MASTER'S THESIS

D-BAND DOWNCONVERSION MIXER DESIGN IN CMOS SOI

Author	Mikko Kaikkonen
Supervisor	Olli Kursu
Second Examiner	Aarno Pärssinen
(Technical Advisor	Mikko Hietanen)

September 2023

ABSTRACT

The current surge in research interest around the sub-THz frequency region comes as a no surprise. The potential for greater data rates and available bandwidths are just a couple reasons why research around these frequencies should be prioritized. Many viable receiver structures have been presented for these frequency regions, but they all have one thing in common: They all include a downconversion mixer. The mixer is a crucial piece in the receiver structure, converting the higher frequency radio frequency (RF) signal to a much lower intermediate frequency (IF) signal using multiplication with a local oscillator (LO) signal. The resulting waveform is much easier to handle for signal processing that comes after. The downconversion should be able to provide a fair amount of gain to the converted signal on a wide range of input signals, measured with the 1dB compression point. The noise figure is also a major consideration for RF-devices, but in the case of the mixer, its importance is not as prevalent as it is for the LNA that precedes it, since the noise of the mixer is attenuated by the gain of the previous stages.

This master's thesis work introduces the basic theory around downconversion mixers, followed by the design of a mixer from schematic level circuit design all the way to the physical layout. The physical design is done using 22nm FDSOI technology, provided by GlobalFoundries. The design is made for a direct conversion receiver using Gilbert cell topology, meaning image rejection is reasonable and depends only on the received signal itself, and good noise and feedthrough performance should be expected in simulations. The mixer is to downconvert a 151 GHz signal down to $0 - 1$ GHz, using an LO signal between $150 - 151$ GHz. Two iterations of the mixer are shown in the end results, the first one being based on the schematic design, and the second one with adjustments made for better performance. While driving a high impedance 500 Ohm load, the second iteration was able to reach a conversion gain of -10.0 dB with a 1dB compression point of 6.4 dBm while dissipating 4.7 mW of power. DSB noise figure was simulated to be 17.3 dB and the LO leakage to the IF output at -27.7 dBm.

Key words: Mixer, Gilbert cell, 22nm, D-band, CMOS-SOI, Downconversion

TIIVISTELMÄ

Nykyinen tutkimuksen keskittyminen millimetriaalto ja THz taajuusalueille ei tule kenellekään yllätyksenä. Suurempien datanopeuksien ja vapaiden taajuuskaistojen potentiaali ovat vain joitain monista hyvistä käytännön syistä, miksi tutkimusta näiden taajuuksien ympärillä priorisoidaan. Monia käytännöllisiä vastaanotinrakenteita on esitetty näille taajuusalueille ja niillä on kaikilla yksi yhteinen tekijä: tajuusmuunnin alemmille taajuuksille. Taajuusmuunnin eli sekoitin on olennainen osa vastaanotinrakenteita, muuntaen korkeamman radiotaajuuden (RF) matalammalle välitaajuudelle (IF) käyttäen taajuuksien sekoittamista paikallisoskillaattorilla (LO). Mikserin ulostulosignaali on signaalinprosessoinnin näkökulmasta paljon käytännöllisempi. Alaspäin taajuusmuuntavan mikserin tulee pystyä vahvistamaan laajaa skaalaa erivahvuisia signaaleja, minkä ylärajaa mittaamme 1 dB kompressiopisteellä. Radiolaitteistossa kohinaluku tulee yleensä myös ottaa huomioon, mutta johtuen mikserin sijainnista vastaanotinketjussa, kohinaluku vaimenee suhteessa sitä edeltävien vahvistuksien verran, eikä siksi ole niin kriittinen.

Tämä diplomityö esittelee lukijalle ensiksi alaspäin muuntavan taajuussekoittimen perusteorian, toisena sen teoreettisen piirikaavion suunnittelun sekä sen simuloinnin tuloksia, ja viimeisenä fyysisen layoutin suunnittelun sekä sen simuloinnin tulokset. Fyysisen layoutin suunnittelu ja simulointi tehdään käyttäen GlobalFoundries 22nm FDSOI teknologiaa. Suunnittelu tehdään suoramuunnosvastaanottimelle käyttäen Gilbertin solu topologiaa, eliminoiden peilitaajuuksista aiheutuvat ongelmat, sekä vähentäen kohinan sekä ei-haluttujen signaalien läpivuotojen vaikutusta. Sekoittimen tulee muuntaa 151 GHz signaali n. 0 – 1 GHz kantataajuudelle käyttäen LO-signaalia taajuusvälillä 150 – 151 GHz. Lopullisissa tuloksissa vertaillaan kahta eri iteraatiota. Ensimmäisenä versiota, joka luotiin alun perin teoriapohjaisen piirisuunnittelun pohjalta, sekä toista versiota, missä useilla parannuksilla mikserin suorituskykyä saatiin parannettua. Korkeaimpedanssista 500 Ohmin kuormaa ajaessa mikseri ylsi -10.0 dB vahvistukseen, 1 dB kompressiopiste oli 6.4 dB kuluttaen 4.7 mW virtaa käytössä. Kohinaluvuksi simuloitiin 17.3 dB, sekä LO signaalin vuodosta IF lähtöön oli -27.7 dBm.

Avainsanat: Sekoitin, Gilbertin solu, 22nm, D-band, CMOS-SOI, alassekoitus

TABLE OF CONTENTS

FOREWORD

First and foremost, I would like to give my biggest gratitude's to Prof. Aarno Pärssinen for not only guiding me along the way on this thesis to get it finished on time, but also by offering me this thesis position in the first place. I would also like to thank Olli Kursu and Mikko Hietanen for their supervising and technical guidance, without which the thesis would probably have never finished on time.

LIST OF ABBREVIATIONS AND SYMBOLS

5G	Fifth generation
6G	Sixth generation
AC	Alternating current
CG	Conversion gain
CMOS IC	Complementary metal-oxide-semiconductor
DC	Direct current
DRC	Design rule check
DSB	Double sideband
EM	Electromagnetic
EMX	Electromagnetic extraction
FET	Field effect transistor
FDSOI	Fully depleted silicon-on-insulator
GHz	Gigahertz
HB	Harmonic balance
HBAC	Harmonic balance AC analysis
HBnoise	Harmonic balance noise analysis
IF	Intermediate frequency
IP2	Second-order intercept point
IP3	Third-order intercept point
LNA	Low-noise amplifier
LO	Local oscillator
LVS	Layout vs schematic
mmWave	Millimeter wave
NF	Noise figure
PDK	Process Design Kit
PEX	Parasitic extraction
RF	Radio frequency
SNR	Signal-to-noise ratio
SSB	Single sideband

α	Amplitude
A	Amplification
C_{gs}/C_{ds}	Gate-Source/Drain-Source capacitance
f/ω_{RF}	RF signal frequency/angular frequency
f/ω_{IF}	IF signal frequency/…
f/ω_{LO}	LO signal frequency/…
f_{Max}	Maximum oscillation frequency
f_t	Unity current gain frequency
g_m	Transconductance
J_{opt}	Optimal current density
K	Gain

L_s/L_g	Source/gate inductance
L	Length
V/P_{RF}	RF signal voltage/power
V/P_{IF}	IF signal voltage/…
V/P_{LO}	LO signal voltage
ω	Angular frequency
W	Width (Transistors)
Z	Impedance

1   INTRODUCTION

The downconversion mixer is one of the core components in a mobile receiver structure, giving access to more realistic specifications, which allow mobile communications to work as a whole. In the recent years with the rise of 5G and 6G research, the demand for devices capable of handling frequencies in the mmWave region has increased exponentially, and as a result, research is being heavily focused on those higher frequency bands [1][2]. Low-cost CMOS silicon-on-insulator technologies have also proven to be a valid candidate for mass production of these high frequency chips, being able to produce lower parasitic, lower power consumption and higher corner frequencies than previously used processes [3][4]. Devices utilizing smaller nanometre scale technology are allowing us to create more compact devices than ever before, although it is also largely enabled by beamforming and antenna array technology [5].

In a direct conversion receiver structure, mixers are devices used to convert an RF-signal from a high frequency (between 110-170 GHz for D-Band) down to a more manageable frequency. Mixers come with performance parameters such as linearity, conversion gain and noise figure, which can all be modified to fit our needs. Out of the three, the linearity is often the most important for receiver performance partitioning, since the mixers ability to drive a high range of amplified RF signals is much more important than its ability to apply gain to it. Additionally, since gain and linearity conflict with each other, we more often see designs that sacrifice gain with the goal of achieving higher linearity. Noise figure is not to be neglected completely, but being preceded by the highly amplifying RF amplifier, mixer noise figure is heavily attenuated and thus works more as an indicator to how well the mixer is operating. Several works of RF-receivers in the D-band have been employed, where downconversion mixers are included [1][2][6][7].

In this thesis, a downconversion mixer using the classic Gilbert cell topology was designed in the D-band. The RF signal of interest is between 150-151 GHz, and the resulting IF frequency should fall between Zero-IF and 1 GHz. Zero-IF receivers have gained popularity, due to their image rejecting properties. Initially, the mixer is unable to reach any kind of reasonable performance in post-layout simulations. Hence, a problem analysis is also presented in this topic, where the solution is explained. After adjustments, the mixer is able to reach a 1 dB compression point of 6.4 dBm with a conversion gain of -10.0 dB.

The frequency step taken is relatively large compared to previous studies, which favoured two-stage downconversion, where the D-band frequency is first downconverted to approximately 10-40 GHz [2][8].

1.1   Goals and motivation

As previous studies have shown, D-band receivers have been shown to function to an acceptable degree in their given frequency band. The opportunity for new bandwidths that allow for faster rates is not one to be given away without fully exploring it first, hence continued research into it is necessary. The downconversion mixer is one of the most vital blocks in the receiver structure, giving ample motivation for fresh research angles at its problems.

Therefore, the goal of this thesis is to design and simulate a downconversion mixer in the D-Band, while identifying and trying to solve any problems that appear throughout the

research. The thesis takes a specific look at the downconversion mixer, from the perspective of a low supply voltage (800mV). The 800mV is a shared supply voltage between all the devices on the final receiver chip, so the mixer must also comply with the requirements. Additionally, it is the maximum recommended voltage supply for the technology used. Typical mixers of the same nature often utilize 50% or even more supply voltage. The target of the thesis is to provide a complete design of the downconverter in the highest layer provided by the process while providing reasonable performance. If performance is not up to par, explanations and improvement ideas should be able to make up for them. Going even further than that, it would be ideal if the mixer would be able to perform with a balanced 1dB compression point of 0 dBm and conversion gain of 0 dB.

1.2   Thesis structure

The thesis begins with a theoretical explanation on mixers. A slow build-up is done from the ideal switch mixer to the more complicated Gilbert cell. In Chapter 2.2, the most essential parameters from the perspective of mixer performance will be described in detail.

Following the theory, the simulation methods for acquiring these parameters will be explained in Chapter 2.3. Finally, then the circuit and layout design will be explained in Chapters 2.4 and 4, and finally the results and its discussion are shown in Chapter 5.

2 DOWNCONVERSION MIXER

The fundamental structure of a RF receiver requires at least one downconversion mixer to assist channel selecting performed after downconversion at a lower frequency. The following chapters will cover general mixer theory and the most essential mixer design parameters. A design methodology around the circuit design of the downconversion mixer used in the work will also be shown, giving multiple angles at the different ways one can parametrize a double-balanced mixer.

2.1 Mixers in general

At its core, a mixer is a device that takes two input signals with varying frequency and outputs a frequency that is either the sum or subtraction of the two after multiplication process in the time domain. This output is referred to as the intermediate frequency (IF). A simple block diagram is illustrated below, showing the basic function of a mixer in both up- and downconversion.
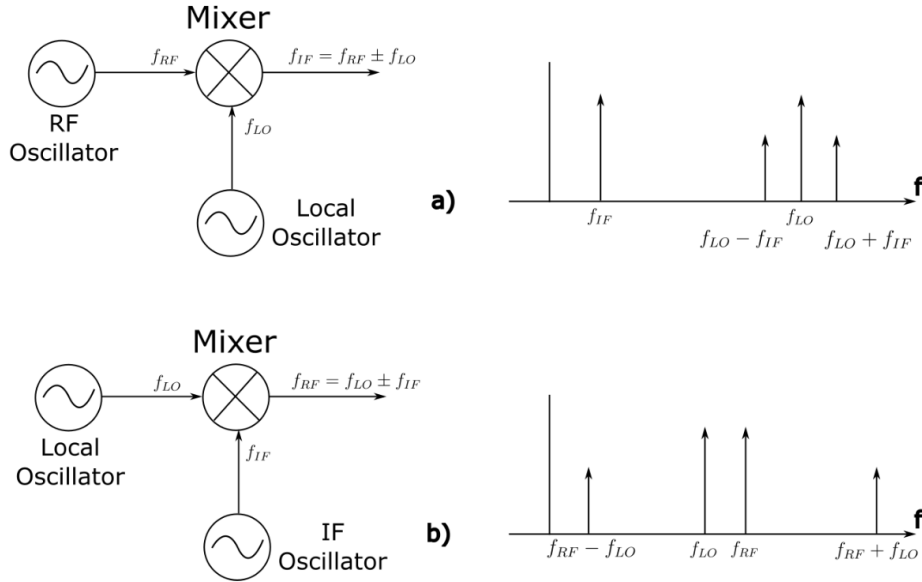

Figure 1. Block diagrams and spectrums for (a) up-conversion and (b) downconversion.

In both cases, we are mixing two frequencies together and then filtering out the wanted product. Focusing on downconversion, we can mark the RF input signal as

$$V_{RF}(t) = \cos(2\pi f_{RF}t),$$

which gets applied to the mixer alongside the LO signal:

$$V_{LO} = \cos(2\pi f_{LO}t).$$

The mixer performs a frequency multiplication at the output:

$$V_{IF}(t) = KV_{RF}(t) * V_{LO}(t) = K\cos(2\pi f_{RF}t) * \cos(2\pi f_{LO}t)$$

$$= \frac{K}{2}(\cos{(2\pi(f_{RF} - f_{LO})t} + \cos(2\pi(f_{RF} + f_{LO})t), \tag{1}$$

where K is the gain applied to the RF signal. The formula shows that the output does include both the sum and subtraction between the two signals. These two products are considered as the sidebands of the mixer. The sum product being the upper sideband and the subtraction product the lower sideband. The principle is the same for up-conversion, but the RF port is now used as the output and the LO signal is mixed with the IF signal [9]. Next, some common topologies for mixers will be discussed.

2.1.1   *The ideal switch mixer*

The structure of a mixer can be simplified by using an ideal switch, which is turned on and off by the signal coming from the LO port.
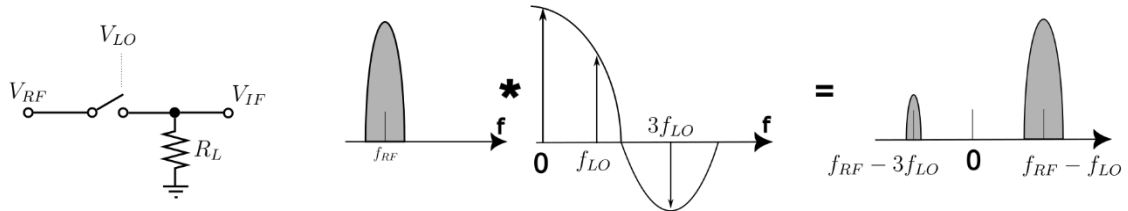

Figure 2. Ideal switch mixer and the principle of frequency multiplication.

Illustrated in the above Figure, we find how the input signal ω_{RF} is modulated (multiplied) by the ω_{LO} signal, which acts as a signal for the switch. The design can be realized using transistors as depicted in Figure 3.
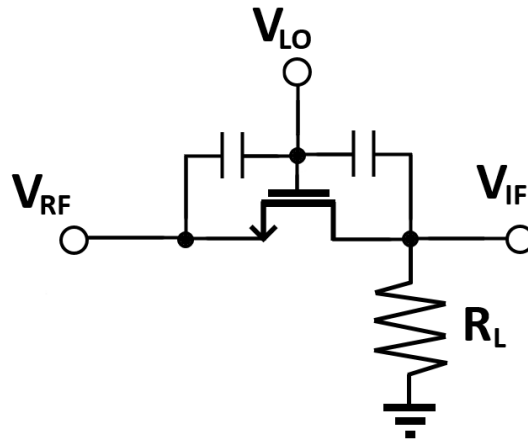
Figure 3. Ideal switch mixer realized using a MOSFET.

