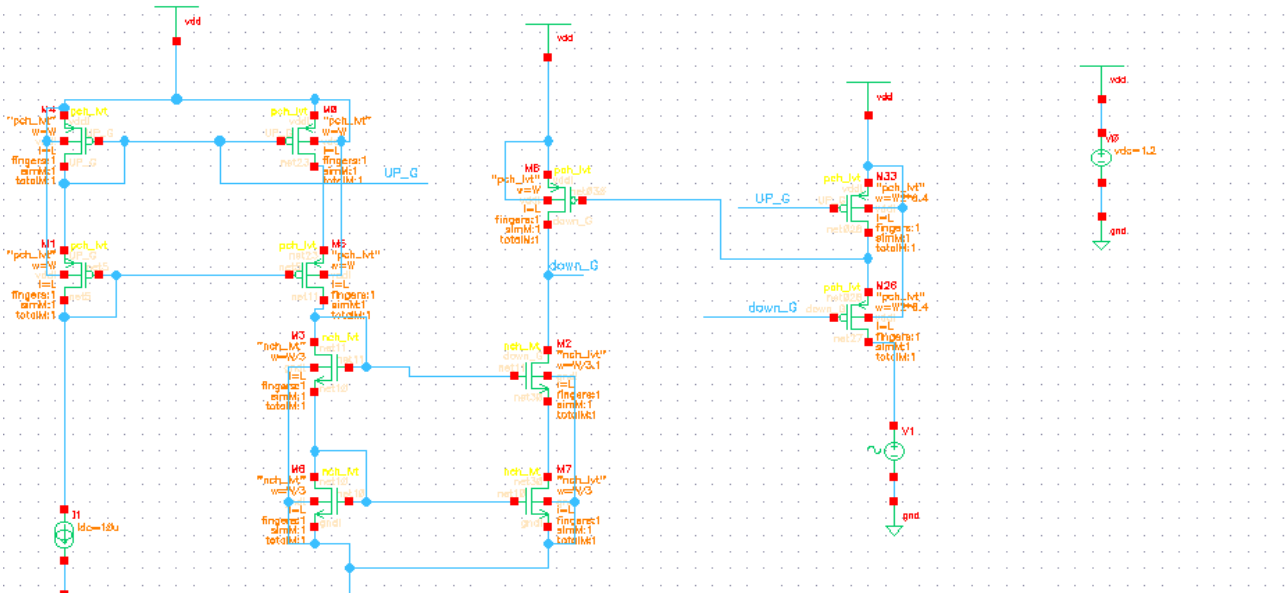


Figure 3.39: Full Detailed current cell with the current mirror

3.8 DAC Buffer Design

At the output of the DAC core, a buffer is implemented in order to enhance the output impedance and accordingly enhance the performance of the converter. The output buffer also isolates the DAC core from the reconstruction filter to prevent any undesired loading between the 2 blocks.

The requirements of the buffer are relaxed regarding the required GBW as the input signal to the buffer will have frequency of 100 KHz (channel bandwidth) with full scale amplitude of 0.2V. Also the buffer will be driving an Op-Amp RC filter, discussed in section 3.8, whose input impedance is estimated to be in the range of tens of KΩs, therefore the architecture chosen for the buffer design is the two stage fully differential Op-Amp shown in Fig.3.40.

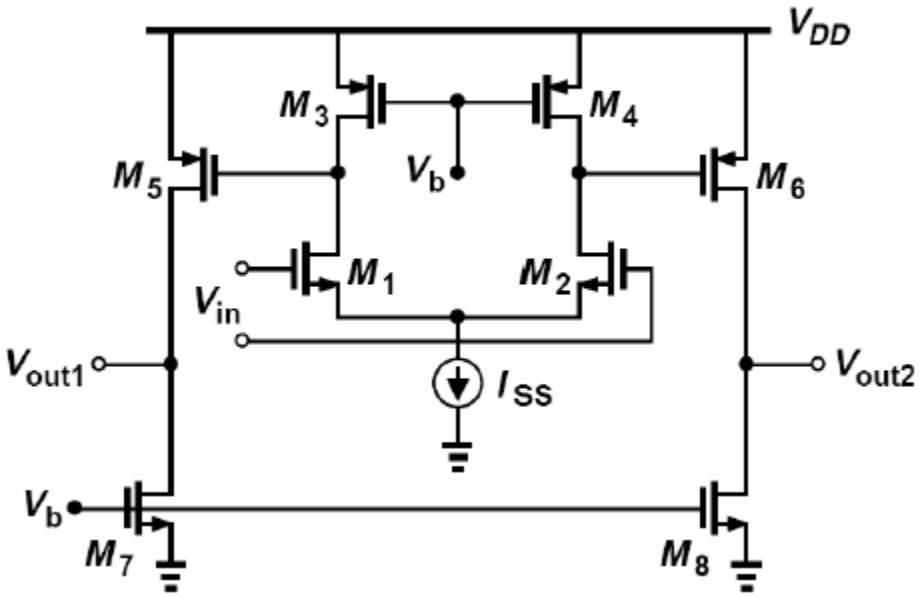


Figure 3.40: Fully Differential two stage Op-Amp

Using this architecture allows a high output swing which gives a good degree of freedom in the design of the filter. A common mode feedback circuit is essential in order to properly define the output common mode voltage level of the buffer. The common mode feedback circuit implemented senses the differential output voltage by 2 resistors. These sensed voltages are summed up in the input node of the error amplifier and then compared with the required common mode level. The action taken by the error amplifier is impacted back on the 2 stage op amp through the Vb node highlighted in Fig.3.40.

The main drawback of using this architecture is the presence of power hungry second stage which needs high current values to pass through it. Another drawback is the presence of 2 poles in the path of the signal degrading the stability of the buffer. In order to overcome the

degradation of the stability specification and increase the achieved phase margin, compensation capacitor is added across the second stage. Adding the compensation capacitor works according to the pole splitting technique of compensation, where the dominant pole is pulled to lower frequency in order to ensure reaching the GBW frequency early enough before the non-dominant pole starts to be in action.

The addition of compensation capacitor across the gate and drain of M5 and M6 transistors in Fig.3.40 (second stage transistors) will make the first stage output node be the dominant pole node as by miller effect, the added compensation capacitor will increase the capacitance of the first stage output node by $A_{v2}*C_{comp}$, where A_{v2} is the second stage gain.

Another effect of the addition of the compensation capacitor is the creation of a right half plane zero which in turn will degrade the stability, so in order to get rid of the effect of the right half plane zero, resistance is added in series with the compensation capacitor. The value of this resistance is equal to 1/gm2 where gm2 represents the trans-conductance factor of the second stage transistor.

Fig.3.41 shows the implemented two stage op amp with its common mode feedback circuit.

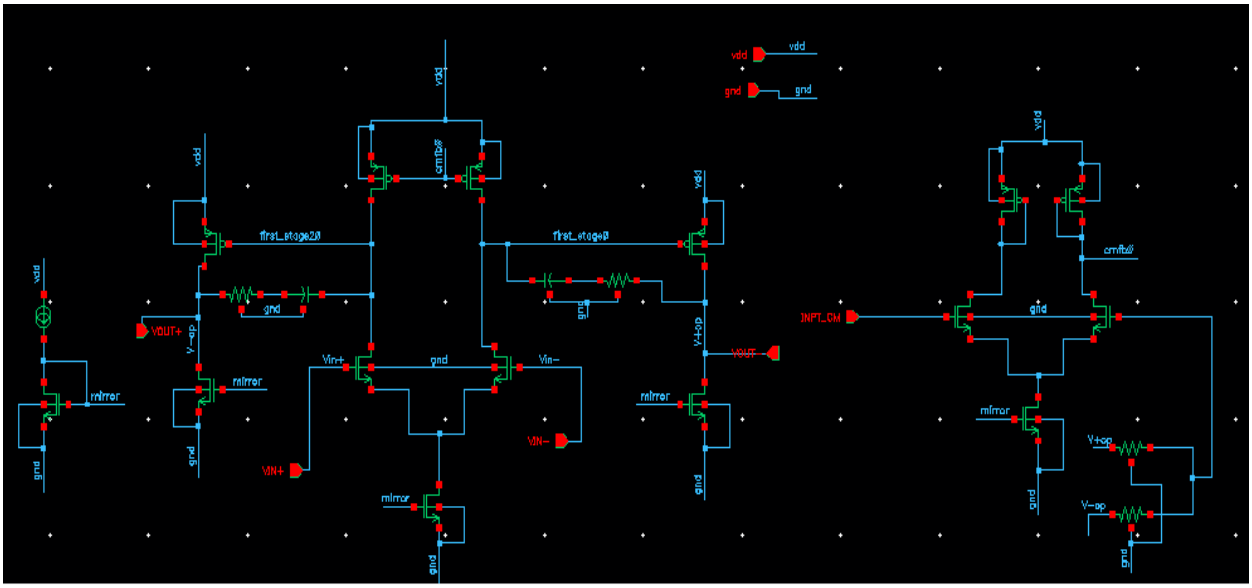


Figure 3.41: Two stage op amp with CMFB circuit

Current consumed in the first stage of the op amp is 150 μA, while that consumed in each branch of the second stage is 350 µA. Compensation capacitor of 800 fF is used in series with 400 ohm resistor for compensation across the second stage. The output common mode level is adjusted at 700 mV. Fig. 3.42 and Fig.3.43 show the AC response of the op amp in both open loop and closed loop configurations respectively. In open loop, the gain of the op amp is 43 dB with bandwidth of 1.62 MHz leading to overall GBW of 180 MHz. Closed loop configuration shows closed loop bandwidth of 168 MHz which is consistent with the fact that GBW ≈ closed loop bandwidth.

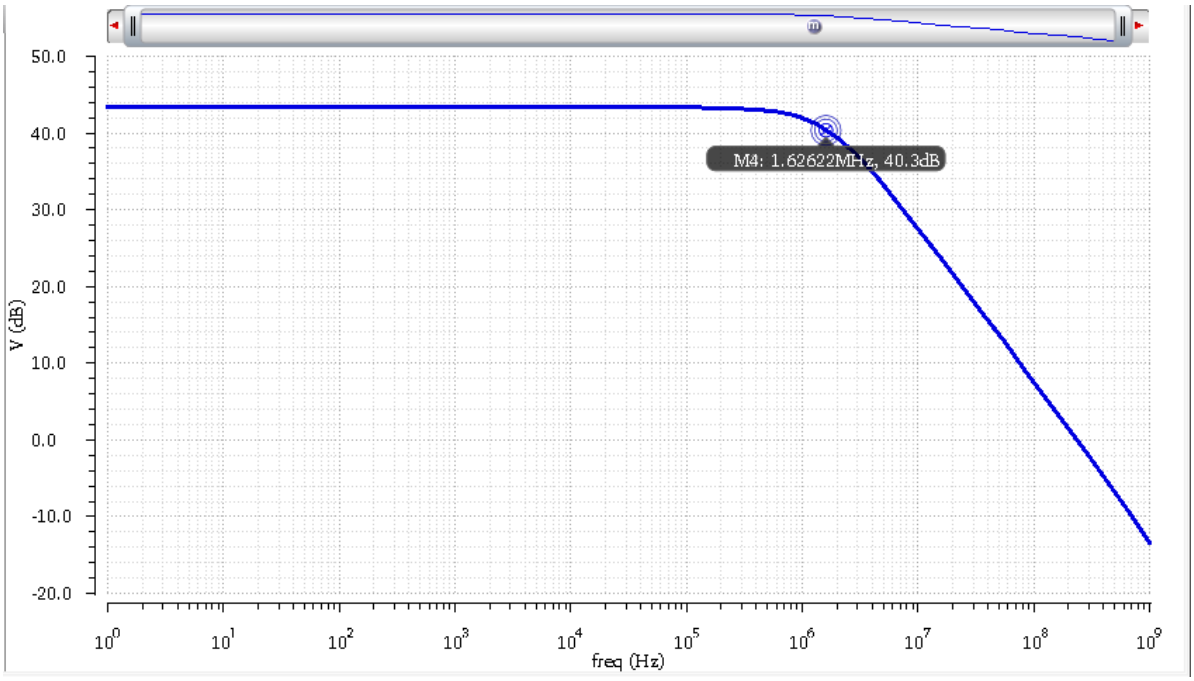


Figure 3.42: Open loop AC response of the op amp

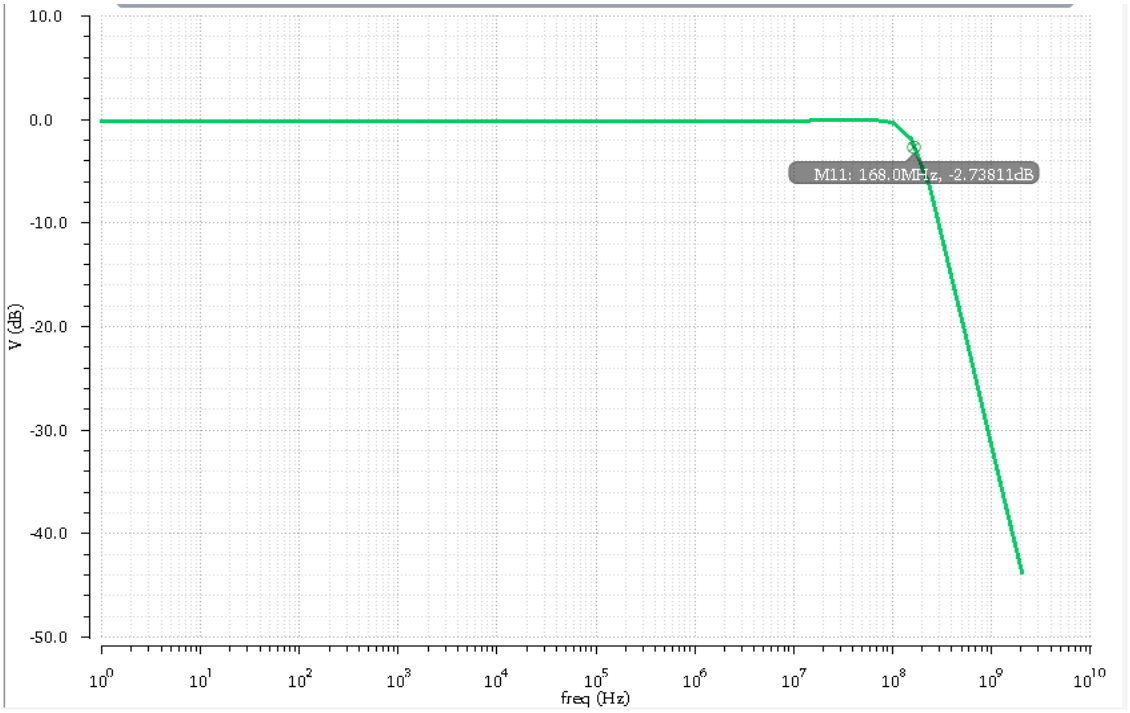


Figure 3.43: Closed loop AC response of the op amp

For testing the stability of the op amp, a cmdm probe was placed at the output of the op amp in order to break the loop at a common point. Using stb analysis on the spectre, PM was calculated to be more than 63^o as depicted in Fig.3.44. Fig. 3.45 and Fig.3.46 give better insight on the linearity specifications of the op amp.

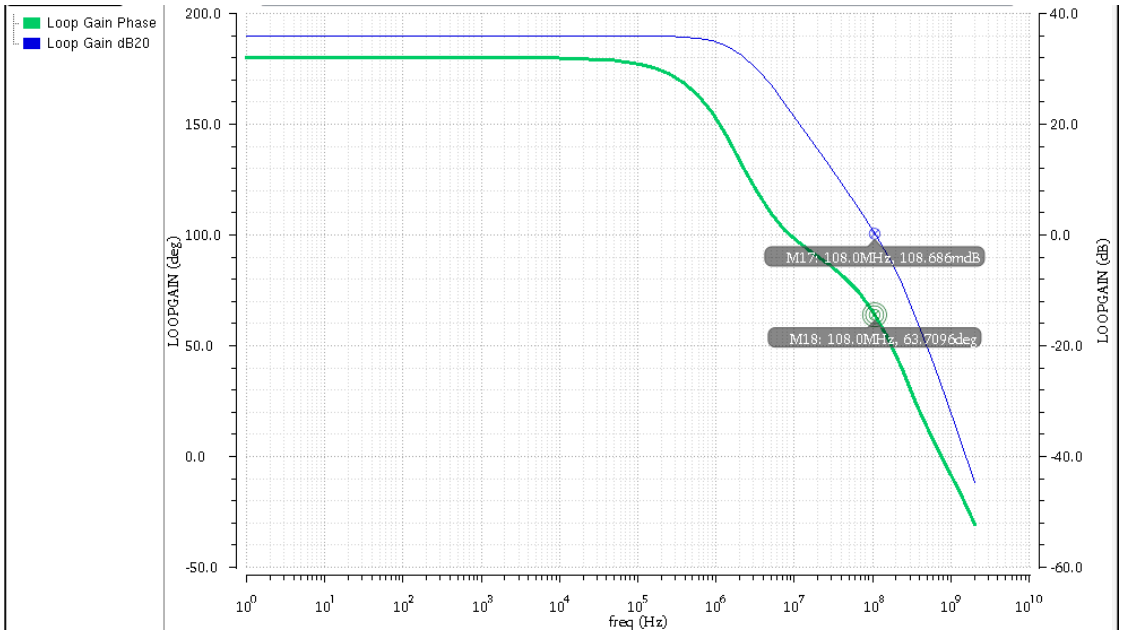


Figure 3.44: Op Amp loop gain and phase

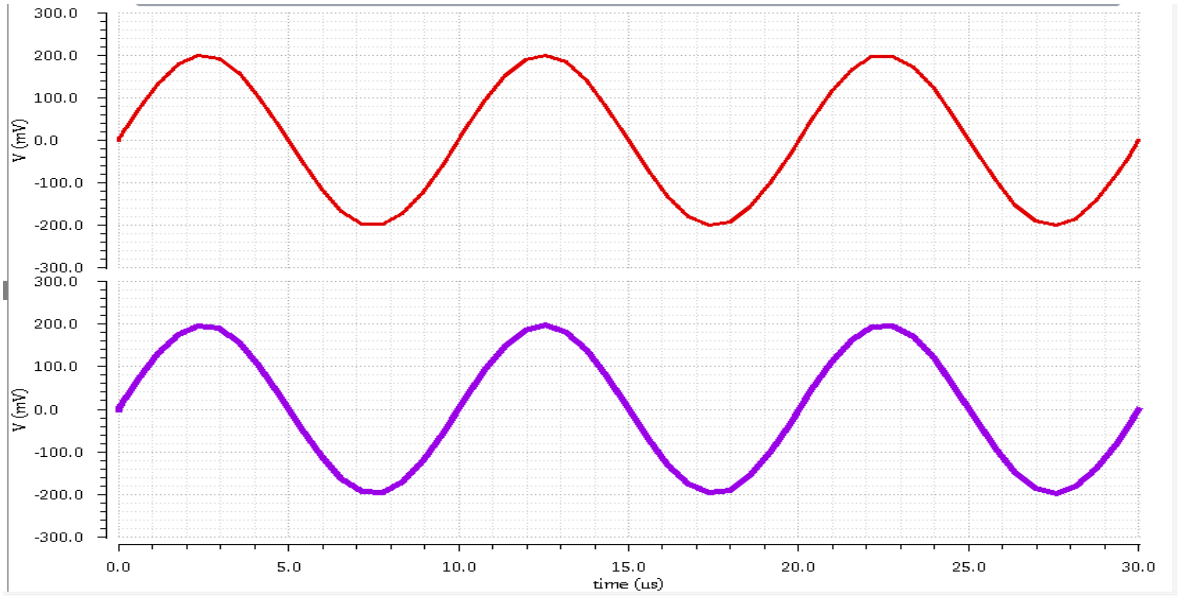


Figure 3.45: Transient input and output signals of the buffer

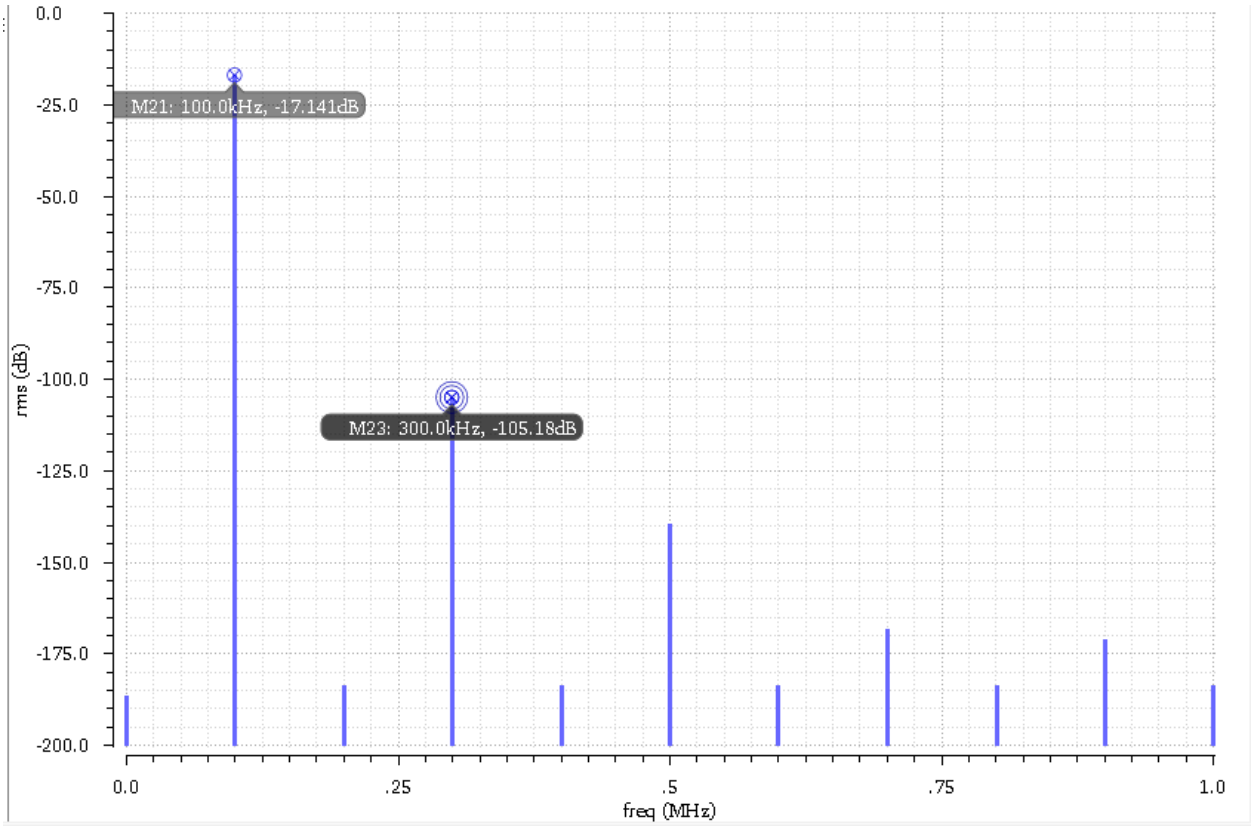


Figure 3.46: Harmonics of the output signal from the buffer. The third harmonic is 88 dB below the fundamental

The output noise of the op amp was simulated and depicted in Fig.3.47. The total integrated noise summary showed integrated noise of about $1.02*10^{-8}\ V^2/Hz$. This noise value gives SNR of about 64 dB.

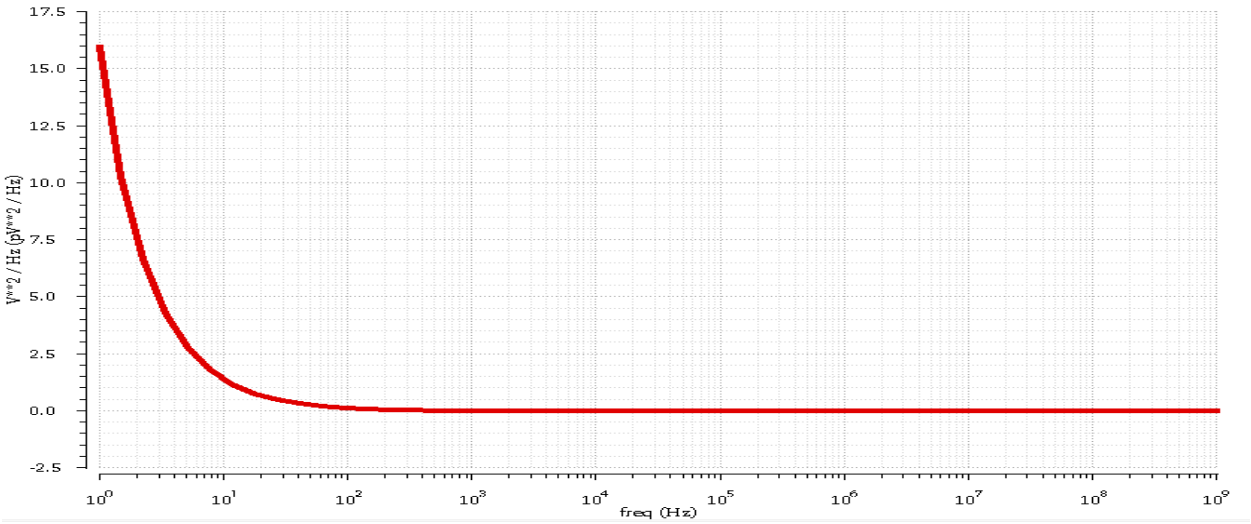


Figure 3.47: Output noise of the op amp

3.9 DAC Reconstruction Filter Design

As stated in section 3.1, designing a digital to analog converter requires the presence of a reconstruction filter to get rid of the unwanted aliases in the spectrum and to smooth the staircase like output signal, eliminating the spurs from the spectrum. The specifications of the required filter are depicted in table 3.4. In this section, the design of the reconstruction filter is presented but before that we would present a brief overview on the filter design.

Table 3.4: Reconstruction Filter required Specifications

Characteristics	**Required**
Type	LPF
Bandwidth (-3 dB point)	120 KHz
Attenuation	>50 dB
Required attenuation frequency (stop band)	500 KHz
NF	<45 dB
P1 dB	-1 dBm

3.9.1 Filter Design Overview

'Filter' is a general term. It is used to separate desired parts from unwanted parts of an original set. In an electrical point of view, filters have a variety of duties. Filters are used to cut a particular range of the frequency spectrum of an electrical quantity out of the overall frequency spectrum to retrieve the interesting information out of the plenty of information channels and noise. They can be used as equalizers or matched filters in transmission channels in order to compensate distortion caused by the transmission channel itself. Filters can also be employed for smoothing purposes at the output of digital-to-analog converters.

Filters are defined mainly by their frequency domain response. Hence, it is more convenient to gain insight on filters' characteristics from frequency domain plots. Fig. 3.48 shows a practical frequency response template for an analog filter. This response divides the frequency spectrum into three regions:

- **Pass band:** Is the range of frequencies that can pass through a filter. This region is determined by the pass band edge frequency(ω_p). The "flatness" in the pass band is measured by the amount of "ripple" that the magnitude of the signal undergoes (A_p). If ripples are very large then the filter -undesirably- changes the frequency contents of the wanted part from the signal.
- **Transition band:** It is the transition between the pass band and the stop band (that includes undesired signal components) it represents how much of the interferer remains alongside the signal. For better frequency selectivity, this band must be sufficiently narrow.
- **Stop band:** The range of frequencies over which the filter does not allow signals to pass. This range is specified by the stop band edge frequency (ω_s). The attenuation (A_s) must be large enough to suppress the interferer signal to a level that is far below the desired signal.

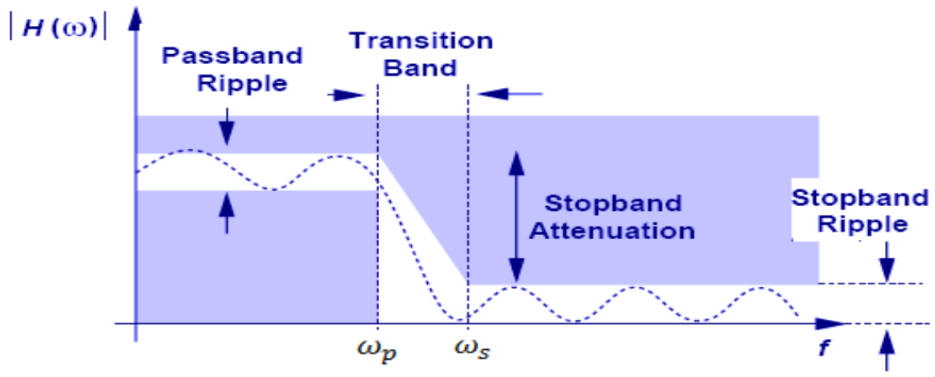


Figure 3.48: Practical Frequency response template for a low pass filter

Filters divide the frequency spectrum into different bands and, hence, shape the transfer characteristic of the filter. As explained earlier, stop-bands denominate frequency ranges, over which some signal components are blocked from the signal. In contrast to this, pass-bands are frequency ranges, over which the signal can pass from the input to the output without blocking. The allocation of the stop-bands and pass-bands specifies the type of the filter. Feasible filter types are the low-pass filter, the high-pass filter, the band-pass filter, the band-stop filter, the all-pass filter and, or the equalization filter.

Low-pass Filter: Low pass filters are filters that pass the signal's frequency components below a cut-off frequency (also called corner frequency) and reject the frequency components above it. Low-pass filters could be considered the mostly used type of filters in communication systems.

High-pass Filter: Opposite to LPF, high-pass filters allow frequency components above the corner frequency to pass and reject components below it.

Band-pass Filter: Band-pass filters can be considered a combination of both low-pass and high-pass filters. It passes signals in a certain bandwidth only. Band pass filters are usually considered the first block in any wireless receiver chain as it's responsible for catching the required channel sent.

Band-stop Filter: Band-stop filters do the inverse response for band-pass filters. Frequencies between two corner frequencies are allowed to pass and any other frequencies are rejected. A special case of band-stop filters is the notch filters whose transfer function contains a zero at certain frequency.

All-pass Filter: Unlike the previous types, All-pass filters do not change the amplitude of the input; rather they only change the phase. Hence, they can be used for phase equalization at the digital demodulators.

Equalization Filter: Equalization filters or equalizers are used for the correction of an uneven frequency response characteristic for smoothing purposes of signals or systems. The transfer function is different from the above considered and is often only usable for one special case of distorted signal.

Figures 3.49 through 3.51 show the frequency response of these different filter types.

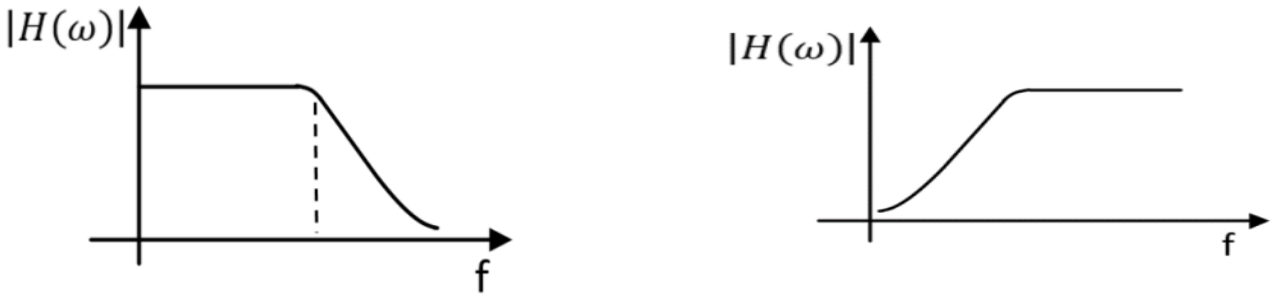


Figure 3.49: The frequency response of low pass filter and high pass filter respectively

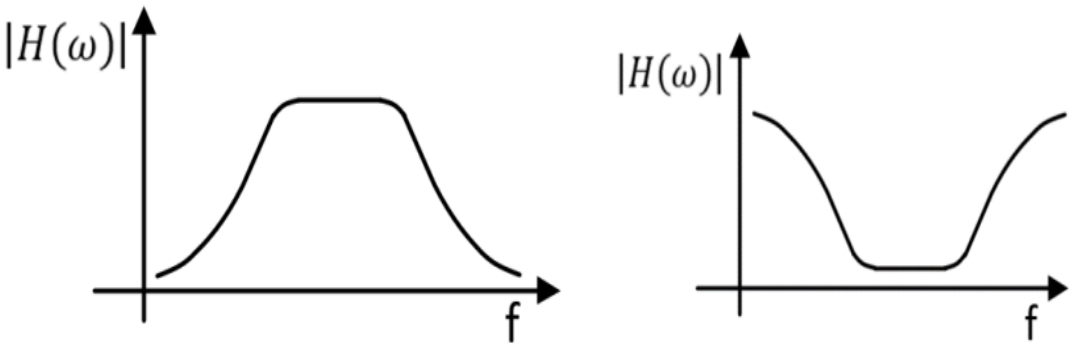


Figure 3.50: The frequency response of band pass filter and ban stop filter respectively

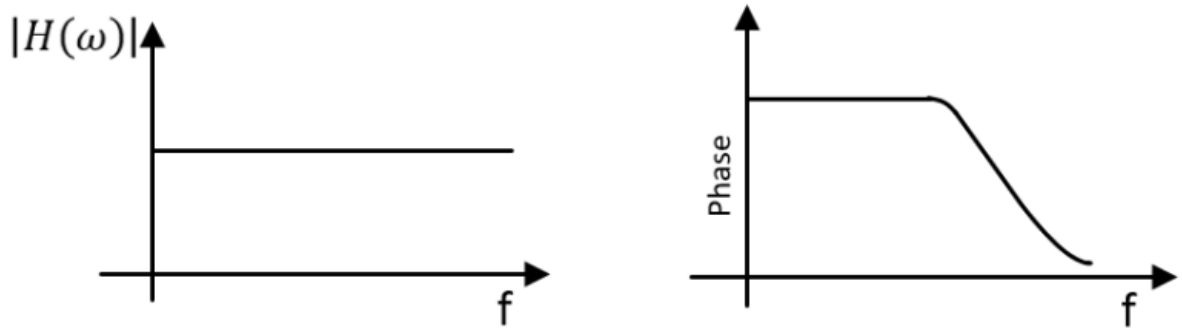


Figure 3.51: The amplitude and phase frequency response of all-pass filter

3.9.2 Filters Mathematical Representation

Considering filters as Linear Time-Invariant (LTI) systems, the relation between its input $x(t)$ and the output $y(t)$ is determined using the convolutional integral between $x(t)$ and the impulse response $h(t)$.

$$y(t) = \int_0^t x(\tau)h(t-\tau)d\tau \tag{3.17}$$

Taking the Laplace transform for both sides in (1.1) gives

$$Y(s) = X(s)H(s) \tag{3.18}$$

where s is the complex frequency variable and $H(s)$ is the filter's transfer function. For most of filters under consideration $H(s)$ is a rational function of s, i.e., the ratio between two real and finite polynomials in s.

$$H(s) = \frac{N(s)}{D(s)} = \frac{a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0}{s^m + b_{n-1}s^{m-1} + \cdots + b_1 s + b_0} = a_n \frac{\prod_{i=1}^{n}(s-z_i)}{\prod_{j=1}^{m}(s-p_i)} \tag{3.19}$$

Where $N(s)$ and $D(s)$ are the numerator and denominator polynomials respectively, m is the filter's order, z_i and p_i are the roots of the numerator and denominator respectively (Also called zeros and poles). For the system to be causal; m should be greater than n. For stability reasons, a_i and b_i must be real and b_i must be positive [15].

If the signal is sinusoidal with angular frequencyω, this allows the substitution of s by $j\omega$ in (3.19).

$$H(j\omega) = |H(j\omega)|e^{-j\phi(\omega)} \tag{3.20}$$

where $H(j\omega)$ can be considered as the Fourier transform of $h(t)$. Equation (3.20) represents the filter's transfer function in terms of its magnitude and phase response. It is a usual practice to represent the magnitude of $H(j\omega)$ in the form

$$A(\omega) = 20\log(|H(j\omega)|) \quad (3.21)$$

This gives the filter gain in decibels. However, in most cases, we talk about the filter attenuation or loss, - $A(\omega)$, also in dB.

3.9.3 Filters Mathematical Approximation

A physical realization of an ideal filter transfer function is impossible. Hence, a frequency scheme is given, which is dependent on many system parameters and trade-offs. Within this frequency scheme the real filter transfer function has to be located. Important selection criteria are for example a fast transition from the pass-band to the stop-band or a minimum filter distortion. As we approach the ideal filter transfer function, we need to put greater effort regarding number of elements, power consumption or costs. Various approximations for the ideal filter transfer functions are realizable. Important and popular approximation functions are Butterworth approximations, which are explained in more detail, Chebyshev and inverse Chebyshev approximations, elliptic or Cauer approximations, and Bessel approximations.

Butterworth Approximation

The Butterworth filter was first described by Stephen Butterworth and is nowadays popular and often used. The Butterworth filter approximation is demonstrated on the normalized low-pass filter. Frequency transformation can be made to realize different types of filters. The normalized Butterworth function of *N*th-order is given by:

$$H(j\omega) = \frac{1}{\sqrt{1+\varepsilon^2\omega^{2n}}} \quad (3.22)$$

The Butterworth low-pass filter has some important characteristics:

- The magnitude function of a Butterworth low-pass filter is monotonically decreasing for $\omega > 0$ and the maximum of $|H(j\omega)|$ is at $\omega = 0$.
- An Nth-order Butterworth low-pass filter has a maximally flat magnitude function.
- A Butterworth filter shows an overshoot in the step response in the time domain, which worsens at rising filter-order N.
- Butterworth filter is characterized by its low group delay.
- The Butterworth response completely avoids ripples in the pass/stop bands at the expense of the transition band slope.

The distribution of the poles and zeros in the s-plane are characteristic for filters and can also be used for filter identification. A normalized Butterworth low-pass filter has its poles in the left half plane on the circumference of the unity circle with the center in the point of origin of the s-

plane. The poles are spread equidistant over the half circle in the left half s-plane. The location of the poles s_k of $H(S)$ in the left half plane can be calculated as:

$$s_k = \sigma_k + j\omega_k = \cos\left(\frac{2k+N-1}{2N}\pi\right) + j\sin\left(\frac{2k+N-1}{2N}\pi\right) \quad (3.23)$$

In case of an even filter order no real pole is existent, an odd filter order has exactly one real pole.

Chebyshev (Equiripple) Approximation

Chebyshev filters offer a frequency response that approximates the ideal low-pass filter more precisely than a Butterworth filter. The transfer function has only poles and lack of any finite zeros. The pass-band of a Chebyshev filter exhibits a pass band ripple, which ranges between two constant values. The ripple amplitude can be adjusted freely. The ripple amplitude is directly proportional to the filter slope in the transition-band (filter selectivity) and the overshoot of the step response in the time domain. The greater the pass-band ripple, the higher the filter selectivity and the overshoot. The number of ripples depends on the order *N* of the filter. For frequencies greater than the cut-off frequency the filter has a monotonically decreasing magnitude function similar to Butterworth filters.

Inverse Chebyshev Approximation

The inverse Chebyshev filter has complementary properties of the Chebyshev filter concerning the magnitude response. In the pass-band the magnitude function is monotonically decreasing for $\omega > 0$ and the equiripple appears in the stop-band. The filter selectivity is not as high as in Chebyshev filters. The zeros in the filter transfer function bring an additional realization effort.

Elliptic or Cauer Approximation

Elliptic filters have an equal ripple in the pass-band and in the stop-band. Therefore elliptic filter is also called double Chebyshev filter. The two ripples are individually adjustable. The transition between pass-band and stop-band shows high filter selectivity. The sharp filter edge is realized by a transfer function using a balanced combination of poles and zeros. The magnitude function of an elliptic filter is the best approximation of the ideal low-pass filter compared to Butterworth and Chebyshev filters.

Bessel Approximation

In contrast to Butterworth or Chebyshev filters, which approximate the magnitude function of an ideal low-pass filter, the Bessel filter approximates the phase response. Hence, the Bessel filter is also called maximally flat group delay filter. In the pass-band the Bessel filter has a distortion free transmission and keeps the group delay constant. The constant group delay in the pass-band results in a step response, which shows no overshoot. The filter selectivity is not as good as in

Butterworth filter structures. The filter order N is the only parameter to adjust a normalized Bessel filter. The value of N defines the phase and the magnitude response. The higher the filter order, the larger the frequency range with constant group delay and the higher the filter selectivity.

3.9.3 Analog Active Filter Topologies

Various filtration techniques and topologies have been developed in the previous decades. Each topology is used for certain applications and frequency ranges. Fig. 3.52 shows a classification of active filters topologies used across different frequency ranges. The widely used active filter topologies are Op amp-RC, Gm-C and Switched-C filters.

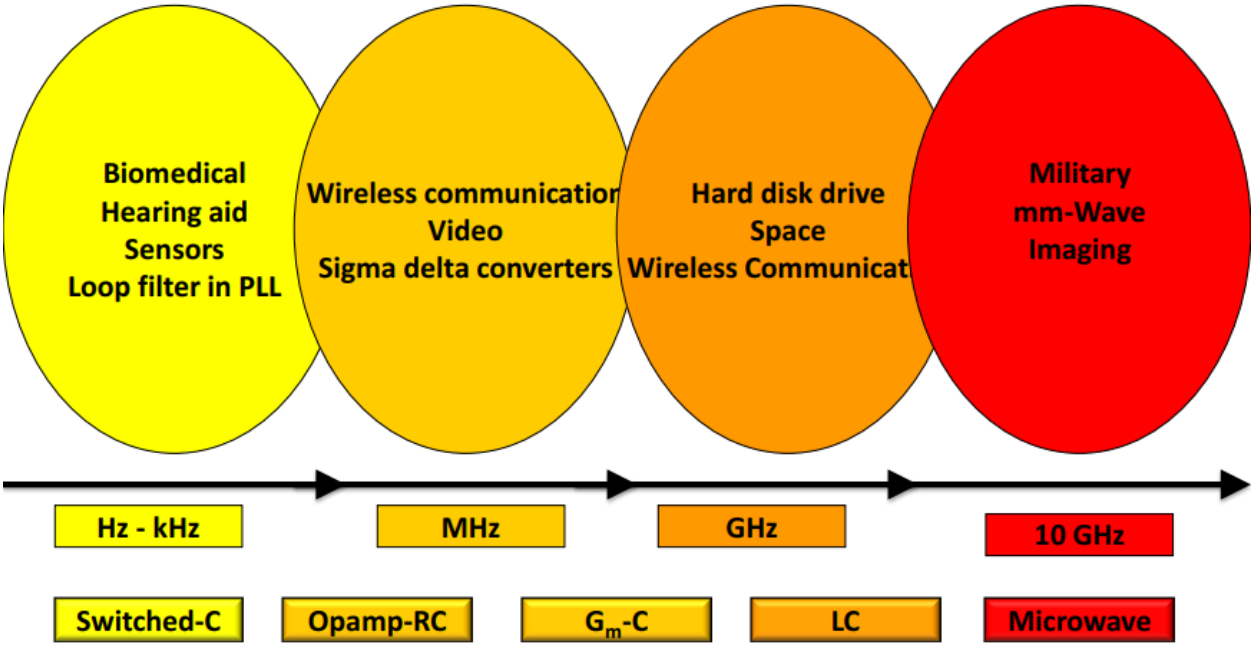


Figure 3.52: Classification of active filters topologies used across different frequency ranges

Opamp RC Filters

The RC active filters are still being used in low-frequency applications. These filters use op amps, resistors and capacitors as basic elements. In integrated circuits, the resistors can be implemented either as diffused or poly-silicon resistors. The absolute accuracy of the resistance is in the order of 30%. The linearity of the poly-silicon resistors is quite good and the accuracy of resistor ratios can be as good as 1%. Nevertheless, the accuracy of the RC products can be as worse as 50%. Even more, the RC product is a strong function of both technology tolerances and temperature variations. In filtering applications, the frequency of the poles and the frequency of the zeros are determined by RC products. Typically, the quality factor of the filter is determined by resistors and/or capacitor ratios. Because the filters are more sensitive to variations in the frequency of the poles than to the quality factor of the poles, the accuracy of the active RC filters is quite low. The accuracy of these filters can be improved if certain kind of tuning, either on chip or externally, is included. The active RC filter approach is almost not used for high-frequency applications because there are some other major disadvantages; e.g. the resistors are implemented in long strips over field oxide or over active area. Therefore, there are distributed capacitors connected to the resistor. These distributed capacitors limit both the precision and the high frequency performance of the filter. Furthermore, due to the use of the resistors buffered op amps are required. The two poles buffered Opamp limits the frequency response of the system and increases dramatically the power consumption.

OTA-C Filters

A more versatile technique is the so-called OTA-C. OTA-C filters have already been used for frequency ranges from audio frequencies (a few kHz) up to very high-frequency (GHz range). The OTA is a voltage-to-current transducer with very high output impedance. The C stands for a capacitor implemented in CMOS technology. The OTA has tunable trans-conductance and can work at high frequencies, which make it most attractive for fully integrated high frequency filter design. It is used as an open loop amplifier in gm-C filter design. Practical OTAs will have finite input and output impedances. At very high frequencies, the OTA trans-conductance will be frequency dependent. These non-ideal characteristics will degrade frequency performances of OTA-C filters. Practical OTAs also exhibit nonlinearity for large signals and have noise, which will affect the dynamic range of OTA-C filters. In fully integrated high frequency gm-C filter design, automatic tuning circuitry is usually included on the same chip to overcome the effects of parasitics, temperature and environment.

Switched-C Filters

In switched capacitor filters, the main characteristics are determined by a clock frequency and by capacitor ratios. In CMOS technologies both parameters can be controlled with an accuracy as high as 0.5%. Even more both parameters, clock frequency and capacitor ratios, are almost independent of the process parameter tolerances and temperature variations. Hence, a major advantage of this technique is the high accuracy of its integrator time constant. Therefore, the use of additional tuning circuits is avoided. These factors make these techniques very attractive for the design of high-performance analog integrated filters. In low frequency applications, the high DC gain and high Gain-Bandwidth Product (GBW) of the CMOS Operational Amplifiers (Op amps) make these filters insensitive to the parasitic capacitors. Furthermore, as the poly-silicon capacitors are quite linear and due to the Op amp local feedback (small ac-signal at its input terminals), the Total Harmonic Distortion (THD) of the switched-capacitor filters is very low, e.g. THD < - 70 dB. Unfortunately, these advantages of switched-capacitor filters are not necessarily maintained for high-frequency applications. This is mainly due to the finite parameters of the Op Amp (finite DC gain and finite GBW), finite resistance of the switches, and clock feed through effect. In high-frequency applications, the Op amp has to be fast enough to settle to the right output within a half clock period. For a settling precision of 0.1 %, the settling time should be higher than the GBW of the Op amp at least by a factor 7. However, due to the additional capacitors connected to the OPAMP output the effective settling time increases and as a result even larger GBWs are required. Typically a factor 10 (or more) instead of 7 is very often used. In order to guarantee the stability of the closed loop system, the second pole of the Op amp should be placed around 3 times higher than GBW. For high-frequency applications it is very difficult to satisfy last constraint because it means that the frequency of the Op amp second pole should be higher than the clock frequency by a factor of 30. The other major limitation of switched-capacitor systems is the relatively low DC gain of the CMOS Op amp. The DC gain of the Op amp has to be high enough to reduce the ac signal at its input terminals. If so, the circuit becomes parasitic insensitive and allows the total charge and discharge of the capacitors.

3.9.4 DAC Reconstruction Filter Implementation

The output of the DAC consists of rectangular pulses summed together. In the frequency domain, these pulses become a weighted sinc function. This sinc function in the output causes spectral images. To reduce the impact of the spectral images, a reconstruction filter is applied to the output of the converter. This filter consists of a multi-order low pass filter. By filtering the output, the pulse-like output is smoothed to appear sinusoidal.
For this design, a Butterworth low pass filter was chosen. This is because the Butterworth filter has a maximally flat response opposed to other common filter types as stated in section 3.9.3.In order to define the required order of the filter; the magnitude frequency response in (3.22) is used.

Using equation (3.22), we can find that the pass band magnitude can be estimated as:

$$Ap\ (in\ dB) = 10\log(1 + \epsilon^2) \qquad (3.23)$$

While the stop band magnitude can be estimated as:

$$As\ (in\ dB) = 10\log(1 + \epsilon^2\ \Omega s^{2n}) \qquad (3.24)$$

where $\Omega s = \frac{stop\ band\ frequency}{pass\ band\ frequency}$, n is the filter order, and ϵ can be set to 1 is set to one in order to have the gain of the filter through the pass band equal to 0 dB.
According to the specs required from the filter, namely having pass band width of 120 KHz and stop band attenuation of 50 dB at 500 KHz, it can be found that Ωs is equal to 4.167. By substituting with the required values, the filter needed is a **4th order low pass filter.**

The 4th order low pass filter can be implemented as 2 successive biquads, each of them is considered a second order filter. For a second order filter, the general transfer function is shown in (3.25):

$$H(s) = \frac{\omega o^2}{s^2 + \frac{\omega o}{Q}s + \omega o^2} \qquad (3.25)$$

According to (3.25), for getting quality factor (Q) greater than ½ , complex poles will certainly be created.

As the filter to be designed has a cut-off frequency (corner frequency) in the range of hundreds of KΩs, it can be found, as obvious in Fig.3.52, that the most proper topology to be used is the **Op amp RC topology**.
Biquads in Op amp RC topology can be implemented either using single amplifier, or two amplifiers like Geffe Biquad or even three amplifiers like Tow-Thomas Biquad. To decide which kind of Biquads to be used, first we need to briefly characterize these 3 Biquad topologies.
The general performance parameters of a biquad are listed as:

- **Parasitic Insensitivity**: This parameter means that the filter performance is not affected by the parasitic capacitances of different nodes within the filter.
- **Small Parameter Sensitivity**: Small process variations on the R and C sections should have low effect on the quality of the filter.
- **Independent Tuning**: It is better to be able to independently tune the gain and bandwidth of the filter.
- **Small Component Spread:** Having components of nearly equal sizes enhances the accuracy of the filter performance.
- **No Op amp input Swing:** It is easier to design an op amp with no input swing which is achieved if the inputs i=of the op amp are connected to virtual ground nodes.

Single amplifier biquads cannot achieve the small parameter sensitivity and the small component spread at the same time unlike the two and three amplifiers biquads. But on the other hand, multiple amplifier biquads consume much power and need large areas to be implemented. So, as the system required to be implemented has constraints on the power consumed, single amplifier biquad was chosen.

The Multiple Feedback Single Amplifier Biquad architecture, shown in Fig.3.53, was chosen as it needs less power, less area and simple op amp design and specifications. The drawback of this architecture is its parasitic sensitivity as the parasitic capacitances can affect the performance of the filter but as this filter will be operated at relatively low frequencies; this draw back will not have any dominant effects. Another drawback of this architecture is its high parameter sensitivity. Changes due to process or temperature can affect several components of the filter degrading its performance. So, to overcome this drawback, programmability phase has been introduced to the filter in order to count for the process and temperature variations. This programmability phase will be discussed in details in next sections.

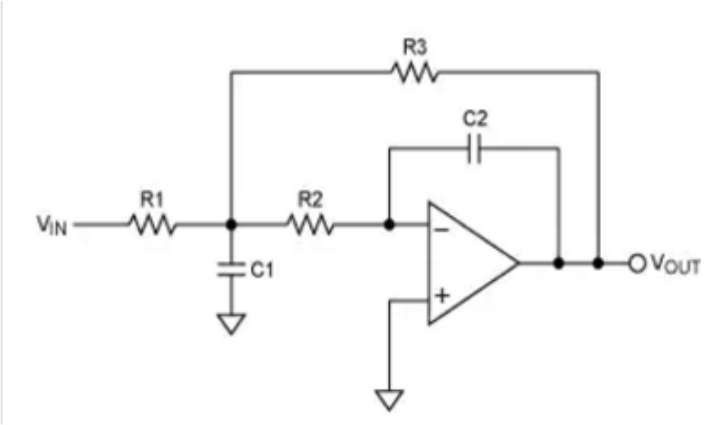


Figure 3.53: Multiple Feedback Biquad

The transfer function of this biquad is obtained as:

$$H(s) = \frac{\frac{-1}{R1.R2.C1.C2}}{s^2+\left(\frac{1}{R1.C1}+\frac{1}{R2.C1}+\frac{1}{R3.C1}\right)s+\frac{1}{R2.R3.C1.C2}} \qquad (3.26)$$

As the order of the filter is 4th order, the poles of its transfer function can be obtained through (3.23) which represents the poles equation according to butter worth normalized approximation.

The poles of this 4th order filter on the s plane are located at:

s1= (-0.3826 + j 0.9238), s2= (-0.3826 - j 0.9238), s3= (- 0.9238 -j0.3826), and

s4= (- 0.9238+j0.3826).

By distributing these 4 poles on the 2 stages and de-normalizing for cut off frequency of 100 KHz (2pi*100KHz), the required transfer functions of each of the 2 biquads are given as:

$$\mathrm{H1(s)} = \frac{(3.94* 10^{11})}{s^2+1161006s+3.94* 10^{11}} \quad , \mathrm{H2(s)} = \frac{(3.94* 10^{11})}{s^2+480663s+3.94* 10^{11}}$$

It can be noticed that the quality factor of the second biquad (1.307) is higher than that of the first biquad (0.54). This is in order to prevent the peaking of the signal in the middle node between the two stages.

Now we have the transfer function of the filter topology and the required transfer functions for the 2 biquads. So the values of resistors and capacitors can be easily estimated as shown in table 3.5. Components used are rppoly_wo resistors and mimcap 3 terminal capacitors.

Table 3.5: Filter Components Values

	R1	R2	R3	C1	C2
First Stage	45 KΩ	23 KΩ	45 KΩ	60 pF	11 pF
Second Stage	75 KΩ	36KΩ	75 KΩ	20 pF	11 pF

Fig.3.54 shows the implemented 4^{th} order butter worth reconstruction filter. The op amps used for its implementation are the same as the one used in the implementation of the buffer in section 3.8.

The achieved cut-off frequency of the filter is 116 KHz as depicted in Fig.3.55 with attenuation of 49 dB at 500 KHz. This corresponds with the required specification of the filter to a great extent.

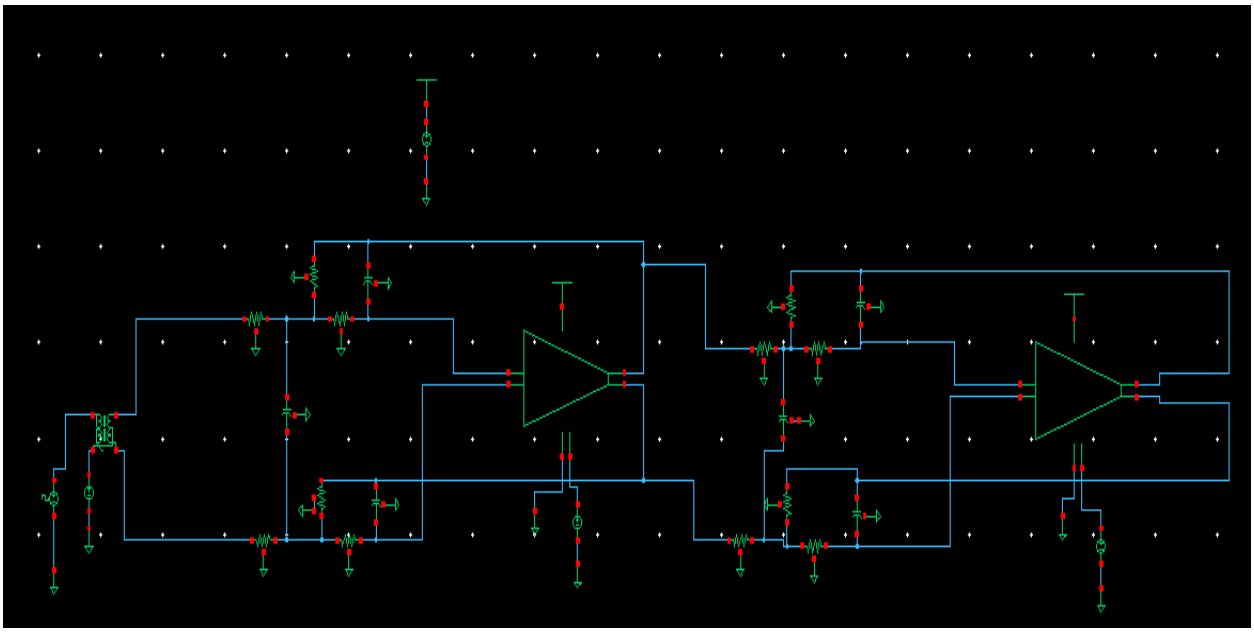


Figure 3.54: Implemented 4th order butter worth filter

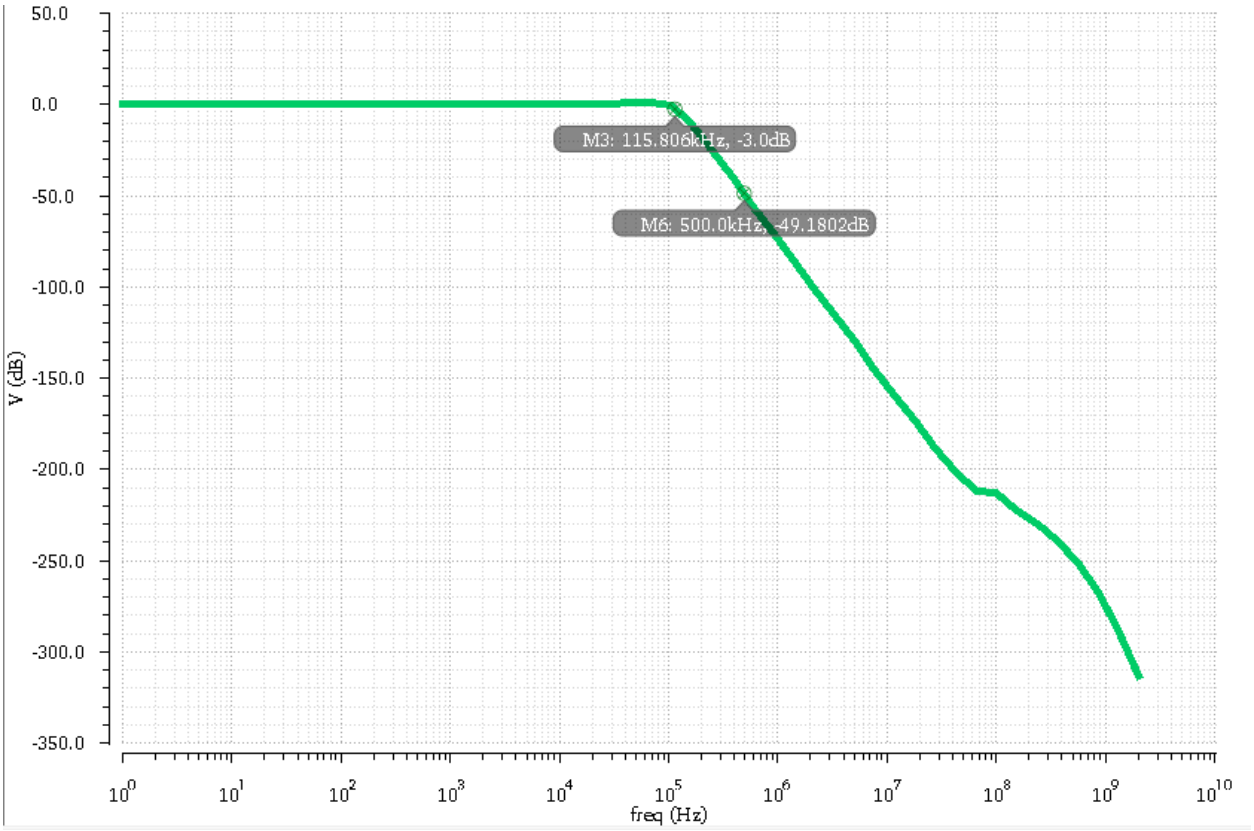


Figure 3.55: AC response of the reconstruction filter

Fig.3.56 shows the transient response of the filter. It can be seen that the filter causes delay of 4μs in the path of the signal. The noise figure of the filter was simulated and it achieved noise figure of 45 dB at the typical corners. By varying the temperature, it was observed, as shown in Fig.3.57 that the noise figure at 100 KHz didn't vary in a great manner.

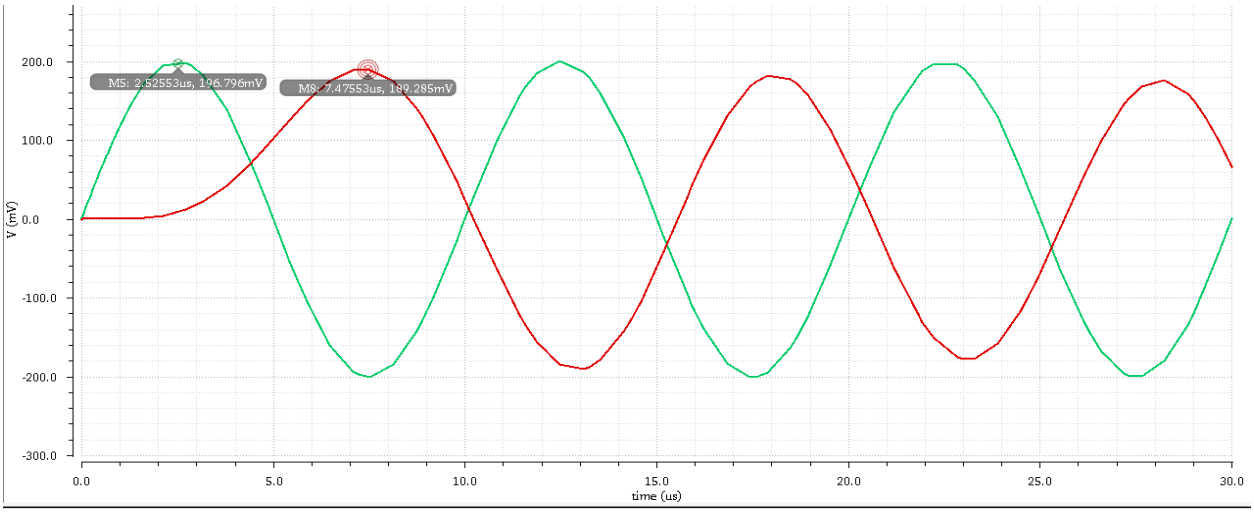


Figure 3.56: Transient response of the reconstruction filter

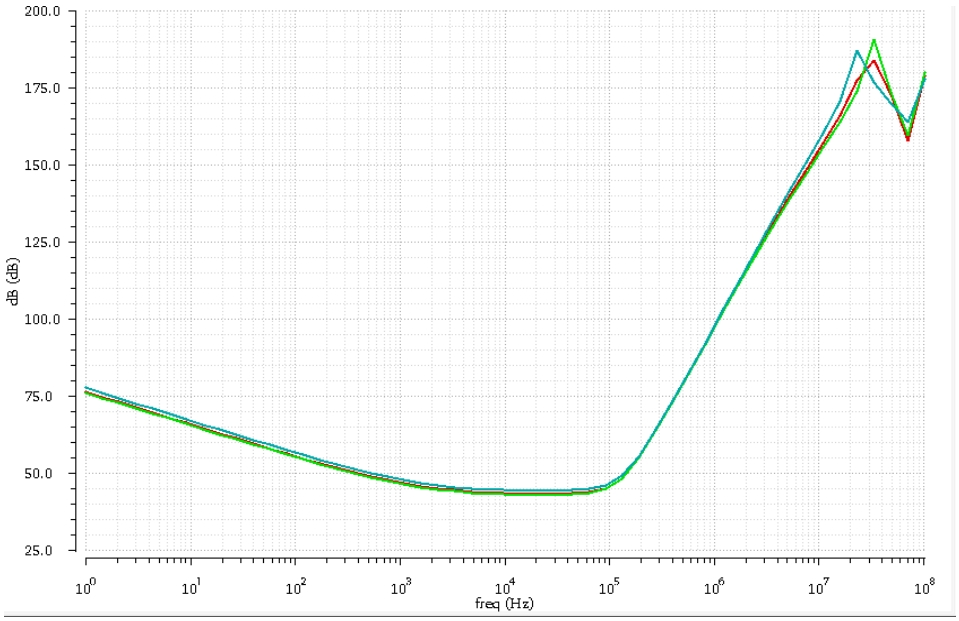


Figure 3.57: Noise Figure of the implemented filter across temperature variations

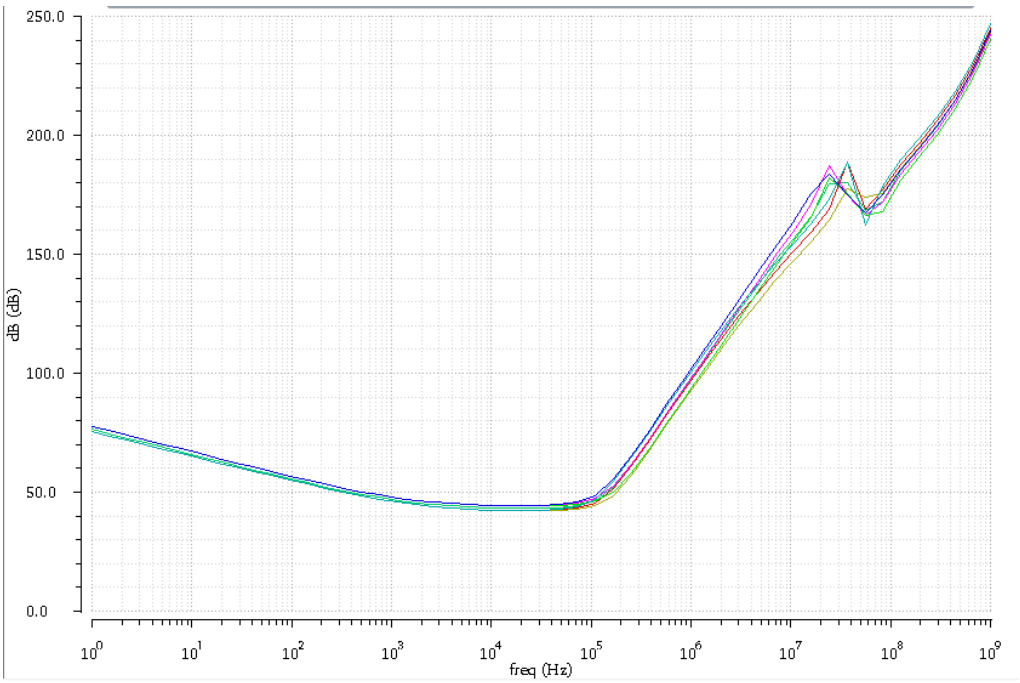


Figure 3.58: Noise Figure of the implemented filter across temperature variations and process variations of the components

The filter was simulated under different process variations as well as different temperatures through corner analysis. The AC response of the filter was affected by the variations especially the variations in the process variations of the passives. Fig.3.59 shows the AC response variations across different corners. Table 3.6 shows the variation in the cut-off frequency and the stop band attenuation at stop band for the main corners.

