University of Padova

Department of
Information Engineering

DEGREE THESIS

Amplifier Design for a Pipeline ADC in 90nm Technology

SUPERVISOR: Prof. Andrea Gerosa

CANDIDATE: Alessandro Michielin

Padova, 5th October 2010

I would like to thank my family for the constant support, prof. Gerosa for the guidance throughout the whole Master Degree, and the nSilition staff; in particular Thierry and Laurent for their continuous technical and logistic support on my internship period.

Summary

Abstract

This paper explains the choices taken for the design of two full differential operational amplifiers. These op amp have been designed for the third and the fifth stage of a pipelined A/D Converter. It shows also the solutions found to reach high gain, wide bandwidth and short settling time, without degrading too much the output swing.

First the operational amplifier specification are extracted starting from the ADC architecture, then the issues related to the sub-micrometrical design are analysed; the different structures tested are then presented and the motivation of the final topology choice are shown. It presents then the op amp schematic implementation, the simulation results and the layout with the 90nm TSMC design kit.

Chapter 1

Introduction

1.0 Introduction

Operational Amplifiers are one of the most widely used building blocks for analog and mixed-signal systems. Nowadays, complementary metal-oxide semiconductor technology has become dominant over bipolar technology for analog circuit design in a mixed-signal system due to the industry trend of applying standard process technologies to implement both analog circuits and digital circuits on the same chip.

While many digital circuits can be adapted to a smaller device level with a smaller power supply, most existing analog circuitry requires considerable change or even a redesign to accomplish the same feat. With transistor length being scaled down to tens of nanometers, analog circuits are becoming increasingly more difficult to improve upon, in fact, if small geometries can improve speed decreasing the parasitic capacitance, the gain can be heavily affected. So, gain enhancement techniques are required, but these methods often require more complicated circuit structures and higher power supply voltage, and may produce a limited output voltage swing or introduce a significant noise contribution.

This thesis summarizes the work produced during a six months stage by nSilition sprl, a fabless company specialized in the design of high performances, low power converters. The design object was a 14bit, 200MS/s ADC. The converter is

implemented with a six stages pipeline architecture; the design is based on switch-capacitor circuitry. Each stage consists of an OTA and a subADC, and stage 5 and stage 3 OTA are the main objects of this dissertation. The devices are implemented through TSMC 90nmRF process technology.

Analog circuit design requires a good understanding of how the system and circuit work. Unlike digital circuitry which works with two distinct states, many parameters are under consideration for analog circuits which work with continuous values. Due to the multi-dimensional variables of an analog circuit, any slight change in the analog configuration like current, voltage, a transistor parameter, a device model, a manufacturing process, or a modified layout may cause significantly different performance. For analog design engineers, a good design methodology including intuition, mathematical methods, and specialized tools are assets. The design tools consists on Virtuoso Front to Back Design Environment for the schematics and layout, Matlab and Excel for the specifications extraction.

All the specifications required have been met.

1.2 Thesis organization

The thesis is organized into eight chapters.

- Chapter 1 introduces the problem.

- Chapter 2 reviews the basic theory of A/D converters and the principle of the pipelining; the main characteristics of the converter of the project are described as well as the methodology used for the extraction of the amplifier specifications in each stage.

- Chapter 3 describes the main side effects related to the use of short channel devices and how they will affect the schematic modelization and the layout.

- In Chapter 4 the design of two differential amplifiers is discussed. The different trade-off between gain, bandwidth and stability are presented, and the chosen solutions are explained.

- In Chapter 5 the designs of the common mode regulators are discussed. Several different architectures are presented, the one chosen is described as well as the modifications applied to the main amplifier to reach the stability specifications.

- Chapter 6 deals on the simulation sets: the testbenches are presented as well as the simulation results.

- In Chapter 7 the layout work is shown. The main source of issues are presented as well as the solution chosen. The whole amplifier layout is shown.

- Chapter 8 analyses the power consumption and describes the possible future works

Chapter 2

Pipeline converters

2.0 Ideal A/D Converter

An analog-to-digital converter performs the quantization of analog signals into a number of amplitude-discrete levels at discrete time points. A basic block diagram of an A/D converter is shown in Fig. 2.1.

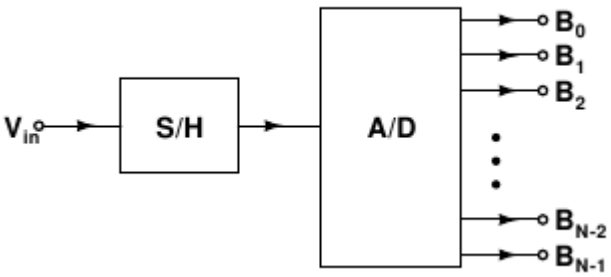

Fig 2.1: *basic A/D converter*

A sample-and-hold amplifier is added to the input to sample the analog input and to hold the signal information at the sampled value during the time needed for the conversion into a digital number. The analog input value V_{IN} is converted into an N-bit digital value using the equation

$$\frac{V_{in}}{V_{ref}} = D_{out} + e_q = \sum_{m=0}^{N-1} B_m 2^m + e_q \qquad (2.1)$$

In the equation, R_{ref} represents a reference value, which may be a reference

voltage, current or charge. B_{N-1} is the most significant bit and B_0 is the least significant bit of the converter. The quantization error e_q represents the difference between the analog input signal V_{in} divided by R_{ref} and the quantized digital signal D_{out} when a finite number of quantization levels is used. Eq. 2.1 can be partly rewritten as

$$D_{out} = \sum_{m=0}^{N-1} B_m 2^m \tag{2.2}$$

The sampling operation of analog signals introduces a repetition of input signal spectra at the sampling frequency and multiples of the sampling frequency. To avoid aliasing of the spectra, the input bandwidth must be limited to not more than half the sampling frequency (Nyquist criterion).

2.1 Pipeline ADC

The pipelined is a popular architecture for modern applications of analog-to-digital converters due to its high sustained sampling rate, low power consumption, and linear scaling of complexity. Figure 2.2 shows a block diagram of a pipelined ADC. The term "pipelined" refers to the stage-by-stage processing of an input sample V_{IN}.

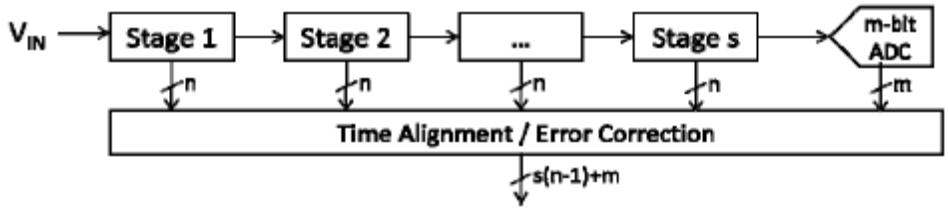

Fig 2.2: *basic pipeline ADC architecture*

In the above diagram, the analog input voltage V_{IN} enters the ADC. Each subsequent pipeline stage of the ADC resolves a certain n number of bits to be contributed to the final conversion output. The number of bits that each stage is responsible for quantizing is usually on the order of 1–5 bits. Simultaneously, after each stage has finished quantizing its input sample to n bits, it outputs an analog residue voltage that serves as the input to the next stage. After s stages of conversion, an m-bit ADC resolves the lower bits of the overall ADC digital output.

Each stage's digital decision is then passed to a digital block that properly time-aligns the output bits and corrects for any errors in each stage. The final digital decision

is then produced.

2.3 MDAC operation

Each stage displayed in the block diagram shown above can be explored further. A typical pipeline stage is displayed in Fig. 2.3.

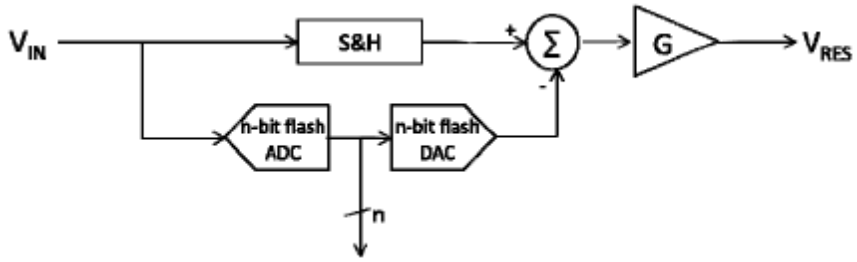

Fig 2.3: *MDAC architecture*

The input voltage is sampled and held in the sample-and-hold circuit embedded in each stage. Subsequently, an n-bit flash ADC quantizes the analog voltage and produces a digital decision of n bits. The digital decision is then fed through an n-bit flash DAC to be re-converted into an analog signal. The summation node presented in the above diagram takes the input voltage from the sample-and-hold circuit and subtracts the DAC voltage from it. This difference voltage is then fed through a gain stage with gain G to produce the residue voltage, the output voltage of this stage. In a typical pipelined ADC implementation, like the one under design, the sample-and-hold circuit and flash DAC are implemented in a single switched-capacitor circuit called a multiplying DAC, or MDAC. The amplification of the residue usually occurs with a closed-loop operational amplifier.

In equation form, the output of each pipeline stage can be described as:

$$V_{res} = G(V_{in} - DV_{res}) \tag{2.3}$$

The residue voltage, V_{RES}, becomes the input voltage to the next stage. The digital decisions versus input voltage and the residues versus input voltage of a typical pipelined ADC are displayed in Fig 2.4.

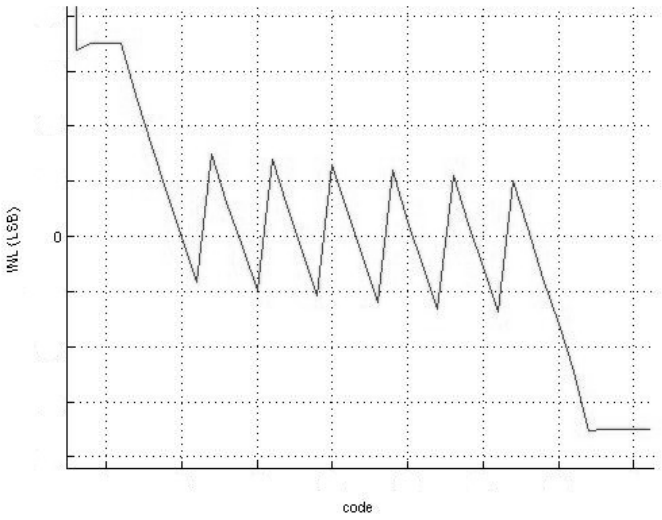

Fig 2.4: *output of a MDAC*

In Fig 2.4 the input voltage is swept through the whole operative range. The residue represents the amplified remainder from the subtraction of the DAC output voltage from the stage input voltage.

The pipelined ADC theory of operation is that each stage is responsible for quantizing a certain set of bits that will eventually become integrated into the final conversion output. For the generalized pipeline ADC described previously, each stage is responsible for quantizing n bits of the input sample. The final ADC output consists of a weighted sum of each stage's digital decision. The weightings are determined by the interstage gains, or the gains of the residue amplifiers within each stage. The final output is weighted according to:

$$x = \frac{D(0)}{G_0} + \frac{D(1)}{G_0 G_1} + \ldots + \frac{D(N_s - 1)}{G_0 G_1 \ldots G_{N_s - 1}} \tag{2.4}$$

where D(i) and G_i represent the digital decision and the residue amplifier gain of each pipeline stage. The above equation suggests that later stages have a smaller weight in the final ADC output. This is indeed the case, as later stages resolve the lower bits of the overall conversion.

In the above example, $D(N_s\text{-}1)$ represents the digital decision made by the final flash ADC, responsible for resolving the least significant bits of the output.

As mentioned before, each stage in a generalized pipelined ADC is responsible for resolving n bits of the ADC output, while the final flash ADC is responsible for

quantizing the m least significant bits of the ADC output. It is evident from the serialized operation of the pipelined ADC that some sort of time-alignment and error-correction circuitry is required for aligning each stage's digital decision to produce the final output.

2.4 Pipelined ADC Performance Characteristics

In general, the pipeline architecture enables the implementation of relatively high-resolution ADCs without sacrificing processing speed or power draw. Additionally, the linear complexity scaling inherent to the pipeline architecture makes the implementation of higher-resolution pipeline ADCs more manageable than with another ADC architecture.

The architecture of the pipelined ADC enables it to have a high throughput rate. This is evident in that pipelined ADCs can have sampling rates of a few MSps up to 200Msps, like the device discussed here. The reasoning for this is that the sample-and-hold circuit can begin processing the next analog input voltage sample as soon as the DAC, summation node, and gain amplifier have finished processing the previous sample. This pipelining action allows a high sustained sampling rate. Additionally, since each stage is only responsible for quantizing a low number of bits relative to the overall resolution of the pipeline ADC, each stage processes each sample relatively quickly.

The architecture of the pipelined ADC also allows it to scale linearly as complexity increases. In the generalized pipeline ADC discussed earlier, each stage has a small flash ADC that performs the quantization of the input sample. These flash ADCs are comprised of many comparators that are responsible for quantizing the sample. For an n-bit flash ADC, 2^n comparators are needed to perform the conversion. In a pipeline ADC, higher overall resolution is obtained effectively by adding additional small flash ADCs in the form of having more stages.

A 14-bit pipeline ADC with 6 stages, 2.8 bits per stage, is implemented using only 42 comparators. This is in stark contrast to a 14-bit pure flash ADC, which would require $2^{14} = 16384$ comparators in order to quantize the sample. The complexity in a pipeline ADC scales linearly and not exponentially, as is the case in a flash ADC. It also follows that fewer required comparators translates to much less power dissipation and power draw, another advantage of the pipeline architecture.

Although the pipeline ADC allows for high speed, lower power dissipation, and low complexity, there are still tradeoffs. For instance, the serialized nature of the conversion process means that there is a significant time delay between the sample that enters the first sample and hold of the first stage and when the digital alignment circuitry produces the correct output code.

Each stage in a pipeline ADC delays the data output by approximately one additional clock cycle. This data latency has to be accounted for when implementing a pipelined ADC.

Even in spite of these tradeoffs, the pipelined ADC architecture enables an ADC to have relatively high resolution, high speed, and low power dissipation, all with very few tradeoffs.

2.5 Double sampling tecnique

The property of the successive ADC stages working in opposite clock phases can be exploited by sharing the operational amplifier, the comparators and the all the logic part between two parallel component ADCs. This approach uses the double-sampling concept of switched capacitors circuits.

By using this technique, the equivalent sampling rate is doubled, but still the power dissipation remains almost the same as for an ADC having traditional single sampled pipeline stages with a half sample rate. The area can be reduced up to 40%. In contrast, the complexity of the pipeline stage is increased and more clock signals with different phases are needed.

Scheme of the double sampling multiplying D/A converter is shown in fig 2.5.

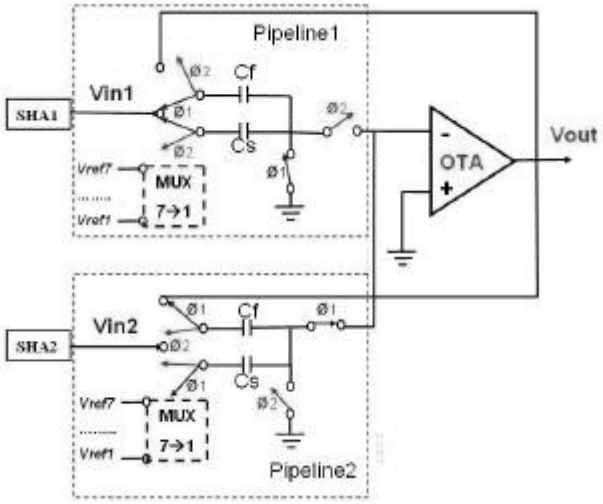

Fig 2.5: *double-sampling MDAC architecture*

The capacitors of two parallel channels working on opposite clock phases share the same amplifier. While the pipeline1 samples the Vin1 signal onto the Cs and Cf capacitors independently of the amplifier, the pipeline2 switches to the amplification phase.

Two important side effects are caused by the amplifier sharing. First, the amplifier load capacitance is increased and affects its bandwidth. Second, the amplifier input offset is never reset; this can be tolerated by an adequate amplifier open loop DC-gain. The second one is not so critical in this design because of the differential architecture used, and thereby a symmetric compensation is possible.

2.6 Derivation of the OperationalAmplifier Parameters

To calculate the DC-gain of the amplifier in a multiplying D/A converter it is necessary to deal with the resolution; instead, for the slew rate and GBW specifications, the sampling speed of the A/D converter is the key parameter.

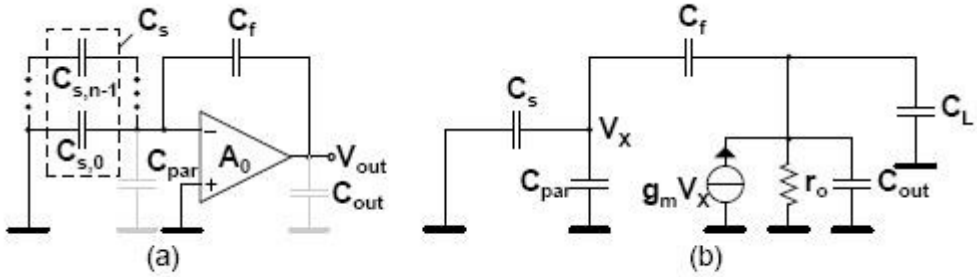

Fig 2.6: *hold mode MDAC*

The topology presented in fig 2.6 is assumed to be in the hold mode as single-ended for simplicity. However, all the calculations are performed for a fully differential topology. In this configuration, the input signal is sampled to the sampling capacitors

$$C_s = \sum_{j=0}^{n-1} C_{s,j} \tag{2.5}$$

and feedback capacitor C_f.

2.6.1 Open loop DC-Gain

The settling error at the output of the operational amplifier in a multiplying D/A converter, resulted from the finite open loop DC gain $A_0 = g_m r_o$ is approximately given by

$$\varepsilon_0 = \frac{1}{A_0 \cdot f} \tag{2.6}$$

where f is the feedback factor

$$f = \frac{C_f}{C_f + \sum_{j=0}^{n-1} C_{s,j} + C_{par}} \tag{2.7}$$

which can be approximated in case of $C_s, C_f \gg C_{in}$ to equal

$$f \approx \frac{1}{2^{B_i}} \tag{2.8}$$

Assuming that the errors $\varepsilon_{A0,i}$, caused by the finite DC-gain in all the $m = k$-1 stages with a resolution of $B_i + r$ bits, are the only error sources, the total input error of a N-bit A/D converter is

$$\varepsilon_{tot} = \sum_{i=1}^{m} \frac{\varepsilon_{A_{0,i}}}{\prod_{l=1}^{i} 2^{B_l}} \tag{2.9}$$

Applying some substitutions, the total error at the input can be rewritten

$$\varepsilon_{tot} = \sum_{i=1}^{m} \frac{2^{B_i}}{A_{0,i} \cdot \prod_{l=1}^{i} 2^{B_l}} \tag{2.10}$$

where $A_{0,i}$ is the open loop DC-gain of the amplifier in the ith stage. At the same time, the total error at the ADC input must be less than LSB/2, which corresponds to $\varepsilon_{tot} < 1/2^N$ for an *N*-bit ADC. The inequality for the dimensioning of the amplifier open loop DC-gains becomes in general case

$$\sum_{i=1}^{m} \frac{2^{B_i}}{A_{0,i} \cdot \prod_{l=1}^{i} 2^{B_l}} < \frac{1}{2^N} \tag{2.11}$$

2.6.2 Gain Bandwidth

The successive pipeline stages operate in opposite clock phases, which gives a settling time of a half of the clock cycle (T/2). The settling time is determined first by the slew rate (SR) and finally by the gain bandwidth of the amplifier, as indicated in Fig. 2.7. Again, the MDAC topology of Fig. 2.6(a) is considered as fully differential.

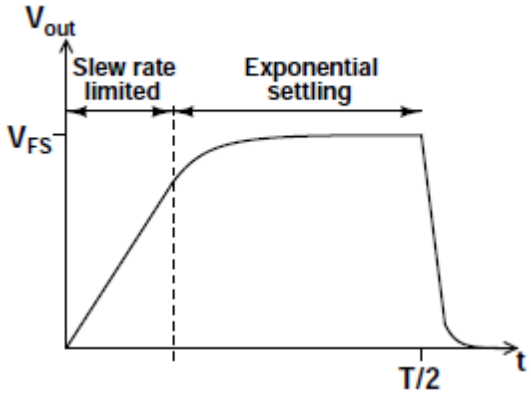

Fig 2.7: *settling of the output*

The most commonly used OTAs can be modeled with a single-pole small-signal model of Fig. 2.6(b). The GBW frequency of an OTA is related to the transconductance g_m by

equation

$$GBW = \frac{g_m}{2\pi C_{L,tot}} \tag{2.12}$$

where the total load capacitance $C_{L;tot} = C_L + C_{out}$ includes the parasitic output capacitance C_{out}. Using the symbols of Fig. 2.6(b), the corner frequency for the settling in the hold mode is

$$\omega_{-3dB} = \frac{gm}{C_{L,H}} \cdot f = \frac{g_m}{\frac{C_{L,tot}}{f} + \sum_{j=0}^{n-1} C_{s,j} + C_{par}} \tag{2.13}$$

It was arbitrarily chosen to reserve one third of the settling time for the SR limited part and two thirds for the GBW limited exponential settling. The error ε_t caused by the incomplete exponential settling during $T/3 = 1/(3f_S)$ is given by

$$\varepsilon_\tau = e^{\frac{-\omega_{-3dB} \cdot 1}{3f_s}} = e^{\frac{g_m}{(\frac{C_{L,tot}}{f} + \sum_{j=0}^{n-1} C_{s,j} + C_{par}) 3f_s}} \tag{2.14}$$

In order to fulfill the resolution requirement, the settling error must be less than LSB/2, this case reduced to the input of the stage i, which results in a condition

$$\varepsilon_{\tau,i} < \frac{1}{2^{N_i}} \tag{2.15}$$

where N_i is the resolution of the remaining back-end pipeline including the ith stage.
By combining Eq 2.8, 2.14, and 2.15, and solving the amplifier transconductance g_m yields

$$g_m > 3\ln 2 \cdot 2^{B_i} \cdot N_i \cdot f_S \cdot k C_{L;tot} \tag{2.16}$$

where the constant k>1 is the ratio between the effective load capacitance in the feedback configuration $C_{L;H}$ and in open loop $C_{L;tot}$, resulting in

$$k = \frac{C_{L,H}}{C_{L,tot}} = 1 + \frac{C_f(\sum_{j=0}^{N-1} C_{s,j} + C_{par})}{(C_L + C_{out}) + (C_f + \sum_{j=0}^{N-1} C_{s,j} + C_{par})} \tag{2.17}$$

On the other hand, the transconductance is related to the width W, length L, and drain current I_D of the transistor by

$$g_m = \sqrt{2\mu C_{ox} \frac{W}{L} I_D} \tag{2.18}$$

where μ is the mobility and C_{ox} the gate oxide capacitance. By substituting Eq. 2.18 into

Eq. 2.16, a condition for the minimum drain current of one transistor of the amplifier input differential pair I_D can be derived, and this can be expressed in terms of the minimal gain bandwidth

$$GBW > \frac{3\ln 2}{2\pi} \cdot N_i f_s \cdot \left(2^{B_i} + \frac{\sum_{j=0}^{n-1} C_{s,j} + C_{par}}{C_{L,tot}} \right) \tag{2.19}$$

An interesting special case occurs when all the pipeline stages are identical, having $B_i=B$ with equal correlated settling errors $\varepsilon_{t;i} = \varepsilon_t$ being the only error sources. The total error reduced to the input of an N bit pipeline ADC is given by

$$\varepsilon_{tot} = \varepsilon_\tau \cdot \frac{1 - \frac{1}{2^{mB}}}{2^B - 1} \tag{2.20}$$

Again, for an N bit ADC it must hold that $\varepsilon_{tot} < 1/2^N$. By combining this to Eqs. 2.8, 2.14 , and 2.20 , for the transconductance g_m holds

$$g_m > 3(N\ln 2 - \ln(2_B - 1)) \cdot 2^B \cdot f_S \cdot kC_{L;tot} \tag{2.21}$$

2.6.3 Slew Rate

The slew rate of a single-stage OTA, like a folded cascode amplifier, is linearly dependent on the maximal current I_{max} charging and discharging the load capacitance.

To assure symmetrical slewing of the output, the currents of the output stages have to be equal to the current of the input stage, which indicates $I_{max} = 2I_D$. In a pipeline stage, the load capacitance during the slewing depends on the capacitor charging in the previous operation phase. In the worst case, the total load capacitance is $C_{L;tot}+C_f$. Using the symbols of Fig. 2.6(b), the slew rate is given in this case by

$$SR = \frac{I_{max}}{C_L + C_{out} + C_f} = \frac{2I_D}{C_{L,tot} + C_f} \tag{2.22}$$

For a worst-case slewing of the differential full-scale voltage $V_{pp;diff}$, the SR limited part being one third of the settling time, holds the condition

$$\frac{T}{6} \cdot SR > V_{FS,diff} \tag{2.23}$$

Substituting Eq. 2.22 into the inequality of Eq. 2.23, the minimum drain current set by the slew rate is given by

$$ID > 3 f_S \cdot V_{FS;diff} \cdot (C_{L;tot} + C_f) \tag{2.24}$$

2.6.4 Noise considerations

The errors in each stage come from different sources:

- the finite gain of the amplifier;
- the incomplete settling;
- the mismatch in the capacitances or in the transistors.

It can be demonstrated that the last source is not critical for a stage number lower than 6; moreover the mismatch in the capacitor can be corrected by calibration. Furthermore, considering an equal contribution from all the sources, the allowed error increase from stage to stage by a factor equal to the interstage gain; this implies that the largest contribution will come from the first stages. So the capacitors can be scaled down from stage to stage with a factor equal to the square of the interstage gain, a relation coming from the area dependency model:

$$\sigma^2\left(\frac{\Delta C}{C}\right)=\frac{A_C^2}{WL}=\frac{A_C^2 C_{ox}}{C} \quad , \tag{2.25}$$

down to a minimum dictated by other constraints.

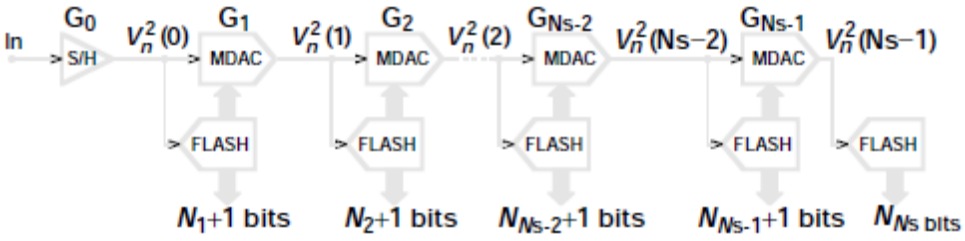

Fig 2.7: *noise through the pipeline*

Concerning the input referred noise

$$V_{n,inref}^2=\frac{V_n^2(0)}{G_0^2}+\frac{V_n^2(1)}{G_0^2 G_1^2}+\ldots+\frac{V_n^2(N_s-1)}{G_0^2 G_1^2 \ldots G_{N_s-1}^2} \tag{2.26}$$

where N_s is the number of stages and G_x is the residual gain of stage x.

The thermal noise comes mainly from the switches and the amplifiers. It can be modelized in a MDAC as following:

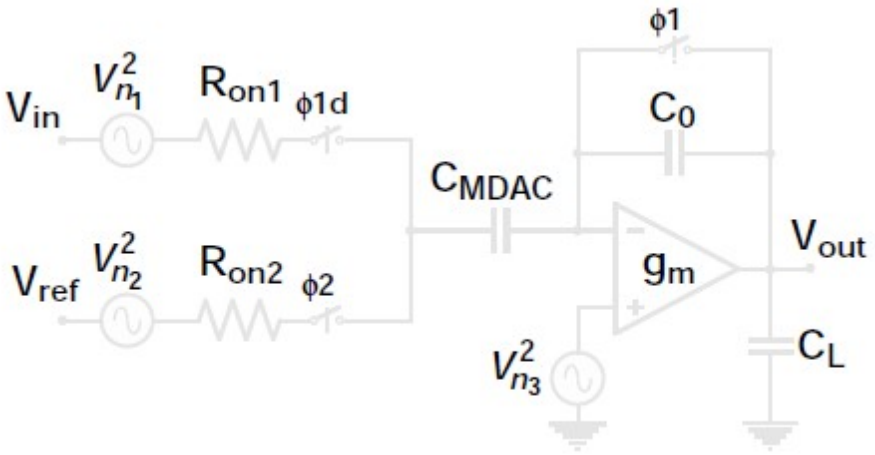

Fig 2.8: *noise source modelization*

Considering:

$$G=\frac{C_{MDAC}}{C_0} \tag{2.27}$$

$$R_{on2}=\frac{1}{\alpha g_m}$$

with α a design variable,

$$R_{eq}\approx\frac{2}{3g_m}$$

$$B_{eq}\approx\frac{\pi}{2}\frac{g_m}{2\pi C_{eq}}$$

$$C_{eq}=C_{MDAC}+C_L+\frac{C_{MDAC}C_L}{C_0}$$

it is possible to express the noise at the output as:

$$V_{n,outref}^2=\frac{kT}{C_{MDAC}}+(4kTR_{on2}+4kTR_{eq})\cdot B_{eq}\cdot G^2 \tag{2.28}$$

where the first addend represents the sampled and held contribution, the first addend inside brackets is the broadband contribution from the switches during the amplification phase, and the second addend inside brackets is the contribution coming from the amplifier.

Chapter 3

Design issues in short channel devices

3.0 Introduction

As the technology scales beyond 100-nm sizes, the traditional design approach needs to be modified to take into account the increased process variation, interconnect processing difficulties, and other physical effects.

It can be experienced a significant increase in gate tunnelling current, due to the thin oxide. Subthreshold leakage and gate tunnelling are no longer second-order effects. If these effects are not taken care of, the result will be a dysfunctional devices, especially for digital circuits, but also the analog environment will be heavily changed.

Typically, processor designers tape out their design when the verification confidence level is high enough. Debug continues on silicon, which is usually several orders of magnitude faster and would result in getting a product to market sooner. Now, due to the increased mask cost and longer fabrication turnaround time, the trade-off to arrive at the most cost-effective product and shortest time to market will certainly be different [28].

The transistor figure of merit is now deviating from the reciprocal of the gate length. Furthermore, global wiring is not scaling, whereas wire resistance below 0.1 μm is increasing exponentially. This is due primarily to surface scattering and grain-size

limitations in a narrow trench, resulting in carrier scattering and mobility degradation. The gate dielectric thickness is approaching atomic dimensions and at 1.2 nm in the 90nm node is about five atomic layers of oxide. Source–drain extension resistance is getting to be a larger proportion of the transistor "on" resistance. Source–drain extension doping has been increased significantly, and the ability to reduce this resistance has to be traded off with other short-channel effects, such as hot-carrier injections and leakage current due to band-to-band tunnelling. Source–drain diffusions are getting so thin that implants are at the saturation level and resistance can no longer be reduced unless additional dopants can be activated.

The main effects related to the reduced dimensions of the devices are the following:

-the current losses

-the mobility degradation

-the threshold shift

-the gate capacitance shift

-the g_{ds} degradation

3.1 Current losses

The device under design has no consumption particularly stringent specification; of course power must be minimized, but, since it is not the main goal, the leakage in devices like the amplifier can be neglected (it is not the case instead for the switches, the bootstrap or other other circuital elements, but, since it is not object of this dissertation, the problem will be ignored).

The mail losses are related on:

-gate tunnelling through the oxide

-junction losses

-hot carrier current

-gate induced drain leakage

3.1.1. Tunnelling

The electron wave function, for oxides below 8-10nm, can spread up to the anode; in particular, depending on the electric field, several cases are possible:

a) $E_{ox} > \varphi_B / t_{ox}$: the voltage drop is higher than the barrier. In this case, since the tunnelling is an energy conservative process, the electron sees a triangular potential barrier, having an effective depth of $t_{eff} = \varphi_B / E_{ox}$, as depicted in fig 3.1.

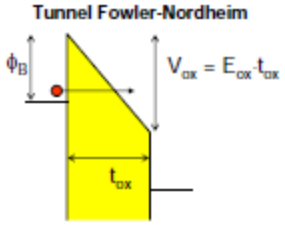

Fig 3.1: *Fowler-Nordheim tunnelling*

This process, named Fowler-Nordheim Tunnel, implies that the electron reaches the conduction band of the oxide and crosses it up to the anode, following the relation:

$$J_{FN} = A \cdot E_{OX}^2 \cdot \exp\left(\frac{-B}{E_{OX}}\right) \tag{3.1}$$

where B a constant related on the barrier height and on the type of oxide.

b) $E_{ox} < \varphi_B / t_{ox}$: the voltage drop is lower than the barrier height and the electron sees a trapezoidal barrier; it jumps directly from the conduction band of the cathode to the one of the anode, so this process is named Direct Tunnel.

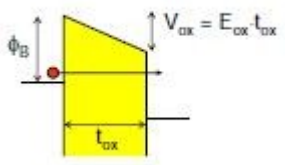

Fig 3.2: *Fowler-Nordheim tunnelling*

The expression is similar to the Fowler-Nordheim case:

$$J_D \simeq A \cdot E_{OX}^2 \cdot \exp\left(\frac{-B}{E_{OX}} \cdot \left(1 - \left(1 - \frac{q \cdot E_{OX} \cdot t_{OX}}{\varphi_B}\right)^{3/2}\right)\right) \tag{3.2}$$

c) the Hole Valence Band tunnelling, where the process is similar to the previous cases

but involving holes instead of electrons; although it is less probable, since the barrier is intrinsically higher (5eV).

d) the Electron Valence Band tunnelling, where the hypothesis is that the electron has an energy at the same level of free energy level in the anode. The gate voltage required is around 1.5V, but, when an electron crosses the oxide, it sees a barrier of about 4.2eV, so this happens only for an oxide thickness lower than 2-3nm.

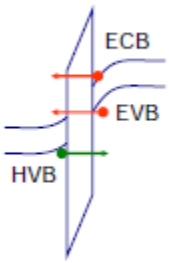

Fig 3.3: *different tunnelling*

e) the Band-to-band tunnelling; when the doping is high, the charge space region drops: if also the potential between source and drain is high, then the band bending is strong. This means that the electrons from the cathode see a triangular voltage barrier Eg high; so the electron can jump in the conduction band by tunnelling.
The expression is similar to the one for the Fowler-Nordheim; the difference is in the carrier concentration and in the state function density:

$$J_{b-b} = \frac{A \cdot E \cdot V_d}{\sqrt{E_g}} \cdot \exp\left(\frac{-\beta \cdot E_g^{3/2}}{E}\right) \qquad (3.3)$$

3.1.2. GIDL

It happens typically when the gate is grounded and the drain is at Vdd. The MOS is off, there is no channel and the substrate is in accumulation. All the silicon surface is in accumulation, so it behaves like a high doped p semiconductor, where the Fermi level is close to the valence band. Then, at the Si/SiO_2 interface, it forms a p+/n/n+ junction which can create a band-to-band tunnel leakage.

