Multi-band/Multi-Mode RF Front-end Receiver for Basestation Applications

Şerban Motoroiu

Supervisors:

Dr. Ing. Leo de Vreede
TU Delft

Ir. Paul Mattheijssen
NXP Semiconductors

Delft University of Technology

Department of Electrical Engineering

The undersigned hereby certify that they have read and recommend to the Faculty of Electrical Engineering, Mathematics and Computer Science for acceptance a thesis entitled “Multi-Band/Multi-Mode RF Front-End Receiver for Basestation Applications” by Serban Motoroiu, in partial fulfillment of the requirements for the degree of Master of Science.
Date: 30.08.2011

Chairman:

Prof. Dr. John R. Long

Advisors:

Dr. Ing. Leo C.N. de Vreede

Ir. Paul Mattheijssen

Committee Members:

Prof. Olexander Yarovyi

Mark van der Heijden

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Leo de Vreede, without whose constant help and patience I would not have been able to bring this thesis to completion in a successful way. His ideas and knowledge have not just shaped my work, but also had a contribution to my development as an engineer. My heartfelt thanks go also to my supervisor from NXP, Paul Mattheijssen, for offering me this project that turned out to be much more interesting and captivating than I had initially foreseen. I hope I did not disappoint them and that they both feel rewarded by the outcome of my work.

Abstract

In the last couple of years, mobile telephony has truly become global, with more than 4 billion users worldwide nowadays (compared to less than 1bn ten years ago). This has represented a boost not just for the mobile phones industry, but also for that of wireless infrastructure. Larger-scale production and increased competition have transformed the manufacturing costs into an important issue for the producers of basestation equipment.

The current state-of-the-art front-end circuits for basestations have extremely high dynamic range, but are narrowband and predominantly implemented in GaAs. One way to minimize costs is to use wideband low-noise amplifiers and mixers that cover all the mobile standards. Another way is to implement the aforementioned circuits in less expensive processes, such as SiGe. And finally, significant cost reduction can also be achieved if some of the constituent blocks (such as low noise amplifiers and mixers) are integrated on the same chip.

The two criteria that describe the dynamic range of a basestation receiver are the (output) 3rd order intercept point, (O)IP_3, and the noise figure (NF). To achieve the demanding linearity specifications (OIP_3 above 30dBm), special linearization techniques (such as out-of-band harmonic cancellation) are employed in the design of the first stage, a low noise amplifier. Since the linearity of the following stages is even more important for the overall OIP_3, a current-commuting NMOS mixer followed by a trans-impedance amplifier has been selected on account of their excellent linearity.

Our down-converter achieves an output 3rd order intercept point higher than 37dBm over a bandwidth from 0.7GHz-3.5GHz (performance unmatched by any of the front-ends available on the market), and has a gain of 18.5dB and a noise figure of 3dB at 2GHz.

Table of Contents

Chapter 1

Introduction

Although members of the same family, the receivers for mobile handsets and those for base stations are rather different when it comes to requirements (linearity, noise figure, power consumption) and topologies. For example, if one looks at a datasheet of a low noise amplifier for base station receivers, one will be struck by the huge linearity figures: 44dBm of third order output intercept point is not uncommon a number at all [1]. And as if this is not incredible enough, the noise figure of such amplifiers must be lower than 1dB (the amplifier presented in [1] has a noise figure of 0.62dB, but figures as low as 0.37dB can be found [2]). Whereas for mobile phones, the 3rd order output intercept points of the front-end LNAs rarely exceed 5dBm, while their noise figures are in the vicinity of 3dB. These demands can be understood if we think of cellular base stations as hubs in mobile phone networks, towards which a myriad of signals are directed. On the one hand they must be capable of rejecting close-by interferers usually of higher level than the desired signals – hence the requirement of excellent third order linearity – and on the other hand they are supposed to sense even the minutest signals emitted by far away users – hence the necessity of very low noise figures.

Following the same line of reasoning, the large interferers – be they nearby or far away – a basestation receiver may experience can saturate its front-end, and therefore lead to the malfunction of the entire system. This has a direct impact on the gain compression figures of the front-end circuits: for example, the LNA in [1] has a P_{1dB} of 27.5dBm.

As for the differences in topologies, it is enough to think about the physical dimensions of a cellular phone and those of a base station. A complete receiver for mobile phones implemented in CMOS technology takes around 1mm^2 of chip area, whereas in the case of base stations a such receiver chain stretches over a couple of meters – from the top of the tower down to the ground level where most of the equipment is housed. This situation is depicted in figure 1.1 [3]. The tower-top amplifier is mounted as close as possible to the antenna to minimize any losses that can lead to noise figure degradation. This amplifier also 'prepares' the signal to travel along a cable whose length is somewhere between a couple of meters and tens of meters. The ground level assembly has a structure similar to the well known mobile handsets receivers.

Contrary to the figures mentioned above, the designers of base station receivers don't have as tough a job as we might have imagined. Most of the products available on the market are implemented in extravagant, and therefore expensive, technologies like for example GaAs, and have current consumptions in the order of tens of mA[1].

[1] The low noise amplifier presented in [1] is implemented in a GaAs p-HEMT technology and has a current consumption of 280mA at 5V supply voltage. The amplifier in [2] draws a lower bias current of 54mA but has an OIP_3 8dB lower than the previous one.

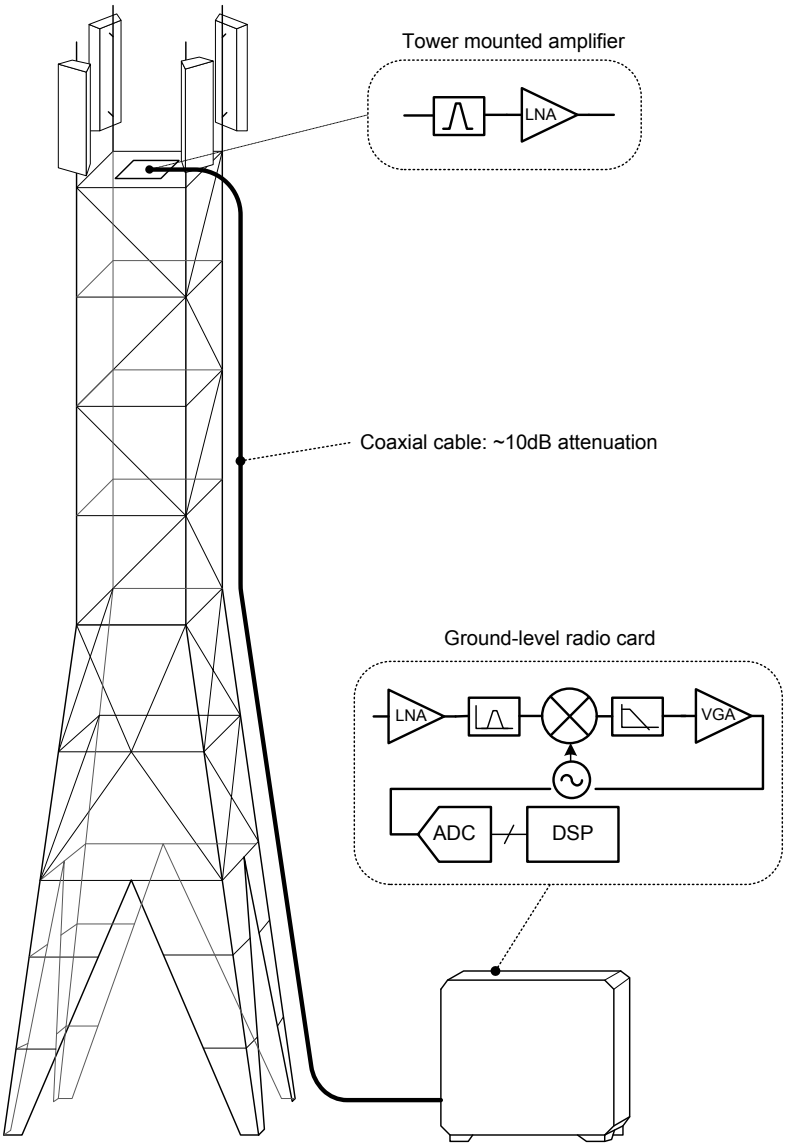


Fig 1.1: Typical structure of a base station

1.1 Research Goals

The goal of this project is to design a front-end circuit (LNA and mixer) that would get as close as possible to the figures mentioned above using a much less expensive SiGe BiCMOS process. There are nonetheless physical limitations particularly with regard to the minimum noise figure a SiGe bipolar transistor is capable of achieving. That is why this project will mainly focus on the RF blocks used on the ground level board – low noise amplifier, mixer, LO interface but excluding here the local oscillator – whose noise requirements are a bit relaxed thanks to the anterior tower-mounted LNA:

To sum everything up, our main objectives are:

- to offer a systematic design approach of an integrated wideband receiver for basestations;
- to investigate if NXP's Qubic SiGe technology offers the possibility of achieving the goals set by a series of requirements specific to BST applications;

- to find circuit design techniques that could enhance the linearity of LNA and mixer and offer broadband operation.

1.2 Outline

The first step in our work consists of a system level analysis of different receivers whose main purpose is to clearly understand how the requirements set on the constituent blocks change when the RF circuits are integrated onto the same chip. We also need to judge if the design specifications are realistic in terms of what is physically feasible with the given technology. Having done this, we can finally choose the best topology for an integrated product, identify the design trade-offs and decide if some of the given requirements can be relaxed, or on the contrary surpassed.

Once the topology has been chosen and the requirements for each block decided, chapter three will describe the design of the first placed block, the low noise amplifier. A small signal analysis will derive the conditions for impedance and noise matching. The next section deals with the large-signal behavior and linearization techniques.

Chapter 4 encompasses everything related to the down-conversion mixer: the mixer itself, the output buffer and the LO interface.

Chapter 5 deals with the practical implementation, including that of the passive elements. The full system simulations presented here will finally validate the decisions made in the previous chapters.

Chapter 2

System Level Analysis

For optimal results, the design of circuits that are supposed to be part of a receiver should be preceded by a thorough system level analysis. This gives the designer the possibility to understand not just the behavior of each circuit in its natural context but also the role of additional blocks – like for instance filters. The interaction that exists between different circuits of a receiver gives rise to limitations that are difficult to recognize and understand by simply looking at the transistor level schematic of each circuit.

We can say that there are two types of parameters: local and global. The former ones can be designed locally because they remain the same when moving form circuit level to system level design. The input reflection coefficient, S_{11}, is one such parameter: if we design a general purpose LNA with a certain S_{11} (assuming good isolation S_{12}) without taking into consideration the entire system, we can be sure that once placed into a receiver, S_{11} will not significantly change (i.e. the entire receiver will have the same input reflection coefficient as the LNA, independently of the reflection coefficients of the other blocks that follow). The global parameters on the other hand are influenced by the behavior of the receiver as a whole. The noise and linearity figures are examples of such specifications. There is no use in – for example – targeting an extremely high third-order output intercept point for the first stage if the next stage has poor linearity performance, since typically the last stage will determine the global linearity figure. Therefore a top level perspective can offer the possibility to relax (or tighten) some of the requirements of each block but also to propose and explore different receiver topologies that befit best a specific circumstance.

2.1 A Generic Basestation Receiver

We will start this system level analysis by looking at the receiver illustrated in the previous chapter, whose electrical scheme has been redrawn here – see figure 2.1 [3]. As previously mentioned in section 1.1, this project focuses *mainly on the ground level RF blocks* – namely the second low noise amplifier and the mixer, but excluding the local oscillator. A succinct explanation of the front-end blocks is given in the coming lines.

Cavity (or Duplex) Filter

This block follows the duplexer and prevents the strong signals generated by the transmitter from leaking into the receiver. Such unwanted signals can lead to the saturation of the receiver or undergo

down conversion onto the desired signal. A duplex filter usually has a pass-band attenuation of approximately 0.5dB and at least 40dB stop-band attenuation, 20MHz away from carrier [4].

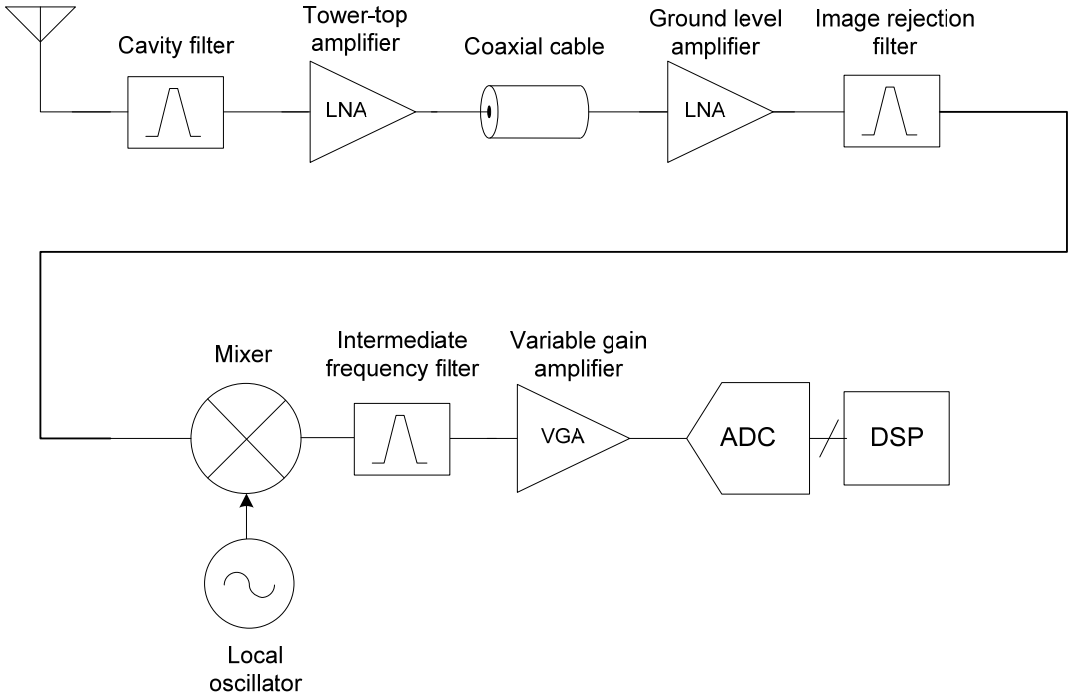


Fig 2.1: A basestation receiver

Tower-Top Amplifier

A low noise amplifier is normally the first active block that follows the antenna. An amplifying stage at the beginning of the chain ensures the noise arising from subsequent (noisy) stages does not degrade the signal-to-noise ratio. Typically, a two or three-stage amplifier can provide enough gain so that the signal level remains decent after passing through a long cable down to the ground level unit.

As an example, the specifications of one of Avago Technologies' low noise amplifiers for the GSM900 standard are listed below [5][1].

Table 2.1: List of specifications for the top-mounted LNA

Gain	38dB
Bandwidth	400MHz – 1.5GHz
Noise Figure	0.51dB
Output IP_3	41.4dBm
$P_{1dB,out}$	23.3dBm
Current Consumption	177mA

Coaxial Cable

As shown in figure 1.1, it connects the tower top unit with the rest of the receiver. With lengths greater than 20m, we expect attenuation as high as 10dB. Some relevant electrical figures are listed in table 2.2 [6].

[1] This is a two-stage low noise amplifier, implemented in a GaAs pHEMT 0.25um technology

Table 2.2: List of specifications for the coaxial cable

DC Resistance (Ω/100m)	
Inner Conductor	1.6
Outer Conductor	1.9
Attenuation (dB/100m)	
@ 0.96GHz	6.84dB
@1.8GHz	9.61dB

Bottom Level Amplifier

The attenuation caused by the connecting cable requires that the signal level be restored by a second low noise amplifier; in addition to this, most of the mixers for basestations are passive and have losses in the order of 8dB. This mixer is preceded by a passive filter that also introduces losses. A gain stage is therefore a must. The noise figures are more relaxed for this amplifier - see table 2.3 [7]

Table 2.3: List of specifications for the radio card LNA

Gain	18dB
Bandwidth	50MHz – 5GHz
Noise Figure @1.9GHz	1.7dB
Output IP_3 @1.9GHz	37dBm
$P_{1dB,out}$	18.2dBm
Current Consumption	52mA

Image Rejection Filter

Signals at $|2f_{LO} - f_{RF}|$ (called image frequency), where f_{LO} is local oscillator's (LO) frequency and f_{RF} is input signal's frequency, can be down-converted onto the useful signal if some measures are not taken. The duplex filter provides some image attenuation but not as much as the GSM standards require. In addition, the noise at the image frequency generated by the LNA can undergo the same process of down-conversion, resulting in a deterioration of the total noise figure. Therefore an IR filter must be placed between the LNA and the mixer. Usually, in basestation applications this is a high quality passive SAW (Surface Acoustic Waves) filter. Some typical numbers taken from a commercial datasheet for WCDMA applications [8] are:

Table 2.4: Some figures of the IIR

Insertion Loss	1.5dB
Pass-band Ripple	2.5dB
Attenuation @ 20MHz offset	28dB
Attenuation @ 200MHz offset	24dB

Mixer

Due to the gain of the preceding LNA, the linearity of this stage is of utmost importance: third order output intercept points in the range of 36dBm are currently considered state-of-the-art. The most representative figures of a passive mixer produced by Analog Devices [9] is are shown in table 2.4.

Table 2.5: Typical specifications for a basestation mixer

Power conversion gain	8.4dB
Input IP3	27dBm
SSB noise figure	9.2dB
Current consumption	190mA

It is fabricated in a BiCMOS process and includes two on-chip passive baluns, for the RF and LO paths. As the last line of this table shows it, the current consumption is not a major concern for this type of applications.

IF Filter

Its role is to eliminate the up-converted products and any parasitic signals coming from the local oscillator. It can be low-pass or band-pass in the case of zero-IF and high-IF receivers respectively.

2.2 Project Specifications

Before embarking on any kind of analyses, our project's requirements must be defined. In order to reduce the costs, a wideband receiver would provide the possibility of using the same circuit for different standards: GSM800, 1800, CDMA and others. Costs minimization can also be achieved by integrating the LNA and mixer onto the same chip, implemented using not GaAs technology but a much cheaper SiGe BiCMOS process. Because at this moment the basestation industry does not have such integrated solutions, the numbers given in the table below (2.6) are defined for stand-alone implementations. As we will see later on, an integrated solution will bring some changes to these requirements.

Table 2.6: Project requirements

	LNA	Mixer	
Bandwidth	-	700 – 4000	MHz
Supply voltage	3.3	5	V
Supply current	60	90	mA
Max input power	-24	-	dB
Noise figure, 1GHz	0.3	-	dB
Noise figure, 3.7GHz	0.6	-	dB
Power gain, 1GHz	18	0	dB
Power gain, 3.7GHz	16	0	dB
OIP_3	45		dBm
IIP_3		32	dBm
Output P_{1dB}	28	9	dBm
Input return loss, S_{11}	-10	-14	dB
Output return loss, S_{22}	-10		dB
Reverse Isolation, S_{12}	-45		dB
LO Input power		0	
LO input return loss		-12	dB

The advantages of an integrated solution over a hybrid implementation can be summarized as follows:

- smaller chip/board area as no pads/pins are required for the LNA output and mixer input;
- no need for LNA output matching and mixer input matching (therefore less design constraints and also wideband matching between the two blocks);
- relaxation of some requirements like, for instance, the output 1dB compression point of the LNA – as will be shown next.

2.4 Choice of Circuits

This project will focus on an architecture made up of LNA and mixer. This corresponds to the front-end shown in the lower part of figure 1.1. Once the idea has been proven, extensions to other circuit blocks are always possible.

As mentioned in the introduction, a front-end receiver for base stations must have a very high dynamic range (DR). On its lower side, the dynamic range is determined by the minimum detectable signal (quantified by the noise figure). In the higher side, the dynamic range is determined by the maximum applicable signal before compression or distortion occurs (spurious free dynamic range).

Having set the requirements on each block, we can calculate the total noise figure, third order input intercept point and 1dB compression point of the receiver. Using Friis's formula for cascaded systems [10], the total noise factor is

$$F_{total} = F_{LNA} + \frac{F_{Mixer} - 1}{G_{LNA}} \cong 0.89\text{dB} \tag{2.1}$$

where F_{LNA} is the LNA noise factor, F_{Mixer} the mixer's double side band noise factor and assuming the mixer has a noise figure $NF_{dsb,mixer}$ of 7dB. The IIP_3 and OIP_3 for the same system can be calculated as [10]:

$$IIP3_{total} = \left[\frac{1}{IIP3_{LNA}} + \frac{G_{LNA}}{IIP3_{Mixer}} \right]^{-1} = 13.8\text{dBm} \tag{2.2}$$

$$OIP3_{total} = IIP3_{total} + G_{total} = 31.8\text{dBm} \tag{2.3}$$

In a similar manner, the output 1dB compression point is expressed as:

$$P_{1dB,total} = \left[\frac{1}{P_{1dB,LNA} G_{Mixer}} + \frac{1}{P_{1dB,Mixer}} \right]^{-1} = 8.95\text{dBm} \tag{2.4}$$

In this case all the linearity figures – $P_{1dB,LNA}$ and $P_{1dB,Mixer}$ – are referred to the output.

These results are summarized in the table 2.7.

As a first well-known conclusion, we can say that the total noise is for the most part determined by the first stage, whereas the second stage sets the linearity figures. Also it must be pointed out that the gain of the first stage plays an equally important role in all the above figures: on the one hand, the higher this gain the lower the impact of the second stage's noise, but on the other hand a high gain in the first stage puts more pressure on the second stage when it comes to linearity. But if there is a trade-off there must be an optimum, and finding this optimum is what the next part is about.

Table 2.7: Calculated values for gain, *NF* and OIP_3 of the receiver (column 4), using the numbers given in columns 2 and 3

	LNA	Mixer	Receiver	
Gain	17[1]	0	17	dB
Noise figure	0.6	7(ssb)[2]	0.89 [dsb]	dB
OIP_3	45	32	32	dBm
Output P_{1dB}	28	9	9	dBm
I_{supply}	60	90	150	mA

Another conclusion drawn by looking at table 2.6 is that the linearity requirement on the LNA can be relaxed from 45dBm down to 32dBm.

[1] An in-between value is used for gain because it was defined at 2 frequencies in table 2.6

[2] As mentioned above, this value was not explicitly given in the list of requirements but was selected as being a typical value found in the industry.

As a final observation, in our calculation the presence of the image rejection filter was not taken into account (i.e. we assumed an ideal filer). In reality, in the case of BST receiver this is an external passive filter that adds losses and requires 50Ω impedance matching. Since our goal is to integrate the LNA and mixer on the same chip, we would like to avoid such external components. The following chapters discuss among others the effect of completely removing this filter or replacing it using different circuit techniques.

2.3 Different Topologies

2.3.1 Current Structure

The front-end shown in figure 2.2 is generally accepted and used in today's basestations for the ground level assembly. It corresponds to the middle part of the complete receiver illustrated in figure 2.1.

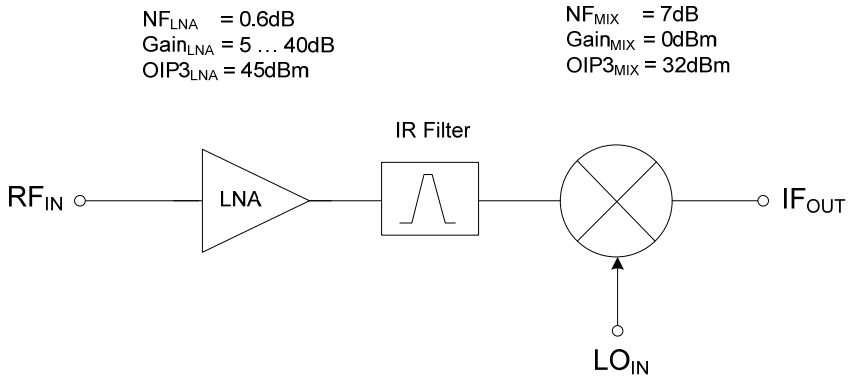


Fig. 2.2: Front-end circuit of the ground level assembly (see figure 2.1). NF_{MIX} is a single-sideband NF

Figure 2.2 also shows some figures used in the simulations whose results are plotted in figure 2.3. Except for those parameters, everything was considered ideal (perfect matching to 50Ω, infinite bandwidth, lossless image rejection filter etc.)

The simulations results confirm the hand calculations: the noise of the mixer becomes insignificant as the LNA's gain exceeds 18dB, while the total output 3rd order intercept point is chiefly limited by the mixer, irrespective of the LNA's linearity[1].

2.3.2 A Step towards Integration: Eliminating the Inter-stage Filter

There is an element in the schematic shown in figure 2.2 that does not get along with our goal of designing a wideband integrated receiver: the image rejection filters used in BST are normally very narrow-band. In addition their high-Q makes them impossible to implement on chip.

Figure 2.4 shows one possible solution: by moving the IR filter in front of the LNA, the wideband character of our receiver is preserved (of course the user will finally use it as a narrow band receiver by placing the aforementioned narrowband filter) without any sacrifice in terms of image rejection.

[1] This statement holds as long as the OIP_3 of the LNA is at least equal to the IIP_3 of the mixer, as we expect to be the case

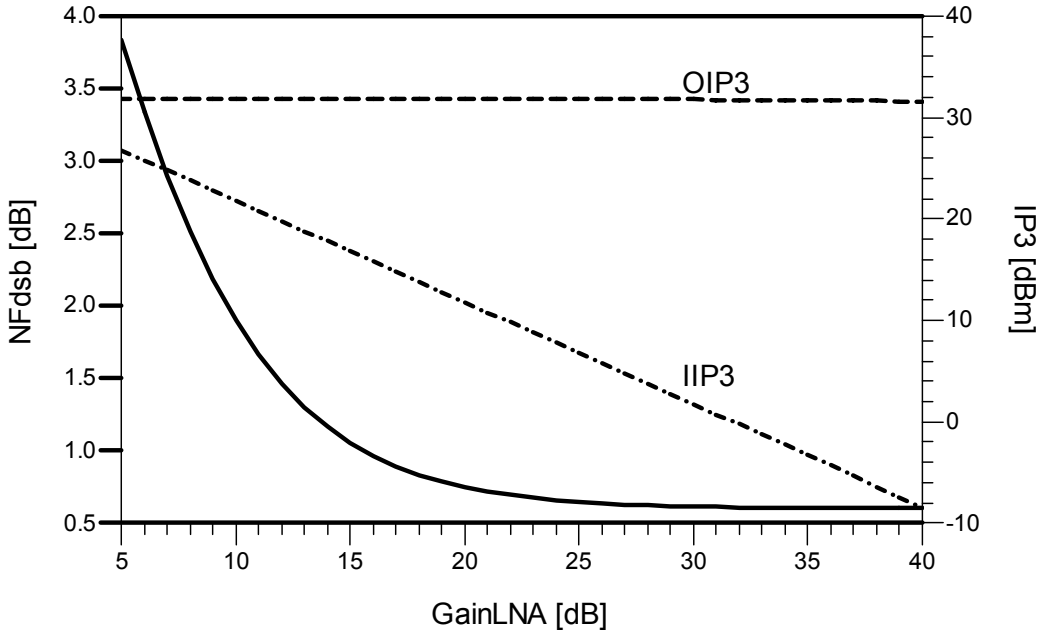


Fig. 2.3. Simulated NF_{total} and IP_3 as function of LNA's Gain

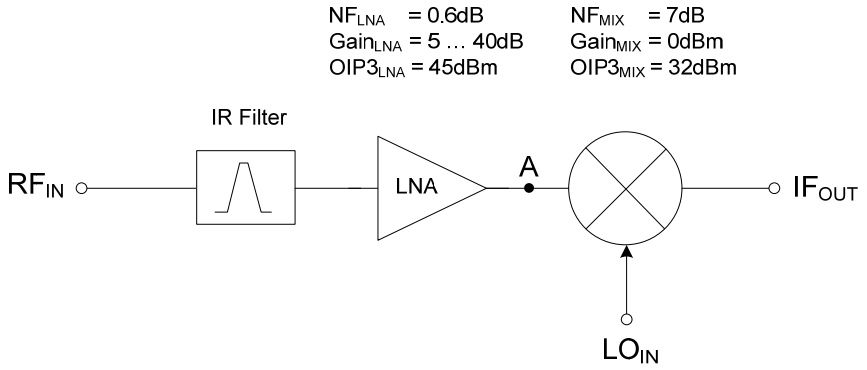


Fig. 2.4: Shifting the IR Filter to the input allow integration of LNA and mixer on the same chip

If the interferers at the image frequency are still rejected by the IR filter, not the same thing can be said about the noise generated by the LNA at that frequency. This situation is illustrated below (figure 2.5).

To better understand the importance of this phenomenon, figure 2.6 proposes a simple schematic used for a few hand calculations. For simplicity we neglect the mixer's noise and consider only the noise of the LNA (therefore $F_{total} \approx F_{LNA}$), which is represented by an input voltage noise source $\overline{v_{n,LNA}^2}$; $\overline{v_{n,s}^2}$ is the noise of an RF signal source of resistance R_S (normally 50Ω).

The RMS noise at the IF output will be obtained by integrating each noise source over its bandwidth:

$$v_{n,outRMS}^2 = \left[v_{n.s}^2(f)\cdot B_1 + v_{n,amp}^2(f)\cdot(B_1+B_2)\right]\cdot G_{LNA}^2 G_{Mixer}^2 \qquad (2.5)$$

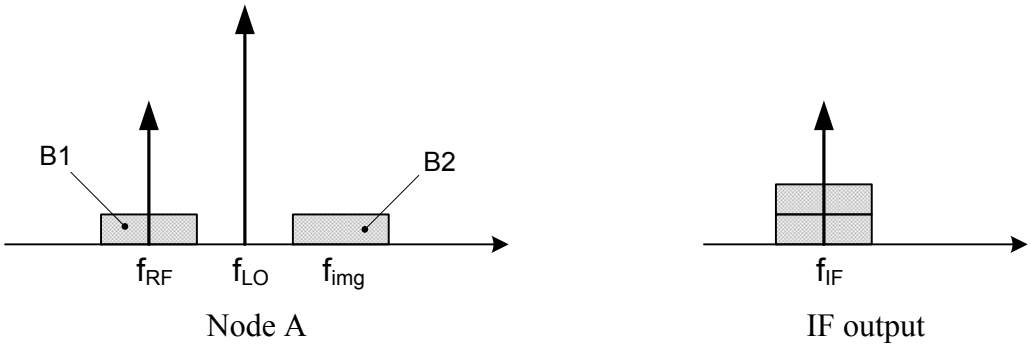


Fig 2.5: The noise at f_{img} is folded by the mixer onto the useful signal

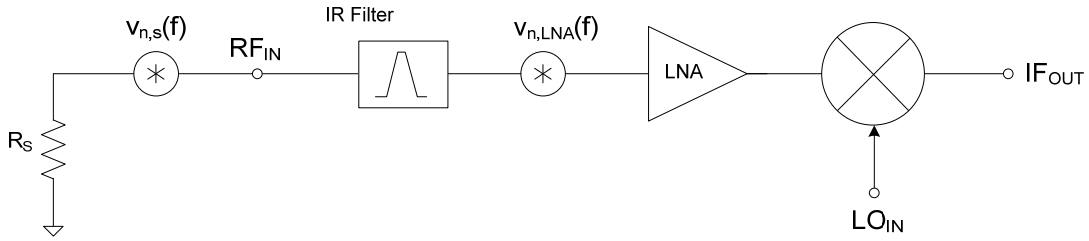


Fig 2.6: Noise sources added to the receiver of figure 2.4

The image rejection filter will remove the noise of the source in the bandwidth B_2; this explains why $v_{n,s}^2(f)$ is multiplied only with B_1. Finally the noise factor can be calculated as follows:

$$F = \frac{SNR_{IN}}{SNR_{OUT}} = \frac{\dfrac{v_{signal,IN}^2}{v_{n,sRMS}^2}}{\dfrac{v_{signal,OUT}^2}{v_{n,outRMS}^2}} = 1 + 2\frac{\overline{v_{n,LNA}^2}}{\overline{v_{n,s}^2}} \qquad (2.6)$$

which is what we more or less expected from the intuitive approach of figure 2.5: the noise of the LNA is doubled, resulting in an irreparable deterioration of the total noise figure. This is confirmed by simulations, as shown in figure 2.7.

The conclusion should now be obvious: even though the receiver is not exposed to the risk of down-converting an unwanted interferer present at the image frequency (which will be taken care of by the input IR filter), an inter-stage filter between the LNA and mixer should exist in order to keep the noise figure to its lowest possible value. The initial solution (figure 2.2) has already proved to be disadvantageous with respect to our requirements that the receiver should be wideband.

2.3.3 IQ Receiver

One solution which has been around for quite some time in the mobile handset receivers relies on complex frequency operation to eliminate the image signals - interferers and noise as well. This is the IQ topology shown in figure 2.8

Normally, the *I* and *Q* paths are continued each by its own A-to-D converter and the summation is done by the digital signal processing unit. In this manner the 90º phase shifter in the signal path and the analog summation block are avoided. For a relevant comparison though, the simulations are performed using ideal elements to represent the aforementioned blocks.

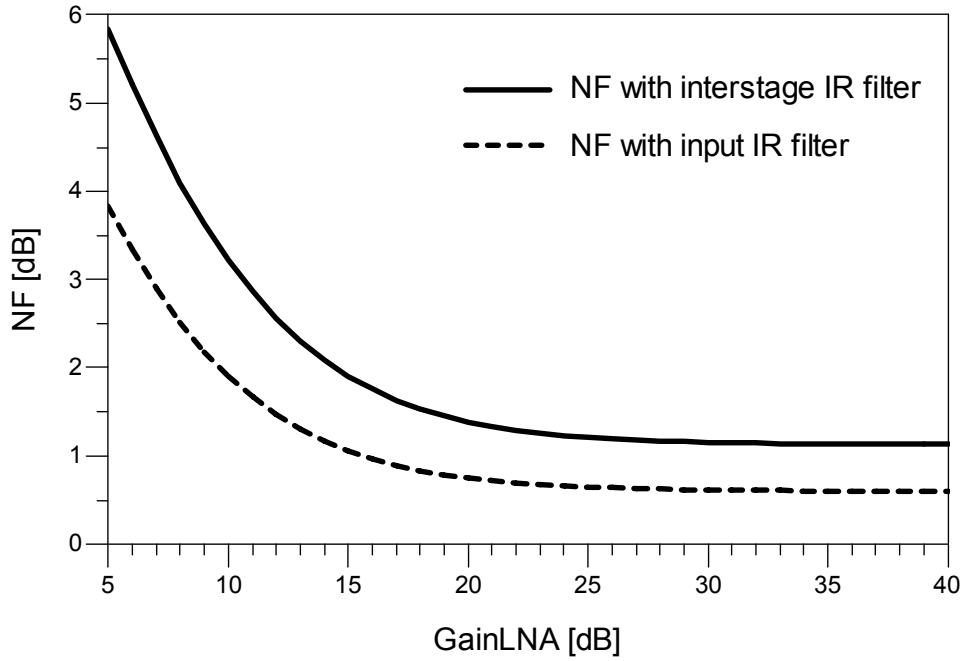


Fig 2.7: The noise figure increases when the filter is shifted in front of the LNA

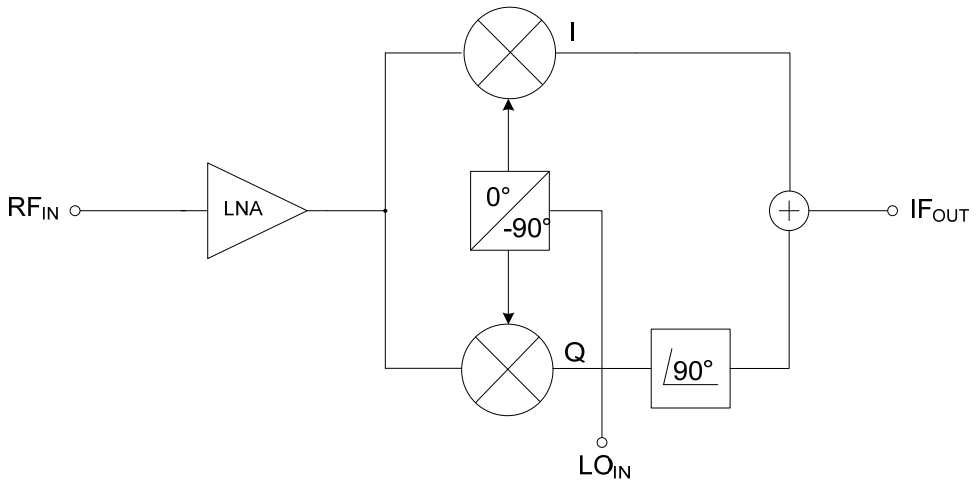


Fig 2.8. Hartley Image Rejection Receiver [11]. For parameter values see figure 2.2 or 2.4

Simulation results for the three structures discussed so far is show in figure 2.9 (compared to the previous situations nothing changed in terms of linearity). The second mixer required by the IQ operation does not introduce so much noise as to require more gain from the first stage: with 20dB of gain, the *NF* figure of the IQ receiver (dotted line) is almost the same as in the case of a single-mixer implementation (dashed line).

