uc3m Universidad Carlos III de Madrid

Bachelor's Degree in Biomedical Engineering 2017-2018

Bachelor Thesis

"DESIGN AND CHARACTERIZATION OF LOW VOLTAGE OPERATIONAL AMPLIFIERS FOR SMART SENSORS USING LOW COST CMOS TECHNOLOGY"

Alexander Martínez Pasek

Tutor: Enrique Prefasi Sen

Leganés, July 2018



This work is licensed under Creative Commons Attribution – Non Commercial – Non Derivatives

TRABAJO DE FIN DE GRADO – BACHELOR THESIS

DESIGN AND CHARACTERIZATION OF LOW VOLTAGE OPERATIONAL AMPLIFIERS FOR SMART SENSORS USING LOW COST CMOS TECHNOLOGY

Autor: Alexander Martínez Pasek

Tutor: Enrique Prefasi Sen

Tribunal Calificador

Presidente: D. Luis Hernández Corporales

Secretario: Dña. Claudia de Molina Gómez

Vocal: Dña. Lorena Cusso Mula

Tras haberse realizado el acto de defensa y la lectura del Trabajo de Fin de Grado el día 5 de Julio de 2018 en la Universidad Carlos III de Madrid, Escuela Politécnica Superior, Leganés, el Tribunal acuerda otorgarle la CALIFICACIÓN de:

VOCAL

SECRETARIO

PRESIDENTE

ABSTRACT

This bachelor thesis brackets the use of different OTA topologies and compares them under the scope of their application as low power comparators and adders for a $\Sigma\Delta$ ADC. This was undertaken under the "*Design and characterization of main building blocks for Medical instrumentation ADCs*" research project and, more specifically, in the "*Design of a Low-IF Sigma-Delta Modulator*" section.

The researched topologies include a folded cascode, telescopic cascode, class A Miller as well as a class AB Miller. The implementation was performed at transistor level of the for all topologies in a 0.18 μ m with original 1.8 V, downscaled to 1.5 V with the goal of reducing power consumption.

Keywords: CMOS integrated circuits, Differential amplifiers, Circuit simulation, Biomedical electronics, Analog integrated circuits.

ACKNOWLEDGEMENTS

It doesn't seem it has been four years already since I started this bachelor degree. Back at the beginning, during the first day of the degree, I did not realize what a huge journey all this would turn out to be. I have met my best friends (Roberto, thank you for all of your support during these years, my fellow engineer!), colleagues and soul mate, countless with whom I have crossed paths and have allowed me to become what I am today. To all of them I want to thank, since this seems the most appropriate place and date.

A mis padres, Juan Carlos y Joanna, así como a mi hermana Selene. Muchísimas gracias por todo el apoyo, desde el fondo de mi corazón. I jak mógł bym zapomnieć o mojej rodzinie w Polsce, dziękuję wam wszystkim za całą troske o mnie.

No puedo olvidar a quienes me inspiraron en seguir el camino de la ciencia allá en los primeros años de instituto. Vivian, Emilia, Juan Carlos, muchísimas gracias por animarme a seguir este camino tan apasionante.

А так же, моей девушке и родственной душе, спасибо тебе за всю твою поддержку и за прекрасное время проведённое вместе. Ты всегда заставляешь меня улыбаться и помогаешь становиться лучше. Я люблю тебя.

Finally, I want to thank all those in the Electronics Department, specially Enrique, for giving me this chance to learn about the thrilling subject that is electronic technology. I cannot forget about Marcos, thank you for inspiring me, Лети высоко мой друг, and Rubén, thank you for withstanding the constant questions I was subjecting you to.

INDEX

1 INTRODUCTION	1
1.1 MOTIVATIONS	1
1.2 BRIEF OVERVIEW OF OPERATIONAL AMPLIFIERS	1
1.2.1 SOURCE CLASSIFICATION	2
1.2.2 I/O CLASSIFICATION	2
1.3 OBJECTIVES	3
1.4 STATE OF THE ART	4
2 CMOS TECHNOLOGY	6
2.1 INTRODUCTION TO CMOS TECHNOLOGY	6
2.1.1 SILICON AND DOPING	6
2.1.2 THE PN JUNCTION	7
2.1.3 THE TRANSISTOR	8
2.1.4 THE MOSFET	9
Triode region	
Saturation region	11
Channel length modulation	
2.1.5 CMOS TECHNOLOGY	14
2.2 SMALL SIGNAL MODEL PARAMETERS	14
2.2.1 TRANSCONDUCTANCE	15
2.2.2 DRAIN-SOURCE RESISTANCE	15
2.2.3 PARASITIC CAPACITANCES	15
2.2.4 SMALL SIGNAL MODEL	
2.3 DESCRIPTION OF THE USED CMOS TECHNOLOGY	
2.3.1 SHORT CHANNEL EFFECTS	
Carrier velocity saturation	17
Mobility degradation	
Drain induced barrier lowering (DIBL)	
Impact ionization and Hot carrier injection	
2.3.2 ANALOG TECHNOLOGY	
2.3.3 DESIGN TIME AND COST	
2.3.4 TRANSISTOR CUT OFF FREQUENCY	
2.3.5 THRESHOLD AND SATURATION VOLTAGE IMPLICATIONS	
2.4 MOTIVATIONS BEHIND THE TECHNOLOGY CHOICE	

3 OTA TOPOLOGIES	23
3.1 INTRODUCTORY CONCENPTS	23
3.1.1 CURRENT MIRRORS	23
3.1.2 CASCODING THE CURRENT MIRROR	24
3.1.3 MILLER'S THEOREM	25
3.2 SYSTEM STABILITY AND COMPENSATION	26
3.2.1 STABILITY	26
3.2.2 FREQUENCY COMPENSATION	30
3.3 SINGLE STAGE TOPOLOGIES	31
3.3.1 FOLDED CASCODE OTA	31
3.3.2 TELESCOPIC CASCODE OTA	34
3.4 TWO STAGE TOPOLOGIES	36
3.4.1 MILLER-COMPENSATED OTA	36
3.4.2 CLASS A AND A/B	40
Class A	40
Class AB	41
Class A vs AB	42
The class AB buffer and its biasing	43
3.5 SINGLE ENDED AND DIFFERENTIAL TOPOLOGIES	44
3.5.1 FROM SINGLE-ENDED TO FULLY DIFFERENTIAL	44
3.5.2 INCREASE IN DYNAMIC RANGE	44
3.5.3 NOISE ISOLATION	45
3.5.4 BIAS POINT VARIATION	46
3.5.5 CMFB CIRCUITS	47
Resistor-Capacitor network CMFB	48
Differential Pair CMFB	49
3.5.6 FULLY DIFFERENTIAL CLASS AB IMPLEMENTATION	50
3.6 COMPARISON	52
3.6.1 FOLDED CASCODE OTA	52
3.6.2 TELESCOPIC CASCODE OTA	53
3.6.3 TWO STAGE MILLER-COMPESNATED OTA	53
3.6.4 TWO STAGE CLASS AB OTA	54
4 DESIGNED CIRCUIT	55
4.1 DESIGN EQUATIONS	55
4.1.1 FOLDED CASCODE	57

4.1.2 TWO STAGE, CLASS A MILLER	60
4.1.3 TWO STAGE, CLASS AB MILLER	62
4.2 SIMULATIONS	66
4.2.1 DC ANALYSIS	66
4.2.2 AC ANALYSIS	66
4.2.3 TRANSIENT ANALYSIS	67
4.2.4 CORNERS AND MONTE CARLO ANALYSES	67
4.3 RESULTS AND DISCUSSION	69
4.3.1 CMFB	69
RC CMFB Circuit	69
Differential Pair CMFB Circuit	69
4.3.2 FOLDED CASCODE	71
4.3.3 MILLER CLASS A	75
4.3.4 MILLER CLASS AB	78
4.4 DESIGN CONCLUSIONS	81
4.4.1 CMFB	81
4.4.2 FOLDED CASCODE	81
4.4.3 TWO STAGE MILLER CLASS A	81
4.4.4 TWO STAGE MILLER CLASS AB	81
4.5 FUTURE WORK	82
5 CONCLUSIONS	83
6 BUDGET AND WORK PLAN	84
7 BIBLIOGRAPHY	86

LIST OF FIGURES

Figure 1.1 – I/O classification of amplifiers	3
Figure 1.2 – Local feedforward topology	4
Figure 2.1 – Silicon bonding in lattice (a), phosphorus doped silicon (b) and boron do	ped
silicon (c)	7
Figure 2.2 – pn junction and depletion region	7
Figure 2.3 – BJTs and FETs	8
Figure 2.4 – nMOS device structure	9
Figure 2.5 – Channel formation in nMOS devices.	. 10
Figure 2.6 – nMOS device in pinch off	. 12
Figure 2.7 – Channel length modulation effect on an nMOS device	. 12
Figure 2.8 – Application of the model in the different working regions	
Figure 2.9 – Small signal model assumptions [8].	. 14
Figure 2.10 – Parasitic capacitances on an nMOS device.	. 16
Figure 2.11 – nMOS small signal model equivalent circuit	. 16
Figure 2.12 – Dependence of threshold voltage on pair length (minimal length for a gi	ven
fabrication technology)	. 21
Figure 2.13 – Trade-offs of analog design in digital CMOS processes [2]	. 22
Figure 3.1 – Simple current mirror	. 23
Figure 3.2 – Small signal equivalent circuit of the simple current mirror	. 23
Figure 3.3 – Cascode current mirror	. 24
Figure 3.4 – Miller effect demonstration	. 25
Figure 3.5 – LTI, open loop system.	. 26
Figure 3.6 – Response of the LTI system to an input.	. 26
Figure 3.7 – LTI, closed loop system	. 27
Figure 3.8 – Bode plot for gain increase	. 27
Figure 3.9 – Pole-zero diagram for equation 3.8 and its closed loop system	. 28
Figure 3.10 – Nyquist diagram for Figure 3.9.	. 29
Figure 3.11 – Gain and frequency normalized plot for closed loop amplifiers with diffe	rent
phase margins [14]	. 30
Figure 3.12 – Single-ended, folded cascode topology	. 31
Figure 3.13 – Small signal equivalent circuit for single-ended folded cascode	. 32
Figure 3.14 – Output resistance for a folded cascode	. 33
Figure 3.15 – Single-ended, telescopic cascode topology	. 34
Figure 3.16 – Two stage amplifier topologies.	. 36
Figure 3.17 – Two stage, Miller-compensated OTA.	. 36
Figure 3.18 – Two stage, Miller-compensated OTA small signal equivalent circuit	. 37
Figure 3.19 – Class A, actively loaded, common source amplifier operation	. 41
Figure 3.20 – Class AB amplifier operation	. 42
Figure 3.21 – Increase in Dynamic Range	. 45

Figure 3.22 – Common mode noise cancellation, (a) noisy half-signals and (b) diff	
Figure 3.23 – Differential pair amplifier [24].	
Figure 3.24 – Stages of a CMFB circuit.	
Figure 3.25 – CMFB loop implemented using a resistor and capacitor CM sense	
Figure 3.26 - CMFB loop implemented using a resistor and capacitor CM sens	
together with the comparison and amplification blocks.	
Figure 3.27 – CMFB loop implemented using a differential pair CM sense block.	
Figure 3.28 – Fully differential class AB output stage half circuit.	51
Figure 4.1 – Transconductance dependencies	55
Figure 4.2 - Veff vs ID.	56
Figure 4.3 – Results for effective voltage at different drain-source currents, Idc,	
nMOS and pMOS devices with the dimensions reported above	
Figure 4.4 – Fully differential, folded cascode OTA with differential pair CMFB n	
Figure 4.5 – Biasing circuit for fully differential, folded cascode OTA	
Figure 4.6 - Fully differential, two stage, Miller-compensated OTA with different	tial pair
CMFB network.	61
Figure 4.7 - Two stage, Miller-compensated, class AB amplifier with DP CMF	B block
and no output CMFB network	62
Figure 4.8 – Differential pair (DP) CMFB circuit used for the first stage with the re	ference
generation circuit.	
Figure 4.9 – Diagram implementation of the RC CMFB circuit for a fully differenti AB.	
Figure 4.10 – Bias circuit branches for the two stage, Miller-compensated, cl amplifier.	
Figure 4.11 – AC analysis results for the RC CMFB circuit implementation	
Figure 4.12 – AC analysis results for the differential pair CMFB circuit impleme	ntation.
Figure 4.13 – Fully differential folded cascode OTA	
Figure 4.14 - AC results as a Bode plot for fully the differential folded cascode C	
Figure 4.15 – Slew rate results for fully differential folded cascode OTA.	
Figure 4.16 – AC results as a Bode plot for fully the differential folded casco	
without the load resistor	
Figure 4.17 – Two stage, Miller-compensated, class A amplifier.	
Figure 4.18 – AC analysis results for two stage, Miller-compensated, class A a	
with load resistor.	•
Figure 4.19 – AC analysis results for two stage, Miller-compensated, class A a	mplifier
without load resistor	77
Figure 4.20 – Two stage, Miller-compensated, class AB amplifier	78
Figure 4.21 - AC Analysis results for the two stage, Miller-compensated, cl	ass AB
amplifier.	79
Figure 4.22 – Slew rate results for the two stage, Miller-compensated, class AB and	•
Figure 6.1 – Work plan applied throughout the thesis.	

LIST OF TABLES

Table 1.1 – Design Parameters for the amplifiers.	3
Table 2.1 – KVL applied to the terminals of an nMOS device.	11
Table 2.2 - Large signal equations for nMOS devices neglecting some second of	order
effects.	13
Table 3.1 – Single-ended, folded cascode equations	34
Table 3.2 - Single-ended, telescopic cascode equations	35
Table 3.3 – Single-ended two stage, Miller compensated OTA equations	
Table 3.4 – Class A and Class AB comparison.	43
Table 3.5 – Comparison of different amplifier topologies.	54
Table 4.1 – Values of k for the 180 nm pMOS and nMOS devices	56
Table 4.2 – DC analysis results for RC network CMFB circuit.	69
Table 4.3 - DC analysis results for differential pair CMFB circuit	70
Table 4.4 – Folded cascode OTA DC analysis, bias results	71
Table 4.5 – Corner process analysis results for the fully differential folded cascode (OTA.
	74
Table 4.6 – Monte-Carlo analysis results for fully differential folded cascode OTA	74
Table 4.7 – Two stage, Miller-compensated, class A amplifier DC analysis results	75
Table 4.8 – Corner analysis results for two stage, Miller class A amplifier	76
Table 4.9 – Two stage, Miller-compensated, class AB amplifier DC analysis results	78
Table 4.10 - Corner analysis results for two stage, Miller class AB amplifier	79
Table 4.11 - Monte-Carlo analysis results for the two stage, Miller class AB amp	lifier.
	80
Table 5.1 – Update topology comparison table.	83
Table 6.1 – Budget of the developed research	84

1.-INTRODUCTION

1.1.-MOTIVATIONS

Ever since the advent of portable electrical devices power consumption has been a paramount issue. Since these systems have to be connected to a battery, a dependency upon the available electric charge exists, and thus there is a limit to its service life.

One of such first devices is considered to be the electrically-powered watch, driven through a pair of solenoids and developed in the late 50s. During the next two decades these devices became transistorized, transforming into electronic wrist-watches. The advent of the 70s brought up the first pocket electronic calculators, and coupled to the rest of computing devices, it was clear that the path to follow was marked by transistors.

As the century approached the 90s, more and more transistors were crammed into the same area with each passing year, following the trend stated by Moore's law. Parallel to this, batteries kept increasing in size and charge capacity while remaining comfortable enough to handle.

It was not until the start of the millennium that Moore's law predicted the density of transistors needed to embed computers in wafers with size in the order of decimeters and below. With this quantity of transistors came a corresponding power consumption that had to be provided by batteries which, at that point, were already on the verge of their capacity/size ratio. Henceforth, an initial true effort was started to scale down voltages and currents in an attempt to scale down power consumption.

Nowadays, the onset of *smart* devices is seeing this effort pushed strongly, trying to reduce power consumption by the means of scaling down the voltage of the used transistor technologies to 1.5, 1.2 and even 0.8 or 0.5 V. As a prospect for the future, this downscaling will allow to develop smaller, lighter (and thus more portable) devices, which allow for a more efficient use of energy.

The use of these efficient devices is appealing in the medical field, and concerns analog microelectronic blocks of systems such as DACs and ADCs. Having this in mind, the thesis was developed in a research project titled "*Design and characterization of main building blocks for Medical Instrumentation ADCs*" at the Microelectronics Design and Applications (DMA) research group at Carlos III University of Madrid.

1.2.-BRIEF OVERVIEW OF OPERATIONAL AMPLIFIERS

The term *amplifier* is used to define a family of electronic devices that perform the same basic task. Throughout the document the term *amplifier* will be used in reference to operational amplifiers. Their function consists on taking a certain signal as an input and, through one or more stages, boost or enlarge it obtaining a scaled output. Of course, this signal can be a current or a voltage, and depending on whether one is converted into another during the amplification, one can think of four different scenarios.

1.2.1.-SOURCE CLASSIFICATION

Take for instance a voltage input and feed it into an amplifier. If the output was to be an amplified voltage signal, one could label the amplifier as voltage-controlled voltage source, a VC-VS. However, if the output was an [amplified] current signal, the device would be now labelled as a voltage-controlled current source, VC-CS for short. Conversely, if the input signal was a current and the output a voltage, it would be labelled as a current-controlled voltage source (CC-VS). Finally, the last combination suggests a current input for a current output, making it a current-controlled current source (CC-CS).

Current and voltage inevitably coexist as Ohm's law states, so how can one talk about a signal being "current" or a signal being "voltage"? The truth is that a signal has two characteristics: the current and the voltage. Only one of them is controlled, being the other one determined by the load through which the signal is being forced. In amplifier applications, depending on the type of the signal and the characteristics of the load, one will use any of the four aforementioned combinations.

A classification intro *controlled sources* was mentioned above. It is true that one can think of an amplifier as a current or voltage source controlled by a certain parameter. Nevertheless, let's analyze the properties one would wish those devices had to achieve that *controlled source* definition.

For instance, take a voltage-controlled voltage source. This implies that given a voltage input, another voltage output is provided, regardless of the demanded current. This is in conjunction with the definition of voltage source, a device that provides a constant voltage despite of the required current. This is achieved by placing a very low impedance at the output (if this resistance was to be null, the voltage drop across the output would be exactly zero). Analogously, when considering the input of the device, one wants it to *sense* the same voltage, no matter the current. This is typically done by applying a very high input resistance.

These are key points one has to bear in mind when designing amplifiers, as depending on the application one or another topology will be used as a function the signal characteristics and the load requirements.

Finally, it is mandatory to introduce some nomenclature. While the vast majority of amplifiers are preferred to have an infinite input impedance (as already mentioned, to sense the input voltage without interfering with it), the output impedance is a factor that attends to the necessities of the design. As such, two types of infinite input impedance amplifiers arise: those who present a low output impedance, and those with a high output impedance. These are referred to as operational amplifiers (Op-Amps or OAs, for short) and operational transconductance amplifiers (OTAs).

1.2.2.-I/O CLASSIFICATION

Another classification can be established based on their type of input/output (IO), amplifiers are segregated into single-ended and differential IOs. An amplifier can be configured into a single-ended input by referencing one of its inputs to ground, being the other the signal carrier. On the other hand, if they are presented with two inputs which are complementary, they are said to be configured into a differential input.

One can already realize that the vast majority of designs make use of differential inputs and that they are used in one fashion or another. However, one cannot configure the outputs as it is done with the inputs in such a straight-forward manner. The topologies used in amplifiers require them to be designed as differential or single-ended outputs themselves, due to the control of the common mode level within the amplification stages.

In short, there exist four different categories:

- Single-ended input, single-ended output, shown in Figure 1.1 (a).
- Single-ended input, differential output.
- Differential input, single-ended output, represented below as Figure 1.1 (b).
- Differential input, differential output, depicted in Figure 1.1 (c).

This leads to the used classification, the first and third categories are termed as differential amplifiers mostly because, as aforementioned, one can convert the differential input into a single-ended one. And fully differential amplifiers that present two differential inputs and a differential output.

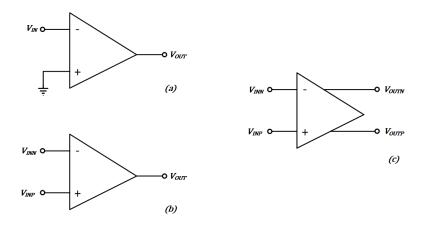


Figure 1.1 – I/O classification of amplifiers.

1.3.-OBJECTIVES

The objective of this bachelor thesis span over the design and characterization of different operational amplifiers at low voltage for a $\Sigma\Delta$ ADC specifically designed for biomedical sensors [1]. The characteristics of this amplifier have to meet a maximum specified in the table below.

	A _{dc}	GxBW	SR	I_{DC}^{max}
	55 dB	200 MHz	15 V/µs	5 mA
l = 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1				

Table 1.1 – Design Parameters for the amplifiers.

The system for which the amplifiers have to be designed is shown in Figure 1.2. The first four amplifiers are acting as integrators, while the fifth and last one acts as an adder. The connected resistances and capacitors are given by the state variables of the system.

Note that the maximum capacitance that the amplifiers will see is of 20 pF, and a minimum resistance of 1 k Ω . This has to be taken into account when designing the

amplifier, as the equivalent load in open loop will amount to the capacitances in parallel to the resistor.

The integrator-configured amplifiers will be designed with the same specifications, in open loop with a 20 pF load and a 1 k Ω resistor. The adder-configured amplifier will be moving the inputs of the flash ADC, modelled by a 3 pF load.

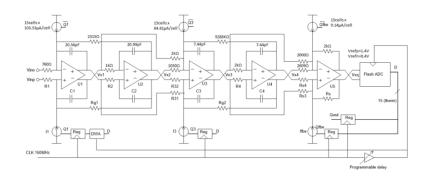


Figure 1.2 – Local feedforward topology.

Nevertheless, the main goal of this thesis will be to assess the differences between the different topologies as well as implementation possibilities both analytically and through simulation results.

1.4.-STATE OF THE ART

It is no mystery that the analog microelectronics field has been approaching its extinction. With the exception of current drivers, ADCs and RF systems, the industry is quickly shifting towards digital design. As a consequence, the available technologies are shrinking in size and being optimized in delay times rather than in the aspect of their linearity.

The explained above implies that analog design hast to be usually undertaken with technologies meant for digital design. This, in turn, causes that the accuracy of the fabricated circuit does not match the simulations which, consequently, do not follow accurately the predictions of the equations that describe the topologies. This adds up to the design and development time of analog systems and causes them to lag behind digital design. This phenomenon is caused by the rapid migration from one technology to the next, smaller one [2].

Nowadays, the field has been revisited due to the rise in consumer portable devices such as smart-devices which have to present an array of sensors and instruments. It is precisely the need of interaction with analog signals in the form of ADCs and amplifiers that still drives the interest in the analog field.