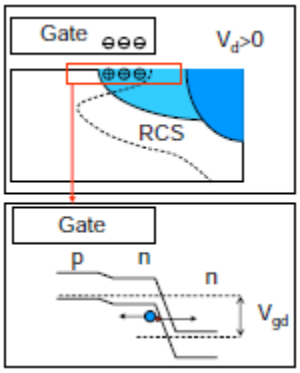

Fig 3.4: *gate induced drain lowering*

3.1.3 DIBL

The current flow in the channel depends on creating and sustaining an inversion layer on the surface. If the gate bias voltage is not sufficient to invert the surface ($V_{GS}<V_{T0}$), the carriers (electrons) in the channel face a potential barrier that blocks the flow.

Increasing the gate voltage reduces this potential barrier and, eventually, allows the flow of carriers under the influence of the channel electric field. In small-geometry MOSFETs, the potential barrier is controlled by both the gate-to-source voltage V_{GS} and the drain-to-source voltage V_{DS}. If the drain voltage is increased, the potential barrier in the channel decreases, leading to drain-induced barrier lowering (DIBL). The reduction of the potential barrier eventually allows electron flow between the source and the drain, even if the gate-to-source voltage is lower than the threshold voltage.

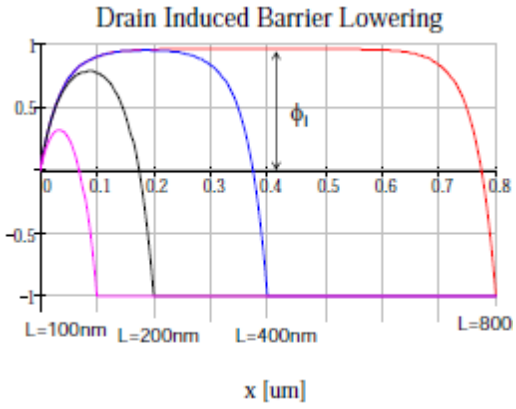

Fig 3.5: *Drain-induced barrier lowering*

3.2 Punchtrough

The expressions for the drain and source junction widths are:

$$x_{dD}=\sqrt{(\frac{2\varepsilon_{Si}}{qN_a})(V_{DS}+\varphi_{Si}+V_{SB})} \tag{3.4}$$

and

$$x_{dS}=\sqrt{(\frac{2\varepsilon_{Si}}{qN_a})(\varphi_{Si}+V_{DB})} \tag{3.5}$$

where V_{SB} and V_{DB} are source-to-body and drain-to-body voltages.

When the depletion regions surrounding the drain extends to the source, so that the two depletion layers merge (i.e., when $x_{dS} + x_{dD} = L$), punchtrough occurs. Punchthrough can be minimized with thinner oxides, larger substrate doping, shallower junctions, and obviously with longer channels.

It can be a destructive effect, so it must be strictly avoided.

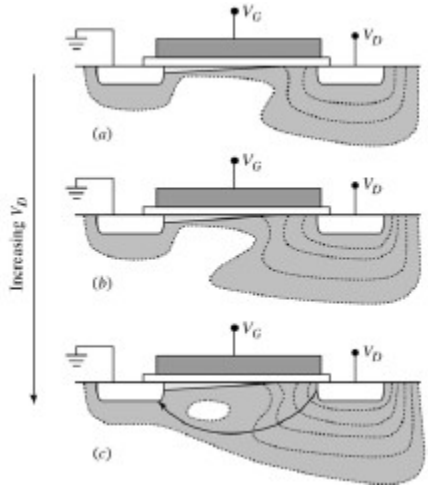

Fig 3.6: *punchtrough*

3.3 Surface scattering

When the channel length becomes smaller due to the lateral extension of the depletion layer into the channel region, the longitudinal electric field component e_y increases, and the surface mobility becomes field-dependent. Since the carrier transport in a MOSFET is confined within the narrow inversion layer, and the surface scattering (that is the collisions suffered by the electrons that are accelerated toward the interface by e_x) causes reduction of the mobility, the electrons move with great difficulty parallel to the interface, so that the average surface mobility, even for small values of e_y, is about half as much as that of the bulk mobility.

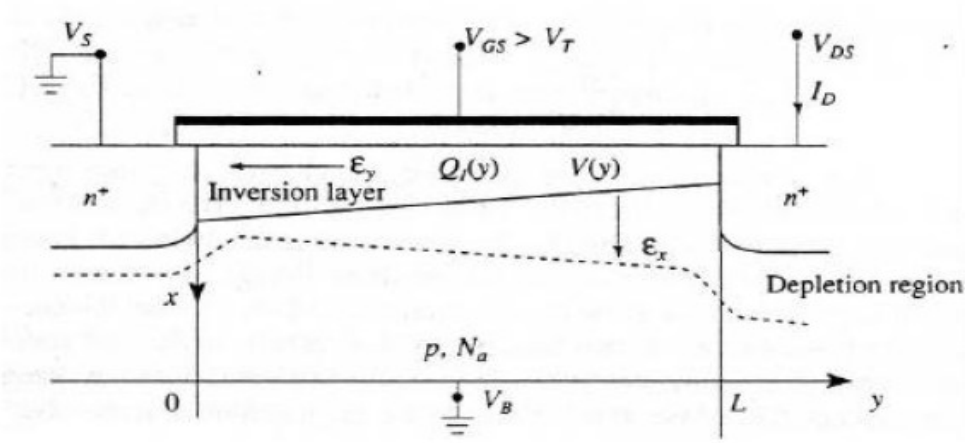

Fig 3.7: *MOS cross section showing the field contributions*

3.4 Velocity saturation

The performance short-channel devices is also affected by velocity saturation, which reduces the transconductance in the saturation mode. At low e_y, the electron drift velocity v_{de} in the channel varies linearly with the electric field intensity.

It can be noted that the drain current is limited by this effect instead of pinchoff. This occurs in shortchannel devices when the dimensions are scaled without lowering the bias voltages.

Using $v_{de(sat)}$, the maximum gain possible for a MOSFET can be defined as

$$g_m = Wc_{ox}V_{de(sat)} \tag{3.6}$$

3.5 Impact ionization

Another undesirable short-channel effect, especially in NMOS, occurs due to the high velocity of electrons in presence of high longitudinal fields that can generate electron-hole pairs by impact ionization, that is, by impacting on silicon atoms and ionizing them.

It happens as follow: normally, most of the electrons are attracted by the drain, while the holes enter the substrate to form part of the parasitic substrate current. Moreover, the region between the source and the drain can act like the base of an npn transistor, with the source playing the role of the emitter and the drain that of the

collector. If the holes are collected by the source, and the corresponding hole current creates a voltage drop in the substrate material of the order of 0.6V, the normally reversed-biased substrate-source pn junction will conduct appreciably. Then electrons can be injected from the source to the substrate, similar to the injection of electrons from the emitter to the base. They can gain enough energy as they travel toward the drain to create new e-h pairs. The situation can worsen if some electrons generated due to high fields escape the drain field to travel into the substrate, thereby affecting other devices on a chip.

3.6 Hot electrons

Another problem, related to high electric fields, is caused by so-called hot electrons. This high energy electrons can enter the oxide, where they can be trapped, giving rise to oxide charging that can accumulate with time and degrade the device performance by increasing V_T and affect adversely the gate's control on the drain current.

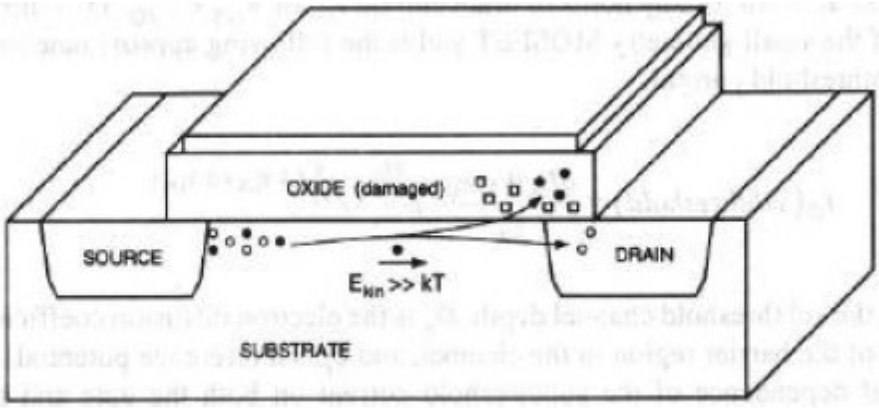

Fig 3.8: *hot-electron damages*

3.7 The modification of the threshold voltage due to Short-Channel Effects (SCE)

The equation giving the threshold voltage at zero-bias

$$V_{T0} = V_{FB} + 2\varphi_F + \left(\frac{1}{C_{ox}}\right)\sqrt{2q \cdot \varepsilon_{Si} \cdot N_A (2\varphi_F)} + \frac{qD_I}{C_{ox}} \qquad (3.7)$$

is accurate in describing large MOS transistors, but it collapses when applied to small-geometry MOSFET. In fact that equation assumes that the bulk depletion charge is only due to the electric field created by the gate voltage, while the depletion charge near n+

source and drain region is actually induced by pn junction band bending. Therefore, the amount of bulk charge the gate voltage supports is overestimated, leading to a larger V_T than the actual value.

The electric flux lines generated by the charge on the MOS capacitor gate electrode terminate on the induced mobile carriers in the depletion region just under the gate. For short-channel MOSFET, on the other hand, some of the field lines originating from the source and the drain electrodes terminate on charges in the channel region. Thus, less gate voltage is required to cause inversion. This implies that the fraction of the bulk depletion charge originating from the pn junction depletion and hence requiring no gate voltage, must be subtracted from the V_T expression.

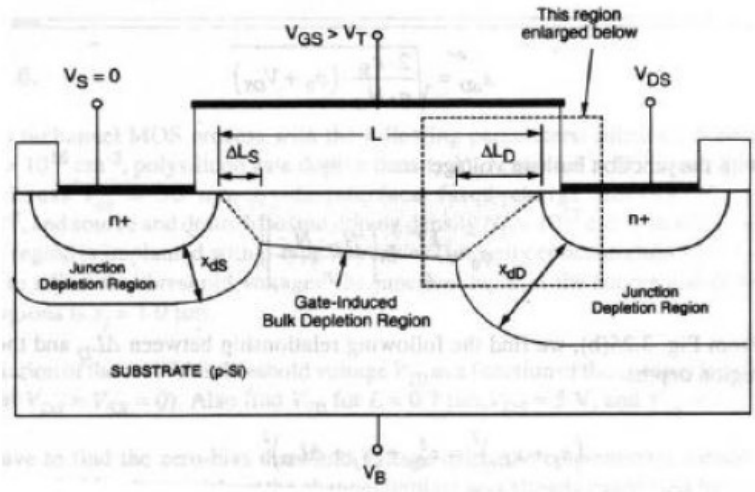

Fig 3.9: *gate-induced bulk depletion region*

The figure shows the simplified geometry of the gate-induced bulk depletion region and the p-n junction depletion regions in a short channel MOS transistor. It can be noted that the bulk depletion region is assumed to have and asymmetric trapezoidal shape, instead of a rectangular shape, to represent accurately the gate-induced charge. The drain depletion region is expected to be larger than the source depletion region because the positive drain-to-source voltage reversed-biases the drain-substrate junction. We recognize that a significant portion of the total depletion region charge under the gate is actually due to the source and drain junction depletion, rather than the bulk depletion induced by the gate voltage. Since the bulk depletion charge in the short channel device is smaller than expected, the threshold voltage expression must be modified to account

for this reduction:

$$V_{T0short\ ch} = V_{t0} - \Delta V_{t0}$$

where V_{T0} is the zero-bias threshold voltage calculated using the conventional long-channel formula and ΔV_{T0} is the threshold voltage shift (reduction) due to the short-channel effect. The reduction term actually represents the amount of charge differential between a rectangular depletion region and a trapezoidal depletion region.

Let ΔL_S and ΔL_D represent the lateral extent of the depletion regions associated with the source junction and the drain junction, respectively. Then, the bulk depletion region charge contained within the trapezoidal region is:

$$Q_{B0} = -\left(1 - \frac{\Delta L_S + \Delta L_D}{2L}\right)\sqrt{4q\epsilon_{Si} N_A \varphi_f} \tag{3.8}$$

To calculate ΔL_S and ΔL_D, the simplified geometry shown in the figure can be useful.

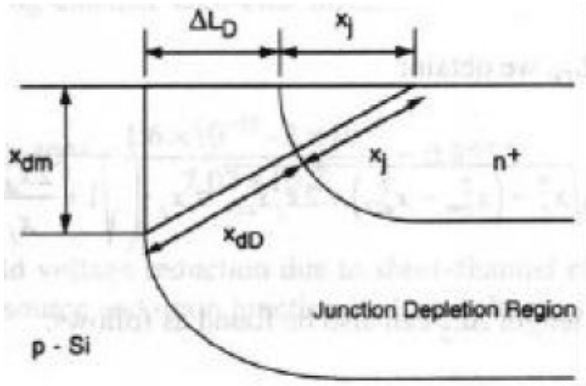

Fig 3.10: *geometry of the depletion region*

Here, x_{dS} and x_{dD} represent the depth of the pn-junction depletion regions associated with the source and the drain, respectively. The edges of the source and drain diffusion regions are represented by quarter-circular arcs, each with a radius equal to the junction depth, x_j. The vertical extent of the bulk depletion region into the substrate is represented by x_{dm}. The junction depletion region depths can be approximated by

$$x_{dD} = \sqrt{\left(\frac{2\epsilon_{Si}}{qN_A}\right)(V_{DS} + \varphi_0)} \tag{3.9}$$

and

$$x_{dS}=\sqrt{\left(\frac{2\epsilon_{Si}}{qN_A}\right)(\varphi_0)} \tag{3.10}$$

with the junction built-in voltage

$$\varphi_0=\frac{kT}{q}\ln\left(\frac{N_D N_A}{n_i^2}\right) \tag{3.11}$$

From figure,

$$\Delta L_D \approx x_j\left(\sqrt{1+\frac{2x_{dD}}{x_j}}-1\right) \tag{3.12}$$

Similarly, the length ΔL_S can also be found as follows:

$$\Delta L_S \approx x_j\left(\sqrt{1+\frac{2x_{dS}}{x_j}}-1\right) \tag{3.13}$$

Now, the amount of the threshold voltage reduction ΔV_{T0} due to short-channel effects can be found as:

$$\Delta V_{T0}=\frac{1}{C_{ox}}\cdot\sqrt{4q\epsilon_{Si}N_A\varphi_F}\cdot\frac{x_j}{2L}\cdot\left[\left(\sqrt{1+\frac{2x_{dD}}{x_j}}-1\right)+\left(\sqrt{1+\frac{2x_{dS}}{x_j}}-1\right)\right] \tag{3.14}$$

The threshold voltage shift term is proportional to x_j/L. As a result, this term becomes more prominent for MOS transistors with shorter channel lengths, and it approaches zero for long channel MOSFET where $L >> x_j$.

3.8 Output conductance reduction

High performance logic devices are optimized for good drive current, low leakage and SCE control which incorporate super halo and double-pocket structures. These structures however, often result in low output resistance, device gain, transconductance-to-drive current ratio and matching properties [1]. Low output resistance is the result of increase ID with VD in saturation regime. Three components are associated with this increase, namely channel length modulation (CLM), drain-induced-barrier-lowering (DIBL) and substrate current body effect (SCBE).

It can be found that gds is most sensitive to LDD dose, halo dose, halo tilt and APT dose and energy.

3.8.1 Lightly doped drain implant

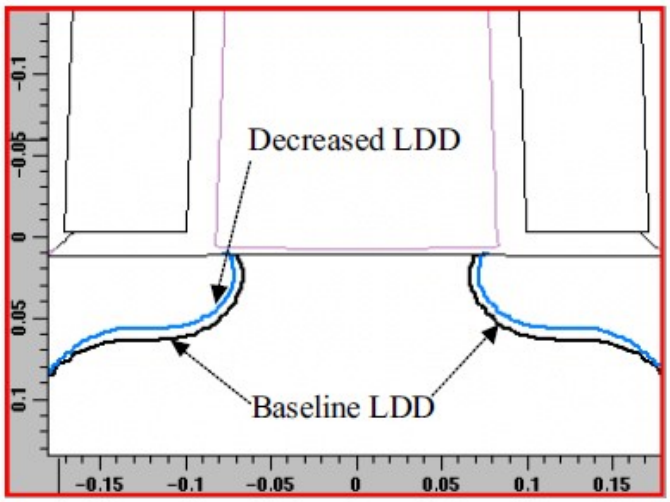

Fig 3.11: *Junction profile for different LDD dose*

Fig 3.12 shows the sensitivity of V_A to LDD dose for a PMOS transistor; V_A is defined as

$I_D/g_{ds} - V_D$, the Early voltage.

If the LDD dose is increased, V_A decreases because the effective channel becomes shorter, as shown in Figure 3.11. Shorter channel length results in larger residual DIBL thus causing output resistance to decrease.

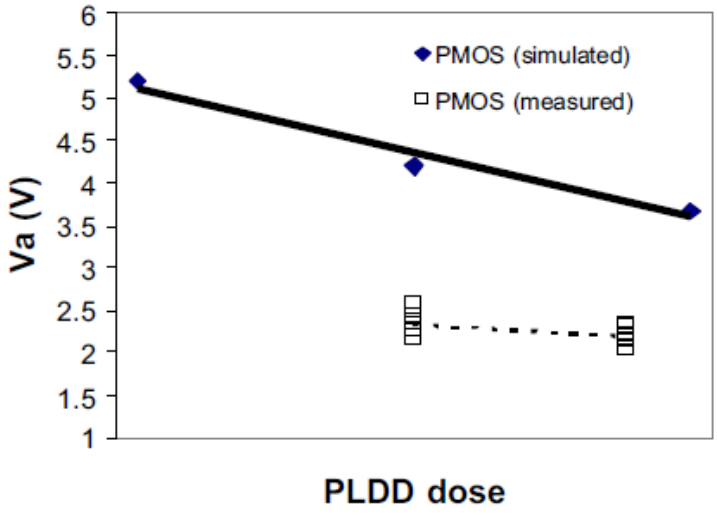

Fig 3.12: *V_A dependence on PLDD dose*

3.8.2 Halo Implant

The halo is a p+ implant applied in the proximity of the source and drain junctions; its purpose is to reduce the effects of charge sharing, DIBL and punchthrough.

Figure 3.13 shows the drawback of this technique: increasing the halo dose increases V_A.

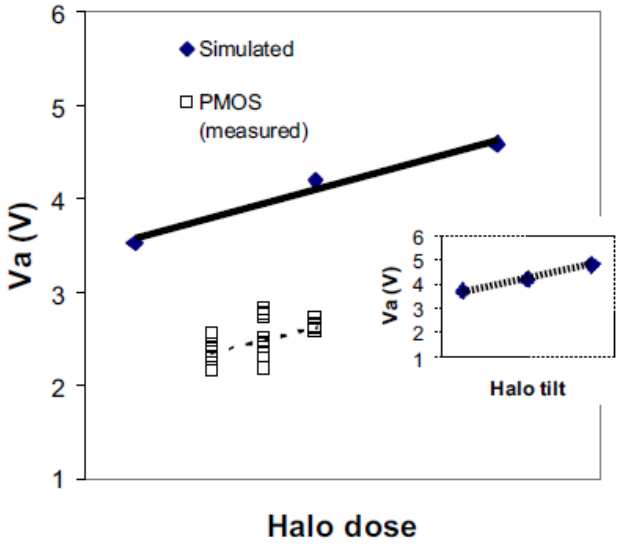

Fig 3.13: *V_A dependence on halo dose*

Increasing the halo tilt angle also increases V_A as shown in the inset: a larger tilt angle places the halo implant almost at the centre of the channel. The additional arsenic in the channel lessens the effect of DIBL as shown in Figure 3.14. However, increasing halo dose also causes I_{dsat} to decrease. Experimental studies have also shown that pocket implant has tradeoff effects on V_A and I_{dsat}.

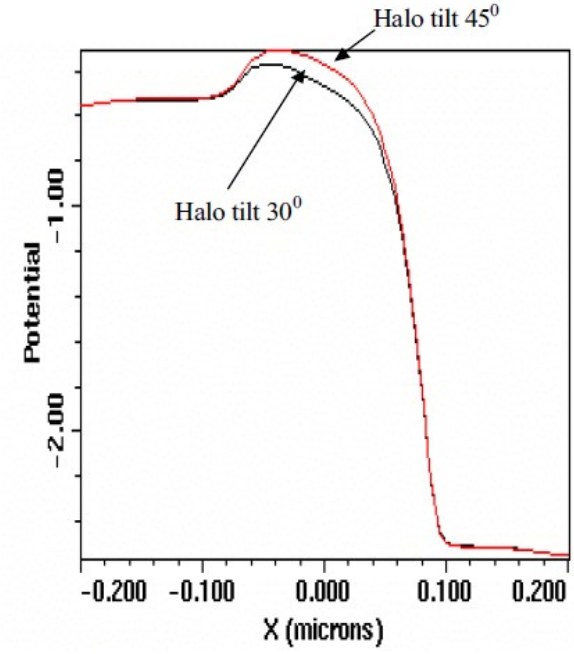

Fig 3.14: *Potential barrier shifts at different halo tilt angle*

3.8.3 Anti Punchthrough Implant

The natural thresholds of the NMOS is about 0V and of the PMOS is about 1.2V. An p implant is used to make the NMOS harder to invert and the PMOS easier resulting in threshold voltages balanced around zero volts.

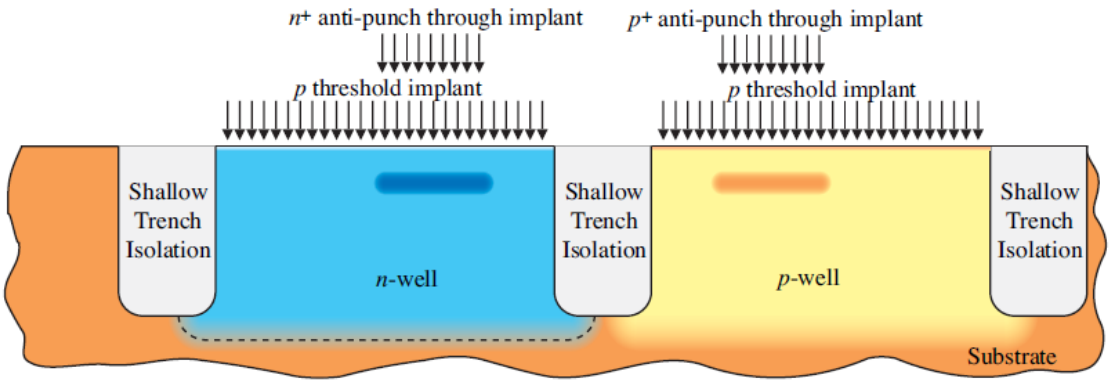

Fig 3.15: *Anti punchthrough implants*

Also an implant can be applied to create a higher-doped region beneath the channels to prevent punch-through from the drain depletion region extending to source depletion region. This technique is typically named anti punchthrough (APT) implant.

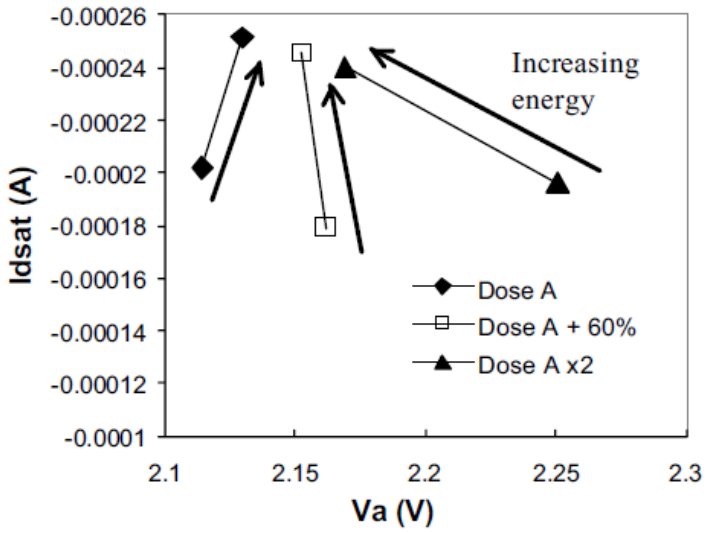

Fig 3.16: *Effects of APT implant on V_A and I_{dsat}*

Fig 3.16 shows the plot of I_{dsat} versus V_A at different APT implant conditions. As APT energy implant is increased, V_A shows a contradict trend depending on the implant dose used. At high dose and high energy, the plot is shifted to the up-left when energy is

increased. However, the plot shifted to the up-right as energy is increased at low APT implant dose and energy. At low dose and energy, increasing the APT implant energy forms super steep retrograde channel, which has positive effects for analog applications.

Chapter 4

Differential Amplifier design

4.0 Introduction

In general, operational amplifiers are amplifiers with an open loop gain high enough to ensure that the closed loop transfer characteristic with negative feedback is approximately independent of the op amp gain. To ensure wide swing and noise immunity, a fully differential architecture is used; so a differential and a common mode behaviour will be investigated

4.1.1 Fifth stage amplifier: differential

The design effort is directed towards the power consumption minimization. Different circuital topologies have been taken in exam: the goal is to reach the following specifications:

Gain	48,29dB
GBW	1177MHz
Dynamic Range	1.1Vpp
Noise (inp_referred)	41,19nV/ $\sqrt{Hz}$
Current Capability	0.28mA

Table 4.1: *OpAmp stage 5 specifications*

The dynamic range specification implies that at least 0.55 over the 1.2V available must be dedicated to the output swing; this limits the number of devices than can be stacked in the output branches and, consequently, the resistance. At the same time, previous considerations demonstrate that the $g_m r_0$ in this technology is low, so it seems to be mandatory to cascode the output.

At the state of the art, three different topologies seem to meet the specifications required:

-the folded-cascode;

-the active-cascode;

-the two-stage architecture.

Working in an purely intuitive way, if the first met the specifications, it would automatically be better than the second, because the power consumption is mainly related on bandwidth, so the boosters would only increase the current consumption.

Finally, assuming that the folded-cascode topology has enough gain, what to do is to choose between a single stage or a two stages architecture.

In an industrial environment, where the human effort is a parameter to take into account in a design, the more efficient methodology is probably to design the single stage device, to study the technology limits, then to design a two stage amplifier for a more sophisticated MDAC (for instance the stage 3 MDAC) and finally to choose which one is more indicated for the project purpose.

4.1.2 Folded cascode architecture

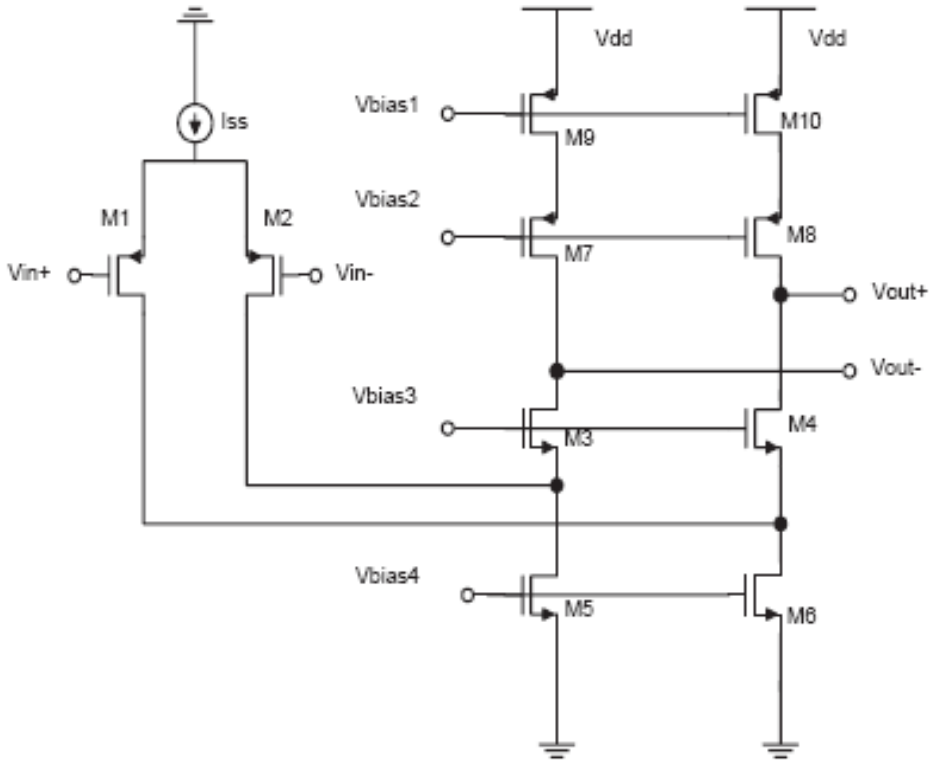

Fig 4.1: *folded cascode architecture*

The most important advantage of the folded structure lies in the voltage output swing because it does not "stack" the cascode transistor on top of the input device. The lower swing of the output is given by

$$V_{min} = V_{ds,sat3}+V_{ds,sat5}, \tag{4.1}$$

and the upper end by

$$V_{max} = V_{dd}-(V_{ds,sat7}+V_{ds,sat9}). \tag{4.2}$$

Thus the peak to peak swing on each side is therefore:

$$V_{swing} = V_{dd}-4*V_{ds,sat}. \tag{4.3}$$

Using the half circuit depicted in fig. 4.2(a), and writing that

$$|A_v|=G_mR_{out}, \tag{4.4}$$

it is possible to calculate the equivalent G_m and R_{out}. As shown in Fig. 4.2(b), the output of the circuit current is approximatively equal to the drain current of M1, as the

impedance seen looking into the source of M3 is much lower than $R_{on1}||R_{on5}$.

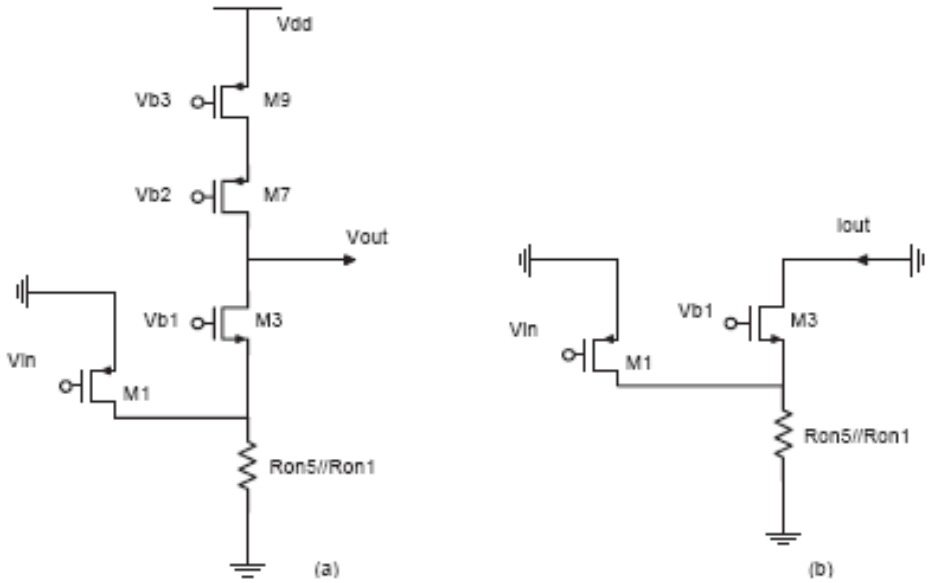

Fig 4.2: *half circuit representation*

The use of a cascode structure allows to reach a very high impedance seen from the output node. In fact, as previously discussed, in 90nm technology the g_{ds} is relatively high, due to physical (the channel length) and technological (the halo structure) aspects.

4.1.3 Cascode structure

To increase the gain of the CMOS stage, the transconductance of the stage can be improved or the output resistance can be enhanced. The output resistance increases in proportion to a decrease in bias current as shown in Eq. 4.5

$$r_{ds} = \frac{1}{\lambda I_{DP}} \tag{4.5}$$

where I_{DP} is the drain pinchoff current and λ is the channel length modulation factor.

Instead the transconductance increases as the square root of the increase in bias current in a relation that can be simplified by the following:

$$g_m = \frac{\partial i_D}{\partial v_{gs}} = \sqrt{2\mu C_{ox}(W/L) I_D} \tag{4.6}$$

It is power efficient to increase the output resistance by lowering the bias current.

Fig. 4.3 shows a single stage amplifier using a conventional cascode connection where the common-gate stage device M2, biased by a voltage supply V_{G2}, is added to the input common-source stage M1. V_{G1}, V_{G2}, and I_{bias} are chosen to make M1 and M2 to operate in their active regions.

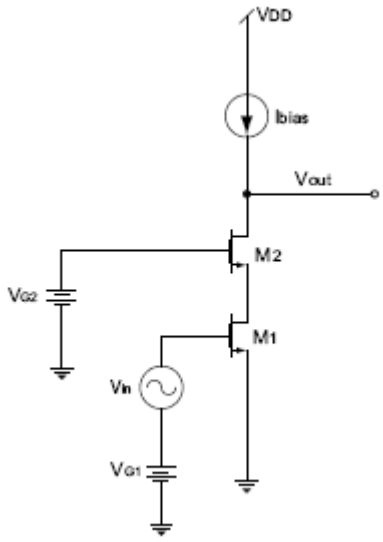

Fig. 4.3: *cascode structure*

Assuming the current source I_{bias} is ideal, the output resistance is

$$r_{out} = r_{ds1} + r_{ds2} + (g_{m2} + g_{mb2})r_{ds1}r_{ds2}. \quad (4.7)$$

The midband voltage gain for the circuit of Fig. 4.3 is

$$A_0 = -[g_{m1}r_{ds1} + g_{m1}r_{ds1}(g_{m2} + g_{mb2})r_{ds2}], \quad (4.8)$$

where g_{m1} and g_{m2} are the transconductance of M1 and M2 individually, r_{ds1} and r_{ds2} denote the drain to source resistance of M1 and M2 at the bias point used, and g_{mb2} represents the transconductance that models the body effect of M2. As indicated by Eq. 4.8, it is clear that the cascode structure can achieve significantly higher voltage gain than a simple MOS stage by providing a higher output resistance. However, this configuration requires that the bias voltage V_{G2} for M2 be $V_T + 2V_{eff}$. The drain of M2 is set higher than V_{G2} in order to allow for the voltage swing. The operation of this cascode connection has limitations for low voltage, low power applications due to the bias voltage requirement and limited output swing. To achieve an even higher gain, more cascode devices can be added in the cascode stack connection to form a "triple cascode". But this further reduces the output swing, so in 1.2V Vdd technology cannot be implemented.

