

Analog Integrated Circuits
TITOLO SERIO

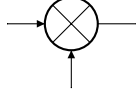
Roberto Rr
chopper19

DATA

Contents

1	Introduction	3
2	Analysis of a Gilbert cell-based multiplier	7
2.1	Gilbert cell overview	7
2.2	Gilbert cell circuit analysis	7
2.3	Mathematical analysis: conversion gain	12
3	Design by hand of a down-converting Gilbert cell	16
3.1	Design specifications	16
3.2	Gilbert cell design flow	17
4	Design by simulation	20
4.1	Gilbert Cell design	20
4.2	Biasing network design	24
5	Layout of the Gilbert Cell	26
5.1	Differential RF stage	26
5.2	Differential LO stages	27
5.3	Current mirror and diode connected transistor	27
5.4	Capacitors	27
5.5	Resistors	28
5.6	Putting together the layout	29
6	Simulation vs schematic comparison	30
7	Conclusions	31

DA SISTEMARE There are many deadlocks and dependencies which have to be taken into consideration while designing the specifications for the transistors in order to satisfy the above design constraints. All the calculated values are changed based on the parametric analysis of transistors in Cadence Virtuoso.



1 Introduction

An analog multiplier, also known as mixer, is a circuit that performs the product between two signals. As we will see, this feature can be exploited to convert informations from a certain frequency band to another by means of the intrinsic non-linear behaviour of this net. Supposing to have two sinusoidal signals

$$\begin{aligned}x_1(t) &= A_1 \cos(\omega_1 t + \varphi_1) \\x_2(t) &= A_2 \cos(\omega_2 t + \varphi_2)\end{aligned}$$

and the ideal mixer shown in **figure one** has that the out-coming signal $y(t)$ is given by:

$$\begin{aligned}y(t) &= x_1(t) \cdot x_2(t) \\&= A_1 A_2 \cos(\omega_1 t + \varphi_1) \cos(\omega_2 t + \varphi_2) \\&= \frac{A_1 A_2}{2} \{ \cos[(\omega_1 - \omega_2)t + \varphi_1 - \varphi_2] + \cos[(\omega_1 + \omega_2)t + \varphi_1 + \varphi_2] \} \\&= A \cos(\omega_{LF} t + \varphi_{LF}) + A \cos(\omega_{HF} t + \varphi_{HF})\end{aligned}$$

where

$$\begin{aligned}\omega_{LF} &= |\omega_1 - \omega_2| \\ \omega_{HF} &= |\omega_1 + \omega_2| \\ \varphi_{LF} &= \varphi_1 - \varphi_2 \\ \varphi_{HF} &= \varphi_1 + \varphi_2\end{aligned}$$

Therefore, it comes that two signals with different frequency allocation are obtained at the output port: the former at $\omega_{LF} < \omega_1, \omega_2$ will be the down-converted component whereas the latter, with $\omega_{HF} > \omega_1, \omega_2$, will be the up-converted one. Moreover, it is possible to select only one of these two signals by properly filtering out the unwanted part. Overall one can read the process as a modulation of an input signal by means of a carrier. Since the mixer is a bidirectional three-port, one can distinguish:

- The high frequency signal, RF. In case of down-conversion this is one input of the circuit, vice-versa it is the output (after filtering).
- The intermediate frequency signal, IF. In case of up-conversion this is one input of the circuit, vice-versa it is the output (after filtering).

- The local oscillator, LO. This is the carrier and it is always an input with known frequency.

Based on the above, it turns out that to mix-up two signals a non-linear device is needed, since the input components are at different frequency with respect to the output and a non-linear relation between voltages and currents appears. In general, mixing can be carried out by time-varying systems that can be implemented using:

passive devices typically switches (diodes and transistors). In this case the mixing process introduces a *loss* since the output power is always less than the input one.

active devices amplifying devices are used in active region providing *gain*.¹ They are more power consuming and less noisy than passive mixers.

Another way to classify mixers is the following:

Single balanced mixers One or both input signals can pass to output, but it is not possible to suppress both of them,

Double balanced mixers Thanks to symmetry in the net both the input and the LO are rejected at the output port. They show better isolation between ports than SBM.

To understand why a transistors can be used, let's consider the non-linear quadratic model for a nMOSFET, supposing to drive that by injecting a two tone signal $v_{GS}(t) = v_{RF}(t) + v_{LO}(t)$ at the gate (down-conversion configuration). One has:

$$\begin{aligned} i_D(t) &= k(v_{GS}(t) - V_{th})^2 \\ &= k(v_{RF}(t) + v_{LO}(t) - V_{th})^2 \\ &= k[v_{RF}^2(t) + v_{LO}^2(t) + 2v_{RF}(t)v_{LO}(t) - 2(v_{RF}(t) + v_{LO}(t))V_{th} + V_{th}^2] \end{aligned}$$

Supposing now $v_{RF} \ll v_{LO}$:

$$\begin{aligned} i_D(t) &\simeq k(v_{LO}(t) - V_{th})^2 + 2k(v_{LO}(t) - V_{th})v_{RF}(t) \\ &= I_D(t) + g_m(t)v_{RF}(t) \end{aligned}$$

hence, the model becomes a *quasi-linear* time-varying *small signal model* and the device operates in the *Small Signal Large Signal* regime (SSLS), because we still have the product of the two input signals through the transconductance. It is worth it to notice that depending on the amplitude of the RF (the LO is always driven in large signal), mixing can be obtained both in linear region or through second order non-linearity (fully non-linear

¹The definition of the gain and loss in mixer will be presented further.

device). A more detailed analysis about how to drive a mixer follows in section 2.

Similarly to linear circuits, even in case of mixers it is possible to define the gain. Actually, gain is defined only in case of linear systems, but this figure of merit is necessary to qualify the performance of the conversion stage. In case of non-monochromatic signals (general case) one defines the *voltage conversion gain* as:

$$A_v = \frac{V_{IF,rms}}{V_{RF,rms}}$$

moreover, one has that the output power is proportional to LO power (that is also called a *pump*), even if this dependency does not appear explicitly.