In the case of this bachelor thesis, the objective focus on designing an OTA as a part of an ADC system. A search powered by IEEE Xplore Digital Library reveals that there are few OTA designed in 0.18 μ m at voltage lower than 1.8 V with the characteristics specified above.

There are several publications that report amplifiers using 180 and 65 nm processes at extremely low voltages [3], [4], [5]. Nevertheless, these present very small unity-gain frequencies (under the 100 MHz range). Older publications [6] have reported results at the tenths of MHz at 3.3 V.

This analysis reveals that the imposed specifications are indeed ambitious at the targeted supply voltage with the goal of optimizing power consumption.

2.-CMOS TECHNOLOGY

The designed amplifier will be part of a microelectronic system integrated on a chip. As such, it has to be designed with the same CMOS technology as the rest of the circuit in order to portray a coherent structure. The manufacturing technology was chosen to be the AMS (Austria Mikro Systeme) 0.18 μ m platform, specifically designed for low power applications.

Along the next sections a brief description will be given on the basics of the CMOS technology and its main parameters.

2.1.-INTRODUCTION TO CMOS TECHNOLOGY

The semiconductor industry is based around one type of material: semiconductors. These are defined as being materials composed by semiconductor elements with an electrical conductivity that lays in between that of materials like metals and insulators. The default semiconductor elements are silicon and germanium, being the first one the dominant in the industry.

2.1.1.-SILICON AND DOPING

Silicon presents four electrons in its valence shell, thus being four electrons short of completing the octet rule. When arranged in a crystalline structure, each silicon atom will try to acquire those four vacant electrons from its immediate neighbors in order to remain energetically stable. Singularly, in a crystal lattice, silicon will be achieving this by sharing each one of its four valence electrons with four other alike atoms (and therefore creating a covalent bond), one electron per adjoining silicon.

One can view each silicon atom inside a silicon crystal lattice as having four free electrons and four vacancies or holes, shown in Figure 2.1 - (a). If some other foreign atom, with a different number of valence electrons, was to be introduced into the crystal lattice, the balance between holes and electrons would be shifted towards one side or another. This process is called doping.

One can introduce an atom with either five or three valence electrons, being phosphorus and boron the most commonly used ones, respectively.

- If phosphorus is added, one extra free electron is introduced per foreign atom into the lattice, which will turn into a charge carrier. This can be seen represented in Figure 2.1 – Silicon bonding in lattice (a), phosphorus doped silicon (b) and boron doped silicon (c).. Note that the crystal is now said to be n-doped (as there is an excess of *negative* charge).
- On the other hand, if a boron atom is introduced with its three free electrons, there will be one fewer electron in the lattice per introduced impurity. This is equivalent to the addition of one extra positive carrier to the lattice as it is shown below, in Figure 2.1 Silicon bonding in lattice (a), phosphorus doped silicon (b) and boron doped silicon (c). (c). Analogously as with phosphorus, the silicon crystal is now said to be p-doped (as there is a lack of negative charge).

The basic building blocks of any transistor circuit hinge around the phenomenon that takes place when two oppositely-doped silicon lattices are found facing each other.

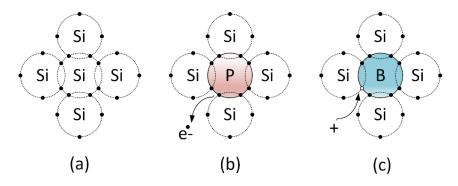


Figure 2.1 – Silicon bonding in lattice (a), phosphorus doped silicon (b) and boron doped silicon (c).

2.1.2.-THE PN JUNCTION

When one half of a silicon crystal is p-doped and the other half is n-doped, the space in region in between them is called the pn junction, which is the area of interest for any diode device. This can be seen in Figure 2.2 - (a).

The n-doped region is rich in free electron carriers, while the p junction is rich in holes, or positive charge carriers. If the conditions are favorable enough, an electron will *jump* into the p-doped region, *filling* a hole (a process known as recombination). This will take place for as long as the potential allows electrons from the n-doped region to jump and recombine with the positive charge carriers in the p-doped region, creating a charge carrier-lacking area known as the depletion region, shown in Figure 2.2 - (b).

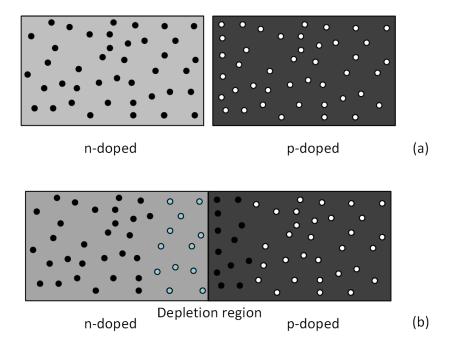


Figure 2.2 – pn junction and depletion region.

As the p-doped region has gained electrons, it now presents a slightly negative charge, while the converse is true for the n-doped region. If a power source is to be connected between the terminals of the silicon crystal, the depletion region will act as an insulator impeding the transit of electrons. Once the voltage between the terminals of the silicon crystal exceeds 0.6 V (the breakdown voltage), the depletion region will collapse and allow the transit of electrons.

2.1.3.-THE TRANSISTOR

In order to dive deeper into the phenomena that govern the operation of transistors, it is necessary to obtain some intuitive understanding of their purpose and working.

It is commonly explained that a transistor behaves like a switch: when some voltage is applied to one of its three inputs, the switch opens letting current flow through the other two terminals. In fact, it is a controlled current source, this is: by applying some input in the form of either a voltage or a current, another current flows through the remaining terminals.

Once known this, one can be introduced to the two main different types of transistors in electronics: the BJTs and the FETs. A transistor can be built by arranging the differently doped regions in several ways. The two most common ways give rise to bipolar junction transistors (BJTs) in Figure 2.3 - (a), and to metal oxide semiconductor field effect transistors (MOSFETs) in Figure 2.3 - (b).

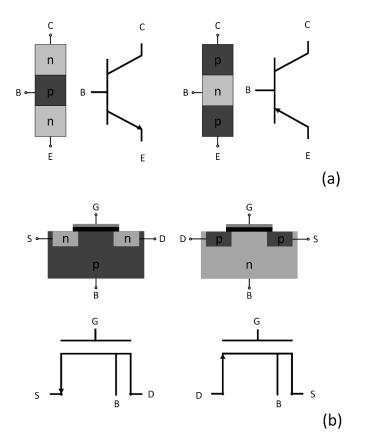


Figure 2.3 – BJTs and FETs.

It is important to note that the two main transistor technologies are the BJTs and the FETs, as already mentioned above. Among the FET transistors there exist different technologies as well like the MOSFETs, JFETs, IGBTs, and many other devices. Nevertheless, the most interesting for CMOS technology are MOSFETs.

BJTs are, in essence, two diodes joined by the base lead. Focusing on an NPN BJT transistor, Figure 2.3 - (a, left). If a voltage higher than the breakdown potential is applied to this terminal (positive to the base, and negative towards the emitter) current will flow through the emitter to the base, as the "diode" becomes forward biased. Moreover, as one applies an additional voltage across collector and emitter, a current will flow between these two terminals. This provides the intuition behind the reason why BJTs are usually referred to as current controlled current sources.

2.1.4.-THE MOSFET

The most commonly used type of MOSFET transistor is the one depicted in Figure 2.3 - (b, left). The device's structure is represented in more detail below.

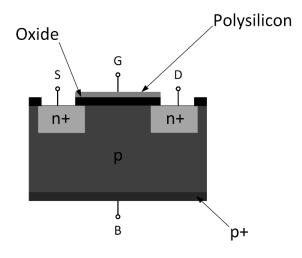


Figure 2.4 – nMOS device structure.

As depicted above, an nMOS-type FET is fabricated on top of a p-doped wafer. Two wells, noted as n-wells, are deposited on the surface of the p-doped crystal. They are referred to as n+, as they are strongly doped. Furthermore, a region called the body is fabricated by strongly p-doping the silicon crystal. Covering the p-region between the n-wells, a thin layer of silicon oxide (SiO₂) is formed onto which another layer of polysilicon is deposited. It is to this film that the gate terminal will be attached. Lastly, note that the oxide layer will behave as an insulator.

Since the technology is fabricated on top of a p-doped crystal, in order to fabricated a PMOS type FET transistor it is necessary to deposit an n-doped region first onto which p-wells will be placed.

The n+ and p interfaces are pn junctions and, consequently, will result in corresponding depletion regions, being equivalent to two diodes placed back-to-back.

TRIODE REGION

When both source and drain are grounded, and a voltage large enough is applied to the gate terminal, the accumulation of positive charges in the polysilicon will induce the accumulation of negative carriers in the p-doped region underneath. Furthermore, when the concentration of negative charges is equal to the concentration of holes in the region, a depletion zone appears between the n-wells. The voltage at which this takes place is referred to as the threshold voltage of the transistor, also known as V_{th} .

If the potential at the gate is increased, and the concentration of electrons becomes that of the holes further away from the oxide layer, a channel is formed as the p-doped region becomes n-type in the proximity of the oxide layer, shown in Figure 2.5.

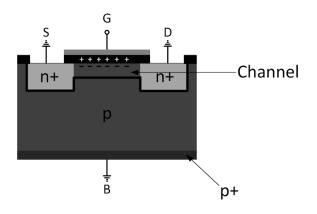


Figure 2.5 – Channel formation in nMOS devices.

It is intuitive that the carrier density of the channel depends on the respective voltages between source, gate and drain. On the other hand, as the potential between gate and source (V_{GS}) increases, so does the *effective* concentration of negative charge carriers in the formed channel. Hence, the voltage that is related to this increase is noted as effective gate-source voltage, $V_{eff} = V_{GS} - V_{th}$.

When the drain potential is increased, a current appears between drain and source. The relationship between this drain-source current is a non-linear function and it is given by equation 2.1 [7].

$$I_D = \mu C_{ox} \frac{W}{L} \left[(V_{GS} - V_{th}) V_{DS} - \frac{V_{DS}^2}{2} \right]$$
(2.1)

Where μ is the mobility of electrons through the semiconductor, C_{ox} is the capacitance of the oxide layer, W is the width of the transistor device and L is the length between the n-wells.

Note that when the channel is starting to form due to the accumulation of charges in the region, the current through the device can be assumed to scale linearly with V_{DS} , since it is small enough to neglect the squared term.

This equation is only valid for a certain range of the gate-source voltage. Until this paragraph, the range is known to be from $V_{GS} > V_{th}$ up until a yet unknown value.

As a final remark, applying Kirchhoff's voltage rule between the terminals of the nMOS device yields the expressions summarized in Table 2.1.

Gate-source	Drain-source	Drain-gate
$V_{GS} = V_G - V_S$	$V_{DS} = V_D - V_S$	$V_{DG} = V_{DS} - V_{GS} = V_D - V_G$
Table 2.1 – KVI applied to the terminals of an nMOS device		

Table 2.1 – KVL applied to the terminals of an nMOS device.

SATURATION REGION

As the drain voltage is increased, there is an accumulation of positive carriers that will deflect the charge density of the channel near the drain terminal, shown in Figure 2.6. This takes place until a voltage that does not allow the sustenance of the channel anymore, this is: until the voltage potential between the gate and the channel reaches the value of the threshold voltage. This has to be expressed mathematically, the channel's voltage at the drain region can be thought as being:

$$V_{GD} = V_{GS} - V_{DS} \tag{2.2}$$

Since V_{DS} is being increased and V_{GS} is assumed to be constant for the time being, then V_{GD} is decreasing steadily. The condition of the channel being narrowed until it can no longer be sustained can be expressed as:

$$V_{GD} \ge V_{th} \tag{2.3}$$

Combining equations 2.2 and 2.3 one obtains the limit of V_{DS} for which any further increase in voltage will yield no further current.

$$V_{GS} - V_{DS} \ge V_{th} \to V_{DS} \ge V_{GS} - V_{th}$$

$$(2.4)$$

At that point the channel is said to be *pinched-off*. Ideally, any further increase in V_{DS} will not increase the maximum current through the channel, which is said to be *saturated*. Thus, equation 2.4 serves as a limiting value between the triode and saturation regions and it is noted as V_{DSat} .

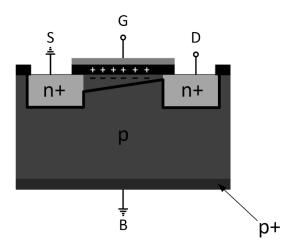


Figure 2.6 – nMOS device in pinch off.

Substituting equation 2.4 into 2.1 yields the relationship of the drain-source current as a function of the transistor's biasing and parameters.

$$I_D = \frac{\mu C_{ox}}{2} \frac{W}{L} (V_{GS} - V_{th})^2$$
(2.5)

CHANNEL LENGTH MODULATION

Equation 2.5 implies that if V_{DS} is increased beyond $V_{GS} - V_{th}$ the current through the device, I_D , stays constant. This is only partially true since the equation was derived neglecting an effect known as channel length modulation.

As V_{DS} is increased beyond $V_{DS,sat}$ the effective length of the channel is reduced. This increases the depletion region (shown as a black line in Figure 2.7) while shortening the channel and increasing the current through the device.

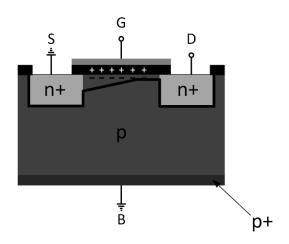


Figure 2.7 – Channel length modulation effect on an nMOS device.

This increase in current can be modeled by modifying equation 2.5 as shown below.

$$I_D = \frac{\mu C_{ox}}{2} \frac{W}{L} (V_{GS} - V_{th})^2 \left[1 + \lambda (V_{DS} - V_{DS,sat}) \right]$$
(2.6)

The parameter that accounts for the *amount of extra current* due to channel length modulation is termed as λ , also called the output impedance constant since it is the factor that decreases the output impedance of the device. It can be modelled as a function of the effective length of the channel, as shown by equation 2.7.

$$\lambda = \frac{\Delta L}{LV_{DS}} \tag{2.7}$$

These equations, also known as power law equations, describe the approximate behavior of transistor for *large signals*, both in the triode and saturation regions, a graph with the applications of the equations can be seen in Figure 2.8, for a constant V_{GS} . It is important to realize that they do not acquaint second order effects such as body effect.

Region	Condition	Equation
Cut off	$V_{GS} < V_{th}$	-
Triode	$V_{GS} > V_{th}, V_{DS} < V_{GS} - V_{th}$	$I_D = \mu C_{ox} \frac{W}{L} \left[(V_{GS} - V_{th}) V_{DS} - \frac{V_{DS}^2}{2} \right]$
Saturation ¹	$V_{DS} \ge V_{GS} - V_{th}$	$I_{D} = \frac{\mu C_{ox}}{2} \frac{W}{L} (V_{GS} - V_{th})^{2} [1 + \lambda (V_{DS} - V_{DS,sat})]$

Table 2.2 – Large signal equations for nMOS devices neglecting some second order effects.

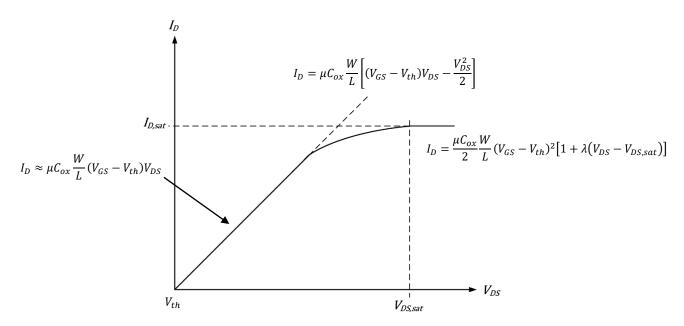


Figure 2.8 – Application of the model in the different working regions.

¹ Channel length modulation has been included into the equation for the current in order to derive the drain-source resistance for the small signal model.

2.1.5.-CMOS TECHNOLOGY

Complementary MOS technology refers to the design technology that uses both nMOS and pMOS FET devices together. It is important to distinguish between what fabrication and design technologies are. CMOS design technology is widely used in the semiconductor industry for the design and fabrication of microprocessors, amplifiers, and all suits of both analog and digital circuits. On top of this, each manufacturer can use another fabrication technology.

As already mentioned before, since both n and p devices have to be fabricated on the same wafer, the pMOS devices need to be placed on top of an n-well deposited on top of the common p-type wafer. How this fabrication is performed, and with what techniques describes the fabrication technology for a CMOS design technology.

CMOS is widely used nowadays due to its numerous advantages like its high input impedance. This high resistance comes from the gate of the device which presents a small capacitance, being the current going through the gate in the order of nano-amperes (a phenomenon known as leakage current). Another one of its most notable advantages is the relatively low power consumption, due to the device's charge and discharge of its parasitic capacitances when current flows through them. Nevertheless, they present a considerable mismatch as well as second order effects that renders them unsuitable for certain applications.

2.2.-SMALL SIGNAL MODEL PARAMETERS

In the last section the power law equations were described, these relations are also known as the large signal model (in the case of the last section, without including second order effects). When designing an analog system in CMOS, one wants to stand on a model that is linear so that everything can be described in terms of linear relationships. This is only possible if the devices are biased for the saturation region, where the relationships between current and voltage can be assumed to be linear around the bias point.

This is why the model is called *small signal*, as one needs to assume that the linearity only holds for small deviations around the bias point. In addition, the relationship between V_{GS} and I_D can be thought of as a line with a slope and an intercept instead of the second order relationship posed by the square law equations, shown in Figure 2.9.

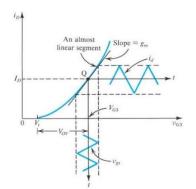


Figure 2.9 – Small signal model assumptions [8].

2.2.1.-TRANSCONDUCTANCE

The slope of the linear relationship mentioned above is known as the gate-source transconductance. The line is defined to be $I_D(V_{GS})$. The slope of that line is the first order derivative, neglecting channel length modulation and body effects:

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \mu C_{ox} \frac{W}{L} (V_{GS} - V_{th})$$
(2.8)

It is important to understand the concept of transconductance, which is the relationship between the output current of a device and the applied voltage.

2.2.2.-DRAIN-SOURCE RESISTANCE

The output resistance in the saturation region, also called drain-source resistance, is an important small signal parameter as well. It is obtained as the inverse of the drain-source transconductance:

$$r_{ds} = r_o = \left(\frac{\partial I_D}{\partial V_{DS}}\right)^{-1} = \frac{1}{\lambda I_{D,sat}}$$
(2.9)

2.2.3.-PARASITIC CAPACITANCES

A capacitance arises whenever there is an accumulation of charge and a consequent voltage difference between them. This means that, whenever there are two electrical components together, unavoidably, an unwanted capacitance will appear. When dealing with transistors, one can assume that the doped regions behave like plates, which are joined in parallel at an interface with other ones. The capacitance of a parallel plate capacitor is described as:

$$C = \frac{\varepsilon_o A}{d} \tag{2.10}$$

Where ε_o is known as vacuum permittivity, *A* is the area of the each plate and *d* is the distance between them. Knowing this, one can identify a set of capacitances in a typical MOSFET device whenever two differently doped regions are confronted. These can be seen in Figure 2.10.

Among all of these capacitances, the ones with the highest values are C_{gs} and C_{gd} . The latter one is often termed as the Miller capacitance when designing amplifier circuits. The capacitances related to the body appear if there is a voltage difference between the body and source (known as the body effect). Nevertheless, assuming that this potential different is null (neglecting the body effect), the capacitances can be understood as nonexistent.

These capacitances can be described, in the saturation region, by the oxide capacitance times the dimension parameters, and following equation 2.10 one can write:

$$C_{gs} \cong \frac{2}{3} C_{ox} L W \tag{2.11}$$

$$C_{gd} = W C_{GDO} \tag{2.12}$$

Where C_{GDO} is described by the channel length and the oxide capacitance. This comes from the fact that the length is not constant in saturation region due to pinch-off and channel length modulation.

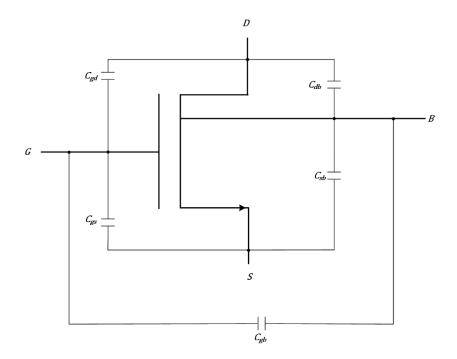


Figure 2.10 – Parasitic capacitances on an nMOS device.

2.2.4.-SMALL SIGNAL MODEL

The small signal model of an nMOS device can be represented through a small signal equivalent circuit. If body effect and channel length modulation are neglected (this is: the body and the source are at the same potential), then this circuit is shown in Figure 2.11.

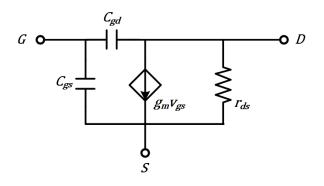


Figure 2.11 – nMOS small signal model equivalent circuit.

This circuit will be useful when analyzing some of the transistor parameters discussed in the upcoming sections. As a final note, the body effect is neglected and, therefore, the capacitances related with the body vanish.

2.3.-DESCRIPTION OF THE USED CMOS TECHNOLOGY

Regarding analog design, CMOS technologies are usually divided among short and long channel processes, having each their respective models. In addition to this, not all technologies are suitable for analog design, as their parameters vary from process to process.

The used CMOS technology presents a minimal channel length of 180 nm, hence not being considered short channel. In order to understand better why it can be considered a good analog technology, one needs to first review some of the concepts that lay behind analog process characterization.

2.3.1.-SHORT CHANNEL EFFECTS

Short channel effects are pronounced phenomena in MOSFET devices whose channel length, L, is in the same order of magnitude as the size of depletion regions of the drain and sources with the substrate.

As transistors shrink in size, the effect of short channel effects is more prominent and the power law equations have to be rewritten. This is particularly problematic when designing analog systems, as the designs are based around the very same models that have to be modified. This is why, in general, most of analog systems are designed with a much higher *L* than the allowed minimum.

It is important to differentiate secondary effects, which are already present in long channel devices but whose magnitude increases as the length is decreased, from strict short channel effects [9]. Two of the most prominent phenomena that worsen with channel length decrease are the saturation of the carrier velocity and their mobility degeneration. Secondly, some of the most notable *strict short channel effects* are the drain-induced barrier lowering (DIBL), impact ionization and hot carrier injection (HCI).

Short channel effects will affect the transconductance parameters, carrier mobility, threshold voltage, induction of parasitic effects, and deviation from the relationships of power law equations.

CARRIER VELOCITY SATURATION

In every transistor there exists a voltage difference across the terminals, this creates electric fields that affect charges flowing through it. One can obtain the horizontal field across the device as:

$$E_H = \frac{V_{DS}}{L} \tag{2.13}$$

This field determines the drift velocity of charge carriers across the channel up to a certain point. One cannot increase the field and expect the carrier velocity to increase infinitely as, eventually, the thermal velocity is reached.