Another aspect to take into account is the rds variation related to the V_{ds} reduction: in fact, as depicted in fig4.4, the gds can drop when cascode V_{ds} comes close to the overdrive value. As previously described, this is a consequence of the Halo implant that reduces the effective channel length.

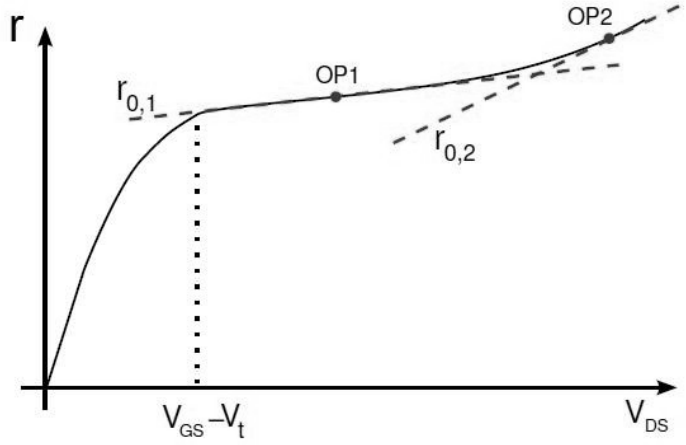

Fig 4.4: *effect of the drain-source voltage on the output resistance*

So the gain of the whole device will be G_mR_{out}, and the bandwidth, or better the dominant pole location will be:

$$\omega_{-3dB} = 1/ R_{out}C_{out} \tag{4.9}$$

which allows to calculate the gain bandwidth product:

$$GBW = \frac{g_{m1}}{2\pi \cdot C_{out}} \tag{4.10}$$

4.1.4 Double input pair

Before deciding which type of transistors to use as input-pair, several aspects must be taken into account:

1) The electrons mobility is considerably higher than the mobility of the holes. g_m, and thereby the gain and unity gain frequency will be higher when using NMOS instead of PMOS-transistors, for the same input capacitance (that is, the W/L ratio).

2) With an NMOS input pair the impedance at the folding points will be lower, due to the intrinsic lower impedance of the PMOS transistors. Both the gain as well as the

phase margin will be lower than if using PMOS at the input, for the same sizes of the transistors.

3) NMOS transistors have lower thermal noise (a parameter related to the g_m).

4) PMOS transistors have lower 1/f noise.

Since high gain is needed, NMOS transistors at the input has been preferred over PMOS. Also, since correlated double sampling will be used, the 1/f noise is reduced, which also make NMOS a better choice from the point of view of the noise.

Another degree of freedom in this circuital topology is the amount of current flowing into the input and the output branches. Typically, once the current capability specification is known, this value is applied at the output, and at the same time at the input to reach the same slewing behaviour when the current is flowing in the two directions relatively the output node.

Intuitively, a small amount of current at the output branch will implies good gain and lower bandwidth, since the MOS effective resistance is inversely proportional to this parameter; on the other side, more current in the input pair will consists in increased bandwidth and gain, since g_m is proportional to the square root of the current:

$$g_m = \sqrt{2\mathrm{k}\frac{W}{L} I_D} \tag{4.11}$$

Here the bottleneck is the gain, so the choice has been to use the minimum amount of current at the output meeting the current capability specifications and increase the one at the input until the gain and bandwidth requirements were met.

This will clearly implies a different behaviour at the output when the signal rises and drops, due to the different current capability available. Simulations demonstrate that this can be not a problem in switching capacitor circuit until the minimum current available is enough to met the slewing specifications.

A drawback can be found in the g_{ds} of the transistors where both the current for the output and the one for the input pair flow. In fact huge transistors will be necessary to have enough V_{ds}, and since the current is high, also the metal width of the connections will be increased. This implies a high parasitic capacitance which, in

parallel with the input pair cascode and the output cascode, determines the secondary pole of the amplifier. This imposes the limit to the maximum over-bias for the input pair.

Another aspect to take into account is that, in the nodes of fig 4.5, besides the sum of

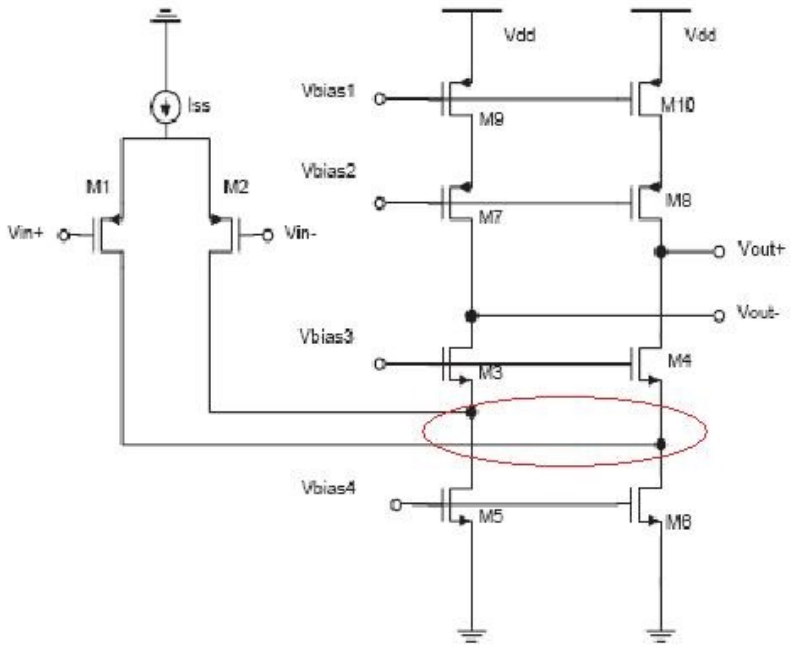

Fig 4.5: *the bottleneck*

the parasitic capacitance, there is also the parallel of the resistances of the mirror transistors with the input pair transistors. Since the node between the input pair and the tail current generator is a virtual ground, and since the input pair devices have a very small channel length to minimize the input capacitance maximizing the g_m, the relative $g_{ds,input\_pair}$ will be also high. So several dB of gain can be lost; the solution chosen has been to stack a cascode transistors between the input pair and the node highlighted in fig 4.5. These devices were chosen identical to the input pair devices, mainly to obtain an easier symmetry when layouting, and biased at V_{dd} so no additional biasing circuitry was needed.

An other peculiarity of the topology lies in the presence of two input pairs. Each of them is driven by a switch connected to a clock signal so that when a pair is turned

on, the other is off.

In fact verifications by simulations show the presence of an unexpected ΔQ at the input of the amplifier, and further investigations demonstrate that it was due to the charges pumped by the bootstrap circuit and by Saap, Sabp, Saan and Sabn switches.

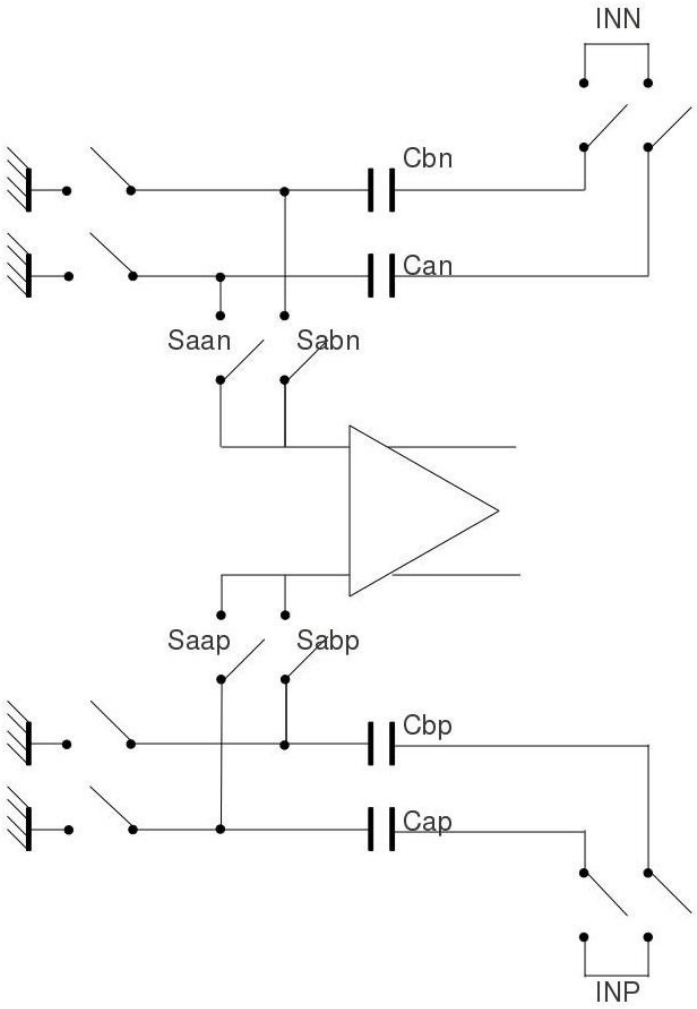

Fig 4.6: *MDAC in double sampling configuration*

In particular, at the moment these switches close, there is a ΔQ pumped from the ground through the parasitic capacitor, and this, of course, generates an error on the value sampled on the capacitor, and an error on the bottom plate voltage of the capacitors. This will imply an extra error later at the end of the amplification phase.

In a single sampling configuration this problem does not exist, because there is no switch in the feedback path of the op-amp.

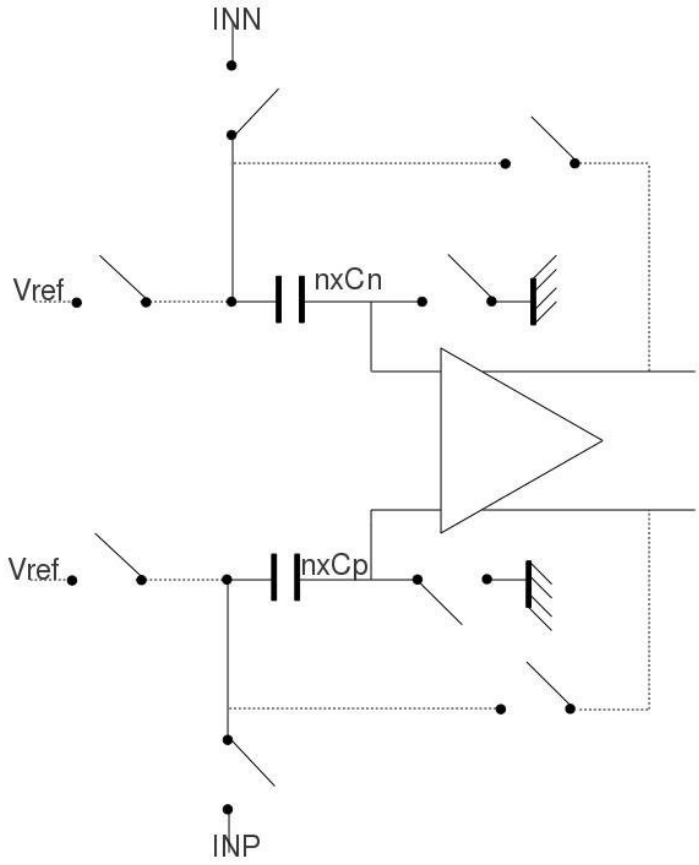

Fig 4.7: *MDAC in single sampling configuration*

From this consideration comes the solution chosen, which is to remove the critical switches by creating a multiple path. This means that the amplifier will have two input pairs, but the extra stored charge is avoided.

This clearly implies the need of extra area and a clock signal, but the power consumption remains nearly unaltered.

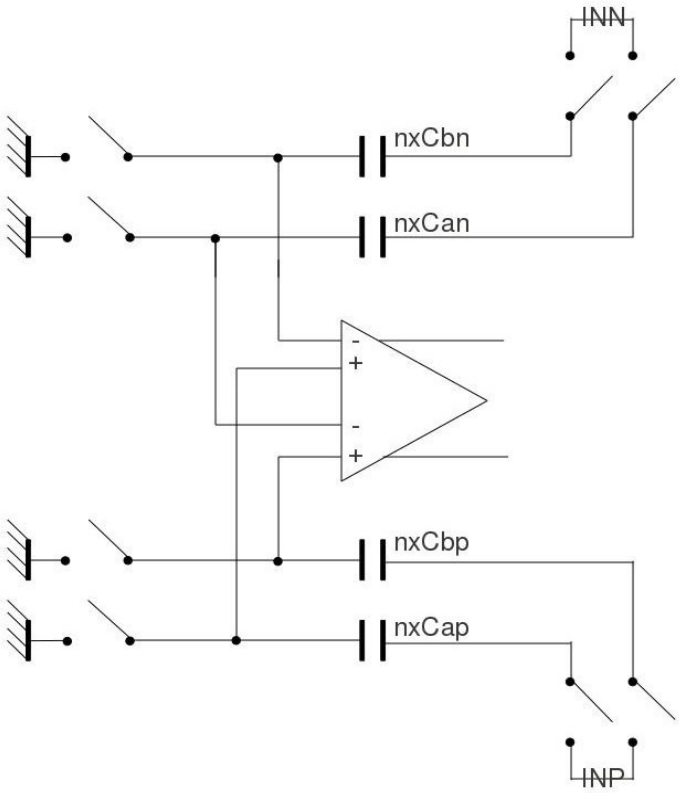

Fig 4.8: *MDAC in double sampling configuration, proposed topology*

4.1.5 Biasing strategy

In the project it is assumed to have a 50uA current source, generated by a bandgap circuit. This means that this current will be proportional to the supply voltage.
What is needed is to bias:

- the transistors used as current source;
- the cascode transistors.

For the first case the 50uA can be simply fed into a device in diode configuration; the voltage at the gate can then be used to bias a scaled version of the same device to reach the current desired.

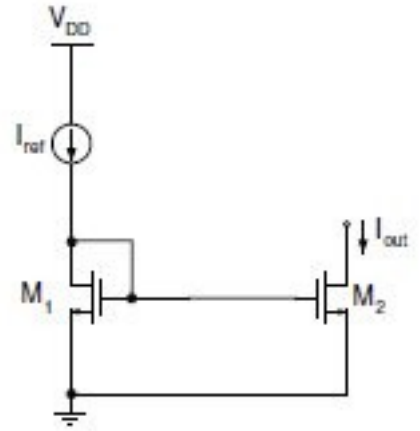

Fig 4.9: *basic current mirror*

The two devices have the same gate voltage, but the V_{ds} can be different. This implies a mismatch in the current (assuming that the gates have the same dimensions), which can be represented:

$$\Delta I = I_2 - I_1 \approx \frac{V_2 - V_1}{r_0} = \frac{\Delta V_{DS,2}}{r_0} \tag{4.12}$$

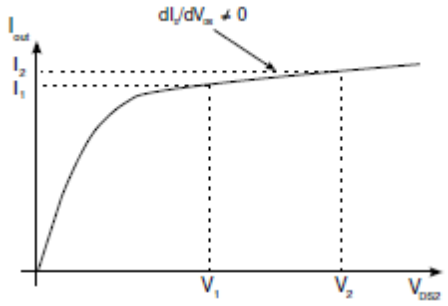

Fig 4.10: *current mismatch due to Early effect*

To avoid this deviation, since every current mirror is part of a cascode, an easy solution is to use the cascode itself to fix the V_{ds}, as shown in fig 4.11.

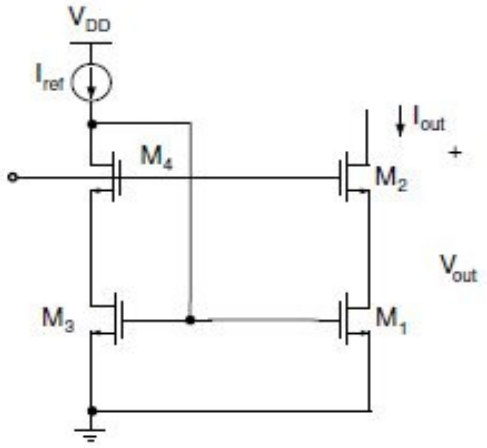

Fig 4.11: *cascoded current mirror*

Using this strategy the topology has robust current mirrors.

What still remains is to bias the cascode transistors. Investigating a particular case, for instance the NMOS cascodes in the output branch, it is clear that the goal is to have a DC voltage high enough to leave M_3 and M_1 in saturation, but low enough to leave M2 also in saturation when Vout is at its minimum, or, better, when the main amplifier operates at its full swing. So the desired voltage is:

$$V_{gs7} = V_{t2} + V_{ov2} + V_{ov1} \tag{4.13}$$

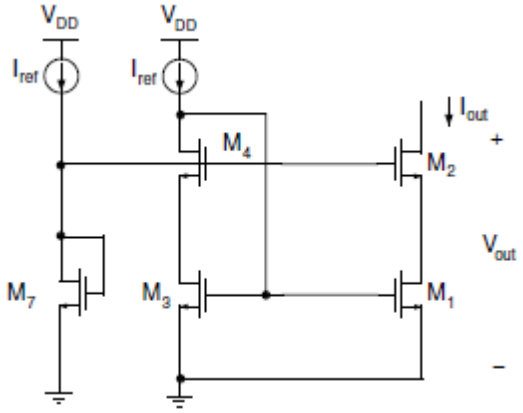

Fig 4.12: *cascode transistor biasing*

In fig 4.12 it is shown a diode connected MOS working as voltage reference; the current is generated by a cascoded current mirror.

Several different solutions have been taken into account; but the diode connected MOS has the advantage that it is simple, small and has a low and adjustable power consumption. Its output node has also a low impedance, which means that the rejection to the supply noise will be good.

4.1.6 Power down switches

It is mandatory to introduce in the design the possibility to power down the circuit even if the supply is still connected. It simply consists in adding to some nodes switches to power down the current sources or the biasing of some transistors so that it is impossible for the current to flow in certain conditions.
In fig 4.13 the power down switches are highlighted

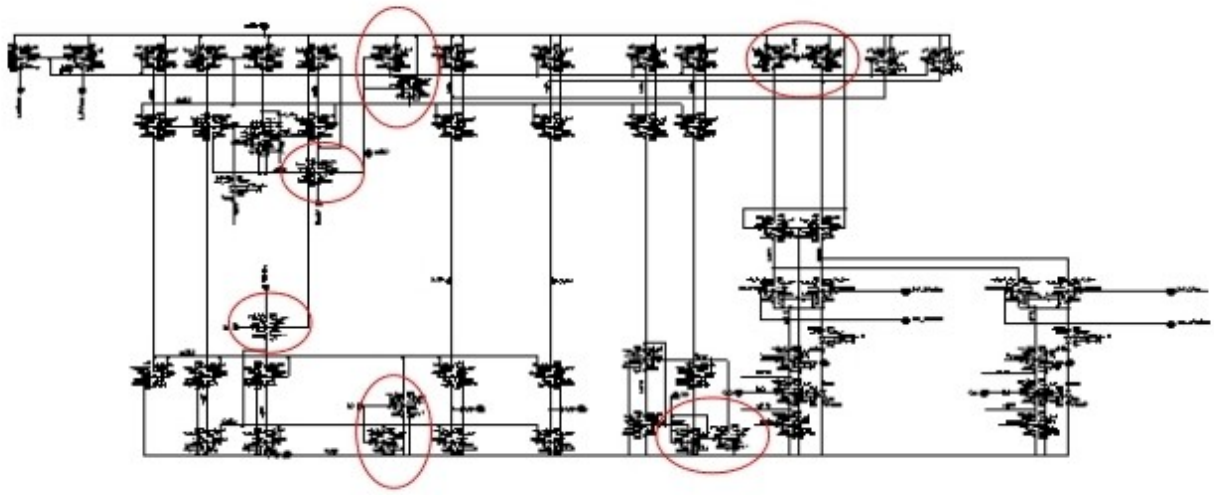

Fig 4.13: *power down switches*

The strategy rely mainly on:
-to stop the current coming from the external current source, obtained applying a switch in series between the current source and the mirror transistor;

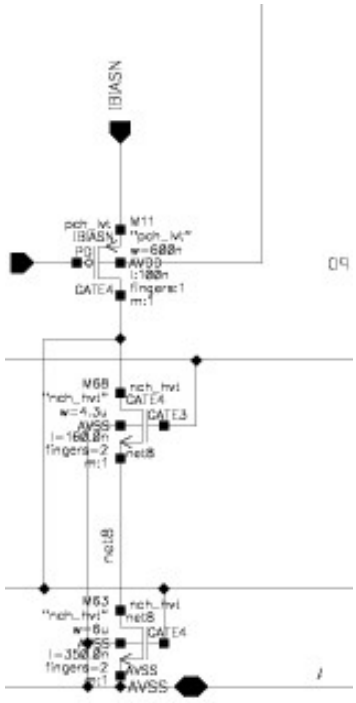

Fig 4.14: *current stopping switch*

-switch off the current mirrors, by shorting the gate voltage to ground or Vdd;

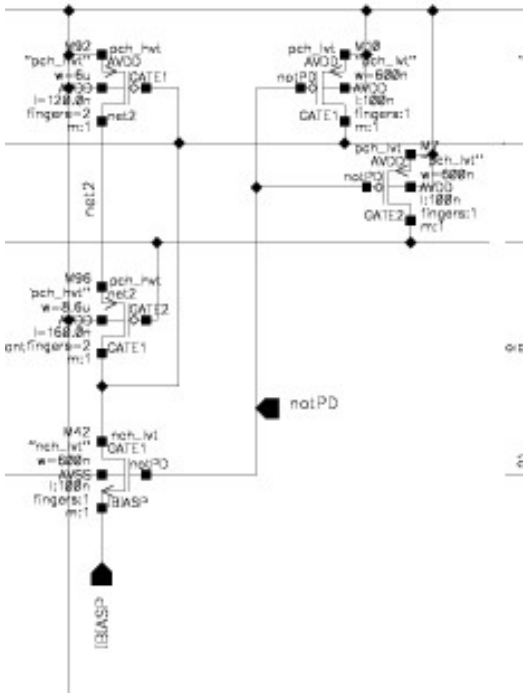

Fig 4.15: *turning off switches for current mirrors*

-shorting the current stored in big capacitance node to a power rail

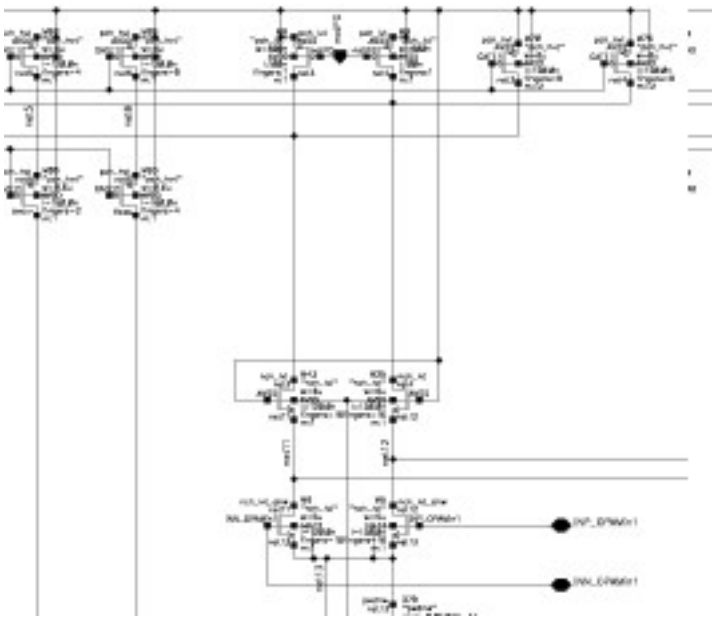

Fig 4.16: *discharge for big capacitance nodes*

The problem arises on the turning on of the circuit. In fact the biasing of the different cascode devices depends on the current generated by the current mirrors, which, at the same time, contain a cascode.

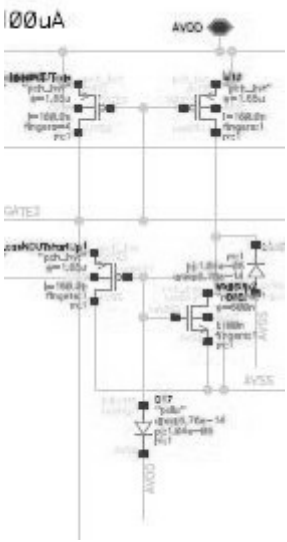

Fig 4.17: *start-up for the cascode bias*

The solution adopted is the one shown in fig 4.17: the circuit guarantees a low current

flowing in the diode connected MOS which bias the cascodes, allowing it to generate some voltage gap which will converge at the desired bias value once Vdd stabilizes.

4.1.7 Noise

Every transistor can be considered as a noise source, so it can be modellized as a current source having a spectral current density:

$$I_{out}^2 = 4kT\gamma g_m + \frac{KFg_m^2}{2\mu WLC_{ox}^2 f} \tag{4.14}$$

For simplicity a single pole model can be considered

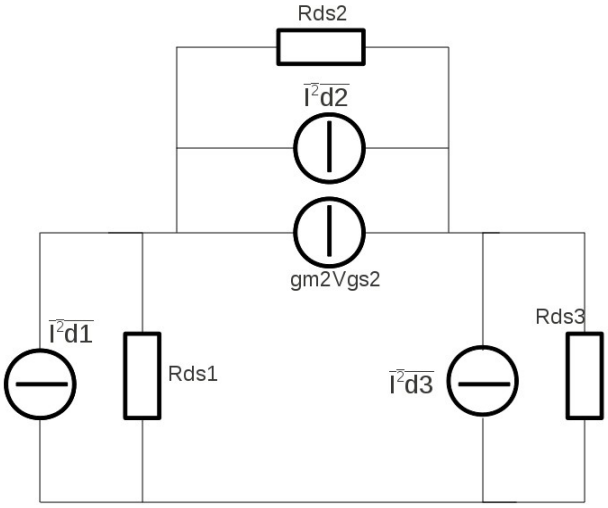

Fig 4.18: *single pole model*

where Rds3 indicates a cascoded structure.

The power spectral density of each device can be referred to the input:

$$V_{n1in}^2 = \frac{I_{ds1}^2(f)}{gm_1^2} \tag{4.15}$$

$$V_{n2in}^2 = \frac{I_{ds2}^2(f)}{gm_1^2 gm_1^2 r_{ds1}^2} \tag{4.16}$$

$$V_{n3in}^2 = \frac{I_{ds3}^2(f)}{gm_1^2} \tag{4.17}$$

Finally the total input referred noise will be the sum of each contribution:

$$V_{nOTA}^2 = \sum_{j=1}^{3} V_{njin}^2 \tag{4.18}$$

4.2.0 Stage 3 amplifier

These are the specifications to reach

Gain	70.79dB
GBW	1415MHz
Dynamic Range	1.1Vpp
Noise (inp_referred)	5.53nV/ $\sqrt{Hz}$
Current Capability	0.42mA

Some considerations: stage 5 op-amp was already at the limit for the technology given: in fact no more gain was affordable. Also the bandwidth was no freely increasable, since the bottleneck at the discussed mirror transistor would make the gain drop if the current was raised.

So a simple single stage amplifier with that kind of architecture can not reach the specifications required.

Also a boosted folded-cascode device would be ineffective, mainly for three reasons:

1) the current to have stable boosters can augment the power consumption up to 40%;
2) the doublet pole-zero which can degrade the slew rate;
3) the noise injected at the output node by the boosters, which have nearly the same amplification than the noise emitted by the input pair transistors.

The third reason is probably the most important, because noise specification, here but especially in the first and in the second MDAC, is the key parameter, or, better, the most difficult specification to achieve.

So it seems the case to use a two stages device. The behaviour concerning the voltage swing is the same as in the previous case: the second stage, which has a topology similar to the output part of stage 5 amplifier, can easily satisfy the requirements, and has the advantage that the gain requirements are less stringent.

Then, supposing that the second stage has a gain between 30 up to 40dB, a value easily reachable using a cascode structure, the first stage will need something around

20mV of swing. So a folded architecture is no more mandatory for neither of the stages, and this means that the current consumption can be reduced since there is a single path between the supply and the ground for each stage.

4.2.1 Telescopic cascode

The first stage is a telescopic cascode OTA, it means that a current source (cascoded, to improve the impedance, as described in the previous chapter) feeds of current an input pair which drives a cascoded load.

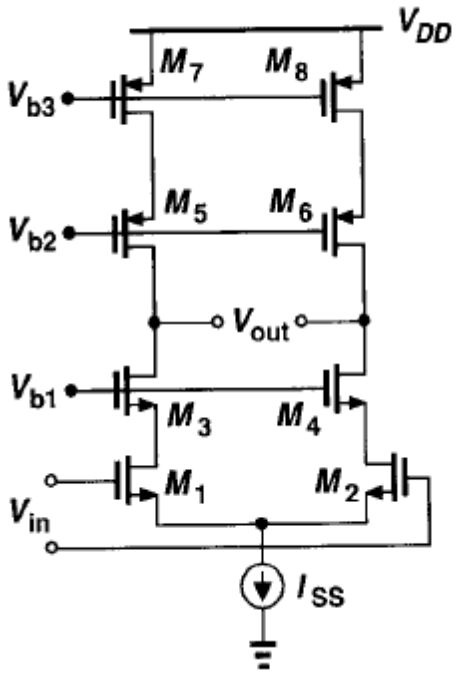

Fig 4.19: *telescopic cascode architecture*

The DC output voltage of the first stage is forced by a negative feedback to a value which, fed at the input of the second stage, bias this transistor to a fixed current value.

The gain characteristics can be easily extracted by the following:

$$A_0 = g_m R_{out} \tag{4.19}$$

where g_m is the transconductance of the input pair transistors and R_{out} is the real part of the impedance seen by the output node, that is

$$R_{out} = (r_{ds1} + r_{ds3} + (g_{m3} + g_{mb3})r_{ds1}r_{ds3}) \,||\, (r_{ds7} + r_{ds5} + (g_{m5} + g_{mb5})r_{ds5}r_{ds7}) \tag{4.20}$$

Instead the frequency behaviour depends on the location of the dominant pole, which is

$$f_{p1} = \frac{1}{2\pi \cdot R_{out} C_{out}} \tag{4.21}$$

which allows to calculate the gain bandwidth product:

$$GBW = \frac{g_{m1}}{2\pi \cdot C_{out}} \tag{4.22}$$

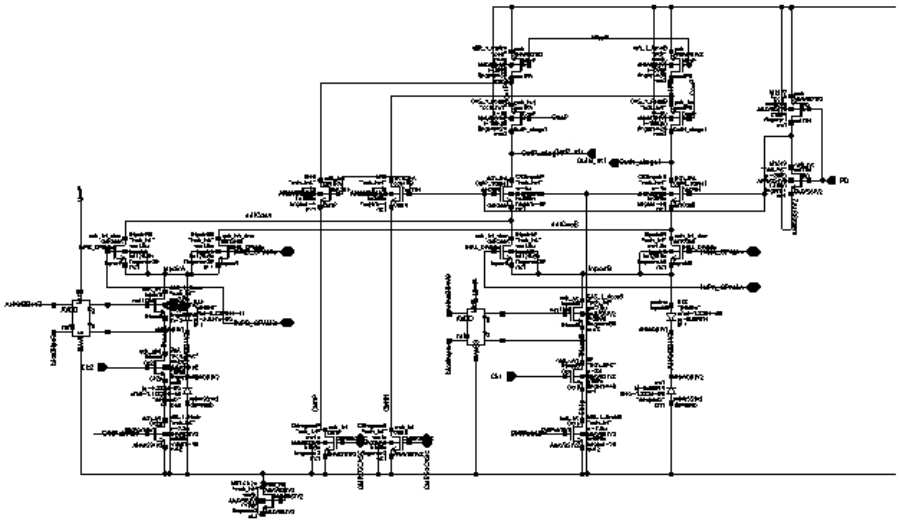

Fig 4.20: *telescopic cascode amplifier schematic*

As previously described, the input pair is doubled and driven by a switch like in stage 5 amplifier.

4.2.2 Output stage

This second stage is a couple of common-source transistors with a cascoded load. The DC gain is, as usual

$$A_0 = g_m R_{out} \tag{4.19}$$

and the dominant pole depends on the load and on R_{out}

$$f_{p1} = \frac{1}{2\pi \cdot R_{out} C_{out}} \quad . \tag{4.21}$$

There is no common path between ground and supply, but the previous stage differential architecture guarantees that the sum of the current in the two output branches remains constant.

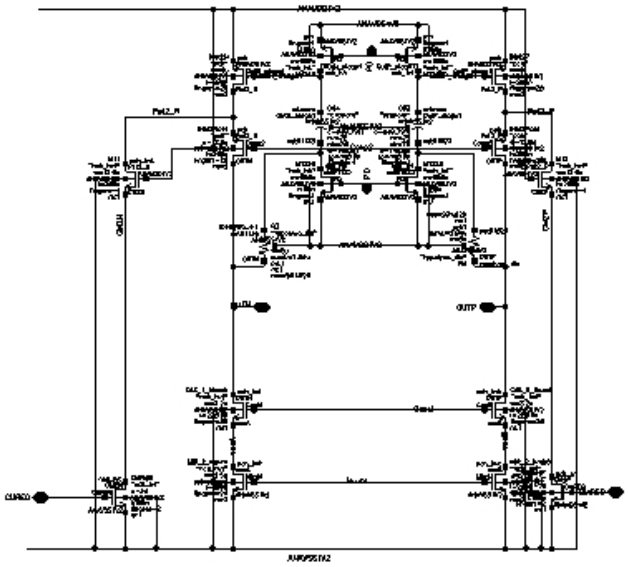

Fig 4.21: *stage 3 amplifier: output stage*

4.2.3 Two stages amplifier compensation

The main difference between a single stage and a two stages amplifier lies in the stability of the architecture itself. In fact, supposing that each stage has an unique dominant pole, it would be necessary a stage having challenging bandwidth characteristics.

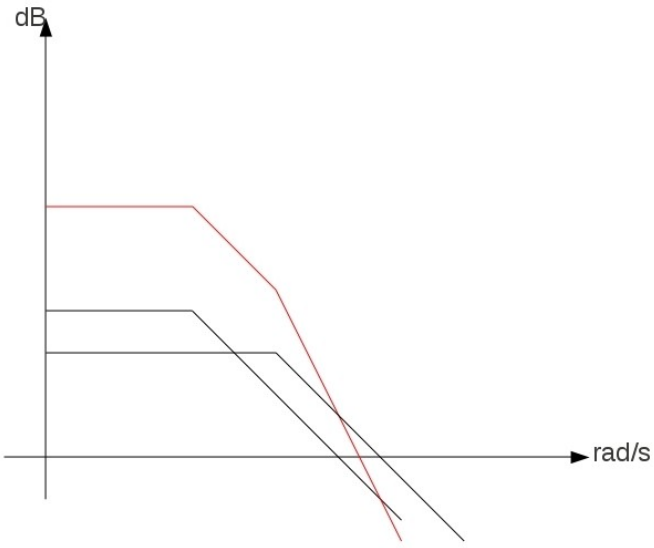

Fig 4.22: *Bode plot of a two stages amplifier*

As depicted in fig 4.21, supposing that each stage has a dominant pole behaviour, the whole system will have a dominant pole at the same frequency of the slowest stage and the secondary one at the frequency of the other pole. So stability in feedback configuration is difficult to achieve, since the two stage have a similar bandwidth, and since each stage secondary poles are not considered.