The output shows the mixer products $\omega_{RF} - \omega_{LO}$ and $\omega_{RF} - 3\omega_{LO}$, the latter which is an unwanted mixing spur caused by the instantaneous switching also being sensitive to the lower amplitude harmonics of the LO port. Thus, the mixer is by nature nonlinear, and the topology can be improved with different techniques to remove unwanted harmonics from the output. The mixer shown is also often called an unbalanced mixer, since all the inputs are one-sided and don't operate with differential signals.

2.1.2 Passive and active mixers

Mixers can be divided into two groups: Passive and active mixers. The fundamental difference between the two is the active mixer possessing the g_m stage which is used to provide an active element which amplifies the IF signal. Therefore, the passive mixer will always provide conversion *loss*, rather than conversion *gain* which the active mixer can provide. On the flipside, the passive mixer will typically have a much wider dynamic range and linearity, which is a by-product of the device reaching saturation more slowly at the output with conversion loss in the system.
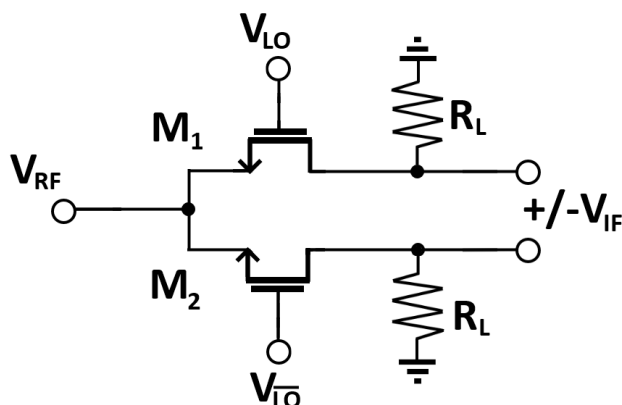
Figure 4. Passive single-balanced mixer.

Figure 4 illustrates how the passive mixer structure is built, where the RF frequency is fed directly into the source of the LO transistor pair. Additional benefits are simplified design. The passive mixer does not require any passive power consumption and thus requires no supply voltage at drain like the active mixer does. Passive mixers were illustrated in all previous examples, for example in Figure 3, where the RF port is fed directly to the source of the LO transistor. The simplicity of the build also means there are less noise sources, minimizing NF.

Figure 5. An active single-balanced mixer

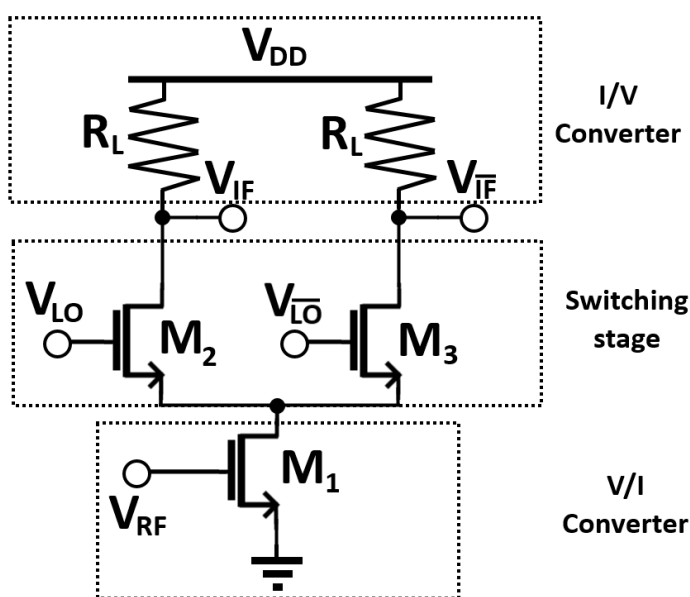Figure 5 shows an active single-balanced mixer. The active mixer consists of three segments: the V/I converter which operates by converting the V_{RF} voltage into a small-signal current in the transistor M_1. The switching stage is then used to steer the current and produce the intermodulation products, which are realized in the I/V converter stage by turning the current back to voltage over the drain resistor.

2.1.3   Single-balanced and double-balanced mixer

The previously shown mixer shown in Figure 3 only utilizes one half of the LO signal, since the mixer will only be active when the switch is turned on during the positive half cycle of the LO signal. Single-balanced mixer topologies aim to fix that by turning the LO port into a differential input, where both positive and negative half-cycle of the waveform can be utilized to switch the mixer on and off. Due to this modification now effectively causing the mixer to be on twice as often, the conversion gain is also doubled. Schematics showing the simplified mixer along with the implementation are presented below in Figure 6.
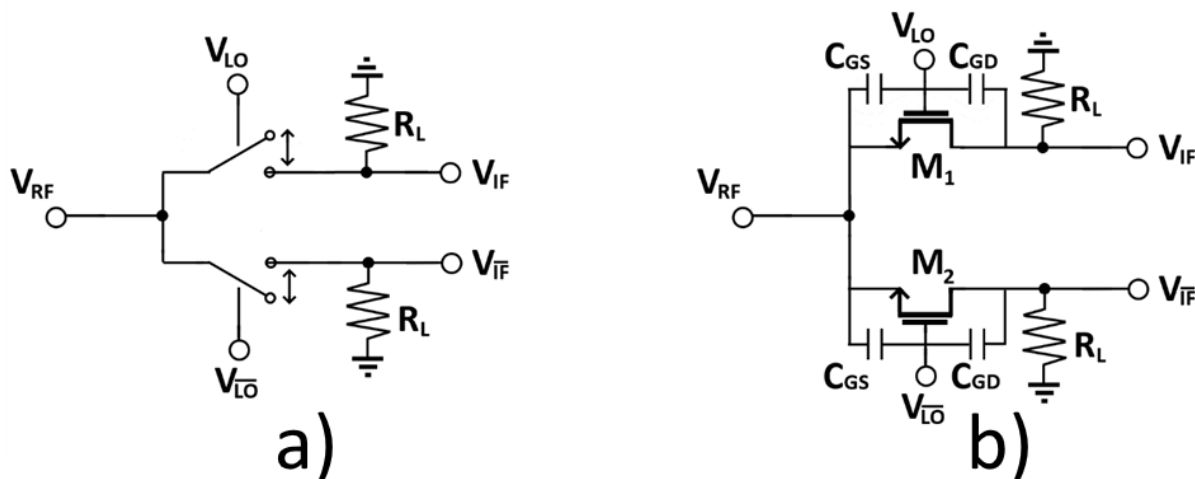
Figure 6. (a) SB-mixer and its (b) implementation with MOSFETs

Additionally, if perfect symmetry is achieved then the problem with RF-LO feedthrough and LO self-mixing is nullified (Explained more in 2.2.4). While this topology has some improvements to the previous one, it suffers from many shortcomings, namely severe LO-IF feedthrough.

The improvement to be made is done by connecting another single-balanced mixer to the first one in a way that causes the LO feedthroughs to cancel each other out in a manner illustrated in Figure 7.

Figure 7. Passive double-balanced mixer

The negative LO port is connected to the positive end of the other SB-mixer, meaning that whatever the feedthrough is, they will cancel each other due to the 180-degree phase difference, The LO-IF feedthroughs will be identical to each other, if perfect symmetry is achieved between the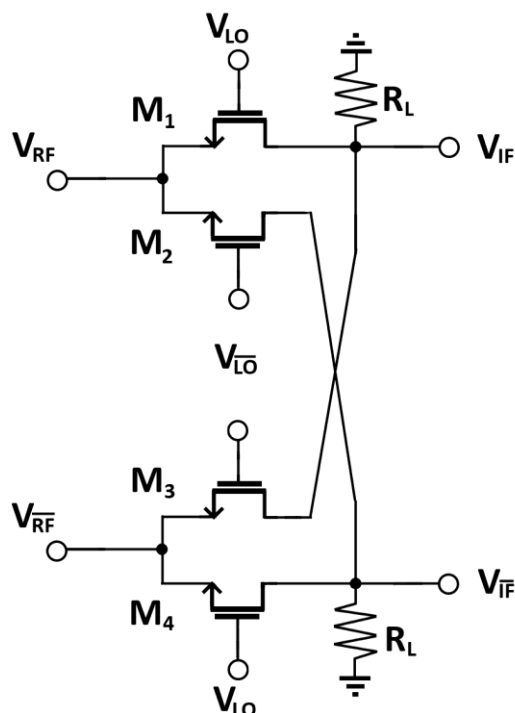 two LO-ports. In an asymmetrical case, the amplitudes and phases of the feedthroughs can change, causing output deteriorating feedthrough of the LO signal. This topology, which will be the one used for this thesis, is also considered the Gilbert cell, proposed in 1968 by Barrie Gilbert [10].

2.2   Mixer parameters

The basic parameters used to describe a mixer's operation will be introduced next. The three most important parameters include linearity, conversion gain and noise figure. Generally, more emphasis is given towards linearity than gain, since the mixer is preferably capable of handling a wider range of input signals, rather than amplifying it. The effect of mixer's noise in the receiver chain is often negligible, due to the mixers placement in the receiver chain allowing it to be attenuated. The mixer's noise figure itself is not negligible. This will be discussed more in the coming chapter 2.2.6.

2.2.1   Linearity

Linearity refers to a mixers (or any amplifiers) upper limit of operation in large signal conditions, before either compression or the third-order intermodulation products start to interfere with the signal amplification or desensitize reception of weak signals. Perfect linearity would mean that the output is always proportional to the input. Linearity often has two main performance parameters that are used to measure it: 1dB compression point and the third-order intercept point.

The hypothetical third-order intercept point is a commonly used method to measure linearity. The idea comes from placing two closely spaced tones ω_1 and ω_2 at the input port:

$$v_i = V_0(\cos \omega_1 t + \cos \omega_2 t), \tag{2}$$

where v_i is the input voltage and V_0 the initial amplitude of the signal. Dissecting the first three terms of the output using Taylor series:

$$v_o = a_0 + a_1 V_0(\cos \omega_1 t + \cos \omega_2 t) + a_2 V_0^2 (\cos \omega_1 t + \cos \omega_2 t)^2$$
$$+ a_3 V_0^3 (\cos \omega_1 t + \cos \omega_2 t)^3 + \cdots$$

$$= a_0 + a_1 V_0 \cos \omega_1 t + a_1 V_0 \cos \omega_2 t + \frac{1}{2} a_2 V_0^2 (1 + \cos 2\omega_1 t) + \frac{1}{2} a_2 V_0^2 (1 + \cos 2\omega_2 t)$$
$$+ a_2 V_0^2 \cos(\omega_1 - \omega_2) t + a_2 V_0^2 \cos(\omega_1 + \omega_2) t$$
$$+ a_3 V_0^3 (\frac{3}{4} \cos \omega_1 t + \frac{1}{4} \cos 3\omega_1 t) + a_3 V_0^3 (\frac{3}{4} \cos \omega_2 t + \frac{1}{4} \cos 3\omega_2 t)$$
$$+ a_3 V_0^3 \left[\frac{3}{2} \cos \omega_2 t + \frac{3}{4} \cos(2\omega_1 - \omega_2)t + \frac{3}{4} \cos(2\omega_1 + \omega_2)t \right]$$
$$+ a_3 V_0^3 \left[\frac{3}{2} \cos \omega_1 t + \frac{3}{4} \cos(2\omega_1 - \omega_2)t + \frac{3}{4} \cos(2\omega_2 + \omega_1)t \right], \tag{3}$$

from which we can make the following discovery: The third-order product increases as the cube of the initial voltage (V_0^3). Third-order products are defined as the frequency components with three components in some combination, for example $(2\omega_1 \pm \omega_2)t$ or $3\omega_1 t$. For small input signals, this won't be a problem, but as the input power increases, the third-order product power will quickly rise. It can also be noted, that the third-order products $(2\omega_1 \pm \omega_2)t$ land relatively close to the ω_1 and ω_2 frequencies, resulting in difficulty in them being filtered out in the system. As visualized in Figure 8, the third-order intercept point is the point at which the cubic third-order product (n=3) overtakes the linear response (n=1). Although as it was mentioned, the point is purely hypothetical and uses theoretical linear curves, in practice both the third-order product and the linear response compress before that [11].

This compression comes from the 1dB compression point, an often used measurement of linearity. The 1 dB compression point is usually around 10 dB lower than the IIP3, and it describes the input power level at which the output power has depreciated by 1 dB from the ideal linear output curve. Illustrated in Figure 8, the linear response ceases being true at some point, when the compression of the system starts taking over and flattening the curve.
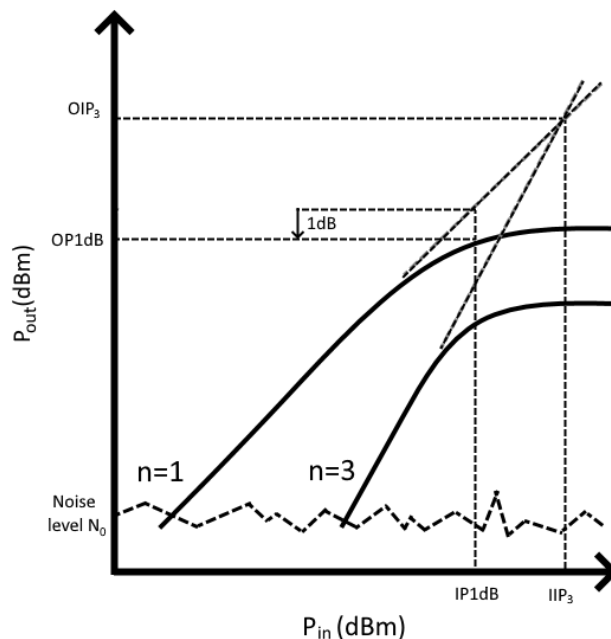
Figure 8. The 1st (n=1) and 3rd (n=3) order input signals saturating at different output powers.

It should also be noted that in direct conversion receivers, there lies the problem of even-order distortion. Similarly, to how IIP3 is tested with a two-tone test, this effect can be realized by placing two strong interferers ω_1 and ω_2 next to the desired channel. Upon going through the receiver structure, they first pass through the LNA creating a low-frequency beat at $\omega_2 - \omega_1$, which by itself is not an issue due to being of low amplitude at a very low frequency. Problems arise when the desired channel is downconverted to the low frequency as well, which lets a fraction of this new "beat" distort the downconverted signal. Majority of the signal is upconverted to a frequency where it doesn't cause issues, but due to feedthrough, some gets through unconverted.

This gives rise to a new variable that is used to measure the RF downconverter linearity, the second-order intercept point or IP2, similar to the IP3, but now instead of the interest being in third-order intermodulation products, it is the second-order intermodulation products, or the beat frequency.
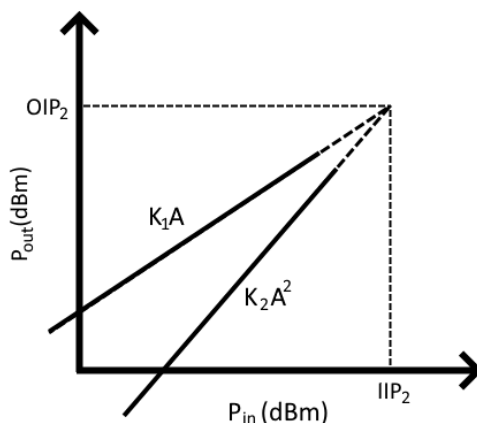

Figure 9. Second-order intercept point (IP2)

The amplitude follows a similar curve as IIP3. The second-order product amplitude increases at a steeper curve than the first-order product, causing an intercept point to happen at some input power.

$$= K_1 A(cos\omega_1 t + cos\omega_2 t) + K_2 A^2 \cos(\omega_1 + \omega_2) t + K_2 A^2 \cos(\omega_1 - \omega_2) t$$
$$V_{out}(t) = K_1 V_{in}(t) + K_2 V_{in}^2(t) \tag{4}$$

Although ideally the second-order intermodulation products in the differential paths of the mixer are in equal phased, or common-mode, meaning they cancel each other out in the output. Asymmetrical design flaws will ultimately result in this ideal case not applying anymore and the IM2 products causing issues. The linearity of the mixer is greatly determined by the linearity of the amplifier stage of the mixer. A well-designed amplifier will be able to fill out the term $V_{DS} = V_{RL}$, meaning the voltage headroom is maximized.

2.2.2 *Noise*

While the upper limit of operation in RF systems is limited by linearity, the lower limit is limited by noise. Without any noise, a receiver would be able to detect infinitely small signals. The noise floor of the system is the absolute minimum above which the signal has to be in order to be detected. In RF circuit design, we have the ability to identify and model the most common sources of noise with voltage and current sources.