Hence, by decreasing L, one increases the field and, at some point, the carriers' velocity due to the electric field reaches the thermal velocities. This has an impact in the power

law equations, as the current expected through the device will differ from the predicted one. This has to be accounted for by placing a parallel resistor between drain and source of the devices in order to model this effect [10].

MOBILITY DEGRADATION

The horizontal field affects the velocity of charge carriers; else ways the vertical field affects their mobility, μ . This field is caused by the potential difference due to the device's gate which forces the carriers closer to the surface, where physical impediment takes place due to imperfections, reducing their mobility. This is accounted for by modifying the mobility parameter to account for the deviations arising from this effect.

DRAIN INDUCED BARRIER LOWERING (DIBL)

The drain induced barrier refers to the potential difference that an electron has to overcome in order to flow across the channel. If the drain voltage is large enough it can induce *punch through*, the extension of the depletion regions of source and drain, forming a unique one. This lowers the threshold voltage and increases the leakage current, one has to take this into account for accurate modelling of the systems.

IMPACT IONIZATION AND HOT CARRIER INJECTION

Both of these effects have to do with the high electric fields as a result of reduced channel lengths, note equation 2.13. Carriers are accelerated to high speed due to strong fields, becoming *hot carriers*.

Impact ionization refers to the effect of hot carriers knocking some electrons in the silicon lattice, causing parasitic effects.

On the other hand, hot carrier injection describes the effect of hot carriers getting trapped between the oxide and the silicon lattice, changing the gate voltage with respect to the channel.

2.3.2.-ANALOG TECHNOLOGY

The most important parameter that describes the performance of analog circuits are transconductance parameters. It basically measures the changes in the output current with respects to variations of the applied voltage across the device, being a good indicator of the device's gain.

The transconductance parameter g_m describes the gain quality of the device, while the output conductance, g_{ds} , describes the operation of the device as a current source (note that if taken the inverse of its expression, one is defining the resistance of the MOS transistor).

When designing amplifiers, and in general analog circuits, their output impedance has to be high while maintaining a rather large gain. Defining the *intrinsic gain* as the quotient between the transconductance and the output conductance, g_m/g_{ds} , one can obtain information about the general device performance for analog design.

Another parameter that has to be taken into account is the technology's matching [2], or how is the reproducibility of a device over the technology, as there are always deviations from transistor to transistor. This is especially critical when designing fully differential amplifiers, as both half circuits need to be as similar as possible to reduce common mode noise.

2.3.3.-DESIGN TIME AND COST

A scalable CMOS technology is defined as any process that presents the possibility to be enlarged with relative ease, referring to the dimensions of the devices.

Scalable technologies are desired for implementation as one can translate the known topologies to the process and then scale them accordingly to match the required specifications. This is of paramount importance in CMOS design as it allows to translate the theoretical topologies and systems easily, and then tailor them accordingly to the design specifications.

Besides the costs related to design, a 0.18 µm technology was chosen as chips for biomedical applications are not produced in quantities large enough to justify the costs of employing cutting edge fabrication technology like 7 nm or 10 nm.

A simplified explanation of this is that chips are fabricated in a fairly standard 7" wafer (without the associated photomask design cost), regardless of the technology. Therefore, the amount of chips one can obtain from a wafer increases with technology shrink down [11]. Nonetheless, the technology jump also comes with an associated cost difference. The amount of chips one can obtain from a wafer fabricated with 18 nm and 0.18 μ m increases tenfold, but the price may increase by a mere 50%. Since the market is not able to sink the offer for the extra chips resulting from using an 18 nm fabrication process, one is presented with an excess of integrated circuits and would need to increase the price to obtain a profit.

The example posed above shows how demand for a certain integrated circuit may influence the choice for the used fabrication technology.

2.3.4.-TRANSISTOR CUT OFF FREQUENCY

When discussing amplifiers, one of the main parameters is the unity gain frequency, or the frequency at which the gain becomes unity. When discussing a MOSFET device, one wants to know what is the frequency at which the current gain falls to unity. This is known as the transition or cut-off frequency, f_T , of a MOSFET. It is found using a small signal model that includes all parasitic capacitances as shown in equation 2.14.

$$f_T = \frac{g_m}{2\pi C_{gs}} = \frac{3\mu}{4\pi L^2} (V_{GS} - V_{th})$$
(2.14)

This is computed by taking into account the parasitic capacitances of the transistor, in Figure 2.11 the most influential ones are shown. Nevertheless, since it is held that $C_{gs} \gg C_{ad}$, then one can neglect them to obtain the transition frequency.

Equation 2.14 reflects the effect of Moore's law upon the transition frequency of transistors. Whenever *L* is decreased, the speed of the transistor is increased by a factor of L^2 , resulting in faster circuits. It is important to note that this conclusion does not take into account the corrections that are necessary to account for short channel effects.

Even though shrinkage of MOS devices makes circuits faster in general, when designing analog systems short channel effects are avoided at all costs, being thus speed a trade-off one should bear in mind.

2.3.5.-THRESHOLD AND SATURATION VOLTAGE IMPLICATIONS

One of the main things one has to take into account when choosing between one technology or another is the saturation, threshold and the rail voltages at which it has to be placed and biased.

The voltage, V_{DD} , is restricted by the junction breakdown of the device. Ultimately, a transistor presents a series capacitor (as already seen). Equation 2.12 describes how the capacitance scales for the dimensions of the transistor. If the specified bias voltage is exceeded, excessive charges build between the *plates* of the capacitors and it breaks down, allowing current to flow while destroying the device. Even so one cannot exceed the bias voltage, powering the transistor at lower bias voltages is possible. This is also true for the gate-source voltage, V_{GS} . An excessive V_{GS} will cause the gate oxide to break down, namely due to the capacitance described by equation 2.11, destroying the device.

On the other hand, the threshold voltage, V_{th} , is the main concern when trying to bias several transistors connected in series, since it is responsible for the main part of the voltage drop across the device, decreasing the amount of available potential for devices downstream. V_{th} depends on several variables like the temperature, the carrier mobility and the dimensions of the transistor. There are several models that propose the relationship of the threshold voltage for this parameter, being the one that described it when the source is shorted to ground noted as V_{tho} , and described by equation 2.15 [12].

$$V_{tho} = V_{FB} + \phi_S \tag{2.15}$$

Where V_{FB} is the flat band voltage, and ϕ_S is surface potential work function. Moreover, V_{FB} depends on the respective semiconductor work functions [13] which, in turn, depend on temperature, electron affinity of the semiconductor and the dimensions of the semiconductor device.

The threshold voltage usually decreases with the shortening of the minimal length of the technology. This dependence can be assumed to be linear until a physical limit is reached, this relationship can be described roughly as shown in Figure 2.12.

Since the threshold voltage is in lockstep with the minimal length throughout technologies before the physical limit is imposed, one can keep a more or less constant number of transistors in series. However, after the limit, the threshold voltage does not decrease so steadily and less and less devices can be fitted in design. This is an appalling effect for

analog design which already relies in the use of cascodes in order to fit more transistors into a design.

Withal this, the saturation voltage, $V_{DS,sat} = V_{GS} - V_{th}$ is also affected, as it decreases abruptly after this limit.

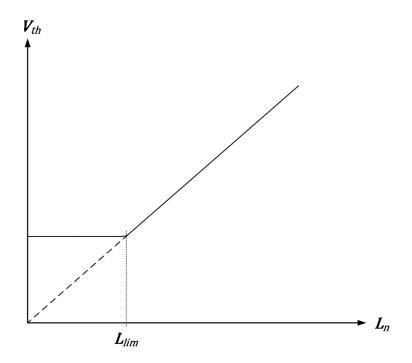


Figure 2.12 – Dependence of threshold voltage on pair length (minimal length for a given fabrication technology).

Finally, it is important to note that this limit, in which the threshold voltage varies negligibly after a given length is reached, is imposed by the physics of the material itself, and not by any design characteristic.

2.4.-MOTIVATIONS BEHIND THE TECHNOLOGY CHOICE

Nowadays choosing a CMOS technology for analog design is extremely troublesome since most of the available processes are optimized for digital control, as already mentioned. Therefore, one needs to stick with digital CMOS technologies and to accept all of the compromises related to it. These parameter trade-offs can be summarized in a simple octagon, shown below in Figure 2.13.

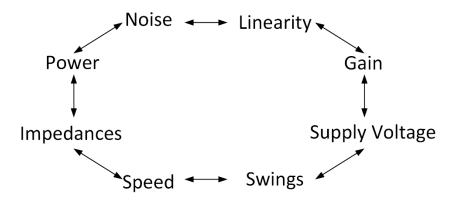


Figure 2.13 – Trade-offs of analog design in digital CMOS processes [2].

Basically, in order to increase one parameter, the other two adjacent parameters will be affected [2]. One of the most importance will be the trade-off between supply voltage, the gain and swing voltages, since the goal is to decrease the power, one will have to decrease current and voltage accordingly. Besides the aforementioned, the compromise between the impedance, speed and power dissipation will be an important point of emphasis as well.

Moving onto the chosen technology, the fact that it is in the midpoint between digital and analog design makes it possible to easily and rapidly integrate analog and digital systems such as ADCs. This comes from the phenomenon that it is small enough to have low delay times (a reasonably high cut-off frequency) while still being large enough to be considered long channel (at > 65 nm minimal length), and thus minimizing short channel effects while maintaining the simpler, long channel models.

Furthermore, it presents a relatively good intrinsic gain compared with purely digital processes, being this ideal for amplifier implementation. Besides, it is a very scalable technology, meaning that the design time and fabrication costs are relatively low in comparison with other technologies.

The size of the process implies that the cut off frequency will not be very high, being it in the range of the GHz, is one of the main compromises of the used technology.

3.-OTA TOPOLOGIES

Amplifiers can be cascaded consecutively, being each amplifier a stage of the total system. This effectively increases the gain of the whole circuit, but at the expense of more power consumption and complexity. Furthermore, this means that the output behavior of the amplifier will be mainly dominated by the output stage, being thus important to choose the final stage topology with care for it to remain an OTA.

3.1.-INTRODUCTORY CONCENPTS

In order to fully grasp the equations and characteristics of the topologies described in this section, it is first necessary to describe some introductory concepts such as current mirrors and the Miller effect.

3.1.1.-CURRENT MIRRORS

These are devices that accept a current input, usually termed I_{REF} , and provide a unityscaled copy of it at their output, noted as I_0 . They are usually used in circuits as a sensing device for a reference current, which is then copied through a supply, or as a means to copy a current within a circuit itself. As they have to work as current sources, they need to present ideally low input and infinite output impedances.

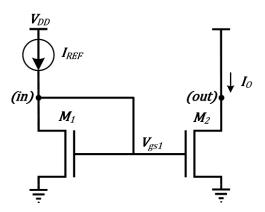


Figure 3.1 – Simple current mirror

Intuitively one may locate these resistances at the nodes noted as (in) and (out) in Figure 3.1 for the input and output impedances, respectively. One can obtain the values for the input and output impedances of a simple current mirror by analyzing the small-signal circuit of the system, represented below in Figure 3.2.

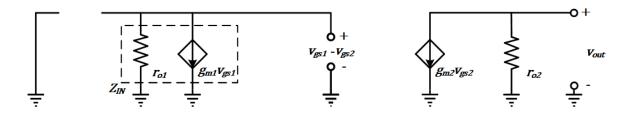


Figure 3.2 – Small signal equivalent circuit of the simple current mirror

Doing this provides an expression for the input impedance:

$$Z_{in} = \frac{1}{g_{m1}} || r_{ds1} \sim \frac{1}{g_{m1}} \text{ (since } r_{ds1} \ll g_{m1})$$
(3.1)

On the other hand, the output impedance is equal to the drain-source resistance of the other nMOS device:

$$Z_{out} = r_{ds2} \tag{3.2}$$

However, the output resistances are too low, and the input resistances too high. This problem will be addressed in the next section.

The usefulness of current mirrors resides in their application as CC-CS. By taking the ratio between the input and output currents, and neglecting channel-length modulation effects, one can manipulate the power law equations of transistors to obtain that:

$$\frac{I_0}{I_{REF}} = \frac{W_2/L_2}{W_1/L_1}$$
(3.3)

This equation shows that it is possible to manipulate the widths, while maintaining the same length, to up-scale or down-scale the current.

3.1.2.-CASCODING THE CURRENT MIRROR

One way to improve the output impedance is to cascode the current mirror, which is done by adding another pair of transistors. The circuit is depicted in Figure 3.3.

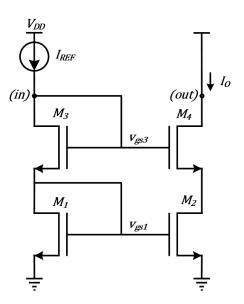


Figure 3.3 – Cascode current mirror

The same type of analysis is performed on this system. From the small signal analysis and neglecting the body effect, one can obtain the output resistance expression:

$$Z_{out} \approx r_{ds4}(1 + r_{ds2}g_{m4}) \cong r_{ds2} \cdot r_{ds4}g_{m4}$$
(3.4)

The output resistance is effectively increased by a factor of $r_{ds4}g_{m4}$. However, this will reduce the available voltage for the signal, or in other words: it reduces the voltage swing.

3.1.3.-MILLER'S THEOREM

The Miller effect, named after John Milton Miller, explains the phenomenon that causes the increase of the input capacitance by a factor of an inverting amplifier's gain. This takes place in any amplifier with a negative feedback loop, consider the amplifier in Figure 3.4 with a certain gain A_v .

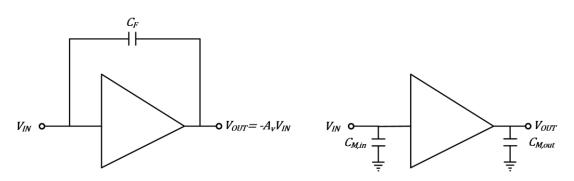


Figure 3.4 – Miller effect demonstration.

The current through the capacitor is provided by the difference between the input and output. However, one can decompose the capacitances as shown in Figure 3.4 (right) and obtain the following expressions:

$$C_{M,in} = C_F (1 + |A_v|) \tag{3.5}$$

$$C_{M,out} = C_F \left(1 + \frac{1}{|A_v|} \right) \tag{3.6}$$

The input capacitance results in the feedback capacitance, C_F , multiplied by a factor of the gain, effectively increasing it at the input and slowing down the operation of the system.

3.2.-SYSTEM STABILITY AND COMPENSATION

In this section, the basis and criteria for the stability of a system will be set as well as the introduction and types of phase compensation will be reviewed.

3.2.1.-STABILITY

An electronic system, and any system in particular, presents a gain and phase response. Let's assume an LTI (linear and time invariant) system, depicted in Figure 3.5.

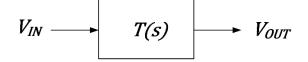


Figure 3.5 – LTI, open loop system.

When an input voltage is applied, the system will provide a modified output. This modification comes from the properties of block T which, in addition, vary with respect to frequency. The gain of a system is the proportion of the input and output magnitudes of the system, while phase is the time shift of the signal. Since the input is a sinusoidal wave, this shift can be measured in terms of degrees or radians. In Figure 3.6 the gain of the system would be 2 (for the given frequency at which the signal is introduced), and a phase shift of $2\pi/3$.

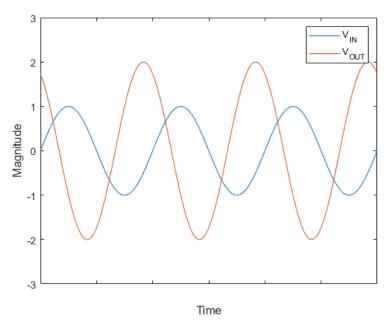


Figure 3.6 – Response of the LTI system to an input.

Since the gain and phase shift vary with respect to frequency Bode plots are used to assess the variations of both parameters over the entire frequency spectrum.

A description of margin is the extra amount of *something* that shields from instability. Therefore, the lower the margin, the more likely the system will be to be unstable. When designing amplifiers, one always talks about phase margin. In order to understand the underlying concept to this, it is important to know beforehand what makes a system unstable.

From control theory, a system described by a transfer function, like the one shown in Figure 3.5, will be unstable if there is at least one pole in the right half plane. This can be also extrapolated to closed loop systems, like the one shown in Figure 3.7. Then, the transfer function of the system becomes equation 3.7.

$$G(s) = \frac{T(s)}{1 + T(s)}$$
(3.7)
$$V_{IN} \longrightarrow T(s) \longrightarrow V_{OUT}$$

Figure 3.7 – LTI, closed loop system

Since instability implies having a non-bound parameter, the system can be unstable only if the gain (this is: the modulus of the transfer function) goes to infinity. This only happens whenever the denominator of the transfer function goes to zero, or mathematically:

T(s) = -1

This is what corresponds to a phase shift of -180° and a gain of 0 dB. If one was to adjust the gain of a system, the Bode plot of the gain would move up or down, depending on the adjustment. Whenever the gain is modified, so is the frequency at which 0 dB is attained, until eventually one hits the value at which the phase shift is -180°, rendering the system unstable. This can be intuitively seen in Figure 3.8. The amount which one would have to modify the gain in order to locate the 0 dB gain at the ω_{180} frequency is called the gain margin.

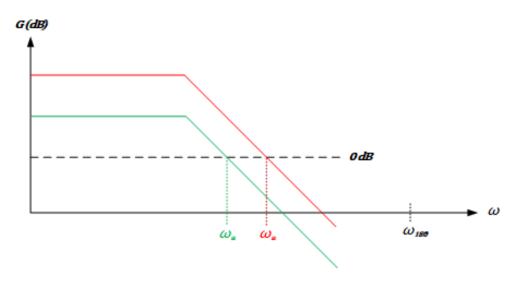


Figure 3.8 – Bode plot for gain increase.

The same concept can be extrapolated to phase margin, one looks for how much phase delay it is necessary, at 0 dB gain, to reach the -180° frequency.

Usually, designing a system in open loop is much easier than doing so in closed loop. But one would like to know if the system is stable in closed loop. Hence, the paramount question arises of how to assure that the closed loop system is stable by looking just at the open loop system.

This is where Nyquist plots are useful. Closing the loop alters the locations and amounts of the poles and zeros, one cannot assess stability looking just at the pole-zero diagram. Let's illustrate this with a simple example, for a second order system with a transfer function shown by equation 3.8.

$$T(s) = \frac{2s+1}{5s^2+4s+1}$$
(3.8)

Then, the pole-zero diagram for this open loop system can be sketched as shown in Figure 3.9 (top). However, if the loop is closed, the transfer function is modified according to equation 3.7, and one has to look for the roots of the denominator. Furthermore, the pole-zero diagram is modified as shown in Figure 3.9 (bottom).

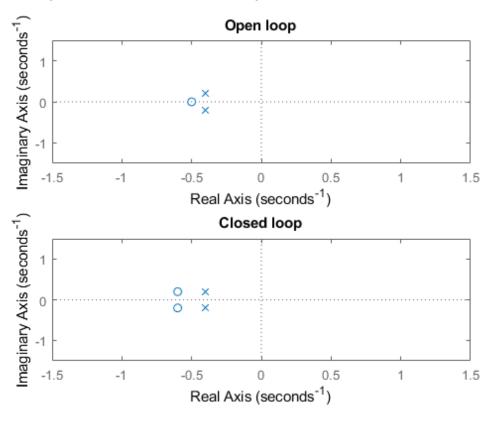


Figure 3.9 – Pole-zero diagram for equation 3.8 and its closed loop system.

This simple example illustrates the difficulty of extrapolating information about poles and zeros from an open loop system to a closed loop system using a pole-zero diagram. However, one can make a transformation to change from this diagram to one which is in the ω domain. This is, by taking points from the imaginary and real chart and fitting them

into the transfer function, one is able to plot them against an ω axis. If one feeds the entire right half plane, the Nyquist diagram is obtained.

Plotting the Nyquist diagrams for both the open loop and closed loop system, as shown in Figure 3.10, shows that by the difference between the poles of these two systems is a translation.

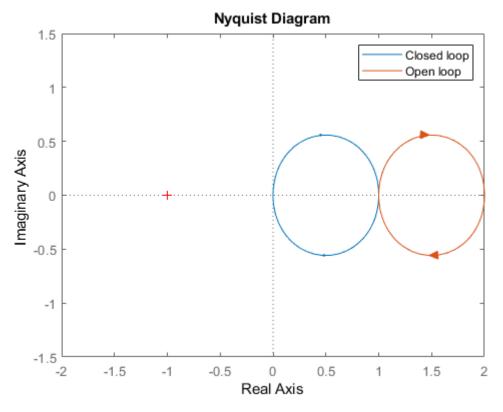


Figure 3.10 – Nyquist diagram for Figure 3.9.

Besides this, if the system happens to encircle the point (-1,0) it will be unstable. This is known as the Nyquist criterion. It is basically a statement that amounts to having poles in the right half plane.

Having this mathematical criterion is very useful, since one can now set the basis for the phase margin so that the amplifier does not become unstable in closed loop. Defining the phase margin as

Phase margin =
$$180^{\circ}$$
 + (phase at which the system gain is 0 dB)

In practice, one designs its system so that it accommodates for unknown factors. When designing an amplifier one has to account for the mismatches in fabrication, inaccuracies of the theoretical models and many other factors. Henceforth, a positive phase margin of 90° is desirable.

To conclude, an illustration on how phase margin affects the gain of the system around the unity gain value is shown in Figure 3.11 [14]. It is straightforward to see that the tighter the phase margin is, the more unstable the system will appear.

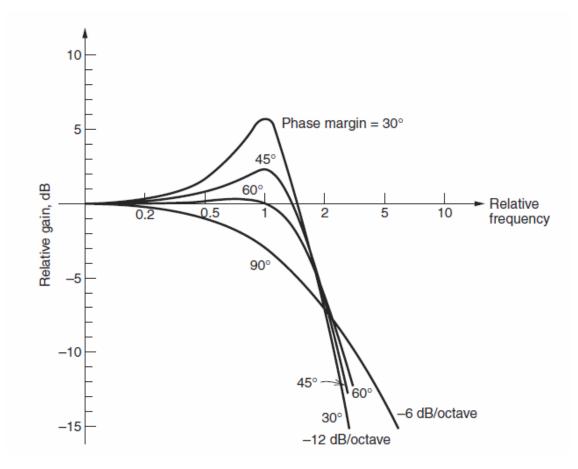


Figure 3.11 – Gain and frequency normalized plot for closed loop amplifiers with different phase margins [14].

3.2.2.-FREQUENCY COMPENSATION

When configuring an amplifier with negative feedback, one is basically trading in open loop gain to obtain a smaller closed loop gain in order to better bound parameters such as distortion, noise and temperature fluctuations, amongst other.