An important figure of merit concerning mixers is given by the -1dB compression point of the V_{in}/V_{out} characteristic. By increasing the value of the RF signal, the amount of harmonics at the output besides the fundamental IF frequency increases as well, eventually saturating the output signals (gain compression or flatness). This is due to the fact that power is no more mainly carried by the fundamental output tone (i.e. the IF tone), but it is rather shared by all the growing up harmonics. In other words the conversion gain stops to be constant, reduces and the output waveform is clipped. Supposing to have input and output impedance matching one has that:

$$\begin{aligned} A_{v,conv} &= \frac{V_{IF}}{V_{RF}} \\ 20 \log_{10} A_{v,conv} &= 20 \log_{10} \frac{V_{IF}}{V_{RF}} \\ 20 \log_{10} V_{IF} &= 20 \log_{10} A_{v,conv} + 20 \log_{10} V_{RF} \\ V_{IF}|_{dB20} &= A_{v,conv}|_{dB20} + V_{RF}|_{dB20} \end{aligned}$$

This suggests both that the ideal output characteristic looks linear in log scale until the gain does not compress and that the -1dB compression point can be defined as the output IF voltage at which:

$$V_{IF}|_{-1dB} = A_{v,conv}|_{dB20} + V_{RF}|_{dB20} - 1dB$$

hence:

$$V_{IF}|_{-1dB} = V_{IF}|_{dB20,ideal} - 1dB$$

Another figure of merit concerns the third order distortions, that can be defined in two condition:

Third Order Distortion in this case the input is a monochromatic signal.

Due to intrinsic non-linearity of the device the output contains frequency components not present at the input. It can be demonstrated that the most important spurious contribution on output fundamental tone comes from the third order term, in fact:

$$I_{D,TOD} \propto \cos^3(\omega_0 t) = \frac{1}{2} \cos(\omega_0 t) + \frac{1}{2} \cos(\omega_0 t) \cos(2\omega_0 t)$$

this presents in-band contribution, moreover the 1dB compression point mainly depends on TOD, then having good linearity is important to reduce this effect.

Third Order Intermodulation this distortion comes from a two-tone input injection ($f_1 = f_0$ and $f_2 = f_0 + \delta f$). Due to intermodulation between tones, spurious frequencies are generated. The most important contribution comes from the third order intermodulation (IM3): $|m| + |n| = 3$ and $mn < 0$ (where $m = \pm 2, \pm 1$ and $n = \pm 2, \pm 1$ are the intermodulation indexes). Taking $m = 2$ and $n = -1$:

$$f_{IM3}|_{m=2,n=-1} = 2f_1 - f_2 = f_0 - \delta f$$

that is clearly an in-band distortion. It can be demonstrated (FONTE) that, in case of same input and output reference impedance, the out and in intermodulation points are related to stage gain and output power of first harmonic:

$$\begin{aligned} OIP_3|_{dB} &= \frac{3}{2}P_{out,1st}|_{dB} - \frac{1}{2}P_{out,IM3}|_{dB} \\ IIP_3|_{dB} &= OIP_3|_{dB} - A_{v,1st}|_{dB} \end{aligned}$$

In case of mixing we change frequency band between input and output, but the previous result holds:

$$f_{IF,IM3}|_{m=2,n=-1} = 2f_{RF1} - f_{RF2} - f_{LO} = f_{IF} - \delta f$$

Also 5th order distortion exist, the same reasonings hold. Even order distortions (2nd, 4th ...) contributes in corrupt the output signals, as we will present further they are not very important though.

DA FINIRE CON FIGURE DI MERITO

2 Analysis of a Gilbert cell-based multiplier

2.1 Gilbert cell overview

A CMOS-based technology Gilbert cell-based mixer is shown in figure **AG-GIUNGI IMMAGINE**. This circuit exploits a differential topology to implement a double-balanced cell, providing:

- Reasonable conversion gain (from one to some tens);
- Good rejection of input frequency components to the output port along with high linearity, thanks to the double-balanced topology if performs²;
- Good isolation between ports is provided by the high suppression of spurious frequency components;
- Possibility of integration thanks to the CMOS architecture.

2.2 Gilbert cell circuit analysis

To analyse the circuit it is necessary to give some informations about its topology, biasing and driving. Looking at **FIGURA DOVE C'E' MIXER COMPLETO** it is possible to identify five main blocks that are described in what follows.

Bias net This net is made up of the current mirror (transistors M1 and M2) and the voltage reference generation branch (transistor M5, R1, R2, R3, R4, R5).

Transistor M1 is the strong branch of the current mirror and it acts as a current sink for the differential pair made up by M3 and M4, fixing the biasing current for the whole circuit. M2 sinks instead the current from the voltage reference section of the bias net. Generally transistors in current mirrors are polarized in saturation region, so that an high output resistance is seen from the stage above: this means that M1 should appear as a good current sink. Given $V_{G1} = V_{GS1}$ the gate voltage in M1 we have that the transistor remains in saturation if:

$$V_{GS} \geq V_{th} \quad (1)$$

$$V_{DS} > V_{od} = V_{GS} - V_{th} \quad (2)$$

The same holds for M2, that is always in saturation condition since it is diode connected (provided 1). In saturation (neglecting channel modulation

²The linearity is meant with respect to the conversion gain.

effects), one has that the drain current only depends on gate voltage and it is given by:

$$I_D = \mu_{n,eff} \frac{C_{ox}}{2} \frac{W}{L} (V_{GS} - V_{th})^2 = \frac{\beta_n}{2} V_{od}^2 \quad (3)$$

hence M1 remains in saturation till when:

$$V_{DS1} \geq V_{od1} = V_{th} - \sqrt{\frac{2I_0}{\beta_n}} \quad (4)$$

therefore to have a low value of the overdrive voltage one should have large transistors, i.e. large W . The output resistance is instead defined by:

$$r_o = \frac{1}{\lambda I_0} \propto L \quad (5)$$

where λ takes count of the channel modulation. It appears that we need long transistors to have better current sinking properties. By looking at FIGURE 2, one has (supposing both M2 and M1 in saturation, and $V_{GS1} = V_{GS2} = V_{DS1} = V_{DS2}$):

$$I_0 = \frac{\beta_{n1}}{2} (V_{GS1} - V_{th})^2 (1 + \lambda_1 V_{DS1})$$

$$I_{REF} = \frac{\beta_{n2}}{2} (V_{GS1} - V_{th})^2 (1 + \lambda_1 V_{GS1})$$

hence, supposing to have equal transistors:

$$\frac{I_0}{I_{REF}} = \frac{W_1/L_1}{W_2/L_2} \quad (6)$$

Ideally the current mirroring is only depending on geometrical parameters, meaning that a very good matching is required. Even if the two devices are very close to each other within the same chip, the possibility to have a variation in parameters that depend on temperature exists, i.e. V_{th} and $\mu_{n,eff}$. Supposing to have equal devices with everything constant but:

$$K_{n1} = K_n + \Delta K_n$$

$$V_{th,1} = V_{th} + \Delta V_{th}$$

one eventually finds that:

$$\frac{I_0}{I_{REF}} \simeq 1 + \frac{\Delta K_n}{K_n} + 2 \frac{\Delta V_{th}}{V_{od}}$$

Then, some of the possible mismatch causes are:

- K_n , that can varies a lot in case of wide circuits;

- V_{od} that usually is set low and varies with V_{th} , whose small variation can produce large differences in the two currents;
- difference between V_{DS1} and V_{DS2} .