The thermal noise of resistors is the most typical source of noise. Thermal noise is a result of excitation of charge carriers in resistors caused by heat coming from the outside. The result is an additional voltage component for the resistance R that follows the power spectral density of

$$\overline{V_n^2} = 4kTR, \tag{5}$$

where k is the Boltzmann constant of $1.38 * 10^{-23} \frac{m^2 kg}{s^2 K}$ and T the ambient temperature (typically assumed to be around 290 Kelvin). From the formula we can also derive the noise as a Norton equivalent parallel current source with a PSD of

$$\overline{I_n^2} = \frac{\overline{V_n^2}}{R} = \frac{4kT}{R}. \tag{6}$$

Active devices can't emit noise either as they have their own specific noise sources. The channel noise is modelled as a current source between the source and the drain of the transistor, with a PSD of

$$\overline{I_{n-mos}^2} = 4kT\gamma g_m, \tag{7}$$

in which g_m is the transconductance of the transistor and γ a transistor generation specific excess noise coefficient. The MOSFETs gate resistance is another major source of thermal noise. For a certain gate strip width W and length L we can determine a resistance

$$R_G = \frac{W}{L} R_\square,$$

(8)

where $R_\square$ is a material specific value for one square of the gate sheet, often simply referred to as the sheet resistance. The drain and the source also generate noise, but to a lesser extent thanks to the physical layouts typically having multiple fingers that reduce the total resistance.

Thermal noise is not the only source of noise, as MOSFETs also exhibit flicker noise ($\frac{1}{f}$ noise), which is modelled as a voltage source in series with the gate. The power spectral density of this noise source can be calculated as a function of the physical parameters of the transistor

$$\overline{V_{n-f}^2} = \frac{K}{WLC_{ox}} * \frac{1}{f}.$$

(9)

Similarly to $R_\square$, K is a process dependent constant. Generally, flicker noise is a non-issue for broadband mixers, since its bandwidth is limited to >1 MHz. However, an exception is made especially for downconverting direct conversion mixers, where flicker noise may be multiplied as a result of the switching stage. Similarly to normal flicker noise, it appears near DC. Noise from the higher frequencies will be downconverted similarly as the signal from the signal from the same bandwidth, in addition noise around different LO harmonics can be downconverted on top of the received signal [3].

2.2.2.1 Noise figure

The noise in an RF system is rarely described using noise currents and voltages previously described, but rather the noise factor F and the noise figure NF of the system. Noise factor is defined as the relation between the input and the output signal-to-noise ratio (SNR) of the RF system as

$$F = \frac{SNR_{in}}{SNR_{out}}.$$

(10)

The noise figure is then defined as the noise factor on a logarithmic scale as

$$NF = 10 \log(F).$$

(11)

The noise figure is more commonly used than the noise factor, since the noise figure can be directly compared to the logarithmic gain and linearity values.

If we recall the downsides of heterodyne structures (receiver structures that downconvert to a nonzero frequency) we remember that these receivers struggle with the problem of image frequency. This image frequency is a copy of the signal on a frequency, that also gets included in the downconversion to ω_{IF}. As depicted in Figure 10, when both frequency bands ω_{RF} and ω_{image} contain noise, we end up doubling the noise at the output, since the output converts both frequencies.
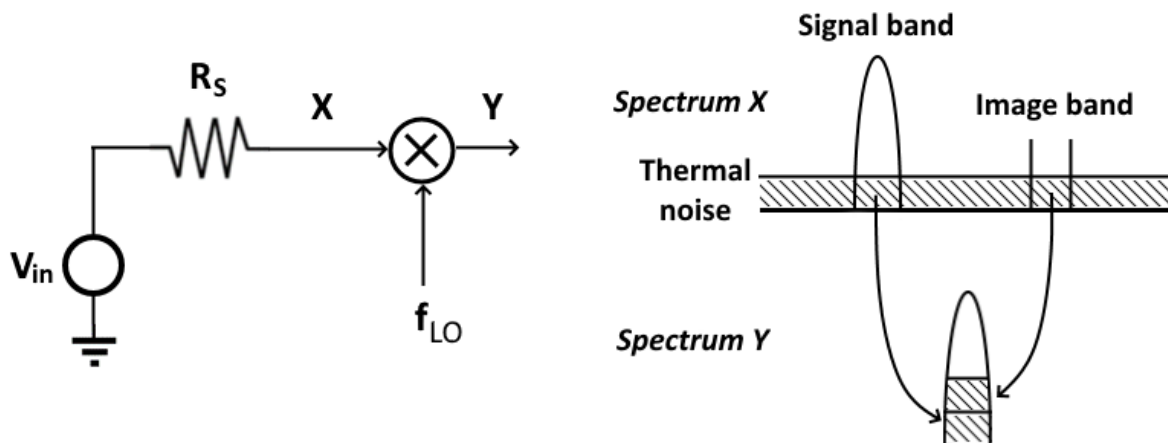
Figure 10. Noise behaviour in double-sideband mixer

This common issue is the cause for even a noiseless mixer to have a noise figure of 3dB, since the noise gets doubled in the output. It should be noted, that in the above case we are talking about single sideband (SSB), where the image band only contains unwanted information. In double sideband (DSB) we can take advantage of this behaviour and include the wanted signal in the image band too, thus amplifying the signal in the output and increasing the SNR. The NF in a SSB system will always be at least 3 dB (two times) higher than of a system that takes advantage of DSB. It is important to realize that this method can only be used in the case of a direct conversion receiver such as ours. In other structures (Basic heterodyne structures) the signal is not included in the lower sideband, causing the DSB noise figure to just be the same as the LSB one, but with double the noise. In downconversion mixers, a single-digit NF is not uncommon to find [2][6][7].

2.2.3 Conversion gain

Typically, a mixer will inherently have some type of conversion loss. This is the case with passive mixers, but development has also been made to optimize active mixers, which are able to provide conversion gain. By definition, conversion gain (CG) refers to the rms-voltage or power ratio between the IF voltage that in the output, and the RF voltage in the input. Figure 5 illustrated how the active mixer can be realized as a type of common-source amplifier, where the amplifier is turned off for the duration that the LO switching transistors are turned on. The gain for a common source amplifier is known to be as

$$A_{cs} = \frac{V_{out}}{V_{in}} = g_{m1}R_D \tag{12}$$

but with the addition of the ideal LO switching (cutting down the time the amplifier is on) the amplification is reduced to

$$A_{mixer} = \frac{V_{out}}{V_{in}} = \frac{V_{IF}}{V_{RF}} = \frac{2}{\pi} g_{m1}R_D \tag{13}$$

The $\frac{2}{\pi}$ component is a direct result of the square wave mixing in the LO ports. It should also be recognized that the conversion gain is also heavily limited by the I_D current that the device

is biased to, and also the available V_{DD} voltage which will limit the voltage swing at the output. This can also be seen as a problem of linearity.

2.2.3.1   LO waveform non-idealities.

Previously it was assumed that the switching happens instantly in the LO ports, or that the LO signal is a perfect square wave. In real cases, the LO port is usually driven by a sinusoidal wave that tries to achieve square-shaped behaviour with a high enough amplitude. The cut-off for the positive and negative side occurs at the saturation and off-state of the transistor, meaning the time spent in the triode and linear region of the transistor is wasted. Figure 11 illustrates how increasing the LO drive can be used to alleviate this issue. As the sinusoidal waves amplitude increases, less time is spent in the switching phase between inactive and active region, increasing the efficiency of the mixer.
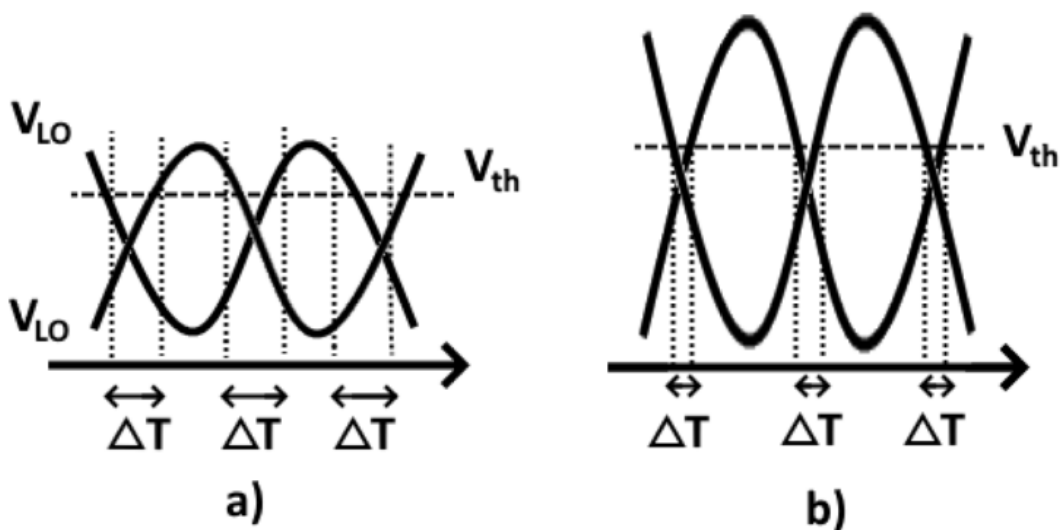

Figure 11. a) Low LO power vs b) High LO power and its effect on inactive period ΔT of the switching mechanism

This creates another balancing issue with the mixer. Driving the LO ports will guarantee behaviour closer to the ideal case of 2.2.3 but will also inadvertently increase the power draw of the entire device. In addition, the implementation of such a high LO power is going to be difficult from a practical viewpoint. In typical cases, and as will be shown in simulations later, the LO power is seen to provide sufficient power at between 0-3 dBm to a matched LO port, after which the benefits are negligible compared to the difficulty caused by the LO increasing power demands. Other commonly used methods to combat the slow switching is by increasing the width of the switching transistors. A higher width transistor will present higher current density, allowing it to flow current more efficiently, but does not come with its own limitations. A higher width transistor will certainly present more parasitic capacitance than a smaller one, and also a higher demand for LO power to be driven.

Another nonideality which will be investigated is phase offset in the LO waveforms. An ideal balanced mixer works at its best efficiency when the two LO waveforms are at a 180-degree phase difference, but differences in the physical layout can easily cause a delay in one of the signals, shifting the LO waveform to an unwanted phase. The asymmetry causes gain

losses, since an offset means there will be time periods where all transistors are in the same state. Having the transistors in the switching quad be in an equal state means no switching can happen, meaning no IF frequency is generated. As was also explained in 2.1.3, asymmetries like this will also mean that the LO signals in a Gilbert cell will not cancel out as effectively anymore, meaning leakage can occur from the RF and LO ports.

2.2.4   Port isolation and feedthrough

As an unfortunate results of parasitic capacitances, mixers suffer from the different ports coupling into each other and corrupting one and another.
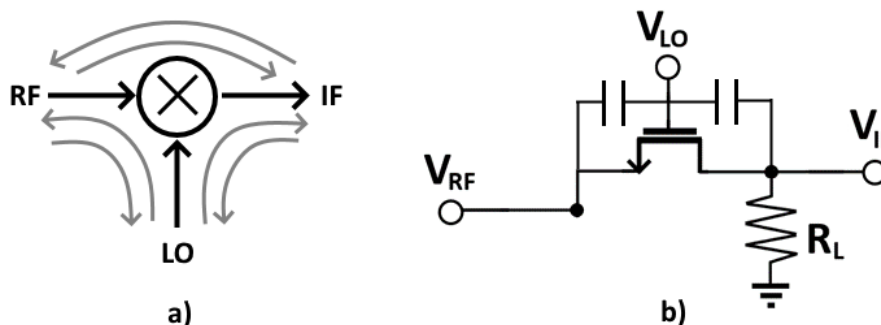

Figure 12. (a) Feedthrough in a mixer, (b) Feedthrough paths in the mixer implemented with MOS transistor

In mixer designs we account for three different types of feedthroughs: RF-LO, LO-IF and RF-IF. The first one of these was mentioned in 2.2.1, as it was one of the causes for the linearity limiting factor IP2. In this work the highest focus is on LO to IF leak, since the RF frequency lands so closely to the LO range that it can be approximated as the same value. Ideally, we would want the IF output to appear a short for the LO frequency, but a 40dB isolation should be achievable with the double-balanced topologies.

   The LO signal leaking towards the RF input gives rise to a phenomenon described as LO self-mixing. When the LO signal finds its way to the RF input, the LO signal starts mixing with itself. A dc-component is generated due to a frequency multiplication of two of the same frequencies. The methods used for preventing feedthrough typically involve the use of buffers, but it can also be seen in mixer topologies that symmetry almost always desired, since the asymmetries are what bring the feedthrough into effect. Single-balanced mixers are much more volatile to LO frequencies being leaked to the IF port, since the switching transistors drain is directly linked to the IF output with no common-mode rejection as in the double-balanced mixer.

2.2.5   Port return loss

The mixer is considered a 3-port device, with two input ports and an output port. One of the most vital elements in an RF device is that these ports are impedance matched. Failure to accomplish will result in what is called a reflection coefficient, defined by the impedance mismatch happening inside the port

$$T = \frac{Z - Z_0}{Z + Z_0}$$

(14)

where Z_0 is the reference impedance, typically 50 Ohms for an RF device. Z is the impedance the input port sees, for example, an RF port seeing the gate of a transistor. The port return losses are analysed for all 3 ports for a mixer device using S-parameters as follows:

- S11 – The input (RF) port return loss
- S22 – The input (LO) port return loss
- S33 – The output (IF) port return loss

The S11 and S33 ports are of highest priority since losses in those will directly result in attenuation in converting signal. As we will find later, a decrease in the LO power will also result in degradation of the conversion gain, but in its case the losses can be compensated for with more power, although a well-designed mixer will likely still account for every port for maximized efficiency.

2.2.6 Mixers in transceiver chains

While the entire receiver is far outside the scope of this thesis, knowing the mixers placement in the receiver chain lets us pick some priorities from the basic parameters.
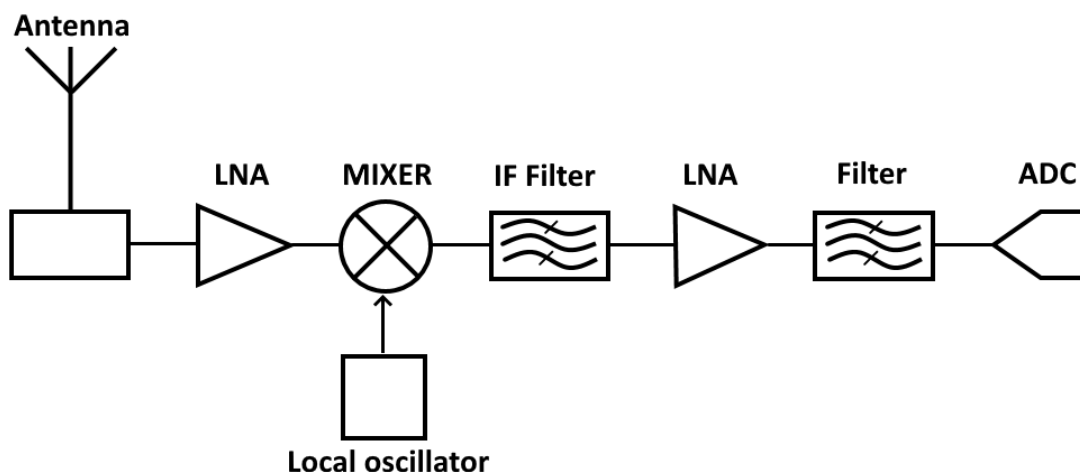

Figure 13. A generic RF receiver chain

Typically the mixer will be very early in the receiver structure, only preceded by the antenna and the low-noise amplifier. The three main parameters which we are interested in (linearity, noise, conversion gain) behave differently when the system is cascaded combined parameter is calculated.

$$F_{tot} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} \ldots \frac{F_n - 1}{G_1 \ldots G_{n-1}} \ldots \tag{15}$$

$$\frac{1}{P1dB_{tot}} = \frac{1}{P1dB_1} + \frac{G_1}{P1dB_2} + \frac{G_1 G_2}{P1dB_3} \ldots \frac{G_1 \ldots * G_{n-1}}{P1dB_n}, \tag{16}$$

$$G_{tot} = G_1 + G_2 + \cdots G_n. \tag{17}$$

The first equation (15) is the commonly known Friis equation used to describe noise factor in a receiver chain [12] In the equation n refers to the order at which the component is seen in the receiver chain, so in the case of Figure 13 the order of the mixer in the signal chain would be three. The values used for calculating are absolute values. Seeing as the noise figure of the third block in the receiver chain is attenuated by both G_1 and G_2 ($G_2 >>> G_1$) we can tell that the noise figure of the mixer will not drastically affect the total noise figure of the system. This means that while the noise figure requirement can be kept fairly high from the perspective of the other components in the chain, it should still be at a reasonable level. A high noise figure can be a telltale sign that something else in the circuit is wrong.