Unfortunately, the current basestation receivers do not allow separate I and Q outputs, with the result that the phase shift and summation need be done in the analog domain. One way of doing this is by means of multi-stage (active) polyphase filters. The great disadvantage, as a literature investigation show, is the huge noise figures of such circuits. A completely passive implementation presented in [12] uses 5 stages to achieve a bandwidth of a little less than 20MHz and 60dB. The downside of this circuit is its huge attenuation of 25dB due to losses of the passive stages, resulting in

unacceptable noise figure[1]. An active version presented in the same place overcomes the problems related to gain, but adds great linearity limitations (the circuit shown there has an IIP3 of 15dBm) again hard to accept from a circuit that will be used as a last stage in our receiver. Further improvements with regard to bandwidth are presented in [13] and [14] (70MHz and 100MHz respectively) but suffer from the problems, i.e. poor dynamic range and high noise)

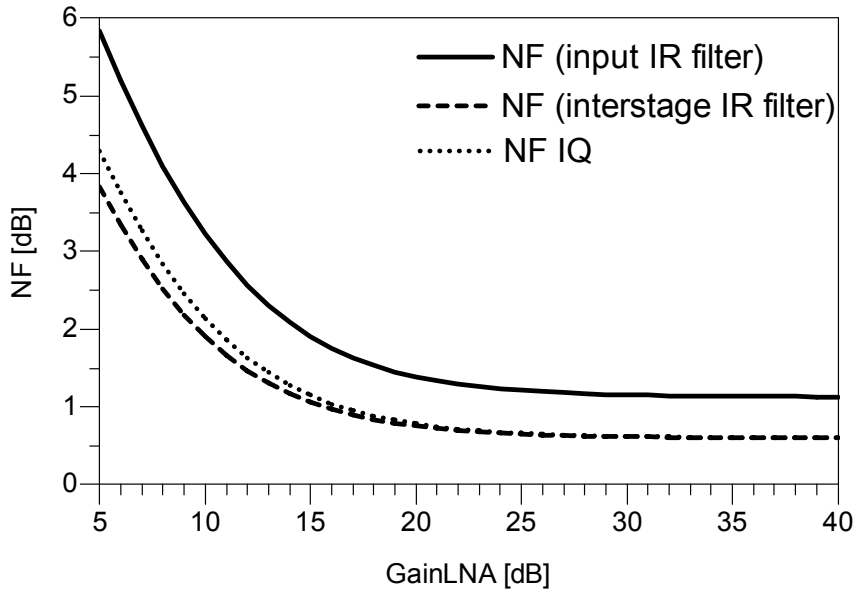


Fig 2.9 Noise Figures for the three cases discussed so far

An alternative – probably the only one given the constraints mentioned above – is to use a Weaver topology. Although this solution offers a more reasonable way to carry out the phase shifting, it too has its own drawbacks as the following discussion will show it.

2.3.4 Weaver [11]

The Weaver topology uses a second pair of IQ mixers to implement the by 90º phase shifting that makes the IF summation possible (figure 2.10).

Figure 2.11 illustrates how the down-conversion takes place: the RF signal at node *A* is down-converted in quadrature to a first intermediate frequency, f_{IF1} (nodes B_I and B_Q), filtered by a band-pass filter (nodes *C*), then down-converted to a final IF to nodes *D*. The signals can now be summed together, as both paths have the same phase (output node IF_{OUT}).

The presence of an interstage bandpass filter centered on f_{IF1} is justified in the same figure: the system is supposed to be immune to the image signals with respect to the first pair of mixers (continuous gray arrow in fig. 2.11) but not to signals at the image frequency of the second pair of mixers (called second image, represented by dashed gray line). This band-pass filter is the only element in receiver that can take care of this unwanted component. The attenuation of the second image depends solely on the quality of this filter, and as we know high-Q filters are impossible to implement in silicon without sacrificing great chip area. In addition, this topology has the disadvantage of requiring a second LO signal, which is incompatible with the current receiver structures.

[1] In a polyphase filter, the next stage is always noisier than the previous stage, due to the fact that each stage that is added must have a pole at a lower frequency than the existent stages. This is usually done by choosing large valued resistors, hence the increase as we add stages.

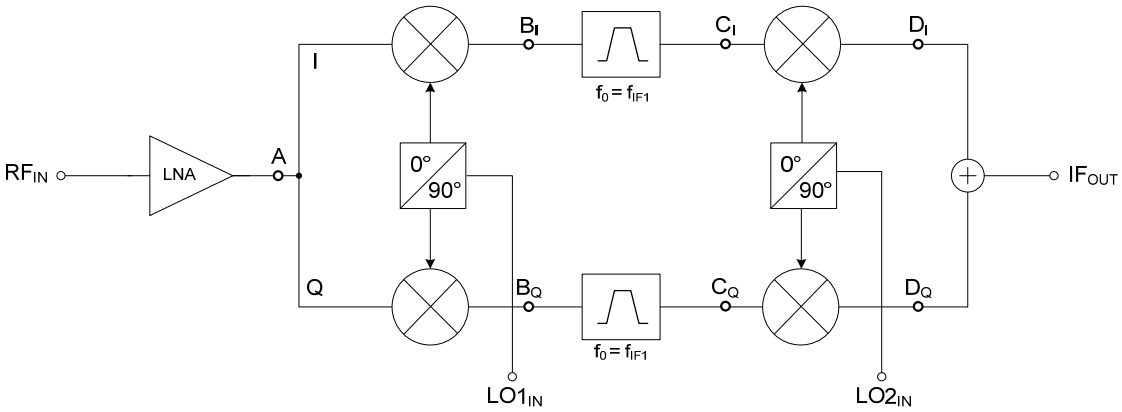


Fig 2.10 Weaver Receiver

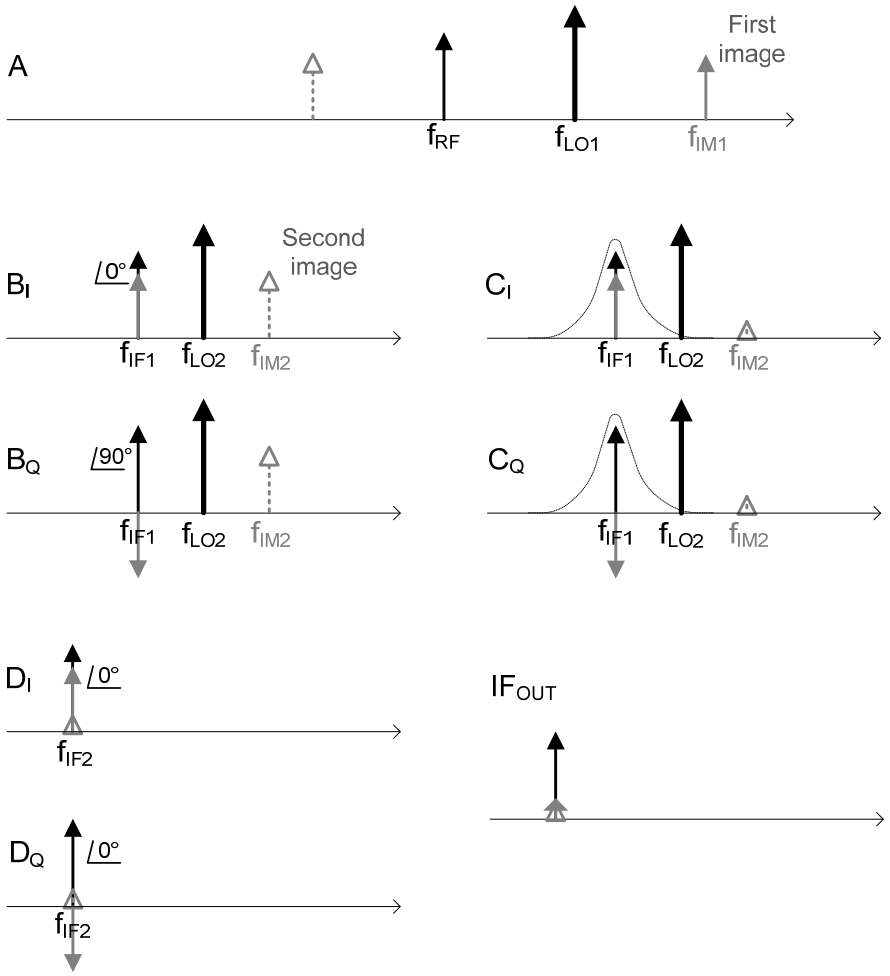


Fig 2.11: Graphical analysis of Weaver architecture

Disadvantages arise also from the increased number of stages: an extra pair of mixers means not only more mismatches (therefore further deterioration of the image rejection ratio, *IIR*) but also more pressure on the first stage when it comes to noise figure, and on the last stage when it comes to linearity. The simulations presented in figure 2.12 show that to counteract the increased noise brought

by the now four mixers, the gain of either the LNA or the first mixers ought to be increased. This on the other hand has a negative impact on IIP_3, even more significant than it was in the case of the simple architecture whose simulation results were shown in figure 2.3.

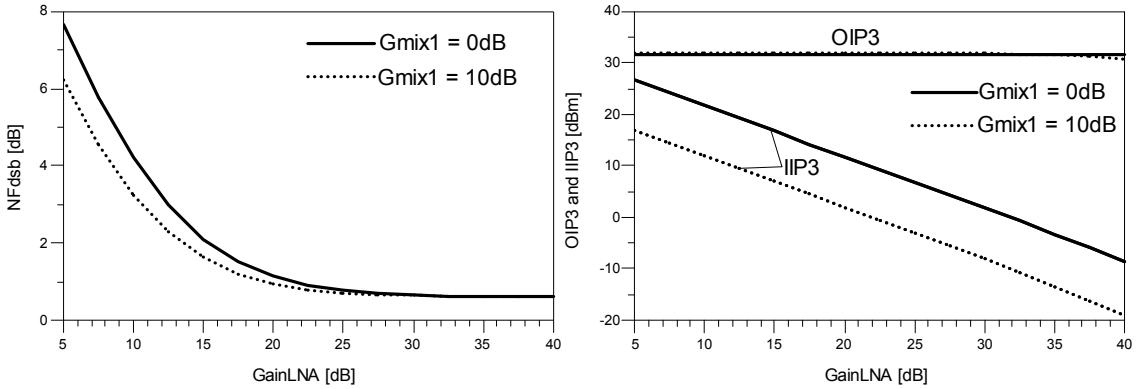


Fig 2.12: Noise (left) and linearity figures (right) for two values of the gain of the first pair of mixers

2.3.5 Comparison

Table 2.8 summarizes the advantages and disadvantages of the topologies discussed above. Albeit the second solution has the poorest noise figure of all, it has been chosen for reasons of simplicity and more importantly compatibility with the current basestation architectures. In addition, its noise behavior will have a diminished overall impact since, as shown in figure 2.1, a very low noise amplifier will be placed in front of our circuit.

Table 2.8: Advantages and disadvantages of the architectures studied in this chapter

	Interstage IR filter (fig 2.2)	Input IR filter (fig 2.4)	Simple IQ (fig 2.8) [1]	Weaver (fig 2.10)
Noise figure	++	--	+	-
Image rejection	+	+	-	--
Integrability	-	++	++	+
Complexity	+	++	-	--
IIP3	+	+	+	--
Compatibility with actual basestations	++	+	--	-

2.4 Choice of Circuits

All the above discussion has put it clearly: the last stage's linearity is a matter of paramount importance. Its noise is not, assuming the first stage has high enough a gain to handle it. That is why decisions as to what circuit should one choose begin with the selection of a mixer topology. On top of that, the input impedance of this block determines the LNA's architecture to a remarkable degree.

Most of the solutions found in literature rely on the Gilbert mixer [15, 16], but their linearity is limited (to about 15dBm of OIP_3) by the input transconductance stage. This is why efforts are directed towards linearizing the input stage [17]. Commercial active mixers rely on high bias currents to achieve 3rd order input intercept points of no more than 28.5dBm [18]. Passive mixers on the other

[1] separate I and Q paths considered in this comparison

hand are known to be having a higher linearity owing to their current domain operation and the excellent switching capabilities of the nMOS transistors [19, 20, 21, 9]. Current domain operation translates into low impedance nodes only (save for the local oscillator input); this situation is obtained through the use of a trans-impedance amplifier (TIA) at the output of the mixer. In practice, the overall linearity is basically limited by the TIA's linearity [9].

Having decided upon a mixer, we can now judge what type of LNA we ought to use. Since it has to convert input voltage to output current it must be a transconductance stage. In addition we want precise, constant gain over a large bandwidth. Applying negative feedback is therefore a must. At this point we limit our reasoning as to what type of feedback should be used only to the output node. Here, the LNA works in the current domain, therefore series feedback will sense this output current and feed it back (as current or voltage) to the input. In addition, series feedback increases the output impedance, making the g_m stage behave even more like a current source – a desired effect. We can venture to make some assumptions vis-à-vis the input feedback: since the input has to be matched to 50Ω and a bipolar transistor in CE configuration has an input impedance higher than this value, shunt feedback can level the input impedance to the desired level. As we will see later in this report, things are a little more complicated than this and a dual loop feedback network will be required to precisely set the input impedance. The first stage of the amplifier presented in [22] is a good example of what the feedback loops should look like.

Figure 2.13 reflects all the decisions made in the lines above. In addition, an LO interface is necessary to transform the power of the LO signal to voltage square waves in order to drive the nMOS switches that form the mixer core and achieve high linearity.

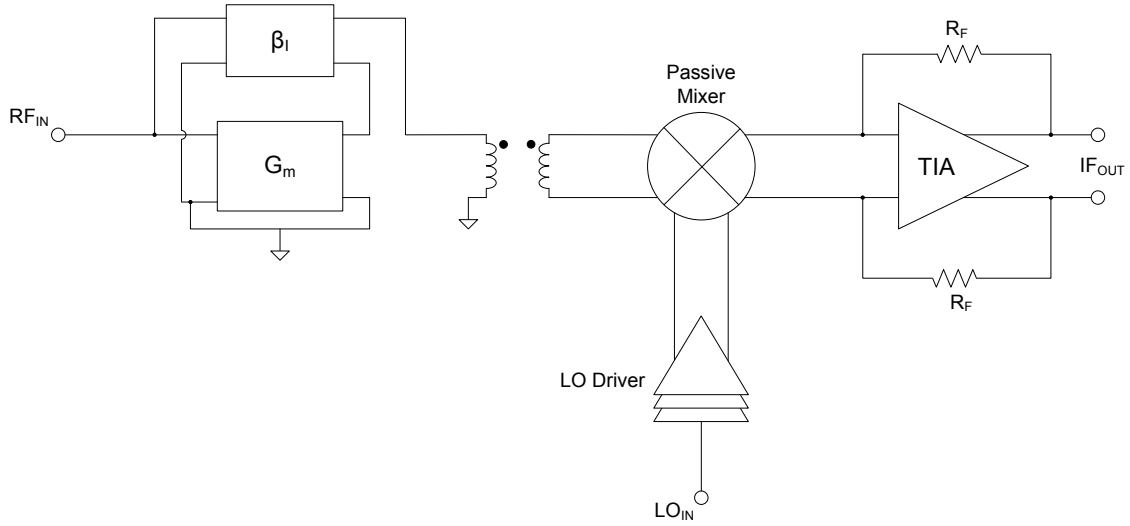


Fig 2.13: Detailed diagram of the chosen implementation

2.5 Device Characterization

Our defined goal is to find the right circuit implementations that could replace the actual circuits designed in GaAs technology. The question is now: is this feasible? What is the best we can hope for from our available devices? We obviously cannot design an amplifier with 0.4dB noise figure if the minimum noise figure the available devices can achieve is higher than 0.4dB.

To understand the physical limitations and avoid chasing impossible goals, the next part gives a short account of the performance of the devices available in the Qubic4x SiGe process.

2.5.1 SiGe npn Transistor

A list of device parameters for a Qubic4x npn transistor (length 20.7µm and width 0.5µm) is given below:

- current gain at DC, $\beta = 360$ (under the condition that the current density: $1.5\text{mA} \leq J_C \leq 5\text{mA}$)
- base-emitter junction transit time, $\tau_F = 0.6\text{ps}$
- base-emitter depletion capacitance, $C_{je} = 90\text{fF}$
- base-collector depletion capacitance, $C_{bc} = 4\text{fF}$
- base resistance, $r_B = 10\Omega$
- emitter resistance, $r_E = 0.4\Omega$
- collector resistance, $r_C = 13\Omega$

A plot of the minimum noise figure (NF_{min} @ 2GHz) (figure 2.14) – shows a lower limit on the noise figure of 0.5dB (for a collector current in the vicinity of 1mA). Adding matching and feedback elements will undoubtedly increase this number, revealing a first limitation of our SiGe process compared to more advanced GaAs processes.

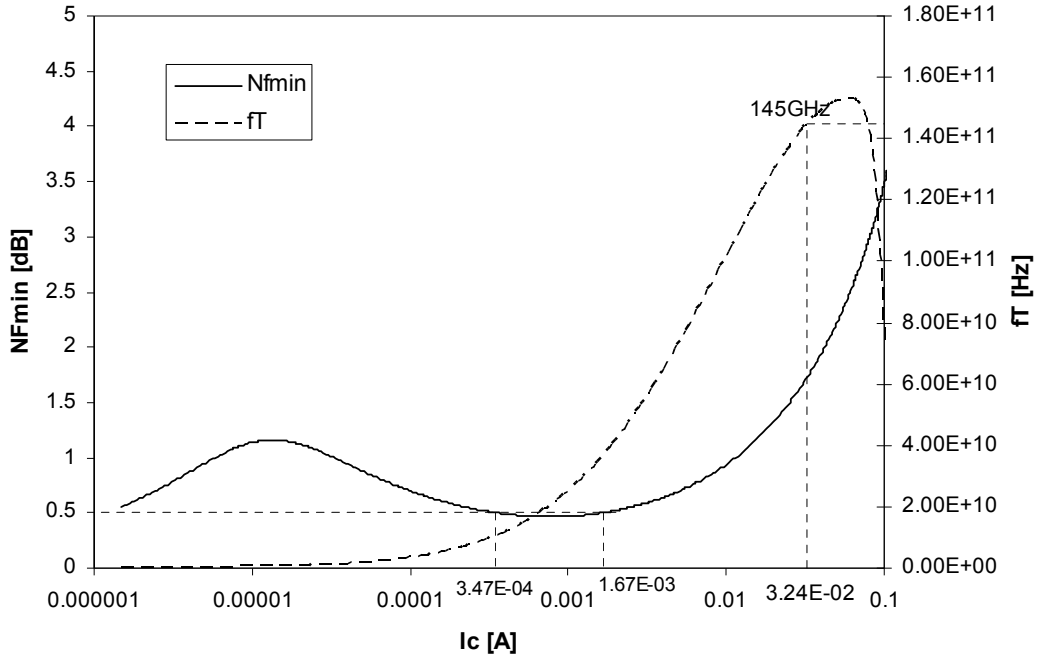


Fig 2.14: Minimum noise figure, NF_{min}, and transit frequency, f_T

To characterize the linearity performance of the same transistor, the setup shown in figure 2.15.a [23] is used. The feedback resistors, R_E and R_F, are sized such that input and output matching to 50 Ω and a power gain of approximately 20dB are achieved [24]:

$$\sqrt{R_F R_E} = R_S = 50\Omega \tag{2.7}$$

$$S_{21} = \frac{R_F}{R_S} \tag{2.8}$$

The device area was set to 0.5µm×20.7µm×36 to accommodate currents as large as 80mA. The DC voltage source V_{BC1} was scaled against V_{BB1} in order to maintain a constant V_{CE1} of 1.8V.

2.5.1 SiGe npn Transistor

The third order intercept points are plotted in figure 2.15.b as function of the collector current of transistor Q_1. We observe almost a flattening in the IP3 curves in the upper region of the collector current range. This shows that our device cannot reach values as high as 20dBm for IP3 per se, unless special techniques of linearization are employed.

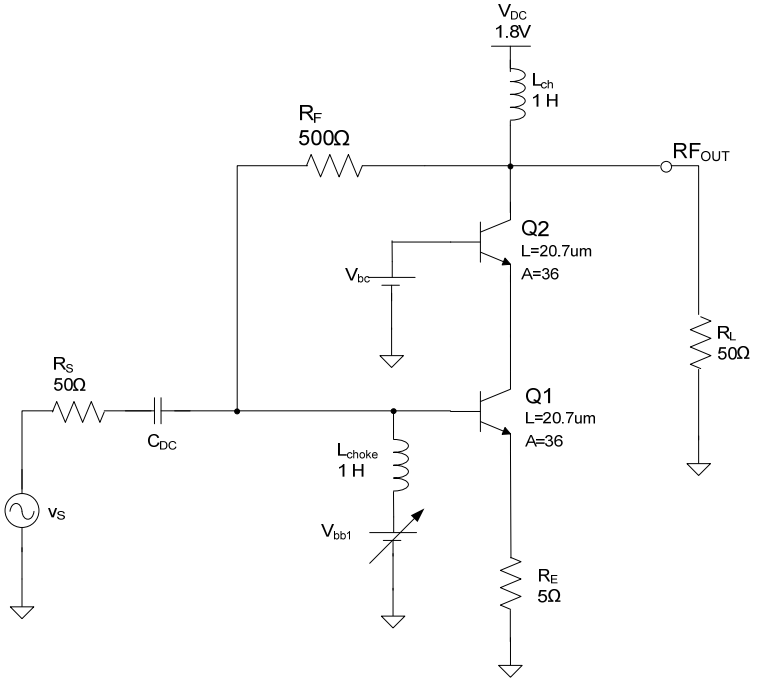
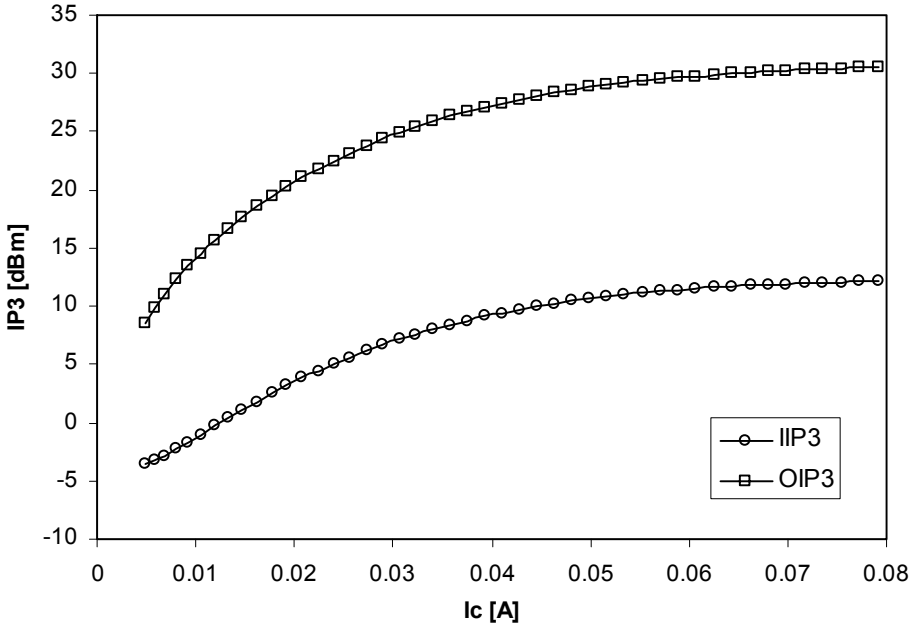


Fig 2.15.(a)



(b)

Fig 2.15: Large-signal characterization: a) simulation setup; b) Input and output IP_3 versus I_C f_{RF} =2GHz, tone spacing 1MHz, P_{RF} = -40dBm

2.5.2 MOS transistor

As for the MOS transistors, we expect them to be employed as switches in the current steering mixer, but also in the local oscillator subsystem, where the LO signal is basically a digital waveform. To quantify its performance in this respect, we use as benchmark a ring oscillator (figure 2.16) made up of minimum-size inverters: the length of both N and PMOS transistors is 0.25µm, the width 1 and 2 µm for the N and PMOS transistors respectively. This is also pointed out in figure 2.5. The generated waveform – plotted in figure 2.17 – shows an oscillation period of 0.307ns, which corresponds to a frequency of 3.257GHz.

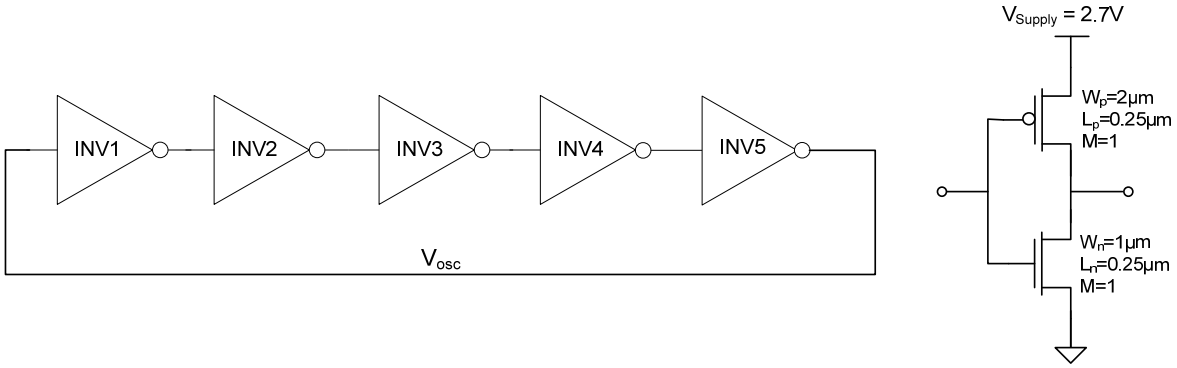


Fig 2.16: Ring oscillator used as benchmark for MOS transistors (left). Sizing also shown (right)

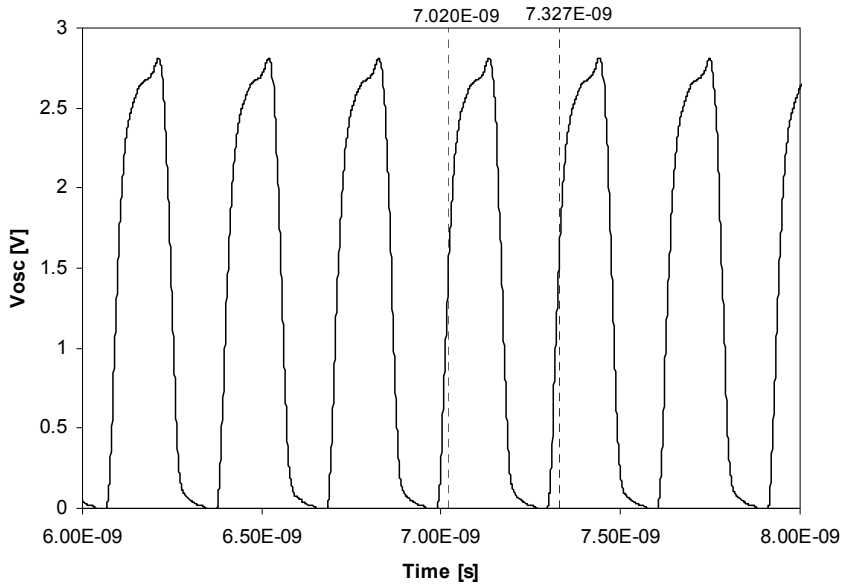


Fig 2.17: V_{osc} has an oscillation frequency of 3.26GHz

2.6 Conclusion

In our view, the simple IQ receiver would have normally been the best choice for a single IC implementation, had it not been for the incompatibility with the current basestation architectures. A less optimal solution – LNA and mixer without IR filter in between – has been selected though, for the main reason of simplicity and in fact small price to be paid.

Chapter 3

Low Noise Amplifier

The structure of the low noise amplifier (LNA) we are going to use in out basestation receiver has already been contoured at the end of the previous chapter: a g_m stage employing dual-loop feedback (series-shunt and series-series). Let us reiterate the reasons why transconductance and feedback are the right decision.

Driving an ideally zero load impedance requires that the output of our amplifier act like a current source. The input of the LNA on the other hand is driven by a power (I-V) source. A transistor in common emitter or common source configuration fits perfectly this role.

Wideband operation demands constant gain and port matching over large bandwidths – in our case 0.7 – 4GHz. Conventional microwave techniques use matching networks that are by nature narrow band or limited to a couple of hundreds of MHz (e.g. LC ladders) making them unsuited for wideband design. Negative feedback not only sets the closed loop gain to a precise value (almost independent of process variation, temperatures a.o.) but it also alters the input and output impedance levels: series feedback increases the impedance whilst shunt feedback decreases it [25], allowing thus good control over the input impedance.

Another known benefit of negative feedback is linearization. As a matter of fact, negative feedback is probably the first thing that comes to a designer's mind when speaking about high linearity amplifiers. It is known from theory that distortion goes down as the loop gain increases, but unfortunately a high loop gain is difficult to achieve at high frequency. For this reason, this approach is of little use in our design and in addition, as we will see later on, negative feedback actually creates some problems as far as stability is concerned.

Integrating the LNA and the mixer on the same chip eliminates the need of 50Ω output and input matching respectively and creates the conditions for a wideband matching at the interface between the two blocks. The only port that has to be impedance-matched is therefore the input of the LNA. Since the input signal is power – thus a combination of I and V, shunt (current feeding to the input) and series (voltage summation at the input) feedbacks at this node will offer the means to achieve impedance matching.

As for the output feedback type, due to the low input impedance of the mixer, the only signal that can be sensed in the output loop is current. Therefore series dual feedback is the only possible choice.

The circuit drawn in figure 3.1 embodies all the decisions made so far.

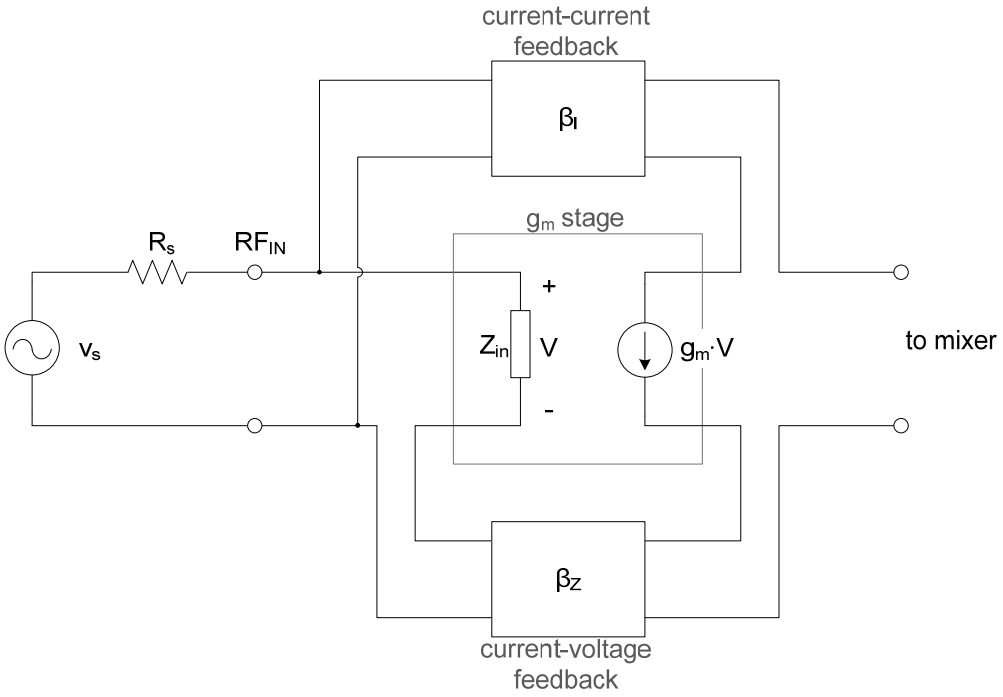


Fig 3.1: Block diagram of the LNA

3.1 Circuit Selection

The low noise requirement is so stringent that special care must be taken in the selection of feedback networks. The series-shunt feedback network (the top box of figure 3.1) will be connected between the collector and base of the transistor and implemented using a transformer, which ideally is a noiseless element. The structure of a transistor – which is a three-terminal device – restricts the choice of the other feedback network to a simple impedance placed in the emitter. A similar structure is presented in [23], with the main difference that that is a narrow-band amplifier while we target operation over a couple of GHz.

The complete schematic of our dual loop feedback amplifier is shown in figure 3.2. The core of this amplifier – a g_m stage – is formed by a common-emitter cascode stage (Q_1 and Q_2). Cascoding is a good method of unilateralizing an amplifier as it eliminates the Miller effect; this translates in a significant increase in bandwidth. In addition, cascoding augments the output resistance of the amplifier (it therefore behaves even closer to an ideal current source).

3.2 Small-Signal Analysis

In the following analyses, the cascode stage will be replaced by the small signal equivalent circuit of a simple transistor without the base-collector capacitance. For reasons of simplicity, the transformer is considered ideal, i.e. $k_m = 1$. A small signal equivalent schematic for the amplifier in figure 3.2 is shown in figure 3.3. It will be used to derive expressions for gain and input impedance.