Assuming as maximum variations:

$$\frac{\Delta K_n}{K_n} = \pm 5\%$$

$$\frac{\Delta V_{th}}{V_{GS} - V_{th}} = \pm 10\%$$

one has:

$$\frac{I_0}{I_{REF}} = 1 \pm 15\%$$

A cascaded voltage divider is connected as M2's load:

- M5 is diode connected and it is used to generate the right gate bias voltage for the gain stage.
- A resistive voltage divider made up of R2 and R4 is used to polarize the mixing stage.
- Resistors R1 and R3 are used to connect the bias net to the gain and mixing stages' gates. They also acts as high impedance for the RF and LO signals coming from outside the circuit and prevents them to enter the bias net.
- Capacitors C1 and C2 shunt possible non-DC disturbances coming from the Gilbert Cell, improving the bias isolation.

Noise issues will not be analysed within this treatise, even if it should be considered for a more complete design though. As a rule of thumb, narrow gate and low overdrive voltages should be used to lower noise contribution. (FONTE) Using a resistive voltage reference can be detrimental in case of voltage supply fluctuation, since a resistive voltage divider cannot reject current variations. However using active devices we would produce much more noise injection within the net though.

Gain stage The gain stage (shown in figure 3) is the linear part of the mixer, it must handle without distortion and corruptions the power coming from the RF signal giving some amplification. The topology is suited for a low noise amplifying purpose (FONTE): a differential stage with source degeneration. The stage is made up of two transistors (M3 and M4) in common source configuration, polarized with current by M1, and two source degeneration resistances R_S .

The RF differential signal modulates the current flowing into each of the two transistors, that have to be polarized in saturation region. For both of them it must be ensured an high output voltage range and a quite low overdrive voltage, along with a large drain-to-source bias voltage (required to maintain the stage in saturation during the swing of the LO stage).

To have an idea on the gain that can be obtained by this stage, one can look at the stage in FIGURE 3. Removing the degeneration resistances one has that the gate-to-gate mesh gives:

$$V_{in,1} - V_{in,2} = V_{GS3} - V_{GS4}$$

from equation 3, recalling that $I_1 + I_2 = I_0$ and squaring:

$$(V_{in,1} - V_{in,2})^2 = \frac{2}{\beta_n} (I_0 - 2\sqrt{I_1 I_2})$$

that once inverted yields:

$$I_1 - I_2 = \sqrt{\beta_n I_0} (V_{in,1} - V_{in,2}) \sqrt{1 - \frac{\beta_n}{4I_0} (V_{in,1} - V_{in,2})^2}$$

Writing the previous equation in term of $\Delta I = I_1 - I_2$ and $\Delta V_{in} = V_{in,1} - V_{in,2}$ and computing the value of the slope of this characteristic, one obtains that the *maximum* differential voltage gain in equilibrium condition $\Delta V_{in} = 0$ is given by:

$$|A_v| = \sqrt{\beta_n I_0} R_{L1} \quad (7)$$

This suggest that it is better to polarize the stage exactly with $V_{GS1} = V_{GS2}$. The complete expression for equation 7 also gives that the differential transconductance is a strongly non-linear function of the gate bias voltage.

As we said before there it is the possibility to have mismatch between the transistors and the loads due to circuit dimensions, temperature distribution and process tolerances. Then, looking at FIGURE 3, $M3 \neq M4$ and $R1 \neq R2$. It can be demonstrated that the offset output voltage due to differences in the circuit parameters and dimensions is:

$$V_{o,offset} = \Delta V_{th} + \frac{V_{GS} - V_{th}}{2} \left(- \frac{\delta R}{2R} \frac{\Delta W/L}{2W/L} \right)$$

To reduce this error a common centroid topology should be used in layout. Besides, it is important to notice that R_{L1} and R_{L2} are both the source output conductance of a common gate stage (mixing stage of the Gilbert cell) and that the actual gain for the RF stage is a current gain.

The small signal equivalent transconductance of the stage, once we added the source degeneration, (APPENDICE?) is given by:

$$G_{m,eq} = \frac{g_m}{1 + g_m R_S} \quad (8)$$

By adding R_S we lower the RF stage's current gain, however we also get a more linear behaviour reducing its dependency with respect to bias point (this fact will be demonstrated later). **LINEARITA' CON OVERDRIVE?? (RAZAVI rf PAG 193)**

The two transistors have the body contact connected to ground, this means that the threshold changes accordingly to the source to body voltage, V_{SB} :

$$V_{th} = V_{th0} + \gamma_B (\sqrt{2\phi_P + V_{SB}} - \sqrt{2\phi_P}) \quad (9)$$

This equation holds for all the transistors in the circuit, other than in current mirror.

Mixing stage This stage is also called *switching stage*, suggesting the way we want to drive transistors M6 to M9 (FIGURE 4). Also this one has a differential topology, however, differently from the RF stage, we drive each pair with complementary LO signals that turn on and off cycle-by-cycle one of the two devices. The driving signal is sinusoidal (similarly one can apply a square wave) and it must be large enough to ensure the abrupt switching of the stage: when a device is on it must be in saturation (not in triode! These are actually *switch-like* transistors), while the other one must be completely off, interdicted. A compromise is needed since:

- small amplitude of LO signals produce slower switching speed, hence part of the power is wasted as common mode signal at the output;
- too large LO signals drive the MOSFET in triode region producing spikes that corrupt the output signal with unwanted feed-through and reduces speed.