2.3   Simulations methods and tools

Typical simulations approach circuits from a steady-state response at a certain frequency. For mixers this approach falls short, as the response we are interested in often spans between multiple frequencies due to intermodulation. The simulation must be able to handle knowledge of a center frequency around which intermodulation harmonics will form.

When using Spectre developed by Cadence Design Systems, there are multiple options and approaches that one can take to complete the simulations for the parameters we are interested in. The small-signal analysis is done using harmonic balance (HB) simulations, accompanied by harmonic balance AC (HBAC) simulations for small-signal analysis. The harmonic balance simulations achieve similar results as steady-state response simulations, but in the frequency-domain rather than time-domain. At the bare minimum, we are interested in two frequencies, RF and IF. Harmonic balance simulations allow us to do more though, and for example let us see the entire IF bandwidth over a range of RF input signals. The noise simulations are performed with harmonic balance noise (HBNoise) simulations and with calculated noise figure over the output frequency range. The parameters we want are simulated in the following ways:

Conversion gain and feedthrough: The harmonic balance simulation is first run as a pre-requisite for running harmonic balance AC simulations. The harmonic balance simulation is run at the center frequency of 150 GHz with high enough harmonics to inspect the frequencies of interest. To avoid problems in analysing downconverted DC signal the harmonic balance AC simulation is run at a range of 150.01 to 151 GHz, while also having the sideband of 10 MHz to 1 GHz selected for inspecting the IF frequency. After the simulation has run, the input and output signal level can be compared, while being mindful that the input frequency was in the first range specified, and the output frequency in the second range specified. Similarly, we can also inspect other frequencies, for example to find out how much a certain frequency component had leaked into the output.

Linearity (P1dB): Harmonic balance is first run at two tones: The RF and LO tone (151 and 150 GHz), included in this simulation is also a sweep for the RF input power. Next, harmonic balance AC is run at a single frequency of 150 GHz, and 1 GHz is selected as the sideband of interest. Spectre can calculate the 1dB compression point using an ideal gain curve at the 1

GHz IF output and comparing it to the results from the RF power sweep. At some power value, the gain will have dropped at least 1 decibel from the ideal curve, which is the compression point.

Noise: The noise of the mixer can be calculated by specifying the input (RF) and the output (IF) ports inside the noise simulation. The output frequency sweep must also be specified, to get the full range of noise inside the bandwidth. After the simulation the integrated noise figure is calculated over the bandwidth to get a noise value.

Port return loss: Port return losses are simulated with the s-parameter (sp) simulations. Specifying the RF port as S1, LO port as S2 and IF port as S3, we can simulate the 3-port system. The simulation can be done for a frequency spectrum to diagnose any mistuned ports, or with a parameter sweep in an attempt to tune a port to a certain frequency. Smith charts were employed in the design to find the right components in the right place, as the impedance and admittance can be manipulated into the correct directions with parallel/series active components.

2.4   Circuit design for upper mmWave

The design of the Gilbert cell was conducted from several different angles. Literature [13] covering the subject with a theoretical approach was taken, particularly useful was the book High Frequency Integrated Circuits by Sorin Voinigescu [14], which helped create a baseline for the mixer. Further optimizations were then made using parameter sweeps to find the optimal values for DC-biases, LO power etc. In the design, emphasis was put on the linearity of the design. Additionally, a matching network with baluns was designed for both RF and LO ports. The Gilbert cell was chosen based on the availability of literature surrounding its strengths and weaknesses [15][16]

It was decided that the mixer will operate at an RF frequency of 151 GHz and an LO frequency of 150-151 GHz, thus, making the output IF frequency to be between 0-1 GHz. Taking into account the image frequency as well, the total bandwidth of the device will be 2 GHz, although actual implementation of this requires I/Q mixers. The mixer will be supplied by a voltage supply of 800mV. Although typically, the bandwidth is not defined so strictly, but rather the 3dB bandwidth of the device is measured after it has been created. Several works describe ways to achieve this [17][18][19], making wideband mixers more than viable.
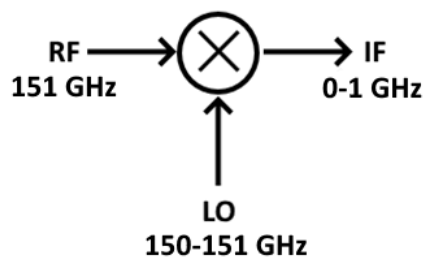

Figure 14. Mixer operation

2.4.1 Design via theory

In this work the design methods and procedures described by Sorin Voinigescu were followed to create a baseline for the design. Voinigescu's book "High Frequency Integrated Circuits" [14] describes many of the design obstacles that come with not only mixers and will likely be used in future improvements to the mixer as well after the thesis.

The first step was to observe any requirements or restrictions we had for the design. For this work, only one requirement was given at the start in addition to frequency range of operation and coarse ideas on gain, noise and nonlinearity. The supply voltage should not exceed 800mV, which already creates limitations. The vertical stacking of multiple MOSFETs very easily creates a bottleneck in the system, which limits the voltage headroom at the output. Secondly, we will draw a schematic loosely based on demands for DC-voltages and currents over the basic structure.
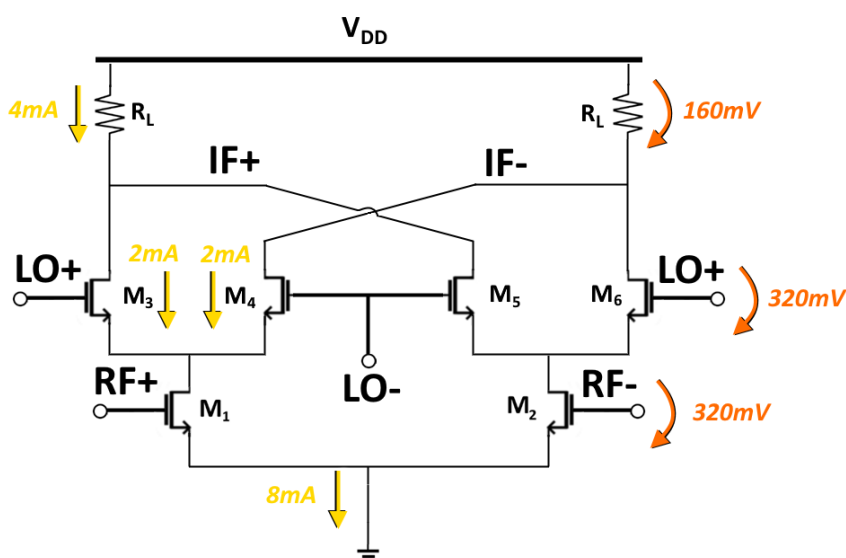

Figure 15. Double-balanced mixer with initial voltage and current biasing.

The voltage drop of 320mV over both stages was decided based on g_m/V_{ds} simulations. A lower V_{ds} would limit the overall current gain in the transistor. With 640mV of the supply voltage spent on the two transistor stages, the leftover voltage is used up on the load resistance as voltage headroom for the output signal. To achieve reasonable resistor values, a DC-current requirement of 4mA in both branches (8mA) total was decided, so the load resistor ends up being $\frac{160mV}{4mA} = 40\Omega$. This design already creates a problem. Since the linearity of the Gilbert cell mixer leans heavily on the linearity of the amplifier stage of the design, maximizing linearity would mean designing the mixer so that the voltage drop over the load resistor is equal to the V_{DS} of the amplifier MOSFET. If we were to meet this requirement, the voltage drop over the LO-transistor in the switching stage would be so low, it would not be able to function properly. If the voltage drops were balanced out to be equal, we would likely not be able to get the amplifier in high enough saturation for constant gain. Voinigescu [14] describes low supply voltages as one of the leading limiting parameters in a Gilbert cell due to the stacking of two transistors.

It is then suggested by [14] to set the current density of the transconductor pair to the minimum NF current density J_{opt} of $0.15 - 0.2 \frac{\mu A}{mm}$ (Based on measurements conducted by Voinigescu and his team [20]) While taking the current demands in consideration, we can deduce that

$$W_{RF} = \frac{I_{Tail}}{2J_{Opt}}.$$

(18)

With a current density of $0.2 \frac{\mu A}{mm}$ we end up with a transistor width of $20 \mu m$. For maximum switching speed, the width of the switching pair is chosen so that the current density is equal to half of the transconductor pair. The width of the switching transistors will be $10 \mu m$, creating a very common mixer design where a 2/1 width split is used between the amplifying and switching phase of the mixer.

The DC bias voltages were chosen in this design to fit the current demands described. With the help of simulated V_{gs}/I_{ds} curves, it was possible to pinpoint the exact voltage locations for achieving these current goals. From there, the LO power was chosen to be high enough to quickly lower the voltage to insert the transistor into inactive mode (0mA) and to the maximum tail current (8mA), so that when the other transistor is perfectly conducting, the other one is completely closed.

This design was merely left as a good starting point for the rest of the design. The simulator results proved that there was a lot of room for improvement. With a LO power of 3dBm, this design was able to achieve conversion gain in the acceptable region of -2.70 dB. The 1dB compression point was found to be much more alarming, only being able to reach a value of -9.82 dBm. The mixer would not be able to function at a wide enough dynamic range.

2.4.2   Design via simulations

While the theoretical approach proved to find a good baseline for the work, the design of a mixer will always vary depending on specifications set by pre-existing devices or other requirements set by the customer or due to small nuances like the type of technology being used. What may have been the best approach for Voinigescu, may not be the most optimal for this work. After the baseline was set, parameter sweeps were conducted from multiple angles to determine the optimal setup for the mixer. Especially the DC bias voltages of the gm and switching stage were of interest, since they can be used to control the conversion gain and linearity of the system.
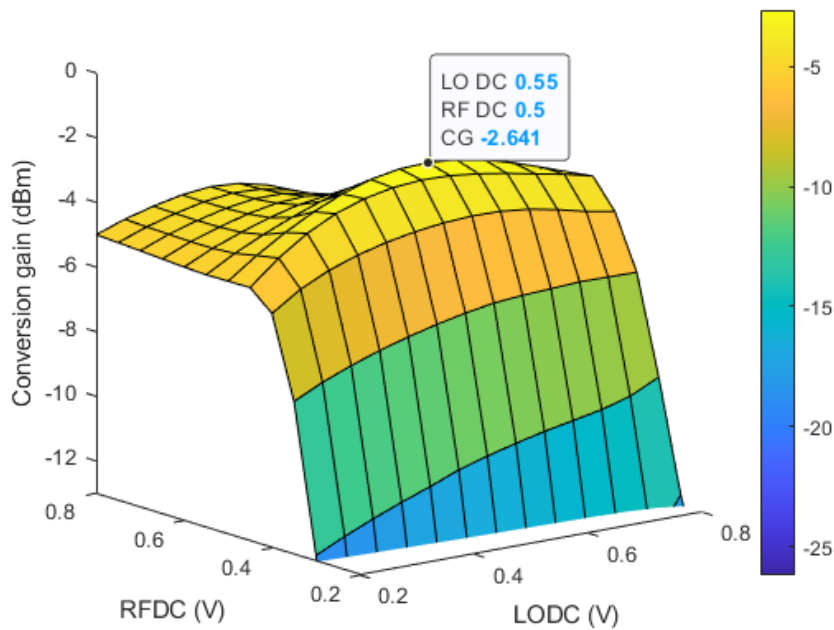
Figure 16. Surface plot of RF and LO DC levels versus conversion gain

The most significant change happens when RF DC bias changes. At low values, the amplifying stage is not able to function, causing the most significant decrease in gain performance. At 0.5 volts RF DC-bias, the amplifier is found to be functioning at its peak performance. The RF bias is often easier to choose than the LO bias, since the RF bias usually only has to fill the requirement of being in the active region.

The LO bias can be more difficult, since finding the optimal point means also taking into consideration how much the LO power swings. From results in Figure 16 we find that the best performance (From gain perspective) is found at 0.55 volts. Since linearity is also of high priority, an additional sweep was conducted at the RF bias of 500 mV to see, how the 1dB compression point acted at different LO DC-bias voltages:
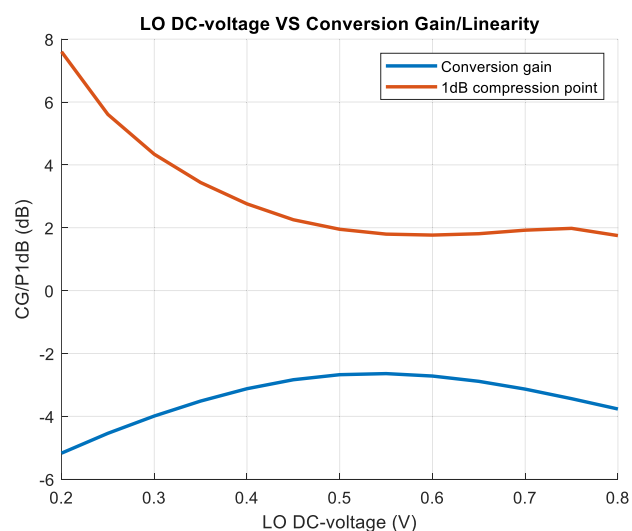

Figure 17. Conversion gain and linearity at various LO biases, when RF bias is at max gain 500mV.

It can be noted how maximum linearity is reached at the lowest conversion gain values, since at lower gain values the output takes longer to reach compression. Figure 17 can be utilized later after the mixer is finished to adjust the parameters to a specific requirement, for example, if a higher linearity was desired, a lower LO DC bias would be picked. For this work, we will remain at the 550mV point for highest gain and good linearity.

2.4.3 LO power considerations.

As was previously discussed in 2.2.3, the LO power plays an important role in the amount of gain the mixer is able to achieve. Ideal behaviour is reached when the LO power is infinitely high. An appropriate value is found by simulating performance parameters in relation to the LO power inserted.
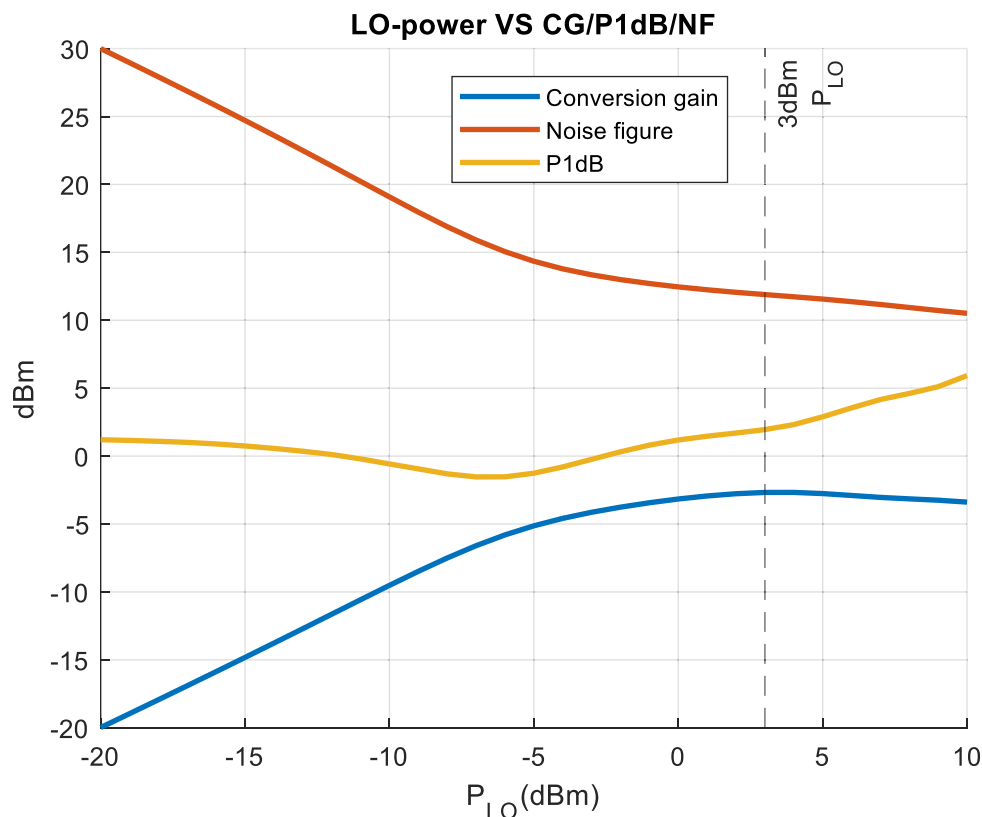

Figure 18. LO power in relation with CG/P1dB/NF

From the simulation results it can be said that the increasing the LO power starts to have diminishing results on the conversion gain around 3 dBm. This is not surprising, as generally the conversion gain is bottlenecked by the output load size and the amplifying stage, rather than the switching speed of the LO transistor. The linearity gains slight improvements even past the 3 dBm mark, but that can be attributed to slower buildup to the compression point due to gain dropping.