Table 3.6: Filter variations across corners

Corner	Cut-off frequency	Stop band Attenuation
Temp = 27, typical-typical	116 KHz	49 dB
Temp = 120, typical-typical	116.2 KHz	48 dB
Temp = -40, typical-typical	114.37 KHz	50 dB
Temp = 27, fast-fast	133.3 KHz	43.45 dB
Temp = 120, fast-fast	134 KHz	42.6 dB
Temp = -40, fast-fast	131 KHz	44.33 dB
Temp = 27, slow-slow	101.8 KHz	54 dB
Temp = 120, slow-slow	103 KHz	53.25 dB
Temp = -40, slow-slow	99.2 KHz	54 dB

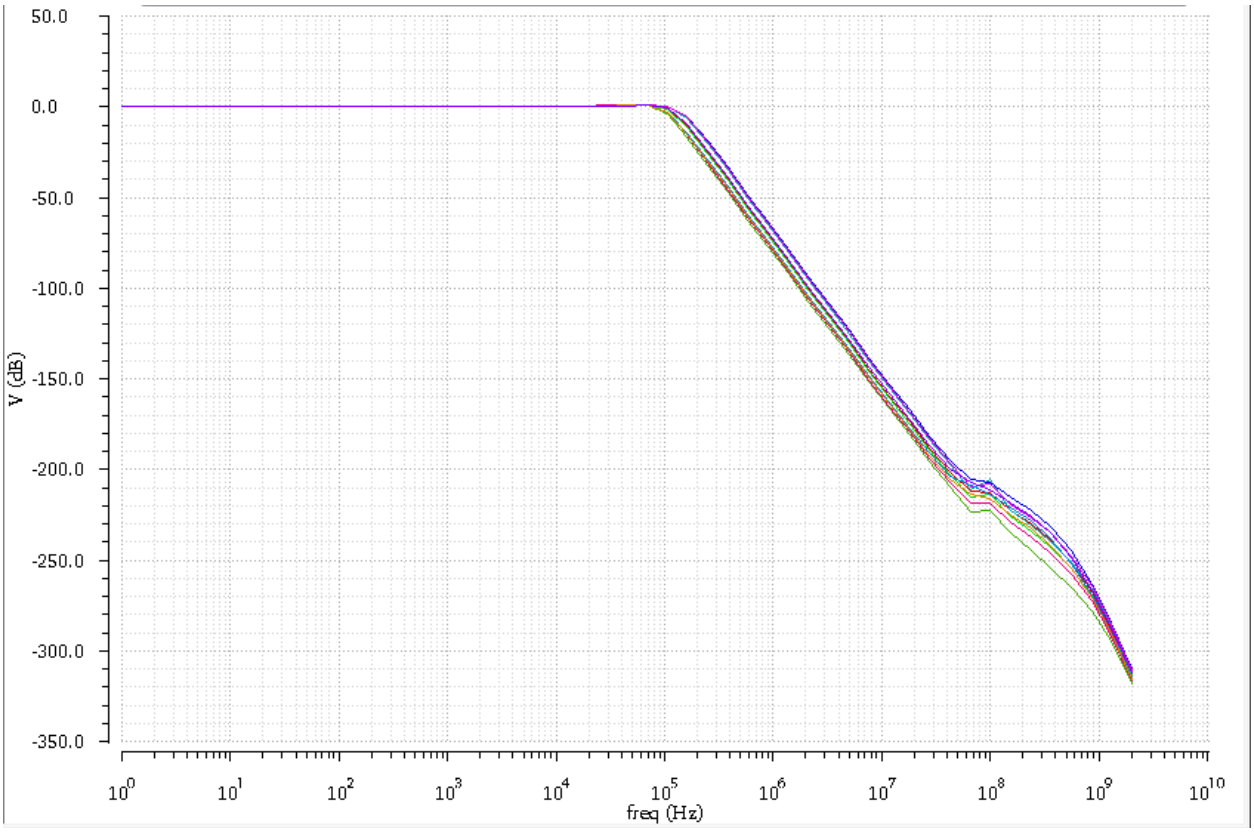


Figure 3.59: Variations in the AC response of the filter across corners

From table 3.6, it is obvious that the bandwidth of the filter increases by increasing the temperature and in the fast-fast corner. The change in the cut-off frequency of the filter is responsible for the change in the attenuation level at stop band. As temperature increases, more current is conducted in the circuit, therefore as current increases, the resistivity of the poly material forming the resistors decreases, leading to higher bandwidth. Same analysis can be done for the fast-fast process variation, because in this variation the threshold voltage decreases, giving chance for more current to be conducted as well.

In order to achieve the system requirements for all cases, the filter needs to be programmable so that it would be able to achieve the required bandwidth and stop band attenuation for all corners. From Fig.3.59 and table 3.6, the variation in the AC response due to change in temperature is small enough to be tolerable, however the variation due to process parameters change is rather large. Therefore, the cases will be classified into 3 groups; one for typical-typical, another one for fast-fast and the last one for small-small.

Programmability is applied on the filter by adding components of different values instead of the previously implemented ones. These new components alongside with the old ones will be connected to the circuit through switches. The control on the switches will be applied from high layers according to the process variation produced from the fab after manufacturing the IC chip.

This control will decide which component will be connected to the circuit and which will not be connected. This programmability technique causes loss in area due to the presence of redundant components; therefore it requires careful implementation in order to change only the smallest possible number of components.

Control is done using 2 bits X0 and X1 as we only have 3 different states. According to the transfer function of the biquad topology used, shown in (3.26), in order to change the cut-off frequency of the biquad, we can change the value of the capacitors or resistors. As this topology has lower number of capacitors than the number of resistors; it was chosen to tune the capacitors.

For the fast-fast corner, the circuit suffers from bandwidth higher than the required which leads to lower attenuation at the stop band. So, to decrease this bandwidth we need to increase the capacitance. Changing C1 capacitor of the first stage from 60 pF to 90 pF was enough to achieve the required performance as shown in Fig.3.60. Therefore, the 2 capacitors (60 pF and 90 pF) are placed and connected to the circuit through two switches controlled by X0.

On the other hand, for the slow-slow corner, the filter suffers from bandwidth lower than the required which leads to undesirable attenuation in the pass band. So, to increase this bandwidth we need to decrease the capacitance. Changing C2 capacitor of the second stage from 11 pF to 7 pF was enough to achieve the required performance as shown in Fig.3.61. Therefore, the 2 capacitors (11 pF and 7 pF) are placed and connected to the circuit through two switches controlled by X1. Fig.3.62 shows the performance of the filter at all main corners after programmability. It shows outstanding uniformity in the performance for all cases.

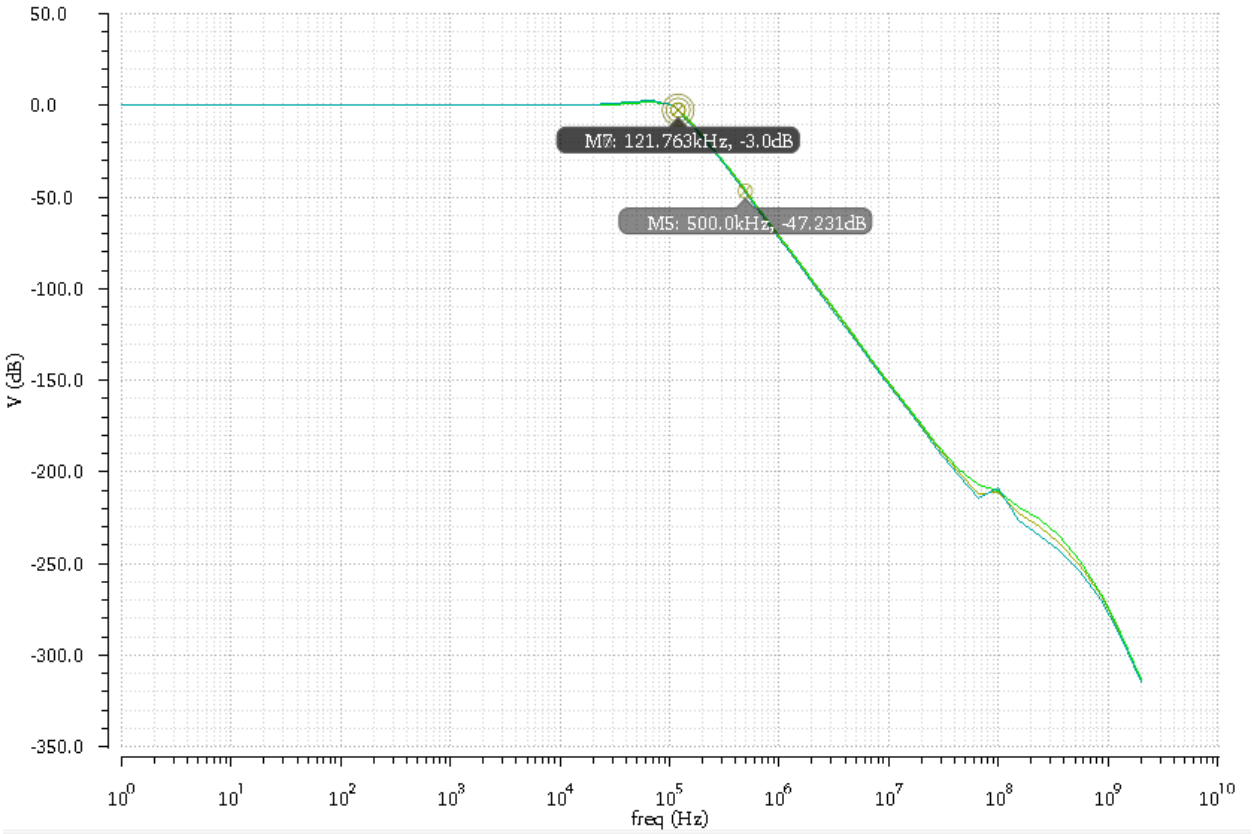


Figure 3.60: Programmable Filter performance at the fast-fast corner

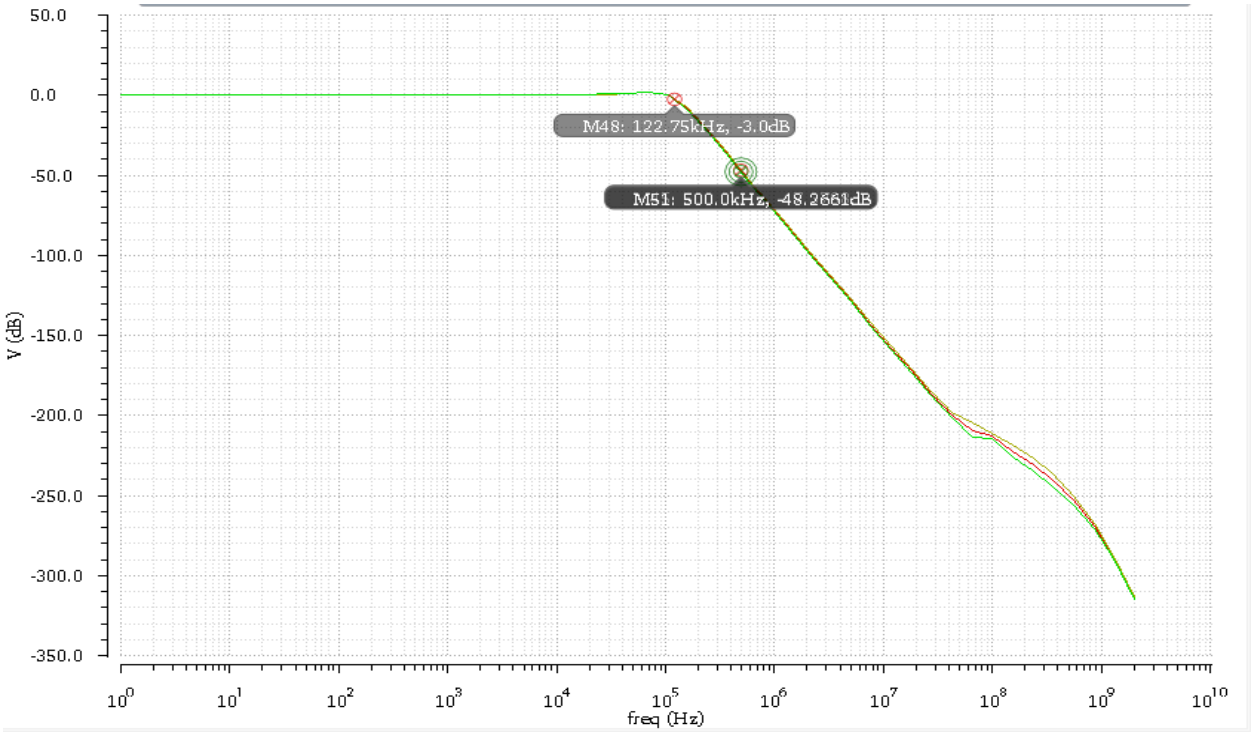


Figure 3.61: Programmable Filter performance at the slow-slow corner

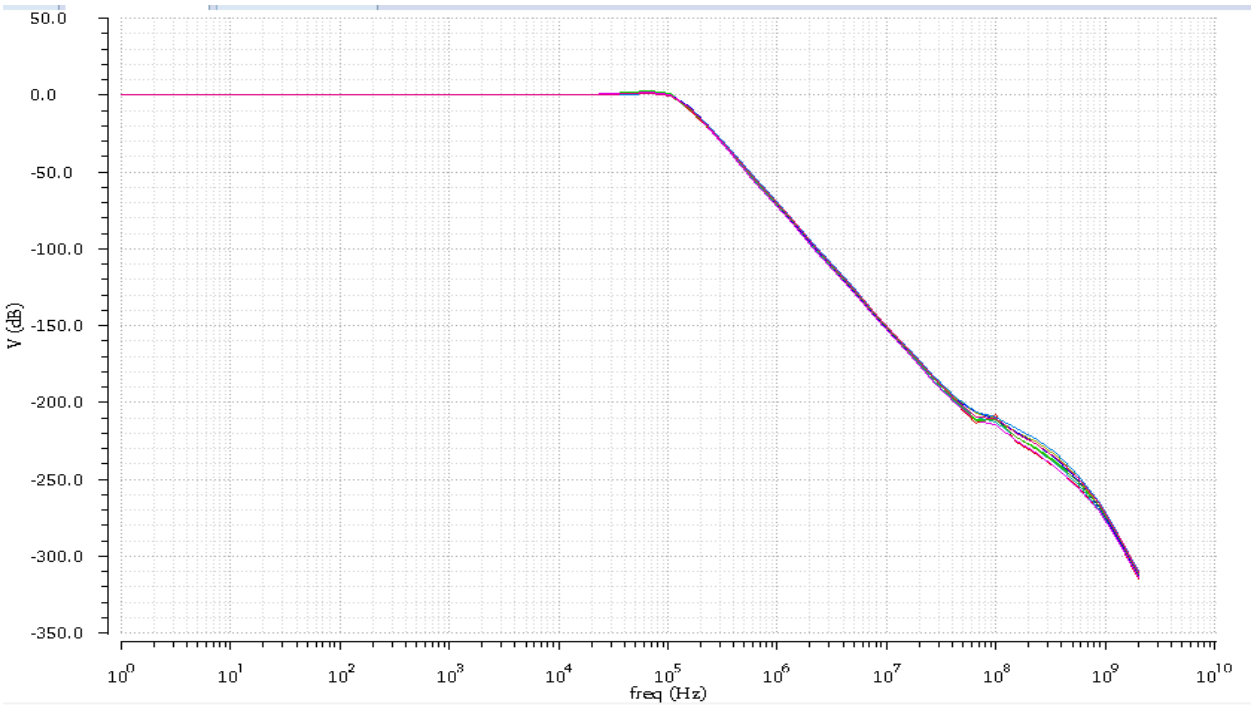


Figure 3.62: Filter performance at all corners after programmability

Concerning the linearity of the filter, P1dB compression point was calculated and it was high enough to pass the full scale output signal of the DAC without any compression. The P1dB point of the filter, as shown in Fig.3.63, is equal to 8.57 dBm which corresponds to input peak voltage of 850 mV which is much greater than the full scale output of the DAC (200 mV), so the filter achieves great linearity performance.

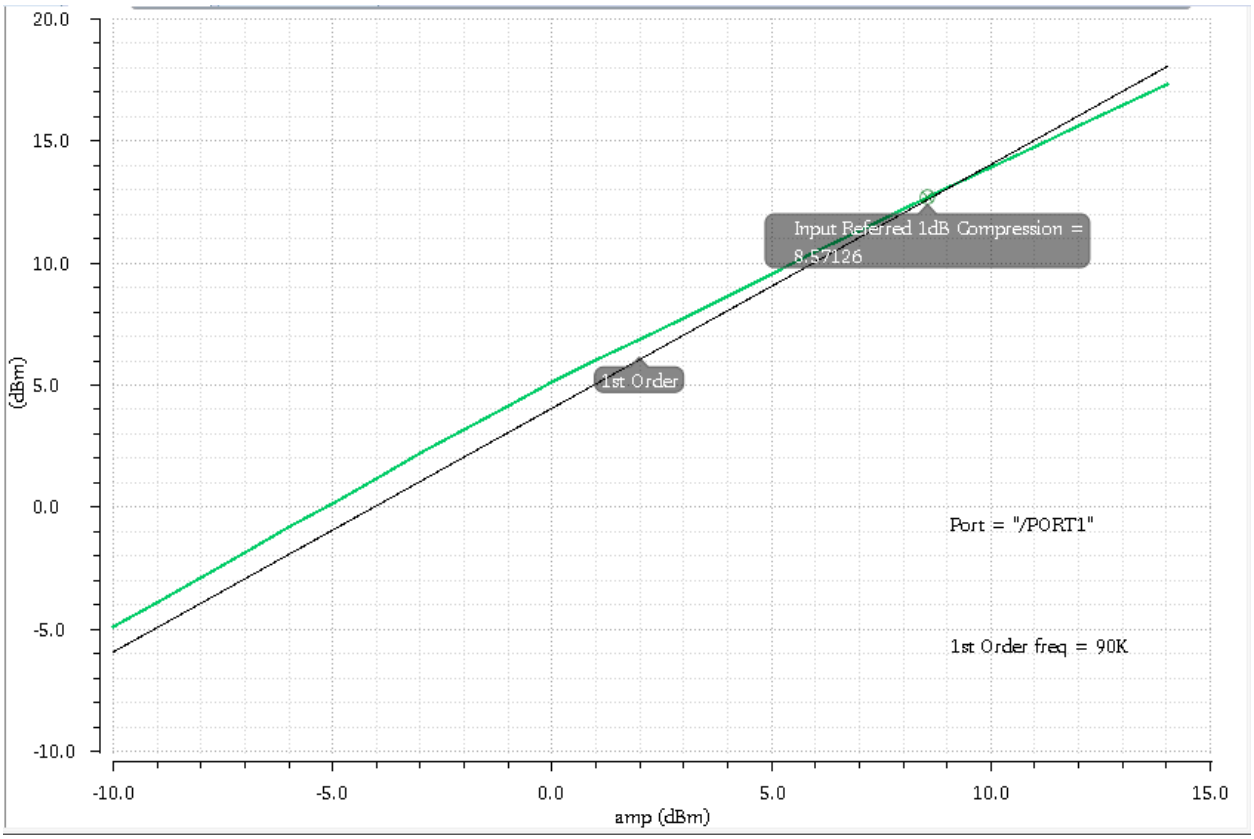


Figure 3.63: Reconstruction Filter P1dB compression point using pss simulation

3.10 Full DAC System Integration Results

Once the DAC was completed, a test bench was created which can be seen in Fig.3.64. To test the dynamic performance of the DAC, an ADC was required. Instead of constructing a 10 bit ADC, an ideal circuit was used from the Cadence ahdlLib. This library contains ideal circuits built using Verilog-A. The model present in the library was for an 8 bit ADC, so its Verilog-A code was modified and extended to be 10 bit ADC. The modified Verilog-A code is placed in Appendix 1.

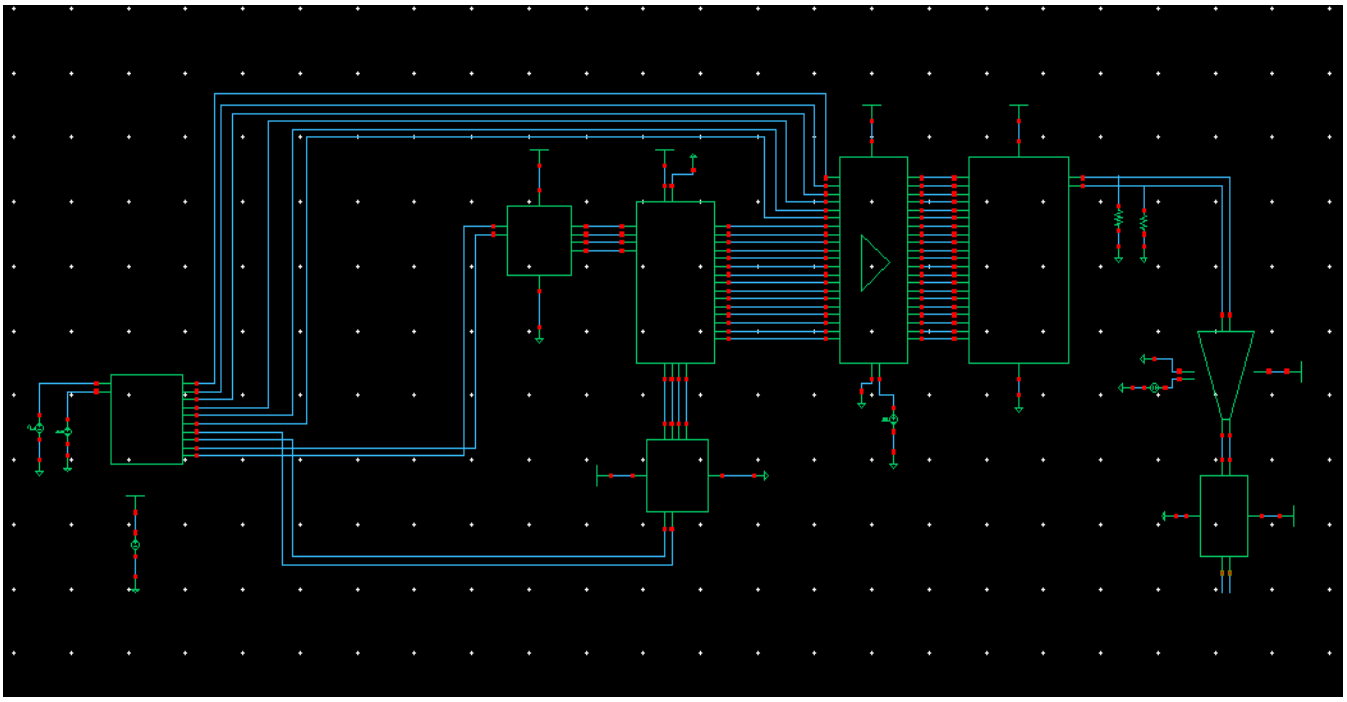


Figure 3.64: Full System Integration for Dynamic performance test

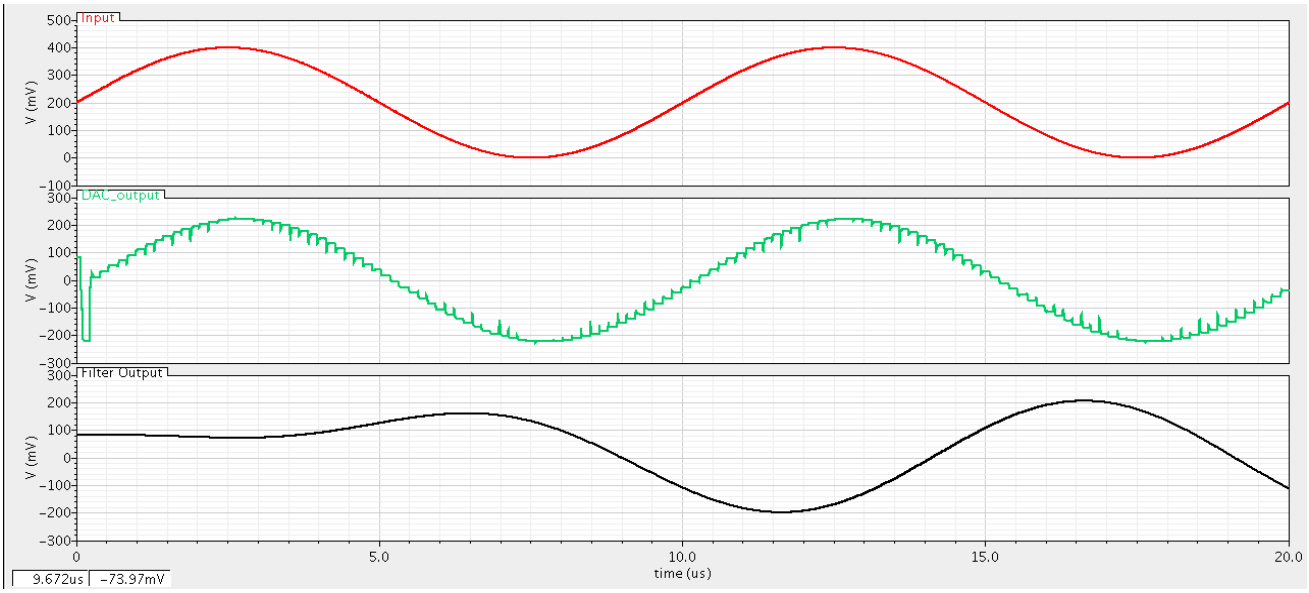


Figure 3.65: Transient response of the DAC system

Fig.3.65 shows the transient response of the DAC integrated system. The input signal to the ideal ADC is plotted along with the DAC output and the reconstruction filter output. It can be seen that the output signal of the DAC suffers from glitches due to its stair-case nature, but the reconstruction filter was able to smooth the signal completely and get rid of the glitches. Also,

the delay of 4 µs due to the filter settling time is clear, this high settling time for the filter is due to its very narrow bandwidth (only 120 KHz).
In order to test the linearity of the DAC and obtain its SFDR; periodic steady state (pss) analysis was applied on the whole DAC system with beat frequency of 100 KHz. The DAC achieved SFDR of 75 dB and the highest harmonic was the third harmonic as shown in Fig. 3.66. Also the total harmonic distortion (THD) was calculated to be -74.5 dB as can be estimated from the result in Fig.3.67.

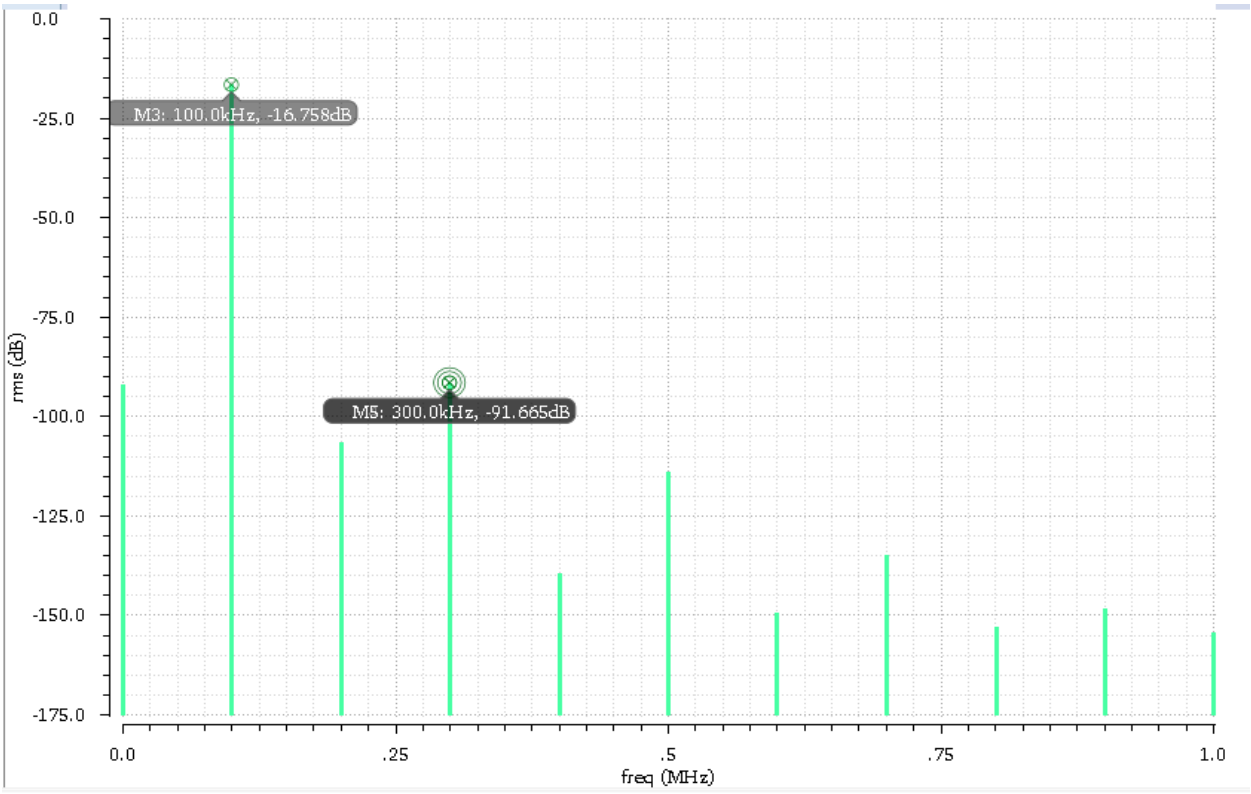


Figure 3.66: DAC system output harmonics

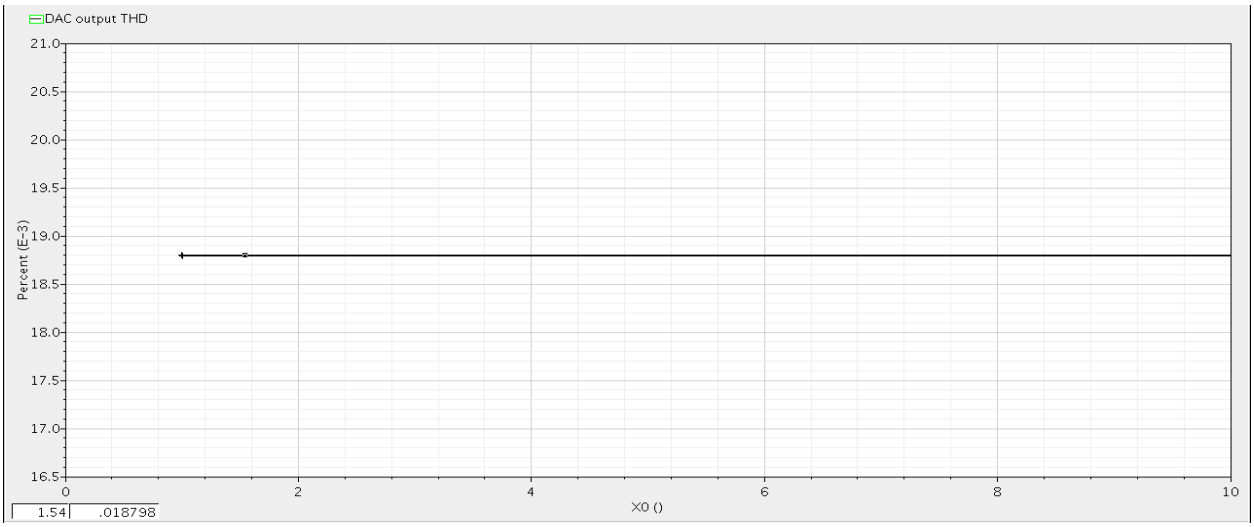


Figure 3.67: DAC system output THD

The output noise level after the DAC was simulated and it showed a noise level of -153 dBm/Hz at offset 500 KHz. A peak appears in the noise response at 100 KHz frequency, as shown in Fig.3.68, this is because this point represents the end of the pass band and at higher frequencies; the filter will start to dominate the response.

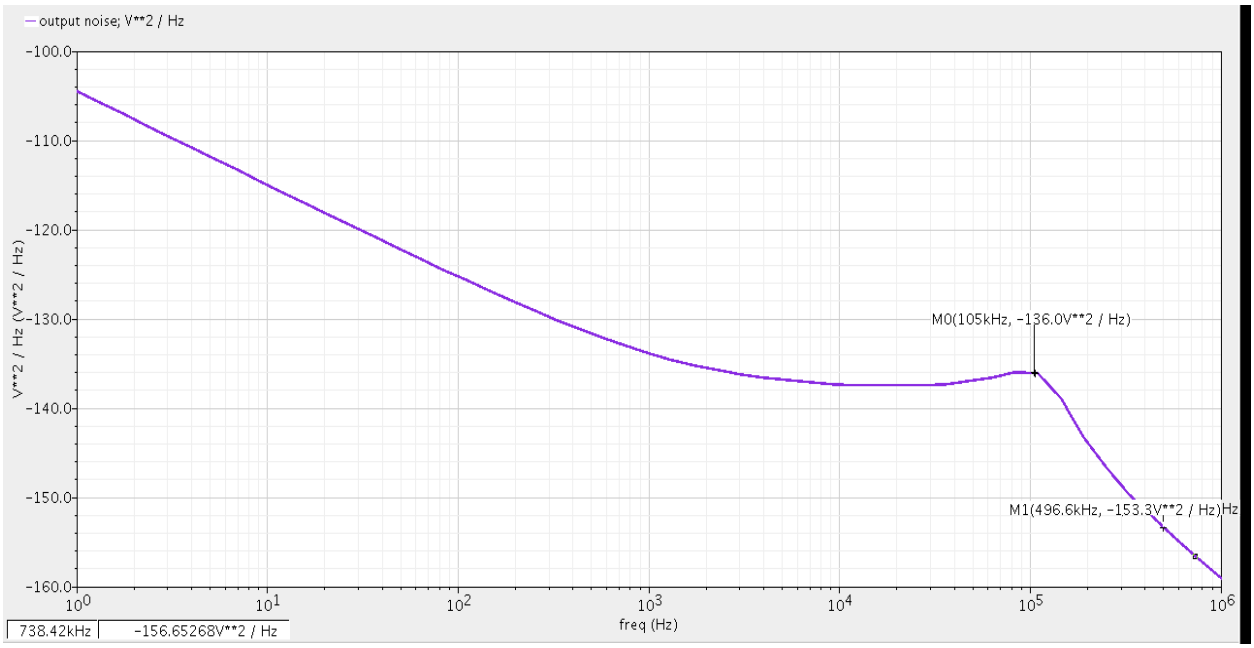


Figure 3.68: DAC system output noise

The output DAC integrated noise in the region from DC to 3.2 MHz (fs/2) was calculated and it was equal to $4.96* 10^{-9}\ V^2/Hz$, as shown in Fig.3.69 and this corresponds to SNR = 10 log $(\frac{0.2^2}{4.96* 10^{-9}}) = 69\ dB$.

Device	Param	Noise Contribution	% Of Total
I33.R26.r4	thermal_noise	2.73408e-10	5.51
I33.R26.r3	thermal_noise	2.73366e-10	5.51
I33.R26.r2	thermal_noise	2.73324e-10	5.50
I33.R26.r1	thermal_noise	2.73282e-10	5.50
I33.R25.r4	thermal_noise	2.71604e-10	5.47
I33.R25.r3	thermal_noise	2.71562e-10	5.47
I33.R25.r2	thermal_noise	2.7152e-10	5.47
I33.R25.r1	thermal_noise	2.71479e-10	5.47
I33.R13.r4	thermal_noise	1.39641e-10	2.81
I33.R13.r3	thermal_noise	1.39635e-10	2.81
I33.R13.r2	thermal_noise	1.3963e-10	2.81
I33.R13.r1	thermal_noise	1.39624e-10	2.81
I33.R8.r4	thermal_noise	1.39084e-10	2.80
I33.R8.r3	thermal_noise	1.39079e-10	2.80
I33.R8.r2	thermal_noise	1.39073e-10	2.80

Integrated Noise Summary (in V^2) Sorted By Noise Contributors
Total Summarized Noise = 4.9657e-09
No input referred noise available
The above noise summary info is for noise data

Figure 3.69: DAC system output integrated noise summary

The rise time of the converter was measured to see how fast the output transitions. A step input was given to the DAC so that the output would change from zero to the full scale range. The measurement was taken when the output settled within 1 LSB of the final step response. The rise time of the converter was measured to be 217 ns as shown in Fig.3.70. It's worth to be mentioned that the test was applied on the output node of the DAC buffer before the filter.

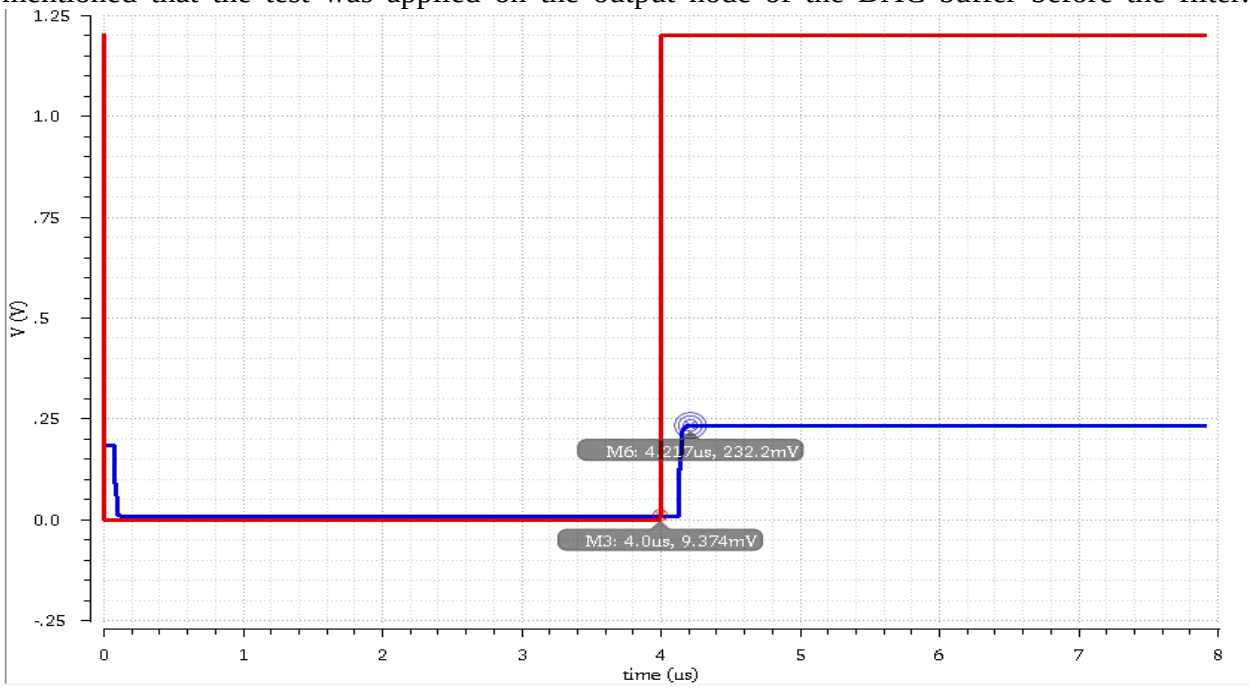


Figure 3.70: DAC rise time simulation

After examining the dynamic performance of the DAC, we move to examining its static performance. The setup for testing the static linearity properties, depicted in Fig.3.71, shows that the input bits are simulated as pulse voltage sources. By making the frequencies of these sources double of each other, we can make these 10 sources simulate all the possible codes for the DAC,

starting from all zeros case to the full scale. The resulting stair-case voltage output of the DAC is depicted in Fig. 3.72

The differential non linearity and the integral non linearity were calculated for all codes by exporting the transient results of the DAC output from Cadence Spectre to Matlab. A Matlab script, found in Appendix 2, was written in order to measure the DNL and INL of the DAC based on their definition stated in section 3.2. The resulting maximum DNL of the converter was measured to be 0.16 LSB at the 608^{th} code. This code corresponds to 1000110000, at this code the 2 highest weighted branches of the current steering DAC core LS segment are on along with 8 branches from the MS segment; therefore it's considered the highest step in the full scale stair case. Considering the INL, its greatest value was 0.96 LSB just slightly below the maximum limit of the INL for proper monotonic DAC performance.

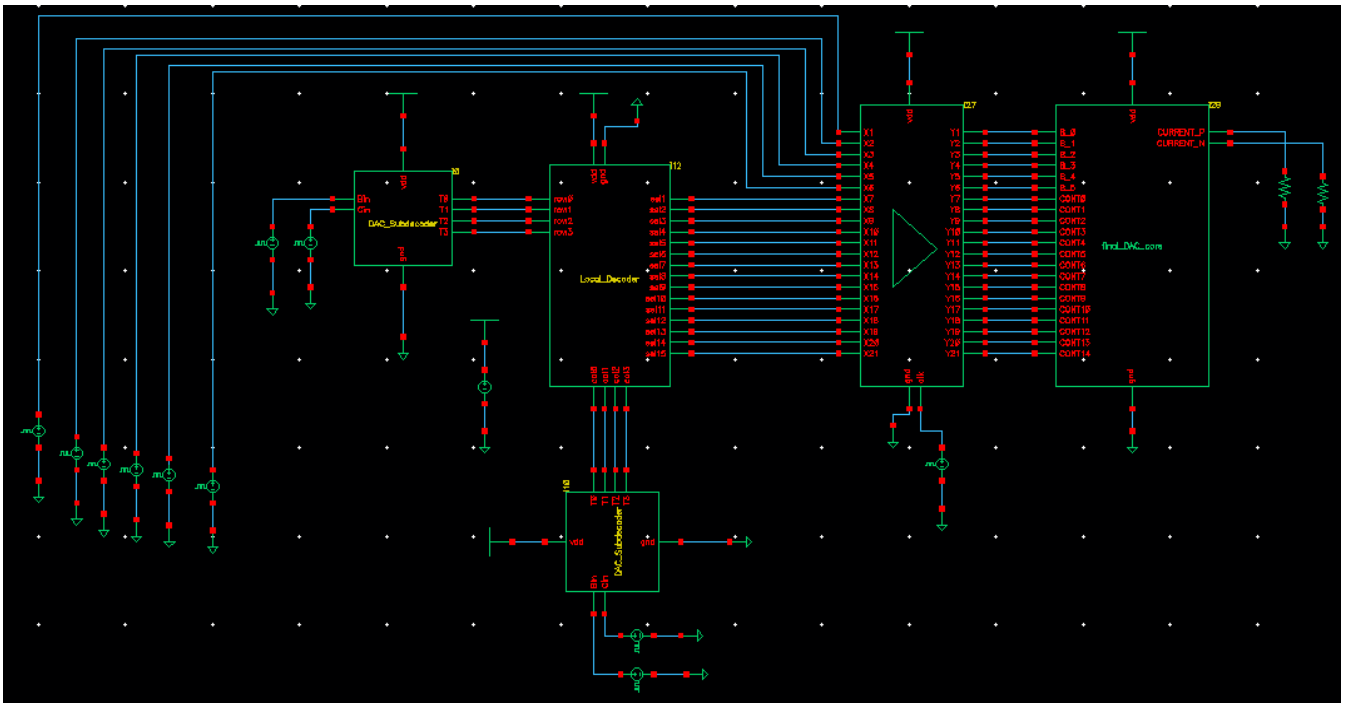


Figure 3.71: Full System Integration for Static performance test

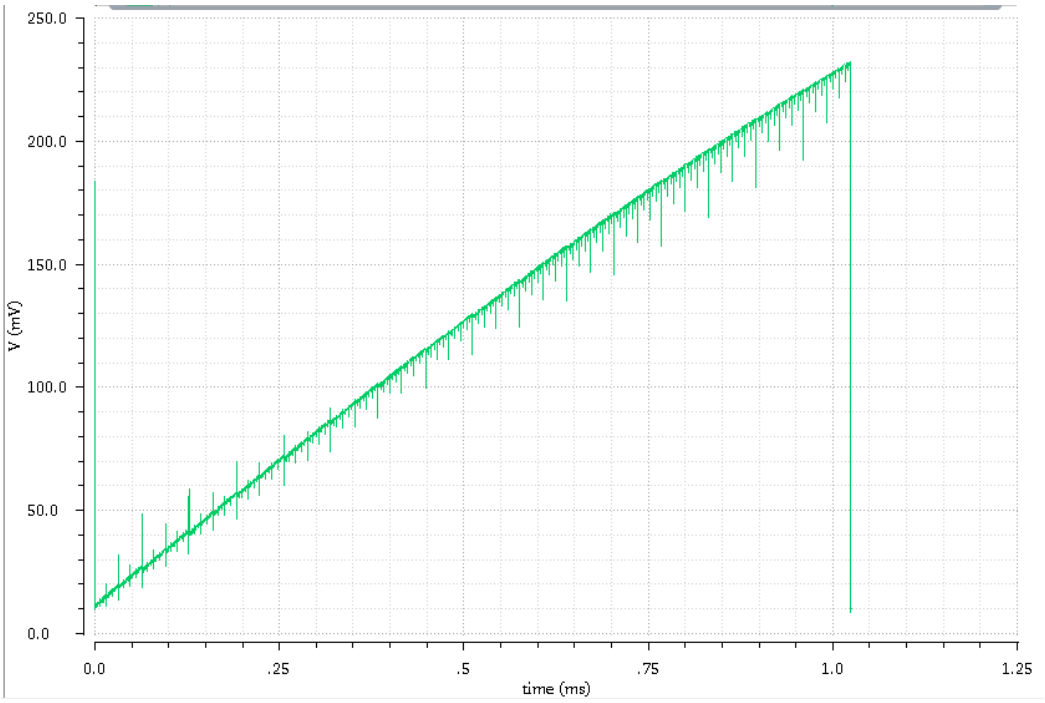


Figure 3.72: Output DAC staircase voltage

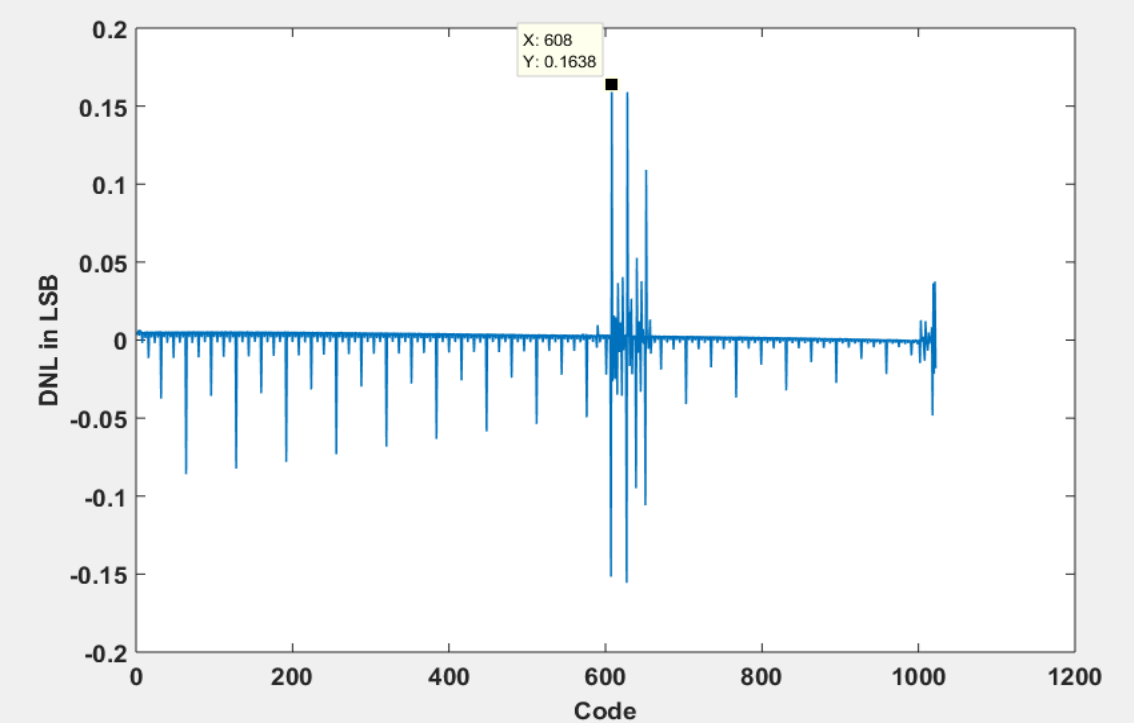


Figure 3.73: DNL of the DAC system for all codes

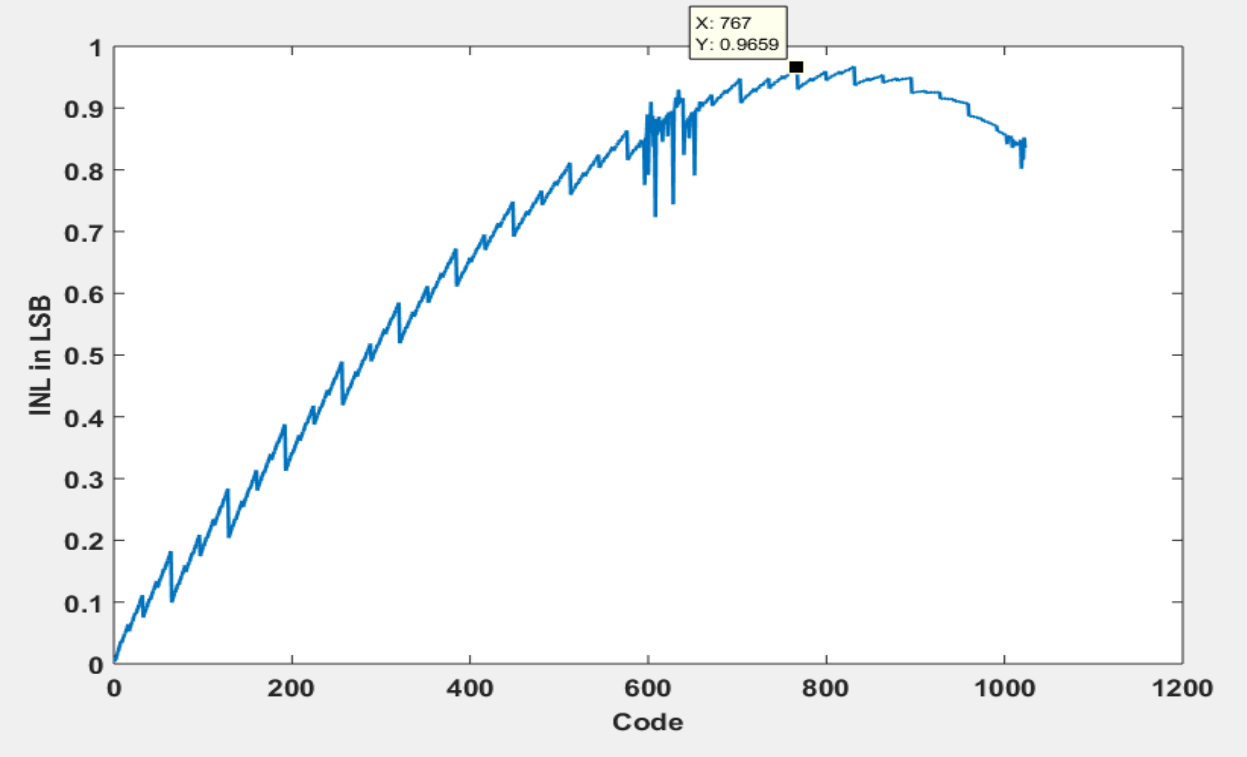


Figure 3.74: INL of the DAC system for all codes

Offset error of the DAC can be easily noticed from the staircase output and is equal to 8.5 mV. However, to calculate the gain error, a linear line having the same start and end points of the DAC staircase was plotted and the gain of two lines was calculated in the middle region by extracting two points from each line. The gain error was calculated to be $15\mu V/LSB$. Fig.3.75 clarifies the calculation of the gain error.

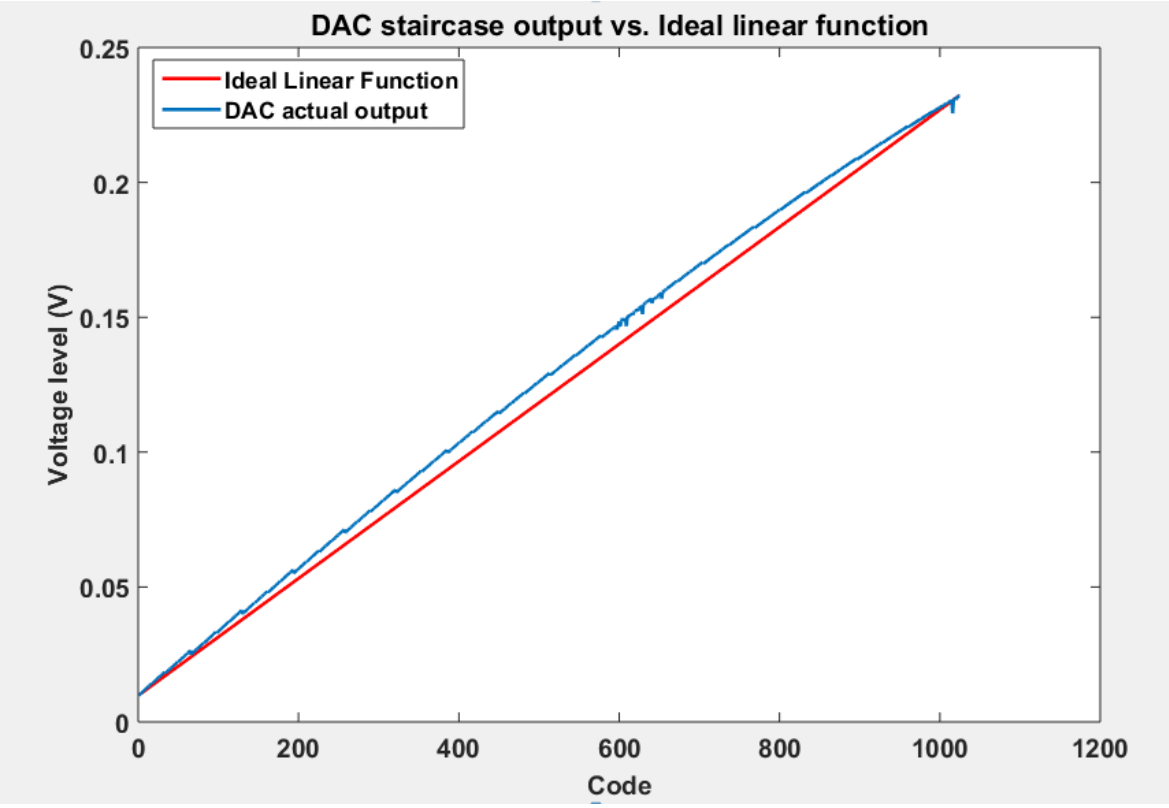


Figure 3.75: Gain error of the DAC system

Finally, we reach to the power dissipation of the DAC system. It can be categorized into the digital sub-circuits, and the analog circuits. The analog power dissipation is based on the current drawn through the load resistor by the converter. Since the full scale output current of the DAC is designed to be 1024 µA, then the power consumed by the DAC current cells is 1.228 mW.
In addition to this, the filter and the DAC output buffer also consumes an amount of current calculated to be 2.9 mA which corresponds to 3.48 mW.

For the digital sub-circuits, the average power dissipation for the thermometer decoder is 83.5 µW, and the local decoder consumes 34 µW. Each latch needs 73.9 µW and as we have 21 latches, total power needed by these latches is 1.55 mW. The switch drivers also sink total amount of current of 80 µA each; so all the switch drivers consume 2mW.

By adding all these power values, it can be found that the full DAC system consumes 8.5 mW.

3.11 NB-IoT Transmitter Envelope path

For the designed transmitter envelope path, the system level design showed the need for a 10 bit DAC similar to that implemented for the phase path. The only difference is the need of a reconstruction filter with different specifications as well as a DC-to DC converter in order to convert the envelope signal generated from the DAC to a higher voltage level to act as a variable supply for the power amplifier to restore the amplitude information of the signal. The DC to DC converter was implemented as an ideal block, the DAC used for the envelope path is exactly the same as the one implemented in previous sections; therefore we will focus on th envelope pathe filter requirements and implementation in this section.

According to the work done in [16], it was found that the bandwidth of the envelope signal for a SC-FDMA QPSK modulated signal can reach 6 times the bandwidth of the phase signal; therefore the required bandwidth of the envelope filter was estimated to be 600 KHz. The stop band attenuation was defined to be 50 dB at 3.2 MHz as this attenuation is enough to get rid of the effect of aliases.

By moving through similar procedure as that explained in section 3.9, it was found that these specifications can be met by designing a 4^{th} order butter worth low pass filter with multiple feedback topology. Table 3.7 show the values of the resistors and capacitors used in this implementation.

Table 3.7: Envelope Filter Components' Values

	R1	R2	R3	C1	C2
First Stage	5.5 KΩ	2.8 KΩ	5.5 KΩ	75 pF	11 pF
Second Stage	14 KΩ	7 KΩ	14 KΩ	12 pF	11 pF

The performance of the filter is depicted in the Figures 3.76 to 3.79. The filter achieved stop band attenuation of -51 dB at 3.2 MHz with total settling time of 0.75 μs. Total noise Figure generated from the filter is 37.6 dB and the P1dB compression point was calculated to be 8.742 dBm

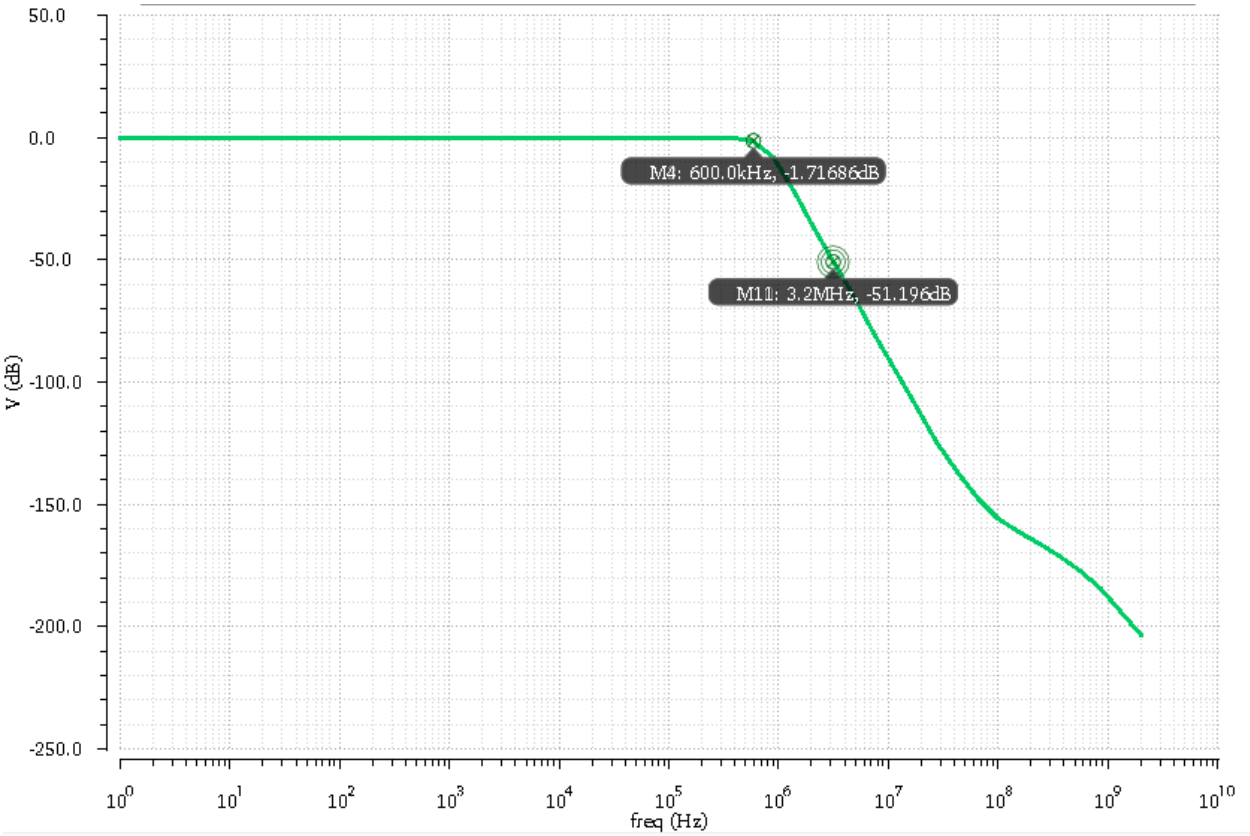


Figure 3.76: Envelope Filter AC Response

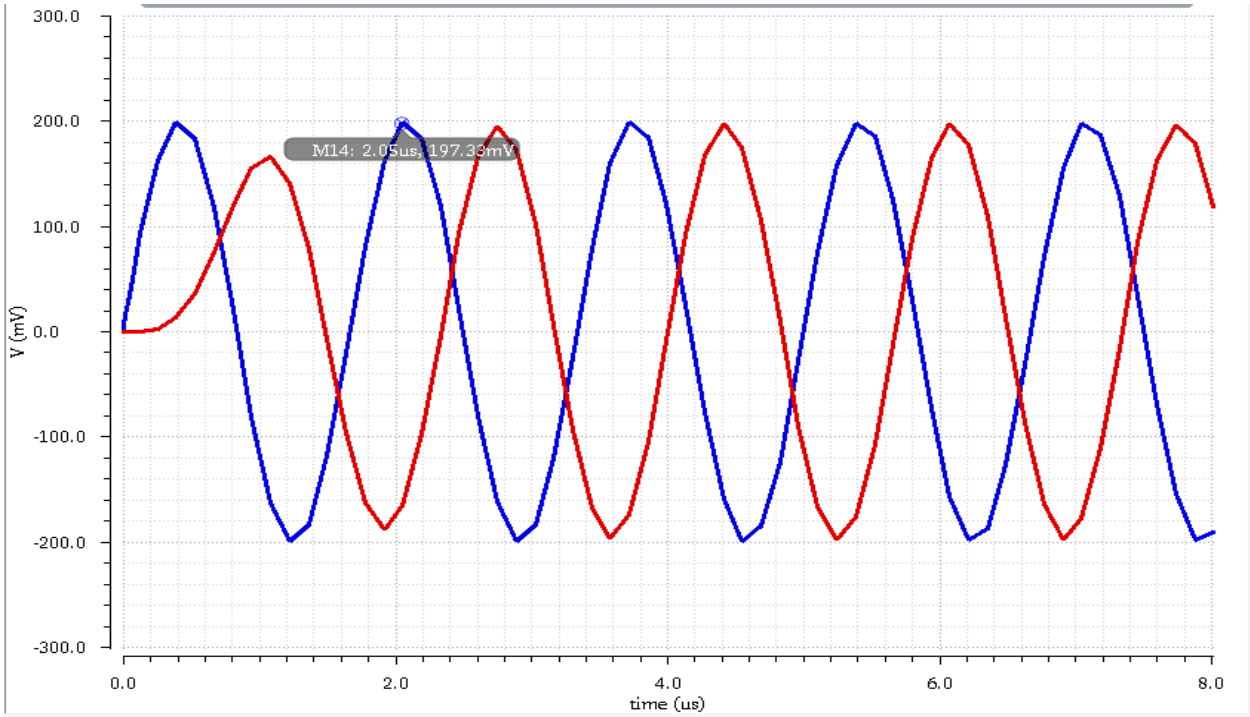


Figure 3.77: Envelope Filter transient Response

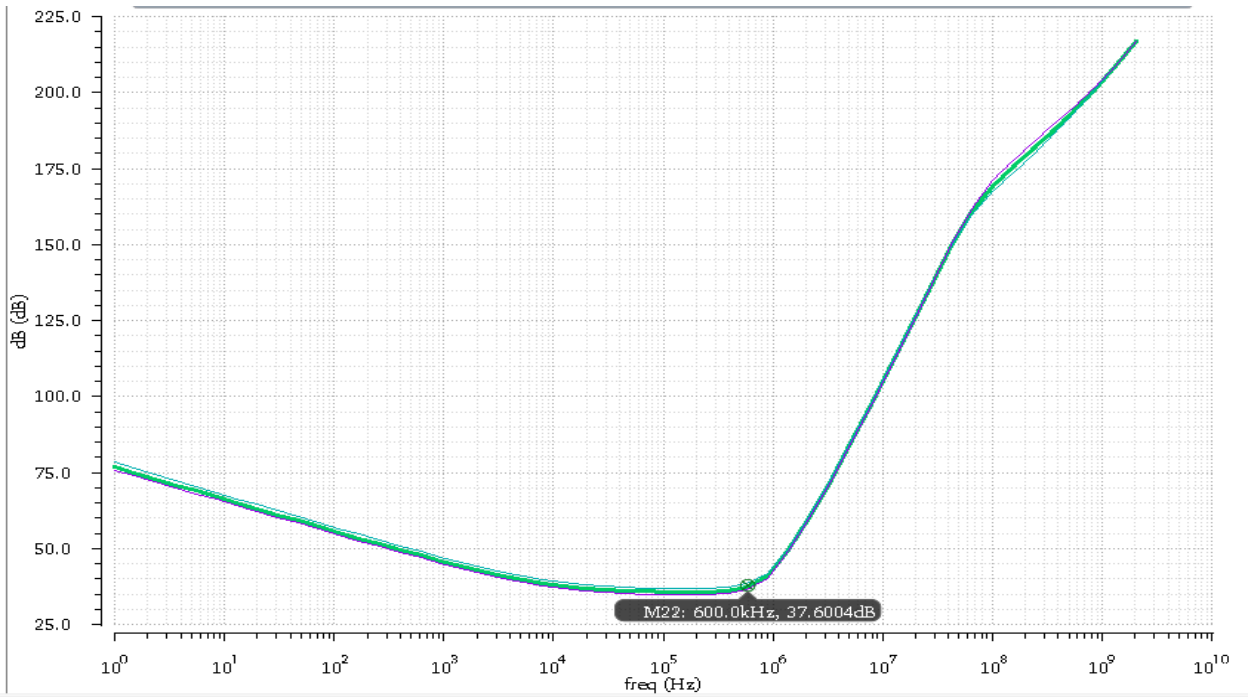


Figure 3.78: Envelope Filter Noise Figure through temperature variations

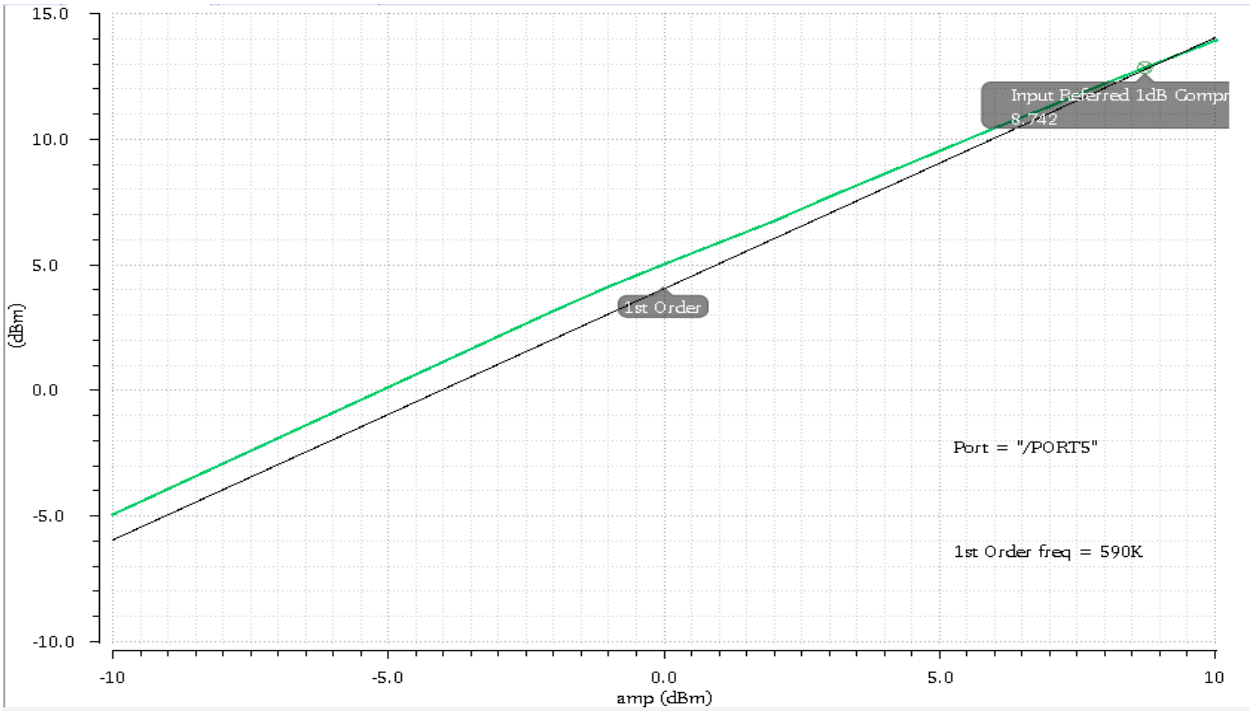


Figure 3.79: Envelope Filter P1dB compression point

Comparing the performance of the envelope filter with the phase filter shows that increasing the bandwidth of the reconstruction filter leads to a significant decrease in its settling time as highlighted in Fig.3.80. Also the increase in the bandwidth lead to the usage of lower resistors in the Op amp RC architecture of filters; this decrease in resistor values lead to lower noise figure.