However, if the feedback loop phase shift is taken into account together with the internal stages delay and it amounts to 360°, the output of the device will reinforce the input creating an oscillatory behavior. This is usually measured by taking the step response of the system and measuring parameters such as the overshoot and the settling time. In order to decrease these values, frequency compensation techniques are employed.

There are several techniques that are used when designing amplifiers which try to improve the aforementioned conditions, being the most popular among them pole splitting [15].

Pole splitting refers to the introduction of a capacitor (usually known as Miller capacitance) which takes advantage of the Miller effect to effectively split the equations of the two most prominent poles in the system from each other. In this way their frequencies can be located with different capacitances.

3.3.-SINGLE STAGE TOPOLOGIES

Single stage topologies refer to systems of CMOS devices that amplify a signal in one single step. This type of topology has usually less gain than cascaded ones, being the first ones simpler to implement.

3.3.1.-FOLDED CASCODE OTA

The folded cascode amplifier is an operational transconductance amplifier (OTA) that takes advantage of the boost in output resistance from the cascoding, effectively increasing its gain over simpler current mirror amplifier topologies. It is a single stage topology, compensated using the load capacitance. The topology makes use of a differential pair (a), a cascode current mirror (b), and an additional cascode gain stage (c) as shown in the Figure 3.12 below. Note that (d) is an active load that serves as a current sink.

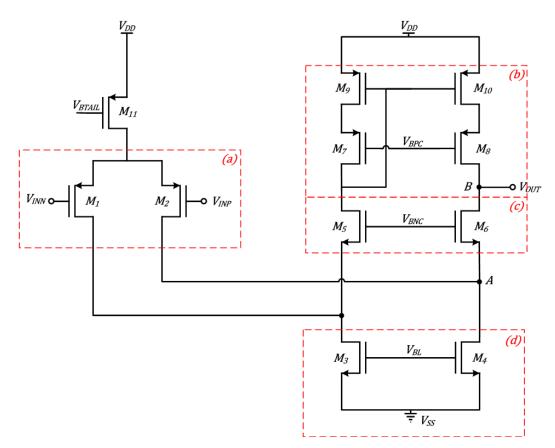


Figure 3.12 – Single-ended, folded cascode topology

The analysis of this circuit is performed under the assumption of a first order system. One can derive the nth order transfer function that describes the frequency response. However, in order for the system to perform as intended, one wants to push all of the unwanted poles and zeroes to high frequencies so that there is a steady 20 dB/dec roll-off between the 3dB-drop and the unity gain frequency. Finding the expressions for the frequency response involves small signal analysis. The small signal equivalent circuit, shown in Figure 3.13, is obtained by shorting all voltage sources to ground [16], while opening the current sources. One can neglect the parasitic capacitances in this analysis, as the load capacitance is much greater than the mentioned ones.

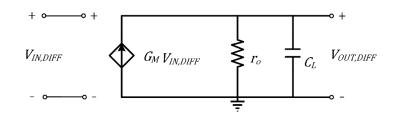


Figure 3.13 – Small signal equivalent circuit for single-ended folded cascode

The frequency response is then given by:

$$\frac{V_{OUT}}{V_{IN}}(s) = G_M \cdot Z_L = \frac{G_M \cdot R_{OUT}}{1 + sC_L R_{OUT}}$$
(3.9)

The DC gain is obtained by picking a null frequency (equivalent to a DC signal) and solving for the expression. In this case it is:

$$|A_{DC}| = G_M \cdot R_{OUT} \tag{3.10}$$

To obtain the unity-gain frequency one has to simply set the transfer function equal to one (or 0 dB), and solve for the frequency.

$$\omega_t = \frac{G_M}{C_L} \tag{3.11}$$

On the other hand, the 3 dB frequency (or the dominant pole) is given by:

$$\omega_{p1} = \frac{1}{R_{OUT}C_L} \tag{3.12}$$

This dominant pole can be derived from the topology itself by following the path of the signal through the different nodes [17]. At node *B*, the signal would encounter the output resistance of the amplifier and the parasitic capacitances from the node to ground, as well as the load impedance applied at the output. Thus, one can find the time constant and construct the pole at that node, which is described by equation 3.12.

Following this time constant discussion, it is possible to identify another pole at node *A*. However, the active load's impedance is high, so the parallel of it with the next stage, whose impedance is rather low, will result in a negligible value. This will locate the pole at the mentioned node at such high frequencies that a first order system behavior can be assumed.

One has to obtain expressions for R_{OUT} and G_M in order to construct the values of the DC gain and the dominant pole frequencies as a function of the device's parameters.

It can be assumed that $G_M \approx g_{m1}$, as the current through M5 will be nearly equal to the current in M1, analyzed from a small signal perspective. This is: the behavior of the current through M5 is dictated by M1.

In order to obtain the output resistance of the amplifier, one has to take into account previous stages [18]. First, the output resistance can be written as the parallel of the output resistance of the cascoded current mirror R_{CCM} , and the output resistance of the cascode stage R_{CSC} .

$$V_{INV} \circ (M_1 \ M_2 \ V_{INV} \ M_2 \ V_{INV} \ M_3 \ M_4 \ M_5 \ M_6 \ M_6$$

$$R_{OUT} = R_{CCM} || R_{CSC} \tag{3.13}$$

Figure 3.14 – Output resistance for a folded cascode

Furthermore, the output resistance of the cascode stage is given in the same way as the cascoded current mirror (applying the concept that drives equation 3.4):

$$R_{CSC} = g_{m5}(r_{o1}||r_{o3})r_{o5}$$
(3.14)

Applying again the concept behind equation 3.4 on the current mirror cascode, and substituting it into equation 3.13 together with equation 3.14 gives the complete expression of the output resistance of the amplifier.

$$R_{OUT} = (g_{m7}r_{o7}r_{o9})||[g_{m5}(r_{o1}||r_{o3})r_{o5}]$$
(3.15)

The slew rate is given by the output current directed through the load capacitance, and thus it is summarized in:

$$SR = \frac{I_{D3}}{C_L} \tag{3.16}$$

Finally, let's summarize all of the obtained equations into a table below.

Parameter	Expression		
$ A_{DC} $	$g_{m1}\{(g_{m7}r_{o7}r_{o9}) [g_{m5}(r_{o1} r_{o3})r_{o5}]\}$		
GxBW	$\frac{g_{m1}}{2\pi C_L}$		
ω_{p1}	$\frac{1}{g_{m1}\{(g_{m7}r_{o7}r_{o9}) [g_{m5}(r_{o1} r_{o3})r_{o5}]\}C_L}$		
SR	$\frac{I_{D3}}{C_L}$		

Table 3.1 – Single-ended, folded cascode equations.

3.3.2.-TELESCOPIC CASCODE OTA

The telescopic topology also takes advantage of a cascoding stage. In fact, it is a simpler iteration of the folded cascode. The only difference between them is the way in which the cascodes are placed, being the telescopic cascode in *series* configuration.

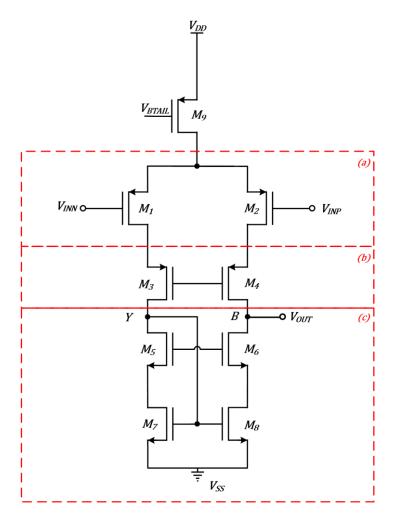


Figure 3.15 – Single-ended, telescopic cascode topology.

Shown in Figure 3.15, the topology consists of a differential pair (a), coupled to a cascode stage (b), and with a cascoded current mirror (c).

Analogously to the folded cascode OTA, the system will be assumed to be of first order. And consequently, the small signal circuit presents the same structure, with other variables. Hence, Figure 3.13 is valid for the description of the telescopic cascode OTA. Since the frequency response will be dominated by the pole at the output of the amplifier, one has to obtain the expression for this pole. The equations that describe the rough behavior of the small signal equivalent circuit have been already shown for the folded cascode OTA, namely equations ranging from 3.9 to 3.12. In order to obtain the expressions for these equations, one has to obtain R_{OUT} and G_M of the system.

The output resistance of the circuit is obtained by looking from node B, and taking the resistance to ground.

$$R_{OUT} = R_{CSC} || R_{CCM} \tag{3.17}$$

Where R_{CCM} follows the same application as in the folded cascode topology, and R_{CSC} differs from the previous example. In the folded cascode the current sink was in parallel with the differential pair, now there is no current sink and equation 3.4 for cascoding can be applied in a straight forward manner.

$$R_{CCM} = g_{m5} r_{o5} r_{o7} \tag{3.18}$$

$$R_{CSC} = g_{m3} r_{o2} r_{o4} \tag{3.19}$$

As a final step, combining equations 3.18 and 3.19 into equation 3.17 will define the output resistance at node *B*.

$$R_{OUT} = (g_{m3}r_{o2}r_{o4})||(g_{m5}r_{o5}r_{o7})$$
(3.20)

Analogously to how the equations for the folded cascode were derived, the $G_M = g_{m1}$. If one combines all this into equations 3.9 to 3.12, a similar table can be constructed as done in the previous section.

Parameter	Expression
$ A_{DC} $	$g_{m1}[(g_{m3}r_{o2}r_{o4}) (g_{m5}r_{o5}r_{o7})]$
GxBW	$\frac{g_{m1}}{2\pi C_L}$
ω_{p1}	$\frac{1}{g_{m1}[(g_{m3}r_{o2}r_{o4}) (g_{m5}r_{o5}r_{o7})]C_L}$
SR	$\frac{I_{D1}}{C_L}$

Table 3.2 - Single-ended, telescopic cascode equations.

3.4.-TWO STAGE TOPOLOGIES

Two stage architectures usually offer a higher gain with respect to single stage topologies, since the stages are cascaded the amplification is increased by a factor of the stage's gain.

Usually, the first stage is designed to present a high input impedance, while the consecutive stages to it are designed to provide the majority of the voltage gain.

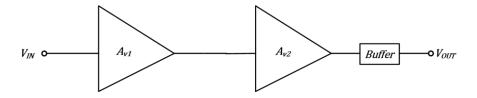


Figure 3.16 – Two stage amplifier topologies.

Furthermore, if there is need to drive a resistive load a buffer can be added, shown in Figure 3.16, with unity gain. The purpose and architecture of this buffer stage will be presented in the next section. Note that the combined gain of the amplifier will be expressed as the product of the stage's gain as shown in equation 3.21.

$$A_{\nu} = A_{\nu 1} \cdot A_{\nu 2} \tag{3.21}$$

3.4.1.-MILLER-COMPENSATED OTA

The Miller two stage amplifier that takes advantage of cascading to provide a high voltage gain. In addition, the amplifier is compensated by using the Miller effect to its advantage, cancelling the effects of a zero with a pole. It is represented in Figure 3.17.

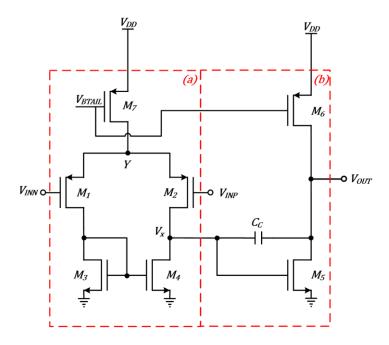


Figure 3.17 – Two stage, Miller-compensated OTA.

The two stages are separated as Figure 3.17 - (a), the first stage, and Figure 3.17 - (b), the second stage.

In order to obtain the frequency response of the circuit, one has to draw its small signal analysis. Since it is a two stage amplifier, and one wants to obtain the dominant frequency of each stage, a second order system needs to be assumed.

In order to draw the small signal equivalent system one can represent the amplifier as the two stages connected through the C_c capacitance. Therefore, one needs to represent the small signal equivalent circuits of each stage separately, connected through the capacitance between the corresponding nodes on each stage (V_x , namely).

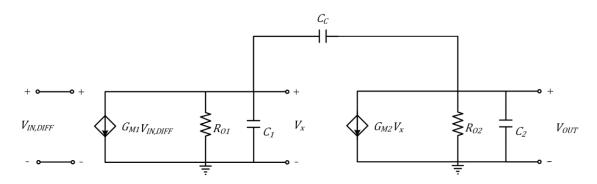


Figure 3.18 – Two stage, Miller-compensated OTA small signal equivalent circuit.

Where C_1 and C_2 are the capacitances from the corresponding output nodes of each stage to ground². The capacitance of the first stage is accounted from node V_x , being thus the parallel of the gate-drain capacitance of M₄, the gate-source capacitance of device M₆ and the gate-drain capacitance of M₂. All other involved capacitances are negligible in comparison with these three.

$$C_1 = C_{gd4} + C_{gs6} + C_{gd2} \tag{3.22}$$

On the other hand, the capacitance of the second stage is given by the parallel parasitic and load capacitances. Since the load capacitance is usually bigger by some orders of magnitude, one can assume that it is given by:

$$C_2 \cong C_L \tag{3.23}$$

Similarly, the output resistances are given for each stage from the output node of the stage to ground. For the output resistance of the first stage, one needs to account for all resistances from node V_x to ground, mainly the parallel drain-source resistances of devices M₂ and M₄.

$$R_{o1} = r_{o2} || r_{o4} \tag{3.24}$$

Analogously, the output resistance of the second stage is the parallel of the drain-source resistances of devices M_6 and M_7 .

² Note that node Y is an AC ground.

$$R_{o2} = r_{o6} || r_{o7} \tag{3.25}$$

In order to obtain the transfer function one has to apply Kirchhoff's laws on the small signal equivalent circuit. This yields:

$$\frac{V_{OUT}}{V_{IN}}(s) = \frac{G_{M1}(G_{M2} - sC_c)R_{o1}R_{o2}}{s^2 R_{o2}R_{o1}[C_1C_2 + C_c(C_1 + C_2)] + s[C_1R_{o1} + C_2R_{o2} + C_c(G_{M2}R_{o2}R_{o1} + R_{o2} + R_{o1})] + 1}$$
(3.26)

In order to obtain the DC gain, one has to pick a null frequency (as it has been done with the previous topologies). By setting s = 0, one obtains:

$$|A_{DC}| = G_{M1}G_{M2}R_{o1}R_{o2} (3.27)$$

To obtain the unity-gain frequency, it is necessary to obtain the frequency at which the gain is one. However, since this is a second order system there will be two solutions, and therefore two possible dominant poles. One has to approach the problem in a slightly different way than the one presented for the folded and telescopic cascode topologies. Solving for the zero implies equating the numerator of the transfer function to a null value, yielding:

$$\omega_Z = \frac{G_{M2}}{C_C} \tag{3.28}$$

The two poles are obtained by assuming that they are real [19], this gives a value for each pole described by equations 3.28 and 3.29. In the case of the first pole, $R_{o1} \sim R_{o2}$, and $C_1 \ll C_2$:

$$\omega_{p1} = \frac{1}{\left[C_1 R_{o1} + C_2 R_{o2} + C_C (G_{M2} R_{o2} R_{o1} + R_{o2} + R_{o1})\right]} \cong \frac{1}{G_{M2} R_{o2} R_{o1} C_C}$$
(3.29)

The other pole is located at a frequency:

$$\omega_{p2} = \frac{G_{M2}C_C}{C_1C_2 + C_c(C_1 + C_2)} \cong \frac{G_{M2}}{C_2}$$
(3.30)

Finally, in order to obtain the frequency response parameters as a function of the devices small signal model parameters, it is mandatory to obtain G_{M1} and G_{M2} . Since the first stage is basically an actively loaded differential pair, the transconductance of the stage is equal to that of the pMOS devices that compose it, $G_{M1} = g_{m1}$.

Analogously, the second stage is an actively loaded common source stage, thus being its gain defined by the output resistance times g_{m5} , and thus $G_{M2} = g_{m5}$. The unity gain frequency will be located at a higher value than the dominant pole and at a lower frequency than the zero and non-dominant pole. In order to achieve this one needs to

choose the Miller, or compensation, capacitor at a reasonable value, thus obtaining a unity gain frequency as described by equation 3.30.

$$\omega_t = \frac{g_{m1}}{C_C} \tag{3.31}$$

Finally, one can add a resistor, R_Z in series with capacitor C_C in order to push, or cancel, the non-dominant pole. When applied the expression, the transfer function will present an impedance (instead of a capacitance) at the numerator multiplying the s term. By solving analogously to equation 3.28:

$$\omega_z = \frac{1}{C_C \left(\frac{1}{g_{m5}} - R_Z\right)} \tag{3.32}$$

In order to cancel the non-dominant pole with the zero the need to present at the same frequency. The value of the nulling resistor, R_Z , at which this happens can be obtained by equating the expressions for the non-dominant pole and the zero:

$$R_Z \approx \frac{C_L + C_C}{g_{m5}C_C} \tag{3.33}$$

The system's frequency response can be summarized in a table as shown below.

Parameter	Expression	
$ A_{DC} $	$g_{m1}g_{m5}(r_{o2} r_{o4})(r_{o6} r_{o7})$	
ω _Z	$\frac{1}{C_C\left(\frac{1}{g_{m5}}-R_Z\right)}$	
ω_{p1}	$\frac{1}{g_{m5}(r_{o2} r_{o4})(r_{o6} r_{o7})C_{C}}$	
ω_{p2}	$\frac{g_{m5}}{C_L}$	
GxBW	$\frac{g_{m1}}{2\pi C_C}$	
SR [20]	$\frac{I_{D7}}{C_C}$	

Table 3.3 – Single-ended two stage, Miller compensated OTA equations.

3.4.2.-CLASS A AND A/B

It is straightforward to note that the Miller-compensated OTA is, in fact, a simple cascode amplifier with a class A second stage. The *class* of an amplifier refers to the type of amplification that is taking place at a transistor level. First, one can distinguish two types of amplification: one in which the transistors are turned on during the whole amplification time, and those in which the transistors switch constantly. However, since the first type is of the most concern for this section, only those will be revised.

As already said, the concerning group of amplifiers will present transistors which are turned on continuously during the amplification time (with respect to a reference). The time for which a transistor is turned on while operating is referred to in terms of one sinusoidal cycle, and is termed as *conduction angle*. Therefore, if a transistor is turned on for the whole duration of a sinusoidal cycle, its conduction angle is 360°. In this category of amplifiers one can distinguish four different types: class A, class B, class AB and class C amplifiers.

CLASS A

A class A amplifier is the simplest in design, being usually configured in a common actively loaded source configuration, shown in Figure 3.19 - (a). The amplifying transistor will be turned on during the whole 360° conduction angle, this is represented in Figure 3.19 - (b). The transistor will be moved around the bias point (in saturation) never stepping into the other regions of operation. The gain of this stage is given by equation 3.34. This was already seen in the previous section for the Miller-compensated two stage amplifier.

$$|A_{\nu}| = g_{m,amp}(r_{o,amp}||r_{o,act})$$

$$(3.34)$$

The impedance looking from the output of the amplifier is, in the case of both transistors, the output resistance, since the output node is connected to their respective drains. Having this in mind, one can obtain the maximum voltage swing is given by equation 3.35.

$$V_{SWING} = V_{DD} - V_{DS} \tag{3.35}$$

Note that it this this continuous operation that makes this class so linear, since there is nearly no distortion across the entire 360° of the conduction angle. However, both the negative and positive slew rates depend on the same transistor, being therefore less efficient in this aspect.

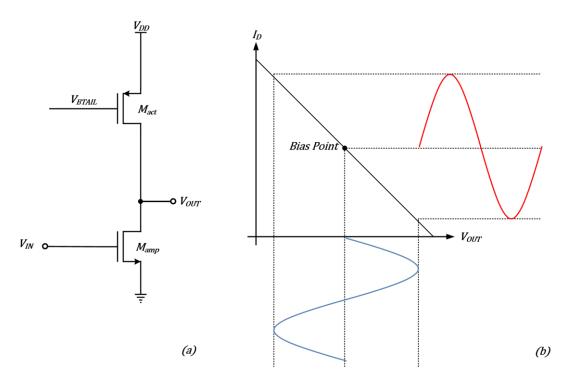


Figure 3.19 – Class A, actively loaded, common source amplifier operation.

CLASS AB

In class AB amplifiers there are two transistors which conduct a bit more than 180° of the conduction angle, shown in Figure 3.20. This is related to how the B) class amplifiers (also called push-pull topologies) operate.

In these latter ones, there are two transistors as well, but each of them are operating exactly through 180° of the conduction angle. This poses a problem when each transistors goes through the bias point and switches off in order to let the other turn on for the rest of the conduction cycle. When this point is crossed there may be place for non-linarites and, therefore, distortion appears.

The class AB amplifiers solve this by turning on each transistor for more than 180° of the conduction cycle, this is represented in Figure 3.20 - (b), overlapping in the regions of the bias point, somewhat redressing class B amplifiers' flaw.

Class AB amplifiers present similar output swing if compared to class A amplifiers. Since there are two transistors turned on, each of them will have more room until they drive towards the cut-off regions for their respective fractions of the conduction cycle.

Relating to the slew rate each transistor operates during a bit more than 180° of the conduction angle, this means that one transistor will be in charge of the positive slewing while the other transistor will be in charge of the negative slewing. In overall, this adds up to efficiency and improves the slew rate of the device.

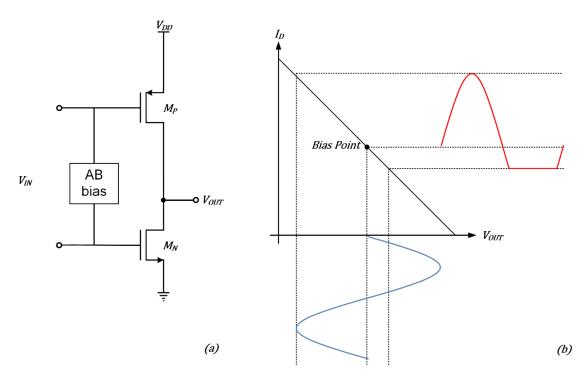


Figure 3.20 – Class AB amplifier operation.

CLASS A VS AB

When designing an output stage or an output buffer one takes into account class A and class AB stages. Both of them present advantages and disadvantages that will impact the design significantly.

Another concerning factor is the biasing that one has to provide to establish the bias point. Class A amplifiers are easy to bias, requiring usually no extra bias circuitry. In opposition to this, class AB amplifiers usually require a more complex biasing requiring of extra circuitry (separate from the bias main bias circuit). This comes due to the fact that they need to be biased so that they conduct for more than 180° of the conduction angle.

Slew rate is yet another factor of importance, since it will determine the behavior of the amplifier's output. As already mentioned, the slew rate of class A amplifiers is considerably worse than that of class AB ones. One can intuitively determine that, since in class A amplifiers only one device is providing with negative and positive slewing, it will be more limited as compared to class AB amplifiers. In these latter ones there is, virtually, a transconductance that will be in charge of the positive slewing, and another for the negative slewing.