2.4.4 Input matching with source degeneration

The input impedance of an LNA or a mixer is capacitive due to the parasitic gate to source capacitance. One common method used in matching to the input is to exploit this behaviour

with an inductor added to the emitter for achieving the wanted impedance at the center frequency. A gate inductor is then used to also cancel out the imaginary part of this impedance, giving us reliable, although narrow banded, matching. This method is maybe more commonly known as an LNA matching method [21][22][23], but since the mixer can be categorized as a common-source amplifier with a switching drain, its benefits can be applied in this work as well. For a MOSFET with inductors L_S and L_G added in series in the source and gate respectively (Figure 19), the input impedance can be approximated as

$$Z_{in}(j\omega) = j\omega(L_S + L_G) + \frac{1}{j\omega C_{gs}} + \frac{g_m}{C_{gs}} * L_S \tag{19}$$

While an equation-based approach can be taken to solve the needed inductor values (assuming C_{gs} is known), in this design the parameters were solved by parameter sweeps and inspecting the change in the S_{11} reflections on the Smith Chart. First, L_S is swept to a value that places it somewhere on the unity circle, and then L_G is swept to bring S_{11} to the center, meaning perfect matching to 50 ohm is achieved. The values in this case were found to be as 80 pH for the gate inductor and 100pH for the source inductor.
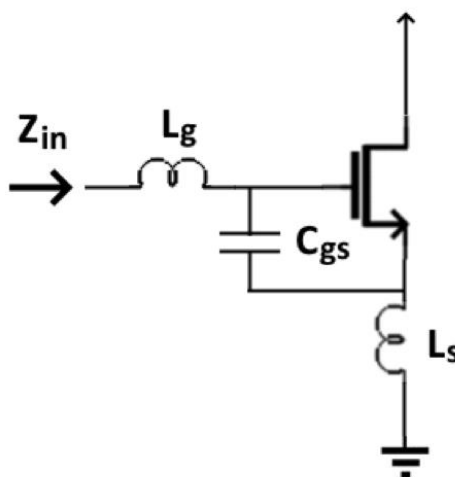

Figure 19. MOSFET with inductors used for resonating with the parasitic capacitance.

With careful small components such as the ones mentioned above, additional care needs to be taken to interface for external world even in probe measurements. The probes used in real-life measurements of the device will add parasitic inductances and capacitances, which are at these values enough to skew with the results. It is therefore of utmost importance, that the device is also simulated with an appropriate probe model attached to it [24].

2.4.5 Output matching

An output matching network is rarely seen in direct conversion receivers, since the frequency band is centered around zero frequency. At these frequencies, components begin to increase in size to a point where they are not in any way practical from an IC-design viewpoint. Additionally, the bandwidth of the output suffers a lot, when the relative size of the bandwidth (From zero to 1 GHz) is large. In upconversion mixers and downconversion

mixers with non-zero IF frequencies, output matching is seen as a much more plausible solution. In direct conversion even wideband signals must be buffered with a broadband amplifier to guarantee matching in measurement environment.

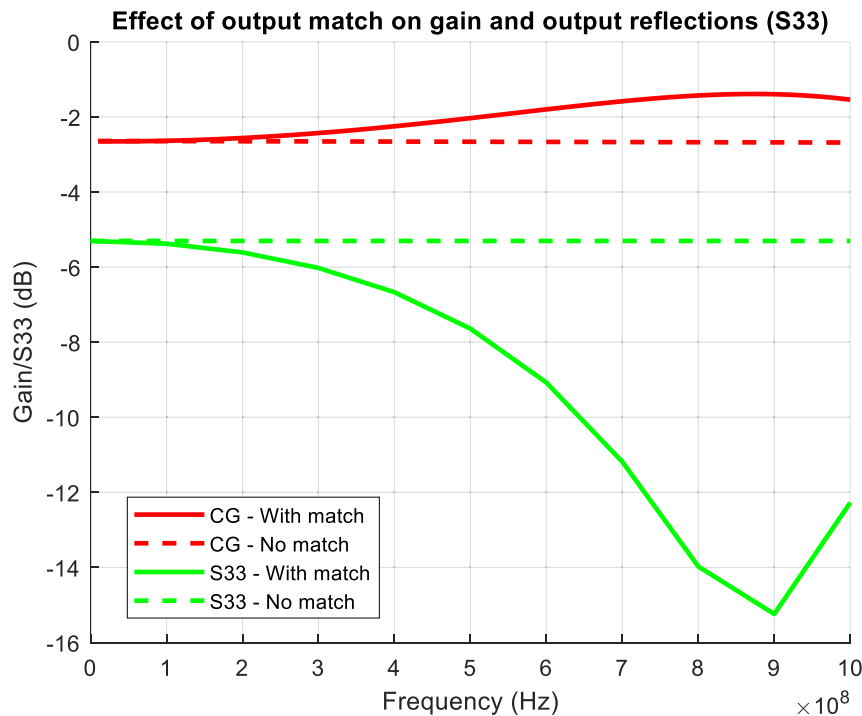**Effect of output match on gain and output reflections (S33)**

Figure 20. Conversion gain and S33 with no output match (Dashed line) and with output match (Full line).

An output matching network was designed by inspecting the output impedance on the smith chart. For proper matching, a series inductor was added for the real impedance part, and a parallel capacitor to AC ground for cancelling out the imaginary parts. For successful matching near the higher end of the IF frequency spectrum, 25 nanohenry inductors would be required. At this point in the work, no viable solution was found to coherently implement this into the physical layout without sacrificing too much surface area, so it will be left only as a simulation.

While this solution doesn't work for downconversion mixers, it is a popular choice for upconversion mixers, where the RF frequency (output) is a large frequency. On the contrary, the upconversion mixers have a much harder time with the input matching described in 2.4.4, due to the low frequency input demanding large inductors [25].

2.4.6   RF and LO baluns

A balun designed by Mr. Hietanen was adopted in this work to connect external signal to differential mixer input. With the input matching with source degeneration being designed with this balun intact, a low reflectance input interface was achieved for the RF port.
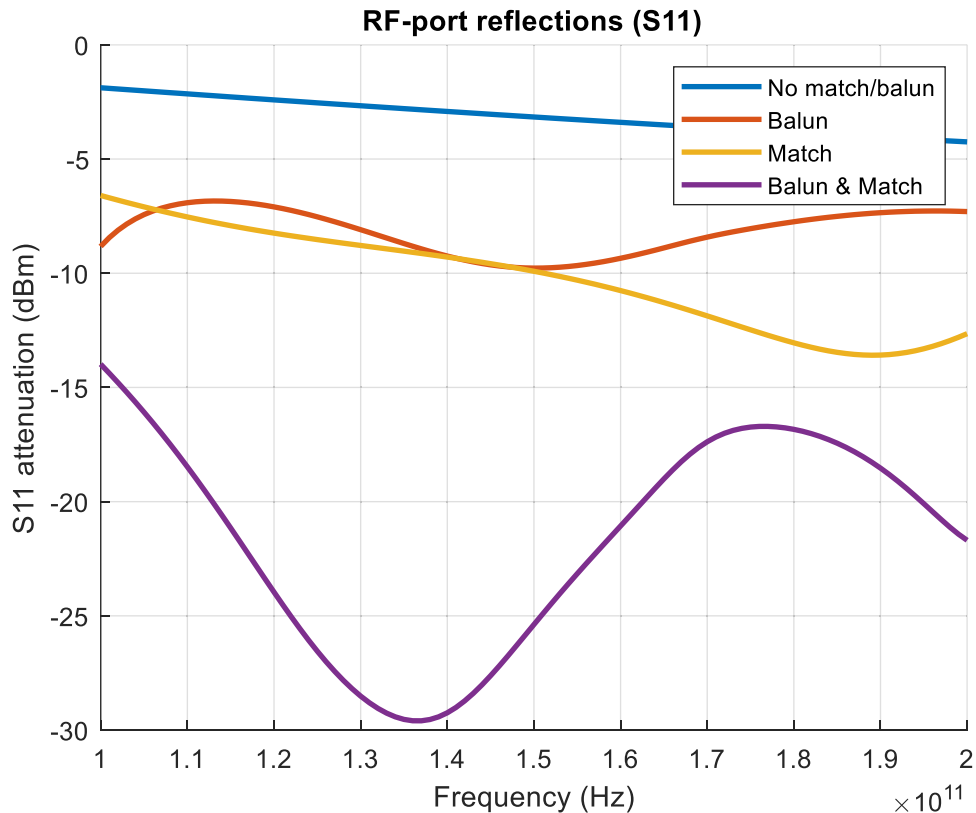

Figure 21. Effect of various setups at the RF input on its self-reflections (S11)

As it was noted in 2.4.3, the LO power is bears a lot of weight when it comes to the most important parameters of the mixer, hence it's important that we also maximize the power transfer efficiency of the LO port. The LO port had its input interface designed using the baluns provided by the GlobalFoundries process design kit (PDK). By first adjusting the imaginary part of the LO port to zero using a gate inductor, the balun could be picked exactly to match the source impedance of the LO port and the impedance seen at the gate. Additional parallel capacitors were added to resonate at the wanted frequency, which resulted in improvements in the reflections at the LO port.
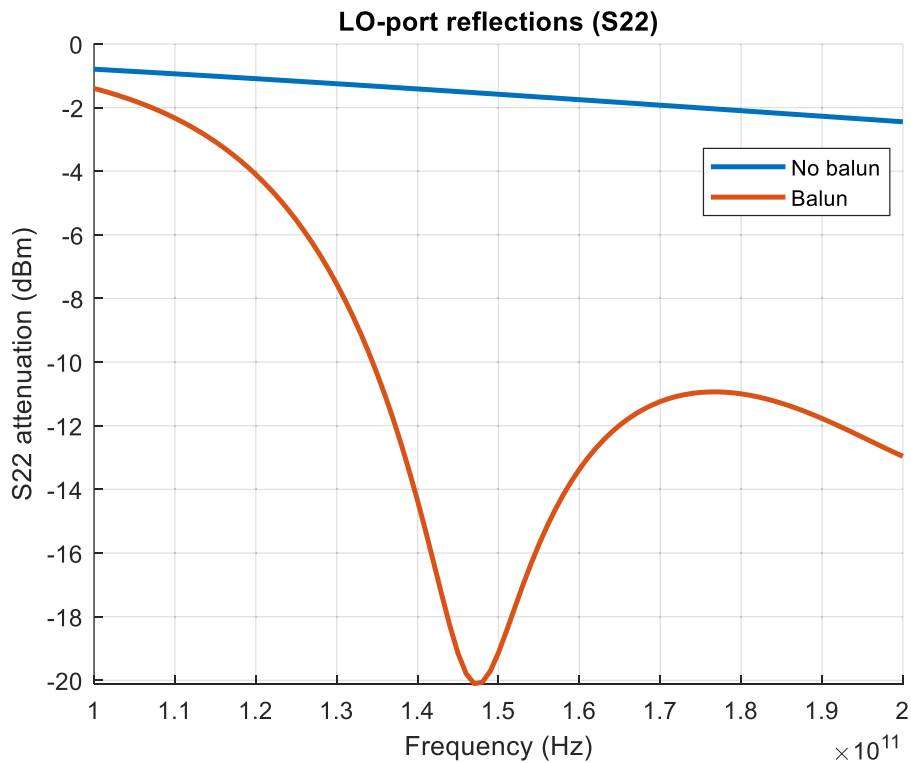
Figure 22. LO port reflections with and without a balun and tuning capacitors.

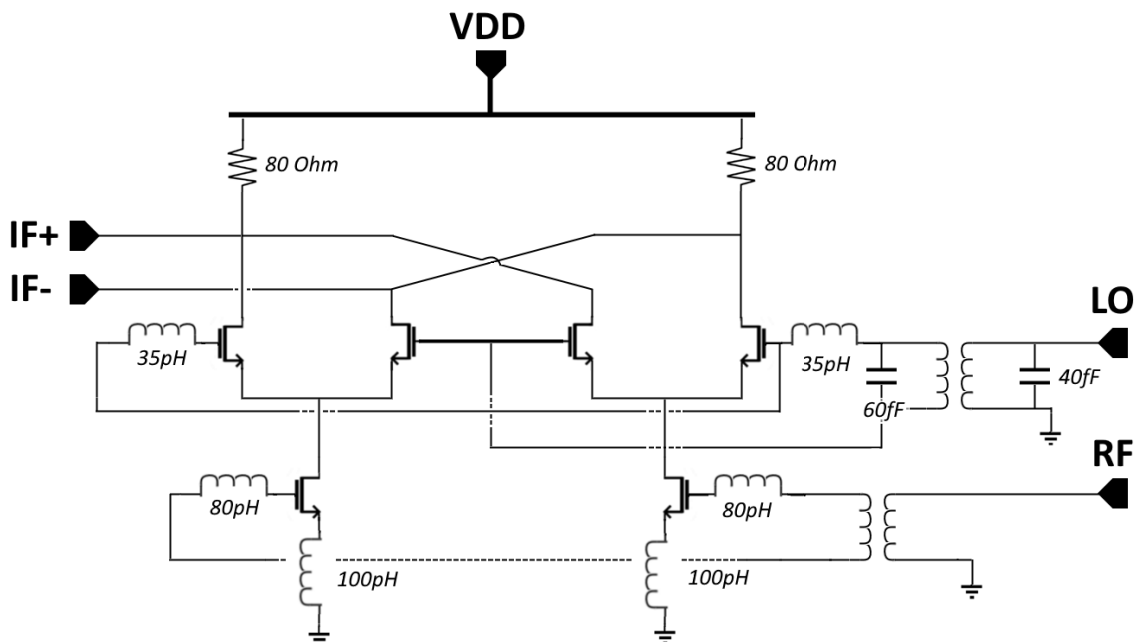### *2.4.7 Summary of circuit design*

Figure 23. Schematic of the final design of the Gilbert cell mixer.

Figure 23 depicts the finalized schematic of the design. The RF port should be driven with a DC-bias of 500 mV, and the LO port with a bias voltage of 550 mV. Table 1 shows how the schematic level simulation parameters changed with the different additions to the circuit. The configurations always include the ones preceding it, meaning the bottom most configuration includes all the ones mentioned above.

Configuration	Conversion Gain*	Linearity (P1dB)*	Noise Figure	LO-to-IF leak	S11 (RF input)	S22 (LO input)	S33 (IF output)
Blank mixer	-9.69 dB	-4.54 dBm	15.3 dB	-296 dBm	-3.21 dB	-1.60 dB	-5.30 dB
+RF Balun	-4.20 dB	-4.69 dBm	15.1 dB	-286 dBm	-9.78 dB	-1.60 dB	-5.30 dB
+Input match	-2.70 dB	1.54 dBm	12.0 dB	-290 dBm	-24.9 dB	-1.60 dB	-5.30 dB
+LO Balun	-2.64 dB	1.80 dBm	11.7 dB	-52.5 dBm	-25.4 dB	-18.4 dB	-5.30 dB
+Output match	-1.54 dB	1.96 dBm	11.7 dB	-52.54 dB	-25.4 dB	-18.4 dB	-12.1 dB

*At 1 GHz

Table 1. Simulated parameters with different additions to the schematic

3   SIMULATION METHOD FOR LAYOUT

The simulation of the layout is always preceded by a three-step procedure.

The first check for testing a layout was to confirm that it complies with the design rule checks set by the PDK creator. If the layout does not comply with the rules, then any following tests are for a design that cannot be created. Common checks include mistakes such as metal layers being unrealistically slim or being close to making contact.

The second test is the LVS (Layout vs Schematic) simulation, where the layout structure created is compared to the schematic it was built based upon. Not only is this check vital for the sensibility of the design, but especially a new designer will easily make mistakes by shorting nets unintentionally. A common way of making this mistake is by pulling a net to a higher layer using a via, while cutting through another net in the process.