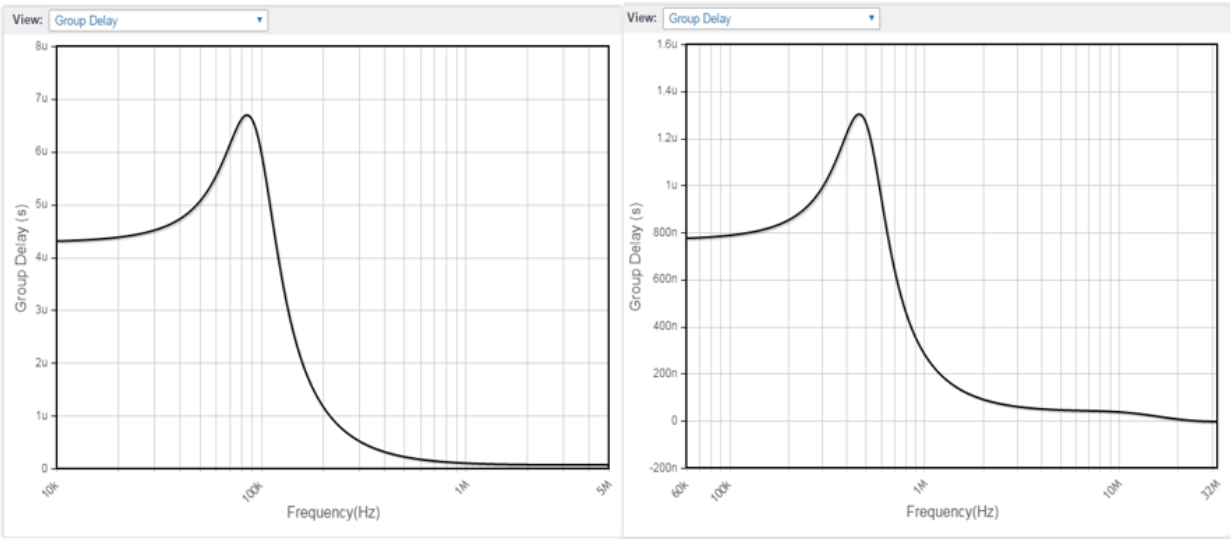


Figure 3.80: Group delay of both the reconstruction filters of the phase and the envelope

3.12 Conclusion

A 10 bit current steering DAC was designed to act as first stage of a NB-IoT transmitter. The DAC was implemented with segmentation ratio of 40% and was able to achieve worst case DNL of 0.164 LSB and worst case INL of 0.97 LSB. The dynamic performance of the DAC was tested and it achieved SFDR of 75 dB and SNR equal to 69 dB corresponding to the output noise level of -153 dBm/Hz. The DAC needs at most 217 ns in order to settle at its final value. The DAC was followed by a reconstruction filter in order to smooth the output signal and achieve the required noise level at offset 500 KHz which represents the PLL bandwidth. The reconstruction filter was implemented as 4th order butter worth low pass filter; it was designed based on the op amp RC topology. The filter achieved band stop attenuation of 50 dB at 500 KHz with noise figure of 45 dB and P1dB compression point of 8.57 dBm. The group delay of the filter at 100 KHz was measured to be 4.2 µs. The overall DAC system consumed less than 8.5 mW for sampling frequency of 6.4 MS/sec.

References

[1] Samiran Halder, Hans Gustat, Christoph Scheytt, and Andreas Thiede. A 20GS/s 8-bit current steering dac in 0.25_m SiGe BiCMOS technology. In Microwave Integrated Circuit Conference, 2008. EuMIC 2008. European, pages 147–150. IEEE, 2008.

[2] Ola Andersson. Modeling and implementation of current-steering digital-to-analog converters. 2005.

[3] Rudy J Van de Plassche. CMOS integrated analog-to-digital and digital-to-analog converters, volume 2. Kluwer Academic Publishers Dordrecht, 2003.

[4] R Jacob Baker. CMOS: circuit design, layout, and simulation, volume 18. JohnWiley & Sons, 2011.

[5] I Myderrizi and A Zeki. Current-steering digital-to-analog converters: functional specifications, design basics, and behavioral modeling. Antennas and Propagation Magazine, IEEE, 52(4):197–208, 2010.

[6] Mikael Gustavsson, J Jacob Wikner, and Nianxiong Tan. CMOS data converters for communications. Springer Science & Business Media, 2000.

[7] C. Toumazou, J. B. Hughes, and N. C. Battersby, *Switched-Currents: an Analogue Technique for Digital Technology,* Peter Peregrinus, Stevenage, UK, 1993.

[8] Walter Allan Kester. Data conversion handbook. Newnes, 2005.

[9] Lin, Yang. *Design of an 8-bit CMOS Unit-Element Current-Steering Digital-to-Analog Converter*. Diss. Thesis of University of Maine, 2008.

[10] Chi-Hung Lin and Klaas Bult. A 10-b, 500-Msample/s cmos dac in 0.6 mm 2. Solid-State Circuits, IEEE Journal of, 33(12):1948–1958, 1998.

[11] T Zaki, T Hassan, and C Scherjon. A compact 12-bit current-steering d/a converter for hdrc image sensors. In Annual ProRISCWorkshop on Circuits, Systems and Signal Processing, Vendhoven, 2010.

[12] Cheng, Kuo-Hsing, Tsung-Shen Chen, and Ching-Wen Kuo. "High accuracy current mirror with low settling time." *Circuits and Systems, 2003 IEEE 46th Midwest Symposium on*. Vol. 1. IEEE, 2003.

[13] Paolo Carbone, Sayfe Kiaei, and Fang Xu. Design, Modeling and Testing of Data Converters. Springer, 2014.

[14] Konstantinos Doris, Arthur HM van Roermund, Domine MW Leenaerts, Domine MW Leenaerts, and Domine MW Leenaerts. Wide-bandwidth high dynamic range D/A converters. Springer, 2006.

[15] Fidler, J. K., T. Deliyannis, and Yichuang Sun. *Continuous-time active filter design*. CRC press, 1998.

[16] Stauth, Jason Thaine, and Seth R. Sanders. *Energy Efficient Wireless Transmitters: Polar and Direct-Digital Modulation Architectures*. Diss. University of California, Berkeley, 2008.

[15] Fidler, J. K., T. Deliyannis, and Yichuang Sun. *Continuous-time active filter design*. CRC press, 1998.

[16] Stauth, Jason Thaine, and Seth R. Sanders. *Energy Efficient Wireless Transmitters: Polar and Direct-Digital Modulation Architectures*. Diss. University of California, Berkeley, 2008.

Chapter 4 : Mixer

Frequency translation is a substantial operation in wireless transceivers. It can be either up-conversion for transmitters or down-conversion for receivers. Great efforts were exerted to come up with variety of implementations for mixers in both active and passive forms. Each of These variants will favor some performance parameters over another and the choice of which circuit to use will be decided based on the need of the designer. In order to judge the performance of mixers- either passive or active - , we have to know their performance parameters [1]. Briefly, they are linearity, conversion gain, port-to-port feedthrough and noise figure. It is also important to mention the concept of single and double balanced mixers and their pros and cons. coming till here, it's time to talk about these parameters in more details.

4.1. Performance Parameters

Conversion gain: What makes Conversion gain unique from the gain of amplifiers is that Conversion gain is defined between two signals that are not at the same frequency due to frequency translation. In addition, it may be a voltage or power conversion gain. Strictly speaking, voltage conversion gain is defined as the ratio of the RMS voltage at the IF port (Output) to the RMS voltage at the RF port (Input). The previous definition is assuming a down-conversion mixer but it can be also applied to the case of up-conversion mixer without loss of generality.

Port-to-port feedthrough: Due to the inevitable intrinsic capacitances of the device, we will suffer from unwanted signal leakage between ports as shown in Fig. 4.1. The severity of port-to-port feedthrough will vary according to the transceiver architecture. For example as shown in Fig. 4.2, LO-RF feedthrough causes DC offsets at the IF port and in case of direct conversion, it may cause Radiation from antenna. RF-LO feedthrough in Direct conversion receivers may cause LO injection pulling which will corrupt the LO spectrum [1]. That is why we use LO buffers. On the other hand, in transmitters we care more about LO-RF feedthrough because it would corrupt the transmitted signal constellation as shown in Fig. 4.3, which will raise EVM.

Noise figure: Generally, Noise figure is used to know how much degradation in SNR. Due to frequency translation in mixers, mixers' noise figure has two definition to account for noise folding in image bands. As shown in Fig. 4.4, Single side band noise figure is used when the desired signal is on one of the sides of ω_{LO} and not both. As shown in figure(3-b) Double side band noise figure is used when signal is on both sides of ω_{LO} (Which is the case in Direct conversion receivers). For noiseless mixers, DSB noise figure is 0 dB while SSB noise figure is 3 dB due to the extra noise component folded back as shown in the Fig 4.5.Thus, we can say that SSB noise figure is more than DSB noise figure by 3 dB.

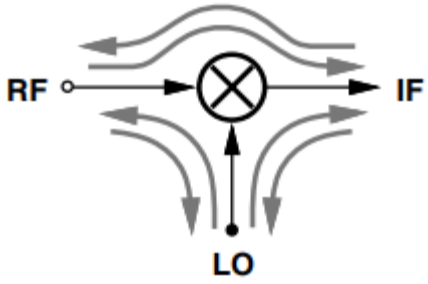


Figure 4.1: Feedthrough across mixer's ports

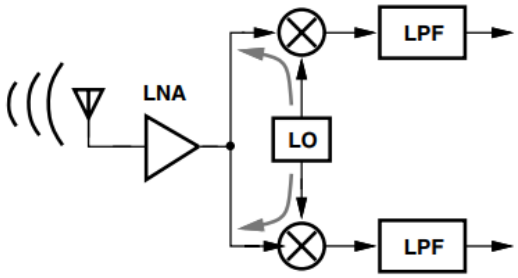


Figure 4.2: LO-RF feedthrough in Homodyne receivers

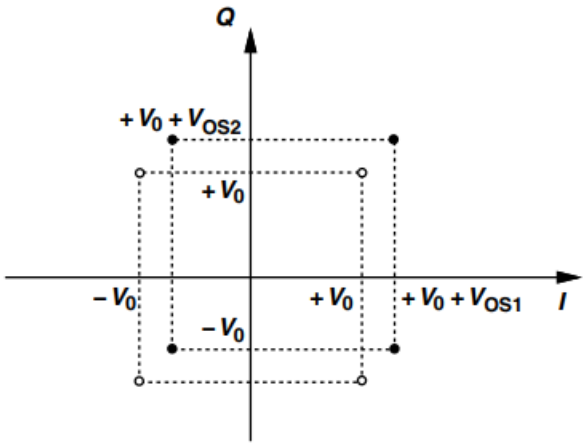


Figure 4.3: DC offset in transmitted constellation due to LO-RF feedthrough

Linearity: In mixers, linearity and noise cannot be treated independently. For convenience, we will investigate both cases of receiver and transmitter. In receivers, we seek to achieve linearity specification without raising NF. Moreover, we may sacrifice NF to increase linearity and compensate this by raising gain of the LNA to make noise of the subsequent stages negligible. On the other hand, in transmitters we favor linearity over noise as we deal with large signal levels. Up-conversion mixer linearity is dependent on the baseband signal swing and type of modulation. 1db compression point is used to quantify how linear is the system as shown in Fig. 4.6.

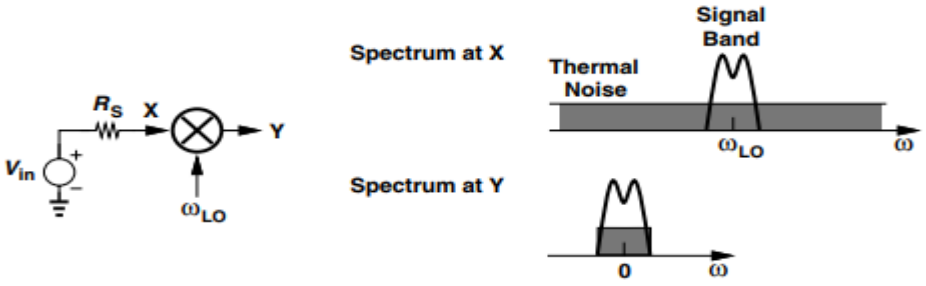


Figure 4.4: Single side band noise figure

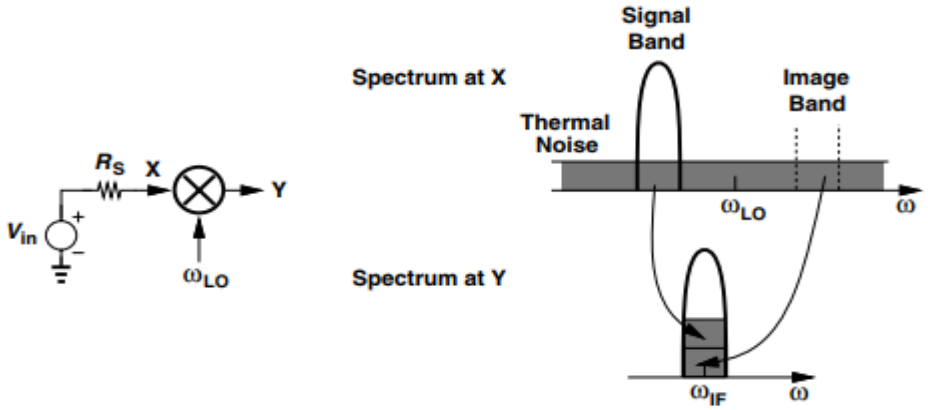


Figure 4.5: Double side band noise figure

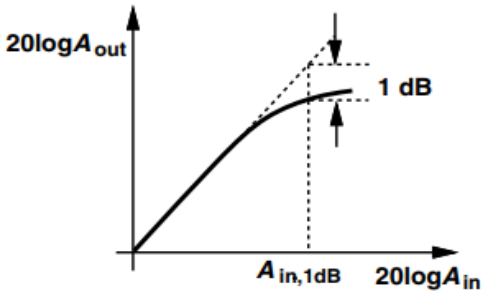


Figure 4.6: 1 dB compression point

4.2. Mixer's Classification

Mixers can be classified into active and passive. Another classification is to categorize them as unbalanced, single balanced or double balanced. An unbalanced mixer is that with both input and LO signal in single ended fashion. The distinction between single and double balanced will be discussed in a later section. It is worth mentioning that the previous two classification are independent. In other words, both passive and active mixers could be unbalanced, single balanced or double balanced.

4.2.1. Single-balanced Vs Double-balanced

By single balanced, we mean that only LO signal is differential while the input signal is not. If we have both LO and input signals in differential fashion, then we have a double balanced mixer. Single balanced mixers are called so because of the balanced LO waveforms, this configuration provides differential outputs even with a single ended RF input, easing the design of subsequent stages. Also, the LO-RF feedthrough at ω_{LO} vanishes if the circuit is symmetric. On the other hand; the single-balanced mixer suffers from significant LO-IF feedthrough. In the double balanced mixer, the circuit operates with both balanced LO waveforms and balanced RF inputs. The advantage of such topology is the cancellation of the LO-IF feedthrough. Figures 4.7 and 4.8 show the implementations of single and double balanced passive mixers respectively.

4.2.2. Active Vs Passive

Mixers can be generally classified as passive topologies, whose transistors do not operate as amplifying devices and active topologies, which achieves conversion voltage gain higher than unity. Both can be realized as a single-balanced or a double-balanced circuit. Passive mixers provide better linearity performance than active mixers but worse noise performance. On the other hand, active mixers provide better isolation between ports as passive mixers incorporate impedance translation from baseband to RF port. Passive mixers have the temptation of having no DC current flowing through it. Hence, ideally, there is no flicker noise. Consequently, passive mixers are good candidates for low power applications. Unfortunately, to enjoy the aforementioned merits, we have to afford that switching is lossy operation. They can give loss ranging from few dBs to around 10 dBs.

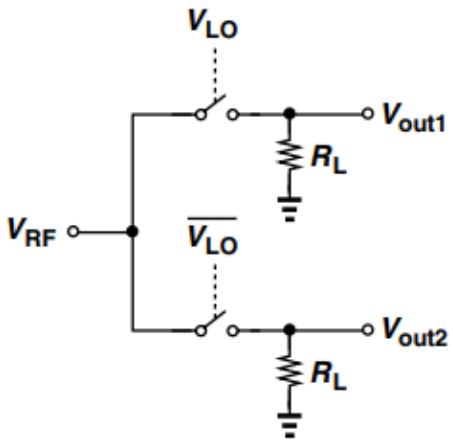


Figure 4.7: Single balanced passive mixer

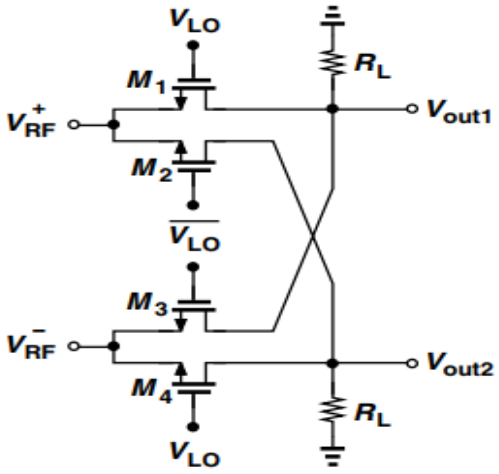


Figure 4.8: Double balanced passive mixer

4.3. Passive Mixer Topologies

Passive mixers can be either in voltage mode or in current mode. The choice of which mode to use is not only dependent on the given specifications, but also it depends on the output of the previous stage whether it is voltage or current. For convenience, in receivers if we have a low noise trans-impedance amplifier, we would better use a current mode mixer. For our case, we will see which is better to use in the next two sections.

4.3.1. Current-mode Passive Mixer

As shown in Fig. 4.9, a current passive mixer is composed of switching quad that commutes the input AC current periodically [2]. Since the output is current from the switching quad, we have to convert it to voltage for the sake of the next blocks. A trans-impedance amplifier is used for this task. Moreover, if the previous stage is not delivering a current, which is the case in the DAC's filter used in this chain, we have to use a Gm stage to convert voltage to current. The switching quad linearity is superior. The same claim holds for the TIA due to the feedback, which creates a low impedance node, making swing at the input too small. Unfortunately, the linearity of Both TIA and switching quad will be degraded dramatically by the humble linearity of Gm stage. Moreover, huge current will be used for TIA and Gm stage. Based on the previous comparative study of advantages and disadvantages of the current mode passive mixer, it will not suitable for this application because we need linear mixer with low power consumption.

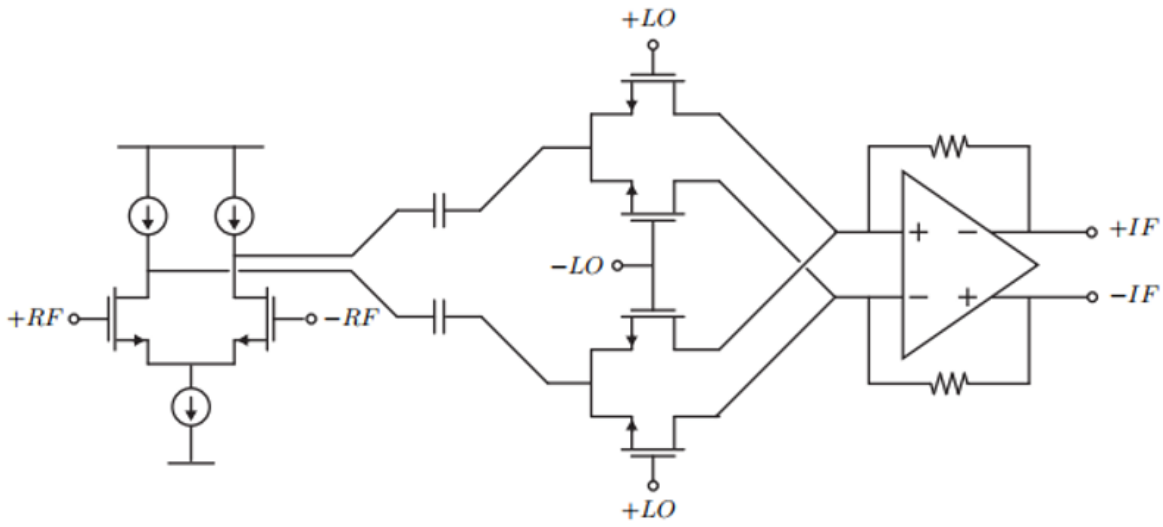


Figure 4.9: Current mode passive mixer

4.3.2. Proposed Passive Mixer

A voltage mode passive mixer would be favored among these topologies [1]. It is composed of switching quad that takes an input voltage. This voltage is used to feed the output nodes in certain sequence. For example, assuming 50% duty cycle, node X will be feed by the positive terminal of the input for half of the LO cycle and by the negative terminal on the other half of the cycle. The output waveform is shown in Fig. 4.11. The advantage of voltage mode mixer is that we do not have to use a Gm

stage before it. This will lead to superior linearity because switching quad can offer high linearity as mentioned in section 4.2.1. Moreover, this topology will give low power solution. More details can be found in [3].

Regarding Conversion gain, an abstraction is made to the mixer as shown in Fig. 4.12. The conversion gain is as shown in (4.1).

$$G_C = 20log_{10}\left[\frac{2}{\pi}\left(\frac{Z_L}{Z_L+R_{ON}||Z_{OFF}}\right) - \frac{2}{\pi}\left(\frac{Z_L}{Z_L+ Z_{OFF}}\right)\right] \quad (4.1)$$

Regarding noise performance, we can say for a 50% duty cycle, each output node is connected to a single switch resistance at a time. This means that power spectral density at output will be 8KT/g_{ds}. This will lead to a noise figure as given by (4.2).

$$SSB\ NF = 10\ log_{10}\left(\frac{1}{g_c^2} + \frac{2}{g_c^2\ g_{ds}R_{source}}\right) \quad (4.2)$$

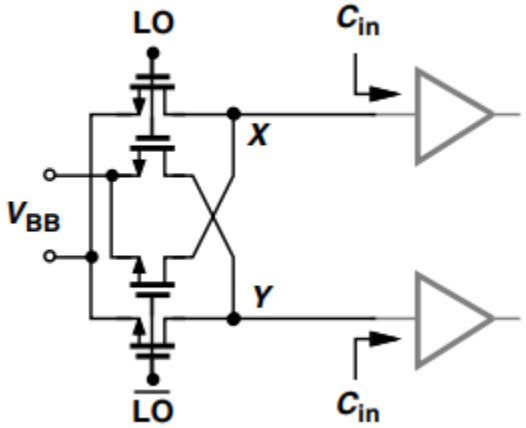


Figure 4.10: Voltage mode passive mixer driving capacitive load

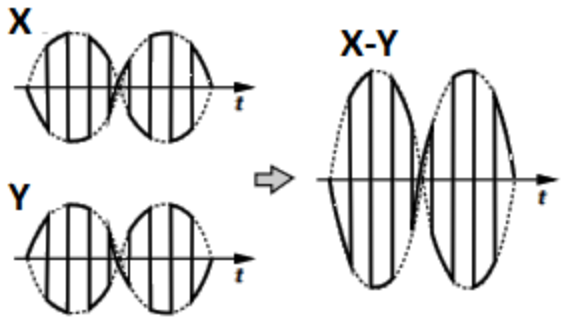


Figure 4.11: Output waveform

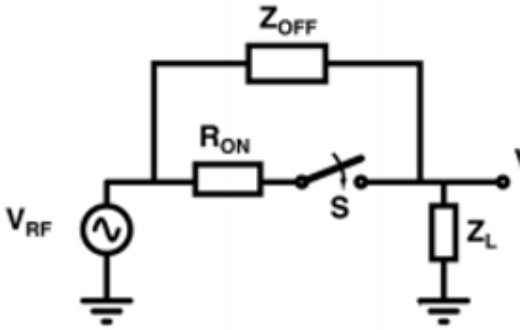


Figure 4.12: Mixer's abstraction as a switch

4.3.2.1. Voltage-mode Mixer's Issues

As we saw in the previous section how voltage mode passive mixer works, we can say that a well-designed switch will lead to good performance. Therefore, we have to mention some issues related) to MOS switches [4].

Voltage Limitation: A great attention must be given to the voltage levels between the device terminals [4]. Each technology will specify the upper limit of the voltage difference between the device terminals to avoid stress and keep it reliable. If we exceeded these limits, we would suffer from irreversible breakdown effects. For example, If the LO is high such that the voltage difference between gate and source is high, the gate oxide will breakdown irreversibly. Moreover, if the voltage between source and drain is too large, the depletion region around the source will get wider until it touches that of the drain and an excessive large current will flow. The latter mentioned issue manifests itself more and more as the technology node advances. Therefore, a great attention should be given to biasing to avoid these issues.

Speed: Since sampling mixers drive capacitive load that is composed of the capacitance of the next stage and its own parasitic capacitance, the input will suffer from a delay due to the finite bandwidth RC filter as shown in Fig. 4.13. This delay will lead to an error where the output will not track the input faithfully. This will exacerbate distortion and reduction in conversion gain. Therefore for good performance, we have to guarantee that the pole at $(r_{sw}\ C_{load})^{-1}$ is sufficiently away from the operating frequency.

Feedthrough: Ideally, using double balanced mixers will remove both feedthroughs from input and LO ports to the output port. However, practically this rejection will be of finite value. Feedthrough is mainly due to the device capacitance as shown in Fig. 4.14. LO to IF feedthrough is given more care where the LO transitions are coupled to the output through a capacitive divider between overlap capacitance Cgd and the load capacitance, causing error in the sampled value at the sampling instants. From (4.3), we can say that driving heavy capacitive load would reduce this effect. Finally, we have to say that using perfectly matched double balanced mixer would make this issue insignificance.

$$\Delta V = (V_{LO,HIGH} + V_{LO,bias}) \frac{W\ C_{ov}}{W\ C_{ov} + C_{Load}} \qquad (4.3)$$

Charge Injection: For the MOS switch to be on, we set the gate to high voltage in order to be enough for making strong inversion and form a conductive channel. The channel is rich in charges and supposed to offer good transmission for the signal without any distortion. When the gate signal goes down, the switch is supposed to be completely off and there is no channel available for signal transmission. Charge injection takes over when the switch is moving from ON to OFF state as shown in Fig. 4.15. Roughly, half of the channel charges finds its way to the input source while the other half will accumulate on the capacitive load [4].We can calculate the voltage offset added due to charge injection using (4.4).

$$\Delta V = \frac{W\ L\ C_{ox}\ (V_{LO,HIGH} + V_{LO,bias} - V_{in} - V_{TH})}{2\ C_{Load}} \qquad (4.4)$$

According to [4], charge injection introduces nonlinearity in addition to dc offsets because the amount of accumulated charges is function of the input signal. Nonlinearity will manifest itself more and more if we considered body effect.

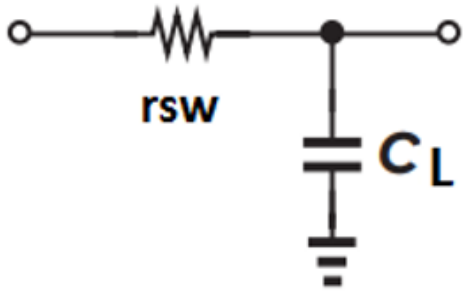


Figure 4.13: Switch ON-state RC section

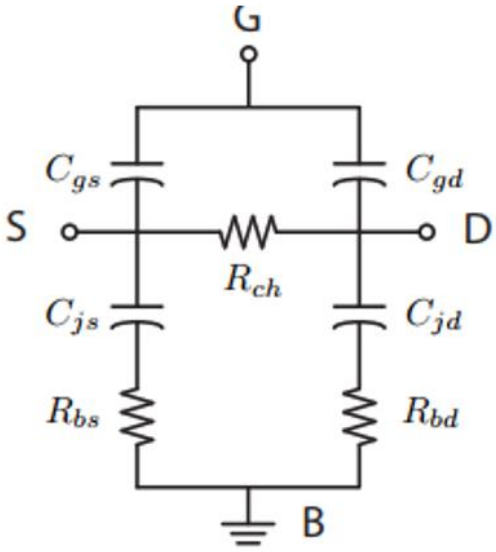


Figure 4.14: Switch model after adding parasitics in ON state

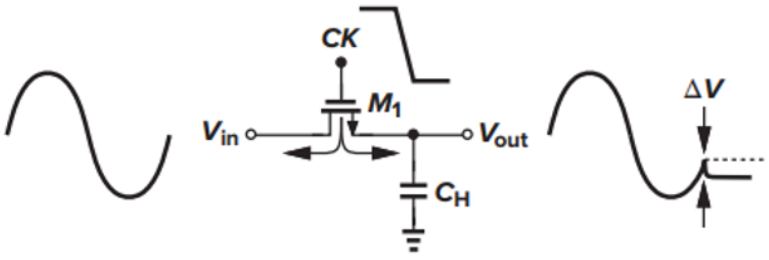


Figure 4.15: Charge injection effect on transmission accuracy

4.3.3. Passive Mixer Results

For consistence, performance parameters that are addressed in section 4.1 are simulated. Results are shown in plots for better visualization.

Conversion Gain: CG is plotted versus LO swing in Fig. 4.16. Conversion gain is about -3.87 dB. From the plot, we can deduce that conversion gain is of well defined value for LO swing equal to or above 750 mVpp. Practically, we have to use rail to rail LO swing to avoid soft switching which will manifest nonidealities that will degrade performance.

Noise Figure: In Fig. 4.17, noise figure is plotted versus frequency at different corners. At typical situation, Noise figure of around 8 dB is achieved. At corners, Noise figure will depart from its value at typical situation, achieving 6.415 dB at best corner and around 10 dB at worst corner.

1-dB Compression Point: In Fig. 4.18, 1 dB compression point is calculated at different corners. At typical situation, P1dB is around 12 dB. Linearity changes at corners achieving 11.4 dBm at worst corner and 12.8 dBm at best corner.

Output Waveform: In Fig. 4.19, mixer's output voltage is plotted versus time for an input of amplitude 160mv. The waveform is near to the waveform shown in figure (7).

LO-Baseband Feedthrough: In Fig. 4.20, LO feedthrough to the baseband port is plotted for each harmonic of the LO signal. It is obvious that isolation is about -230 db. This means that LO signal is highly attenuated when it leaks to the baseband port. Consequently, we won't suffer from offset in the transmitted signal constellation due to self-mixing.

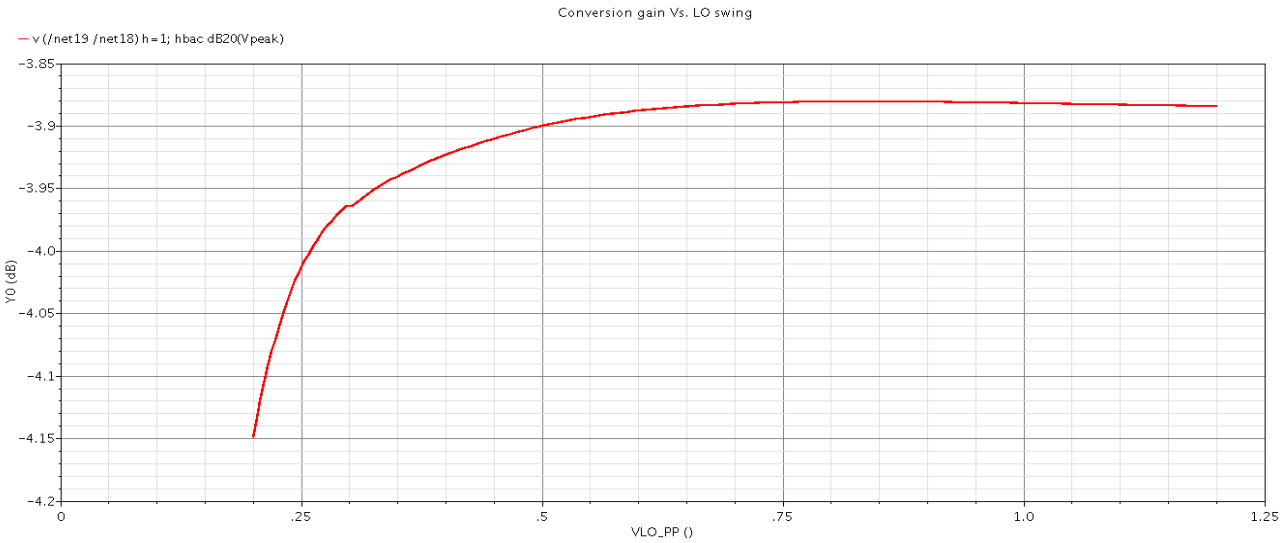


Figure 4.16: Conversion gain versus LO swing

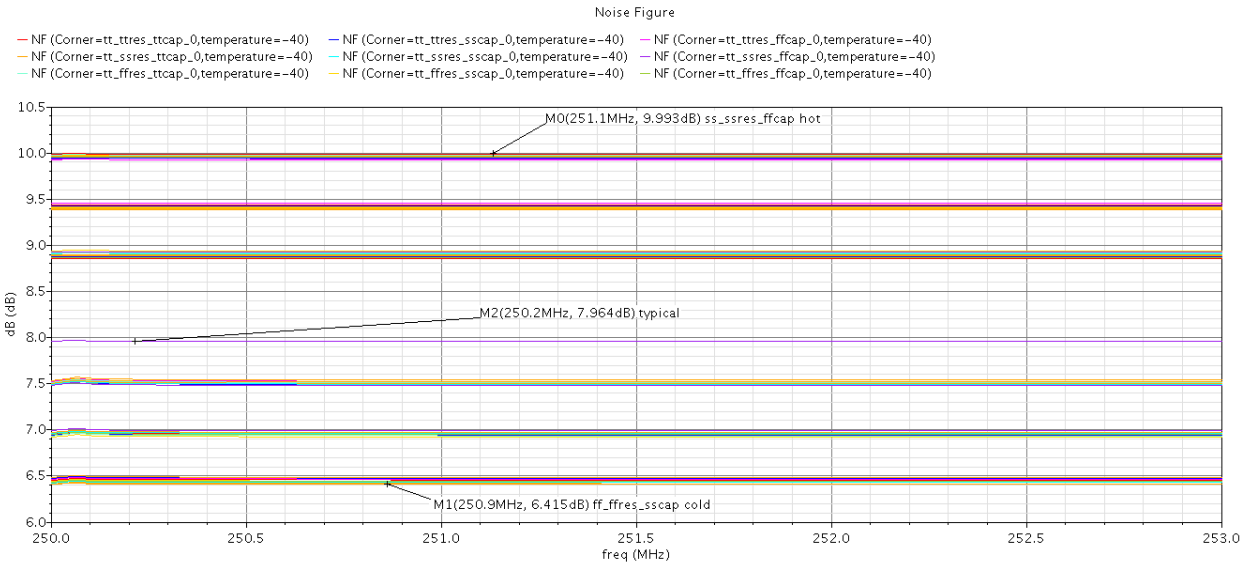


Figure 4.17: Noise figure

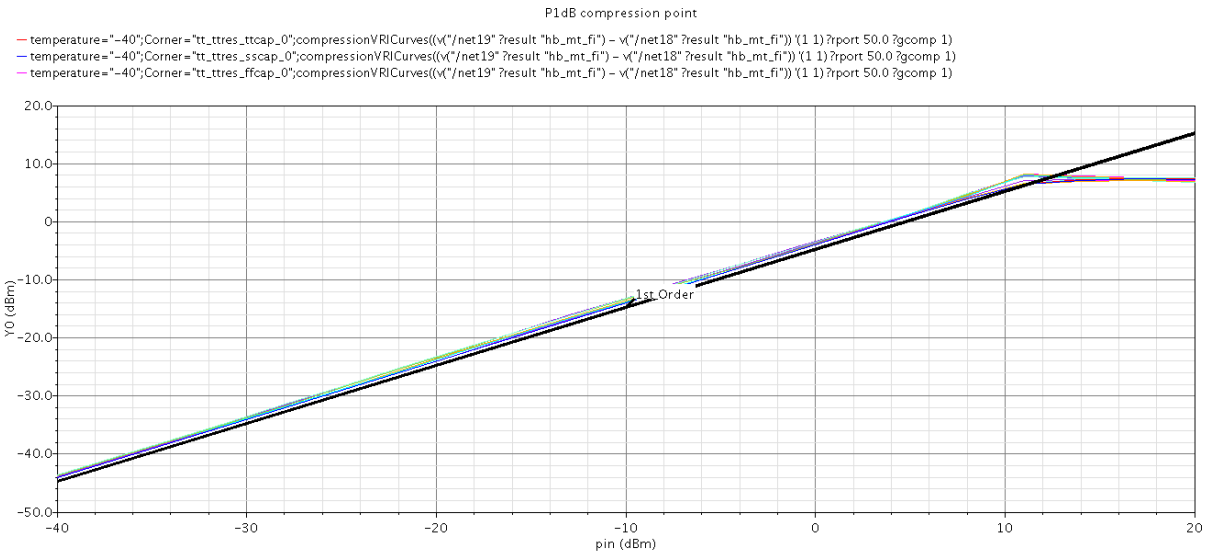


Figure 4.18: 1-dB compression point

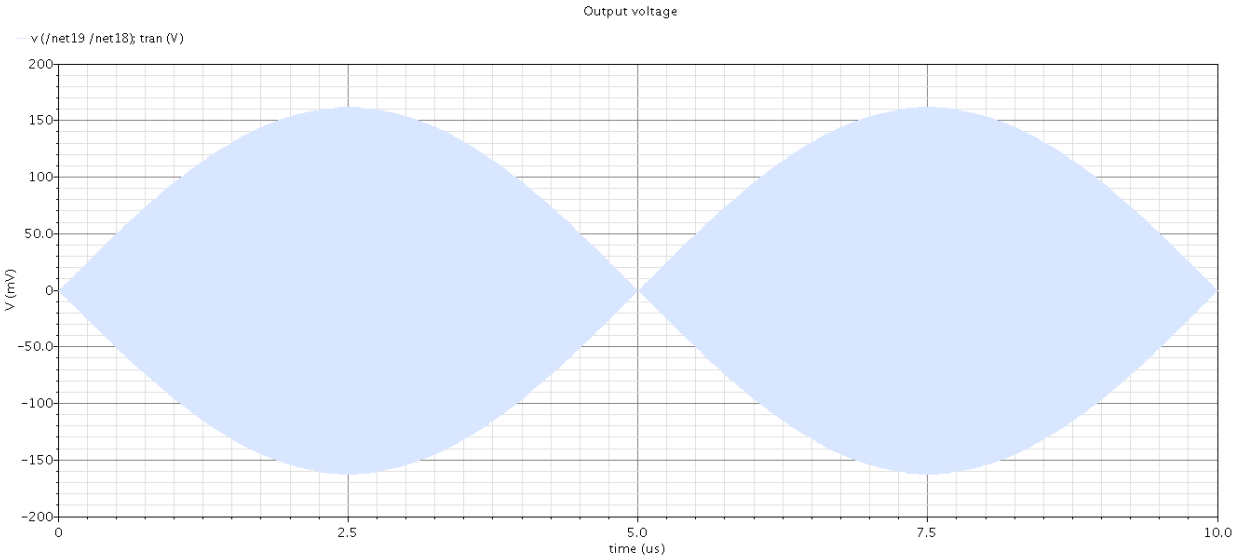


Figure 4.19: Output voltage

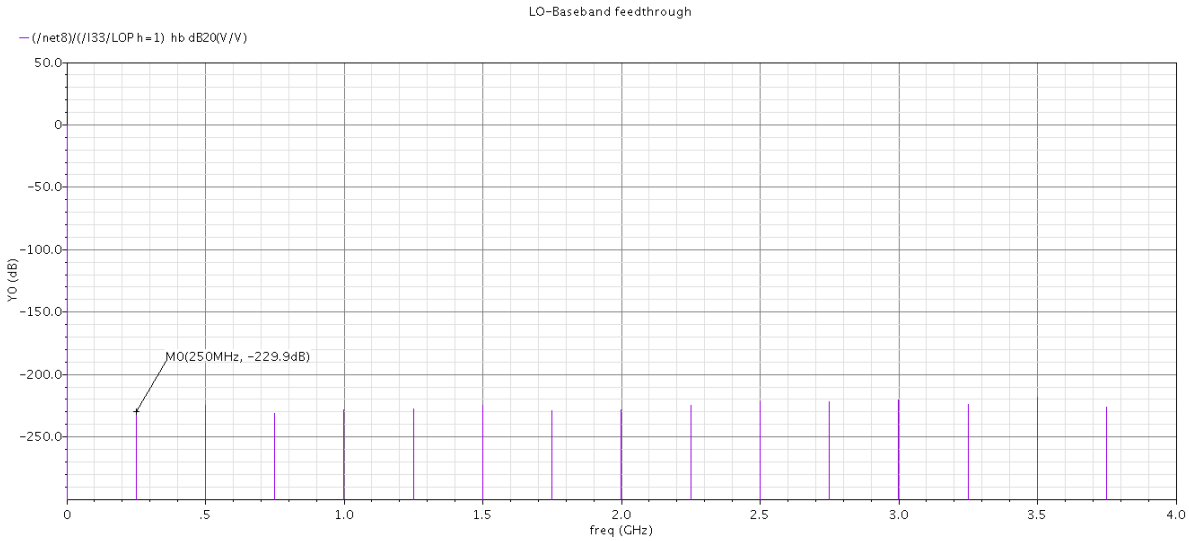


Figure 4.20: LO-Baseband feedthrough

4.4. Quadrature Summer

In quadrature up-converters, we must sum their outputs in order to be processed by the rest of the chain. Hypothetically, if we have two voltage sources, we may sum them by connecting them in series. Unfortunately, quadrature up-conversion mixers will not lend themselves to this idea. On the other hand, summing currents can be made by injecting them into a single node. The previous thoughts can give us a good method for solving this problem. As shown in the Fig. 4.21 we may convert voltages into currents, injecting these currents into nodes and then convert again to voltage.

4.4.1. Voltage Summer Implementation

On the circuit level, a trans-conductor will play the role of converting voltage into current. The most simple and well-known trans-conductor is a MOS device. On the other hand, a trans-impedance amplifier can do the job of converting current into voltage and the simplest implementation would be a resistor. As shown in Fig. 4.22, a quasi-differential pair is used for summing voltage. Thanks to the ground at the source terminal, we can achieve high linearity. After summing, we must have some sort of filtering in order to extract the desired tone from the mixer output. For the sake of this job, we may exploit the load resistance along with a capacitor parallel to it in order to do the job of filtering and summing currents. It is worth mentioning that a great attention should be given during design because the bias point in a strong function of voltage and temperature variations.

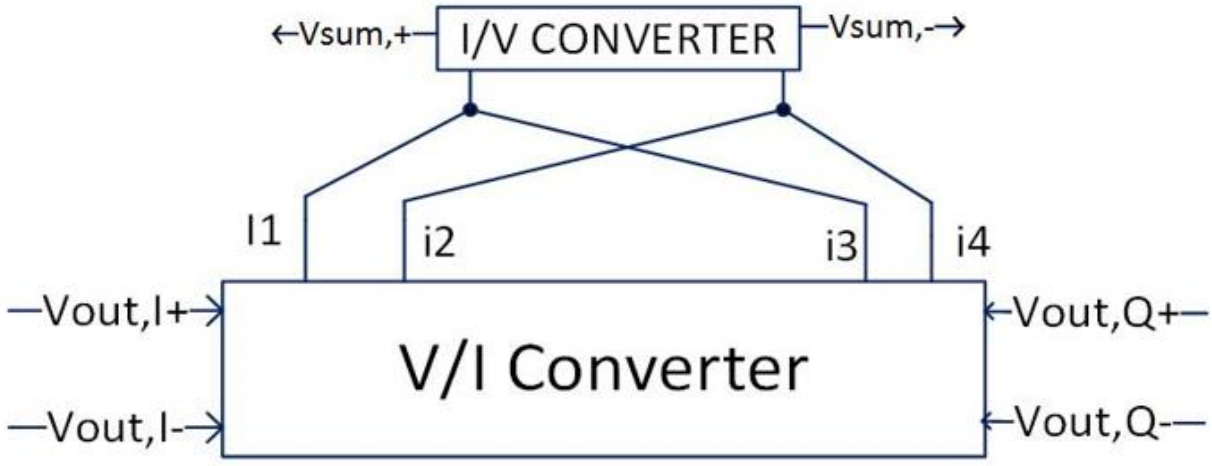


Figure 4.21: Hypothetical implementation of voltage summer

4.4.1. Voltage Summer Results

In this section, gain and linearity are simulated. Plots for results are given for better visualization. As shown in Fig. 4.23, gain has minimum value around -11 dB and maximum value around -4.8 dB at 250MHz. As shown in Fig. 4.24, input referred 1db compression point is calculated at different corners. Its maximum value is around 9.2 dBm and its minimum is around 0 dBm.

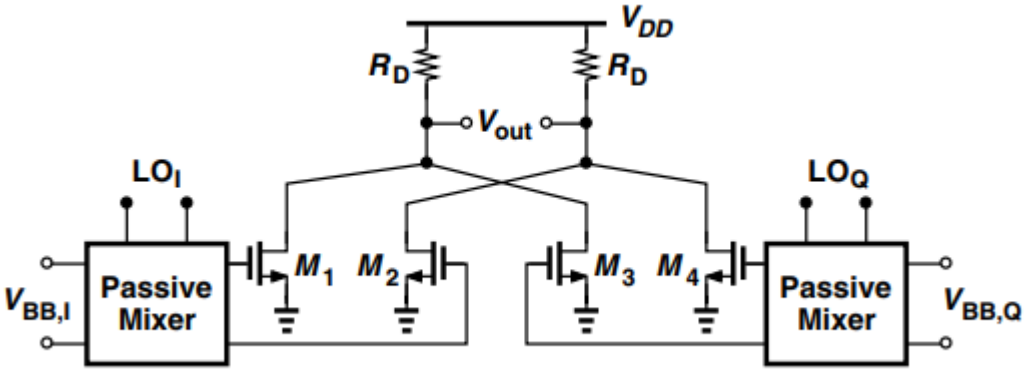


Figure 4.22: Quasi-differential pair as voltage summer

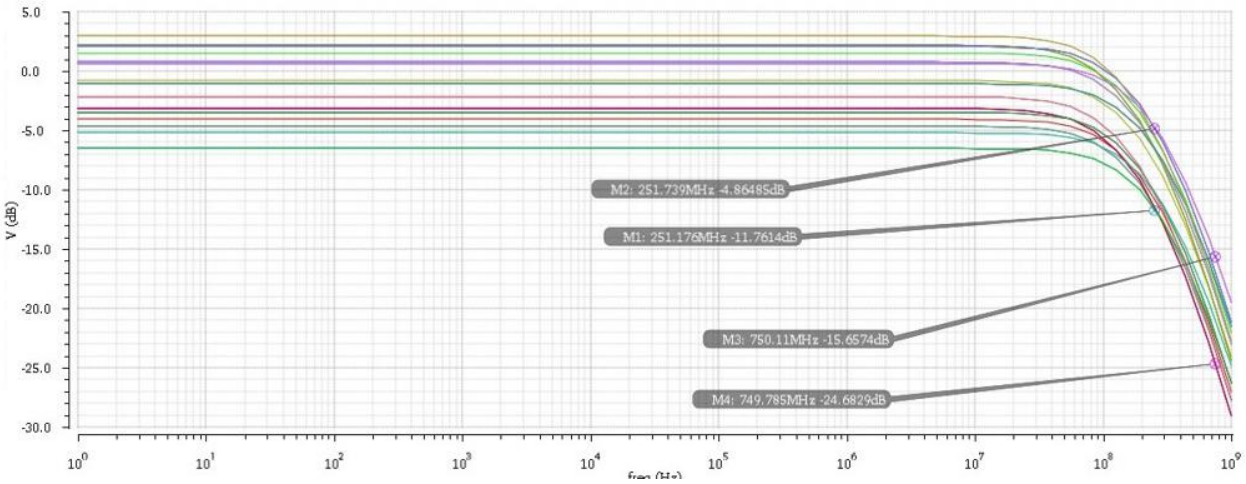


Figure 4.23: Gain across corners

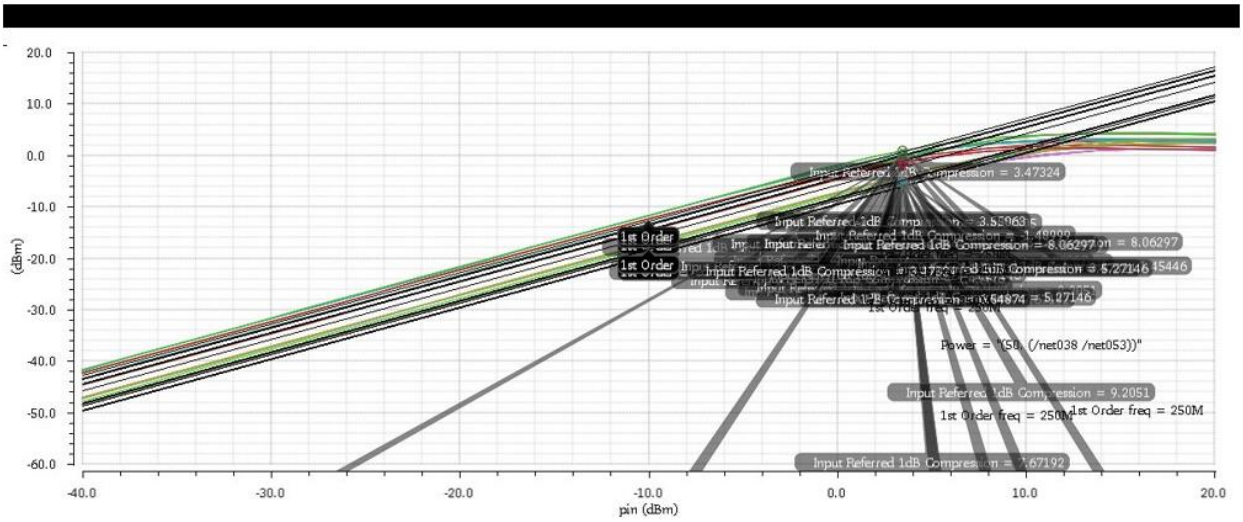


Figure 4.24: 1-dB compression point

4.5. Overall Results

To make sure that system level specifications are achieved. Simulations are carried out for I-Q mixers followed by the summer. From system level design it is required to achieve total conversion gain of -8.5 dB. At typical situation, conversion gain of -9.28 dB is achieved as shown in Fig. 4.25. From system level design, it is required to achieve 1 dB compression point not less than -3 dBm. In Fig. 4.26, 1 dB compression point is calculated at different corners. At typical situation, P 1-dB is around 3 dBm. The 1 dB compression point is about 5.8 dBm at the best corner while it falls to around -2.26 dBm at the worst corner. Additionally, we need to achieve an overall maximum noise figure of 45 dB. Simulations were carried out and noise

figure is plotted versus frequency at different corners in Fig. 4.27 At typical situation, the achieved noise figure is about 36 dB. This value will fall to 32.76 dB at the best corner and rise to 41.38 dB at the worst corner.

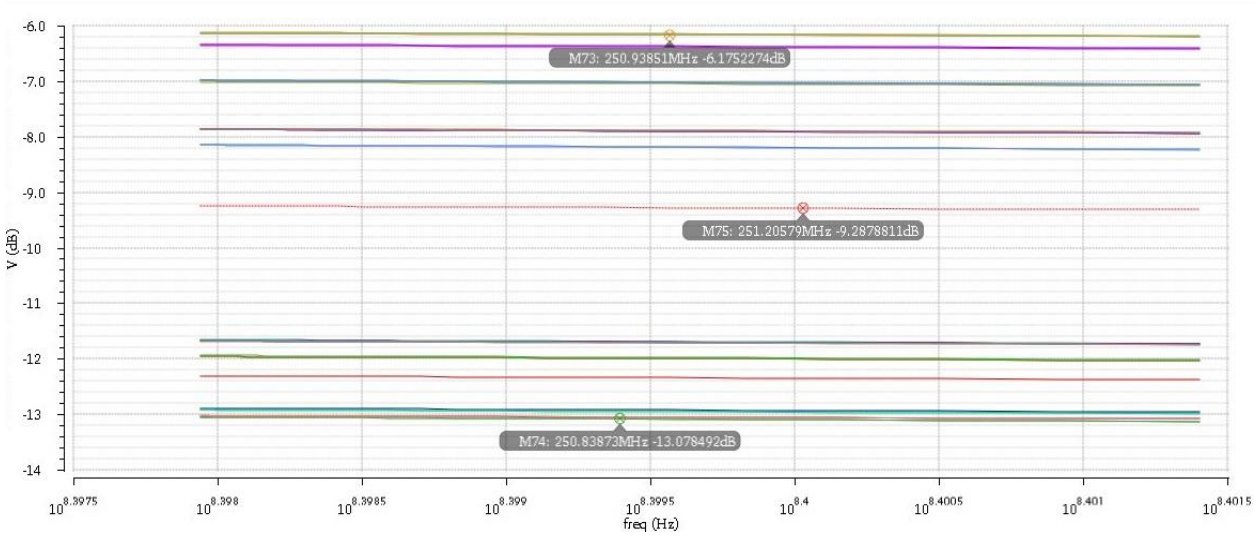


Figure 4.25: Overall gain

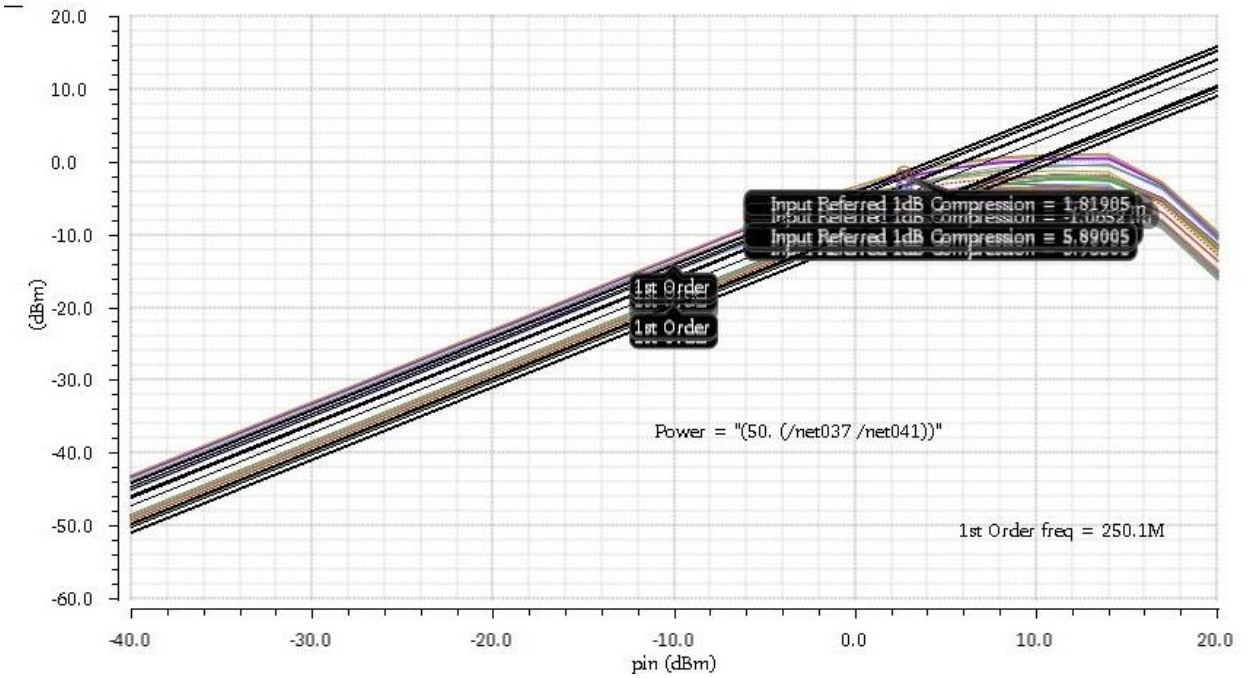


Figure 4.26: Overall 1-dB compression point

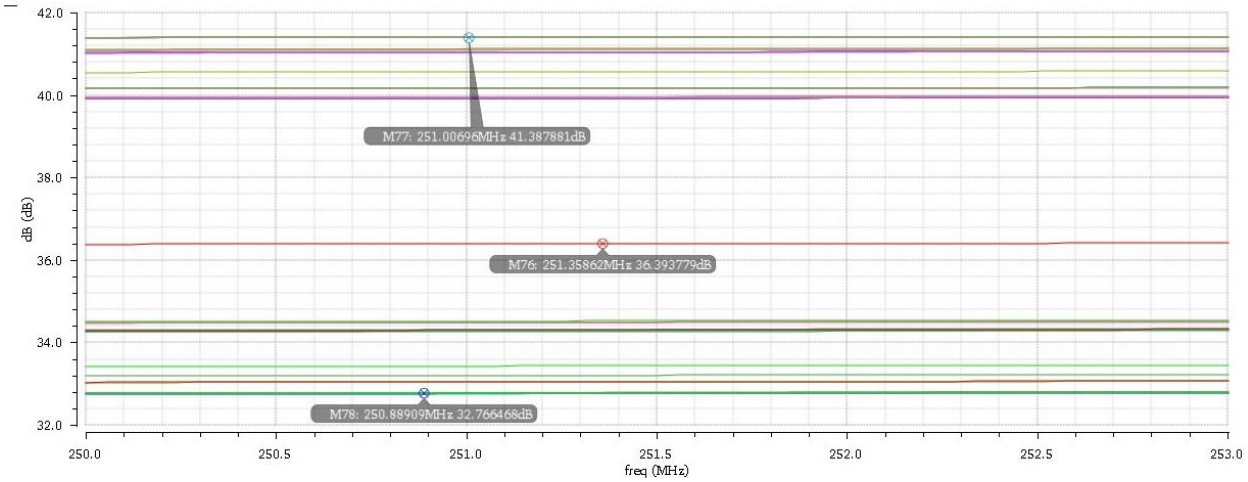


Figure 4.27: Overall noise figure

References

[1] B. Razavi, “RF microelectronics”, Upper Saddle River, NJ: Prentice Hall, 2012

[2] Yun-Yong Yu, Si-Zheng Chen, Tong Li and Hao Min (2016) Design of a Low Power Current-Driven Passive Mixer, IEEE.

[3] K. Komoni and S. Sonkusale, "Modeling, simulation and implementation of a passive mixer in 130nm CMOS technology and scaling issues for future technologies", 2008 51st Midwest Symposium on Circuits and Systems, 2008.

[4] B. Razavi, “Design of analog CMOS integrated circuits”. Boston, Mass. McGraw-Hill.

Chapter 5 : Offset PLL

Phase-Locked Loops are from the most essential and critical building blocks in any RF system. PLL's are used in a variety of applications such as Frequency Synthesis, Frequency Modulation and Demodulation, Clock and Data Recovery and Clock Skew Cancellation. In our application, the PLL is used to precisely up-convert the phase path signal from IF frequency to RF frequency.

5.1. Basic Concepts

A PLL can be simply considered as a negative feedback system for signal's phase. Ideally, PLL consists of a Phase Detector (PD) and a Voltage Controlled Oscillator (VCO) as shown in Fig. 1. A phase detector is a circuit that detects the phase difference between two signals and produce a proportional output voltage corresponding to this phase difference. The VCO in turns generates an output signal with a frequency proportional to the applied voltage. This output frequency can drift due to VCO's phase noise. However, the loop locks this frequency by means of negative feedback thus maintains an accurate and stable output frequency.

5.2. Type-I PLL

From Control theory, a system type is determined by the number of poles which the open loop transfer function has at the origin (at s=0). Type-I PLL is same as the simple PLL shown in Fig. 5.1 except for the addition of a LPF (Loop filter) is to remove the high frequency components added by the PD due to repetitive pulses which can modulate the VCO frequency. Fig. 5.2 shows a simple block diagram representation for Type-I PLL. The output of the loop filter is the average voltage to these pulses and is used as a control voltage to the VCO. Having a lower cut off frequency attenuates the ripples in Vcont. Since the PLL is a phase feedback system, thus the transfer function should be derived from a phase-domain point of view. The phase detector simply subtracts the output phase from the input phase and scales the result by a factor of K_{PD} so as to generate an average voltage. This voltage is applied to the low-pass filter and subsequently to the VCO. Since the phase detector only senses the output phase, the VCO must be modeled as a circuit with a voltage input and a phase output. According to [1], the VCO is modeled as ideal integrator with gain Kvco. The loop filter can be modeled by a simple parallel RC section. The corresponding phase-domain model for Type-I PLL is shown in Fig. 5.3. The transfer function is then given by

$$H(s) = \frac{K_{PD}K_{VCO}}{R_1C_1s^2+s+K_{PD}K_{VCO}} \tag{5.1}$$

By comparing (5.1) to the familiar second order system transfer function

$$H(s) = \frac{{\omega_n}^2}{s^2+2\xi\omega_n s+{\omega_n}^2} \tag{5.2}$$

where ξ is the "damping factor" and ω_n the "natural frequency".

Thus,

$$\xi = \frac{1}{2}\sqrt{\frac{\omega_{LPF}}{K_{PD}K_{VCO}}} \tag{5.3}$$

$$\omega_n = \sqrt{K_{PD}K_{VCO}\omega_{LPF}} \tag{5.4}$$

and $\omega_{LPF} = \frac{1}{R_1C_1}$

Type-I PLL incorporates some drawbacks as mentioned below [1].

- There is a tight relation between the loop stability (damping factor) and the LPF corner frequency.
- The limited acquisition range, if the input frequency and the VCO free running frequency are far apart at the startup, the PLL may never acquire lock.
- Finite static phase error and its variation with input frequency.

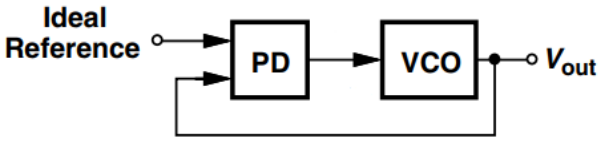


Figure 0.1: Simple PLL

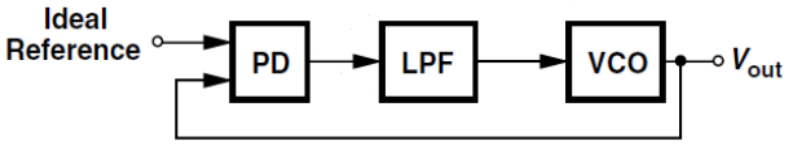


Figure 0.2: Type-I PLL

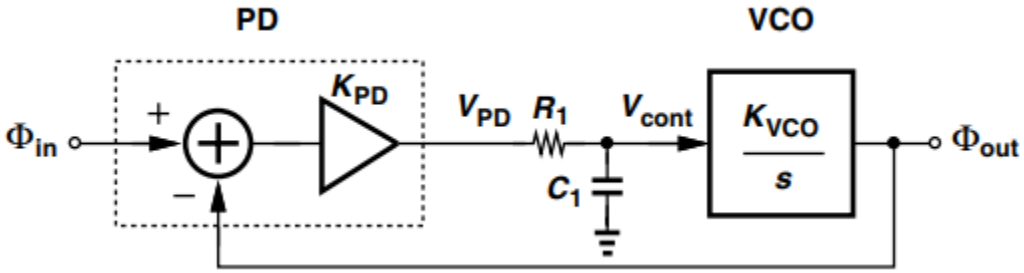


Figure 5.3: Type-I PLL Phase-domain Model

5.3. Type-II (Charge Pump) PLL

One critical drawback of type-I PLL is limited acquisition range. This limitation arises because the PD produces little information if it senses unequal frequencies at its input. To address the acquisition range issue, the PFD (Phase/Frequency Detector) is introduced.

5.3.1. Phase / Frequency Detector (PFD)

Fig. 5.4 conceptually shows the operation of a PFD. The circuit produces *two* outputs, Q_A and Q_B, and operates based on the following principles: (1) a rising edge on *A* yields a rising edge on Q_A (if Q_A is low), and (2) a rising edge on *B* resets *QA* (if Q_A is high). The circuit is symmetric with respect to *A* and *B* (and Q_A and Q_B,). We observe from Fig. 4(a) that, if $\omega_A > \omega_B$, then Q_A produces pulses while Q_B, remains at zero. Conversely, if $\omega_B > \omega_A$, then positive pulses appear at Q_B, and $Q_A = 0$. On the other hand, as depicted in Fig. 2(b), if $\omega_A > \omega_B$, the circuit generates pulses at either Q_A or Q_B, with a width equal to the phase difference between *A* and *B*. Thus, the average value of $Q_A - Q_B$, represents the frequency or phase difference. To arrive at a circuit implementation of the above idea, we surmise that at least three logical states are necessary: $Q_A = Q_B = 0$; $Q_A = 0$, $Q_B = 1$; and $Q_A = 1$, $Q_B = 0$. In addition, to avoid dependence of the output upon the duty cycle of the inputs, the circuit should be realized as an edge-triggered sequential machine. Fig. 5.5 shows a state diagram summarizing the operation.