Switching speed is related to the device cut-off frequency, that for short channel devices is inversely proportional to channel length and linearly dependent to the overdrive voltage:

$$f_T \propto V_{od}/L$$

suggesting that the highest overdrive and the minimum channel length should be chosen. Again, a compromise must be found among the following solutions:

- Setting the available length (the minimum is fixed by the used design kit);
- Selecting higher overdrive voltage one has the possibility to increase f_T , however to keep our switches in saturation and to obtain a large output swing we need to keep small this quantity;
- Widening the gate width one has faster device, but the source to bulk capacitance increases, eventually shunting the RF signal coming from the LNA stage to ground;

- Reducing the current a faster commutation is achieved, along with lower transconductance though.

Even if the switching stage is not linear it is possible to define its gain and in order to do that we analyse one single pair. We recall that the gate LO signal modulates the stage drain current as a square wave, that can be written in term of its Fourier's series:

$$i_D(t) = I_{pk} \left(\frac{1}{2} - \frac{2}{\pi} \sum_{n=1,3,5\dots} \frac{1}{n} \sin(n\omega t) \right)$$

where only odd harmonics are present. The instantaneous gain fo the stage, related to the first harmonic is:

$$A_c|_{switch} = \frac{2}{\pi} \quad (10)$$

that is neither dependent on the bias of the stage nor on the actual device transconductance.³ It is important to notice that equation 10 is exact in case of gate's square wave signals, whereas in case of sinusoidal driving the actual gain will be reduced because of the smoother variation of the drain current waveform.

Load stage The load stage is formed by two resistors R_L whose role is to give the stage enough gain and to set the output bias voltage through the RF drain current. They must be as matched as possible to make the differential outputs of the Gilbert cell equals.

2.3 Mathematical analysis: conversion gain

As already stated a mixer is a non-linear time variant system, hence a descriptive approach based on the behaviour of the cell is well suited to find the conversion gain of the cell.⁴ The following analysis is based on the schematic in **FIGURE 6**.

We suppose all the transistors in gain stage and bias in saturation, then we identify:

- i_3 and i_4 the current through M3 and M4;
- i_6 and i_7 the current through M6 and M7;
- i_8 and i_9 the current through M8 and M9;

³It is however important to remember what stated above, regarding the intensity of the applied gate voltage, that further reduces this quantity in case of too large input signals. Moreover its better to notice that the stage introduces a loss rather than an actual gain.

⁴It is not possible to use neither the superposition of effect nor the Laplace transform to compute the gain due to the presence of the switching pairs.

- i_{o1} the current through R_{L1} ;
- i_{o2} the current through R_{L2} ;
- I_D the bias current;
- i_{od} the output differential bias current;
- $S(t) = \frac{1}{2} - \frac{2}{\pi} \sum_{n=1,3,5,\dots} \frac{1}{n} \sin(n\omega_{LO}t)$ is the switching function of the mixing cell (then we consider the LO stage as it was made up of ideal switches).

The differential input RF signal is given by:

$$v_{RF}(t) = \frac{V_{RF}}{2} \cos(\omega_{RF}t) \quad (11)$$

Currents flowing into the two RF stages are given by:

$$i_3(t) = I_D + i_d(t) \quad (12)$$

$$i_4(t) = I_D - i_d(t) \quad (13)$$

where, from equation 8, the i_d current is a small signal variation given by:

$$\begin{aligned} i_d(t) &= I_{RF} \cos(\omega_{RF}t) \\ &= G_{m3,4} \frac{V_{RF}}{2} \cos(\omega_{RF}t) \end{aligned} \quad (14)$$

It is now possible to define the current flowing into the switching stage recalling that they are discontinuous functions modulated by the switching function $S(t)$:

$$i_6(t) = [I_D + i_d(t)]S(t) \quad (15)$$

$$i_7(t) = [I_D + i_d(t)]S(t - T_{LO}/2) \quad (16)$$

and

$$i_8(t) = [I_D - i_d(t)]S(t - T_{LO}/2) \quad (17)$$

$$i_9(t) = [I_D - i_d(t)]S(t) \quad (18)$$

because M6 and M9 share the same LO signal; the same holds for M7 and M8. Hence the current through R_{L1} and R_{L2} is given by:

$$\begin{aligned} i_{o1}(t) &= I_6(t) + I_8(t) \\ &= [I_D + i_d(t)]S(t) + [I_D - i_d(t)]S(t - T_{LO}/2) \end{aligned} \quad (19)$$

and

$$\begin{aligned} i_{o2}(t) &= I_7(t) + I_9(t) \\ &= [I_D + i_d(t)]S(t - T_{LO}/2) + [I_D - i_d(t)]S(t) \end{aligned} \quad (20)$$

Therefore, the output differential current is:

$$\begin{aligned} i_{od}(t) &= I_{o1}(t) - I_{o2}(t) \\ &= 2i_d(t)S(t) - 2i_d(t)S(t - T_{LO}/2) \end{aligned} \quad (21)$$

Noticing that:

$$\begin{aligned} S(t - T_{LO}/2) &= \frac{1}{2} - \frac{2}{\pi} \sum_{n=1,3,5\dots} \frac{1}{n} \sin(n\omega_{LO}t - T_{LO}/2) \\ &= \frac{1}{2} + \frac{2}{\pi} \sum_{n=1,3,5\dots} \frac{1}{n} \sin(n\omega_{LO}t) \\ &= 1 - \frac{1}{2} + \frac{2}{\pi} \sum_{n=1,3,5\dots} \frac{1}{n} \sin(n\omega_{LO}t) \\ &= 1 - S(t) \end{aligned} \quad (22)$$

and substituting equation 22 in 21⁵ we get:

$$\begin{aligned} i_{od}(t) &= -\frac{8i_d(t)}{\pi} \sum_{n=1,3,5\dots} \frac{1}{n} \sin(n\omega_{LO}t) \\ &= -\frac{8I_{RF}}{\pi} \cos(\omega_{RF}t) \sum_{n=1,3,5\dots} \frac{1}{n} \sin(n\omega_{LO}t) \\ &= -\frac{8I_{RF}}{\pi} \sum_{n=1,3,5\dots} \frac{1}{n} \cos(\omega_{RF}t) \sin(n\omega_{LO}t) \\ &= -\frac{8I_{RF}}{\pi} \sum_{n=1,3,5\dots} \frac{1}{2n} \left\{ \sin[(\omega_{RF} + n\omega_{LO})t] + \sin[(\omega_{RF} - n\omega_{LO})t] \right\} \end{aligned} \quad (23)$$