So a compensation is needed; the solution chosen is to use a "Miller" capacitor, which, intuitively, is a capacitor at the output of the first stage which impedance is amplified by the second stage OTA.

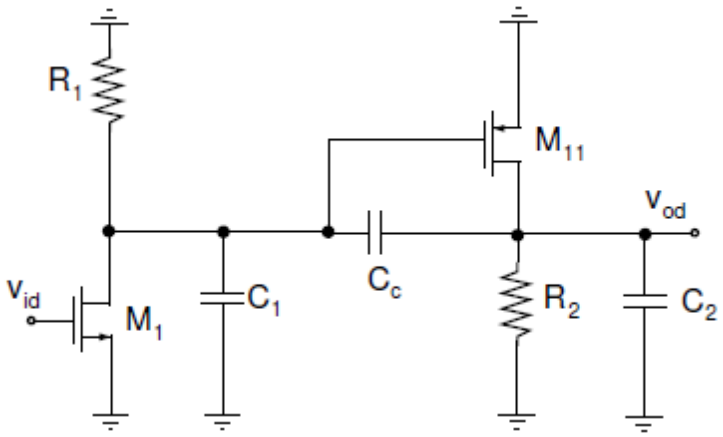

Fig 4.23: *small signal schematic of a two stages amplifier*

If Cc was placed in parallel to C1, then the behaviour of the two stages would be the following:

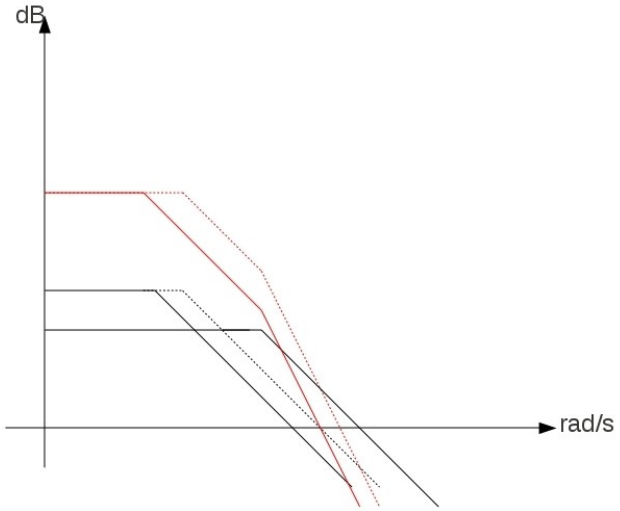

Fig 4.24: *parallel compensation of a two stages amplifier*

Since stability is a parameter linked to the ratio between the secondary pole and the 0dB frequency, a phase margin improvement is clear from the graph (both axes are logarithmic).

Instead, applying Cc in series between the outputs of the two stages, the small signal response becomes:

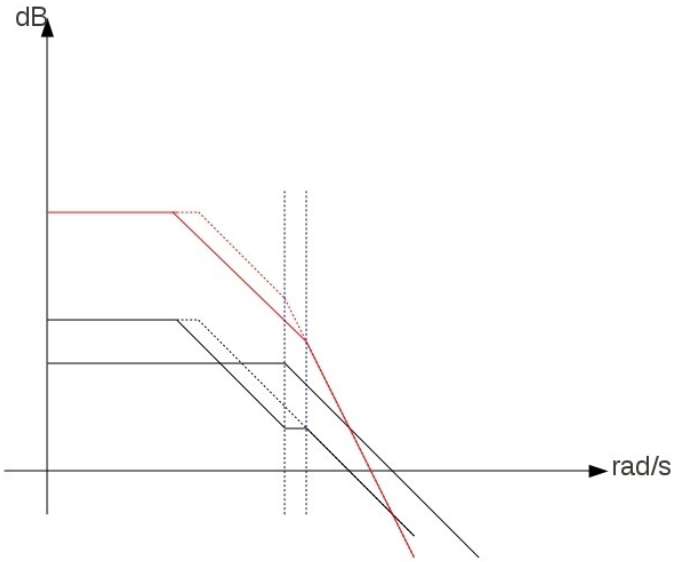

Fig 4.25: *Miller compensation of a two stages amplifier*

which is clearly better than the previous case, because the secondary pole shifts at higher frequencies leaving the 0dB cross unaltered.

Intuitively, what happens is that the effective impedance seen by the first stage is amplified by the second stage; but, at higher frequencies, when the gain of second stage starts to drop at its dominant pole, also the efficient value of the capacitor drops. This is why the first stage seems to have a zero located at the second stage dominant pole.

Also the second stage has a variation respect to the parallel case, in fact its bandwidth increases due to the fact that the efficient value of the Miller capacitor is reduced due to the attenuation imposed by the first stage.

This effect is called “pole splitting”, since in the small signal plot of the whole amplifier the dominant pole move towards DC and the secondary pole increases its value.

An other fact to note is the presence of a zero in the Miller capacitor path. Intuitively,

the problem can be analized as follow:

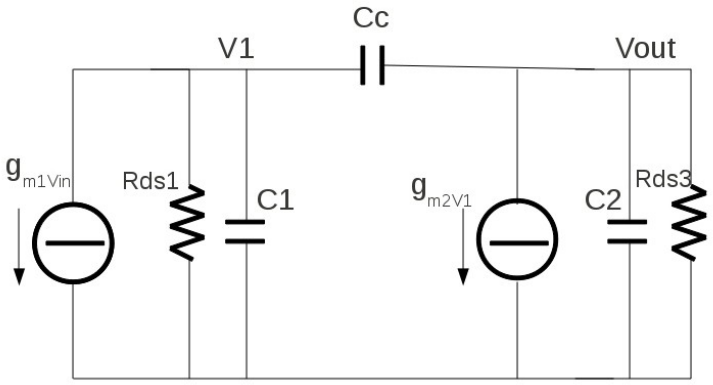

Fig 4.26: *small signal representation of a two stages amplifier*

The two stages gain and bandwidth can be assumed to be very similar: this is mainly due to the fact that the gm is proportional to the square root of the current. This assertion allows to assume that C_c is much bigger than C_1 and C_2, condition necessary to achieve stability. So, for a certain range of frequencies, C_1 and C_2 can be seen as open circuits; instead C_c can be assumed as a short, as shown in fig 4.25:

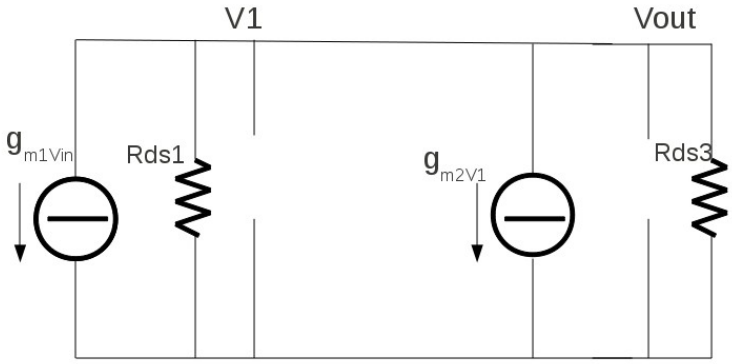

Fig 4.27: *middle band representation of a two stages amplifier*

So the whole second stage can be assumed as a diode connected transistor in parallel to a coscoded load, which means a total impedance of $1/g_{m2}$, making the R_{ds3} contribution uninfluent.

From the current point of view, instead, assuming the previous hypothesis, the system can be represented as follows:

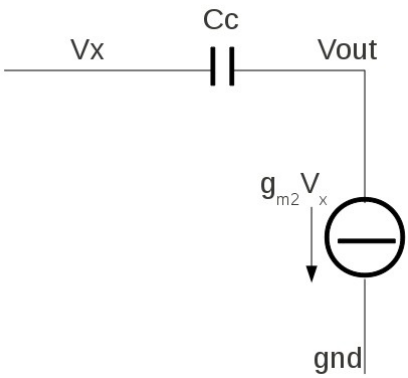

Fig 4.28: *second stage equivalent schematic*

so the zero can be easily found:

$$\omega_{zero} = \frac{g_{m2}}{C_C} \tag{4.23}$$

where all the current generated by the first stage flow through the second stage transconductance.

The problem can be easily solved placing a resistor in series with the Miller capacitor.

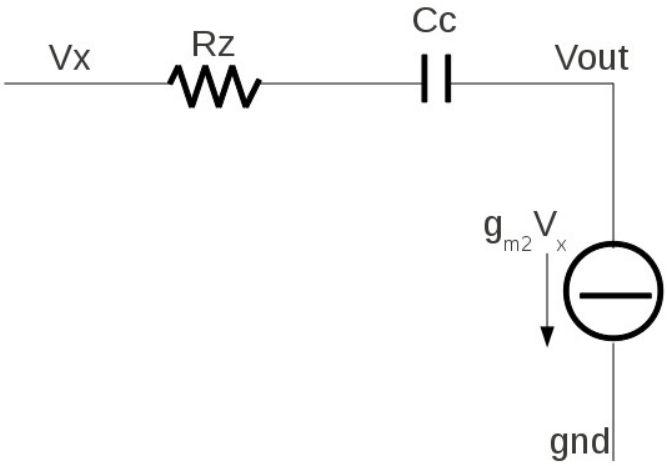

Fig 4.29: *nulling resistor*

So:

$$V_{out} = V_x \cdot (1 - g_{m2} R_z - \frac{g_{m2}}{sC_c}) \tag{4.24}$$

If $R_z = 1/g_{m2}$, then Vout can never be null for any frequency.

Chapter 5

Common mode feedback design

5.0 Introduction

The common mode feedback circuit is a critical component in this kind of architecture: in fact what is needed is a device having

1) a wider bandwidth than the amplifier itself;
2) a low gain to achieve easily the stability;
3) an input dynamic range equal to the output range of the amplifier.

For high bandwidth structures, it is typically used a switched capacitor structure: it has a negative gain and no problems concerning the input range.

But it needs an operative clock that is the double of the one used for the main converter: this means the necessity of a clock generator and can be risky from the noise point of view (it would implies a tone at the double of the operative frequency). So a continuous time feedback regulator is mandatory.

5.1 Stage 5 common mode regulator topologies

The input dynamic range is probably the most critical specification for this device in low voltage technology, so it has been the first taken into account. Different architectures have been investigated.

5.1.1 Inverter based comparator

The first topology considered consists in two inverters having two degeneration resistors, as shown in fig

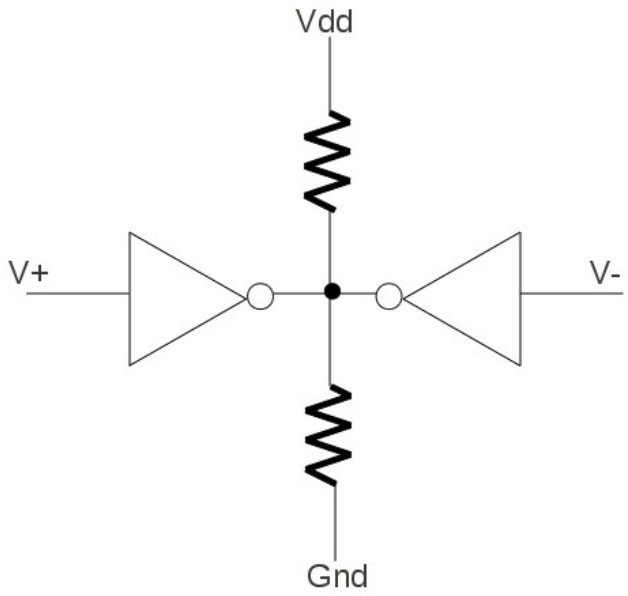

Fig 5.1: *inverter based comparison*

This structure has a wide input range and a gain adjustable by changing the resistors value (which are made using transistors); the current consumption is instead related to the inverters sizes. The voltage value at the central node would be compared by an amplifier to the one produced by the same structure biased at AGND.

The drawback of this topology, however, is that it has only two degrees of freedom, the W/L ratio and the value of the resistors; so, even if it can be good enough for stage 5 amplifier, it seems no possible to use it for the wider bandwidth amplifiers.

5.1.2 Current based comparison

Another solution proposed is the one shown in fig 5.2. Here the current generated by the pair connected to the main amplifier output is fed, through a mirror, into a MOS biased at AGND, which is the desired voltage value for the DC output. M_1 and the couple M_{2+} M_{2-} together have the same effective W/L ratio.

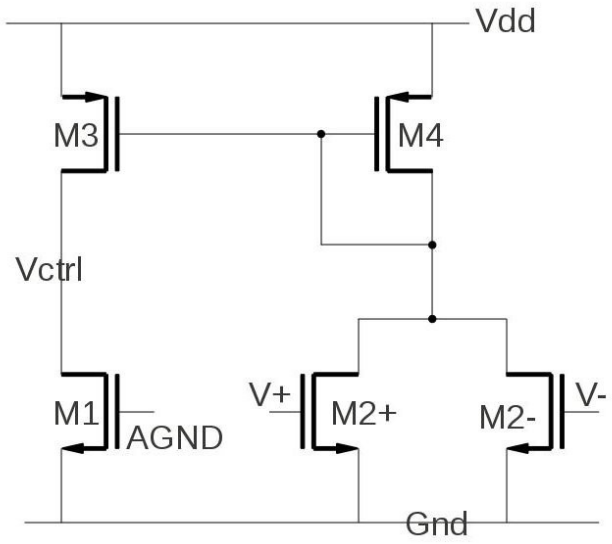

Fig 5.2: *current based comparison*

The main disadvantage of this architecture is that the current consumption is defined only by the device geometries; in particular, since M2+ e M2- has to be in saturation in the whole swing (that means from 0.325 up to 0.875V), their Vth must be low (at maximum 0.3V): so their overdrive must be high, which means also an high current consumption. Simulations and calculations demonstrates that this architecture is not power efficient.

5.1.3 Voltage buffers comparison

A different solution proposed is the use of a buffered Miller amplifier, as depicted in fig 5.3:

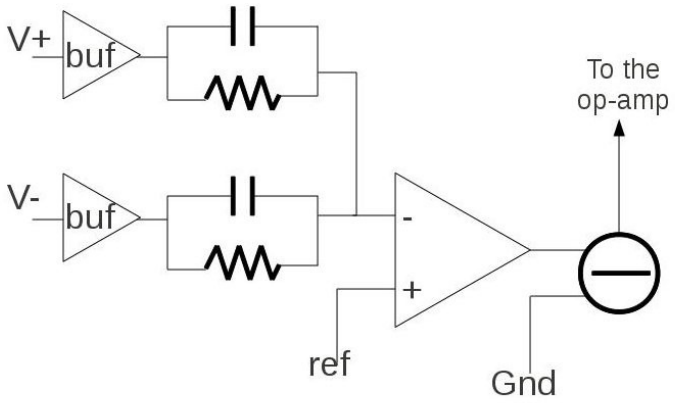

Fig 5.3: voltage buffers solution

The buffer is a simple source followerMOS in common drain configuration made using a native transistors. Its purpose is to widen the input range of the amplifier, which would not be enough instead.
Native transistors are devices made using a very low doping, or leaving the silicon intrinsic. So they are big devices, but with the advantage that the V_{th}, can be very low or eventually negative (see Appendix B).

The current can be tuned by the current mirror which bias the native MOS, so the bandwidth over the power consumption can be optimized. A capacitor has also been added to introduce a zero in the common mode path: this can reduce the current flowing through the buffer to achieve the same bandwidth, so its dimensions can be minimum, reducing also the capacitive load for the main amplifier.

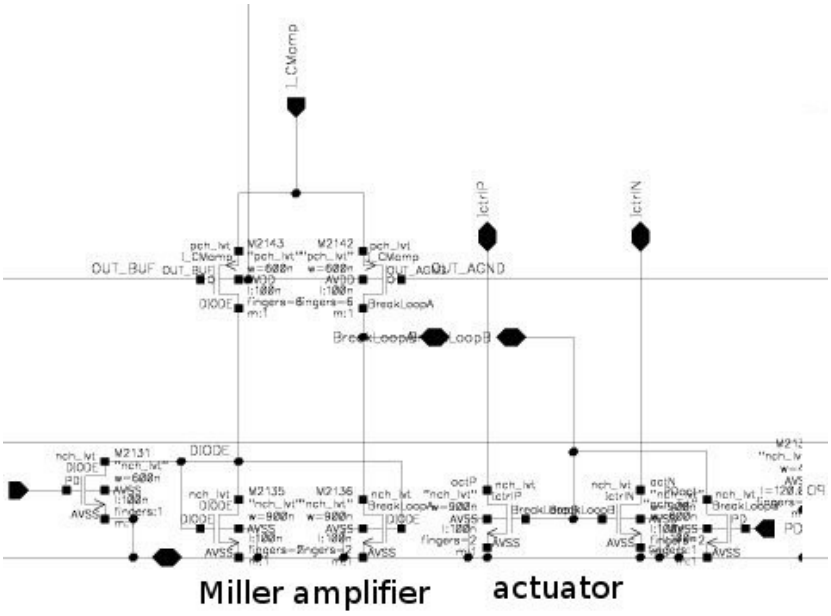

Fig 5.4: *amplifier and actuator*

Finally a simple Miller, single ended amplifier is used to regulate the actuator for the common mode feedback: it compares the voltage provided by the sensor to the voltage provided by another sensor, identical to the first one, but biased at AGND.

Riassuming, the differential output of the OTA is converted into a common mode voltage by the sensor, compared to the desired value by the amplifier and then fed into

an actuator, which is a simple MOS converting the previous voltage into a current.

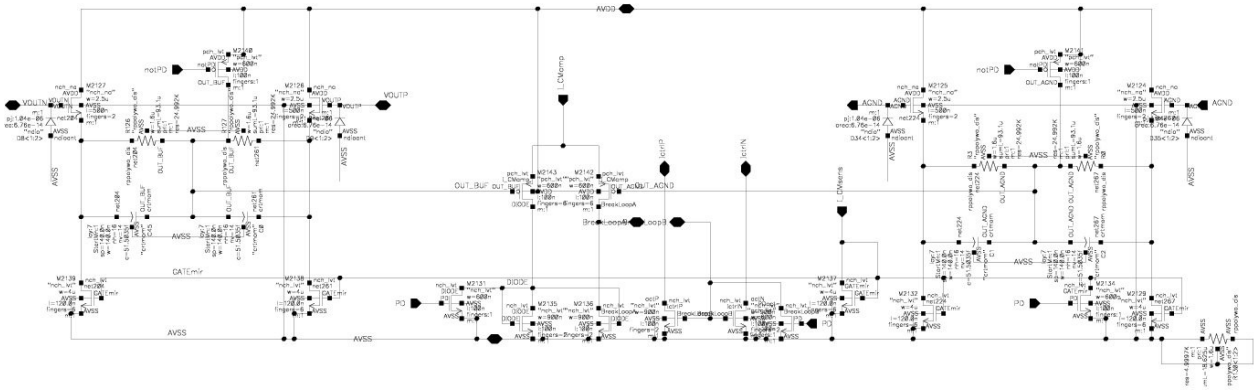

Fig 5.5: *stage 5 common mode regulator schematic*

As previously said, the common mode open loop gain must be as low as possible (the lower bound can be the voltage offset for the DC behaviour and the CMR specification for the AC behaviour) to satisfy the stability requirement, so the g_m of the actuator must be the lowest possible.

The solution chosen is to split the current flowing in the output branches of the main amplifier in two parts: a DC current provided by a mirror, and an other current regulated by the actuator.

The ratio of these two values is related on the difference of the current magnitude flowing in the output branches at the full dynamic; in few words, there must be some current flowing in the actuator when the amplifier is completely unbalanced, otherwise the common mode regulation would not be effective anymore.

Another advantage coming from this current splitting is linked to the fact that the MOS actuator can have a very small size (but big enough to guarantee saturation), so the load of the Miller amplifier (the comparator block) can be minimum, a fact that permits to reach stability in a easier way.

5.2 Stage 5 common mode rejection

As previously said, one of the purposes of the common mode regulator is to attenuate the common mode signals which can affect the behaviour of the fully differential amplifier. But at the same time the gain must be minimized in the loop, to reach stability

without affecting the current consumption. Clearly these two affermations contraddict each other: in fact the common mode rejection is directly proportional to the loop gain.

So, to improve the rejection, the solution chosen has been to improve the intrinsic rejection of the main amplifier. By definition, the CMRR is:

$$CMRR = A_{dm}/A_{cm}$$

To calculate the common mode gain in a differential structure it is convenient to split the input pair:

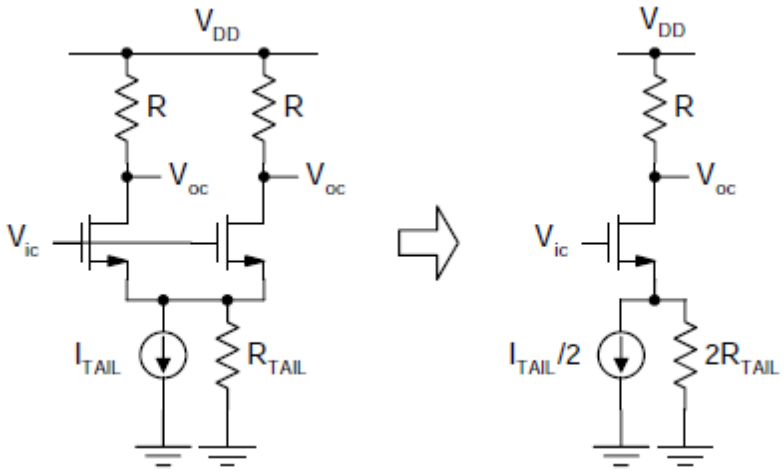

Fig 5.6: *common mode signal equivalent schematic*

So,

$$A_{cm} = \frac{g_m}{1+2g_m R_{tail}} \cdot (R \| r_0[1+2g_m R_{tail}]) \qquad (5.1)$$

which means that the only parameter available to manage the common mode rejecton is R_{tail}: increasing R_{tail}, the rejection is improved.

This is why the current source for the input pair has been cascoded:

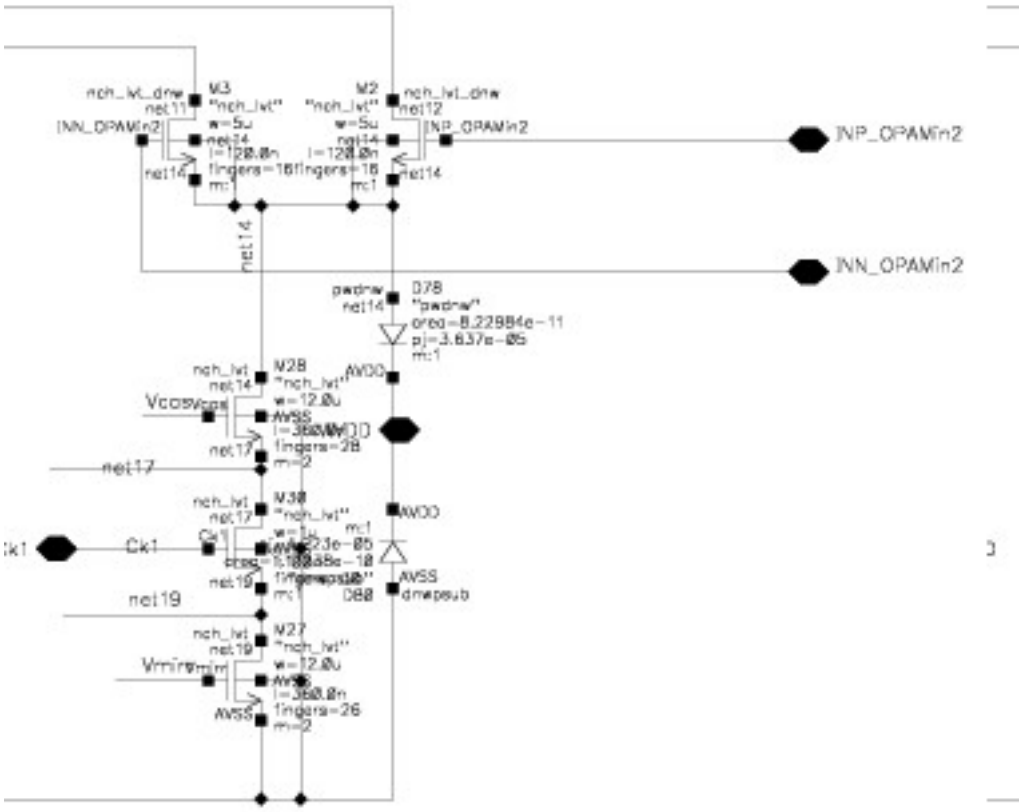

Fig 5.7: *cascoded tail resistance*

5.3 Stage 3 commmon mode feedback regulator

The architecture used for the blocks of the common mode regulator is the same used for the single stage amplifier: two buffers to sense the common mode voltage, a comparator and an actuator.

The main problem rises from the second stage of the main amplifier: there, in fact, the common mode signal is amplified instead of being rejected, since there is no common path for both the positive and negative signals, so there is no intrinsic common mode rejection.

In fact, in the first stage, the common mode attenuation will be proportional to the impedance of the current generator which feeds the input pair; instead in the second stage the same signal will have an amplification:

$$A_{CM} = g_{m2}r_{02} \tag{5.2}$$

There is also a second aspect to take into account: when the OTA is connected in negative feedback (or, at least, in a configuration detected by the differential signal perceives as a negative loop), there is a positive path for the common mode through the capacitors, as shown in fig 5.8:

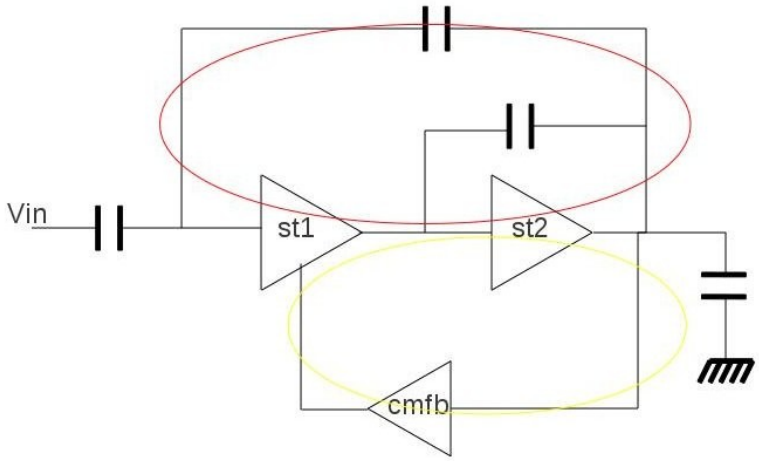

Fig 5.8: *different common mode loops*

In fact, through the yellow path in fig 5.8, the signal crosses three inverting devices, the two stages of the amplifier and the CM Miller amplifier. Instead, along the red path, the inverting blocks are only two.

This happens because the (negative) amplification introduced by the capacitive path:

$$feedback\ factor = \frac{C_f}{C_L + C_f + C_S} \tag{5.3}$$

is bigger than the one of the regulator path. In fact, although the Miller amplifier can have a gain of approximately 15dB, the actuator behave like an attenuator: the MOS used is very small, and its load is a mirror current source which, since its bias is in saturation but near the liner region, has a very high g_{ds}. The connection was made there, and not at the output node, because this actuator must act as a current source, and if the load was too high, there would be some Miller effect on the C_{gd}, which would deteriorate the bandwidth behaviour of the common mode loop, already critical.

So a solution could be to increase the gain of the regulator, but it is not possible because of the stability requirement. In fact, to obtain a stable loop and at the same time a CMR high enough for, at least, the whole bandwidth of the ADC (that would be something like 20dB of attenuation for 200MHz), a comparative block having much more bandwidth than the main amplifier is needed (and this means that the common mode regulator would consume approximately as much as the OTA, that is not acceptable).

The first improvement is to reduce the common mode gain in the main amplifier. Since it is not possible on the second stage (leaving the architecture as previously designed), it must concern only the first stage.
There, as already discussed, the rejection is proportional to the equivalent resistance of the tail transistors:

$$A_{cm} = \frac{g_m}{1+2g_m R_{tail}} \cdot (R \| r_0 [1+2g_m R_{tail}]) \tag{5.4}$$

A simple solution is to boost the cascode:

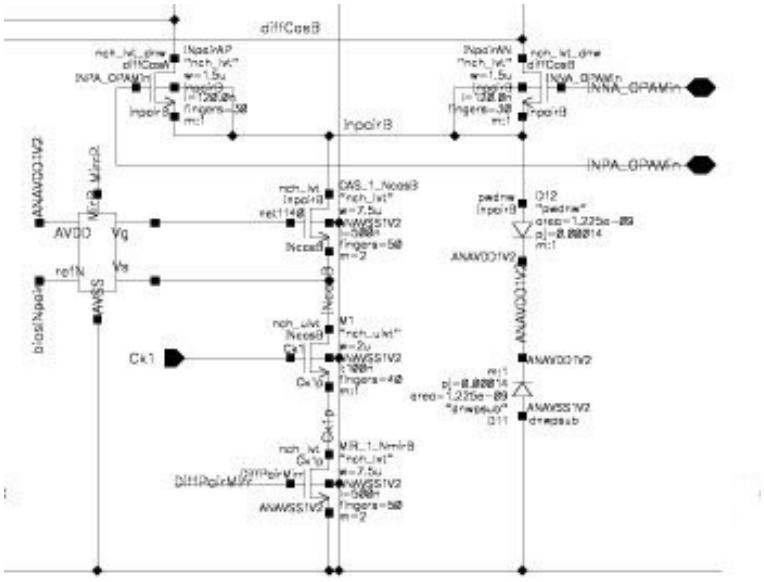

Fig 5.9: *boosted tail resistance*

In this way the first stage rejection improves, so the gain of the second stage can be

partially compensated. The booster architecture is the following:

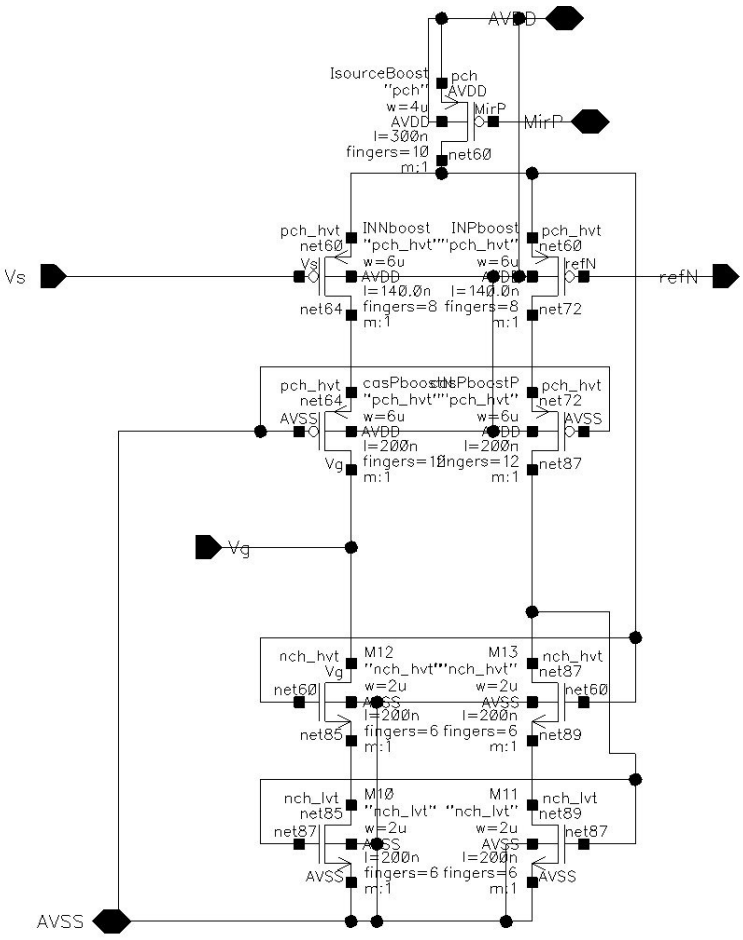

Fig 5.10: *booster implementation*

It is a telescopic cascode single ended architecture which compares the input voltage to a reference value, which is the biasing gate voltage for the cascode transistor.

The second improvement is to split the whole common mode loop in two, one for every stage of the amplifier. In this way the total gain of the loop would be the sum of the gains of the two loops, and it is convenient, since gm has a square root relation with current, and a single amplifier would cost much more power.

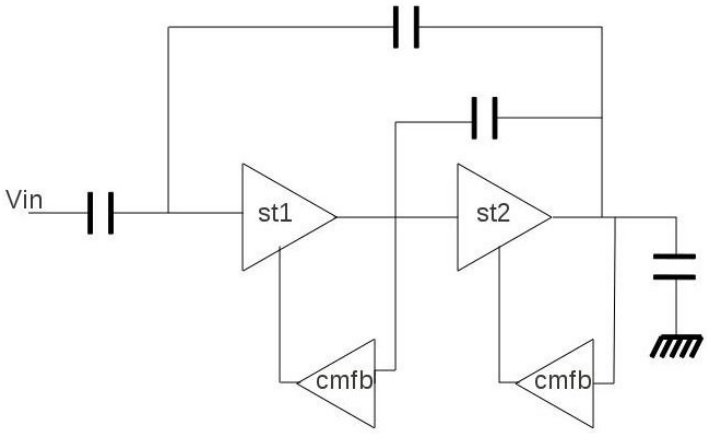

Fig 5.11: *common mode regulators*

Both the loops have the same architecture as in stage 5. The first, in particular, instead of taking AGND as the reference parameter, uses a value extracted from a dummy architecture that replies the bias voltage needed by the input of the second stage, as shown in fig 5.12.

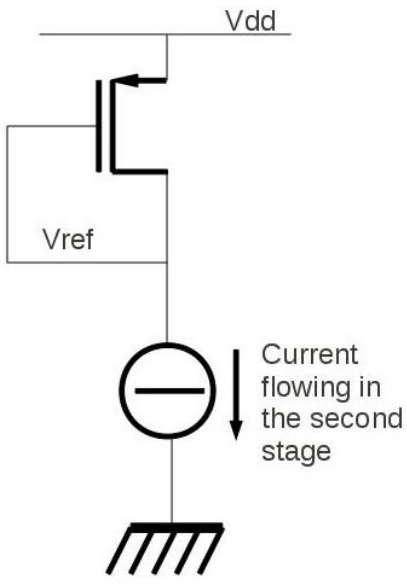

Fig 5.12: *reference voltage for the first CM loop*

Of course, the current generator and the diode connected transistor are scaled version of

the effective one to reduce the power consumption.