The third step involves creating a non-ideal circuit of the design to detect any layout issues. This is done in two ways depending on the components we are working with:

1. For very tightly bundled active component layers like the transistor core, a parasitic extraction is employed. A PEX (Parasitic Extraction) model is created of the layout, which contains the parasitic resistances and capacitances of the traces. Common problems that are being detected are parasitic capacitances causing leakage, due to traces being near one and another. Parasitic resistances are especially considered at the transistors gate, as it was mentioned how the R_g of the transistor plays a large role in its performance.
2. For higher layer metals, an EMX (Electromagnetic Extraction) model is created, which largely changes the behaviour of the device as unideal electromagnetic behaviour is considered. Traces get inductive properties and gain resistance, causing unexpected behaviour to occur, if the designer is not careful.

The initial simulations of the layout are done using the same methods that are described in 2.3. With the addition of the parasitic resistances and capacitances, we can identify changes in the mixer behaviour based on our layout design decisions.

4   PHYSICAL LAYOUT

Arguably the most important phase in designing an integrated circuit is the physical layout of the device. On a schematic level, everything works ideally no matter how wires are routed or overlap each other. With physical traces, everything has an effect on the overall performance on the mixer. So extreme care is required to preserve the integrity of the original design. Physical traces have inductivity and impedance, meaning symmetry is a requirement for fast signals where phase difference must be kept at a minimum. The mismatches are grouped into extrinsic and intrinsic factors, the latter being induced by the transistor devices itself, and the former circuit traces and devices around the transistor device [26]. Studies have also been conducted on the effects on mismatches specifically on mixers, seeing how a mismatch in the LO or RF ports deteriorate the IF amplitude [27]. Outside of mixers, there are many studies that focus on the methods for cancelling process gradients (errors) caused by layout variability [28][29]

4.1   Devices used in the layout.

Both active and passive devices used in the design were provided by the GlobalFoundries millimetre-wave component library. Such devices were tested to be appropriate for a D-band device prior to being used. The exception to this was made in the RF balun stage, where a pre-designed and measured balun with pads on the highest metal layer was provided.

4.1.1   Active devices

With transistors in the millimetre-wave region, it is crucial that the device can operate in the high frequency region. With devices becoming increasingly smaller and smaller, there appears a contradiction when you begin to consider that as the ports of the device decrease in size, the metal trace width will also have to decrease, increasing the effective parasitic resistance in the port. This is the most prevalent in the gate of the transistor, where the effect of R_g has been studied and found to have a detrimental effect on the f_{Max} and f_t of the device. Both parameters eventually causing a cut-off in the device performance [30]. This work briefly compared the effects of R_g at the core level, but several works have done more detailed investigations of not only R_g, but also parasitic capacitances spawning from poor layout decisions [31][32]. Besides that, some of the most common mistakes can come from poor grounding, as they can easily limit the operation of high-frequency devices [33]. The transistor model used are the six-terminal *nfet_rf_6t* models provided by GlobalFoundries. The transistor uses multiple metal layers for the gate, allowing for lower R_g connection and also a back gate biasing option, which can be used to lower the threshold voltage of the transistor if needed.

4.1.2   Passive devices

Passive devices in this case include the inductors, capacitors and resistors of the device. With the exception of the library inductor components, these reside in the lower metal layers.

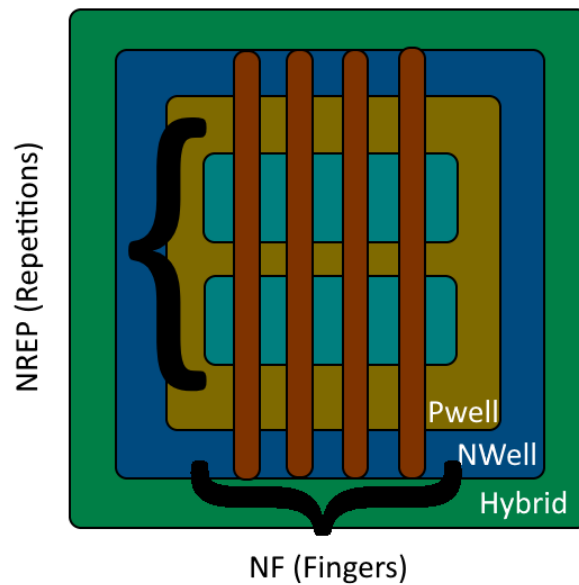#### *4.1.2.1   Capacitors and inductors*

Figure 24. Capacitor model used in the layout.

The capacitors used are structured with an array of polysilicon fingers nf and cells $nrep$ over an N-well layer. Adjusting these values and the width/length of the capacitor, a wide range of useful capacitance values in the millimetre wave region can be achieved. The n-well connection is used as a back gate connection to control the capacitance tuning region.
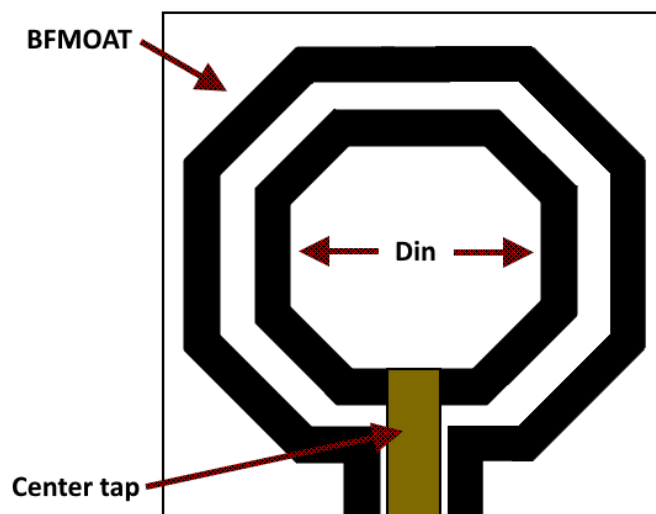

Figure 25. The inductor used in the layout.

The octagonal library inductors are used to provide inductance for the matching networks. While drawing metal strips by hand is also a way of controlling the inductance in the system, the library components come with pre-measured Q-values. Minimizing Q-values is important, as at the end of the day it minimizes resistive losses. The BFMOAT layer is used in further defining the substrate layer.

Two different types of baluns were used in the design. The RF side uses a premade balun provided by the designer of the previous block (LNA) of the receiver chain. The balun has been tested to provide wideband attenuation in D-band, making it suitable for any device in this frequency range.
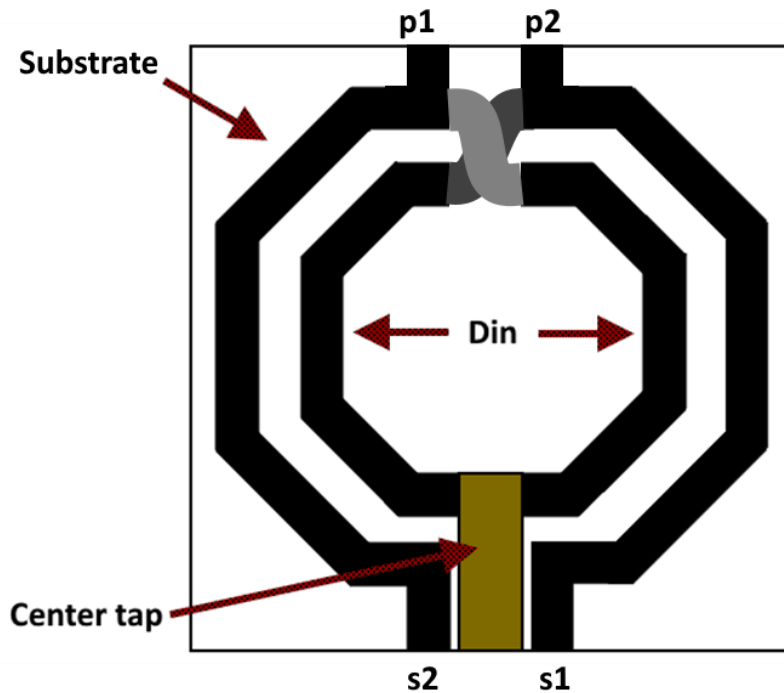

Figure 26. Balun device used in the LO port.

The balun used on the LO side is provided by the GlobalFoundries PDK. The stacked balun allows for multiple turn ratios, while also optionally providing a center tap. The center tap is useful in our case, since it can be used to provide the input with AC-isolated DC-voltage. When the appropriate balun has been selected for the wanted frequency, tuning capacitors can be added to resonate for an even greater benefit to the S-parameters of the LO port.

Figure 27. Various 3D models of provided components.

4.1.2.2   Resistors

The resistors used in the work are unsilicided N+ diffusion resistors provided by GlobalFoundries.
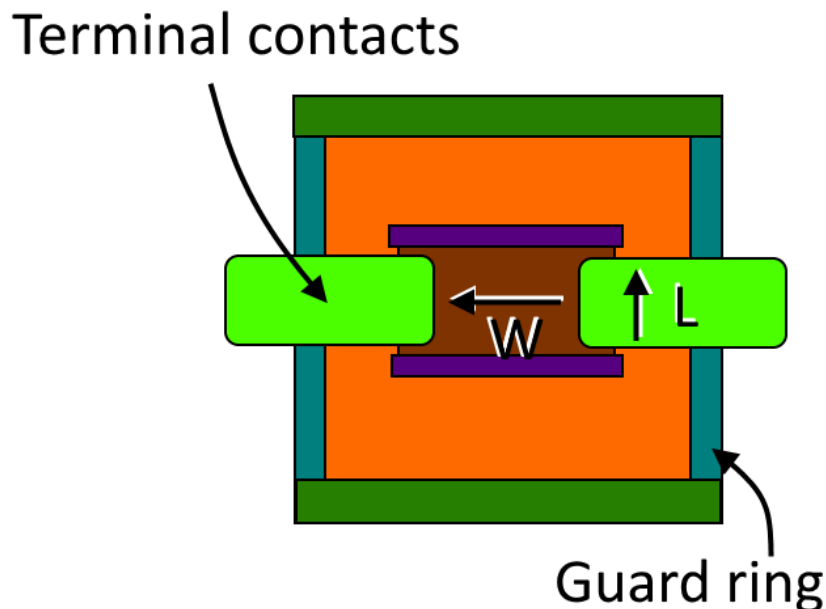

Figure 28. Resistor model used in the design.

The width and the lengths of the physical contacts and the widths between them could be increased, increasing, or decreasing the resistance. The length of the whole device can be extended all the way to 50 micrometres, at which point the highest resistance values are acquired.

4.2   Mixer core layout

The innermost layer of the design, the transistor core was the first block to be designed. The transistor core included the two MOS devices used for amplifying and the four MOS devices used in the switching quad. A short comparison study between layout strategies was made, where in the first layout the transistor gates were connected together through a wide gate connection, and in the second iteration the gates were connected symmetrically. The symmetrical connection sacrificed some of the gate width, since a connection could not be made through the center of both transistors due to different polarities.
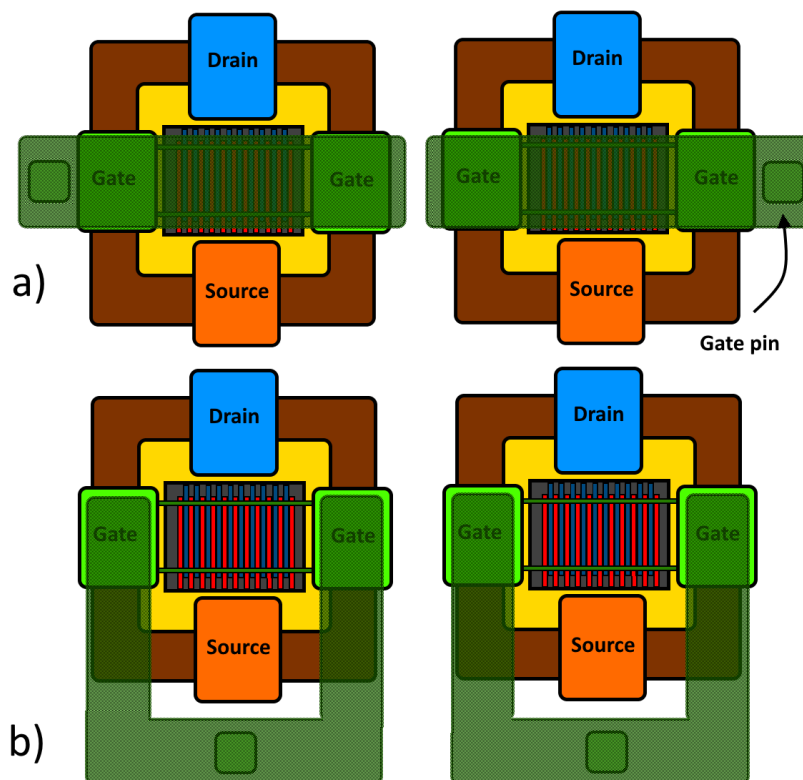
Figure 29. Transistor gate layouts for a) Maximum gate width and b) symmetry

After evaluating both options with simulations, it was found that the differences in the gm-stage were marginal whether symmetry or gate width were used, changes in conversion gain being less than a tenth of a decibel. At this stage in the design, asymmetry on such a small level also didn't affect the LO stage by a lot, achieving $0.12dB$ higher gain in the wider gate width setup. Regardless, the wider gate width could only be utilized for the gm-stage, since asymmetric design created complications with further extending the LO metal traces in a rational manner. The effects of asymmetry are marginal at this stage, since the trace lengths are so short they don't have a large impact on the signals yet.

4.3   Higher metal layers

The higher metal layer (Or the EMX simulated layer) contains the inductors, baluns and the connections to the top-level aluminium pads in the physical layout. In the layout, the RF signal is inserted from the bottom and the LO signal from the left side. The IF signals are then received from the top of the layout. Figures 31 and 34 better illustrate this division between different layers of the design and how they are simulated.

Although great care was taken in not creating unnecessary coupling between lines, it was found in initial results that using the component values shown in Figure 23 resulted in the mixer to be resonating at the incorrect frequency. The gate inductor was changed to a much larger 300pH inductor, and the source inductor to a smaller one.
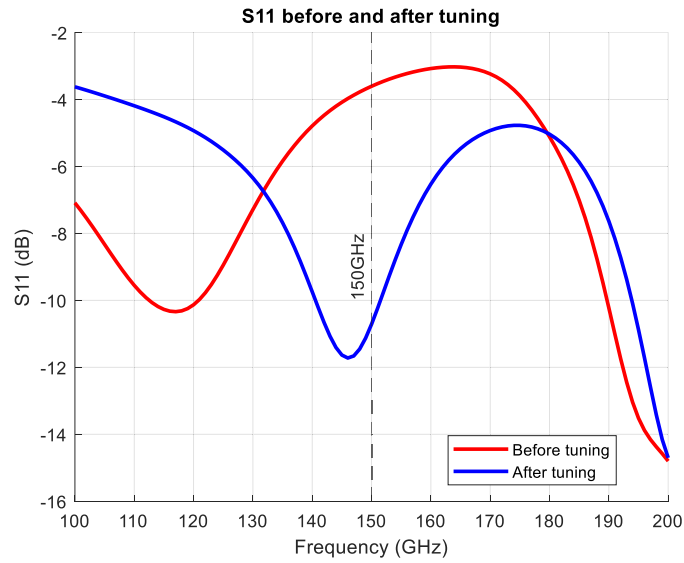
Figure 30. S11 (RF port) tuning after higher layer layout simulations

The reflections in the S1 (RF port) were reduced by close to 7 dB. A significant improvement, but as results will show later, there are more variables that affect the final results than just the s-parameters.

4.4   Post layout simulation results

In the post layout simulation results the results will be split in the following way as described in earlier chapters:

- **Schematic:** The results only contain ideal schematic results

- **PEX Core:** The results contain the parasitic extracted transistor core, consisting of the lower-level metals (M1-2 and C1-5 metals). This namely includes the transistor core, but also capacitors and resistors that reside in the lower metal layers as well.