Considering only the fundamental tone of the LO signal (other frequency terms are out-of-band with respect to IF) and multiplying by $R_L = R_{L1} = R_{L2}$ one finds that the differential output voltage is given by:

$$\begin{aligned} v_{od}(t) &= -\frac{4I_{RF}R_L}{\pi} \{ \sin[(\omega_{RF} + \omega_{LO})t] + \sin[(\omega_{RF} - \omega_{LO})t] \} \\ &= -\frac{4I_{RF}R_L}{\pi} [\sin(\omega_{HF}t) + \sin(\omega_{IF}t)] \end{aligned} \quad (24)$$

then, referring to the IF component:

$$\begin{aligned} v_{od,IF}(t) &= -\frac{4I_{RF}R_L}{\pi} \sin(\omega_{IF}t) \\ &= -\frac{2G_{m3,4}R_L V_{RF}}{\pi} \sin(\omega_{IF}t) \end{aligned} \quad (25)$$

⁵ $\cos(x)\sin(y) = \frac{1}{2}\sin(x+y) + \frac{1}{2}\sin(x-y)$

that, considering the IF and RF *rms* values eventually yields the Gilbert cell mixer conversion gain:

$$A_{vC} = \frac{2}{\pi} \frac{R_L}{\frac{1}{g_{m3,4}} + R_S} \quad (26)$$

It is now possible to highlight that:

- in first approximation the conversion gain does not depends on the amplitude of the LO signal;
- both LO and RF components are rejected at the output;
- it is possible to keep only the IF component by filter out the HF signal;
- the degeneration resistance R_S improve the stage's linearity. In fact, if properly chosen, one has: $A_{vC}|_{g_{m3,4} \ll R_S} \simeq \frac{2}{\pi} \frac{R_L}{R_S}$. Besides to linearity, the conversion gain is less sensitive with respect to the RF stage bias point;
- no informations about frequency behaviour appear with this kind of analysis.

3 Design by hand of a down-converting Gilbert cell

3.1 Design specifications

The design specifications follow:

- The use technology is the AMI $0.6\mu\text{m}$ by MOSIS. **No constraints regarding maximum gate width are considered. This technology is nowadays old-fashioned, hence this project has only didactic purposes.** The Cadence Design Environment will be used to carry on the project;
- the supply voltage is $V_{DD}=5\text{V}$;
- we are design a down-conversion mixer. The analysis will be carried on considering two monochromatic signals $f_{RF}=110\text{ MHz}$ and $f_{LO}=100\text{ MHz}$, producing a wanted baseband signal at $f_{IF}=10\text{ MHz}$ (the HF component is considered as filtered out);
- all the transistors should work in saturation;
- the conversion gain is chosen accordingly to the previous specification equal to $G_{vC}=15\text{dB}_{20}$.

The most important parameters used in the design are reported in the following table: ⁶

Table 1

Parameter Name	Value	Unit
A_{vC}	5.6	
V_{DD}	5	V
I_0	5	mA
V_{th0}	0.709	V
K_n	116	$\mu\text{A}/\text{V}^2$
I_{dss}	466	$\mu\text{A}/\mu\text{m}$
ϕ_P	0.7	V
γ_B	0.5	V
L_{min}	0.6	μm

It is necessary to notice that the following calculation is carried on considering the easiest between the physic-based models for a MOSFET, i.e. the level 1 model and its aim is to make the reader to understand the

⁶MOSFET's parameters are intended at $T=27^\circ\text{C}$

design choices that follow in the next section. Moreover, the used model relies on a level 3 model, that takes count of short channel effects (we are using sub-micrometric devices) and non-ideality such as the carrier velocity saturation.

3.2 Gilbert cell design flow

Current mirror bias Both M1 and M2 must be in saturation. We choose the same current to flow in the weak and strong branches in order to have equal structures (this aids the circuit symmetry and matching, allowing us to use interdigitated structure). We impose:

$$I_{D1} = I_{D2} = I_0 = 5mA \quad (27)$$

$$V_{od1} = V_{od2} = 0.4V \quad (28)$$

$$L_1 = 3L_{min} = 1.8\mu m \quad (29)$$

The current magnitude is chosen thinking about the mixer used in a transmitter receiving chain, after an hypothetical power stage. The overdrive voltage's value is chose in order to keep turned on the stage even with strong fluctuation of the drain node. However, to have the same current mirrored we must have:

$$V_{GS1} = V_{GS2}$$

$$V_{DS1} = V_{DS2}$$

$$W_1 = W_2$$

suggesting that this kind of current source can work properly only with small signal variation coming from the RF stage. Now, by using 3, neglecting channel modulation effects and inverting:

$$W_1 = \frac{2I_0}{K_n V_{od1}^2} L_1 = 969.8\mu m \quad (30)$$

Both M1 and M2 are not affected by body effect, then:

$$V_{th1} = V_{th2} = V_{th0}$$

hence:

$$V_{GS2} = V_{DS2} = V_{th0} + \sqrt{\frac{2I_0 L_1}{K_n W_1}} = 1.1V \quad (31)$$

The design for the voltage reference bias net is reported in section 4.

Load stage and gain choice As previously stated we chose $A_{vC}=4$. Since the whole differential structure must be symmetrical, current I_0 is evenly split into M3 and M4: this is the same current flowing also into the two loads. In order to give the LO stage **enough output swing** we impose one third of the supply voltage must drop on the load. Therefore, we can calculate the load resistance's value:

$$R_L = \frac{\frac{2}{3}V_{DD}}{I_0/2} = 1.33k\Omega \quad (32)$$

Gain stage bias The design of the RF bias begins with the choice of its transconductance, fixed by A_{vC} , R_L and R_S . We consider M3 and M4 equal and in saturation. The gate length for this stage is set

$$L_3 = L_4 = 3L_{min} = 1.8\mu m \quad (33)$$

The degeneration resistance's value is chosen equal to 10Ω , hence:

$$V_{R_S} = \frac{R_S}{I_0/2} = 25mV \quad (34)$$

Now, by inverting equation 26:

$$g_{m3} = \frac{\pi}{2} \frac{1}{\frac{R_L}{A_{vC}} - R_S} = 6.9mS \quad (35)$$

We notice that we have many variables that are involved and that directly affect the design feasibility:

- high values of R_S decrease the RF stage transconductance, requiring large transistors with low overdrives. Low values give instead more transconductance and linearity reducing the output dynamic though;
- increasing the conversion gain we have the same effects reported before, however mixing signals is a strongly inefficient operation, then we cannot reduce too much this quantity;
- the choice of the load is important to fix the output dynamic and must be chosen carefully.