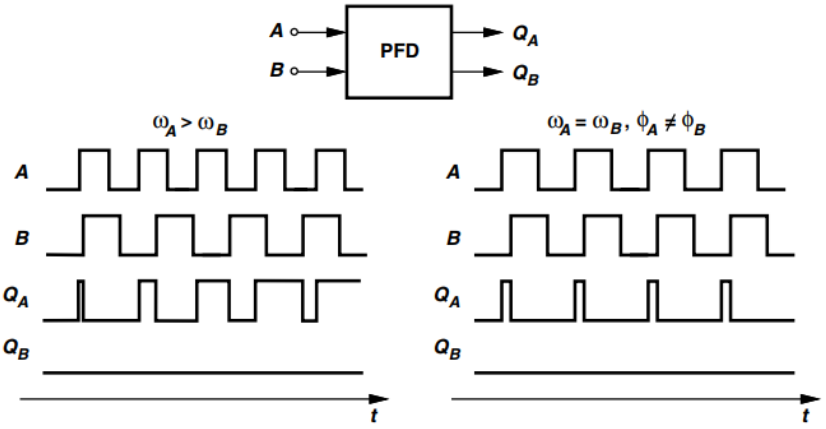


Figure 0.4: Response of a PFD to inputs with unequal (a) frequencies, or (b) phases

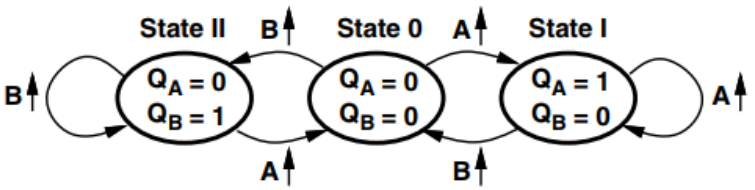


Figure 0.5: State Diagram of the PFD

5.3.2. Charge Pump

To address the finite static error in phase and the dependency between system stability and LPF corner frequency, Charge Pumps are introduced. A charge pump sinks or sources current for a limited period. Fig. 5.6 shows a conceptual view of the charge pump. When up is high and S1 is on, I1 charges C1. When down is high and S2 is on, I2 discharges C1. Nominally, I1=I2=Ip such that in locked condition when up and down pulses arrive simultaneously, no current flows in C1 and Vout is kept constant. The operation of the charge pump is shown in Fig. 5.7. During the Up cycle, a pulse of width ΔT on Q_A turns S1 on for a period of ΔT charging C1 by ΔT.I1/C1. During the Down cycle, a pulse of width ΔT on Q_B turns S2 on for a period of ΔT discharging C1 by ΔT.I2/C1. Fig. 5.7 shows the integration behavior of the charge pump, hence its transfer function is

$$\frac{V_{cont}}{\Delta\phi}(s) = \frac{I_p}{2\pi C_1}\frac{1}{s} \qquad (5.5)$$

In addition, the total transfer function of the charge pump PLL is given by

$$H(s) = \frac{\frac{I_p \ K_{VCO}}{2\pi C_1 s \ \ s}}{1+\frac{I_p \ K_{VCO}}{2\pi C_1 s \ \ s}} = \frac{I_p K_{VCO}}{2\pi C_1 s^2 + I_p K_{VCO}} \qquad (5.6)$$

Equation (5.6) shows that the open-loop transfer function has two poles at s=0, thus called Type-II PLL. Moreover, the system is unstable due to the absence of coefficient of s in the characteristic equation. The system can be stabilized if one of the two integrators (VCO and Charge Pump) becomes lossy. A series resistance is added with C1 to introduce this effect as shown in Fig. 5.8. The closed loop transfer function is now given by

$$H(s) = \frac{\frac{I_p K_{VCO}}{2\pi C_1}(R_1 C_1 s+1)}{s^2+\frac{I_p K_{VCO}}{2\pi} R_1 s+\frac{I_p K_{VCO}}{2\pi C_1}} \qquad (5.7)$$

Similar to Type-I PLL, the transfer function in (5.7) is compared to the second order system transfer function to get the damping coefficient and the natural frequency.

$$\xi = \frac{R_1}{2}\sqrt{\frac{I_p K_{VCO} C_1}{2\pi}} \qquad (5.8)$$

$$\omega_n = \sqrt{\frac{I_p K_{VCO}}{2\pi C_1}} \qquad (5.9)$$

Interestingly, as C1 increases (so as to reduce the ripple on the control voltage), so does ξ —a trend opposite of that observed in type-I PLLs.

For the continuous time approximation to be valid (i.e. the charge pump is approximated correctly as an integrator), The bandwidth of the loop filter should be much less than reference frequency (generally 0.1 reference frequency is sufficient)

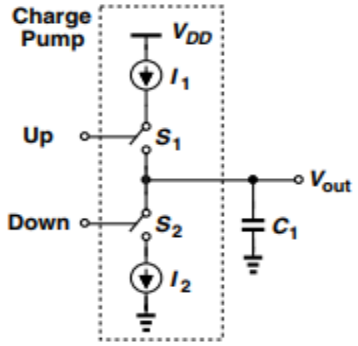


Figure 5.6: Conceptual Charge Pump View

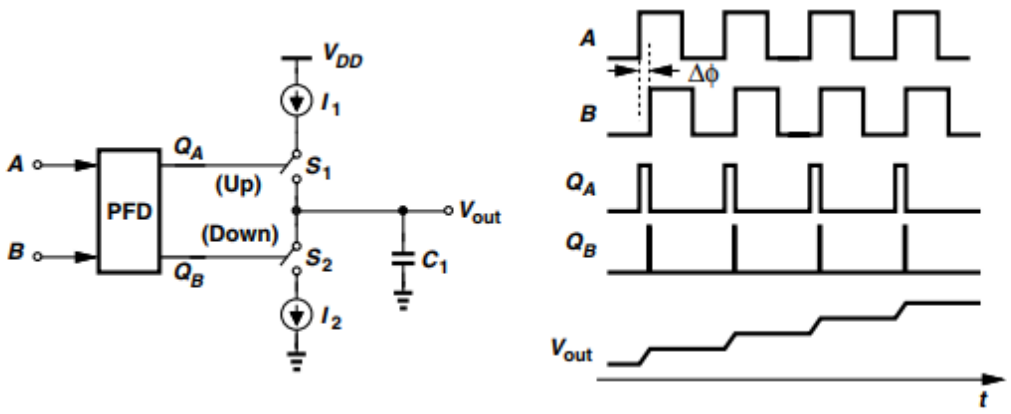


Figure 5.7: Charge Pump Operation

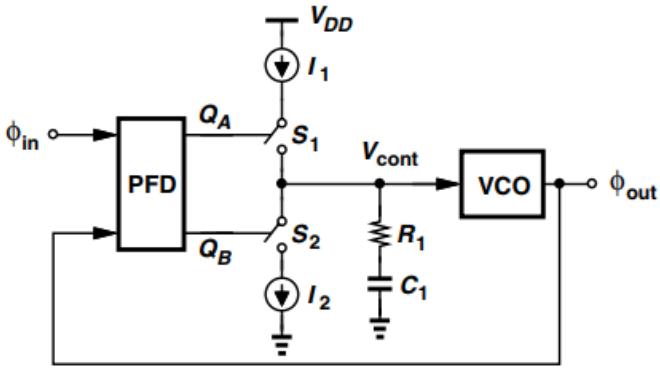


Figure 5.8: Charge Pump PLL

5.4. Higher Order Loops

The loop filter consisting of R1 and C1 in Fig. 8 does not maintain a smooth control voltage as it does not suppress the ripples resulting from current passing through R1 even in the lock condition. Fig. 9 demonstrates the ripple resulting from Up/Down pulses arriving every Tin seconds with a small skew equals ΔT.

A solution to smooth these ripples is to introduce a capacitor C2 tied directly from control line to ground introducing a low impedance path for the unwanted charge pump current as shown in Fig. 5.10.

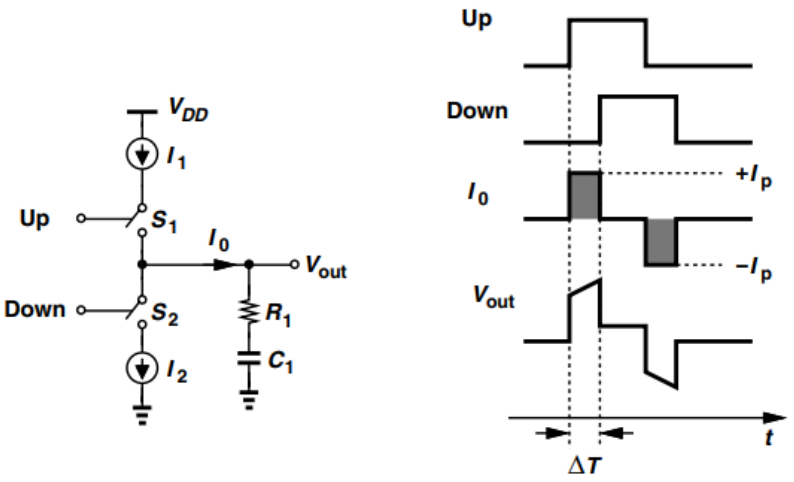


Figure 0.9: Effect of skew between up and down pulses

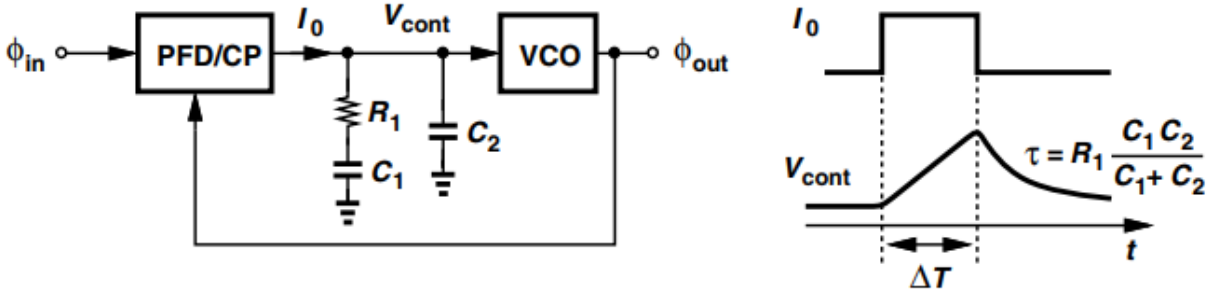


Figure 0.10: Addition of second capacitor to loop filter

5.5. Offset PLL

Offset PLL (OPLL) depends on mixers to perform feedback frequency-division operation. As explained before, implementation of the divider using a mixer is to keep the phase information without change while translating the VCO output frequency to the PFD/CP input frequency [1]. The main difference between a PLL and an OPLL is that the frequency modulation of the reference input is reproduced at the VCO output without scaling [2]. Fig. 11 shows a block diagram of the proposed OPLL in this work. A divide-by-2 divider is used and the VCO is run at twice the PA frequency to avoid injection pulling between the PLL and the PA [1, 3]. Each block will be presented in details as well as OPLL system design equations and parameters in the following subsections.

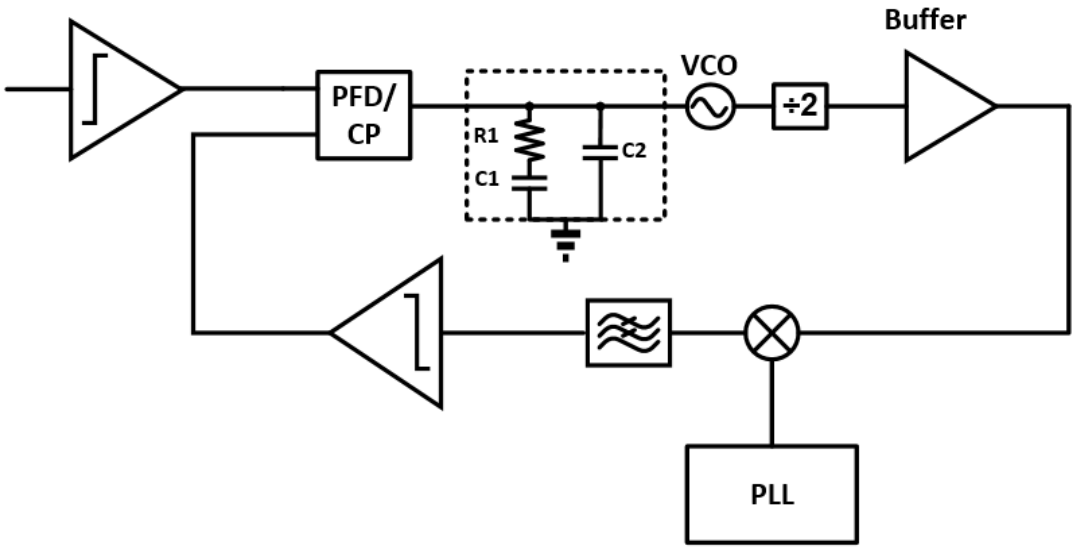


Figure 0.11: Proposed PLL

5.5.1. OPLL Bandwidth

Phase locked loops perform low-pass filtration for reference frequency and/or phase variations [1]. This behavior can be utilized to reduce the signal's frequency and/or phase noise. However, PLL has a high-pass behavior regarding the VCO phase noise. Hence, a trade-off arises between reference spurs rejection and VCO noise rejection. Additionally, tight PLL bandwidth increases the loop settling time and hence, introduces more phase error [2]. Based on the previous discussion, PLL's bandwidth should be chosen in order to achieve a good compromise between good noise suppression, hence relaxing noise requirements on baseband blocks, and settling behavior and phase error. Therefore, the PLL's bandwidth is chosen to be 500 KHz.

5.5.2. OPLL System Design

In this subsection, the OPLL system design methodology is illustrated. The target is to reach the required specs of each block. A summary is presented in Table 5.1 The OPLL in Fig. 5.11 is considered a unity feedback loop as the feedback divider only multiplies the frequency not the phase. The following design equations are from [1]. Equations (5.8) and (5.9) are rewritten again for consistence.

$$\xi = \frac{R1}{2}\sqrt{\frac{IpC1Kvco}{M}} \tag{5.10}$$

$$\omega_n = \sqrt{\frac{IpKvco}{M*C1}} \tag{5.11}$$

$$\omega_{3-dB}{}^2 = \omega_n{}^2[1 + 2\xi^2 + \sqrt{(1 + 2\xi^2)^2 + 1)}] \tag{5.12}$$

$$\text{PM} = \tan^{-1}[4\xi^2(1 + \frac{1}{32\xi^2})] - \tan^{-1}[4\xi^2\frac{C_{eq}}{C_1}(1 + \frac{1}{32\xi^2})] \tag{5.13}$$

$$C_{eq} = \frac{C_1C_2}{C_1+C_2} \tag{5.14}$$

Where M is the divide ratio of the divide-by-2 divider following the VCO, i.e. M=2.

As explained in the previous subsection, the loop bandwidth (ω_{3-dB}) is chosen to be 2π*500 KHz. The charge pump current can be chosen within the range from few microamps to few milliamps. Here it is chosen to be 1.6 μA. The VCO gain is set to be 250 MHz/V and C1 equals to 100 pF. Substituting by these values in (5.11) gives a natural frequency $\omega_n = 2\pi * 225\ Krad/s$. Now ξ can be achieved by substituting in (5.12). Hence, $\xi = 0.8485$. Additionally, the loop filter resistance's value is obtained by substituting in (5.10). Thus, $R_1 = 12\ K\Omega$. By choosing the PLL's phase margin to greater than 60, the value of C_2 can now be obtained by substituting in (5.13) and (5.14) respectively. Therefore, for $C_2 = 5\ pF$, $PM = 63.45^o$.

In addition, the VCO Phase noise is determined according to the spectral emission mask specification. Hence, the VCO phase noise at an offset of 500 KHz -85 dBc/Hz.

Finally, the total integrated phase noise corresponds to an Error Vector Magnitude (EVM) of 7.9 % which meets the standard specs according to the following equation [1]

$$EVM = \sqrt{4 * S_0 * f} \quad (5.15)$$

Where S_0 is the phase noise at the loop BW in rad^2/Hz and f is the loop BW. As will be evident in the next chapter, a much better phase noise was achieved and a better EVM.

Table 5.1: OPLL Parameters Values Summary

Parameter	Value
R1	12 KΩ
C1	100 pF
C2	5 pF
Ip	1.6 μ A
Kvco	250 MHz/V
Bandwidth	500 KHz
Phase Margin	63.45

5.5.3. PFD

In this subsection, the PFD circuits architectures, design procedure and simulation results are presented. A practical PFD implementation have some non-idealities compared to the explained theory in subsection 5.3.1.

5.5.3.1. Dead Zone

Charge pump switches need a finite pulse to turn on the switches. Narrow phase error produces short pulses that cause the charge pump not to respond to this phase error. Fig. 5.12 shows the effect of dead zone on the PFD transfer function. To resolve this issue, the up and down pulse widths are made greater than switch on-time with an overlap in the up/down pulses to keep the control voltage constant in lock. Fig. 5.13 shows the overlap in the up/down pulses.

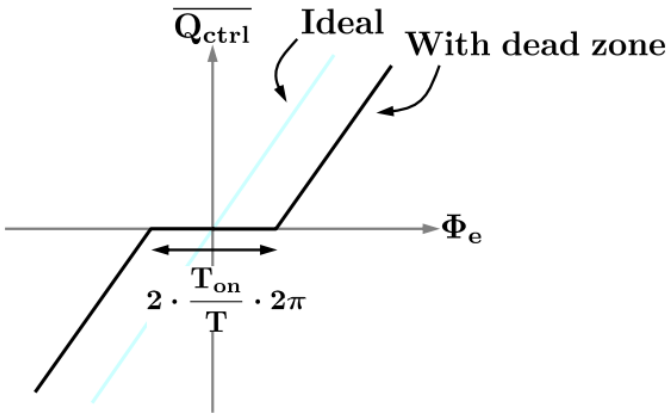


Figure 0.12: PFD transfer function with dead zone

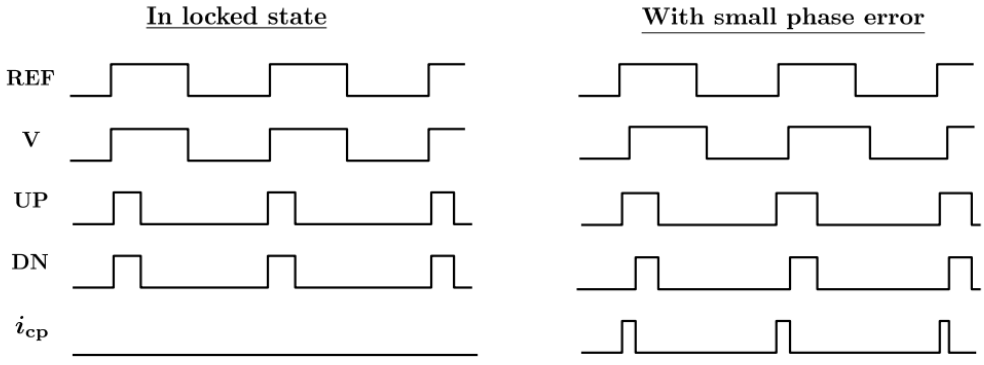


Figure 0.13: Up/down overlap to reduce the dead zone

5.5.3.2. Maximum Operating Frequency

The maximum frequency (f_{max}) is the frequency at which the PFD can still operate properly. There is a tradeoff between the maximum operating frequency and the dead zone reduction. Moreover, f_{max} is equal to $1/2T_{rst}$. However, in our case a T_{rst} of 150 ps makes the maximum operating frequency of 3 GHz which is much above our operating frequency of 250 MHz. Moreover, increasing T_{rst} increases the noise contribution of the charge pump. Thus a compromise between the noise contribution and the acceptable phase error is done.Up and Down Skew and Width Missmatch

5.5.3.3. Up and Down Skew and Width Missmatch

An arrival time mismatch of ΔT translates to two current pulses of width ΔT, height I_p, and opposite polarities that are injected by the charge pump at each phase comparison instant. Owing to the short time scales associated with these pulses, only C2 in Fig. 5.14 acts as a storage element, producing a pulse on the control line. These pulses produce spurs at offset frequencies equal to multiple of the reference frequency [1]. Equation (5.16) shows the amplitude of this spur relative to the carrier.

$$\frac{A_{side}}{A_{carrier}} = \frac{1}{2\pi}\frac{\Delta T I_p}{C_2} T_{res} K_{VCO} \tag{5.16}$$

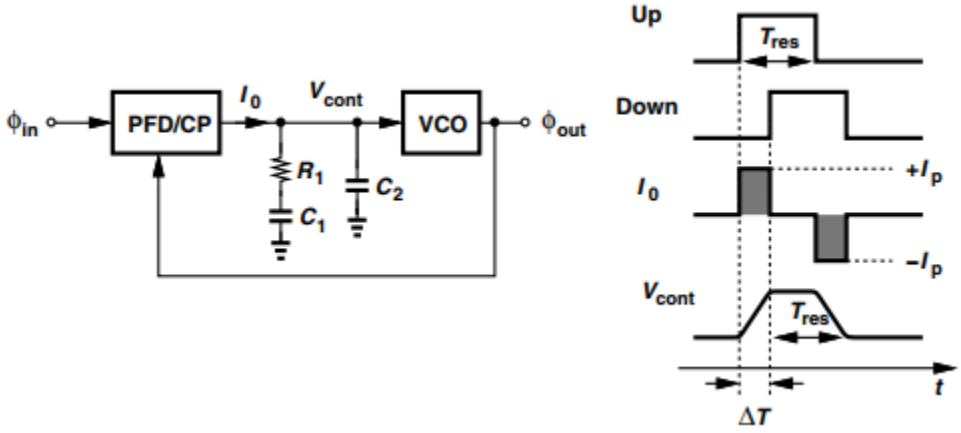


Figure 5.14: Effect of Up and Down skew on V_cont for a second-order filter

5.5.3.4. Literature Review

Various architectures were introduced in the literature as in [4,5]. The PFD in [4] achieves the least dead zone; but requires a duty cycle of 50% which is not guaranteed. The NAND PFD in [5], which is shown in Fig. 5.15, and the Tri-State PFD in [1], which is shown in Fig. 5.16, are other architectures for the PFD. Both architectures can be optimized to reduce the dead zone and can be used at a high frequency. The NAND based PFD can operate at a higher frequency as it does not reset the output through a feedback signal. However, the Tri-State PFD speed was sufficient for our design. The chosen topology is the Tri-State.

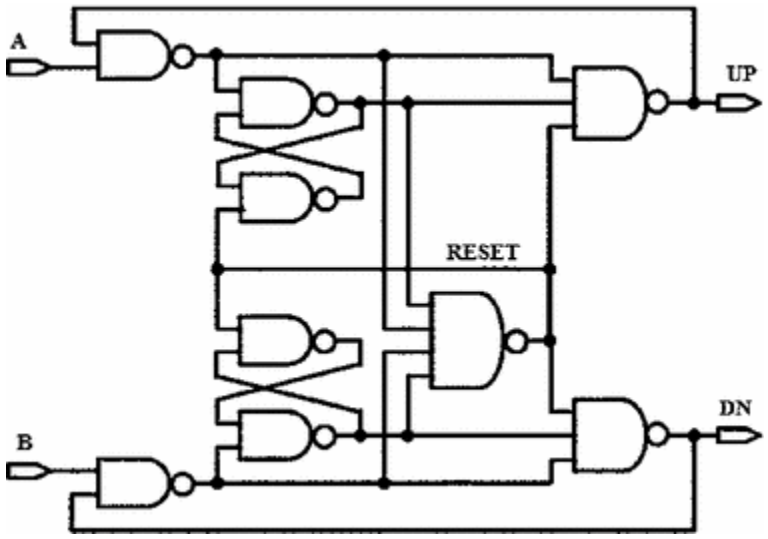


Figure 5.15: NAND PFD

The D flip flop whose input D is connected to VDD can be implemented as a two SR Latch Flip Flop. The Latch is designed using two NOR gates in a positive feedback. The NOR based SR latch is chosen instead of the NAND based SR latch as it is an active high topology making it easier to implement. The logical implementation is shown in Fig. 5.17. Figure 5.18 shows the complete implementation of the PFD. The AND gate is implemented using a NAND gate and an inverter. A symmetric NAND gate as depicted in Fig. 5.19 is used to reduce up/down mismatch [5].The NOR gate is depicted in Fig. 5.20. The up pulse is passed by an inverter to be able to turn on the PMOS switch in the charge pump. The down pulse is passed through a transmission gate such that the up and down pulses almost have the same delay until they reach the charge pump switch. This reduces the clock skew but does not eliminate it completely.

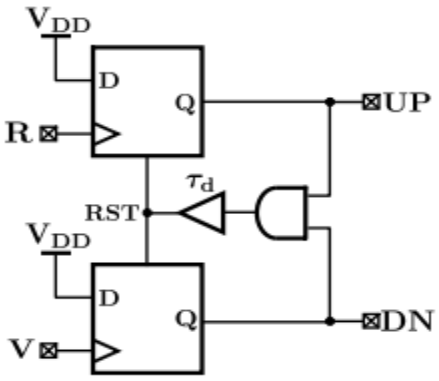


Figure 0.16: Tristate PFD

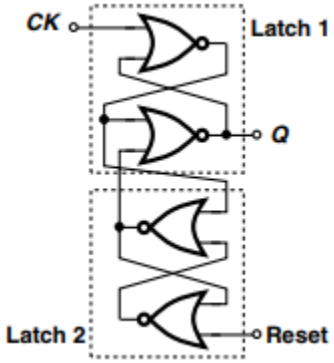


Figure 0.17: SR Latch Logical implementation

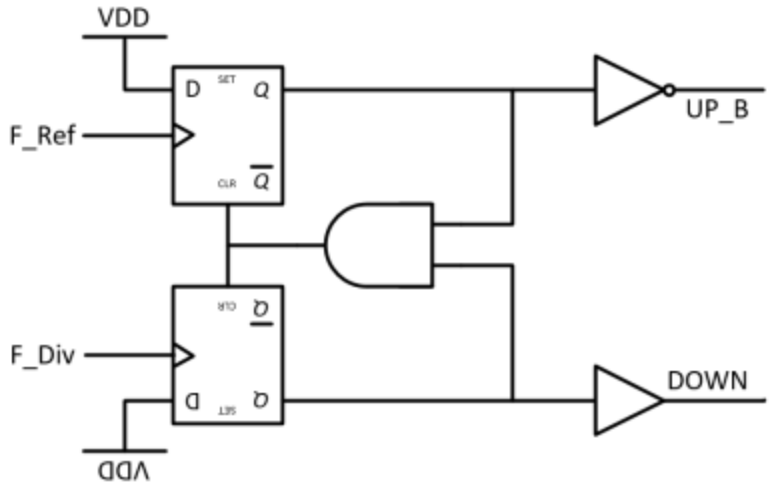


Figure 5.18: Tristate PFD

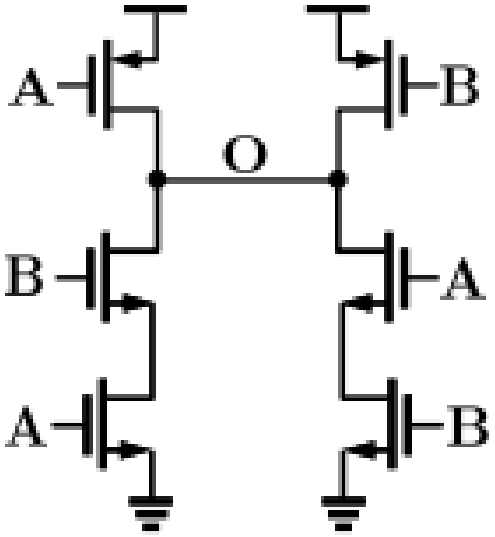


Figure 5.19: Symmetric NAND

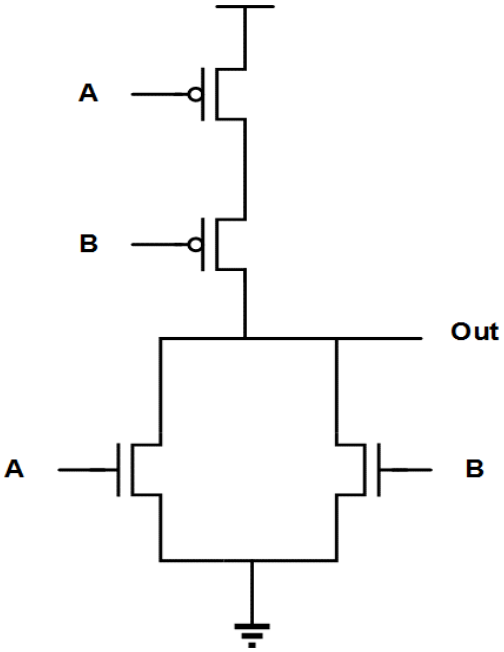


Figure 5.20: NOR gate implementation

5.5.3.5. Simulations Results

The PFD consumes 16.428 μw average power from 1.2 V supply. The output of the PFD is shown in Fig. 5.21 where the feedback frequency is higher than the reference frequency.

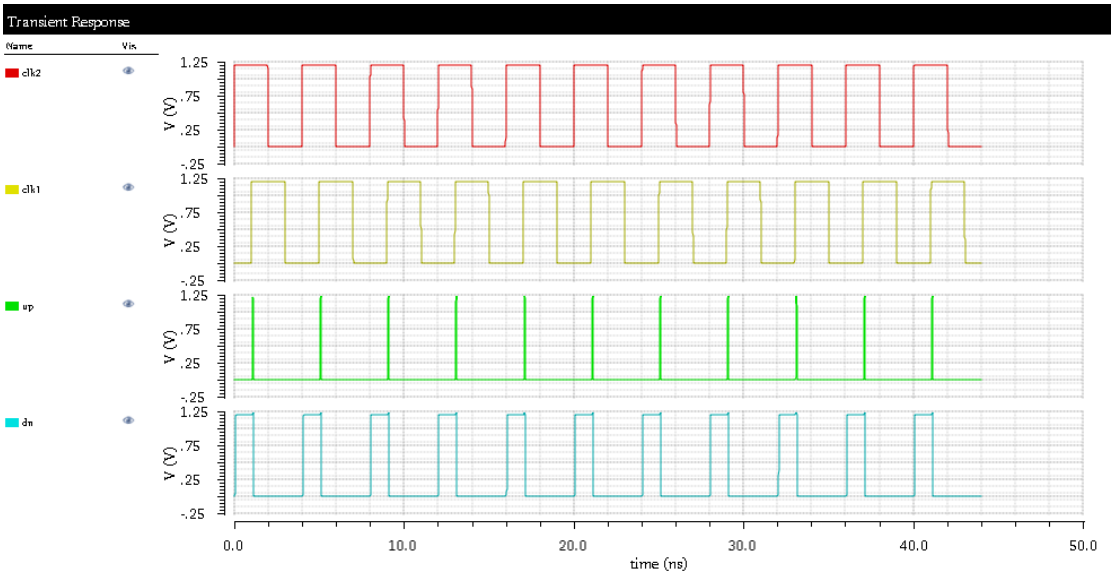


Figure 0.21: PFD Transient Simulation

5.5.4. Charge Pump

The designed CP is presented in this subsection. Similar to the PFD, this subsection starts by listing the non-idealities which evolve with the practical CP implementation compared to the explained theory in section 5.3.2.

5.5.4.1. Charge Sharing

Charge sharing occurs depending on the location of the two switch transistors. This effect can be pronounced in the drain switched charge pump depicted in Fig. 5.22(a).

When both switches S1 and S2 are OFF, Node X is pulled up to VDD while node Y is pulled down to GND and VCONT is a floating node. For non-ideal narrow pulses there is a series of short periods when both switches are on at the same time. This will cause V_x to decrease while V_y increases, this results in a deviation in output VCONT due to charge sharing between C_x,C_y and C_L.

The solution to this problem is to hold $V_x$and $V_y$at constant voltages during switching. The implementation of this solution will be shown in the chosen charge pump topology.

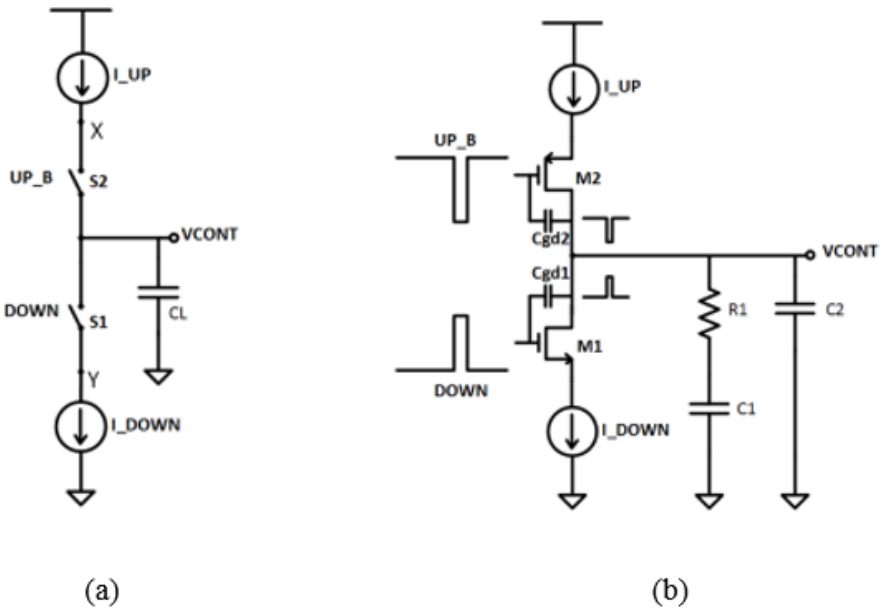


Figure 5.22: (a) Drain Switched Charge Pump (b) Clock feedthrough

5.5.4.2. Charge Injection and Clock Feedthrough

Clock Feedthrough relates to the gate-drain overlap capacitance of the switches. As shown in Fig. 5.22(b), the UP and DOWN pulses couple through Cgd1 and Cgd2 respectively and reach VCONT causing undesired ripples. Proposed solution is to place the switches near supply rails as in source switched topology.

Charge Injection occurs when the switching transistors, M1 and M2, carry a certain amount of mobile charge in their inversion layers when they are on. When the switches turn off, they dispel this charge through their source and drain terminals. If the source of the drain of the switch is next to the output node, it causes undesired ripples, which causes the output to deviate from its desired value.
If the current values IUP and IDN are not exactly same, or there is some delay between the controlled signals UP and DN, then there will be a natural phase error between reference frequency and output frequency of the VCO even if the PLL is in locked state. This is achieved by keeping VCONT constant by keeping $Q_{Charged} = Q_{Discharged}$ at the filter at all times as shown in (5.17).

$$Q_{Charged} = Q_{Discharged} = I_{UP} \times T_{UP} = I_{DOWN} \times T_{DOWN} \quad (5.17)$$

For the designed charge pump, a current mismatch of 5% is required to reduce spurs and phase offset.

5.5.4.3. Proposed Architecture

The proposed charge pump is shown in Fig. 5.23 [6]. The unity gain amplifier is used to keep the current sources' drain voltage constant across switching. Thus, eliminating charge sharing problem. The unity gain buffer is designed as a simple two stage OpAmp placed in a unity feedback. The OpAmp schematic is shown in Fig. 5.24. The OpAmp is designed with a PMOS input stage to reduce the flicker noise at the charge pump output. The OpAmp achieves a DC gain of 61.5 dB. The phase margin for the OpAmp in unity feedback is 64.90^{o}. The OpAmp loop gain and phase is plotted in Fig. 5.25. The switches are designed as transmission gate to reduce the effect of charge injection as the injected carriers find another path which is the second transistor of the transmission gate [7] as in Fig.5.26.

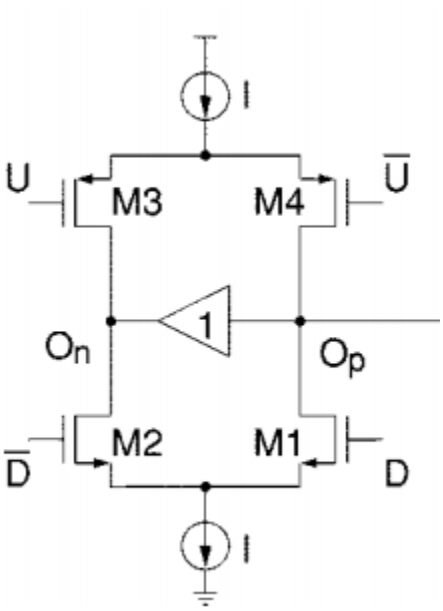


Figure 0.23: Charge pump Design

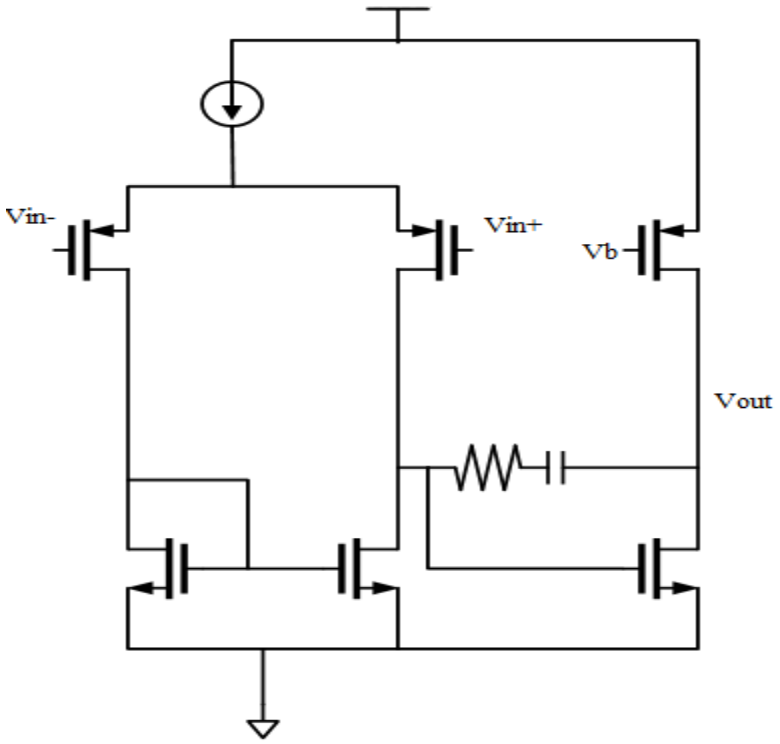


Figure 0.24: OpAmp Schematic

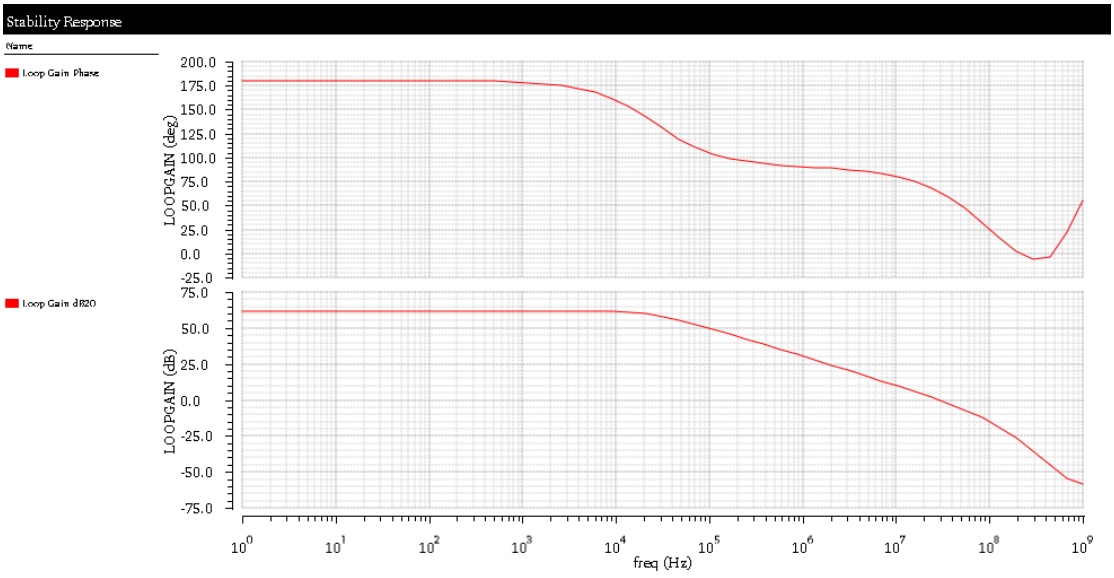


Figure 5.25: Charge Pump OpAmp Loop Gain and Phase

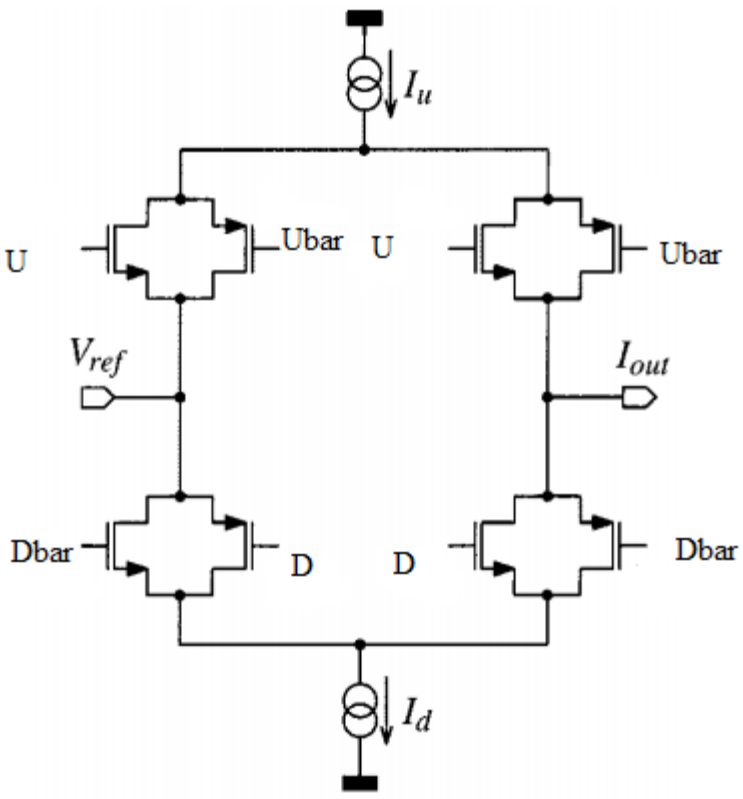


Figure 0.26: Use of Transmission gate switches to reduce Charge Injection

5.5.4.4. Static Simulations Results

With both up and down currents on, a mismatch of +/- 5% could be tolerated across corners to reduce spur level. The simulation results across corners is shown in Fig. 5.27. This small mismatch was achieved by using a cascode current mirror with low threshold voltage transistors. The threshold voltage for these transistors is in the range of 170m. Thus, a compliance voltage of 270 mV is achieved with an overdrive voltage of 50 mV.

The Charge pump noise is the dominant noise contributor inside the loop bandwidth. The output noise current is shown in Fig. 5.28, which achieves a noise of -255 dBA/Hz at an offset of 500 KHz. The output noise current density depicted in Fig. 5.29 shows a variation of 3.5 dBA/Hz in the worst corner which is acceptable as the achieved noise meets the standard's mask spec by a margin.

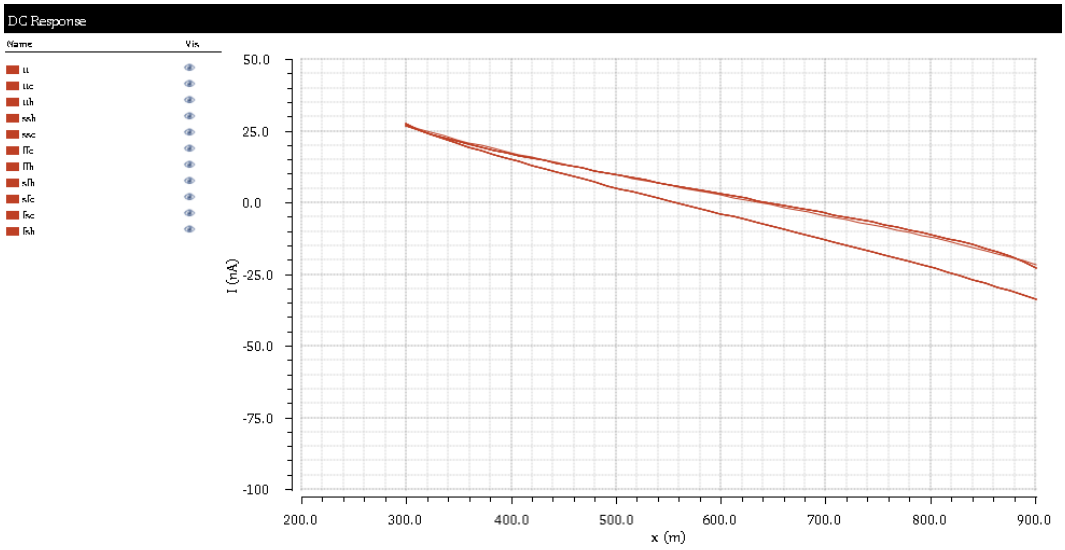


Figure 0.27: Mismatch Currents across Corners

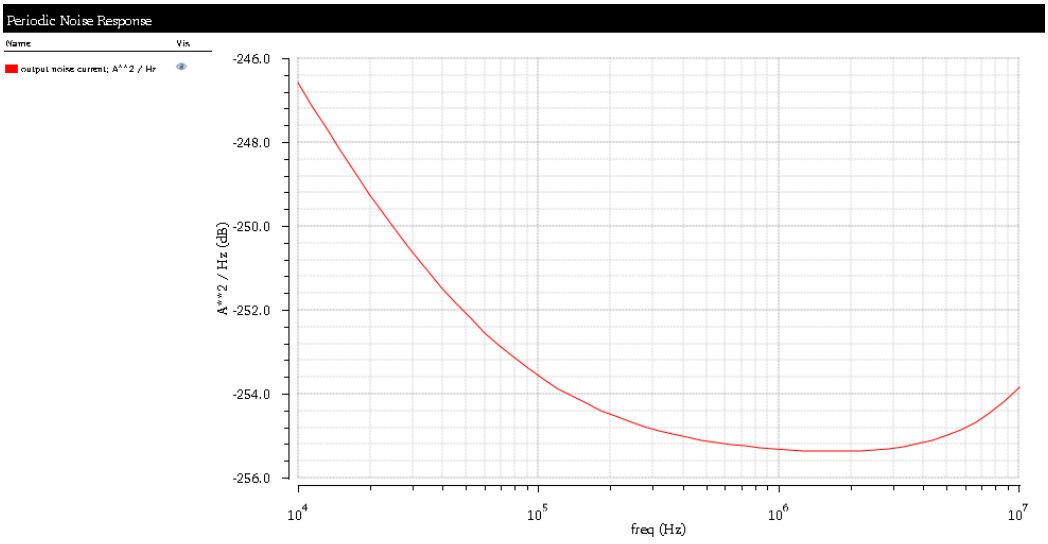


Figure 0.28: Charge Pump Output Noise Current

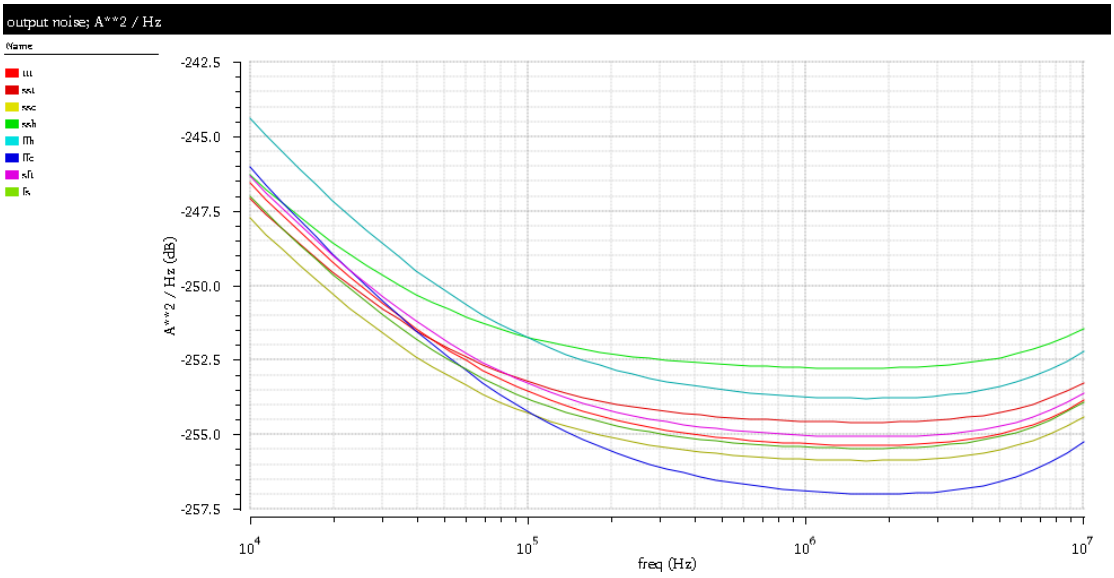


Figure 0.29: Charge Pump Output Noise Current across Corners

5.5.4.5. Dynamic Simulations Results

The dynamic simulations are done to show the PFD/CP characteristic. In the graph shown in Fig. 5.30, a time of 4ns corresponds to a delta phase of 0^o, a time of 0ns corresponds to a delta phase of -2π and a time of 8 ns corresponds to a delta phase of $2\ \pi$. The charge pump at phase offset near 2π and -2π changes its direction as the delay added in the PFD reset path to reduce the charge pump dead zone, causes the characteristic to clip near the end. This effect reduces the PFD/charge pump speed. However, in our case, the speed is sufficient as the added delay is 150 Ps and this translates to a maximum frequency of 3GHz, which is above the reference frequency [8]. Fig. 5.31 shows the derivative for the PFD/CP characteristic to check for the PFD/CP linearity as their linearity affects the PLL spurs.

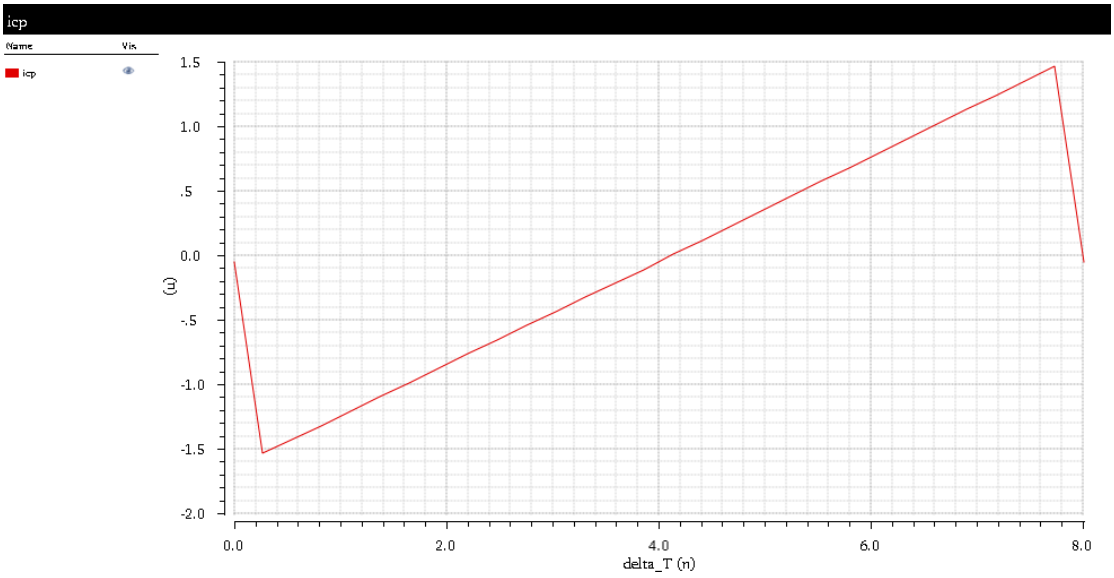


Figure 0.30: PFD/CP Characteristic

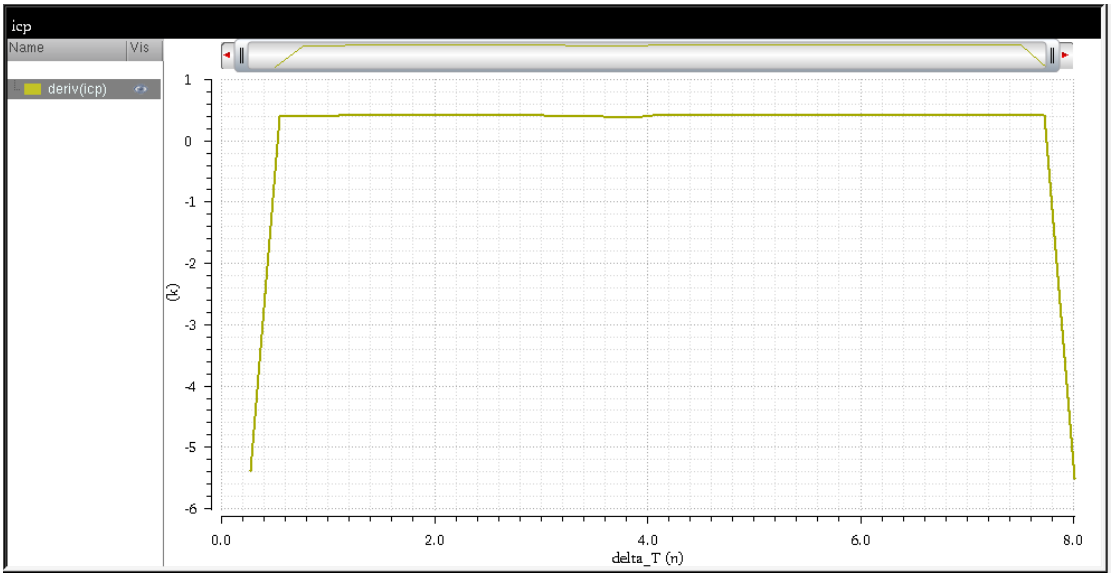


Figure 5.31: shows the PFD/CP Linearity

5.5.5. VCO

Oscillators are considered from the most essential and critical blocks in RF systems either in the receive or the transmit path. Most oscillators must be tuned over a certain frequency range. Hence, an electronically controllable oscillator is very important and of a great use. In this section the fundamental concepts and the results of the designed VCO are illustrated. An oscillator can be viewed from two different point of views: Feedback view and One-Port view. Both will be discussed in details in the following subsections.

5.5.5.1. Performance Parameters

An oscillator used in an RF transceiver must satisfy some of requirements. These requirements are related to required system specifications and interface specifications. In this section, the oscillator performance parameters and their role in the overall system are explained.

Frequency Range: The oscillator should have a continuous tuning range across the band.

Output Voltage Swing: Oscillators must produce sufficiently large output swings to ensure nearly complete switching of the transistors in the subsequent stages.

Phase Noise The spectrum of an oscillator in practice deviates from an impulse and is "broadened" by the noise of its constituent devices. Called "phase noise," this phenomenon increases evm and in band emissions. The required phase noise from the VCO and divider cascade is (-85 dBc at 500KHz offset).

5.5.5.1. Feedback View of Oscillators

An oscillator may be viewed as a "badly-designed" negative-feedback amplifier. This amplifier is designed to have a negative phase margin to start up then its phase margin should be zero to sustain its oscillations. From the poles point of view, the oscillator should have poles in the right half pane (RHP) to start up. As the output increases, the poles should move to the $j\omega$ axis to sustain the oscillator's oscillations [9]. To analyze the oscillator from this point of view, consider the linear negative-feedback system shown in Fig. 5.32 where the equation describing it is (5.18)

$$\frac{Y}{X}(s) = \frac{H(s)}{1+H(s)} \tag{5.18}$$

If at $H(s = j\omega_1)$ becomes equal to -1, the gain from the input to the output goes to infinity, allowing the circuit to amplify a noise component indefinitely. That is, the circuit can sustain an output at ω_1.Since H(s) is a complex function, the condition $H(j\omega_1) = -1$ can equivalently be expressed as

$$|H(s = j\omega_1)| = 1 \tag{5.19}$$

$$Arg\big(H(s = j\omega_1)\big) = 180 \tag{5.20}$$

which are called "Barkhausen's criteria" for oscillation which is a necessary but not sufficient condition for oscillation [1]. For a noise component at ω_1 to "build up" as it circulates around the loop with positive feedback, the loop gain must be at least unity. Fig. 5.33 illustrates the "start-up" of the oscillator. An input at ω_1 propagates through H(s); emerging un-attenuated but inverted. The result is subtracted from the input, yielding a waveform with twice the amplitude. This growth continues with time.

If the magnitude is greater than unity and the phase equals 180 degrees, the oscillations keep growing and only cease due to the circuit non-linearities. This is equivalent to the poles becoming on the jw axis instead of the RHP.

Figure 0.32: Negative feedback system

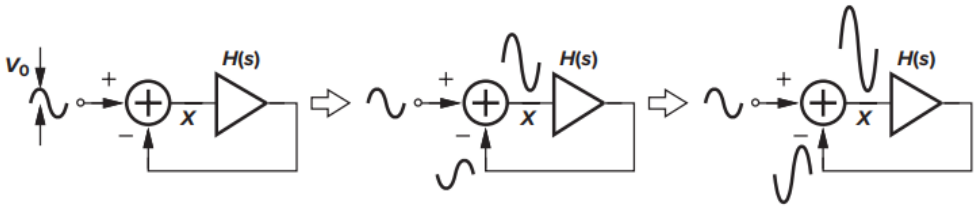


Figure 0.33: Oscillation buildup

5.5.5.2. One-Port View of Oscillators

An alternative perspective views oscillator as two one-port components, namely, lossy resonator and an active circuit that cancels the loss. This perspective provides additional insight. Suppose, as shown in Fig. 5.34(a), a current impulse $I_0\delta(t)$ is applied to a lossless tank. The impulse is entirely absorbed by $C1$ generating a voltage of $I_0 = C_1$. The charge on $C1$ then begins to flow through $L1$; and the output voltage falls. When V_{out} reaches zero, $C1$ carries no energy but $L1$ has a current equal to $L1dV_{out} = dt$ which charges $C1$ in the opposite direction, driving V_{out} towards its negative peak. This periodic exchange of energy between $C1$ and $L1$ continues indefinitely, with an amplitude

given by the strength of the initial impulse. The waveform in Fig. 5.34(a) shows the oscillations after applying the current impulse. The lossy tank depicted in Fig. 5.34(b), behaves similarly except that R_p drains and burns some of the capacitor energy in every cycle, causing an exponential decay in the amplitude. We therefore surmise that, if an active circuit replenishes the energy lost in each period, then the oscillation can be sustained. Fig. 5.35 shows a simple cross coupled oscillator that uses the idea of the one port view of the oscillator. Fig. 5.36(a) shows the series combination of two identical tanks that can be expressed as a single tank. We arrive at the circuit depicted in Fig. 5.36(b). The oscillator can be viewed as a lossy resonator (2L1, C1/2, and $2R_p$) tied to the port of an active circuit (M1, M2, and I_{SS}), expecting that the latter replenishes the energy lost in the former. That is, Z1 must contain a negative resistance. From small signal analysis, Z1 can be proved to be equal to $-2/g_m$. This negative resistance can be used to compensate for $2R_p$ that causes the energy loss in the tank. This result can be used to get the oscillation startup condition. For oscillations to be sustained

$$\frac{-2}{g_m} = 2 * R_p \tag{5.21}$$

Thus,

$$g_m = R_p \tag{5.22}$$

For startup, g_m is greater than $3R_p$ is taken as a rule of thumb to ensure starting.

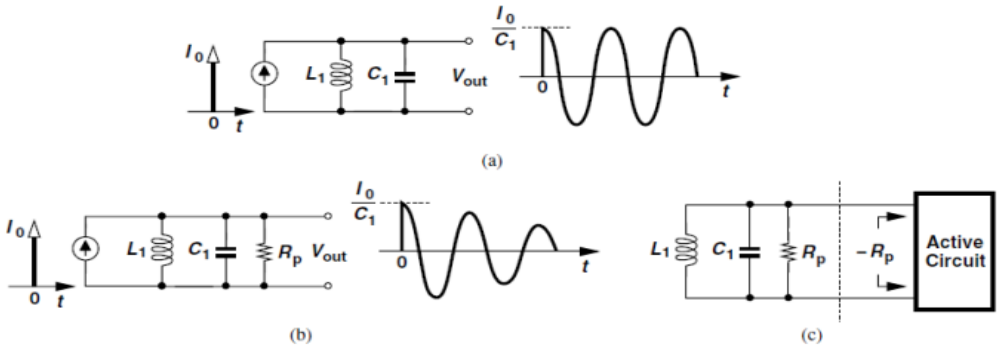


Figure 0.34: (a) Lossless LC Tank (b) Lossy LC Tank (c) LC Tank with Negative Resistance

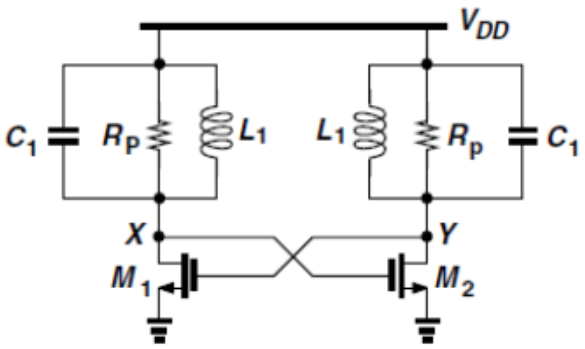


Figure 0.35: Simple Cross-Coupled Oscillator

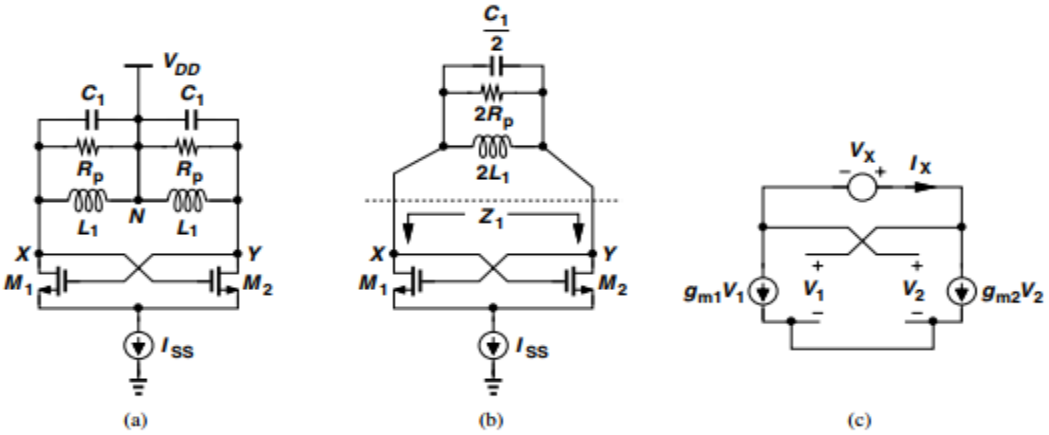


Figure 0.36: (a) Cross coupled oscillator (b) load tanks merged (c) small signal model to get Zin

5.5.5.3. Cross-Coupled VCO Topologies

An LC oscillator is chosen as a result of the low Kvco needed. Ring oscillators are designed for a large Kvco in the range of GHz [10]. Moreover, to achieve the same phase noise, ring oscillators consume much more current [11].

5.5.5.3.1. Tail-Biased VCO

With differential *VX* and *VY*, we surmise that *M*1 and *M*2 can operate as a differential pair if they are tied to a tail current source. Shown in Fig. 6, the resulting circuit is robust and can be viewed as an inductively loaded differential pair with positive feedback. The oscillation amplitude grows until the pair experiences saturation, i.e., *M*1 and *M*2 enters triode region. M_{V1} and M_{V2} are the varactors. The varactors are implemented as accumulation mode MOS varactors [12]. The cross-coupled VCO of Fig. 5.37 suffers from a narrow tuning range. *C*1 - denotes parasitic capacitance - tend to limit the effect of the varactor capacitance variation. The tuning range obtained from the C-V characteristic depicted in Fig. 5.38 may prove prohibitively narrow, as he capacitance range corresponding to *negative VGS* (for *Vcont* > *VDD*) remains unused. It is better to work in the linear region where Vgs ranges from -VDD/2 to VDD/2 to have better tuning range and linearity.

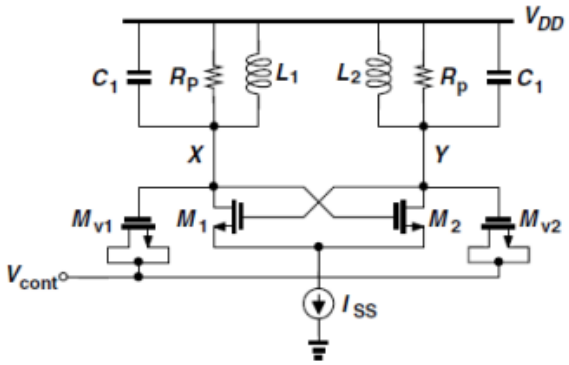


Figure 0.37: Tail biased VCO

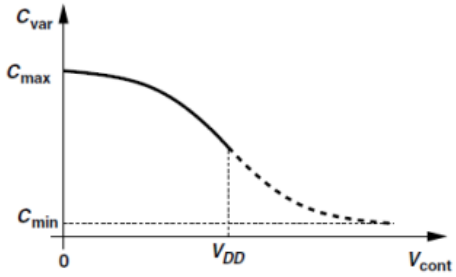


Figure 0.38: Varactor C-V characteristics of a tail biased VCO

5.5.5.3.2. Top-Biased VCO

Figure 5.39 shows top-biased topology. Unlike the tail-biased configuration, this circuit defines the bias currents of *M*1 and *M*2 by a top current source, I_{DD}. The common mode (CM) level is simply given by the gate-source voltage of the transistors carrying a current of $I_{DD}/2$. This CM level can be chosen to be equal to $V_{DD}/2$ such that as *Vcont* varies from 0 to V_{DD}; the gate-source voltage of the varactors, *VGS of the varactor;* goes from $+V_{DD}=2$ to $-V_{DD}=2$; sweeping almost the entire capacitance range from *Cmin* to *Cmax.* The circuit producing *Vcont* (the charge pump) can handle only the voltage range from *V*1 to *V*2; yielding a capacitance range from *C1* to *C2* as shown in Fig. 5.40. While providing a wider range than its tail-biased counterpart, the topology of Fig. 5.39 suffers from a higher phase noise as the current source directly modulates the varactors. This effect does not occur in the tail-biased oscillator because the output CM level is "pinned" at V_{DD} by the low dc resistance of the inductors. Another drawback is that the CM level is determined by device parameters that vary greatly with process and temperature.

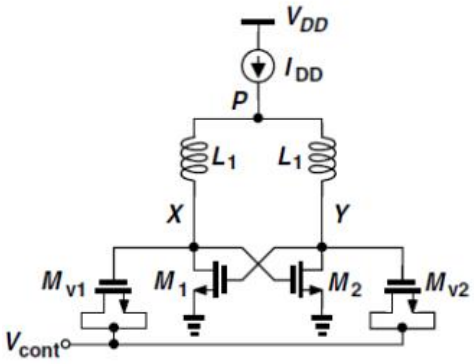


Figure 0.39: Top Biased VCO

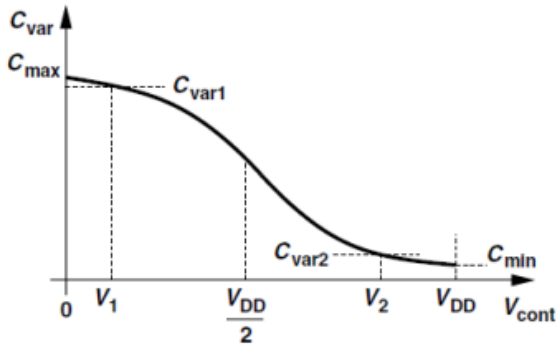


Figure 0.40: Varactor C-V Characteristic of a Top Biased VCO

5.5.5.3.3. CMOS VCO

Another VCO topology that naturally provides an output CM level approximately equal to VDD/2 is shown in Fig. 5.41. The circuit can be viewed as two back-to-back CMOS inverters, except that the sources of the NMOS devices are tied to a tail current, or as a cross-coupled NMOS pair and a cross-coupled PMOS pair sharing the same bias current. Proper choice of device dimensions and I_{SS} can yield a CM level at X and Y around $V_{DD}/2$; thereby maximizing the tuning range.

In this circuit, the bias current is "reused" by the PMOS devices, providing a higher transconductance. Nevertheless, a more important advantage of the above topology over tail biased and top-biased is that it produces twice the voltage swing for a given bias current and inductor design. Thus, achieving a better phase noise for the same bias current. The circuit of Fig. 5.41 nonetheless suffers from two drawbacks. First, for |VGS3| + VGS1 + VI_{SS} to be equal $_{VDD}$; the PMOS transistors must typically be quite wide, contributing significant capacitance and limiting the tuning range. Using a smaller inductor, means a lower resistance and a higher gm needed for startup. Second, as in the top-biased topology, the noise current of the bias current source modulates the output CM level and hence the capacitance of the varactors, producing phase noise [1].