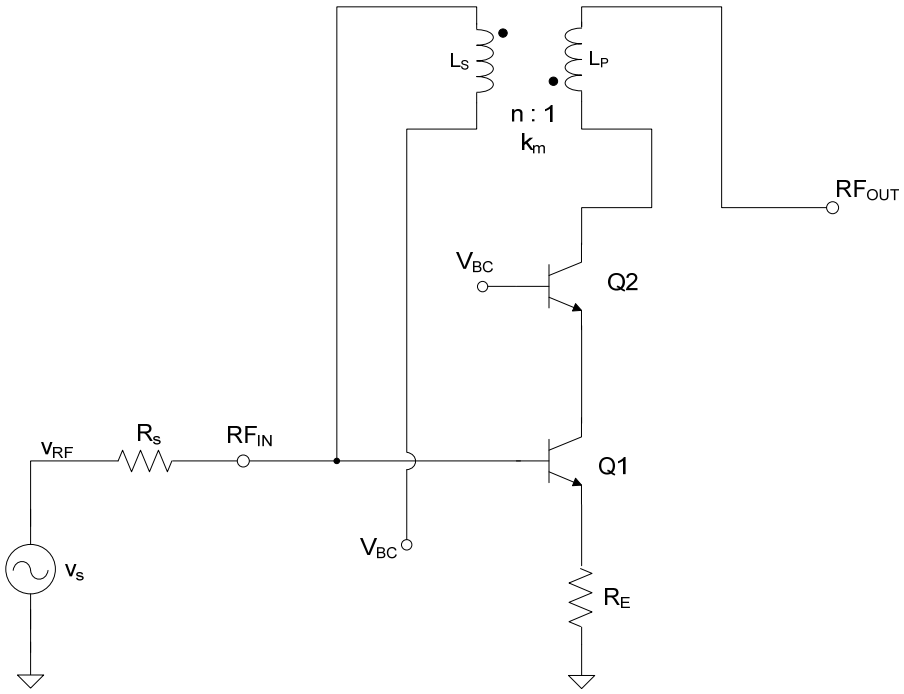


Fig 3.2: Dual-loop feedback amplifier

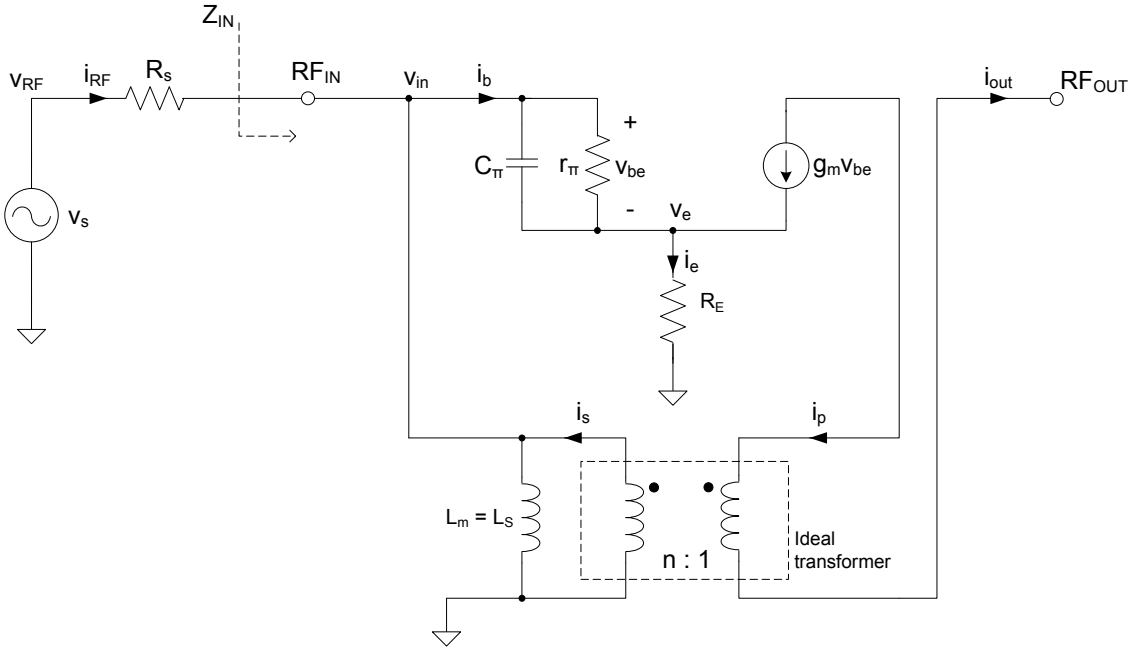


Fig 3.3 Small signal equivalent circuit of the amplifier shown in figure 3.2

First let us define some quantities that will be used throughout these analyses. The relation between the currents in the primary and secondary windings of the transformer is determined by the turns ratio *n* [10]:

$$i_S = \frac{i_P}{n} \tag{3.1}$$

A degenerated transistor of transconductance g_m and emitter impedance Z_E can be equated with a non-degenerated stage of transconductance $g_{m,eff}$ defined as [25]:

$$g_{m,eff} = \frac{g_m}{1 + g_m Z_E} \tag{3.2}$$

3.2.1 Input Impedance

Figure 3.4 shows the equivalent circuit seen looking into the RF_{IN} node. The first element (from left to right) is the self-inductance of the transformer translated unmodified from figure 3.3; the value of the middle resistance is derived from equation (3.1) as:

$$-i_S = -\frac{i_P}{n} = \frac{g_{m,eff} \cdot v_{in}}{n} \rightarrow R = \frac{v_{in}}{-i_S} = \frac{n}{g_{m,eff}}$$

The right-hand side impedance is the effect of gyration of R_E from emitter to base (β-transformation) [10]. Since at high frequency Z_π is dominated by C_π,

$$\beta(j\omega) = \frac{g_m}{j\omega C_\pi} \tag{3.3}$$

and the equivalent impedance is a capacitor of value $C_\pi/g_m R_E$ in series with R_E shown in the same figure. As a further simplification, R_E can be neglected.

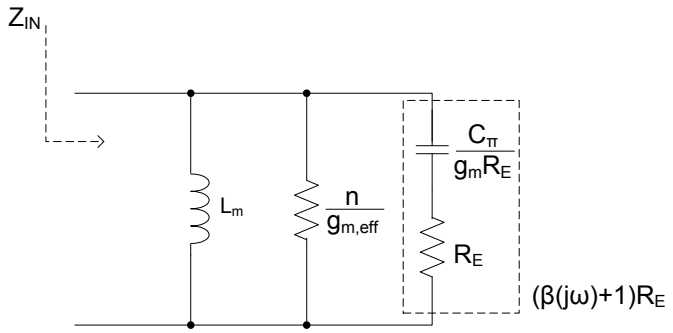


Fig 3.4: Equivalent circuit of input impedance Z_{IN}

The input impedance is now a parallel combination of an inductance, resistance and capacitance

$$Z_{IN} = j\omega L_m \,||\, \frac{n}{g_{m,eff}} \,||\, \frac{g_m R_E}{j\omega C_\pi} \tag{3.4}$$

Matching (to a 50Ω impedance) consists of resonating the complex elements and setting the real part to R_S. The first condition results in a resonance frequency

$$f_{res} = \frac{1}{2\pi\sqrt{L_m \frac{C_\pi}{g_m R_E}}} \tag{3.5.a}$$

It is worth mentioning that although matching relies on LC resonance, the result is wideband since the quality factor of such resonator is very low. The condition imposed on the real part is given by equation (3.5.b), derived using (3.2).

$$\frac{n}{g_{m,eff}} = R_S \Leftrightarrow \frac{n(g_m R_E + 1)}{g_m} = R_S \tag{3.5.b}$$

As equations (3.5) show, there are basically four independent design parameters: the transformer's turns ratio, n and self-inductance, L_m, the transconductance g_m and degeneration resistance R_E. Further

conditions will be imposed by noise, as simultaneous impedance and noise matching must be achieved, and gain.

3.2.2 Gain

The transmission matrix of the LNA illustrated in figure 3.3 is given by

$$\begin{pmatrix} v_{IN} \\ i_{RF} \end{pmatrix} = \begin{pmatrix} 0 & -\dfrac{1}{g_{m,eff}} \\ \infty & -\dfrac{1}{n} - \dfrac{1}{\beta(j\omega)} - \dfrac{1}{g_{m,eff}\, j\omega L_m} \end{pmatrix} \begin{pmatrix} v_{out} \\ i_{out} \end{pmatrix} \tag{3.6.a}$$

where $g_{m,eff}$ is given by (3.2). At resonance (of L_m with C_π), and considering β at that frequency high enough, expression (3.6.a) simplifies to:

$$\begin{pmatrix} v_{IN} \\ i_{RF} \end{pmatrix} = \begin{pmatrix} 0 & -\dfrac{1}{g_{m,eff}} \\ \infty & -\dfrac{1}{n} \end{pmatrix} \begin{pmatrix} v_{out} \\ i_{out} \end{pmatrix} \tag{3.6.b}$$

3.2.3 Noise and Noise Matching

Instead of relying on correlation matrixes to derive expressions for noise matching, we will use classic methods to calculate the input referred sources of noise – voltage and current equivalent sources alike. Figure 3.5 shows the schematic with noise sources added. The expressions for each of these sources are given in the list below:

$$\begin{aligned} \overline{v_{n,B}^2} &= 4kTr_b \\ \overline{i_{n,C}^2} &= 2qI_C \\ \overline{i_{n,B}^2} &= 2qI_B \\ \overline{i_{n,E}^2} &= \frac{4kT}{R_E} \end{aligned} \tag{3.7}$$

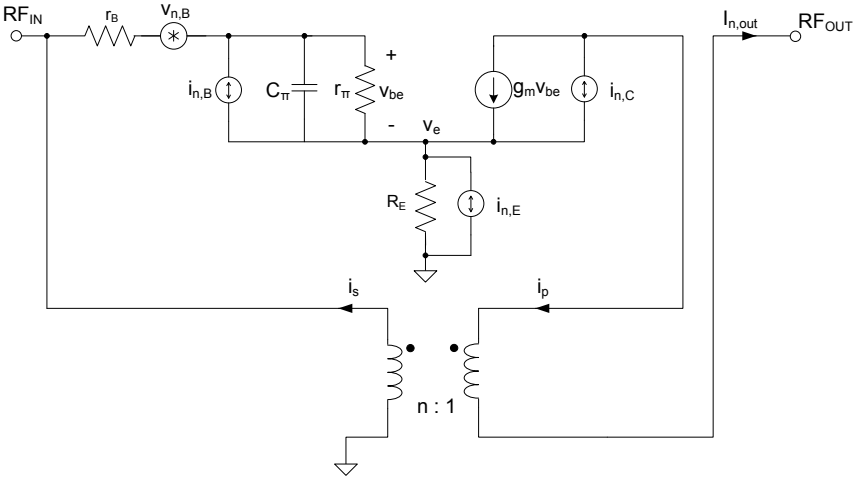


Fig 3.5: Equivalent circuit for noise analysis

The power spectral density of the input referred noise voltage source, $\overline{v_{n,eq}^2}$, obtained by short-circuiting the input, has the following expression

$$\overline{v_{n,eq}^2} = \overline{v_{n,B}^2} + \frac{\overline{i_{n,E}^2} + \overline{i_{n,B}^2} + \overline{i_{n,C}^2}}{g_{m,eff}^2} \tag{3.8}$$

while the power of the equivalent current source, $\overline{i_{n,eq}^2}$, (obtained with an open input) is expressed as

$$\overline{i_{n,eq}^2} = \overline{i_{n,B}^2} + \frac{\overline{i_{n,E}^2} + \overline{i_{n,B}^2} + \overline{i_{n,C}^2}}{n^2} \tag{3.9}$$

The noise factor is defined as the signal-to-noise ratio at the output to the signal-to-noise ratio at the input:

$$F = \frac{SNR_{Out}}{SNR_{In}} = 1 + \frac{\overline{v_{n,eq}^2} + \overline{i_{n,eq}^2} R_S^2}{4kTR_S} \tag{3.10}$$

To determine the optimum input impedance for minimum noise figure ($R_{S,opt}$) we take the derivative of (3.10) with respect to R_S.

$$\frac{\partial F}{\partial R_S} = 0 \rightarrow R_{S,opt} = \frac{\overline{v_{n,eq}^2}}{i_{n,eq}^2} \tag{3.11}$$

A series of simple manipulations leads to

$$R_{S,opt} = \frac{n^2}{g_m^2}\left(\frac{2R_E r_B + \left(1+\frac{1}{n}\right)g_m R_E + 2}{g_m R_E\left(n+\frac{1}{n}\right) + R_E + 2}\right) \tag{3.12}$$

3.2.4 Noise and Input Impedance Matching

For simultaneous noise and impedance matching, we impose an additional constraint on the value of $R_{S,opt}$ as given by equation (3.12), that is:

$$R_{S,opt} = R_S = 50\Omega \tag{3.13}$$

where R_S is the generator resistance, usually equal to 50Ω.

The condition for impedance matching is set by equation (3.5.b), rewritten below:

$$n\left(R_E + \frac{1}{g_m}\right) = R_S \tag{3.14.a}$$

In our design we expect that $R_E >> 1/g_m$, and consequently (3.14.a) simplifies to:

$$nR_E = R_S \tag{3.14.b}$$

As a design example, let us assume we can design a transformer that has an effective turns ratio[1] of 10. The required degeneration resistor must consequently have 5Ω. To achieve noise matching, we can control two parameters: g_m – via the DC collector current I_C, and r_B – by scaling the device ($r_B = r_{B0}/scale$, where r_{B0} is the value of base resistance of an unscaled transistor).

[1] The effective turns ratio is given by the product between the physical number of turns, N, and the coupling coefficient, k_m.

Figure 3.6 shows a plot of $R_{S,opt}$ (given by equation (3.12)) as function of the collector current, for different values of the transistor scale. The intrinsic base resistance, r_{B0}, was taken as 10Ω, based on the data given at the end of chapter 2 for a Mextram transistor of length 20.7μm. We can see from that figure the significant contribution of r_B, and the importance of scaling the device in controlling $R_{S,opt}$.

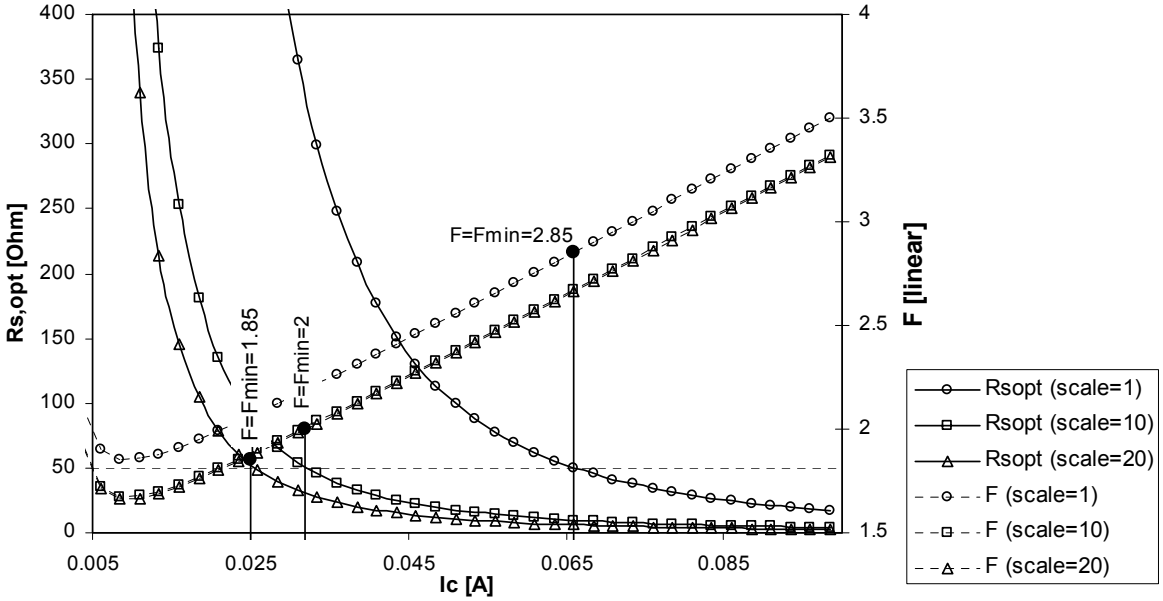


Fig 3.6: Optimum noise resistance, $R_{S,opt}$ (eq. (3.12)), and related noise factor F (eq. (3.10)) as function of collector current, I_C, and scale

We see from this figure that achieving simultaneous noise and impedance matching at higher currents becomes increasingly difficult, since for these higher current levels to a very small value for the scaling factor is required. This translates into large collector current densities, way above the capabilities of regular SiGE BJTs (for example for a collector current of 70mA, *scale* = 1, and consequently the current density coincides with I_C). In addition, the same figure shows that minimum noise figure is usually achieved at low current densities (compare also in figure 2.14 the value of 1mA around which NF has a minimum, with 32.4mA where f_T is maximum; usually, currents higher than this value are not recommended). This goes against some designers' desire to improve the linearity by increasing the collector current, which in this case would hurt the noise performance.

3.3 Large Signal Analysis

3.3.1 Out-of-band Emitter Tuning

[23] proves that by carefully choosing the terminal impedances at the frequencies that relate to the 2nd order products (baseband – Δs, and double-frequency – $2s$), transconductance g_m and area, the IM_3 products of a common-emitter stage can be nullified – technique called out-of-band third-order distortion cancellation.

Without taking on anew the derivation presented in [23], we will make use of the conclusion of that analysis. To achieve third-order harmonic cancellation the following conditions must be fulfilled:

$$r_{2F} = \frac{1 - r_{BB}}{2r_{BB}} \tag{3.15.a}$$

$$C_R = \frac{g_m \tau_F}{C_{jE} + C_{bc}} = r_{2F} \tag{3.15.b}$$

where r_{BB} and r_{2F} refer to resistances at the baseband and broadband harmonic frequencies, defined as

$$r_{BB} = g_m \left(\frac{R_{b_{BB}}}{\beta} + R_{e_{BB}} \right) \tag{3.16.a}$$

$$r_{2F} = g_m \left(\frac{R_{b_{2F}}}{\beta} + R_{e_{2F}} \right) \tag{3.16.b}$$

where $R_{b,BB}$ is the resistance $R_b = R_S + r_b$ at the baseband frequency and $R_{b,2F}$ the same resistance at the second-harmonic frequency. $R_e = R_E + r_e$ is defined in a similar manner for *BB* and *2F* (r_b and r_e are the intrinsic base and emitter resistances); C_R is the ratio of the diffusion and depletion capacitances. For clarity, figure 3.7 shows the large-signal model of a CE stage in which the aforementioned elements are annotated.

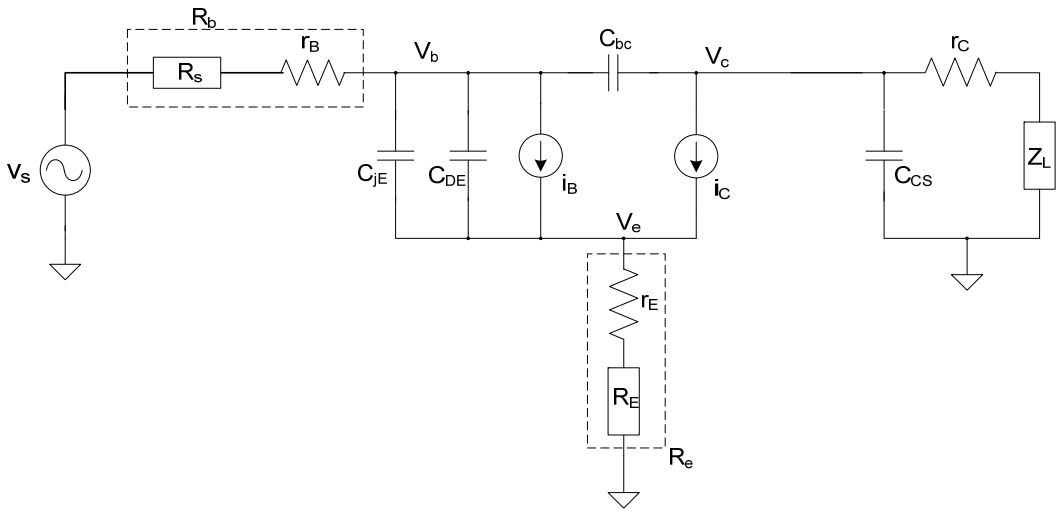


Fig 3.7: Large-signal model of a CE stage [23]

To avoid any kind of asymmetry in the IM_3 products, the baseband and 2nd harmonic resistances are usually chosen equal and of value 0.5. The value of C_R will consequently be affected:

$$r_{BB} = r_{2F} = g_m \left(\frac{R_b}{\beta} + R_e \right) = 0.5 \tag{3.17.a}$$

$$C_R = \frac{g_m \tau_F}{C_{jE} + C_{bc}} = 0.5 \tag{3.17.b}$$

For low frequency cancellation, fulfilling condition (3.17.a) will suffice. As the operation frequency increases, the distortion caused by the base-emitter diffusion capacitance becomes more significant. The analysis in [23] shows that the frequency at which this nonlinear capacitance becomes dominant over the nonlinear collector current differs when the tuning is done on the base resistance (thus controlling the value of R_b) as compared to emitter tuning through the resistance R_E. The first situation occurs at frequencies below the GHz region whereas the latter is much more

broadband[1]. To minimize the degradation in IP_3 by caused by the base-emitter capacitance the device must me sized such that condition (3.17.b) is satisfied.

3.3.2 Effect of (Shunt) Feedback on Harmonic Cancellation

Applying shunt feedback – as in figure 3.2 – to the input of a CE stage will alter the conditions set by equations (3.15) (or (3.17)) due to a modification in the effective current gain, β. To illustrate how β changes when a current is fed back to the input node, we will use the simple circuit shown in figure 3.8.

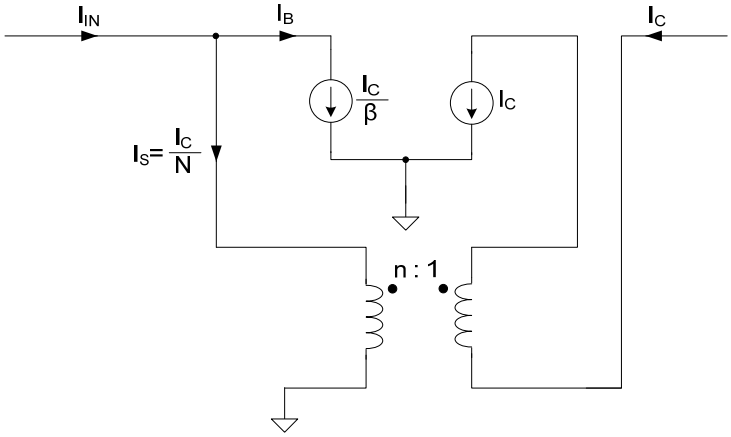


Fig 3.8: Simplified CE stage with current-current feedback implemented by means of a transformer

The current gain – defined as $\beta' = \frac{I_{IN}}{I_B}$ - is determined by the parallel combination of the transistor's base current, $I_B = \frac{I_C}{\beta}$, and the current that flows through the secondary winding of the transformer, $I_B = \frac{I_C}{N}$, where N is the transformer's turns ratio. The result is:

$$\beta' = \frac{N\beta}{N+\beta} \cong N \tag{3.18}$$

Since $N << \beta$, to satisfy equation (3.17.a) g_m must be decreased; we therefore draw the conclusion that when current feedback is applied, harmonic cancellation occurs at significantly lower currents compared to the no-feedback situation.

To verify the statements made in this and previous paragraphs, we use ideal transistors to implement cascode amplifier; the distortion mechanisms taken into account are illustrated in the large-signal equivalent circuit shown in figure 3.9 For reasons of simplicity, we only consider the distortion due to the exponential behavior of the collector current (therefore the base emitter diffusion capacitance was ignored). The transistor was modeled as a nonlinear voltage-controlled current source of coefficients g_m, g_{m2}, g_{m3} given as [24]:

[1] The drop in IP3 improvement due to cancellation occurs at $\frac{3}{2}\omega_T / \beta$ in the case of base tuning and $\frac{3}{2}\omega_T$ in the case of emitter tuning.

$$g_m = \frac{I_C}{V_T}$$
$$g_{m,2} = \frac{I_C}{2V_T^2} \tag{3.19}$$
$$g_{m,3} = \frac{I_C}{6V_T^3}$$

where I_C is the DC collector current and V_T the thermal voltage. The collector current is therefore expressed as:

$$i_c = g_m v_{be} + g_{m,2} v_{be}^2 + g_{m,3} v_{be}^3 \tag{3.20}$$

The base current is obtained by simply dividing i_c to β; in our schematic β is set to 100.

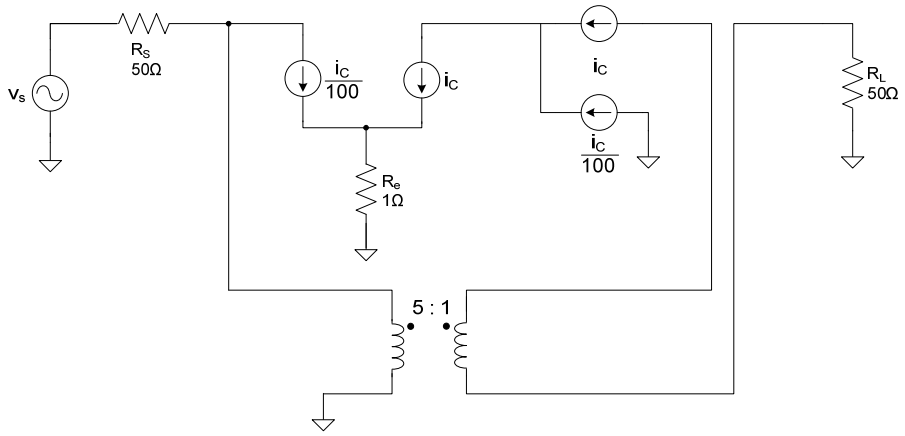


Fig 3.9: Large-signal schematic of cascode amplifier employing current-current feedback. The value of i_C is given by eq. (3.20)

Simulation results for IIP_3 versus the collector DC current are shown in figure 3.10 for two cases: when no feedback is applied (continuous line with rectangles), IIP_3 peaks at approximately 12mA; when negative feedback is added, we observe a shifting of the value of the current at which

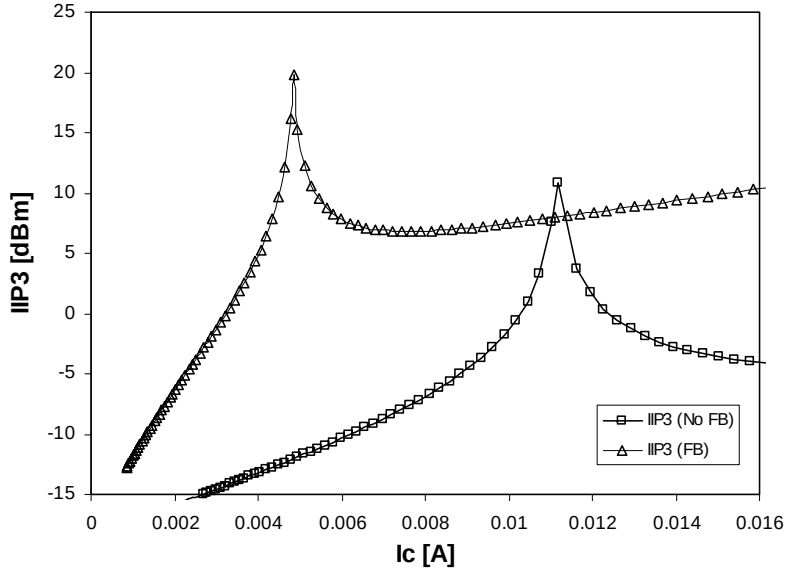


Fig 3.10: 3rd order input intercept point, IIP_3 versus collector current I_C of the circuit from figure 3.9

cancellation occurs. Using the component values annotated in figure 3.9 and equation (3.18), the current gain β' is calculated as having a value of 4.76. Because of this value, the term between brackets in equation (3.17.a) becomes 3.2 times larger than before (when no feedback was applied and β = 100); this in turn translates to an I_C necessary to satisfy the condition for cancellation 3.2 times smaller than before. This is confirmed by figure 3.10 where we can see the peak occurring at about 4.5mA – continuous line with triangles.

As for the high-frequency condition – equation (3.17.b) – as g_m is scaled down to satisfy the condition on the resistive termination, the depletion capacitance C_{jE} must be adjusted accordingly by decreasing the area of the transistor. This problem is studied by adding C_{DE} and C_{jE} to our ideal transistor. The schematic in figure 3.9 is updated to include a non-linear capacitance – C_{DE} – modeled by the following coefficients:

$$C_D = \tau_F g_m \tag{3.21.a}$$

$$C_{D,2} = \tau_F g_{m,2} \tag{3.21.b}$$

$$C_{D,3} = \tau_F g_{m,3} \tag{3.21.c}$$

and a linear capacitance – C_{jE} – in parallel with C_{DE}, simply modeled as

$$C_{jE} = scale \cdot C_{jE,0} \tag{3.22}$$

where $C_{jE,0}$ is the capacitance of an unscaled (unity) transistor. In our simulations τ_F was set to 0.6ps and $C_{jE,0}$ to 100fF. These changes are pictured by figure 3.11.

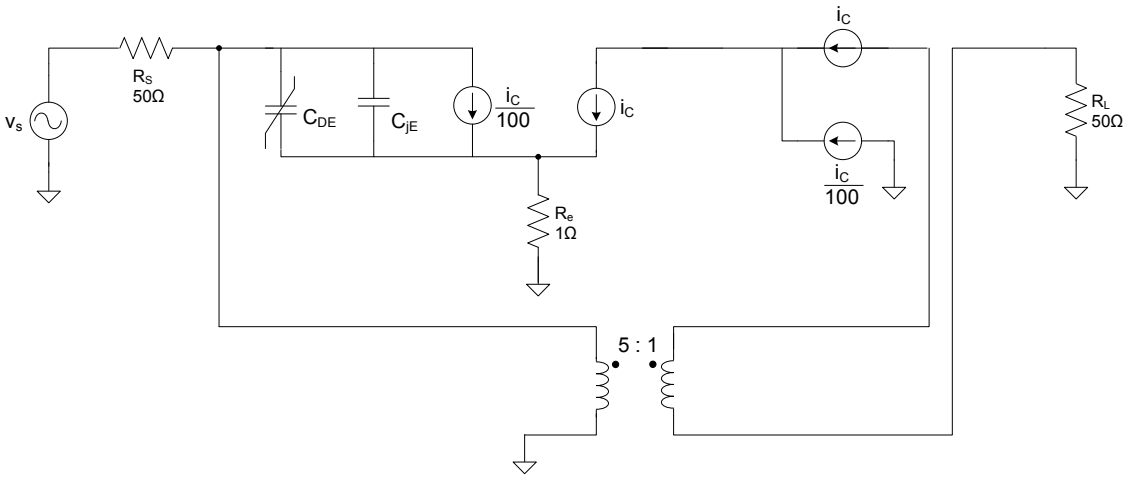


Fig 3.11: Distortion due to non-linear base-emitter capacitance is added to the schematic form figure 3.9

Fig 3.12 presents the simulation results, again for three cases: with and without feedback, and with different transistor sizing (set by the parameter *scale*), as imposed by equation (3.17.b).

This effect is even more dramatic in reality, as simulations run on a Mextram transistor amplifier show it – see figure 3.13. The differences in the optimum current for cancellation between figure 3.13 and 3.12 are mainly due to the intrinsic emitter resistance that increases the effective value of R_e and push the peak to lower values. This contribution becomes even more significant as the device is scaled down to adjust the capacitive part of equation (3.17), in which case the peak occurs at even lower values of I_C.

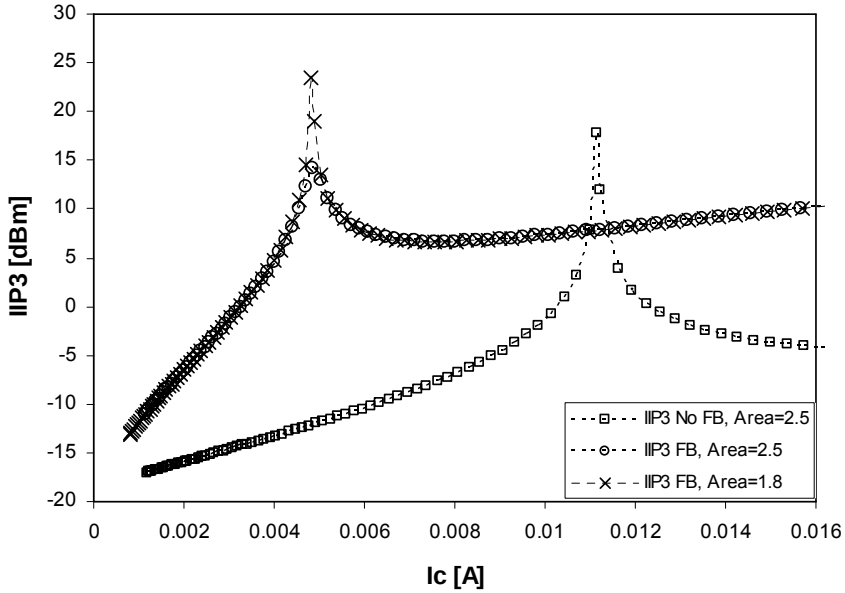


Fig 3.12: Input IP3 of the amplifier of figure 3.10 for three cases: no feedback, scale = 2.5; feedback, scale = 2.5; feedback, scale = 1.7

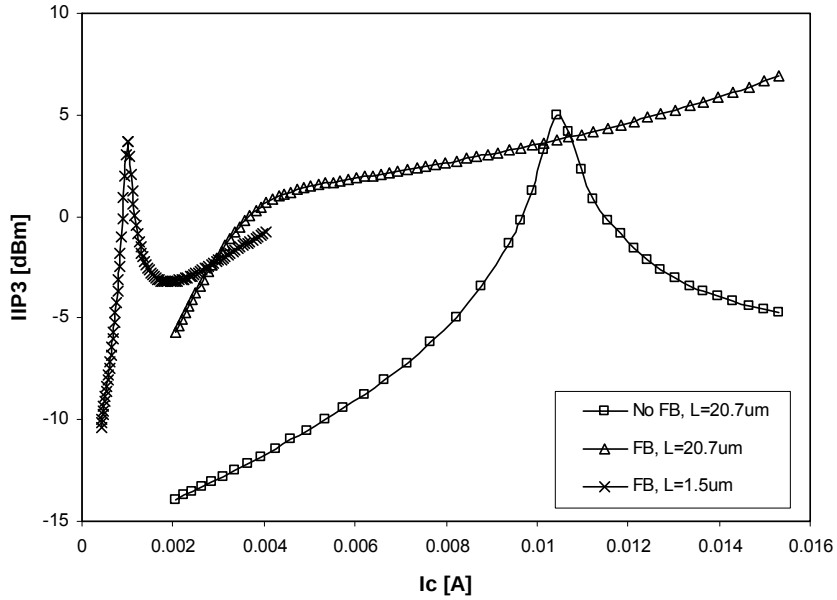


Fig 3.13: IIP_3 of a Mextram transistors amplifier

3.3.3 Using the Feedback through C_{bc} to Make the IM3 Cancellation Independent of Scaling

At some point downsizing the transistor (to satisfy condition (3.17.b)) becomes impractical, since there are other constraints (such as maximum current density, noise matching, maximum f_T, etc.) on the current density of the transistor (from a length of 20.7μm down to 1.5μm, as we can see from the legend of figure 3.13).

A reexamination of equation (3.15.b) – rewritten below for convenience – may prove useful if we want to overcome the abovementioned issue.

$$C_R = \frac{g_m \tau_F}{C_{jE} + C_{bc}} = 0.5 \tag{3.23}$$

The summation of C_{bc} to C_{jE} indicates that the author in [23] assumed the device was unilateral, in which case the two capacitors can be considered as connected parallel and as having currents of the same sign[1]. This situation changes in the case C_{bc} undergoes a Miller effect: since the voltage swing at the collector is larger than at the base (therefore voltage gain) the current through C_{bc} will no longer flow from the base as before but to the base, with the result that C_{bc} must be subtracted from C_{je} in equation (3.23). Equation (3.23) is rewritten now to include the Miller effect:

$$C_R = \frac{g_m \tau_F}{C_{jE} - A_V C_{bc}} = 0.5 \tag{3.24}$$

where A_V is the voltage gain from collector to emitter. This offers the possibility to keep the area of the device larger than the value imposed by (3.23) but still achieve cancellation, since C_R is kept equal to 0.5 (by controlling A_V). Freedom of choosing the size of the device means we can bias it to a higher cut-off frequency, which will also compensate any loss in bandwidth caused by the Miller effect on C_{bc}.

A circuit implementation of the aforementioned method is shown next in figure 3.14 for a cascode stage. The voltage gain A_V is approximately

$$A_V = g_m R_C \tag{3.23}$$

which means that by controlling R_C we can compensate the increase in C_{jE}.

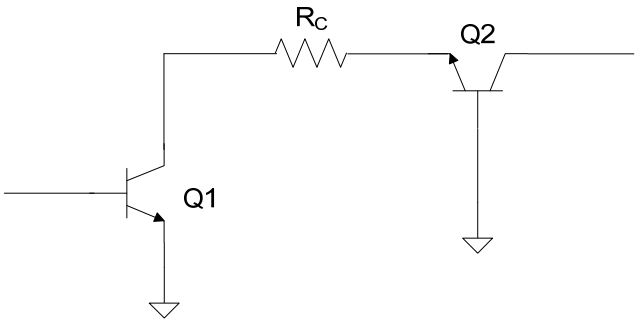


Fig 3.14: The Miller effect can be used to obtain harmonic cancellation without having to scale the device

To verify our statements, the test amplifier from figure 3.11 is updated to include C_{bc} and the collector resistance, R_C. The new model is illustrated in figure 3.15 while simulations results are presented further in figure 3.16.

If no R_C is added, the device has to be scaled to *Area* = 1.1 – simple continuous line. Next, the device area is deliberately untuned to a higher value (3), with the consequence that the cancellation becomes less perfect – dashed line. Finally, R_C is added and adjusted such that IIP_3 peaks again at a value of 25dBm – no line, rectangles only.

The values for $C_{jE,0}$ and $C_{bc,0}$ used in the simulations above are 100fF and 50fF respectively. For an area of 1.1, the denominator of equation (3.23) is 165pF. Setting the area to 3 increases this value to 450fF. In order to still satisfy the condition imposed by (3.21) we have to get rid of almost 300fF. According to (3.24), $A_V = 6$ will create a Miller capacitance of the desired value (300fF). With $g_m = 0.2$S (at $I_C = 5$mA, as read from figure 3.16), R_C has to be 30Ω. In reality, as figure 3.16 shows, R_C must have a value larger than this to account for the drop in gain at RF frequencies.