By inverting equation 7:

$$W_3 = g_{m3}^2 \frac{2L_3}{K_n I_0} = 148.5\mu m \quad (36)$$

We have body effect due to the source to body voltage, shared by M3 and M4:

$$V_{SB3} = V_{DS1} + V_{R_S} \quad (37)$$

then, from equation 9:

$$V_{th3} = V_{th0} + \gamma_B(\sqrt{2\phi_P + V_{SB3}} - \sqrt{2\phi_P}) = 0.957V \quad (38)$$

from which it comes that:

$$V_{od3} = \sqrt{\frac{I_0}{K_n W_3/L_3}} = 0.723V \quad (39)$$

$$V_{GS3} = V_{th3} + V_{od3} = 1.731V \quad (40)$$

Unfortunately the value of the overdrive voltage looks large (**FORTE**), requiring us to fix $V_{DS3,4}$ high enough to avoid the stage to be turned off by LO's output swing. We impose to have half the supply voltage to drop on the gain stage and current mirror, having:

$$V_{DS3} = \frac{1}{2}V_{DD} - V_{DS2} = 1.4V \quad (41)$$

To ensure saturation of M3 and M2 we set:

$$V_{G3} = V_{GS3} + V_{RS} + V_{DS1} = 2.805V \quad (42)$$

Mixing stage bias Also in this case we require transistors from M6 to M9 to be in saturation (they are equal and biased the same way). We reduce as much as possible the overdrive voltage: in this way the stage is kept nearby the threshold fostering a the commutation of the switches. Having small overdrives we produce large transistors then, to limit this we set:

$$L_6 = L_{min} \quad (43)$$

$$V_{od6} = 150mV \quad (44)$$

By using 3 and taking in count the body effect we get:

$$W_6 = 1.1mm \quad (45)$$

$$V_{SB6} = V_{DS1} + V_{RS} + V_{DS2} \quad (46)$$

$$V_{th6} = 1.18V \quad (47)$$

This result does not consider short channel effects, that produces a different threshold dependency on source to bulk voltage. However, considering ideal behaviour we have that the M6 to M9's bias gate voltage is:

$$V_{GS6} = 1.326V \quad (48)$$

$$V_{G6} = V_{GS6} + V_{DS1} + V_{RS} + V_{DS2} = 3.855V \quad (49)$$

4 Design by simulation

4.1 Gilbert Cell design

After simulating the design made with Level 1 model, simulation results proven to be not enough accurate to have a correctly behaving circuit. For this reason a new design was carried on taking advantage of the simulated trans-characteristic of each stage.

In literature (**FONTE**) an optimum value for transistors width in mixers is proposed:⁷

$$W_{opt} = \frac{1}{3\omega L C_{ox} R_g} \quad (50)$$

where ω is the stage's working frequency (i.e. f_{LO} or f_{RF}), L is the channel length, C_{ox} is the MOSFETS's total oxide capacitance and R_g the generator output resistance (we set this value equal to 50Ω). Values coming from this formula are actually too large ($\geq 1mm$), therefore we chose to pursue another approach to begin the design.

The starting specification was to maximize g_{m3} of the input RF stage, with some reasonable hypotheses on transistor dimensions and power consumption.

$$\begin{aligned} L &= 3L_{min} = 1.8\mu m \\ 50\mu m &\leq W_3 \leq 500\mu m \\ I_0 &\approx 5mA \end{aligned}$$

Further, suppositions on bias voltages were made. For example, a voltage equal to $\frac{3}{5} \cdot V_{dd}$ was supposed to drop between the RF stage's drain and ground, equally divided between M1 and M3 (neglecting the very small voltage drop on R_S).

$$V_{SB3} = V_{DS1} = V_{DS3} = 1.5V$$

This way allows us to take into account the body effect on threshold voltage of transistor M3. Then we plotted the current and transconductance of the transistor as a function of V_{GS} and V_{DS} , that are shown in pictures 1a and 1b.

⁷More generally, this gives a compromise between power dissipation and noise in LNA.

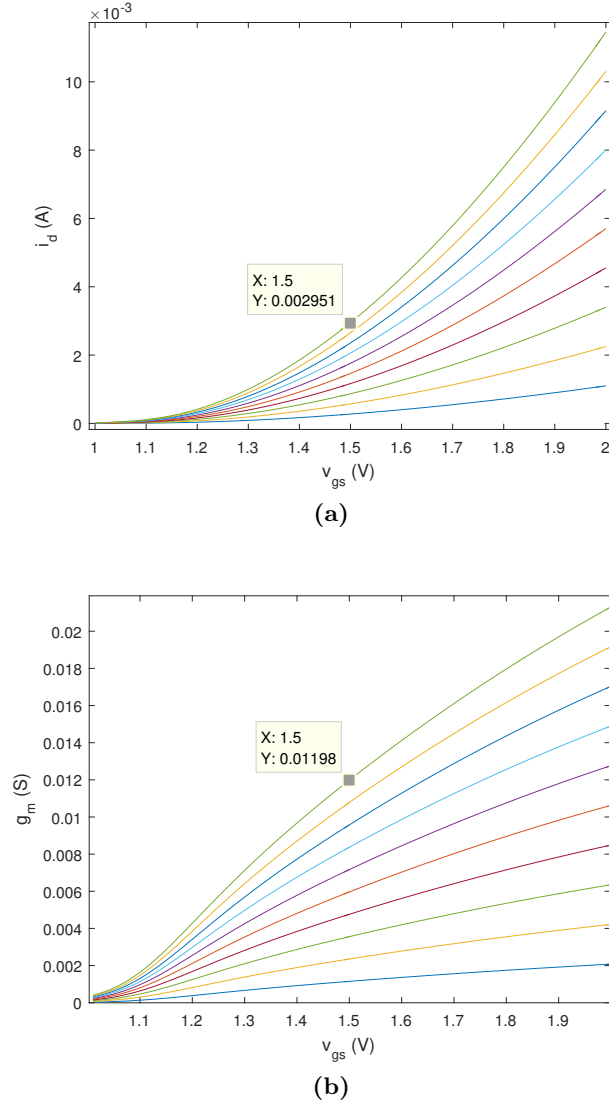


Figure 1: a) Drain current versus gate-source voltage of the RF stage, with W varying from $50\mu m$ to $500\mu m$. b) Transconductance versus gate-source voltage of the RF stage, with W varying from $50\mu m$ to $500\mu m$.