Output voltage swing is another consideration that one has to take into account. The voltage swing of class A amplifiers is generally bigger than that of class AB ones, since they fit less transistors than the latter in the same bias voltage.

Farther from a voltage analysis, the current sourcing capabilities of these two configurations are radically different. For an increased current sourcing from an external load (modelled as a small resistance), a class A operation stage will be sourcing all the

bias current, from M_{act} in Figure 3.19, into the load. Conversely, if there is a capacitor connected as a load together with the resistor, the current would come from the capacitor into M_{amp} . This poses a great waste of current in the stage, since it is not being employed in anything useful. However, when it comes to the class AB buffer, one can think of it as one of devices turning off, and not allowing any current to be sourced through them. This poses another challenge, since the transistors must be placed in such bias point that the aforementioned takes place.

Lastly, all this information can be summarized in an organized manner, presented below in Table 3.4.

	Class A	Class AB
Voltage Swing	$V_{DD} - V_{DS}$	$V_{DD} - V_{DS}$
Slew Rate	One g_m	Two g_m
Linearity	Linearity No distortion Low distortion on bia	
Biasing	Simple	Complex, extra circuitry

Table 3.4 – Class A and Class AB comparison.

THE CLASS AB BUFFER AND ITS BIASING

Regarding the polarization of a class AB buffer, one needs to find the bias point for which one transistor, M_P in Figure 3.20, starts shutting down while the other begins to turn on, namely M_N as the input signal requires.

In a single-ended topology there are several ways described in the literature to achieve this biasing. One of the most popular ones is the Monticelli bias scheme, also termed as quiescent current biasing, involves the usage of floating current sources to generate the appropriate DC level for the class AB stage or buffer as demanded by the input signal [21]. Other methods for biasing a single-ended class AB buffer include the setup of the class AB stage into common drain configuration, and the use of error amplifiers [22].

Separately, when the amplifier topology is fully differential there is an important advantage: the differential signals are guaranteed to be 180° out of phase. Nevertheless, this will be addressed when differential topologies are discussed in the upcoming sections.

3.5.-SINGLE ENDED AND DIFFERENTIAL TOPOLOGIES

A single-ended amplifier provides an output signal which is meaningful when measured against a fixed voltage which, in the vast majority of cases, is the ground potential. This is, in essence, equivalent to probing a circuit between one potential and ground.

On the other hand, differential amplifiers provide two output signals. In contrast to the single-ended device output, a differential signal is measured between two nodes which are equal in magnitude and amplitude, but opposite phase with respect to a fixed voltage (which is, as in single-ended operation, the ground potential). There will be a central DC value shared between both differential parts of the signal, called the common mode level, or CM.

3.5.1.-FROM SINGLE-ENDED TO FULLY DIFFERENTIAL

The equations derived in sections 3.3.- and 3.4.- for the folded cascode, telescopic and two stage Miller-compensated amplifiers were obtained from the single ended topologies. It is important to realize that the only difference between single-ended circuits and their fully differential counterparts is that, in the first ones, only one branch of the circuit is used being the other symmetrical one tied through a current mirror. This is, in fact, the reason why common mode control is not needed, since copying the current ensures that they stay at the same common mode level.

Nevertheless, the path each differential signal covers is identically the same in singleended and fully differential topologies, meaning that the power consumption, bandwidth and models are exactly the same as in the single-ended case [21]. The only difference is the increase in voltage swing, as already mentioned when discussing the increase in dynamic range.

3.5.2.-INCREASE IN DYNAMIC RANGE

Differential topologies provide a larger voltage swing (also called dynamic range). It is straight forward to determine the reason why. Since each of the differential signals are equal in magnitude but opposite in phase, when taking their difference one is adding the dynamic ranges of each of the differential signals. This effectively doubles the dynamic range of the differential signal with respect of each differential input, represented in Figure 3.21 – Increase in Dynamic Range.

This also results in a higher SNR (signal-to-noise ratio), with a 3 dB increase with respect to the single-ended topologies [21], [22].

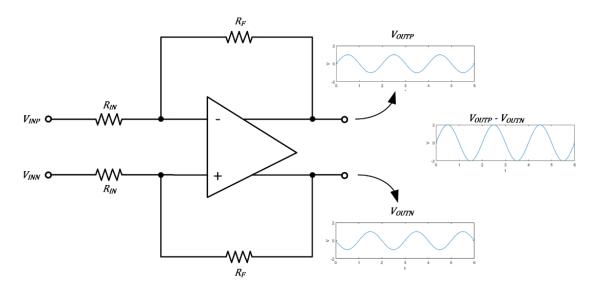


Figure 3.21 – Increase in Dynamic Range

3.5.3.-NOISE ISOLATION

The main reason behind the use of fully differential (differential in both the input and output) amplifiers is that of their relative isolation from noise. Noise can be present due to capacitive coupling between signal lines in small, power supply noise, etc... An example that allows to visualize refers to the introduction of random noise in the lines that carry the signal. On account of the two signals carrying complementary information, the difference between them will remain intact as the effects of noise cancel out. This is represented below in Figure 3.22.

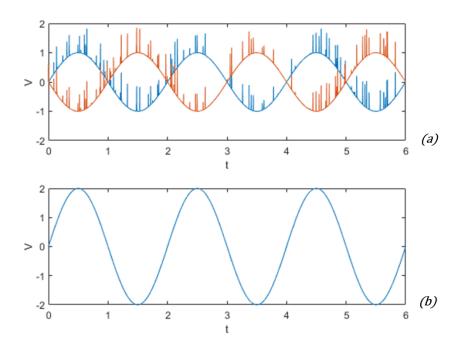


Figure 3.22 – Common mode noise cancellation, (a) noisy half-signals and (b) differential output.

This takes place under the assumption that the amplifier is balanced, and no defects take place during the fabrication process. If this supposition does not hold, then the common mode noise is different on each half-circuit, causing a differential component at the output.

On top of the aforementioned noise due to mismatches in each half-circuit, there is also random noise inherent to each half-signal. Nonetheless, this is countered through the increase in dynamic range mentioned above. Strictly speaking, the presence of random noise is not significant enough compared to the increase in the SNR ratio.

3.5.4.-BIAS POINT VARIATION

If left unattended, small variations in the common mode level will eventually drive the transistors towards the linear or cut-off regions. Even if the transistors are not driven off from the saturation region, these small changes in the common mode voltage will disturb the biasing of the MOS devices, leading to changes in the small signal gain.

These perturbations will change the node voltages, changing the drain-source currents of the involved transistors. This, in turn, affects the values of their respective transconductances, thus changing the said small-signal parameters [23].

The main culprit is the finite output impedance of the current sources used in the circuits. Take the amplifier in Figure 3.23, its common mode gain is given by equation 3.36.

$$A_{CM} = -\frac{R_D g_m}{1 + 2g_m R_S}$$
(3.36)

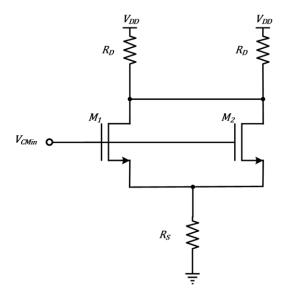


Figure 3.23 – Differential pair amplifier [24].

Intuitively, if the current source output impedance approached infinity, the common mode gain would tend towards zero. However, since this impedance cannot be infinite, the amplifier will always present a common-mode gain.

3.5.5.-CMFB CIRCUITS

The matters argued in the previous section force differential amplifiers to have circuitry, known as common mode feedback (CMFB) circuits, which set the common mode output to prevent the aforementioned effects. The CMFB network monitors the differential outputs and compares them with a voltage reference, set outside of the amplifier. By taking the difference and multiplying by a gain, one can regulate the bias point of the system to set a common CM voltage in both outputs, regardless of the common mode noise.

The goal of the CMFB circuit is to set the sensed common mode voltage to the level of the applied reference, making their difference effectively zero. In intermediate steps, the result of this difference has to be modified and applied to the amplifier so that one can apply negative feedback. Hence, for any differential amplifier, the common block diagram of the CMFB network consists of a CM detector, a compare (subtraction) stage, an amplification stage, and the implementation stage. It is this last block which varies in concept from one topology to another.

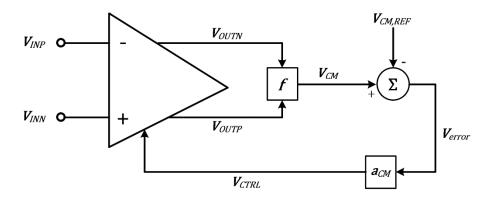


Figure 3.24 – Stages of a CMFB circuit.

The block diagram of such CMFB circuit is shown in Figure 3.24. First, a block *f* senses the common mode of the amplifier's output. This is then compared to an outside reference (note that the comparison is a difference between the two potentials), what will provide the drift from the wanted common mode level. Finally, the obtained error voltage has to be multiplied by a common mode gain, obtaining V_{CTRL} , and negatively fed back into the amplifier to adjust for the difference. Note that, since the CMFB network feeds to the amplifier, every common mode compensation network will present a different expression for its total gain, A_{CM} , depending on the applied topology. Furthermore, this total A_{CM} gain differs from the CMFB network gain, a_{CM} , but it is a function of it due to negative feedback, $A_{CM} = a_{CM}a_{CM,int}$.

Regarding the stability of the system, the CMFB network presents its own frequency response since it is an amplifier itself. Likewise, the CMFB loop can be compensated with the output load capacitance being its unity gain frequency defined as:

$$GxBW_{CM} = \frac{A_{CM}}{2\pi C_L} = \frac{a_{CM}a_{CM,int}}{2\pi C_L}$$
(3.37)

In order to compensate the CMFB loop without inquiring into the phase margin of the differential path, the gain of the CMFB amplifier has to fulfill that $a_{CM} \leq 1$ [25].

On top of this, one needs to guarantee that the CMFB can keep up with the DM operation for all the $GxBW_{DM}$, meaning that its unity gain frequency will have to be higher than the latter one:

$$GxBW_{CM} > GxBW_{CM} \tag{3.38}$$

The implied by the equation above is usually met at once, since the CMFB loop does not move the same amount of current as the DM one, the transistors are smaller and, therefore, faster.

Having the previous in mind, it is only necessary to keep the gain of the CMFB stage below 0 dB so that the system remains stable.

There are several ways to implement the described CMFB, being the most popular ones the resistor/capacitor network and the differential pair CMFB systems, for continuous-time amplifiers in $\Sigma\Delta$ ADCs.

RESISTOR-CAPACITOR NETWORK CMFB

In this case the f block, or CM sense block, is implemented through a voltage divider, shown in Figure 3.25. Note that in this case the sensed voltage is simply:

$$V_{CM} = \frac{V_{OUTN} + V_{OUTP}}{2}$$
(3.39)

This voltage is the one that will be used in later stages of comparison and amplification. Note that this modifies the frequency response of the CMFB loo by means of the resistor having the effect of introducing a pole and the capacitors a RHP zero [26]. There are two ways to mitigate this effect; on the one hand one can increase the value of the resistance in order to push the zero to high frequencies, or by means of choosing a value for the capacitors such that they cancel the effect of the pole at the given frequencies.

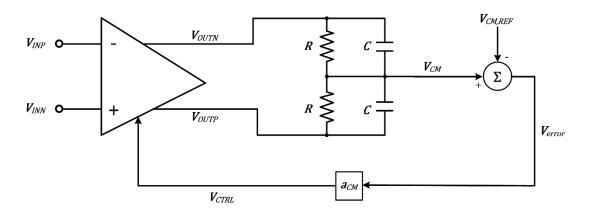


Figure 3.25 – CMFB loop implemented using a resistor and capacitor CM sense block.

The comparison and amplification of the signal can be implemented in a single block as depicted in Figure 3.26.

Figure 3.26 - CMFB loop implemented using a resistor and capacitor CM sense block together with the comparison and amplification blocks.

The signal mean is fed into the input of a differential pair, effectively subtracting the common mode reference from it. Note that the signal which is applied back to the amplifier through negative feedback can be analytically derived as:

$$V_{CTRL} = a_{CM} (V_{CM} - V_{CM,REF}) + V_{GS3}$$
(3.40)

DIFFERENTIAL PAIR CMFB

The system is shown in Figure 3.27 represents an implementation of a differential pair CMFB network.

Assuming that both V_{OUTN} and V_{OUTP} have the same CM level one can defined the currents through M₂ and M₃ are simply half of the bias current each, since the differences in both signals cancel out. This will translate into the current through the active load not deviating from the bias value, set by the current mirror.

On the other hand, if there are small differences between these two, a current deviation will from the reference value will appear. The current through M_6 and M_7 will be respectively:

$$-I_{D6,7} = \frac{I_{REF}}{2} + g_{m3,4} \frac{\left(V_{OUTP,N} - V_{CM,REF}\right)}{2}$$
(3.41)

(Note that current is negative since the used transistors are p-channel MOS devices). These two currents join at the M_8 active load, effectively adding together. Therefore, the

voltage that appears across the active load will be given as a function of the drain-source current that crosses it.

$$I_{CTRL} = I_{REF} + g_{m3,4} \left(\frac{V_{OUTP} + V_{OUTN}}{2} - V_{CM,REF} \right) = I_{REF} + g_{m3,4} \left(V_{CM} - V_{CM,REF} \right)$$
(3.42)

Analogously to the expression for V_{CTRL} in the capacitor-resistor CMFB network, the voltage that will appear due to the current mismatch will be the one given by the active load's voltage drop due to the bias current, and due to the deviation.

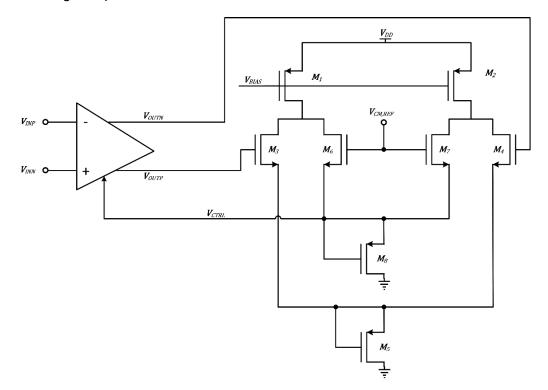


Figure 3.27 – CMFB loop implemented using a differential pair CM sense block.

3.5.6.-FULLY DIFFERENTIAL CLASS AB IMPLEMENTATION

As already discussed in previous sections, when implementing a fully differential class AB output stage or buffer the differential nature of the signals poses a great advantage. By being both differential signals 180° out of phase, one can be ensured that when one of the signals is rising, the other one is falling. This allows to save the bias circuit of the class AB as depicted on next page, in Figure 3.28.

In a quiescent state, the outputs of the previous stage are at the common mode level, being the current in both branches equal. Note that going through the path of $V_{OUTP,S1}$ there is an nMOS transistor which is part of a current mirror (M_{P1} and M_{POUT}). Having this in mind, it is possible to see that the current through M_{N1} will be copied to M_{NOUT} through this very same trail.

However, when $V_{OUTP,S1}$ is rising, $V_{OUTN,S1}$ is falling. This means that the current through M_{N1} will start to increase (differing from the bias one). While this is taking place, transistor M_{NOUT} will present a decrease in voltage at its input, turning it off eventually.

On the other hand, when $V_{OUTP,S1}$ is falling, $V_{OUTN,S1}$ is rising. This translates into the voltage at the gate of M_{N1} decreasing, causing it to turn off at some point. If this takes place, M_{POUT} shuts off and the stage remains in a state where it can put no current into the output.

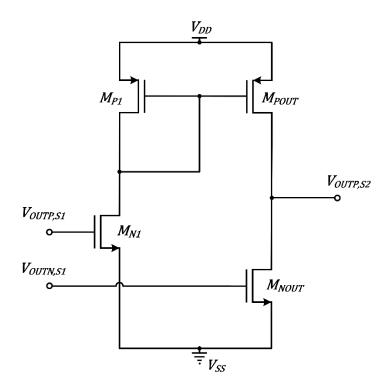


Figure 3.28 – Fully differential class AB output stage half circuit.

Nonetheless, there is a major problem with this circuit. As it relies on the common mode voltage level to set the quiescent state current, it is fundamentally important to monitor the CM at the output of the first stage, this adding to the complexity of the CMFB network that one has to implement in the final system. As a final note, the circuit presented in Figure 3.28 is only one half circuit, being the other one symmetrical to the presented one.

3.6.-COMPARISON

In this section a summary of the advantages and disadvantages of the analyzed topologies will be presented.

3.6.1.-FOLDED CASCODE OTA

Starting with the folded cascode OTA, it is a single-stage and easy to bias circuit. It is load compensated, meaning that the frequency compensation is achieved through the value of the load capacitance and, therefore, it is suited for capacitive loads. Another buffer stage will be needed in order to drive a resistive load.

The output voltage swing of each half circuit is given by the bias voltage minus the voltage drops across the transistors located at the output. Hence, the total output swing is twice this amount due to the increase in dynamic range of fully differential amplifiers.

$$V_{SWING} \propto 2 \left(V_{DD} - 4 V_{DS,sat} \right) \tag{3.43}$$

Having this output voltage swing is not desirable as when decreasing V_{DD} the threshold available voltage needed to keep further transistors in saturation decreases.

The gain is, however, not as high as with a two stage amplifier. It has a high single stage gain, but it is expectably lower than a multiple stage amplifier. The common mode control is added simply on the cascoded current mirror, being thus easy to implement as well.

Finally, as discussed in section 3.3, there is a non-dominant pole at node A (where the folding takes place). This pole, although assumed to be far away in frequency so that it does not affect the phase margin, is at a lower value than other non-dominant poles from other topologies such as the telescopic cascode OTA. It can be compared to the unity gain frequency of transistor M_5 , from equation 2.14, it is given by:

$$f_{T5} = \frac{g_{m5}}{2\pi C_{gs5}} \tag{3.44}$$

The non-dominant pole is defined by equation 3.31:

$$f_{p2} = \frac{g_{m5}}{2\pi C_{NA}}$$
(3.45)

Where C_N is the parallel of all parasitic capacitances from node A to ground. It can be expressed as $C_{NA} = C_{gs5} + C_{db1} + C_{db3}$ and, since they are all in the same order of magnitude, one can further express it as $C_N \approx 3C_{gs5}$. This will result in the non-dominant pole frequency being one third of the cut off frequency of device M₅, as described by equation 3.46.

$$f_{p2} \approx \frac{f_{T5}}{3}$$
 (3.46)

3.6.2.-TELESCOPIC CASCODE OTA

The telescopic cascode amplifier is very similar in characteristics to the folded cascode. However, since the output voltage swing is now smaller, due to one extra transistor in the output design. This voltage swing , analogously to the folded cascode, is proportional to V_{DD} minus the voltage drop across the mentioned transistors.

$$V_{SWING} \propto 2 \left(V_{DD} - 5 V_{DS,sat} \right) \tag{3.47}$$

The voltage swing is smaller if compared to the folded cascode, since one transistor more has to be taken into account. Hence, the amplifier will perform even worse in low-voltage applications.

It is a load compensated amplifier as well, this will mean, analogously to the folded cascode, that the amplifier will be suited to drive capacitive loads. In order to supply a resistive load an output buffer will be needed. In addition to this, the gain in similarity to the one of the folded cascode. The common mode can be set similarly by controlling the bias voltage of the cascoded current mirror.

Regarding stability, since node A does not exist in a telescopic cascode (its equivalent is, in fact, a much lower impedance node than the one in the folded cascode one), there will be less concerns as to a reduced phase margin, being in this sense superior to the folded cascode topology. This non-dominant pole is located at node Y, and it can be derived in an analogous way as it has been done for the folded cascode topology. However, now the capacitance from node Y to ground is given by the parallel of two parasitic capacitances, instead of three as it was the case of the folded cascode. This effectively locates the non-dominant pole at half of M_5 cut-off frequency.

$$f_{p2} \approx \frac{f_{T5}}{2} \tag{3.48}$$

3.6.3.-TWO STAGE MILLER-COMPESNATED OTA

This topology takes advantage of stage cascading, being superior in gain to the single stage topologies. However, each single stage will have a lower gain than a folded or telescopic cascode amplifiers alone. The voltage swing is greater in this topology, making it better suited to low voltage technologies. Analogously to the other two topologies, the voltage swing is given by:

$$V_{SWING} \propto 2 \left(V_{DD} - 3 V_{DS,sat} \right) \tag{3.49}$$

It is a Miller-compensated amplifier with a class A output. This means that the frequency compensation takes advantage of the Miller effect in order to force the amplifier to behave as a first order system. The class A output allows for an easy CM control, in a similar way to single-stage amplifiers. Having two stages it will be able to drive a resistive load (by sacrificing the gain of the second stage).

3.6.4.-TWO STAGE CLASS AB OTA

Any two stage amplifier with a class AB output stage will present a similar gain as a two stage Miller-compensated OTA, but differs depending on the used stage topology. However, the main disadvantage is the complicated biasing which, in turn, will result in a much more complicated CM control.

The output voltage swing of the AB stage is similar to that of the two stage Millercompensated amplifier, since they present a similar output structure.

It will have to be Miller-compensated as well, hence stability will not be such a paramount issue as with the folded and telescopic cascode amplifiers.

	Folded Cascode	Telescopic	Two stage Miller-	Two stage class
		Cascode	compensated	AB
Gain	$G_{M1}R_o$	$G_{M1}R_o$	$G_{M1}G_{M2}R_{o1}R_{o2}$	$G_{M1}G_{M2}R_{o1}R_{o2}$
Voltage Swing	$2(V_{DD}-4V_{DS,sat})$	$2(V_{DD}-5V_{DS,sat})$	$2(V_{DD}-3V_{DS,sat})$	$2(V_{DD}-3V_{DS,sat})$
Compensation	Load capacitance	Load capacitance	Miller-compensation	Miller-compensation
CMFB	Simple	Simple	Simple	Complex
Bias	Intermediate	Intermediate	Simple	Complex
Stability	$f_{p2} \approx \frac{f_{T5}}{3}$	$f_{p2} \approx \frac{f_{T5}}{2}$	$f_{p2} \approx \frac{g_{m2}}{2\pi C_L}$	_
Current	4 <i>I</i> _{diff}	2I _{diff}	$(2+n)I_{diff}$	_
consumption			<i>n</i> = 1, 2	

Table 3.5 – Comparison of different amplifier topologies.

4.-DESIGNED CIRCUIT

Once all of the theoretical fundamentals have been reviewed, the design equations have to be posed. These statements will relate the lengths, widths as well as voltages of the system devices with the amplifier parameters.

The steps undertaken in analog design start usually by determining the current that has to pass through the system, which is usually coupled to the slew rate of the system. Once the current is known, one has to determine the transconductance parameters of the devices that make up the system. In turn, the transconductance is given by the GxBW of the system (for both two and single stage amplifiers). When the transconductance is known one has to determine the effective voltage, the dimensions and currents through the device in order to achieve that transconductance.