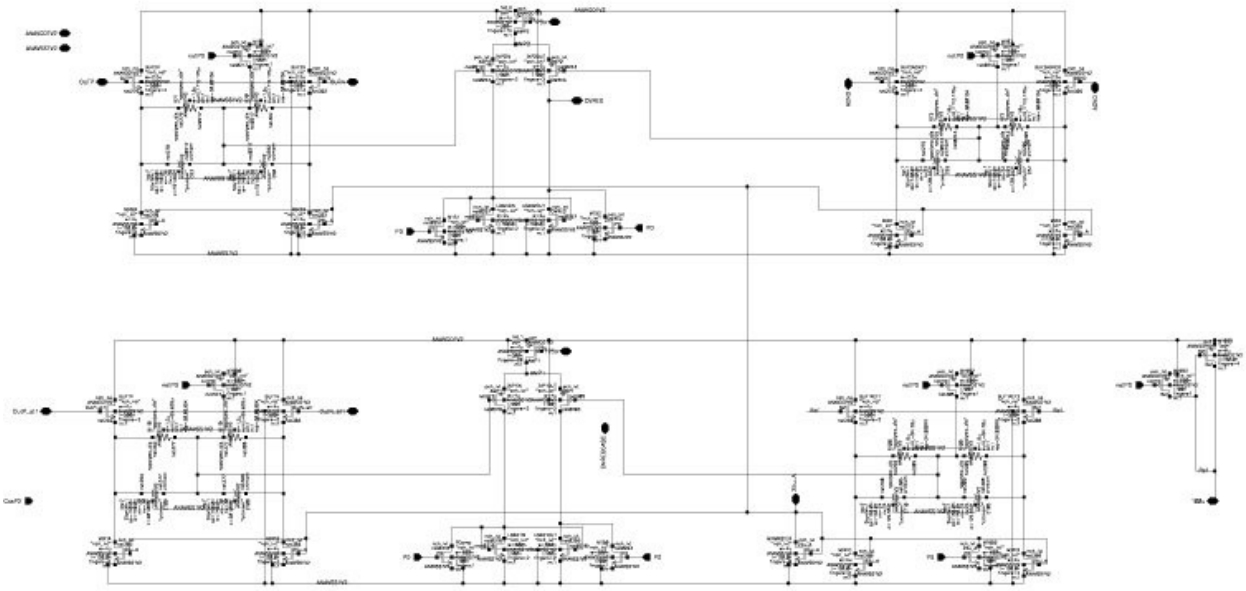

Fig 5.13: *complete common mode regulator schematic*

Chapter 6

Amplifier characterization

6.0 Introduction

To verify the proper behaviour of the device, several parameters have been taken into account. The whole characterization set has been performed considering the corners setup provided by the foundry at a temperature compatible with the specifications and at the full span of the voltage supply; in fact variations in fabrication process, ambient temperature and supply voltage affect the electrical performance of the transistors. For example, a higher temperature and a lower supply voltage make the transistor operate slower. This is why the operation of the circuit has been verified by simulating the design in slow (SS) corner, typical corner (TT) and fast corner (FF), and also by simulating the design with fast NMOS and slow PMOS corner (FN), and slow NMOS and fast PMOS corner (SN).

6.1 Reusability

In an industrial environment a key point is the reusability of the design and the efficiency in sharing the components between different team members. So it is important to use the same strategy in making the symbols to optimize the work during the different design steps.

In fig 6.1 the symbol used to represent the complete amplifier is depicted:

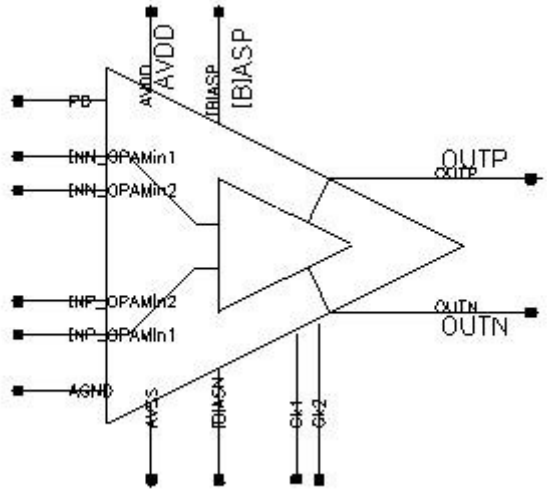

Fig 6.1: *complete amplifier symbol*

This symbol contains the differential amplifier and also the common mode feedback regulator, which are represented in fig 6.2; splitting differential and common mode part allows to have an easier debug during the design process.

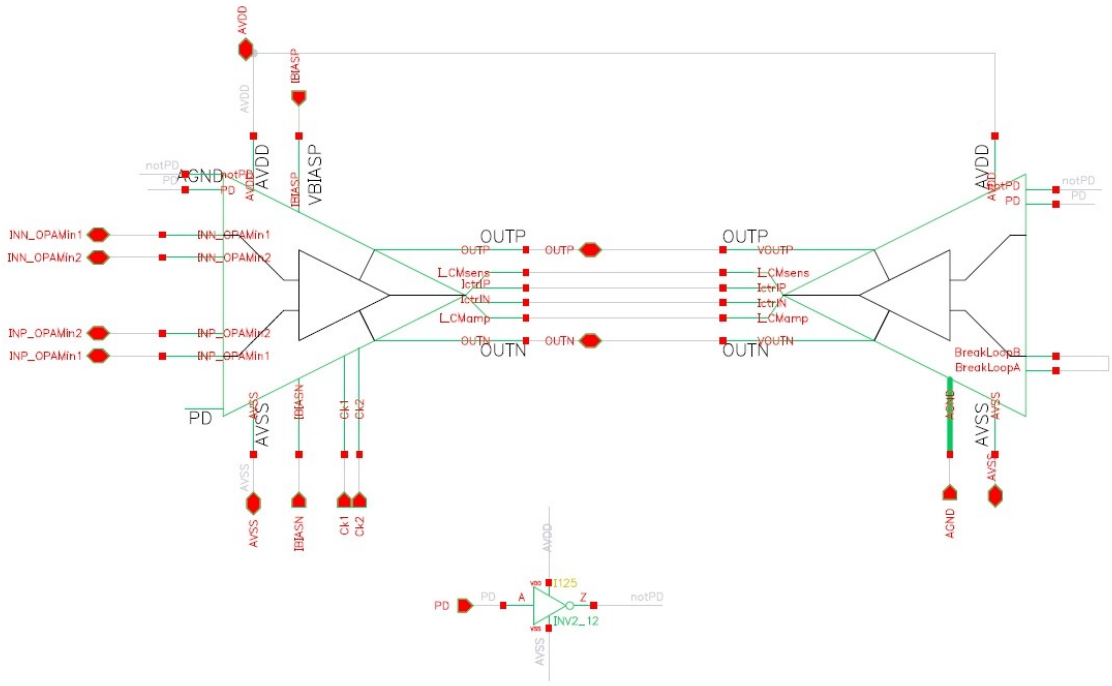

Fig 6.2: *differential and common mode symbols*

6.2.1 Stage 5 op-amp AC behaviour

The first testbench relies on the DC gain, the bandwidth and the stability of the amplifier. As depicted in fig 6.3 the main goal is to obtain the environment in which the device is supposed to operate. As already said, it is important to use of modular hierarchy in the symbols to make it simpler to interface other blocks in successive steps of the design.

This justify the use of dummy blocks, in order to simulate the effective load of the device: in fact, in the converter, there will be a chain of stacked amplifier, so a dummy amplifier can be an efficient way to simulate the effective impedance seen.

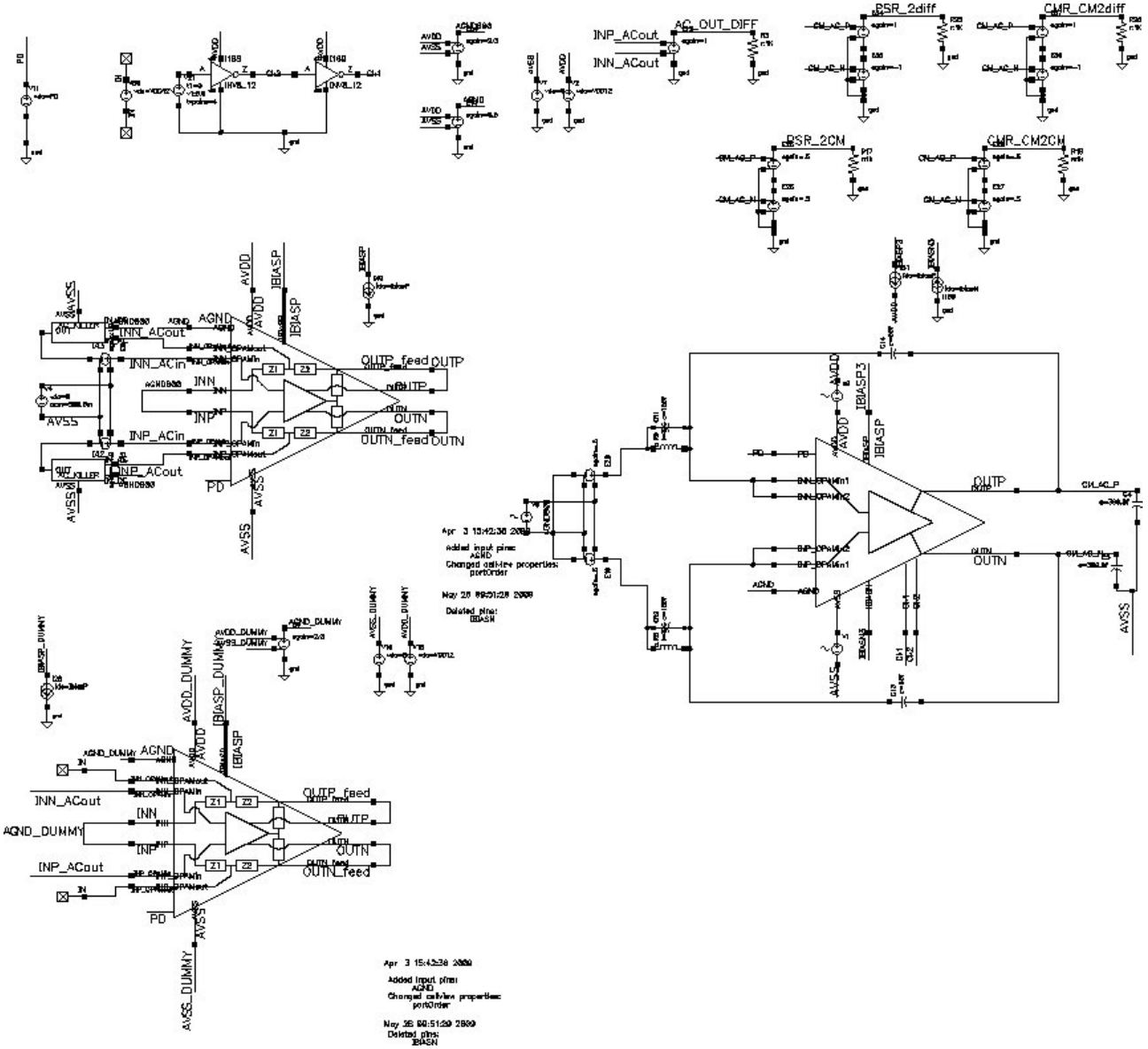

Fig 6.3: *AC behaviour testbench*

In table 6.1 the AC behaviour is plotted, considering all corners at the minimum and maximum of the temperature range.

Corners	gainBwProd(Hz)	gain(dB) openloop @100kHz
Alltyp1	1,12E+009	57,24
Alltyp2	1,25E+009	56,71
Alltyp3	1,36E+009	57,88
Alltyp4	9,43E+008	55,5
Alltyp5	1,02E+009	56,71
TffRtypCtyp1	1,24E+009	55,62
TffRtypCtyp2	1,35E+009	56,73
TffRtypCtyp3	9,34E+008	**53,98**
TffRtypCtyp4	1,00E+009	55,1
TfnspRtypCtyp1	1,23E+009	56,86
TfnspRtypCtyp2	1,33E+009	57,97
TfnspRtypCtyp3	**9,30E+008**	55,76
TfnspRtypCtyp4	9,97E+008	56,93
TsnfpRtypCtyp1	1,27E+009	56,49
TsnfpRtypCtyp2	1,38E+009	57,74
TsnfpRtypCtyp3	9,56E+008	54,9
TsnfpRtypCtyp4	1,03E+009	56,21
TssRtypCtyp1	1,26E+009	57,54
TssRtypCtyp2	1,36E+009	58,87
TssRtypCtyp3	9,51E+008	56,17
TssRtypCtyp4	1,02E+009	57,7

Table 6.1: *stage 5 AC characterization*

It can be noted that the DC gain is not too much sensitive to the corner variation, it is around 6dB for the whole lot, and that the secondary pole is enough close to the 0dB intercept. This means that the device will be stable, in particular these are the values extracted for the phase margin:

Corners	PhaseMargin (°)		Corners	PhaseMargin (°)
Alltyp1	70,14		TfnspRtypCtyp3	70,98
Alltyp2	68,03		TfnspRtypCtyp4	72,77
Alltyp3	69,6		TsnfpRtypCtyp1	72,38
Alltyp4	70,68		TsnfpRtypCtyp2	71,19
Alltyp5	71,9		TsnfpRtypCtyp3	75,58
TffRtypCtyp1	75,32		TsnfpRtypCtyp4	73,89
TffRtypCtyp2	77		TssRtypCtyp1	61,98
TffRtypCtyp3	77,25		TssRtypCtyp2	**61,54**
TffRtypCtyp4	78,71		TssRtypCtyp3	67,56
TfnspRtypCtyp1	68,71		TssRtypCtyp4	65,6
TfnspRtypCtyp2	70,8			

Table 6.2: *stage 5 amplifier stability*

The way the phase margin is calculated is the following. At first, all the signal sources are turned off; then, in an arbitrary point of the loop of interest, a block named AC_killer is added.

As depicted in fig this is an ideal block which cuts the whole AC behaviour leaving unchanged the DC value of the signal.

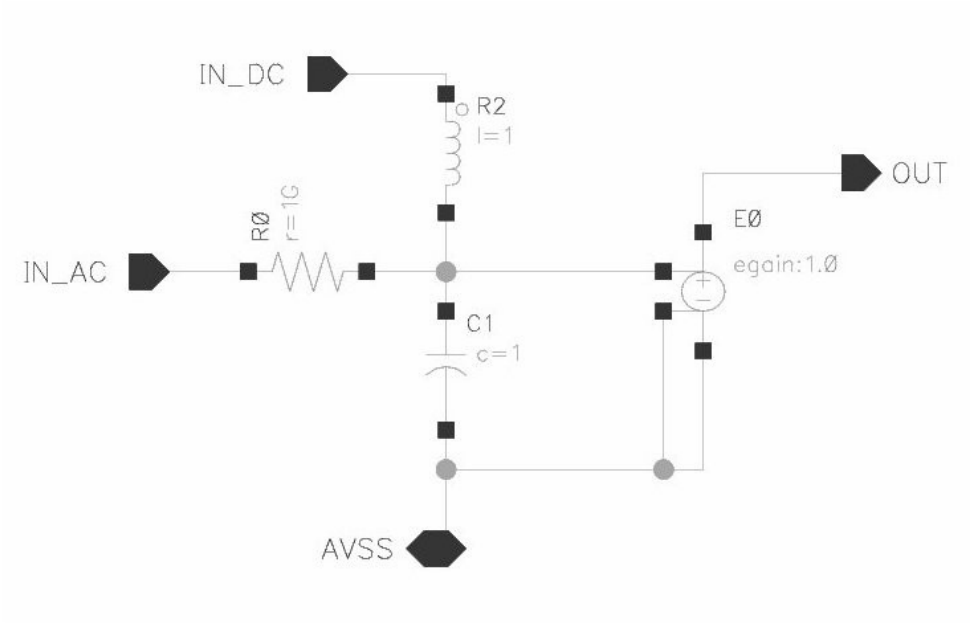

Fig 6.4: *the "AC_killer" block*

The input of this device is connected at the same point of the loop to a dummy amplifier; in this way the impedance seen in the loop under exam remains approximatively unchanged.
The output instead is used to bias a small-signal source and then fed in the point of the topology where the loop was broken.

So it is possible to plot the AC behaviour of the loop without affecting neither the internal impedances, neither the bias point of the devices, in a certain way recreating the same condition in which the real amplifier is expected to operate. In fig 6.5 the Bode plot is shown; the device seems to be robust among the whole corner lot, since the variations affect the gain and the secondary pole location, but leaving it still inside the specifications required.

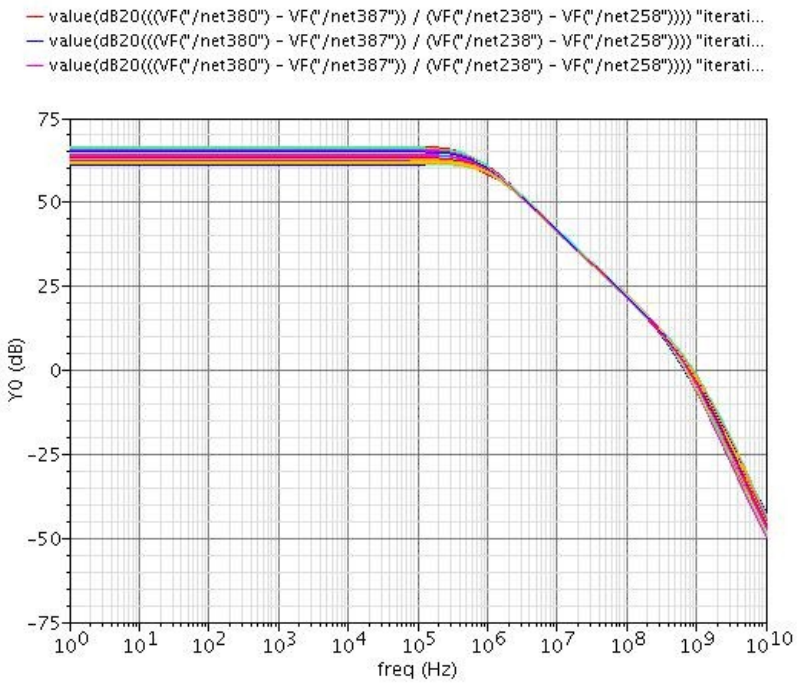

Fig 6.5: *stage 5 Bode plot*

6.2.2 Stage 5 CM behaviour

The parameters that the common mode controller must guarantee are an acceptable DC output offset, a sufficient common mode error suppression and the stability of the regulator itself.

Corners	DC offset [V]	Effective_swing
Alltyp1	0,0034	1,3992
Alltyp2	0,0079	1,4164
Alltyp3	0,0044	1,6304
Alltyp4	0,0056	1,1708
Alltyp5	-0,0002	1,3876
TffRtypCtyp1	0,0045	1,4484
TffRtypCtyp2	0,0004	1,6664
TffRtypCtyp3	-0,0043	1,2272
TffRtypCtyp4	-0,0118	1,4256
TfnspRtypCtyp1	0,0059	1,4124
TfnspRtypCtyp2	0,0024	1,6264
TfnspRtypCtyp3	0,0019	1,1820
TfnspRtypCtyp4	-0,0039	1,3996
TsnfpRtypCtyp1	0,0114	1,4124
TsnfpRtypCtyp2	0,0063	1,6348
TsnfpRtypCtyp3	0,0122	1,1508
TsnfpRtypCtyp4	0,0033	1,3788
TssRtypCtyp1	0,0112	1,3816
TssRtypCtyp2	0,0074	1,5984
TssRtypCtyp3	**0,0153**	**1,1168**
TssRtypCtyp4	0,0082	1,3376

Table 6.3: *stage 5 offset and swing*

The offset is inversely proportional to the common mode loop gain, and, since its maximum absolute value is about 15mV, its main effect is to reduce the effective output voltage swing.

In fig 6.6 the maximum ripple is shown, it is obtained forcing the amplifier to represent a sinusoidal wave at full swing. Clearly the different mean values of these sine are linked to the fact that V_{AGND} (which is the DC voltage output) depends on V_{dd}, in fact $V_{AGND} = V_{dd}/2$ by definition.

In table the common mode loop gain and the phase margin are shown:

Corners	loop_gain @100kHz (dB)	phaseMargin (°)
Alltyp1	39,08	70,14
Alltyp2	39,73	68,03
Alltyp3	40,63	69,6
Alltyp4	37,16	70,68
Alltyp5	37,92	71,9
TffRtypCtyp1	37,38	75,32
TffRtypCtyp2	37,89	77
TffRtypCtyp3	34,18	77,25
TffRtypCtyp4	34,43	78,71
TfnspRtypCtyp1	40,06	68,71
TfnspRtypCtyp2	40,62	70,8
TfnspRtypCtyp3	37,61	70,98
TfnspRtypCtyp4	38	72,77
TsnfpRtypCtyp1	37,35	72,38
TsnfpRtypCtyp2	39,67	71,19
TsnfpRtypCtyp3	**33,99**	75,58
TsnfpRtypCtyp4	36,47	73,89
TssRtypCtyp1	40,39	61,98
TssRtypCtyp2	42,35	**61,54**
TssRtypCtyp3	37,31	67,56
TssRtypCtyp4	39,55	65,6

Table 6.4: *stage 5 CM AC behaviour*

The way these values have been extracted is depicted in fig 6.7; the goal is to break the common mode loop in the amplifier in feedback configuration, and sense the transfer function between a common mode input and a point anywhere in the loop**.** As in the differential case, an AC_killer block has been used, and a dummy block has been inserted to recreate the effective impedance seen where the loop was broken. The ratio between the input of the AC_killer and the output of the AC source represent the loop behaviour.

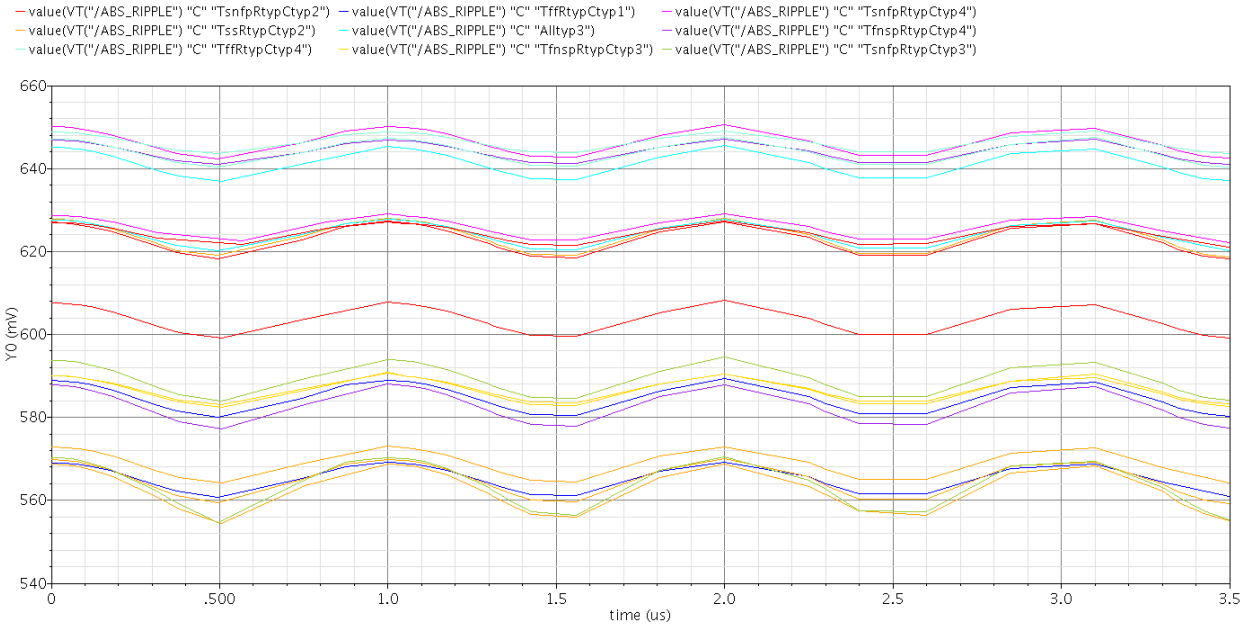

Fig 6.6: *stage 5 DC ripple*

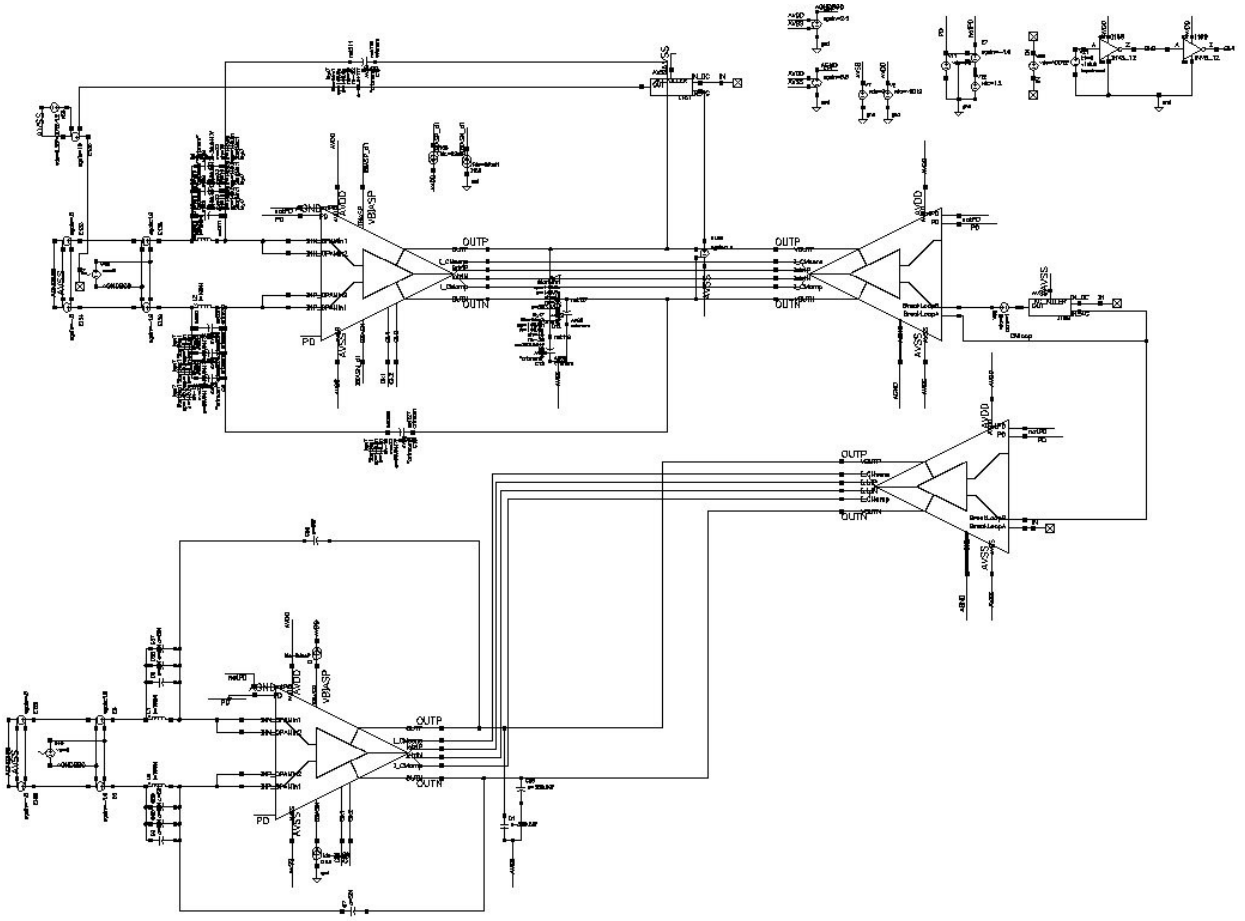

Fig 6.7: *common mode testbench*

Finally it is possible to plot the CMR by applying a common mode signal at the input and sensing the output:

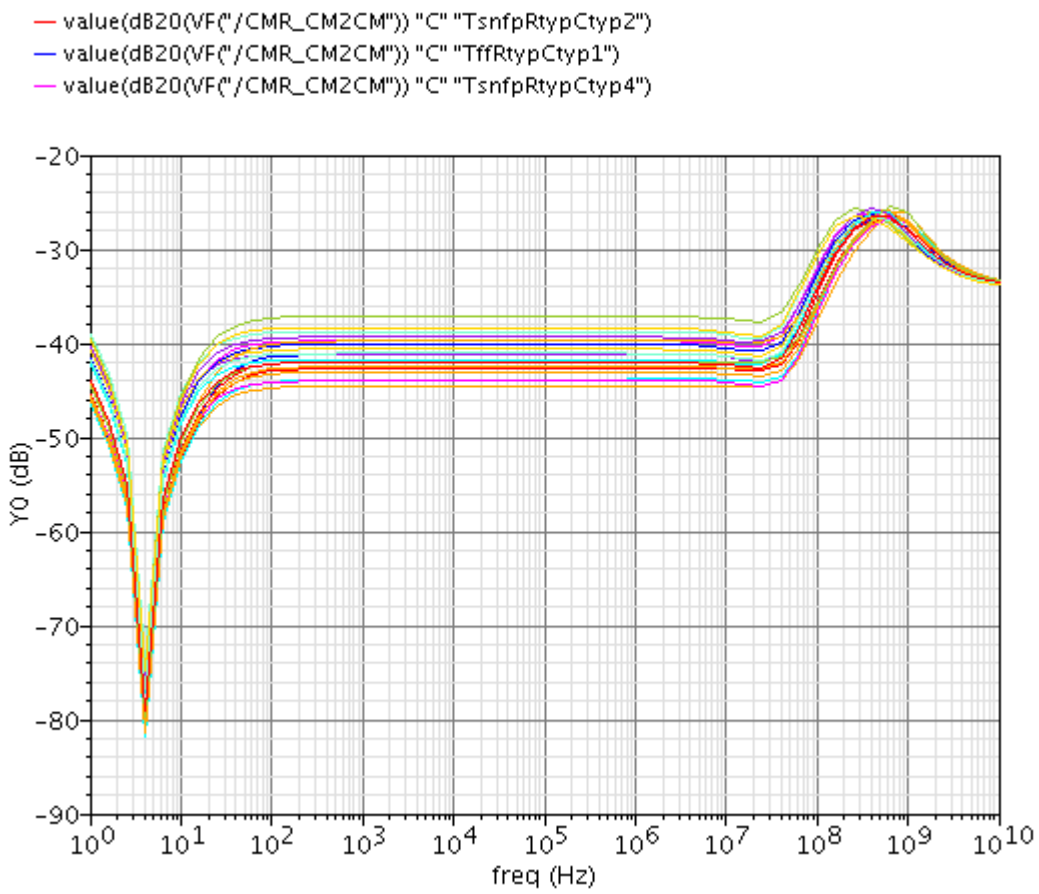

Fig 6.8: *stage 5 common mode rejection*

The rejection is good, which means that it is about -40dB up to tenth of MHz and then it increases up to -25dB. This is due to a combination of the limited band of the comparing block in the common mode loop, and to the parasitic capacitance in the tail transistors.

Another parameter to test is the PSR, extracted by applying a signal source at the voltage supplies, and sensing the output. The PSR is plotted in fig 6.9.

There is the same peak seen in the CMR plot, it depends on the fact that the supply is perceived as a common mode signal and so it is processed in the same way.

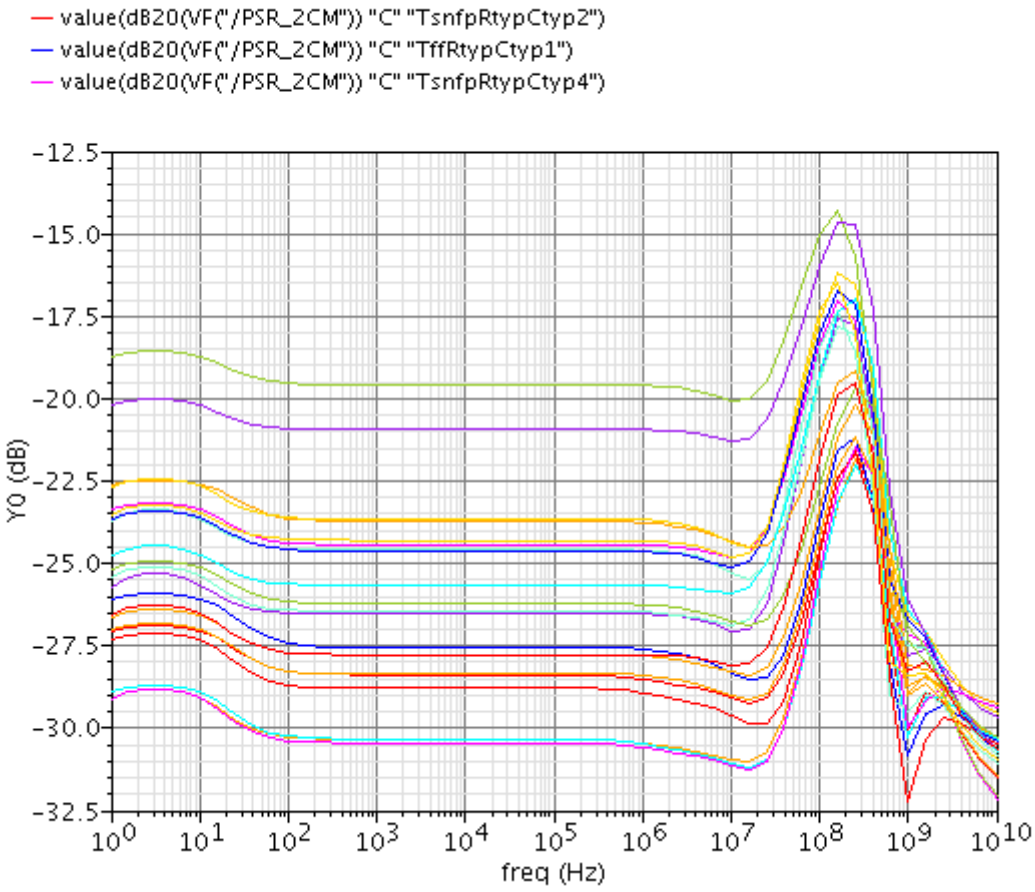

Fig 6.9: *stage 5 power supply rejection*

6.2.3 Max dynamic configuration

As previously described, the worst case for the INL extraction is located at the extreme points of the voltage gap. This means that the differential outputs of the amplifier will be completely unbalanced, reducing the value of R_{out} according to the previous considerations, and forcing the gain to drop.