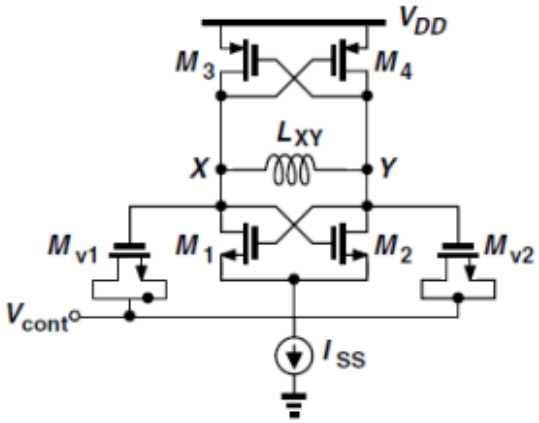


Figure 0.41: CMOS VCO

5.5.5.4. Chosen VCO Topology

The chosen topology is shown in Fig. 5.42 [1]. The tail biased nmos VCO was used but with some modifications. Coupling capacitors Cs1 are used to bias the gate of the varactors at $V_{DD}/2$ to gain maximum tuning range. *CS*1 and *CS*2 must be *much greater* than the maximum capacitance of the varactors; *Cmax;* so that the capacitance range presented by the varactors to the tanks does not shrink substantially. *CS* = 10*Cmax* reduces the capacitance range by 10% [1]. The coupling capacitor was chosen 10 times Cmax in our design. As *R*1 and *R*2 appear approximately in parallel with the tanks, their value must be chosen much greater than *the inductor resistance (Rp) as* a tenfold ratio lowers the Q by about 10%. Second, noise on the mid-supply bias; *Vb;* directly modulates the varactors and must therefore be minimized. Rb was chosen 10 times Rp. Simulations are done to see if this value meets the phase noise requirement. In applications where a wider tuning range is necessary, discrete tuning may be added to the VCO so as to achieve a capacitance range well beyond Cmax/Cmin of varactors. Illustrated in Fig. 5.43(a), the idea is to place a bank of small capacitors, each having a value of C_U; in parallel with the tanks and switch them in or out to adjust the resonance frequency. Vcont can also be viewed as a "fine control" and the digital input to the capacitor bank as a "coarse control." Figure 5.43(b) shows the tuning behavior of the VCO as a function of both controls. The fine control provides a continuous but narrow range, whereas the coarse control shifts the continuous characteristic up or down. The capacitor bank will be connected to nodes x,y in Fig. 5.42.
One disadvantage is that the on-resistance; Ron; of the switches that control the unit capacitors degrades the Q of the tank and degrades phase noise. Using wider switches introduce a larger capacitance from the bottom plate of the unit capacitors to ground, thereby presenting a substantial capacitance to the tanks when the switches are off.

The solution is to place the main switch; S1; between nodes A and B as in Fig. 5.44 (a) so that, with differential swings at these nodes, only half of Ron appears in series with each unit capacitor [Fig. 5.44 (b)].

This allows a twofold reduction in the switch width for a given resistance. Switches S2 and S3 are minimum size devices, merely defining the CM level of A and B. The discrete tuning curves in Fig. 5.44(b) should have an overlap between them such that the VCO does not fail to cover the range between each two consecutive curves. The number of curves is chosen such that the VCO operates in the linear region of Kvco.

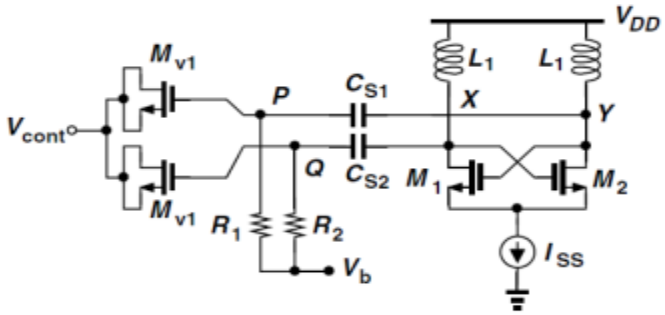


Figure 5.42: Tail biased VCO with coupling capacitors

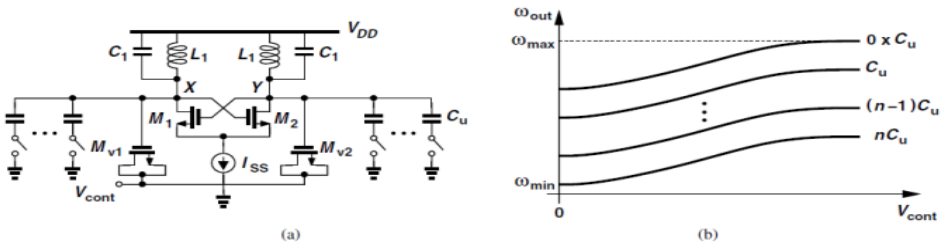


Figure 0.43: (a) Discrete tuning using switched capacitors (b) Discrete Curves

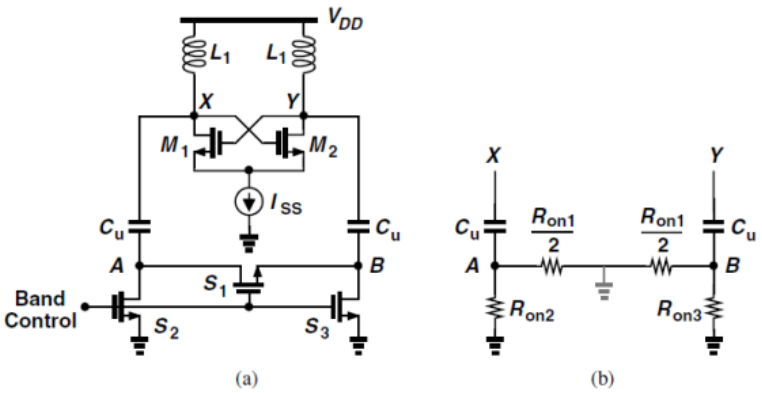


Figure 0.44: (a) Use of a floating switch (b) Equivalent Circuit of the switch

5.5.5.5. Phase Noise

An ideal oscillator produces a perfectly-periodic output of the form $x(t) = A\cos(\omega_c t)$. The zero crossings occur at exact integer multiples of Tc = $2\pi/\omega_c$. In reality, however, the noise of the oscillator devices randomly perturbs the zero crossings as shown in Fig. 5.45. Consequently, the impulse is "broadened" to represent this perturbation in the zero crossings as depicted in Fig. 5.46 (b).
Various models were done for the phase noise to expect the effect of the noise of each device on the output phase noise such as Leeson's model [13] or the model introduced by Hajimiri [14]. The equation used in our analysis is (5.23) as it gives insight to the designing process [1].

$$S(\Delta w) = \frac{\pi^2}{Rp}\frac{KT}{Iss^2}\left(1 + \frac{3\Upsilon}{8}\right)\frac{wo^2}{4Q\Delta w^2} \qquad (5.23)$$

The equation shows to get a better phase noise, one can increase the quality factor of the tank or increase I_{SS}. However, increasing Iss does not always decrease the phase noise. This can be seen as I_{SS} increases, M1 and M2 of Fig. 5.42 enter the triode region and can be modelled as a resistance. At this case, the tail current capacitance becomes in series with the resistance of M1 and M2 and degrade the quality factor of the tank. This explanation is depicted in Fig 5.47. The vco bias current is designed for minimum gm to decrease its phase noise contribution [1].

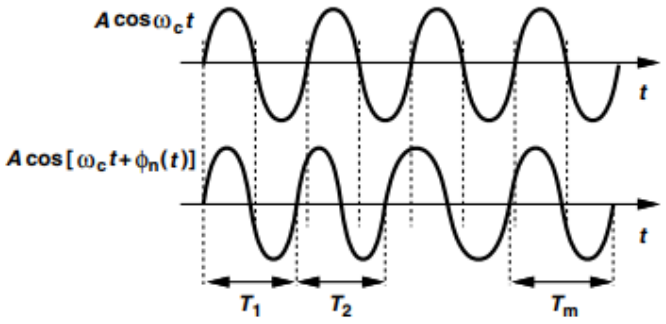


Figure 5.45: Effect of Phase Noise on Oscillators

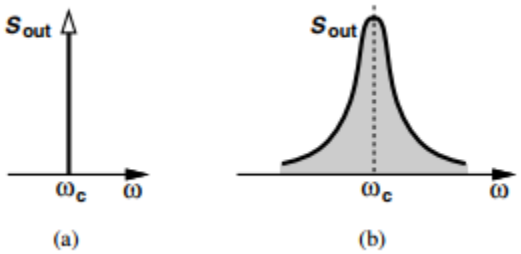


Figure 5.46: (a) Ideal Spectrum (b) Spectrum of a Noisy Oscillator

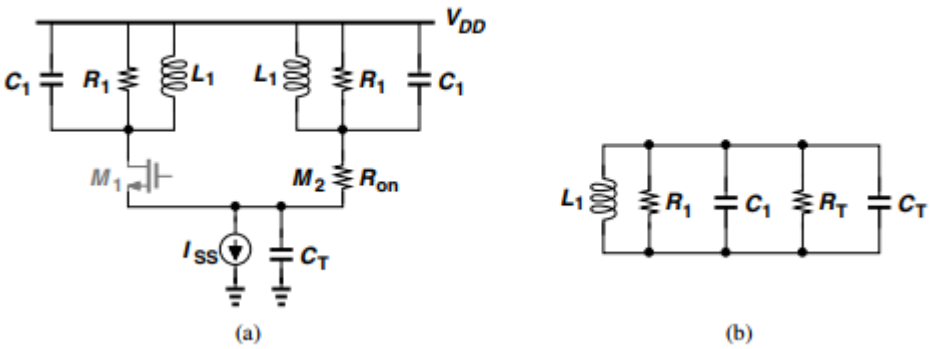


Figure 5.47: (a) One transistor going into triode region (b) Equivalent tank circuit

5.5.5.6. Design Equation

The VCO is made to run at twice the carrier frequency to reduce pulling by the PA. The VCO should have a frequency range from 3420 MHz to 3570 MHz and taking a margin to cover process variations the range becomes from 3170 MHz to 3820 MHz. L1 has a parallel resistance due to the finite Q of the inductor. For a Q of 15.272 and an inductance of 2.35nH from the design kit:

$$R_p = QL\omega_o \quad (5.24)$$

where ω_o is 2π*3.48 Grad/s. This translates to Rp of 797 ohm. For sustained oscillations, gm of M1 and M2 should be 1.2 mS and to maintain oscillations a margin three times this value is chosen. Cs is chosen equal to 10 times C_{varmax} I_{SS} is chosen to get the required swing of 700 mV pp from the relation in (5.25):

$$V_{PP} = (4/\pi) * I_{ss} * R \quad (5.25)$$

The output swing changes with frequency and process corners as evident from (5.24) that Rp depends on frequency and the inductor value. The VCO is designed to deliver the required swing for the divider at the worst corner. Since the oscillation frequency of the circuit is:

$$f_{osc} = \frac{1}{\sqrt{(LC)}} \quad (5.26)$$

From f_{max} = 3.82 GHz, we get C_{min} and from f_{min} = 3.17 GHz, we get C_{max}.Considering parasitics, we can write $C_{min}$and C_{max}as following:

$$C_{min} = C_{parMOS} + C_{parInd} + C_{varMin} + C_{Load} + C_{Const} \quad (5.27)$$

$$C_{max} = C_{parMOS} + C_{parInd} + C_{varMax} + C_{Load} + C_{Const} + C_{CapBank} \quad (5.28)$$

These values are used to achieve a kvco of 250 MHz.

Inductor parasitic capacitance is calculated from the inductor self resonance. The cap bank is a 4 bit binary weighted cap bank with Cu, 2Cu and 4Cu and 8 Cu.

The number of calibration bits is chosen such that the variation in Kvco is reduced across the frequency range. Thus, keeping the PLL dynamics constant. Using a MATLAB code in appendix 3 that shows the variation in the phase margin and the loop bandwidth that corresponds to Kvco variation.

5.5.5.7. Simulation Results

The transient simulation with the startup is shown in Fig. 5.48. The transient Simulation across corners is depicted in Fig. 5.49. The swing at the worst corner is 750 mVpp to ensure the correct operation of the divider across all corners. The swing varies as Rp varies with the inductor value and the operating frequency.

This effect can be deduced from (5.24), (5.25) and (5.26). The variation of Kvco with the control voltage is depicted in Fig. 5.50. Figures 5.51,5.52 and 5.53 show the 16 curves at different corners. At the fast corner, the curves shift upwards as the capacitance and inductance decrease. At the slow corner, the curves shift downward. Figure 5.54 shows the phase noise of the vco which achieves a phase noise of -109 dBc at an offset of 1MHz which achieves the system spec. Figure 5.55 shows the phase noise across corners. The simulations show a variation of only 1dB across corners. The vco consumes 2.2 mW from 1.2V supply.

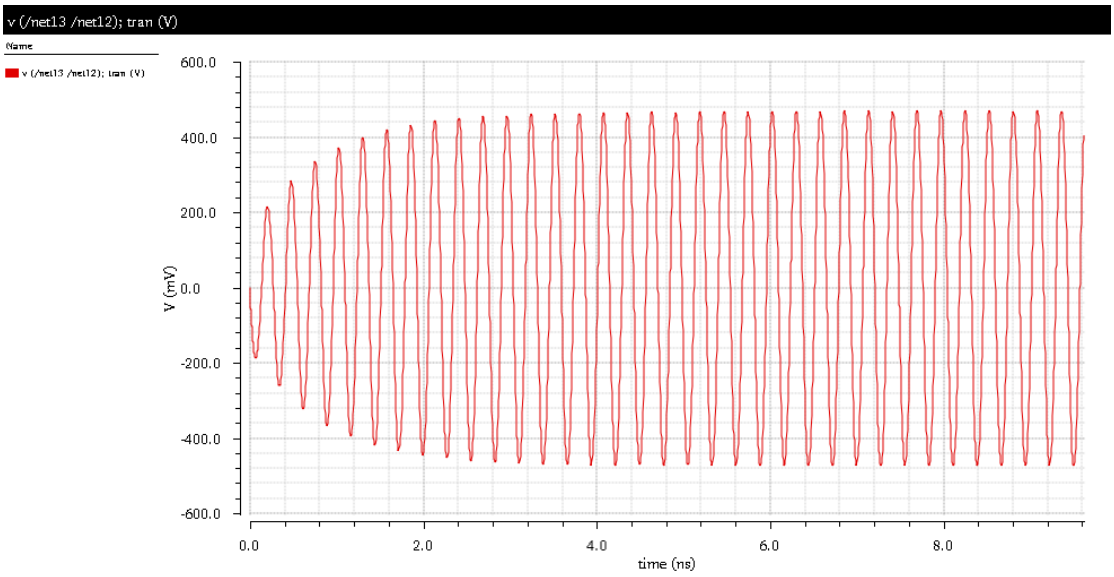


Figure 5.48: VCO Transient Simulation

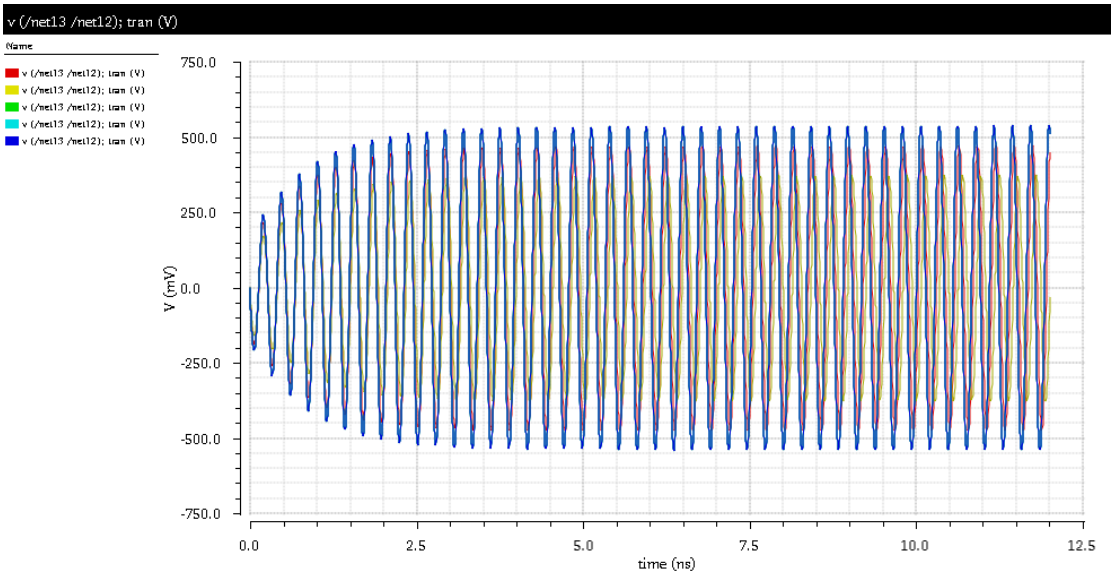


Figure 0.49: Transient Simulation across Corners

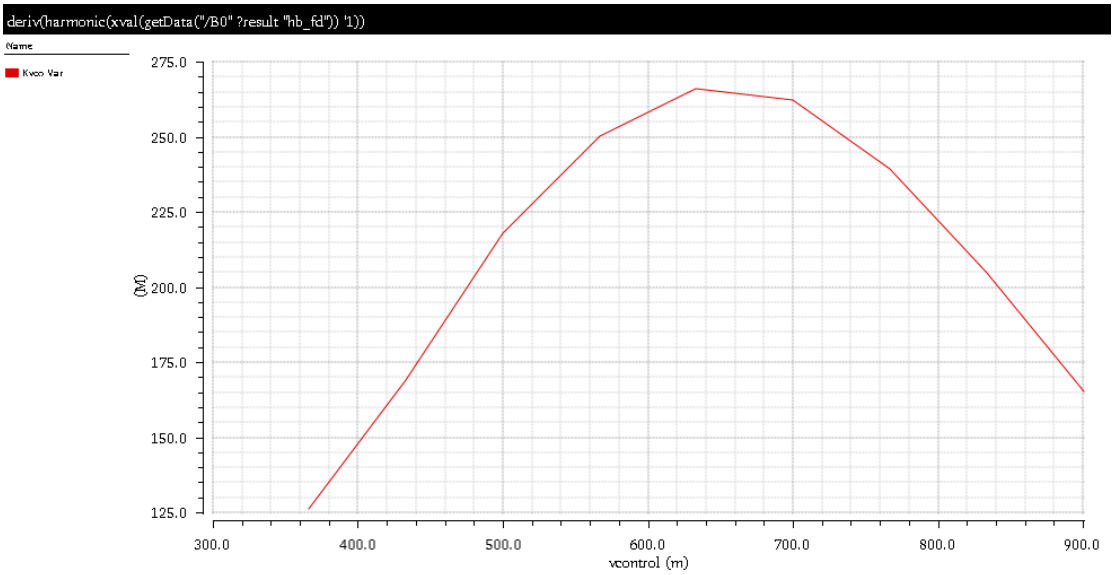


Figure 0.50: Kvco Variation across Vcontrol

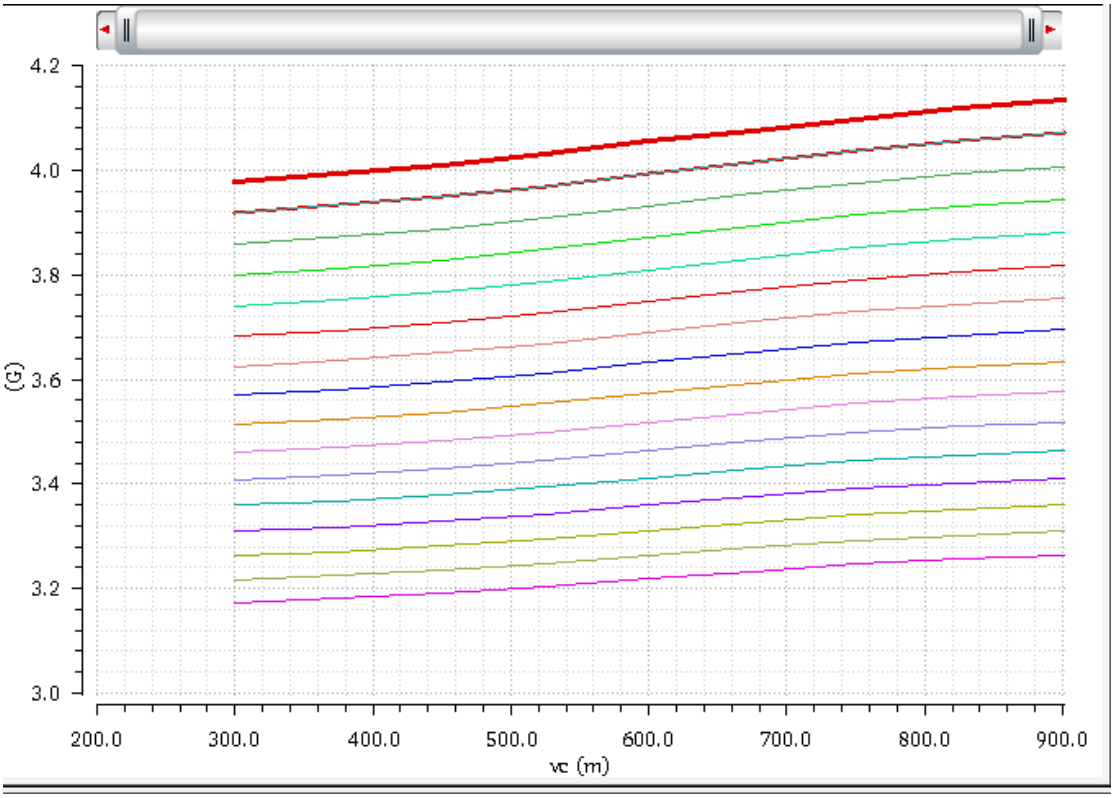


Figure 0.51: VCO curves at TT corner

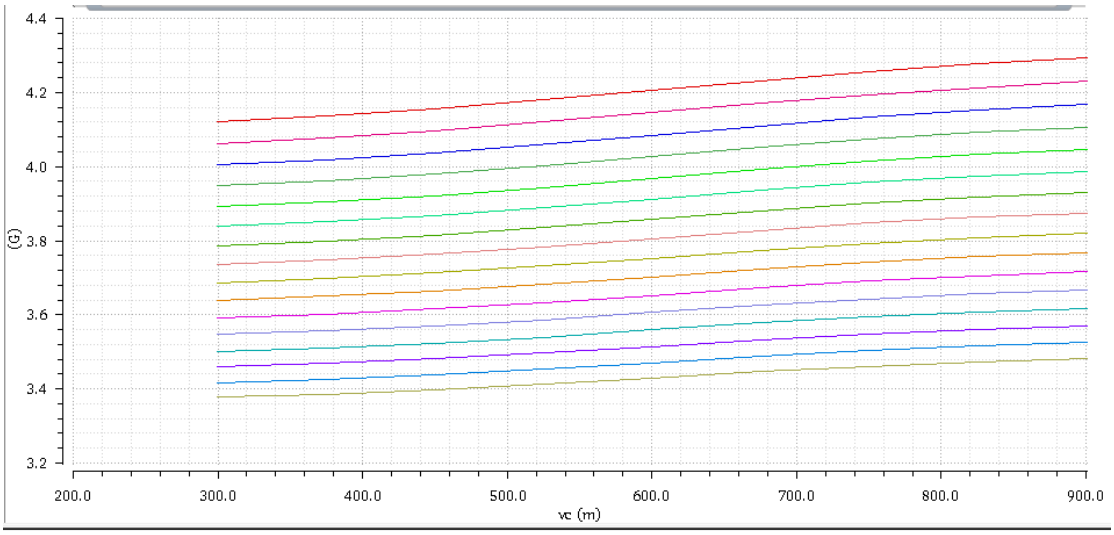


Figure 0.52: VCO curves at FF corner

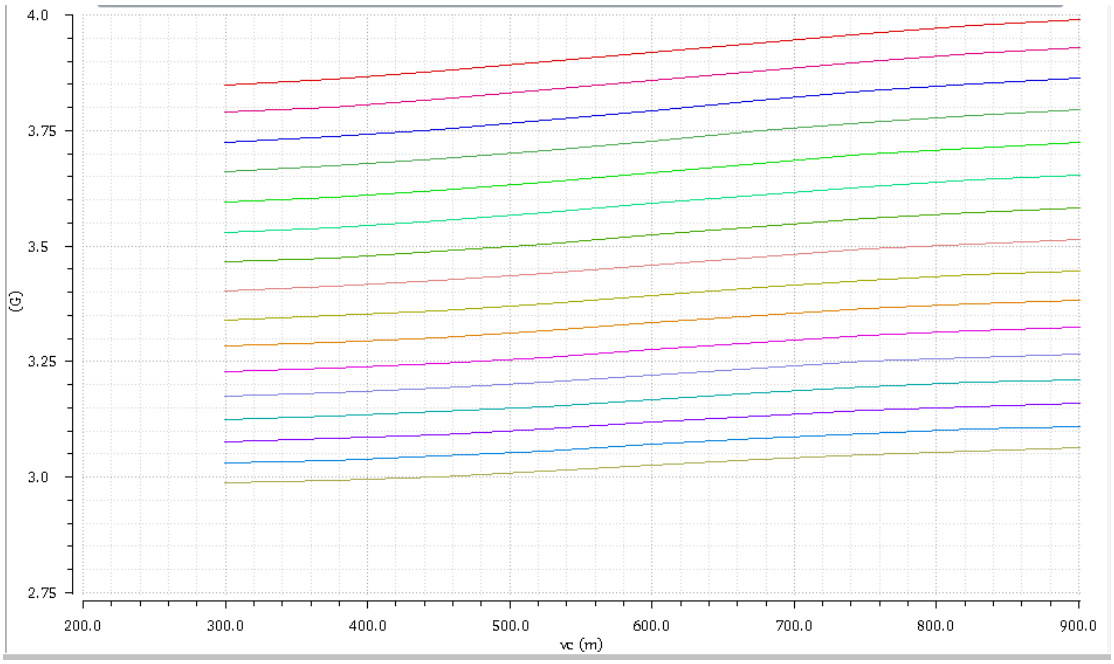


Figure 5.53: VCO curves at SS corner

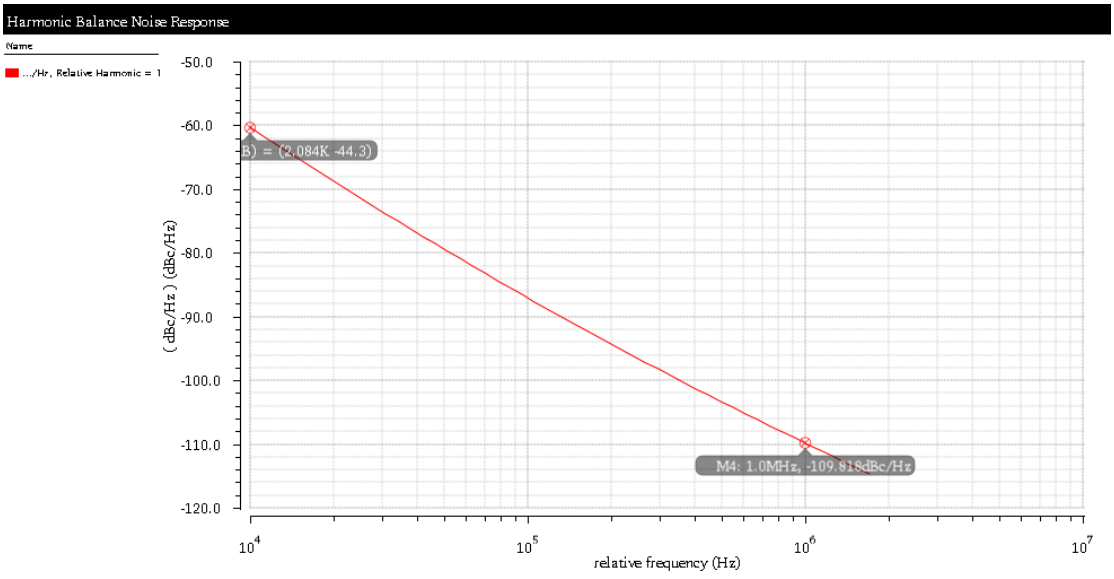


Figure 0.54: VCO Phase noise

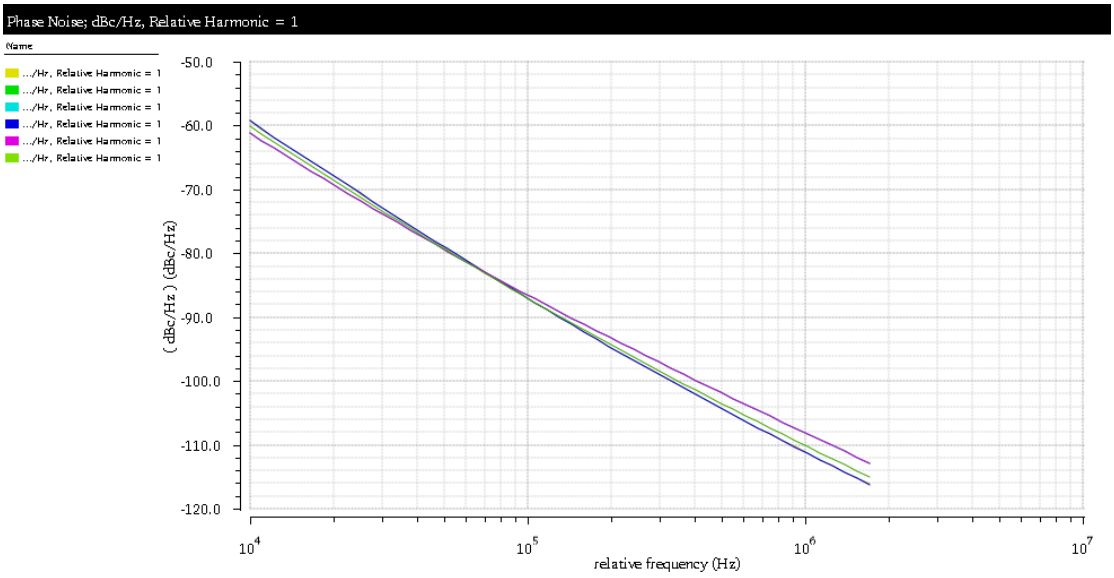


Figure 0.55: VCO Phase Noise across Corners

5.5.6. Current Mode Logic (CML) Divider

Affording the fastest circuits, current-steering logic, also known as current-mode logic (CML), operates with moderate input and output swings. CML circuits provide differential outputs necessary to drive differential circuits. CML derives its speed from the property that a differential pair can be rapidly enabled and disabled through its tail current source. Figure 5.56 shows a CML latch topology [1]. In figure 5.57, there is a regenerative pair in the right branch. It is a positive feedback for memory operation. This CML latch has two modes of operation: Follow Mode, as current is directed through the differential amplifier that passes input signal, and Hold Mode, where current is shifted to cross-coupled pair. The divide by 2 circuit is based on the Johnson counter which are two latches connected in a negative feedback as shown in Fig. 5.57. Figure 5.58 shows the operation of CML divider when the clock is high [15]. For the left latch, current is directed into differential amplifier portion of latch, and the latch output follows input from right latch. For the right latch, current is directed into cross-coupled pair portion of latch, and output is held by this regenerative pair. When clock becomes low, the operation is exchanged between the two latches; Left latch output voltage is held, while the right latch output follows the left latch input. When clock becomes high again, same process repeats on left side noting the voltage polarity is now flipped. There are different modifications in CML dividers. For example, PMOS loads may replace resistors at the expense of speed, as the triode transistors add its parasitic capacitance and resistance that result in slower operation. Diode connected PMOS can also replace resistor load, but at the expense of large headroom $|V_{TP} + V_{OD}|$ and linearity. There are other different topologies that produce rail-to-rail outputs, which cannot be used in our case, as this divider's output will be fed to a mixer.

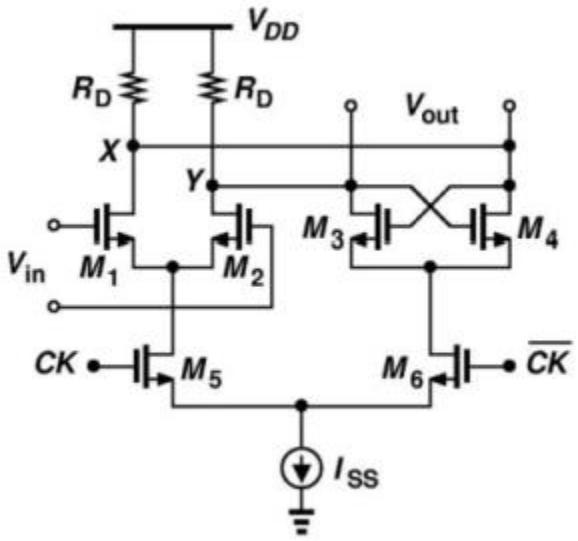


Figure 0.56: CML Latch

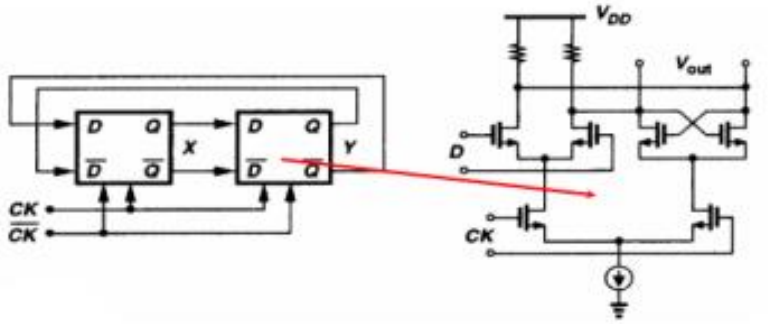


Figure 5.57: Divider based on CML Latch

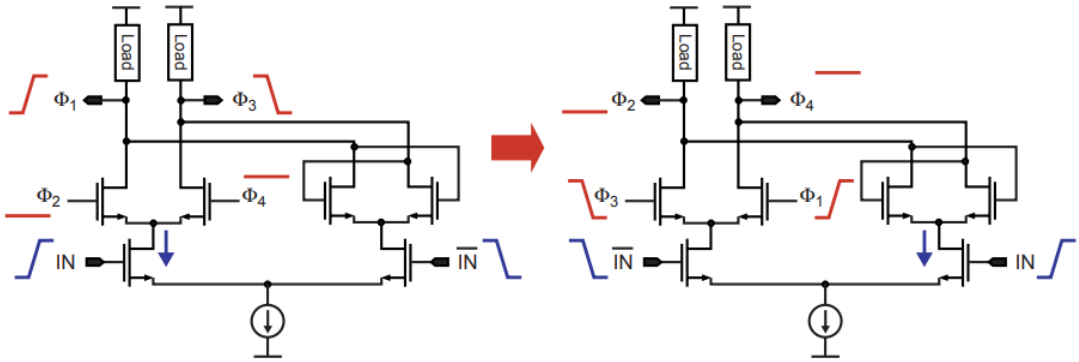


Figure 5.58: CML Divider operation

5.5.6.1. Performance Parameters

Self-Oscillation Frequency: The frequency at which the divider would oscillate with zero swing clock input. The self-oscillation frequency should be in the range of the output range where it should divide correctly which is from 1.7 GHz to 1.9 GHz.

Frequency range: The range of frequencies where the divider should divide correctly by 2. The input frequency ranges from 3.4 Ghz to 3.8 GHz.

The complete divider schematic is shown in Fig. 5.59.

The output common mode was chosen to be 800mV. The output common mode is equal to Vdd – 0.5*IR, where Vdd is equal to 1.2 and R is equal to 800 ohm and Iss is chosen 1.1 mA. The sizing of M5 and M6 is chosen large enough to ensure complete current steering with the input swing coming from the vco. The sizing of M1,M2 is chosen to steer completely the current in M5. The sizing depends on the output required swing as it will be fed back to the latch which is equal to 400mV. M3 and M4 are two cross coupled transistors acting as regenerative transistors to get the required swing at the output nodes. The sizing is done such that gm of M3 and M4 is greater than 1/RD to build up small swings at the output.

In designing the divider, the divider acts as a ring oscillator and it should oscillate at frequency close to the frequency range at the divider output. The self-oscillation is shown in Fig. 5.63. The Divider self oscillates at 1.982 GHz. The divider should be able to divide correctly for an input frequency from 3.42 GHZ to 3.57 GHZ in the locked condition to cover the band of interest.

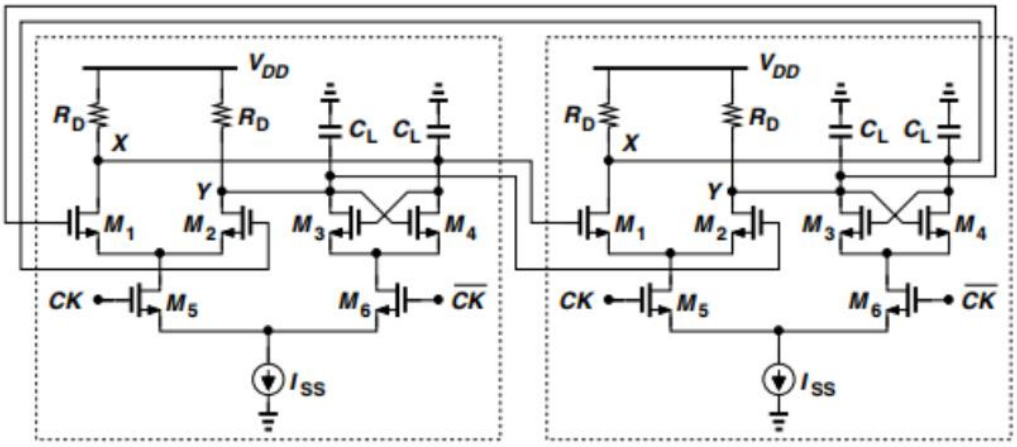


Figure 0.59: Complete CML Divider

5.5.6.2. Simulations Results

The output transient of the divider for a 3.5GHz input is shown in Fig. 5.60. The transient output across corners is depicted in Fig. 5.61. Figure 5.62 shows the output frequency across corners which is 1.75G across all corners. The self-oscillation of the divider at zero differential input is shown in Fig. 5.63. The output of the divider for an input frequency of 3.4 GHz is shown in Fig. 5.64.
The output of the divider for an input of 3.6 GHz is shown in Fig. 5.65. When the loop starts, the VCO starts from the mid curve in the calibration process until the loop settles at the middle of one of the vco curves . The divider consumes 2.64 mW.

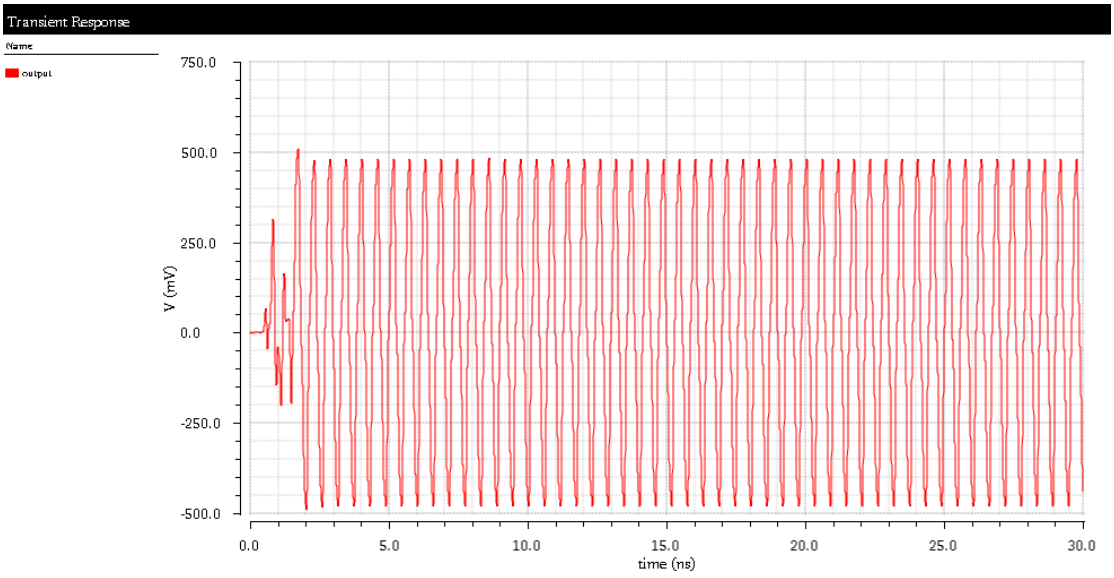


Figure 0.60: Divider Output for 3.5 GHz Input

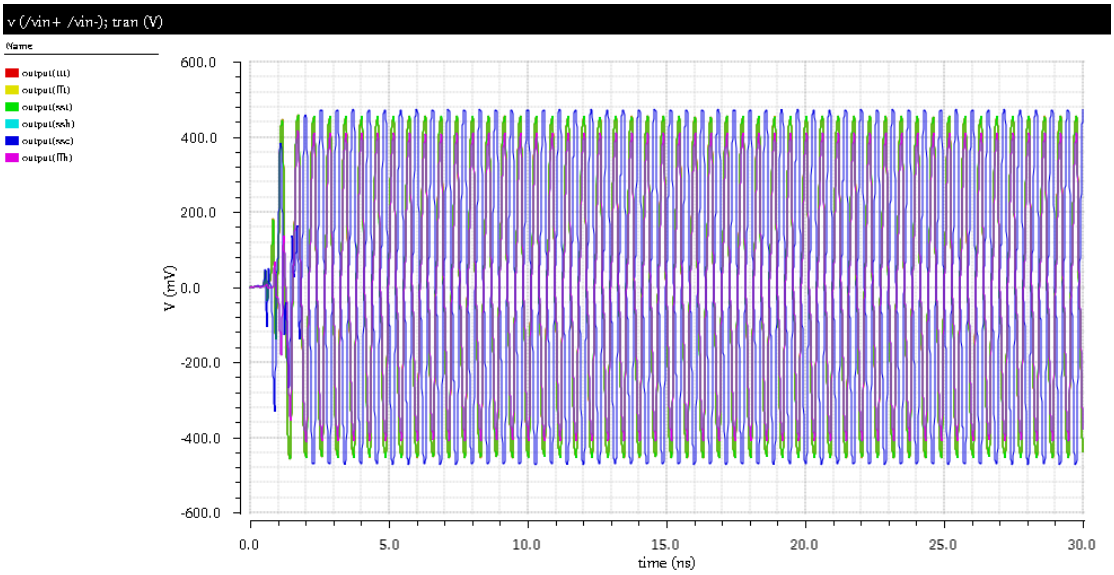


Figure 0.61: Divider Output for 3.5 GHz Input across Corners

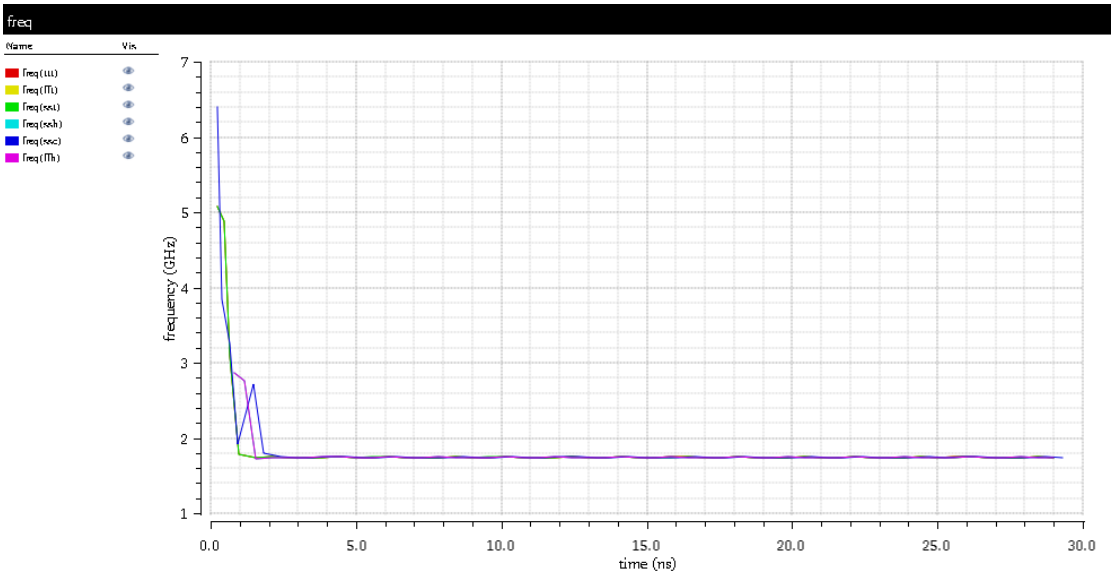


Figure 0.62: Divider Output Frequency for 3.5 GHz Input

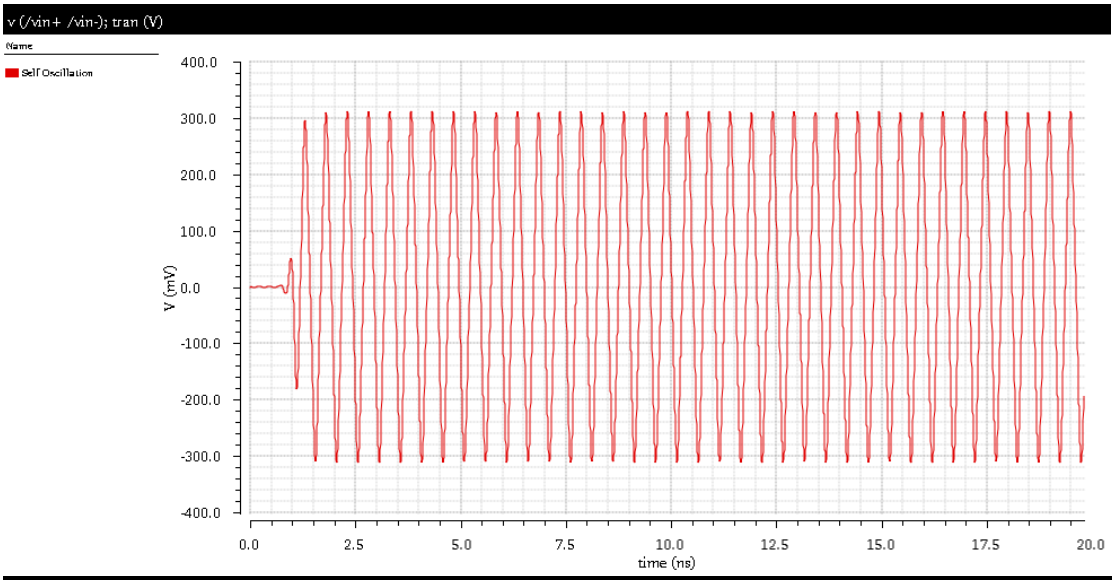


Figure 0.63: Divider Self Oscillation

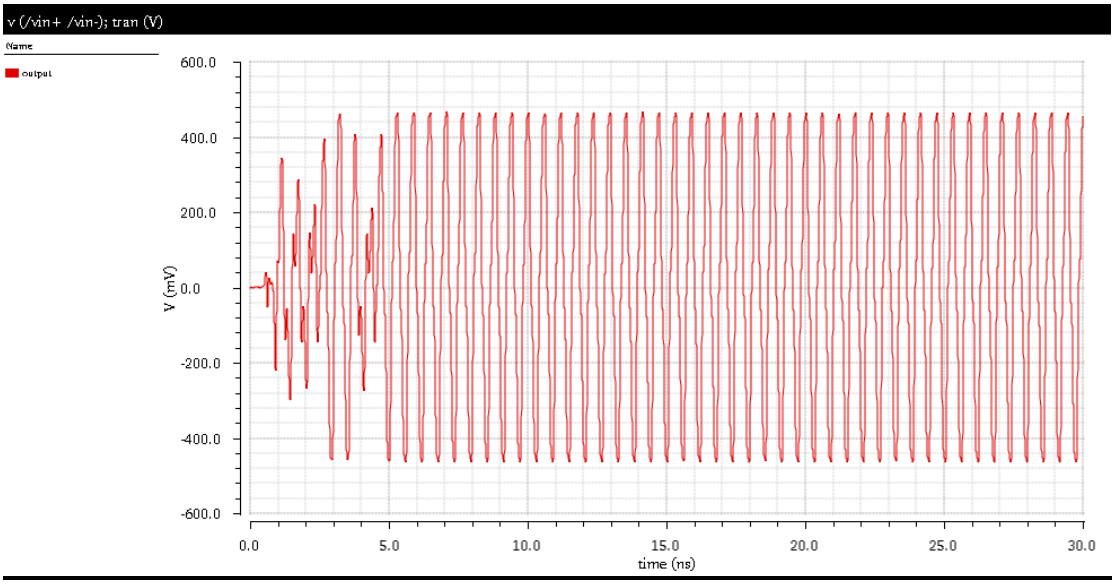


Figure 0.64: Divider Output at 3.4 GHz Input

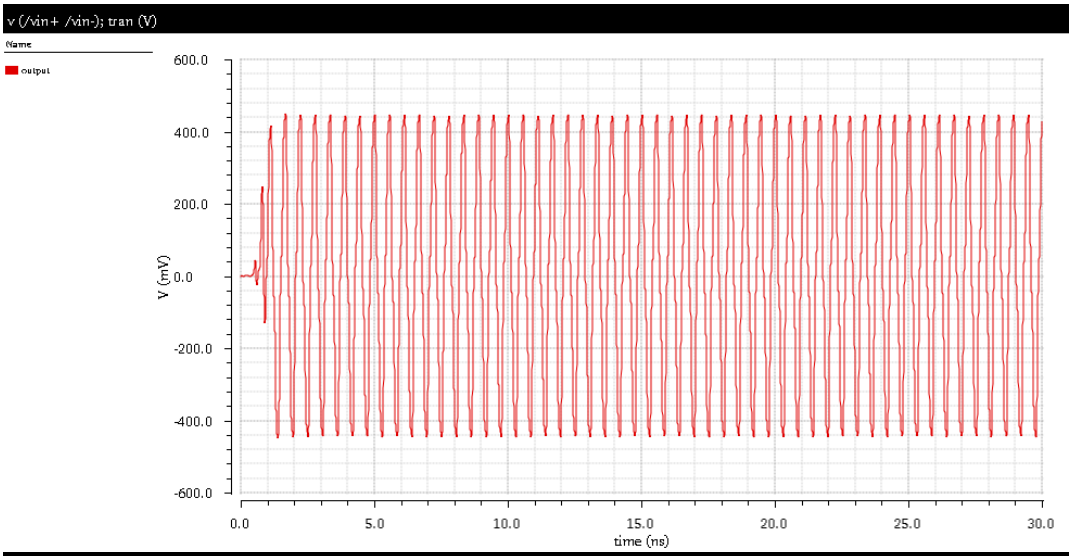


Figure 5.65: Divider Output at 3.6 GHz Input

5.5.6.3. CML Buffer

A CML buffer was designed to eliminate the loading effect of the mixer on the divider. The buffer is a simple common source amplifier as depicted in Fig. 5.66. A poly current is used (Reference current that is proportional to 1/R) to keep the output common mode level constant across corners to avoid signal non-linearities. Figure 5.67 shows the transient simulation in the typical case. In Fig. 5.68, the transient simulation across corners is depicted and it shows low dependence on the process variations. The swing changes by only 50 mV across corners due to the buffer gain variation but the common mode is constant due to the use of the poly current. The buffer consumes 1.2 mW from 1.2 V supply.

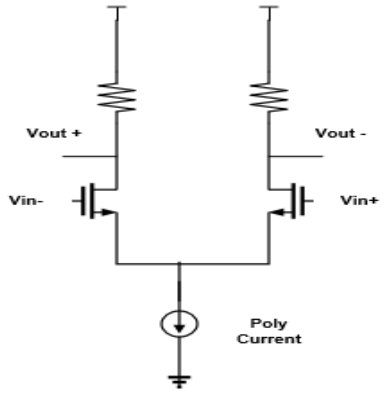


Figure 0.66: CML Buffer

5.5.6.3.1. Simulations Results

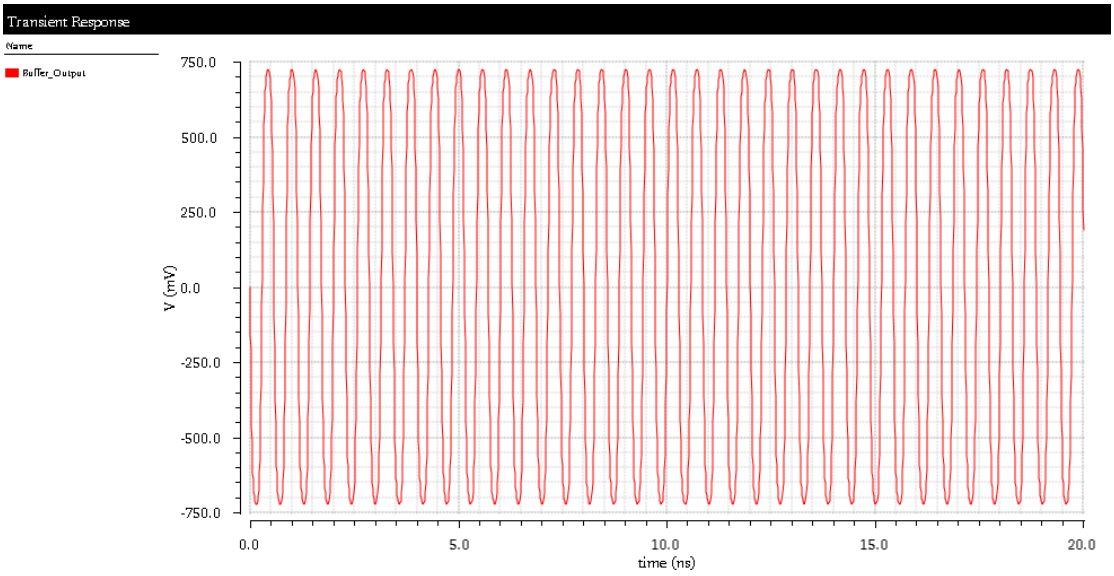


Figure 0.67: Transient Simulation in Typical Case

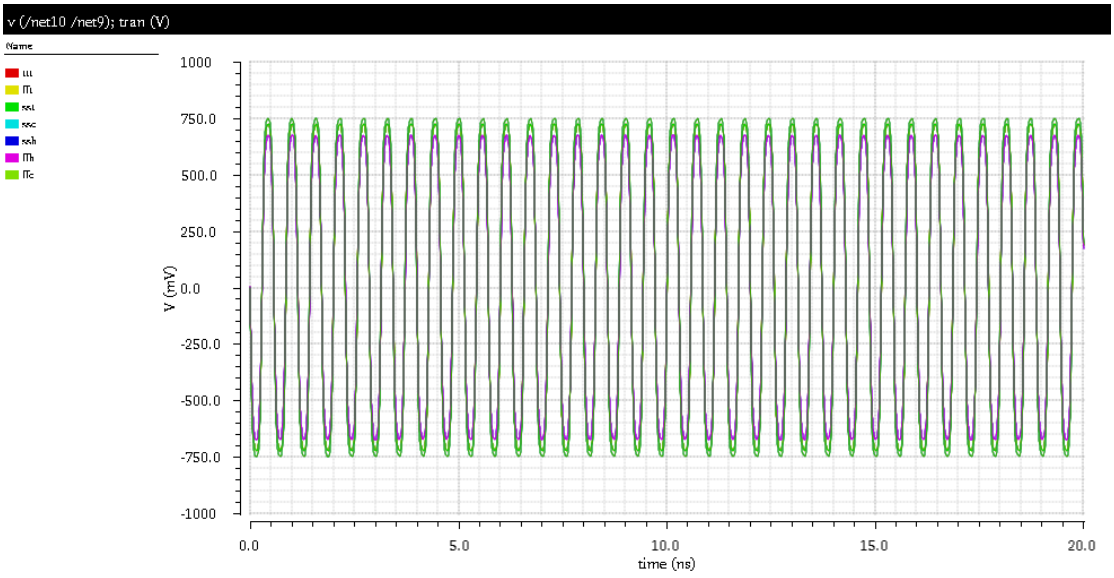


Figure 0.68: Transient Simulation across Corners

5.5.6.3.2. VCO, Divider and Buffer Cascade

The transient simulation of the cascade is shown in Fig. 5.69 with the loaded with both the cml buffer of the PA that is used to convert the PLL output into a square wave and loaded by the mixer in the feedback path. The harmonics of the cascade are shown in Fig. 5.70. The VCO is made to run at 3.383 GHz. The output is at half the frequency which is 1.6915 GHz. The third harmonic distortion is 20 dB. However, this harmonic will be attenuated by the PLL PA matching network as it is very far from the carrier frequency. In the PLL feedback path, the filter will only take the first harmonic after passing through the mixer. This harmonic had no effect on the system performance.

The spec of -85 dBc/Hz at 500 KHz on the phase noise is on the cascade of the VCO, divider and the output buffer which are inside the loop. As shown in Fig. 5.71, the achieved phase noise is -116.7 dBc/Hz at an offset of 1MHz and -110 dBc/Hz at an offset of 500 KHz. The achieved phase noise of the cascade is better than the phase noise of the VCO alone. This is due to the presence of the divider which divides the phase and frequency by 2. Thus it achieves a 6 dB better phase noise.

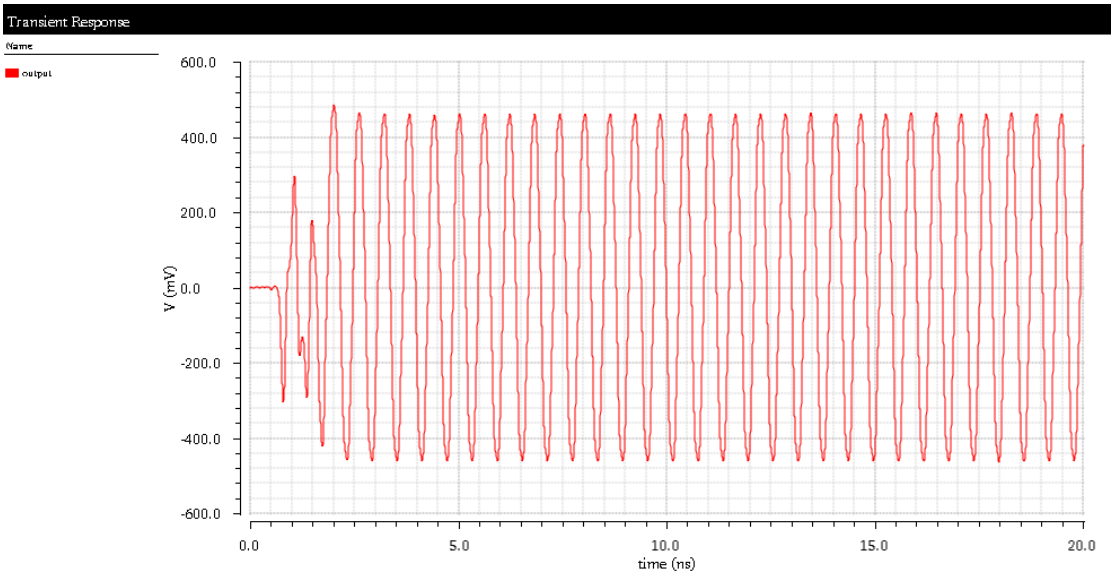


Figure 0.69: Cascade Transient Output

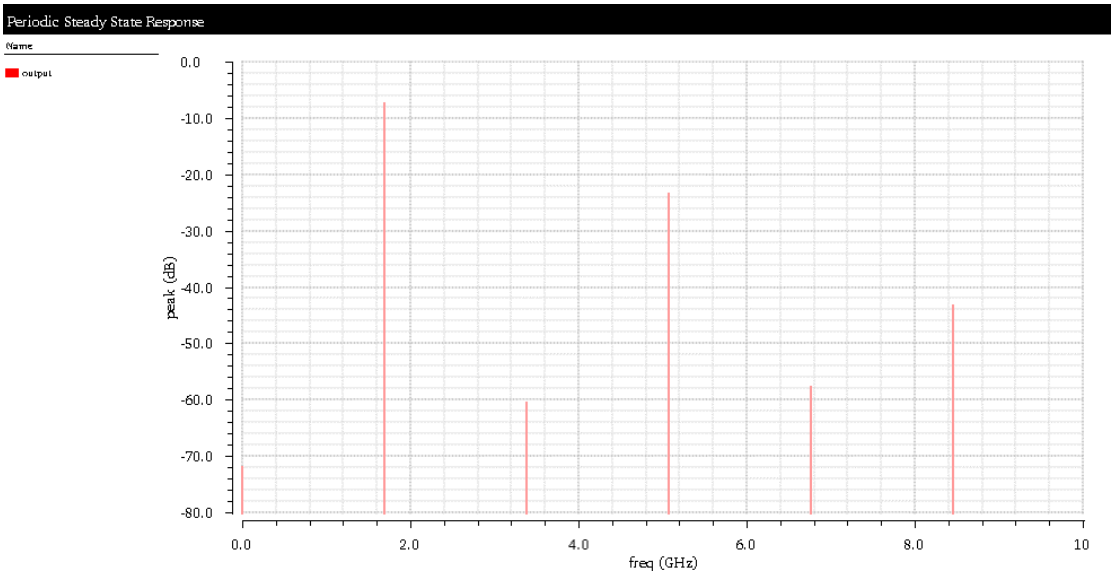


Figure 5.70: Harmonics of the cascade

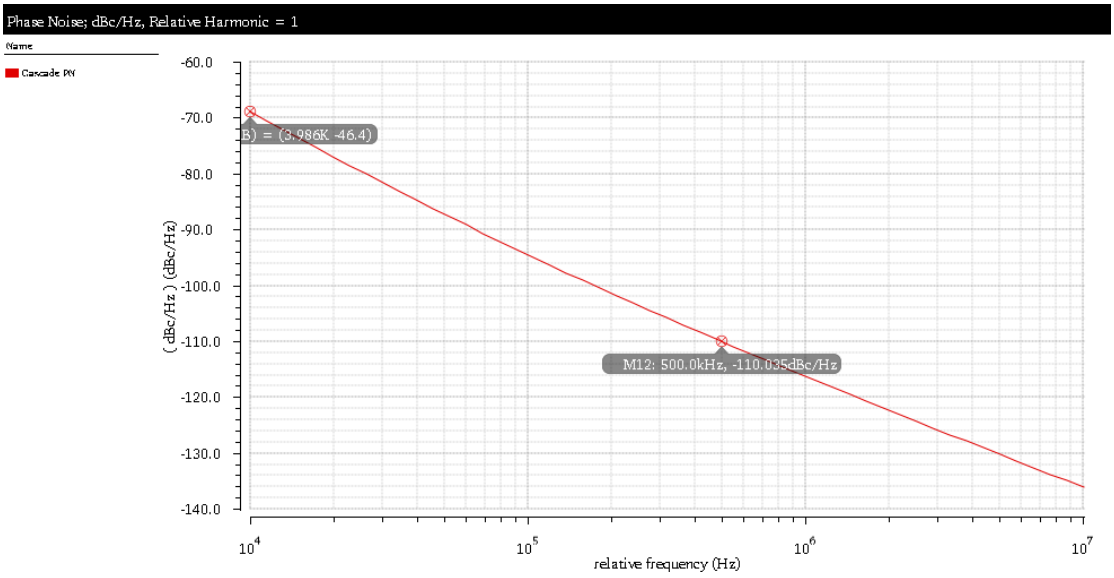


Figure 0.71: Cascade Output Phase Noise

5.5.7. Offset Mixer

Offset PLL depends on mixers to perform feedback frequency division operation. As explained before, implementation of the divider using a mixer is to keep the phase information without change while translating the VCO output frequency to the PFD/CP input frequency [1]. Mixers usually classified as up-conversion mixers used in transmitters and down conversion mixers in receivers. In our scenario the mixer is a down conversion mixer but it is used in an offset PLL transmitter as a frequency divider. This scenario implies different considerations to be taken upon extracting the design specifications. The most important specification is the LO-RF and IF-RF feedthrough as the RF port is connected to the output of the VCO and also the input of the PA, thus poor isolation between ports could lead to violating the specified spectral emission mask. Linearity is specified according to the RF signal power level, thus the 1 dB compression point should be chosen few dBs higher than the RF signal power. Cascade analysis is made for the three blocks in the PLL feedback path. This analysis leads to 2 dB conversion gain and 25 dB noise figure specs on the mixer. Based on the previous discussion, the mixer is chosen to be active mixer to provide good ports isolation. Active mixers may suffer from linearity issues but these issues are not critical in our situation due to the existence of the limiter.

5.5.7.1. Literature Review

The very basic active mixer is the Gilbert cell mixer, the double-balanced version is shown in Fig. 5.72. Double balanced mixers provide better isolation than single balanced. The main idea is to convert the RF voltage to current through the input transconductance pair, then commutate this current through the M3-M6 and convert it to output voltage through the load resistance. The main drawback of this architecture is the large needed voltage headroom which is not consistent with low voltage applications. Also, the mixer's linearity is mainly dominated by the input transconductors linearity. The input transconductors linearity is somehow proportional to the overdrive voltage. To increase the overdrive voltage of a transistor one can increase the bias current, which is undesired for low power applications, or decrease the transistor's aspect ratio which will impact the noise performance as the flicker noise is proportional to the transistor's width and length. Many enhancement techniques are introduced in literature to overcome the traditional Gilbert cell mixer drawbacks. Current-bleeding technique proposed in [16] reduces the bias current passing from the RF stage to the LO stage and thus decreases the voltage drop across the load resistors and consequently allows low voltage operations. According to [17], this technique is limited to narrow band applications. Another enhancement technique targeting low voltage operation is the bulk injection topology [18]. This technique requires extra fabrication efforts due to the need for a triple well process, also it suffers from poor noise performance [17]. Another promising technique is what called folded mixer topology [19, 20].

5.5.7.2. Proposed Designed

Based on the above discussion and literature review the proposed mixer is chosen to follow the same folded current-reuse self-bias topology introduced in [17]. Fig. 5.73 shows a circuit diagram for the proposed mixer. R1 and R2 are providing self-biasing for the RF stage and thus eliminate the need for extra biasing circuitry.

5.5.7.3. Simulation Results

In this section, the simulation results for the proposed mixer are presented. Simulations are done based on [21, 22]. Fig. 5.74 shows the testbench used for testing the mixer's performance. Also, it shows the current consumption which is about 680 uA which is considered a very low current consumption. Table 5.2 shows a summary of the achieved specs.

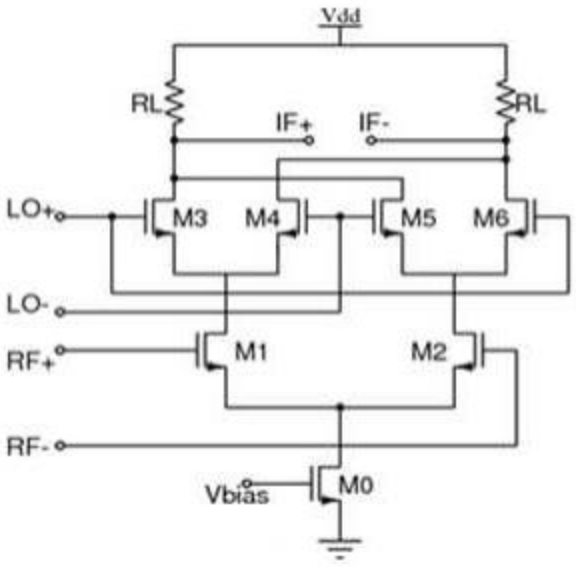


Figure 0.72: Gilbert Cell Mixer

Conversion Gain Vs LO Power

As shown in Fig. 5.75, the LO power is swept versus the mixer's conversion gain to find a compromise between required conversion gain and low LO power. A 0-dBm LO power will achieve the required gain with a margin for corners.