- **EMX Layer:** The results contain electromagnetic extracted models of the higher-level metals, used for the baluns, inductors, and top-level connections to pads.
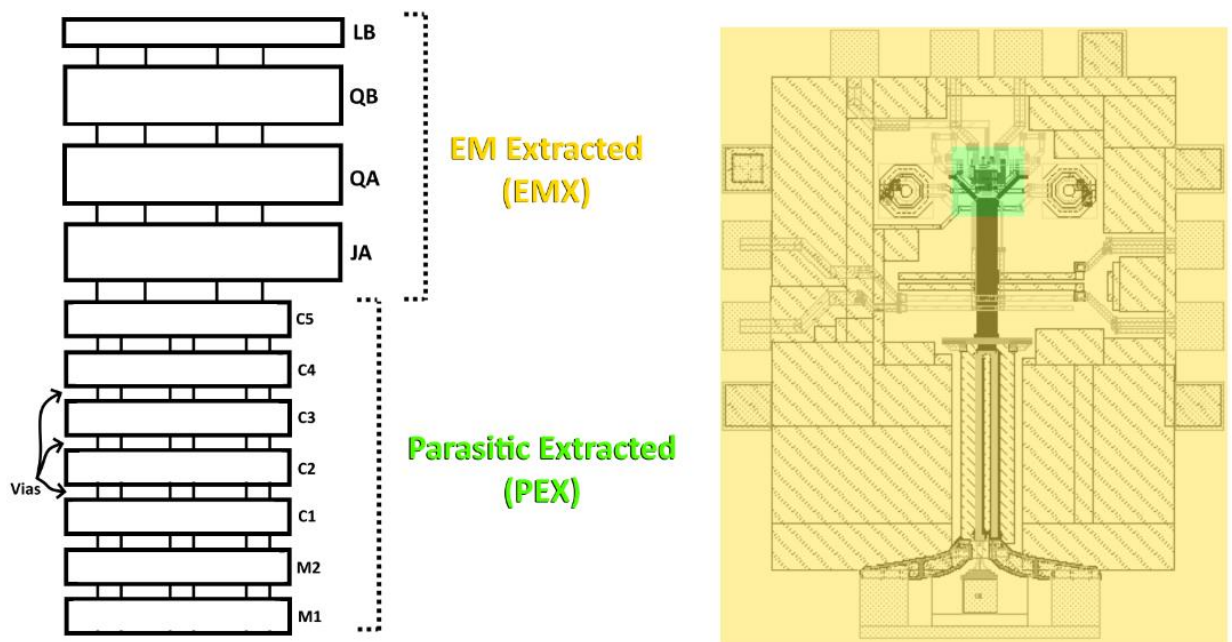

Figure 31. Metal stack and the extracted regions in the layout (PEX and EMX)

4.4.1   First iteration of layout

From a tuning and port return loss standpoint the mixer appeared to be working fine, as illustrated by Figure 32. The changes in the s-parameters were studied as hierarchies were added into the design, starting from the bottom with the transistor core, all the way up to the top-level metals. Between the two are the baluns and the inductors used for matching.
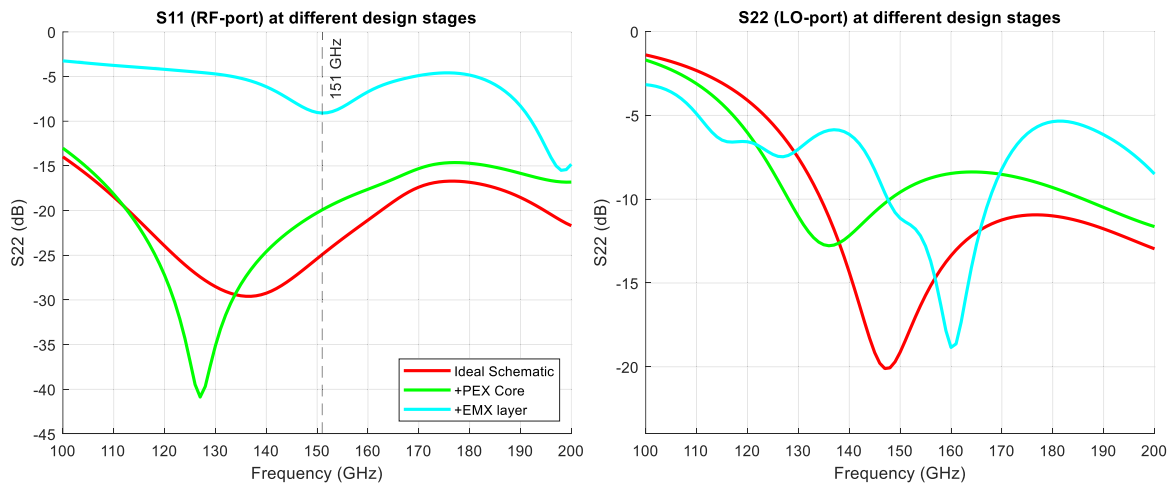
Figure 32. S11 (RF port) and S22 (LO port) at various design stages.

While the performance does heavily drop at the higher frequencies especially in the case of the RF port, the tuning is still very accurate around the 150 GHz region where it should be. In the LO port, even better tuning was seen, as the S22 parameter was tuned to a slightly higher frequency, yet still achieving -12 dB of attenuation. However, the performance parameters simulated clearly showed that the tuning was not the only parameter that mattered for performance.

Configuration	Conversion Gain*	Linearity (P1dB)*	Noise Figure	LO-to-IF leak	S11 (RF input)	S22 (LO input)	S33 (IF output)
Schematic	-2.64 dB	1.80 dBm	11.7 dB	-52.5 dBm	-25.4 dB	-18.5 dB	-5.30 dB
+PEX Core	-3.85 dB	1.78 dBm	12.3 dB	-43.8 dBm	-20.2 dB	-9.40 dB	-5.31 dB
+EMX Layer	-30.2 dB	4.84 dBm	30.8 dB	-26.0 dBm	-9.02 dB	-11.6 dB	-2.25 dB

Table 2. Initial post layout results

The reason for the lack of performance was found after transient waveforms were investigated. There were three main shortcomings that were revealed which caused a detrimental decrease in performance when the EMX was taken into account.
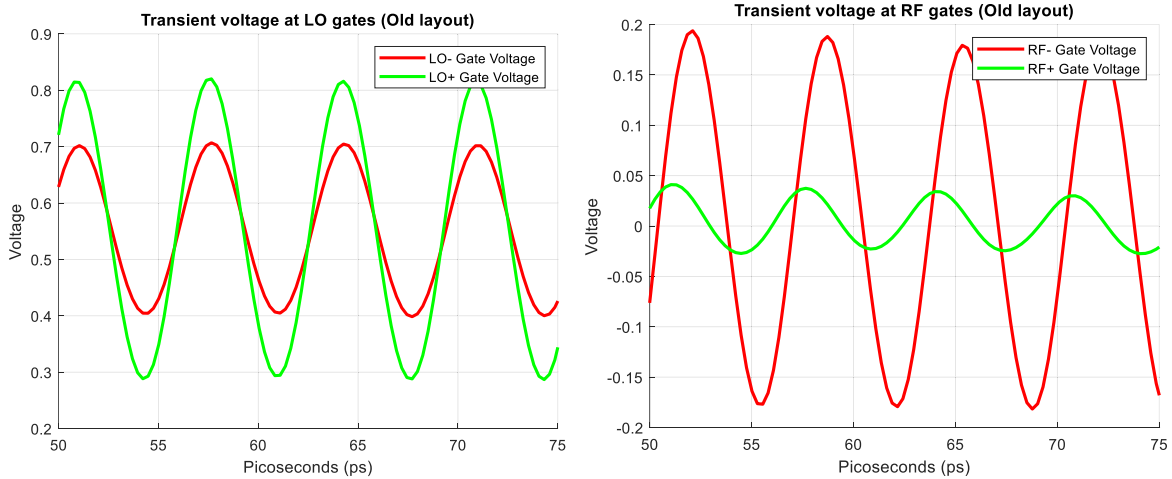
Figure 33. (Left) Gate voltages at the LO transistors (Right) Gate voltages at the RF transistors.

As per discussed in 2.2.3.1, the waveform losses and phase errors can be a great cause of decreases in performance. Figure 33 illustrates one of the worst possible scenarios for the LO signal, where the LO signals in the first iteration were in the same phase, meaning that the switching function efficiency was halved. The RF waveform shows similar phase errors but to a lesser extent. More noticeably, the RF signal is completely useless to the mixer due to the fact that it is not getting the DC-bias voltage which it should be getting. In 2.4.2 it was shown that for maximum gain, 500mV of DC-bias was needed, and that lower values quickly reduced the gain in the system.

Not shown in the transients but can be seen in the results is also a common mistake in layout design where a designer trying to work fast accidentally leaves a small gap in the metals. In this first iteration, such mistake was made in one of the IF transmit lines, visible in Table 2 as the S33 parameter dropping by approximately 3dB, signifying that only half of the IF voltage was getting out of the system.
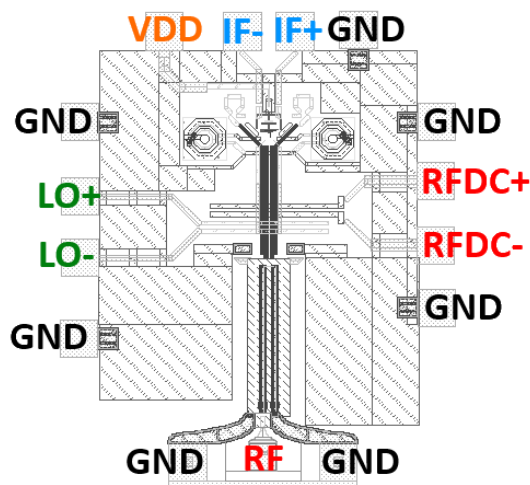
4.4.2   Second iteration of design

Figure 34. Second layout iteration with labelled pads.

In the layout iteration the following changes were made: The LO balun was removed, as it was causing significant drops in the LO power and also problems with phase errors. Secondly, the symmetry between the LO and RF branches were improved using dummy traces. The symmetry in RF and LO metals was made difficult due to the placement of pins in the core of the transistors.
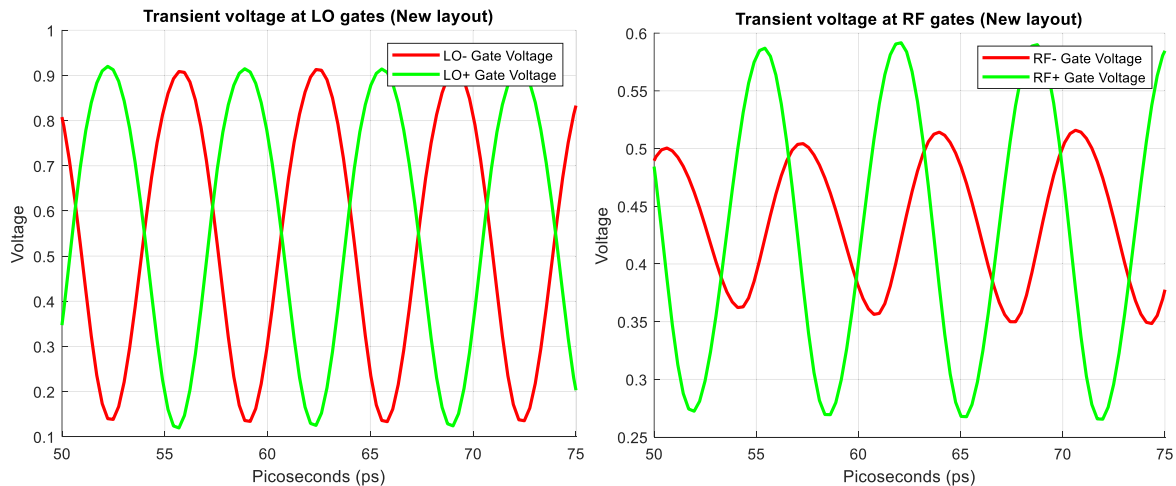
Figure 35. (Left) Gate voltages at the LO transistors (Right) Gate voltages at the RF transistors (Second iteration).

The clearest improvement is found in the amplitude improvements, where not only were both of the input signals preserved better, but their errors also reduced. The RF amplitude error dropped to near non-existent and the LO amplitude error to a just 2 dBm. The RF phase-error also decreased from the changes, but unfortunately the same cannot be said for LO, which saw a fairly remarkable increase.

Figure 36. LO and RF port s-parameters in second iteration.

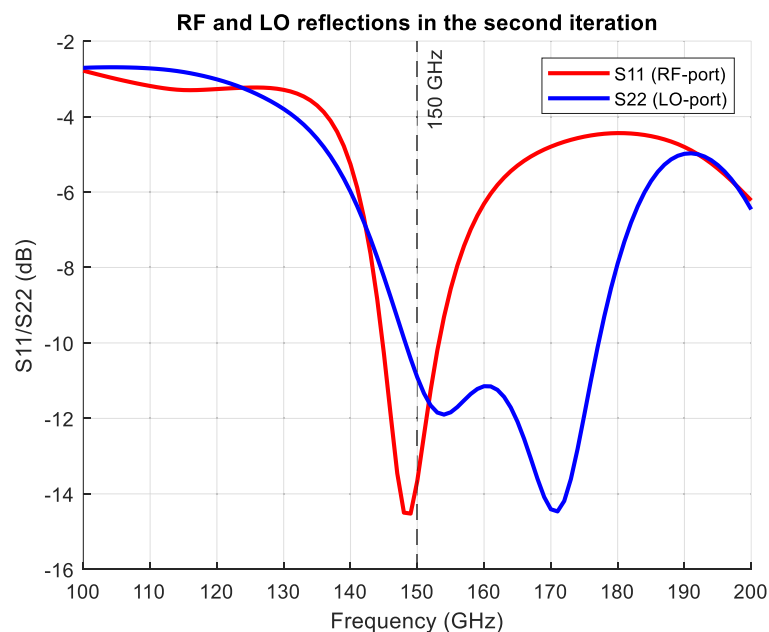The tuning was fixed after the layout was changed to a simplified version. After the layout was adjusted with better techniques, the inductor values could be returned to values closer to the initial ones. The inductors from the process design kit were found to have the best Q-values between 50-200 picohenries, hence the 300 picohenry inductor in the first iteration was in itself causing some losses, even if it provided tuning in the right frequency.

*at 1 GHz

Configuration	Conversion Gain*	Linearity (P1dB)*	Noise Figure	LO-to-IF leak	S11 (RF input)	S22 (LO input)	S33 (IF output)
Schematic	-2.64 dB	1.80 dBm	11.7 dB	-52.5 dBm	-25.4 dB	-18.4 dB	-5.30 dB
+PEX Core	-3.85 dB	1.78 dBm	12.3 dB	-43.8 dBm	-20.2 dB	-9.40 dB	-5.31 dB
+EMX Layer	-10.0 dB (+20.2)	6.40 dBm (+1.56)	17.3 dB (-13.5)	-27.7 dBm (-1.7)	-13.7 dB (-4.7)	-11.3 dB (+0.3)	-5.34 dB (-3.12)

Table 3. Post simulation results at different stages (Second iteration) and its improvements (compared to the first iteration).