The largest width $W_3 = 500\mu m$ (the green curve) was chosen, along with $V_{GS3} = 1.5V$, accordingly to the current and transconductance's value, eventually taking the highest values:

$$g_{m3} = 11.9mS$$

$$\frac{I_0}{2} = 2.9mA$$

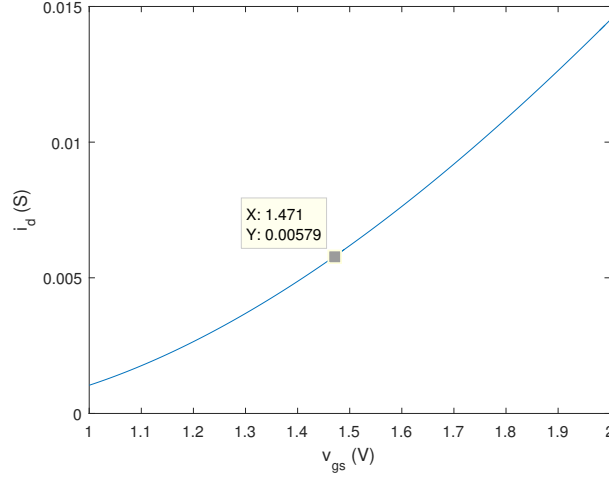


Figure 2: Drain current versus gate-source voltage of the bias transistor M1, with $W=373\mu m$

We moved then to current sink design, M1 and M2. Given the data:

$$\begin{aligned} V_{R_S} &= 2.9mA * 10\Omega = 29mV \\ V_{DS1} &= 1.5V - V_{R_S} = 1.471V \\ V_{GS1} &= 1.471V \end{aligned}$$

This set of values leads to a width equal to $W_1 = 373\mu m$ in order to have $I_0 = 2 \cdot 2.9mA = 5.8mA$ (figure 2).

Finally the LO stage was designed, along with the load resistance R_L starting from the wanted conversion gain, lowered to $A_v = 4$ with respect to what we decided in the previous section. Given the analytic expression of voltage conversion gain 26 (**FORTE**):

$$A_v \approx \frac{2}{\pi} \left(\frac{R_L}{R_S + \frac{1}{g_{m3}}} \right) = 4$$

The value of R_L is then determined

$$R_L = A_v \cdot \left(\frac{\pi}{2} \cdot \frac{1}{g_{m3}} + R_S \right) = 577\Omega \quad (51)$$

The drain-to-source voltage of the LO stage, V_{DS6} , can be thus evaluated as follows:

$$\begin{aligned} V_{SB6} &= V_{DS1} + V_{R_S} + V_{DS3} \\ &= 1.47V + 0.029V + 1.5V = 3V \end{aligned}$$

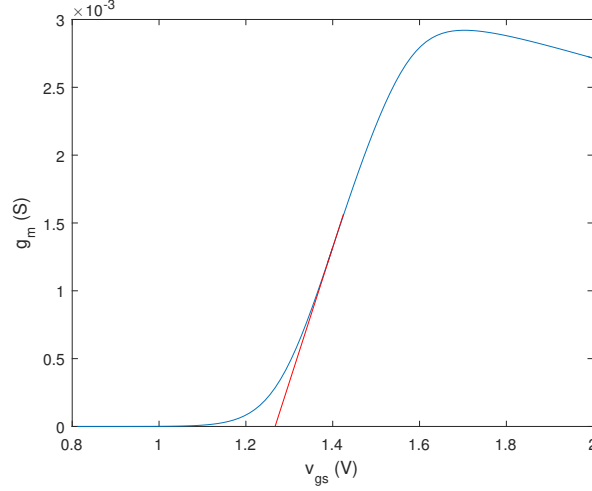


Figure 3: Extrapolation of M6 threshold voltage from transconductance versus the V_{GS} curve. The threshold happens to be at $1.27V$.

and

$$V_{RL} = 2.9mA \cdot 577\Omega = 1.673V$$

$$V_{DS6} = V_{dd} - V_{RL} - V_{SB6} = 327mV$$

The LO gate bias voltages must be slightly above the threshold, in order to let the transistors turn on and off with small LO signal variations, deviating rapidly current coming from the RF stage. To accomplish this, the transistor's threshold was extracted from simulation, using the g_m graph as a function of V_{GS} (figure 3). Having $V_{th6} = 1.27V$, a very small overdrive of $\Delta V_6 = 60mV$ was chosen, for the reason reported before, and to assure the stage is not working in triode region. This leads to:

$$V_{GS6} = V_{th6} + \Delta V_6 = 1.33V$$

Once fixed the node's voltages and the current flowing in the transistor ($2.9mA$, due to the fact that only two transistors of LO stage conduct simultaneously), transistor width W_6 was swept to fulfil precisely this biasing. The followed procedure is the same displayed in figure ?? and the resulting found dimensions are:

$$W_6 = 170.3\mu m$$

$$L = L_{min} = 0.6\mu m$$

Only for this stage the minimum channel length was taken, in order to keep small these transistors regardless overdrive small value.

4.2 Biasing network design

From the previous section, the bias network specifications required by current sink, RF and LO stage are suddenly imposed:

$$V_{G1} = 1.471V$$

$$V_{G3} = 3V$$

$$V_{G6} = 4.33V$$

The bias network circuit employed is visible in **figure**.

The mirroring ratio for the current sink transistor M1 was chosen to be 1 as chosen in section 3. With a transistor length $L_2 = 1.8\mu m$, and being $V_{GS2} = V_{DS2} = 1.471V$ imposed, W_2 has been trimmed in the simulator to sink a current equal to $I_0 = 5.8mA$, resulting in:

$$W_2 = 373\mu m$$

In the same way transistor M5 width was set, in order to have a gate voltage $V_{G5} = 3V$ when stacked above M2:

$$W_5 = 130.45\mu m$$

$$L_5 = 0.6\mu m$$

Minimum length was taken in this case to reduce the transistor's size. The sum of R_2 and R_4 must be such that with total current of $5.8mA$, the voltage drop across them is such that $V_{G5} = 3V$:

$$R_2 + R_4 = \frac{V_{dd} - V_{G5}}{I_0} = 344\Omega$$

The partition between them must give a bias voltage to LO stage of $4.33V$. This brings to

$$R_2 = 229\Omega$$

$$R_4 = 115\Omega$$

Resistors R_1 and R_3 were chosen to be large enough to isolate biasing network from RF and LO signals respectively, and form a low pass filter along with capacitors C_1 and C_2 . Resistors were thus chosen of the order of some $k\Omega$.