[1] For simplicity in calculations, [21] set the collector impedance at baseband and second harmonic to zero, which implies no voltage swing at the collector and therefore unilaterality

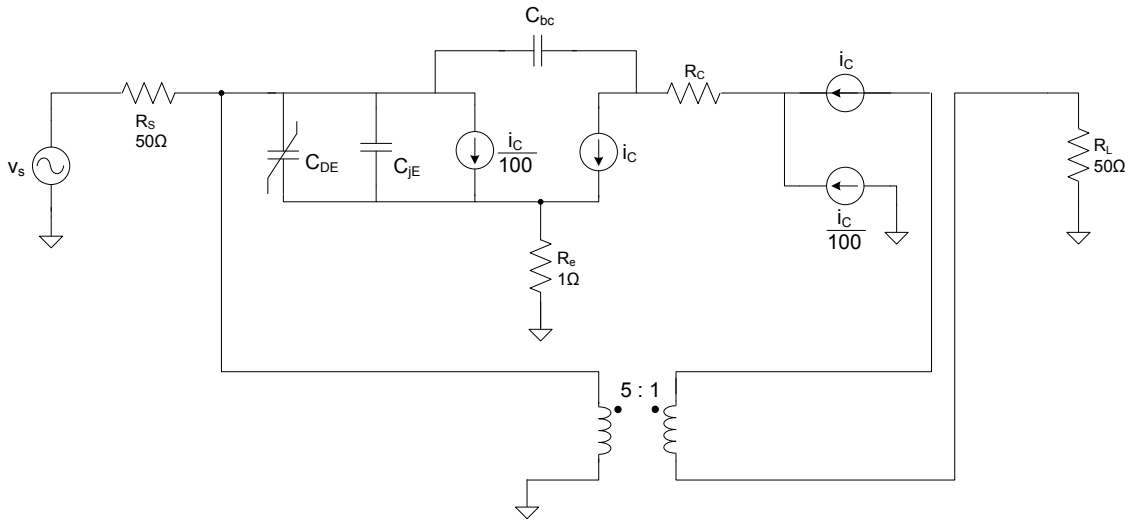


Fig 3.15: C_{bc} and R_C are added to the large signal model of figure 3.10

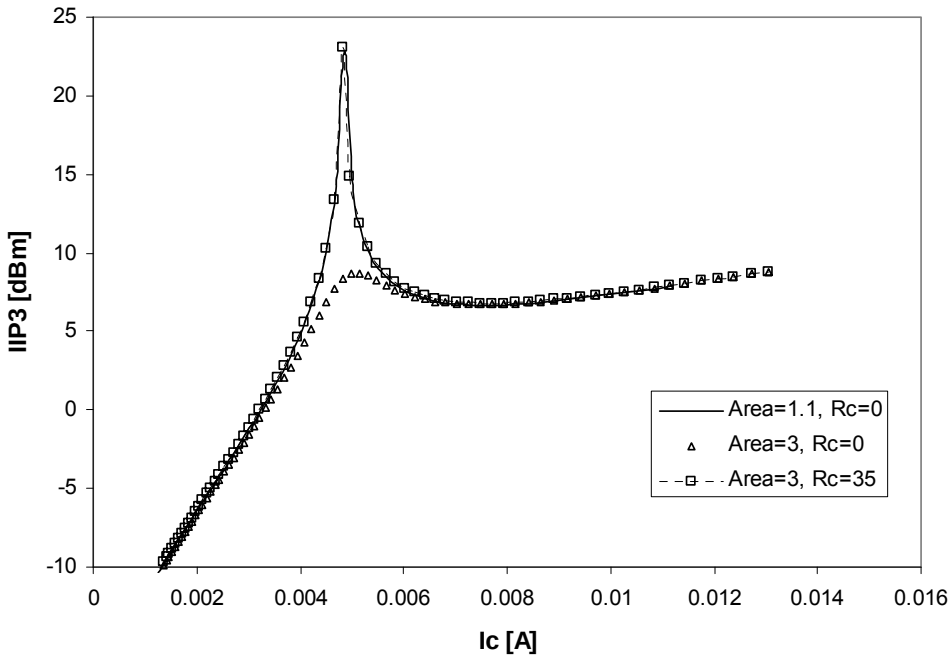


Fig 3.16: IIP_3 vs I_C for three cases (Area=1.1, and Area=3 and $R_C = 0$ and $R_C = 20$).
Adding C_{bc} required a re-scaling of the device as the denominator in equation (3.17.b) changed (for the case where $R_C = 0$)

3.3.4 Increasing the Bias Current while Preserving the IM3 Cancelation

Two things push the current at which IM_3 cancellation occurs down to low values:

- high values of the emitter resistance, R_e (according to equation (3.17.a). Equation (3.5.b) states that the input impedance of our shunt-series feedback amplifier is *directly*

proportional to the transformer's turns ratio, *n,* and degeneration resistor, R_e. Since n[1] is limited by physical limitations related to IC implementations, we must rely on R_e to set Z_{in} to 50Ω.

For example, if $N = 4$ (the physical number of turns) and $k_m = 0.5$ (coupling coefficient), then $n = 8$ and R_e must be set to approximately 6Ω.

- applying shunt feedback, as the previous section has proved it.

Figure 3.12 shows that for $R_e = 1\Omega$ and feedback applied the optimum current for harmonic cancellation is around 6mA. For a value of $R_e = 6\Omega$, we expect that cancellation be produced at a current 6 times smaller (approximately 1mA). Much too low a value for our design, which must also meet targets related to gain and P_{1dB}.

To reconcile the two needs – IM_3 cancellation and input matching – we have to separate them in two distinct problems:

- Z_{in} is determined by the broadband emitter resistance $R_{e,2F}$[2], through the (approximate) relationship:

$$Z_{in} \approx n \cdot R_{e,2F} \tag{3.24}$$

(derived from (3.5.b)), where n is the effective turns ratio of the transformer.

- g_m is for the most part determined by the baseband emitter resistance, $R_{e,BB}$. To observe this dependence we rewrite below equation (3.15.a), that sets the condition for harmonic cancellation, as:

$$r_{2F} = \frac{1 - r_{BB}}{2r_{BB}} \rightarrow g_m R_{2F} = \frac{1 - g_m R_{BB}}{2 g_m R_{BB}} \tag{3.25}$$

where

$$R_{BB} = \frac{R_{b_{BB}}}{\beta} + R_{e_{BB}} \tag{3.26.a}$$

$$R_{2F} = \frac{R_{b_{2F}}}{\beta} + R_{e_{2F}} \tag{3.26.b}$$

An important observation has to be done here: the DC bias voltage of the transistor is usually applied to the base terminal by means of an RF choke, which also functions as short-circuit for the baseband harmonics. In this case, equation (3.26.a) simplifies to

$$R_{BB} = R_{e_{BB}} \tag{3.27}$$

Solving equation (3.25) for g_m leads to the following solution:

$$g_m = \frac{-R_{BB} + \sqrt{R_{BB}^2 + 8 R_{2F} R_{BB}}}{4 R_{2F} R_{BB}} \tag{3.28}$$

Equation (3.28) is plotted in figure 3.17 as a function of R_{BB} (or $R_{e,BB}$) for different values of $R_{e,2F}$. It shows combinations of $R_{e,BB}$, $R_{e,2F}$ and g_m that lead to 3rd order harmonic cancellation. The condition imposed by equation (3.17.b) in which $R_{e,BB}$ is equal to $R_{e,2F}$ is just a particular case that generates an ultra-wideband solution for the 3rd order harmonics cancellation.

The obvious solution is to use a low $R_{e,BB}$ to get the possibility of using a large bias current, and a properly high $R_{e,2F}$ to ensure input matching. Figure 3.18 shows two possible ways of implementing a degeneration resistor of different values at baseband and high frequency respectively. The maximum current for cancellation is limited by the intrinsic emitter and base resistances of the transistors, which are included in $R_{e,BB}$.

[1] n is the effective turns ratio, which is determined by the physical turns ratio, N, and the coupling coefficient k_m, according to the formula $n = N/k_m$.

[2] To be more precise, Z_{in} is determined by R_F, which is the emitter resistance at the fundamental frequency F. In a wideband design $R_F = R_{2F}$ though.

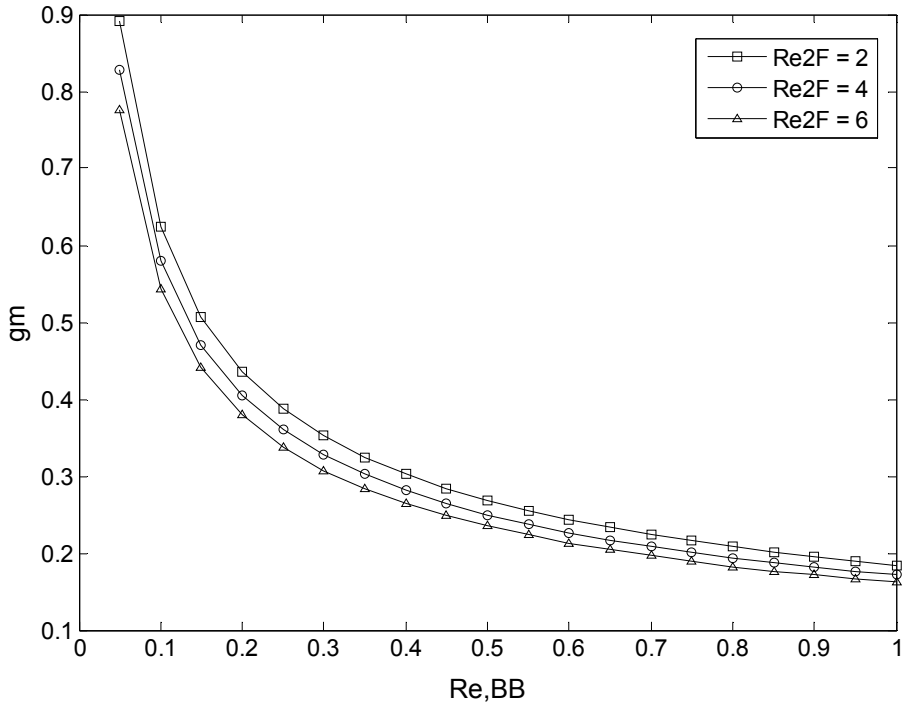


Fig 3.17: Plot of g_m versus the baseband emitter resistance $R_{e,BB}$. $\beta = 100$, $R_S = 50\Omega$.

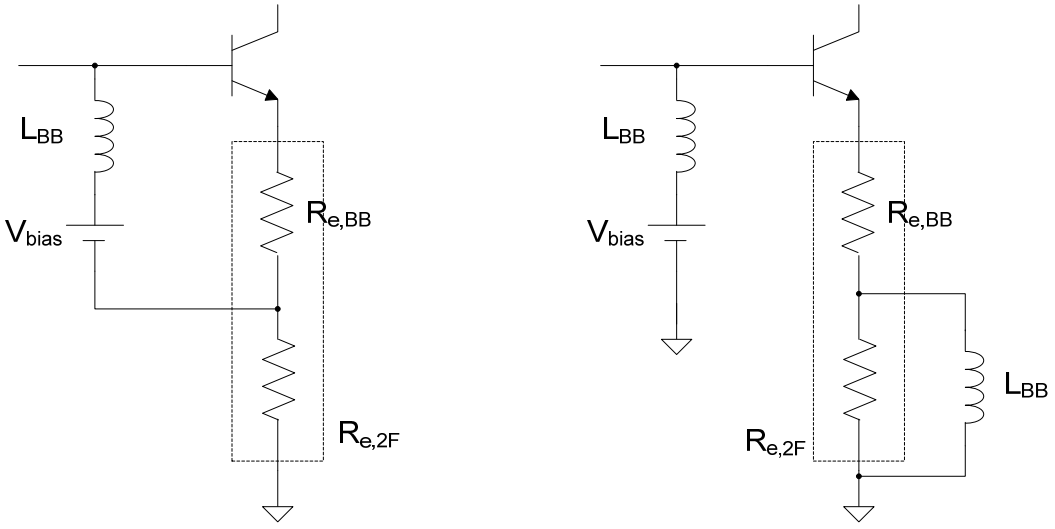


Fig 3.18: Two implementations for R_e, if $R_{e,BB} \neq R_{e,2F}$. L_{BB} must be a short for the baseband harmonic and an RF choke for the RF signals

3.4 Design Example

A schematic of an LNA that embodies all the points discussed so far – shunt-series negative feedback through the use of a transformer; harmonic cancellation by means of C_{bc} tuning and distinct baseband and high-frequency degeneration resistor – is shown in figure 3.19.

The solution presented in figure 3.18.a is preferred over the one in 3.18.b because it shows less sensitivity to the actual baseband impedance value of L_{BB}. For example an L_{BB} of 50nH has at

1MHz (1MHz being the baseband frequency of the second order component, $f_2 - f_1$) an impedance of 0.3Ω. Placing this impedance on the emitter side (like in figure 3.18.b) will affect the total impedance seen by the second harmonic (R_{BB} as given defined by equation (3.26.a)) to a greater extent than if the same impedance were placed on the base side (figure 3.18.a), where it will be divided by β when gyrated to the emitter (as shown by the same equation (3.26.a)).

Since we want to design an LNA for basestation applications, it is desirable that the DC collector current be as large as possible. For this reason, $R_{e,BB}$ from figure 3.18.a is entirely determined by the intrinsic emitter resistance (0.04Ω for a transistor area, A = 10) and was not added to figure 3.19 since it is part of the transistor Q_1.

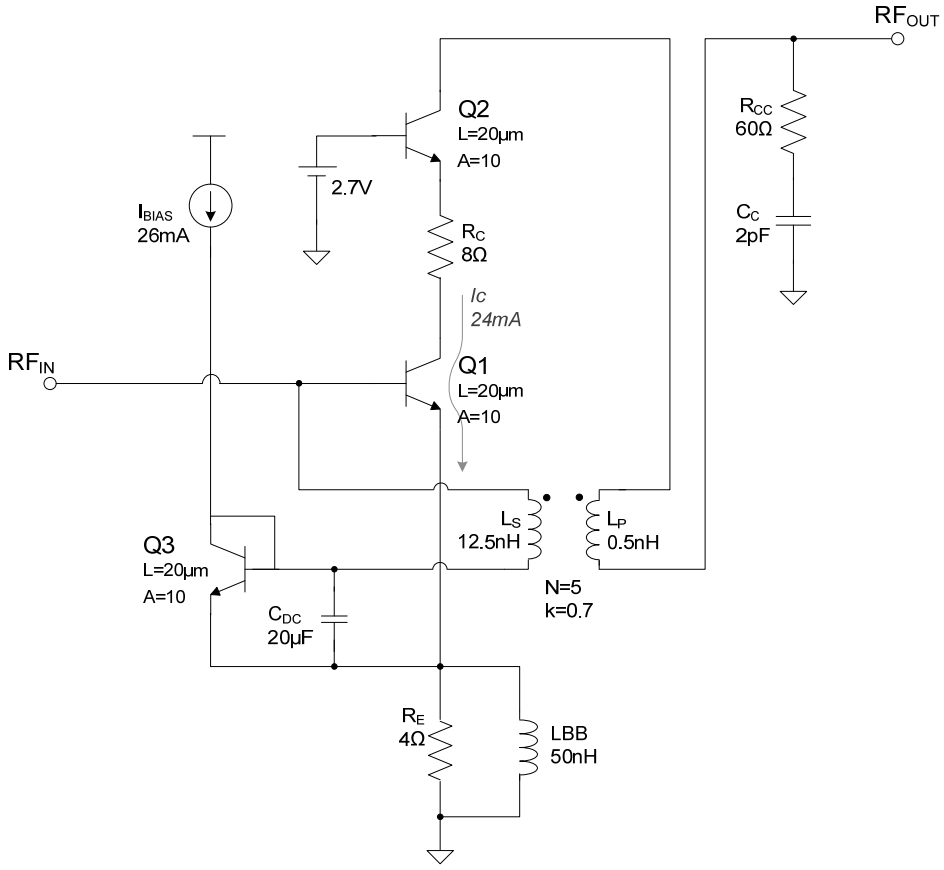


Fig 3.19: Schematic of the proposed LNA

In addition to figure 3.18.a a baseband short (L_{BB}) is placed in parallel to R_E to neutralize any voltage drop caused by the current of the baseband harmonic, which can create asymmetry between the IM_3 components. The voltage source shown in figure 3.18.a is replaced by a diode-connected transistor (Q_3) and a decoupling capacitor, of high enough capacitance to act as a short for the baseband harmonics.

To simulate the above amplifier, a trans-impedance amplifier was employed to create a low impedance node at the output of the LNA (see figure 3.20). The supply voltage is fed through the primary winding of the output balun and has a value of 3.7V to accommodate the cascode stage ($V_{CE1} \approx V_{CE2} \approx 1.85$V). In an attempt to further recreate the environment in which our LNA will operate, a capacitor of 1pF was added to each of the balun output to account for the loading due to the gate-to-source capacitances of the mixer.

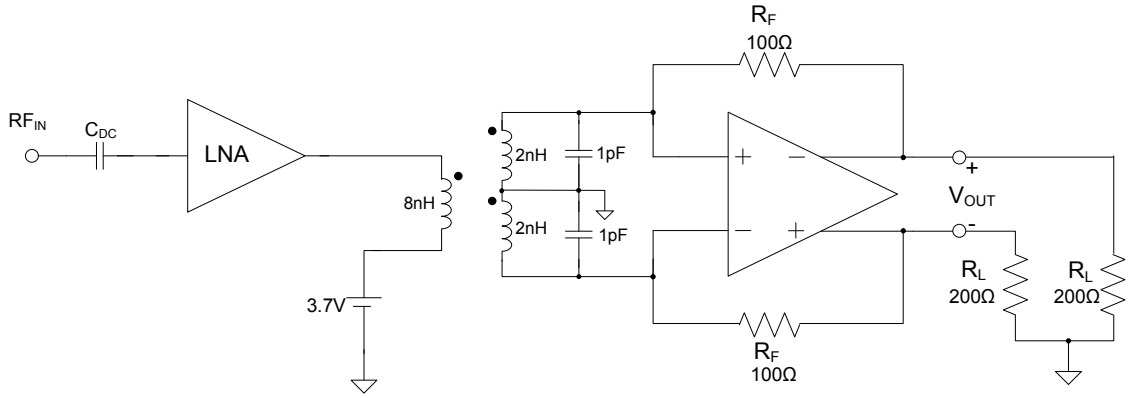


Fig 3.20: The simulation setup tries to reproduce the conditions in which the LNA of figure 3.18 will operate

Figure 3.19 also shows the biasing conditions: a bias current of 26mA was chosen based on the plot shown in figure 3.21, where IIP_3 is shown as a function of the bias current for $R_C = 8$ and when no R_C is used. This results in a collector current of 24mA, annotated in figure 3.19. We observe a slight asymmetry between the low-side (denoted by IIP_{3L}) and high-side (IIP_{3H}) 3rd order products, caused as mentioned in [23] and also in the lines above by the use of distinct $R_{e,BB}$ and $R_{e,2F}$. Biasing the device at the point where $IIP_{3L} = IIP_{3H}$ (26mA) is an attempt to minimize this imbalance.

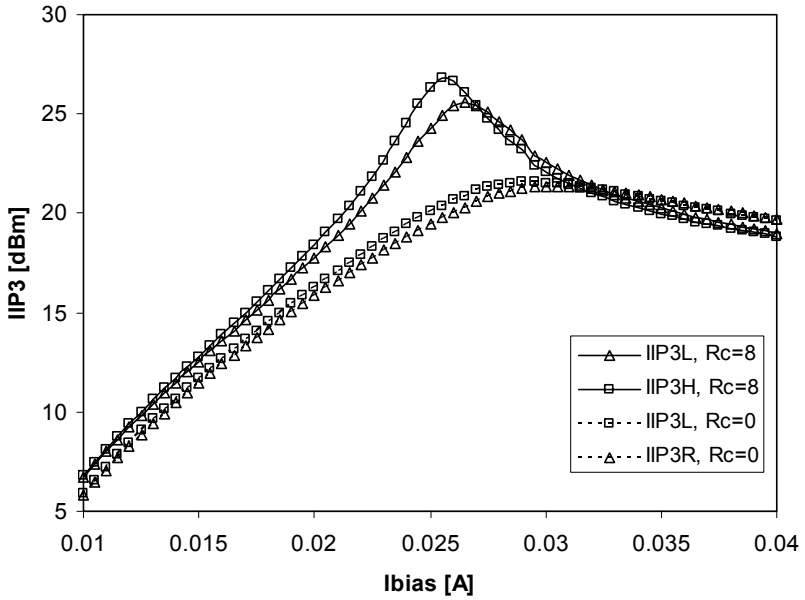


Fig 3.21: The biasing currents are chosen such that IIP_3 is maximized. $f_{RF} = 2$GHz, $\Delta f = 1$MHz.

S-Parameters and large-signal simulations are presented next in figures 3.22 and 3.23 respectively.

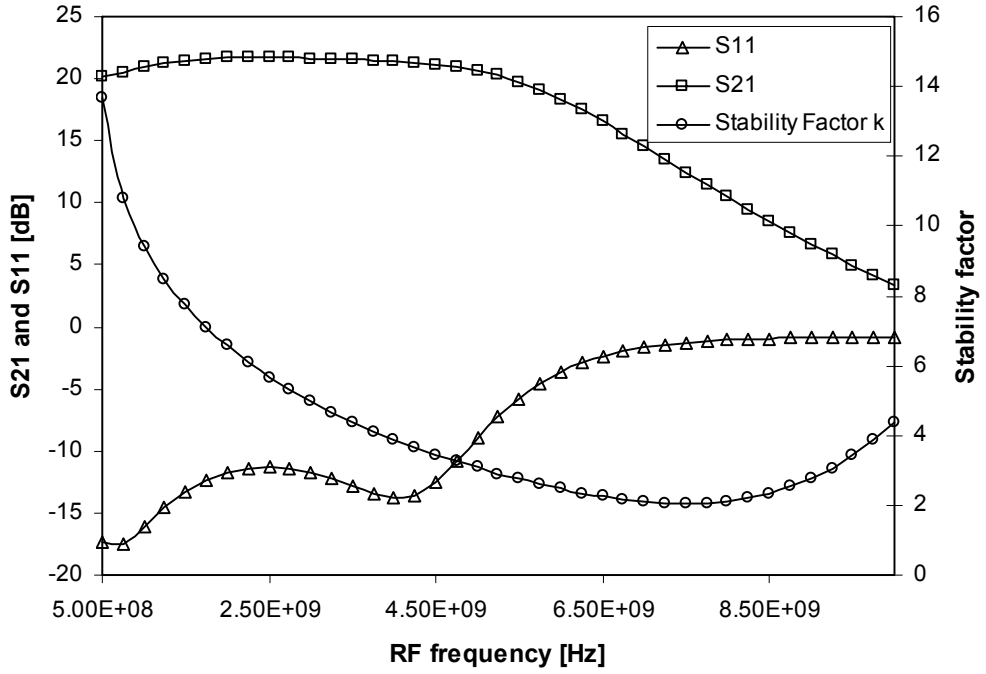


Fig 3.22.a: Small signal simulation results
(gain, reflection coefficient, stability factor)

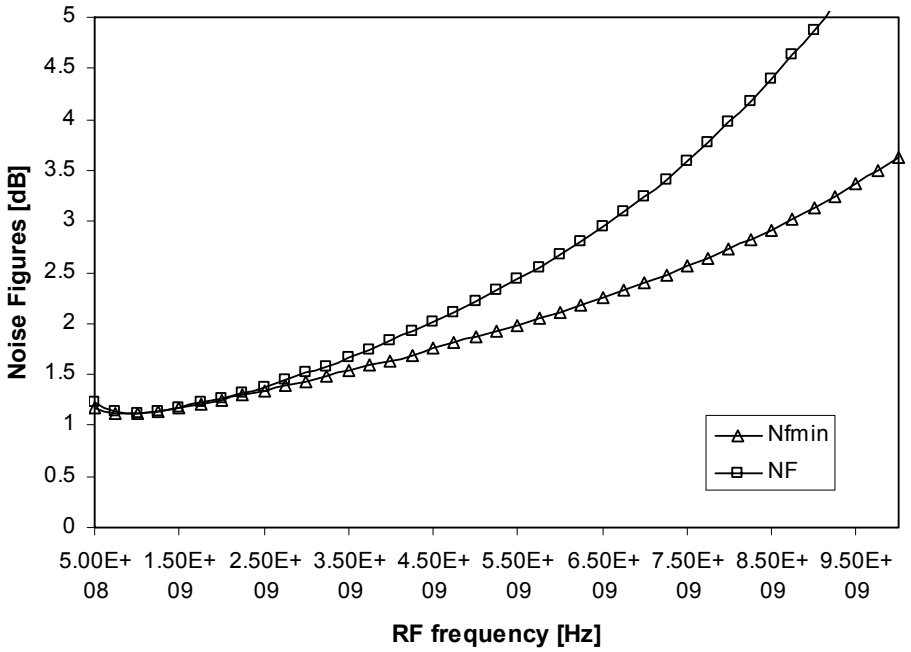


Fig 3.22.b: Small signal simulation results – Noise figures

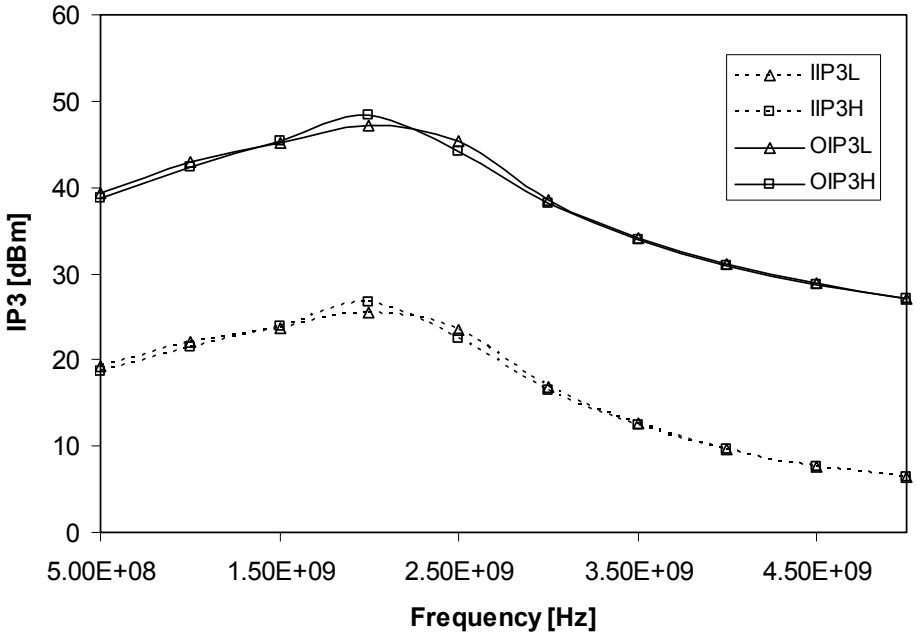


Fig 3.23: Input and output IP3 versus RF frequency

We have been able to demonstrate that low-noise and very high linearity can coexist in the same amplifier. This has been possible through the use of special linearization techniques that allow high values of OIP_3 at reasonably low collector currents. As figure 3.6 and the related discussion make it clear, low currents make noise and impedance matching possible, without putting unrealistic demands on transistors.

Chapter 4

Down-Conversion Mixer

As we have already emphasized on many occasions, as we move further away from the antenna the linearity becomes more and more of a problem. On the same time the requirements on noise can gradually be relaxed. This 'supremacy' of linearity over noise affects the design strategy in a significant manner. The third-order output intercept point is not yet another requirement the mixer designer will take care of sometime in the future, but a fundamental issue that should be considered from the very beginning, even before a topology has been chosen. Sticking to this line of thinking, it is easy to discard the Gilbert mixer (remember at this stage we have not yet selected our LNA): a g_m stage would pose serious limitations when it comes to linearity and bring little improvements to the noise figure because it would be anyway preceded by a low noise amplifier. The natural choice is to avoid another amplifying stage and connect the LNA directly to the stage responsible for the mixing operation itself, i.e. the switching core. The passive CMOS mixer has already an established reputation of being very linear due to the excellent switching performance of deep sub-micron MOS transistors

Figure 4.1 presents the complete top level schematic of our mixer. The inputs – RF and LO ports – are single-ended while the output – the IF port – is differential. The conversion from single-ended to differential signals is implemented using passive baluns. In the case of the RF path, the input balun is placed at the input of the mixers and is therefore connected directly to the LNA's output. Its primary winding is used to provide the supply voltage for the amplifier; the secondary winding is used to bias the input stage of the transimpedance amplifier through the mixer core. The balun used in the LO driver accomplishes the same functions as the RF input balun.

Each of the blocks labeled in figure 4.1 will be extensively discussed in the following sections. We will start with the nMOS switching core, as this is the central block that sets the general course of our design: it determines the necessity of using a transimpedance amplifier – discussed thereafter, and the way the LO subsystem should behave.

4.1 NMOS Switching Core

The mixer core – shown in figure 4.2 – is fundamentally a four-quadrant multiplier. The multiplication is implemented by steering the input RF current alternatively between the positive and negative outputs.

This current domain operation is – as mentioned in chapter 2 – the determining factor in the excellent linearity of this mixer. Keeping the voltage swings all over the circuit as low as possible (ideally no voltage swing) minimizes the distortion due to the nonlinear gate-to-drain and gate-to-

source capacitances and channel length modulation (both phenomena occur if a voltage swing is developed over C_{gd} and C_{gs}, and the channel respectively).

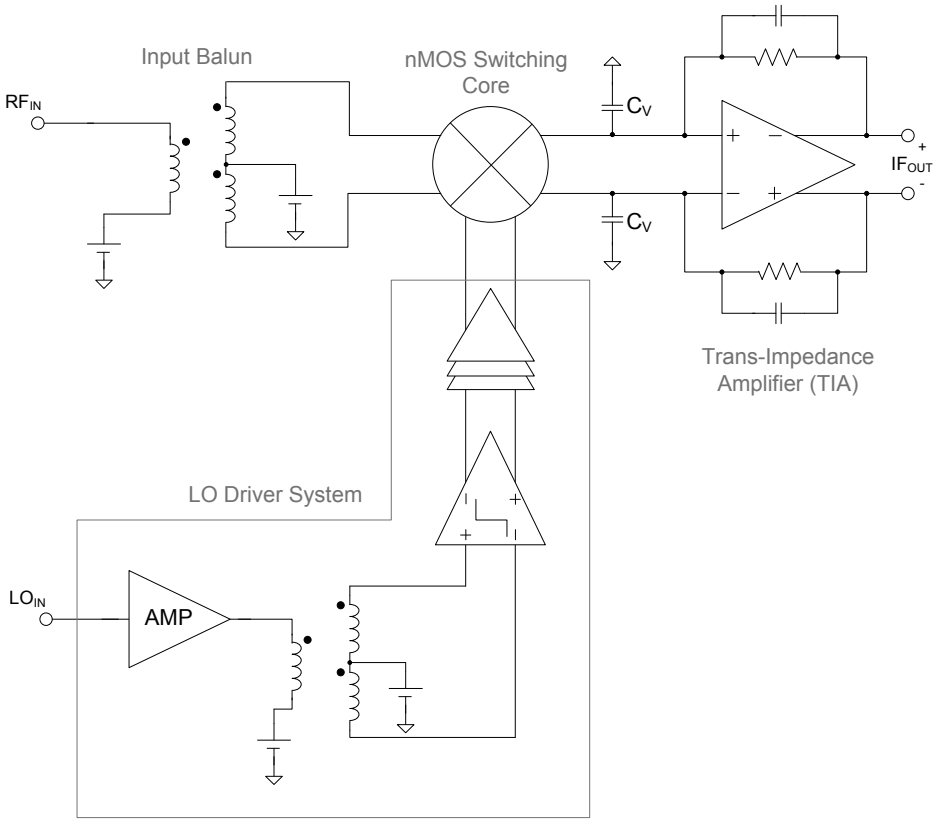


Fig 4.1: Block diagram of the mixer

For convenience, a DC current is allowed to flow through the mixer, which is a noticeable deviation from a passive mixer. This DC current flows into the output terminals of the mixer core (as shown in figure 4.2) and has a value equal to the DC current flowing through the feedback network of the TIA (R_F) minus the bias current of the input pair of the same amplifier.

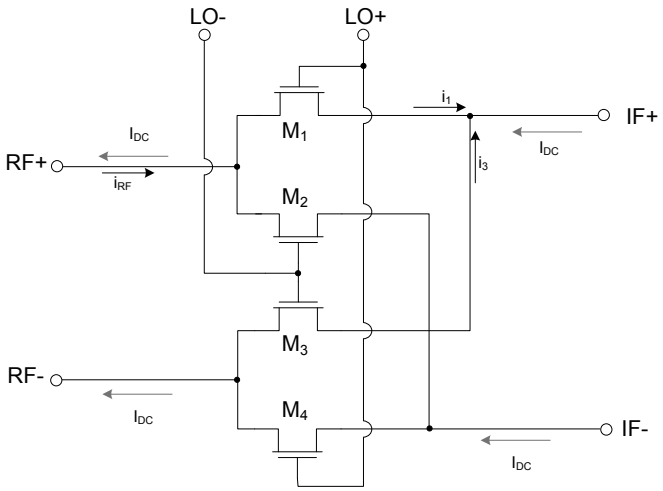


Fig 4.2: Current switching mixer. A DC current flows through the transistors

4.1.1 Sizing

There is a direct compromise between linearity and power consumption: on the one hand devices with a large gate width can provide improved switching behavior (as a consequence of lower channel resistance, R_{ON}), hence higher linearity and lower noise. But on the other hand, larger area devices result in an increased gate capacitance and require a 'stronger' output stage in the LO driver.

Normally, if the transistors are wide enough the limitation in linearity comes from other stages (more precisely the TIA following the mixer) and not from the switching core itself.

4.1.2 Local Oscillator Waveforms

We know very well that a MOS transistor behaves like a switch if it is driven by a square wave, but we need a more detailed understanding regarding the precise requirements on this square waveform. For instance one might ask oneself if this differential square wave should be either overlapping or non-overlapping, and in the latter case what should be the filling factor.

Non-Overlapping or Overlapping

Since the mixing technique we employ here is nothing more than a sampling operation, a designer with a basic knowledge in sample-and-hold circuits would say that the 'clock' wave forms (i.e. LO signal) should not overlap. Otherwise, for a short period of time (the overlapping time) all the transistors would be open resulting in a direct connection between the positive and negative output nodes. This lack of isolation will have negative consequences mainly on gain (not all the current reaches the desired output port) and noise (the noise of the TIA will be visible at the input).

But the main drawback of non-overlapping waves turns out to be related to linearity. Figure 4.3 compares the voltage levels at the RF positive and negative inputs (RF+ and RF- in figure 4.2) in a transient simulation for both cases (non- and overlapping LO waveforms). When the mixer is driven by ideally non-overlapping waveforms – illustrated in figure 4.3.a – we observe unusual spikes in the voltage levels at the RF+ and RF- nodes (the sources of the nMOS transistors – as labeled in figure 4.2). The explanation lies in what happens to those nodes during the non-overlapping period: for this short period of time all four transistors are turned off with the direct result that the input becomes a high-impedance node; pumping current (the LNA's output signal) into a high-impedance node causes that node's voltage to go up to an undefined value. This explanation is confirmed by simulations using overlapping waveforms. In this latter case, since not for a single moment are the transistors off, no high-impedance node is created and thus no sudden jumps in the voltage levels occur. Remember that large voltage swings in our current steering mixer are highly unwanted because they trigger distortion mechanisms related to non-linear capacitances and channel length modulation.

Unwanted phenomena related to the LO signal

The problems related to the presence of a large digital signal in our circuit (the LO signal) are fundamentally caused by the coupling through the gate-source and gate-drain capacitances of the switches. We see from figure 4.2 that the nodes RF+ (and RF-) and IF+ (and IF-) are common mode nodes with respect to the LO signal. In the ideal case no influence from the LO signal should be noticed (this is actually one of the benefits of using a four quadrant mixer). We know from the digital circuits that glitches below the ground and above the power supply levels can occur as a result of switching and imperfections in the LO wave. These glitches lead to common-mode signals, and therefore cannot be neutralized by the four quadrant structure employed here. This is illustrated in

figure 4.4, which mainly shows how glitches and non-zero asymmetric rise and fall times generate unwanted signals in the sources and drains of the transistors.