Conversion Gain Versus RF Frequency

The conversion gain is plotted versus frequency as shown in Fig. 5.76. Fig. 5.77 shows the same simulation but across different corners. In our case, the lower the gain the better the noise performance of the PLL, thus the system will tolerate the gain variations across corners.

Noise Figure

Mixer's noise figure is simulated across different corners as shown in Fig. 5.78. The mixer's achieves noise figure less than 26.6 dB in the typical corner. This value varies by around ± 3 dB cross all corners which can be tolerated by the system due to the better noise figure achieved by the limiter as shown later.

Linearity

As mentioned before, mixer's linearity does not have significant importance, as it will be followed by a limiter, which will clip the amplitude of the signal to 0 and 1.2v. Alternatively, the 1-dB compression point should be close to the input power level to avoid unwanted spurs. Figure 5.79 shows the 1-dB compression point simulation of the mixer. The mixer achieves 0 dBm compression point. Fig. 5.80 illustrates the IIP3 of the mixer. The mixer achieves 8.85 dBm IIP3.

Ports Isolation

Isolation performance between ports is critical in our case especially the feedthrough to the RF port as it is connected directly to the PA's input. Figures 5.81 to 5.84 shows the isolation performance among various ports.

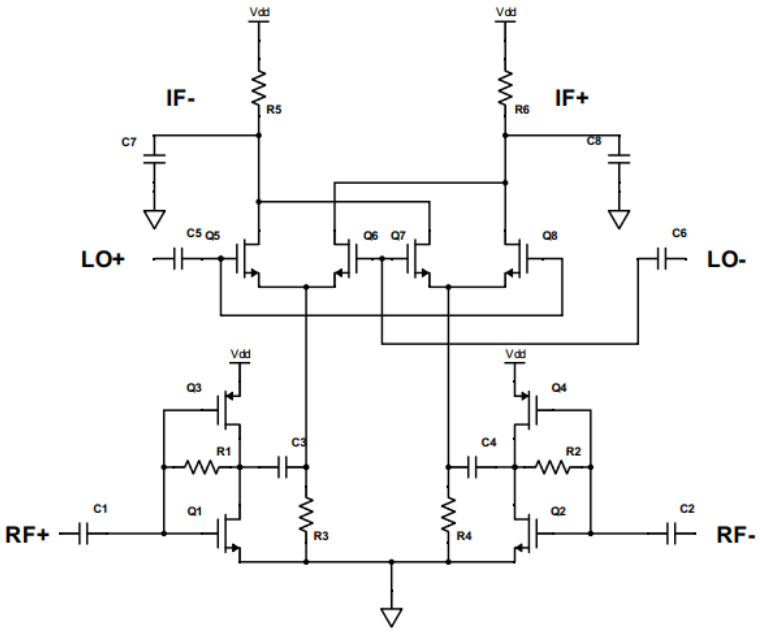


Figure 0.73: Proposed Mixer Circuit Schematic

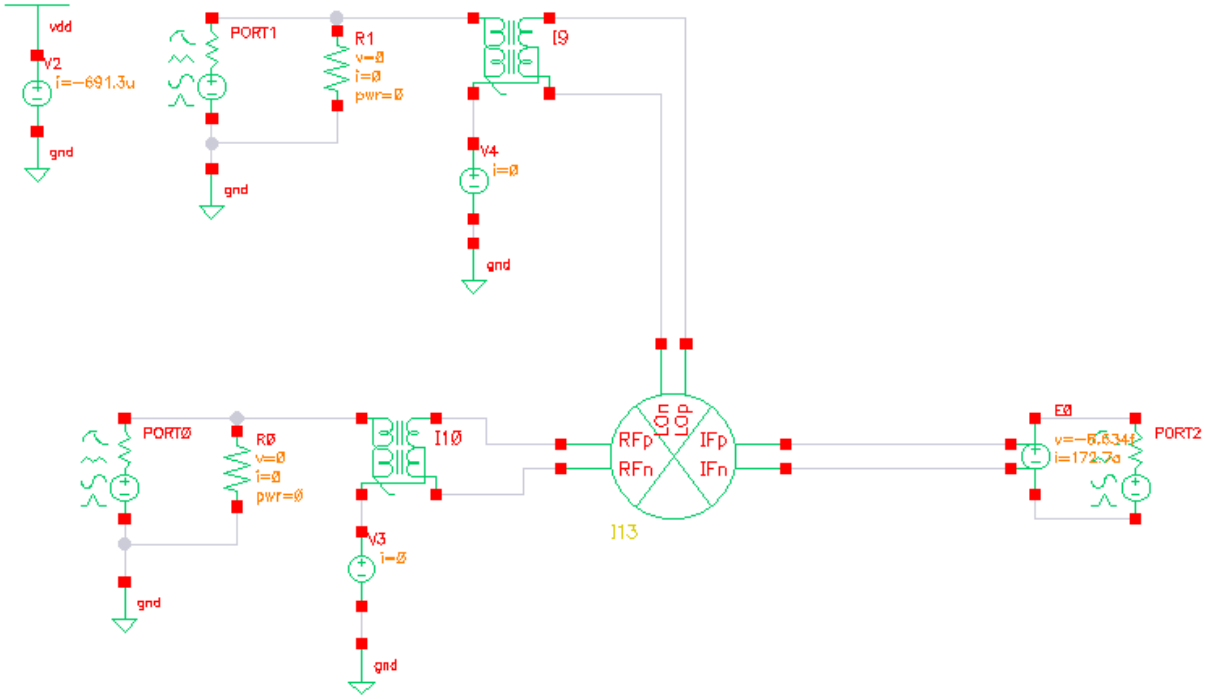


Figure 0.74: Mixer's Testbench

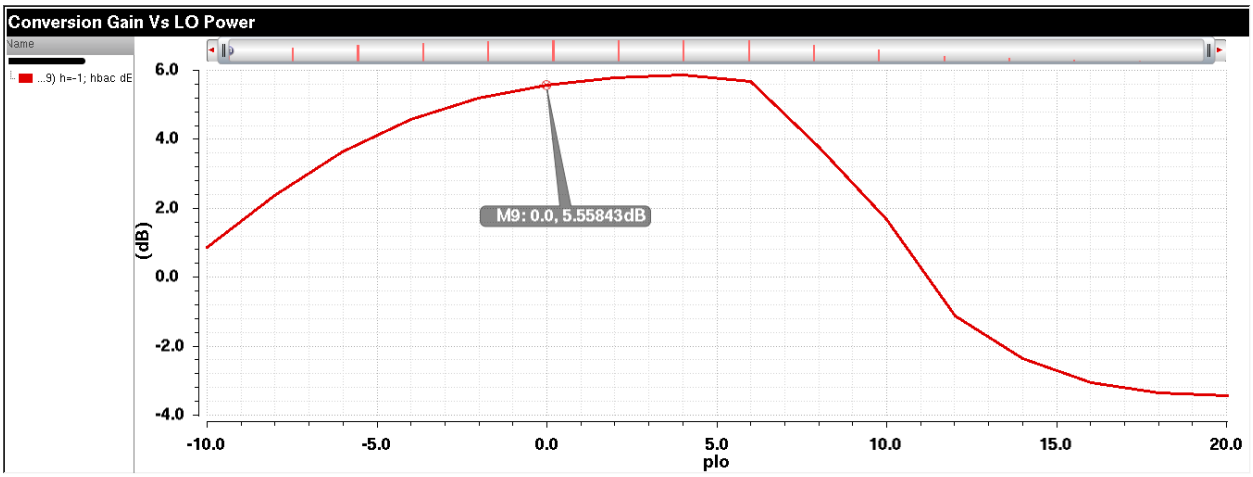


Figure 0.75: Conversion Gain Vs LO Power

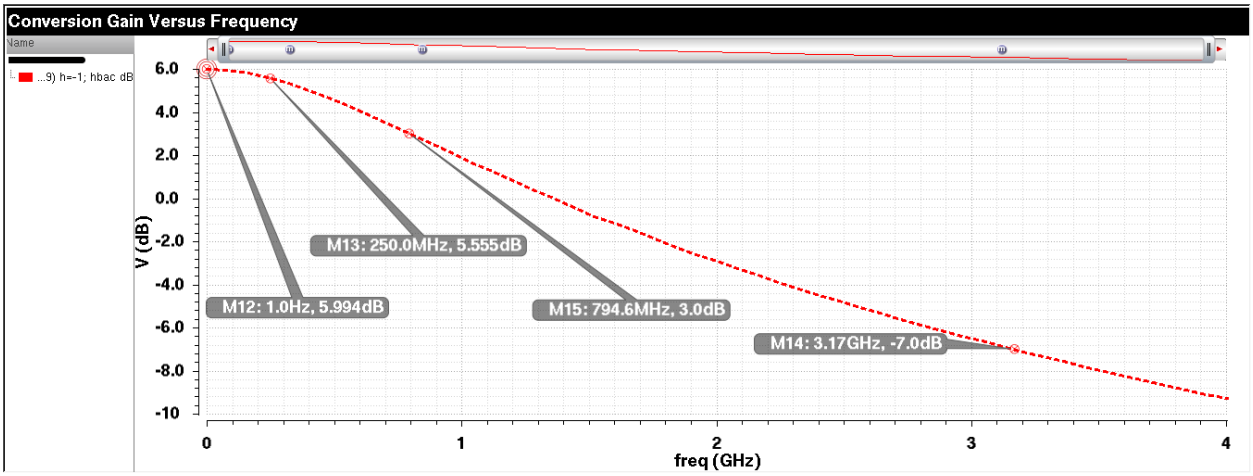


Figure 5.76: Conversion Gain Versus Frequency

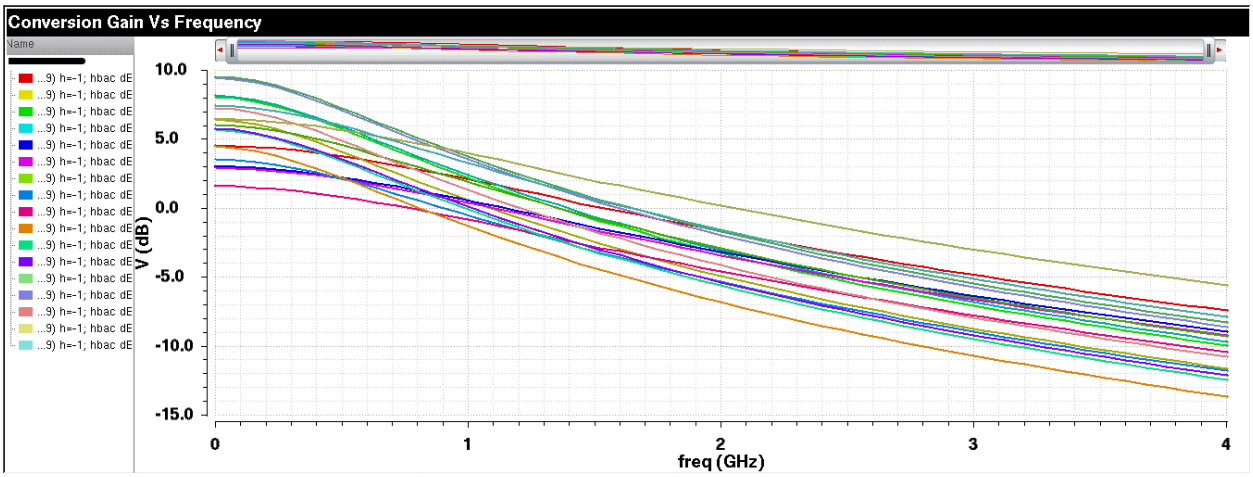


Figure 0.77: Conversion Gain Versus Frequency across Corners

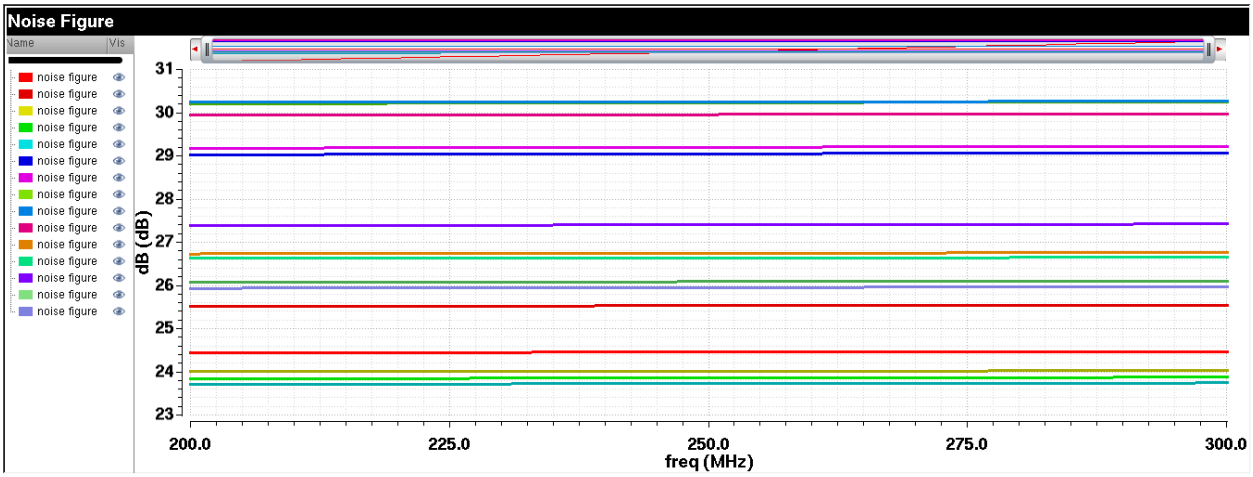


Figure 5.78: Noise Figure across Corners

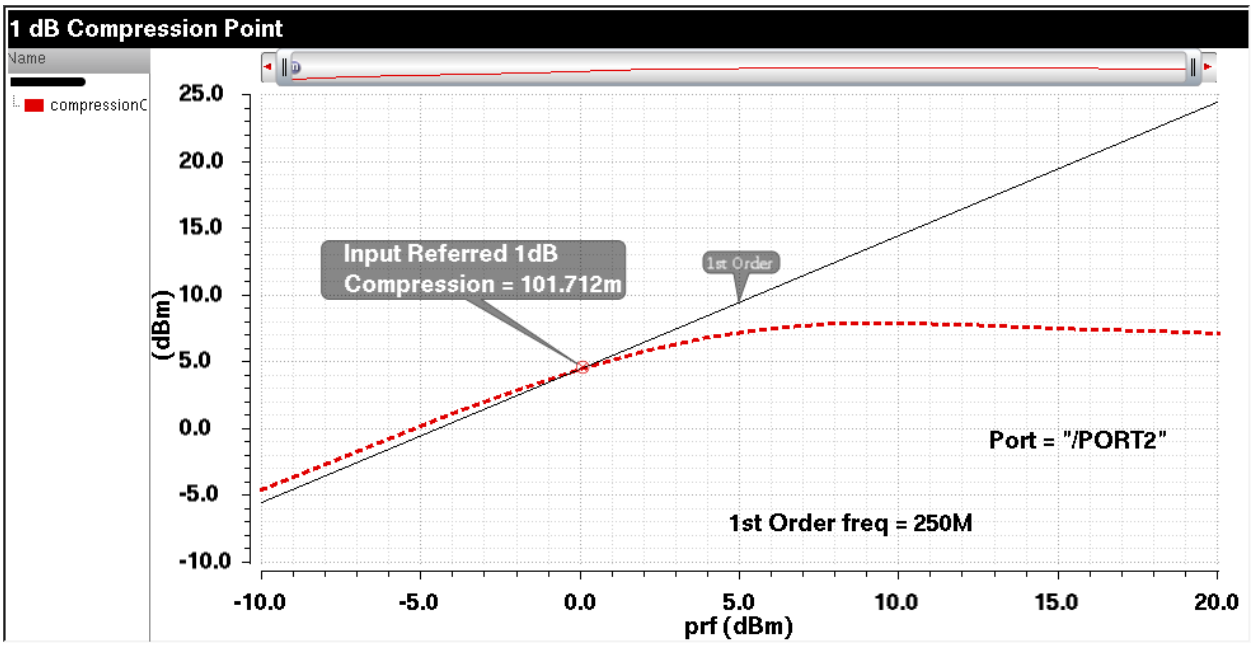


Figure 0.79: 1-dB Compression Point

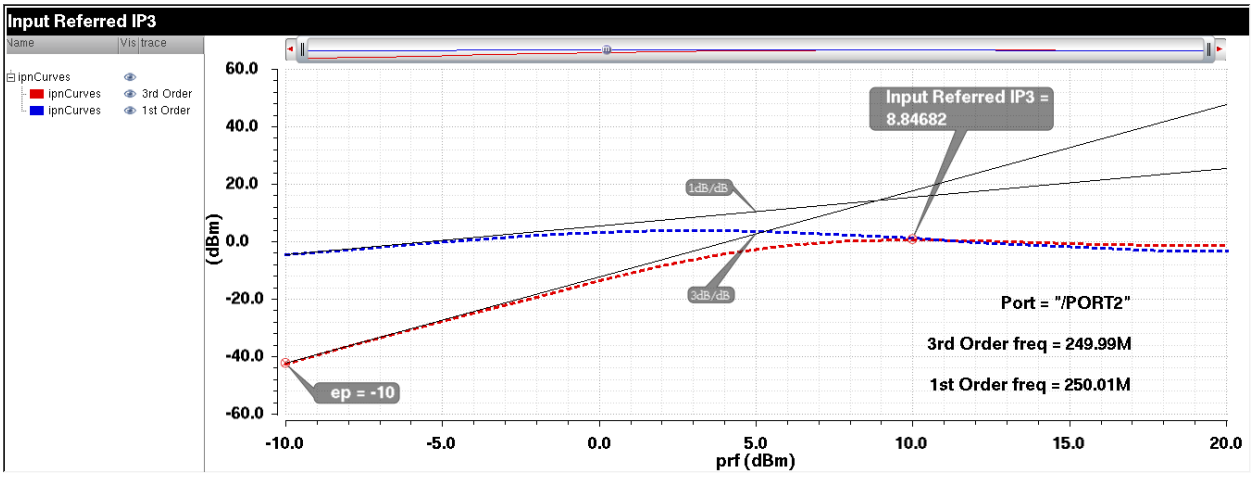


Figure 0.80: Input Referred IP3

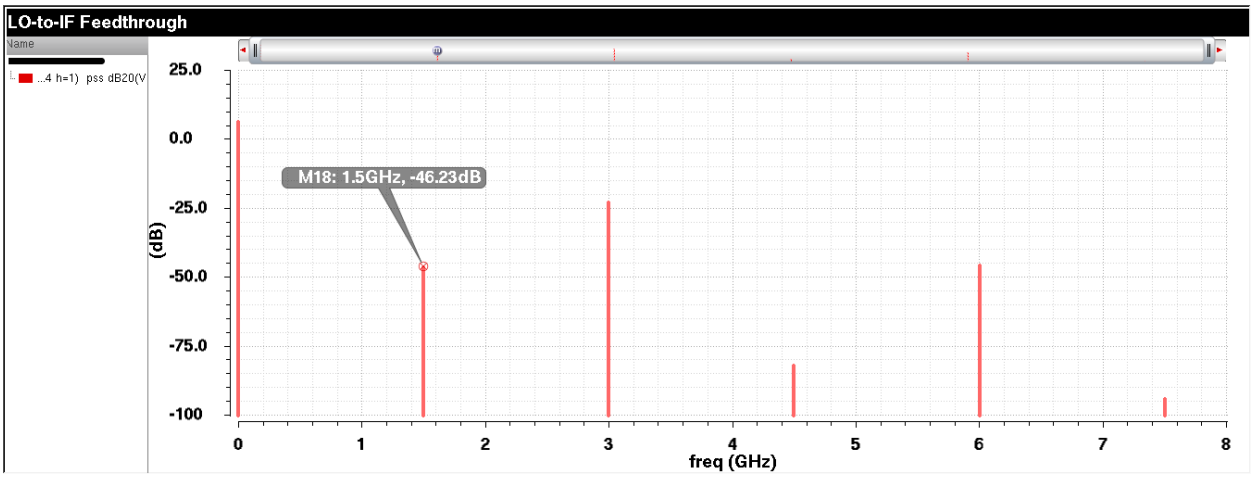


Figure 0.81: LO to IF Feedthrough

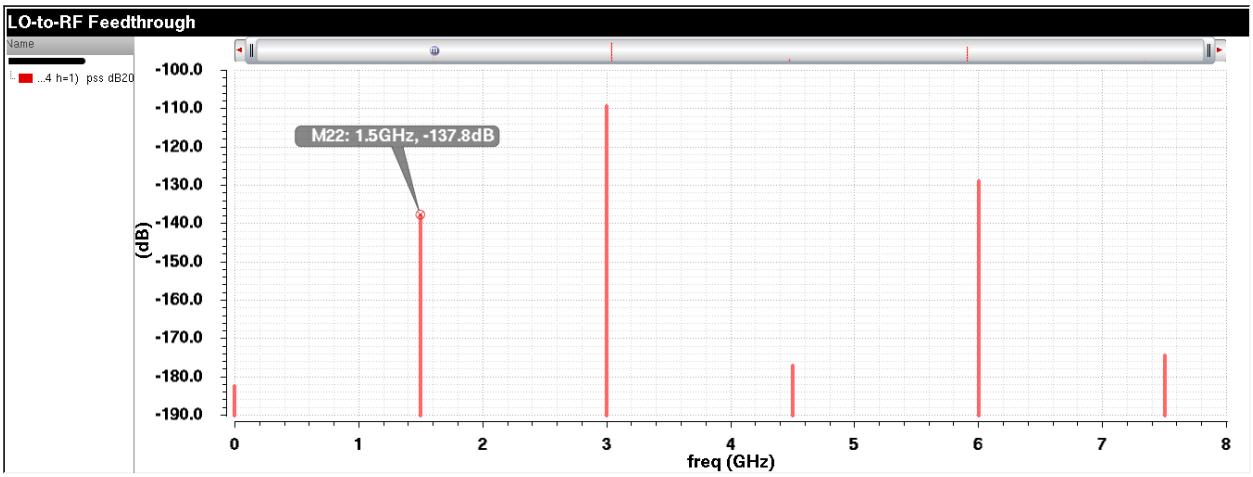


Figure 0.82: LO to RF Feedthrough

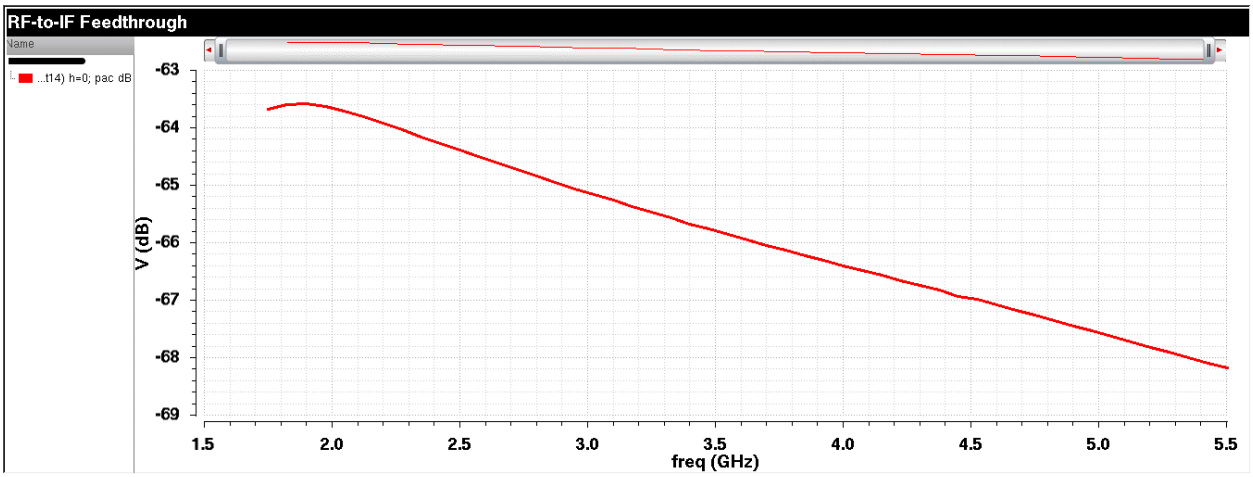


Figure 0.83: RF to IF Feedthroug

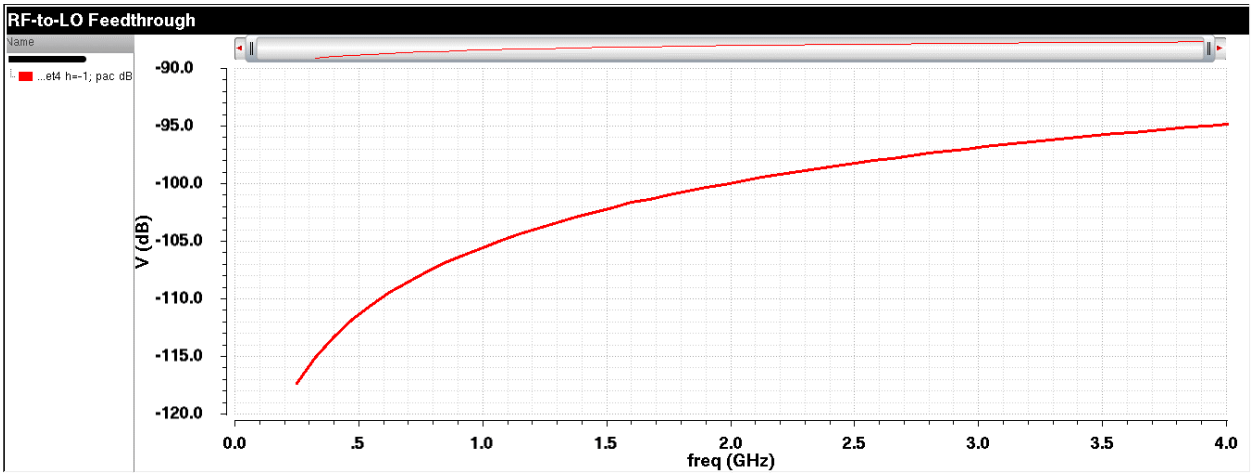


Figure 5.84: RF to LO Feedthrough

Table 5.2: Mixer's Achieved Specs

Spec	Value
Bandwidth	800 MHz
Gain	3.57 dB @250 MHz
Noise Figure	26.7 dB
Current	680 Ua
P1dB	0 dBm
IIP3	8.85 dBm
LO-IF Feedthrough	-46 dB
LO-RF Feedthrough	-137 dB
RF-IF Feedthrough	-64 dB
RF-LO Feedthrough	-100 dB

5.5.8. Filter

The targeted filter is required to reject the sum of the LO and RF components and pass the difference between them. According to the chosen band, the RF frequency (f_{RF}) ranges from 1710 – 1785 MHz. The PLL reference is chosen to be fixed at 250 MHz, this also requires the LO frequency (f_{LO}) to be tunable in the range from 1460 to 1535 MHz. This implies that the required frequency to be rejected is also variable in the range of 3170 to 3320 GHz. Thus, as a worst case for the design, the designed filter is simulated with its input as the sum of two sinusoids at 250 MHz and 3.17 GHz respectively. According to the previous discussion and what is mentioned in chapter 4, OTA-C filters is considered to be the most suitable active filter topology in the operating frequency range. To ensure good rejection and to comply with the spectral emission mask specified by the standard, a 4th order low-pass filter following Butterworth approximation is designed.

5.5.8.1. Design Procedure

OTA-C, also called Gm-C, filters mainly depend on the operational transconductance amplifier. The relationship between the output current and the input voltage of the transistor is a nonlinear function. In the most simple case, it is an exponential function for the bipolar transistor and a quadratic function for the MOS transistor. The most important aspects in the design of voltage-to-current transducers for continuous-time filters are: linearity, phase response, noise level and power consumption [23]. In our scenario, the filter will be followed by a limiter, which clips the amplitude of the filter's output signal. This follows that no linearity specs will be required on the filter. As a low-power application, the main goal will be minimizing the consumed power by the filter. Various OTA techniques are available. For simplicity, a 5-T OTA with resistive CMFB is chosen for the targeted filter.

The implementation of the filter starts from obtaining the appropriate transfer function. The normalized transfer function for Butterworth approximation for different orders is found in [24]. The transfer function for 4th order filter is given by

$$H(S) = \frac{1}{(S^2+0.7537S+1)(S^2+1.84776S+1)} \tag{5.29}$$

Frequency transformation is done to denormalize this transfer function and transform the cut-off frequency from 1 rad/sec to the required cut-off frequency ω_c. This achieved by replacing every S by$\frac{S}{\omega_c}$. To get the cut-off frequency, some other filter parameters must be specified. Figure 5.85 shows a possible magnitude response for a real approximation for a low pass filter.

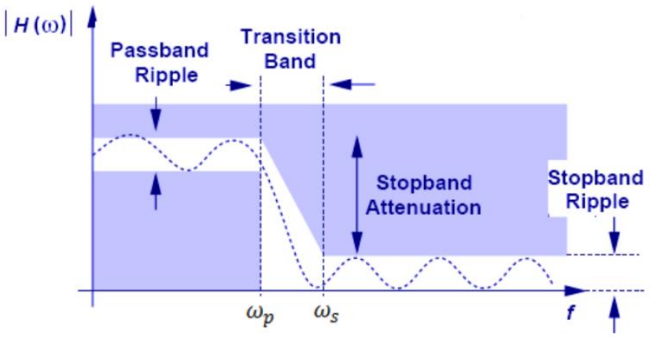


Figure 0.85: Low pass approximation

For Butterworth approximation, the parameters in Fig.5.85 are related to the filter order and cut-off frequency as follows [23]:

$$A_p = 10\log\left[1+\left(\frac{\omega_p}{\omega_c}\right)^{2n}\right] \quad (5.30)$$

$$A_s = 10\log\left[1+\left(\frac{\omega_s}{\omega_c}\right)^{2n}\right] \quad (5.31)$$

A_p: Maximum passband attenuation

A_s: Minimum stopband attenuation

ω_p: Passband edge frequency

ω_s: Stopband edge frequency

Solving (5.30) by letting $A_p = 1\,dB$ and ω_p=300MHz thus ω_c is obtained 355MHz.

We now need to synthesize the denormalzied transfer function and find a circuit diagram that could perform the transfer function. A possible biquad implementation is found in [25] and will be followed in this design. The chosen biquad is shown in Fig. 5.86.

The transfer function of this biquad is as follows

$$H(s) = \frac{\frac{g_{m1}g_{m3}g_{m4}}{g_{m4}\quad C_1C_2}}{s^2+\frac{g_{m2}}{C_1}s+\frac{g_{m3}g_{m4}}{C_1C_2}} \quad (5.32)$$

Comparing it to the common second order low pass filter transfer function

$$H(s) = \frac{{\omega_0}^2}{s^2+\frac{\omega_0}{Q}s+\omega_0} \quad (5.33)$$

Thus, a relation between the circuit parameters: gm's and the capacitors, and the filter parameters Q and ω_0 can be obtained.

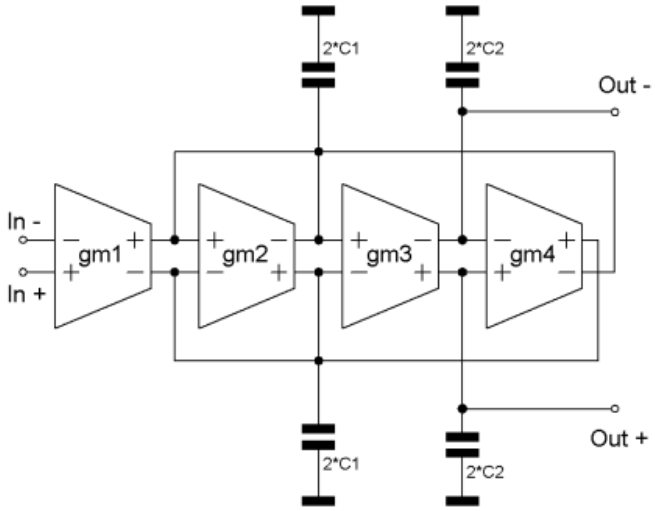


Figure 0.86: Gm-C Biquad

$$\omega_0{}^2 = \frac{g_{m3}g_{m4}}{C_1C_2} \qquad \frac{\omega_0}{Q} = \frac{g_{m2}}{C_1} \tag{5.34}$$

Choosing $g_{m1} = g_{m3} = g_{m4} = g_m$ and C1=C2=C leads to

$$\omega_0 = \frac{g_m}{C} \qquad Q = \frac{g_m}{g_{m2}} \tag{5.35}$$

By letting C=100fF, all transconductors values can now be obtained. Table 5.3 summarizes the derived circuit elements of both biquads.

Table 5.3: Biquads Elements Values

	Biquad 1	Biquad 2
Gm	193.4 uA/V	193.4 uA/V
Gm2	357.41 uA/V	147.969 uA/V
C	100 Ff	100 Ff

5.5.8.2. Simulation Results

In this section, the simulation results of the designed filter are presented.

Gm Cells

Figures 5.87 to 5.89 shows the DC simulations of the designed Gm cells. The OTA transconductance is approximately equal to the input pair transconductance. The designed OTA is a single pole system, which is supposed to be normally stable. However, a stability analysis is made for the "gm" cell. Figure 5.90 shows the open loop gain and plots for the gm cell. It is clear that the OTA is stable with large phase margin.

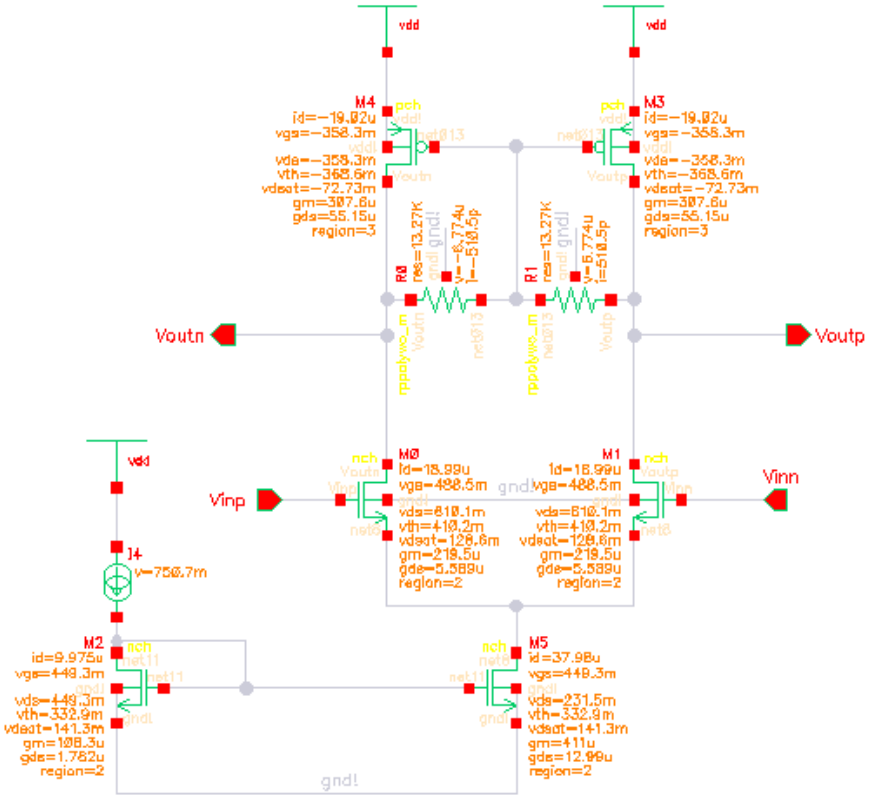


Figure 0.87: Schematic of the "gm" cell

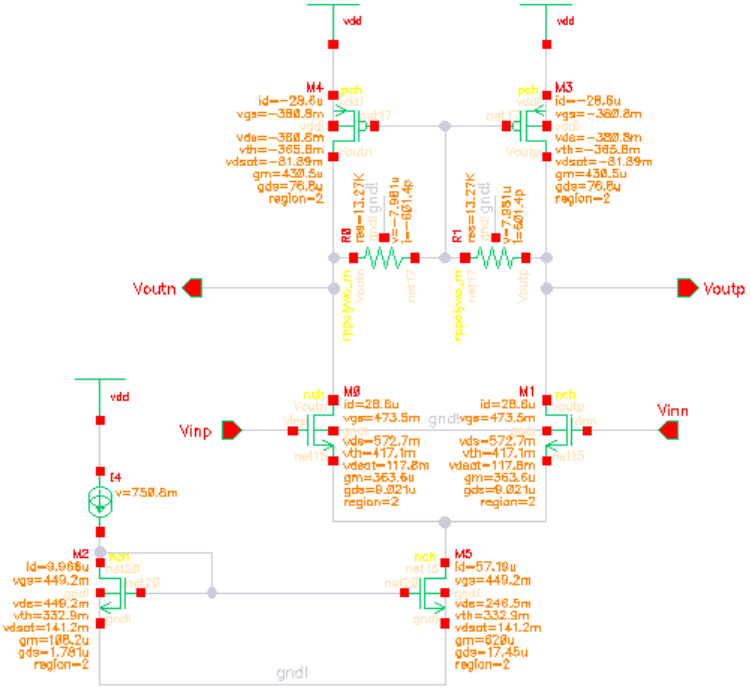


Figure 0.88: Schematic of the "gm21" cell

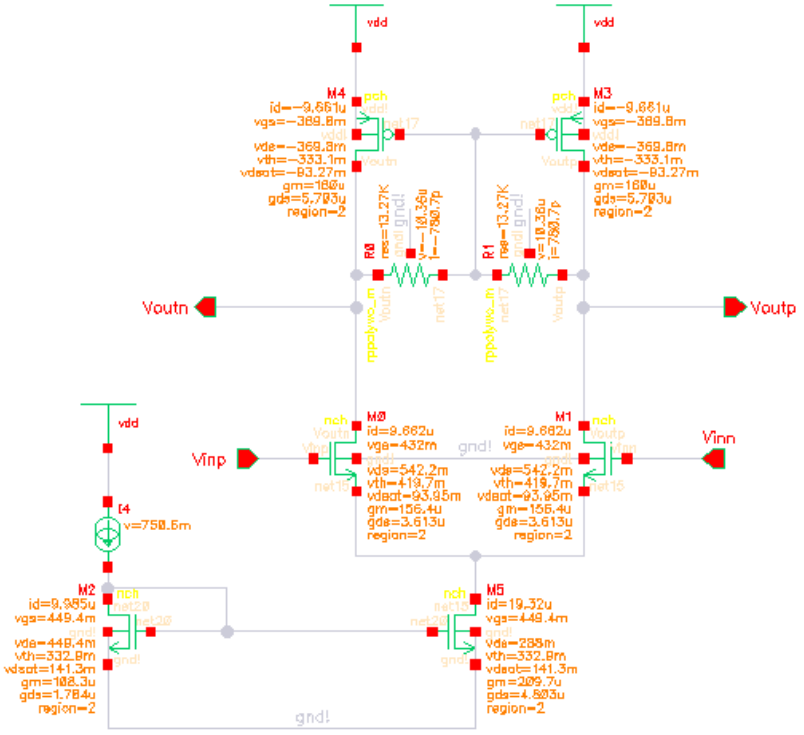


Figure 0.89: Schematic of the "gm22" cell

Filter Biquads

Figure 5.91 shows the filter's biquads. Figure 5.92 illustrates the testbench for the designed filter. AC, transient and noise simulations are carried out across different process and temperature corners. It is clear from Fig. 5.92 that the filter consumes only 300 uA (excluding the reference bias currents of the OTAs). As shown in Fig. 5.93, the filter achieves more than 60 dB attenuation at the required stopband frequency. The transient response in Fig. 5.93 shows that the filters output is around 50 mvpd which is well above the limiter's sensitivity. The filter's input waveform is shown in Fig. 5.94. Figure 5.95 shows the frequency of the output waveform calculated using Cadence Virtuoso's calculator function *freq*. The noise figure of the filter is required to be 35 dB to achieve the required phase noise to comply with spectral emission mask. Figure 5.96 shows the filter's noise figure across corners which is clear that the filter achieve the required spec by a margin except for the hot corner (125 degC) at which the noise figure is about 36 dB. However, the noise figure spec can benefit from the low noise figure achieved in the limiter.

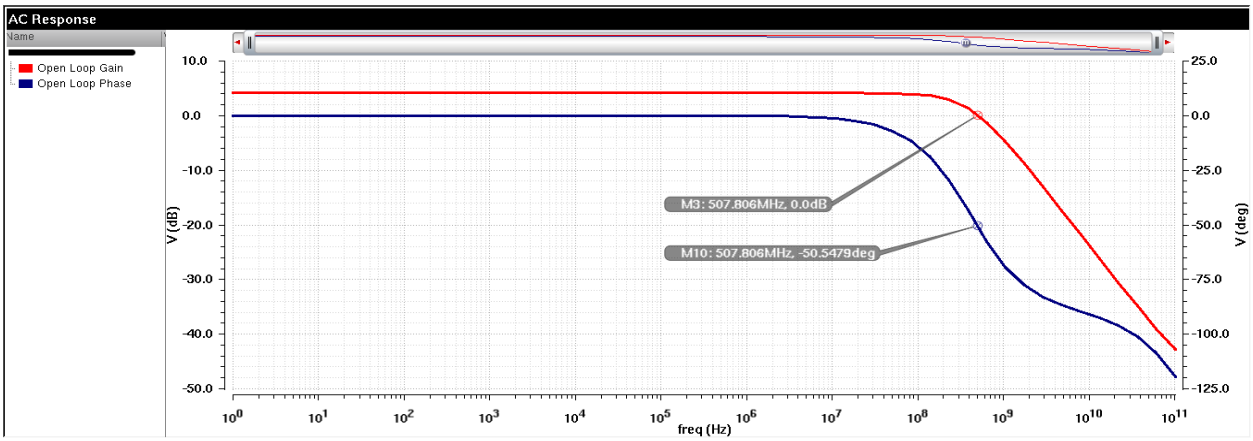


Figure 0.90: Gm Open Loop Gain and Phase

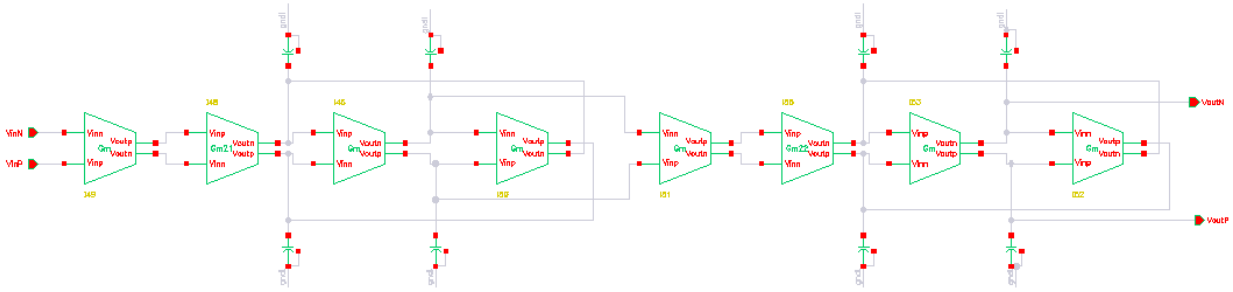


Figure 0.91: Filter's Biquads

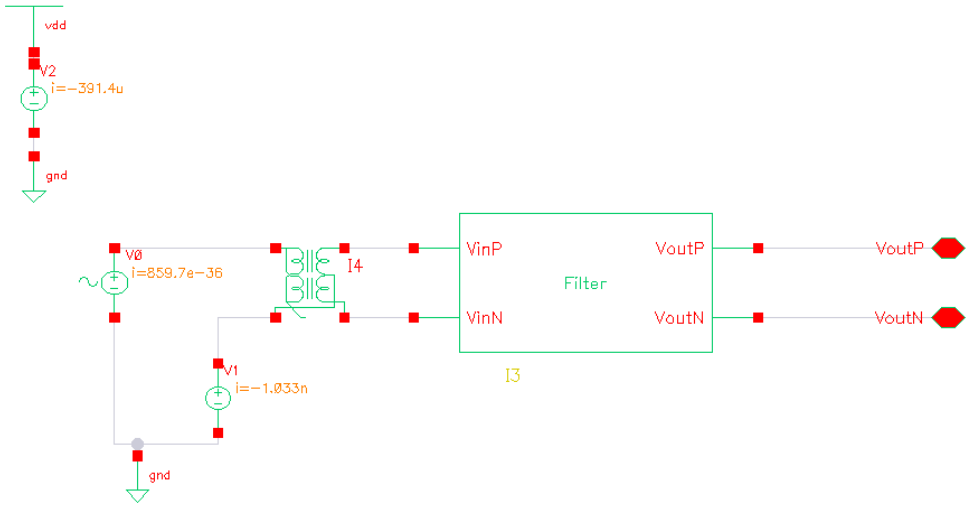


Figure 0.92: Filter's Testbench

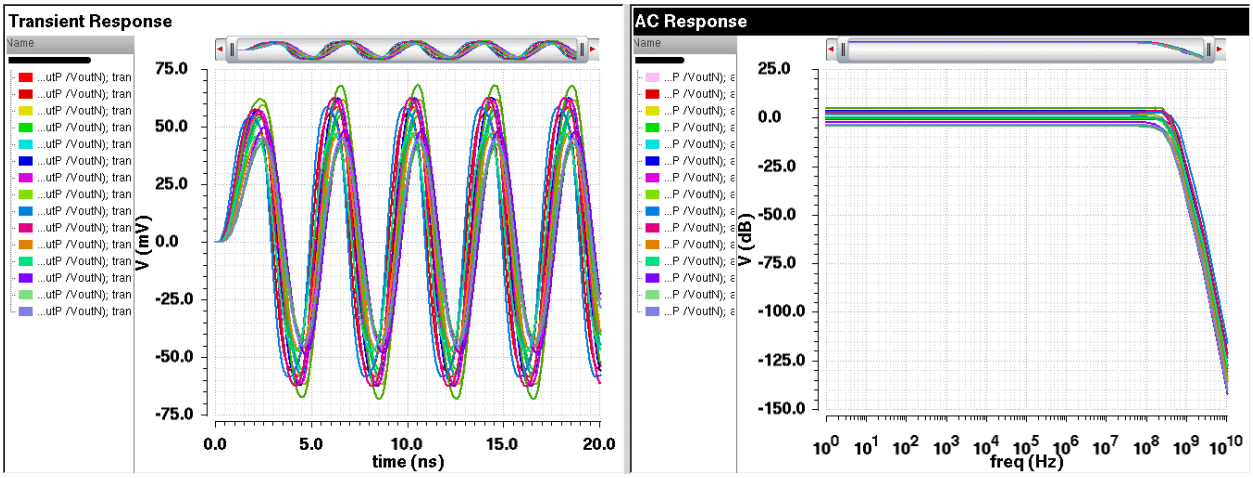


Figure 0.93: Filter's AC and Transient Response across Corners

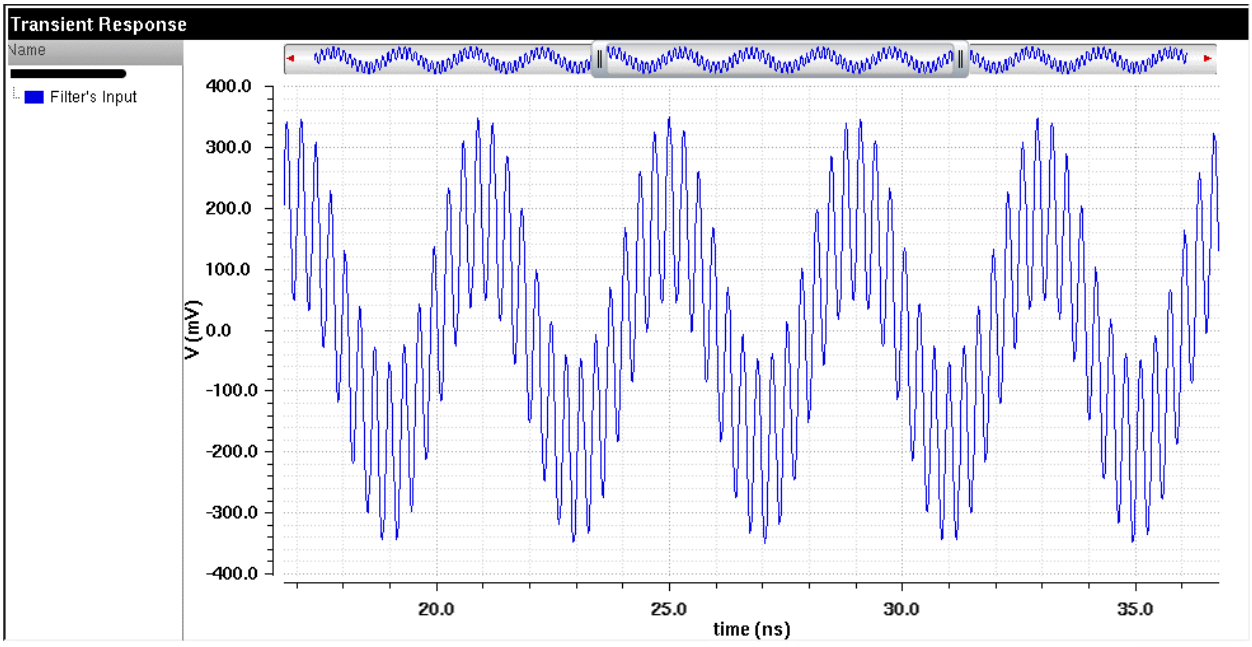


Figure 0.94: Filter's Input Waveform

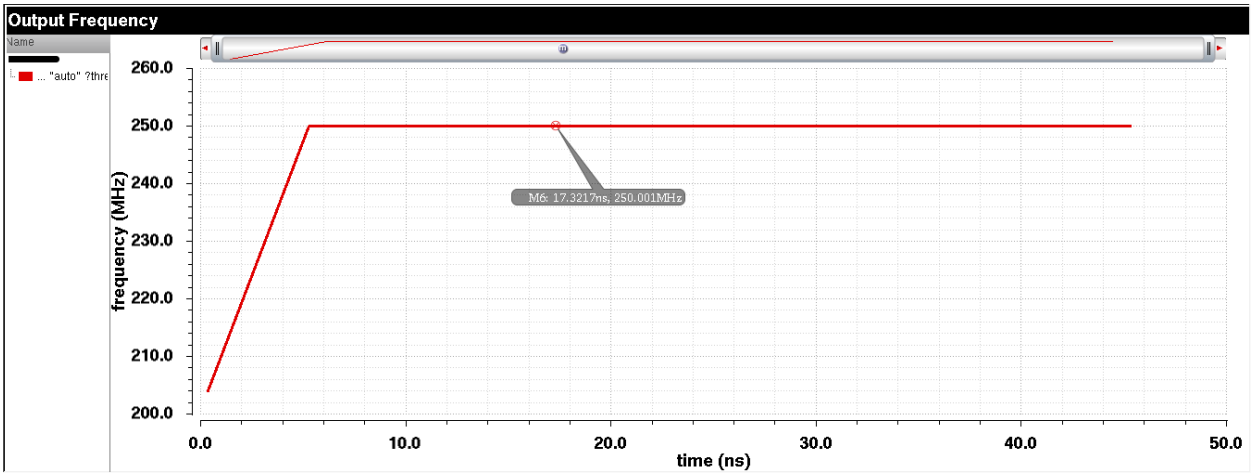


Figure 0.95: Filter's Output Frequency

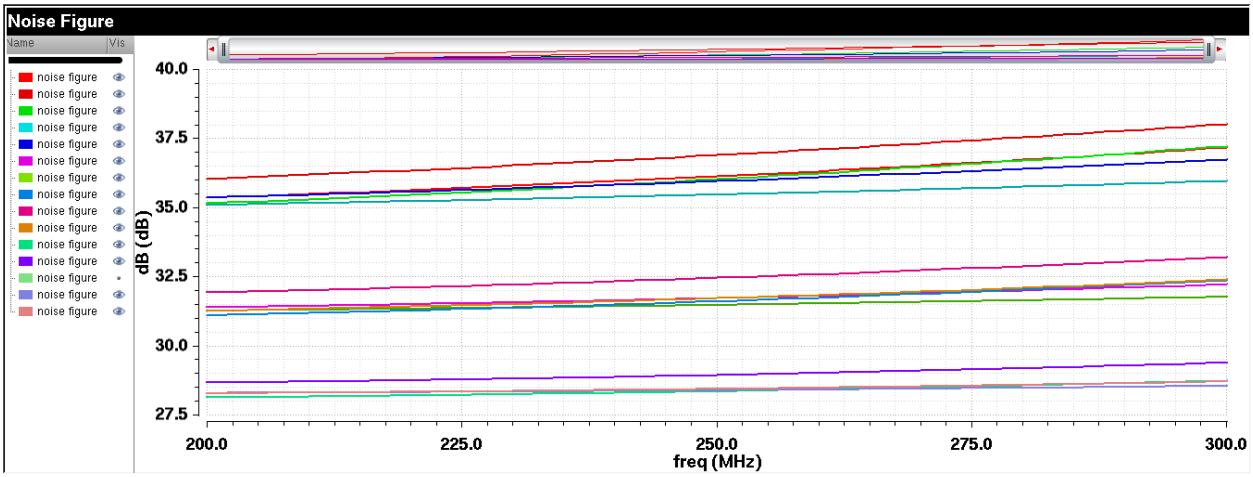


Figure 0.96: Filter's Noise Figure

5.5.9. Limiter

Also called “Limiting Difference Amplifier”, is a circuit which clips the signal’s amplitude producing a square wave signal. Limiters are mostly used in optical receivers [26-28] and low IF FSK receivers [29, 30]. In the first case, the limiter’s bandwidth is very large as optical communications support very high data rates and thus very high bandwidth which is in the order of GHz. While in low IF receivers the bandwidth required does not exceed few tens of MHz. Our case is considered as an in-the-middle case compared to both previously mentioned cases as the limiter’s bandwidth should cover the IF frequency bandwidth which is 250 MHz. This bandwidth can be easily achieved especially with the short-channel 65 nm technology used

5.5.9.1. Design Requirements

In the proposed transmitter architecture two limiters are needed, one after the IQ modulator to extract the phase modulated signal and the other is in the feedback path of the Offset PLL. The output swing from the IQ modulator will be around 50mv this requires at least a limiter’s gain of 30 dB. This gain cannot be achieved from a single amplification stage. A good design which compromises between gain, area and noise performance could use typically 3 to 4 stages [31]. Also, a 27 dB noise figure is required such that the input referred noise figure of the limiter does not degrade the system phase noise and thus violating the spectral emission mask. Current consumption should be minimized as we are dealing with a low power communication standard. Also, an offset cancellation technique to eliminate the DC offset which could cause a great problem with this high gain. Three main offset cancellation schemes are found in the literature.

The first one is using *feedback* technique [31]. The DC offset at the output of the last stage is sensed by an RC LPF and is subtracted at the input of the first stage. This technique could suffer from stability issues at high frequencies. Another method is *feedforward* technique [32]. Here, The DC offset at the output of each stage is extracted and amplified as well as the signal by two identical amplifiers and then forwarded and subtracted from the output of the subsequent stage. In a fully differential scenario this method will consume large area overhead by adding two extra amplifiers and two RC filters to each stage. Finally, AC coupling technique can be used to block the DC of each preceding stage. This technique requires large capacitors and resistors values which cannot be implemented on chip if the frequency of operation is low as the case of low IF FSK receivers. A best compromise in our case is using AC coupling technique as the IF frequency is quite large thus the capacitors and resistors values will be reasonable and can be implemented on chip. Table 5.4 summarizes the required specs.

Table 0.4: Limiter's Required Specs

Spec	Value
Bandwidth	250 MHz
Input Sensitivity	50mv
Output Level	Rail-to-Rail (0-1.2v)
Gain	30 dB
Noise Figure	20 dB
Input	Differential
Output	Single Ended

5.5.9.2. Design Methodology and Results

Based on the previous discussion, a 3-stage limiter is designed with the first two stage as differential pairs with resistive loads and the last stage is a single-ended output single stage op Amp. An inverter is added after the last stage to provide a rail-to-rail output across all corners. The inverter allow the required gain to be relaxed as the inverter can easily produce a rail-to-rail output swing. Figure 5.97 shows a circuit diagram of the proposed limiter. The testbench used for simulations is shown in Fig. 5.98. Figures 5.99 to 5.100 show the transient, AC and noise simulations across various corners. The frequency of the limiter's output is shown in Fig. 5.101. Table 5.5 shows a summary for the achieved specs.

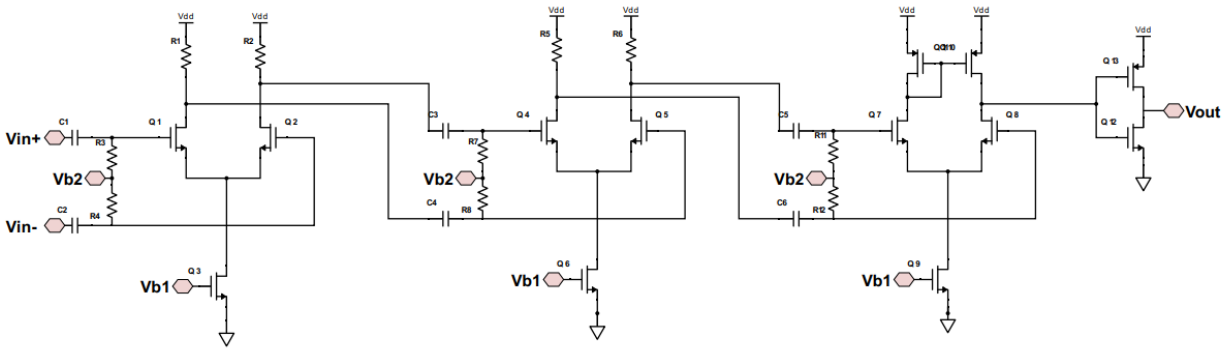


Figure 0.97: Proposed Limiter Circuit Diagram

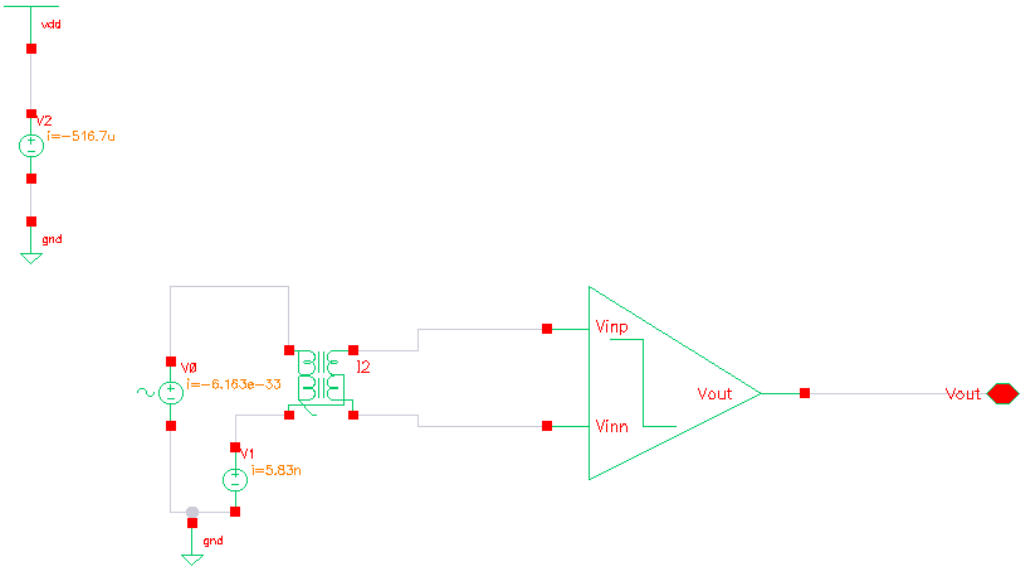


Figure 0.98: Limiter's Testbench

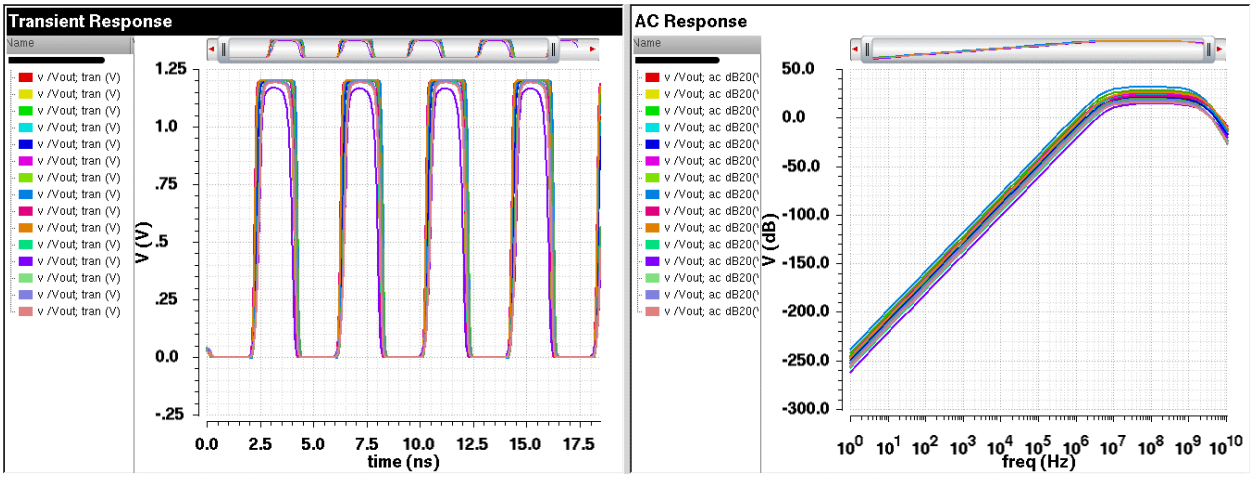


Figure 5.99: Limiter's AC & Transient Response across Corners

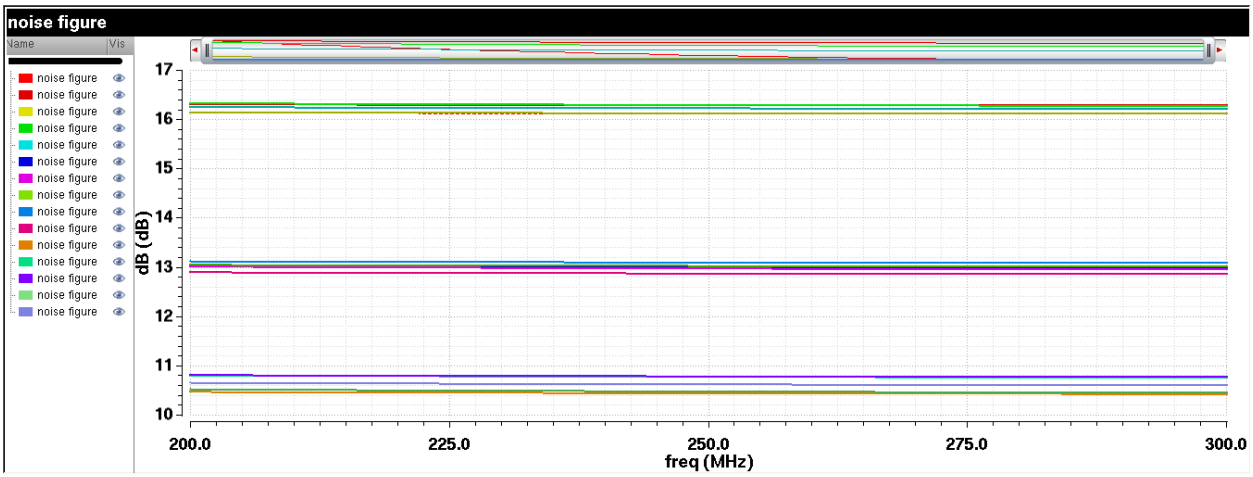


Figure 0.100: Simulated Noise Figure across Corners

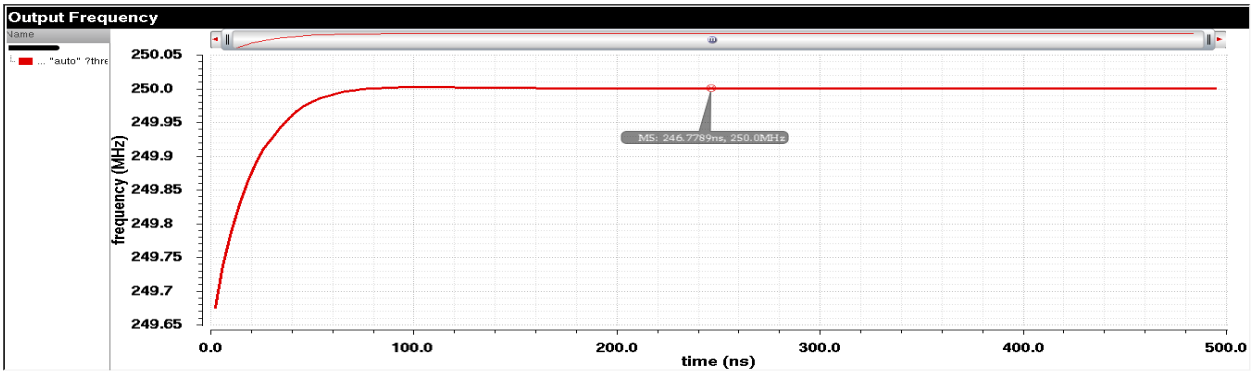


Figure 5.101: Limiter's Output Frequency

Table 0.5: Limiter's Achieved Specs

Spec	Value
Bandwidth	1.3 GHz
Output Level	Rail-to-Rail (0-1.2v)
Gain	21 dB
Noise Figure	13 dB
Current Consumption	500 uA

5.5.10. Feedback Divider Blocks Integration

The PLL's feedback divider blocks are integrated and noise and transient simulations are performed. Figure 5.102 shows the testbench for testing the three blocks. Figure 5.103 shows the waveform at the output of each block. In Fig. 5.104, the phase noise of this path is simulated. The achieved phase noise at 500 KHz offset is -128 dBc/Hz which is well below the required spec for the PLL's phase noise at this offset (-85 dBc/Hz).