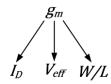


Figure 4.1 – Transconductance dependencies.

However, the drain current, effective voltage and dimensions ratio ultimately depend on the current as well. This is one of the main reasons behind the use of simulation software in microelectronic design, since systems become almost impossible to plan at transistor level with so many dependencies.

4.1.-DESIGN EQUATIONS

A common step to all of the designed circuits will be to set the effective voltage to a certain value. Since it affects the transition (or cut off) frequency of the MOS devices it needs to be set beforehand. Following equation 2.9, if a $V_{eff} = 115 \text{ mV}$ is set, the maximum transition frequency (at lowest channel length) is predicted in the range of the 10ths of GHz for n-transistors and in the range of GHz for p-transistors, at minimal channel length. Note that, in order to obtain a value for the unity gain frequency it is necessary to know C_{ox} . This, in turn, can be computed by simulating a transistor and obtaining values for C_{gs} . Applying equation 2.11, one can solve for C_{ox} , provided with the dimensions of the simulated device.

In order to estimate the device ratio, a value for μC_{ox} needs to be known. Since this value varies for each technology, it is mandatory to obtain them by means of a simulation. This can be done be running a simulation of both a pMOS and nMOS device, with known dimensions. The μC_{ox} parameter is then computed in MATLAB using equation 4.1.

$$k = \mu C_{ox} = 2I_D \frac{L}{W \cdot V_{eff}^2}$$
(4.1)

The results are represented in Table 4.1.

$k_p (A/V^2)$	$1.9 \cdot 10^{-4}$
$\sigma_p (A/V^2)$	$1.6 \cdot 10^{-5}$
$k_n \left(A/V^2 \right)$	$3.6 \cdot 10^{-4}$
$\sigma_n (A/V^2)$	$1.5 \cdot 10^{-4}$

Table 4.1 – Values of k for the 180 nm pMOS and nMOS devices.

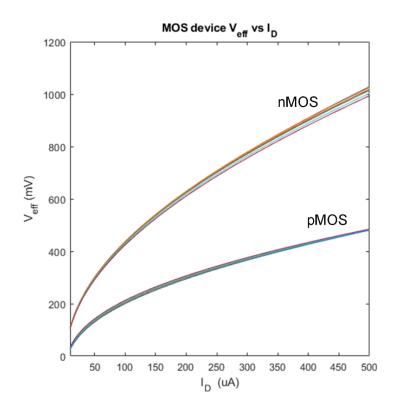


Figure 4.2 - V_{eff} vs I_D .

Another useful design equations relates the transconductance to the current that flows through the device, one can relate them easily. This is done by solving for the effective voltage, V_{eff} , in equation 2.6 and substituting it in equation 2.5.

$$g_m = \sqrt{2I_D \mu C_{ox} \frac{W}{L}} = \sqrt{2I_D k \frac{W}{L}}$$
(4.2)

One can solve for the ratio of dimensions from the equation above, once the transconductance is known, the dimensions of the device will be set according to equation 4.3.

$$\frac{W}{L} = \frac{g_m^2}{2I_D k} = \frac{2I_D}{kV_{eff}^2}$$
(4.3)

4.1.1.-FOLDED CASCODE

The first step is to determine the current through each of the devices. In order to account this, one needs to impose the criteria for the GxBW and SR, from Table 3.1:

$$GxBW = \frac{g_{m1}}{2\pi C_L} \to g_{m1} = 2\pi C_L \cdot GxBW$$
(4.4)

$$SR = \frac{I_{D3}}{C_L} \to I_{D3} = SR \cdot C_L \tag{4.5}$$

As the slew rate has is imposed to be $15 V/\mu s$, and the load capacitance is 20 pF, then the current through device M_3 has to be, applying equation 4.5, at least $I_{D3} = 300 \mu A$.

The current ratio in the amplifier will be 1:1 in order to avoid complex scaling. Thus, the current through devices M_1 and M_2 has to be half that of M_3 , namely: $I_{D1} = I_{D2} = \frac{I_{D3}}{2}$.

The dimensions for devices M_1 and M_2 are the same and given by substituting the expression in equation 4.4 into 4.3, yielding:

$$\left(\frac{W}{L}\right)_{1,2} = \frac{4\pi^2 (C_L \cdot GxBW)^2}{2I_{D1,2}k_p}$$
(4.6)

Imposing a load capacitance of 20 *pF* and a GxBW = 300 MHz on equation 4.4 will reveal that the necessary transconductance for the differential pair is $g_{m1} = 37.7 mA/V$. Substituting this value in equation 4.3 will result in the dimension ratio for the pMOS differential pair, being this $W/L \approx 12500$.

In order to accomplish this, the needed effective voltage, V_{eff} , would be so low that the device would be driven into the subthreshold region. Although operation in this region is possible, the models derived in previous sections predict the behavior in the saturation region. This is a very extreme case, meaning that the topology will not be able to drive such a high capacitive load unless the currents are scaled up to the range of mA.

Moreover, scaling up current or the dimensions will translate into an increase in transconductance only up to a certain point. Hence, it is not viable to try scaling the amplifier's current and dimensions to achieve such high transconductances. On top of all this, an increase in transconductance will translate in a drop of output resistance, which will trade away for the gain of the device. All this contributes to the transconductance increase stagnating regardless of how much drain current and dimensions scale up.

Several transistors will be arranged in parallel, each of them carrying a $10 \ \mu A$ with a fraction of the transconductance calculated above. This will allow for a much easier design of the single individual transistors and a later current upscaling.

With this in mind, for a pMOS device, the ratio of dimensions is computed using equation 4.3 to be $(W/L)_p \sim 20$, and $(W/L)_n \sim 9$ for nMOS devices. Nevertheless, in order to

obtain an n-transistor with the same effective voltage, a simulation was run for the ndevices and dimensions were obtained at which the effective voltage was the same as the pMOS devices. The chosen transistors had dimensions $(18.8/1)_p$ and $(0.88/0.2)_n$, both around the 115 mV specified effective voltage, shown in Figure 4.3.

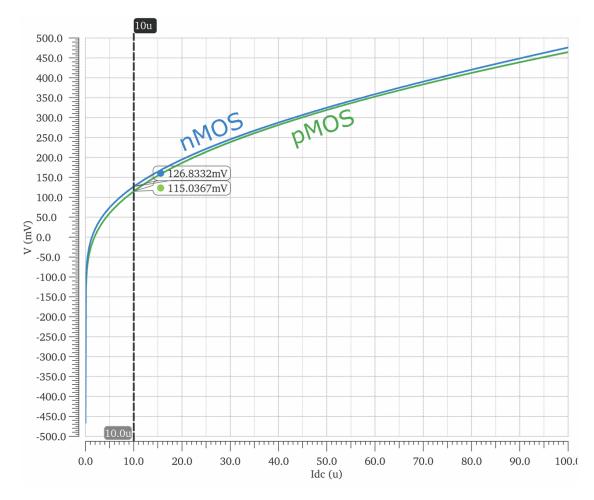


Figure 4.3 – Results for effective voltage at different drain-source currents, I_{dc}*, for both nMOS and pMOS devices with the dimensions reported above.*

The transistors were designed equally for the entire system, while scaling them for different currents in the different stages of the system. The transistors at the differential pair were an exception, as these were maximized for transconductance since it determines the GxBW of the amplifier.

The first approach to the design was to build a single-ended system and obtain a stable operating point, this system is exactly the same as the one portrayed in Figure 3.12. The system was designed by adding current sources first, in place of current mirrors to avoid complexity in the design. Once a successful bias point for the system was established, the current sources were replaced with current mirrors and the corresponding biasing circuit.

Once the operating point of the system was verified to be correct, the fully differential system with a virtual CMFB was designed. The virtual CMFB consisted on blocks that compared the CM level of the outputs to an outside reference, and then amplified by a CM gain. Once the system was verified to perform as desired, the virtual CMFB circuit was replaced by a differential pair CMFB network. The final fully differential circuit is shown in Figure 4.4. Finally, the system was optimized to meet the specifications in Table 1.1.

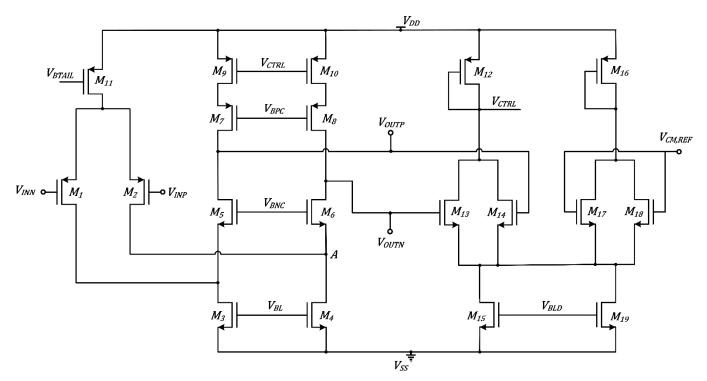


Figure 4.4 – Fully differential, folded cascode OTA with differential pair CMFB network.

The bias circuit was designed to generate V_{BL} , V_{BLD} , V_{BPC} , V_{BTAIL} and V_{BNC} . A bias circuit sets the respective bias voltages by using a network of active loads that emulate the ideal voltage drop across the transistors. In the case of the circuit depicted in Figure 4.4 the bias circuit requires three different branches for the different potentials that one has to apply across it, being shown in Figure 4.5.

Since the reference current was provided to be $100 \ \mu A$ and the base transistors for $10 \ \mu A$, devices M_{B1} and M_{B2} were designed with a multiplicity of 10 (this is: ten transistors in parallel) so that contiguous branches provide the desired $10 \ \mu A$ reference voltage.

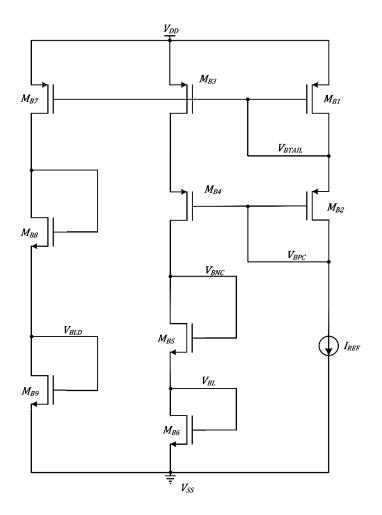


Figure 4.5 – Biasing circuit for fully differential, folded cascode OTA.

4.1.2.-TWO STAGE, CLASS A MILLER

In order to design the two stage, Miller-compensated OTA, the same design procedure as the one used for the folded cascode was followed. In order to compare them, the same transistors and CMFB circuit was used as with the previously designed amplifier.

Applying the criteria for the *GxBW* and *SR* into equations from Table 3.3:

$$GxBW = \frac{g_{m1}}{2\pi C_c} \to g_{m1} = 2\pi C_c \cdot GxBW$$
(4.7)

$$SR = \frac{I_{D7}}{C_C} \to I_{D7} = SR \cdot C_C \tag{4.8}$$

The same equations apply to the differential pair of the two stage, Miller-compensated OTA as to the folded cascode. The load capacitance of $C_L = 20 \ pF$ and a Miller capacitor of $C_c = 3 \ pF$ were chosen. In order to reach a $GxBW = 300 \ MHz$ the transconductance of the differential pair needs to be, at least, $g_{m1} = 5.64 \ mA/V$ as predicted by equation

4.7. Furthermore, for a slew rate of $SR = 15 V/\mu s$, the current through device M_7 should be, at least, $I_{D7} = 45 \mu A$ as predicted by equation 4.8.

The same strategy was adopted as with the folded cascode; this is: the same ratios for the devices were chosen in order to upscale the current by placing transistors in parallel, and the differential pair was also maximized for transconductance. However, the fact that the slew rate is determined by the current through device M_7 means that it cannot be readily optimized in later iterations of the design, as it is tied to the transconductance of the differential pair. The designed circuit can be seen in Figure 4.6 below.

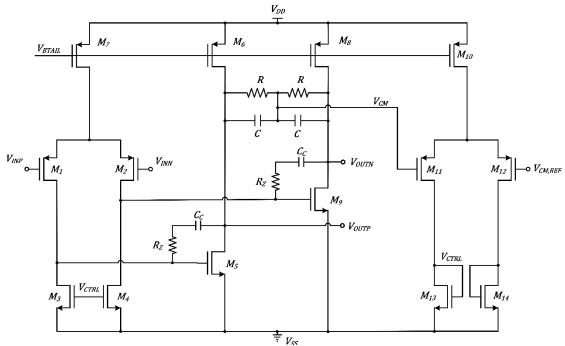


Figure 4.6 – Fully differential, two stage, Miller-compensated OTA with differential pair CMFB network.

Analogously to the folded cascode OTA, one needs a bias circuit in order to set the value of V_{BTAIL} . This is achieved by a simple, one branch bias circuit, consisting on a simple current mirror. As with the previous design, the current reference was $100 \,\mu A$, so the multiplicity of device M_{B1} was set to 10, in order to provide the needed $10 \,\mu A$ reference.

Finally, the value for R_Z was computed after the first DC run, once the transconductance of the corresponding device was known.

4.1.3.-TWO STAGE, CLASS AB MILLER

After obtaining the results for the above-described topologies, a two-stage, class AB Miller-compensated OTA that effectively attained to the specifications was designed in collaboration with the PhD students.

This two-stage amplifier needed a high first-stage gain, so cascoded topologies such as the folded and telescopic cascode were picked as candidates. In order to provide an effective way to reduce the power consumption, a class AB was suited as the second stage.

The amplifier was designed in four blocks, with the first stage, the CMFB circuit of the first stage, as well as the class AB output stage with its output CMFB loop. The first stage consisted on a telescopic amplifier in order to obtain a high gain, as already mentioned. Consequently, a differential pair (DP) CMFB network was used to monitor the CM voltage level at the input of the second stage. The class AB was biased as described in section 3.5.6, with a RC CMFB network to adjust the output common mode voltage.

The designed circuit is presented in Figure 4.7, note that the DP CMFB circuit is shown as a block and that the output CMFB as well as bias circuits are missing.

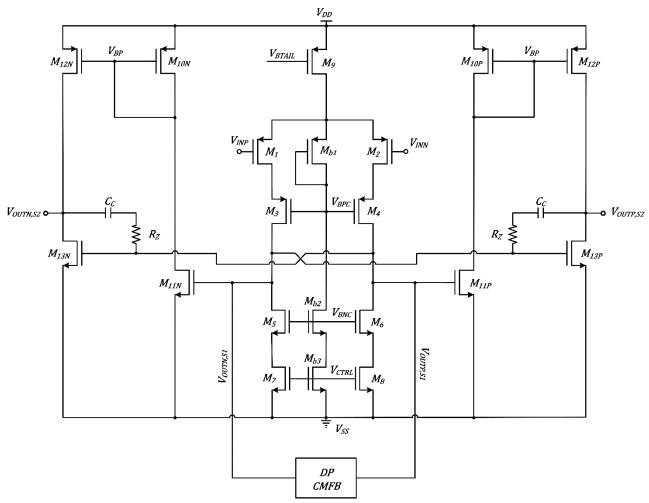


Figure 4.7 – Two stage, Miller-compensated, class AB amplifier with DP CMFB block and no output CMFB network.

The DP CMFB was the same one depicted in Figure 3.27, the complementary version of the one used in the folded cascode topology, since the control point are nMOS transistors. Another reason to use this CMFB circuit is that, as already discussed, the DP CMFB does not load the first stage, allowing for the correct implementation of Miller-compensation through C_C and R_Z . The reduced input dynamic range will not influence the stability of the system since the variations at the output of the first stage are assumed to me low enough that they fit into the dynamic range of the DP.

The only major difference consisted on the reference voltage, which was set through the reference current, as it cannot be set to the 750 mV. It will be, in fact, different. The implementation of this stage is portrayed in below. Furthermore, a wide swing cascode was introduced with respect to the CMFB shown in Figure 3.27, in order to emulate the voltages closely. The implemented DP CMFB circuit is shown in Figure 4.8.

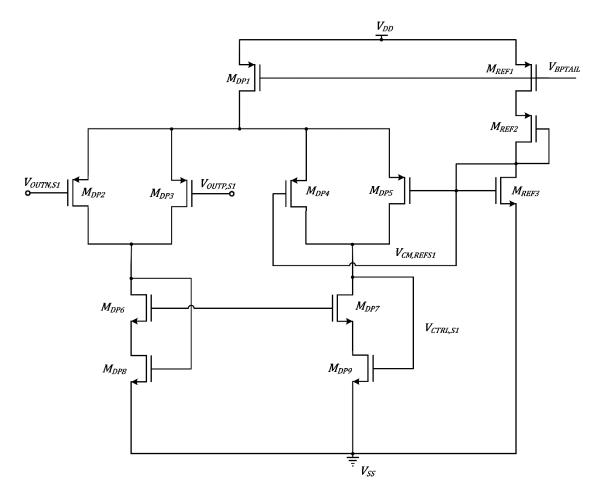


Figure 4.8 – Differential pair (DP) CMFB circuit used for the first stage with the reference generation circuit.

The next challenge was to implement the output stage CMFB circuit. Since it is the output of the amplifier, the common mode variations will be bigger than those in the output of the first stage. This implies that a DP CMFB might not be stable, as the variations in the CM may exceed its input range. Since loading is no longer such a paramount issue as in the first stage, an RC CMFB can be used.

The RC network has to be implemented so that it senses the common mode, outside of the class AB stage. A strategy to achieve this is by separating both M_{10P} and M_{10N} into two different transistors. This will allow to sense the CM voltage and feed it into the next stages of the CMFB, obtaining $V_{CTRL,S2}$. The circuit is depicted in Figure 4.9.

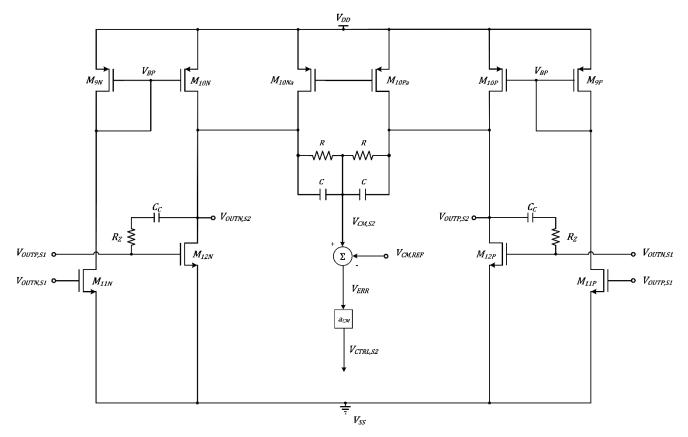


Figure 4.9 – Diagram implementation of the RC CMFB circuit for a fully differential class AB.

The bias circuit of the system consists of five different branches, the first branch generates the bias voltage for the pMOS cascodes are formed by devices M_3 and M_4 . It is embedded in the amplifier branch in order to obtain a better voltage reference.

The three other branches (represented in Figure 4.10, next page) generate the tail bias voltage, V_{BTAIL} (namely through transistors M_{B1} and M_{B4} . From there on the currents are copied to next branches, providing the bias voltages for the wide swing cascode formed by devices M_6 , M_7 , M_8 and M_9 (through devices M_{B7} and M_{B8}), as well as for the embedded bias circuit devices. The fourth branch provides the bias voltage for the RC CMFB circuit, V_{BLD} , through device M_{B9} .

Finally, the fifth branch provides the bias voltage is the one providing the common mode voltage reference for the CMFB circuit that monitors the output of the first stage, represented in Figure 4.8.

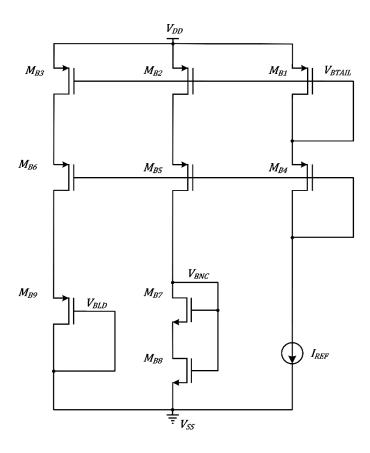


Figure 4.10 – Bias circuit branches for the two stage, Miller-compensated, class AB amplifier.

4.2.-SIMULATIONS

There is a set of five simulations that have been performed to evaluate the performance of each system. The approach to the simulation was gradual, each change in the design was based on the results of the previous analysis type.

Each simulation is coupled to two cells which consist of schematics files. These two cells are termed schematic (SCH) and test bench (TB) cells. On top of this, each cell will present several views, the SCH cell will present the schematic view of the amplifier itself and a symbol view. On the other hand, the TB cell will consist on the schematic cell of the test bench that uses the amplifier from the SCH cell through the symbol view and a ADE L, or XL, simulation view that contains all the simulation parameters and results.

4.2.1.-DC ANALYSIS

The DC analysis is a tool that allows to find the operating point of the analyzed circuit. Since the MOSFET devices present models with non-linear equations, it is a difficult task to undertake this procedure by hand. As such, numerical methods are employed by software like SPICE or Spectre that allow to come up with an operating point solution in a relatively short amount of time.

The operating point analysis is the first step of analog design, since it provides with the information of the steady state of the system upon which further design decisions will be taken. If the DC analysis is erroneous, the consequent steps in the design will be flawed as well.

The behavior of the circuit with respect to time can be described in terms of its current i(v(t)) and charge variation over time $\frac{d}{dt}q(v(t))$, which depend on voltage and charge variations over time, and a given input u(t) [27]. Since DC analysis involves no variations over time, the derivatives can be assumed to be zero, and the mathematical description of the model equation becomes the one in shown in equation 4.9.

$$i(v_{DC}) + u_{DC} = 0 (4.9)$$

The methods used in pursue of the solution are usually based on the Newton-Raphson method. In addition to this, one has to be aware that DC analysis is based on the large signal behavior of the circuit.

4.2.2.-AC ANALYSIS

Once the operating point has been set correctly in the DC analysis, the second step is usually an AC study. This kind of evaluation involves all sorts of frequency domain tests based on estimating the frequency response on the system through phasor-based analysis [28].

This phasor-based analysis is performed on the small-signal model of the circuit which is constructed in the first step of the simulation. Then the model is linearized around the operating point (this is the solution obtained in the DC analysis) using Taylor expansions and solving for the system using phasor-based analysis (shown in equation 4.10).

$$(G+j\omega C)v = -u \tag{4.10}$$

Where G and C are the conductance and capacitance of the small signal model, respectively. Note that since AC analysis disregards any dependencies with time, some effects are not modelled and one cannot asses the stability of the system.

For amplifier design the AC analysis arises as a very useful tool to approach the frequency response through the gain magnitude and phase outputs in a Bode plot.