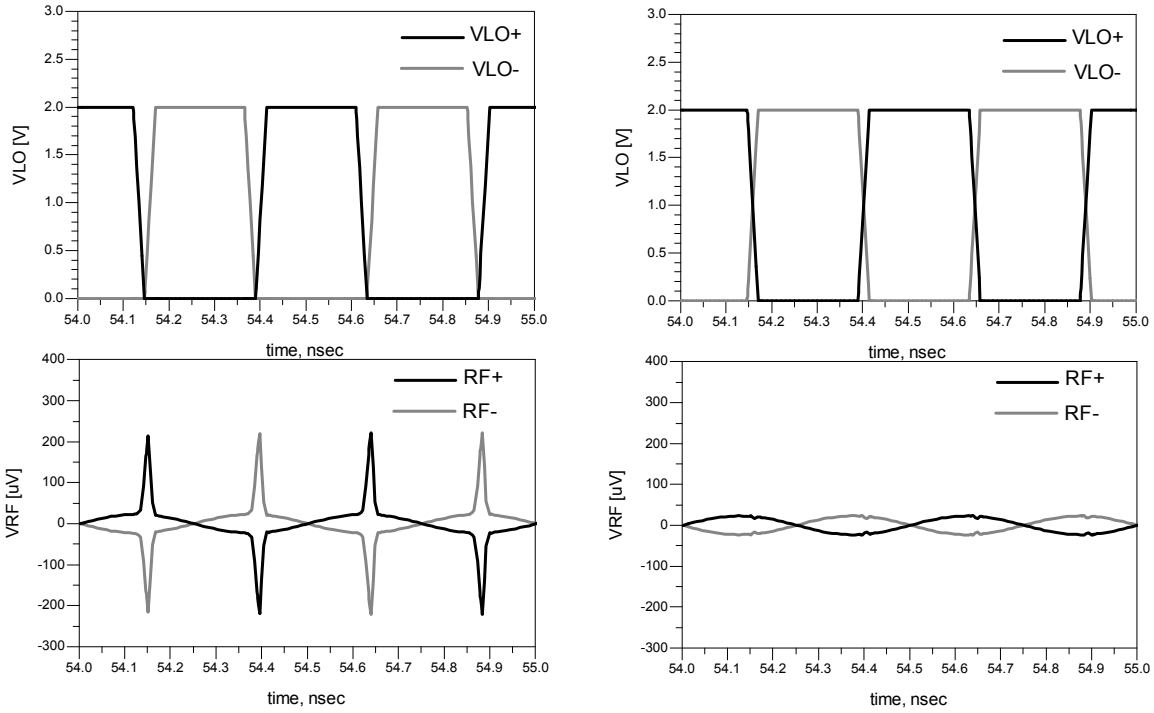


Fig 4.3: Input RF signal of a generic mixer driven by a) ideal non-overlapping, and b) overlapping LO square waves

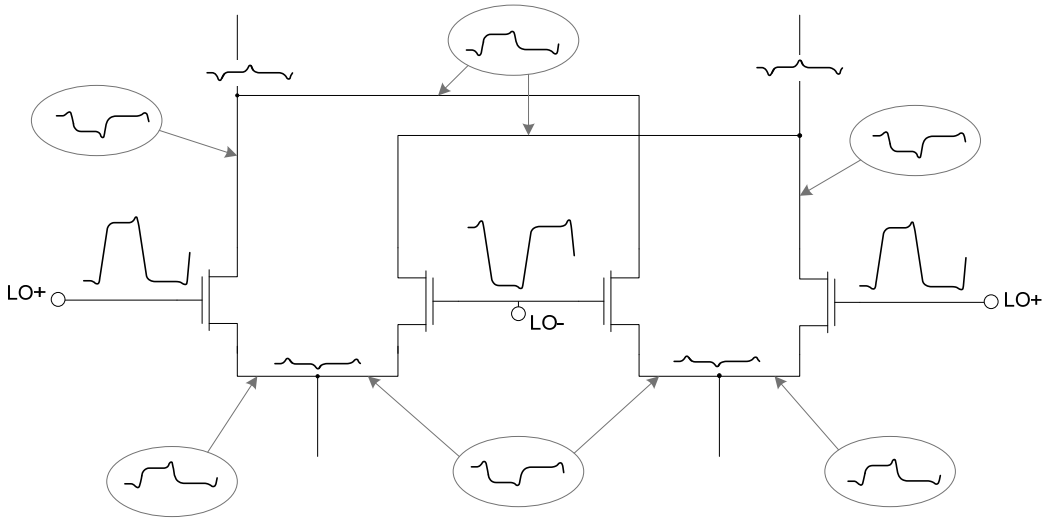


Figure 4.4: Non-ideal LO waveforms and coupling generate common mode signals at the input and output nodes. In clouds are shown would-be signals before summation takes place. The spikes around the supply voltage have different magnitudes compared to those around ground

There are two important consequences that result from this coupling:

- when it takes place via the gate-to-source capacitance, C_{gs}, the coupled LO signal will appear at the input; this will in turn lead to self mixing of the LO signal that will further produce signals at DC and $2f_{LO}$.

- the gate-to-drain capacitance, C_{gd}, will bring forth a common mode voltage swing of frequency f_{LO} at the IF outputs. The impact of this high-frequency signal is twofold: first as we have already made it clear, voltage swing activates distortion mechanisms related to nonlinear gate-to-drain capacitances and channel length modulation[1]; second, it will create linearity problems for the TIA, which as we will see later on, is the limiting block when it comes to the linearity of this mixer.

Usually this coupling issue is tackled by keeping the area of the switches as small as possible; but since we do not want to scale down the switches in order to keep the linearity high, other solutions must be discovered.

In the case of input-coupled LO signal, ensuring a low impedance path for the common mode signal should bring down the resulting voltage level of any such signals. This can be easily achieved through the use of a center-tapped input balun as already illustrated in figure 4.1. The secondary windings of this balun are in fact two coupled inductors, which as we know from [10] create in the ideal case a zero-impedance path for signal when driven in common mode and a high-impedance path for the differential signals.

As for the other side of this problem (coupling due to gate-to-drain capacitance), [28] suggests attenuating the unwanted high-frequency components by means of large capacitors placed before the TIA. This explains the presence of those two capacitors between the mixer and the output amplifier – see figures 4.1 and 4.2. C_v and C_{gd} create in fact a voltage divider (independent of frequency) that attenuates the LO signal that reaches the input of the trans-impedance amplifier. In addition, the R_{ON} resistance of the switches and the same capacitor C_v create a 1st order RC filter for the up-converted RF signals. The attenuation of any undesired signal (all but the IF signal) before reaching the trans-impedance amplifier brings significant improvements to this block in terms of linearity.

As a final word on the coupled LO, since this is a strong common-mode signal (present at the input of the TIA) it will pose stringent requirements on the common mode rejection capabilities of this amplifier.

4.2 Trans-Impedance Amplifier

The output trans-impedance amplifier (TIA) acts as a buffer between the output of the mixer (a low impedance node) and the outside world (usually an analog-to-digital converter). Its block diagram is shown in figure 4.5 while its basic functions are enumerated down below:

- summing circuit for the currents resulting form the mixing process (e.g. i_1 and i_3 in figure 4.5; see also figure 4.2); this summing action requires that the TIA have (ideally) zero input impedance (Z_{IN} in fig 4.5).
- I-to-V conversion through the feedback impedance Z_F; the is an impedance of value:

$$G_Z = \frac{v_{out}}{i_{RF}} = Z_F \tag{4.1}$$

(for notations see figure 4.5 and 4.2);

- filter for the up-converted products: $Z_F = R_F||C_F$ forms a 1st order RC filter that attenuates signals at $f_{RF}+f_{LO}$;

The list of functions above and the discussion about coupled common mode signals carried out in the previous section can help us establish the requirements on the differential amplifier that forms the core of this block:

- high enough loop-gain to ensure very low input impedance;
- good CMRR to withstand large CM signals coming from LO;

[1] This applies to C_{gs} as well

- high linearity, since it is the last stage in the receiver.

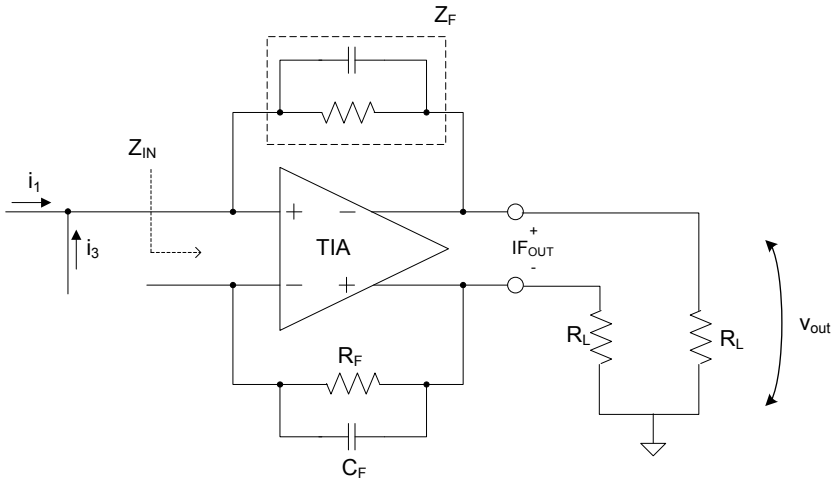
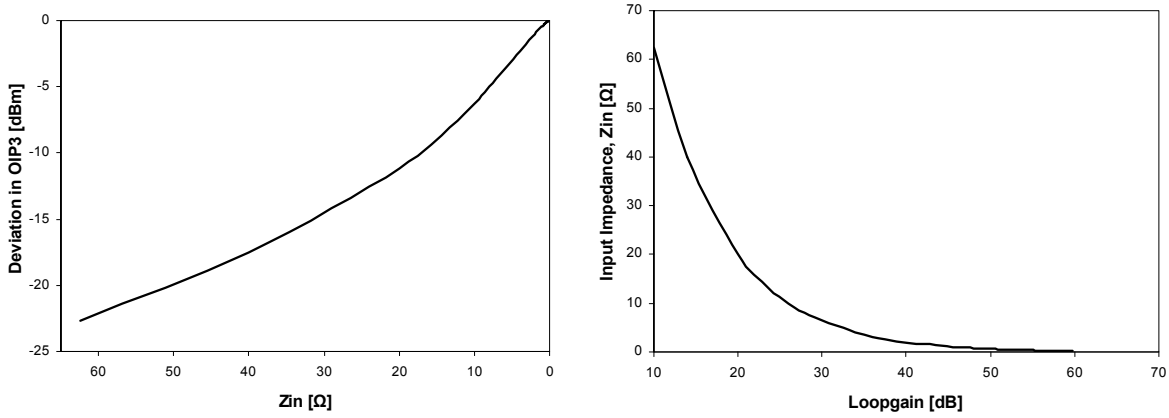


Fig 4.5: Top view of the transimpedance amplifier

4.2.1 Selecting a Topology

Gain

First let us examine to what extent the open-loop gain affects the closed-loop input impedance (Z_{in} in fig. 4.5) and subsequently the linearity of the mixer. Figure 4.6 shows the results of a simulation in which the mixer core was the only 'real' element, while the TIA and LO driver were taken as ideal blocks. In this manner we can investigate the impact the input impedance has on the linearity of the switching quad without its being obscured by e.g. the differential amplifier's linearity. Figure 4.6.a indicates that as long as the input impedance is below 1Ω no sensible deterioration in OIP_3 is noticed; figure 4.6 shows that in order to achieve this number, the loop-gain should be higher than approx 45dB.



(a) (b)

Fig 4.6: The impact of the loop-gain over input impedance (left), and over OIP3 (through Zin) shown (right) as relative deviation from the ideal case of infinite loop-gain

A low frequency, an open loop gain of 45dB can normally be achieved by a single-stage amplifier, if buffered from the load and feedback resistances (otherwise the strong loading effects of these low value resistances will bring the loop gain down). This decision is endorsed also by the

constraint regarding the bandwidth of this amplifier – approximately 500MHz – which more or less rules out a two-stage design.

Linearity

Good common-mode rejection ratio and fully differential structure are the two reasons why linearization techniques based on harmonic cancellation cannot be used here. [23] does show a differential pair that employs out-of-band cancellation but it has two drawbacks that make it impossible to be implemented in our circuit:

- common mode rejection is ensured by input and output baluns which work only at RF frequency, and not in the frequency range we are dealing with here.
- it uses a low value resistance in the common emitter node. This again goes against our desire of having a high impedance element (i.e. current source) that can dramatically improve the CMRR.

As a conclusion to this discussion, the only method of linearization consists in increasing the collector current, whose benefits are twofold: on the one hand it directly improves the linearity of a BJT [24], and on the other hand it increases the loop-gain, which results in the minimization of distortion.

4.2.2 Circuit Implementation

Fully differential amplifiers require additional circuits in order to set the common mode levels of the internal nodes. These circuits – called common-mode feedback (CMFB) amplifiers – are therefore in charge not just with setting the DC levels but also with stabilizing any common mode variations that can occur within the useful bandwidth. Figure 4.7 shows two ways of implementing this common mode feedback and their corresponding practical implementations [25, 29].

To understand the necessity of CMFB let us for the moment ignore the CM sense circuit shown in the circuits on the left hand side of figure 4.7. In order to have a correct DC output level, e.g. $V_{DD}/2$, the current provided by I_{load} must match perfectly to the collector current of the input transistors. Even the minutest variation in any of them will send the output DC level to either V_{DD} or V_{GND}. This happens because the output is a high impedance node. This reasoning applies not just to the DC but to any common mode signals. We consequently say that this type of circuits has no immunity against common mode signals. The CMFB amplifiers work on whichever of the two aforementioned currents (I_{load} and I_C) by adjusting them so that they match perfectly.

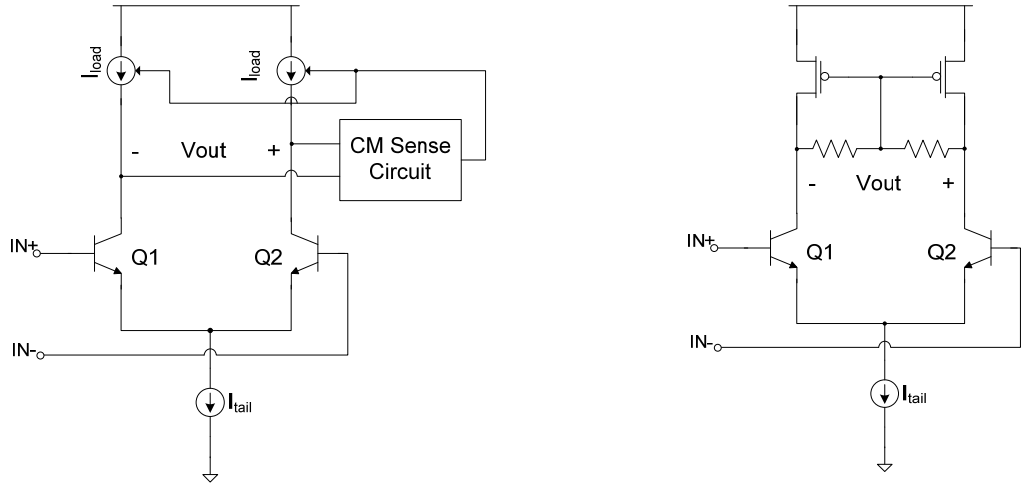
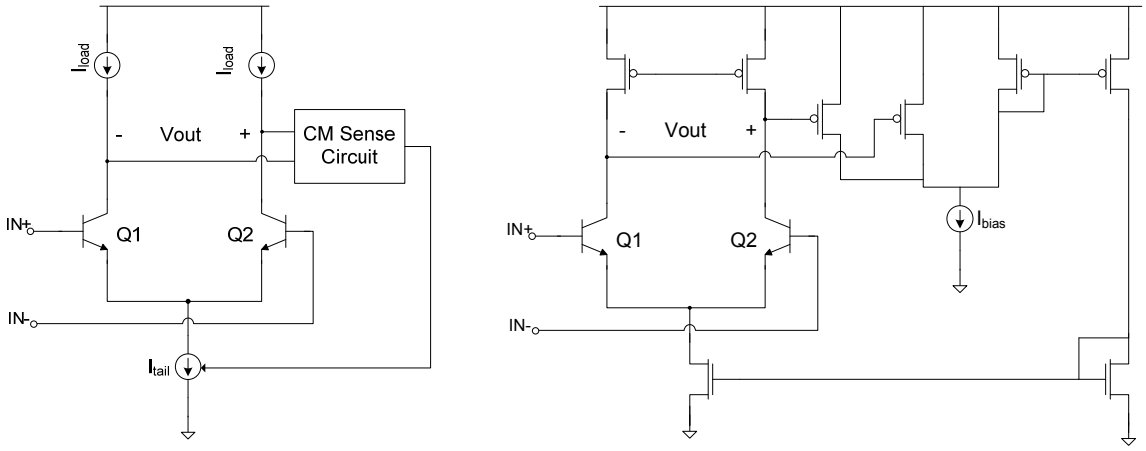


Fig 4.7 (a) CMFB applied to the active loads



(b) CMFB applied to the tail current source

Fig 4.7: Two consecrated ways of employing CMFB (left) and possible practical implementations (right)

The right hand side column in figure 4.7 shows common implementations of the sense circuit and current sources. The pMOS transistors in figure 4.7.a are active loads for the differential pair formed by Q_1 and Q_2, but behave as amplifiers for the common mode signals. This is probably the simplest way of implementing a CMFB circuit.

The biggest disadvantage of the diff-amps shown in figure 4.7 (and in fact of any differential amplifier that employs CMFB) is the increased number of devices that play an active role in the general functioning of the circuit, i.e. amplification. The conclusion is that any common mode feedback should by any means be avoided.

The amplifier shown in figure 4.8 can be considered a step back in complexity when compared to the circuits in figure 4.7 that use active loads, but actually offers the solution to the abovementioned distortion issue.

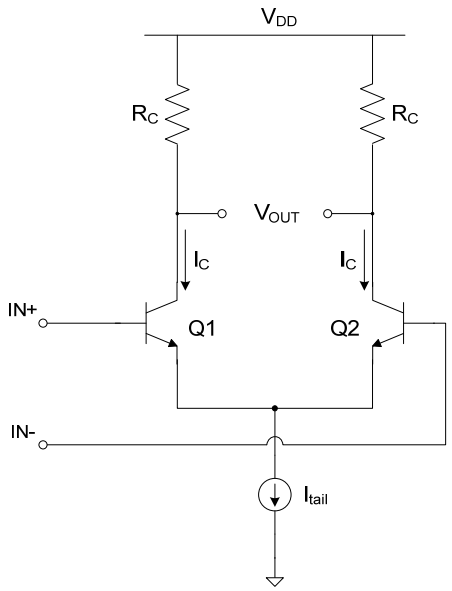


Fig 4.8: Passive loads offer a simple way of setting the CM levels

The output common mode level is unambiguously set to $V_{DD}–I_{tail}R_C/2$ [25]. The gain of this differential pair is not necessarily much lower than in the case where active loads were used. This is because at high current levels the output resistance of the MOS transistors is anyway reduced. Add to

this the fact we ought to use minimum channel lengths (in our case 0.25μm) and we end up with output resistances of a couple of kΩ. We have already determined a minimum open loop gain necessary to ensure small input impedance of approx 45dB, and decided upon large collector current to guarantee good linearity (e.g. I_C = 30mA is probably a good estimate). We can now figure out what the minimum value of R_C must be:

$$A_{OL} = g_m R_C = 45\text{dB} \cong 180 \underset{I_C = 30mA}{\rightarrow} R_C = 150\Omega \tag{4.2}$$

150Ω is really an easily implementable value. The problem though comes from the voltage room available: the drop over this resistance will be of 4.5V! Much too great a number for a 5V supply voltage, especially when 150 is just a starting point and much higher values will be used to account for any variations that may occur.

The solution is to offer different currents for the input transistors and the load resistors: a large collector current for the input pair to maximize the linearity and low currents for the loads to minimize the voltage drop. This is done by diverting a great deal of I_C away from the passive loads, by means of what is called 'bleeding sources' used in Gilbert mixers [30], as shown in figure 4.9. The output voltage is now defined as:

$$V_{OUT} = V_{DD} - \left(\frac{I_{tail}}{2} - I_{bleed} \right) R_C \tag{4.3}$$

and can remain unchanged if R_C and I_{bleed} are simultaneously increased.

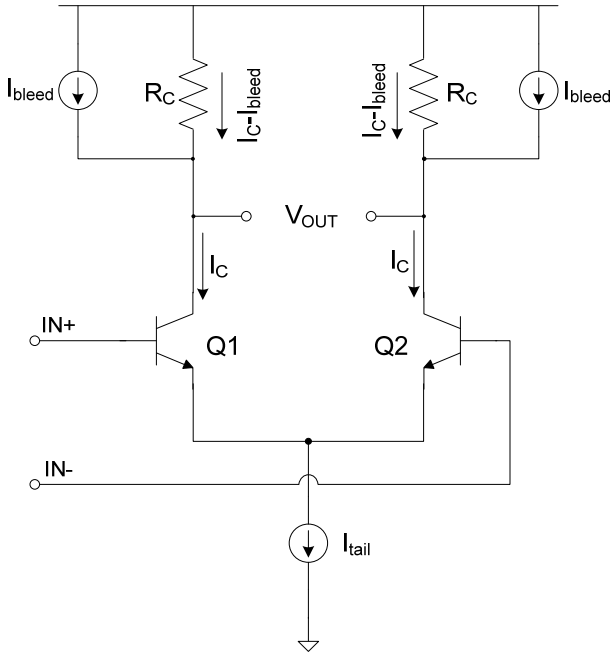


Fig 4.9: Differential pair with current bleeding sources

In the ideal case, the linearity is not affected by the presence of the bleeding current sources since they are used only for biasing. In reality, pMOS transistors will be used as current sources as shown in figure 4.10. They will create distortion through the gate-to-drain nonlinear capacitance (C_{gd}) and channel length modulation. This distortion is however less significant than in the case the pMOS transistors were used as part of a CMFB circuit.

To illustrate the advantage this solution brings, we compare the amplifier from 4.7.b (CMFB implemented using pMOS active loads) with the one shown in figure 4.9 for two cases: ideal and pMOS bleeding current sources. The simulation setup is drawn in figure 4.10: our trans-impedance amplifier is driven by an ideal current source controlled by a 50Ω voltage source.

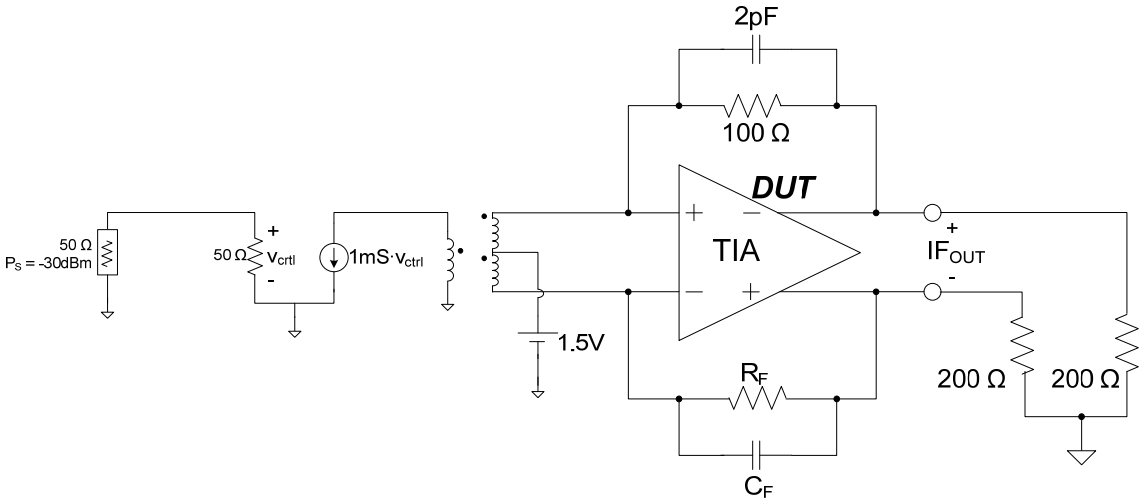


Fig 4.10: Simulation setup used for comparison between CMFB and non-CMFB TIAs

Complete schematics of the three amplifiers under test shown above include an output buffer to isolate the output nodes of the differential pair from the low value load resistors – see figures 4.11. The tail current was set in all the cases to 60mA, while the buffer draws a total current of 80mA (high enough such that its linearity does not affect the overall performance). In the case of the bleeding source amplifiers, the current that flows through each load resistance, R_C (of value 2kΩ), is approximately 1.5mA.

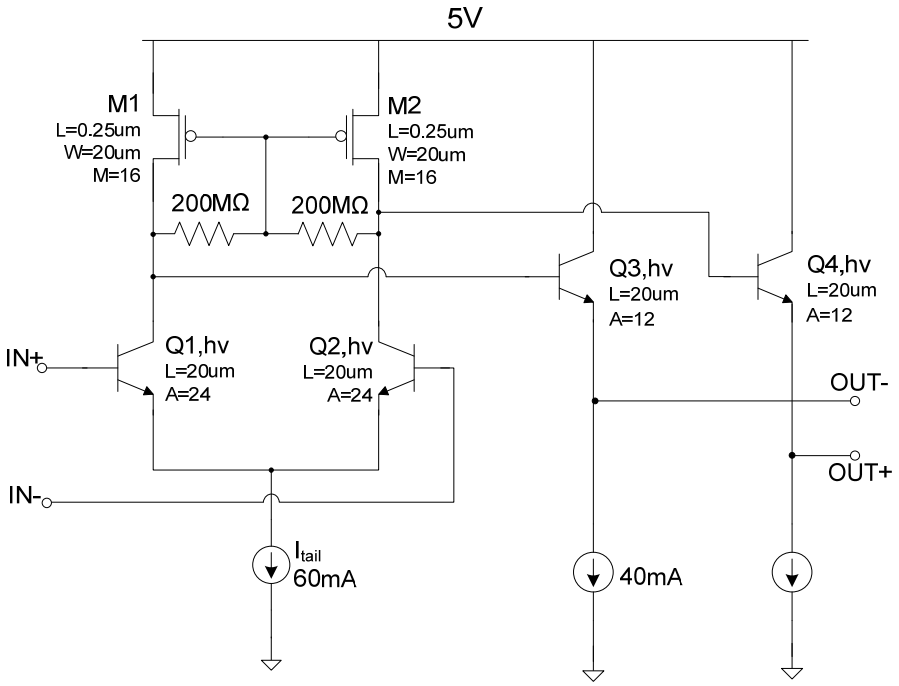


Fig 4.11.(a): CMFB implemented using the active loads

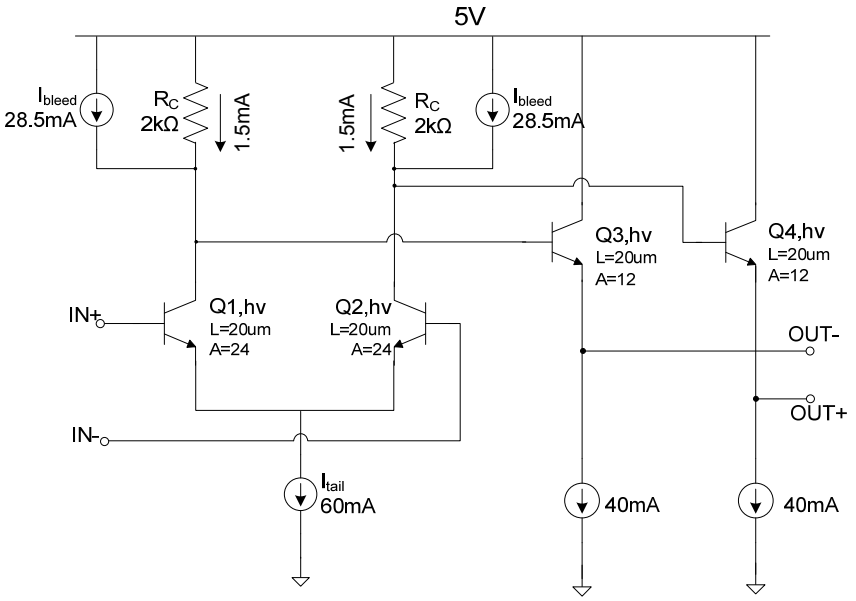
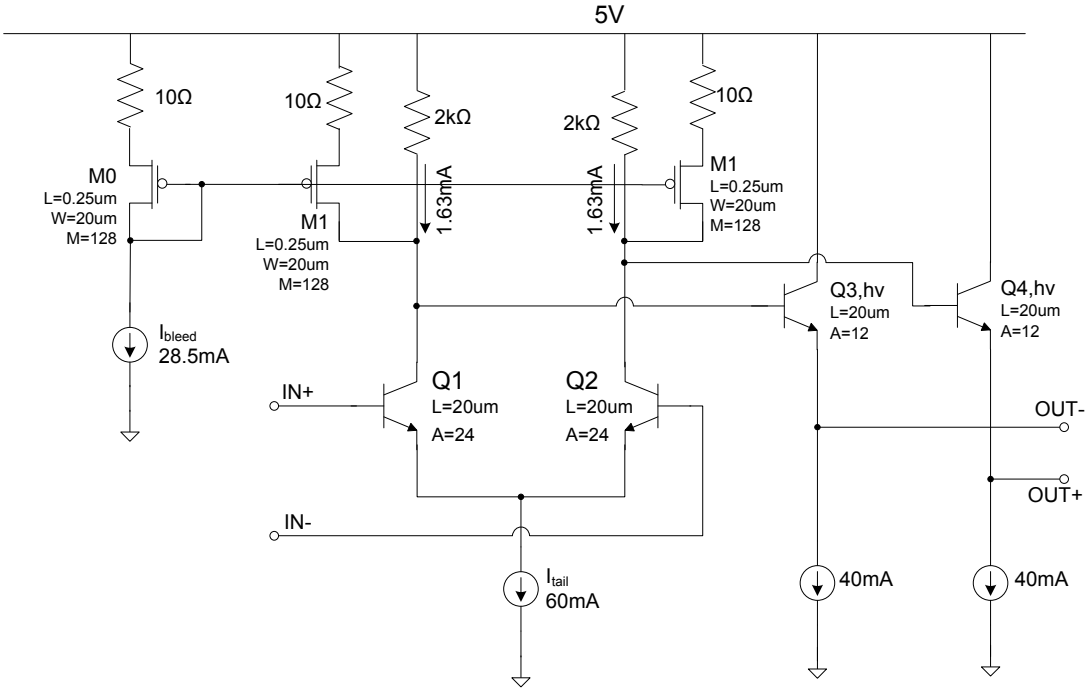


Fig 4.11.(b) Ideal bleeding sources (no CMFB circuit)



(c) pMOS transistors are used as bleeding sources

Fig 4.11: Implementations of the transimpedance amplifier shown in fig 4.10

Table 4.1 summarizes the results by comparing the output 3rd order intercept points. As we can see, avoiding the use of CMFB increases the OIP_3 by 10dBm when compared to the ideal case. As mentioned before, the active loads do contribute to distortion. They degrade the linearity figure by 4dBm compared to the case where ideal bleeding sources were used.

Table 4.1: OIP3 of the circuits shown in fig 4.11

pMOS CMFB – fig 4.11.a	34dBm
No CMFB, ideal bleeding sources – fig 4.11.b	44dBm
No CMFB, pMOS bleeding sources – fig 4.11.c	40dBm

4.3 LO Subsystem

The complete block diagram of this structure is redrawn separately in figure 4.12. Since it mostly deals with square waves, its design is not as demanding as in the case of the analog circuits. The inside and the purpose of each block in figure 4.12 from right to left are analyzed in the next sections.

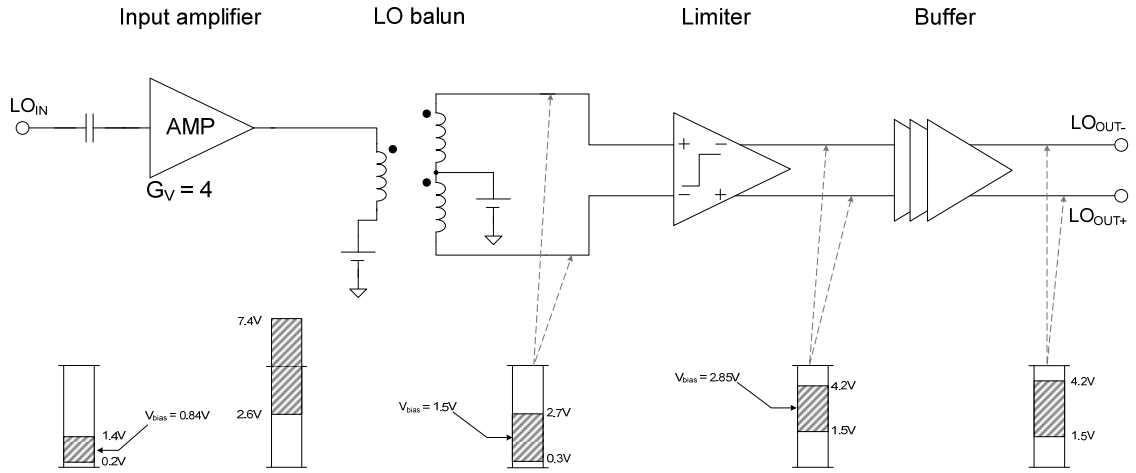


Fig 4.12: LO driver system (the voltage ranges are shown for single-ended signals)

4.3.1 Buffer

A buffer stage is used to drive the LO inputs of the switching quad. This is necessary because the limiting stage is unable to deliver good square waves if its load gets too large. This buffer therefore carries out the transition between a small capacitive load – the input capacitance of the buffer itself, and a large load capacitance, i.e. the gates of the nMOS transistors that form the switching core.

Using a chain of inverters, as shown in figure 4.13, offers the additional benefit of regenerating the waveforms produced by the limiter.

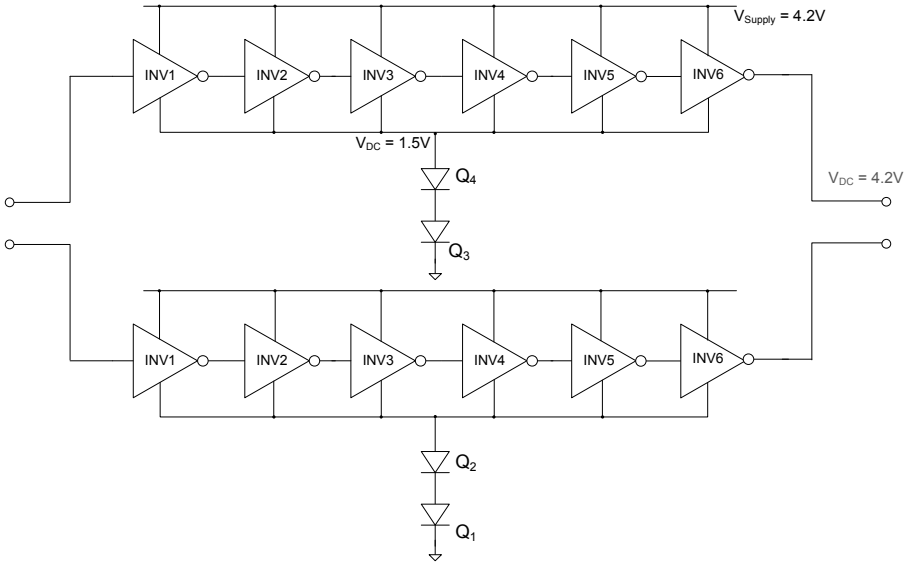


Fig 4.13: LO buffer made up 2 chains of inverters (6 each).
Concrete figure for W and L are shown in chapter 5

The digital circuit design has a well established theory regarding the sizing of such chain in order to minimize the delay. This can be used as a starting point in our design, but further optimization should take care not of the delay, but of the squareness of the wave. This is so because we are interested in improving the switching behavior of the mixer's transistors and not in transmitting a digital signal across the LO path.

A level shifter – implemented in figure 4.13 by two diodes connected in series – brings the lower supply rail – the '0' logic level – to a level of approximately 1.5V. This is needed as the drains and sources of the transistors inside the mixer are biased at this voltage value. The upper limit of 4.2V (shown in figure 4.1) is determined by the source-to-drain punch-through voltage (2.75V for our process). To keep the consistency, the input voltage (i.e. output of the limiter) must be in the same range.

4.3.2 Limiting Amplifier

The limiter is the first step in generating good square waveforms. It eases the demands on the subsequent inverters by generating square wave of quick transistions. The schematic, shown in figure 4.14 comprises a differential pair responsible for the limiting action, and an output stage that shifts the voltage down by one V_{BE}.