So, to extract the effective INL behaviour, it is necessary to characterize the amplifier in this particular bias configuration. It can be easily achieved by the use of a DC feedback made by ideal devices. The role of this loop is to force the DC output to be a constant:

$$V_{+,DC} + V_{-,DC} = V_{max\_swing}$$

while the DC input is unbalanced by the loop itself, as can be seen in fig

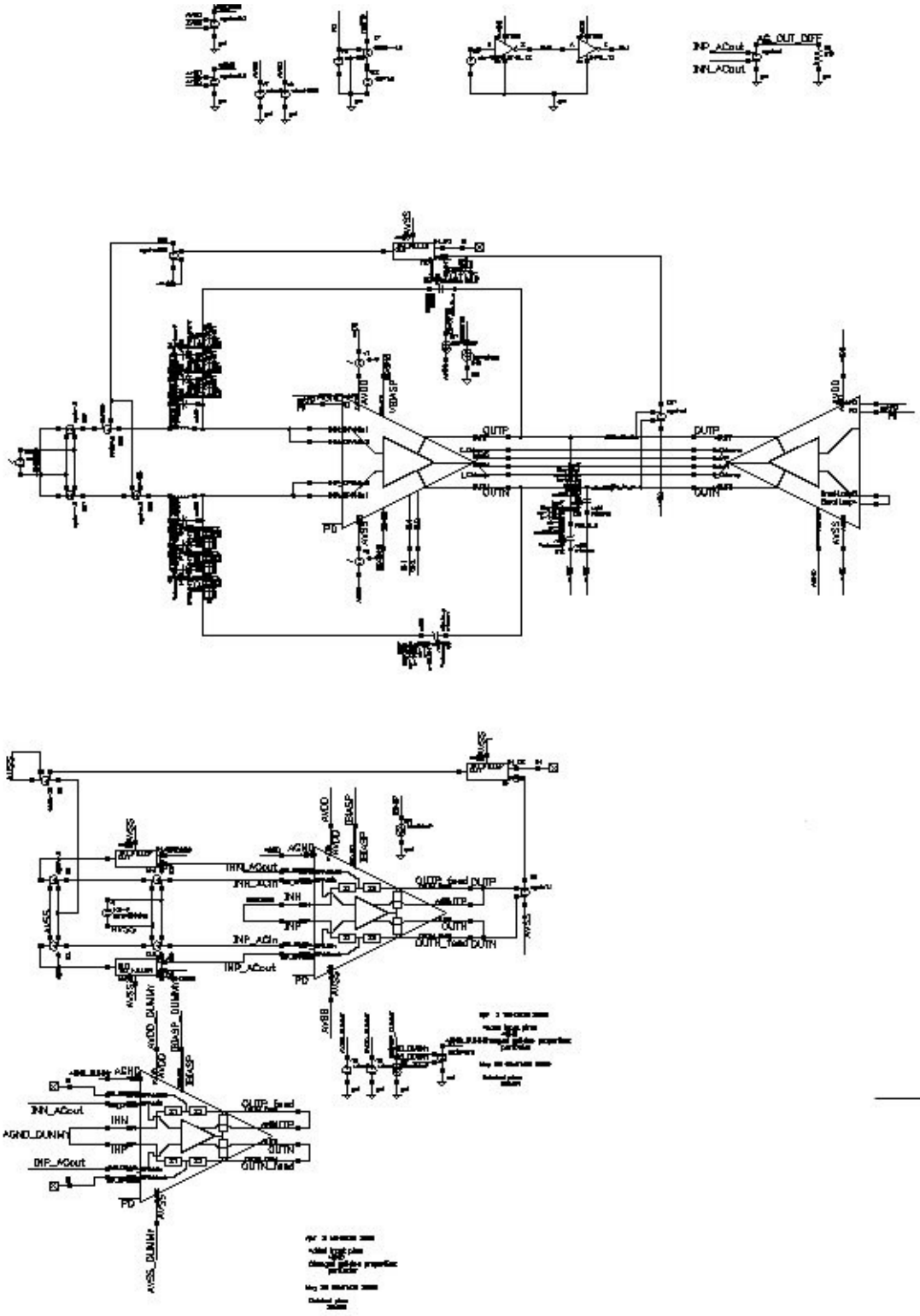

Fig 6.10: ***Max*** *dynamic configuration testbench*

The values extracted are still in spec

Corners	BW @maxSwing	PhaseMarg @maxSwing	Gain @maxSwing
Alltyp1	3,14E+008	60,96	54,69
Alltyp2	3,44E+008	61,63	54,66
Alltyp3	3,76E+008	60,39	56,17
Alltyp4	2,57E+008	63,07	50,83
Alltyp5	2,87E+008	61,28	53,45
TffRtypCtyp1	3,40E+008	62,03	53,71
TffRtypCtyp2	3,72E+008	60,88	55,1
TffRtypCtyp3	**2,52E+008**	63,61	**50,06**
TffRtypCtyp4	2,82E+008	62,04	52,23
TfnspRtypCtyp1	3,43E+008	60,63	54,75
TfnspRtypCtyp2	3,75E+008	59,48	56,2
TfnspRtypCtyp3	2,56E+008	62,23	51,02
TfnspRtypCtyp4	2,86E+008	60,52	53,57
TsnfpRtypCtyp1	3,45E+008	62,69	54,51
TsnfpRtypCtyp2	3,77E+008	61,33	56,11
TsnfpRtypCtyp3	2,56E+008	64,3	50,24
TsnfpRtypCtyp4	2,87E+008	62,33	53,05
TssRtypCtyp1	3,46E+008	61,65	55,27
TssRtypCtyp2	3,78E+008	**60,21**	57,04
TssRtypCtyp3	2,57E+008	63,62	50,5
TssRtypCtyp4	2,88E+008	61,41	53,84

Table 6.5: *stage 5 max_dyn behaviour*

There are approximatively no variations in bandwidth from the previous characteristics; instead there is a gain drop of about 4dB.

6.2.4 Noise

The noise figure and the relative integral can be extracted using the tools provided by the design tools: the tools calculate the noise generated by each device and represents it at the input and at the output of the amplifier designed.

The tools can estimate the noise power spectral density figure; here it is only reported the integral among the noise bandwidth.

Corner	Input_ref (nV/sqrt(Hz), 100-20G)	output_rms (mVrms)
Alltyp1	1,86E-008	1,12E-003
Alltyp2	2,07E-008	1,24E-003
Alltyp3	2,07E-008	1,24E-003
Alltyp4	1,79E-008	1,07E-003
Alltyp5	1,80E-008	1,08E-003
TffRtypCtyp1	2,38E-008	1,42E-003
TffRtypCtyp2	**2,39E-008**	1,43E-003
TffRtypCtyp3	1,93E-008	1,16E-003
TffRtypCtyp4	1,96E-008	1,17E-003
TfnspRtypCtyp1	1,93E-008	1,15E-003
TfnspRtypCtyp2	1,93E-008	1,16E-003
TfnspRtypCtyp3	1,68E-008	1,01E-003
TfnspRtypCtyp4	1,70E-008	1,02E-003
TsnfpRtypCtyp1	2,21E-008	1,33E-003
TsnfpRtypCtyp2	2,21E-008	1,33E-003
TsnfpRtypCtyp3	1,90E-008	1,14E-003
TsnfpRtypCtyp4	1,91E-008	1,14E-003
TssRtypCtyp1	1,82E-008	1,09E-003
TssRtypCtyp2	1,81E-008	1,09E-003
TssRtypCtyp3	1,67E-008	9,97E-004
TssRtypCtyp4	1,67E-008	1,00E-003

Table 6.6: *stage 5 noise*

6.2.5 Montecarlo

The Montecarlo tool simulates the deviations introduced into the silicon by the process variations. It allows to check the robustness of the device against mismatches.

In fig 6.11 the testbenches used for the simulations are shown. A feedback loop is added: its goal is to eliminate the differential offset before being fed into the input by the capacitive path, otherwise it would be amplified by the OTA forcing the output to clip.

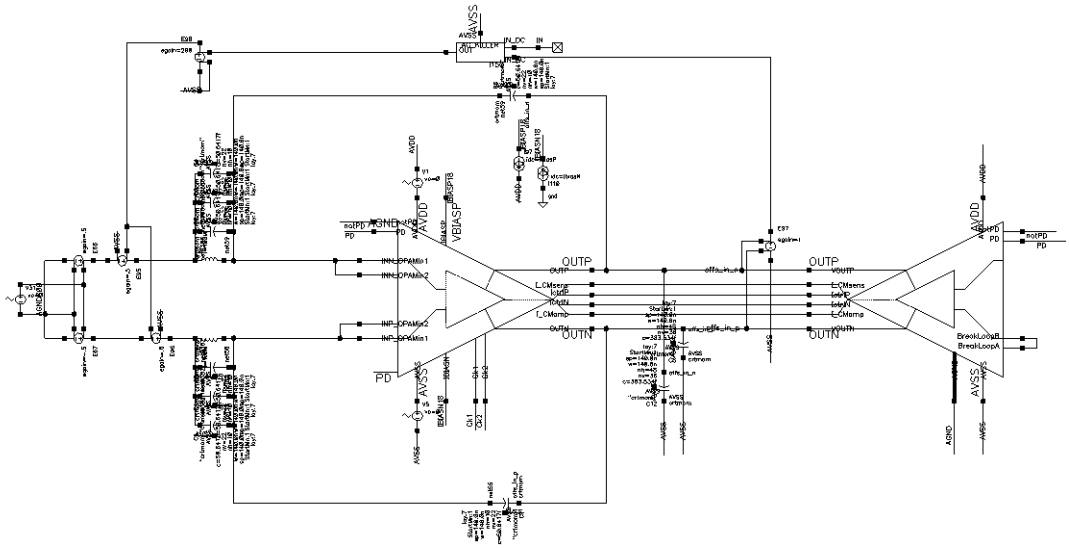

Fig 6.11: *Montecarlo testbench*

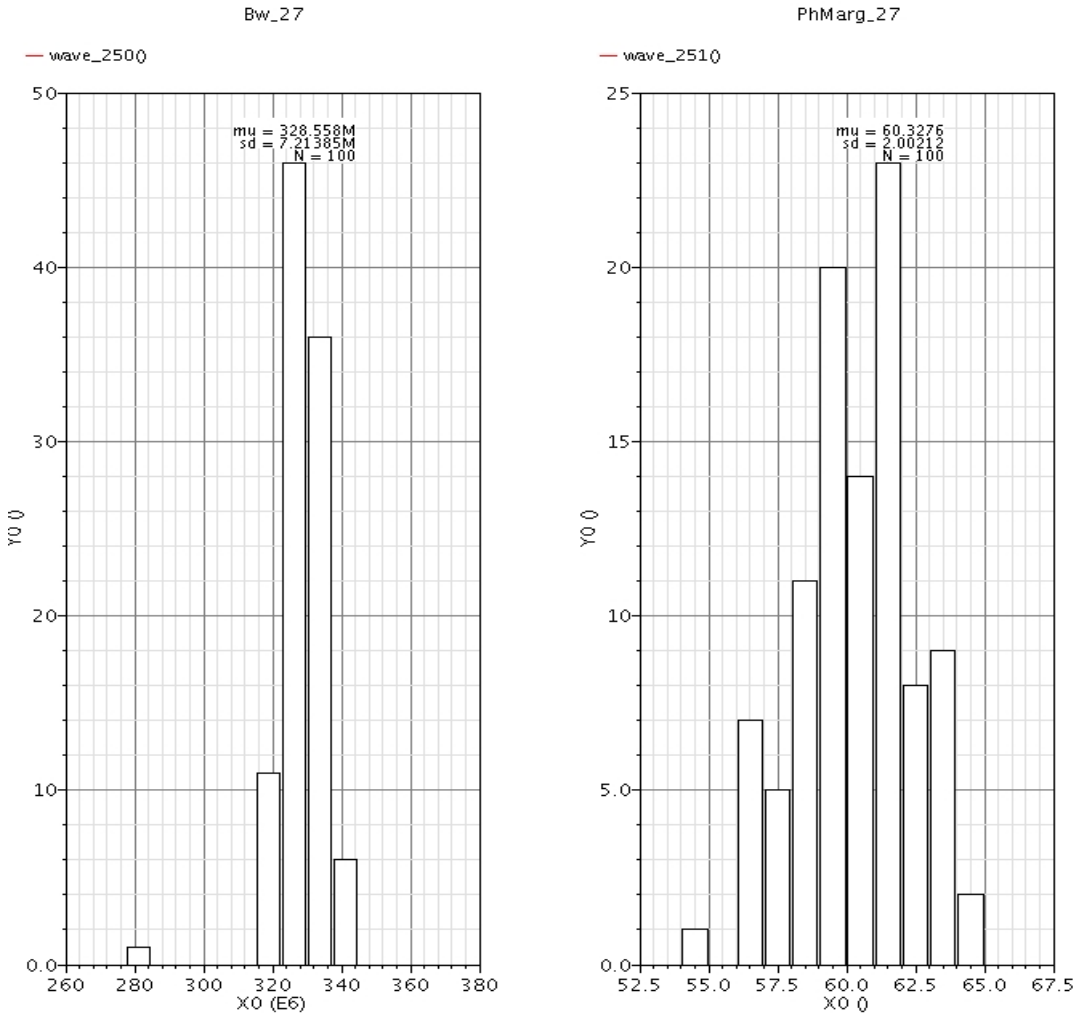

Fig 6.12: *bandwidth and phase margin variations in Montecarlo simulation*

Bandwidth and phase margin are enough robust against mismatch

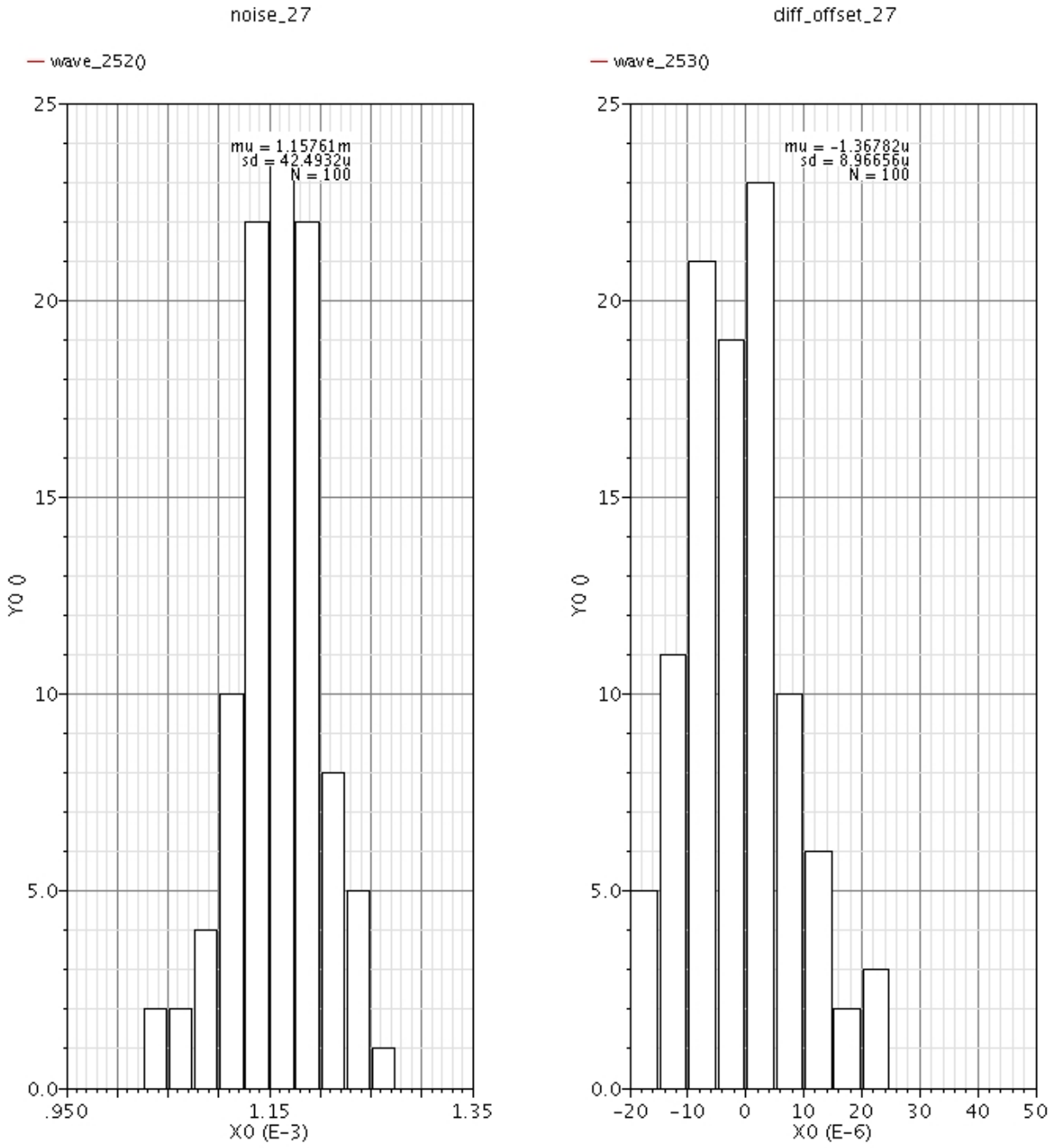

Fig 6.13: *noise and offset variations in Montecarlo simulation*

6.2.6 Start-up and switch down

The specifications require that the device can be turned on at -40° degree. The simulation is made by applying a 1us ramp at the power supply from 0V to Vdd.

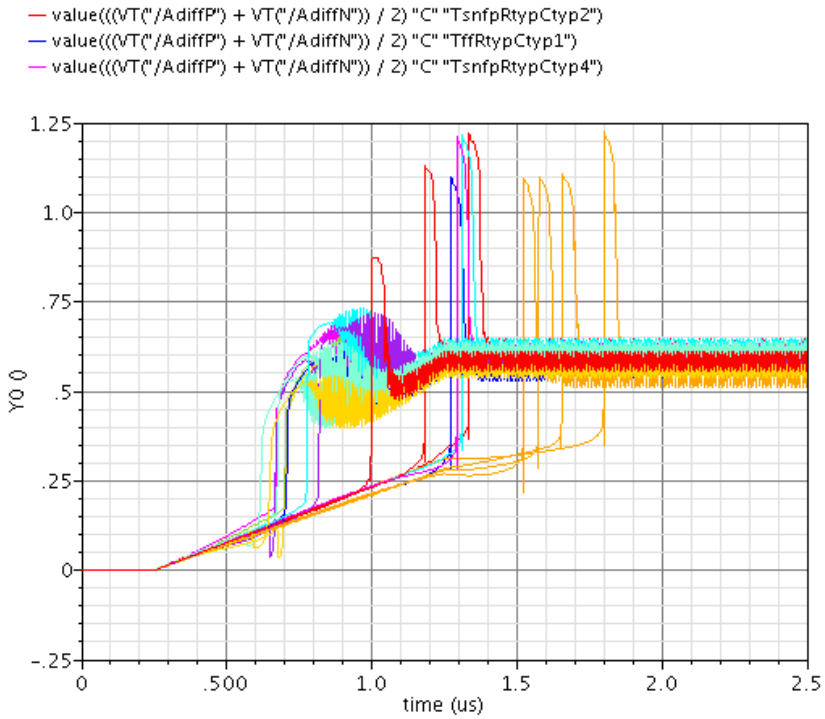

Fig 6.14: *ramp-up simulation*

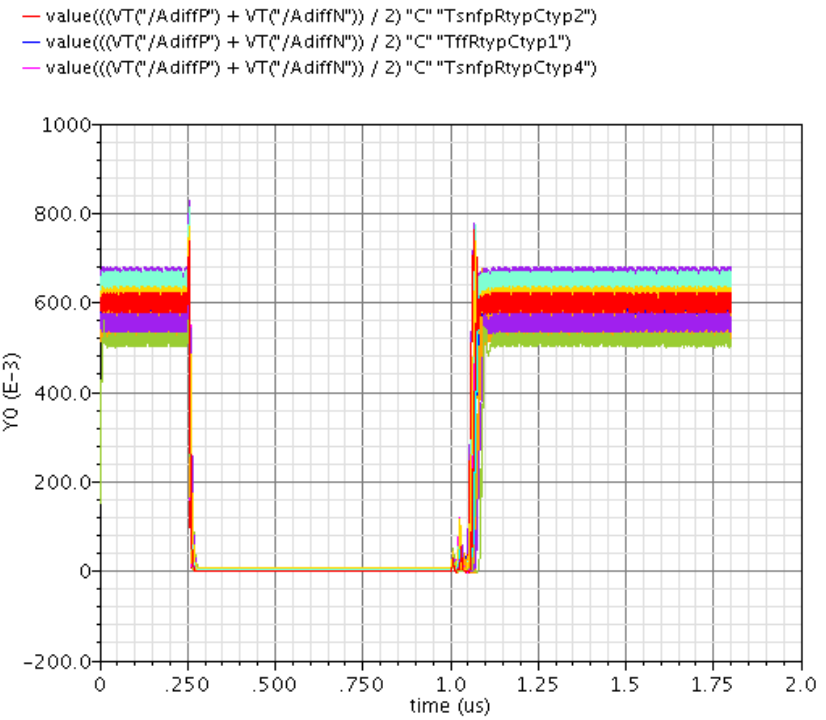

Fig 6.15: turn-off, turn-on simulation

6.2.6 INL simulation

To check the effectivity of the device designed, it is important to simulate if it works properly inside the environment it was made for.

The testbench realizes the complete MDAC as it will be printed on the silicon, only the subADC is substituted by a behavioural model to avoid huge simulation times (therefore the time machine needed is more than a day). The input of the system is fed by a ramp; the output obtained is then elaborated by a Matlab routine (see Appendix A).

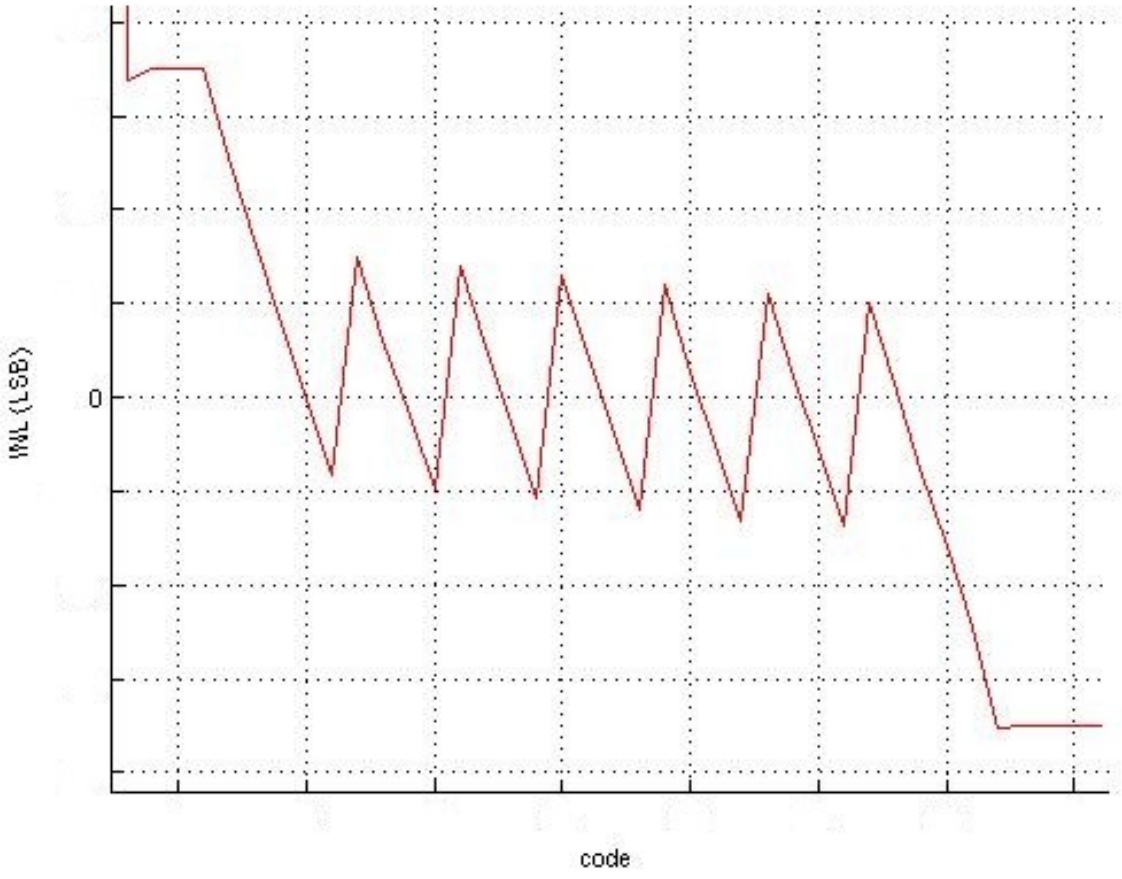

Fig 6.16: *INL simulation result*

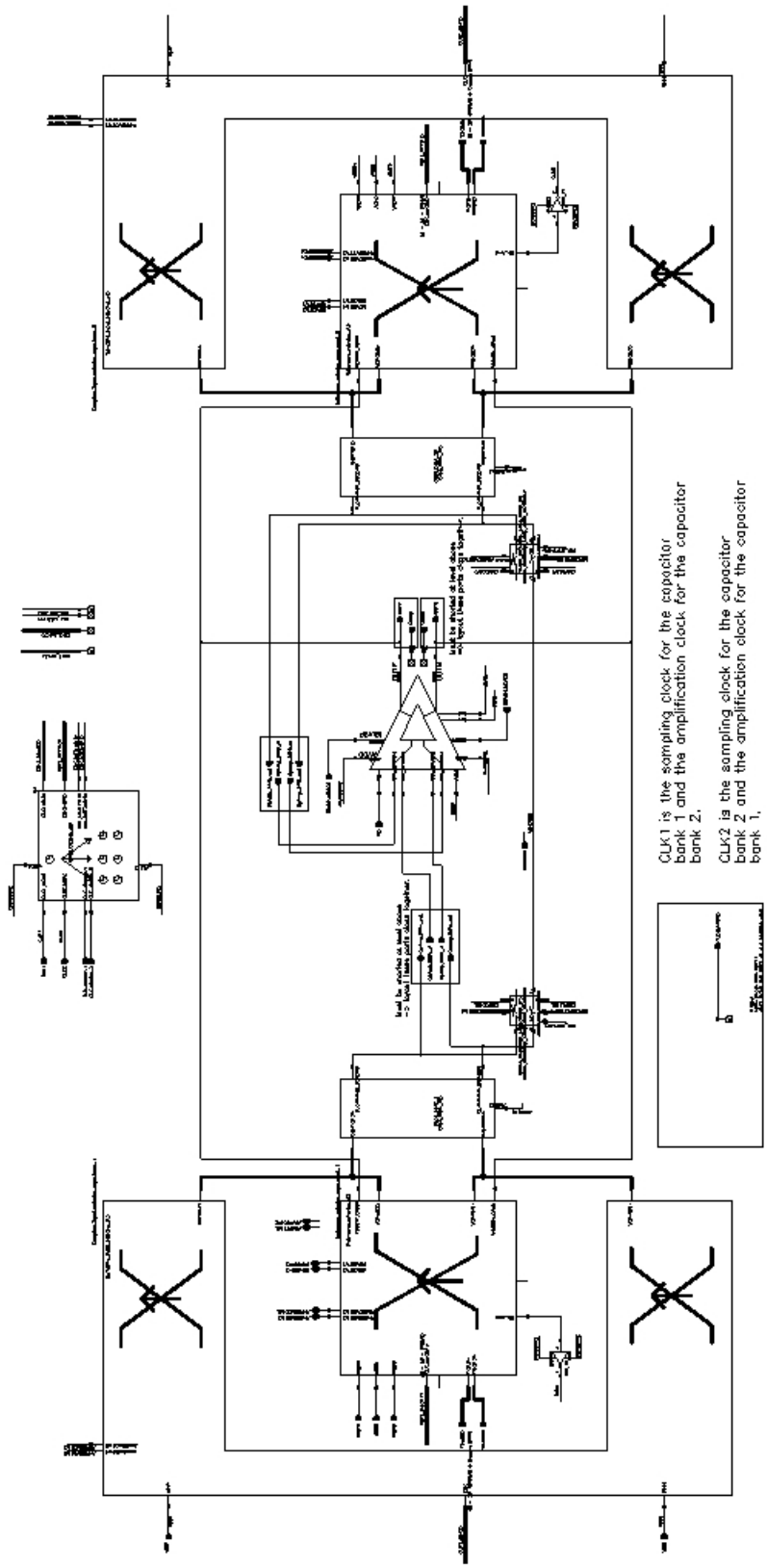

Fig 6.17: *INL testbench*

6.3.0 Stage 3 characterization

Only the results of the characterization set are reported, since the methodology is the same as in stage 5 amplifier.

6.3.1 AC behaviour

CORNER	bandwidth(Hz) 0dB	gain(dB) openloop @100Hz	phaseMargin(°)
Alltyp1	3,13E+008	89,49	61,67
Alltyp2	3,71E+008	90,96	61,43
Alltyp3	4,09E+008	92,54	63,03
Alltyp4	2,47E+008	85,23	61,28
Alltyp5	2,76E+008	86,67	61,9
TffRtypCtyp1	3,67E+008	87,51	62,92
TffRtypCtyp2	4,06E+008	89,17	64,66
TffRtypCtyp3	2,44E+008	80,58	63,6
TffRtypCtyp4	2,74E+008	82,09	64,17
TfnspRtypCtyp1	3,63E+008	86,28	61,46
TfnspRtypCtyp2	4,00E+008	88,69	63,26
TfnspRtypCtyp3	**2,41E+008**	**80,71**	61,72
TfnspRtypCtyp4	2,70E+008	83,17	62,41
TsnfpRtypCtyp1	3,74E+008	92,92	61,47
TsnfpRtypCtyp2	4,13E+008	94,19	63,07
TsnfpRtypCtyp3	2,50E+008	85,43	61,13
TsnfpRtypCtyp4	2,79E+008	86,34	61,77
TssRtypCtyp1	3,71E+008	94,3	60,25
TssRtypCtyp2	4,07E+008	95,75	61,96
TssRtypCtyp3	2,49E+008	89,43	**59,6**
TssRtypCtyp4	2,77E+008	90,81	60,31

Table 6.6: *stage 5 noise*

There are 12dB of variation in gain among corners, the double as in the previous case as expected, since the stages are two and the deviations are uncorrelated.