4.2.3.-TRANSIENT ANALYSIS

Transient analysis simulates the circuit time behavior over a pre-defined interval, meaning that it is an unsteady state assay unlike the steady state (time invariant) DC and AC analysis.

Furthermore, AC and DC analyses relied on solving algebraic equations given by Kirchhoff's laws (DC) and phasor analysis in the complex plane (AD). Transient analysis deals with differential equations, the proposed model reads as shown by equation 4.11.

$$i(v(t)) + \frac{d}{dt}q(v(t)) + u(t) = 0$$
(4.11)

It is very useful to determine time-related parameters such as stability or slew rate, in the case of amplifiers.

4.2.4.-CORNERS AND MONTE CARLO ANALYSES

Once the DC, AC and transient analyses are passed accordingly for a designed system, one needs to evaluate statistically how will the fabricated circuit perform. These two types of analysis belong to the family of experimental design (DOE), which intent to model the experimental variations of the design by controlling a set of predefined variation parameters. In the case of microelectronic design, one wants to predict the statistical variation of the manufactured designs in order to account for them adjusting and redesigning the system.

The corner process analysis will simulate the variation of the fabrication parameters of the design on the silicon wafer. So far, the DC, AC and transient analyses have been performed with a *nominal model* which assumes that the fabrication defects and variations stay under the Gaussian norm. The corner process analysis will apply extreme models that lay further in the tails (or corners) of the Gaussian distribution, simulating the worst and best case scenarios. Depending on how the analysis is performed, one can talk about front end of line (FEOL) and back end of line (BEOL) corners. The only different between FEOL and BEOL that the latter take into account the via and metal layers while FEOL only takes into account the devices.

The applied corner process analysis was the pre-layout FEOL analysis. It presents three corners, namely the fast, nominal and slow models for each transistor type (pMOS and nMOS). This was performed so that the specifications can be held in each corner scenario.

The setup models for the corner analysis were termed as follows.

- *cmostm*: refers to the nominal case, the one used for the DC, AC and transient simulations.
- *cmoswp*: worst case power, meaning that both nMOS and pMOS devices will operate at their maximum speeds.
- *cmosws*: also referred to as worst case speed, where both nMOS and pMOS devices will perform slowly.
- *cmoswo*: refers to the worst case one. In this model the pMOS devices operate slowly, while the nMOS devices are fast.
- *cmoswz*: worst case zero, is the convers of *cmoswo*, being in this model the pMOS devices the fast ones, while the nMOS devices operate slowly.

There exist other corner models for temperature, however these models are not applied to the analysis.

Along with the corner analysis, Monte Carlo simulations sample randomly the system. For instance, when tossing a stone one can analyze the problem of how far will it land from the tossing point by using the formulae derived for a parabolic shot and a set of known initial conditions. Of course this approach neglects drag and other second order effects.

Nevertheless, it is intuitive that the stone will land at a given distance from the pitcher. A Monte Carlo simulation will take into account these second order effects, which are usually random in nature, and produced a convergent solution to the given problem. In essence, a Monte Carlo simulation will introduce a number of points with random differences in the input values, within acceptable limits, allowing for a glimpse of the overall statistical behavior of the system. In microelectronic design, these input values are usually directed to model the mismatch between different transistors as well as imperfections in the system.

In general corner and Monte-Carlo analyses are paired together. The corner process analysis offers a global view of how mismatches in the fabrication technology affect the system. On the other hand, the Monte-Carlo simulations allows for the acquirement of much more localized information about the mismatches, allowing for a local view of how the technology mismatches affect the system.

4.3.-RESULTS AND DISCUSSION

In this section the results of the mentioned simulations as well as the setup of the test benches will be presented and discussed.

4.3.1.-CMFB

Both CMFB circuits, presented in section 3.5, were characterized. The followed design procedure consisted on picking transistors as in section 4.1. The CMFB circuits are represented in Figure 3.26 and Figure 3.27. The biasing circuit consisted in a simple current mirror with a $100\mu A$ current reference.

The first analysis consisted on a DC sweep to find and verify a correct bias point for the system at the picked transistor dimensions. An AC analysis should reveal the a_{CM} gain to be less than 0 dB as well as an insight to the phase margin.

RC CMFB CIRCUIT

The first circuit to be analyzed was the resistor-capacitor network CMFB circuit, the DC results are represented in Table 4.2.

	Mult.	$L(\mu m)$	$W(\mu m)$	Region	$V_{eff}\left(mV ight)$	$V_{DS}(mV)$	$V_{DS,sat}(mV)$
M1	1	1.0	3.5	2	167.9	835.9	192.8
М3	1	0.3	1.0	2	46.8	484.9	131.9
M5	1	1.0	18.9	2	114.2	179.2	165.1

Table 4.2 – DC analysis results for RC network CMFB circuit.

The frequency response of the circuit was obtained with and without the use of capacitors, and the analysis was performed by feeding an AC offset signal to the CMFB circuit inputs, together with the CM voltage reference and a 3 pF load. Results can be seen in Figure 4.11.

The DC gain of the system is $a_{CM} = -5.54 \ dB < 0 \ dB$, translating to around $a_{CM} = 0.53 \ V/V < 1$. Hence, the requirement of the system for its stability is fulfilled.

DIFFERENTIAL PAIR CMFB CIRCUIT

The same test bench was assembled for the differential pair CMFB and similar tests were run to evaluate the performance and later compare the two implementations.

The DC analysis is summarized in Table 4.3. An AC analysis was performed in the same manner as done for the RC network CFMB circuit, the only difference being the unique run since it lacks the RC common mode sense network. Results of the AC analysis can be seen in Figure 4.12.

The differential pair CMFB circuit presents a gain of $a_{CM} = -4.31 \, dB < 0 \, dB$, what amounts to roughly $a_{CM} = 0.61 \, V/V < 1$ fulfilling the stability requirement as well.

	Mult.	$L(\mu m)$	$W(\mu m)$	Region	$V_{eff}(mV)$	$V_{DS}(mV)$	$V_{DS,sat}(mV)$
M1	1	1.0	18.9	2	100.6	508.1	160.6
М3	1	0.2	0.88	2	26.3	802.0	135.7
M5	1	1.0	4.15	2	131.7	189.9	164.0

Table 4.3 - DC analysis results for differential pair CMFB circuit.

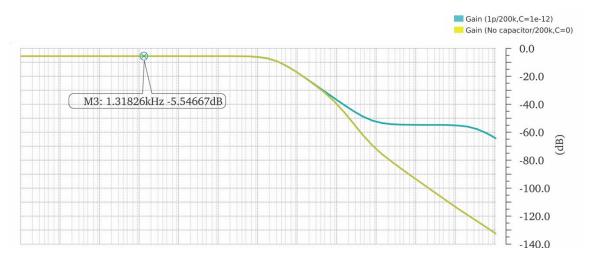


Figure 4.11 – AC analysis results for the RC CMFB circuit implementation.



Figure 4.12 – AC analysis results for the differential pair CMFB circuit implementation.

4.3.2.-FOLDED CASCODE

The DC analysis of the folded cascode OTA yielded a favorable bias point of all the transistors, being all of them in the saturation region (represented in the simulations as region 2). The results of the simulation are presented in Table 4.4. Note that in the table only one half-circuit is represented, since they are symmetrical.

	Mult.	$L(\mu m)$	$W(\mu m)$	$V_{eff}(mV)$	$V_{DS}(mV)$	$V_{DS,sat}(mV)$	$g_m(mA/V)$	$\boldsymbol{r_o}(\boldsymbol{k}\Omega)$
М1	372	0.19	3.50	98.4	1029.0	166.1	35.12	0.631
М3	496	1.00	4.15	131.7	212.1	164.0	40.66	0.416
M5	124	0.20	0.88	53.2	476.1	476.1	10.41	2.365
M7	124	1.00	18.80	126.1	625.3	172.2	15.51	8.285
M9	124	1.00	18.80	131.7	186.6	174.9	14.93	1.354
M11	744	1.00	18.80	113.6	259.0	166.4	67.71	0.713

Table 4.4 – Folded cascode OTA DC analysis, bias results.

The circuit is presented below, in Figure 4.13. The DC solution provided by the simulator shows that all the transistors are in the saturation region, as their $V_{DS} > V_{DS,sat}$. Additionally, all devices appear to have a similar effective voltage, with M_7 as an exception. Knowing this, one can assume that they are all operating in moderate inversion.

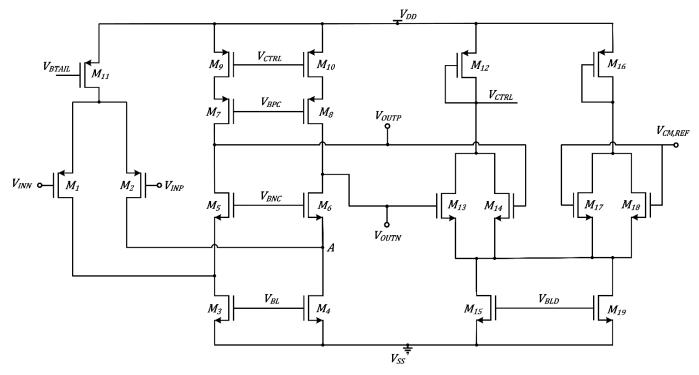


Figure 4.13 – Fully differential folded cascode OTA.

Note as well that M_{11} has to always have twice the multiplicity of M_1 , since the quiescent current will split equally through the differential pair. Furthermore, a low channel length was chosen in order to maximize transconductance and get as much GxBW as possible. And likewise, the multiplicity of M_3 will be the multiplicity of M_5 plus M_1 in order to keep the current ratio. Having said this, the current through the differential pair is 7.2 mA, already above the requirement, while the one through M_5 is 0.75 mA.

The small signal parameters, such as the transconductance and the output resistance are on the extremes. Since so much current is needed, the transconductance is huge, while this means that the output resistance of the transistors is low.

The next analysis was an AC simulation of the open loop system and the result are shown in Figure 4.14 as a Bode plot. The dominant pole is located at $f_{p1} \cong 8 MHz$, with a steady 20 *dB/dec* roll off until the unity gain bandwidth, located at $u_T \approx 219.9 MHz$ and a gain of 29.3 *dB*. This reduction in gain comes from the load resistor placed to simulate the system's response. Since the gain is given by equation 3.8, one can see that it involves the output resistance. When connecting the 1 $k\Omega$ load in parallel to the capacitor, one is effectively decreasing the output resistance of the system, reducing thus the gain.

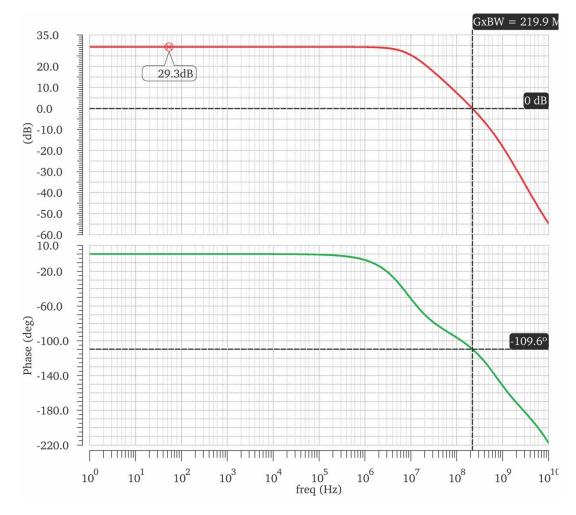


Figure 4.14 - AC results as a Bode plot for fully the differential folded cascode OTA with a the load resistor.

The phase margin can be computed as 180° plus the phase shift at GxBW, being it in this case $PM = 70.4^{\circ}$, and remaining mostly stable. Nevertheless, the shape of the phase response suggests that the second, non-dominant, pole is close to the dominant one. A reason behind this may be that since the transistors are so big, and one can approximate

through equation 3.34 as $f_{p2} \approx \frac{f_{T5}}{3}$, the cut off frequency of the transistors is lower and, as such, so is the non-dominant pole.

A transient analysis will allow to compute the slew rate. Since the SR is a large signal parameter, the used test bench includes the amplifier in voltage follower configuration with ideal square signals as inputs. The slew rate is given by the slope of the output signal (red) in Figure 4.15 with a value of around 50 MV/s, or alternatively, $50 V/\mu s$.

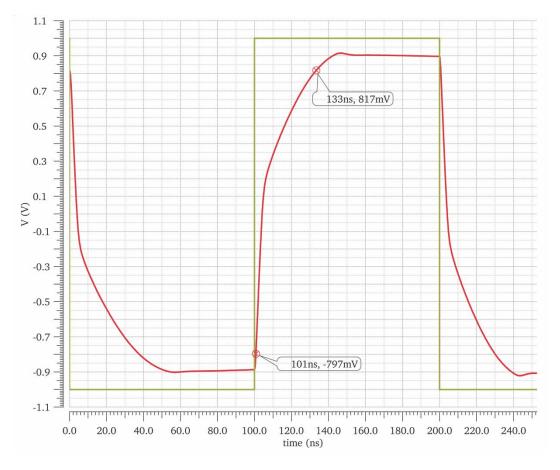


Figure 4.15 – Slew rate results for fully differential folded cascode OTA with load resistor.

The system was simulated again without the load resistor to reflect how the load resistor affects the frequency response of the OTA through a Bode plot, shown in Figure 4.16. It is possible to see how the gain has increased considerably, from the 29.3 *dB* to 50.48 *dB*. Furthermore, the *GxBW* has not shifter considerably at 217.8 *MHz*, meaning that the dominant pole has displaced to lower frequencies, around $f_{p1} \cong 0.88 \text{ MHz}$. This is expected, as the analytical description for this pole involves the output resistance, which has increased. Additionally, the shape of the Bode plot suggests that the non-dominant pole has shifted towards higher frequencies.

As a final analysis for this topology, a corner process and Monte-Carlo analyses were performed, to check that the circuit was performing properly at the achieved bias with the load resistor. From the corner analysis one can conclude that the design is properly biased, this is: all the transistors are in the saturation region. By analyzing the Monte-Carlo analysis, one can check for local variations from the sample population, which are

small due to the low standard deviation. The results of these two analyses are represented in Table 4.5 and Table 4.6 below.

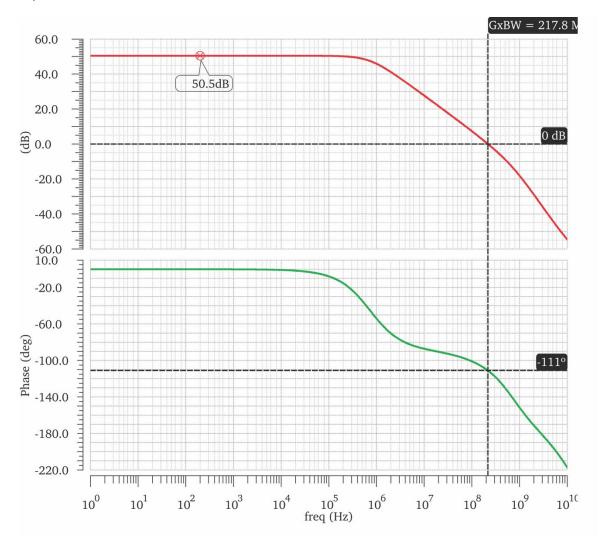


Figure 4.16 – AC results as a Bode plot for fully the differential folded cascode OTA without the load resistor.

Corner	cmostm	cmoswp	cmosws	cmoswo	cmoswz
$A_{DC}(dB)$	29.32	29.61	26.77	28.3	28.53
GxBW (MHz)	219.6	240.7	168.7	204.5	205.1
Phase Margin (deg)	70.15	73.05	72.70	74.20	70.20
$I_{DC}(mA)$	10.24	10.46	10.09	10.36	10.20

Table 4.5 – Corner process analysis results for the fully differential folded cascode OTA.

Output	Min.	Max.	Mean (µ)	Std. Dev (σ)
$A_{DC}(dB)$	27.92	30.21	29.24	0.38
GxBW (MHz)	184.7	243.8	218.2	10.15
Phase Margin (deg)	67.76	72.94	70.18	0.90
$I_{DC}(mA)$	9.68	10.83	10.24	0.226

Table 4.6 – Monte-Carlo analysis results for fully differential folded cascode OTA.

4.3.3.-MILLER CLASS A

For the DC analysis of the two stage Miller a favorable bias point was achieved for all the transistors, being all of them in the saturation region, as $V_{DS} > V_{DS,sat}$. The results of the simulation are presented in Table 4.7. Note that in the table only one half-circuit is represented, since they are symmetrical as presented for the folded cascode.

	Mult.	$L(\mu m)$	$W(\mu m)$	$V_{eff}(mV)$	$V_{DS}(mV)$	$V_{DS,sat}(mV)$
M1	124	0.18	4.00	83.9	738.3	160.7
М3	124	1.00	4.15	128.9	498.1	162.6
M5	124	1.00	4.15	128.0	752.4	162.1
M6	124	1.00	18.80	114.1	747.6	166.7
M7	248	1.00	18.80	113.6	263.6	166.4

Table 4.7 – Two stage, Miller-compensated, class A amplifier DC analysis results.

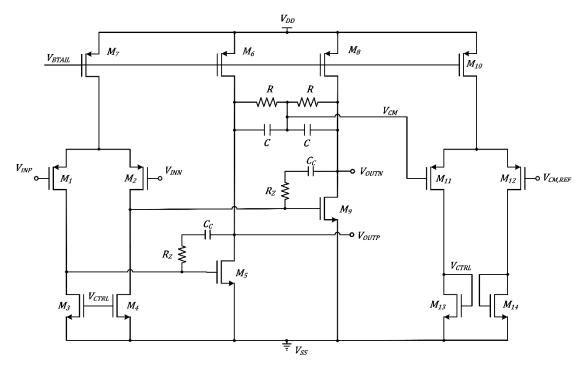


Figure 4.17 – Two stage, Miller-compensated, class A amplifier.

The open loop AC analysis with the complete load (both the capacitor and resistor) results are shown below as a Bode plot with the marked GxBW and $|A_{DC}|$ in Figure 4.18. It is important to note although it reaches the GxBW specification, the phase margin reveals that the stability is rather poor. Furthermore, from the shape of the phase response, it is possible to foresee that dominant and non-dominant poles are rather close to each other.

This results from the nature of pole splitting through the Miller-compensation, already described in previous sections. However, since the non-dominant pole is located at $\omega_{p2} = g_{m5}/C_L$, and the load capacitance is in the range of tenths of pF, the non-dominant pole will be always present at frequencies in the range of the hundreds of MHz.

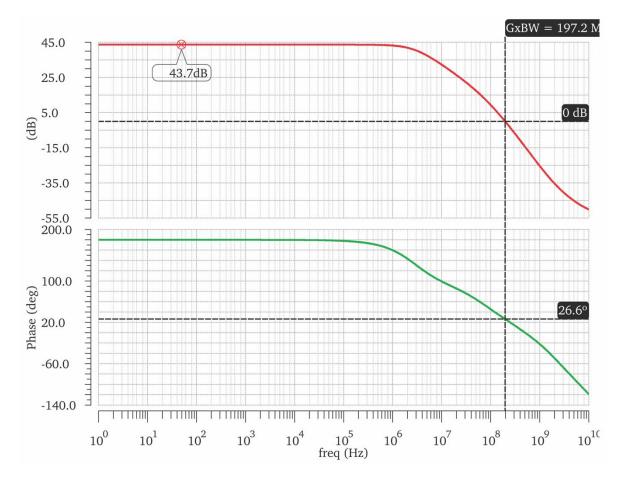


Figure 4.18 – AC analysis results for two stage, Miller-compensated, class A amplifier with load resistor.

The DC gain of the amplifier is located at 43.7 dB with the use of a load resistor, a value much higher than in the folded cascode topology. This sprouts from the fact that this is a two stage system, and the gain is given by the product of the gains of both stages. Since the load resistor is *shielded* from the first stage, only the gain in the second stage is reduced, allowing for a much higher value overall than one stage-topologies. Note that this is the basic operation of a buffer.

Another AC analysis was run to assess the performance of the amplifier without the load resistor, being the results reported in Figure 4.19. Note that the gain decrease is lower than in the previous case for the folded cascode topology.

The corner analysis was performed as well, being the results displayed in Table 4.8. This test reveals that the designed circuit is well biased for all process fabrication cases.

Corner	cmostm	cmoswp	cmosws	cmoswo	cmoswz
$A_{DC}(dB)$	40.96	39.08	44.47	40.71	40.76
GxBW (MHz)	323.0	371.0	279.4	315.0	323.2
Phase Margin (deg)	27.04	29.25	24.01	29.93	24.75
$I_{DC}(mA)$	5.116	5.190	5.072	5.076	5.188

Table 4.8 – Corner analysis results for two stage, Miller class A amplifier.

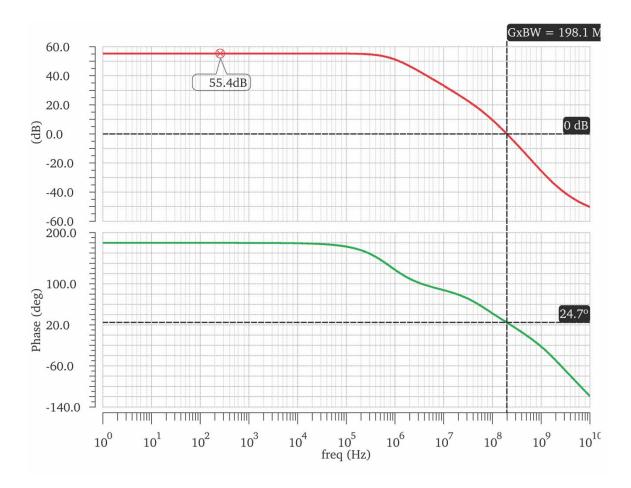


Figure 4.19 – AC analysis results for two stage, Miller-compensated, class A amplifier without load resistor.

4.3.4.-MILLER CLASS AB

The design approach to the amplifier was developed in collaboration with the PhD students. After the optimization process, the circuit devices (without taking into account the bias circuits and CMFB networks) presented the dimensions and results stated in Table 4.9, below.

	Mult.	$L(\mu m)$	$W(\mu m)$	$V_{eff}(mV)$	$V_{DS}(mV)$	$V_{DS,sat}(mV)$
M1	1	0.20	63.03	99.5	211.0	166.4
М3	15	0.20	4.20	105.7	401.6	168.2
M5	3	0.20	2.04	165.6	387.9	193.8
M7	3	0.20	2.04	172.7	261.5	195.9
M9	45	1.00	12.99	157.4	238.1	187.4
M10	15	0.20	2.00	192.7	596.8	209.5
M11	3	0.20	2.04	168.2	903.2	193.8
M12	100	0.20	2.00	196.4	750.0	211.4
M13	24	0.20	2.04	165.9	750.0	192.7

Table 4.9 – Two stage, Miller-compensated, class AB amplifier DC analysis results.

The DC analysis reveal that all the devices remain in the saturation region and within the same inversion-curve location.