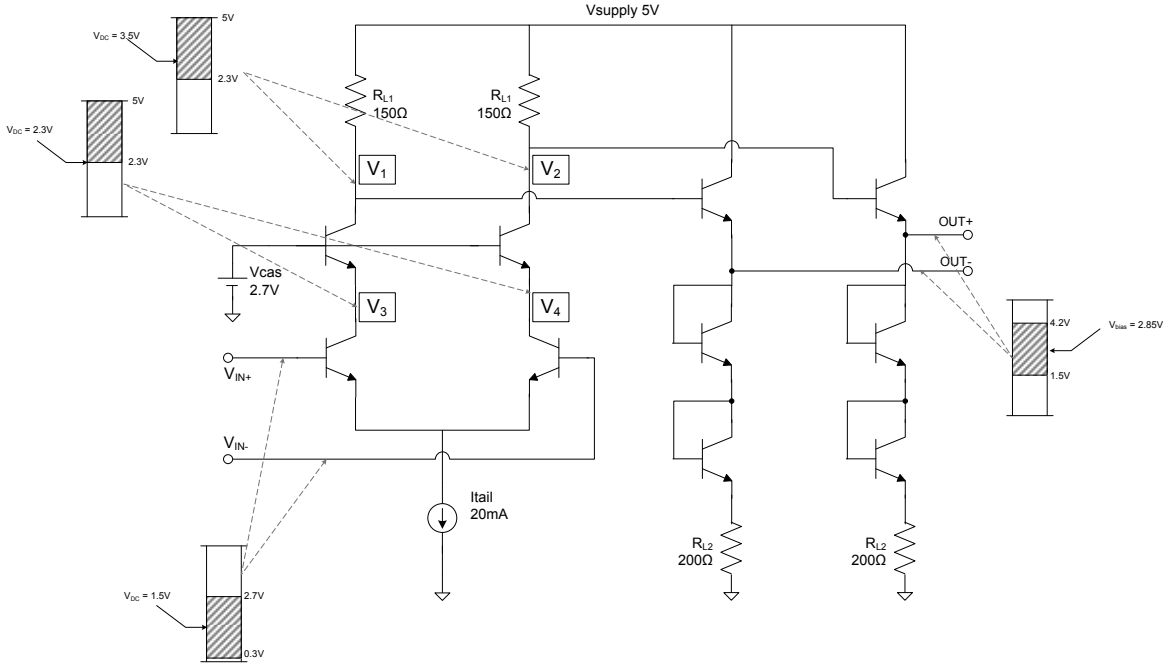


Fig 4.14: Limiting amplifier

The DC output level of the differential pair (V_1 and V_2 in figure 4.14) is set at V_{DD}-$V_{LO}/2$; with a supply voltage of 5V and a maximum allowable swing of 2.7, the DC voltage should be around 3.65V, while the minimum voltage of 2.3V. The role of the cascode transistors is to make sure the collectors of the input transistors (Q_1 and Q_2) are kept at a (reasonably) constant value. The value of this voltage determines the maximum allowable input voltage as

$$V_{in,max} = V_{BC} + V_{3,4} \tag{4.4}$$

This condition states the input voltage must not exceed the collector current by a junction voltage, and therefore prevents the input transistors from entering the saturation region by opening the base-collector junction. We would like to set this voltage as high as possible, but unfortunately $V_{3,4}$ can not exceed the aforementioned value of 2.3V and as result $V_{in,max}$ 3.1V. This is not a real disadvantage,

since the project requirements translate in a voltage swing at the input of the limiter of 0.3…2.7V (this is shown also in figure 4.12). All the above reasoning vis-à-vis the voltage levels is shown in figure 4.14.

4.3.3 Input Amplifier

Because the limiter is by nature a non-linear block (i.e. large-signal operation only), it cannot provide adequate input matching (S_{11} is a small signal parameter). An input amplifier takes care of matching and in addition offers some (voltage) gain that helps the limiter (the higher the voltage swing at the input of the limiter the shorter the rise and fall times of the output square wave will be), and compensates for eventual losses due to the interstage balun.

Figure 4.15 shows a facile way of obtaining both 50Ω input matching and precise voltage gain. A gain of 4 has already been marked in figure 4.12. We derived this value by using the reasoning illustrated in sections 4.3.1 and 4.3.2 and the given specification related to the LO signal. The power at the local oscillator port is bound between -3 and 6 dBm. The corresponding voltage (over a 50Ω impedance) is 224 … 631mVpeak. This explains the values pointed out in figure 4.12.

The requirements in terms of noise (the LO wave finally becomes a digital wave, unaffected by noise) and gain (4 V/V) are not stringent at all. The voltage gain of this inverting amplifier goes by the well known formula for inverting amplifier:

$$G_V = -\frac{R_2}{R_1} \tag{4.5}$$

With R_1 already set to 50Ω, the value of R_2 is constrained to 200Ω for a gain of 4.

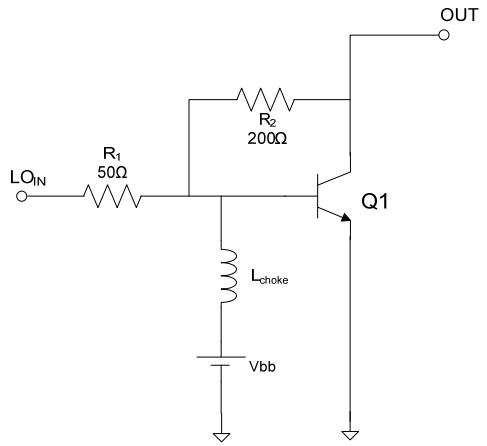


Fig 4.15: LO input amplifier

4.4 Conclusion

The central meaning of this chapter was to show that beyond a certain point, high linearity figures become a question of details. We started from consecrated circuits – e.g. passive mixer – and techniques – the use large collector currents to improve linearity, but what finally helped us achieve good overall linearity were small fixings: the use of center-tapped balun and large bypass capacitors to remove unwanted common-mode (such as LO-coupled voltages) and up-converted signals before they reach the TIA; bleeding current sources to avoid the employment of CMFB circuits; a chain of inverters made possible the use of large switches in the mixer core.

Chapter 5

Simulation Results

This section presents completely annotated schematics and comprehensive simulation results accompanied by succinct clarifications. It also includes comparisons with industry relevant products, and as conclusion possible measures to improve the current design.

5.1 Circuit Schematics

The complete block diagram is redrawn in figure 5.1 for convenience. Details related to the practical aspects of each of the blocks shown in this picture are presented in the next sections.

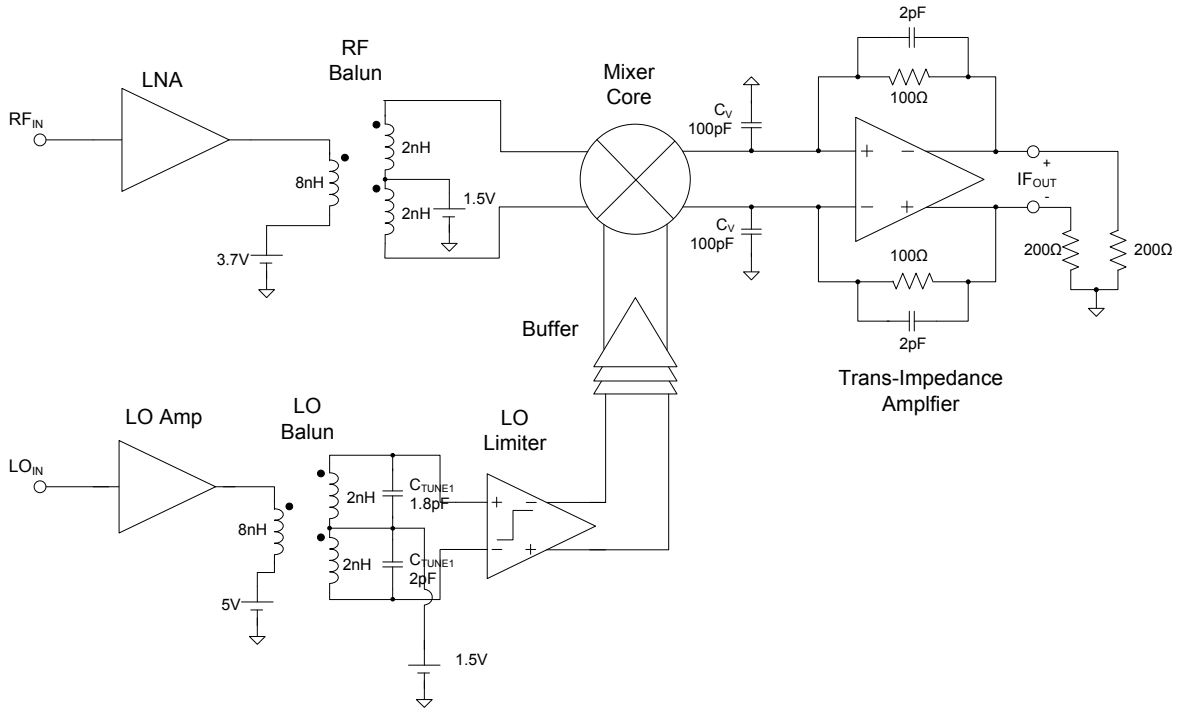


Fig 5.1: Complete top level diagram of the front-end

5.1.1 Low Noise Amplifier

A schematic of the LNA was previously introduced in figure 3.19. The final schematic – shown below in figure 5.2 – uses a Momentum-extracted model for the feedback transformer (shown later on in figure 5.3), and includes the most relevant parasitic elements – like inductances in the input and emitter paths.

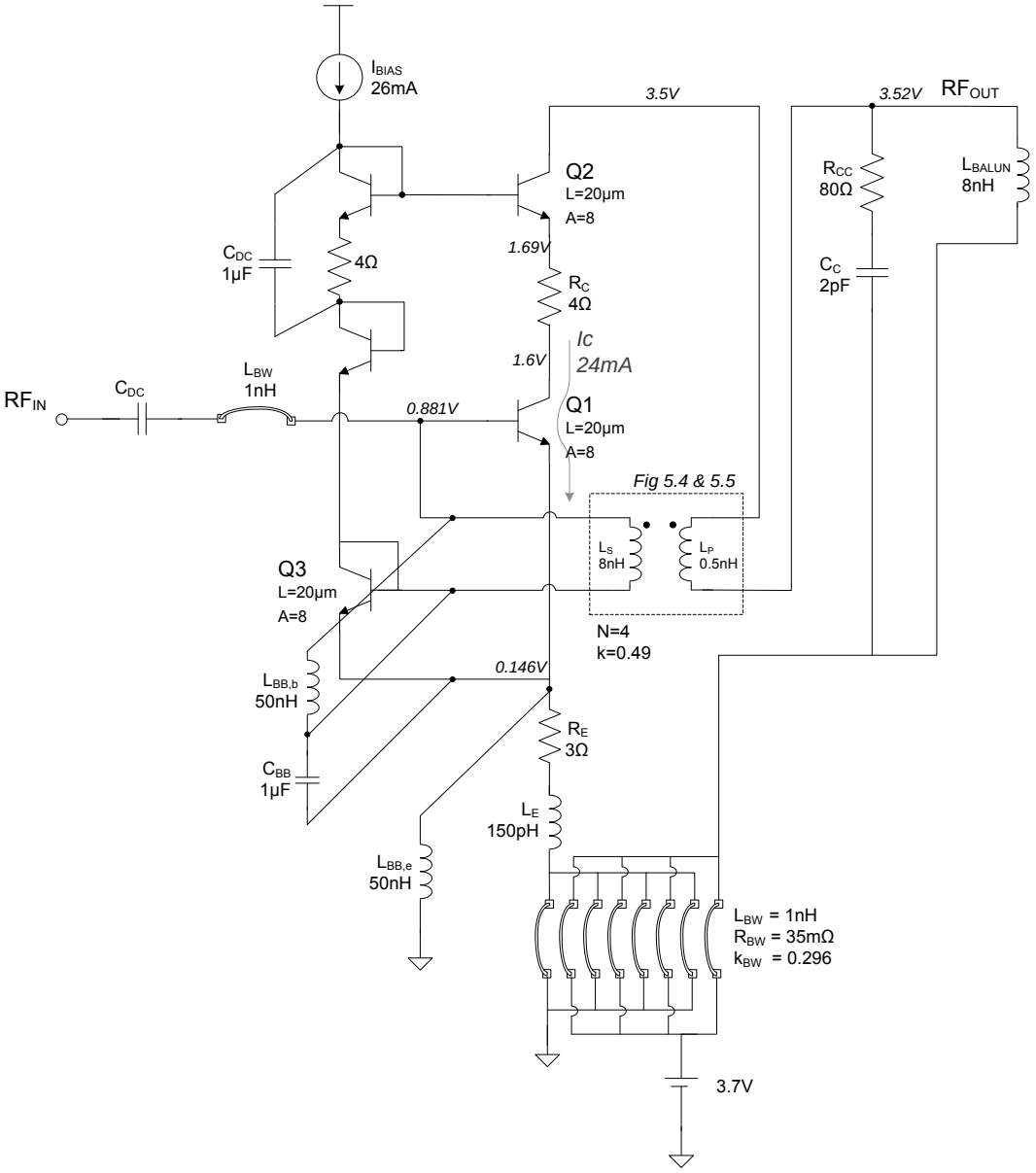


Fig 5.2: Full schematic of the LNA, including bondwires.
The elements connected with oblique lines are off-chip components

Feedback Transformer

The most important limitations are created by the feedback transformer, whose inter-winding capacitance – C_{X1} and C_{o1} in figure 5.3.b – undergo a Miller effect. This will result in significant

losses in terms of bandwidth and input matching if no measures are taken. Although the output of the amplifier is in the current domain, and therefore (supposedly) acts like a low-impedance node, in reality the leakage inductance and parasitic resistance of the RF balun (denoted L_{BALUN} in figure 5.2), and the input resistance of the mixer (shifted to the primary side of the RF balun) will cause voltage swings that will augment the Miller effect. The measure we took to minimize the inter-windings capacitance was to increase the spacing between the primary and secondary wires to 18μm, as the magnified area in figure 5.3.a shows. From the same figure we can also notice that the primary winding is made wider than the secondary one – 10μm compared to 8μm respectively – in an attempt to maximize the ratio L_P/L_S (and consequently the turns ratio).

Bondwire inductance

Generally speaking, the most problematic parasitic for single ended amplifiers is by far the inductance in the emitter or source of the amplifier due to bondwires. Using more wires in parallel, with the hope that this will minimize the total inductance, works up to a certain point – the mutual inductance between two adjacent bondwires will set a lower limit on the value of the total inductance, according to the following formula [31]:

$$L_{BW,eff} = \frac{L_{BW} + M_{BW}}{2} \tag{5.1}$$

where it was assumed the two bondwires have equal inductance, L_{BW}, and $M_{BW} = k_{BW} \cdot L_{BW}$.

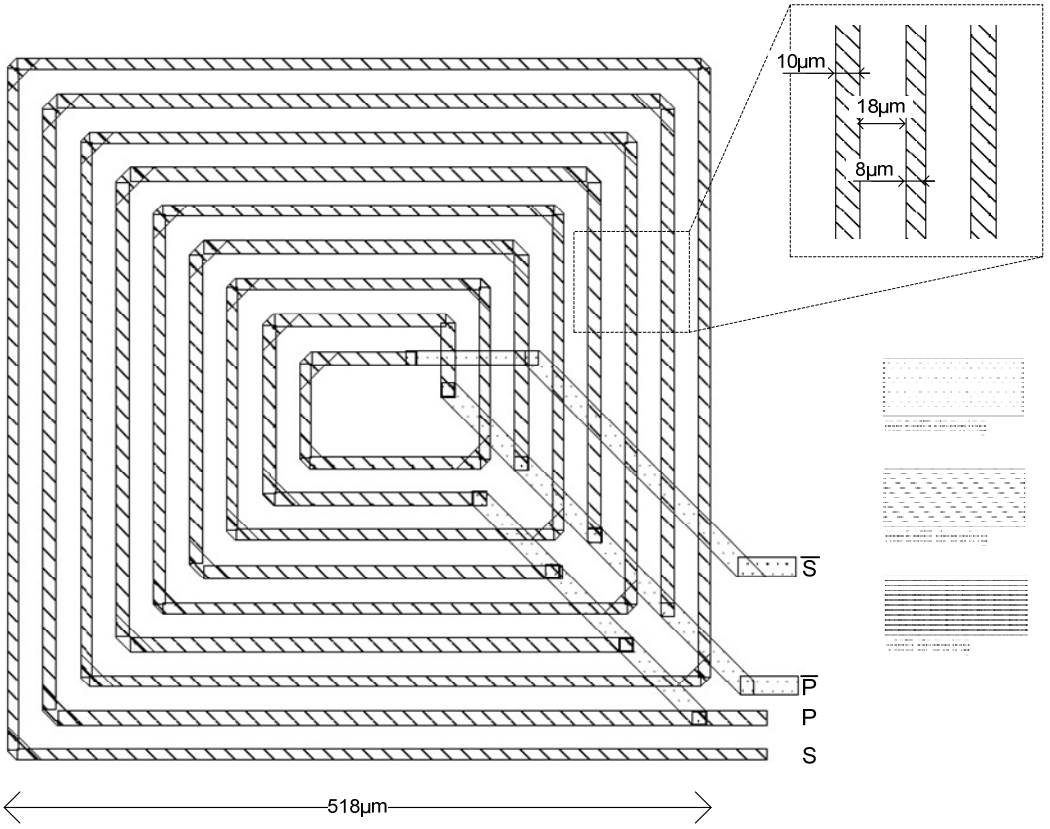
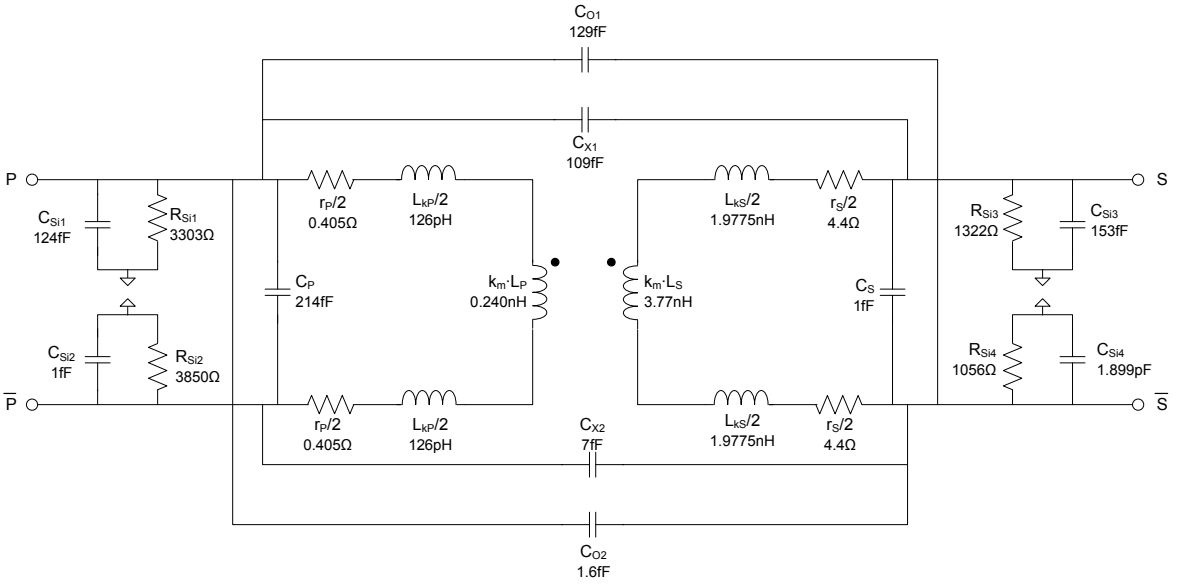


Fig 5.3 (a)



(d)

Fig 5.3. Feedback transformer: a) Layout view, b) Details about the substrate structure;
c) transfer function shows a self-resonance frequency of 10GHz
d) Equivalent circuit for Spectre simulations.
N = 4, k_m = 0.488, L_P = 0.4922nH, L_S = 7.725nH

One way to get rid of the mutual inductance is to interpose wires through which currents of equal signs but opposed polarity flows. In this case the coupling, k_{BW}, and consequently M_{BW}, are negative. This is usually done in differential circuits, where currents are guaranteed to have the same magnitude and opposed directions. What we have done in our circuit – as shown in figure 5.2 – does not fully comply with the above requirements: the currents have opposite directions but not the same magnitude – the emitter current of Q_1 is of course different to (roughly) the collector current of Q_2. The effect is nonetheless significant and works in our favor. The only thing left to do is to minimize the term ($L_{BW} - M_{BW}$) by placing a number of bondwires in parallel (4 for the ground terminal and 4 for the supply voltage in figure 5.2).

As a final note on this topic, the input bondwire inductance has been incorporated in the input matching network and therefore creates no disturbance.

IM_3 *cancellation*

A couple of remarks specific to the implementation of figure 5.2 must be made in addition to what has been said in section 3.3 and 3.4. Although the secondary winding of the feedback transformer is expected to act as a baseband short (L_{BB} in figure 3.17.a), due to the monolithic implementation its parasitic resistance has a significant (unwanted) contribution – 8.8Ω as figure 5.3.d shows. To ensure a proper baseband short, an external inductance, $L_{BB,b}$, is added in parallel with the secondary winding. R_C was reduced to 4Ω (compared to 8Ω in sub-chapter 3.4) to minimize the Miller effect on the base-collector capacitance of Q1. This was necessary because, as mentioned in the lines above, the transformer too has a capacitance that undergoes a Miller effect – situation unaccounted for in section 3.4.

5.1.2 RF and LO Baluns

When used in the RF path, the leakage inductance of the primary winding (of value $(1-k_{1,2})\cdot L_P$) of our balun (shown in figure 5.4) causes an unwanted voltage swing at the output of the LNA, of magnitude

directly proportional to f_{RF}. The parasitic resistance r_P too has a contribution in this respect. Because of that, maximizing the coupling factor, $k_{1,2}$, and minimizing L_P (which will reduce the absolute value of the leakage inductance, as well as the parasitic), were two issues of prime consideration. As figure 5.4 points out, the wire width was set to 8μm (a compromise between area, coupling and parasitic resistance), while the distance between two wires was kept to its minimum, i.e. 3μm. The top metal layer (sixth) was used because of its lower resistance (0.0105Ω/square).

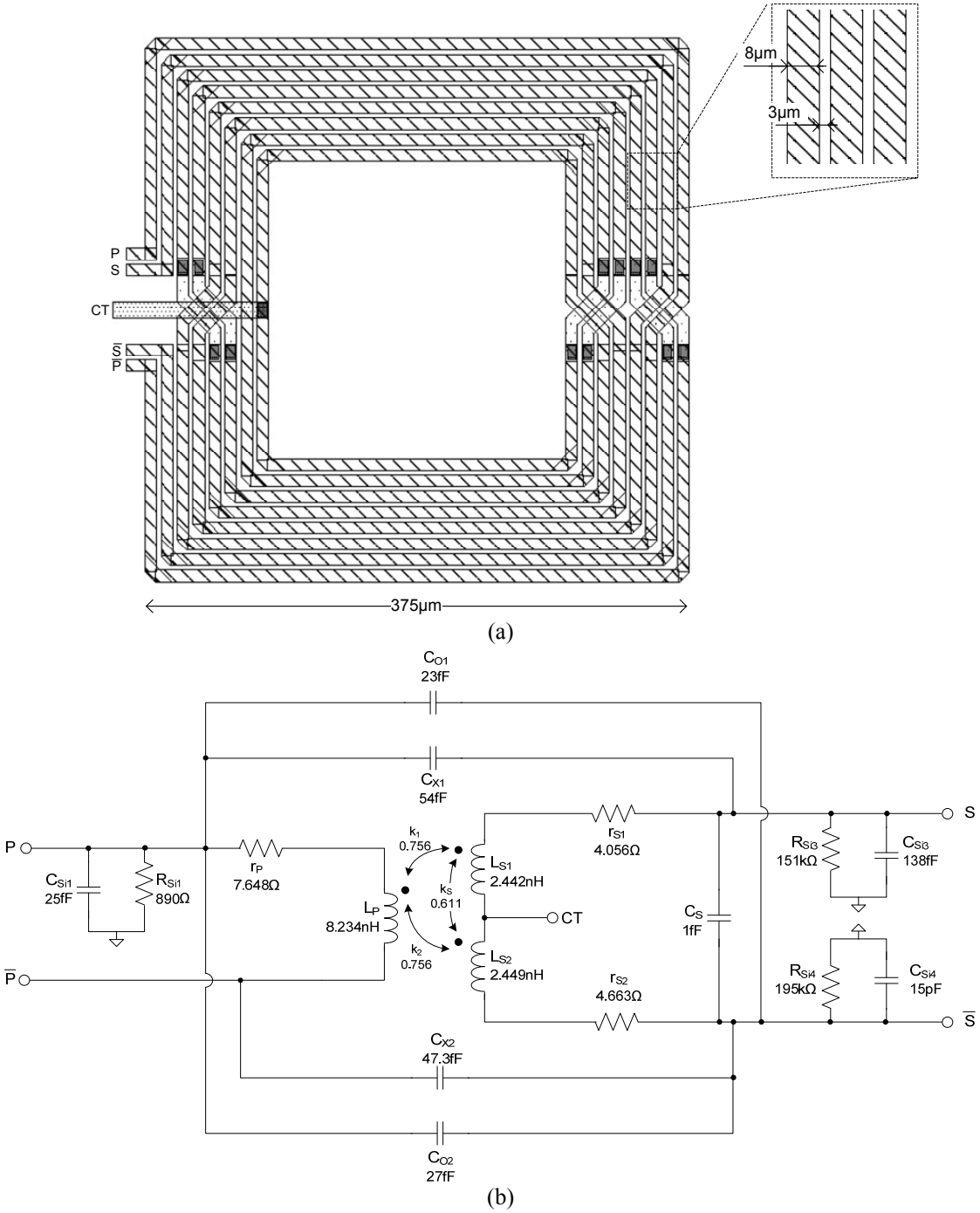


Fig 5.4: RF (and LO) balun: a) Layout view (CT denotes the central tap of the secondary winding); b) Equivalent circuit. Some impedances that model the losses to the substrate are short-circuited to ground and therefore not added to the model (it is the case for ports P' and CT).

Interesting to say, when a balun is used in the current domain, resonating the leakage inductance has (in theory) no advantage. On the contrary, it might be rather disadvantageous. If we drive the input of the balun by an (ideal) current source, the leakage inductance is in series with this current source and therefore has no influence on the current that is transferred from the primary to the secondary. Adding a tuning capacitance in parallel with any of the windings will influence the transfer function since a current will be deviated to ground through the tuning capacitor.

In the case of the LO balun, tuning is required since in this case the balun is employed to pass voltages. As shown in figure 5.1, tuning capacitors have been added to each of the secondary windings. They have different values to compensate for the imbalance in the transfer functions, S_{21} and S_{31}[1], as shown in figure 5.5.

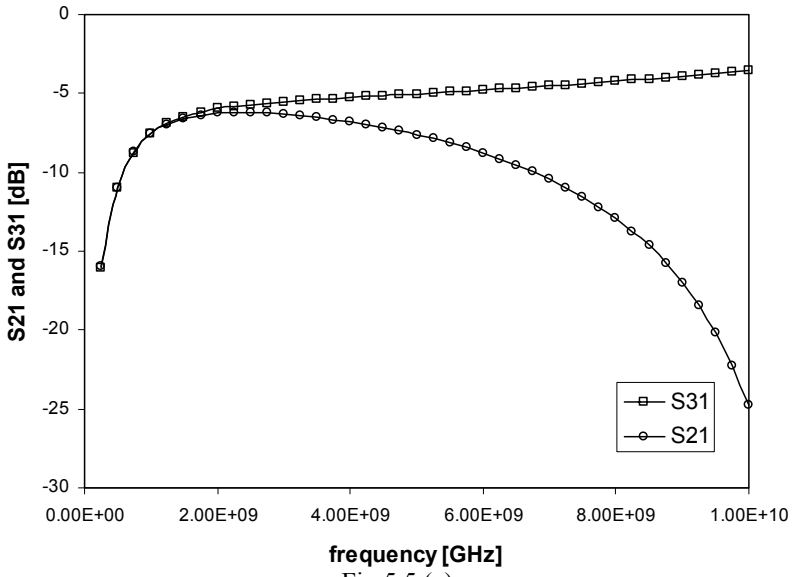
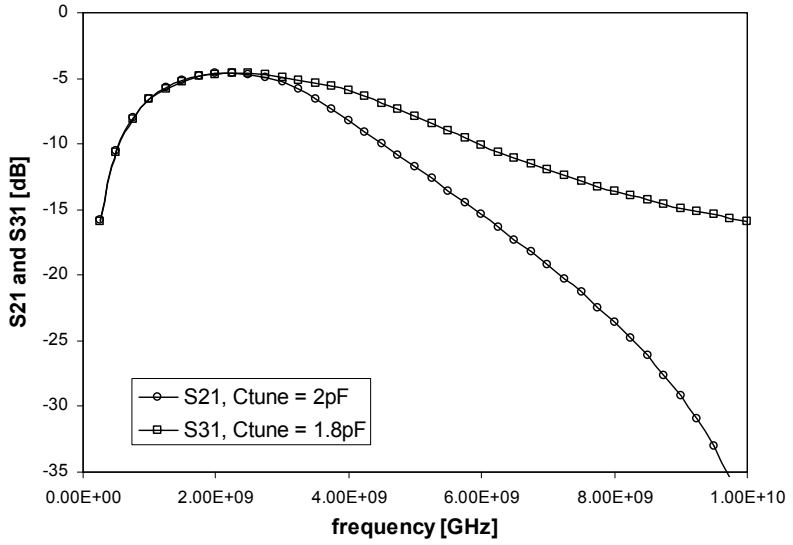


Fig 5.5 (a)



(b)

Fig 5.5: The transfer functions form primary to the two secondaries,
(a) without and (b) with tuning capacitors added

[1] S_{21} is the transfer function of an inverting transformer, while S_{31} is of a non-inverting amplifier. The former has a zero in the transfer function while the latter, by its nature, overcomes this problem and has a more extended bandwidth [10]. This explains why S_{21} rolls off more abruptly than S_{31} does.

5.1.3 Mixer Core

Each transistor in figure 5.6 (sized as shown in the upper left corner) has a gate capacitance of 51fF. This results in a load of 102fF seen by each of the LO inputs. The DC voltage present at the RF inputs is actually (as said before) passed by the mixer to the trans-impedance amplifier, where it biases the bases of the input differential pair. Using the mixer to set the common-mode voltage eliminates any possible problems related to the CM behavior of the TIA.

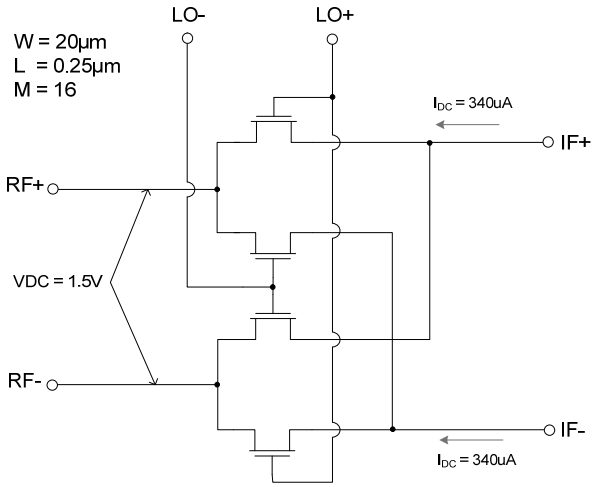


Fig 5.6: Schematic of the mixer

5.1.4 Trans-Impedance Amplifier

The right-hand side of the schematic shown in figure 5.7 corresponds to the amplifier shown in figure 4.11.c. The left side is an actual implementation of the circuit used to generate I_{bleed} (as labeled in figure 4.11.b).

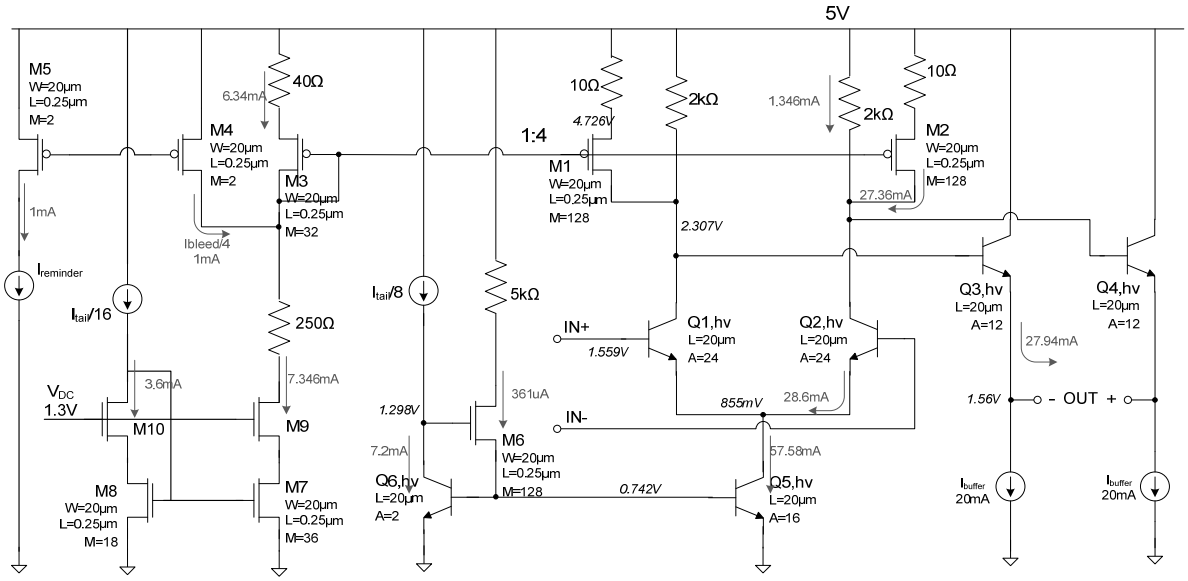


Fig 5.7 : Transimpedance amplifier: schematic of the differential opamp

5.1.5 LO Input Amplifier

Compared to figure 4.15, the LO amplifier shown in figure 5.8 includes the effect of bondwires.

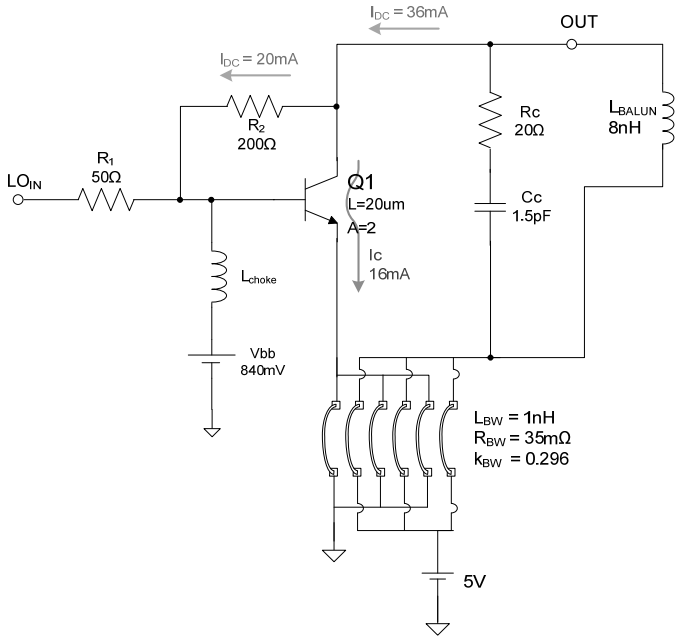


Fig 5.8: Annotated schematic of the input LO amplifier

We applied the same technique as the one showed in figure 5.2, with the difference that instead of using 8, we used 6 intertwined bondwires. The currents that flow through these bondwires do have opposite signs, but their magnitudes are far from being equal. The neutralization of the mutual inductance is less significant than in the case of the LNA, but it does help although to a much smaller extent. The voltage gain versus frequency – see figure 5.9 – drops at frequencies above 3GHz precisely because of the effect of parasitic inductance on the emitter's side.