6.3.2 CM behaviour

As previously described there are two different common mode loops. They are simulated in two steps

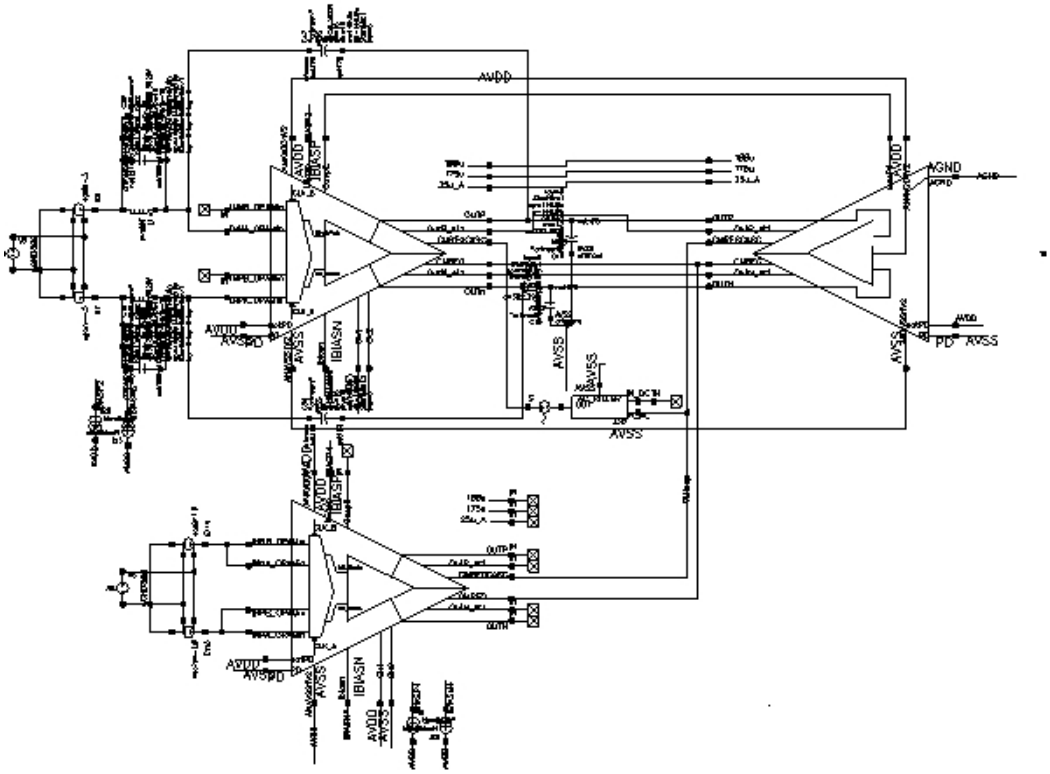

Fig 6.18: *first testbench for double CM loop simulation*

The first loop, the one regulating the common mode behaviour of the first stage has the following characteristics:

CM	Cm loop1_gain @10kHz (dB)	phaseMargin (°)	CM DC offset stage1
Alltyp1	45,95	69,25	2,45E-003
Alltyp2	47,99	67,23	3,11E-003
Alltyp3	48,78	67,63	1,33E-003
Alltyp4	40,3	71,34	3,86E-003
Alltyp5	40,64	71,71	1,35E-003
TffRtypCtyp1	44,54	75,06	7,98E-004
TffRtypCtyp2	45,2	76,08	-1,63E-003
TffRtypCtyp3	**34,39**	77,53	-3,27E-003
TffRtypCtyp4	34,65	78,55	-6,53E-003
TfnspRtypCtyp1	49,57	67,37	2,60E-003
TfnspRtypCtyp2	50,23	68,23	7,95E-004
TfnspRtypCtyp3	43,16	71,26	3,25E-003
TfnspRtypCtyp4	43,54	71,81	6,30E-004
TsnfpRtypCtyp1	45,7	69,24	3,71E-003
TsnfpRtypCtyp2	47,03	68,26	1,74E-003
TsnfpRtypCtyp3	35,74	73,16	4,23E-003
TsnfpRtypCtyp4	36,26	73,09	1,60E-003
TssRtypCtyp1	50,19	59,3	4,53E-003
TssRtypCtyp2	51,74	**57,61**	2,82E-003
TssRtypCtyp3	44,5	63,16	**6,96E-003**
TssRtypCtyp4	45,42	62,62	4,61E-003

Table 6.7: *stage 7 CM behaviour of the first loop*

The second one, which regulates the output of the whole amplifier:

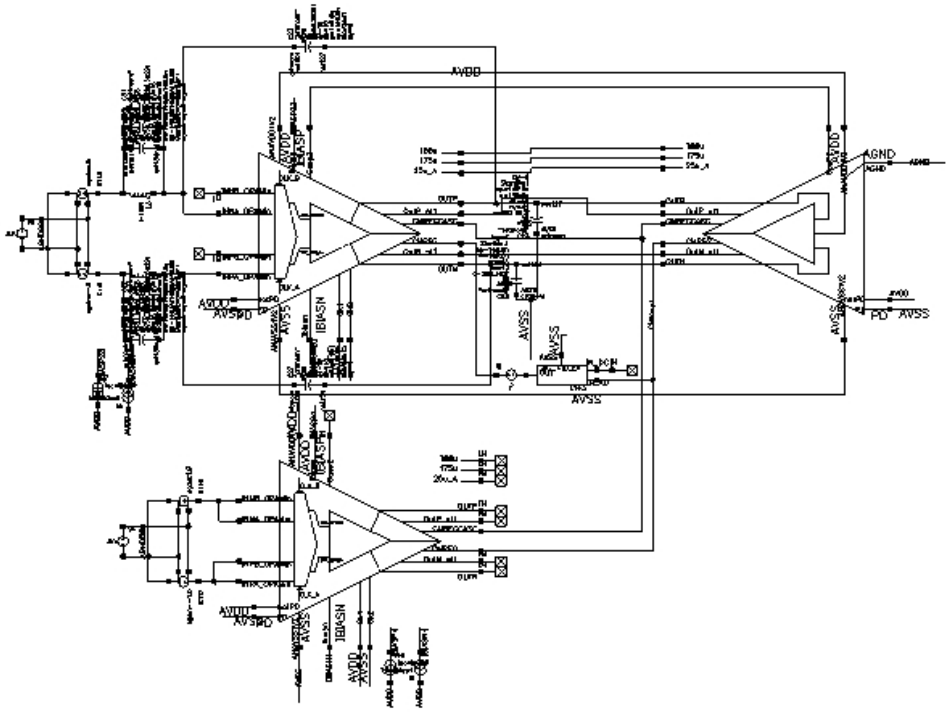

Fig 6.19: *second testbench for double CM loop simulation*

CM	Cm loop2_gain @10kHz (dB)	phaseMargin (°)
Alltyp1	46,15	70,89
Alltyp2	46,7	69,22
Alltyp3	47,87	71,15
Alltyp4	44,22	70,98
Alltyp5	45,25	73,29
TffRtypCtyp1	44,3	79,05
TffRtypCtyp2	45,39	82,07
TffRtypCtyp3	**41,92**	82,83
TffRtypCtyp4	42,74	85,47
TfnspRtypCtyp1	46,37	69,5
TfnspRtypCtyp2	47,44	72,11
TfnspRtypCtyp3	43,83	70,42
TfnspRtypCtyp4	44,79	73,56
TsnfpRtypCtyp1	46,53	70,02
TsnfpRtypCtyp2	48,08	70,64
TsnfpRtypCtyp3	44,41	71,79
TsnfpRtypCtyp4	45,56	73,27
TssRtypCtyp1	48,81	60,81
TssRtypCtyp2	50,18	61,5
TssRtypCtyp3	46,32	**60,59**
TssRtypCtyp4	47,47	62,09

Table 6.8: *stage 7 CM behaviour of the second loop*

The offset at the output:

CM	Vout_cm-V(AGND) (V)	CM Vout_DC	Effective_swing
Alltyp1	2,88E-003	0,6029	1,4140
Alltyp2	2,78E-003	0,5728	1,4664
Alltyp3	2,77E-003	0,6328	1,6688
Alltyp4	2,64E-003	0,5726	1,1468
Alltyp5	3,39E-003	0,6334	1,3436
TffRtypCtyp1	2,55E-003	0,5726	1,4800
TffRtypCtyp2	2,44E-003	0,6324	1,6864
TffRtypCtyp3	6,85E-003	0,5769	1,1676
TffRtypCtyp4	**7,63E-003**	0,6376	1,3608
TfnspRtypCtyp1	2,20E-003	0,5722	1,4656
TfnspRtypCtyp2	1,88E-003	0,6319	1,6724
TfnspRtypCtyp3	1,60E-003	0,5716	1,1556
TfnspRtypCtyp4	1,96E-003	0,6320	1,3508
TsnfpRtypCtyp1	3,24E-003	0,5732	1,4668
TsnfpRtypCtyp2	3,53E-003	0,6335	1,6700
TsnfpRtypCtyp3	3,89E-003	0,5739	1,1380
TsnfpRtypCtyp4	5,00E-003	0,6350	1,3356
TssRtypCtyp1	2,80E-003	0,5728	1,4488
TssRtypCtyp2	3,07E-003	0,6331	1,6516
TssRtypCtyp3	1,94E-003	0,5719	**1,1092**

Table 6.9: *stage 7 CM offset*

Since the booster attenuates the common mode amplification, the CMR is high, approximatively 75dB in the worst case:

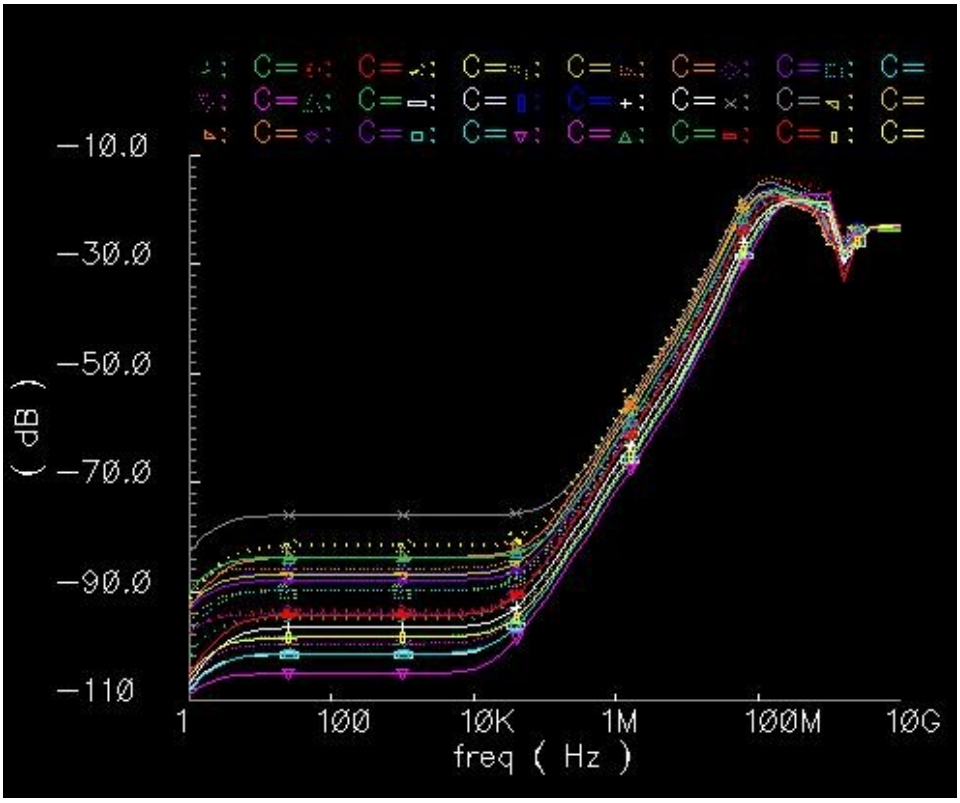

Fig 6.20: *stage 3 CMR*

Also the PSR is good, around 30dB

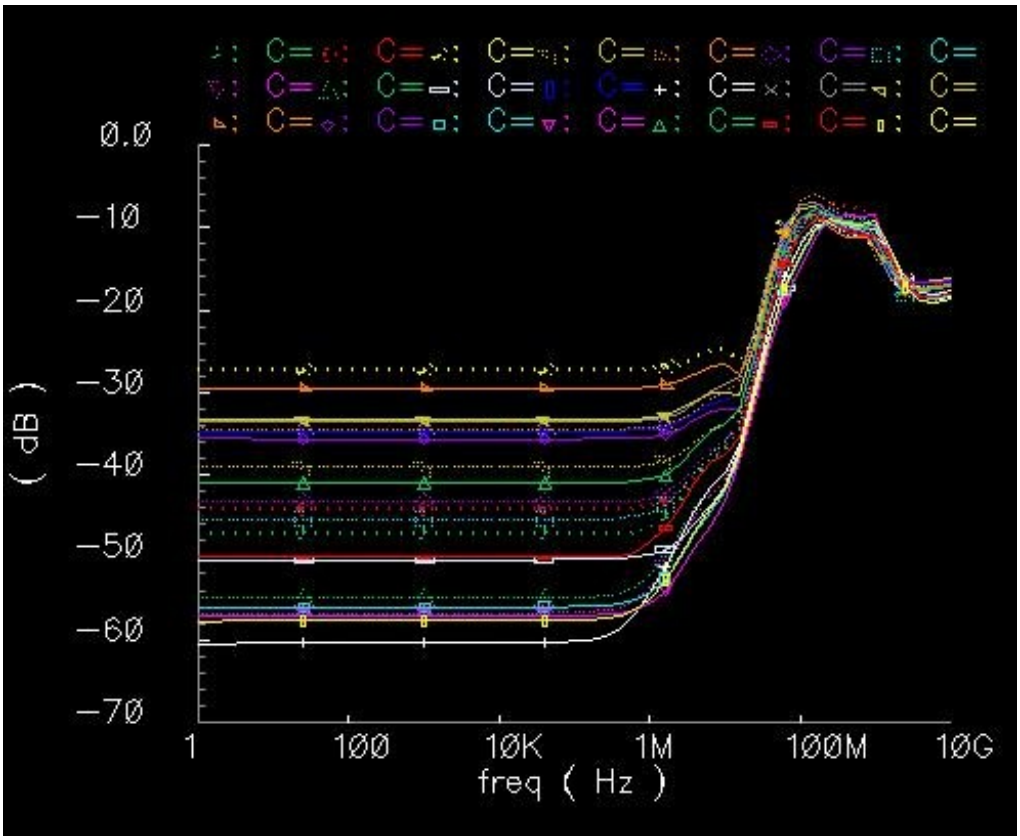

Fig 6.21: *stage 3 PSR*

6.3.3 Max dynamic configuration

Corners	BW @maxSwing	Gain @maxSwing	phaseMargin(°) @maxSwing
Alltyp1	3,12E+008	86,39	61,55
Alltyp2	3,71E+008	88,71	61,27
Alltyp3	4,09E+008	90,53	62,85
Alltyp4	2,44E+008	79,29	61,04
Alltyp5	2,73E+008	82,24	61,78
TffRtypCtyp1	3,67E+008	85,34	62,81
TffRtypCtyp2	4,05E+008	87,23	64,56
TffRtypCtyp3	2,41E+008	75,11	63,35
TffRtypCtyp4	2,71E+008	77,91	64,03
TfnspRtypCtyp1	3,63E+008	84,02	61,36
TfnspRtypCtyp2	4,00E+008	86,69	63,17
TfnspRtypCtyp3	**2,38E+008**	**74,83**	61,6
TfnspRtypCtyp4	2,67E+008	78,75	62,39
TsnfpRtypCtyp1	3,74E+008	90,69	61,27
TsnfpRtypCtyp2	4,13E+008	92,2	62,82
TsnfpRtypCtyp3	2,47E+008	79,39	60,77
TsnfpRtypCtyp4	2,76E+008	81,88	61,54
TssRtypCtyp1	3,71E+008	91,88	60,05
TssRtypCtyp2	4,08E+008	93,58	61,7
TssRtypCtyp3	2,46E+008	82,6	**59,29**

Table 6.10: *stage 3 behaviour at max dynamic*

6.3.4 Noise

The noise figure and the relative integral can be easily extracted using the tools provided by the design software:

Corners	Input_ref (nV/sqrt(Hz))	Output noise (Vrms, 100Hz-10GHz)
Alltyp1	6,05E-009	4,67E-004
Alltyp2	5,87E-009	4,53E-004
Alltyp3	5,92E-009	4,57E-004
Alltyp4	6,21E-009	4,79E-004
Alltyp5	6,25E-009	4,82E-004
TffRtypCtyp1	6,24E-009	4,82E-004
TffRtypCtyp2	6,30E-009	4,86E-004
TffRtypCtyp3	6,39E-009	4,93E-004
TffRtypCtyp4	**6,44E-009**	**4,97E-004**
TfnspRtypCtyp1	5,77E-009	4,45E-004
TfnspRtypCtyp2	5,82E-009	4,49E-004
TfnspRtypCtyp3	6,14E-009	4,74E-004
TfnspRtypCtyp4	6,18E-009	4,77E-004
TsnfpRtypCtyp1	5,96E-009	4,60E-004
TsnfpRtypCtyp2	6,02E-009	4,65E-004
TsnfpRtypCtyp3	6,26E-009	4,83E-004
TsnfpRtypCtyp4	6,30E-009	4,86E-004
TssRtypCtyp1	5,59E-009	4,31E-004
TssRtypCtyp2	5,64E-009	4,35E-004
TssRtypCtyp3	6,06E-009	4,68E-004
TssRtypCtyp4	6,09E-009	4,70E-004

Table 6.11: *stage 3 noise characterization*

6.3.5 Montecarlo

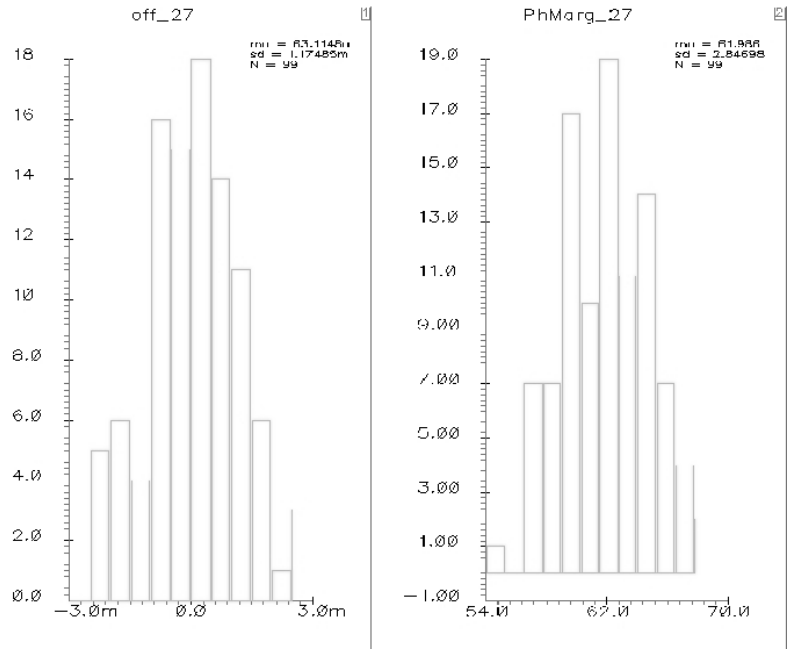

Fig 6.22: *phase margin and offset variations in Montecarlo simulation*

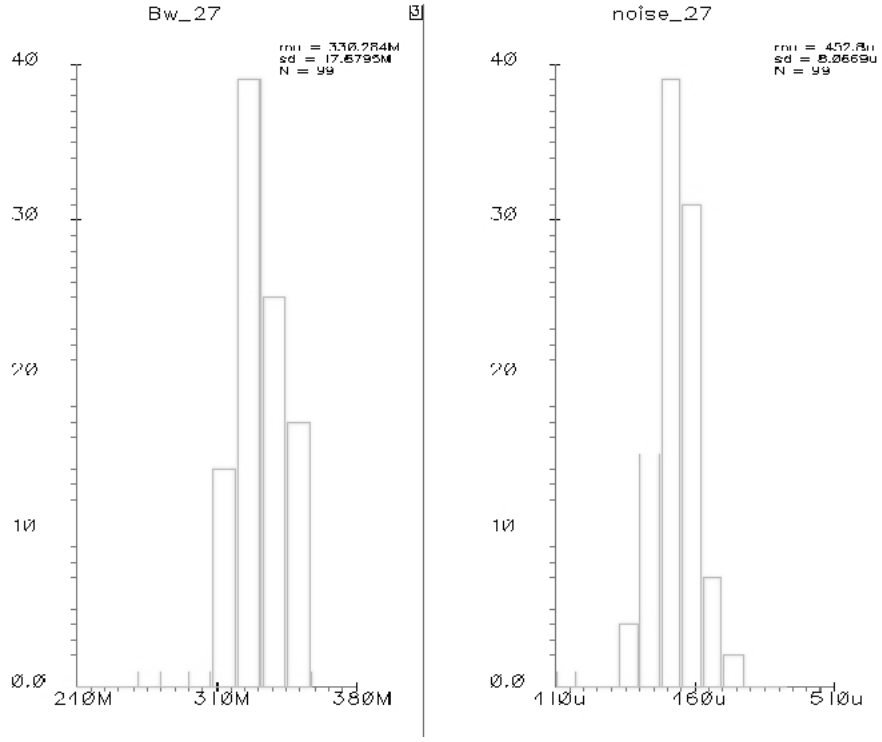

Fig 6.23: *bandwidth and noise variations in Montecarlo simulation*

Chapter 7

Layout

7.0 Introduction

Once the tests to check against process variation are performed, the second step is to make the design robust against manufacturing effects. This step relies mostly on the layout design.

Several aspects must be taken into account: in fact, with the scaling of the devices, the subwavelength gap widens, making it harder to print most structures; some structures are even harder to print, leading to lithographical distortions which in some cases result in yield loss as well as performance degradation; interconnect manufacturing issues represent the largest yield detractor in nano-CMOS processing. A design put together without design for manufacturability in mind can result in copper erosion and dishing, changing the designed characteristics affecting electromigration and timing. Certain wiring patterns can result in high yield loss due to shorts. Open via is another major yield detractor in copper technology. Interconnect density variation causes interlayer dielectric thickness variation, resulting yield loss due to underpolish metal shorts as well as unexpected timing due to variation of capacitive parasitics.

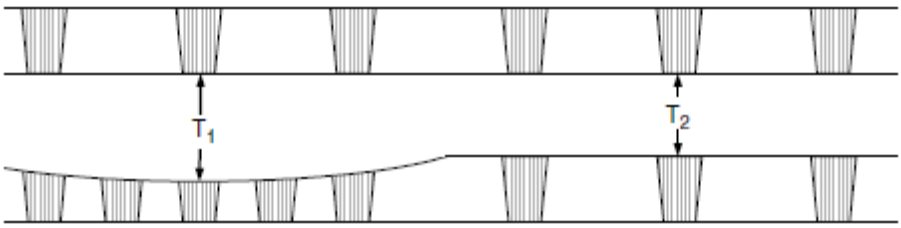

Fig 7.1: *wire density variation*

Antenna problems can lead to yield loss due to gate damage and in some cases, degrade transistor performance by inducing early negative bias temperature instability or Vth shifts. So the use of antenna diodes becomes mandatory, and this implies an additional parasitic capacitance as well as increasing the risk of latch-up.

7.1 Main sources of variations

The technology scaling enables exponential improvement of digital circuit performance and functions on a chip, but on the other hand, has made analog design more challenging on many fronts. Table 1 in appendix C summarizes the modelling challenges that can affect analog designs. Analog circuits require good device matching; listed below are the main sources of matching problems.

- Asymmetry (leads to misalignment sensitivity)
- Small geometries (narrow-width effects; short-channel effects; larger Vth variation)
- Proximity effects (well proximity; poly proximity; microloading etch effects)
- Position of well and ground taps (body effect differential)
- Horizontal and vertical effects
- Temperature differential
- STI stress effects
- Diffusion and poly flaring (strong design influence in the nano-CMOS regime)
- Mirror layout effects (capacitance; Rsd; misalignment)
- Random dopant fluctuation
- Dopant channelling through gate
- Poly-L variation; Leff variation

- Degradation due to antenna effect
- Hot-carrier injection
- Metal density variation (thickness variation; capacitance variation)

So there are several effects to take into account; also, since the number of devices is quite high, the more performing solution seems to use a “greedy” strategy, or better, to apply local optimizations and then, at the end, optimize the global layout.

7.2 Interdigit structure

Where the matching is critical, the use of interdigit structure is a good choice. It consists in alternating wherever possible same length, same width, fingered transistors.
In particular, the input pair in a differential architecture is extremely sensitive to any unbalanced deviation; in fact any mismatch which can modify the behaviour of a transistor in comparison to the other, would appear at the output amplified by the gain of the amplifier itself.

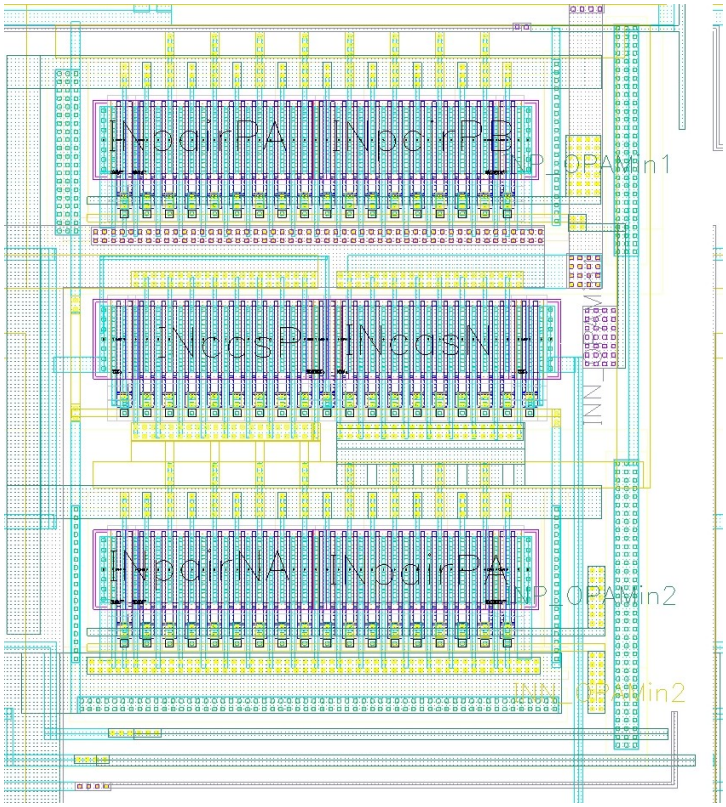

Fig 7.2: *Input pair layout*

As can be seen in fig 7.2, the design is oriented towards symmetry; the left part of a component must be as much similar as possible to the right part. The goal is to make the mechanical stress introduced by the STI oxide and by the metal routing the same on both the devices of the input pair.

In particular, since defects can exist in singular spot of the silicon crystal, the interdigiting strategy allows to minimize the variation, since the defect is better distributed among the two transistors. This technique is applied also on current mirrors, where matching is also fundamental.

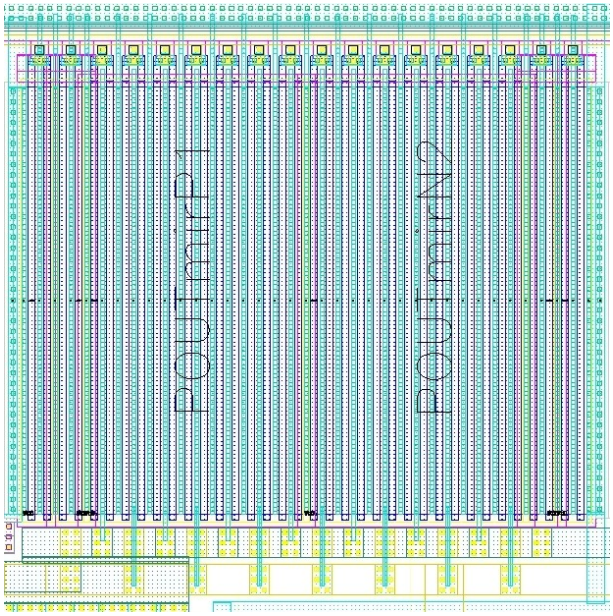

Fig 7.3: *particular of intedigited mirror layout*

7.3 Antenna effect and antenna diodes

The gate oxide underneath the poly is thin, between 1 and 2nm. If the charges accumulated on the poly is sufficiently large, the charges accumulated can damage the gate oxide. This is known as process antenna effect.

The maximum amount of charges that can be accumulated on the poly is proportional to the area of the poly, or better, the charges are accumulated on the perimeter side-wall area of the poly, which can be calculated as the perimeter of the poly multiplied by the thickness of the poly. Thus, an effective layout practice to prevent process antenna violation is to stay within the antenna ratio design rule of the given technology. In particular:

• Minimize the use of poly for routing

• Minimize the use of poly to connect the gates together

A different widely used solution is to place diodes to protect the poly from antenna ratio violation. Antenna diode is only effective in preventing antenna violation from metal routing, and does not help in antenna violation due to poly. The reason is simple. The diodes are made from diffusions, but the poly is deposited onto the wafer before the diffusions are implanted into the wafer.

An other advantage coming from the use of antenna diodes arises during the silicon fabrication process, instead. In fact, several steps in the foundry are related on plasma (for instance etching and deposition). Charge from the electrons and ions can be collected by conductive material on the wafer, and a net charge accumulation could lead to a change of the potential of the conducting material – until that potential itself is big enough to open up the "charge drainage" path to balance out the collection from plasma. If the drainage path is through gate oxide, charge can be trapped in the oxide, leading to many side effects, including shift of device threshold, creation of interface states which leads to earlier breakdown of the oxide, mobility degradation, worse sub-threshold slope, etc.

So the simplest solution is to provide an alternative path for the "drainage path", a so called antenna diode; in normal operation the diodes are reversely biased and since they have minimum size, they have only a minuscule impact on total capacitance.

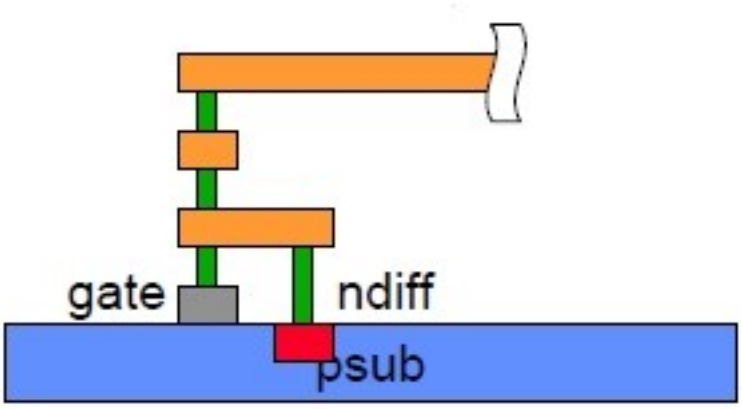

Fig 7.4: *cross section of antenna diode protection*

During processing, even being reversely biased, because of the elevated wafer temperature (200°C plus) and of the reduced breakdown voltage, they can eventually provide a discharge path.

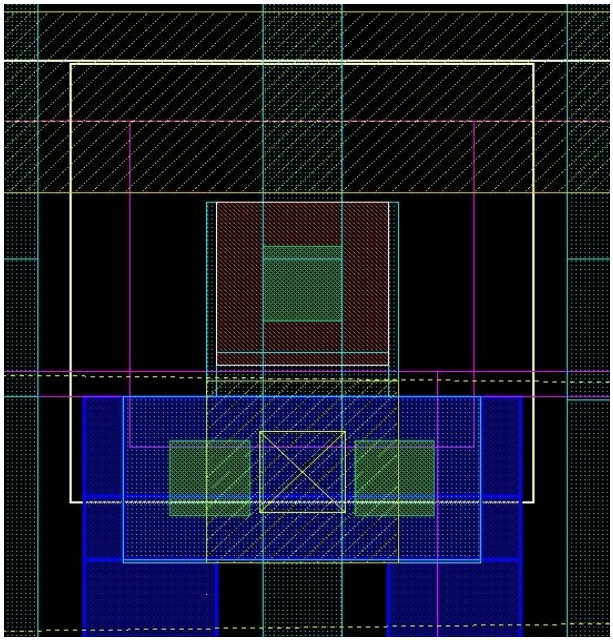

Fig 7.5: *antenna diode layout*

7.4 Dummy transistors

As long as the geometry is reduces, the device behaviour gets more and more sensitive on the mechanical stress. This is why the technological library is provided of different features to take into account second order effects during the design.

One of the more relevant in an analog environment is the LOD effect: it simulates the effect of the STI on the matching of the transistors. The STI (shallow trench isolation) introduces a silicon spot which is in a non-uniform state of superficial stress; this implies an impact on the device performance, adding a I_{dsat} and a V_{th} offset, which are a strong function of the layout.

In fig 7.6 an axial representation of the stress is depicted:

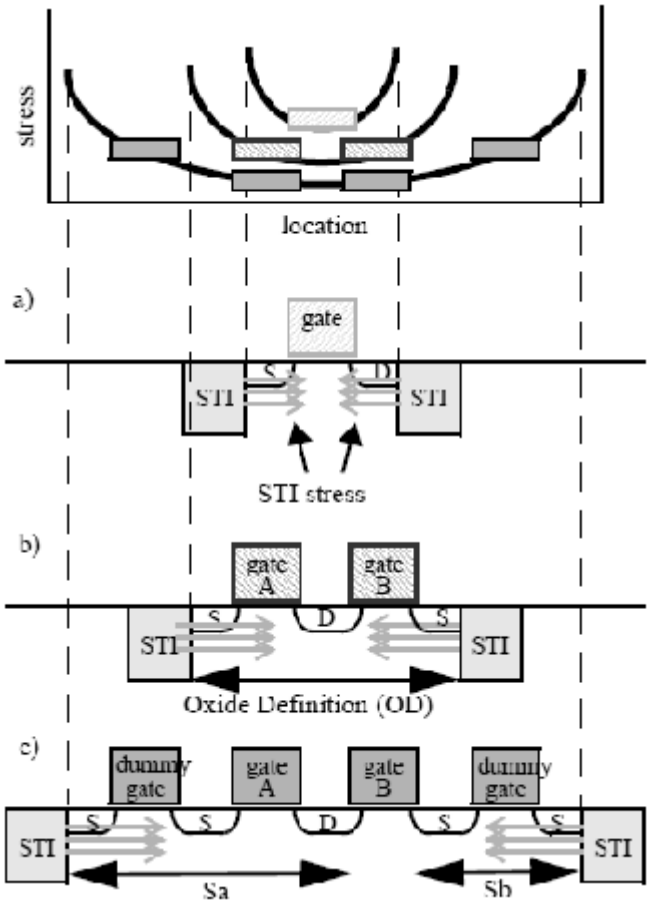

Fig 7.6: *STI stress*

This stress is function of the geometry, and can be qualitatively described by Sa and Sb, which are the distance from the gate to the edge of the OD (oxide definition) on both sides of the device:

$$Stress = \frac{1}{Sa + \frac{L}{2}} + \frac{1}{Sb + \frac{L}{2}} \tag{7.1}$$

This effect is very important when designing current mirrors, differential pairs or any other structure based on ratio and symmetry. The solution chosen relies on multifinger devices and on the use of dummy transistors; in fact, it can be shown by simulations that the use of a single dummy device appears to be very effective, but the use of two fingers dummy transistors even better since it offers marginal differential return from the values expected and the ones experienced. So the adopted methodology is based on the use of doubly fingered dummy blocks placed on both sides of every critical device.

An additional improvement consists in merging different devices together, using same-width fingers, and placing the most sensitive devices in the middle of the array, as shown in fig 7.7:

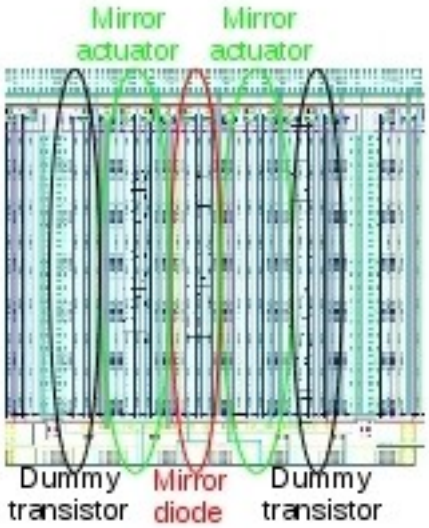

Fig 7.7: *array of fingered devices in a current mirror*

7.5 Electromigration

The electromigration consists in a mass transport of electrons in metals where these metals are stressed at high current densities. This may result in a change of the conductor dimension, eventually causing the creation of either voids or hillocks in the affected regions.

This process is intrinsically linked to time: in an industrial design, one of the specifications to take care of is the lifetime of the device.

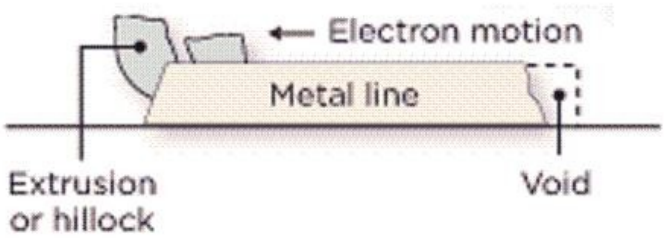

Fig 7.8: *effects of electromigration*

Physically, the copper or aluminium interconnects are polycrystalline; while conducting current, the electrons interact with the atoms in the lattice, forcing them to migrate in the direction of the main flow. At the end this process consists in a material transport, which mainly occurs at the metal-dielectric interface and at the boundaries of the grains.

After a sufficient amount of time, atoms are deposited, leading to the generation of hillhocks and the build-up of mechanical stress around the hillhock area. While these hillhocks can cause shorts with neighbouring interconnects, the build-up of mechanical stress can lead to cracks in the surrounding insulation layers. Subsequently, material migrations towards these cracks can generate the so called whiskers which may also introduce shorts to neighbouring wires. Also, voids can reduce the conductivity over time, which can lead increase the resistivity or to interconnect failures. It may be noted that these processes are self-accelerated effect cycle.

The foundry provides a documentation reporting study and experimental studies concerning these effects. It also provides the constants to calculate, through the “Black's Law”, the mean time to failure (MTTF):

$$MTTF = \frac{A}{j^n} \cdot \exp\left(\frac{E_a}{kT}\right) \tag{7.2}$$

where A is a cross section area dependent constant, j is the current density, E_a is the effective activation energy of the electromigration process and n is a scaling factor.

Calculating power rail for metal1 in tsmc90

TechConstant	Time [hours]	Temperat [°K]	AverageCurrent [mA]
2,89E-007	1,00E+005	3,83E+002	1,00E-001

Jmax for M1 [mA	Wmin[um]
2,00E+000	6,01E-002

	Irms[mA]X5°C	Irms[mA]X2°C	Irms[mA]X1°C
	1,40E+000	8,82E-001	6,24E-001
Wmin[um]	8,54E-002	6,93E-002	6,46E-002

Ipeak,max acce	Peak duration[s]	Period[s]
1,38E+001	2,10E-011	1,00E-009

Table 7.1: *routine to calculate the interconnections lifetime*

To optimize the human time dedicated a self calculating routine was created: it extracts

the minimum width of a connection to ensure a desired lifetime, express as the maximum augment of resistance acceptable.

7.6 General consideration for layouting

Here is reported a general vademecum to take care of during the layout process.