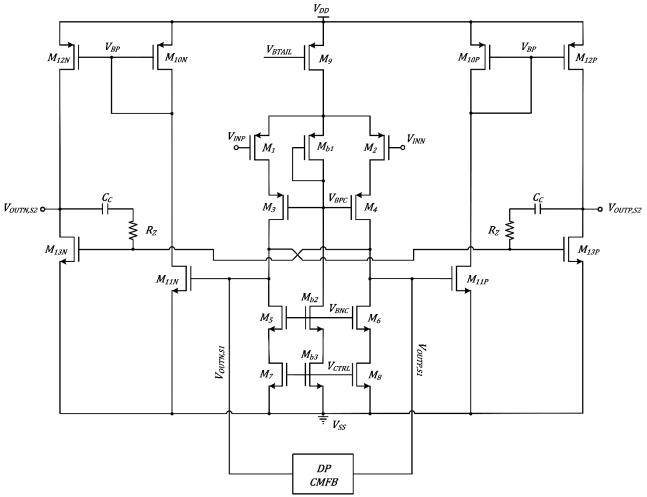


Figure 4.20 – Two stage, Miller-compensated, class AB amplifier.

The AC analysis was run in the same configuration as for the tests above, being the results shown in Figure 4.21. The gain is at 64.1 *dB*, with a steady 20 *dB/dec* roll off until the unity gain frequency, located at $u_T \approx 255.7 \text{ MHz}$. The phase margin is of 47.5°, indicating that the system is moderately stable.

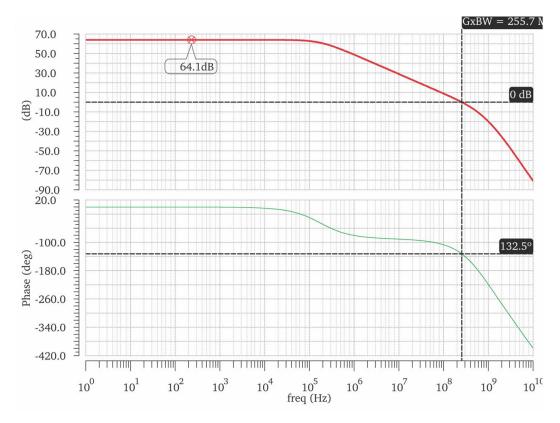


Figure 4.21 – AC Analysis results for the two stage, Miller-compensated, class AB amplifier.

A transient analysis with the amplifier in a voltage buffer configuration, with $1 k\Omega$ resistor loads, was performed in order to obtain the slew rate. The results are shown in Figure 4.22 in the next page. The slew rate was computed as the slope between 10% and 90% of the output signal value, being $SR \cong 18.5 V/\mu s$. Note that only the positive half signal is presented.

The corner process and Monte-Carlo analyses were performed as for the previous topologies, with the reported results presented in Table 4.10. Overall, the results are positive, being the gain not maintained only when model *cmoswo* is applied. Nevertheless, the amplifier still is operational, with a higher phase margin, but with worse power consumption and gain.

Corner	cmostm	cmoswp	cmosws	cmoswo	cmoswz
$A_{DC}(dB)$	64.08	60.54	60.62	26.62	65.65
GxBW (MHz)	47.51	48.72	46.13	60.32	40.82
Phase Margin (deg)	255.7	271.5	310.9	213.3	234.6
$I_{DC}(mA)$	3.95	3.98	3.98	18.7	3.74

Table 4.10 – Corner analysis results for two stage, Miller class AB amplifier.

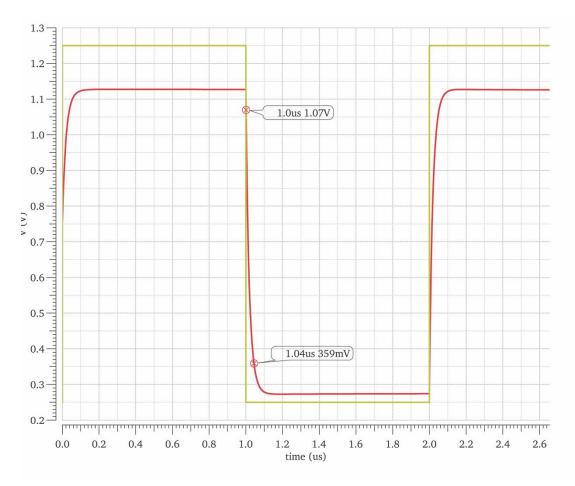


Figure 4.22 – Slew rate results for the two stage, Miller-compensated, class AB amplifier.

Finally, the Monte-Carlo analysis was performed, with results reported in Table 4.11 below.

Output	Min.	Max.	Mean (µ)	Std. Dev (σ)
$A_{DC}(dB)$	60.95	69.40	64.53	1.68
Phase Margin (deg)	41.71	53.41	46.9	1.729
GxBW (MHz)	218.1	294.3	253.2	13.4
$I_{DC}(mA)$	3.17	5.03	4.00	0.37

Table 4.11 – Monte-Carlo analysis results for the two stage, Miller class AB amplifier.

As it can be observed, the standard deviation suggests non-significant local variations among 100-point sample population, confirming the robustness of the designed circuit.

4.4.-DESIGN CONCLUSIONS

In this section, the main conclusions driven from the results in the previous section are presented and summarized.

4.4.1.-CMFB

The first decision that has to be taken was the CMFB circuit choice. By taking a look at the results obtained from the simulation of the RC CMFB implementation, one can realize that the addition of a capacitor will partially cancel the effect of the pole introduced by the resistors. This can be seen in Figure 4.11, where the Bode diagram is shown for the system with and without the capacitor.

The choice of the CMFB circuit did not only take into account the results presented in the previous section. Since the RC implementation is basically an impedance at the output of the amplifier, it will load the system. Nonetheless, it presents a much higher output range [ref]. Conversely, the differential pair implementation will have a much shorter output range, given by the gate-source potential at which the differential pair transistors leave the saturation region.

The final design choice came down to the speed of the differential pair implementation and, in lower grade, due to the unnecessary loading of the amplifier by the RC implementation. The drawn conclusions are that the DP CMFB circuit should be used in differential circuits which do not have great common mode variations (and if great bandwidths are needed), while the RC CMFB circuit ought to be used when the variations in the common mode level are expected to be high.

4.4.2.-FOLDED CASCODE

A single stage folded cascode was not an efficient way of driving such a high capacitive impedance and such low resistive load since the gain is reduced greatly. A lot of current is spent to increase the gain attenuated by the load resistor.

Nevertheless, it high gain when configured with a capacitive load (through which is also compensated) makes it a great candidate for a first stage implementation in a multi-stage system.

4.4.3.-TWO STAGE MILLER CLASS A

The class A Miller amplifier presented too little gain on the first stage when the voltage was reduced from 1.8 to 1.5. Topologies like the folded cascode (described above) or a telescopic cascode are needed to increase this first stage gain.

Further current was wasted in the second stage due to class A operation, being current sank through either the load resistor or the current sink transistor respectively at any time of operation. This makes the topology not suitable for efficient, low power (and high demand) operation.

4.4.4.-TWO STAGE MILLER CLASS AB

The two-stage Miller, class AB amplifier is an adequate solution to drive the demanded load efficiently at low voltage. It maintains a high gain thanks to the first stage telescopic

cascode topology, while being able to efficiently source current as a result of the class AB output stage.

Nevertheless, it is complex to bias, having a considerable bias circuit that consumes some marginal power. Furthermore, the topology requires more transistors than other topologies, increasing the area on the chip.

4.5.-FUTURE WORK

It is possible to save more power by lowering the voltage to 1.2 V, however the first stage may not fit inside this voltage. Potentially, one can optimize the first stage to work at such potential.

Another feasible option is to change to a folded cascode topology for the first stage, similar to the designed one.

5.-CONCLUSIONS

Along the span of this bachelor thesis, several topologies have been analyzed and presented with the analytical expressions for their quiescent, frequency and transient response. These equations were put into practice with the implementation of the respective circuits in a 0.18 μ m, 1.8 V technology, scaled down to 1.5 V.

Several problems appear when a technology is used below its corresponding voltage, as the threshold voltage does not scale accordingly. This means that there is less voltage range to fit in transistors, and some topologies can be discarded if the use of wide swing cascodes is not considered.

Another issue with the technology is that, although nowadays it can be considered a long channel process, in practice second order effects can be perceived.

A recurring topic with respect to the analytical models derived for CMOS is that they rarely correspond to the simulation results. Since the simulation software uses the models extracted from the experimental data (this is: the already fabricated transistors), several effects are already accounted for. These effects are not modelled by the presented analytical equations and, therefore, they differ from the simulation data. As such, the derived equations for each topology are a mere tool to understand the response of the system and guide the researcher intuitively during the implementation process.

Nevertheless, the main conclusion one can draw from this bachelor thesis relates to the reasons behind the implementation of a given topology to drive a corresponding load. Several topologies were explored and their main drawbacks to drive the proposed load were identified. Table 3.5 presented an analytical comparison between the different discussed topologies in that matter. With the results obtained in the previous section, one can update it in order to summarize the main points of the paper in Table 5.1.

	Folded Cascode	Telescopic	Two stage Miller	Two stage class
		Cascode	class A	AB
Gain	45 – 60 dB	45– 60 dB	55 – 70 dB	55 – 70 dB
Voltage Swing	$2(V_{DD}-4V_{DS,sat})$	$2(V_{DD}-5V_{DS,sat})$	$2(V_{DD} - 3V_{DS,sat})$	$2(V_{DD} - 3V_{DS,sat})$
Compensation	Load capacitance	Load capacitance	Miller-compensation	Miller-compensation
Loads	Mainly capacitive	Mainly capacitive	Any load	Any load
CMFB	Simple	Simple	Simple	Complex
Bias	Intermediate	Intermediate	Simple	Complex
Stability	$f_{p2} \approx \frac{f_{T5}}{3}$	$f_{p2} \approx \frac{f_{T5}}{2}$	$f_{p2} \approx \frac{g_{m2}}{2\pi C_L}$	_
Current	4 <i>I</i> _{diff}	2I _{diff}	$(2+n)I_{diff}$	_
consumption			<i>n</i> = 1, 2	

Table 5.1 – Update topology comparison table.

6.-BUDGET AND WORK PLAN

The budget for the research involved in this thesis includes the usage of a Cadence Virtuoso Schematics and ADE XL license, as well as the paid hours of the research internship assistant.

	Unitary price (€)	Work hour pricing (€ / h)	Utilization in thesis	Total hours
Cadence XL and Schematics License	3600	0.6	60%	111.6
Research intern	930	5	90%	186.0
Total				996.96€

Table 6.1 – Budget of the developed research.

A work plan was established in order to address the workflow of the project. One of the main goals in the re-design of the $\Sigma\Delta$ ADC was to devise and implement a more efficient system. Part of the power consumption resided in the local feedforward topology prior to the flash ADC, presented in Figure 1.2. The implemented amplifiers were class A Millers, being their current in the order of 8 mA.

The first task was to explore the different topologies and their main issues when implemented as integrators and adders, with the proposed loads (which, in turn, were set due to the state variables).

Once a topology was chosen it was characterized through a DC analysis to determine if it was correctly biased. If the results were negative, the issues were assessed through the power law relations, and the transistors were modified accordingly. This was repeated until the results of the DC analysis were convergent.

The next step is the characterization of the frequency and time response through an AC and transient analysis. If the results were negative, the issues were identified and the analytical models for the corresponding topology were used to guide the re-design of the circuit. This was iterated until a positive solution was reached.

Lastly, the robustness of the circuit was determined through corner process and Monte-Carlo analyses. If the results were not favorable, another topology was chosen that attained for the shortcomings of the currently explored one. On the other hand, if the results were positive, the layout of the circuit was ready to be overtaken. All this workflow is summarized in Figure 6.1.

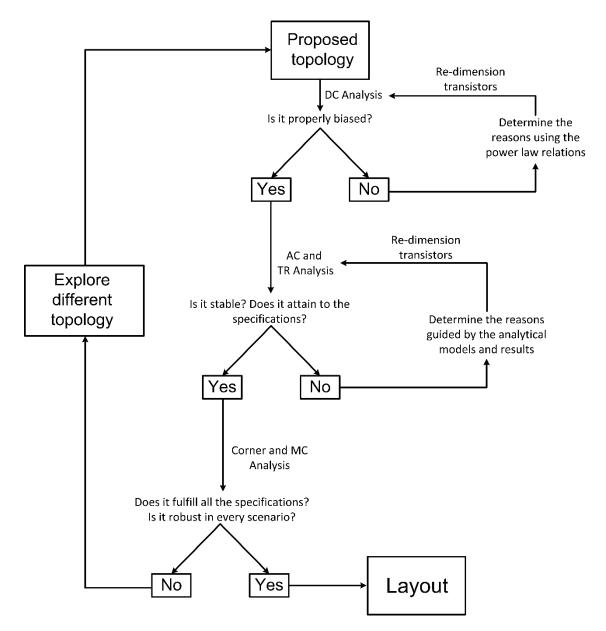


Figure 6.1 – Work plan applied throughout the thesis.

7.-BIBLIOGRAPHY

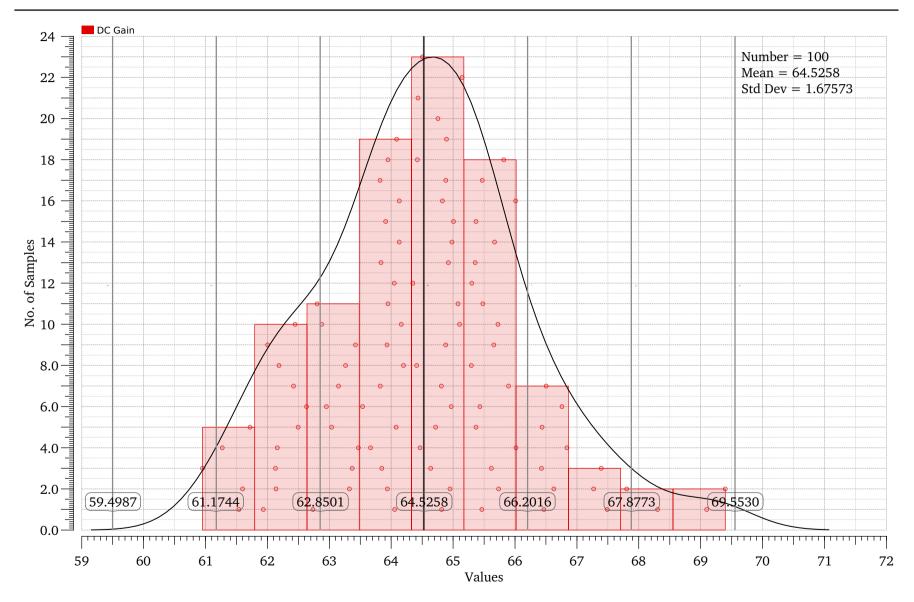
- [1] M. Domínguez Suárez, "Diseño del layout y caracterización de un circuito integrado CMOS para aplicaciones biomédicas : convertidor analógico digital sigma delta de doble rampa," Madrid, 2013.
- [2] B. Razavi, "CMOS technology characterization for analog and RF design," Proceedings of the IEEE 1998 Custom Integrated Circuits Conference (Cat. No.98CH36143), Santa Clara, CA, 1998, pp. 23-30.
- [3] T. W. Kim, K. C. Woo, S. K. Hwang, M. J. Kim and B. D. Yang, "A-250mV supplyvoltage 65dB-gain OTA with an enhanced bandwidth and a reduced compensationcapacitor," 2018 International Conference on Electronics, Information, and Communication (ICEIC), Honolulu, HI, 2018, pp. 1-4.
- [4] T. Sharan, V. Bhadauria and P. Chetri, "Bulk-driven feed-forward compensated sub-threshold low-voltage OTA to drive high capacitive load," 2017 International Conference on Innovations in Electronics, Signal Processing and Communication (IESC), Shillong, 2017, pp. 121-126.
- [5] Jamuna G, S. Yellampalli and Swetha S, "Design and implementation of telescopic OTA in 8 bit second-order continuous-time band-pass Sigma-Delta ADC," 2014 International Conference on Electronics, Communication and Computational Engineering (ICECCE), Hosur, 2014, pp. 1-7.
- [6] S. Mallya and J. H. Nevin, "Design procedures for a fully differential folded-cascode CMOS operational amplifier," in IEEE Journal of Solid-State Circuits, vol. 24, no. 6, pp. 1737-1740, Dec 1989.
- [7] D. A. Jones and K. W. Martin, "Integrated-Circuit Devices and Modelling," in Analog Integrated Circuit Design, NY, John Wiley & Sons, 1997, pp. 71-74.
- [8] A. S. Sedra and K. C. Smith, "MOS Field Effect Transistors (MOSFET)," in Microelectronics Circuits, NY, McGraw Hill, 2004, p. 288.
- [9] R. S. Muller and T. I. Kamins, "MOS Field-Effect Transistors I: Physical Effects and Models," in Device Electronics for Integrated Circuits, NY, Wiley & Sons, 1987, p. 448.
- [10] P. R. Gray, P. J. Hurst, S. H. Lewis and R. G. Meyer, "Models for Integrated-Circuit Active Devices," in Analysis and Design of Analog Integrated Circuits, NY, John Wiley & Sons, 2001, pp. 59-63.

- [11] B. Sviličić and A. Kraš, "CMOS technology: challenges for future development," Pomorstvo: Journal of Maritime Studies, vol. 20, no. 2, pp. 97-104, 2006.
- [12] R. J. Baker, "The MOSFET," in CMOS Circuit Design, Layout and Simulation, NJ, Wiley & Sons, 2010, pp. 133-140.
- [13] S. Dimitrijev, "The Threshold Voltage and the Body Effect," in Principles of Semiconductor Devices, NY, Oxford, 2012, pp. 304-307.
- [14] P. R. Gray, P. J. Hurts, S. H. Lewis and R. G. Meyer, "Frequency Response and Stability of Feedback Amplifiers," in Analysis and Design of Analog Integrated Circuits, NY, John Wiley & Sons, 2001, p. 632.
- [15] A. Garimella and P. M. Furth, "Frequency compensation techniques for op-amps and LDOs: A tutorial overview," 2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS), Seoul, 2011, pp. 1-4.
- [16] D. A. Johns and K. W. Martin, "Folded-Cascode Opamp," in Analog Integrated Circuit Design, NY, John Wiley & Sons, 2012, p. 268.
- [17] B. Razavi, "Association of Poles with Nodes," in Design of analog CMOS integrated circuits, NY, McGraw-Hill, 2001, pp. 118-119.
- [18] B. Razavi, "One-stage Op Amps," in Design of analog CMOS integrated circuits, NY, McGraw-Hill, 2001, pp. 169-171.
- [19] D. A. Johns and K. W. Martin, "Basic Opamp Design and Compensation," in Analog Integrated Circuit Design, NY, John Wiley and Sons, 2012, pp. 242-249.
- [20] M. Yavari, N. Maghari and O. Shoaei, "An accurate analysis of slew rate for twostage CMOS opamps," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 52, no. 3, pp. 164-167, March 2005.
- [21] D. M. Monticelli, "A quad CMOS single-supply op amp with rail-to-rail output swing," in IEEE Journal of Solid-State Circuits, vol. 21, no. 6, pp. 1026-1034, Dec 1986.
- [22] P. R. Gray, P. J. Hurst, S. H. Lewis and R. G. Meyer, "CMOS Class AB Output Stages," in Analysis and Design of Analog Integrated Circuits, NY, John Wiley & Sons, 2001, pp. 382-397.
- [23] D. A. Johns and K. W. Martin, "Fully Differential Amplifiers," in Analog Integrated Circuit Design, NY, John Wiley & Sons, 2012, p. 282.
- [24] P. R. Gray, P. J. Hurst, S. H. Lewis and R. G. Meyer, "Fully Differential Operational Amplifiers," in Analysis and Design of Analog Integrated Circuits, NY, John Wiley & Sons, 2001, p. 810.

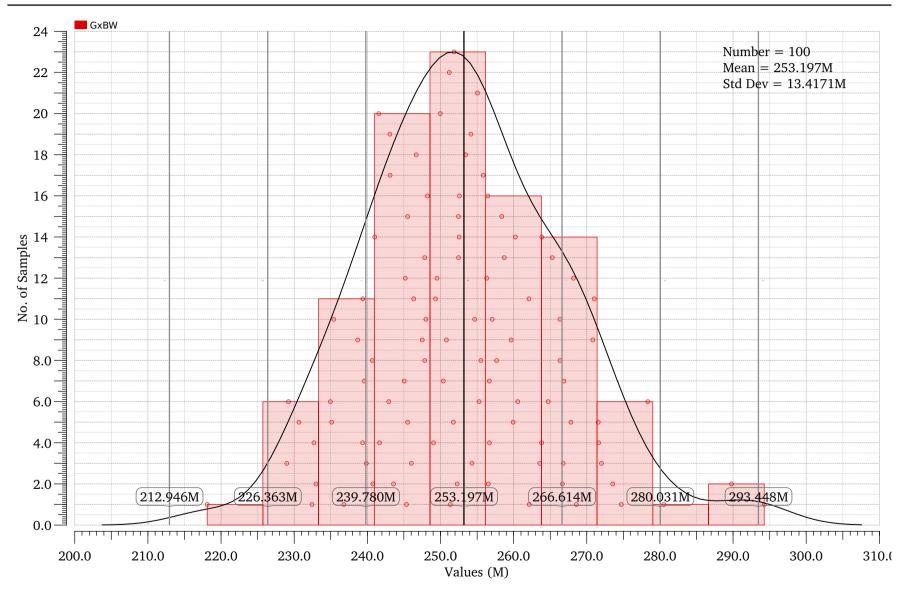
- [25] D. A. Johns and K. W. Martin, "Fully Differential Opamps," in Analog Integrated Circuit Design, NY, John Wiley & Sons, 2012, p. 281.
- [26] B. Razavi, "Common-mode Response," in Design of analog CMOS integrated circuits, NY, McGraw-Hill, 2001, pp. 303-305.
- [27] R. J. Baker, "Compensating the CMFB Loop," in CMOS Circuit Design Layout and Simulation, NJ, Wiley & Sons, 2010, p. 871.
- [28] P. R. Gray, P. J. Hurst, S. H. Lewis and R. G. Meyer, "CMFB Circuits," in Analysis and Design of Analog Integrated Circuits, NY, Wiley & Sons, 2001, pp. 814-815.
- [29] K. S. Kundert, "Circuit Simulation," in The designer's Guide to SPICE and Spectre, NY, Springer US, 1995, pp. 1-14.
- [30] K. S. Kundert, "AC Analyses," in The Designer's Guide to SPICE and Spectre, NY, Springer US, 1995, pp. 51-55.

ANNEX A - MONTE CARLO HISTOGRAMS FOR MILLER CLASS AB

DC Gain







3W

Phase Margin