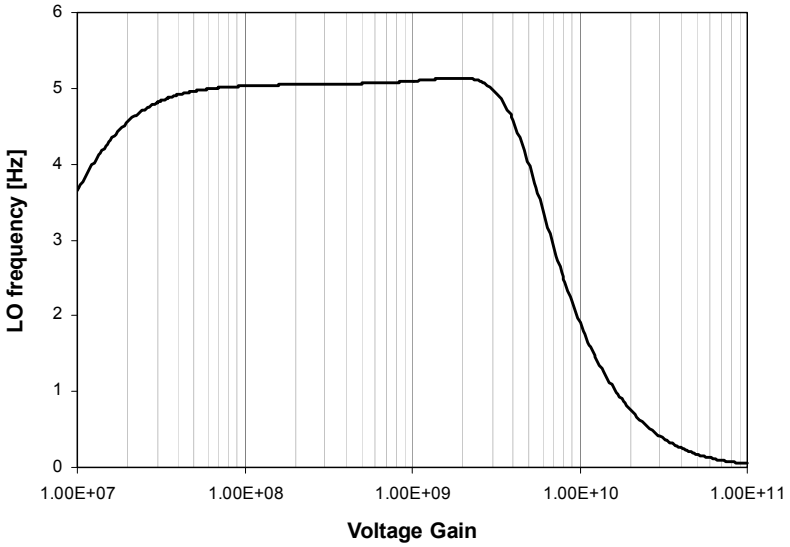


Fig 5.9: Voltage gain (from LO_{IN} to OUT as labeled in fig 5.8) versus LO frequency

5.1.6 LO Limiting Amplifier

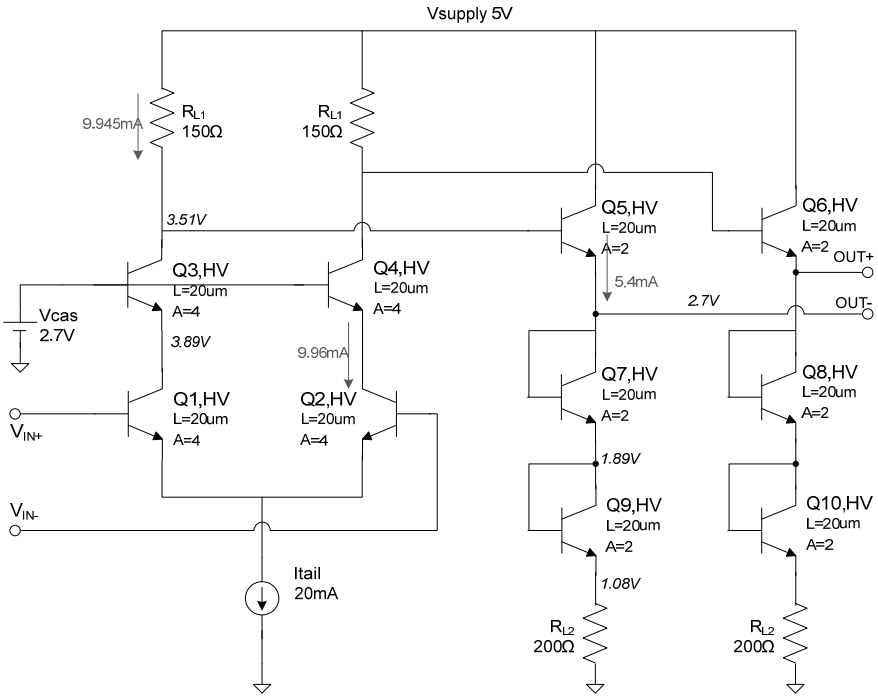


Fig 5.10: Annotated schematic of the limiting stage

The limiting stage (fig 5.10) did not suffer any changes in topology compared to figure 4.14. Its fully differential structure is insensitive to bondwire inductance, which for this reason was not added in our simulations. The 4 diode-connected transistors in the output stage ($Q_{7,8,9,10}$) serve as level shifters in case voltage levels different to 1.5V…4.2V are desired (For example 0.75V…3.45V, or 0V…2.7V).

5.1.7 LO Buffer

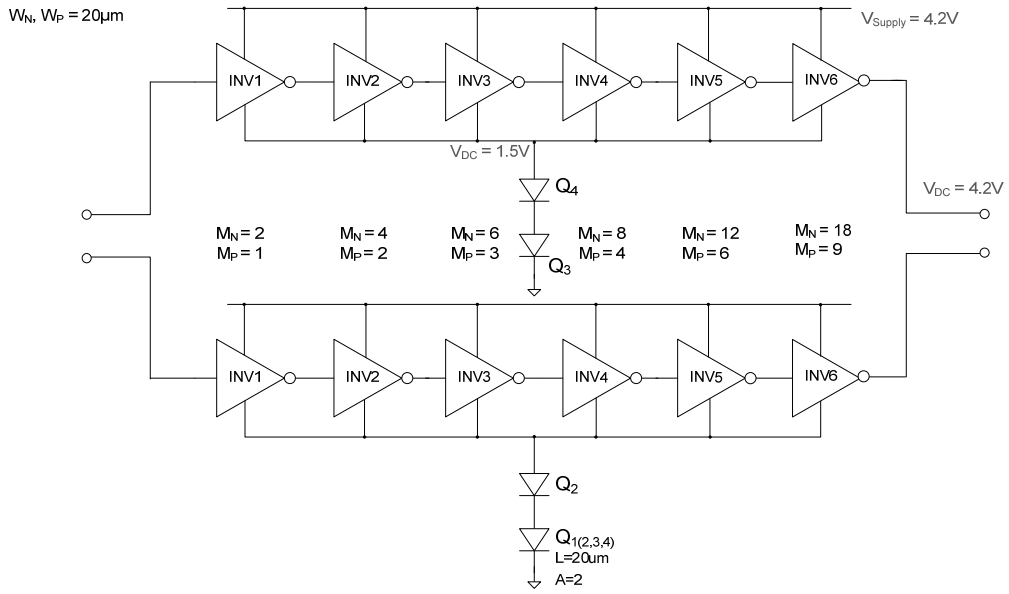


Fig 5.11: Schematic of the LO buffer

5.2 Simulation Results

To quantify the performance of our receiver, small-signal and large-signal simulations were run versus the input RF frequency, the LO power level, the output (intermediary) IF frequency, input power level, and finally tone spacing.

Versus RF *frequency*

The receiver shows a total gain, S_{21}, of 19.6dB at 1GHz and 16.1 at 3.5GHz, an input reflection coefficient, S_{11}, lower than -10dB up to 3.5GHz, and a single-side band noise figure, NF_{SSB}, lower than 5.5dB in the bandwidth 1-4GHz, with a minimum of 4.7dB at 2GHz (see figure 5.12 for small-signal simulation results). At RF frequencis higher than 4.5GHz the logic gates – i.e. the inverters used as LO buffers – stop switch correctly. There are three reasons for this: by far the most important is a technology-related limitation. Frequencies higher than 3GHz are already too high for our 0.25µm process, as the ring oscillator analyzed in figures 2.16 and 2.16 illustrates it. Secondary reasons are the bandwidths of the LO balun (see figure 5.5.b) and the LO input amplifier (figure 5.9): the magnitude of the voltage that reaches the first pair of inverters is smaller above 4GHz as compared to lower frequencies. In addition, the difference between S_{21} and S_{31} in figure 5.5.b becomes significant above 4GHz. That explains the abnormal behavior observed in figures 5.12 and 5.13.

Figure 5.12.b also shows a would-be double-side band noise figure, which can have no practical significance since no image-rejection filter can be placed between the LNA and mixer to remove the noise at the image frequency. However, it does give us a clue as to the noise contribution of the mixer and TIA, by comparing this NF_{DSB} with the noise figure of the LNA (alone) plotted in figure 3.22.b. For example, we can see from figure 3.22.b. that the noise figure at 2GHz is 1.27dB, while figure 5.12.b shows that adding the mixer and TIA increases it by 0.53dB.

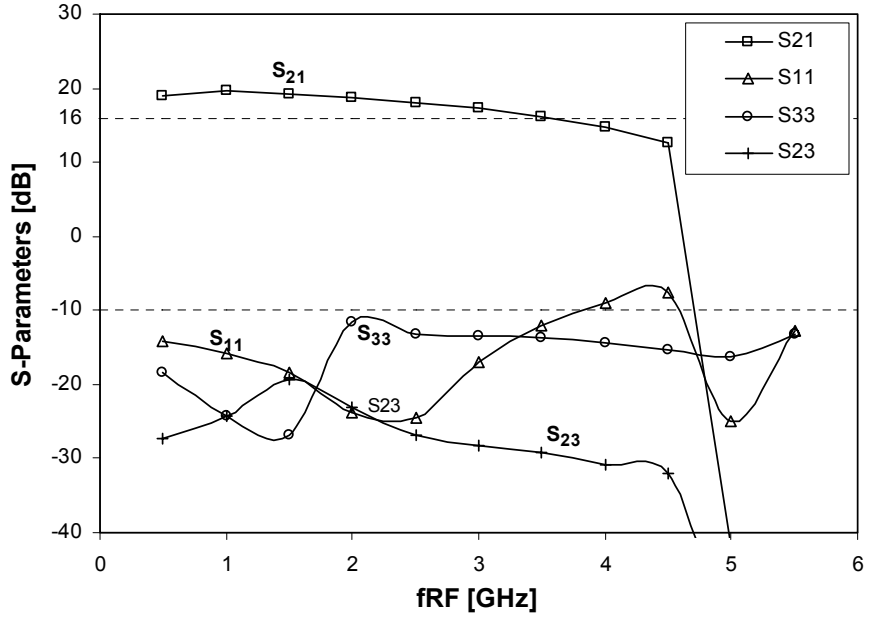
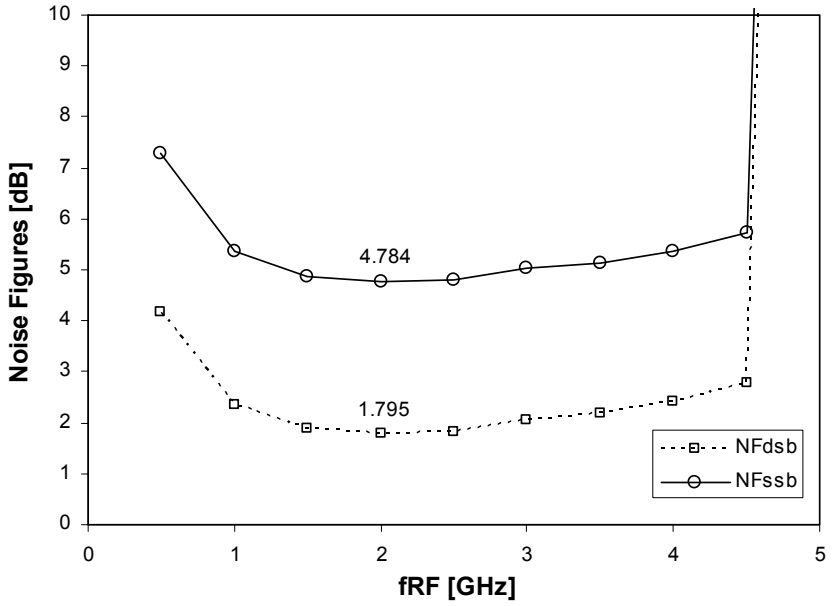


Fig 5.12 (a)



(b)

Fig 5.12: S-Parameters versus input frequency, f_{RF} [1]

As for linearity (figure 5.13), the receiver shows an input IP_3 higher than 18dBm throughout the bandwidth (0.7-4GHz), comparable or in some cases higher than that of some commercial LNAs, introduced in chapter 1 and 2[2].

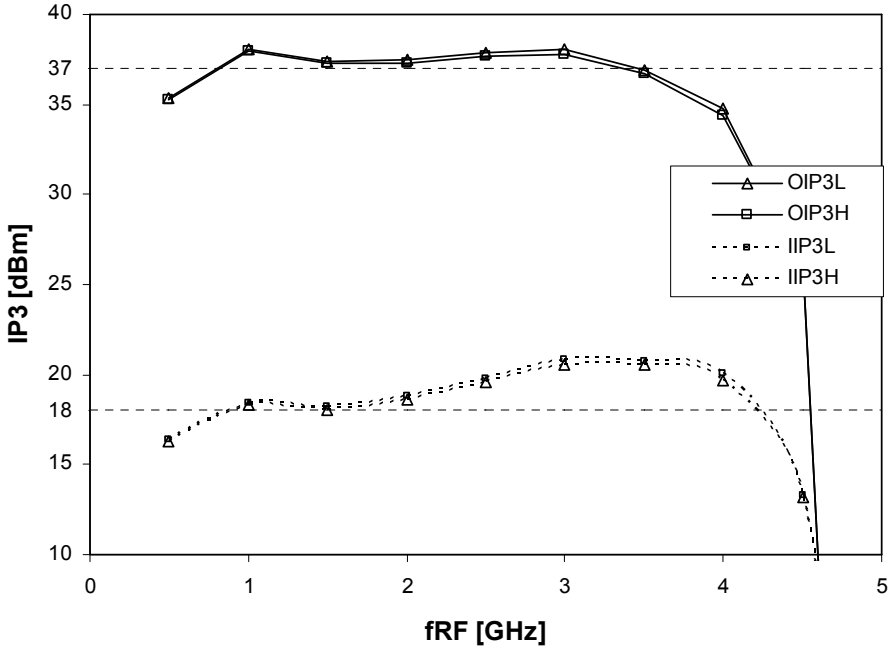


Fig 5.13: Third-order intercept points as function of RF frequency

[1] The LO reflection coefficient, S_{33}, and LO-to-IF leakage, S_{23}, are actually simulated at $f_{LO} = f_{RF} + f_{IF}$, where f_{IF} = 20MHz

[2] [2] has an input IP3 of 19dBm, [1] of 12.7dBm, and [5] of 3.4dBm

Versus LO *Power*

Figures 5.14 and 5.15 show that the power of the LO signal has a very small influence on the behavior of the receiver, as long as it stays smaller than 5dBm. This happens because the voltage that reaches the LO inputs of the mixer has always the same magnitude (thanks to the inverters that produce the same rail-to-rail voltage swing), independently of the LO power. Levels higher than 5dBm though will put the limiter out of order, because, as figure 4.12 shows, the voltage at the input of this block exceeds the allowed range.

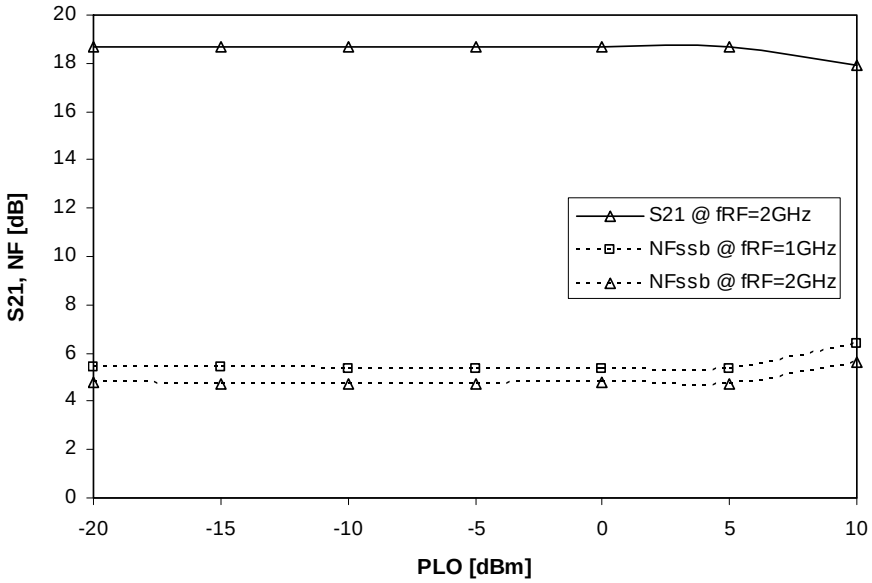


Fig 5.14: Gain and noise figure as function of LO input power

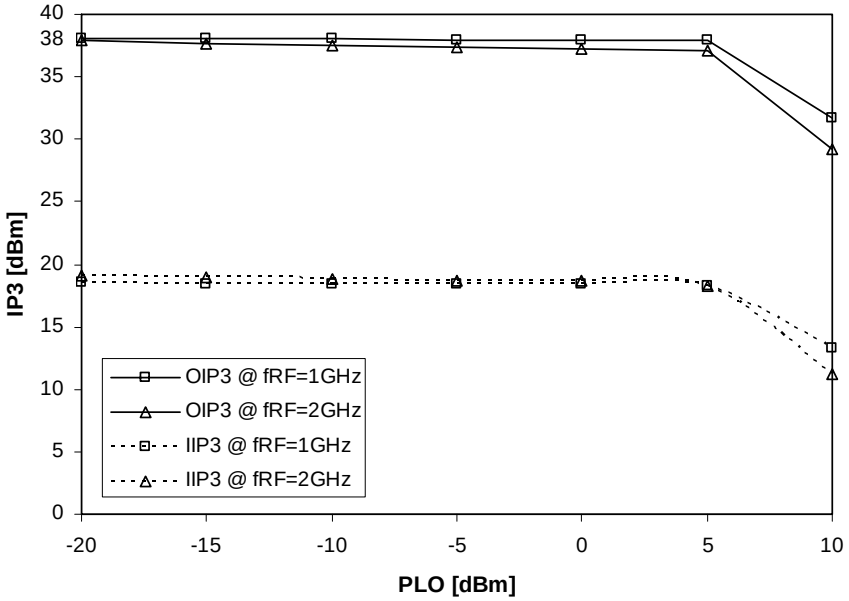


Fig 5.15: IP3 is (almost) insensitive to LO level for $P_{LO} \leq 5$dBm

Versus IF *frequency*

The gain and the noise figure are shown below, in figure 5.16, as function of the output frequency, f_{IF}. The feedback network of the trans-impedance amplifier (made up of $C_F||R_F$) acts as a 1st order RC

filter, with a 3dB frequency at 530MHz. This can be observed in the plot of S_{21} (for f_{RF} = 2GHz), shown in the same figure.

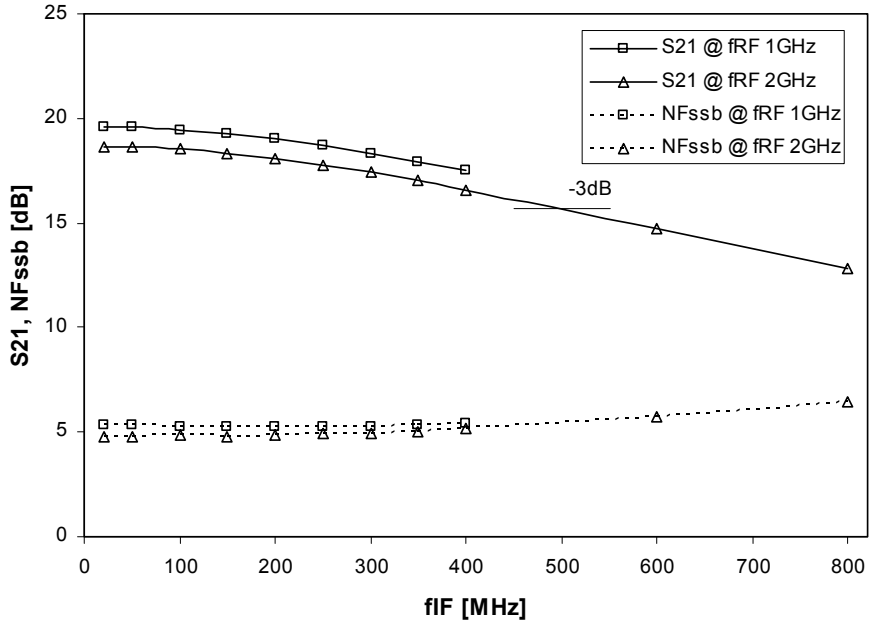


Fig 5.16: Gain and noise figure versus the intermediary frequency, f_{IF}

The plot above also reveals another aspect, this time related to the 1/*f* noise of the NMOS transistors that make up the mixer core. Because we allowed a DC current to flow through these transistors, we expected to see the 1/*f* noise contributing to the total noise figure. We can now say that its part is insignificant, otherwise we would see a drop in NF_{ssb} in figure 5.16, as f_{IF} increases. To validate this assumption, we made the mixer passive by adding a capacitor before each of the mixer inputs, and replaced the 1.5V DC source connected to the center tap of the RF balun (see figure 5.1) with a 1uF capacitor. The results, plotted in figure 5.17, show an insignificant deviation of the single-sideband NF from the previous state when the mixer was "active"

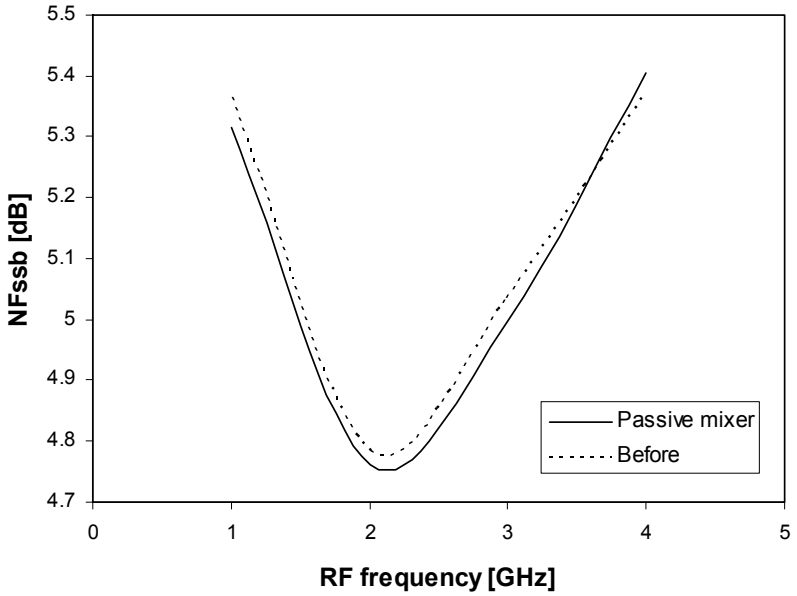


Fig 5.17: NFssb does not change considerably if the mixer is made passive

The linearity figures versus the intermediary frequency f_{IF} (figure 5.18) see a decrease as f_{IF} goes up. This is mainly caused by the fact that the linearity of the trans-impedance amplifier worsens as the operating frequency increases, as other distortion mechanisms (especially the non-linear base-emitter capacitance of the input pair) come into play in addition to the exponential current distortion. Other reasons (by far of less significant impact) can be attributed to the increase in the input impedance of the TIA (see also figure 4.6) as f_{IF} becomes higher.

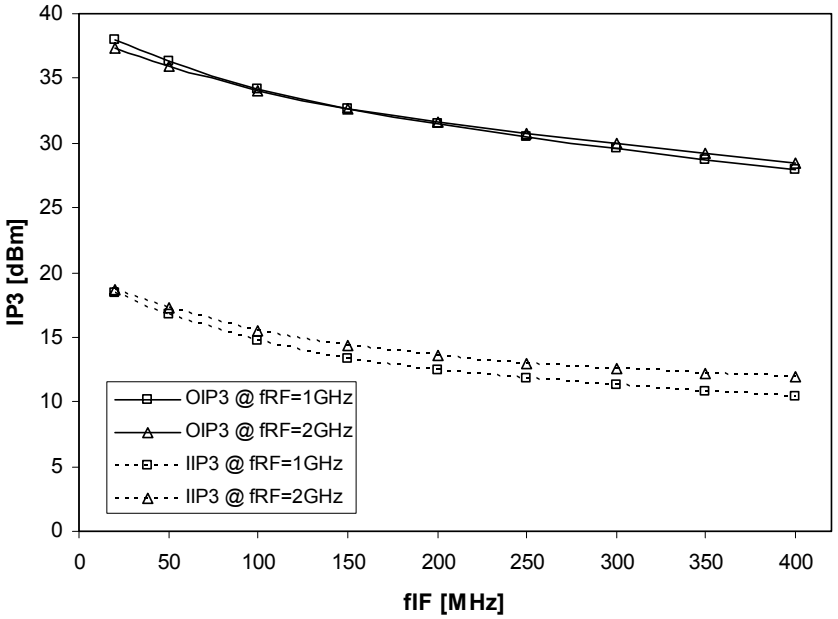


Fig 5.18: Input and output IP_3 versus f_{IF}

Versus Tone Spacing

As the tone spacing becomes larger, we observe from figure 5.19 how the IM3 products become more asymmetric. This problem is generated by the use of distinct *BB* and 2*F* emitter resistances (according to [23]), but also by the biasing inductance $L_{BB,b}$.

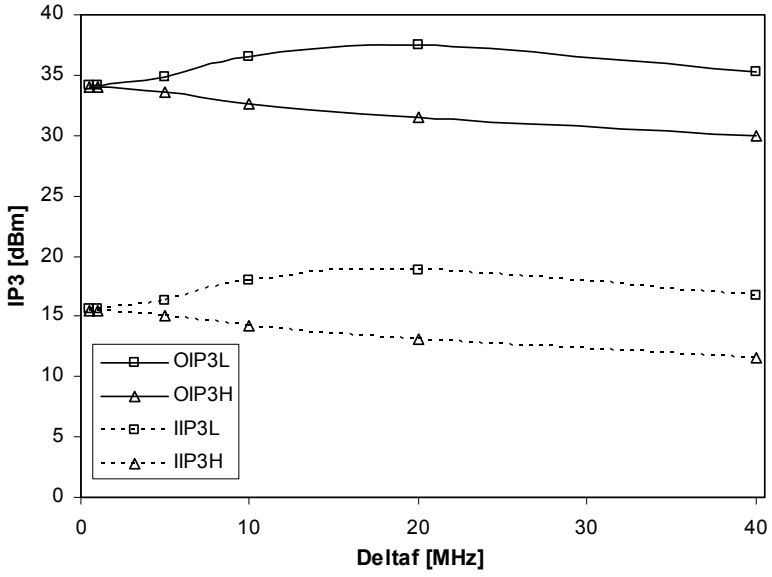


Fig 5.19: OIP_3 versus RF tone spacing

1dB Compression Point (figure 5.20)

An output P_{1dB} of 11dBm corresponds to a voltage of 2.24V over a (differential) load resistance of 400Ω. This means a voltage of 1.12V (or 1.58V) at each of the outputs of the trans-impedance amplifier (see figure 5.7). This is precisely the voltage available at the output node[1], and reveals the limitation in terms of compression comes from the output stage of the TIA. Steps toward increasing P1dB should consist in increasing the DC voltage level at the output of the differential pair of this amplifier. Another issue mentioned in 5.1.2 regards the compression point of the LNA, and has to do with the leakage inductance of the RF balun. An improvement is possible here if higher supply voltages are used.

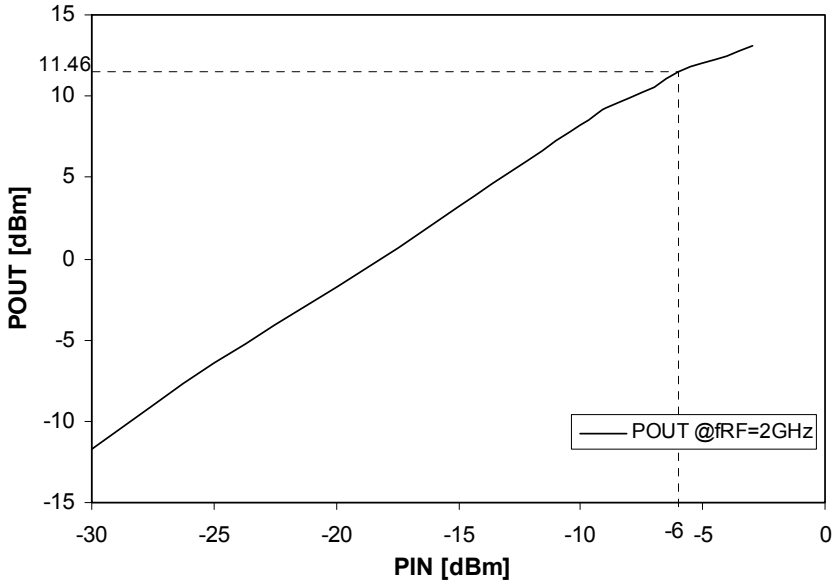


Fig 5.20: 1dB compression point

Effect of Temperature

The following pictures compare the gain (figure 5.21), noise factor (fig 5.22), and output IP_3 (figure 5.23) at -40, 27 and 90 centigrade. A remark must be added here: since our project did not include the design of a PTAT biasing circuit, we employed ideal current sources.

[1] In our simulations, ideal sources are employed to bias the output stage of TIA; this explains why the output can swing as low as 0V (if not lower). Real sources will undoubtedly reduce the dynamic range.

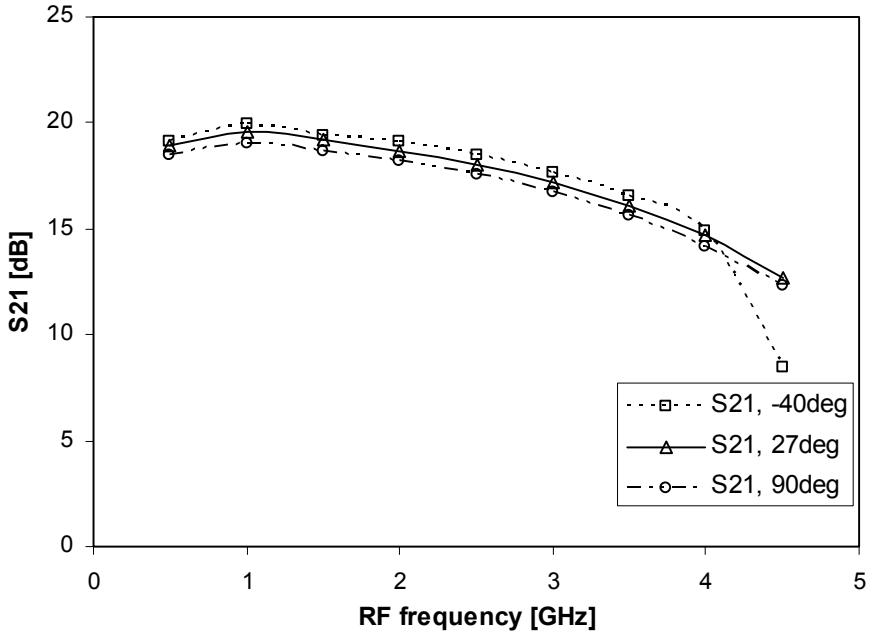


Figure 5.21: Gain versus RF frequency for different temperatures

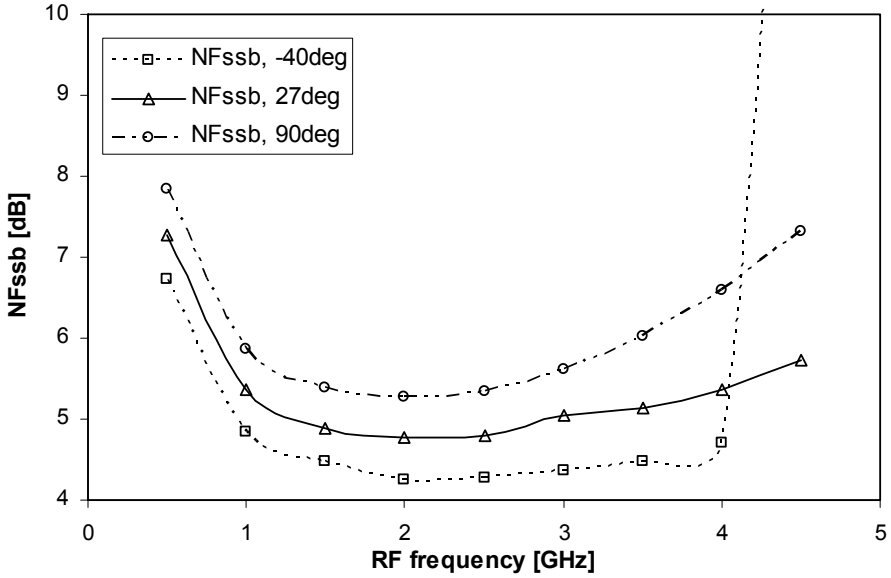


Figure 5.22: Noise figure versus RF frequency for different temperatures

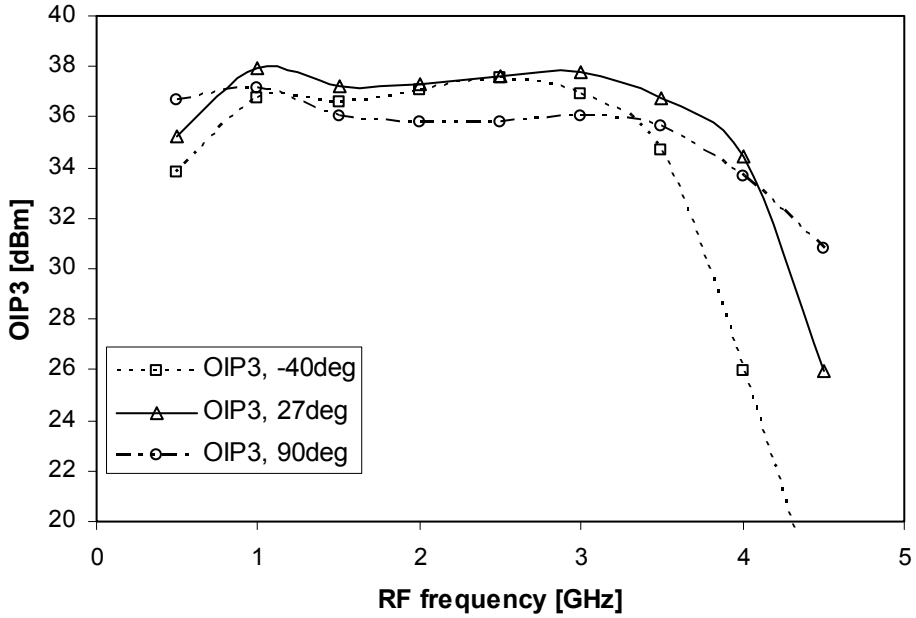
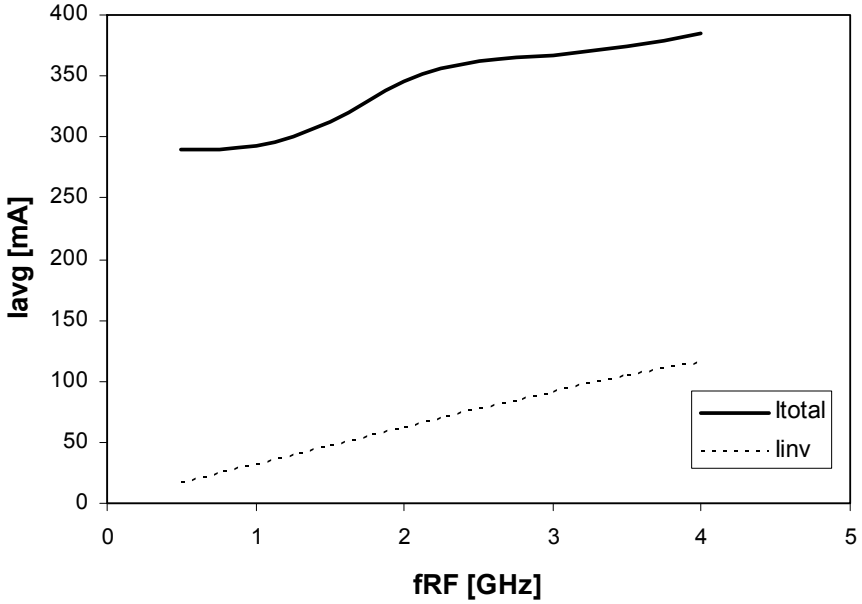


Figure 5.23: OIP_3 versus RF frequency for different temperatures

Dynamic Current Consumption

The total current drawn by our circuit shows an increase versus frequency, mainly due to the inverters chain. This is pointed out in figure 5.24, where the current drawn by the inverters is plotted beside the total current. This behavior is expected from a basically digital circuit, whose current depends on the number of switches per second (which increases versus frequency). The major contributor throughout the bandwidth is the TIA though (more than 33%) and it is directly related to its linearity. Therefore any power savings can be made at the expense of linearity. Table 5.1 gives the figures for each block.