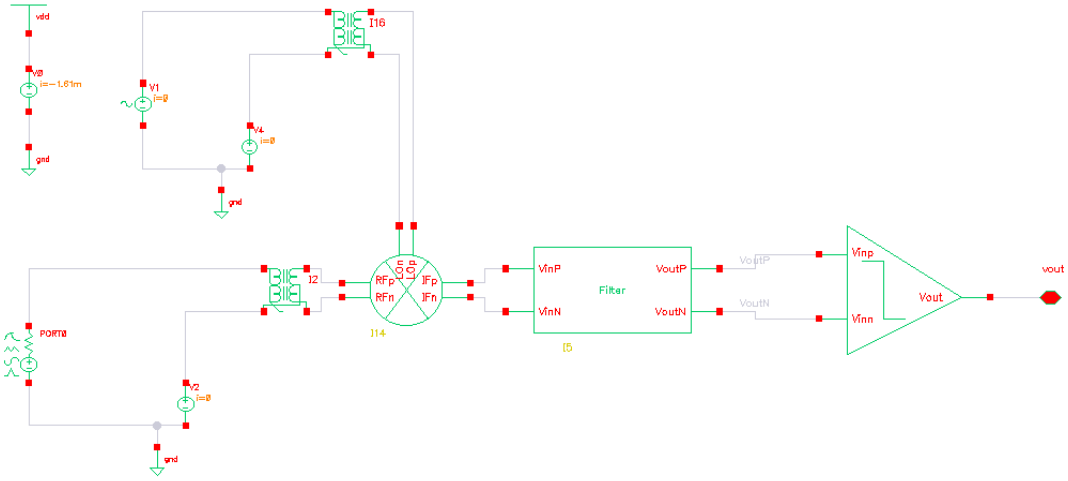


Figure 0.102: Feedback Divider Testbench

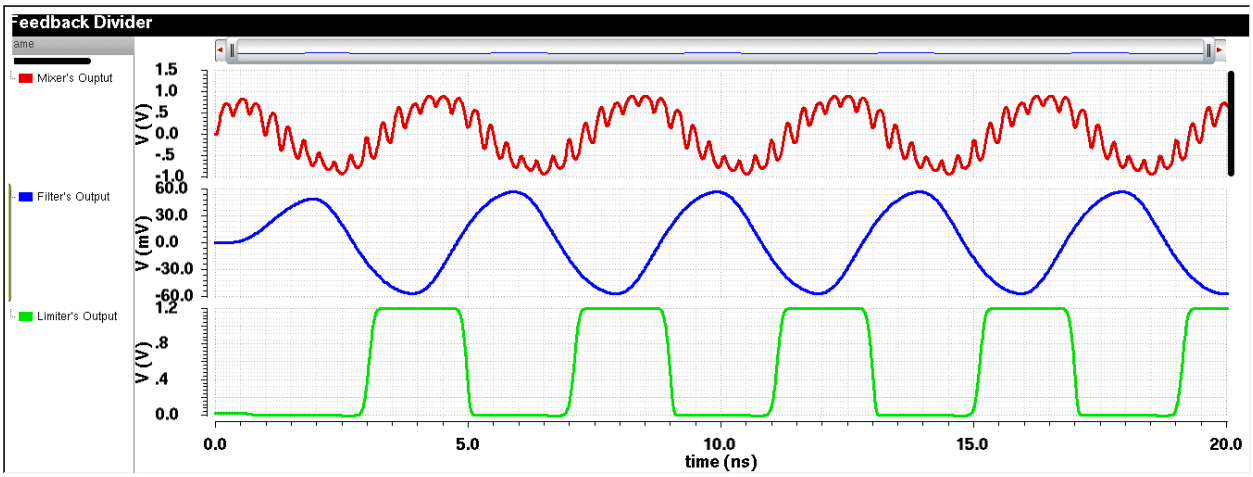


Figure 5.103: Feedback Divider Output Waveforms

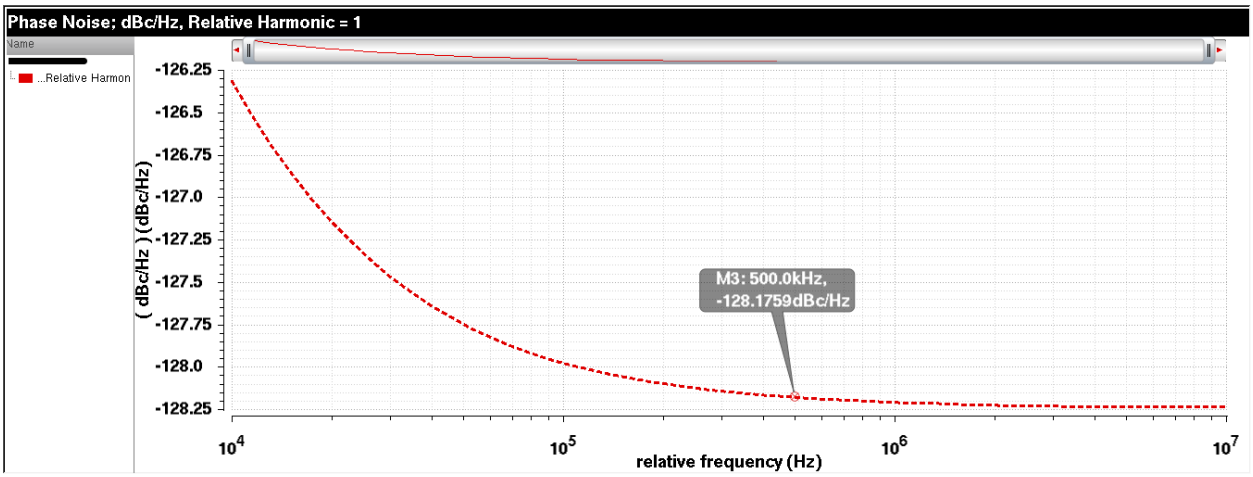


Figure 5.104: Feedback Divider's Phase Noise

5.5.11. OPLL Simulations

The control voltage during lock is shown in Fig. 5.105. The feedback from the divider path and the reference to the PLL are plotted on the same graph in Fig. 5.106 to show the tracking with minimum phase error. The phase noise of the PLL is simulated using a MATLAB code shown in appendix 4. The code is based on the noise model and transfer functions derived in [12, 33]. The phase noise meets the standard's mask spec by a margin of 16 dB at an offset of 400 KHz and a margin of 13 dB at an offset of 600 KHz from the carrier. Figure 5.107 shows the mask and PLL phase noise in dBc/Hz . The PLL consumes a total power of 6.78 mW.

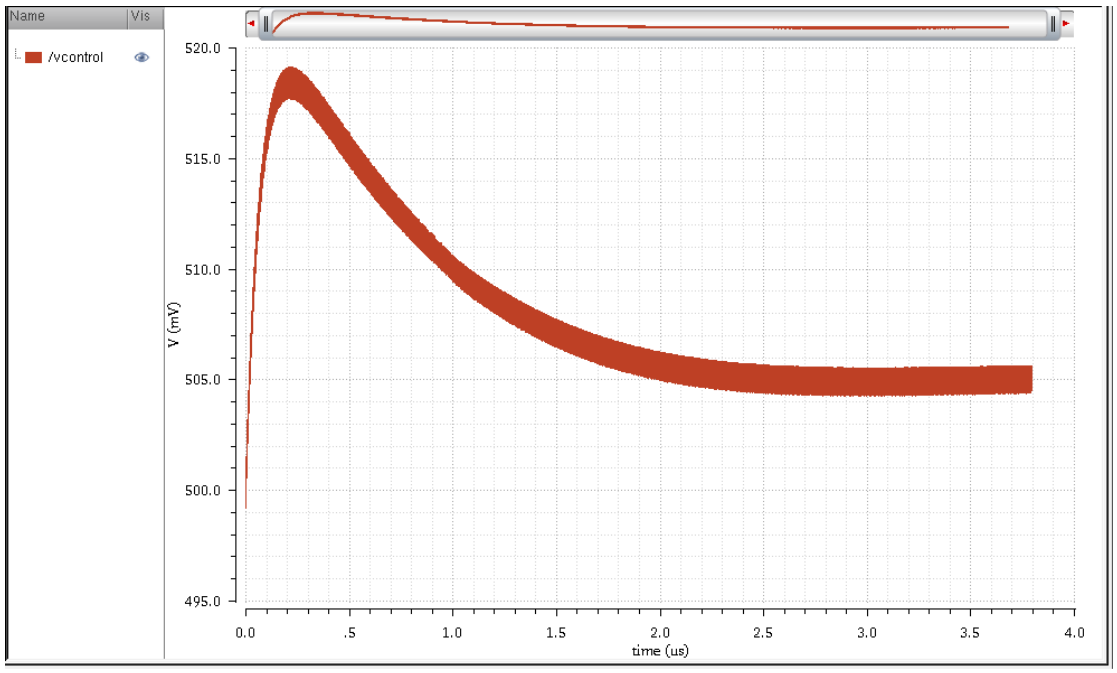


Figure 0.105: PLL Control Voltage Lock

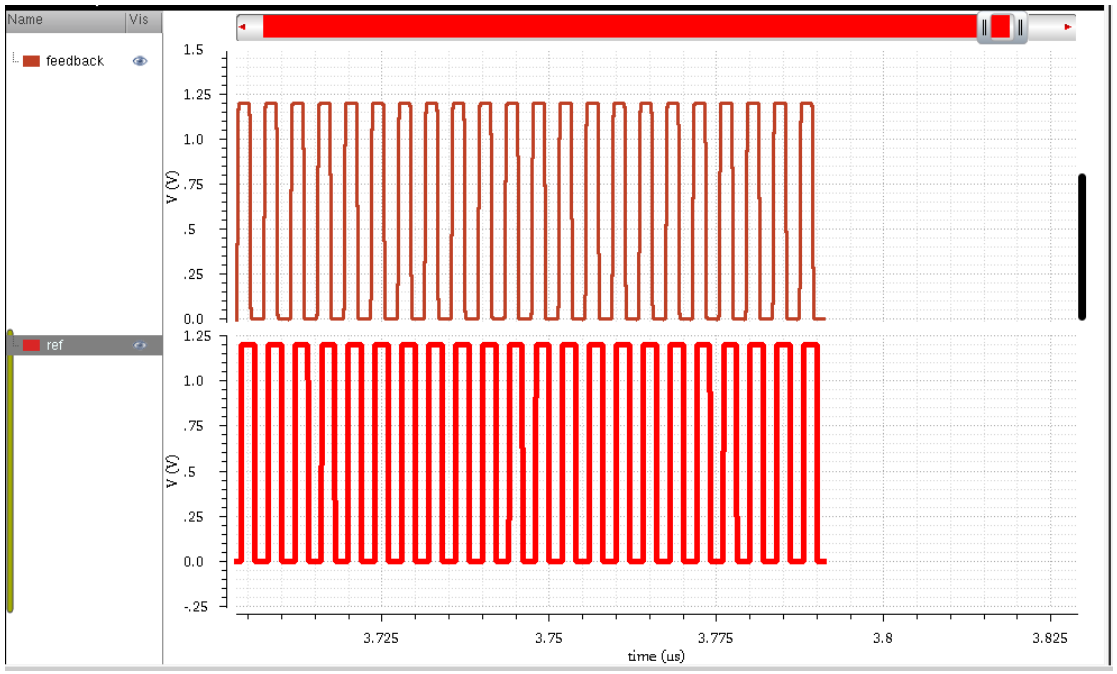


Figure 0.106: PLL Phase Tracking

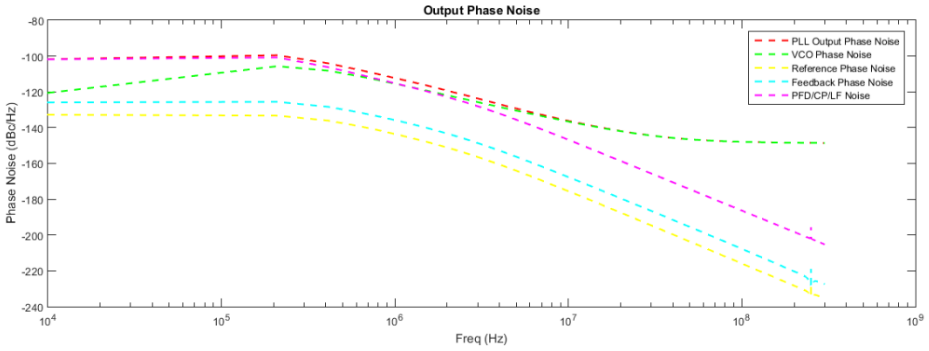


Figure 0.107: PLL Phase Noise

References

[1] B. Razavi, “RF microelectronics”, Upper Saddle River, NJ: Prentice Hall, 2012

[2] T. Yamawaki, M. Kokubo, K. Irie, H. Matsui, K. Hori, T. Endou, H. Hagisawa, T. Furuya, Y. Shimizu, M. Katagishi and J. Hildersley, "A 2.7-V GSM RF transceiver IC", IEEE Journal of Solid-State Circuits, vol. 32, no. 12, pp. 2089-2096, 1997.

[3] B. Razavi, "A study of injection locking and pulling in oscillators", IEEE Journal of Solid-State Circuits, vol. 39, no. 9, pp. 1415-1424, 2004.

[4] Johansson, H. O. (1998). A simple precharged CMOS phase frequency detector. IEEE Journal of Solid-State Circuits, 33(2), 295-299.

[5] P. Hanumolu, ECE 599: Phase locked Loops - I, Class notes for ECE 599, Department of Electrical and Computer Engineering, Oregon State University.

[6] Ping-Yuan Deng and Jean-Fu Kiang, "A 5-GHz CMOS Frequency Synthesizer With an Injection-Locked Frequency Divider and Differential Switched Capacitors", IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 56, no. 2, pp. 320-326, 2009.

[7] B. De Muer and M. Steyaert, "A CMOS monolithic ΔΣ-controlled fractional-N frequency synthesizer for DCS-1800", IEEE Journal of Solid-State Circuits, vol. 37, no. 7, pp. 835-844, 2002.

[8] M. Soyuer and R. Meyer, "Frequency limitations of a conventional phase-frequency detector", IEEE Journal of Solid-State Circuits, vol. 25, no. 4, pp. 1019-1022, 1990.

[9] N. Nguyen and R. Meyer, "Start-up and frequency stability in high-frequency oscillators", IEEE Journal of Solid-State Circuits, vol. 27, no. 5, pp. 810-820, 1992.

[10] F. Kallel, M. Fakhfakh, M. Loulou, S. Oumansour and M. Sbaa, "Modelling the frequency sensitivity Kvco of a ring oscillator", 2008 15th IEEE International Conference on Electronics, Circuits and Systems, 2008.

[11] A. Abidi, "Phase Noise and Jitter in CMOS Ring Oscillators", IEEE Journal of Solid-State Circuits, vol. 41, no. 8, pp. 1803-1816, 2006.

[12] B. Razavi, "Phase-locking in high-performance systems". Hoboken: Wiley-Interscience, 2003.

[13] D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," Proc. IEEE, vol. 54, pp. 329–330, Feb. 1966

[14] A. Hajimiri and T. H. Lee, "A General Theory of Phase Noise in Electrical Oscillators," IEEE J. of Solid-State Circuits, vol. 33, pp. 179–194, Feb. 1998

[15] M. Perrott, "Short Course On Phase-Locked Loops and Their Applications", 2008. [Online]. Available: http://www.cppsim.com/PLL_Lectures/day2_pm.pdf

[16] S.-G. Lee and J.-K. Choi, "Current-reuse bleeding mixer," Electron. Lett., vol. 36, no. 8, pp. 696–697, Apr. 2000

[17] Hampel, S., Schmitz, O., Tiebout, M. and Rolfes, I. (2010). Inductorless Low-Voltage and Low-Power Wideband Mixer for Multistandard Receivers. IEEE Transactions on Microwave Theory and Techniques, 58(5), pp.1384-1390.

[18] G. Kathiresan and C. Toumazou, "A low voltage bulk driven downconversion mixer core," in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS),vol. 2. Jul. 1999, pp. 598–601.

[19] F. Mahmoudi and C. Salama, "8 GHz, 1 V, high linearity, low power CMOS active mixer," in Proc. IEEE Radio Freq. Integr. Circuits Symp. Dig., Jun. 2004, pp. 401–404.

[20] V. Vidojkovic, J. van der Tang, A. Leeuwenburgh, and A. van Roermund, "A high gain, low voltage folded-switching mixer with current-reuse in 0.18 m CMOS," in Proc. IEEE Radio Freq. Integr. Circuits Symp. Dig., Jun. 2004, pp. 31–34.

[21] SpectreRF Workshop, Mixer Design Using SpectreRF, Product Version 6.0, November 2005.

[22] SpectreRF Workshop, Mixer Design Using SpectreRF, Product Version 14.1, September 2014.

[23] Silva-Martinez, M. Steyaert and W. Sansen, High-performance CMOS continuous time filters. Boston: Kluwer academic publ., 1993.

[24] A. Poularikas, The handbook of formulas and tables for signal processing. Boca Raton, Fla.: CRC Press, 1999.

[25] J. E. Kardontichik, "Introduction to the Design of Transconductance Capacitor Filters", Kluwer Academic Publishers, 1992

[26] S. Galal, and B. Razavi, (2003). "10-Gb/s limiting amplifier and laser/modulator driver in 0.18-μm CMOS technology". IEEE Journal of Solid-State Circuits, 38(12), pp.2138-2146.

[27] Garcia del Pozo, J., Celma, S., Otín, A., Lope, I. and Urdangarín, J. (2010). "1.8V–3GHz CMOS limiting amplifier with efficient frequency compensation". Microelectronics Reliability, 50(12), pp.2084-2089.

[28] Sackinger, E. and Fischer, W. (2000). "A 3-GHz 32-dB CMOS limiting amplifier for SONET OC-48 receivers". IEEE Journal of Solid-State Circuits, 35(12), pp.1884-1888.

[29] Chen, Y., Wu, Y. and Huang, P. (2007). "A 1.2-V CMOS Limiter / RSSI / Demodulator for Low-IF FSK Receiver". 2007 IEEE Custom Integrated Circuits Conference.

[30] Po-Chiun Huang, Yi-Huei Chen and Chorng-Kuang Wang (2000). "A 2-V 10.7-MHz CMOS limiting amplifier/RSSI". IEEE Journal of Solid-State Circuits, 35(10), pp.1474-1480.

[31] B. Razavi, (2012). "Design of Integrated Circuits for Optical Communications (2nd Edition)". Wiley.

[32] Duy-Dong Pham, James Brinkhoff, Kai Kang, Chyuen-Wei Ang, Fujiang Lin, Chun-Huat Heng. (2009). "Feedforward technique for offset cancellation in broadband differential amplifiers". Proceedings of the 2009 12th International Symposium on Integrated Circuits, pp.429-432.

[33] K. Kundert. (2002). "Predicting the phase noise and jitter of PLL-based frequency synthesizers". [Online]. Available: www.designers-guide.com

Chapter 6 : Power Amplifier

6.1. Introduction

Power consumption is an increasingly important issue in highly integrated wireless systems. While advances in semiconductor technology have driven continuous integration of features and services into portable devices, power consumption is now a major limiting factor on computational complexity and the ability to communicate over long distances. In portable communication devices, the wireless transmitter is often the dominant source of power consumption, such that in recent years there has been a major effort to improve the power efficiency of transmitter circuits, especially the power amplifier (PA).

In addition to power consumption, it is now apparent that energy consumption is an important metric for transmitter circuits. Energy consumption more accurately predicts the battery life, especially when a portable system operates with a wide range of output power.

A power amplifier is a key element to build a wireless communication system successfully. Power amplifiers are required to be more power efficient and reconfigurable. To minimize interference and spectral re-growth, transmitters must be linear. Power amplifiers are the most power-hungry building block of RF transceivers and pose difficult design challenges.

6.2. General Considerations

6.2.1. Efficiency

Since PAs are the most power-hungry block in RF transceivers, their efficiency is critical. The efficiency of the PAs is defined by two metrics:

- The drain efficiency " η " is defined as:

$$\eta = \frac{P_L}{P_{DC}} \qquad (6.1)$$

where P_L the average power delivered to the load and P_{DC} the average power drawn from the supply voltage.

- Power added efficiency " PAE " is defined as:

$$\text{PAE} = \frac{P_L - P_{in}}{P_{DC}} \qquad (6.2)$$

where P_{in} is the average input power.

6.2.2. Linearity

For PA there is no small signal but large signal so the non-linearity means a crazy amount of distortion. PA nonlinearity leads to two effects:

- High adjacent channel power as a result of spectral regrowth.
- Amplitude compression (AM/PM).

.

A modulated waveform

$$x(t) = A(t)\cos[\omega_c t + \varphi(t)] \qquad (6.3)$$

is said to have a constant envelope if A(t) does not vary with time. Otherwise, we say the signal has a variable envelope. Constant- and variable-envelope signals behave differently in a nonlinear system. This is an undesirable effect in nonlinear PAs. Called "spectral regrowth," this effect corrupts the adjacent channels. We have mentioned that variable-envelope signals require linear PAs, whereas constant-envelope signals do not [1]. The Figure 6.1 illustrated the spectrum regrowth due to two input signal with different modulation schemes on non-linear PA.

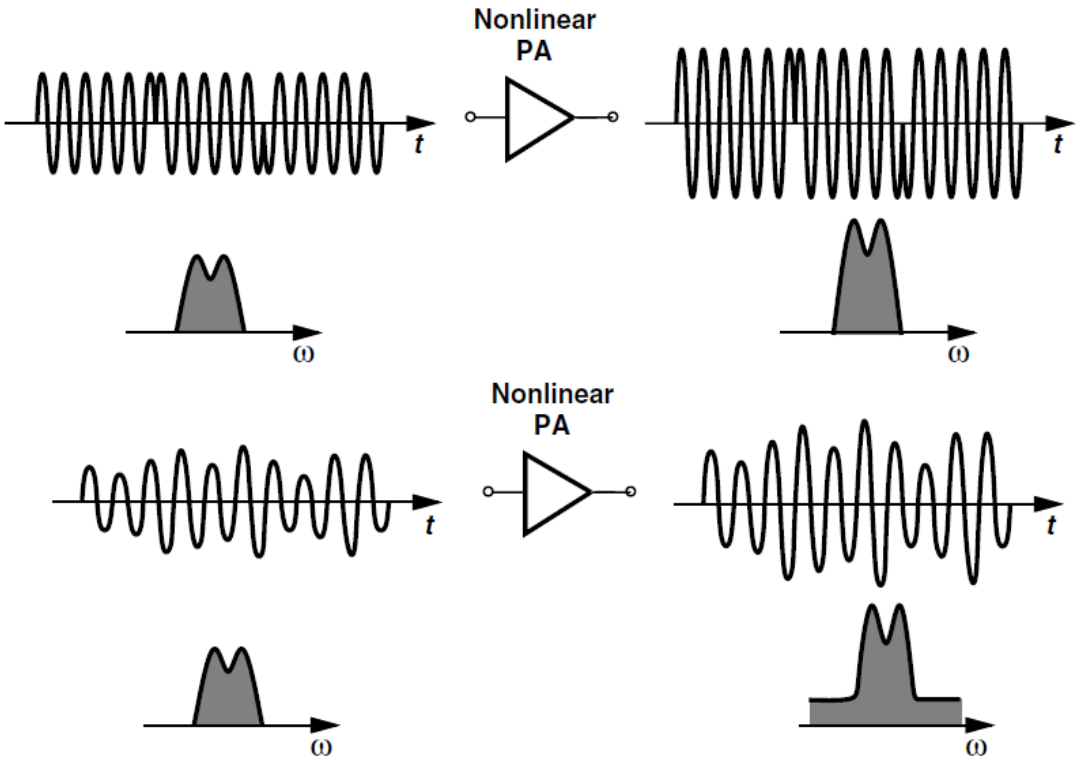


Figure 6.1: Amplification of constant- and variable-envelope signals and the effect on their spectra.

6.2.3. Cascode Output Stages

For linear Classes efficiency calculations, the drain waveform is assumed to have a peak-to-peak swing of nearly 2VDD. However, if VDD is chosen equal to the nominal supply voltage of the process, the output transistor experiences breakdown or substantial stress. One can choose VDD equal to half of the maximum tolerable voltage of the transistor, but with two penalties:

(a) The lower headroom limits the linear voltage range of the circuit, and

(b) The proportionally higher output current (for a given output power) leads to a greater loss in the output matching network, reducing the efficiency.

A cascode output stage relaxes the above constraints by shielding the input transistor as V_X rises, keeping the drain-source voltage of M1 less than $V_b - V_{TH2}$. V_X swings by about 2VDD and V_Y by about $V_b - V_{TH}$ as shown in Figure 6.3.

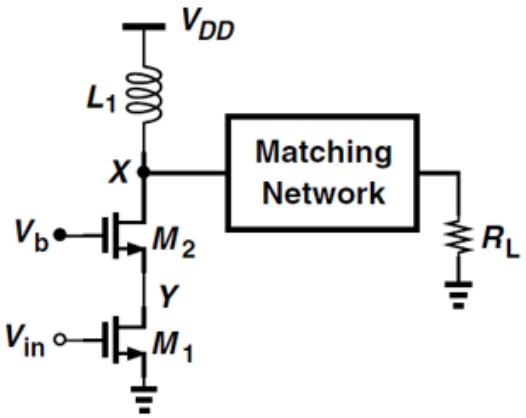


Figure 6.2: Cascode PA

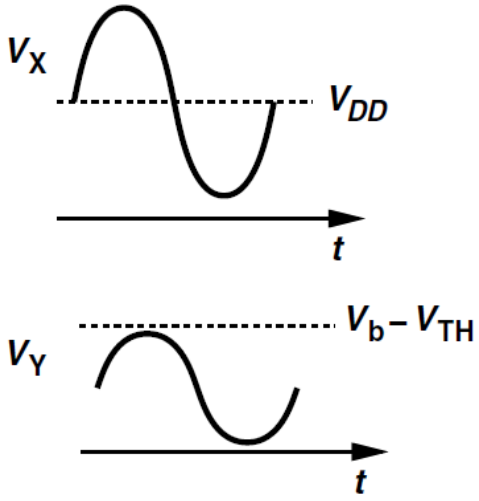


Figure 6.3: Voltage of node X and Y

6.2.4. Another Power Amplifier Considerations

- PAs drive large voltages/currents into small load impedances. Thus matching networks are critical. Any loss in the matching network has a severe impact on the efficiency of the amplifier.
- Heat generation is high. We need to carefully provide heat sinks to keep the junction temperatures as low as possible. Due to the interface with the external "off-chip" world, packaging and board parasitics are very important.
- The spectral "leakage" and harmonic generation in a PA must be kept to the minimum in order to minimize interference to other users.

6.3. Classification OF Power Amplifier

Research into wireless power amplifiers (PAs) has long been focused on improving the tradeoffs between efficiency and linearity. Switching power amplifiers (Classes D/E etc.) have higher efficiency compared with their linear counterparts (Classes A/AB/B etc.), but cannot be directly applied in systems which employ spectrally efficient modulation. It is for this reason a great deal of research has focused on breaking this tradeoff.

There are many classes of power amplifier, distinguished by the mode of operation of the active device, the current conduction angle, and the topology of resonant network at the output.

Classes A, AB, B are considered linear power amplifiers, since the amplitude of the output signal is proportional to the amplitude of the input signal.

Classes C, D, E, and F are considered nonlinear since input−output amplitude linearity is not inherent, or it is not possible to modulate the output amplitude from the input terminal.

6.3.1. Linear Classes

Linear power amplifier classes are characterized by RF power amplifier in which the transistor is operated as a voltage- or current-dependent current source. This allows these classes to amplify variable-envelope signals properly.

6.3.1.1. Class A

The Class A power amplifier is a linear amplifier. A linear amplifier is supposed to produce an amplified replica of the input voltage or current waveform. It provides an accurate reproduction of both the envelope and the phase of the input signal as shown in Figure 6.5. The transistor in the Class A amplifier is operated as a dependent-current source [2].

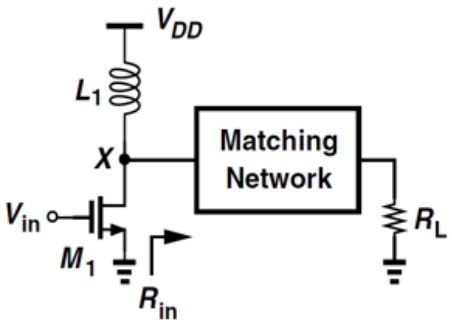


Figure 6.4: Class A

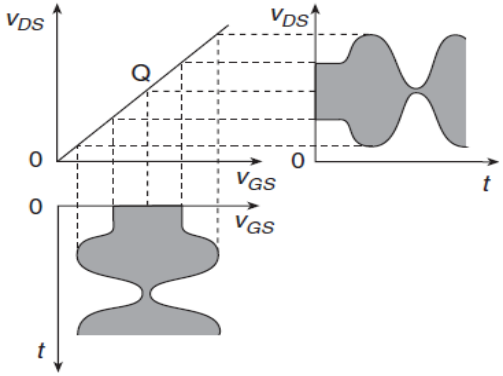


Figure 6.5: Amplification of AM (Class A)

Concept of operation

- The conduction angle is defined as the percentage of the signal period during which the transistor(s) remain on multiplied by 360◦. In class A stages, the conduction angle is 360◦because the output transistor is always on.

- The operating point of the transistor in the Class A amplifier is above the transistor threshold voltage such that the gate-to-source voltage waveform is above for the entire voltage swing. Therefore, the transistor never enters the cutoff region.

- Power loss in the transistor of the Class A RF power amplifier is high.
  The Class A power amplifier dissipates the maximum power at zero output power. The maximum drain efficiency of Class A RF power amplifiers with the RF choke is 50%.

6.3.1.2. Transformer less Class B

Concept of Operation

The Class B RF power amplifier consists of a transistor and a parallel-resonant circuit. The transistor is operated as a dependent current source. The conduction angle of the drain or collector current in the Class B power amplifier is 180◦. The parallel resonant circuit acts like a bandpass filter and selects only the fundamental component. The efficiency of the Class B power amplifier is higher than that of the Class A power amplifier.

The circuit of a Class B RF power amplifier is shown in Figure 6.6. It consists of a transistor, parallel-resonant circuit, and RF choke. The operating point of the transistor is located exactly at the boundary between the cutoff region and the active region The dc component of the gate-to source voltage is equal to the transistor threshold voltage.

Class B amplifier works with a single devices by sending half sinusoid current pulses to the load. The device is biased at the edge of conduction. The load voltage is sinusoidal because a high Q RLC tank shunts harmonics to ground.

In a single transistor version, the “minus” pulse is in fact delivered by the RLC tank as shown in Figure 6.7. The Q factor of the tank needs to be large enough to do this. This is analogous to pushing someone on a swing (for any second order system). You only need to push in one direction, and the reactive energy stored will swing the person back in the reverse direction.

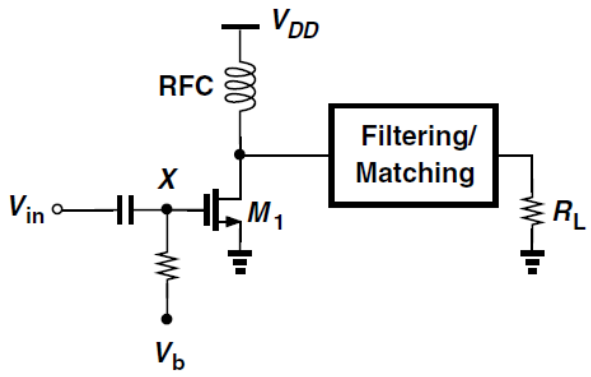


Figure 6.6: Transfer less Class B

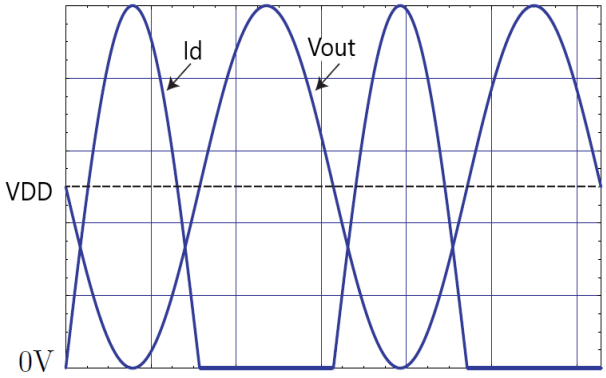


Figure 6.7: Class B waveforms

6.3.1.3. Class AB

The term “class AB” is sometimes used to refer to a single ended PA whose conduction angle falls between 180◦ and 360◦in which the output transistor turns off for less than half of a period. From another perspective, a class AB PA is less linear than a class A stage and more linear than a class B stage.

6.3.1.4. Class C

The circuit of the Class C power amplifier is the same as that of the Class B RF power amplifier. The operating point of the transistor is located in the cutoff region. The dc component of the gate-to-source voltage is less than the transistor threshold voltage. Therefore, the conduction angle of the drain current is less than 180◦. Voltage and current waveforms in the Class C power amplifier are shown in Figure 6.9. The only difference is the conduction angle of the drain current, which is determined by the operating point.

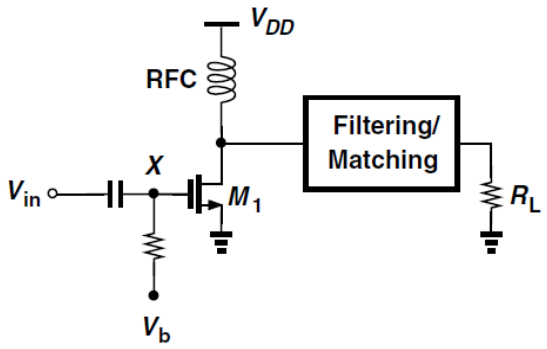


Figure 6.8: Class C

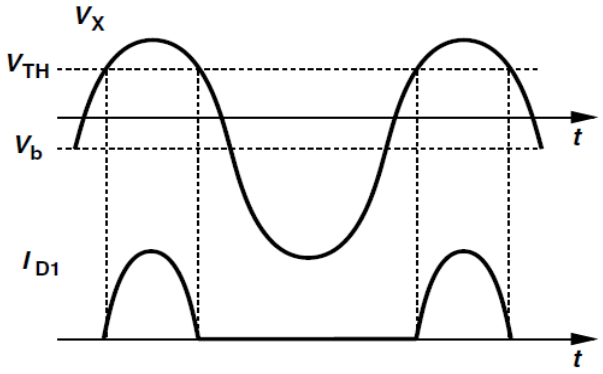


Figure 6.9: Class C waveforms

Efficiency of the Class AB, B, and C Amplifiers

The efficiency of the Class AB, B, and C amplifiers is given by:

$$\eta = \frac{1}{4}\frac{\theta - \text{Sin}\,\theta}{\text{Sin}(\theta/2) - (\theta/2)\,\text{Cos}(\theta/2)} \qquad (6.4)$$

The formula assumed that the drain current of transistor in Figure 6.8 is to be the peak section of a sinusoid and the drain voltage a sinusoid having a peak amplitude of VDD.

The output power of the Class AB, B, and C amplifiers is given by:

$$P_{out} \;\alpha\; \frac{\theta - \text{Sin}\,\theta}{1 - \text{Cos}(\theta/2)} \qquad (6.5)$$

The maximum efficiency of 100% is often considered a prominent feature of class C stages. However, another attribute that must also be taken into account is the actual power delivered to the load.

The Pout falls to zero as θ approaches zero. In other words, for a given design, a class C stage provides a high efficiency only if it delivers a fraction of the peak output power (the power corresponding to full class A operation).

The small conduction angle dictates that the output transistor be very wide so as to deliver a high current for a short amount of time. In other words, the first harmonic of the drain current must be equal in the two cases [1].

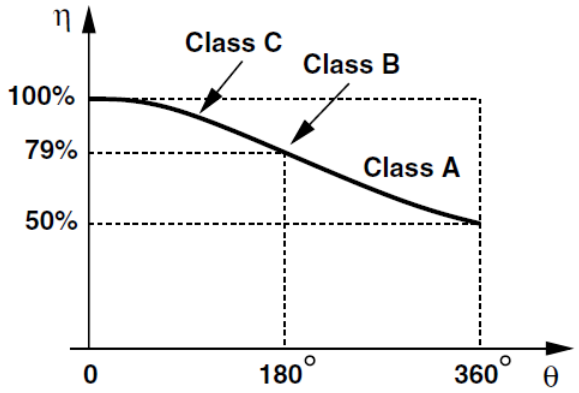


Figure 6.10: Efficiency as a function of conduction angle.

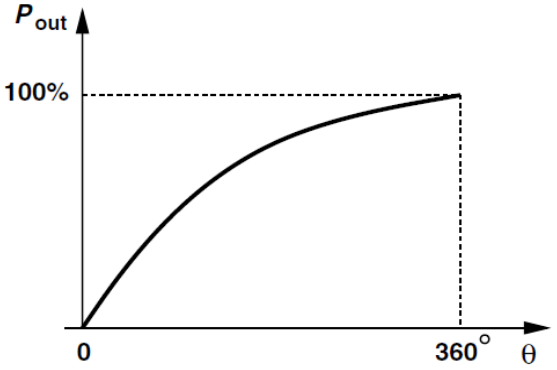


Figure 6.11: Output power as a function of conduction angle

6.3.2. Non-Linear Classes

Switching power amplifiers can achieve higher efficiency than linear amplifiers because the active device operates as a switch. The active device is usually driven

hard enough that it operates as a resistor in the ‘on’ state and conducts no current in the ‘off’ state.

6.3.2.1. Class D

The basic Class-D PA is depicted in Figure 6.12. The input signal and the drain voltage are both square waves. If this square wave was applied directly to the load resistor R_l , significant harmonic power would be wasted. Hence, a resonant tank consisting of L_o and C_o is inserted in series with in order to allow only sinusoidal current through R_l .

A series LCR filter only allows the first harmonic of voltage to flow into the load. Since current only flows through a device when it’s fully on (ideally V_{ds} = 0), little power dissipation occurs in the devices.

There is ideally no overlap between the current and voltage waveforms as shown in Figure 6.14, thereby in principle achieving 100% efficiency.

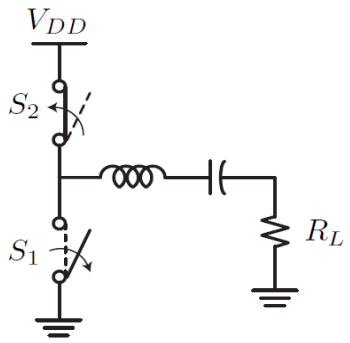


Figure 6.12: Basic idea of Class D

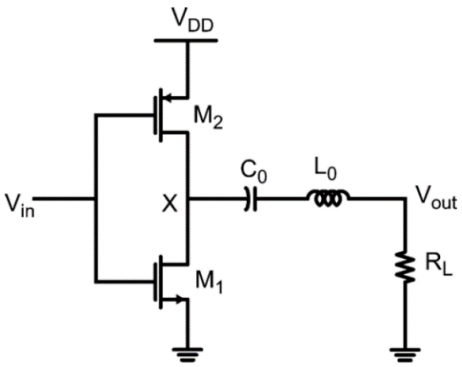


Figure 6.13: Class D

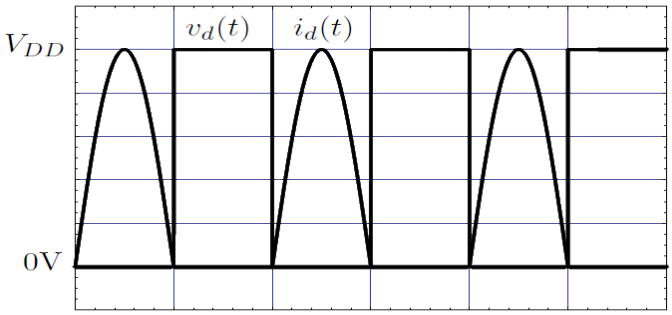


Figure 6.14: Class D waveforms

6.3.2.2. Class F

An ideal class F shapes the waveform of voltage into a square wave, while the current is shaped as a half current waveform. The wave shaping is done by adding odd harmonics to voltage and adding even number of harmonics to current waveform.

As shown in Figure 6.15, the circuit isolates the fundamental resonant load from the drain at the third harmonic, allowing the drain voltage to contain enough third harmonic to create a more square like waveform as shown in Figure 6.17. It can be shown that just adding a third harmonic boost the efficiency from 78% (Class B) to 88%.

As shown in Figure 6.16, the circuit a quarter wave transformer converts the low impedance at the load (due to the capacitance) at harmonics of the fundamental to a high impedance for all odd harmonics. Even harmonics are unaltered as they see a $\lambda/2$ line. In theory then we can create a perfect square wave at the drain of the transistor and thus achieve 100% efficiency.

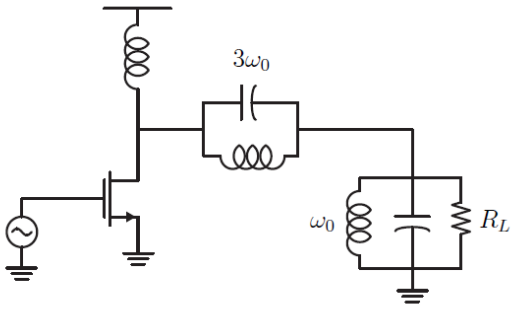


Figure 6.15: Class F

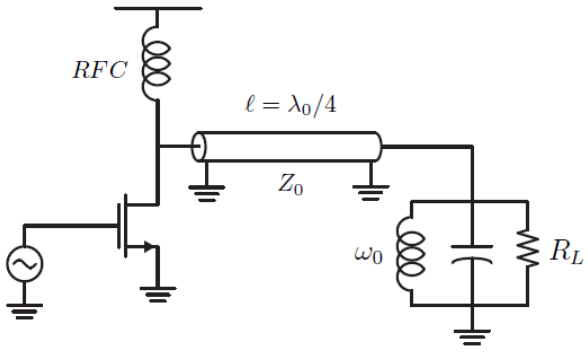


Figure 6.16: Class F with a quarter wave transformer

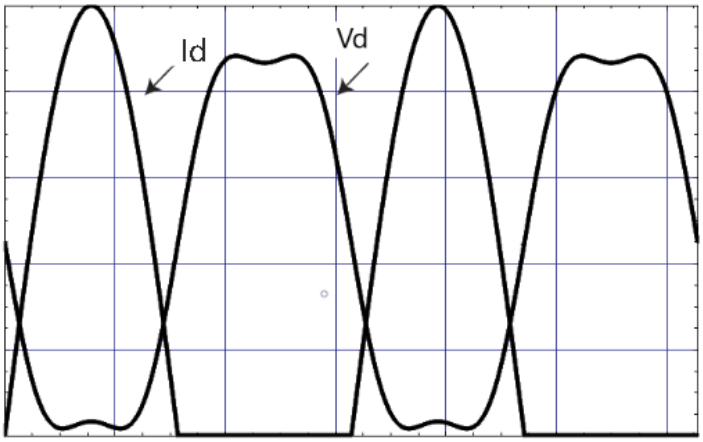


Figure 6.17: Class F waveforms

6.3.2.3. Class E

In Class E amplifiers, the transistor is operated as a switch. Class E power amplifiers are the most efficient amplifiers known so far. The current and voltage waveforms of the switch are displaced with respect to time, yielding very low power dissipation in the transistor. Since the switch current and voltage waveforms do not overlap during the switching time intervals, switching losses are virtually zero, yielding high efficiency.

The output network is designed to minimize power loss by preventing overlapping V and I waveforms. Ideally, the active device turns on when the drain−source voltage is zero, achieving zero voltage switching. This reduces power loss in the system by minimizing the loss incurred from switching the drain capacitance of the active device. There are two types of Class E power amplifiers

(1) Class E zero-voltage-switching (ZVS) power amplifiers.
(2) Class E zero-current-switching (ZCS) power amplifiers.

The basic circuit of the Class E ZVS power amplifier is shown in Figure 6.18. It consists of power MOSFET operating as a switch, L-C-R series-resonant circuit, shunt capacitor C1 and choke inductor RFC.

The switch turns on and off at the operating frequency. The transistor output capacitance, the choke parasitic capacitance, and stray capacitances are included in the shunt capacitance C1. For high operating frequencies, all of the capacitance C1 can be supplied by the overall shunt parasitic capacitance.

The resistor R is an Ac load. The choke inductance RFC is assumed to be high enough so that the ac current ripple on the dc supply current can be neglected. A small inductance with a large current ripple is also possible.

The device is configured similarly to the linear PA device, except that the gate signal drives the device between triode and cutoff regions. As shown in Figure 6.19, when the device is ‘on,’ the drain−source voltage is constrained to near zero. When the device is ‘off,’ the drain− source voltage is constrained by the dynamics of the output network as shown in Figure 6.20.

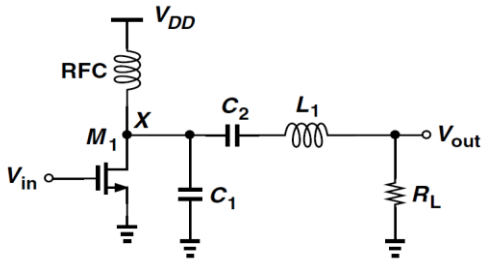


Figure 6.18: Class E

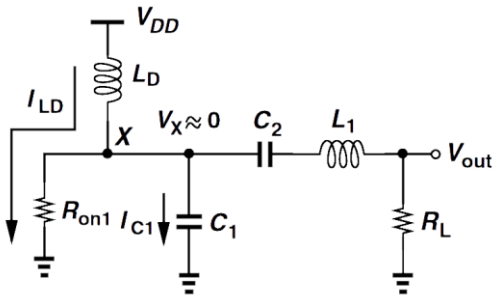


Figure 6.19: The transistor is on

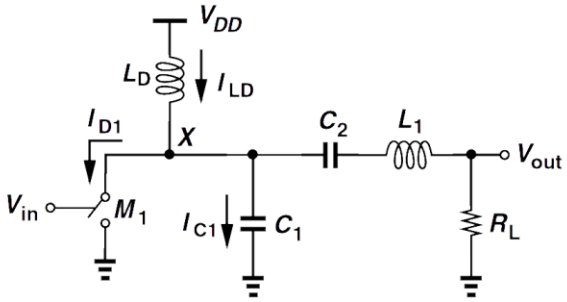


Figure 6.20: The transistor is off

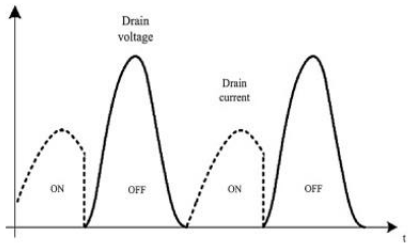


Figure 6.21: The Class E wave forms

As shown in Figure 6.21, No time overlap between transistor's drain source voltage and drain current and the first derivative of VDS is zero at the moment device turns on. The requires that both the switch voltage and its first derivative are zero when the switch closes. The requirement for a zero first derivative is not absolutely necessary to obtain 100% efficiency. However, this property makes the amplifier less sensitive to component variations

Table 6.1: Key concerns and tradeoffs in non-linear classes

Class	Main features
Class F	• The efficiency of class F depends on the number of harmonics added, as the number of harmonics increase the efficiency increase. • Adding infinite number of harmonic makes the design of output network too complex, most designers uses finite number of harmonics and sacrifice some losses as using transmission line may be inconveniently long or even inapplicable in fully on-chip integration . • Class F main trade is between output network complexity and efficiency. • The maximum efficiency of an ideal PA increase from 50% of class A to 70.7, 81.6, 86.6 and 90.4 as harmonics added [3].
Class D	For gigahertz-range applications the Class-D amplifier has some drawbacks [4]: • The drain capacitors of the transistors are not a part of the matching network and need to be charged and discharged every cycle, leading to power dissipation equal $C\ V_{DD}{}^2\ f$ • Class-D PA is the that it requires the use of a PMOS device, which has higher on-resistance than its NMOS counterpart. Typically, to achieve similar on-resistance as the NMOS switch, the PMOS switch size needs to be increased two to three times, thereby increasing input capacitance significantly and making driver design more challenging. • The power $C\ V_{DD}{}^2\ f$ required to drive the switches also increases proportional to C_{gs} since we have to burn power to drive the inverter.
Class E	• Class-E amplifiers have become popular since they can overcome the $C\ V_{DD}{}^2\ f$ loss associated with the drain parasitics of the transistor. Like Class D, Class E is also capable of achieving 100% efficiency. • The main advantage of Class-E design is that, when the switch is turned on, the voltage across it is zero, and therefore the drain parasitic capacitance of the switch need not be discharged. This is referred to as zero voltage switching. • In fact, the parasitic drain–source capacitance of the switch can be absorbed into the shunt capacitor required in a Class-E PA. The Class-E theory requires that both the switch voltage and its first derivative are zero when the switch closes. The requirement for a zero first derivative is not absolutely necessary to obtain 100% efficiency. However, this property makes the amplifier less sensitive to component variations. • In spite of being an attractive architecture, Class E suffers from some practical issues in fully integrated designs: ❖ It requires multiple passive components, whose finite quality factor quickly degrades the achievable efficiency. Multiple passives also mean increased silicon area and higher cost, which are undesirable for a portable solution. ❖ The Class-E amplifier also uses a series inductor, which is not readily compatible with baluns or transformers that are in any case required for differential PAs to interface to a single-ended antenna [4].

6.4. Linearization techniques

PAs designed for a high efficiency suffer from considerable nonlinearity. For relatively low output power levels we may simply back off from the PA's 1-dB compression point until the linearity reaches an acceptable value. The efficiency then falls significantly but the absolute power drawn from the supply may still be reasonable. For higher output power levels, however, a low efficiency translates to a very large power consumption.

6.4.1. Back-off

Transmit less power to achieve higher linearity by avoiding the saturation nonlinearity

Main features

- Back-off is the simplest linearization technique.
- The required back off depends on modulation type, AM-to-AM and AM to-PM distortion levels.
- Efficiency is sacrificed to achieve desired linearity.
- Low cost and no added complexity.

A great deal of effort has been expended on linearization techniques that offer a higher overall efficiency than back-off from the compression point does. As we will see, such techniques can be categorized under two groups: those that require some linearity in the PA core, and those that, in principle, can operate with arbitrarily nonlinear stages. We expect the latter to achieve a higher efficiency.

6.4.2. Predistortion

Compensate for amplifier gain and phase variation over power range by applying an inverse nonlinearity to the input signal prior to entering the amplifier as shown in Figure 6.22.

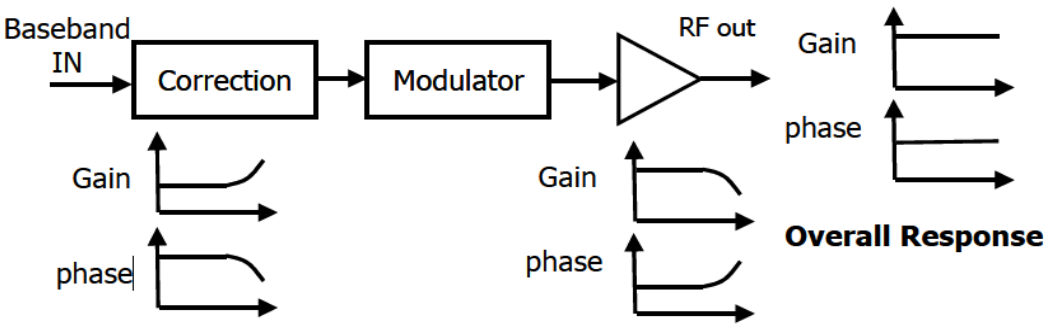


Figure 6.22: Realization in baseband

Main features

- Look up table is perfectly suited for open loop predistortion.
- Process and temperature variations as well as long term drifts usually mandate the need for adaptation.

- Adaptive predistortion systems can be complex.
- Variations in the antenna impedance somewhat affect the PA nonlinearity, but predistortion provides a fixed correction [5].

6.4.3. Feed forward

Correct the nonlinearity by subtracting an estimate of nonlinearity induced artifacts from the output as shown in Figure 6.23 [5].

Main Feature

- Requires a linear auxiliary PA to amplify the error signal.
- Requires matching of the delays and attenuation/gain behaviors in order to function correctly.
- Efficiency is low due to losses in the output delay line, output power combiner and the auxiliary amplifier.
- Sensitivity to parameter variations over temperature and time usually mandates the need for complex adaptive schemes.
- the analog delay elements introduce loss if they are passive or distortion if they are active, a particularly serious issue as it carries a full-swing signal.

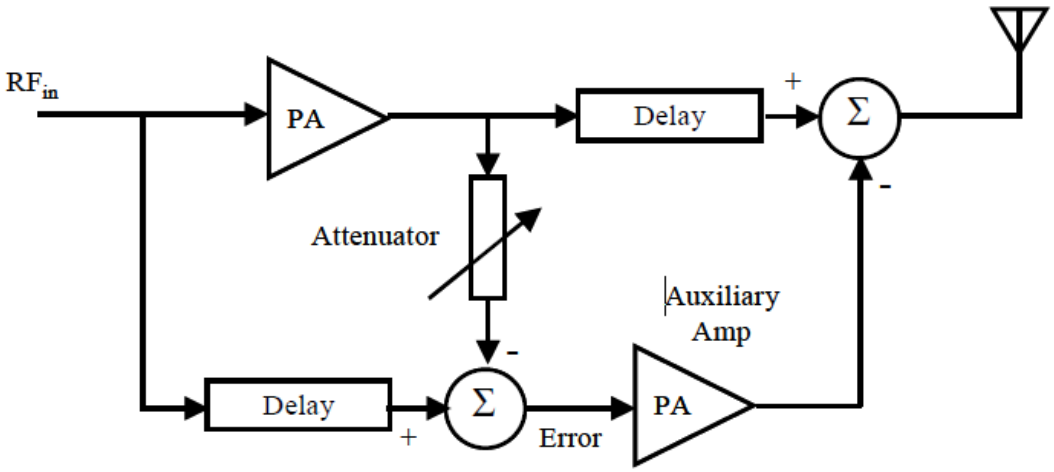


Figure 6.23: Feedforward linearization

6.4.4. Polar Feedback

Using feedback from the output, detect magnitude and phase information and use it to correct the signal fed into the amplifier by comparing it with that of the desired input signal as shown in Figure 6.24 [5].

Main features

- Two feedback loops: One for phase and one for magnitude. Stability is a concern.
- Bandwidth of the envelope loop should be reasonably larger than the envelope maximum frequency.
- When AM-to-PM is manageable, phase feedback may be eliminated.
- High performance envelope detection is required.

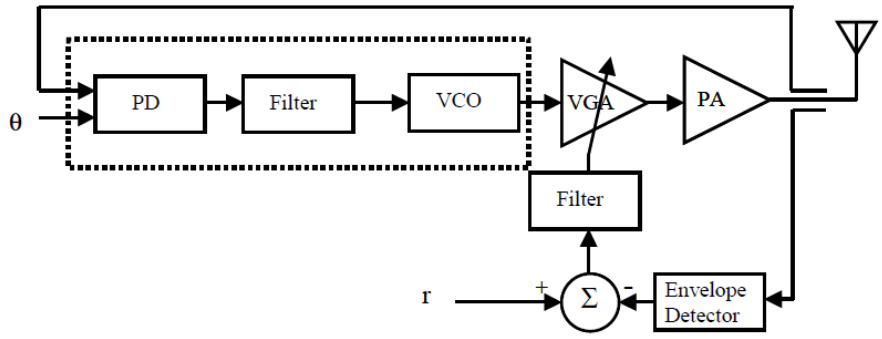


Figure 6.24: Polar feedback Linearization

6.4.5. Cartesian Feedback

Feedback a sample of the output, demodulate and detect in-phase and quadrature components and use them to correct the signals fed into the amplifier by comparing them with those of the desired signal as shown in Figure 6.25 [5].

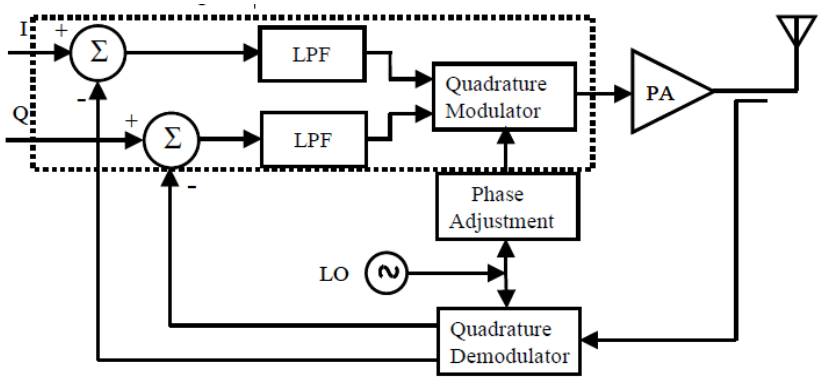


Figure 6.25: Cartesian Feedback Linearization

Main features

- Two feedback loops on I and Q. Stability is a concern.
- I and Q phases have to be aligned
- Loop bandwidth has to be reasonably higher than I and Q channel bandwidths.

6.4.6. Linear amplification using nonlinear components (Out Phasing)

Modulate the amplitude by combining two out-phased constant envelope amplified signals whose phase difference contains the amplitude information as shown in Figure 6.26 [5].

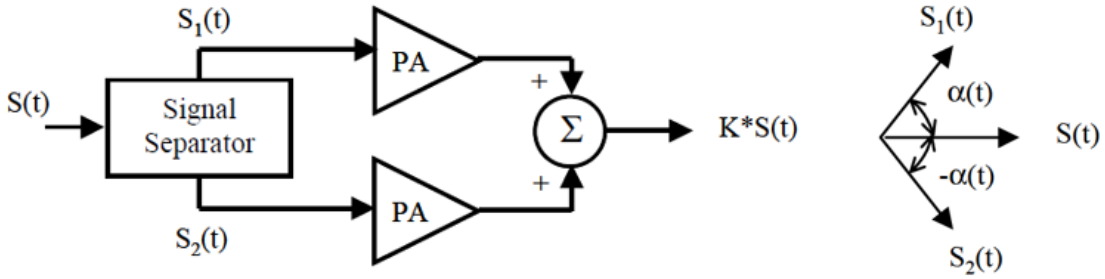


Figure 6.26: Basic Out Phasing

The equation can be expressed as:

$$S(t) = \text{b(t)}\, Cos(wt + \varphi(t)) \quad (6.6)$$

$$S_1(t) = \text{A}\, Cos(wt + \varphi(t) + \alpha(t)) \quad (6.7)$$

$$S_2(t) = \text{A}\, Cos(wt + \varphi(t) - \alpha(t)) \quad (6.8)$$

$$\alpha(t) = Cos^{-1}\,(b(t)/\varphi(t)) \quad (6.9)$$

Main features

- When fully matched hybrid combiners used, efficiency is proportional to the average output power due to the loss in the hybrid.
- Lossless combining on the other hand stresses the individual amplifiers reducing their efficiency, and degrades the overall linearity.
- Not efficient for modulations with large peak to average ratio
- Difficult to implement the signal separator with analog components.

6.4.7. Polar Modulation

Polar Transmitters main idea is based on the Khan envelope elimination and restoration (EER) technique.

The principle of an EER transmitter is to split the input signal into two paths. A baseband path contains the envelope of the input signal and a radio frequency (RF) path contains a constant-envelope phase modulated signal as shown in Figure 6.27.

The PA output waveform that swings from 0 to K VDD will take the shape of the envelope. Polar modulation is an entire transmitter architecture that differs from EER in the generation of envelope and phase signals, while no difference exists in reconstruction.

Rather than up converting I and Q components, adding them and applying the result to the envelope detector and limiter, polar modulation transmitters have ready envelope and phase signals as their inputs from the digital domain as shown in Figure 6.28 [6].

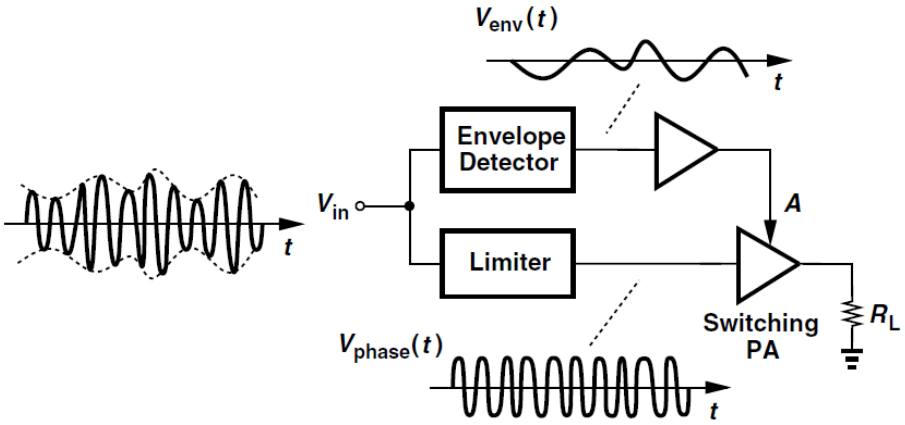


Figure 6.27: EER Block diagram

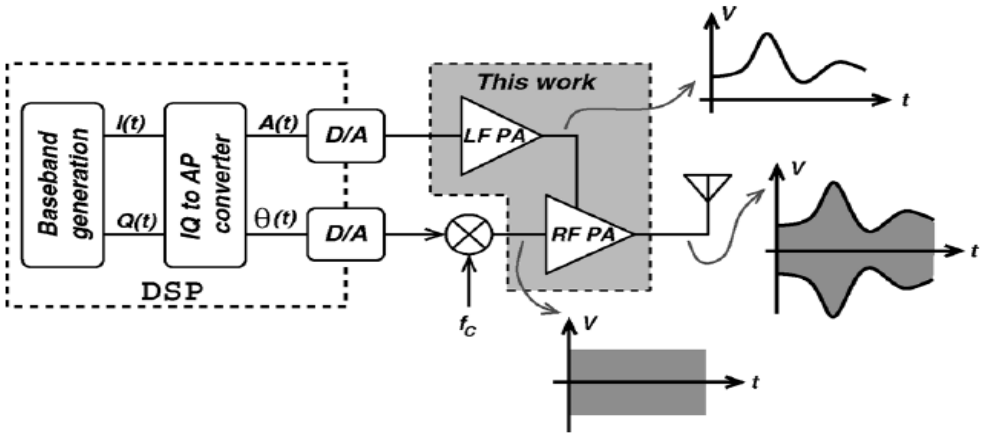


Figure 6.28: Polar transmitter block diagram

Main features

- Requires a very low loss, power efficient and fast supply modulator, such as class S supply modulators or sigma delta DC-DC converters.
- Supply variations result in AM-to-PM distortion.
- Delay mismatch between the envelope and RF phase paths can result in distortion.
- Envelope feedback can be used to improve the linearity.
- A variation of this technique that uses amplifiers with additional back-off to guarantee linearity is called envelope tracking

Research into wireless power amplifiers (PAs) has long been focused on improving the tradeoffs between efficiency and linearity. Switching power amplifiers (Classes D/E etc.) have higher efficiency compared with their linear counterparts (Classes A/AB/B etc.), but cannot be directly applied in systems which employ spectrally efficient modulation. It is for this reason a great deal of research has focused on breaking this tradeoff.

The main benefit of this idea is that the phase signal is a constant envelope signal, which makes it possible to use efficient nonlinear PA's to amplify it. This is increasing overall efficiency compared to linear power amplifier classes.

Polar modulation is one such system-level solution. However, analog polar modulation systems suffer from efficiency–bandwidth tradeoff in the supply modulators and hence have not become very popular for wide-bandwidth systems like wireless LAN (WLAN).

Table 6.2: Key concerns and tradeoffs in Polar modulation systems

Issue	Proposed Solution
Delay Mismatch Delay between envelope and phase paths cause distortion such as jitter and ACPR. (spectral regrowth)	• Adding a delay compensation unit in the design • A delay compensation unit is inserted in the phase path (which is usually faster) or the phase signal is delayed in the DSP domain to compensate for delay mismatch.
Envelope Detector Linearity Linear envelope detection of modulated RF signals over a wide enough dynamic range is extremely challenging.	• Envelope will be calculated in the DSP However, their implementation requires a change to the base band chips, which poses a different set of challenges.
Overall efficiency Designers have to be particularly careful not to lose the higher efficiencies they gain from operating the PA's to the power they consume in the limiter, switching power supply, PLL and filter.	The efficiency of PA = 53% The PA consumes 46 mW to support 25 mW (14 dBm) and the total power consumption equal 87 mW.
Efficiency–bandwidth tradeoff Analog polar modulation systems suffer from efficiency–bandwidth tradeoff in the supply modulators.	Our system is NB IoT with a very low BW.

6.5. Phase Path Results

As shown in Figure 6.29, the PA chain consists of

- Current Mode Logic (CML)
- Driver (Chain of Inverters)
- PA core (Class E)

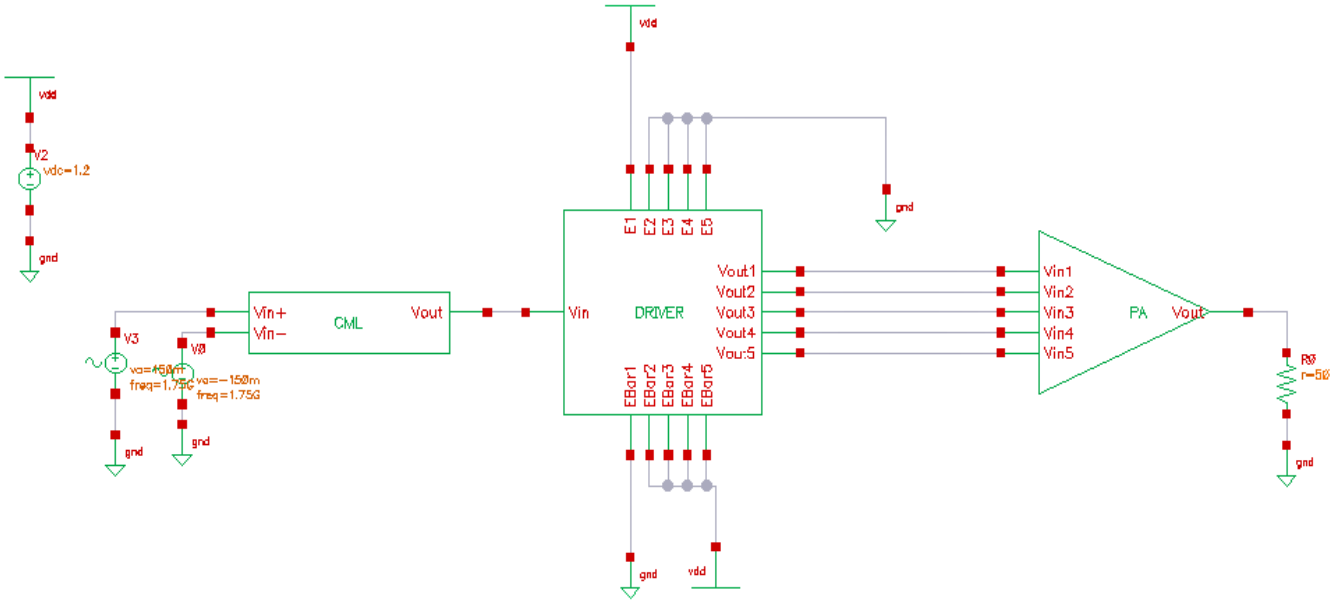


Figure 6.29: PA Blocks

6.5.1. Current Mode Logic (CML)

As shown in Figure 6.30, the schematic of the CML with high pass filter with cutoff 159 MHz to block the dc offset from PLL 900 mV and bias the dc this point with dc offset 600 mV. The total current 320 µA and the total power consumption at CML is 384 µW.

The main features of CML

- Convert the signal from differential into a signal ended .
  The Class-E amplifier uses a series inductor, which is not readily compatible with baluns or transformers that are in any case required for differential PAs to interface to a single-ended antenna.
- Guarantee that the duty cycle is 50% . All eqns of class E depend on this Assumption.
- The swing of CML is rail to rail signal .
  Due to using Class E, a CML with is chosen as pre-amplifier because it converts small input signal into a rail-rail signal with large swing.

The first graph in Figure 6.31 shows the single ended input from PLL with amplitude 150 mV and Dc offset 900mv and the second graph shows the relation between the Differential input and the single ended output after inverter.