- Minimum-spaced and minimum-width wires must be avoided wherever possible to minimize erosion distortion of the signal lines, which increases resistivity and degrades timing that is not comprehended by the tools.
- Wide wires may require more space, since the walls of wide trenches have a tendency to collapse, causing shorts. The sidewall incline of wider wires is also greater and can result in shorts to neighbouring wires.
- Nwell proximity effects can cause as much as a 50-mV Vth shift for NMOS and a 20-mV Vth shift for PMOS. Attention must be paid to the placement of matched devices where the orientation and space to the well are identical.
- Limiting the degrees of freedom in a layout, such as by having all transistors oriented the same way, can dramatically improve process control and optimization.
- Design uniformity and the use of tiled devices guarantee identical devices, which helps in device matching.
- Constraining poly pitch and the use of dummy devices to guarantee the neighbourhood desired makes the lithographic processes easier and results in better poly-CD control.
- Symmetry in critical layout and the use of precision rules will help to ensure that the end caps have ample diffusion overlap.
- The use of multiple contacts and vias has a considerable impact on yield. It is better to use more structured design methodology where random layout patterns are not allowed.
- Precision or analog design rules should be used with analog cells.

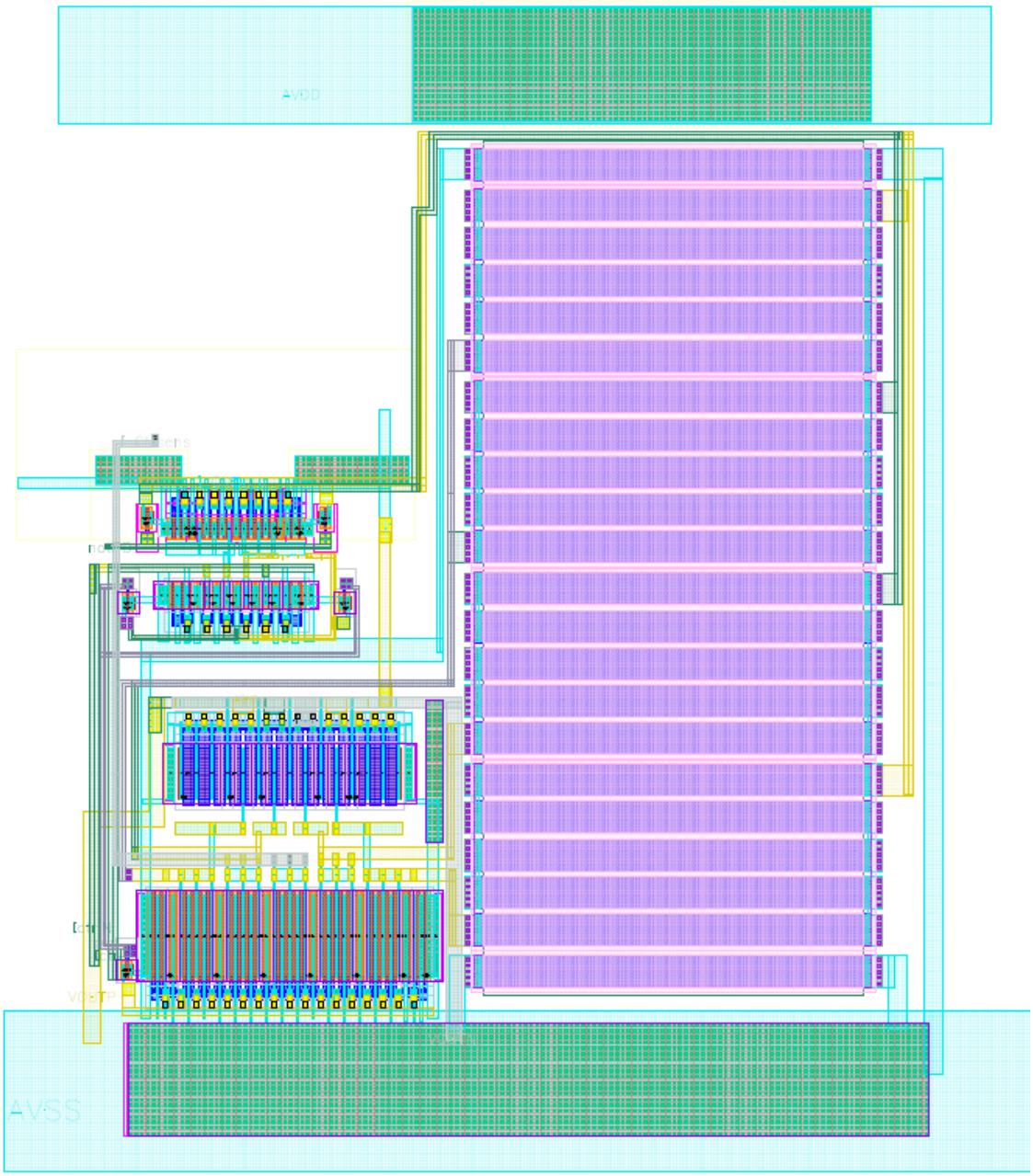

Fig 7.9: *common mode regulator layout*

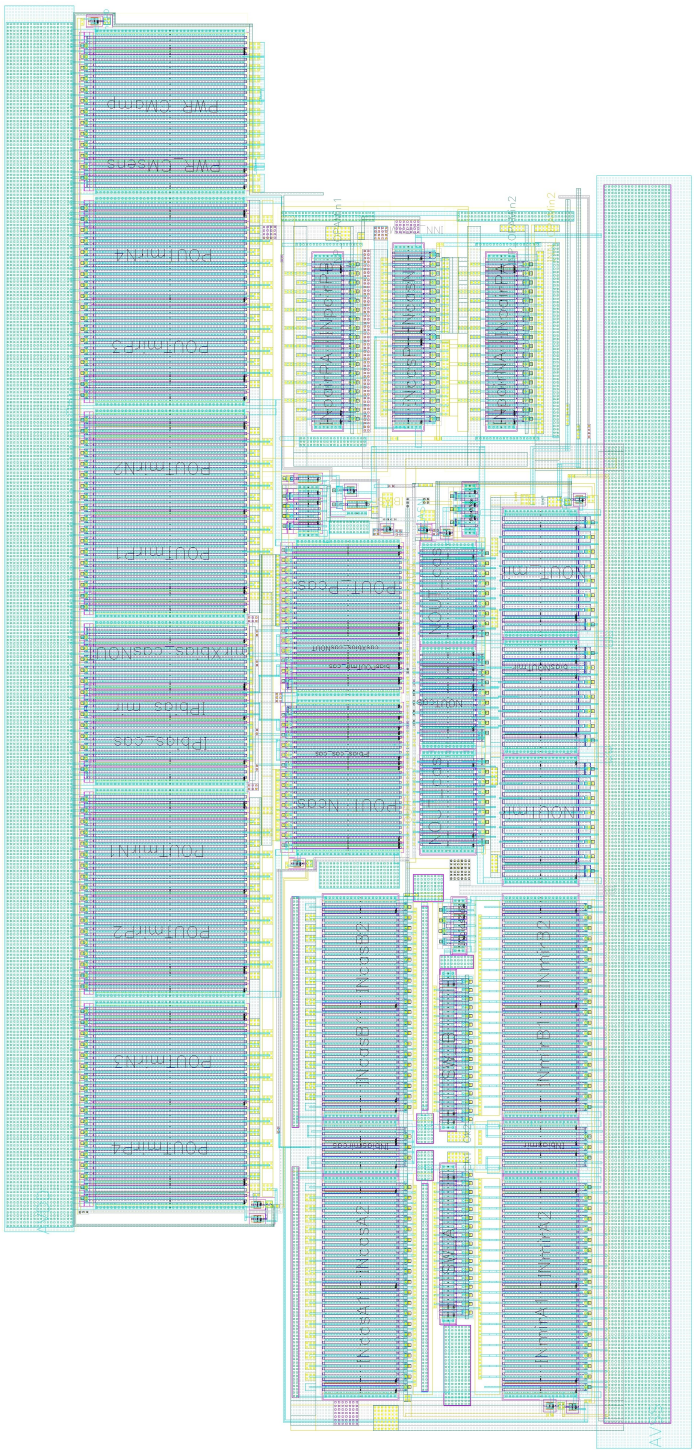

Fig 7.10: *differential amplifier layout*

Chapter 8

Conclusion

8.0 Introduction

All the specifications required were met.

Table 8.1 shows the power consumption among corners of both the amplifiers discussed, in particular the power used by the whole device and the one taken by the common mode regulator.

CORNER	st3Power_total (W)	st3Power_DC_CMdriv	st5Power_total (W)	st5Power_CMdi
Alltyp1	5,27E-003	8,75E-004	2,73E-003	1,24E-004
Alltyp2	4,69E-003	7,69E-004	2,43E-003	9,41E-005
Alltyp3	5,84E-003	9,70E-004	2,99E-003	1,27E-004
Alltyp4	4,73E-003	7,87E-004	2,48E-003	1,21E-004
Alltyp5	5,89E-003	9,95E-004	3,05E-003	1,59E-004
TffRtypCtyp1	4,82E-003	8,05E-004	2,46E-003	1,15E-004
TffRtypCtyp2	6,02E-003	1,02E-003	3,03E-003	1,55E-004
TffRtypCtyp3	4,96E-003	8,41E-004	2,52E-003	1,54E-004
TffRtypCtyp4	**6,20E-003**	**1,07E-003**	**3,10E-003**	**2,02E-004**
TfnspRtypCtyp1	4,68E-003	7,57E-004	2,44E-003	1,07E-004
TfnspRtypCtyp2	5,84E-003	9,64E-004	3,01E-003	1,45E-004
TfnspRtypCtyp3	4,72E-003	7,77E-004	2,49E-003	1,40E-004
TfnspRtypCtyp4	5,90E-003	9,95E-004	3,07E-003	1,84E-004
TsnfpRtypCtyp1	4,70E-003	7,78E-004	2,42E-003	8,40E-005
TsnfpRtypCtyp2	5,85E-003	9,76E-004	2,98E-003	1,13E-004
TsnfpRtypCtyp3	4,73E-003	7,92E-004	2,46E-003	1,07E-004
TsnfpRtypCtyp4	5,90E-003	9,94E-004	3,03E-003	1,41E-004
TssRtypCtyp1	4,61E-003	7,41E-004	2,40E-003	7,99E-005
TssRtypCtyp2	5,73E-003	9,32E-004	2,96E-003	1,07E-004
TssRtypCtyp3	4,61E-003	7,49E-004	2,45E-003	9,97E-005
TssRtypCtyp4	5,74E-003	9,45E-004	3,01E-003	1,31E-004

Table 8.1: *power consumption*

8.1 Future Work

Today, only MDAC from stage 3 up to stage 6 have been completely designed. Still remain stage 1 and 2 amplifiers, since the other components for the MDAC have been already designed.

In table 8.1 the whole specifications set is reported. It is clear that designing the amplifiers for the first stages will be challenging, since the gain required is nearly 100dB for a GBW of 2.5GHz, but probably the most stringent requirement is the 1.58 $nV/\sqrt{Hz}$ of input referred noise.

Previously it was described as the boosting techniques spoil the noise behaviour; moreover, if the booster are applied at the second stage of a two stages amplifier, it could be possible that, having enough gain in the first stage, the noise specification will be met.

		Input buffer	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7
Top Level	Supply voltage (V)	1,2							
	Dynamic range (Vpdiff)				0,55				
	Sampling rate (Mspl/s)	n.a.				200			
	num of bits	0	2	2	2	2	2	2	2
	Total num of bits	0	2,81	2,81	2,81	2,81	2,81	2,81	2,81
	max out error (mV)	0,061	0,244	0,977	3,906	15,625	62,5	250	1000
	Guard time (ps)	500							
	Max INL	0,5							
	Rel INL contribution (%)	n.a.	16,67	16,67	16,67	16,67	16,67	16,67	0
MDAC	Rel noise contr to total noise	2,84	87,34	6,55	1,75	1,09	0,22	0,22	0
	kT/C noise (mVrms)	n.a.	0,24	0,26	0,55	0,78	0,78	0,78	n.a.
	Ron noise (mVrms)	n.a.	0,08	0,09	0,1	0,16	0,15	0,1	n.a.
	Max op-amp out noise (mVrms)	0,02	0,24	0,55	0,26	2,26	4,16	16,85	n.a.
	Total stage out noise	0,02	0,4	0,44	0,92	2,89	5,16	20,65	n.a.
	Capacitor unit value (pF)	n.a.	0,523	0,438	0,143	0,072	0,072	0,072	n.a.
	Cap load seen by the output (pF)	4,66	4,85	1,48	0,58	0,52	0,53	0,24	n.a.
Op-Amp MDAC	Min op-amp DC gain (dB)	linearity	97,24	85,09	70,79	59,42	48,29	34,25	n.a.
	Min GBW (MHz)		2480	2160	1415	1290	1177	746	n.a.
	Iout capability (mA)		4,07	1,11	0,42	0,33	0,28	0,11	n.a.
	noise BW (MHz)	314	658	580	492	413	334	249	n.a.
	Input ref noise (nV/Hz^0.5)	1,03	1,58	1,86	5,53	22,7	41,19	227,25	n.a.
	Op-amp input capacitance (pF)	n.a.	0,25	0,15	0,055	0,55	0,1	0,04	n.a.

Otherwise, a different option is a three stages amplifier, since it was previously demonstrate that a 30dB and more of gain per stage is easily achievable; so at this point the problems seem to be related mainly on bandwidth.

The technology shows severe limitations for the output resistance achievable; at the same time there are limitations also on the maximum current density, which implies the need of parallelism for several applications, adding further capacitance which can spoil the intrinsic speed of short channel technologies.

APPENDIX A: Matlab routine

The routine used in Matlab to simulate the INL for each MDAC are reported.

MDAC.m

```
function [ve_out]=MDAC(ve_in,ve_Qstate,ve_C,ve_Vref,sc_A0,sc_Gain_CL_var_rel,
sc_Cp)
% model : MDAC without nonidealities, except the capacitor mismatch (if taking in
account in vector C)
% and the AGND error.

%ve_out : output signal of the MDAC
%ve_in : input signal of the MDAC
%ve_C : vector with all the capacitor values (the last capacitor is the feedback
capacitor)
%ve_Qstate : state coming out from the quantizer (values from 0 to Nstate-1).
%ve_Vref : reference voltage value (VrefP = +Vref; AGND = 0; VrefN = -Verf).
%sc_A0 : OPAM open loop gain
%sc_Gain_CL_var_rel : Close Loop Gain error (relative). Vout = Vout_ideal * (1 +
sc_Gain_CL_var_rel)
%sc_Cp : OPAM input parasitic capacitance. Not yet implemented.

if nargin < 3
   error('MADC : not enough input arguments')
elseif nargin == 3
   ve_Vref = 1;
   sc_A0 = 1e100;
   sc_Gain_CL_var_rel = 0;
   sc_Cp = 0;
elseif nargin == 4
   sc_A0 = 1e100;
   sc_Gain_CL_var_rel = 0;
   sc_Cp = 0;
elseif nargin == 5
   sc_Gain_CL_var_rel = 0;
   sc_Cp = 0;
elseif nargin == 6
   sc_Cp = 0;
end

if length(ve_Vref) == 1
   ve_Vref = ve_Vref * ones(1,length(ve_in));
elseif length(ve_Vref) ~= length(ve_in)
   error('MADC : length(ve_Vref) ~= length(ve_in)')
end
```

```
%transform the vector Qstate on a line vector if necessary
[l,c]=size(ve_Qstate);
if l>c
   ve_Qstate=ve_Qstate';
end
if min(l,c)~=1 %test if it is a vector
   error('MDAC : input state is not a vector')
end

%transform the vector C on a line vector if necessary
[l,c]=size(ve_C);
if l>c
   ve_C=ve_C';
end
if min(l,c)~=1 %test if it is a vector
   error('MDAC : input C is not a vector')
end

%transform the vector ve_in on a line vector if necessary
[l,c]=size(ve_in);
if l>c
   ve_in=ve_in';
end
if min(l,c)~=1 %test if it is a vector
   error('MDAC : input ve_in is not a vector')
end


% Switch matrix construction : the matrix SW give the states of the switches connected
to the capacitors
% in function of the quantizer output state (vector Qstate). The line i correspond to the
quantizer state i-1.
% The first column give the states of the switches connected to the capacitor 1... A value
of -1 means that
% it is the switche connected to -Vref that is closed, 0 it is the switche connected to
AGND and 1 it is the
% switche connected to +Vref.
sc_Nstate = 2*size(ve_C,2)-1;
if log(sc_Nstate+1)/log(2) ~= round(log(sc_Nstate+1)/log(2)) %number of different
state possible. Only 3,7,15,31,63...
   % are possible. This is due to the error correction and to the fact that in digital all
number are in base 2
   % (gain must be a power of 2).
   warning('MDAC :  the number of states does not correspond to a normal value. The
vector C (capacitor) is not correctly sized.')
end

ma_SW = zeros(sc_Nstate,size(ve_C,2)-1); %initialisation of the Switch matrix. In this
```

```
matrix the Qstate k correspond
% to line k+1 (first line have the index 1 and not 0)
ma_SW(1,:)=-1;
for i=2:sc_Nstate
   ma_SW(i,1:(size(ve_C,2)-floor(i/2)-1)) = -1;
   ma_SW(i,(size(ve_C,2)-floor(i/2)):(size(ve_C,2)-1)) = 1;
   if round(i/2)==i/2 % i is even
      ma_SW(i,size(ve_C,2)-i/2)=0;
   end
end

ve_C_sampling=ve_C(1:size(ve_C,2)-1); %generate a capa vector without the
feedback capacitor
ve_Qstate=ve_Qstate+1; %add 1 to Qstate in order to have a correspondance between
the SW line and the state
ve_out=((sum(ve_C)*ve_in'-
(ma_SW(ve_Qstate',:)*ve_C_sampling'.*ve_Vref'))/ve_C(size(ve_C,2)))';
%this formula is a direct application of the charge conservation.

%Error due to finite OPAM gain :
ve_out=ve_out*sc_A0/(length(ve_C)+sc_A0);
ve_out=ve_out*(1 + sc_Gain_CL_var_rel);
```

ADC_Quan.m

```
function ve_out_decimal =
ADC_Quan(ve_in_quantizer,sc_Nstate,ve_Vref,ve_offset,sc_sigma_noise,sc_hysteresis)

%Model : ideal flash quantizer foreseen for an architecture with error correction. Non-
idealities added : comparator
%offset, comparator noise. Hysteresis not yet implemented.


%ve_in_quantizer : input signal. Must be greater than -ve_Vref and lower than ve_Vref
(hard cliping above these limits).
%sc_Nstate : number of different output state possible for the quantizer. The only value
possible are 3,7,15,31.... = 2^n - 1.
% This is due to the fact that the stage gain must be a power of 2 (on each stage we
resolve/assign a integer
% number of bits n => we must amplify the residue by 2^n) and this model is valid only
for a structure
% like the classical 1.5 bits (i.e. with 2 reference and 1 analog ground). cf Lewis in
JSSC March 92.
% The output states are 0,1 .... ,sc_Nstate-1
%ve_Vref : Reference value. Here we make the assumption (without losing any
generality) that the references
```

```
% are -ve_Vref, +ve_Vref and 0.
%ve_offset : input differential offset of the comparator (coming from mismatch). One
offset per comparator.
% The first comparator is this of the lowest comparison threshold. It is supposed that
there is Nstate - 1 comparator.
%sc_hysteresis : half value of the misdecision interval (the same for each comparator).
% Hysteresis not yet implemented.
%sc_sigma_noise : scalar given the


%Example 1.5 bits
%
%    ^ OUTPUT_MDAC [-ve_Vref,+ve_Vref]
%    |
%                  /
%                 /
%        /|   /|   /
%       / |  / |  /      -> INPUT [-ve_Vref,+ve_Vref]
%      /  | /  | /
%     /   |/   |/
%    /
%   /
%
% First segment code 00, second segment 01, third 10.
%
%First state is in = [-ve_Vref,-ve_Vref+3*ve_Vref/(sc_Nstate+1)[
%Second state is in = [-ve_Vref+3*ve_Vref/(sc_Nstate+1),-ve_Vref+5*ve_Vref/
(sc_Nstate+1)[
%Third state is in = [-ve_Vref+5*ve_Vref/(sc_Nstate+1),-ve_Vref+7*ve_Vref/
(sc_Nstate+1)[
%.....
%sc_Nstate state is in = [-ve_Vref+((sc_Nstate*2)-1)*ve_Vref/(sc_Nstate+1),+ve_Vref]

if nargin < 2
   error('ADC_Quan : not enough input arguments')
elseif nargin == 2
   ve_Vref = ones(1,length(ve_in_quantizer));
   sc_hysteresis = 0;
   ve_offset = zeros(1,sc_Nstate-1)
   sc_sigma_noise = 0;
elseif nargin == 3
   sc_hysteresis = 0;
   ve_offset = zeros(1,sc_Nstate-1)
   sc_sigma_noise = 0;
elseif nargin == 4
   sc_hysteresis = 0;
   sc_sigma_noise = 0;
elseif nargin == 5
```

```
   sc_hysteresis = 0;
end

if length(ve_Vref) == 1
   ve_Vref = ve_Vref * ones(1,length(ve_in_quantizer));
end

if log(sc_Nstate+1)/log(2) ~= round(log(sc_Nstate+1)/log(2)) %number of different
state possible. Only 3,7,15,31,63... are possible.
   %This is due to the error correction and to the fact that in digital all number are in
base 2 (gain must be a power of 2).
   warning('ADC_Quan :  the number of states does not correspond to a normal value.
The variable sc_Nstate is not correctly sized.')
end

%transform input to a line vector in the case where it is a column vector
[l,c]=size(ve_in_quantizer);
if l>c  % column vector
   ve_in_quantizer=ve_in_quantizer';
end
if min(l,c)~=1 %test if it is a vector
   error('ADC_Quan : input ve_in_quantizer is not a vector')
end

%transform offset vector to a line vector in the case where it is a column vector
[l,c]=size(ve_offset);
if l>c  % column vector
   ve_offset=ve_offset';
end
if size(ve_offset,2) ~= sc_Nstate - 1
   error('ADC_Quan : offset vector have not the correct number of element')
end

% ma_threshold = -ve_Vref' * ones(1,sc_Nstate-1) + ve_Vref' * [3/(sc_Nstate+1):2/
(sc_Nstate+1):(2*sc_Nstate-1) ...
%    /(sc_Nstate+1)];
ma_threshold = -ve_Vref' * ones(1,sc_Nstate-1) + ve_Vref' * [3/(sc_Nstate+1):2/
(sc_Nstate+1):2*sc_Nstate ...
   /(sc_Nstate+1)];
ma_threshold = ma_threshold + ones(length(ve_Vref),1) * ve_offset; %add constant
offset (mismatch + systematic) to
% the ideal treshold.
ma_threshold = ma_threshold + randn(length(ve_Vref),sc_Nstate-1)*sc_sigma_noise;
%add a noise to the threshold
% (represent the comparator input noise)

ve_out_decimal = sum(floor((sign((ve_in_quantizer' * ones(1,sc_Nstate-1)) -
ma_threshold) + 1)/2),2)';
```

INL.m

```
function ve_INL=INL(se_filename_out,se_filename_in, sc_resolution, sc_QuantState,
Vref, i)

%se_filename_out : is simulation result filename. The results must be in 2 column.
%   The first column is the time and the second the MDAC output
%se_filename_in : is simulation result filename. The results must be in 2 column.
%   The first column is the time and the second the MDAC input
%The input and the output are supposed include between -Vref and +Vref.

%sc_resolution : is the resolution in bit needed at the output of the MDAC.
%sc_QuantState : is the number of state of the Quantizer. Must 2^n-1 where n is an
integer.
%Vref : is the reference value (VrefP = +Vref and VrefN = -Vref)

% single data ----------------------------------------------------
 [ve_time,ve_MDAC]=textread(se_filename_out,'%f, %f','headerlines',1);
 [ve_time1,ve_in1]=textread(se_filename_in,'%f, %f','headerlines',1);
 ve_in=ve_in1(1:length(ve_MDAC));


%Corner routine-------------------------------------------------
%M = csvread(se_filename_out);
%N = csvread(se_filename_in);


code=[0:1:256]*4;
code(257)=1023;
%ve_in=(code/2^10-0.5)*2;

%plot(ve_in)
ve_quant = ADC_Quan(ve_in,sc_QuantState,Vref);
%plot(ve_quant)
ve_capa = ones(1,(sc_QuantState+1)/2);
ve_outideal = MDAC(ve_in,ve_quant,ve_capa,Vref)

figure(1)
%clf;
hold on;
plot(ve_outideal,'r');
grid on;
plot(ve_MDAC,'g');

%transform the vector ve_in on a line vector if necessary
```

```
[l,c]=size(ve_MDAC);
if l>c
   ve_MDAC=ve_MDAC';
end

ve_INL=(ve_MDAC(1:length(ve_outideal))-ve_outideal)*2^(sc_resolution-1)/Vref;
figure(2);
%clf
hold on;
plot(ve_INL(3:length(ve_INL)),'r');
grid on;
xlabel('code');
ylabel('INL (LSB)');

end
```

APPENDIX B: Native MOSFET

A transistor that has not undergone the channel doping process is termed a native transistor, and has a lower threshold voltage because it must rely on the intrinsic background or body of the transistor to set the threshold voltage. The typical native transistor threshold voltage can range from 0.1V to 0.3V

Low-Vth natural threshold voltage transistors can be fabricated as an option in a dual poly gate CMOS process. A process designed for 1.2V operation may be further simplified by eliminating the steps required for hot-carrier reduction i.e., LDD (Lightly doped Drain) implant and formation.

Figure B.1 shows the measured I_D and g_m versus gate voltage for a 20/0.4um nMOSFET biased at V_{DS} = 0.1V.

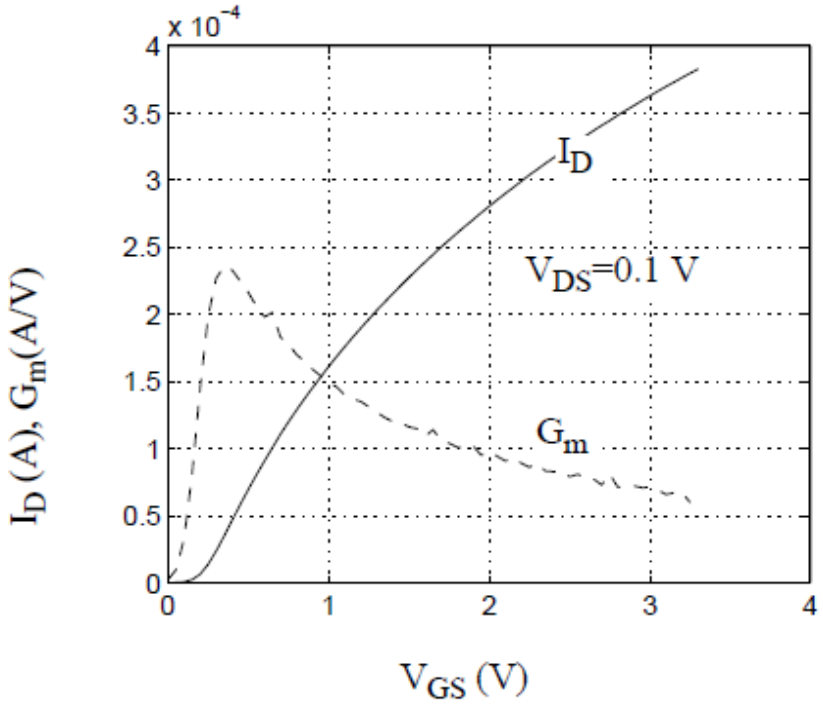

Figure B.1: *Measured I_D(A) and g_m(A/V) versus V_{GS}(V) for "natural" nMOSFET*

In Fig B.2 are illustrated the measured IDS-VDS characteristics for a natural threshold nMOSFET.

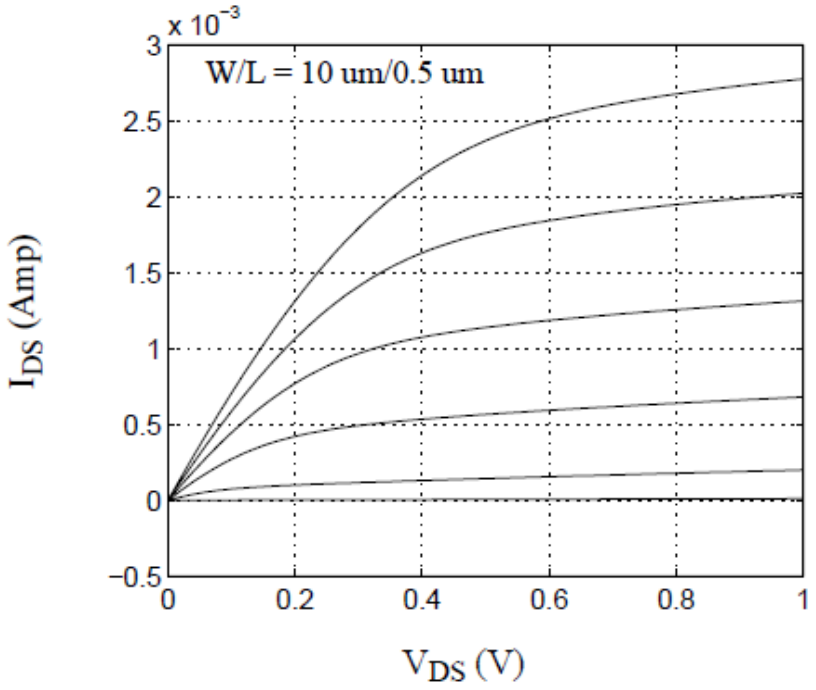

Figure B.2: *Measured nMOSFET I_{DS}-V_{DS} characteristics. $I_{DS} \sim (V_{GS}-V_{th})^{1.27}$*

APPENDIX C: Summary of modeling issues

Here is reported a summary of the main sources of effects coming from the miniaturization of the devices.

Parameter	Reason for effect	Synopsis of effect
RSC	Halo implants (technology, physical device effect)	Reverse short channel effect due to lateral nonuniform doping; when channel length varies, Vth varies
DITS	Halo implants (technology, physical device effect)	Drain induced threshold-voltage shift, due to change in DIBL for long channel length devices when the halo implant's influence on the channel diminishes
Early voltage and output resistance	Halo implants (technology, physical device effect)	Change in DIBL for long channel device similar to above
Poly depletion	Ultrathin gate oxide (technology, physical device effect)	Poly depletion is getting significant for ultrathin gate oxide, which accounts for about 8nm increase in equivalent oxide thickness for most devices, less for predoped poly
Gate tunnel current	Ultrathin gate oxide (technology, physical device effect)	Direct tunneling from gate to channel occurs due to ultrathingate oxide
Mobility-dopant dependence	Halo implants (technology, physical device effect)	Mobility improves with reduction in dopants
Linear proximity effects	Dense, isolated	Partly due to lithographic effects and partly to etch microloading effects, also due to dopant scattering from the poly, causing systematic dopant variation as a function of poly-line space of the design
Nonlinear proximity effects	Optical proximity correction	Subwavelength lithography requires resolution extension
GIDL	Band-to-band tunneling	High field in the drain to gate causes band-to-band tunneling, due to high junction doping and abrupt junction
Diffusion and poly flaring	Technology and layout effects	Subwavelength lithography causes flaring of diffusion and poly, causing device variations of small geometry devices and proximity of poly contact pads to diffusion edge
Well proximity	Devices at the edge of the well	Lateral scattering of well implant atoms out of the resist, which leads to threshold voltage increase for device close to the well edge
STI stress	Proximity effects to STI to device channel	STI stress reduces electron mobility but increases hole mobility, thus effecting Idsat

Bibliografia

1. A. Cester, “Nanoelettronica”, appunti dalle lezioni, 2008

2. A. Gerosa, *“Circuiti Integrati per l'eleborazione dei Segnali”,* appunti dalle lezioni, 2008

3. A. Bevilacqua, “*Progettazione di circuiti integrati analogici”*, appunti dalle lezioni, 2007

4. H. W. L. R. Jacob Baker and D. E. Boyce, “*CMOS Circuit Design, Layout and Simulation*” NY: IEEE Press on Microelectronic Systems, 1998.

5. D. A. Jones and K. Martin, “*Analog Integrated Circuit Design*” NY: John Wiley & Sons, Inc., 1997

6. J. Vital, *Pipeline Analogue-to-Digital Converters*, lectures at ChipIdea Conference

7. P.R. Gray and R.G. Meyer, "*MOS Operational Amplifier Design - A Tutorial Overview*", IEEE Journal of Solid State Circuits, Vol. SC-17, No. 6, pp. 969-982, DIC 82

8. B. Razavi, “*Design of Analog CMOS Integrated circuits*”, McGrawHill, 2001

9. Gray, Hurst, Meyer, Lewis, “*Analysis and design of analog integrated circuits*”, Wiley, 4th ed.

10. Wong, Mittal, Cao, Starr, “*Nano-CMOS Circuits and physical design*”, Wiley-Interscience, 2005

11. L. Wu, "*Low-Voltage Pipeline A/D Converter*", Thesis, Oregon State University

12. L. Sumanen, "*Pipeline Analog-to-Digital Converters for Wide-Band Wireless Communications*", Thesis, Helsinki University of Technology, 2002

13. V. Abramzon, "*Analog to digital converters for high speed links*", PhD Thesis, Stanford University, 2008

14. K. Sockalingam and R.Thibodeau, "*10-Bit 5MHz Pipeline A/D Converter*", Thesis, ECE Umaine, 2002

15. T. Khanna, "*Equation-based Hierarchical Optimization for a Pipelina ADC*", Thesis, Cornell University, 2005

16. Andrew M. Abo and Paul R. Gray, "*A 1.5V, 10-bit, 14 MS/s CMOS pipeline analog-to-digital converter*" IEEE J. Solid-State Circuits, vol. 34, no. 5, pp. 599-605, May 1999

17. *G. Moore, "No exponential is forever, keynote"* IEEE International Solid-State Circuits Conference, 2003

18. Cline and Paul R. Gray, *"A power optimized 13-b 5Msamples/s pipelined analog-to digital converter in 1:2m CMOS"* IEEE J. Solid-State Circuits, vol. 31, no. 3, pp. 294-303, March 1996

19. F. D'Agostino, D. Quercia, "*Short-Channel Effects in MOSFETs"*, project, 2000

20. A. Lakhlef, A. Benfdila, H. Achour and K. Bennamane, "*Characterization of the Short Channel Effect on the Threshold Voltage in Deep Sub-micron MOSFETs*", African Physical Review, pp 18-21, 2008

21. K. Bult, G. Geelen, *"A Fast Settling CMOS Op-Amp for SC Circuits with 90dB DC Gain*", IEEE J. Solid-State Circuits, vol. 25, no. 6, pp. 1379-1384, December 1990