5.24: I_{total} and $I_{inverters}$ versus RF frequency

Table 5.1: Dynamic current consumption of each block

f_{RF}	$I_{Inverter}$ [mA]	$I_{Limiter}$ [mA]	I_{LNA} [mA]	I_{TIA} [mA]	I_{LOAMP} [mA]	I_{Total} [mA]
1	31.94	32.37	56.06	131.1	41.88	293.35
2	62	32.37	56.06	131.1	63.89	345.42
3	91.1	32.37	56.06	131.1	56.01	366.64
4	115.1	32.37	56.06	131.1	50.33	384.96

Layout

Finally figure 5.25 shows a possible layout, useful to estimate the area occupied by our receiver to 1300µm × 1500µm. The passive components make up for more than 50% of the chip area.

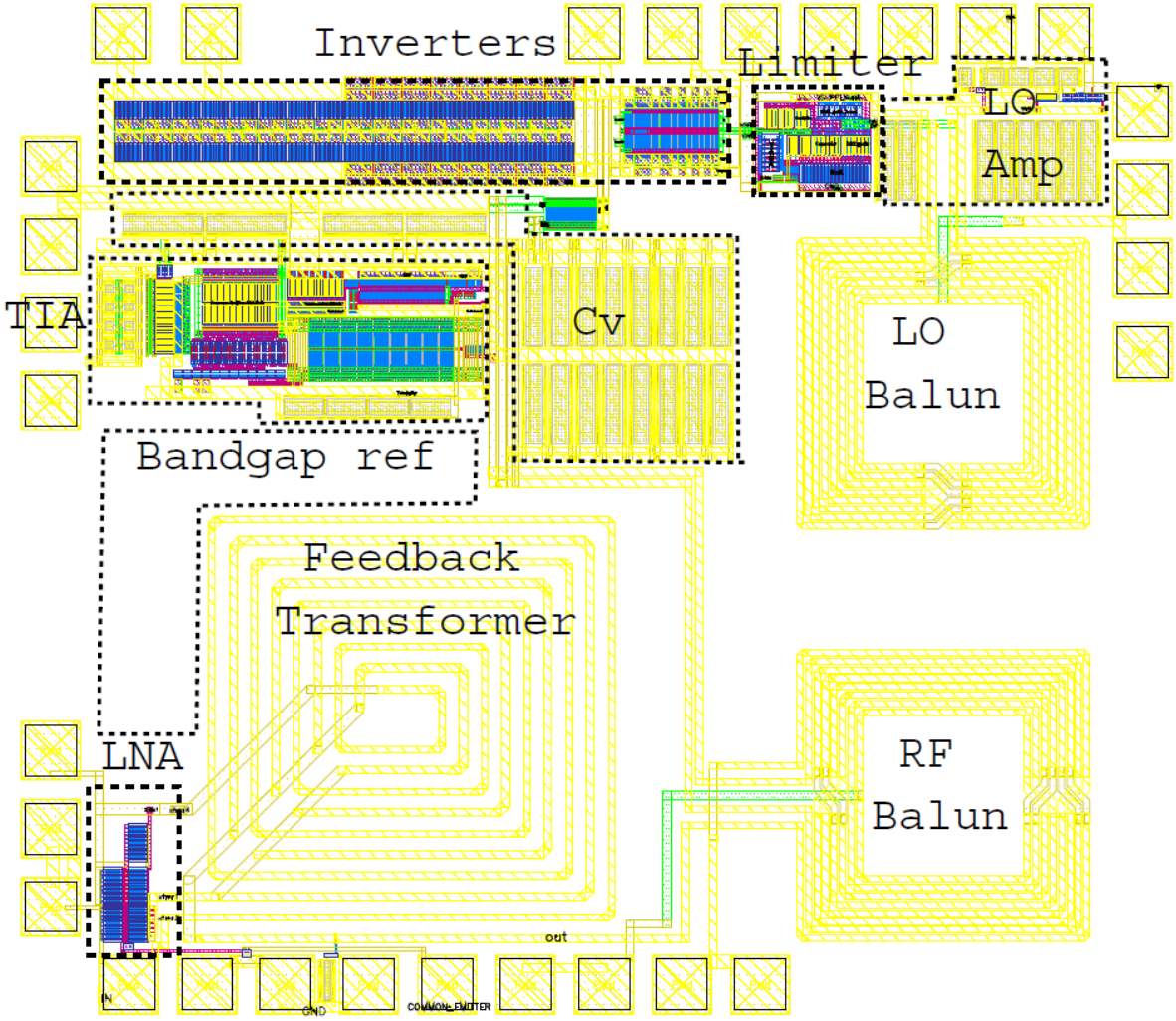


Fig 5.25: A possible layout of the entire receiver

5.3 Adding the Input Image-Rejection Filter

A most important observation must be made at this point. The solution proposed in section 2.3.2 includes an image-rejection filter at the beginning of the chain. Because we wanted to get a clear picture about the performance of our blocks, no external filter has been included in our simulations. Referring back to figure 2.5, if no filter is added in front of the LNA, it is not just the noise of this block at the image frequency that is folded onto the desired (down-converted) signal, but also that of the signal source. The single-sideband noise figure obtained in figure 5.12.b has the following expression (cf. equation (2.6)):

$$NF_{SSB} = 2 + 2\frac{\overline{v_{n,LNA}^2}}{\overline{v_{n,s}^2}}$$

Placing an image rejection before the LNA removes the contribution of the signal source resistance (v_{ns} in figure 2.6) at the image frequency. This noise figure, denoted NF_{min}, described by equation (2.6), is plotted in figure 5.26, alongside NF_{SSB} and NF_{SDB} as plotted in figure 5.12.b.

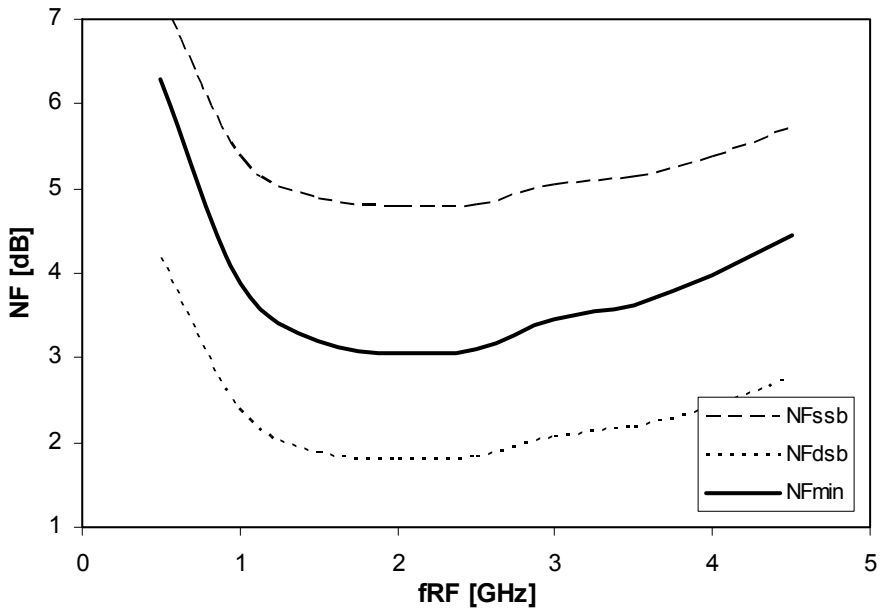


Fig 5.26: Adding an IR filter before the LNA improves the noise figure

5.4 Comparison

Since on the market there are no equivalent products (single chip LNA and mixer) to make a relevant comparison, we selected the best existing low-noise amplifiers and mixers (with respect to noise and linearity) and built a receiver. For a schematic, the reader is asked to go to figure 2.2. TriQuint [32] has a single-chip down-converter for UMTS, but the blocks are unconnected to one another and require external matching networks and filters that make it narrowband. The results are plotted in figures 5.27, 5.28 and 5.29. As a wideband[1] solution, we used a 50MHz to 6GHz LNA from Avago

[1] The constituent LNA and mixer are wideband, but the receiver as a whole is narrowband due to the interstage filter. In other words, we can use the same LNA and mixer throughout the given bandwidth, but for each frequency we need a different IR filters

(MGA-53543) and a 10MHz to 6GHz active mixer from Analog Devices (ADL5801). Other combinations also use LNAs from Avago Technologies and mixers from Analog Devices, but these are one-standard receivers (i.e. for center frequency of 800MHz, 1900MHz or 2400MHz). All the above receivers include an ideal image-rejection filter of 1.3dB in-band attenuation (except of course for our receiver and the one made by Triquint). That explains the difference between our *NF* (NF_{min} as show in thick line in figure 5.26) and the noise figures of the other receivers. In addition to this reason, the Avago's LNAs are GaAs PHEMT devices with extremely low noise figures.

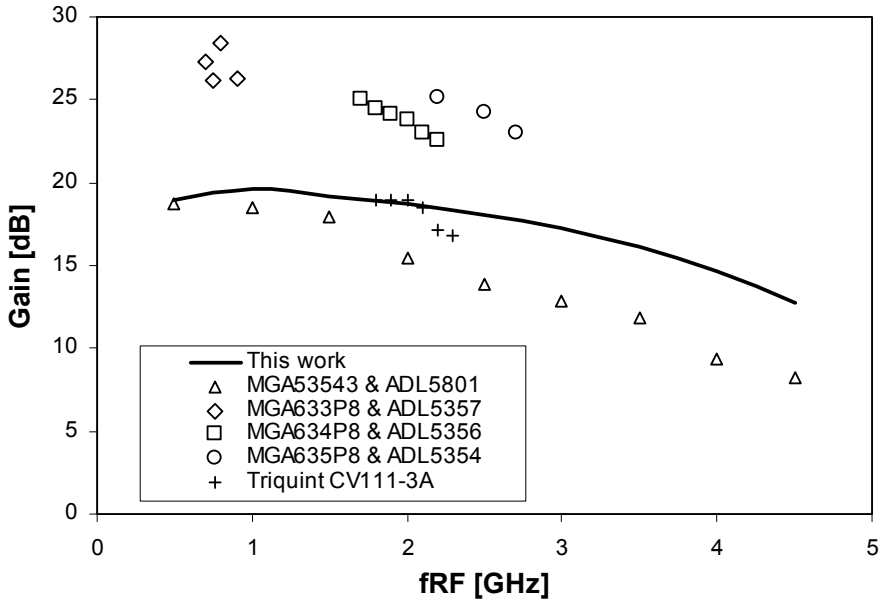


Fig 5.27: Gain of our receiver is compared to that of different receivers (composed by us, except Triquint's)

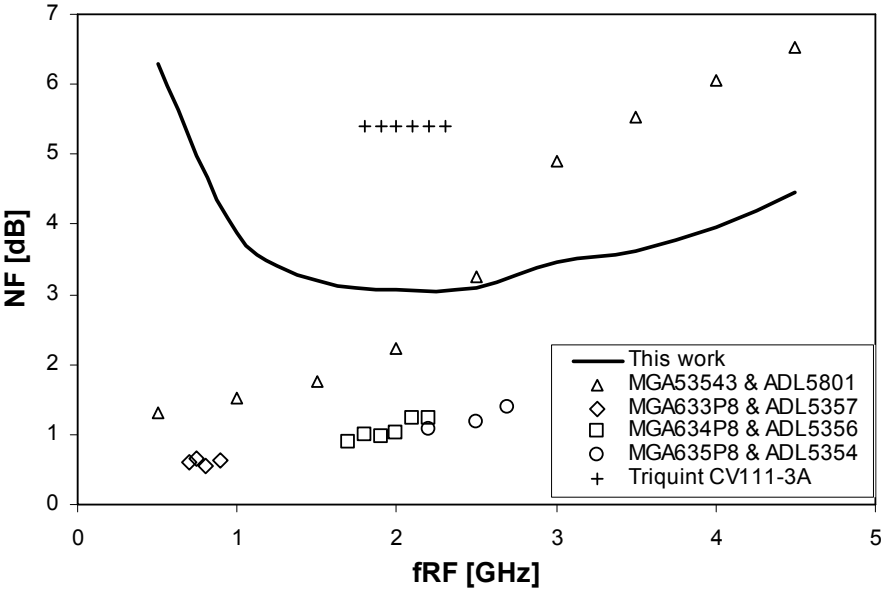


Fig 5.28: Comparison between noise figures

If the noise figure of our receiver is higher than that of the others (due to the inexistence of an image-rejection filter between the LNA and mixer), the linearity figures (see figure 5.29) on the other hand are significantly better than the rest, except (partially) for one designed for 1.9GHz operation.

Although the LNAs have extremely high linearity figures (around 40dBm of OIP_3), the mixers figures on the other hand are limited, and therefore determine the overall IP_3 figures.

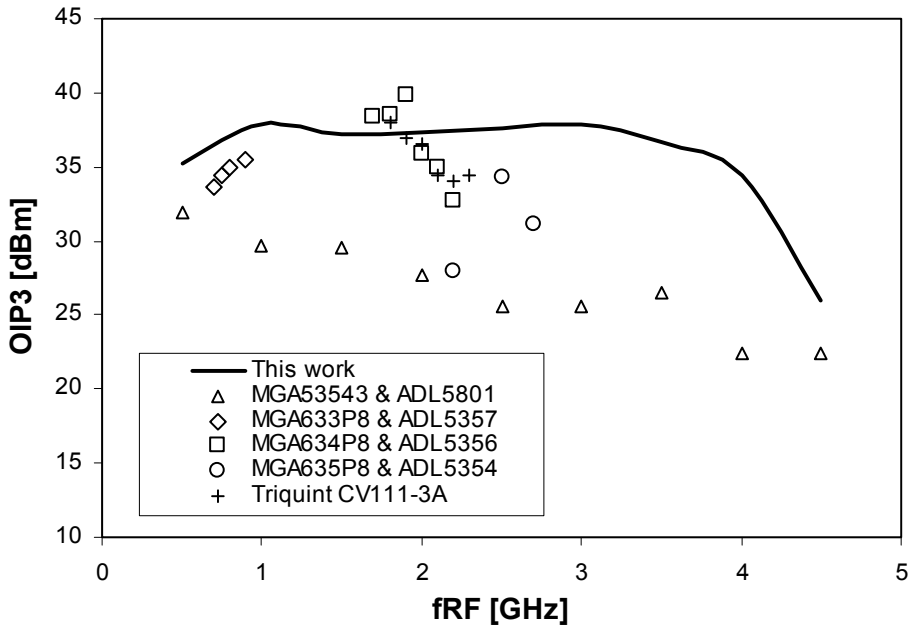


Fig 5.29: OIP3 versus RF frequency (comparison)

Table 5.2 offers some figures for the blocks used in our comparison

Table 5.2: Specifications for the blocks used for generating figures 5.27, 5.28, 5.29

Part	*BW* [GHz]	*Gain* [dB]	*NF* [dB]	OIP_3 [dBm]	$P_{1dB,out}$ [dBm]	I_{supply} [mA]
LNAs						
Avago MGA-53543[1]	0.05-6	15.4	1.5	39	18.6	54
Avago MGA-633P8[2]	0.45-2	18	0.37	37	22	54
Avago MGA-634P8[3]	1.5-2.3	17.4	0.44	36	21	48
Avago MGA-635P8[4]	2.3-4	18	0.56	35.9	22	56
Mixers						
Analog Devices ADL5801	0.01-6	1.8	9.75[5]	30.3	15.1	130
Analog Devices ADL5357	0.5-1.7	8.6	9.1[5]	35.2	18.8	185
Analog Devices ADL5356	1.2-2.5	8.2	9.9[5]	39.2	19.2	350
Analog Devices ADL5354	2.2-2.7	8.6	10.6[5]	34.7	19.2	350
Down-converter						
Triquint CV111-3A	1.9-2.2	20	5.4[6]	37.9	20.7	360

[1] Specifications at 1.9GHz
[2] At 900MHz
[3] At 1900MHz
[4] At 2500MHz
[5] Single-sideband noise figure
[6] No plot was offered for NF, so we consider it constant and equal to 5.4dB over the entire bandwidth

5.5 Conclusion

When we chose to remove the IR filter (that was placed between the LNA and mixer), we stated that the degradation in terms of noise figure would not be in fact significant, because a tower-top amplifier (placed in front of our ground-level receiver) would set the overall noise figure. To verify this assertion, we simulated our front-end together with the tower-top amplifier. The simulation setup is illustrated in figure 5.30 and corresponds to parts of figure 2.2. For the tower-top amplifier we used a generic LNA. The numbers for gain, OIP_3 and NF are averages of the figures that can be found on the market

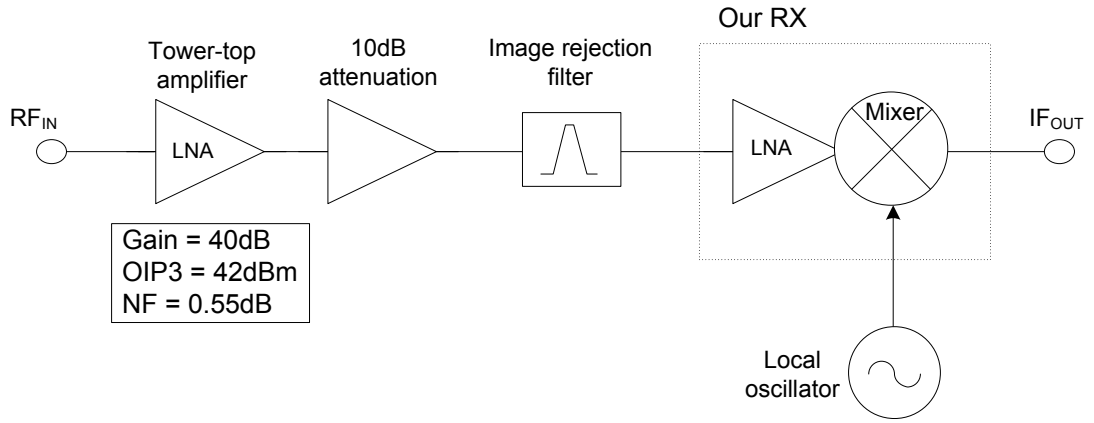


Fig 5.30: Our single-chip solution is simulated as part of a complete front-end

The results, plotted in figure 5.31 confirm that the noise of our receiver becomes almost unnoticeable once it is placed after a very low noise amplifier. The good figures for linearity are also preserved – compare figure 5.32 with 5.13 or 5.28.

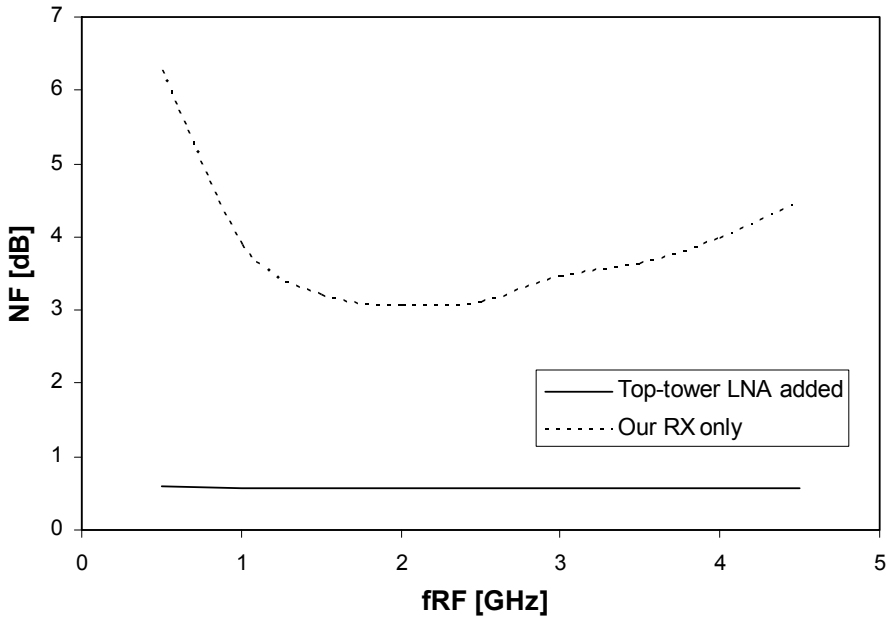


Fig 5.31: Total noise figure of the complete front-end introduced in figure 5.30

While the noise figures can be fixed in one way or the other, not the same can be said about the 1dB compression point. By comparing the P_{1dB} of our receiver (11.4dBm according to figure 5.20) with the ones of the mixers mentioned in table 5.2[1], we observe that it is outperformed by its competitors. This is, as already mentioned, mainly due to the voltage DC levels at the output of the TIA, which are not at $V_{DD}/2$. Anyway, the 11.4dBm value is according to tables 2.6 and 2.7 in agreement with our project requirements.

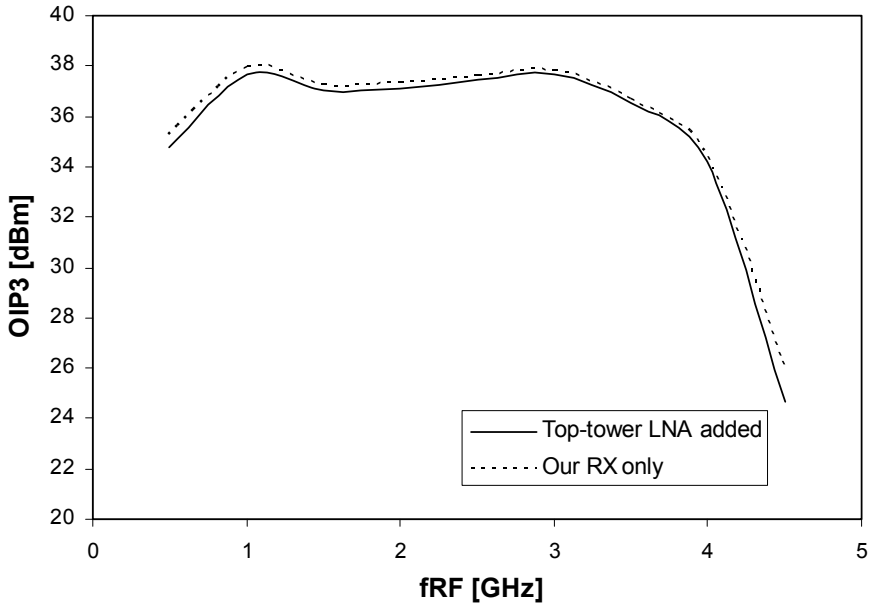


Fig 5.32: OIP_3 of our RX determines the overall linearity

[1] The mixers produced by Analog Devices (and used here for comparison) use open-collector transistors for the output stage, which in combination with large biasing inductor offers very large voltage swings, hard to compete with using source followers, as we do.

Chapter 6

Conclusions and Future Directions

Review of work

This report has illustrated the design flow of an integrated receiver for basestation applications. The most important design criteria on which we based our analyses at every step of our design have been linearity (through the 3rd order intercept points) and noise. An initial analysis of the existing receiver architectures allowed us to understand the compromises (with respect to the two abovementioned requirements) inherent to our decision to integrate the LNA and the mixers on the same chip. Since our down-converter is not supposed to be placed at the forefront of the receiving chain, we decided to put linearity on the first place, and allowed small sacrifices when it came to noise. This decision sat behind the first steps we made when we selected the transistor level implementations of the mixer and LNA. Additional refinements and techniques (such as out-of-band cancellation) gave us the possibility to maximize the linearity, but also to improve the noise performance.

The results we obtained proved how correct this design strategy has been: our receiver achieves very high output 3rd order intercept points (above 37dBm) over a large bandwidth (0.8 – 3.5GHz). Compared to this performance, the best receivers available on the market show poorer linearity figures (OIP_3 equal to 35dBm on average) and their operation is restricted to narrow bandwidths.

The poor noise figure on the other hand is the consequence of integration (that made us eliminate the interstage IR filter) and the constraints imposed by the actual basestation structure (which do not allow for an IQ topology), rather than of inadequate design. The last section of the previous chapter has proved that, as part of a complete receiver system, our circuit does not degrade the overall noise figure

Our solution should be attractive not just because it offers better linearity but also because its wideband character, which makes it suitable for multiple mobile standards. This latter argument and the integration of the LNA and the mixer on the same chip can bring significant benefits in terms of costs and time to market.

Possible improvements

Noise. Our receiver performs very well as far as linearity is of concern. Any further improvements must in consequence target a lower noise figure. We have already seen that the noise of the mixer and TIA has a significant contribution to the total noise figure. This has happened because, as we moved from an ideal feedback transformer to a monolithic implementation, we were forced to reduce the number of turns in order to minimize the inter-winding capacitance. Due to this the gain of the LNA was reduced, and consequently the noise of the subsequent stages became more significant. Making the physical distance between windings even greater will lower the coupling factor, and as a result augment the closed-loop gain. In addition, this move can increase the bandwidth (due to a decrease in

the inter-winding capacitance) but also reduce the value of R_E that is necessary in order to achieve 50Ω input matching[1] (with the result that the minimum noise figure becomes smaller).

As an example let us assume the gain of the LNA is increased such that the noise of the mixer is made insignificant. The double-sideband noise figure now coincides with that of the LNA (as plotted in figure 3.22.b). The minimum achievable noise figure, NF_{min}, is plotted in figure 6.1 compared to the NF_{min} we have in our actual receiver (see also section 5.3). We also added a third case, in which we assumed the noise figure of the LNA is somehow further reduced by 0.3dB. We can now understand why attention should focus on improving the gain of the first stage, but also on decreasing its noise.

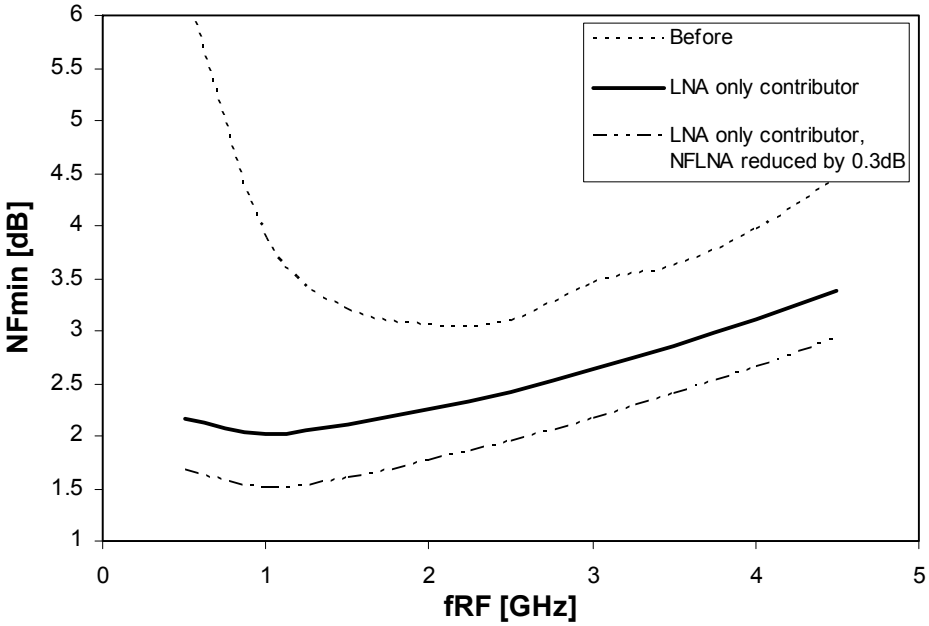


Fig 6.1: NF_{min} of the receiver for three cases: a) actual situation,
b) the noise of the second stage is made negligible by an increase in the LNA gain,
c) the noise figure of the LNA is reduced by 0.3dB

A different approach to minimizing the noise figure implies more radical changes. Interposing an off-chip image-rejection filter between the LNA and the mixer will bring the noise figure down to its lowest possible value (i.e. the double-sideband noise figure).
This situation is illustrated in figure 6.2: the LNA and mixer are still integrated on the same chip, but their output and input ports respectively must be matched to 50Ω. The consequence is that current operation between these nodes is no longer possible (this might prove disadvantageous for linearity as well), and the LNA structure must be changed. A good starting point in this direction can be the amplifier introduced in figure 2.15.a; low noise figures are achievable by this amplifier if the feedback resistors are carefully selected (very large R_F and very small R_E). More difficult might prove to ensure 50Ω input impedance for the mixer, since the current topology has fundamentally only low impedance nodes (the input impedance of the mixer is in the order of a couple of Ω). One way to tackle this problem is to increase the turns ratio of the RF balun. However, practical considerations put a limit on the maximum number of turns, and additional ways of increasing the input impedance must be investigated (such as increasing the ON resistance of the switches that make up the mixer

[1] The condition for input matching is $50\Omega = n \cdot R_E$, where $n = N \cdot k_m$, is the effective turns ratio (see equation (3.14.b))

core, increasing the input impedance of the TIA, etc.). Care must be taken in this case so that the linearity is not significantly deteriorated.

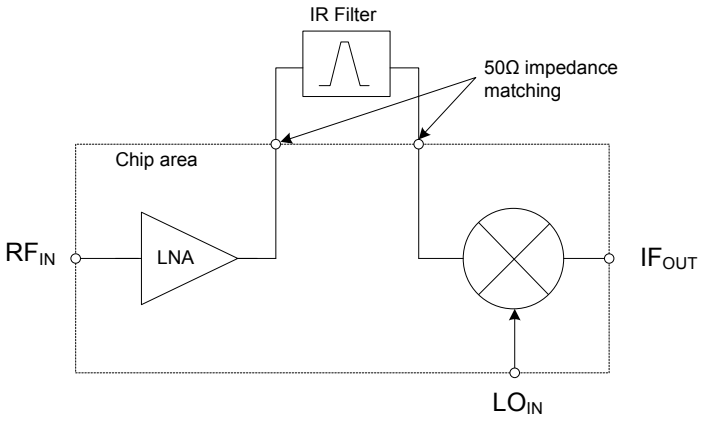


Fig 6.2: Adding an inter-stage IR filter results in minimum noise figure, but requires impedance matching

1dB Compression Point. Another point open to inprovement in our design was related to the 1dB compression point. As mentioned earlier on, responsible for this compression restriction is the trans-impedance amplifier. A readjustment of the DC levels of the op-amp can fix this issue.

Power consumption. To minimize the supply current, attention should focus especially on the LO system. Here the current consumption can be lowered by eliminating the DC current that flows through the feedback resistor; increasing the value of this resistor can reduce the dynamic power consumption, but attention must be paid not to drive the this block into saturation. Another step toward power reduction can be made by removing the limiter altogether and connect the inverters directly to the LO balun. The last option in this direction relates to the inverter chain itself, which has a current consumption that increases with the frequency. Small adjustments to the transistor sizing can save some mA's without too much sacrifice in terms of linearity.

A silicon implementation of the current design is the ultimate test for the techniques developed in this report. Steps have been made towards this goal, but a final revision of the current layout is a must. However, with the layout provided we could get a clear picture about the area occupied by our receiver, which finally sets the production costs.

Bibliography

[1] Avago Technologies ALM-1222. 1.8-2.2GHz Low Noise, High Gain, High Linearity Balanced Amplifier Module.

Datasheet available at: http://www.avagotech.com/docs/AV02-1470EN.

[2] Avago Technologies MGA-633P8. Ultra Low Noise, High Linearity Active Bias Low Noise Amplifier.

Datasheet available at http://www.avagotech.com/docs/AV02-2329EN.

[3] Application Note 2021. Specifications and Measurement of Local Oscillator Noise in Integrated Circuit Base Station Mixers.

Link: http://www.maxim-ic.com/app-notes/index.mvp/id/2021.

[4] Datasheet Powerwave LGP146nn. 1900 MHz Duplex filter Outdoor with DC Bypass.

[5] Avago MGA-13116. High Gain, High Linearity, Very Low Noise Amplifier.

Datasheet available at: http://www.avagotech.com/docs/AV02-2877EN.

[6] LS Cable, LHF 12D (1/2). Low Loss Flexible Foam Dielectric Feeder.

RF Brochure available at: http://www.lsrts.com/downFile.php?type=catalogue&seq=2.

[7] Avago Technologies MGA-53543. 50MHz to 6GHz High Linear Amplifier.

Datasheet available at: http://www.avagotech.com/docs/AV02-0455EN

[8] TriQuint Part no 855859, 1960MHz RF SAW Filter.

Datasheet:
http://www.triquint.com/prodserv/more_info/proddisp.aspx?prod_id=855859.

[9] Analog Devices ADL5355 1200 MHz to 2500 MHz Balanced Mixer with LO Buffer, IF Amp, and RF Balun.

Datasheet: http://www.analog.com/static/imported-files/data_sheets/ADL5355.pdf.

[10] J. Long, “RF IC Design” Lecture Notes, Delft University of Technology, 2009-2010.

[11] B. Razavi, “RF Microelectronics”, Prentice Hall, 1998.

[12] F. Behbahani, Y. Kishigami, J. Leete, A. Abidi, “CMOS Mixers and Polyphase Filters for Large Image Rejection”, IEEE Journal of Solid-State Circuits, Vol. 36, No. 6, June 2001.

[13] K. Linggajaya, D. Manh Anh, M. Jin Guo, Y. Kiat Seng, “A New Active Polyphase Filter for Wideband Image Reject Downconverter”, ICSE 2002 Proc. 2002, Penang, Malaysia.

[14] K. Komoriyama, E. Yoshida, M. Yashiki, H. Tanimoto, “A Very Wideband Fully Balanced Active RC Polyphase Filter Based on CMOS Inverters in 0.18μM CMOS Technology”, IEEE 2007 Symposium on VLSI Circuits Digest of Technical Papers.

[15] X. Wang, A. Dengi, S. Kiaei, "A High IIP3 X-Band BiCMOS Mixer for Radar Applications", Proceedings of the 2004 International Symposium on Circuits and Systems, 2004. ISCAS '04.

[16] T. Tikka, J. Ryynänen, M. Hotti and K. Halonen, "Design of a High Linearity Mixer for Direct-Conversion Base-Station Receiver", 2006 IEEE International Symposium on Circuits and Systems, 2006. ISCAS 2006. Proceedings.

[17] L. Sheng and L.E. Larson, "An Si-SiGe BiCMOS Direct-Conversion Mixer with Second-order and Third-order Nonlinearity Cancellation for WCDMA Applications", IEEE Transactions on Microwave Theory and Techniques, Vol. 51, No. 11, November 2003.

[18] Analog Devices ADL5801, "High IP3 10MHz to 6GHz, Active Mixer".

Datasheet: http://www.analog.com/static/imported-files/data_sheets/ADL5801.pdf.

[19] H. Khatri, P.S. Gudem, L.E. Larson, "Distortion in Current Commutating Passive CMOS Downconversion Mixers", IEEE Transactions on Microwave Theory and Techniques, Vol. 57, No. 11, November 2009.

[20] J. Crols and M. Steyaert, "A Full CMOS 1.5GHz Highly Linear Broadband Downconversion Mixer", Twentieth European Solid-State Circuits Conference, 1994. ESSCIRC '94.

[21] Analog Devices ADL5367, "500MHz to 1700MHz Balanced Mixer, LO Buffer and RF Balun".

Datasheet: http://www.analog.com/static/imported-files/data_sheets/ADL5367.pdf

[22] M.T. Reiha, J.R. Long and J.J. Pekarik, "A 1.2 V Reactive-Feedback 3.1-10.6 GHz Ultrawideband Low-Noise Amplifier in 0.13 μm CMOS", IEEE Radio Frequency Integrated Circuits (RFIC) Symposium, 2006.

[23] M.v.d. Heijden, "RF Amplifier Design Techniques for Linearity and Dynamic Range", Doctoral thesis, Delft University of Technology, 2005.

[24] L.d. Vreede, "Microwave Circuit Design", Lecture notes, Delft University of Technology, 2009-2010.

[25] B. Razavi, "Design of Analog CMOS Integrated Circuits", McGraw-Hill, 2001

[26] P. Wambacq, W. Sansen, "Distortion Analysis of Analog Integrated Circuits", Kluwer Academic Press, 1998

[27] K. Laken, W. Sansen, "Design of Analog Integrated Circuits And Systems", McGraw-Hill

[28] J. Crols, M. Steyaert, "A Full CMOS 1.5 GHz Highly Linear Broadband Downconversion Mixer", Twentieth European Solid-State Circuits Conference, 1994. ESSCIRC '94.

[29] W. Sansen, "Analog Design Essentials", Springer, 2006

[30] Lee, S.G., and Choi, J.K.: 'Current-reuse bleeding mixer', Electron. Letters, 13th april 2000, Vol. 36, No. 8, pp. 696–697

[31] A. Niknejad, "Electromagnetics for High-Speed Analog and Digital Communication Circuits", Cambridge University Press, 2007

[32] TriQuint CV111-3A, UMTS-band High Linearity Downconverter.

Link: http://www.triquint.com/prodserv/more_info/default.aspx?prod_id=CV111-3A