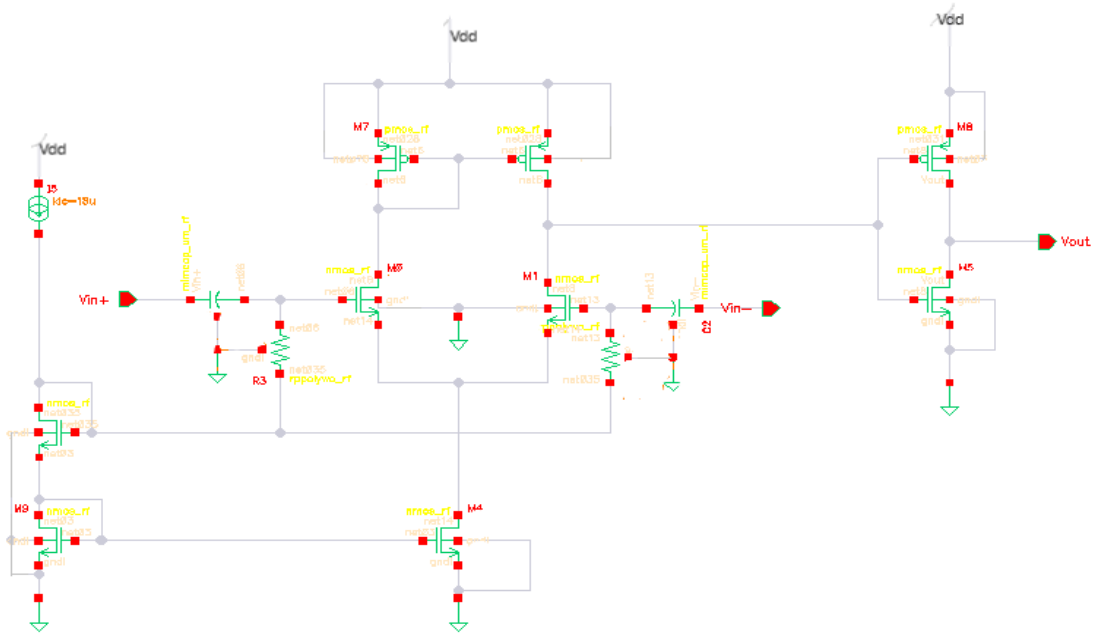


Figure 6.30: Schematic of CML

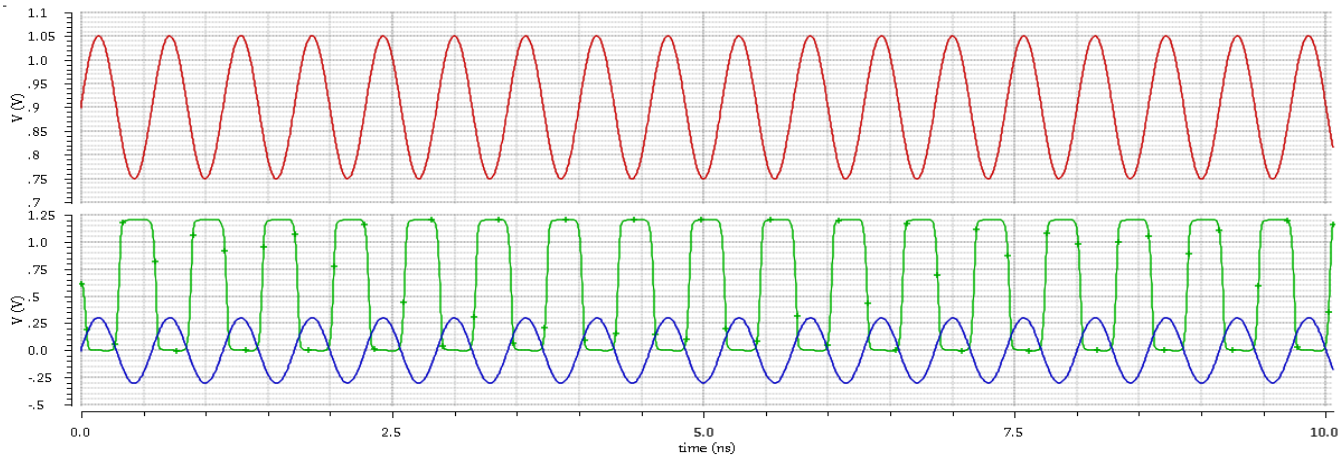


Figure 6.31: Differential Input and Single ended Output after CMOS Inverter

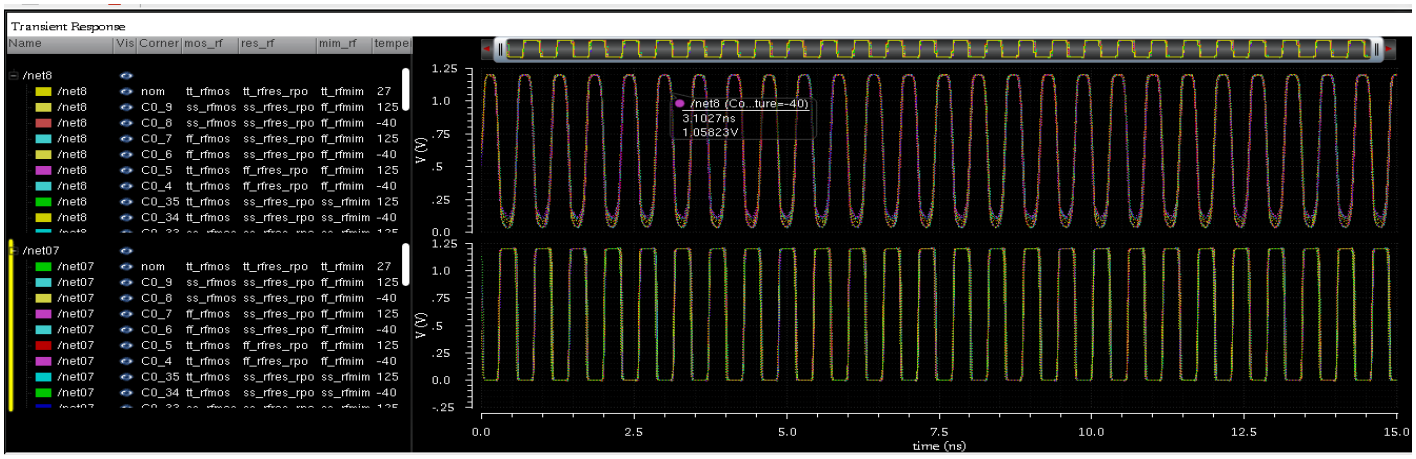


Figure 6.32: CML Output across Corners

6.5.2. Inverters

The main benfit of chain of inverters is driving high capacitance input from Class E. As shown in Figure 6.33 and 6.34, Inverter's delay is a function of the ratio between its external load capacitance and its input capacitance. The ratio is called the effective fan-out F:

$$F = f^N = \frac{C_L}{C_{in}} \qquad (6.10)$$

To minimize the delay, Size of each stage is the geometric mean of two neighbors

$$f = \frac{C_{gin,j+1}}{C_{gin,j}} = \frac{C_{gin,j}}{C_{gin,j-1}} \qquad (6.11)$$

When each stage has the same effective fanout, each stage has the same delay.
When N is large, the first component dominates (intrinsic delay). If N is too small, the effective fanout of each stage becomes large, and second component is dominant.

For a given load, C_L and given input capacitance C_{in}, the optimum number of stages:

$$N = \ln(F)$$

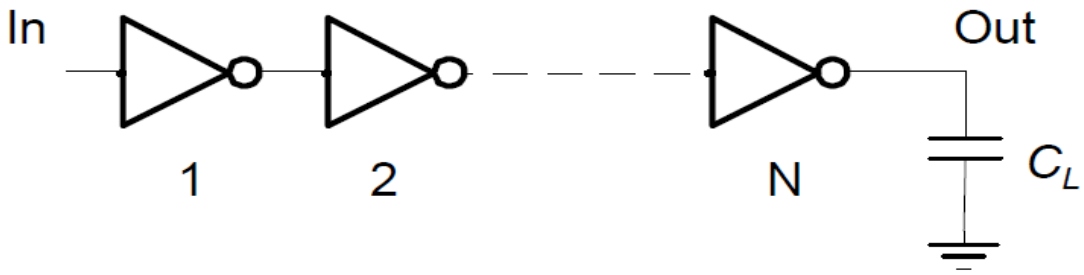


Figure 6.33: Chain of Inverters

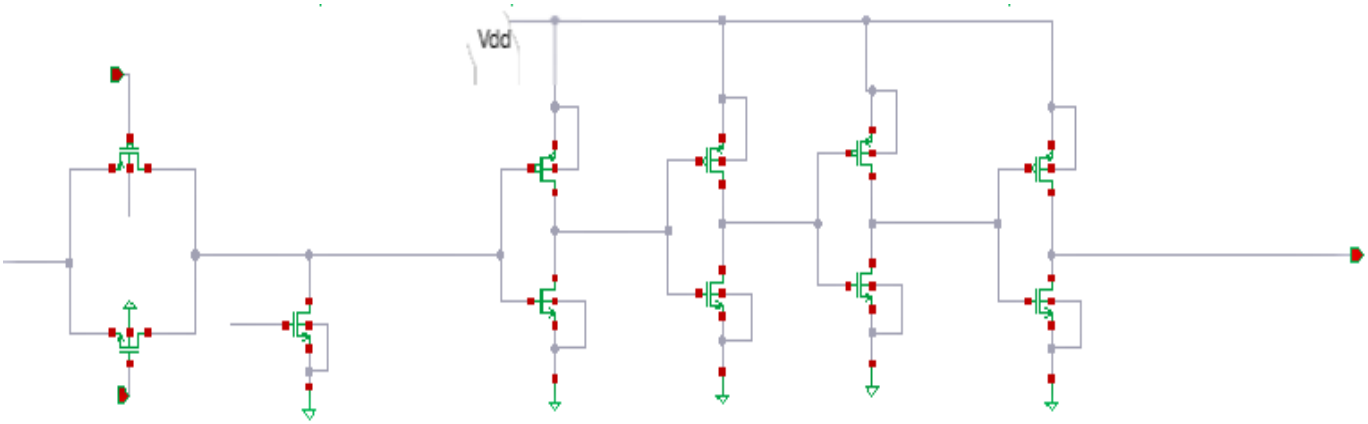


Figure 6.34: Chain of Inverters

Transmission Gate

As shown in Figure 6.35, the transmission gate works as a buffer in ON state but it works as high impedance with pull down network when this path is not used.

The size of transmission gate has a trade-off between loading capacitance on driver and the effective transmission of the RF signal through switch.

The pull down network is connected with enable bar, if this is path is selected then the enable bar equal zero which mean that this is path is short circuit. If this is path is not selected then enable bar equal VDD which mean that the transmission gate is a high impedance and the output node is forced to be ground.

As shown in Figure 6.36, there is only one path is selected and other paths are forced to be ground (almost Zero).

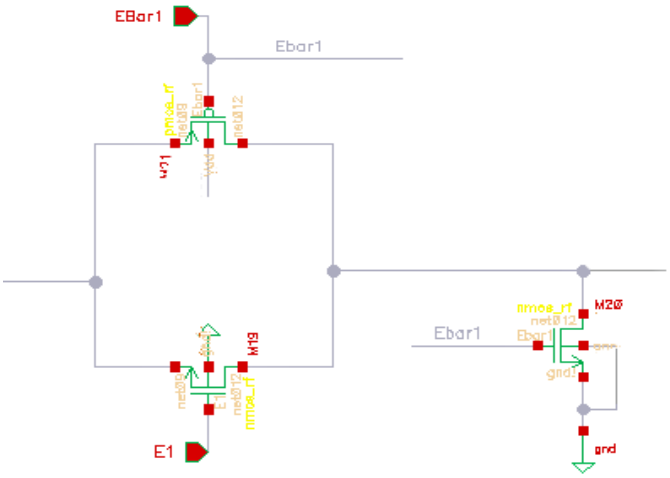


Figure 6.35: Transmission Gate

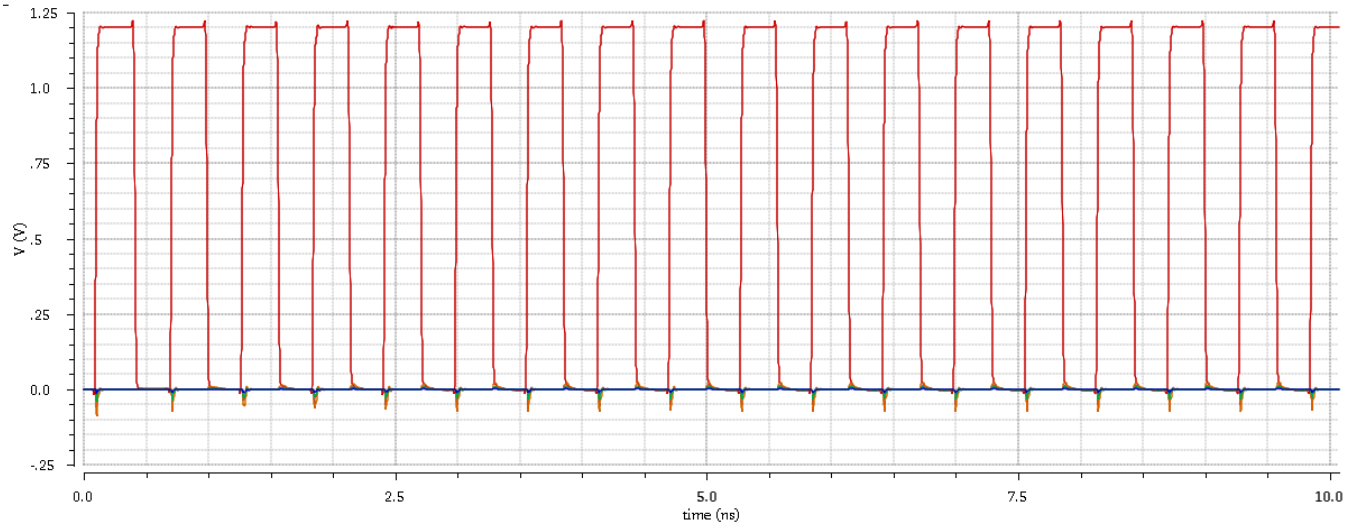


Figure 6.36: Only one path is selected

Due to different input capacitance from PA paths, the optimum number of stages is different which mean the delay of each path is varying according to the number of stages of inverters. The design according to this is approach is not a practical.

The second approach is to get the optimum number of stages for the first path (maximum load capacitance) then get the optimum size of this number of stages and for other paths use the same number of stages with scaling ratio in the transistor size equal to the scaling ratio of capacitance load.

As shown in Figure 6.37 and 6.38, the same delay for all paths.

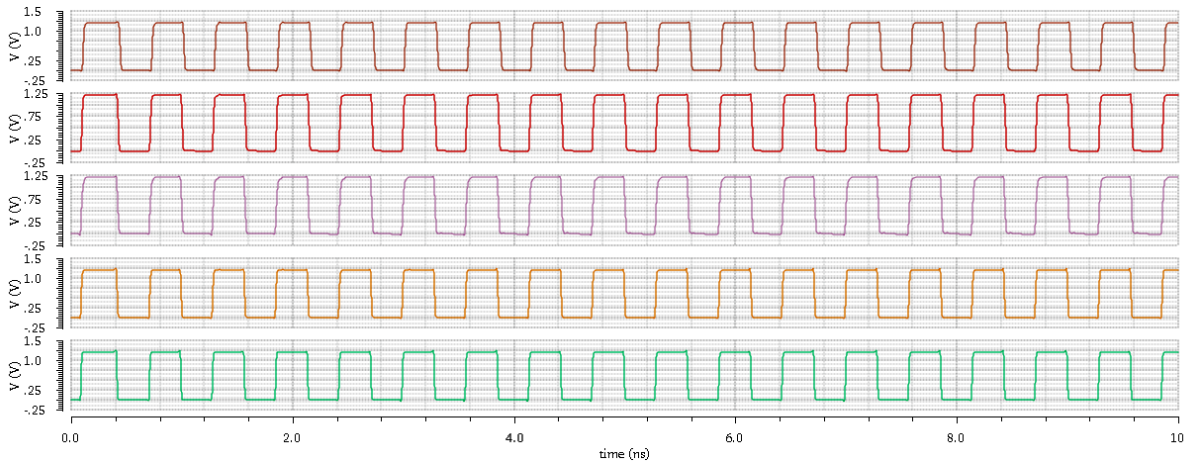


Figure 6.37: All paths are selected

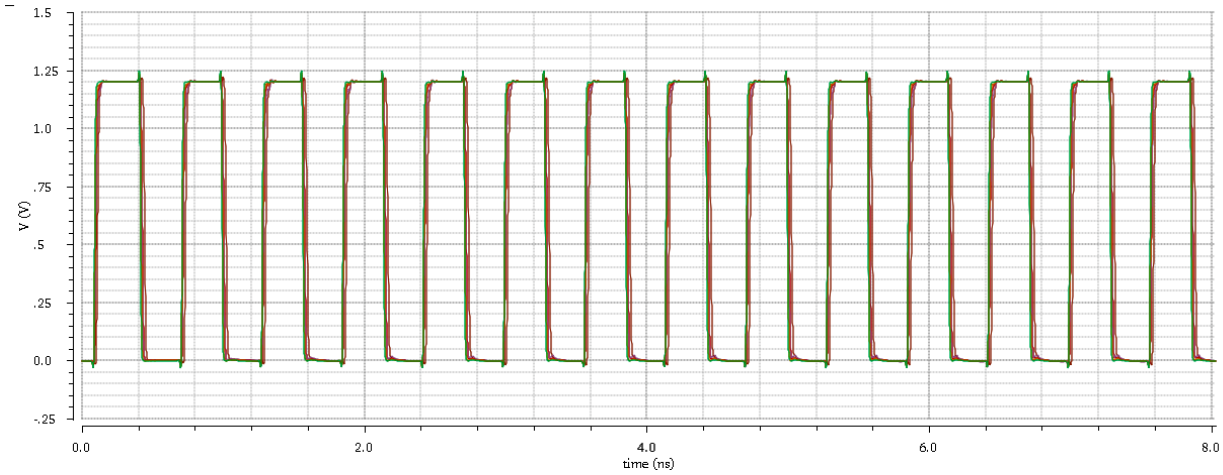


Figure 6.38: All paths are selected

6.5.3. PA (Class E)

Class E power amplifier is one of the nonlinear classes that can theoretically achieve 100% efficiency.

There are two states of Class E:

- **ON state**

The device is configured similarly to the linear PA device, except that the gate signal drives the device between triode and cutoff regions. As shown in Figure 6.39, when the device is ‘on,’ the drain−source voltage is constrained to near zero.

- **OFF state**

When the device is ‘off,’ the drain− source voltage is constrained by the dynamics of the output network and the current flow through the shunt capacitor as shown in Figure 6.40.

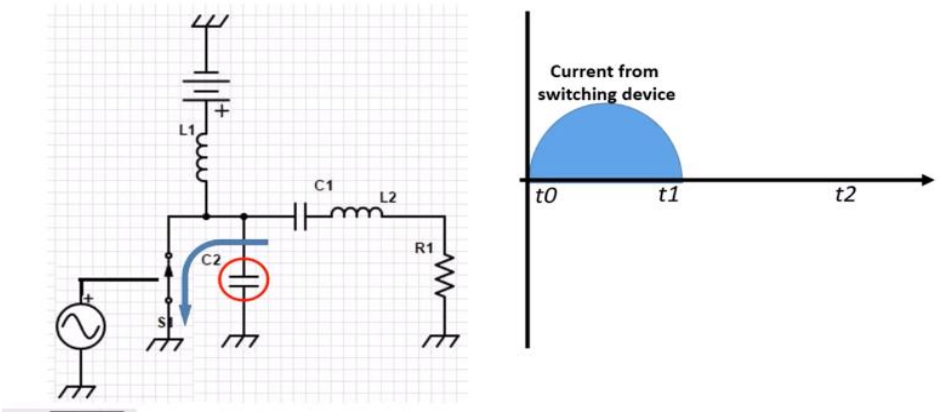


Figure 6.39: ON state of Class E

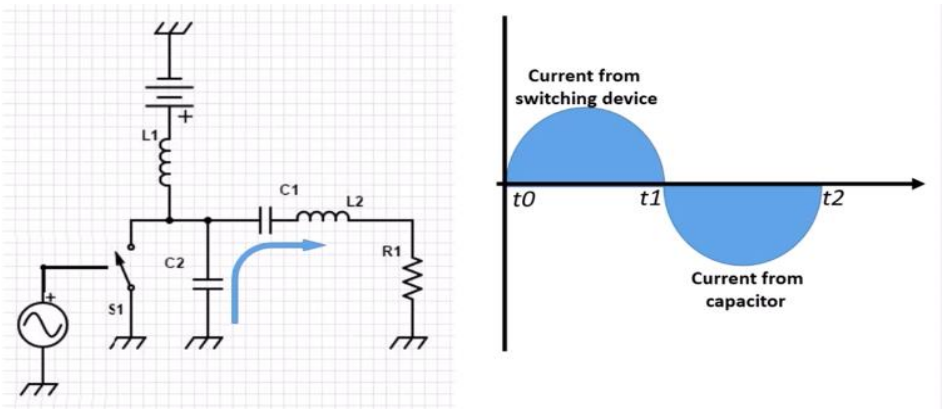


Figure 6.40: OFF state of Class E

The current and Voltage equations:

$$I_{sw} = I_{DC} + sin(\theta - \emptyset) \quad 0 < \theta < \pi \tag{6.12}$$

$$I_C = I_{DC} + sin(\theta - \emptyset) \quad \pi < \theta < 2\pi \tag{6.13}$$

$$V_{sw} = \frac{1}{C}\int_{\pi}^{2\pi} I_C \; d\theta \tag{6.14}$$

V_{sw} is the integral of ac current during the OFF state as shown in Figure 6.41 and 6.42.

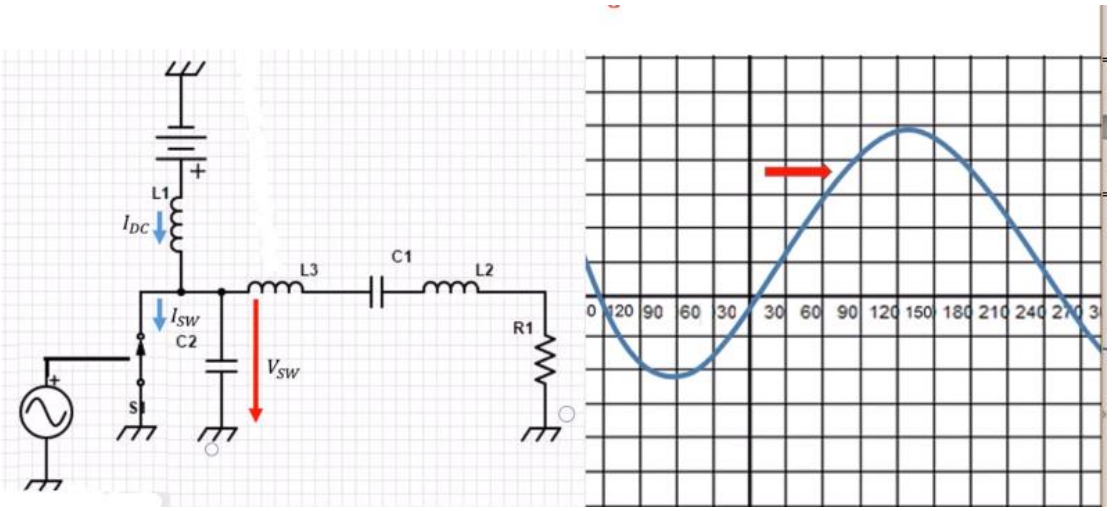


Figure 6.41: Class E waveform

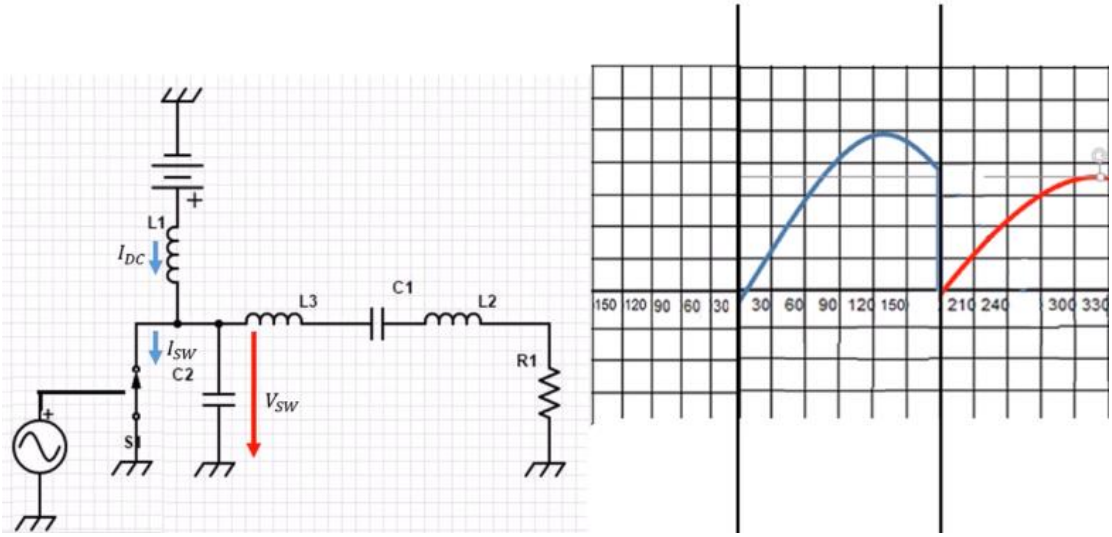


Figure 6.42: Class E waveform

This is equations introduced by Sokal 1975 [7]:

$$C_1 = \frac{1}{w_c R_l (Q_L - 1.1525)} \tag{6.15}$$

$$C_2 = 0.1836 \frac{1}{w_c R_l} \tag{6.16}$$

$$L_3 = 1.1525 \frac{R_l}{w_c} \tag{6.17}$$

$$P_L = 0.5768 \frac{VDD^2}{R_l} \tag{6.18}$$

Using ADS, this design equation is related together and there is threshold for each design parameter, the final value as shown in Figure 6.43.

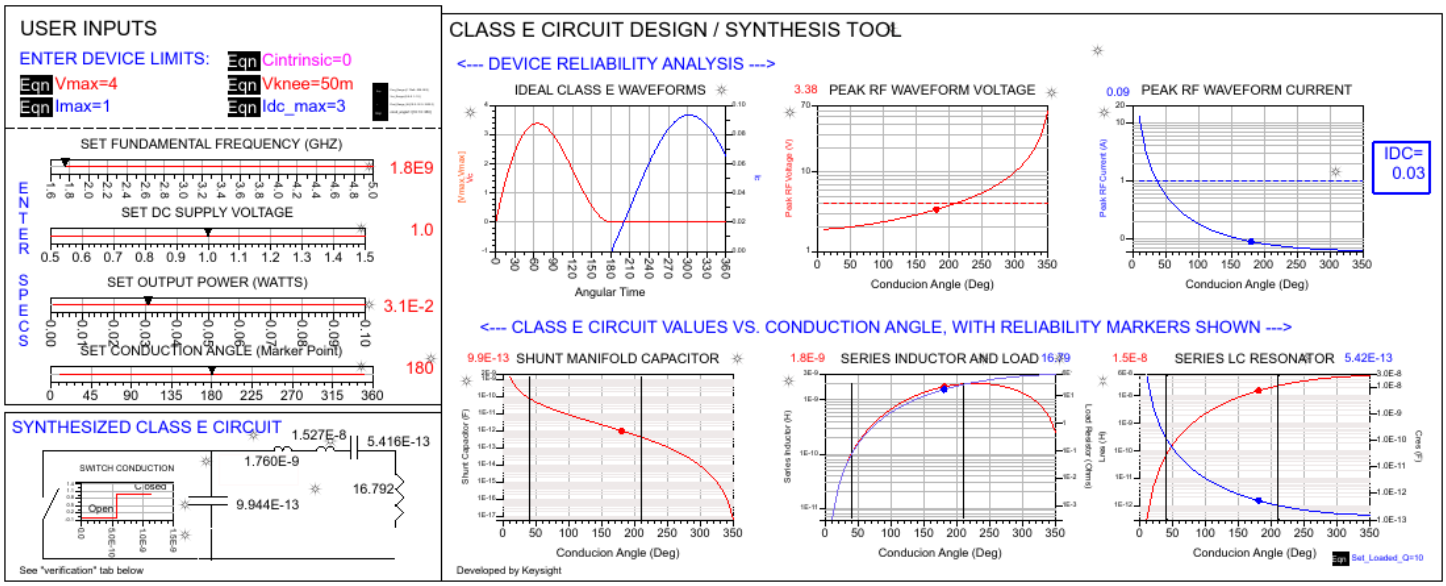


Figure 6.43: Class E circuits design

The value of choke inductor is $L_1 = \frac{R_{DC}\, Q}{w_c}$

where $R_{DC} = 1.7337\, R_l$ and the quality factor =12 and $L_2$is related to C_1as LC tank with resonance frequency w_c.

Table 6.3: The value of the components

Component	Value
C_1	541 f F
C_2	994 f F
L_2	15 n H (off Chip)
L_3	1.7 n H
L_1	33 n H (off Chip)
R_l	17 ohm

As shown in Figure 6.44 and 6.45, the value of off chip Inductor and its quality factor and self-resonance frequency**.**

muRata

⚠Note • Please read rating and ⚠CAUTION (for storage, operating, rating, soldering, mounting and handling) in this catalog to prevent smoking and/or burning, etc.
• This catalog has only typical specifications. Therefore, please approve our product specifications or transact the approval sheet for product specifications before ordering.

O05E.pdf
Apr.28,2017

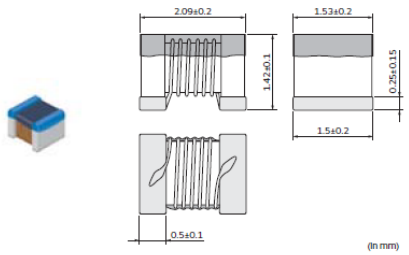
RF Inductors

LQW2BAS_00 Series 0805 (2015) inch (mm)

Inductors for Power Lines

Inductors for General Circuits

Appearance/Dimensions



Packaging

Code	Packaging	Minimum Quantity
K	ø330mm Embossed Taping	8000
L	ø180mm Embossed Taping	2000
B	Packing in Bulk	500

Figure 6.44: RF Inductor

Rated Value (☐: packaging code)

Part Number	Inductance	Inductance Test Frequency	Q (min.)	Q Test Frequency	Rated Current	Max. of DC Resistance	S.R.F.* (min.)
LQW2BAS2N8J00☐	2.8nH ±5%	250MHz	80	1500MHz	800mA	0.06Ω	12200MHz
LQW2BAS3N0J00☐	3.0nH ±5%	250MHz	65	1500MHz	800mA	0.06Ω	12200MHz
LQW2BAS5N6J00☐	5.6nH ±5%	250MHz	65	1000MHz	600mA	0.08Ω	5900MHz
LQW2BAS6N8J00☐	6.8nH ±5%	250MHz	50	1000MHz	600mA	0.11Ω	5600MHz
LQW2BAS7N5J00☐	7.5nH ±5%	250MHz	50	1000MHz	600mA	0.14Ω	4800MHz
LQW2BAS8N2G00☐	8.2nH ±2%	250MHz	50	1000MHz	600mA	0.12Ω	4400MHz
LQW2BAS8N2J00☐	8.2nH ±5%	250MHz	50	1000MHz	600mA	0.12Ω	4400MHz
LQW2BAS10NG00☐	10nH ±2%	250MHz	60	500MHz	600mA	0.10Ω	4300MHz
LQW2BAS10NJ00☐	10nH ±5%	250MHz	60	500MHz	600mA	0.10Ω	4300MHz
LQW2BAS12NG00☐	12nH ±2%	250MHz	50	500MHz	600mA	0.15Ω	4000MHz
LQW2BAS12NJ00☐	12nH ±5%	250MHz	50	500MHz	600mA	0.15Ω	4000MHz
LQW2BAS15NG00☐	15nH ±2%	250MHz	50	500MHz	600mA	0.17Ω	3200MHz
LQW2BAS15NJ00☐	15nH ±5%	250MHz	50	500MHz	600mA	0.17Ω	3200MHz
LQW2BAS18NG00☐	18nH ±2%	250MHz	50	500MHz	600mA	0.20Ω	3100MHz
LQW2BAS18NJ00☐	18nH ±5%	250MHz	50	500MHz	600mA	0.20Ω	3100MHz
LQW2BAS22NG00☐	22nH ±2%	250MHz	55	500MHz	500mA	0.22Ω	2600MHz
LQW2BAS22NJ00☐	22nH ±5%	250MHz	55	500MHz	500mA	0.22Ω	2600MHz
LQW2BAS24NG00☐	24nH ±2%	250MHz	50	500MHz	500mA	0.22Ω	2400MHz
LQW2BAS24NJ00☐	24nH ±5%	250MHz	50	500MHz	500mA	0.22Ω	2400MHz
LQW2BAS27NG00☐	27nH ±2%	250MHz	55	500MHz	500mA	0.25Ω	2580MHz
LQW2BAS27NJ00☐	27nH ±5%	250MHz	55	500MHz	500mA	0.25Ω	2580MHz
LQW2BAS33NG00☐	33nH ±2%	250MHz	60	500MHz	500mA	0.27Ω	2150MHz
LQW2BAS33NJ00☐	33nH ±5%	250MHz	60	500MHz	500mA	0.27Ω	2150MHz

Figure 6.45: RF Inductor

As shown in Figure 6.46, there are five path to support the full range of Power from -40 dBm to 14 dBm.

The first stage support from 14 dBm to $-$ 5 dBm, the second stage support from-5 dBm to $-$ 15 dBm, the third stage support from-15 dBm to $-$ 23 dBm, the fourth stage support from-23 dBm to $-$ 30 dBm and the fifth stage support from-30 dBm to $-$ 40 dBm.

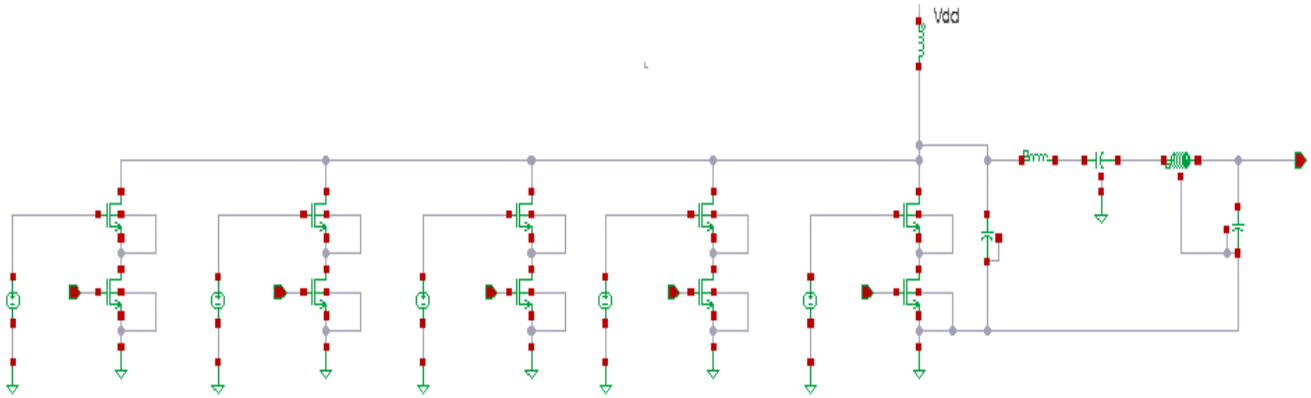


Figure 6.46: PA Schematic

As shown in Figure 6.47, there is a very small overlapped between Drain Voltage and Current so the efficiency with ideal choke coil is 86% and after added a real coil with a finite series resistance, the efficiency dropped to 53%.

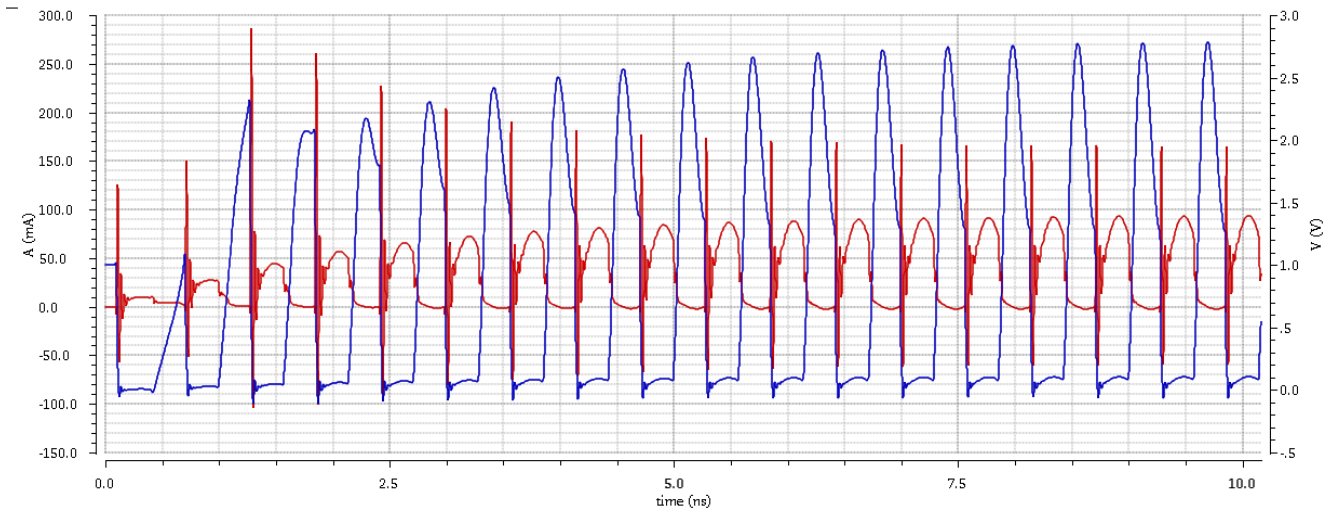


Figure 6.47: Drain Voltage and Current

One of Class E issues is the high value of the drain voltage in OFF state equal 3.56 VDD, which raises the concerns about breakdown and reliability issues. The solution of this issue is using

cascade device. A shown in Figure 6.48 the difference between any two terminal does not exceed the supply voltage.

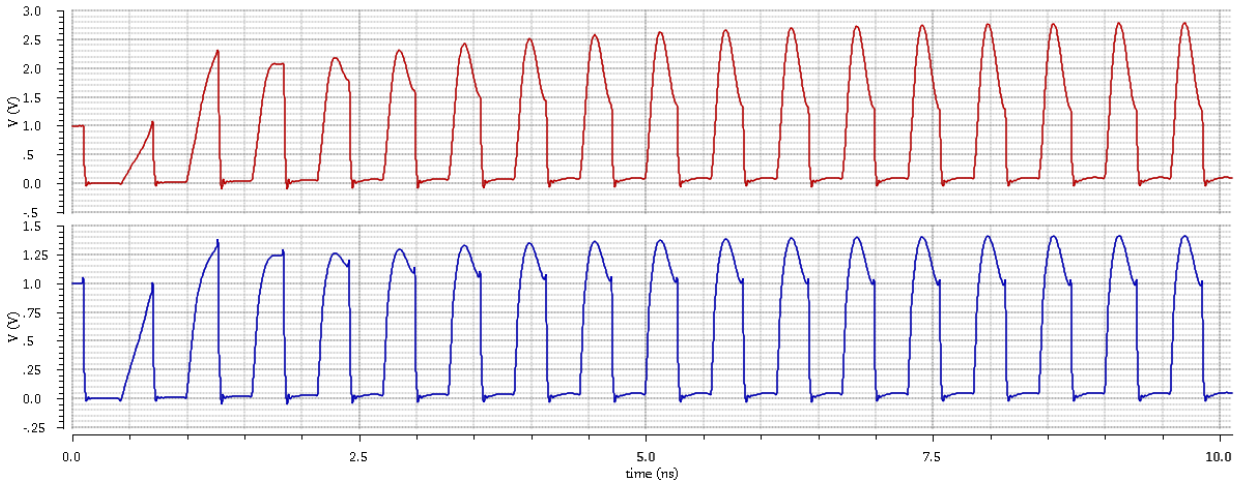


Figure 6.48: Cascode Voltage and Internal node

As shown in Figure 6.49, the relation between the input signal of class E and the output on 50 ohm (14 dBm).

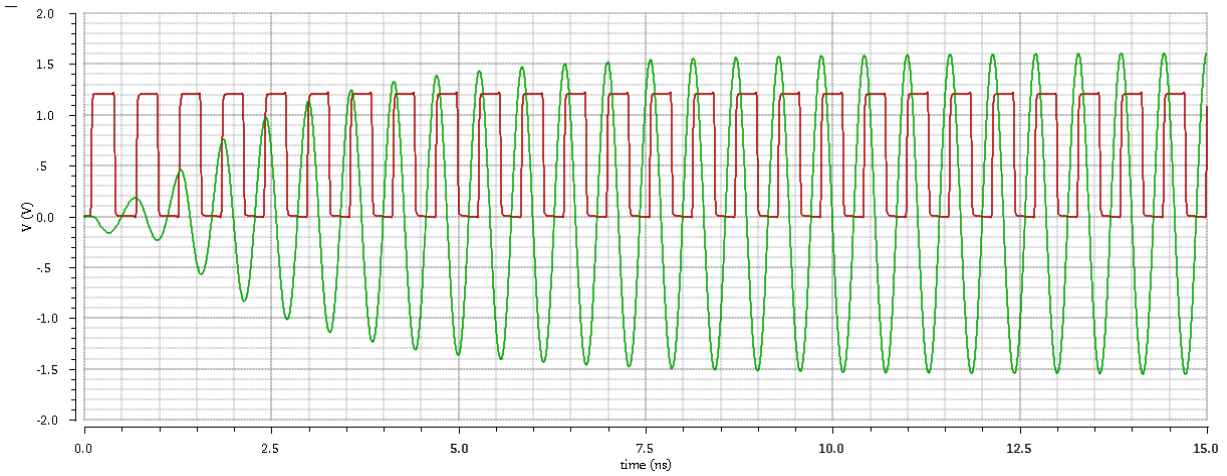


Figure 6.49: I/P and O/P

Using parametric analysis to get the relation between the VDD_{PA} and the output power in mW and the derivative of this relation to get the linear range, as shown in Figure 6.50.

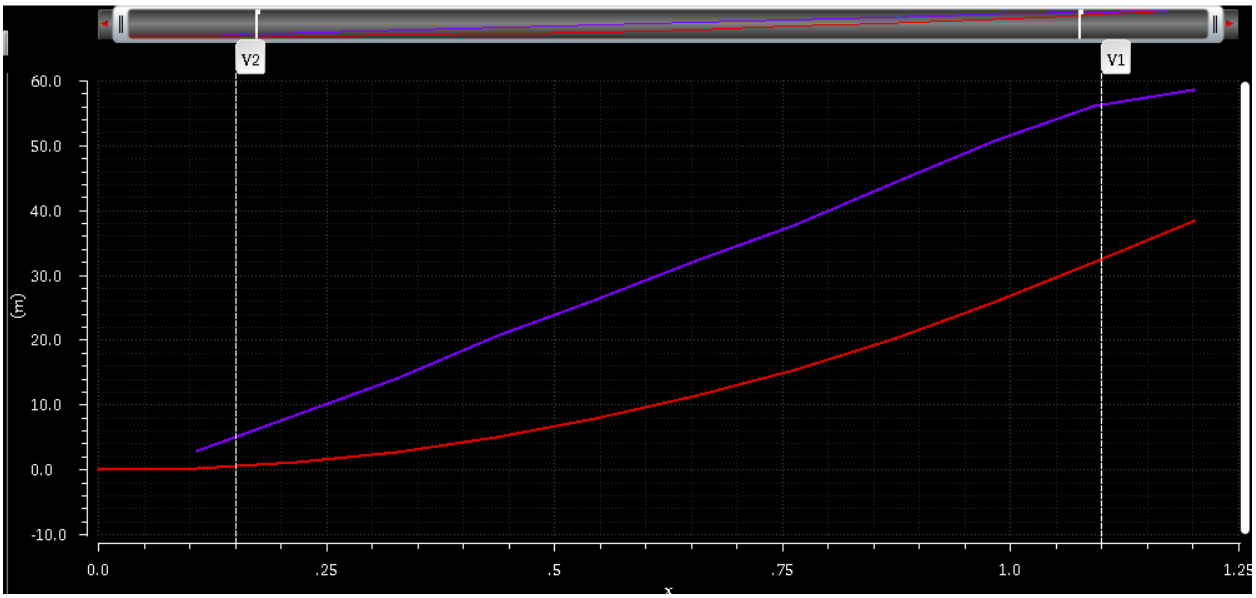


Figure 6.50: Output Power VS VDD_PA and its derivative

References:

[1] Behzad Razavi, "RF microelectronics" second edition.

[2] Marian Kazimierczuk , "RF Power Amplifier " second edition, Wiley.

[3] Frederick H. Raab, "Maximum efficiency and output of class- F power amplifiers" IEEE transactions on circuits and system, Vol. 49, No.6, pp.1162-1166, 2001.

[4] Debopriyo Chowdhury,LuYe, Elad Alonand Ali M. Niknejad,"An Efficient Mixed-Signal 2.4-GHz Polar Power Amplifier in 65-nm CMOS Technology" IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 46, NO. 8, AUGUST 2011

[5] Mehdi F. Soltan, ," High Efficiency Linear Power Amplifiers: Analysis, Linearization and Implementation" , June 2004

[6] Seth R. Sanders, Jason Thaine Stauth ," Energy Efficient Wireless Transmitters: Polar and Direct-Digital Modulation Architectures" , February 4, 2009

[7] Nathan O. Sokal and Alan D. Sokal "Class E – A new class of high efficiency tuned signal-ended switching power amplifiers " IEEE J. Solid-State Circuits , vol. SSC-10, pp.168-176, June 1975.

Chapter 7 : System Simulations

In this chapter, the whole transmitter is integrated and simulated. The PLL control voltage showing lock is shown in Fig. 7.1. First, the amplitude is varied to show the effect of the amplitude variation on the PA output. This result is shown in Fig.7.2. The upper part of the figure shows the PA output while the lower part shows the envelope it should track which it tracks successfully. A QPSK symbol at a single subcarrier was used as an input to the transmitter. The four possible outputs are shown in Fig. 7.3 on the same graph. A zoomed version of the output is depicted in Fig. 7.4 with the delay between the zero crossings shown on the graph. The constellation is plotted in Fig. 7.5. There is a small deviation in the constellation points about 2 degrees. This comes from the simulator accuracy as the output period is in the order of 100 ps while the simulator accuracy in the delay calculations is in the order of 10 nS. This is evident in Fig. 7.3. The output phase noise is plotted with the mask on the same graph in Fig. 7.6. The transmitter output spectrum is plotted in Fig. 7.7. The delay mismatch between the phase and amplitude paths was studied. The phase path reconstruction filter has a higher delay of 4.2 us than the envelope path reconstruction filter. This is due to the lower bandwidth of the phase path filter. The phase path has an extra delay in the IQ mixers and limiter till the filter output reaches the PLL reference input. This delay was calculated to be equal to 17ns. The PLL settling time that adds further delay is 67.1486 us. The PA drivers adds a delay of 0.1103 ns. This means that a delay of 71.36571 us will be inserted before the envelope path in the digital domain as it is the faster path. A comparison with the state of the art transmitters is shown in table 7.1. The transmitters in [1], [2] and [3] were tested and fabricated after testing. The proposed architecture consumes less power. However, the second PLL providing the LO to the offset PLL feedback mixer still needs to be designed. Moreover, the DC/DC converter that will support the PA supply will be designed. After designing these blocks, the transmitter efficiency is expected to degrade. The transmitter power will increase but it is expected to be kept below [1], [2] and [3]. However, this comes at the expense of the larger area as the proposed design contains two PLLs and two VCO inductors while the reported transmitters rely mainly on digital circuits which reduce the area considerably. The transmitter in [2] supports a much higher output power (23 dbm) which translates to a higher power consumption. The achieved evm was 1% but it is expected to degrade from the envelope path, the IQ mismatches and the integration with the digital part due to the finite quantization error.

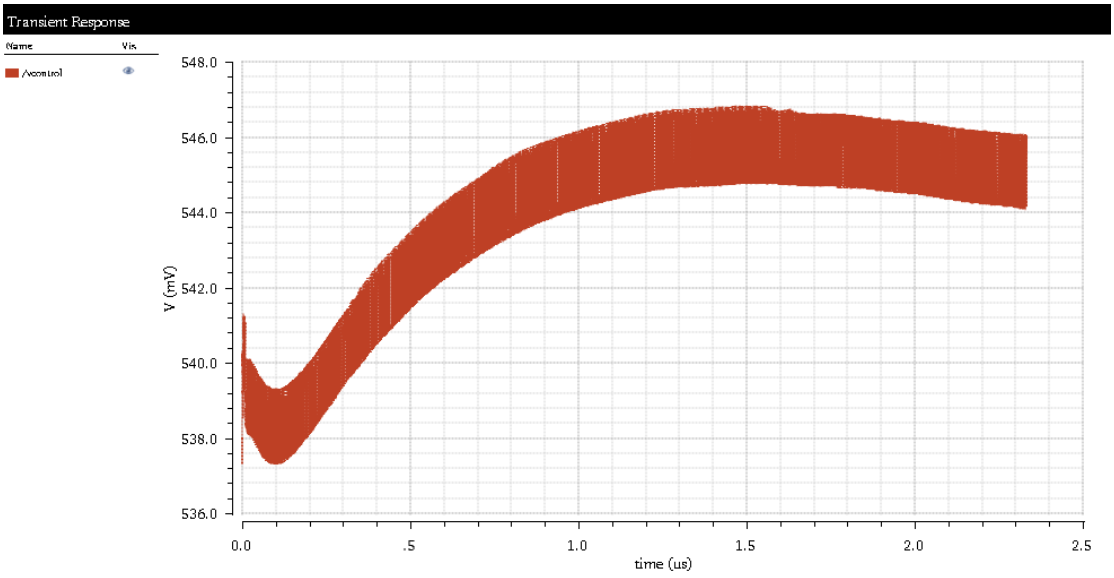


Figure 7.1: PLL Control Voltage in Lock

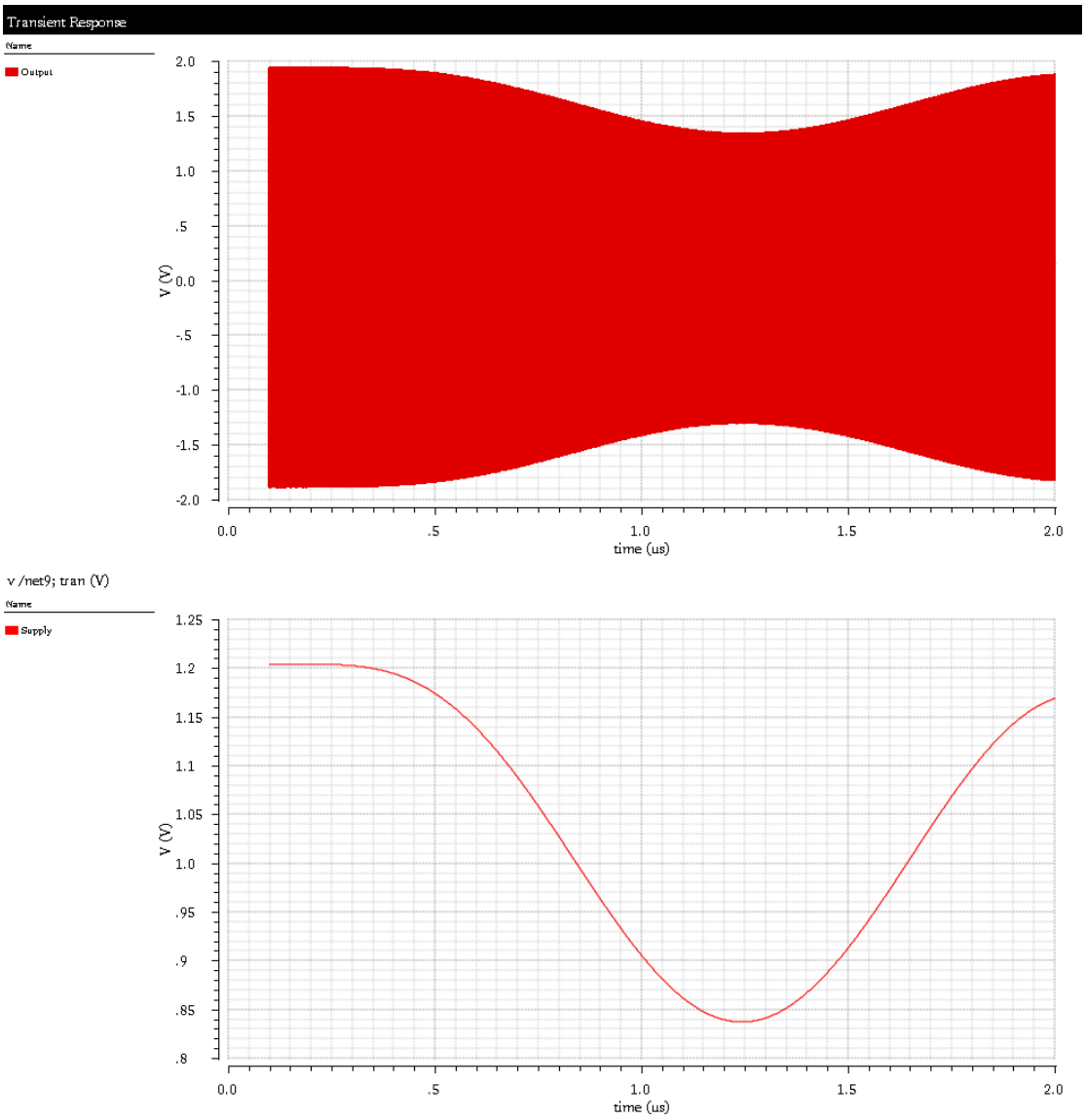


Figure 7.2: Envelope Variation at PA Output

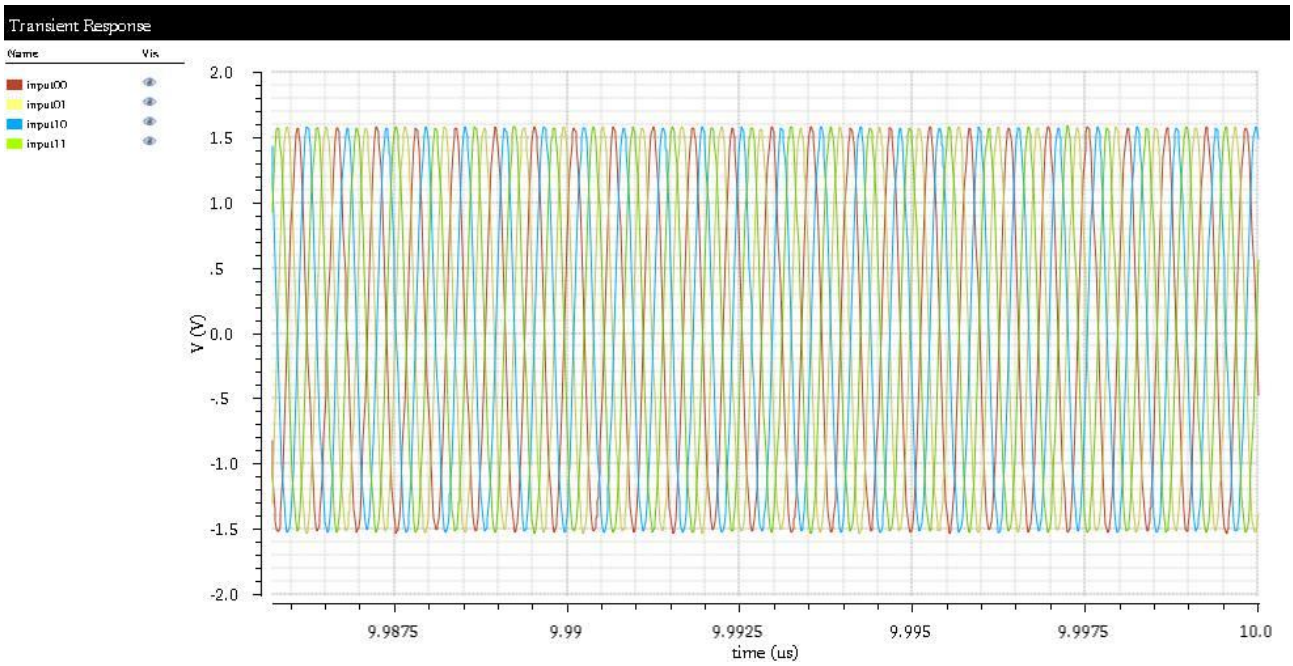


Figure 7.3: Four QPSK Outputs

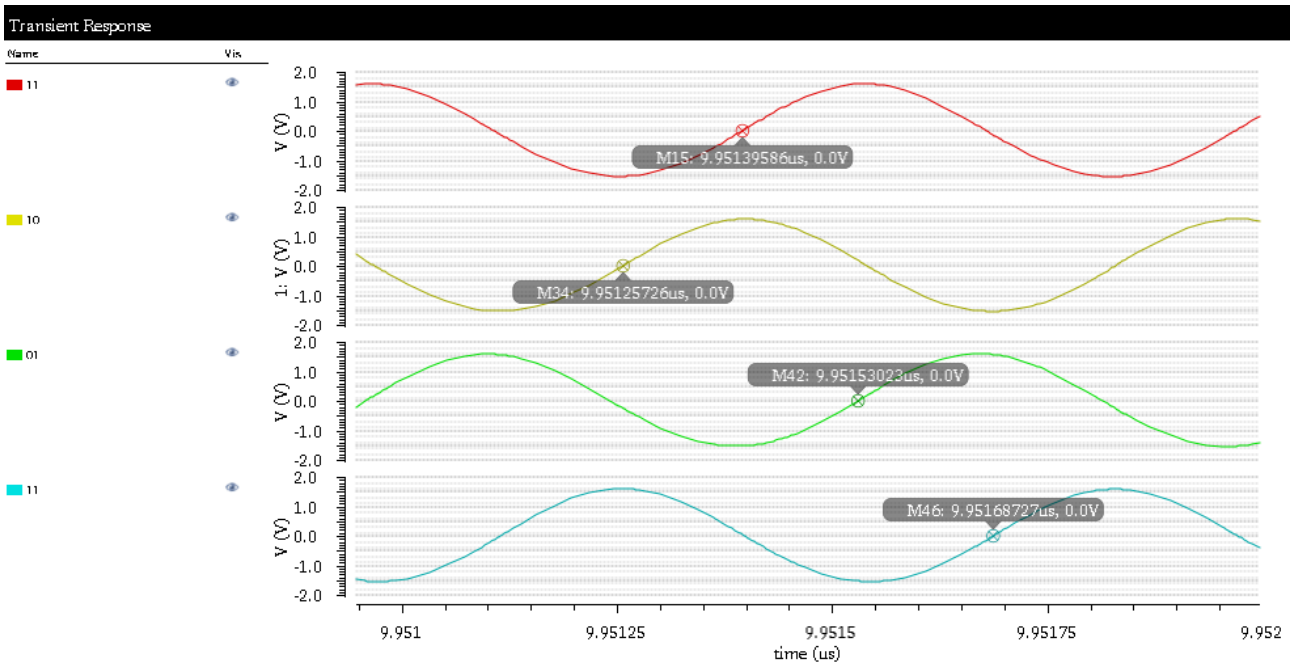


Figure 7.4: Zoomed QPSK Outputs

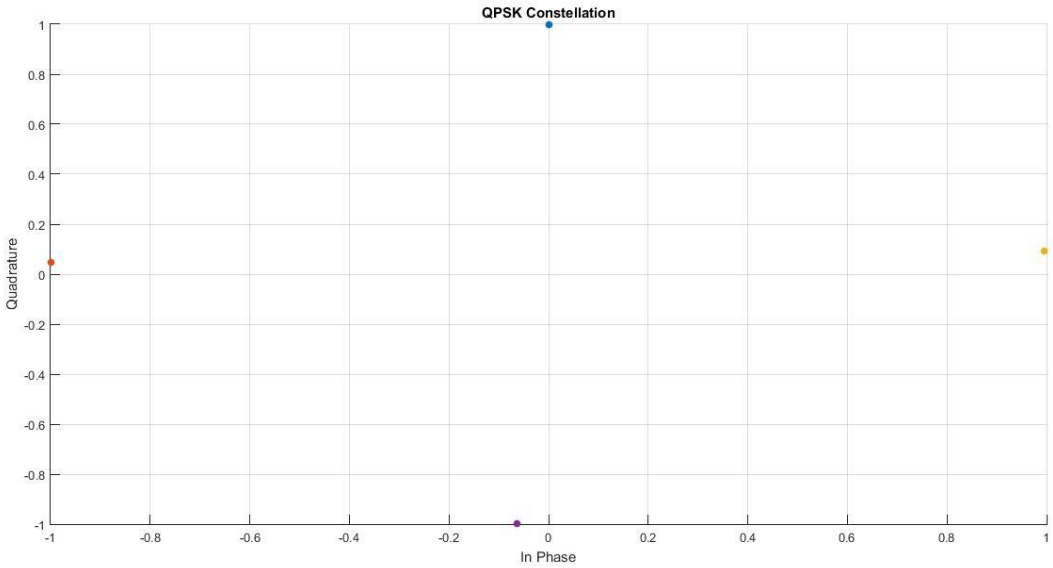


Figure 7.5: QPSK Constellation

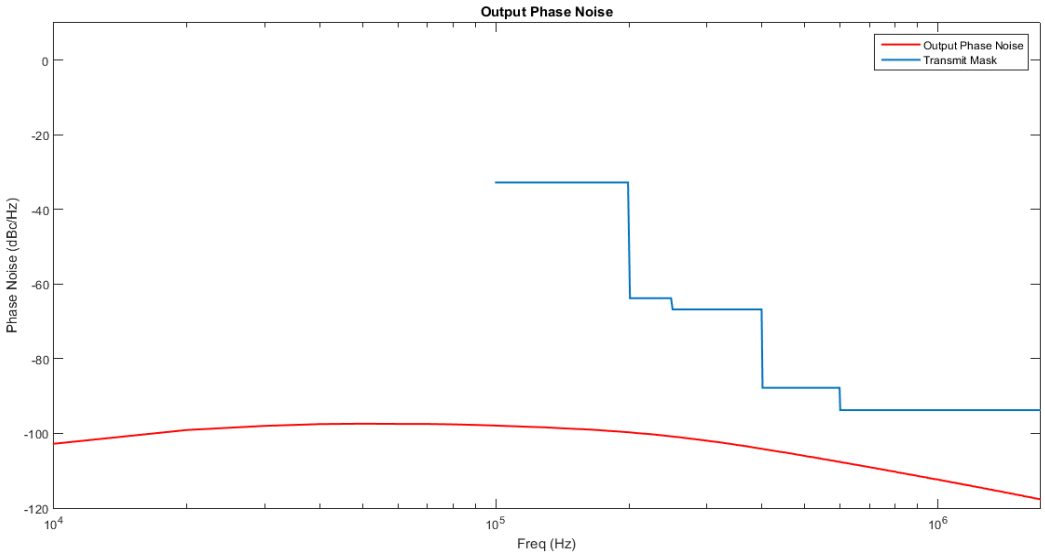


Figure 7.6: Transmitter Output Phase Noise

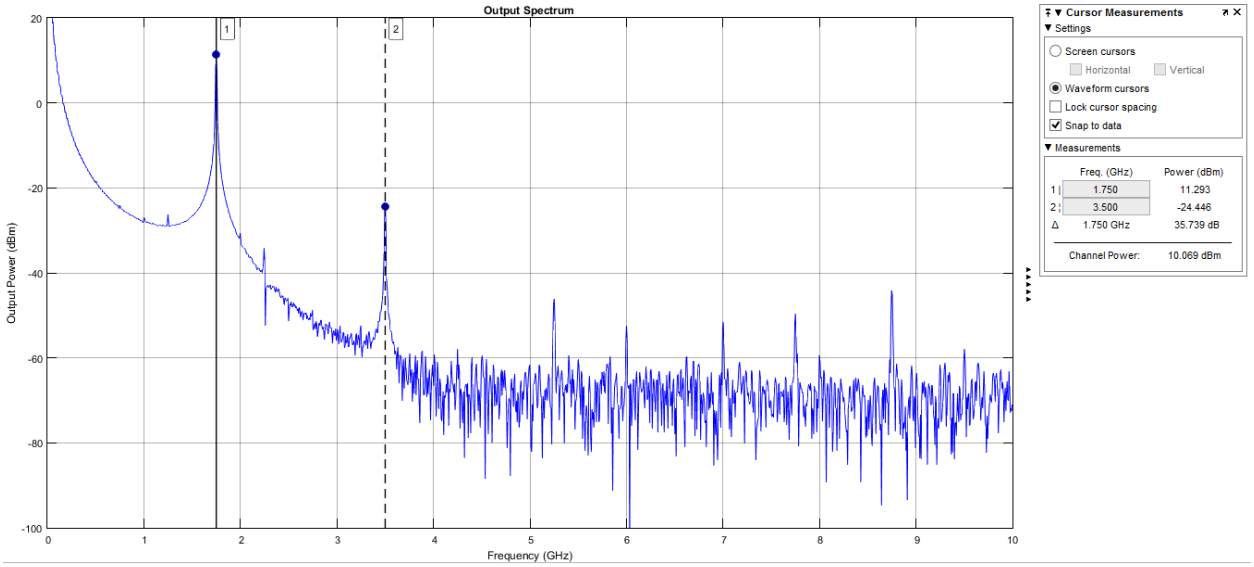


Figure 7.7: Transmitter Output Spectrum

Table 7.1: Performance Comparison With The State-Of-The-Art Transmitters

Index	[1]	[2]	[3]	This work
Technology	65 nm	180 nm	55 nm	65 nm
Frequency (GHz)	2.4	0.75-0.96	2.4	1.75
Pout	14	23.2	11/8	14
Power (mW)	150	475.8	127.3	87.6
PA efficiency	44	50.1 %	-	53 %
EVM	3.9%	3.78%	5.7 rms	1%

References

[1] D. Chowdhury, L. Ye, E. Alon, and A. M. Niknejad, "An efficient mixed-signal 2.4-GHz polar power amplifier in 65-nm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 46, no. 8, pp. 1796–1809, 2011.

[2] Z. Song, X. Liu, X. Zhao, Q. Liu, Z. Jin, and B. Chi, "A Low-Power NB-IoT Transceiver With Digital-Polar Transmitter in 180-nm CMOS," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 64, no. 9, pp. 2569–2581, 2017.

[3] Y. S. Shih, H. L. Chu, W. K. Hong, C. C. Hung, A. Tanzil, Y. L. Huang, J. Y. Chen, L. H. Hung, L. C. Cho, J. Cao, Y. C. Huang, Y. L. Hsueh, and Y. H. Chung, "A 55nm, 0.6mm2Bluetooth SoC integrated in cellular baseband chip with enhanced coexistence," *Proc. 2013 IEEE Asian Solid-State Circuits Conf. A-SSCC 2013*, pp. 193–196, 2013.

Chapter 8 : Conclusion

In this work, a polar transmitter based on the offset PLL architecture for NB IOT was introduced and implemented in 65 nm CMOS psocess. The transmitter covers the 1.8 GHz band . The transmitter consumed only 87.6 mW from 1.2 V supply. The noise transmitted at an offset of 100 Khz is -89.9 dbc. A max output power level of 14 dbm is supported. The transmitter achieves an rms evm of 1%. The achieved PA efficiency is 53%.

8.1. Future Work

The next step for our transmitter is to design the two remaining blocks which are the second PLL and the DC/DC converter to provide the PA supply. Integration with the digital part will be done. Then layout to the transmitter blocks will be done to be able to fabricate and test our transmitter. Also, supporting higher power classes defined by the standard is considered.