$$R_1 = R_3 = 30k\Omega$$

From here, the minimum capacitance to filter out LO signal can be evaluated in order to have a pole at least one decade before working frequency. The

output resistance of M5 has been considered negligible with respect to R_1 and R_3 , leading to an equivalent resistance for the low pass filter equal to:

$$R_{eq} = R_1 // R_3 // R_2 // R_4 = 76.4\Omega$$

from which the resulting capacitance:

$$C_1 \geq 10 \cdot \frac{1}{2\pi \cdot R_{eq} \cdot f_{lo}} = 20.8pF \quad (52)$$

$C_1 = C_2$ was chosen for simplicity. Finally a CAD optimization over them was carried on to maximize the gain and minimizing the capacitors' dimensions, leading to the final chosen value of:

$$C_1 = C_2 = 25pF \quad (53)$$

5 Layout of the Gilbert Cell

The overall layout was carried on in place with a view to:

- reducing technology gradients using common centroid, interdigitated or multi-finger structures;
- placing components all in the same direction for a uniform error affection;
- reducing border effects with dummy elements;
- minimizing encroachment errors with same-length transistor fingers;
- limiting substrate noise with guard rings;
- avoiding connections which could lead to crosstalk;
- minimizing metal changes in the routing process;
- aiming a compact structure.

5.1 Differential RF stage

A careful attention has been paid to the layout of differential pairs like the one of RF stage. Since it's desirable to have the most symmetric input stage a common centroid configuration has been employed, following a structure sketched as in **FIGURE**.

This minimizes errors due to process gradients within the planar structure. Dummy elements were also placed next to the four transistors at the edges to guarantee a uniform dopant concentration of the inner transistors.

Dummy MOSFETs' drain and source are tied together and then short circuited to source. The transistor length of the differential pair has been divided to form multi-finger transistors, such that the whole structure results as square and compact as possible. This was intended again for the purpose of reducing gradients and with a glance forward to the overall mixer layout's compactness. The final layout of differential RF pair is shown in figure.

Drain, source, and ground nets disposal is highlighted in **figure**.

5.2 Differential LO stages

The two differential LO stages have been designed trying to prevent the same mismatch causes seen for the RF stage. Two common centroid structures were drawn at first. Being the transistor wide, multi-finger structure was employed also here to have an almost-square layout along with dummy elements. Attention has been paid in so that the two LO pairs could be easily stack above the RF stage with the same width.

Secondly, we opted for a merging of the two common centroid LO layouts in a more compact tiled layout. The gate disposal simplifies the access to these nets from the signal capacitors in the global layout. Final LO layout and highlighted nets are shown in **figure**.

5.3 Current mirror and diode connected transistor

Current mirror employs an interdigitated structure similar to the one visible in **figure**.

Since fingers are all the same size the current ratio is independent from encroachment width. Dummy transistors was included to reduce border effects, and source connections on both top and bottom sides have been added for a small source resistance. The weak branch transistor's drain has been tied together to the gate with an array of contacts. The final layout for current mirror is visible in **figure** and highlighted nets in **figure**.

Concerning the diode-connected transistor M5, interdigitated structure was implemented. Dummy transistors are placed aside whereas an array of contacts to minimize series resistance between gate and drain has been added. Final layouts and highlighted nets are visible in **figure**.

5.4 Capacitors

Bias capacitors have been designed using two configurations of common centroid layout.

Bias capacitors Bias capacitors layout structure is drawn in **figure**. Even if no exact ratio between the two capacitors is needed, a modular geometry, made by equal squares, was used. This minimizes undercut effects, and the square shape reduces border errors. In order to reduce the area, capacitor units were designed using poly1 and poly2 (elec) layers, which give the maximum capacitance per unit area available with this technology kit. A dummy capacitor frame was placed around to minimize border effect. All the capacitors lie above an n-well, whose purpose is to minimize fringing field leakage. This n-well is tied to V_{DD} potential. The final layout with highlighted nets is visible in **figure**.

Signal capacitors Since these capacitors resulted to be very space demanding due to capacitance value, a multi-layer capacitor was tried to be built alternating poly and metal layers. Unfortunately this idea was soon abandoned due to the *rule 11.6 of SCMOS* rules, which does not allow unrelated metal layer over poly2. The most practical solution proven to be using poly1-to-poly2 capacitance in a common centroid structure. This time matching is more important with respect to bias capacitors, because we want signal coming to the differential stages pass through a strongly symmetrical

paths. The common centroid structure has been implemented and it is shown in **figure**. Since this is a high frequency path to reduce series resistance and inductance, metal plates with contacts were placed all over the square unit capacitors. Dummy short circuited elements were positioned all around the capacitor as well. An additional n-well connected to V_{DD} surrounds the capacitor to reduce fringing field leakage. The overall capacitance layout with highlighted nets is visible in **figure**.

5.5 Resistors

Two kinds of resistors have been implemented in this layout, both of them are described in the following paragraphs.

N-well resistors Resistors fabricated in n-well have worse tolerance, but an higher sheet resistance. That's the way this kind of resistor was chosen to implement biasing resistors R_1 and R_3 , which are, as a matter of fact, very large in absolute value, but their tolerance is not too important. We can thus have smaller layout adopting n-well layout. Resistors were split in modular structures with the common centroid fashion to reduce differences between R_1 and R_3 . Due to the fact that this resistors are very close to the mixing stage, a guard ring connected to ground has been included to prevent substrate noise being injected through the biasing network. The final layout obtained for R_1 and R_3 resistor is visible in **figure**.

Poly silicon resistors Since other resistors are smaller than R_1 and R_3 , poly has been employed to build them, taking advantage of their higher precision but smaller sheet resistance. All of them were drawn in common centroid structure, except for R_4 (due to its small value), and dummy elements on both sides were always used. Final layouts for these resistors are shown in **figure**.

5.6 Putting together the layout

The layout has been organized in order to minimize the occupied surface area. The final configuration can be seen in **figure**. Components have been placed to shorten interconnections, emphasize symmetries for matching and to make paths as equal as possible for high frequency signals. Body contacts where placed all around the circuit where possible, and then connected to ground, in order captures free charges in the substrate reducing noise.

6 Simulation vs schematic comparison

7 Conclusions