The device was found to be consuming only around 4.7mW (Including the LO power draw), hence it cannot generate a large amount of voltage if the output load is low.
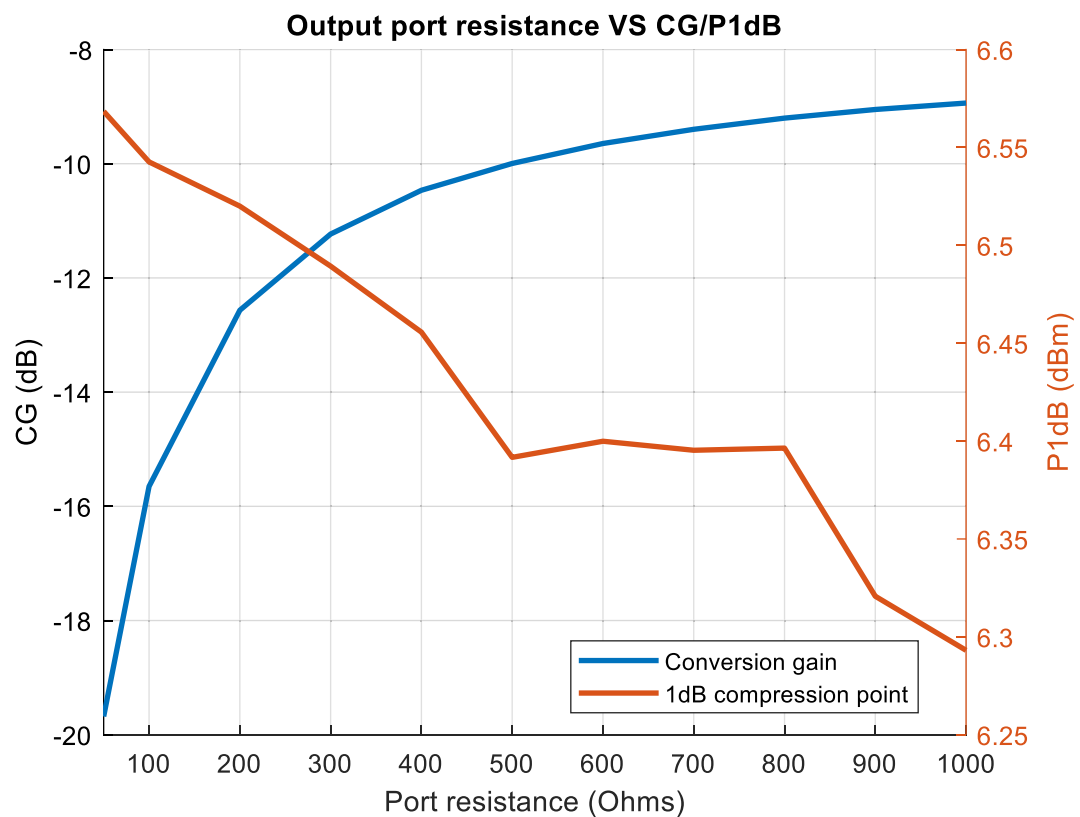

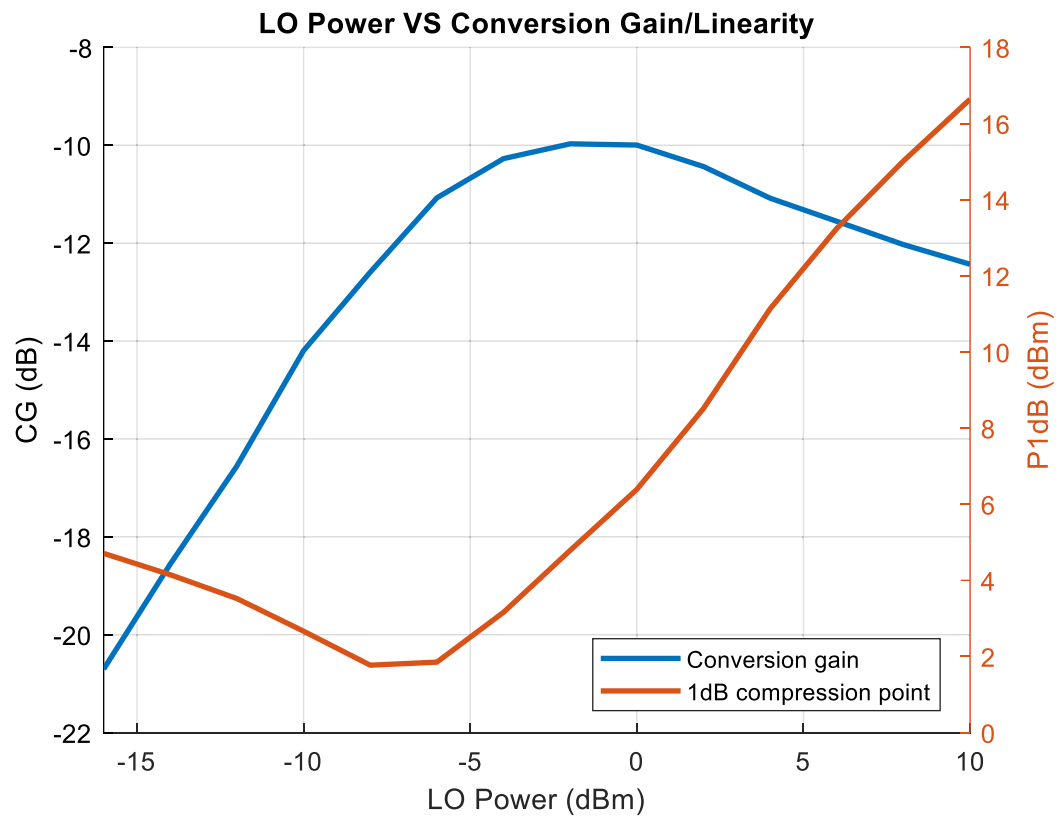Figure 37. Output port resistance VS CG/P1dB

Figure 38. LO power vs CG/P1dB in the second layout iteration.

5   DISCUSSION

The results showed great improvements in the downconversion mixer after improvements to the layout were made. It could be noted that after power delivery to the ports was fixed and the phase errors in the LO ports were eliminated, significant improvements were found. The gain improved by 20.2dB and the linearity by 1.56dB. The noise figure also saw a significant decrease, going down to just 17.3dB from almost 31dB.

Room for improvement can be found in the transient waveforms, as it can be noted that the RF signals are out of phase and there is an amplitude imbalance between the two. Asymmetries are born from layout mistakes, such as when one of the transmission line passes another transmission line, while the other one does not [34]. The transient waveforms also hint that signal strength is lost before reaching the core of the transistors, which should be investigated before any schematic level changes are made.

The finalized mixer was relatively simple. Many improvements to the Gilbert cell topology have been proposed, which could specifically also be used in this work. The low current in the mixer was mainly a result of other issues inside the mixer, but additional improvements could be made using a current bleeding method[35]. On the contrary, an increased current in the system will also mean higher current in the transistors, which can easily lead to voltage headroom issues with low supply voltages.

Several other methods for low power supply mixers have been proposed as well. One of which uses transformer coupled RF signals, which has been characterized as a method that results in a lower power supply demand for a set linearity performance, at the cost of noise and conversion loss [36][37]. This is potentially a great method if linearity is to be prioritized. Another low power supply method involves using folded LC-tanks to get better resonating at lower supply voltages [38].

With any of the suggestions listed above being implemented, the results could get close to reaching the schematic values. In its current state, the downconversion mixer has very good linearity, but it is expected that when the gain increases, the linearity will drop.

Technology	Input Frequency (GHz)	Conversion Gain (dB)	Linearity P1dB (dBm)	Noise Figure (dB)	Power draw (mW)	LO-power (dBm)	Chip area (mm2)	Supply voltage (V)
This work 22nm	150-151	-10.0	6.4	17.3	4.7	3	0.34	0.8
[6] 22nm	123-146	2	-12.5	8.2	20	-5	N/A	1.6
[7] 45nm	134-149	6	-8.5	8.8	14	3	1.44	1
[39] 28nm	113-127	-11	-17.2	33.5	6	-9	0.11	1

Table 4. Comparison table with other works of the same nature.

6   CONCLUSION

A downconversion mixer was designed and simulated. The initial layout made based on the schematics performed poorly due to layout, being only able to perform with a conversion gain of -30.2 dB, 1dB compression point of 4.8 dBm and a noise figure of 30.8 dB. After adjustments were made to the physical layout to fix lossy components and improve amplitude and phase errors due to asymmetry in a second iteration, the conversion gain improved to - 10.0 dB. The linearity measured with 1dB compression point was 6.4 dBm. The noise figure dropped down to 17.3 dB.

It can be said that the research goals were partially filled, since the mixers main issues were identified and as a result, a downconversion mixer with reasonable performance parameters was created. Unfortunately the ambitious gain and linearity goals of 0 dB could not be achieved this time, but suggestions were made that could move the results to that direction. The work does not end here, since its optimization will continue after the thesis, ultimately leading to a tape-out in the future. Likely layout improvements will be prioritized over circuit design tricks in order to achieve decent performance before the finalized version is brought into the physical world from a tape out. Simulations against process and temperature variables will also be conducted.

7   REFERENCES

[1] A. A. Farid, A. Simsek, A. S. H. Ahmed and M. J. W. Rodwell, "A Broadband Direct Conversion Transmitter/Receiver at D-band Using CMOS 22nm FDSOI," 2019 IEEE Radio Frequency Integrated Circuits Symposium (RFIC), Boston, MA, USA, 2019, pp. 135-138.

[2] C. Wang and G. Rebeiz, "A 2-Channel 136-156 GHz Dual Down-Conversion I/Q Receiver with 30 dB Gain and 9.5 dB NF Using CMOS 22nm FDSOI," 2021 IEEE Radio Frequency Integrated Circuits Symposium (RFIC), Atlanta, GA, USA, 2021, pp. 219-222.

[3] Terrovitis M.T. & Meyer R.G. (1999) Noise in current-commutating CMOS mixers. IEEE Journal of Solid-State Circuits 34, pp. 772–783.

[4] Adan A.O., Naka T., Kagisawa A. & Shimizu H. (1998) SOI as a mainstream IC technology. In: 1998 IEEE International SOI Conference Proceedings (Cat No.98CH36199), pp. 9–12.

[5] W. Roh et al., "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," in IEEE Communications Magazine, vol. 52, no. 2, pp. 106-113, February 2014.

[6] A. A. Farid, A. Simsek, A. S. H. Ahmed and M. J. W. Rodwell, "A Broadband Direct Conversion Transmitter/Receiver at D-band Using CMOS 22nm FDSOI." 2019 IEEE Radio Frequency Integrated Circuits Symposium (RFIC), Boston, MA, USA, 2019, pp. 135-138.

[7] S. Li and G. M. Rebeiz, "A 134-149 GHz IF Beamforming Phased-Array Receiver Channel with 6.4-7.5 dB NF Using CMOS 45nm RFSOI," 2020 IEEE Radio Frequency Integrated Circuits Symposium (RFIC), Los Angeles, CA, USA, 2020, pp. 103-106,

[8] Y. Yang, S. Zihir, H. Lin, O. Inac, W. Shin and G. M. Rebeiz, "A 155 GHz 20 Gbit/s QPSK transceiver in 45nm CMOS," 2014 IEEE Radio Frequency Integrated Circuits Symposium, Tampa, FL, USA, 2014, pp. 365-368.

[9] H. Nam, C. J. Lee, D. Kim, S. -K. Kim, D. Y. Lee and I. Na, "A D-Band High-Linearity Down-Conversion Mixer for 6G Wireless Communications," in IEEE Microwave and Wireless Technology Letters, vol. 33, no. 5, pp. 579-582, May 2023.

[10] B. Gilbert, "A new wide-band amplifier technique," in IEEE Journal of Solid-State Circuits, vol. 3, no. 4, pp. 353-365, Dec. 1968.

[11] David M. Pozar (2012) Microwave Engineering (4th Edition), Wiley 2011.

[12] H. T. Friis, "Noise Figure of Radio Receivers," Proc. IRE, vol. 32, pp. 419–422, July 1944.

[13] Razavi B. (2011) RF Microelectronics (2nd Edition) (Prentice Hall Communications Engineering and Emerging Technologies Series). Prentice Hall Press, USA, 2nd edition.

[14] Voinigescu S. (2013) High Frequency Integrated Circuits (The Cambridge RF and Microwave Engineering Series). Cambridge University Press, UK.

[15] Keng Leong Fong and R. G. Meyer, "Monolithic RF active mixer design," in IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, vol. 46, no. 3, pp. 231-239, March 1999.

[16] S. -F. Lu and J. -C. Guo, "5.5 GHz Low Voltage and High Linearity RF CMOS Mixer Design," 2008 European Microwave Integrated Circuit Conference, Amsterdam, Netherlands, 2008, pp. 171-174

[17] A. Ismail and A. Abidi, "A 3-10-GHz low-noise amplifier with wideband LC-ladder matching network." IEEE Journal of solid-state circuits 39.12 (2004): 2269-2277.

[18] P. Qin and Q. Xue, "Compact wideband LNA with gain and input matching bandwidth extensions by transformer." IEEE microwave and wireless components letters 27.7 (2017): 657-659.

[19] M. Vigilante and P. Reynaert, "20.10 A 68.1-to-96.4 GHz variable-gain low-noise amplifier in 28nm CMOS." 2016 IEEE International Solid-State Circuits Conference (ISSCC). IEEE, 2016, pp. 360-362.

[20] S. P. Voinigescu, T. Chalvatzis, K. H. K. Yau, A. Hazneci, etc "SiGe BiCMOS for analog, high-speed digital and millimetre-wave applications beyond 50GHz," IEEE BCTM Digest, pp. 223–230, October 2006.

[21] H. Fouad, K. Sharaf, E. El-Diwany and H. El-Hennaway, "An RF CMOS modified-cascode LNA with inductive source degeneration," Proceedings of the Nineteenth National Radio Science Conference, Alexandria, Egypt, 2002, pp. 450-457.

[22] H. I. Wu, R. Hu and C. F. Jou, "Complementary UWB LNA Design Using Asymmetrical Inductive Source Degeneration," in IEEE Microwave and Wireless Components Letters, vol. 20, no. 7, pp. 402-404, July 2010.

[23] Lee H. & Min B. (2017) Ka-band up-conversion mixer with inductive degeneration. In: 2017 IEEE International Symposium on Radio-Frequency Integration Technology (RFIT), pp. 211–213.

[24] Hao Xu, D. R. Jackson and J. T. Williams, "Comparison of models for the probe inductance for a parallel-plate waveguide and a microstrip patch," in IEEE Transactions on Antennas and Propagation, vol. 53, no. 10, pp. 3229-3235, Oct. 2005.

[25] A. Verma, K. K. O and J. Lin, "A low-power up-conversion CMOS mixer for 22-29-GHz ultra-wideband applications," in IEEE Transactions on Microwave Theory and Techniques, vol. 54, no. 8, pp. 3295-3300, Aug. 2006.

[26] D. Coffing and E. Main, "Effects of offsets on bipolar integrated circuit mixer even-order distortion terms," in IEEE Transactions on Microwave Theory and Techniques, vol. 49, no. 1, pp. 23-30, Jan. 2001.

[27] Jun Wang & Wong A.K.K. (2001) Effects of mismatch on CMOS double-balanced mixers: a theoretical analysis. In: Proceedings 2001 IEEE Hong Kong Electron Devices Meeting, pp. 85–88.

[28] C. C. McAndrew, "Layout Symmetries: Quantification and Application to Cancel Nonlinear Process Gradients," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 36, no. 1, pp. 1-14, Jan. 2017.

[29] A. Hastings, "The Art of Analog Layout", Upper Saddle River, NJ, USA: Prentice-Hall, 2001.

[30] Razavi B., Ran-Hong Yan & Lee K.F. (1994) Impact of distributed gate resistance on the performance of mos devices. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications 41, pp. 750–754.

[31] C. Liang and B. Razavi, "A layout technique for millimeter-wave PA transistors," 2011 IEEE Radio Frequency Integrated Circuits Symposium, Baltimore, MD, USA, 2011, pp. 1-4.

[32] D. Zhao and P. Reynaert, "A 60-GHz Dual-Mode Class AB Power Amplifier in 40-nm CMOS," in IEEE Journal of Solid-State Circuits, vol. 48, no. 10, pp. 2323-2337, Oct. 2013.

[33] U. Gogineni, H. Li, J. A. del Alamo, S. L. Sweeney, J. Wang and B. Jagannathan, "Effect of Substrate Contact Shape and Placement on RF Characteristics of 45 nm Low Power CMOS Devices," in IEEE Journal of Solid-State Circuits, vol. 45, no. 5, pp. 998-1006.

[34]   K. Voudouris and J. M. Noras, "Effects of amplitude, phase, and frequency imperfections on the performance of a direct conversion receiver (DCR) for personal communication systems," in IEEE Microwave and Guided Wave Letters, vol. 3, no. 9, pp. 313-315, Sept. 1993.

[35] L. A. NacEachern and T. Manku, "A charge-injection method for Gilbert cell biasing," Conference Proceedings. IEEE Canadian Conference on Electrical and Computer Engineering (Cat. No.98TH8341), Waterloo, ON, Canada, 1998, pp. 365-368 vol.1.

[36] J. R. Long and M. A. Copeland, "A 1.9 GHz low-voltage silicon bipolar receiver front-end for wireless personal communications systems," in IEEE Journal of Solid-State Circuits, vol. 30, no. 12, pp. 1438-1448.

[37] S. T. Nicolson et al., "A Low-Voltage SiGe BiCMOS 77-GHz Automotive Radar Chipset," in IEEE Transactions on Microwave Theory and Techniques, vol. 56, no. 5, pp. 1092-1104, May 2008.

[38] E. Abou-Allam, J. J. Nisbet and M. C. Maliepaard, "Low-voltage 1.9-GHz front-end receiver in 0.5-/spl mu/m CMOS technology," in IEEE Journal of Solid-State Circuits, vol. 36, no. 10, pp. 1434-1443, Oct. 2001.

[39] C. J. Lee and C. S. Park, "A D-Band Gain-Boosted Current Bleeding Down-Conversion Mixer in 65 nm CMOS for Chip-to-Chip Communication," in IEEE Microwave and Wireless Components Letters, vol. 26, no. 2, pp. 143-145, Feb. 2016.